

Epoch-based Entropy for Early Screening of Alzheimer's Disease

N. Houmani^{*,†,§}, G. Dreyfus^{*,†,¶} and F. B. Vialatte^{*,†,||,**}

^{*}ESPCI ParisTech, PSL Research University

10 rue Vauquelin, 75005 Paris, France

[†]SIGMA (SIGnal processing and MACHine learning) Laboratory

10 rue Vauquelin, 75231 Paris Cedex 05, France

[‡]Brain Plasticity Laboratory, CNRS UMR 8249

10 rue Vauquelin, 75231 Paris Cedex 05, France

[§]Nesma.Houmani@espci.fr

[¶]Gerard.Dreyfus@espci.fr

^{||}Francois.Vialatte@espci.fr

Accepted 12 August 2015

Published Online 9 November 2015

In this paper, we introduce a novel entropy measure, termed epoch-based entropy. This measure quantifies disorder of EEG signals both at the time level and spatial level, using local density estimation by a Hidden Markov Model on inter-channel stationary epochs. The investigation is led on a multi-centric EEG database recorded from patients at an early stage of Alzheimer's disease (AD) and age-matched healthy subjects. We investigate the classification performances of this method, its robustness to noise, and its sensitivity to sampling frequency and to variations of hyperparameters. The measure is compared to two alternative complexity measures, Shannon's entropy and correlation dimension. The classification accuracies for the discrimination of AD patients from healthy subjects were estimated using a linear classifier designed on a development dataset, and subsequently tested on an independent test set. Epoch-based entropy reached a classification accuracy of 83% on the test dataset (specificity = 83.3%, sensitivity = 82.3%), outperforming the two other complexity measures. Furthermore, it was shown to be more stable to hyperparameter variations, and less sensitive to noise and sampling frequency disturbances than the other two complexity measures.

Keywords: Alzheimer's disease; EEG signal; complexity measures; entropy; hidden Markov models; stationary multichannel epochs.

1. Introduction

With the unprecedented ageing of the Western population, dementia and neurodegenerative disorders have become a major societal concern. Alzheimer's disease (AD) is the most common form of dementia; it affects 11% of the world population aged over 65¹ and 50% of the persons aged over 85. The number of individuals with AD is expected to reach 115 million in 2050.² AD is characterized by irreversible brain damages, associated with memory impairments and

a wide range of cognitive dysfunctions. The causes of AD are not identified; however, the impairments in cognitive functions reflect the spread of this pathology from medial-temporal to parietal brain areas.³⁻⁶ Besides, at the early stage, symptoms of AD are often dismissed as normal consequences of aging.

The early detection of AD has three main interests. First, the patient and his caregivers can understand the daily consequences of the disease. Second, the diagnostic can help them anticipate the future.

^{||}Corresponding author.

^{**}F.B. Vialatte was in SIGMA laboratory when the research reported in this manuscript was prepared.

Finally, even if treatments cannot stop the disease yet, early therapeutic interventions may delay its evolution.^{7,8}

Electroencephalography (EEG) has been considered recently in several studies as a potential tool for diagnosing AD. The main advantage of EEG is its relatively low cost, and its high temporal resolution that allows analyzing fast dynamics in the cortex. Nevertheless, diagnosing AD with EEG signals at the early stage remains a challenge.^{5,9–12} This is due to the complex nature of EEG signals that must be modeled as nonstationary, nonlinear and multidimensional time series.^{13–15}

The application of complexity theory to biosignals can provide some relevant insights for clinical applications.¹⁶ Indeed, several studies have highlighted that one of the major effects of AD is the reduction in complexity of the EEG signal compared to that of healthy subjects.^{14,15,17} However, it is not always easy to detect such effects because of the large inter-variability between AD patients. Several methods can be used in order to assess the complexity of EEG signals. The fractal dimension⁷⁴ was applied to EEG signals in Refs. 14 and 18–20 as a potential marker of AD. The correlation dimension^{19,20} and the first positive Lyapunov exponent²¹ were frequently used.^{15,22–27} The correlation dimension ($D2$) reflects the number of independent variables that are necessary to describe the dynamics of the system. The Lyapunov exponent ($L1$) describes the divergence of trajectories starting at nearby initial states. It has been found that EEG signals from AD patients exhibit lower values of such measures (hence lower complexity) than signals from age-matched normal subjects^{15,24} in almost all EEG channels. $D2$ and $L1$ were found useful for detecting changes between different brain states. Nevertheless, these two measures are computationally expensive, since they involve the reconstruction of a phase space trajectory reconstruction.¹⁹

Therefore, alternative solutions suitable for sparse data were introduced. They rely on quantifying signal complexity from the point of view of information theory. In this context, the complexity of a signal refers to its unpredictability: irregular signals are more complex than regular ones since they are more unpredictable. In that framework, several measures were proposed to assess the complexity of the signal; most of them rely on the concept of entropy²⁸:

sample entropy,^{29,30} Tsallis entropy,³¹ approximate entropy,^{32,33} multi-scale entropy,^{34,35} and Lempel–Ziv complexity.³⁶

These measures, however, are not suitable for the analysis of EEG data; they share two main drawbacks. First, such measures were computed on whole EEG sequences without addressing the problem of their nonstationarity. The assumption of stationarity is generally not true with physiological data. In order to apply complexity measures to a nonstationary time series such as EEG, one should take some precautions: double check for nonstationarity effects,³⁷ and remove nonstationarities.³⁸ Multi-scale entropy, though powerful, is also not suitable for studying EEG signals due to its linear extraction of scales.³⁵ In this context, some studies concluded that EEG time series are quasi-stationary: in Ref. 39, the authors suggested that EEG can be described as a piecewise stationary process, i.e. EEG data can be segmented into stationary segments with different probabilistic characteristics. In Ref. 40, the authors claimed that the EEG signal could be modeled as a sequence of quasi-stationary segments (epochs) separated by abrupt transitions. Other studies^{41–43} identified quasi-stationary states in EEG, called “microstates”. These states are supposed to reflect coherent neural activities. Also, in Ref. 44, the author suggested that perception is based on sequences of stationary patterns demarcated by discontinuities.

Secondly, such measures did not consider the EEG signal as a multidimensional time series: the prevailing paradigms extract information from EEG signals by averaging them over channels. The EEG being a multidimensional signal provided by a number of electrodes (channels), it is of high potential interest to exploit its spatio-temporal nature through techniques that can take into account inter-channel relations. However, alternative methods were used for studying the EEG background activity. For instance, mutual information analysis,^{15,45} functional graphs,^{46,47} measures of synchrony⁴⁸ (e.g. Pearson correlation coefficient and Granger causality) and measures of coherence^{6,15,49–51} were used for assessing information transmission between different brain areas.

The purpose of the present work is to use the concept of entropy²⁸ to characterize EEG complexity in order to discriminate AD patients, at the early stage,

from healthy subjects. As mentioned above, different variants of entropy were used previously; they differ in theoretical foundations and computational complexity; their efficiency is problem dependent. Our contribution consists of the introduction of a novel entropy measure derived from a refined characterization of the local statistical properties of the EEG signal, by means of a Hidden Markov Model (HMM).⁵² This new entropy measure, hereinafter referred to as epoch-based entropy, is computed on stationary epochs of multi-channel data, and takes into account the nonstationarity and multi-dimensionality of EEG data. We will show that this modeling approach for EEG improves the analysis of the underlying neuronal dynamics. We extend our preliminary results⁵³ and demonstrate the effectiveness of the proposed complexity measure for AD screening.

Two alternative complexity measures are also used as ground truth to assess the effectiveness of our approach: Shannon’s entropy,^{28,54} which has the same mathematical background as the proposed measure, and the correlation dimension, which is inherited from chaos theory⁷³ and frequently used in the literature for AD diagnosis.^{15,22,55,56} The classification accuracy of each measure for the discrimination of AD patients from healthy subjects are estimated using a linear classifier designed on a developmental dataset, and subsequently tested on an independent test set. The robustness of the measures to variations in the levels of noise and different sampling rates is also investigated.

The remainder of the paper is organized as follows. In Sec. 2, the proposed complexity measure is presented after a reminder of the correlation dimension and of Shannon’s entropy. In Sec. 3, we describe the EEG databases used for the experiments. In Sec. 4, we present and analyze the results. Discussion on the results and conclusions are stated in Sec. 5.

2. Definition of the Three Complexity Measures for EEG Analysis

2.1. Correlation dimension

Correlation dimension (denoted D_2) was primarily designed as a measure of the complexity of dynamic systems^{19,20,56}; it was widely used to investigate the nonlinear dynamics of human EEG.^{15,55,57–61} A signal that exhibits a high complexity has a high correlation dimension value, which indicates that many

degrees of freedom are needed to describe such a signal.

Consider a univariate time series $x(1), x(2), \dots, x(K)$ with sampling period T . The time series can be embedded into an m -dimensional space by defining, for each sample i , an m -dimensional vector $\mathbf{V}_i = \{x(i), x(i + \tau), \dots, x(i + (m - 1)\tau)\}$, where the lag τ is an integer with the condition $i \leq K - (m - 1)\tau$. Thus, in the embedding space, the time series is represented by a trajectory. Following Grassberger and Procaccia,^{19,20} the quantity $C_i^m(r)$ is defined as the proportion of points of the trajectory that lie within a sphere of radius r centered on point i , and it is subsequently averaged over the whole trajectory:

$$C^m(r) = \frac{1}{N - (m + 1)\tau} C_i^m(r). \quad (1)$$

$C^m(r)$ is thus the average proportion of pairs of points of the trajectory that fall within a distance r of each other. It is shown in Refs. 60 and 61 that, for sufficiently high embedding dimension, this quantity (known as the correlation integral) grows with r as r^{D_2} where D_2 is the correlation dimension; therefore, the correlation dimension can be estimated as the slope of the graph of $\log(C^m(r))/\log r$ when $r \rightarrow 0$. A small correlation dimension results in a slow growth of the correlation integral, due to the fact that most points of the trajectory in embedding space are very close to each other: the time series has low complexity. By contrast, a large correlation dimension results in a rapid growth of the correlation integral, due to the fact that the points of the trajectory in embedding space are far apart: the time series has high complexity.

The above procedure requires the choice of three quantities: the embedding dimension m , the lag τ , and the range of variation of r .

The above computation of correlation dimension is applied to univariate time series (single-channel EEG). In order to exploit the multi-channel EEG recordings as mentioned in Sec. 1, each EEG time series should be considered separately. This method may be followed for signals that are uncorrelated. This is not the case of EEG signals, where some correlation effects exist at least at the region level. Additionally, the analysis of single EEG channels may imply a loss of relevant information related to inter-channel variability. We therefore propose in this work to concatenate the multi-channel EEG signals

into a single vector, and then to compute the correlation dimension of that vector. This solution has the advantage of expressing the complexity of a set of EEG time series of a subject by a single value resulting from the analysis of multi-channel EEG signals.

2.2. Shannon's entropy

The concept of information entropy was introduced by Shannon in 1948,⁵⁴ and is generally referred to as "Shannon's entropy".²⁸ Entropy has different meanings that depend on the application. In physics, entropy is a measure of disorder: the higher the disorder, the larger the entropy of the system under consideration. In information theory, entropy is a measure of uncertainty: it quantifies the predictability of future realizations of the signal time series based on the probability distribution of past realizations.

Shannon's entropy of a discrete random variable Z with K possible realizations $z_i, i = 1, \dots, K$, is defined by:

$$H(Z) = - \sum_{i=1}^K p_i \cdot \log_2(p_i), \quad (2)$$

where p_i is the probability of outcome z_i , with $\sum p_i = 1$. Entropy is thus a measure of how uniformly distributed the random variable Z is: if there are K outcomes, the Shannon entropy is maximum, equal to $\log_2(K)$, if $p_i = 1/K \forall i$ point Base 2 logarithm is used in order to express entropy in bits.

Shannon's entropy is thus the mean number of bits needed to describe the random variable. Based on this interpretation, entropy can also quantify the complexity of time series: a signal of high complexity is constructed from a large number of elementary patterns and thus presents a high entropy value. Applied on a univariate EEG signal, Shannon's entropy measures the complexity of the EEG signal based on the probability distribution of amplitude values observed in the signal. For a multivariate EEG analysis and comparable results with the correlation dimension, we proceed following the same approach as that described in Sec. 2.1: we concatenate the multi-channel EEG signals of a subject into one single vector and then compute the Shannon entropy of the concatenated vector. Thus, the complexity of a set of EEG time series of a subject is expressed by a single value of Shannon's entropy.

2.3. Epoch-based entropy measure

Numerous entropy measures were proposed in the literature for quantifying the complexity of EEG signals.^{19,21,29-35} Despite the common name of entropy, these measures have different mathematical backgrounds, hence describe different properties in the signal.

As defined in Sec. 2.2, the entropy of a random variable depends only on its probability density value. The present work presents a new entropy measure based on the fundamental assumption that the EEG signal is piecewise stationary, i.e. can be viewed as being stationary at the time scale of an epoch.³⁹⁻⁴⁴ In this context, HMMs,^{52,62} which are extensively used in speech recognition for instance, are good candidates for estimating complexity in piecewise stationary signals: they can segment the EEG signals into stationary epochs, and at the same time perform a local estimation of the probability density on each epoch.

A HMM is a probabilistic model that can be used to describe the evolution of observable events or signal realizations, called "observations", which depend on internal factors that are not directly observed, called "hidden states".⁶² Emission probabilities are the conditional distributions of the observed variables from a specific state. For continuous observations (such as EEG signals), the emission probabilities are continuous. Each state can output an observation based on the observation probability distribution. Therefore, the use of HMM in this framework is also motivated by the fact that HMM's structure is adapted for modeling neural dynamics underlying the observed EEG signals. Such statistical modeling for EEG signals was already applied in the literature, for instance for sleep staging applications^{63,64} or for motor imagery classification.⁶⁵

In the present work, as in our preliminary study,⁵³ EEG signals are modeled by a continuous left-to-right HMM (Fig. 1): the topology of the model allows transitions from each state to itself and to its immediate right-hand neighbors only. The states of the HMM correspond to the stationary parts of the EEG signal, and the transitions of the HMM correspond to the variations of the signal. We thus consider the EEG signal of a given subject as a succession of epochs, obtained by segmenting such a signal via the Viterbi algorithm⁶² using the corresponding

subject's HMM. Viterbi's algorithm is a widely used algorithm in HMMs to find the best state sequence. It can be viewed as a modified forward algorithm: instead of summing up the probabilities from all different paths, the optimal path, called "Viterbi path",^{52,62} is chosen. Thus, each obtained epoch corresponds to a state of the HMM and contains a given number of observations (sample points). For each epoch \mathbf{S}_i the probability density function is modeled by a mixture of M Gaussian functions considering a diagonal covariance matrix for each multivariate Gaussian.

Then each observation z in a given epoch \mathbf{S}_i is considered as a realization of a random variable Z_i that follows a given observation probability distribution $P_i(z)$ modeled by the Gaussian mixture. Consequently, a random variable is associated to each stationary epoch of the signal (Fig. 1), and the entropy $H^*(Z_i)$ of the considered epoch \mathbf{S}_i is that of an ensemble of realizations of Z_i :

$$H^*(Z_i) = - \sum_{z \in \mathbf{S}_i} P_i(z) \cdot \log_2 P_i(z). \quad (3)$$

The EEG sampling period (typically 8 ms) is small with respect to the epoch length (typically 250 ms). Therefore, although Z_i is a discrete variable, one takes advantage of the continuous emission probability law estimated on each epoch by the HMM.

By averaging the entropy over all the epochs of the EEG signal of the considered subject, an entropy-based complexity value $\text{EpEn}(Z)$ of the considered

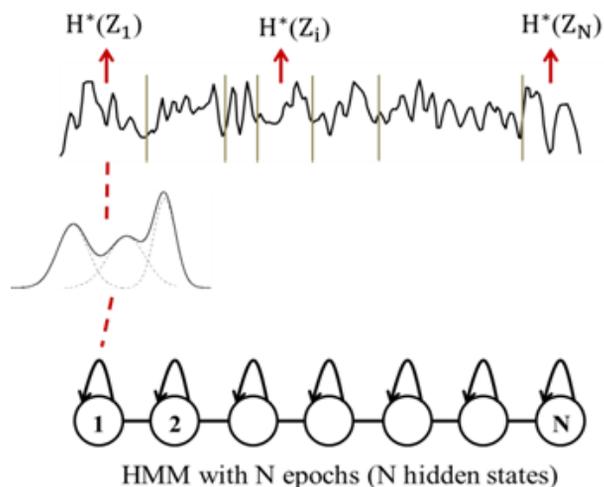


Fig. 1. Epoch-based entropy computation of a univariate EEG signal.

signal, called "epoch-based entropy", is obtained as:

$$\text{EpEn}(Z) = \frac{1}{N} \sum_{i=1}^N H^*(Z_i). \quad (4)$$

The use of HMMs is further motivated by the multi-channel EEG analysis, since EEG data are often correlated time series from multiple electrodes on the scalp: HMMs can manage multidimensional signals by applying multivariate probability density functions on such signals. Hence, they are appropriate for modeling the inter-relations between EEG time series recorded from multiple electrodes. In this case, for each subject, an HMM is trained on a set of D EEG time series recorded from D electrodes.

At time t , a hidden state emits a D -dimensional observation vector. By applying the Viterbi algorithm, N epochs are generated for each EEG signal and the entropy $H^*(Z_i)$ of each epoch \mathbf{S}_i is computed considering the probability density estimated by the HMM on the observations of the D epochs \mathbf{S}_i (see Fig. 2).

Although all N epochs are matched between EEG channels, the model does not constrain these epochs to be of equal length for all channels: this is a valuable feature of the model, because stationary epochs

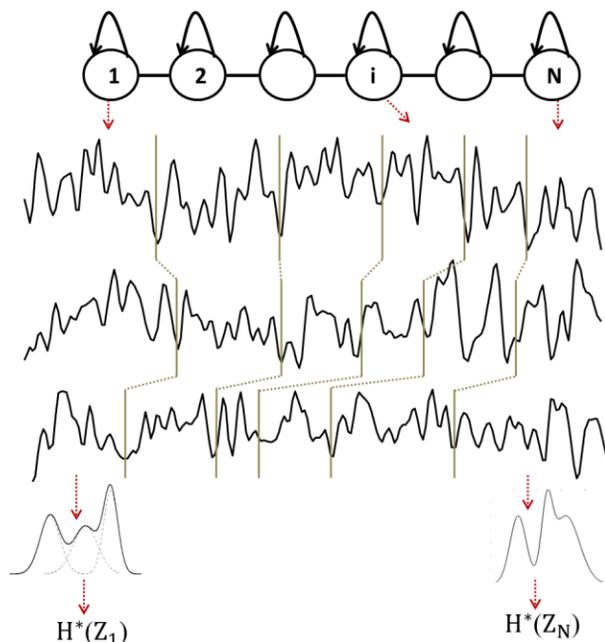


Fig. 2. Illustration of multi-channel ($D = 3$) EEG signal modeling with HMM. The model is feedforward, therefore hidden states are visited only once — epochs and states are matched.

of EEG data do not necessarily have the same duration for all channels. The HMM is feedforward, consequently epochs and hidden states are matched (hidden states are visited only once). Finally, by averaging the entropy over all the N epochs, an epoch-based entropy value associated to the multi-channel EEG of the considered subject is computed.

2.4. Illustration of epoch-based entropy mechanisms

In this section, we illustrate the mechanisms of epoch-based entropy for measuring the complexity of multivariate piecewise stationary EEG signals. To this end, the epoch-based entropy on three EEG signals (shown in Fig. 3) is computed, considering them first separately in a univariate analysis, and subsequently as pairs of signals for a multivariate analysis. Each EEG signal is taken from a healthy subject, sampled at 128 Hz.

Table 1 shows that the three signals exhibit different epoch-based entropy values, hence have different complexities. These results are in agreement with the visual impression about signal complexity: higher entropy is associated with high “irregular” or “complex” signals.

Considering pairs of EEG signals, epoch-based entropy detects the statistical dependencies between channels, as shown in Table 2. The proposed entropy reflects both intra-channel complexity (complexity over time) and the inter-channel complexity (spatial

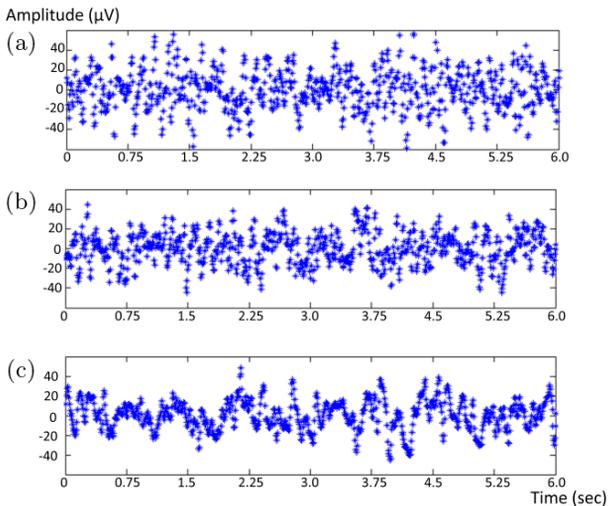


Fig. 3. Examples of three EEG signals of different complexities.

Table 1. Epoch-based entropy values computed on the three EEG signals (a, b, and c) when considered separately.

Signals	a	b	c
Epoch-based entropy	12.1	10.6	8.8

Table 2. Epoch-based entropy computed on pairs of the three signals. Pairs of identical signals ([a;a], [b;b], and [c;c]) have lower complexities than pairs of different signals ([a;b], [a;c], and [b;c]).

Signals	a	b	c
a	10.8	16.1	15.9
b	/	10.5	13.5
c	/	/	8.6

complexity, or heterogeneity between the overall signals).

For instance, when computing epoch-based entropy on identical signals ([a;a], [b;b], and [c;c]), the entropy takes a lower value than when the signal is considered alone (Tables 1 and 2). This is due to the fact that there is no inter-channel difference. Consequently the combined distributions are more regular, leading to a decrease of the entropy value of the multivariate signal. However, this combined entropy is still nonzero since entropy takes also into account intra-channel disorder.

The algorithm estimates complexity using the probability distributions (see Fig. 2). Regularity between channels means in our case that the channels have similar probability distributions. Without noise, the distributions are specific to each channel. With added Gaussian white noise of high amplitude, the channels converge towards more similar (and Gaussian) distributions.

This also holds true when entropy is computed on two signals of different complexities. As an example, when computing entropy on the least complex signal (c) with a signal of higher complexity (a or b), entropy increases as well as the difference increases between signals (inter-channels) and also over time for each signal (intra-channel). When computing entropy on the most complex signal (a) combined

with a signal of lower complexity (b or c), entropy also increases.

To summarize, the statistical estimation of entropy with HMM allows one to quantify the complexity of multivariate EEG signals at two levels simultaneously: at the time level, complexity corresponds to entropy in piecewise stationary epochs of EEG signals over time; at the spatial level, complexity corresponds to the heterogeneity of piecewise stationary epochs between multi-channel EEG signals.

3. Methods

3.1. Databases

We used three datasets, containing EEG recordings of healthy subjects and AD patients at rest and with closed-eyes conditions.^{53,66} These datasets differ in terms of EEG setups and recording conditions. They were produced within the BIOPATTERN European project (University of Plymouth, Plymouth, UK). This project is aimed at developing a generalized, clinical EEG model for AD diagnosis and symptom quantification by analyzing subject bio-profiles and clinical measures recorded from hospitals in different EU countries.⁶⁷ The local institutional ethics committees approved this research and informed consent was obtained from all subjects and caregivers prior to recording and experimentation.

3.1.1. Dataset A

This dataset originates from Derriford Hospital, Plymouth, UK. It contains EEG data of 24 healthy subjects (aged 69.4 ± 11.5 years) and 17 patients diagnosed with mild form of AD (aged 77.6 ± 10 years). It was shown that no significant effect stemmed from discrepancy of age between the two groups.⁶⁷ Patients underwent a battery of neuroimaging and cognitive tests. EEG signals were recorded during 4 min at a sampling frequency of 256 Hz, later down-sampled to 128 Hz. 21 electrodes were placed on the scalp according to the Maudsley System (Fp₂, F₇, F₃, Fz, F₄, F₈, A₁, T₃, C₃, Cz, C₄, T₄, A₂, T₅, P₃, Pz, P₄, T₆, O₁, O₂, Fp₁).

3.1.2. Dataset B

This dataset contains EEG data of five age-matched healthy subjects (aged 76.6 ± 5.6 years) and five AD patients (aged 78.8 ± 2.4 years). Patients

were diagnosed with early stage, mild form AD according to the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA),⁶⁸ and the Diagnostic and statistical manual of mental disorder, 4th edition (DSM IV) criteria. They underwent general medical, neurological, and psychiatric tests. EEG signals were recorded during 1 min at a sampling frequency of 128 Hz using 21 electrodes (Fp₁, Fp₂, F₇, F₃, Fz, F₄, F₈, A₁, T₃, C₃, Cz, C₄, T₄, A₂, T₅, P₃, Pz, P₄, T₆, O₁, O₂) placed according to the 10–20 international system at the University of Malta.

3.1.3. Dataset C

This dataset is obtained from the Ecological University of Bucharest. It consists of three healthy subjects (aged 73.5 ± 2.2 years) and eight age-matched AD patients (aged 75 ± 3.4 years). Patients were diagnosed with a mild form of AD using psychometric tests, neuroimaging, and clinical examinations. EEG signals were recorded during 10 to 20 min at a sampling frequency of 512 Hz using 22 electrodes (Fp₁, Fp₂, F₇, F₃, Fz, F₄, F₈, A₁, T₃, C₃, Cz, C₄, T₄, A₂, T₅, P₃, Pz, P₄, T₆, O₁, O₂, Oz) disposed according to the International Federation of Clinical Neurophysiology standards for digital recording of clinical EEG. In this work, EEG signals of this dataset were down-sampled to 128 Hz similarly to datasets A and B.

3.2. Study design

Classification was performed by single-feature Linear Discriminant Analysis (LDA), using each complexity measure as input feature. The performance of each classifier was assessed following a consistent protocol: a development subset containing EEG data of dataset B (five AD patients and five healthy subjects) and dataset C (eight AD patients and three healthy subjects) was used. On these 21 subjects, the three complexity measures were computed, and each of them was used as feature of an LDA classifier; the performance of each classifier was estimated by a Leave-One-Out procedure (LOO, see for instance Ref. 71), and the threshold that gave the best correct classification rate was selected. The optimal values of the hyperparameters needed for a reliable estimation of the complexity measures (e.g.

number of Gaussians and epochs for the epoch-based entropy) were also selected based on the correct classification rate on the development dataset. In other words, the development dataset is used both as a training set and a validation set (using LOO cross-validation). The robustness to hyperparameter variations is detailed in Sec. 4.3.

Finally, in the test phase, the three measures were computed for each subject of the dataset A, using the optimal hyperparameters found on the development dataset. Then, by considering each complexity measure as a single input feature of the LDA classifier, the correct classification rates were estimated on the test dataset based on the optimal thresholds found on the development dataset for each complexity measure. Consequently, the dataset A was used as an independent test dataset — which is indeed the case, as the acquisition sites and experimental conditions for collecting dataset A differ from datasets B and C.

Brain regions were defined arbitrarily, using sets of channels located in regions susceptible to be sensitive for changes due to AD. We defined five regions of interest: frontal (F_{p1} , F_{p2}), occipital (O_1 , O_2), parietal (T_6 , P_4 , P_z , P_3 , T_5), temporal (T_6 , T_4 , T_5 , T_3 , F_7 , F_8), and parieto-temporal (F_7 , F_8 , T_3 , T_4 , T_5 , P_3 , P_z , P_4 , and T_6). We did not optimize these regions of interest. The interested reader can report to Sec. 4.2.3 for simulations of stability to changes in these channel sets. The concatenation order applied for Shannon’s entropy and D_2 was decided arbitrarily according to the electrode placement order. There was no significant effect when changing this order.

4. Experimental Results

4.1. *Epoch-based entropy, Shannon’s entropy and correlation dimension for AD screening*

In this section, we investigate the effectiveness of the proposed epoch-based entropy measure for the discrimination of AD patients from healthy subjects on the test dataset. We compare it to both Shannon’s entropy and correlation dimension in terms of classification accuracy.

Table 3 presents the correct classification rate and area under the curve (AUC) values, per brain region, with the three complexity measures on the

test dataset. To obtain an epoch-based entropy value per person and per brain region, the HMMs were trained on a set of EEG signals collected by electrodes of the considered brain area, as explained above in Sec. 2.3. For comparison purposes, Shannon’s entropy and correlation dimension were computed per person and per brain region by concatenating EEG time series of the considered brain region.

Results clearly show that Shannon’s entropy is the worst complexity measure in terms of discrimination between AD patients and healthy subjects, with an accuracy of 41.4% for all brain regions. This rate corresponds to the *a priori* probability of the class of AD patients (17 AD patients among the 41 subjects of the test dataset): the classifier based on Shannon’s entropy classifies all subjects of the test dataset as AD patients (specificity = 0%, sensitivity = 100%). By contrast, the best discrimination is provided by epoch-based entropy when parietal and temporal regions are considered: a correct classification rate of almost 83% is obtained (specificity = 83.33%, sensitivity = 82.35%).

For a refined comparative analysis, Fig. 4 shows the box plots of complexity values obtained on healthy subjects and AD patients for each complexity measure, on both development and test datasets. AD patients have lower values of Shannon’s entropy and correlation dimension than healthy subjects. However, epoch-based entropy values of healthy subjects are lower than those of AD patients, contrary to two other complexity measures.

This inverse behavior stems from the specific properties of the epoch-based entropy measure: AD induces a reduction in complexity of EEG signals, but also an increase of irregularity between EEG channels. As epoch-based entropy quantifies both effects simultaneously (disorder over time and spatial disorder), it behaves differently from other two complexity measures.

Moreover, it appears clearly that the box plots of epoch-based entropy of healthy subjects are significantly different from those of AD patients (Mann–Whitney $p = 5.5 \times 10^{-7}$ computed on all datasets), slightly less significant for correlation dimension ($p = 5.7 \times 10^{-5}$) and not significant for Shannon’s entropy ($p = 0.064$). This accounts for the fact that the discrimination efficiency of epoch-based entropy is higher than that of the other two measures.

Table 3. AUC values and correct classification rates (Acc, in %) on the test dataset with each complexity measure per brain region.

Brain regions	Epoch-based entropy		Correlation dimension		Shannon’s entropy	
	AUC	Acc	AUC	Acc	AUC	Acc
Frontal	0.75	63.4	0.82	70.7	0.68	41.4
Occipital	0.83	58.5	0.84	75.6	0.73	41.4
Parietal	0.89	73.2	0.85	58.5	0.75	41.4
Temporal	0.89	63.4	0.74	60.9	0.72	41.4
Parieto temporal	0.88	82.9	0.82	75.6	0.74	41.4
All	0.90	82.9	0.80	78.1	0.74	41.4

It has been shown in Ref. 69 that complexity measures and spectral measures can be strongly correlated. However, since the proposed entropy measure uses local information (from short time segments), we expected there would be no correlation. In order to confirm this hypothesis, we estimated the correlations of relative power in the theta and alpha ranges with the proposed entropy measure. Those correlations were nonsignificant for both estimates (theta: $R^2 = 0.06$, F-score = 2.30, $p = 0.14$; alpha: $R^2 = 0.07$, F-score = 2.78, $p = 0.10$).

In addition, the box plots of Shannon’s entropy and correlation dimension values on healthy subjects (or AD patients) are very different in the development and test datasets. This is even more conspicuous for AD patients: the box plots of epoch-based entropy on both test and development datasets are quite similar on AD patients (Fig. 4(a)), although the datasets originate from different sites and were obtained with different electrode placements, as explained in Sec. 3. Therefore, choosing the decision threshold on the development dataset does not affect significantly the performance on the test dataset with epoch-based entropy.

For Shannon’s entropy (Fig. 4(c)), the distribution of healthy subjects and AD patients are completely different on test and development datasets. Therefore, the selected optimal threshold on the development dataset is not adapted to the test dataset, and thus leads to bad classification rate on the test dataset, as observed above.

It has thus far been demonstrated that epoch-based entropy is more reliable for AD screening than the other two complexity measures. In the next section, we investigate the robustness of the proposed measure with respect to variations in experimental condition that may occur during EEG acquisition.

4.2. Effect of noise and sampling rate on complexity measures

Several technical issues during EEG acquisition may affect the analysis of EEG signal, which is known to have a low Signal-to-Noise Ratio (SNR). Actually, EEG recordings are strongly affected by different sources of noise and disturbances, be they physiological (e.g. neuromuscular noise) or electronic. Inter-subject variability is partly due to differences of SNR. Therefore, measures that would not be robust to variations of SNR are prone to inaccuracies depending on experimental conditions. Sampling frequency is another potential technical issue; given that EEG signal is digitized, low sampling frequency may lead to information loss. Clinical EEG databases are digitized with different sampling rates, which can constitute a limit for the generalization of a screening method. In the following, we study the effect of white noise and sampling frequency on the three complexity measures when applied to EEG data.

4.2.1. Effects of added white noise

In this section, we analyze the influence of white noise on the complexity estimations obtained by epoch-based entropy, correlation dimension, and Shannon’s entropy, on the signals of the test dataset, independently from sampling rate.

The three complexity measures are first computed on the original signals of the test dataset. They are subsequently computed on six datasets containing noisy data generated by adding six levels of Gaussian white pseudo-random noise to the original data, thereby generating six signals with different SNR values (20dB, 15dB, 10dB, 5dB, 1dB, -10dB). As white noise is generated randomly, 100 different realizations of noisy data are produced for each SNR

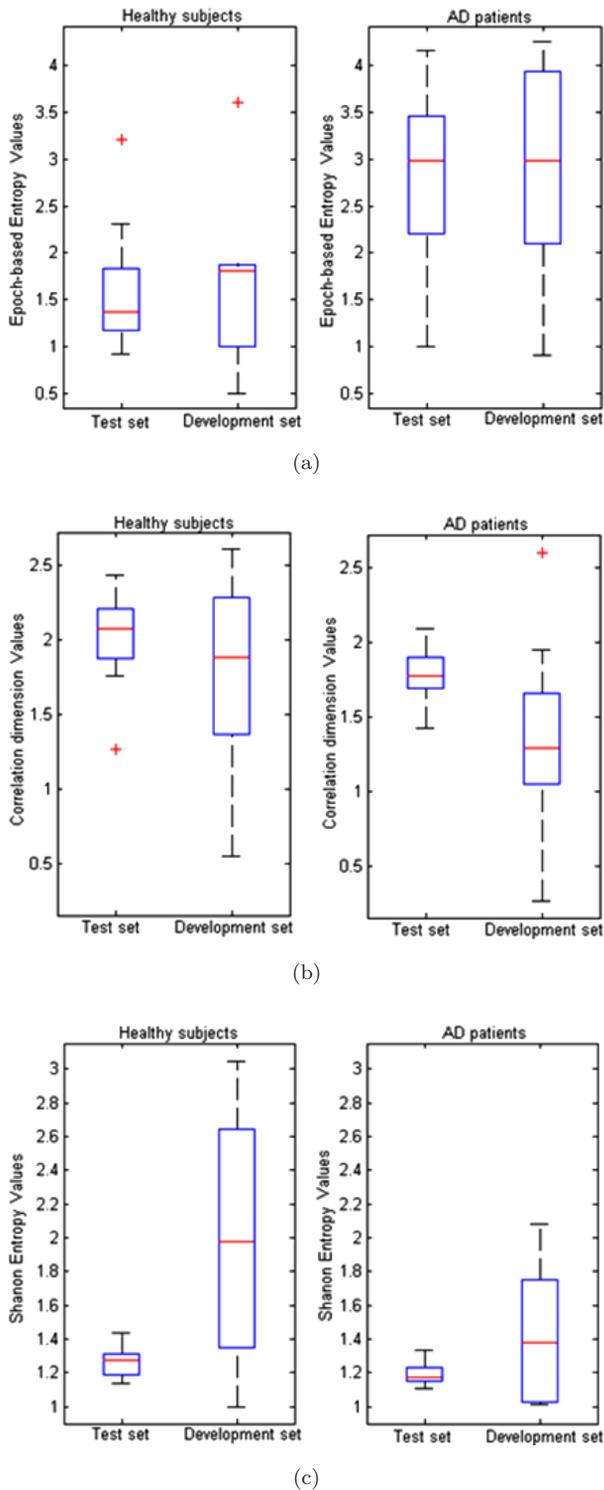


Fig. 4. Box plots of (a) epoch-based entropy, (b) correlation dimension, and (c) Shannon's entropy values measured on parieto-temporal region. Values for healthy subjects (left) and AD patients (right) are shown for both the test and the development datasets.

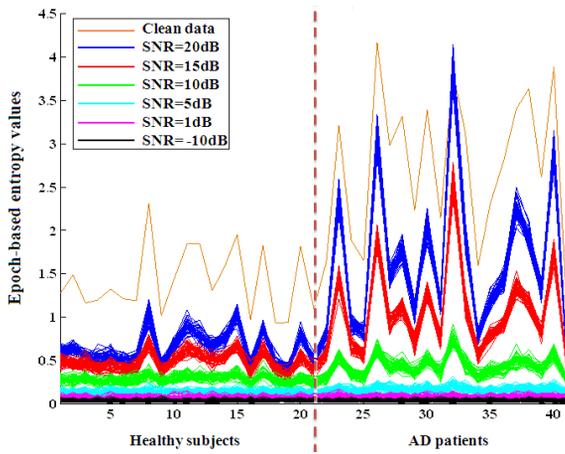
value in order to obtain a reliable estimate of classification accuracy from these data. The classification is performed from these data using the optimal threshold obtained on the original development dataset (datasets B and C, without added noise). Consequently, the obtained results are representative of the robustness of the classifier to variations of noise when classifying new independent examples.

Figure 5 displays the values of the three complexity measures obtained on healthy subjects and AD patients of the test dataset considering the original clean data and the six levels of noisy data. It is first observed that increased noise in the data causes an increase in correlation dimension and Shannon's entropy values, but a decrease in epoch-based entropy values. This inverse behavior of epoch-based entropy is due to its properties, as explained in Sec. 2.3 and observed in Sec. 4.1: increased noise in the multidimensional EEG data would result in more regularity between channels since the original data will be hidden by the added noise. This phenomenon leads to a decrease of the epoch-based entropy.

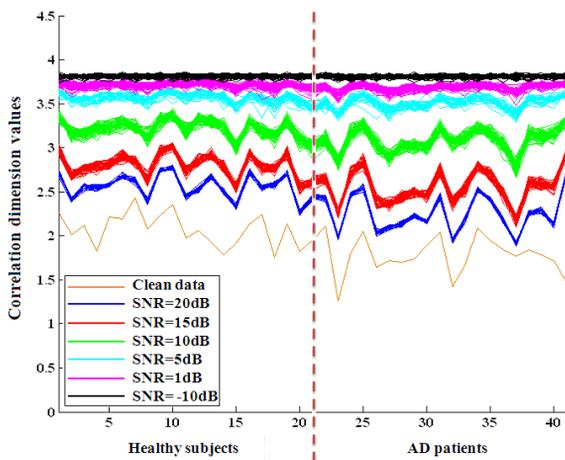
Figure 5 also shows that the gap in values between the seven datasets is more important with Shannon's entropy and correlation dimension. In addition, the difference between healthy subjects and AD patients is larger with epoch-based entropy (Fig. 5(a)), even in the presence of noise with $\text{SNR} = 10 \text{ dB}$. These first results indicate that epoch-based entropy is less sensitive to noise than the other two complexity measures.

Figures 6(a)–6(c) show the seven ROC curves obtained for classification by LDA, using respectively epoch-based entropy, correlation dimension, and Shannon's entropy as a single feature, on the original signals of the test dataset and on the six levels of noisy signals obtained from the test dataset.

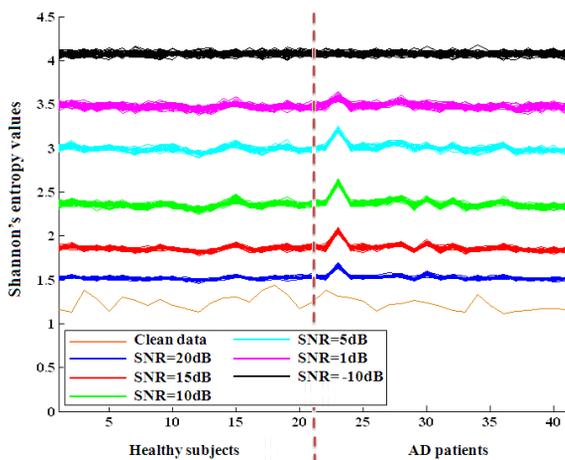
Note that ROC curves obtained on noisy signals are smoother than those obtained on the original signals due to averaging over 100 noise realizations. Specificity and sensitivity correspond to good classification of healthy subjects and AD patients, respectively. Figure 7 displays the evolution of the AUC and the correct classification rates obtained for each complexity measure on all signals. It is clear that when adding noise to the original signals, classification performance decreases for all measures, but to different extents.



(a)

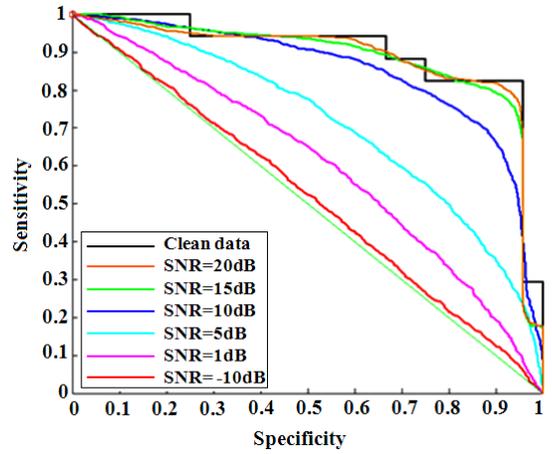


(b)

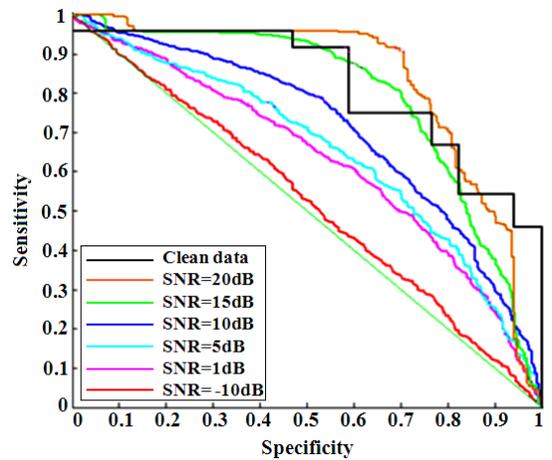


(c)

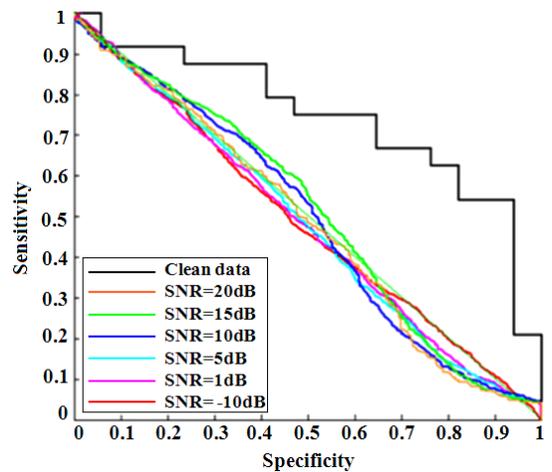
Fig. 5. Values of (a) epoch-based entropy, (b) correlation dimension, and (c) Shannon's entropy for the 41 subjects of the test dataset, on the clean data and the six noisy data with different SNR.



(a)

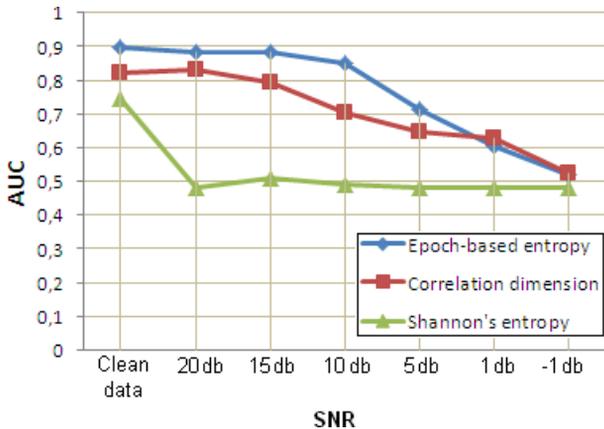


(b)

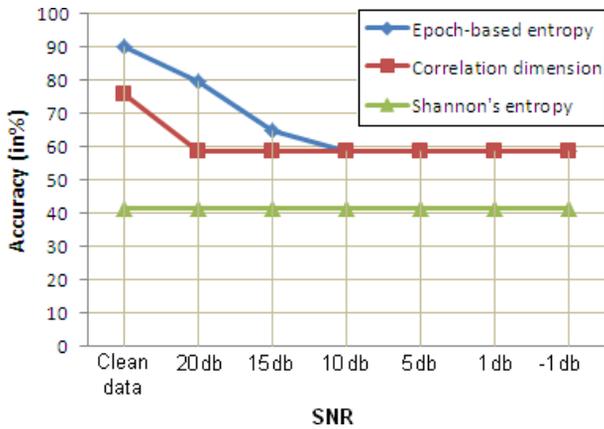


(c)

Fig. 6. ROC curves for classification, using the classifier designed on the development dataset, of the original signals of the test dataset and on the same signals with added noise, from: (a) epoch-based entropy, (b) correlation dimension, and (c) Shannon's entropy.



(a)



(b)

Fig. 7. AUC (top) and correct classification rate (bottom) for the three complexity measures on the clean data and the six noisy data with different SNR.

First consider the classification results obtained with epoch-based entropy and correlation dimension. In both cases, the performance degrades gracefully as the amount of noise in data increases, albeit more markedly for correlation dimension than for epoch-based entropy. Figures 6(a), 6(b) and 7(a) show that for $\text{SNR} \geq 10$ dB, a smooth variation of AUC values appears with epoch-based entropy. Similarly, in terms of accuracy, which is dependent on the threshold estimated on the original development dataset (clean data), Fig. 7(b) shows that performance degrades rapidly with correlation dimension: the correct classification rate drops to 58% with D_2 when $\text{SNR} = 20$ dB, while, with epoch-based entropy, the same classification rate is achieved when $\text{SNR} = 10$ dB.

Note that for $\text{SNR} \leq 10$ dB, correlation dimension and epoch-based entropy both reach an accuracy of 58.5%, which corresponds to an arbitrary classification of all subjects as healthy (specificity = 0%, sensitivity = 100%): when adding white noise to data, the EEG signals of AD patients become “similar” to those of healthy subjects with respect to epoch-based entropy or correlation dimension.

Finally, for Shannon’s entropy, we obtain a correct classification rate of 41.46% on the original signals as well as for all levels of added noise. This is due to the fact that the threshold value estimated on the development dataset is not efficient on the test dataset as explained in Sec. 4.1. The influence of noise on that measure is apparent on the AUC values: by contrast to other two measures, with Shannon’s entropy, performance drops significantly in terms of AUC on data affected by low noise ($\text{SNR} = 20$ dB).

4.2.2. Sampling frequency effects

This section studies the influence of sampling frequency on the three complexity measures. To this end, the original signals of the test dataset, sampled at 128 Hz, were down-sampled at different frequencies, generated by multiplying the original sampling frequency by seven different multiplicative factors: $\mu = 0.9, 0.8, 0.7, 0.6, 0.5, 0.4,$ and 0.3 (an interpolation procedure is applied when required). We used an anti-aliasing low-pass FIR filter with a Kaiser window in cases when interpolation was required. Results of classification performance as a function of sampling frequency, in terms of AUC and correct classification rate, are shown in Figs. 8(a) and 8(b), respectively.

At each sampling frequency, the optimal hyperparameter configuration for complexity measures and the optimal decision threshold that leads to the best accuracy are estimated on the development dataset (robustness to hyperparameter changes in the computation of D_2 and epoch-based entropy is investigated in the next section). The obtained configurations and thresholds for each sampling frequency are subsequently used to estimate the accuracy on the test dataset.

Figures 8(a) and 8(b) show that the evolution of classification performance as a function of sampling frequency are essentially similar, for classifiers based on epoch-based entropy and correlation dimension.

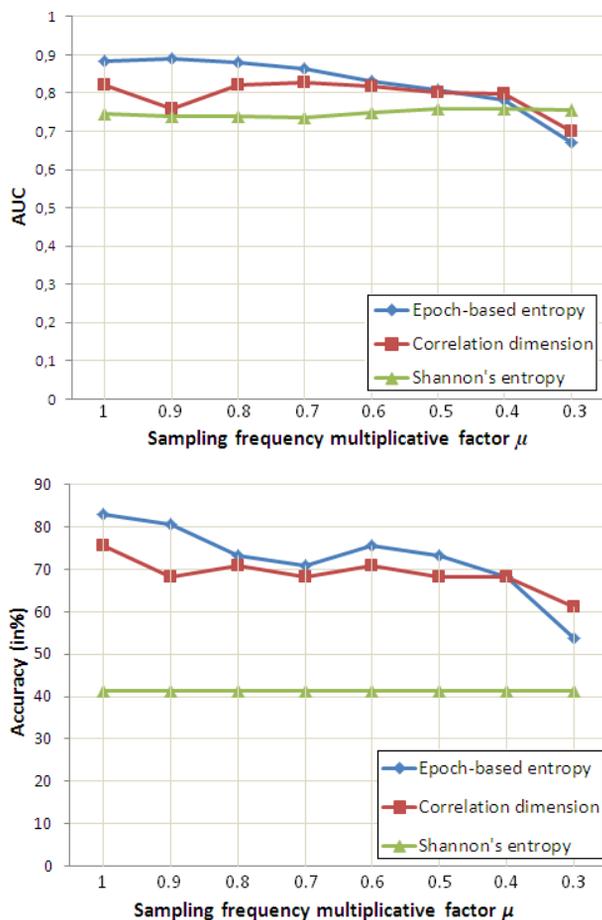


Fig. 8. AUC (top) and Accuracy (bottom) as a function of sampling rate with the three complexity measures.

Shannon's entropy provides poor results irrespective of frequency.

All the obtained results demonstrated that AD screening based on epoch-based entropy is less sensitive to noise and sampling rate than AD diagnosis based on the Shannon entropy and correlation dimension. For a complete comparison between the two best complexity measures, the effects on classifier performance of changing hyperparameters used in their calculation is investigated in the next section.

4.2.3. Regional stability: channel grouping

Epoch-based entropy was computed on *a priori* chosen sets of channels. We evaluate here the robustness of the measure to this channel selection.

The largest set of channels was defined in the parieto-temporal region, where we had selected nine channels: F₇, F₈, T₃, T₄, T₅, P₃, P_z, P₄, and T₆. We

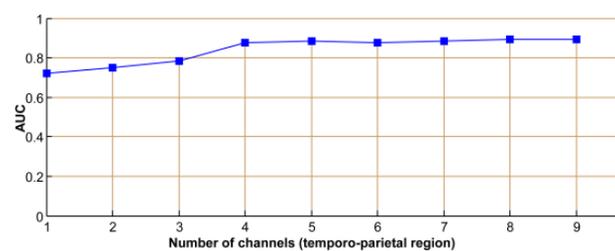


Fig. 9. AUC depending on the number of channels in the parieto-temporal region.

estimated the variations of the AUC when removing iteratively one electrode from this list. The classification performance is globally stable, and drops when less than four channels are used (Fig. 9). The effect of spatial correlations on the entropy values is visible here, as the cross-channel information improves the classification accuracy. Similar effects are observed in other regions.

4.2.4. Temporal stability

We evaluated the stability of epoch-based entropy using a test-retest procedure. We divided the signals of all subjects in the test dataset ($N = 41$), and computed the measure on the first half of the signal, and compared it with the measure applied on the second half of the signal (see Fig. 10). The resulting Pearson correlation score was highly significant ($r^2 = 0.84$, $F = 204.75$, $p < 10^{-6}$), demonstrating the measure stability.

4.3. Effect of hyper-parameters variation

Hyperparameters are involved in the computation of correlation dimension and epoch-based entropy. We evaluated the stability of D_2 and epoch-based entropy to changes in these hyperparameters. As indicated in Sec. 4.1, the optimal hyperparameter configuration is selected by looking for that giving the highest accuracy on the development dataset. The sensitivity of classifiers based on correlation dimension and epoch-based entropy to variations of the hyperparameters is investigated in this section.

For computing epoch-based entropy, the number of epochs N and the number of Gaussians M per epoch are required (Sec. 2.3). Figure 11 displays the AUC values computed with epoch-based entropy as a function of the number of epochs and the number of Gaussians. As expected, a sharp decrease of

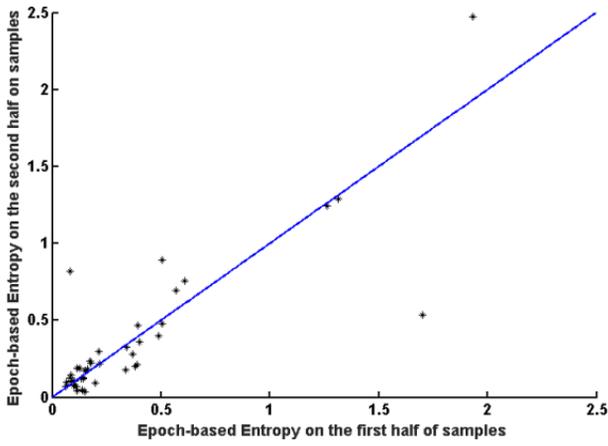


Fig. 10. Scatterplot on epoch-based entropy values computed on the first half of the signals versus epoch-based entropy measured on the second half.

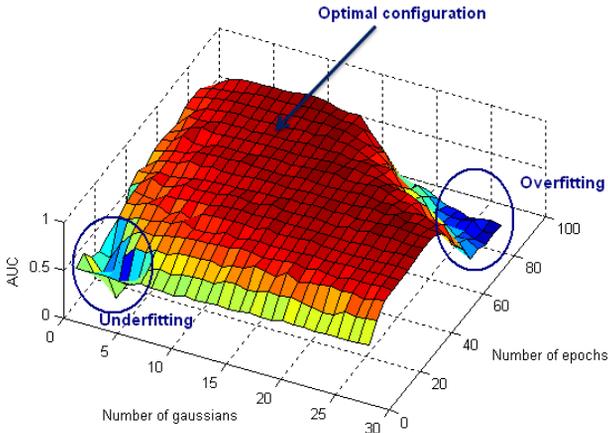
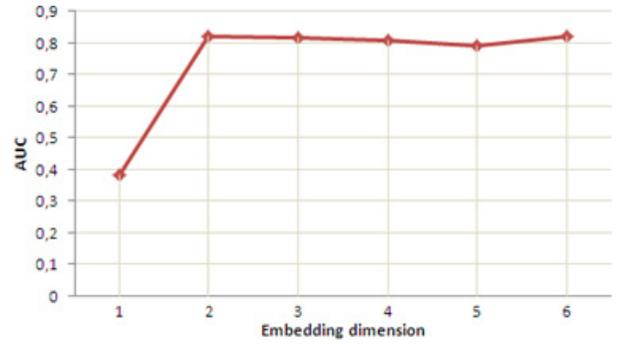


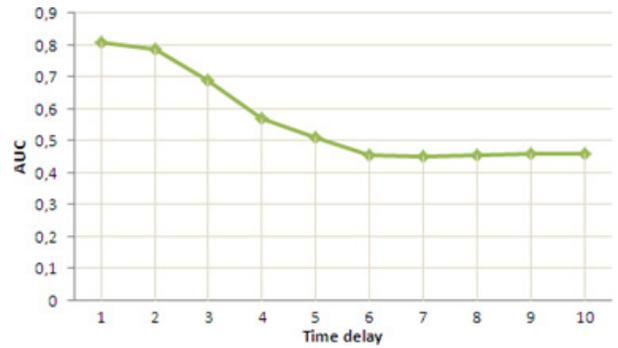
Fig. 11. AUC values on the test dataset for different values of number of Gaussians and epochs.

AUC occurs for small numbers of Gaussians and of epochs (leading to underfitting), and for high numbers of Gaussians and epochs (leading to overfitting). The presence of a large plateau around the optimal value shows that the performance of a classifier based on epoch-based entropy is reasonably insensitive to change of the hyperparameters involved in the definition of epoch-based entropy. The hyperparameters estimated on the development set are far from overfitting and underfitting conditions.

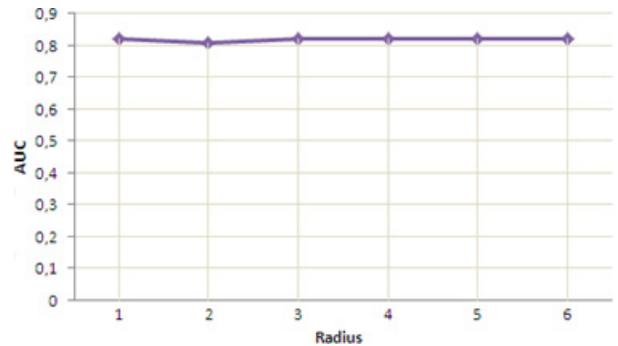
Computing D_2 requires the selection of hyperparameters m , r , and τ . These are commonly referred to as “embedding dimension”, “radius”, and “time delay”, respectively (Sec. 2.1). Figure 12 shows the AUC values obtained with correlation dimension as a function of each of the three parameters involved



(a)



(b)



(c)

Fig. 12. AUC values on the test dataset for different values of (a) embedding dimension, (b) time delay, and (c) radius.

in computing D_2 . Starting from the optimal configuration that was found on the development dataset (embedding dimension $m = 4$, time delay $\tau = 1$, radius $r = 2$), these hyperparameters are varied separately. In the range of hyperparameter values investigated, results show that the AUC is essentially independent of r , and is sensitive to m and τ .

For a better insight into the influence of the embedding dimension m and the time delay τ , Fig. 13 shows the AUC as function of these two parameters,

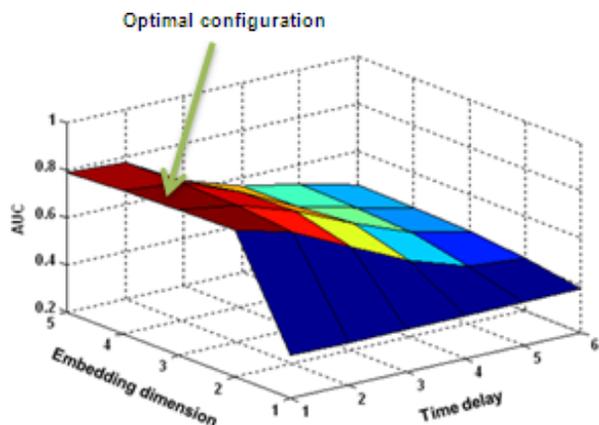


Fig. 13. AUC values on the test dataset for different values of embedding dimension and time delay.

with $r = 2$. A significant performance degradation is observed when $m = 1$, for all values of τ : too small values of m lead to errors because points that are really far apart in time on the EEG signal end up as close neighbors in the phase space. For $m > 1$, the AUC decreases significantly with τ .

5. Discussion and Conclusion

The purpose of this study was to investigate the potential application of a new entropy-based complexity measure to the discrimination of AD patients from healthy subjects on the basis of multi-channel EEG signals. Termed “epoch-based entropy”, this new complexity measure is computed on piecewise stationary epochs using a HMM, which performs local density estimation at the epoch level. This method was tested on multi-channel EEG data in order to illustrate the ability of the method to take into account inter-channel relations.

The originality of the new measure lies on the fact that it estimates the complexity of EEG signals locally over time (as done by classical complexity measures), and also spatially by estimating the inter-channels complexity. These two complexity measures are merged into a single figure. Many investigations on AD screening were based on the reduction in complexity of the EEG signal in AD patients; other studies were based on the reduction of homogeneity between EEG channels in AD patients.^{36,37} However, these two factors were usually measured separately.

By comparing epoch-based entropy measure to two alternative complexity measures, correlation dimension and Shannon’s entropy, it was shown that

the proposed measure is a more reliable feature for AD screening than the other two, on our experimental data. A significantly higher classification rate was obtained with epoch-based entropy than with other two measures. Many studies pointed out the fact that the amount of data required for reliable assessments of D_2 is beyond experimental reach for physiological data,^{70,71} which may explain in part our improved results.

Earlier studies reported classification results in the 75–90% range.^{10,13,14,16–18,23–25,27–29,37} However, note that most studies do not use proper cross-validation designs (either not using an independent test set, or optimizing hyperparameters on the whole database). Without a proper independent test set, accuracies are usually over-estimated, as they do not address properly the risks of overfitting. In this work, an accuracy of 83% is reached with epoch-based entropy on an independent test dataset when parietal and temporal regions are considered. This finding is consistent with clinical knowledge^{3,4,15,72}: parietal and temporal regions are the first affected regions in the early stage of AD. Our measure could therefore be used for the screening of patients. A next step would be to investigate databases including other neurodegenerative pathologies (such as Lewy Body dementia, vascular dementia, FTD, etc.), in order to estimate the potential of this method as a diagnostic tool.

The distinct separation between AD patients and healthy subjects with epoch-based entropy provides further evidence that EEG analysis can be used to detect cognitive abnormalities due to AD. Moreover, our observations are consistent with the fact that AD is associated to the loss of neurons and to impairments in the coordination of neural activity.

Shannon’s entropy was the worst complexity measure in terms of discrimination between AD patients and healthy subjects. The classification achieved is not better than a random classification, and statistical analyses do not show significant differences between the AD patients and the control subjects (p -value > 0.05). The poor classification performances can be explained in part by the mismatch between Shannon’s entropy values estimated on the development dataset and those estimated on the test set. In this case, the optimal decision threshold selected on the development set becomes not appropriate for a deployment on the test

dataset and thus leads to poor classification results on test data. This result highlights the fact that Shannon’s entropy is very sensitive to EEG recording conditions.

Simulations were conducted to assess the robustness of the proposed measure to disturbances in the experimental conditions by artificially varying the SNR or the sampling rate in the test set. Here again, epoch-based entropy was more stable and less sensitive to low SNR (as low as 10 dB) or decreased sampling rates (as low as 51.2 Hz) than the two other measures. This explains its higher generalization performance on the test set. It was further shown that classifiers based on epoch-based entropy are reasonably insensitive to variations of the hyperparameters involved in its computation. This is in contrast to correlation dimension-based classifiers.

In conclusion, our study provides a novel method for analyzing the dynamics of neural activity in patients with AD. Further experiments should be conducted on additional subjects to assess the generalization of our method to clinical usage. Furthermore, in this work, entropy was computed on EEG time series locally at the epoch level, and then averaged over all the epochs for the sake of parsimony. It might be interesting to keep the entropy values per epoch; thus EEG signal would be associated to a sequence of entropy values. This would characterize more finely how the EEG signal fluctuates over time. Moreover, the proposed entropy measure was developed for the analysis of EEG data in the time domain; it could be extended to characterize the underlying brain dynamics in the frequency domain by filtering the EEG signal into frequency ranges prior to the evaluation of the entropy measures.

Acknowledgments

We would like to thank C. Latchoumane and J. Jeong for providing us with the three datasets used in the present study.

References

1. C. P. Ferri, M. Prince, C. Brayne, H. Brodaty, L. Fratiglioni, M. Ganguli, K. Hall, K. Hasegawa, H. Hendrie, Y. Huang, A. Jorm, C. Mathers, P. R. Menezes, E. Rimmer and M. Sczufca, Alzheimer’s disease international, Global prevalence of dementia: A delphi consensus study, *Lancet* **366**(17) (2005) 2012–2017.

2. L. Lili, D. Ruau, R. Chen, S. Weber and A. J. Butte, Systematic identification of risk factors for Alzheimer’s disease through shared genetic architecture and electronic medical records, in *Pac. Symp. Biocomput.* (2013), pp. 224–235.
3. H. Braak and E. Braak, Neuropathological staging of Alzheimer-related changes, *Acta Neuropathol.* **82**(4) (1991) 239–259.
4. J. Pantel, P. Schönknecht, M. Essig and J. Schröder, Distribution of cerebral atrophy assessed by MRI reflects pattern of neuropsychological deficits in Alzheimer’s dementia, *Neurosci. Lett.* **361** (2004) 17–20.
5. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer’s Disease and models of computation: Imaging, classification, and neural models, *J. Alzheimer’s Dis.* **7** (2005) 187–199.
6. Z. Sankari and H. Adeli, Probabilistic neural networks for eeg-based diagnosis of alzheimer’s disease using conventional and wavelet coherence, *J. Neurosci. Methods* **197** (2011) 165–170.
7. M. Hashimoto, H. Kazui, K. Matsumoto, Y. Nakano, M. Yasuda and E. Mori, Does donepezil treatment slow the progression of hippocampal a trophy in patients with Alzheimer’s disease? *Am. J. Psychiatry* **162**(4) (2005) 676–682.
8. C. Pietrzik and C. Behl, Concepts for the treatment of Alzheimer’s disease: Molecular mechanisms and clinical application, *Int. J. Exp. Pathol.* **86**(3) (2005) 173–185.
9. J. Dauwels, F. B. Vialatte and A. Cichocki, Diagnosis of Alzheimer’s disease from EEG signals: Where are we standing? *Curr. Alzheimer Res.* **7**(6) (2010) 487–505.
10. W. Woon, A. Cichocki, F. B. Vialatte and T. Musha, Techniques for early detection of Alzheimer’s disease using spontaneous EEG recordings, *Physiol. Meas.* **28** (2007) 335–347.
11. S. Hulbert and H. Adeli, EEG/MEG and Imaging based diagnosis of Alzheimer’s disease, *Rev. Neurosci.* **24**(6) (2013) 563–576.
12. H. Adeli and S. Ghosh-Dastidar, *Automated Eeg-Based Diagnosis of Neurological Disorders — Inventing The Future of Neurology* (T. & F. CRC Press, Ed., Boca Raton, Florida, 2010).
13. T. Schreiber and H. Kantz, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
14. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A spatio-temporal wavelet-chaos methodology for eeg-based diagnosis of Alzheimer’s disease, *Neurosci. Lett.* **444**(2) (2008) 190–194.
15. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer’s disease: Models of computation and analysis of EEGs, *Clin. EEG Neurosci.* **36**(3) (2005) 131–140.
16. T. A. Holt, *Complexity for Clinicians* (Radcliffe Medical Press LTD, London, UK, 2004).

17. J. Jeong, EEG dynamics in patients with Alzheimer's disease, *Clin. Neurophysiol.* **115**(7) (2004) 1490–1505.
18. A. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-Chao Methodology for EEG-based diagnosis of Alzheimer's disease, *Alzheimer Dis. Assoc. Disorders* **25**(1) (2011) 85–92.
19. P. Grassberger, Generalized dimensions of strange attractors, *Phys. Lett. A* **97**(6) (1983) 227–230.
20. P. Grassberger and I. Procaccia, Measuring the strengeeness of strange attractors, *Physica D* **9** (1983) 189–208.
21. Y. B. Pesin, Characteristic Lyapunov exponents and smooth ergodic theory, *Russian Math. Surv.* **32**(4) (1977) 55–114.
22. J. Jeong, J. H. Chae, S. Y. Kim and S. H. Han, Non-linear dynamic analysis of the EEG in patients with Alzheimer's disease and vascular dementia, *J. Clin. Neurophysiol.* **18**(1) (2001) 58–67.
23. J. Jeong, S. Y. Kim and S. H. Han, Non-linear dynamical analysis of the EEG in Alzheimer's disease with optimal embedding dimension, *Electroencephalogr. Clin. Neurophysiol.* **106**(3) (1998) 220–228.
24. T. Takahashi, Complexity of spontaneous brain activity in mental disorders, *Prog. Neuropsychopharmacol. Biol. Psychiatry* **45** (2013) 258–266.
25. B. Jelles, R. L. Strijers, C. Hooijer, C. Jonker, C. J. Stam and E. J. Jonkman, Nonlinear EEG analysis in early Alzheimer's disease, *Acta Neurol. Scand.* **100**(6) (1999) 360–368.
26. B. Jelles, J. H. van Birgelen, J. P. Slaets, R. E. Hekster, E. J. Jonkman and C. J. Stam, Decrease of nonlinear structure in the EEG of Alzheimer patients compared to healthy controls, *Clin. Neurophysiol.* **110**(7) (1999) 1159–1167.
27. T. Yagyu, J. Wackermann, M. Shigeta, V. Jelic, T. Kinoshita, K. Kochi, P. Julin, O. Almkvist, L. O. Wahlund, I. Kondakor and D. Lehmann, Global dimensional complexity of multichannel EEG in mild Alzheimer's disease and age-matched cohorts, *Dement Geriatr. Cogn.* **8**(6) (1997) 343–347.
28. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd edn. (John Wiley & Sons, 2006).
29. D. E. Lake, J. S. Richman, M. P. Griffin and J. R. Moorman, Sample entropy analysis of neonatal heart rate variability, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **283** (2002) 789–797.
30. D. Abasolo, R. Hornero, P. Espino, D. Alvarez and J. Poza, Entropy analysis of the EEG background activity in Alzheimer's disease patients, *Physiol. Meas.* **27**(3) (2006) 241–253.
31. T. De Bock, S. Das, M. Mohsin, N. B. Munro, L. M. Hively, Y. Jiang, C. D. Smith, D. R. Wekstein, G. A. Jicha, A. Lawson, J. Lianekhammy, E. Walsh, S. Kiser and C. Black, Early detection of Alzheimer's disease using nonlinear analysis of EEG via Tsallis entropy, *Biomedical Sciences and Engineering Conf.* (2010) 1–4.
32. D. Abasolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez and R. de la Rosa, Analysis of regularity in the EEG background activity of Alzheimer's disease patients with approximate entropy, *Clin. Neurophysiol.* **116**(8) (2005) 1826–1834.
33. S. Pincus, Approximate entropy as a measure of irregularity for psychiatric serial metrics, *Bipolar Disord.* **8**(5) (2006) 430–440.
34. M. Costa, A. L. Goldberger and C. K. Peng, Multiscale entropy analysis of complex physiologic time series, *Phys. Rev. Lett.* **89**(6) (2002) 068102.
35. J. Escudero, D. Abásolo, R. Hornero, P. Espino and M. López, Analysis of electroencephalograms in Alzheimer's diseasepatients with multiscale entropy, *Physiol. Meas.* **27** (2006) 1091–1106.
36. D. Abasolo, R. Hornero, C. Gomez, M. Garcia and M. Lopez, Analysis of EEG background activity in Alzheimer's disease patients with Lempel-Ziv complexity and Central Tendency Measure, *Med. Eng. Phys.* **28** (2006) 315–322.
37. J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath and N. Stergiou, he appropriate use of approximate entropy and sample entropy with short data sets, *Ann. Biomed. Eng.* **41**(2) (2013) 349–365.
38. C. Chatfield, *The Analysis of Time Series Data: An Introduction*, 6th edn. (CRC, Chapman & Hall, Boca Raton, 2003).
39. G. Praetorius and H. M. Bodenstein, Feature extraction from the electroencephalogram by adaptive segmentation, *Proc. IEEE* **65**(5) (1977) 642–652.
40. A. Y. Kaplan, A. A. Fingelkurts, A. A. Fingelkurts, S. V. Borisov and B. S. Darkhovsky, Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges, *Signal Process.* **85**(11) (2005) 2190–2212.
41. D. Brandeis, D. Lehmann, M. Michel and W. Mingrone, Mapping event-related brain potential microstates to sentence endings, *Brain Topogr.* **8**(2) (1995) 145–159.
42. D. Lehmann and W. Skrandies, Reference-free identification of components of checkerboard-evoked multichannel potential fields, *Electroenceph. Clin. Neurophysiol.* **48** (1980) 609–621.
43. A. Bauer and H. Flexer, *Discovery of Common Subsequences in Cognitive Evoked Potentials*, Vol. 1510, ed. M. Q. J. M. Zytkow (Springer, Berlin, 1998), pp. 309–317.
44. W. Freeman, A cinematographic hypothesis of cortical dynamics in perception, *Int. J. Psychophysiol.* **60** (2006) 149–161.
45. J. Jeong, J. Gore and B. Peterson, Mutual information analysis of the EEG in patients with Alzheimer's disease, *Clin. Neurophysiol.* **112** (2001) 827–835.

46. M. Ahmadlou, H. Adeli and A. Adeli, New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *J. Neural Transm.* **117**(9) (2010) 1099–1109.
47. M. Ahmadlou and H. Adeli, Functional community analysis of brain: A new approach for EEG-based investigation of the brain pathology, *Neuroimage* **58**(2) (2011) 401–408.
48. J. Dauwels, F. B. Vialatte, T. Musha and A. Cichocki, A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG, *Neuroimage* **49**(1) (2010) 668–693.
49. Z. Sankari, H. Adeli and A. Adeli, Wavelet coherence model for diagnosis of Alzheimer's disease, *Clin. EEG Neurosci.* **43**(3) (2012) 268–278.
50. J. Escudero, S. Sanei, D. Abásolo and R. Hornero, Regional coherence evaluation in mild cognitive impairment and Alzheimer's disease based on adaptively extracted magnetoencephalogram rhythms, *Physiol. Meas.* **32**(8) (2011) 1163–1180.
51. Z. Sankari, H. Adeli and A. Adeli, Intrahemispheric, interhemispheric and distal EEG coherence in Alzheimer's disease, *Clin. Neurophysiol.* **122**(5) (2011) 897–906.
52. L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition* (Prentice Hall, 1993).
53. N. Houmani, F. B. Vialatte, C. F. Latchoumane, J. Jeong and G. Dreyfus, Stationary epoch-based entropy estimation for early diagnosis of Alzheimer's disease, *12th Low Voltage Low Power Conf., IEEE FTFC 2013*, Paris, France (2013) 1–4.
54. C. E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* **27** (1948) 379–423.
55. B. Jelles, P. Scheltens, W. M. Van der Flier, E. J. Jonkman, F. H. da Silva and C. J. Stam, Global dynamical analysis of the EEG in Alzheimer Disease: Frequency-specific changes of functional interactions, *Clin. Neurophysiol.* **119**(4) (2008) 837–841.
56. A. Tsolaki, D. Kazis, I. Kompatsiaris and V. Kosmidou, Electroencephalogram and Alzheimer's disease: Clinical and research approaches, *Int. J. Alzheimer's Dis.* **2014** (2014).
57. C. Silva, I. R. Pimentel, A. Andrade, J. P. Foreid and E. Ducla-Soares, Correlation dimension maps of EEG from epileptic absences, *Brain Topogr.* **11**(3) (1999) 201–209.
58. N. Talebi, A. M. Nasrabadi and T. Curran, Investigation of changes in EEG complexity during memory retrieval: The effect of Midazolam, *Cogn. Neurodyn.* **6**(6) (2012) 537–546.
59. H. Yang, Y. Wang, C. J. Wang and H. M. Tai, Correlation dimensions of EEG changes during mental tasks, *Eng. Med. Biol. Soc., 2004. IEMBS '04. 26th Annual Int. Conf. IEEE*, Vol. 1 (2004), pp. 616–619.
60. A. Mekler, Calculation of EEG correlation dimension: Large massifs of experimental data, *Comput. Methods Prog. Biomed.* **92**(1) (2008) 154–160.
61. Y. Shen, E. Olbrich, P. Achermann and P. Meier, Dimensional complexity and spectral properties of the human sleep EEG, *Clin. Neurophysiol.* **114** (2003) 199–209.
62. L. R. Rabiner and B. H. Juang, An introduction to hidden Markov models, *IEEE ASSP Magazine*, January 1986.
63. A. Flexer, G. Gruber and G. Dorffner, A reliable probabilistic sleep stager based on a single EEG signal, *Artif. Intell. Med.* **33**(3) (2005) 199–207.
64. V. M. Dadok, H. E. Kirsch, J. W. Sleight, B. A. Lopour and A. J. Szeri, A probabilistic framework for a physiological representation of dynamically evolving sleep state, *J. Comput. Neurosci.* **37**(1) (2014) 105–124.
65. D. Lederman and J. Tabrikian, Classification of multichannel EEG patterns using parallel hidden Markov models, *Med. Biol. Eng. Comput.* **50**(4) (2012) 319–328.
66. C. F. Latchoumane, F. B. Vialatte, J. Solé-Casals, M. Maurice, S. R. Wimalaratna, N. Hudson, J. Jeong and A. Cichocki, Multiway array decomposition analysis of EEGs in Alzheimer's disease, *J. Neurosci. Methods* **207**(1) (2012) 41–50.
67. G. Henderson, E. Ifeakor, 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 Trans. Biomed. Eng.* **53**(8) (2006) 1557.
68. G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price and E. M. Stadlan, Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease, *Neurology* **34**(7) (1984) 939–944.
69. J. Dauwels, K. Srinivasan, M. Ramasubba Reddy, T. Musha, F. Vialatte, C. Latchoumane, J. Jeong and C. Cichocki, Slowing and loss of complexity in Alzheimer's EEG: Two sides of the same coin? *Int. J. Alzheimer's Dis.* **2011** (2011) 539621.
70. L. Smith, Intrinsic limits on dimension calculation, *Phys. Lett. A* **133**(6) (1988) 283–288.
71. J. P. Eckmann and D. Ruelle, Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems, *Physica D* **56** (1992) 185–187.
72. A. Brun and E. Englund, Regional pattern of degeneration in Alzheimer's disease: Neuronal loss and histopathological grading, *Histopathology* **5** (1981) 459–564.
73. H. G. Shuster, *Deterministic Chaos*, 2nd edn. (VCH-Verlag Publishers, New York, 1988).
74. B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman and Company, 1982).