Using time causal quantifiers to characterize sleep stages.

Diego M. Mateos [1,2,*], Jaime Gómez-Ramírez[3], and Osvaldo A. Rosso[4]

[1]Facultad de Ciencia y Tecnología. Universidad Autónoma de Entre Ríos (UADER). Oro Verde, Entre Ríos, Argentina.
[2]Instituto de Matemática Aplicada del Litoral (IMAL-CONICET-UNL), CCT CONICET, Santa Fé, Argentina.
[3]Centre for Research in Neurodegenarative Diseases. Fundación Reina Sofía, Valderrebollo, 5 Madrid, Spain.
[4]Departamento de Física, Universidade Federal de Alagoas (UFAL), Maceió, AL, Brazil.
[*]Corresponding author: Diego M. Mateos, mateosdiego@gmail.com.

Abstract

Resent studies have found that sleep is closely connected with mood, playing substantial role in daily cognitive performance and memory. Sleep is a dynamic activity presenting different stages which progresses cyclically throughout night. It is very important to maintain these cycles for healthy body function when the person is awake. The construction of mathematical models of sleep that help us characterize the different sleep stages with an optimal use of sleep recording time is of paramount importance. For this purpose, in this work, we analyzed single-channel EEG signals from 17 healthy subjects. The signals were quantified using the permutation vector approach and applied five different information measures: i) Shannons entropy, ii) MPR statistical complexity, iii) Fisher information, iv) Renyì Min-entropy and v) Lempel–Ziv complexity over the signals. The results show that it is possible to quantify and classify the different sleep stages. In addition, we combine these measures in complexity/information-entropy space, extracting relevant information from the signals impossible to obtain using the measures separately. Information theory-based measures allow us to characterize the underlying dynamics of EEG time series, identifying the different sleep stages and defining electrophysiological biomarkers able to distinguish between pathological and normal conditions.

1 Introduction

We think in sleep as our daily period of rest and recovery from the stresses of everyday life; however, research is revealing that sleep is a dynamic activity during which many processes vital to health and well–being take place. New evidence shows that sleep is essential to helping maintain mood, memory and cognitive performance ([1–3]). Sleep plays also a pivotal role in the normal function of the endocrine and immune systems ([4]). The incidence of sleep problems in different pathologies is uncontroversial and studies shows a connection between sleep duration and health conditions such as obesity, diabetes, hypertension and depression among other ([5–7]).

Sleep is an active physiological process, while metabolism generally slows down during sleep, all major organs and regulatory systems continue to function normally. In general, sleep can be categorized into two different functioning modes: *rapid eye movement* (REM) sleep and *non-REM* sleep (NREM).

REM sleep is an active period of sleep marked by intense brain activity. Brain waves are fast and desynchronized similar to those in the waking state, this is also the sleep stage in which most dreams occur. Usually sleepers pass through four stages: sw1, sw2, sw34 and REM sleep. These stages progress cyclically from sws1 through REM returning to sw1 in a loop. A complete sleep cycle takes an average of 90 to 110 minutes, with each stage lasting between 5 to 15 minutes. It is very important that these cycles are maintained for healthy body function in awake state ([8]).

NREM sleep is characterized by a reduction in a physiological activity, the brain waves get slower and have greater amplitude, breathing and heart rate slow down and blood pressure drops. The NREM phase comprises three stages: stage 1 (sw1) characterizing by drowsiness or transition from being awake to falling asleep slowing down the brain waves and muscle activity. Stage 2 (sw2) is a period of light sleep during which eye movement stop. Brain waves become slower with occasional bursts of rapid waves (12-14 Hz) called *sleep spindles*, coupled with spontaneous periods of muscle tone mixed. Stages 34 (sw34), are characterized by the presence of slow brain wave called delta waves (0.5-4 Hz) interspersed with smaller, faster waves ([9]). In these phases sleep is deeper, with no eye movement, decreased muscle activity though muscles keep their ability to operate, resembling a coma state.

Polysomnography studies are performed to investigate problems in the sleep cycles i.e. 12 hours of multi-physiological registers such as electroencephalography (EEG), electrocardiogram (ECG) or electrooculography

(EOG) among others. These studies are usually performed in medical institutions requiring the patient to stay overnight, becoming an issue for the patient, especially when they are children. Moreover, the detection of the different sleep cycles is done visually (over the electrophysiological signals tracing) and is implemented by specialized doctors, which is time-consuming activity that relies upon the expertise of the person analyzing the signal.

Automating sleep stage classification will thus have many advantages including objective quantitative assessment and cost reduction. The existing algorithms for sleep classification can be broadly divided into two groups, multi-parameters (multi-channel EEG, ECG, etc) and single-channel EEG based analysis.

Multi-parameters algorithms tend to have higher performance ([10]), but are costly from the patient's perspective, for example in terms of lack of comfort ([11]) or the sleep disturbances induced by an excessive number of wire connections during the recording process ([12]). Single channel EEG are therefore more suitable when having in mind patient's comfort, especially given the large amount of time required for the study. According to the evidence ([13]), EEG signals are almost sufficient for reliable scoring. Single-channel EEG based analysis is a cheap way of automatic sleep scoring. Since there is only one signal, it is necessary to extract as much information as possible from it. For this purpose, the aim is to find the best quantifiers to use as a feature extractor. Many features were used in the literature, based in time domain ([14, 15]), frequency domain ([16, 17]), time-frequency domain ([18, 19]) and non-linear feature ([17, 20, 21]). However the large amount of these features, most do not have a physiological interpretation. This makes the parameter setting are carried out through trial and error. Because of this, in this work we propose the use of measures based on information theory as characterizers because they are easily interpreted as functional of the PDF probability distribution produced by the EEG signal.

Bandt and Pompe propose the concept of *permutation entropy* for quantifying the complexity of a system behind a time series ([22]). This methodology has been applied for investigating EEG in different contexts as such as epilepsy ([23]), coma ([24]), anesthetics ([25]), and particularly in sleep ([26–28]), showing better results to traditional analyzes and representing a robust approach to time series analysis based on counting ordinal patterns. The complexity of the brain would represent the amount of "information" contained in the organism, in the sense that it quantifies the dynamical features of the temporal patterns due to functional interactions produced by the underlying structural network. Information theory measures can capture the degree to which the neural system integrates specialized information.

In this work, we analyzed single–channel EEG signals from 17 subjects with a 12 hours polysomnography recorded from each subject. The signals were pre-processed by experts and cut into subsequences corresponding to different sleep stages. We used five time causal information quantifiers tu study the EEG signals: i) Permutation entropy ([22]), Permutation Min-entropy ([29]), MPR statistical complexity ([30]), Fisher information ([31]) and Permutation Lempel-Ziv complexity ([32]). In addition, the signals were analyzed using complexity / information vs entropy planes with the aim to extract non-accessible information through the classical use of these quantifiers. Through these analyzes we have been able to characterize and classify the signals according to the different sleep state. These quantifiers are shown as an interesting alternative to be used as biomarkers for the detection of sleep-associated pathologies.

The paper is organized as follows: In section 2 we explain the data analysed and we give a brief introduction to the information quantifiers used in this work. In the section 3 we study the distribution of signal patterns in the different sleep stages. Then, we analyze the signals using the fives information measure and compare the results. Finally, we analyze the EEG signal through three different complexity / information vs entropy planes. The discussion of the results, possible applications and future research are developed in section 4.

2 Methods

2.1 Electrophysiological recordings

For this study we used a single-channel EEG signal from a polysomnography recording belonging to 17 subjects taken from the Physionet databank, *The Sleep-EDF Database [Expanded]* ([33, 34]). The data are freely available at ([35]). The study group was made up with 5 men and 12 women, 25–38 years of age at the time of the recordings. This polysomnogram (PSGs) collection with accompanying hypnograms (expert annotations of sleep stages) comes from two studies (detail in [34, 36]). The recordings are whole-night polysomnographic sleep recordings containing EEG (from Fpz–Cz electrode locations). The EEG signals were each sampled at 100 Hz. The sleep stages were classified in the five stages above-mentioned: *Awake close eyes, REM, sw1, sw2* and *sw34*.

The EEG recordings were preprocessed with a band–pass filter between 0.5 to 30 Hz. The signals were cut into sleep stages subsequences following the annotation made by the expert. Table 1 shows information about the subjects and recordings characteristics used in this work.

Subject	Gender	Age	N_e of segment for each epoch				
			Awake	REM	sw1	sw2	sw34
1	F	33	1997	125	58	250	220
2	M	34	1885	215	59	373	298
3	F	32	3680	346	201	1022	201
4	F	31	1824	176	92	660	96
5	F	26	1907	163	94	545	95
6	M	29	1971	179	184	402	115
7	F	25	2008	209	61	485	57
8	F	32	1957	199	45	400	131
9	F	34	1773	270	137	514	94
10	M	34	1780	180	114	616	114
11	F	31	2069	102	56	407	136
12	F	35	2007	187	90	417	129
13	M	30	1958	198	89	403	162
14	F	39	1904	168	84	392	222
15	F	25	1985	131	68	262	360
16	M	26	1739	232	19	561	170
17	F	25	1079	265	81	512	107

Table 1: Subjects information analyzed in this work. N_e represents the number of segments for the same sleep stage. All segments have the same length, $L = 3000$ data points ($t = 30$ seconds of recording).

2.2 Brief introduction to the information measures

For the study of a physical phenomena it is necessary to have a sequence of measurements related to them. These sequences are usually given in the form of time series which allow to extract information on the underlying dynamical systems under study. A time series (i.e. an EEG recording) can be associated to a probability distribution function (PDF) and use information theory quantifiers to characterize its properties. Next we will introduce the basic notions of the information measures used in this work.

Shannon entropy

Given a time series $\mathcal{X}(t) \equiv \{x_t; t = 1, \cdots, N\}$, with N the number of observations, the Shannon's logarithmic information measure (*Shannon entropy*) ([37]) of the associated PDF, $P \equiv \{p_i; i = 1, \cdots, M\}$ with $\sum_{i=1}^{M} p_i = 1$, and M the number of possible states is defined by:

$$\mathcal{S}[P] \;=\; -\sum_{i=1}^{M} p_i \log(p_i) \;. \tag{1}$$

When we have total certainty that our system is in the state i the probability $p_i = 1$ and this functional is equal to zero. In contrast, when the probability distribution is uniform, $P_u \equiv \{p_i = 1/M; \forall i = 1, \cdots, M\}$, our knowledge about the system is minimum (all the states have the same probability) and the entropy reach its maximum.

Renyi Min-entropy

In 1961, A. Rènyi [38] was looking for the most general definition of information measures, that would preserve the additivity for independent events, and was compatible with the axioms of probability. He defines his generalize entropy as:

$$\mathcal{R}_q[P] \;=\; \frac{1}{1-q} \;\log\left(\sum_{i=1}^{M} (p_i)^q\right) \tag{2}$$

where the order q ($q \geq 0$ and $q \neq 1$) is bias parameter: $q < 1$ privileges rare events, while $q > 1$ privileges salient events. The Shannon entropy $\mathcal{S}[P]$ is recovered in the limit as $q \to 1$. When $q \to \infty$ the Rènyi entropy $\mathcal{R}_q$ converge to the *Min-entropy* $\mathcal{R}_\infty$

$$\mathcal{R}_\infty[P] \;=\; \min_i(-\log p_i) = -\log \max_i p_i. \tag{3}$$

The Min-entropy is the smallest entropy measure in the family of Rènyi entropies. In this sense, it is the strongest way to measure the information content of a discrete random variable. It can be shown that $\mathcal{R}_\infty[P]$

is a function of the highest probability only ([39]). In particular, Min-entropy is never larger than the Shannon entropy.

Fisher Information

The Shannon entropy $\mathcal{S}$ is a measure of "global character", however in small local regions it lacks sensitivity to strong changes in the PDF. For that case the *Fisher information* $\mathcal{F}$ is a more effective quantifier ([31, 40]):

$$\mathcal{F}[f] \;=\; \int \frac{|\vec{\nabla} f(x)|^2}{f(x)} dx \tag{4}$$

which constitutes a measure of the gradient content of the distribution f (continuous PDF), thus being quite sensitive even to tiny localized perturbations. The Fisher information could be interpreted as the amount of information that can be extracted from a set of measures, as the ability to estimate a particular parameter and also as a measure of the state of disorder of a system ([40, 41]), its most important property being the so-called Cramer–Rao bound ([42]). In equation (4), the gradient operator, $\vec{\nabla}$, significantly influences the contribution of minute local f -variations to the Fisher information value, so that the quantifier is said to be local. Local sensitivity is useful in scenarios whose description necessitates an appeal to a notion of "order" ([43–45]).

For Fisher information measure computation (discrete PDF), we follow the proposal for Dehesa and coworkers in ([46]) which define $\mathcal{F}$ for a discrete distribution as:

$$\mathcal{F}[P] \;=\; 4 \sum_{i=1}^{N-1} (\sqrt{p_{i+1}} - \sqrt{p_i})^2 \;. \tag{5}$$

A system in a very ordered state will have a very narrow PDF, the Shannon entropy is $\mathcal{S} \to 0$ and the Fisher information measure $\mathcal{F} \to \mathcal{F}_{max}$. On the other hand, when the system under study lies in a very disorder distribution (flat PDF) $\mathcal{S} \to \mathcal{S}_{max}$, $\mathcal{F} \to 0$. From this simple example we can extrapolate that the Fisher information measure and the Shannon entropy behavior behave reciprocally ([47]).

2.2.1 MPR Statistical Complexity

The processes present in nature are not, at least directly, quantified by randomness measures, that should make us to look for other types of measures of statistical or structural complexity to understand system dynamics represented by their time series ([48]).

If we take the two extreme -a perfect order (i.e., a periodic sequence) and a maximal randomness (i.e white noise)- both are very easily (number of bits [49]) to describe and their complexity should be closed to zero in both cases. At a given distance from these two extremes, it lies a wide range of possible ordinal structures. The statistical complexity measures allows one to quantify this array of behavior ([50]). In our work we consider the MPR statistical complexity introduced in ([30]), because is able to quantify critical details of dynamical processes underlying the data set.

Based in the Lopez-Ruiz et al. work [51], this statistical complexity measure is defined through the functional product form

$$\mathcal{C}[P] \;=\; \mathcal{Q}_J[P, P_e] \cdot \mathcal{H}[P] \tag{6}$$

of the normalized Shannon entropy:

$$\mathcal{H}[P] \;=\; \mathcal{S}[P]/\mathcal{S}_{max} \;, \tag{7}$$

with $\mathcal{S}_{max} = \mathcal{S}[P_e] = \log N$, $(0 \leq \mathcal{H} \leq 1)$ and the disequilibrium $\mathcal{Q}_J$ defined in term of the *Jensen–Shannon divergence*. That is,

$$\mathcal{Q}_J[P, P_u] \;=\; \mathcal{Q}_0 \; \mathcal{J}[P, P_u] \;, \tag{8}$$

with:

$$\mathcal{J}[P, P_u] \;=\; \mathcal{S}[(P + P_u)/2] - \mathcal{S}[P]/2 - \mathcal{S}[P_u]/2 \;, \tag{9}$$

the above-mentioned Jensen–Shannon divergence and $\mathcal{Q}_0$, a normalization constant $(0 \leq \mathcal{Q}_J \leq 1)$, are equal to the inverse of the maximum possible value of $\mathcal{J}[P, P_e]$. This value is obtained when one of the components of the PDF, P, say p_n, is equal to one and the remaining p_j are equal to zero.

The statistical complexity depends on two different probability distributions, the one associated with the system under analysis, P, and the uniform distribution, P_u. The distance between this two probability distributions are measured using the Jensen–Shannon divergence ([52]). Furthermore, it has been shown that for a given value of $\mathcal{H}$, the range of possible $\mathcal{C}$ values varies between a minimum $\mathcal{C}_{min}$ and a maximum $\mathcal{C}_{max}$, restricting the possible values of the statistical complexity in a given complexity-entropy plane ([53]). Because of this, statistical complexity measure, gives important additional information related to the correlational structure between the components of the physical system.

Lempel-Ziv complexity

Now, we will see a different way to analyze a sequence; in this case, it is not based on the time series $\mathcal{X}(t) \equiv \{x_t; t = 1, \cdots, N\}$ PDF, but in the way that x_t are repeated along the sequence. This complexity is based on the Kolmogorov complexity idea what is the minimal "information" contained in the sequence [42]. To estimate the complexity of $\mathcal{X}(t)$ we will use the Lempel and Ziv scheme proposed in 1976 [54]. In this approach, a sequence $\mathcal{X}(t)$ is parsed into a number $\mathcal{W}$ of words by considering any subsequence that has not yet been encountered as a new word. The Lempel–Ziv complexity c_{LZ} is the minimum numbers of word $\mathcal{W}$ used to reconstruct the information from the original time series. For example, if we have the sequence 10011011001010001011 can be parsed in 7 words: $1 \cdot 0 \cdot 01 \cdot 101 \cdot 1100 \cdot 1010 \cdot 001011$ in this case the complexity is $c_{LZ} = 7$. An easy way to applied the Lempel–Ziv algorithm was develop in [55]. The Lempel-Ziv complexity can be normalized based in the length N of the discrete sequence and the alphabet length (α) as:

$$\mathcal{C}_{LZ} = \frac{c_{LZ}[log_\alpha N]}{N}$$

.

2.3 Time series discretization using Bandt-Pompe approach

The study and characterization of time series $\mathcal{X}(t)$ by recourse to information theory tools assume that the underlying PDF is given a priori. Similar problem arise in the Lempel-Ziv complexity context, where is necessary to have a discrete time series to analyse. In the literature there are many methods to quantify continuous time series such as binarization, histograms or wavelet among others. However, an effective method that emerges naturally is the one introduced by Bandt and Pompe called permutation vectors ([22]).

The permutation vectors method is based on the relative values of the neighbors belonging to the series, and in consequence take naturally the time causality in which these values had been generated. To understand the idea, let us consider a real-valued discrete-time series $\mathcal{X}(t) = \{x_t \in \mathbb{R}\}$, and let $D \geq 2$ and $\tau \geq 1$ be two integers. They will be called the embedding dimension and the time delay, respectively. From the original time series, we introduce a D-dimensional vector $\mathbf{Y}_t^{(D,\tau)}$:

$$\mathbf{Y}_t^{(D,\tau)} \to (x_{t-(D-1)\tau}, \ldots, x_{t-\tau}, x_t) \qquad \text{with} \qquad t \geq (D-1)\tau \ . \tag{10}$$

There are conditions on D and τ in order that the vector $\mathbf{Y}_t^{(D,\tau)}$ preserves the dynamical properties of the full dynamical system. The components of the phase space trajectory $\mathbf{Y}_t^{(D,\tau)}$ are sorted in ascending order. Then, we can define a *permutation vector*, $\mathbf{\Pi}_t^{(D,\tau)}$, with components given by the original position of the sorted values in ascending order of the component of $\mathbf{Y}_t^{(D,\tau)}$. Each one of these vectors represents a pattern (or motif). There are $D!$ possible patterns. To clarify all these technicalities, let us show how all this works with an example. Suppose we have a continuous series such as $\mathcal{X}(t) = \{0.32,\ 1.8,\ 5.4,\ 0.25,\ 1.7\}$ and take the parameters $D = 3$ and $\tau = 1$. The embedding vectors $\mathbf{Y}_t^{(D,\tau)}$ in this case are defined as $\mathbf{Y}_1^{(3,1)} = (0.32, 1.8, 5.4)$; $\mathbf{Y}_2^{(3,1)} = (1.8,\ 5.4,\ 0.25)$; $\mathbf{Y}_3^{(3,1)} = (5.4,\ 0.35,\ 1.7)$, and the respective permutation vectors are $\mathbf{\Pi}_1^{(3,1)} = (0,\ 1,\ 2)$, $\mathbf{\Pi}_2^{(3,1)} = (1,\ 2,\ 0)$ and $\mathbf{\Pi}_3^{(3,1)} = (2,\ 0,\ 1)$.

Regarding the selection of the parameters, Bandt and Pompe suggested working with $3 \leq D \leq 6$ and specifically considered an embedding delay $\tau = 1$ in their paper ([22]). Nevertheless, it is clear that other values of τ could provide additional information. It has been recently shown that this parameter is strongly related, if it is relevant, to the intrinsic time scales of the system under analysis ([56–58]). In this work we used $D = 4$ and $\tau = 1$ for all the analysis, however similar results were obtained for $D = 5$ and $\tau = 1$.

The Bandt and Pompe approach applied to information quantifier where used in many works in the past. For each information measures described above, we have their counterpart based on permutation vectors quantification obtaining: i) Permutation entropy ([22]), ii) Permutation Min-entropy ([56]), iii) MPR statistical complexity ([59]), iv) Permutation Fisher information ([45]) and v) Permutation Lempel–Ziv complexity ([32]). To understand how to applied the Lempel-Ziv complexity to the permutation vector we will show an example. Suppose we have a continuous series $\mathcal{X} = \{2.5,\ 0.7,\ 1.1,\ 3.1,\ 2.7,\ 1.4,\ 2.3,\ 1.5,\ 1.1,\ 0.7,\ 0.8\}$ and we use the BP parameter $D = 2$ and $\tau = 1$, the embedding vectors are $\mathbf{Y}_1^{(2,1)} = (2.5, 0.7)$, $\mathbf{Y}_2^{(2,1)} = (0.7, 1.1)$, $\mathbf{Y}_3^{(2,1)} = (1.1, 3.1)$, $\mathbf{Y}_4^{(2,1)} = (3.1, 2.7)$, $\mathbf{Y}_5^{(2,1)} = (2.7, 1.4)$, $\mathbf{Y}_6^{(2,1)} = (1.4, 2.3)$, $\mathbf{Y}_7^{(2,1)} = (2.3, 1.5)$, $\mathbf{Y}_8^{(2,1)} = (1.5, 1.1)$, $\mathbf{Y}_9^{(2,1)} = (1.1, 0.7)$, $\mathbf{Y}_{10}^{(2,1)} = (0.7, 0.8)$. Since $D = 2$ the alphabet has a length equal to $D! = 2$, if the permutation vectors $\mathbf{\Pi}_0 = (0,\ 1)$ is represented by 0 and $\mathbf{\Pi}_1 = (1,\ 0)$ by 1 our continuous sequence is quantified as 1001101110 applying the algorithm explain in 2.2.1 we can parsing in $1 \cdot 0 \cdot 01 \cdot 101 \cdot 101$ obtaining $\mathcal{C}_{LZ} = 5$. For simplicity we use an example with parameter $D = 2$, since for $D > 2$ the length of the sequence must be much greater than $D!$. However, as shown in [32] this method can be extended to any D.

3 Results

Seventeen subjects were recorded through a 12 hours polysomnography. The signals were cut in sub–signals correspond to the different periods of sleep stages (*Awake, REM, sw1, sw2* and *sw34*). Figure 1 shows an example of an EEG raw signal (above) and the power spectrum (below) for one typical subject (subject 1) in different sleep stages. All the information about the number of epochs for each patient are described in table 1. All the signals were quantified using the Bandt and Pomp aproach with the parameters $D = 4, 5$ and $\tau = 1$. These embedding dimension is enough to the capture causality information of the ordinal structure of the time series ([22]).

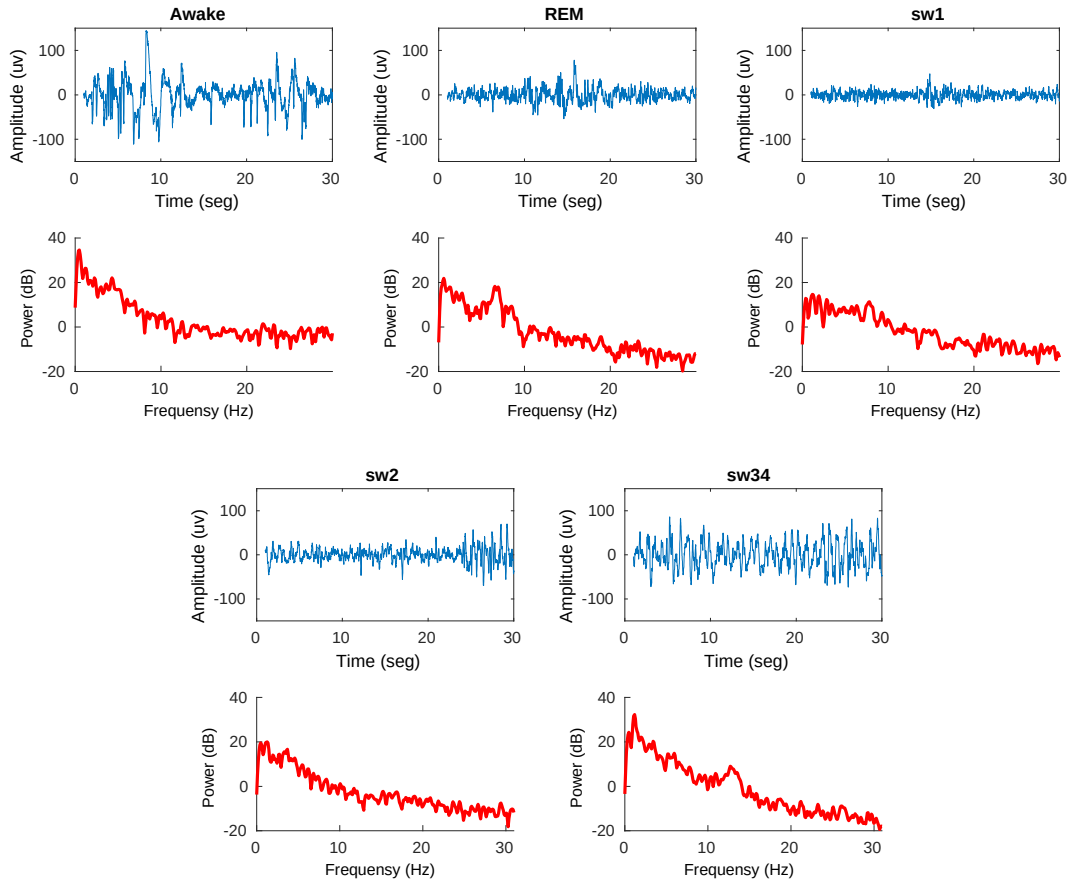


Figure 1: Example of the EEG signals and their respective power spectrum for each sleep stage: Awake, REM, sw1, sw2, sw34.

First, we studied the permutation vectors ($\mathbf{\Pi}^{(D,\tau)}$) distribution in each sleep stage. Figure 2 shows the permutation patterns distribution correspond to one subject (subject 1), the bars represent average overall epoch for the same stage. In Figure 3 we show the same analysis as before but the average is overall 17 subjects considering in this work. The result obtain in both cases are similar.

For subjects in awake state, the patterns distribution tends to be flat or uniform, this means that the occurrence frequency for the different motifs is similar for all the cases. In sw1, specific patterns begin to appear, this trend is accentuated in sw2, where the patient already enters a deeper sleep state. For sw34 two particular patterns arise (#1 and #24), they have a high frequency of occurrence in relation with the rest which are mostly close to zero. This behavior is given because, in an awake state, EEG signals have many superimposed frequencies, being similar to white noises ([27]). These, resulting in the necessity to use many different motifs to mapping the

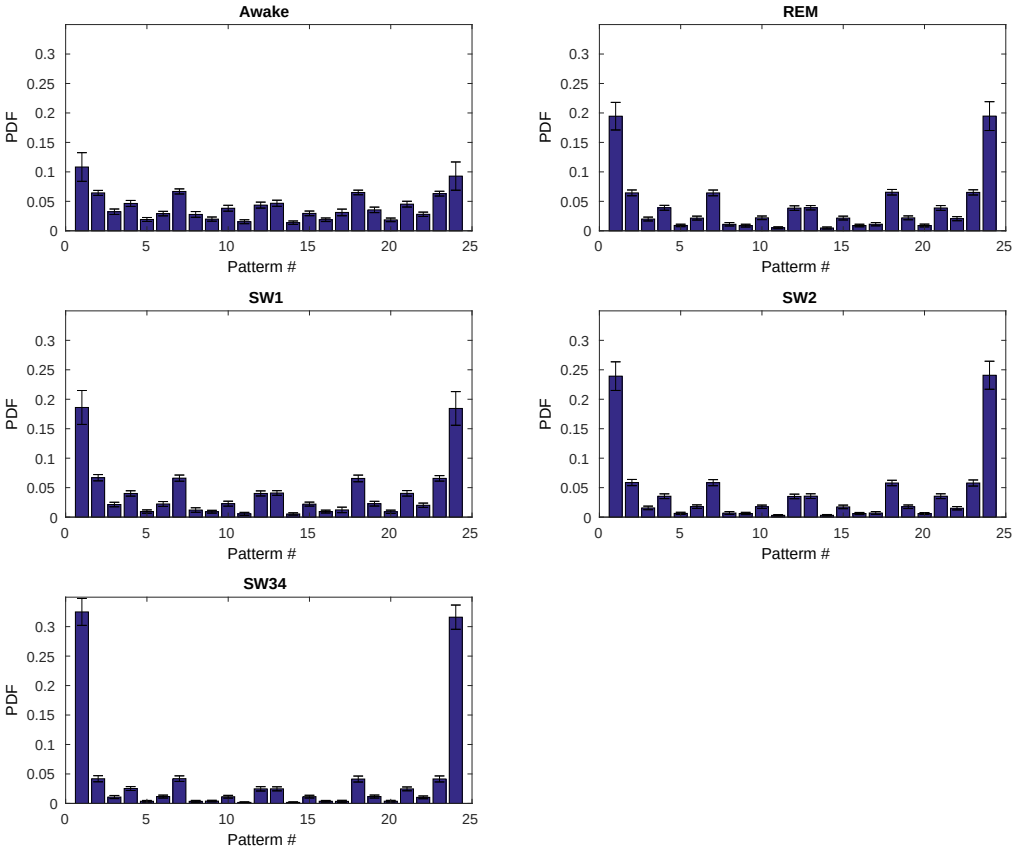


Figure 2: Average Bandt-Pompe PDF for the different sleep states corresponding to subject 1 (table:1), for the case $D = 4$ (pattern length) and $\tau = 1$ (time lag). The pattern histogram are the mean value and standard deviation calculated over all segment. Similar result were found for $D = 5$ and $\tau = 1$.

signal, tending the PDF patterns to be uniform. As the person falls asleep, he goes from awake to drowsy states (sw1, sw2) and finally deep sleep sw34. In these states arise low frequency waves persistent in time. Particularly in sw1 alpha waves (8-12 Hz) start to decrease, increasing theta waves (4-7.5 Hz) and starting to appear Low Voltage Mixed Frequency (LVMF) waves. In sw2 theta waves are more prominent in time than sw1. Sw34 is the deepest stage presenting more than 50 percent delta waves (0-4 Hz) activity ([60]). This tendency to exhibit lower frequencies results in the loss of PDF permutation patterns uniformity, because fewer patterns are required to map them. This is clear in sw34 where the predominant patterns are #1 and #24, which correspond to the ascending $\mathbf{\Pi}_1^{(3,1)} = [012]$ and descending $\mathbf{\Pi}_{24}^{(3,1)} = [210]$ motif, respectively. The REM state have similar PDF to sw1 due to REM present theta wave, low-amplitude and mixed-frequency EEG activity, comparable with are present in sw1 and sw2.

The following analysis was the application of the five measures introduced in section 2. Figure 4 despite boxplot for A) MPR statistical complexity $\mathcal{C}_{MPR}$, B) permutation entropy $\mathcal{H}$, C) Fisher information $\mathcal{F}$, D) permutation Lempel–Ziv Complexity $\mathcal{C}_{LZ}$ and E) permutation Min-entropy $\mathcal{R}_\infty$. In the Appendix A we present the values for each patient and the average over all patients. The five measures show a good distinction between the different sleep states except in the case between REM and sw1. $\mathcal{C}_{MPR}$ has lower values for subjects in an awake state and increases as the subjects enters a deeper sleep state. We obtain similar results for $\mathcal{F}$ but the values corresponding to the awake state present less dispersion. Conversely, $\mathcal{H}$, $\mathcal{C}_{LZ}$ and $\mathcal{R}_\infty$ are largest for the awake state and decrease for REM, sw1 and sw2 being a minimum for sw34. $\mathcal{H}$ and $\mathcal{C}_{LZ}$ have less dispersion for the values of all states in comparison with the other quantifiers. In all cases we cannot make a distinction between sw1 and REM state.

Since in the awake state the probability of all the motif (P_i) is similar (see figure 4A), the distance to uniform PDF (P_u) is minimized, giving a low value for $\mathcal{C}_{MPR}$. As the subject enters deeper sleep states, the PDF lose

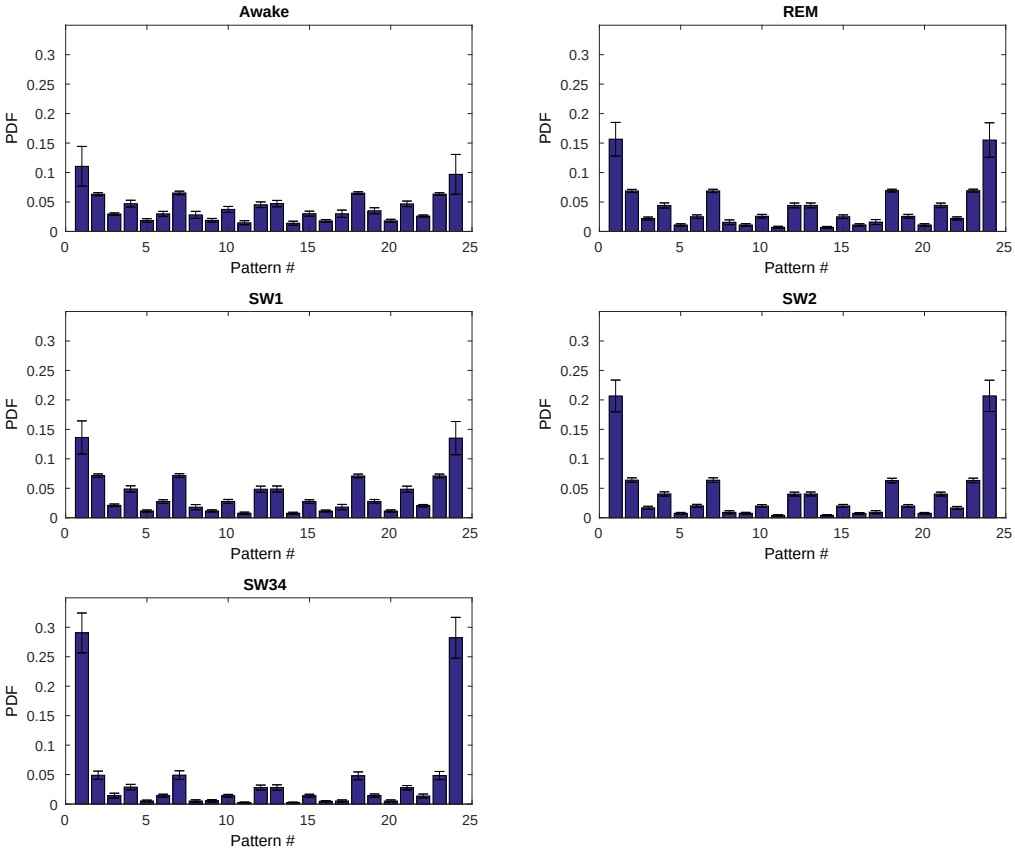


Figure 3: Average Bandt-Pompe PDF for the different sleep states considerate in detail the present work, for the case $D = 4$ (pattern length) and $\tau = 1$ (time lag). The BP pattern histogram is the mean value and standard deviation calculated over all subject and all segments see (table 1). Similar result was found for $D = 5$ and $\tau = 1$.

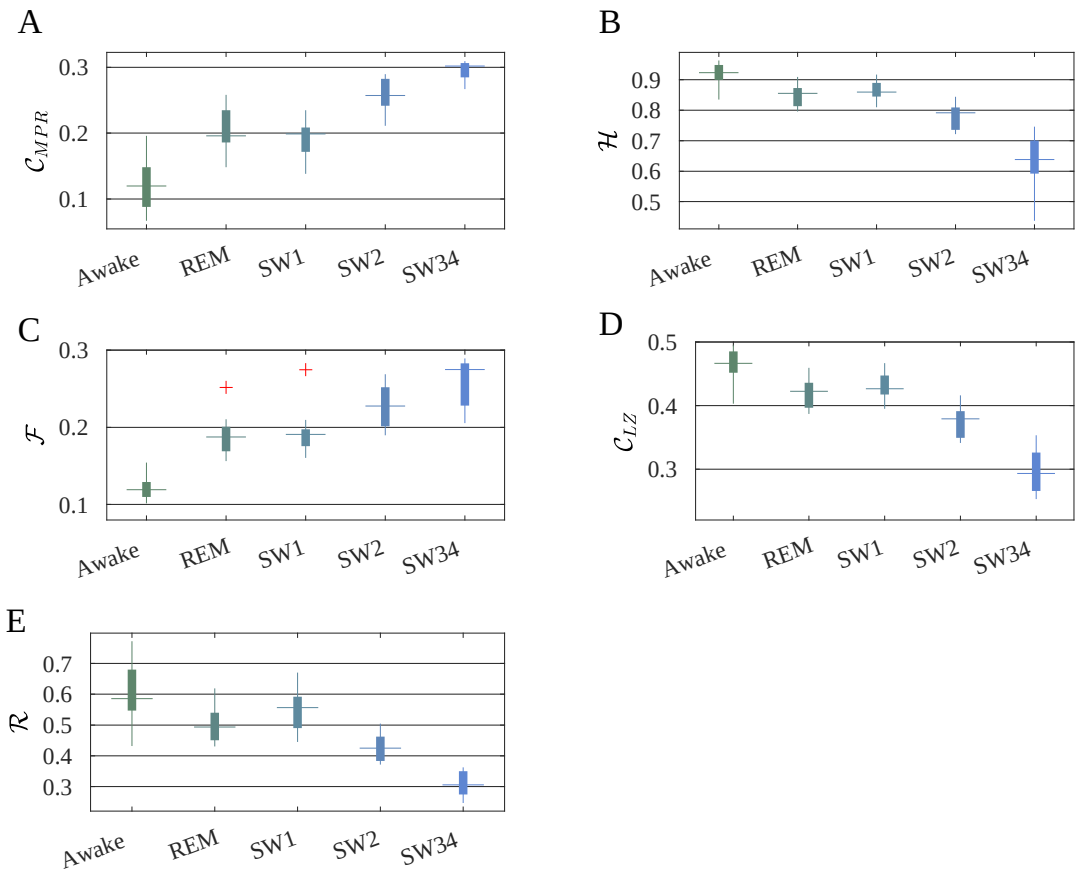


Figure 4: Information quantifiers for different sleep Stages. A)Despite the MPR statistical complexity ($\mathcal{C}_{MPR}$), B) normalized Shannon entropy ($\mathcal{H}$), C) Fisher information ($\mathcal{F}$), D) permutation Lempel-Ziv complexity ($\mathcal{C}_{LZ}$), E) permutation Min-Entropy ($\mathcal{R}_\infty$). The boxplot were estimated considering 17 patient with all their epoch (see table 1). For all the measures used there is a clear differentiation between sleep states, except for the case of REM and sw1. The Bandt-Pompe parameter used were $D = 4$ and $\tau = 1$, similar result we found used $D = 5$ and $\tau = 1$.

uniformity increasing the $\mathcal{C}_{MPR}$ value. Similar result we obtain for Fisher information, due to the PDF variability at the local level becomes more pronounced during deep sleep states. Moreover, uniform PDF gives maximum values for entropy measures such $\mathcal{H}$ or $\mathcal{R}_\infty$. When the PDF tent to a specific values of P_j, the entropies decrease as occurs in sw2 and sw34 states. In awake, the EEG variability results in the necessity to use more information to encode the signal giving a large value of $\mathcal{C}_{LZ}$. As sleep become deeper the complexity decrease because the signal become more repetitive. The quantifiers incapacity to distinguish between sw1 and REM is because both cases the permutation vectors PDF are similar (see 3). However, the $\mathcal{R}_\infty$ mean value in the sw1 state is greater than the average values of REM stage.

MPR permutation complexity is a measure derived from nonlinear systems theory that, when is applied to EEG signals, can be indicative of the dynamical complexity of electrocortical activity. However, to obtain further understanding of our results, we need to plot the localization in the causality complexity-entropy plane the results obtained in the present work. Figure 5 displays all the results considered here in the $\mathcal{H} \times \mathcal{C}_{MPR}$ location. The continuous lines represent the curves of maximum and minimum statistical complexity, $\mathcal{C}_{MPR}^{max}$ and $\mathcal{C}_{MPR}^{min}$, respectively, as functions of the normalized Shannon entropy ([53]).

Figure 5 shows the mean values distribution for each sleep state in a complexity-entropy plane ($\mathcal{C}_{MPR} \times \mathcal{H}$). To show that entropy-complexity values are not caused by random behavior of the signal, the recording for all patient were randomized and analysed. The mean value of the random series was represented in the plane (red star) and lies in the extreme maximum entropy and minimum complexity. Signals that present this king of behaviours (have no correlation as white noise) lives in this area of plane ([59]). Awake signals have more complexity and

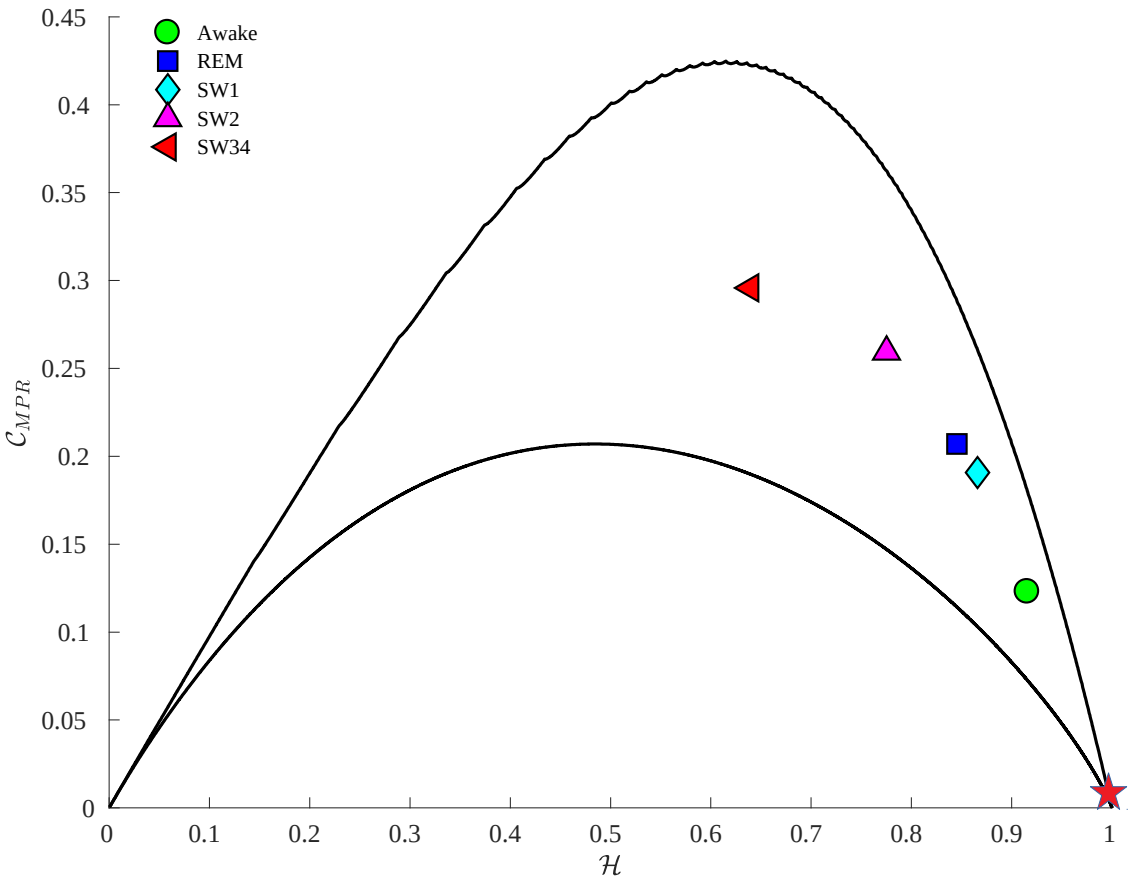


Figure 5: Entropy–Complexity plane ($\mathcal{H} \times \mathcal{C}_{MPR}$) for the 5 different sleep state. Each dot represents the mean value over 17 subject take in the study using the permutation parameter $D = 4$ and $\tau = 1$. The awake states are located in a region with greater entropy and less complexity. As the subjects are introduced in deeper sleep stages (sw1, sw2 and sw34) the complexity increases and the entropy decreases. REM is found in intermediate values of complexity and medium-high entropy. All states are clearly discriminated from the randomized signals (red star) which are at the end of the plane. Similar results were obtained for the parameter $D = 5$ and $\tau = 1$.

less entropy than random signals but still remind in the plane location where noises signals with low correlation lies ([59]). When sleep state change to sw1, sw2 and sw34, the complexity tend to increase and the entropy to decrease, moving the values to the center of the plane.

Using the same idea of entropy-complexity plane to study the sleep dynamics, we analyze the signal through the Fisher– Shannon plane, $\mathcal{H} \times \mathcal{F}$. This plane allows to quantify global versus local characteristic of the time series generated by dynamical processes ([45]). The plane can distinguish between regions corresponding to different sleep states, separating the non-REM from the REM and awake state. However, the REM and sw1 states are indistinguishable from each other, especially for Fisher information values. The awake stage presents low $\mathcal{F}$ and high $\mathcal{H}$. As the sleep go deeper the Non-REM stages are distinguished by having high values of $\mathcal{F}$ and low values of $\mathcal{H}$. By using the $\mathcal{H} \times \mathcal{F}$ plane, a better distinction of the states is observed than using each measure separately.

Finally, the signals were analyzed using a third plane which combines statistical measures (Shannon entropy) with an algorithmic measure (Lempel-Ziv complexity). This $\mathcal{H} \times \mathcal{C}_{LZ}$ plane, has been used to analyze EEG signals in different states of consciousness, giving excellent results characterizing different consciousness states such as awake, seizure or coma ([61]). Observing Figure 7, the plane clearly distinguishes all sleep states. The complexity-entropy values live on a straight line over the plane, this is because for a random, stationary and ergodic –which is the case for EEG under certain conditions ([27])– the Shannon entropy rate tends to the Lempel-Ziv complexity ([54]). The complexity and entropy decrease as the subjects fall asleep, being minimal for sw34. Unlike the other planes, in this case the values of REM and sw1 are clearly differentiable. This show the

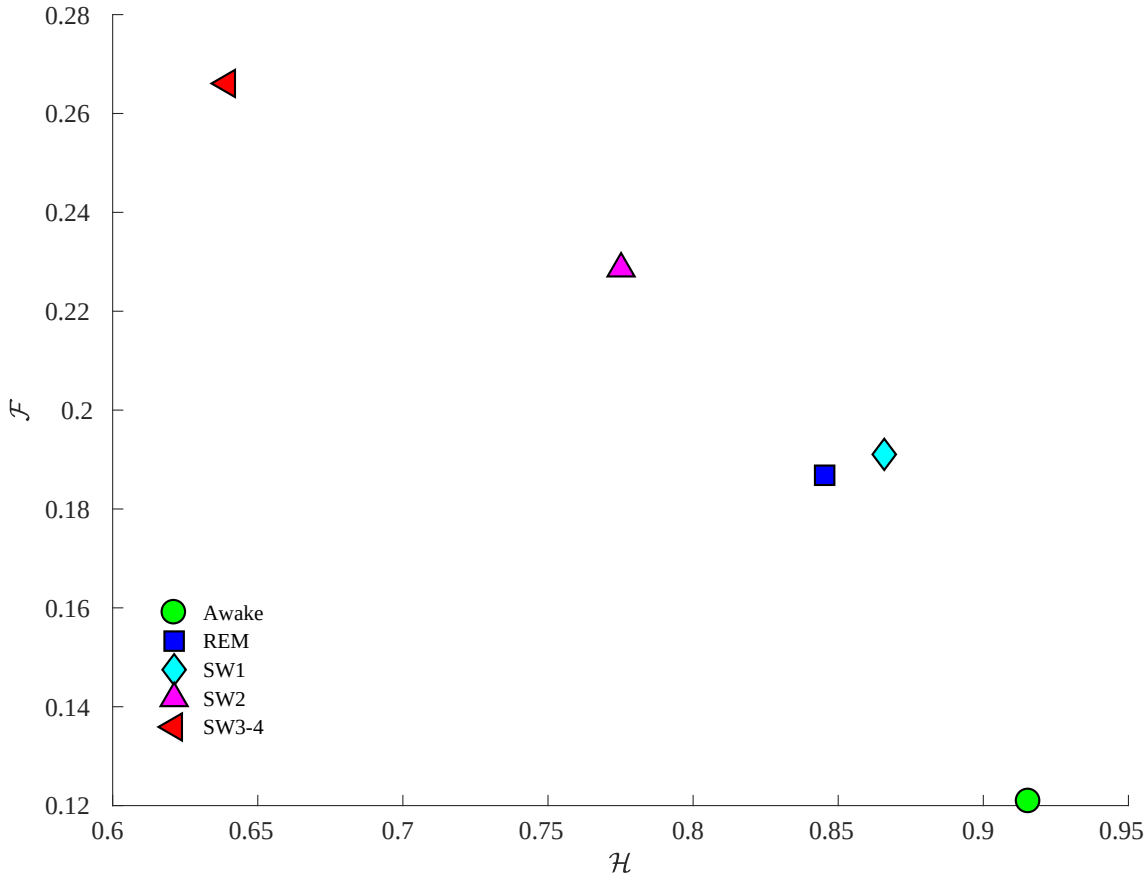


Figure 6: Shannon–Fisher information plane ($\mathcal{F} \times \mathcal{H}$) for the 5 different sleep state. Each dot represent the mean value over 17 subject take in the study using the permutation parameter $D = 4$ and $\tau = 1$. The awake values lies in high entropy low information region. REM and sws1 states are located in a region with greater entropy and meddle Fisher values. As the subjects remain in deeper sleep stages the Fisher values start to increases and their entropy decreases. Similar result were obtained for the parameter $D = 5$ and $\tau = 1$.

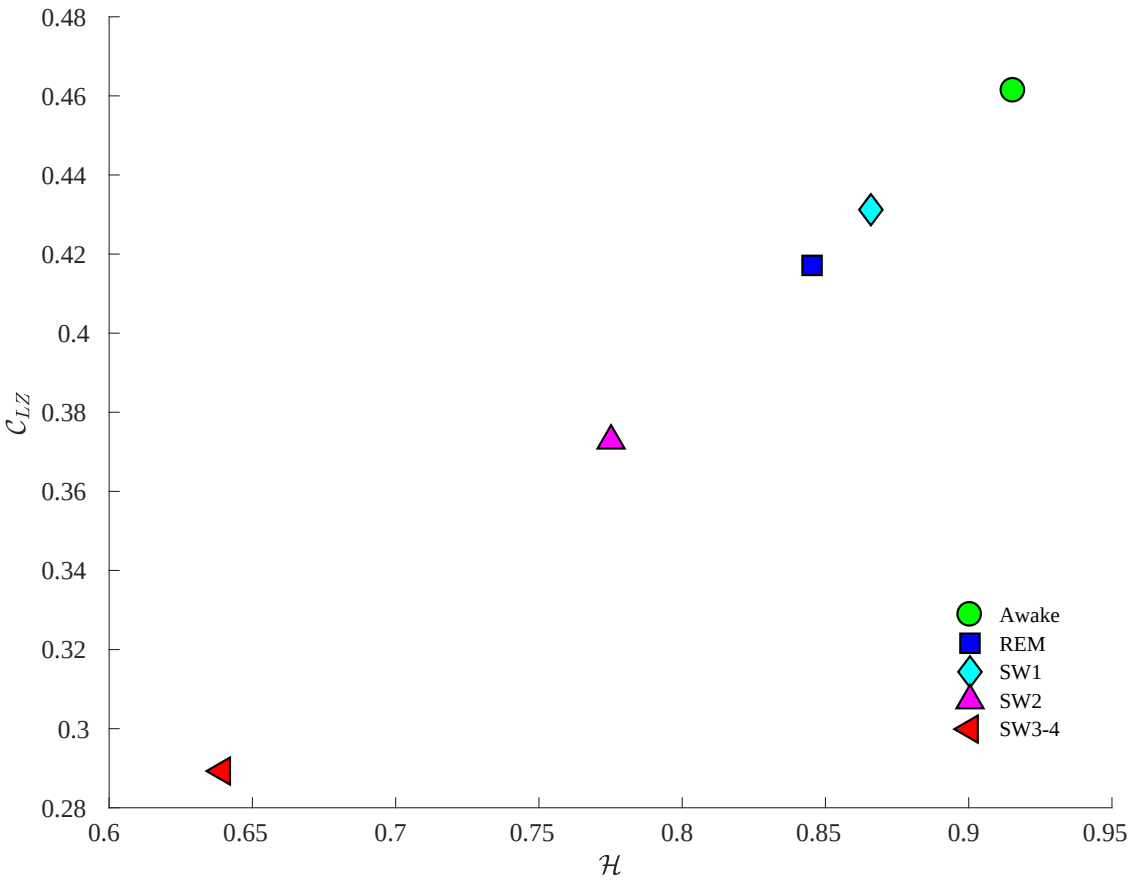


Figure 7: Shannon entropy - Lempel-Ziv complexity plane ($\mathcal{C}_{LZ} \times \mathcal{H}$) for the 5 different sleep state. Each dot represents the mean value over 17 subject take in the study using the permutation parameter $D = 4$ and $\tau = 1$. Awake stage present the maximum value of entropy and complexity. As the sleep enter in deeper stage the complexity and entropy decrease been minimum in sw34. Similar results were obtained for the parameter $D = 5$ and $\tau = 1$.

importance to combine measures for extracting new information about the signal. Similar results were found in intracranial EEG recording ([61]).

4 Discussion

The study of sleep stages are of vital importance to understand both healthy and pathological conditions. Dysfunctions in sleep cycles generate problems such as fatigue, muscular pains and attention problem among others. Because of that, the detection and proper characterization of sleep stages is of paramount importance. In this work we have been able to characterize classify the different sleep stages easy and fast, using tools coming from the Information Theory.

When the brain is in a wakefulness state, it needs to handle internal and external information, this requires the globality of its functions. Because of this, the electrical activity becomes more complex with greater variability as observed in the brains waves. This variability in the signal makes that a greater number of motif are required to mapping the signal, resulting that the permutation vectors PDF tent to the uniformity. This is translated into high values of entropy both Shannon and Renyì, such as in high Lempel-Ziv complexity. These results are consistent with pass research ([26], [62]). MPR complexity shows that signals in the awake state tend to be more similar to white noises corroborating the exposed by Bandt in ([27]). Given the reduced variability in the permutation vectors PDF at the local level in this state, Fishers information present minimal values.

As sleep progresses, most of the vital signals decrease (respiration, blood pressure, etc). The cortex become more inactive, so the EEG become slower and more repetitive with a marked predominance of delta waves.

This makes that the number of patterns required to mapping the signal decreased. Some motifs become more important than others, specially those that mapping the complete ascent or descent of the signal. This is reflected in a decrease in the signal entropy both $\mathcal{H}$ and $\mathcal{R}_\infty$. Due to the signal becomes more repetitive, the number of code required to reconstruct it become smaller, giving a drop in the permutation Lempel-Ziv complexity. Similar results have been obtained by using Shannon entropy and Lempel-Ziv complexity with different quantification techniques ([62, 63]). This constant behavior of the signal in the No-REM states generates a brake in the uniformity PDF permutation patterns resulting in an increase in MRP complexity values. In these states the Fisher information increase because of the difference between $P(\Pi_i)$ and $P(\Pi_{i+j})$ becomes more pronounced (see Figure 2).

As can be seen in Figure 4 most of the values for REM state are similar than those in sw1. This is expected considering that the EEG of REM sleep contains frequencies that are present in the "awake" state and in the lighter stages of sleep sw1 ([26]). Despite the similarities of the REM and awake / light sleep EEG, there are still enough differences between them such that statistically different values are obtained. The presence of 11-16 Hz activity (sleep spindles) in sw2, and more abundant alpha activity (8-13Hz) in REM sleep means that these two stages present activity at an overlapping frequency range, which explains the proximity of the information values obtained for these two stages. Difficulty in separating sw2 and REM sleep has also been encountered using other measures ([64]). However, when analyzing the signals through the complexity/information-entropy planes, we see that the sw2 and REM states tend to have a better differentiation, especially for the $\mathcal{C}_{LZ} \times \mathcal{H}$ plane.

The use of complexity/information-entropy planes provides a better data visualization, giving information about the signals which are impossible to get applying each measure individually . The three planes presented in this work shows a clear differentiation over all sleep states. In particular, they allow us to differentiate sw2 from REM, which is not possible when analyzing the quantifiers separately. These results are given because the combination of different approach, for example: i) statistical and algorithmic measures ($\mathcal{C}_{LZ} \times \mathcal{H}$), ii) local and global measures ($\mathcal{F} \times \mathcal{H}$) or iii) disequilibrium and disorder ($\mathcal{C}_{MPR} \times \mathcal{H}$).

This work makes a substantiated case for the use of information-theoretic tools for the analysis of the one–chanel EEG signals, complementing the visual inspection adopted by electrophysiologists. The information quantifiers used in this work have proven highly effective for the detection and classification of sleep states. Provided the possibility of working with short sequences, its simplicity and calculation speed become it a tool with high potential to use in real time analysis for one-single channel EEG. This feature would allow to develop comfortable and simple devise, to apply in a first stage of sleep studies, without the need to go to a clinical institution. Given the importance that sleep cycles have in peoples health, in future work we will focus on applying these analytical methods for the study of pathological EEGs. To that end, data from patients with insomnia, circadian rhythmic sleep-awake and parasomnia among others, will be explored using these techniques to validate them as biomarkers. Our forthcoming work is to employ these measures as features in unsupervised classifiers, for example support vector machine with mixed penalties.

References

[1] R. M. Benca, M. Okawa, M. Uchiyama, S. Ozaki, T. Nakajima, K. Shibui, and W. H. Obermeyer. Sleep and mood disorders. *Sleep medicine reviews*, 1(1):45–56, 1997.

[2] R. Stickgold. Sleep-dependent memory consolidation. *Nature*, 437(7063):1272, 2005.

[3] J. J. Pilcher and A. S. Walters. How sleep deprivation affects psychological variables related to college students' cognitive performance. *Journal of American College Health*, 46(3):121–126, 1997.

[4] P. A. Bryant, J. Trinder, and N. Curtis. Sick and tired: does sleep have a vital role in the immune system? *Nature Reviews Immunology*, 4(6):457, 2004.

[5] K. L. Knutson, K. Spiegel, P. Penev, and E. Van Cauter. The metabolic consequences of sleep deprivation. *Sleep medicine reviews*, 11(3):163–178, 2007.

[6] P. Lusardi, A. Zoppi, P. Preti, R. M. Pesce, E. Piazza, and R. Fogari. Effects of insufficient sleep on blood pressure in hypertensive patients: a 24-h study. *American journal of hypertension*, 12(1):63–68, 1999.

[7] A. Wirz-Justice and R. H. den Hoofdakker. Sleep deprivation in depression: what do we know, where do we go? *Biological psychiatry*, 46(4):445–453, 1999.

[8] R. D. Ogilvie. The process of falling asleep. *Sleep Medicine Reviews*, 5(3):247–270, 2001.

[9] R. Boostani, F. Karimzadeh, and M. Nami. A comparative review on sleep stage classification methods in patients and healthy individuals. *Computer methods and programs in biomedicine*, 140:77–91, 2017.

[10] S. Khalighi, T. Sousa, G. Pires, and U. Nunes. Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels. *Expert Systems with Applications*, 40(17):7046–7059, 2013.

[11] A. Malhotra, M. Younes, S. T. Kuna, R. Benca, C. A. Kushida, J. Walsh, A. Hanlon, B. Staley, A. I. Pack, and G. W. Pien. Performance of an automated polysomnography scoring system versus computer-assisted manual scoring. *Sleep*, 36(4):573–582, 2013.

[12] S. Liang, C. Kuo, Y. Hu, Y. Pan, and Y. Wang. Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models. *IEEE Transactions on Instrumentation and Measurement*, 61(6):1649–1657, 2012.

[13] A. Flexer, G. Gruber, and G. Dorffner. A reliable probabilistic sleep stager based on a single EEG signal. *Artificial Intelligence in Medicine*, 33(3):199–207, 2005.

[14] K. Šušmáková and A. Krakovská. Discrimination ability of individual measures used in sleep stages classification. *Artificial intelligence in medicine*, 44(3):261–277, 2008.

[15] B. Hjorth. EEG analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3):306–310, 1970.

[16] F. Chapotot and G. Becq. Automated sleep–wake staging combining robust feature extraction, artificial neural network classification, and flexible decision rules. *International Journal of Adaptive Control and Signal Processing*, 24(5):409–423, 2010.

[17] B. Weiss, Z. Clemens, R. Bódizs, Z. Vágó, and P. Halász. Spatio-temporal analysis of monofractal and multifractal properties of the human sleep EEG. *Journal of neuroscience methods*, 185(1):116–124, 2009.

[18] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, and H. Dickhaus. Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier. *Computer methods and programs in biomedicine*, 108(1):10–19, 2012.

[19] T. Schluter and S. Conrad. An approach for automatic sleep stage scoring and apnea-hypopnea detection. In *2010 IEEE International Conference on Data Mining*, pages 1007–1012. IEEE, 2010.

[20] U. R. Acharya, E. C. Chua, K. Chua, L. Min, and T. Tamura. Analysis and automatic identification of sleep stages using higher order spectra. *International journal of neural systems*, 20(06):509–521, 2010.

[21] B. Koley and D. Dey. An ensemble system for automatic sleep stage classification using single channel EEG signal. *Computers in biology and medicine*, 42(12):1186–1195, 2012.

[22] C. Bandt and B. Pompe. Permutation entropy: a natural complexity measure for time series. *Physical Review Letters*, 88(17):174102, 2002.

[23] D. M. Mateos, J. M.. Diaz, and P. W. Lamberti. Permutation entropy applied to the characterization of the clinical evolution of epileptic patients under pharmacological treatment. *Entropy*, 2014.

[24] L. Zhu, G. Cui, J. Cao, A. Cichocki, J. Zhang, and C. Zhou. A Hybrid System for Distinguishing between Brain Death and Coma Using Diverse EEG Features. *Sensors*, 19(6):1342, 2019.

[25] D. Jordan, G. Stockmanns, E. F. Kochs, S. Pilge, and G. Schneider. Electroencephalographic order pattern analysis for the separation of consciousness and unconsciousnessan analysis of approximate entropy, permutation entropy, recurrence rate, and phase coupling of order recurrence plots. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, 109(6):1014–1022, 2008.

[26] N. Nicolaou and J. Georgiou. The use of permutation entropy to characterize sleep electroencephalograms. *Clinical EEG and Neuroscience*, 42(1):24–28, 2011.

[27] C. Bandt. A new kind of permutation entropy used to classify sleep stages from invisible EEG microstructure. *Entropy*, 19(5):197, 2017.

[28] C. Kuo and S. Liang. Automatic stage scoring of single-channel sleep eeg based on multiscale permutation entropy. In *2011 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 448–451. IEEE, 2011.

[29] L. Zunino, . Olivares, and O. A. Rosso. Permutation min-entropy: An improved quantifier for unveiling subtle temporal correlations. *Epl*, 109(1), 2015.

[30] M. T. Lamberti, P. W..and Martin, A Plastino, and O A Rosso. Intensive entropic non-triviality measure. *Physica A*, 334(1):119–131, 2004.

[31] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222:309–368, 1922.

[32] S. Zozor, D. M. Mateos, and P. W. Lamberti. Mixing Bandt-Pompe and Lempel-Ziv approaches: another way to analyze the complexity of continuous-state sequences. *The European Physical Journal B*, 87(5):107, 2014.

[33] A. L. Goldberger and Others. Components of a new research resource for complex physiologic signals, physiobank, physiotoolkit, and physionet, american heart association journals. *Circulation*, 101(23):1–9, 2000.

[34] B Kemp, A H Zwinderman, B Tuk, H. A. Kamphuisen, and J. J. Oberye. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9):1185–1194, 2000.

[35] Physionet databanck. Sleep-EDF Database Expanded. \url{https://physionet.org/content/sleep-edfx/1.0.0/}.

[36] M. S. Mourtazaev, B Kemp, A. H. Zwinderman, and H. A. Kamphuisen. Age and gender affect different characteristics of slow waves in the sleep EEG. *Sleep*, 18(7):557–564, 1995.

[37] C. E. Shannon and W. Weaver. *The mathematical theory of communication.* University of Illinois press, 1998.

[38] A. Rényi and Others. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics.* The Regents of the University of California, 1961.

[39] S. E. Grigorescu and J. M. Petkov. Texture analysis using Renyi's generalized entropies. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, volume 1, pages I—-241. IEEE, 2003.

[40] B. R. Frieden. *Science from Fisher information: a unification.* Cambridge University Press, 2004.

[41] A. L. Mayer, C. W. Pawlowski, and H. Cabezas. Fisher information and dynamic regime changes in ecological systems. *ecological modelling*, 195(1):72–82, 2006.

[42] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, 2nd edition, 2006.

[43] O. A. Rosso, L. De Micco, A. Plastino, and H. A. Larrondo. Info-quantifiers' map-characterization revisited. *Physica A: Statistical Mechanics and its Applications*, 389(21):4604–4612, 2010.

[44] F. Olivares, A. Plastino, and O. A. Rosso. Ambiguities in Bandt–Pompe's methodology for local entropic quantifiers. *Physica A: Statistical Mechanics and its Applications*, 391(8):2518–2526, 2012.

[45] F. Olivares and O. A. Plastino, A.and Rosso. Contrasting chaos with noise via local versus global information quantifiers. *Physics Letters A*, 376(19):1577–1583, 2012.

[46] P. Sánchez-Moreno, R. J. Yánez, and J. S. Dehesa. Discrete densities and Fisher information. In *Proceedings of the 14th International Conference on Difference Equations and Applications. Difference Equations and Applications. Istanbul, Turkey: Bahçesehir University Press*, pages 291–298, 2009.

[47] F. Pennini and A. Plastino. Reciprocity relations between ordinary temperature and the Frieden-Soffer Fisher temperature. *Physical Review E*, 71(4):47102, 2005.

[48] D. P. Feldman and J. P. Crutchfield. Measures of statistical complexity: Why? *Physics Letters A*, 238(4):244–252, 1998.

[49] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.

[50] D. P. Feldman, C. S. McTague, and J. P. Crutchfield. The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(4):43106, 2008.

[51] R. Lopez-Ruiz, H. L. Mancini, and X. Calbet. A statistical measure of complexity. *Physics Letters A*, 209(5-6):321–326, 1995.

[52] I. Grosse, P. Bernaola-Galván, P. Carpena, R Román-Roldán, J. Oliver, and H. Stanley. Analysis of symbolic sequences using the Jensen-Shannon divergence. *Physical Review E*, 65(4):41905, 2002.

[53] M. T. Martin, A. Plastino, and O. A. Rosso. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A: Statistical Mechanics and its Applications*, 369(2):439–462, 2006.

[54] A. Lempel and J. Ziv. On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 22(1):75–81, jan 1976.

[55] F. Kaspar and H. G. Schuster. Easily calculable measure for the complexity of spatiotemporal patterns. *Physical Review A*, 36(2):842–848, 1987.

[56] L. Zunino, M. C. Soriano, I Fischer, O. A. Rosso, and C. R. Mirasso. Permutation-information-theory approach to unveil delay dynamics from time-series analysis. *Physical Review E*, 82(4):46212, 2010.

[57] M. C. Soriano, L. Zunino, O. A. Rosso, I. Fischer, and C. R. Mirasso. Time scales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis. *IEEE Journal of Quantum Electronics*, 47(2):252–261, 2011.

[58] L. Zunino, M. C. Soriano, and O. A. Rosso. Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach. *Physical Review E*, 86(4):46210, 2012.

[59] O. A. Rosso, H. Larrondo, M. T. Martin, A. Plastino, and M. A. Fuentes. Distinguishing Noise from Chaos. *Physical Review Letters*, 99(15):154102, 2007.

[60] N. R. Carlson. *Physiology of behavior*. Allyn & Bacon, 1994.

[61] D. M. Mateos, R. Guevara Erra, R. Wennberg, and J. L. Perez Velazquez. Measures of entropy and complexity in altered states of consciousness. *Cognitive Neurodynamics*, pages 1–12, 2017.

[62] L. Li and R. Wang. Complexity analysis of sleep EEG signal. In *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, pages 1–3. IEEE, 2010.

[63] N. Burioka, M. Miyata, G. Cornélissen, and et al Halberg. Approximate Entropy in the Electroencephalogram during Wake and Sleep. *Clinical EEG and Neuroscience*, 36(1):21–24, 2005.

[64] Q. Noirhomme, M. Boly, V. Bonhomme, P. Boveroux, C. Phillips, P. Peigneux, Soddu, and Others. BISPECTRAL INDEX CORRELATES WITH REGIONAL CEREBRAL BLOOD FLOW DURING SLEEP. *Archives Italiennes de Biologie*, 147(1/2):51–57, 2009.

A Tables with all quantifiers values

Below we present the tables with the average values for the different information measures over each patient and as a whole. The average values were calculated over all epoch belonging to each sleep state.

Subject	Awake	REM	SW1	SW2	SW34
1	0.9229	0.7980	0.8095	0.7295	0.5967
2	0.9470	0.8563	0.8470	0.7938	0.6899
3	0.8365	0.8312	0.8540	0.7353	0.5944
4	0.8344	0.8372	0.8594	0.7355	0.6032
5	0.9102	0.8014	0.8708	0.7217	0.5711
6	0.8729	0.8175	0.8280	0.7423	0.5856
7	0.9351	0.8912	0.8857	0.8085	0.4372
8	0.9328	0.8750	0.8891	0.7912	0.7126
9	0.9188	0.8863	0.9104	0.8095	0.6967
10	0.8878	0.8554	0.8503	0.7950	0.7108
11	0.9523	0.8614	0.8858	0.8219	0.6895
12	0.9207	0.8007	0.8335	0.7558	0.5745
13	0.9247	0.8203	0.8506	0.7453	0.6385
14	0.9509	0.8716	0.8912	0.8011	0.7187
15	0.9030	0.7953	0.8380	0.7335	0.6199
16	0.9511	0.8628	0.8995	0.8145	0.6853
17	0.9629	0.9088	0.9168	0.8438	0.7459

Table 2: Shannon entropy values ($\mathcal{H}$) for each subject analyzed in this work. The table shows the average value over all segments belonging to the same sleep stage. The parameter used were $D = 4$ and $\tau = 1$.

Subject	Awake	REM	SW1	SW2	SW34
1	0.1177	0.2422	0.2349	0.2824	0.3056
2	0.0888	0.1954	0.2064	0.2438	0.2808
3	0.1959	0.2213	0.2065	0.2822	0.3090
4	0.1955	0.2153	0.1994	0.2809	0.3081
5	0.1351	0.2425	0.1854	0.2872	0.3057
6	0.1686	0.2233	0.2125	0.2719	0.2978
7	0.1047	0.1668	0.1740	0.2384	0.3018
8	0.1091	0.1845	0.1737	0.2572	0.2853
9	0.1206	0.1752	0.1486	0.2425	0.3079
10	0.1518	0.1958	0.1987	0.2453	0.2850
11	0.0825	0.1942	0.1753	0.2289	0.2802
12	0.1201	0.2439	0.2224	0.2770	0.3059
13	0.1197	0.2324	0.2074	0.2868	0.3078
14	0.0856	0.1864	0.1648	0.2484	0.2832
15	0.1472	0.2581	0.2323	0.2894	0.3044
16	0.0822	0.1941	0.1630	0.2367	0.2942
17	0.0670	0.1484	0.1381	0.2111	0.2668

Table 3: MPR Statistical Complexity values ($\mathcal{C}_{\mathcal{MPR}}$) for each subject analyzed in this work. The table shows the average value over all segments belonging to the same sleep stage. The parameter used were $D = 4$ and $\tau = 1$.

Subject	Awake	REM	SW1	SW2	SW34
1	0.1110	0.1985	0.1977	0.2409	0.2782
2	0.1014	0.1605	0.1745	0.1944	0.2154
3	0.1542	0.1994	0.2096	0.2543	0.2875
4	0.1528	0.1928	0.1960	0.2495	0.2805
5	0.1268	0.2105	0.1880	0.2575	0.2892
6	0.1362	0.1701	0.1761	0.2167	0.2559
7	0.1118	0.1618	0.1759	0.2017	0.4430
8	0.1192	0.1740	0.1908	0.2343	0.2315
9	0.1109	0.1874	0.1939	0.2275	0.2835
10	0.1231	0.1562	0.1605	0.2000	0.2265
11	0.1114	0.1666	0.1766	0.1961	0.2144
12	0.1267	0.2090	0.2055	0.2511	0.2828
13	0.1199	0.2050	0.1973	0.2688	0.2751
14	0.1062	0.1701	0.1639	0.2208	0.2287
15	0.1391	0.2518	0.2747	0.2672	0.2748
16	0.1028	0.1924	0.1975	0.2163	0.2511
17	0.1016	0.1699	0.1693	0.1895	0.2054

Table 4: Fisher Information values ($\mathcal{F}$) for each subject analyzed in this work. The table shows the average value over all segments belonging to the same sleep stage. The parameter used were $D = 4$ and $\tau = 1$.

Subject	Awake	REM	SWS1	SWS2	SWS34
1	0.4664	0.3884	0.3949	0.3468	0.2700
2	0.4861	0.4249	0.4187	0.3860	0.3230
3	0.4048	0.4048	0.4222	0.3489	0.2667
4	0.4030	0.4088	0.4265	0.3493	0.2712
5	0.4576	0.3901	0.4348	0.3412	0.2527
6	0.4289	0.3986	0.4077	0.3532	0.2620
7	0.4762	0.4467	0.4442	0.3906	0.1199
8	0.4743	0.4370	0.4469	0.3790	0.3352
9	0.4614	0.4415	0.4593	0.3930	0.3211
10	0.4403	0.4224	0.4189	0.3855	0.3348
11	0.4870	0.4273	0.4428	0.4021	0.3201
12	0.4643	0.3893	0.4109	0.3612	0.2541
13	0.4695	0.4028	0.4207	0.3521	0.2932
14	0.4879	0.4354	0.4486	0.3859	0.3402
15	0.4553	0.3869	0.4132	0.3491	0.2811
16	0.4850	0.4270	0.4536	0.3976	0.3192
17	0.4950	0.4596	0.4667	0.4160	0.3531

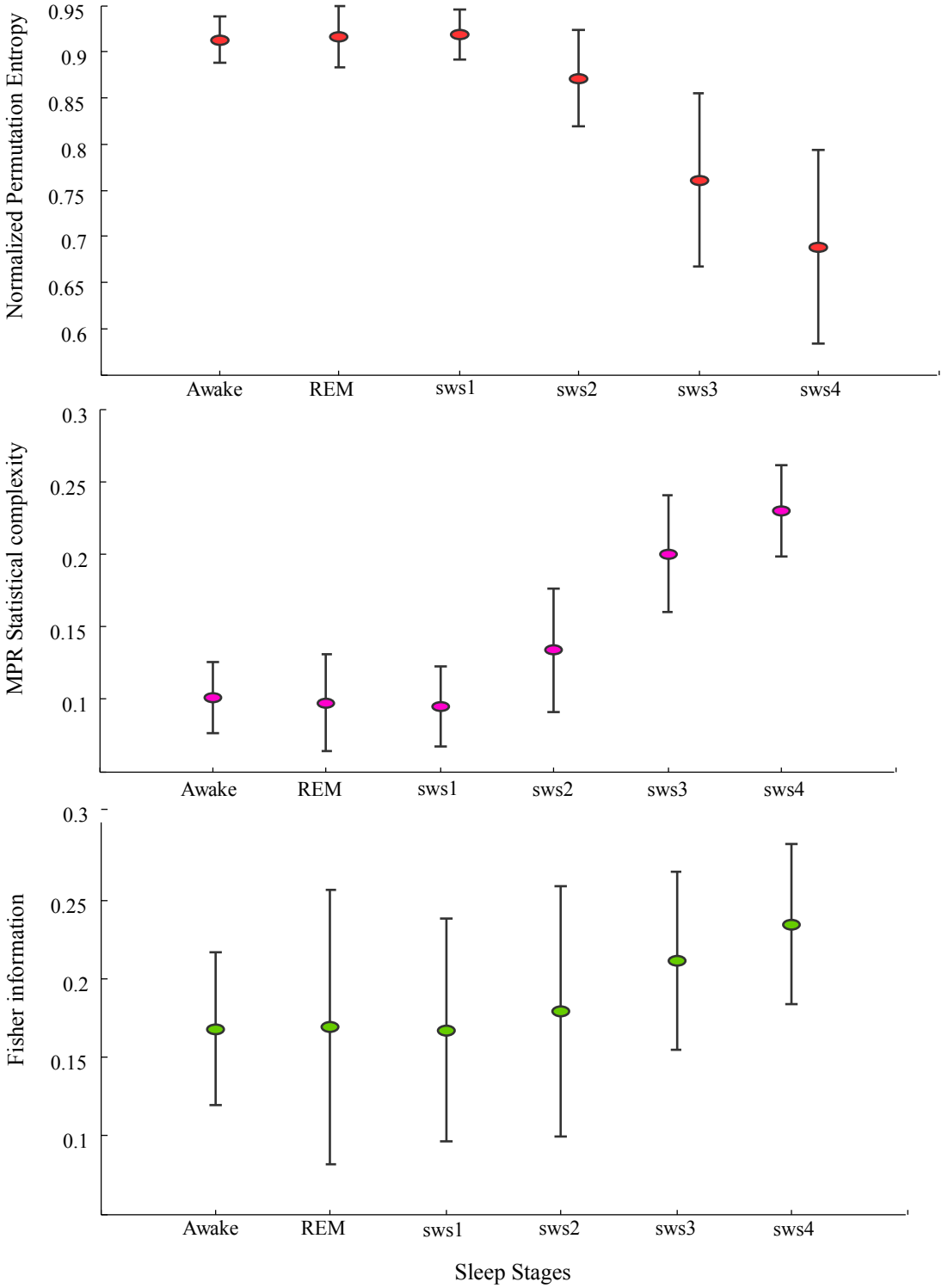
Table 5: Permutation Lempel-Ziv complexity values ($\mathcal{C}_{\mathcal{LZ}}$) for each subject analyzed in this work. The table shows the average value over all segments belonging to the same sleep stage. The parameter used were $D = 4$ and $\tau = 1$.

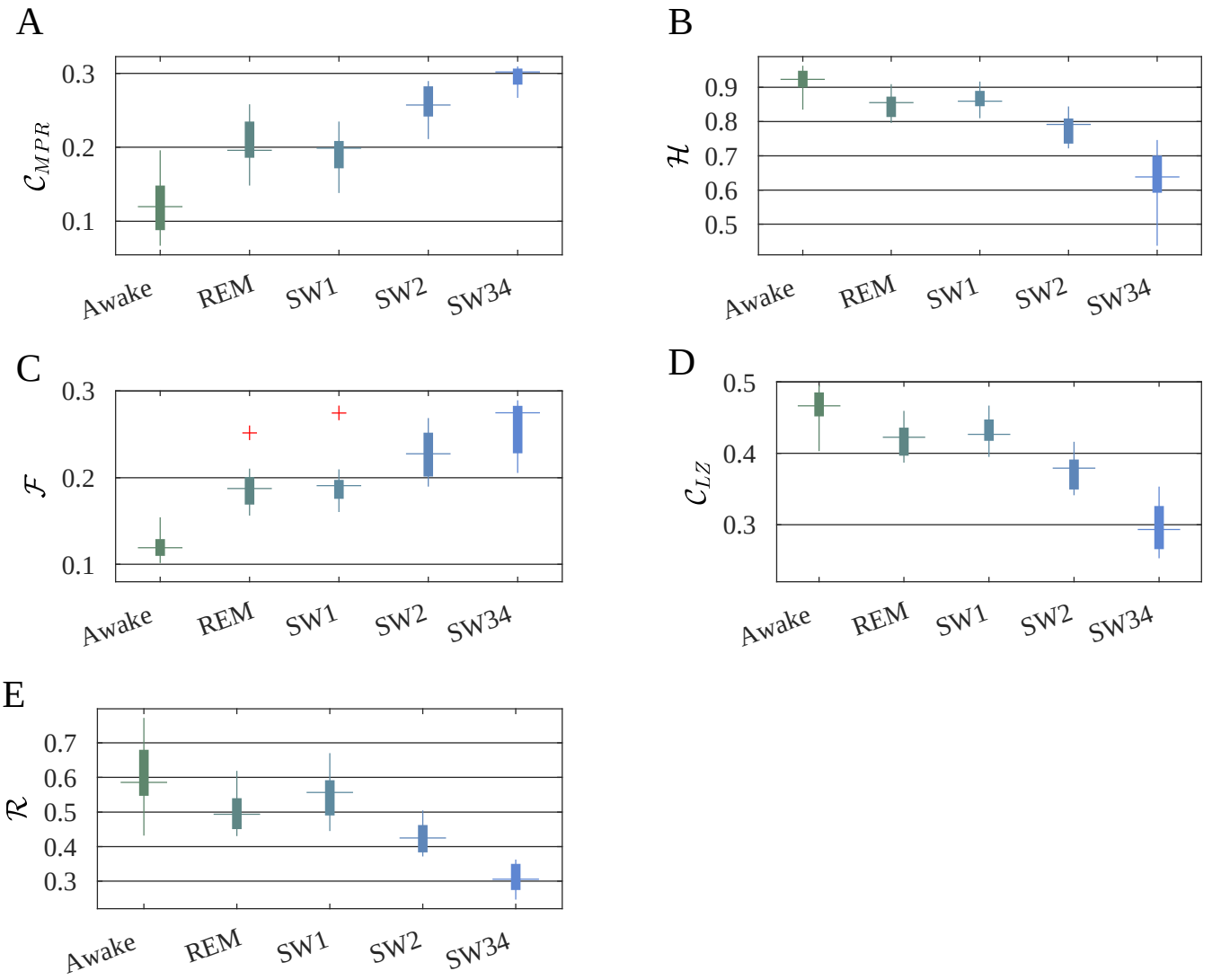
Subject	Awake	REM	SW1	SW2	SW34
1	0.5847	0.4303	0.4451	0.3714	0.2768
2	0.6856	0.5010	0.4887	0.4245	0.3265
3	0.4358	0.4649	0.5162	0.3833	0.2770
4	0.4318	0.4707	0.5194	0.3821	0.2810
5	0.5608	0.4352	0.5567	0.3721	0.2618
6	0.4834	0.4398	0.4717	0.3789	0.2655
7	0.6208	0.5683	0.5745	0.4573	0.2464
8	0.6176	0.5425	0.5958	0.4555	0.3521
9	0.5601	0.5623	0.6583	0.4721	0.3599
10	0.5062	0.4933	0.4883	0.4343	0.3494
11	0.6907	0.5093	0.5840	0.4717	0.3243
12	0.5858	0.4352	0.4899	0.4102	0.2646
13	0.6016	0.4661	0.5200	0.4102	0.3055
14	0.7017	0.5393	0.5910	0.4603	0.3572
15	0.5651	0.4545	0.5707	0.3857	0.2905
16	0.6783	0.5231	0.6349	0.4684	0.3309
17	0.7722	0.6188	0.6705	0.5045	0.3624

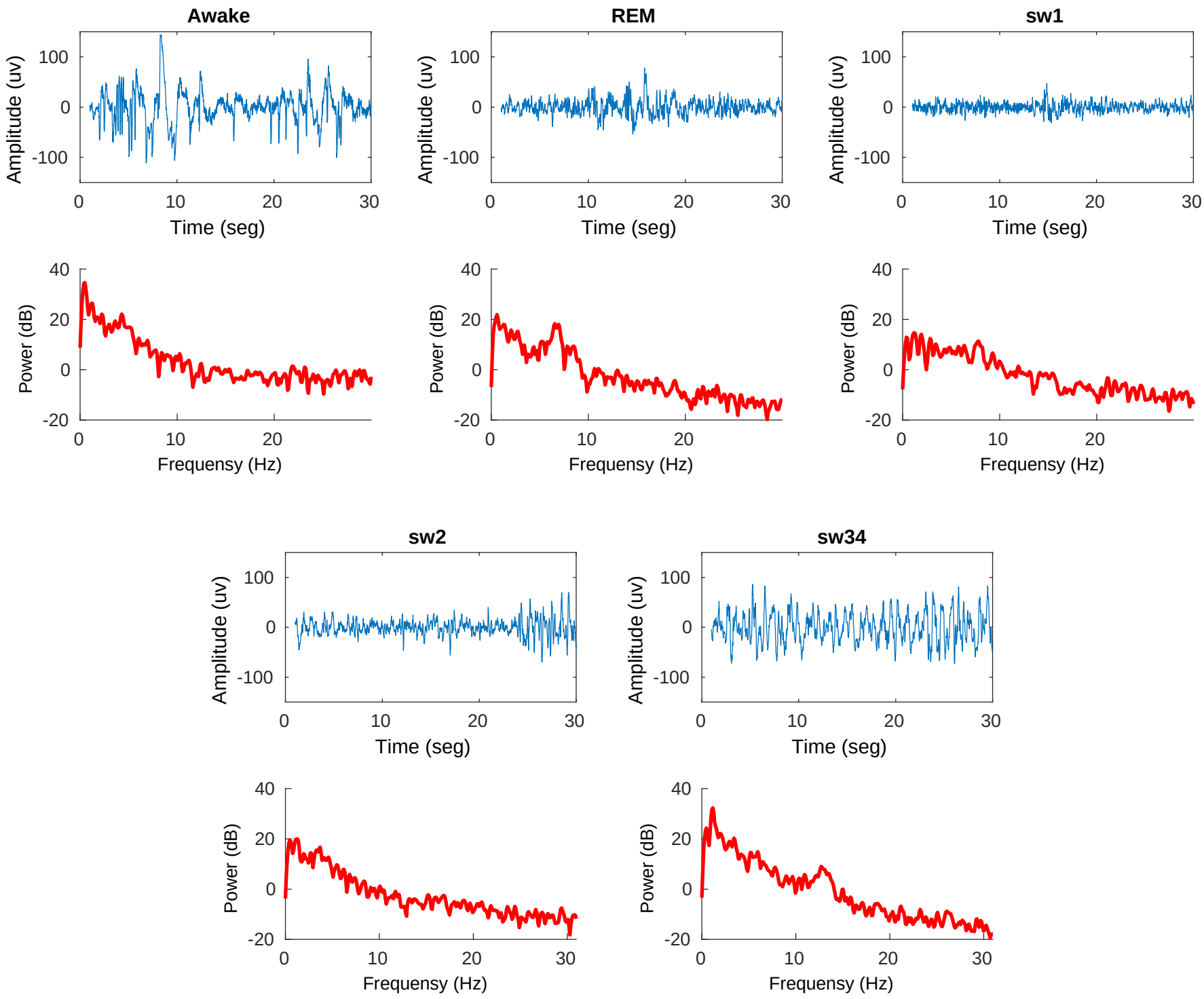
Table 6: Permutation Min-entropy values ($\mathcal{R}_\infty$) for each subject analyzed in this work. The table shows the average value over all segments belonging to the same sleep stage. The parameter used were $D = 4$ and $\tau = 1$.

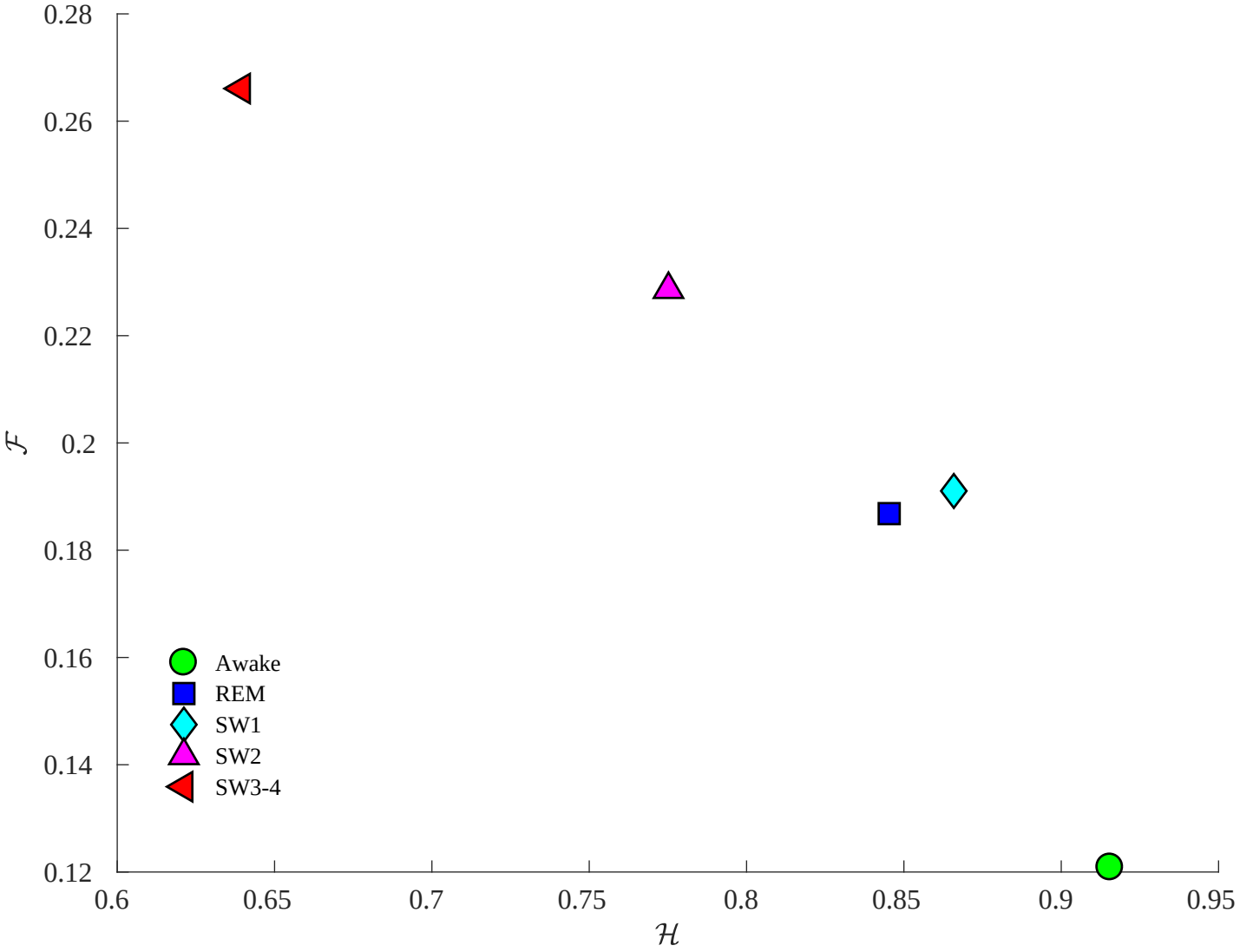
Measure	Awake	REM	SW1	SW2	SW34
$< \mathcal{H} >$	0.9155	0.8453	0.8659	0.7752	0.6395
$< \mathcal{C}_{MPR} >$	0.1231	0.2070	0.1908	0.2594	0.2959
$< \mathcal{F} >$	0.1209	0.1868	0.1910	0.2286	0.2661
$< \mathcal{C}_{LZ} >$	0.1209	0.1868	0.1910	0.2286	0.2661
$< \mathcal{R}_\infty >$	0.5931	0.4973	0.5515	0.4260	0.3077

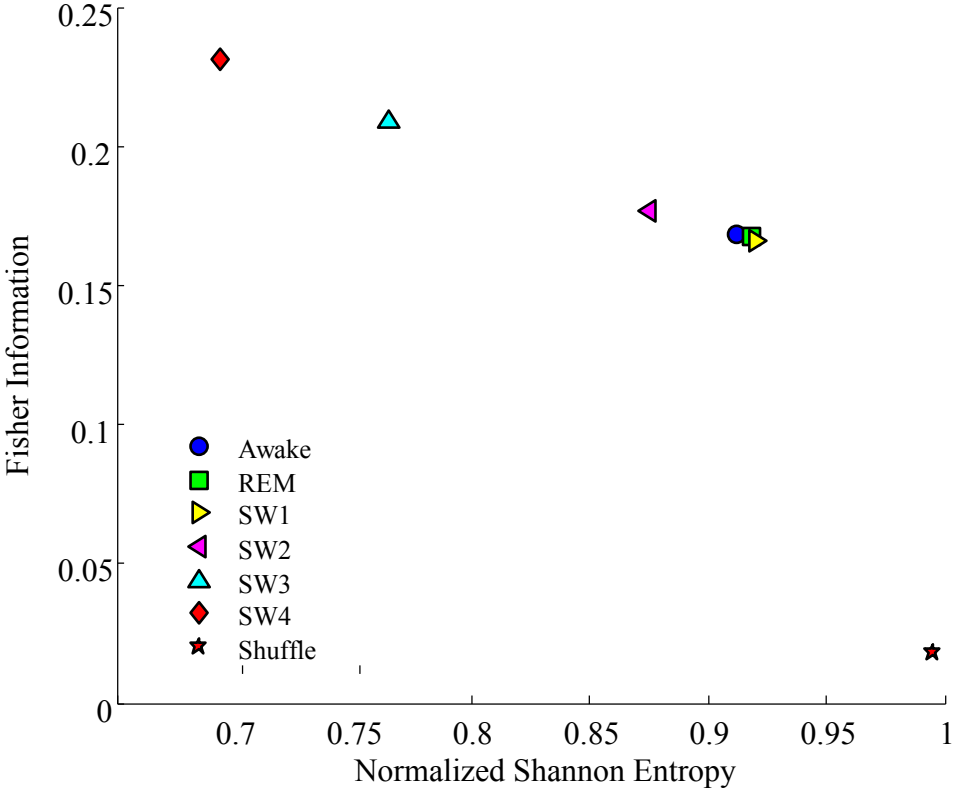
Table 7: Mean values for all information quantifiers applied in this word belong to each sleep state. They were calculated over all the segments and all the patients analyzed in this work. The parameter used were $D = 4$ and $\tau = 1$.

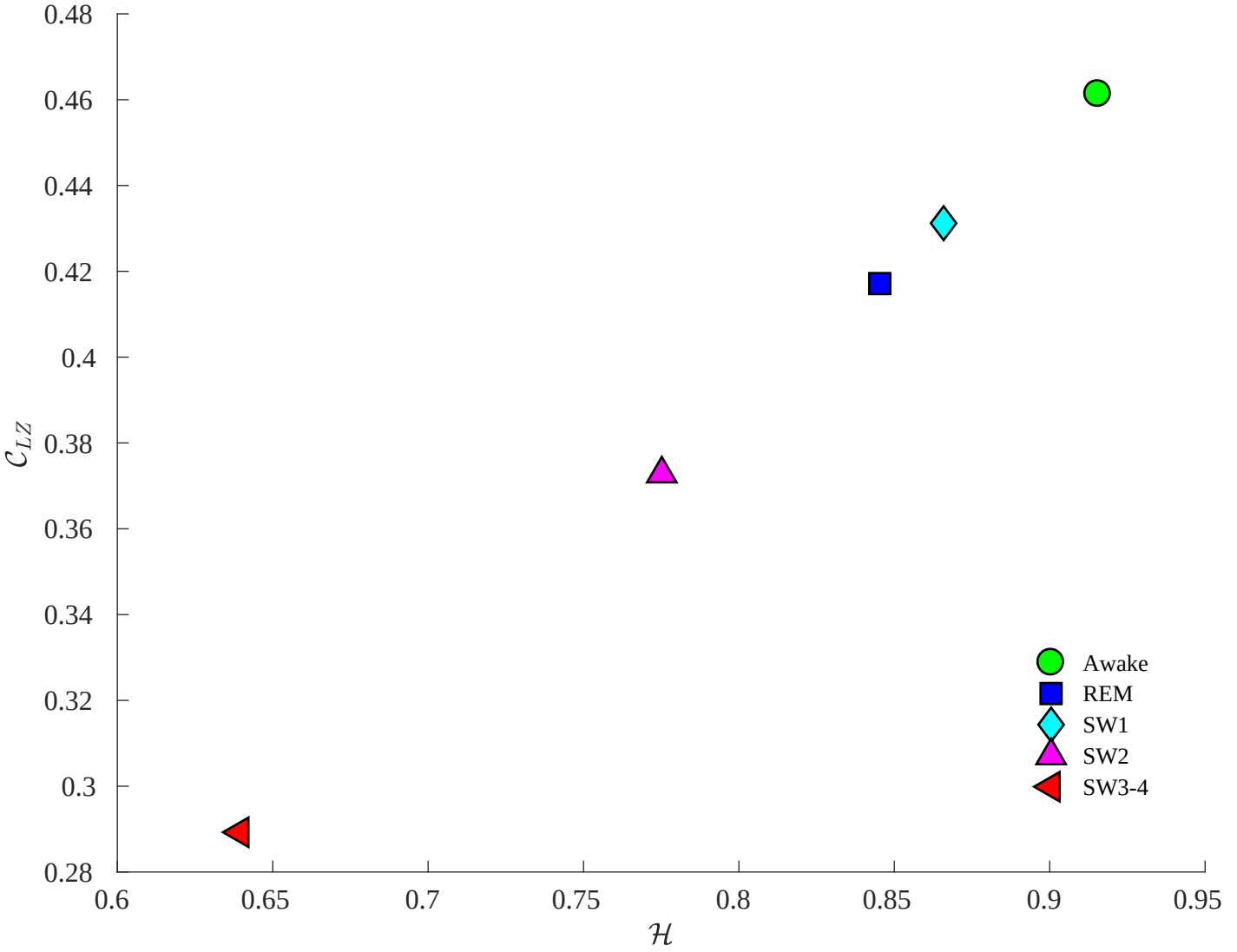

Normalized Permutation Entropy
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
Awake
REM
sws1
sws2
sws3
sws4
MPR Statistical complexity
0.3
0.25
0.2
0.15
0.1
Awake
REM
sws1
sws2
sws3
sws4
Fisher information
0.3
0.25
0.2
0.15
0.1
Awake
REM
sws1
sws2
sws3
sws4
Sleep Stages

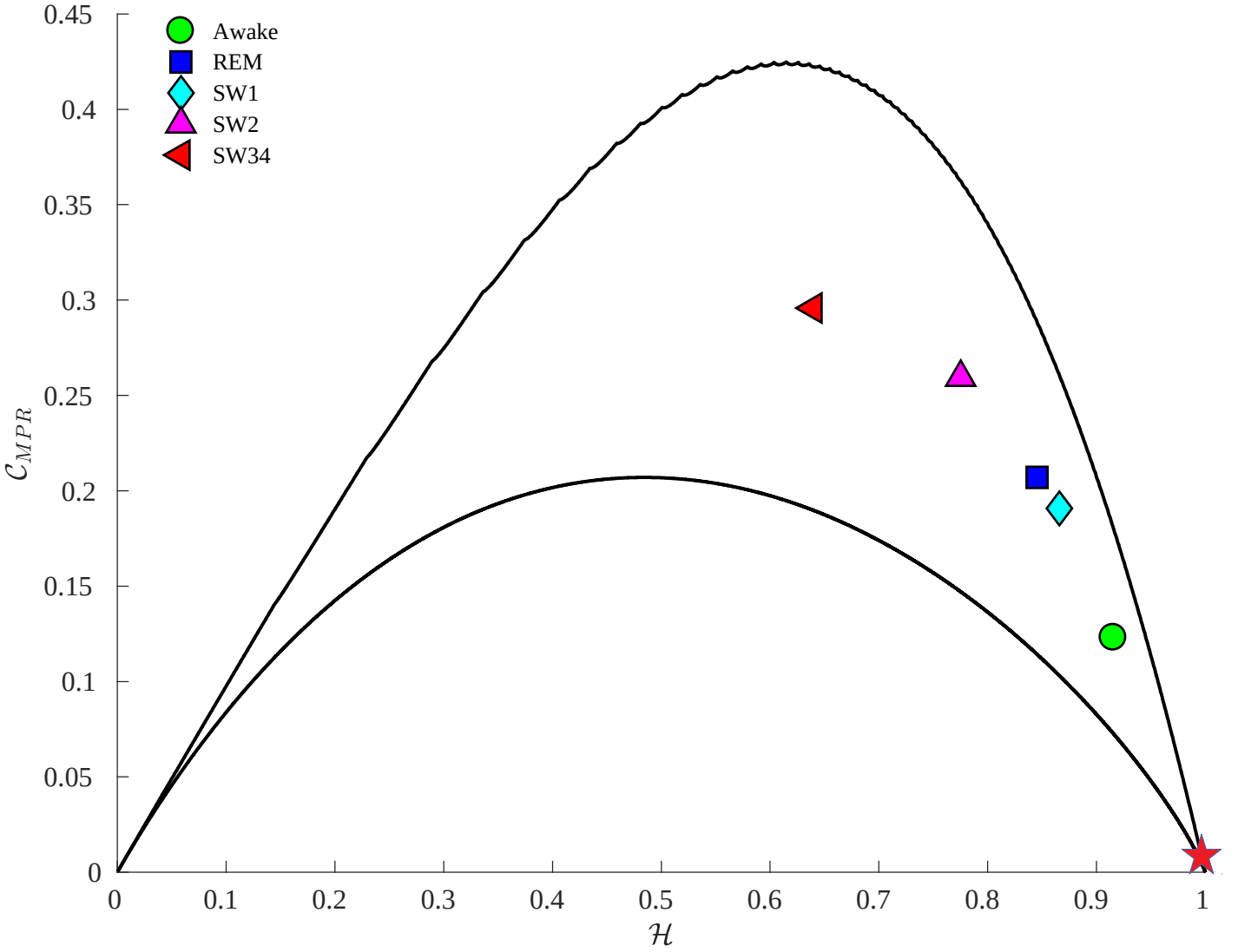

A
0.3
0.2
0.1
$\mathcal{C}_{MPR}$
Awake
REM
SW1
SW2
SW34
B
0.9
0.8
0.7
0.6
0.5
$\mathcal{H}$
Awake
REM
SW1
SW2
SW34
C
0.3
0.2
0.1
$\mathcal{F}$
Awake
REM
SW1
SW2
SW34
D
0.5
0.4
0.3
$\mathcal{C}_{LZ}$
Awake
REM
SW1
SW2
SW34
E
0.7
0.6
0.5
0.4
0.3
$\mathcal{R}$
Awake
REM
SW1
SW2
SW34

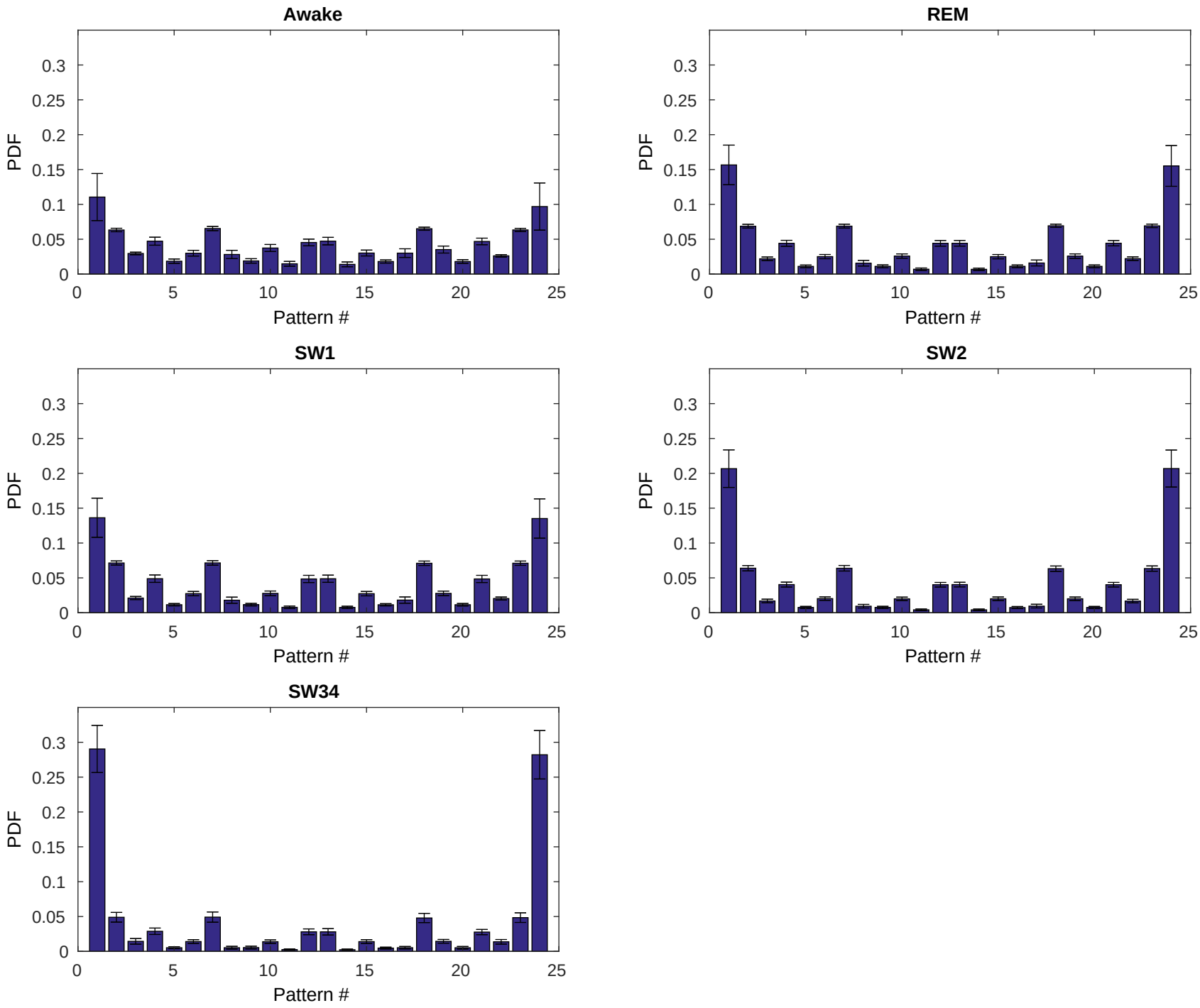

Awake
REM
sw1
sw2
sw34
Amplitude (uv)
Time (seg)
Power (dB)
Frequensy (Hz)
100
0
-100
40
20
-20
0
10
20
30

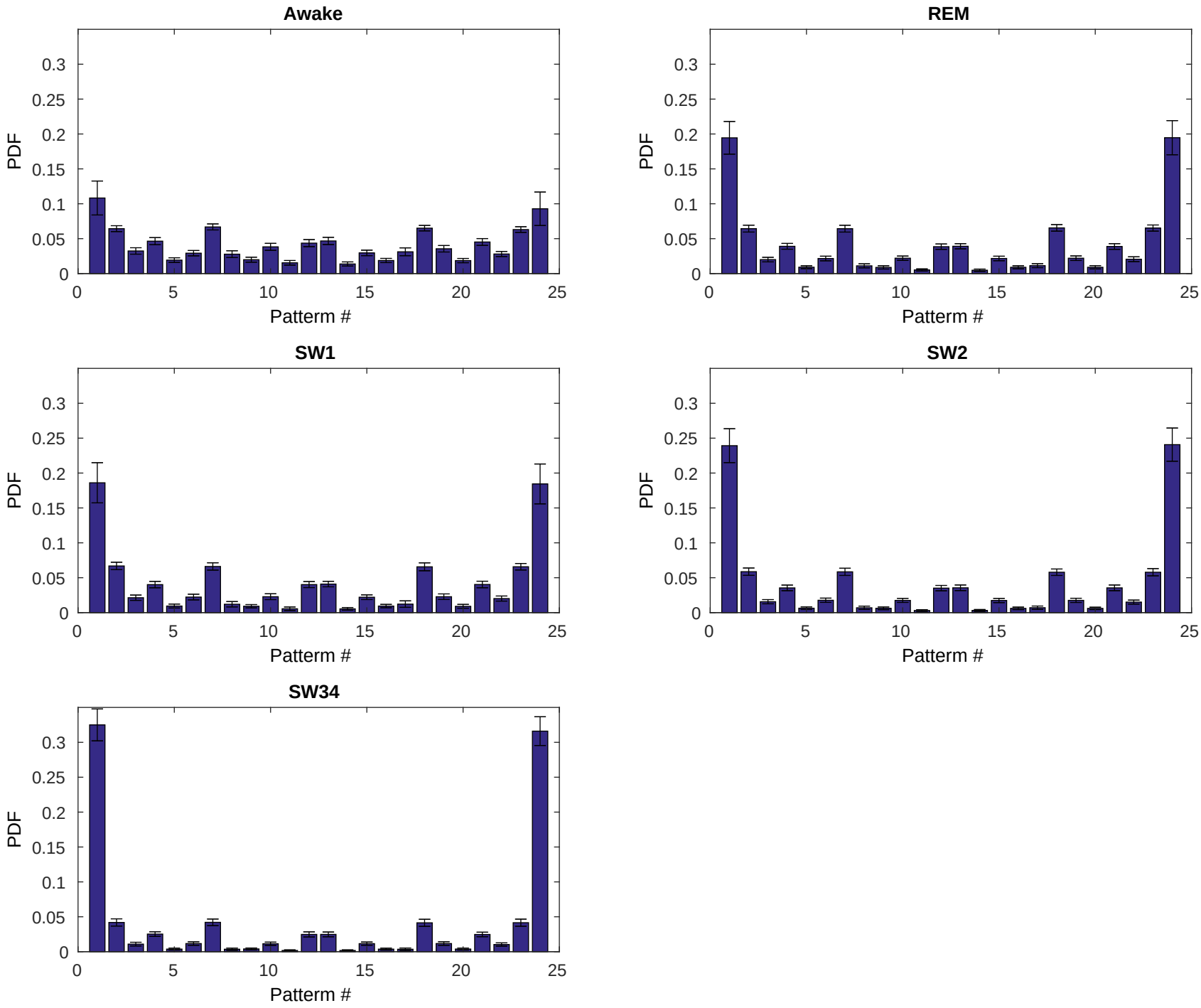

0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
$\mathcal{F}$
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
$\mathcal{H}$
Awake
REM
SW1
SW2
SW3-4


0.25
0.2
0.15
0.1
0.05
0
Fisher Information
Awake
REM
SW1
SW2
SW3
SW4
Shuffle
0.7
0.75
0.8
0.85
0.9
0.95
1
Normalized Shannon Entropy


0.48
0.46
0.44
0.42
0.4
0.38
0.36
0.34
0.32
0.3
0.28
$\mathcal{C}_{LZ}$
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
$\mathcal{H}$
Awake
REM
SW1
SW2
SW3-4


Awake
REM
SW1
SW2
SW34
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
$\mathcal{C}_{MPR}$
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
$\mathcal{H}$


Awake
REM
SW1
SW2
SW34
PDF
Pattern #
0
0.05
0.1
0.15
0.2
0.25
0.3
0
5
10
15
20
25


Awake
REM
SW1
SW2
SW34
PDF
Patterm #
0
0.05
0.1
0.15
0.2
0.25
0.3
0
5
10
15
20
25