

## Entropy-Based Algorithms in the Analysis of Biomedical Signals

Marta Borowska<sup>1</sup>

<sup>1</sup> Department of Materials and Biomedical Engineering, Faculty of Mechanical Engineering, Bialystok University of Technology, Poland

**Abstract.** Biomedical signals are frequently noisy and incomplete. They produce complex and high-dimensional data sets. In these mentioned cases, the results of traditional methods of signal processing can be skewed by noise or interference present in the signal. Information entropy, as a measure of disorder or uncertainty in the data, was introduced by Shannon. To date, many different types of entropy methods have appeared with many different application areas. The purpose of this paper is to present a short overview of some methods of entropy analysis and to discuss their suitability for use in the analysis of biomedical signals.

### Introduction

Over the last few years, entropy has emerged as an appropriate complexity measure in the study of time series from biological systems, such as the brain, the heart, the muscles, or the uterus. Biological systems are characterized by complex dynamics. The physiologic signal contains useful information, which allows scientists to view activity, redundancy, or noise information. Biomedical signal processing methods have become more sophisticated and extract information from signals that is not apparent from visual observation of the signal alone. In order to conclude that something is wrong or that the patients involved have a disease, further processing is necessary. The algorithms used in biomedical signal processing try to deal with the unique nature of physiologic signals such as nonstationarities, event detection, and disturbances.

The concept of entropy is used in many fields of science, including: information theory, statistical mechanics, chaos theory, neural networks, mathematical linguistics, and taxonomy. Considering different approaches, entropy can be used as a measure of disorder or uncertainty in a system. If

the degree of disorder is low, systems become organized. The ideal system is when everything is in complete order and the entropy value is zero. If the degree of disorder is high, then the system lacks stability. The change of entropy is inversely related to the change of free energy.

The notion of entropy was introduced in the nineteenth century by Rudolph Clausius in the context of thermodynamics (Clausius, 1850). It is connected with the second law of thermodynamics, which states that the change of the entropy in the volume element (at an infinitesimal change of state) is equal to the ratio of the heat state changes in temperature:

$$\Delta S = \int \frac{dQ}{T}$$

A simple system for which  $S$  increases is a dish containing gas. Initially, the molecules are in the mid-volume range and it is possible to locate them. At some point, the molecules begin to fill the entire volume of the vessel. In such a system, disorder increases and the molecules are not separated from the other parts of the vessel. The increased state of disorder decreases the amount of knowledge about the state of the system.

Boltzmann presented a statistical interpretation of thermodynamic entropy (Boltzmann, 1896). He linked the macroscopic property of entropy with the microscopic state of a system. The probabilistic study of entropy is presented in Boltzmann's law:

$$S = k \ln W$$

where  $S$  is entropy,  $k$  is Boltzmann's constant, and  $W$  is probability, determined by the configurational properties of a statistical system.

The third major form of entropy was introduced in communication theory through the work of Shannon et al. (1949). For more details regarding theoretical development of entropy, refer to Ben-Naim (2008).

The possibility of using entropy for determining the degree of disorder within a system resulted in the definition of various forms of entropy (Figure 1). In the next sections, some of the applications of these forms will be highlighted. Several previous reviews have summarized the implementation of information entropy in medicine (Holzinger et al., 2014; Zanin et al., 2012) and provide more detail about these applications.

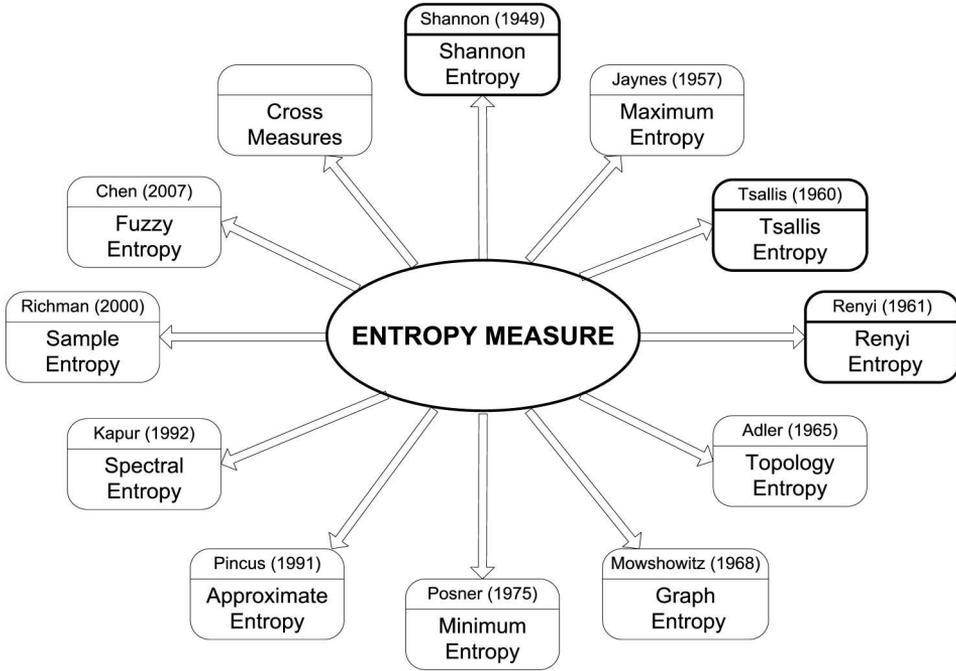


Figure 1. Different forms of entropy, presented in chronological order

## Entropy Methods

*Shannon entropy.* Information entropy was proposed by Shannon et al. (1949). Information entropy  $S$  of a random variable  $X$  that takes the values  $x_1, x_2, \dots, x_N$  is defined as:

$$S_{en} = \sum_{i=1}^n p(x_i) \log_a \frac{1}{p(x_i)} = - \sum_{i=1}^n p(x_i) \log_a p(x_i), \quad a > 1$$

where  $p(x_i)$  are probabilities of acceptance by the random variable  $X$  values  $x_i$ .

Shannon entropy is characterized by a degree of uncertainty associated with the occurrence of the result. A higher value of the entropy gives a more uncertain outcome and is more difficult to predict. Generally, the entropy of  $X$  is a measure of expected uncertainty obtained during the measurement of that variable.

Shannon entropy may be used globally, taking all data into account, or locally, around certain points. This measure can provide additional information about specific events, for example outliers or rare events.

*Renyi and Tsallis entropy.* Renyi entropy (Rényi, 1970) and Tsallis entropy (Tsallis et al., 1998) are generalizations of Shannon entropy that depend on a parameter. If  $p(x_i)$  is a probability distribution on a finite set, its Renyi entropy of order  $\alpha$  is defined as:

$$R_{en} = \frac{1}{(1 - \alpha)} \ln \sum_{i=1}^n p(x_i)^\alpha$$

where  $0 < \alpha < \infty$ . Renyi entropy approaches Shannon entropy as  $\alpha \rightarrow 1$ :

$$\lim_{\alpha \rightarrow 1} R_{en} = - \sum_{i=1}^n p(x_i) \ln p(x_i)$$

The most convincing uses of Renyi entropy seem to involve the limiting cases  $R_0 = \lim_{\alpha \rightarrow 0} R_\alpha$  and  $R_\infty = \lim_{\alpha \rightarrow \infty} R_\alpha$ . These are known as max-entropy and min-entropy, respectively.

Tsallis entropy is defined as:

$$T_{en} = \frac{1}{(\alpha - 1)} \left( 1 - \sum_{i=1}^n p(x_i)^\alpha \right)$$

Tsallis entropy has been used in statistical mechanics and in computer science, in pattern recognition, for example.

*Approximate entropy.* Pincus (1991) introduced approximate entropy, which is useable to quantify regularity in data without knowledge about a system. In the first step of approximate entropy, the vectors are constructed in a pseudo-phase:

$$\begin{aligned} y(i) &= [x(i), \dots, x(i + m - 1)] \\ y(j) &= [x(j), \dots, x(j + m - 1)] \end{aligned}$$

from a time series of  $N$  data points  $x(1), x(2), \dots, x(N)$ ,  $m$  is the embedding dimension.

Approximate entropy can be defined as:

$$App_{en} = \Phi^m(r) - \Phi^{m+1}(r)$$

where

$$\Phi^m(r) = \frac{1}{(N - m + 1)} \cdot \sum_{i=1}^{N-m+1} \ln C_i^m(r),$$

and

$$C_i^m(r) = \frac{(\text{number of } y(j) \text{ such that } d[y(i), y(j)] \leq r)}{N - m + 1}$$

and  $r$  is a filtering level, while  $d$  represents distances of points. Usually,  $r = 20\%$  of the standard deviation of the amplitude values and  $m = 2$ .

$App_{en}$  was introduced to avoid problems in the finite length of a real time series and in the need to distinguish the nature of the generating systems (Fusheng et al., 2001; Oczeretko et al., 2005). Larger values of approximate entropy correspond to more complexity and irregularity in the data.

The disadvantages of  $App_{en}$  are that it is strongly dependent on the record length and is often lower than expected for short records. Another disadvantage is that  $App_{en}$  lacks relative consistency (Richman et al., 2000).

*Sample entropy.* To overcome the disadvantages of  $App_{en}$ , sample entropy ( $Samp_{en}$ ) was proposed to replace  $App_{en}$  by excluding self-matches (Richman et al., 2000), thereby reducing the computing time by one-half in comparison with  $App_{en}$ . The advantage of  $Samp_{en}$  is that it is largely independent of record length and displays relative consistency.

In order to compute the  $Samp_{en}$ , we must define  $\Phi^m(r)$  and  $\Phi^{m+1}(r)$ . The probability  $\Phi^m(r)$  that two sequences match for  $m$  points is computed by counting the average number of vector pairs for which the distance is lower than the tolerance  $r$ :

$$\Phi^m(r) = \frac{1}{N - m} \cdot \sum_{i=1}^{N-m} C_i^m(r)$$

Similarly,  $\Phi^{m+1}(r)$  is defined for an embedding dimension of  $m + 1$ . The  $Samp_{en}$  is calculated as:

$$Samp_{en} = \ln \frac{\Phi^m(r)}{\Phi^{m+1}(r)}$$

Sample entropy is independent of the length of recording and displays relative consistency under various conditions.

*Fuzzy entropy.* Fuzzy entropy ( $Fuzzy_{en}$ ), a measure of time series regularity, was proposed by Chen et al. (2007). Similar to the two existing related measures,  $App_{en}$  and  $Samp_{en}$ ,  $Fuzzy_{en}$  is the negative natural logarithm of the conditional probability that two similar vectors for  $m$  points remain similar for the next  $m + 1$  points. Importing the concept of fuzzy sets, the vectors' similarity is fuzzily defined in  $Fuzzy_{en}$  on the basis of ex-

ponential functions and their shapes. Besides possessing the good properties of *Samp<sub>en</sub>* superior to *App<sub>en</sub>*, *Fuzzy<sub>en</sub>* also succeeds in giving the entropy definition in the case of small parameters. The method can also be applied to noisy physiological signals with relatively short datasets.

In the first step of fuzzy entropy, the vectors constructed in the pseudo-phase are transformed into:

$$\begin{aligned}\bar{y}(i) &= [x(i) - \bar{x}(i), \dots, x(i + m - 1) - \bar{x}(i)] \\ \bar{y}(j) &= [x(j) - \bar{x}(j), \dots, x(j + m - 1) - \bar{x}(j)]\end{aligned}$$

where  $\bar{x}(i)$  is the mean value of  $y(i)$ :

$$\bar{x}(i) = \sum_{k=0}^{m-1} \frac{x(i+k)}{m}$$

In the next step, the fuzzy membership matrix is defined as:

$$D_{i,j}^m = \mu \left( d \left( \bar{y}_i^m, \bar{y}_j^m \right) \right)$$

with the fuzzy membership function:

$$\mu(x) = e^{-\left(\frac{x}{r}\right)^n}$$

Finally, fuzzy entropy is defined as:

$$Fuzzy_{en} = \ln \frac{\Phi^m}{\Phi^{m+1}}$$

where

$$\Phi^m = \frac{1}{N-m} \sum_{i=1}^{N-m} \sum_{j=1, j \neq i}^{N-m} \frac{D_{i,j}^m}{N-m-1}$$

*Fuzzy<sub>en</sub>* uses a fuzzy relationship function, which leads to a weaker influence of the threshold parameter choice.

Fuzzy entropy led to fuzzy measure entropy, which introduced a distinction between local entropy and global entropy (Liu et al., 2013).

*Spectral entropy.* Entropy can be interpreted as a measure of uncertainty about an event at frequency  $f$ . Spectral entropy (Kapur et al., 1992) uses the Fourier transformation method, in which the power spectral density (PSD)  $\hat{P}(f)$  can be obtained. The PSD represents the distribution of power as a function of frequency. So, normalization of  $\hat{P}(f)$  yields a proba-

bility density function. Using the definition of Shannon's entropy, spectral entropy can be defined as:

$$Sp_{en} = \sum_{i=f_l}^{f_h} p_i \log(p_i)$$

where  $[f_l, f_h]$  is the frequency band.

Spectral entropy is usually normalized  $Sp_{en}/\log N_f$ , where  $N_f$  is the number of frequency components in the range  $[f_l, f_h]$ .

*Permutation entropy.* The complexity of time series can also be quantified by using symbolic dynamics. The permutation method was proposed by Bandt et al. (2002) to map a continuous time series onto a symbolic sequence. The result of the statistics of the symbolic sequences is called permutation entropy.

Generally, each time series is associated with a probability distribution  $\Pi$ , whose elements  $\pi_i$  are the frequencies connected with  $i$  possible permutation patterns, where  $i = 1, \dots, m!$ . Permutation entropy ( $P_{en}$ ) is defined as:

$$P_{en} = - \sum_{i=1}^{m!} \pi_i \ln \pi_i$$

Normalized permutation entropy can be defined as:

$$P_{en_{norm}} = - \frac{1}{\log_2 m!} \sum_{i=1}^{m!} \pi_i \log_2 \pi_i$$

A related information measure is normalized Kullback-Leiber entropy (Frank et al., 2006), which is defined as:

$$KL_{en} = 1 - P_{en_{norm}}$$

$KL_{en}$  quantifies the distance between the ordinal probability distribution and uniform distribution.

Permutation entropy is the temporal information contained in the time series and has the qualities of simplicity, robustness, and low computational cost.

## **Biomedical Applications**

The concept of entropy has been used in many diverse applications within the biomedical domain. The aforementioned entropy measures have been applied in physiological time series, for example, with electroencephalo-

graphic signals (EEG), electrocardiographic signals (ECG), electromyographic signals (EMG), and electrohysterographic signals (EHG).

Cornforth et al. (2013) used Renyi entropy to detect cardiac autonomic neuropathy (CAN) in diabetes patients. Multi-scale entropy was applied to the ECG recording. They obtained significant differences between controls and early CAN. Their results suggest that it is easy to distinguish between people with early CAN and controls. This is an important finding and gives hope for a simple and relatively non-invasive test for people developing this disease.

Graff et al. (2012) investigated the usefulness of entropy measures, namely approximate entropy, sample entropy, fuzzy entropy, and permutation entropy, calculated for short ECG series in distinguishing healthy subjects from patients with congestive heart failure (CHF). They found that with a reduction of the data set length up to 250 RR intervals, values of entropy can remain significantly different in patients with CHF compared to healthy individuals.

Akareddy et al. (2013) improved approximate entropy and used it in EEG signal classification for epileptic seizure detection. Their results showed that the accuracy of the proposed method was better than the existing method for epileptic seizure identification.

Sharma et al. (2015) used different entropy measures, namely: average Shannon entropy, average Renyi's entropy, average approximate entropy, average sample entropy, and average phase entropies on intrinsic mode functions for the automated identification of focal EEG signals. Their results also confirmed the usefulness of entropy measures to differentiate focal and non-focal EEG signals.

Ferlazzo et al. (2014) evaluated permutation entropy extracted from EEG recordings during both interictal and ictal phases in patients with typical absences (TAs) and healthy subjects. In patients with TAs, a higher randomness in fronto-temporal areas associated with high permutation entropy levels and a lower randomness in posterior areas associated with low permutation entropy levels occurred. Based on this, they claimed that permutation entropy seemed to be a useful tool to disclose abnormalities of cerebral electric activity.

Avilov et al. (2012) also used permutation entropy on EEG recordings but to detect epileptic seizures. They showed that the PE of a sick patient is twice larger than in healthy patients, and that the PE increases for time ranges where the appearance of epileptiform activity takes place. Li et al. (2014) were the next researchers who used permutation entropy to measure the changes in EEG signals during absence seizures. Their experi-

mental results showed that the mean value of permutation entropy decreases from the seizure-free to the seizure phase and provides evidence that absence epilepsy can be effectively distinguished.

Liang et al. (2015) compared twelve entropy indices, namely response entropy, state entropy, three wavelet entropy measures (Shannon, Tsallis, and Renyi), Hilbert-Huang spectral entropy, approximate entropy, sample entropy, fuzzy entropy, and three permutation entropy measures (Shannon, Tsallis and Renyi) in monitoring depth of anesthesia and detecting burst suppression. They found that Renyi permutation entropy performed best in tracking EEG changes associated with different anesthetic states, and that approximate entropy and sample entropy performed best in detecting burst suppression.

Zhang et al. (2012) presented the possibility of using sample entropy for muscle activity onset detection by processing surface EMG against ECG artifacts. A sample entropy threshold was used for detection of muscle activity and performed significantly better.

Diab et al. (2014) investigated the performance of four non-linearity detection methods: statistics (Time reversibility), predictability (Sample Entropy, Delay Vector Variance), and chaos theory (Lyapunov Exponents). Applied to real uterine EMG signals, they were used to distinguish between pregnancy and labor contraction bursts. Their results confirmed the sensitivity of sample entropy.

Alamedine et al. (2014) proposed several selection methods in order to choose the best parameters to classify contractions in uterine EHG signals for the detection of preterm labor. One of them was sample entropy, which was potentially most useful to discriminate between pregnancy and labor contractions.

Garcia-Gonzalez et al. (2013) also investigated the differences in the contractions generated by women that chose to have a vaginal delivery as opposed to those who elected to have caesarean section. They used sample entropy to calculate the irregularity of manually selected contractions of the EHG time series and confirmed that sample entropy could provide an index to evaluate the quality of the active phase of labor at term.

## **Conclusions**

Entropy measures have been used in analyzing short, sparse, and noisy medical time series. These methods solve important problems in time series analysis, such as: classifying different dynamics, identifying missing points in

time series, predicting appearance events, determining time scales, quantifying similarities between time series, or identifying directionality and causality. Furthermore, entropy measures can be extended to multi-variety and multi-scale systems (Humeau-Heurtier, 2015). Besides the advantages of using entropy measures, there are a lot of unsolved problems. One of them is how to use entropy measures for the classification of pathological and non-pathological data. In studies concerning biomedical data, a common aim is to distinguish between two states of a system. There is little knowledge concerning how to solve the problems of classification and selecting appropriate data ranges to use. This is directly related to the choice of parameter selection concerning the threshold value  $r$ , which has an influence on the value of entropy. There are different parameter sets used, but until now all possible combinations have not been tested and no consensus has been reached. This parameter is dependent on the data and its type.

### Acknowledgments

This paper was supported by statutory funds S/WM/1/2014 from the Białystok University of Technology.

### REFERENCES

- Akareddy, S. M., & Kulkarni, P. K. (2013). EEG signal classification for epilepsy seizure detection using improved approximate entropy. *International Journal of Public Health Science (IJPHS)*, 2(1), 23–32.
- Alamedine, D., Diab, A., Muszynski, C., Karlsson, B., Khalil, M., & Marque, C. (2014). Selection algorithm for parameters to characterize uterine EHG signals for the detection of preterm labor. *Signal, Image and Video Processing*, 8(6), 1169–1178.
- Avilov, O., Popov, A., Kanaikin, O., & Kyselova, O. (2012). Permutation Entropy Analysis of Electroencephalogram. *Signal*, 100, 200.
- Bandt, C., & Pompe, B. (2002). Permutation entropy: a natural complexity measure for time series. *Physical review letters*, 88(17), 174102.
- Ben-Naim, A. (2008). *A Farewell to Entropy: Statistical Thermodynamics Based on Information*. S. World Scientific.
- Boltzmann, L. (1896). *Vorlesungen über Gastheorie* (Vol. 1). Leipzig: J. A. Barth.
- Chen, W., Wang, Z., Xie, H., & Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 15(2), 266–272.
- Clausius, R. (1850). On the motive power of heat, and on the laws which can be deduced from it for the theory of heat. *Poggendorff's Annalen Der Physick*, LXXIX, 368, 500.

- Cornforth, D. J., Tarvainen, M. P., & Jelinek, H. F. (2013, July). Using renyi entropy to detect early cardiac autonomic neuropathy. In *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE* (pp. 5562–5565).
- Diab, A., Hassan, M., Marque, C., & Karlsson, B. (2014). Performance analysis of four nonlinearity analysis methods using a model with variable complexity and application to uterine EMG signals. *Medical engineering & physics*, 36(6), 761–767.
- Ferlazzo, E., Mammone, N., Cianci, V., Gasparini, S., Gambardella, A., Labate, A., Aguglia, U., et al. (2014). Permutation entropy of scalp EEG: A tool to investigate epilepsies: Suggestions from absence epilepsies. *Clinical Neurophysiology*, 125(1), 13–20.
- Frank, B., Pompe, B., Schneider, U., & Hoyer, D. (2006). Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. *Medical and Biological Engineering and Computing*, 44(3), 179–187.
- Fusheng, Y., Bo, H., & Qingyu, T. (2001). Approximate entropy and its application to biosignal analysis. *Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling*, 2, 72–91.
- Garcia-Gonzalez, M. T., Charleston-Villalobos, S., Vargas-Garcia, C., Gonzalez-Camarena, R., & Aljama-Corrales, T. (2013, July). Characterization of EHG contractions at term labor by nonlinear analysis. In *Engineering in Medicine and Biology Society (EMBC), 35th Annual International Conference of the IEEE* (pp. 7432–7435).
- Graff, B., Graff, G., & Kaczkowska, A. (2012). Entropy measures of heart rate variability for short ECG datasets in patients with congestive heart failure. *Acta Physica Polonica B Proc. Suppl*, 5, 153–158.
- Holzinger, A., Hörtenhuber, M., Mayer, C., Bachler, M., Wassertheurer, S., Pinho, A. J., & Koslicki, D. (2014). On entropy-based data mining. In *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics* (pp. 209–226). Berlin Heidelberg: Springer.
- Humeau-Heurtier, A. (2015). The Multiscale Entropy Algorithm and Its Variants: A Review. *Entropy*, 17(5), 3110–3123.
- Kapur, J. N., & Kesavan, H. K. (1992). *Entropy optimization principles with applications*. New York: Academic Press.
- Li, J., Yan, J., Liu, X., & Ouyang, G. (2014). Using permutation entropy to measure the changes in EEG signals during absence seizures. *Entropy*, 16(6), 3049–3061.
- Liang, Z., Wang, Y., Sun, X., Li, D., Voss, L. J., Sleight, J. W., Li, X., et al. (2015). EEG entropy measures in anesthesia. *Frontiers in computational neuroscience*, 9.

- Liu, C., Li, K., Zhao, L., Liu, F., Zheng, D., Liu, C., & Liu, S. (2013). Analysis of heart rate variability using fuzzy measure entropy. *Computers in Biology and Medicine*, 43(2), 100–108.
- Oczeretko, E., Kitlas, A., Swiatecka, J., Borowska, M., & Laudanski, T. (2005). Nonlinear dynamics in uterine contractions analysis. In G. Losa, D. Merlini, T. Nonnemacher, & E. Weibel (Eds.), *Fractals in Biology and Medicine* (pp. 215–222). Basel: Birkhäuser Verlag.
- Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6), 2297–2301.
- Rényi, A. (1970). Probability theory. In *North-Holland Series in Applied Mathematics and Mechanics* (Vol. 10).
- Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6), H2039–H2049.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.
- Sharma, R., Pachori, R. B., & Acharya, U. R. (2015). Application of Entropy Measures on Intrinsic Mode Functions for the Automated Identification of Focal Electroencephalogram Signals. *Entropy*, 17(2), 669–691.
- Tsallis, C., Mendes, R., & Plastino, A. R. (1998). The role of constraints within generalized nonextensive statistics. *Physica A: Statistical Mechanics and its Applications*, 261(3), 534–554.
- Zanin, M., Zunino, L., Rosso, O. A., & Papo, D. (2012). Permutation entropy and its main biomedical and econophysics applications: a review. *Entropy*, 14(8), 1553–1577.
- Zhang, X., & Zhou, P. (2012). Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes. *Journal of Electromyography and Kinesiology*, 22(6), 901–907.