

Review

# The Entropy Universe

Maria Ribeiro <sup>1,2,\*</sup> , Teresa Henriques <sup>3,4</sup> , Luísa Castro <sup>3</sup> , André Souto <sup>5,6,7</sup> , Luís Antunes <sup>1,2</sup> ,  
Cristina Costa-Santos <sup>3,4</sup>  and Andreia Teixeira <sup>3,4,8</sup> 

- <sup>1</sup> Institute for Systems and Computer Engineering, Technology and Science (INESC-TEC), 4200-465 Porto, Portugal; lfa@fc.up.pt
  - <sup>2</sup> Computer Science Department, Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal
  - <sup>3</sup> Centre for Health Technology and Services Research (CINTESIS), Faculty of Medicine University of Porto, 4200-450 Porto, Portugal; teresasarhen@med.up.pt (T.H.); luisacastro@med.up.pt (L.C.); csantos@med.up.pt (C.C.-S.); andreiasofiat@med.up.pt (A.T.)
  - <sup>4</sup> Department of Community Medicine, Information and Health Decision Sciences-MEDCIDS, Faculty of Medicine, University of Porto, 4200-450 Porto, Portugal
  - <sup>5</sup> LASIGE, Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisboa, Portugal; ansouto@fc.ul.pt
  - <sup>6</sup> Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisboa, Portugal
  - <sup>7</sup> Instituto de Telecomunicações, 1049-001 Lisboa, Portugal
  - <sup>8</sup> Instituto Politécnico de Viana do Castelo, 4900-347 Viana do Castelo, Portugal
- \* Correspondence: maria.ribeiro@inesctec.pt

**Abstract:** About 160 years ago, the concept of entropy was introduced in thermodynamics by Rudolf Clausius. Since then, it has been continually extended, interpreted, and applied by researchers in many scientific fields, such as general physics, information theory, chaos theory, data mining, and mathematical linguistics. This paper presents *The Entropy Universe*, which aims to review the many variants of entropies applied to time-series. The purpose is to answer research questions such as: How did each entropy emerge? What is the mathematical definition of each variant of entropy? How are entropies related to each other? What are the most applied scientific fields for each entropy? We describe in-depth the relationship between the most applied entropies in time-series for different scientific fields, establishing bases for researchers to properly choose the variant of entropy most suitable for their data. The number of citations over the past sixteen years of each paper proposing a new entropy was also accessed. The Shannon/differential, the Tsallis, the sample, the permutation, and the approximate entropies were the most cited ones. Based on the ten research areas with the most significant number of records obtained in the Web of Science and Scopus, the areas in which the entropies are more applied are computer science, physics, mathematics, and engineering. The universe of entropies is growing each day, either due to the introducing new variants either due to novel applications. Knowing each entropy's strengths and of limitations is essential to ensure the proper improvement of this research field.

**Keywords:** entropy measures; information theory; time-series; application areas



**Citation:** Ribeiro, M.; Henriques, T.; Castro, L.; Souto, A.; Antunes, L.; Costa-Santos, C.; Teixeira, A. The Entropy Universe. *Entropy* **2021**, *23*, 222. <https://doi.org/10.3390/e23020222>

Academic Editor: Miguel Rubi  
Received: 15 January 2021  
Accepted: 8 February 2021  
Published: 11 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Despite its long history, to many, the term entropy still appears not to be easily understood. Initially, the concept was applied to thermodynamics, but it is becoming more popular in other fields. The concept of entropy has a complex history. It has been the subject of diverse reconstructions and interpretations making it very confusing and difficult to understand, implement, and interpret.

Up to the present, many different types of entropy methods have emerged, with a large number of different purposes and possible application areas. Various descriptions and meanings of entropy are provided in the scientific community, bewildering researchers, students, and professors [1–3]. The miscellany in the research papers by the widespread use of entropy in many disciplines leads to many contradictions and misconceptions involving

entropy, summarized in Von Neumann’s sentence, “Whoever uses the term ‘entropy’ in a discussion always wins since no one knows what entropy really is, so in a debate, one always has the advantage” [4,5].

Researchers have already studied entropy measurement problems, but there are still several questions to answer. In 1983, Batten [6] discussed entropy theoretical ideas that have led to the suggested nexus between the physicists’ entropy concept and measures of uncertainty or information. Amigó et al. [7] presented a review of only generalized entropies, which from a mathematical point of view, are non-negative functions defined on probability distributions that satisfy the first three Shannon–Khinchin axioms [8]. In 2019, Namdari and Zhaojun [9] reviewed the entropy concept for uncertainty quantification of stochastic processes of lithium-ion battery capacity data. However, those works do not present an in-depth analysis of how entropies are related to each other.

Several researchers, such as the ones of reference [10–19], consider that entropy is an essential tool for time-series analysis and apply this measure in several research areas. Nevertheless, the choice of a specific entropy is made in an isolated and unclear way. There is no complete study on the areas of entropies application in the literature to the best of our knowledge. We believe that a study of the importance and application of each entropy in time-series will help researchers understand and choose the most appropriate measure for their problem.

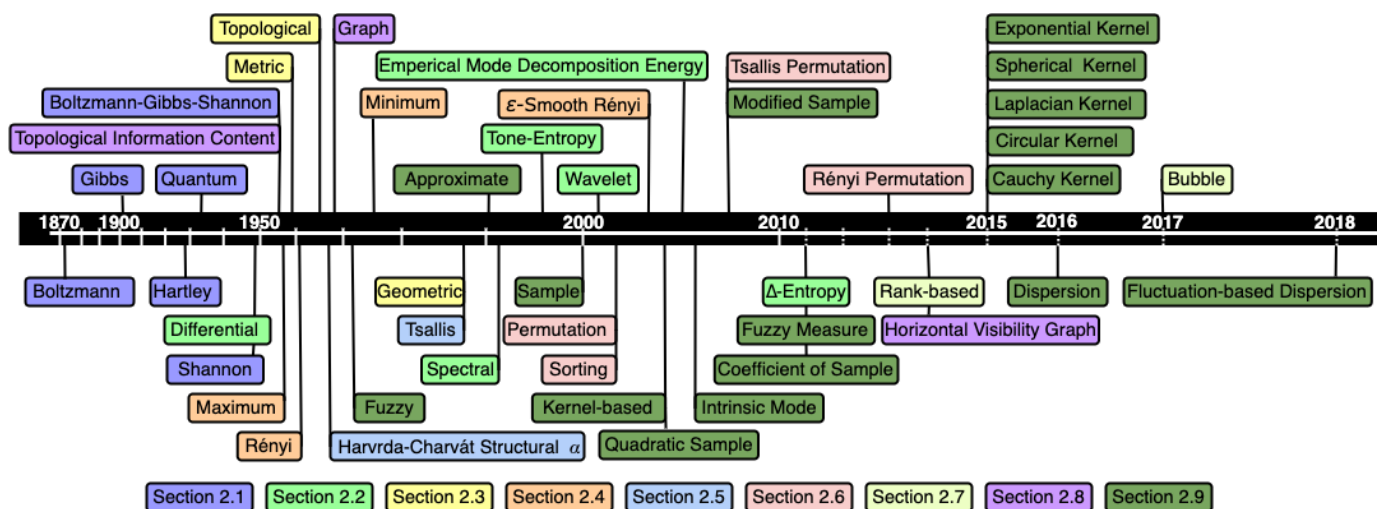
Hence, considering entropies applied to time-series, the focus of our work is to demystify and clarify the concept of entropy, describing:

1. How the different concepts of entropy arose.
2. The mathematical definitions of each entropy.
3. How the entropies are related to each other.
4. Which are the areas of application of each entropy and their impact in the scientific community.

## 2. Building the Universe of Entropies

In this section, we describe in detail how we built our universe of entropies. We describe the entropies, mathematical definitions, the respective origin, and the relationship between each other. We also address some issues with the concept of entropy and, its extension to the study of continuous random variables.

Figure 1 is the timeline (in logarithmic scale) of the universe of entropies covered in this paper. Although we have referred to some entropies in more than one section, in the Figure 1, the colors refer to the section where each entropy has been defined.



**Figure 1.** Timeline of the universe of entropies discussed in this paper. Timeline in logarithmic scale and colors refer to the section in which each entropy is defined.

Boltzmann, Gibbs, Hartley, quantum, Shannon and Boltzmann-Gibbs-Shannon entropies, the first concepts of entropy are described in Section 2.1. Section 2.2 is dedicated to entropies derived from Shannon entropy such as differential entropy (Section 2.2.1), spectral entropy (Section 2.2.2), tone-entropy (Section 2.2.3), wavelet entropy (Section 2.2.4), empirical mode decomposition energy entropy (Section 2.2.5), and  $\Delta$ -entropy (Section 2.2.6).

Kolmogorov, topological, and geometric entropies are described in Section 2.3. Section 2.4 describes the particular cases of Rényi entropy (Section 2.4.1),  $\epsilon$ -smooth Rényi entropy (Section 2.4.2), and Rényi entropy for continuous random variable and the different definition of quadratic entropy (Section 2.4.3). Havrda-Charvát structural  $\alpha$ -entropy and Tsallis entropy are detailed in Section 2.5 while permutation entropy and related entropies are described in Section 2.6.

Section 2.7 is dedicated to rank-based and bubble entropies while topological information content, graph entropy and horizontal visibility graph entropy are detailed in Section 2.8. In Section 2.9, are described the approximate, the sample, and related entropies such as the quadratic sample, coefficient of the sample, and intrinsic mode entropies (Section 2.9.1), dispersion entropy and fluctuation-based dispersion entropy (Section 2.9.2), fuzzy entropy (Section 2.9.3), modified sample entropy (Section 2.9.4), fuzzy measure entropy (Section 2.9.5), and kernel entropies (Section 2.9.6).

### 2.1. Early Times of the Entropy Concept

In 1864, Rudolf Clausius introduced the term entropy in thermodynamics from the Greek word *Entropiein* for transformation and change [20]. The concept of entropy arises, providing a statement of the second law of thermodynamics. Later, statistical mechanics provided a connection between thermodynamic entropy and the logarithm of the number of microstates in the system's macrostate. This work is attributed to Ludwig Boltzmann and the Boltzmann entropy [21],  $S$ , was defined as

$$S = k_B \cdot \ln(W) \quad (1)$$

where  $k_B$  is the thermodynamic unit of the measurement of the entropy and is the Boltzmann constant,  $W$  is called the thermodynamic probability or statistical weight and is the total number of microscopic states or complexions compatible with the macroscopic state of the system. The Boltzmann entropy, also known as Boltzmann-Gibbs entropy [22] or thermodynamic entropy [23], is a function on phase space and is thus defined for an individual system. Equation (1) was originally formulated by Boltzmann between 1872 and 1875 [24] but later put into its current form by Max Planck in about 1900 [25].

In 1902, Gibbs published *Elementary Principles in Statistical Mechanics* [26], where he introduced Gibbs entropy. In contrast to Boltzmann entropy, the Gibbs entropy of a macroscopic classical system is a function of a probability distribution over phase space, i.e., an ensemble. The Gibbs entropy [27] in the phase space of system  $X$ , at time  $t$ , is defined by the quantity,  $S_G(t)$ :

$$S_G(t) = -k_B \int_X f_t(x) [\log f_t(x) - 1] dx \quad (2)$$

where  $k_B$  is the Boltzmann's constant and  $f_t$  is the density of the temporal evolution of the probability distribution in the phase space.

Boltzmann entropy was defined for a macroscopic state of a system, while Gibbs entropy is the generalization of the Boltzmann entropy for over an ensemble, which is over the probability distribution of macrostates [21,28].

In 1928, the electrical engineer Ralph Hartley [29] proposed that the amount of information associated with any finite set of entities could be understood as a function of the set's size. Hartley defined the amount of information  $h(X)$  associated with the finite set  $X$  as the logarithm to some base  $b$  of the  $X$ 's size, as shown in Equation (3). This amount is known as Hartley entropy and is related to the particular cases of Rényi entropy (see Section 2.4.1).

$$h(X) = \log_b |X| \quad (3)$$

In 1932, Von Neumann generalized Gibbs entropy to quantum mechanics, and it is known as Von Neumann entropy (quantum Gibbs entropy or quantum entropy) [30]. The quantum entropy was initially defined as Shannon entropy associated with the density matrix's eigenvalues, but sixteen years before Shannon entropy was defined [31]. Von Neumann entropy is the quantum generalization of Shannon entropy [32].

In 1948, the American Claude Shannon published "*A mathematical theory of communication*" in July and October issues of Bell System technical journal [33]. He proposed the notion of entropy to measure how the information within a signal can be quantified with absolute precision as the amount of unexpected data contained in the message. This measure is known as Shannon entropy (SE) and is defined as:

$$SE(X) = - \sum_i p_i \cdot \ln(p_i) \quad (4)$$

where  $X = \{x_i, i = 1, \dots, N\}$  is a time-series,  $0 \ln 0 = 0$  by convention and  $p_i$  represents the probability of  $x_i, i = 1, \dots, N$ . Therefore  $p_i > 0 \forall i$  and  $\sum p_i = 1$ .

In 1949, at the request of Shannon employer, Warren Weaver, the paper of Shannon is republished as a book [34], preceded by an introductory exhibition by Weaver. Weaver's text [35] attempts to explain how Shannon's ideas can extend far beyond their initial objectives to all sciences that address communication problems in the broad sense. Weaver is sometimes cited as the first author, if not the only author of information theory [36]. Nevertheless, as Weaver himself stated: "*No one could realize more keenly than I do that my own contribution to this book is infinitesimal as compared with Shannon's*" [37]. Shannon and Weaver, in the book "*The Mathematical Theory of Communication*" [38], referred to Tolman (1938), who in turn attributes to Pauli (1933), the definition of entropy Shannon used [39].

Many authors [7,40,41] used the term Boltzmann-Gibbs-Shannon entropy (BGS), generalizing the entropy expressions of Boltzmann, Gibbs, and Shannon and following the ideas of Stratonovich [42] in 1955:

$$S_{BGS}(p_1, \dots, p_w) = -k_B \sum_{i=1}^W p_i \cdot \ln(p_i) \quad (5)$$

where  $k_B$  is the Boltzmann constant,  $W$  is the number of microstates consistent with the macroscopic constraints of a given thermodynamical system and  $p_i$  is the probability that the system is in the microstate  $i$ .

In 1957, Khinchin [43] proposed an axiomatic definition of the Boltzmann entropy, based on four requirements, known as the Shannon-Khinchin axioms [8]. He also demonstrated that Shannon entropy is generalized by

$$H(X) = -k \sum_i p_i \cdot \log(p_i) \quad (6)$$

where  $k$  is a positive constant representing the desired unit of measurement. This property enables us to change the logarithm base in the definition, i.e.,  $SE_b(X) = \log_b SE_a(X)$ . Entropy can be changed from one base to another by multiplying by the appropriate factor [44]. Therefore, depending on the application area, instead of the natural logarithm in Equation (4), one can use logarithms from other bases.

There are different approaches to the derivation of Shannon entropy based on different postulates or axioms [44–47].

Figure 2 represents the origin of the universe of entropies that we propose in this paper, as well as the relationships between the entropies that we describe in this section.

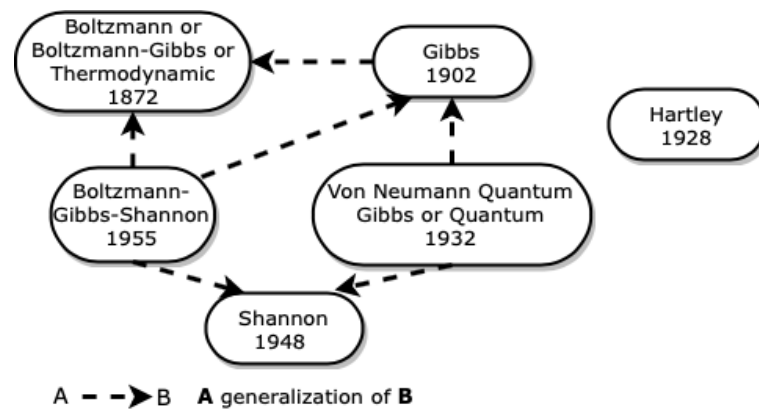


Figure 2. Origin of The Entropy Universe.

2.2. Entropies Derived from Shannon Entropy

Based on Shannon entropy, many researchers have been devoted to enhancing the performance of Shannon entropy for more accurate complexity estimation, such as differential entropy, spectral entropy, tone-entropy, wavelet entropy, empirical mode decomposition energy entropy, and  $\Delta$ -entropy. Throughout this paper, other entropies related to Shannon entropy have been analyzed, such as Rényi, Tsallis, permutation, and dispersion entropies. However, due to its context, we decided to include them in other sections.

In Figure 3, we represent how the entropies described below are related to each other.

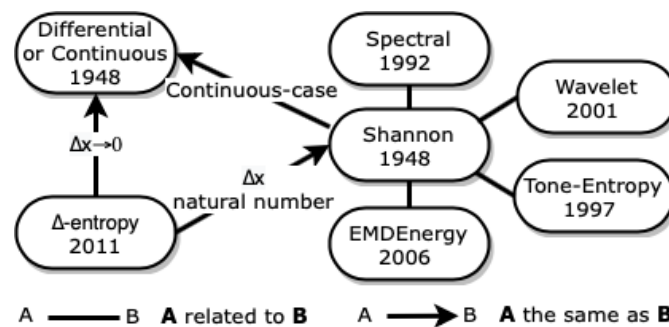


Figure 3. Entropies related to Shannon entropy.

2.2.1. Differential Entropy

Shannon entropy was formalized for discrete probability distributions. However, the concept of entropy can be extended to continuous distributions through a quantity known as differential entropy (DE) (also referred to as continuous entropy). The concept of DE was introduced by Shannon [33] like a continuous-case extension of Shannon entropy [44]. However, further analysis reveals several shortcomings that render it far less useful than it appears.

The information-theoretic quantity for continuous one-dimensional random variables is differential entropy. The DE for a continuous time-series  $X$  with probability density  $p(x)$  is defined as:

$$DE(X) = - \int_S p_i \cdot \ln(p_i) \tag{7}$$

where  $S$  is the support set of the random variable.

In the discrete case, we had a set of axioms from which we derived Shannon entropy and, therefore, a collection of properties that the measure must present. Charles Marsh [48] stated that “continuous entropy on its own proved problematic,” the differential entropy is not a limit of the Shannon entropy for  $n \rightarrow \infty$ . On the contrary, it differs from the limit of the Shannon entropy by an infinite displacement [48]. Among other problems, continuous

entropy can be negative, while discrete entropy is always non-negative. For example, when the continuous random variable  $U$  is uniformly distributed over the interval  $(a, b)$ ,  $p(u) = 1/(b - a)$  Equation (7) results in:

$$DE(U) = \ln(b - a). \quad (8)$$

The entropy value obtained in Equation (8) is negative when the length of the interval is  $< 1$ .

### 2.2.2. Spectral Entropy

In 1992, Kapur et al. [49] proposed spectral entropy (SpEn) that uses the power spectral density  $\hat{P}(f)$  obtained from the Fourier transformation method [50]. The power spectral density represents the distribution of power as a function of frequency. Thus, the normalization of  $\hat{P}(f)$  yields a probability density function. Using the definition of Shannon entropy, SpEn can be defined as:

$$SpEn = \sum_{i=f_l}^{f_h} p_i \cdot \log(p_i) \quad (9)$$

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

### 2.2.3. Tone-Entropy

In 1997, Oida et al. [51] proposed the tone-entropy (T-E) analysis to characterize the time-series of percentage index (PI) of heart period variation. In this paper, the authors used the tone, and the entropy to characterize the time-series PI. Considered a time-series  $X = \{x_i, i = 1, \dots, N\}$ , PI is defined as:

$$PI(i) = \frac{(x_i - x_{i+1})100}{x_i}. \quad (10)$$

The Tone is defined as the first-order moment (arithmetic average) of this PI time-series as:

$$Tone = \frac{1}{N-1} \sum_{i=1}^N PI(i). \quad (11)$$

Entropy is defined from the PI's probability distribution using Shannon's Equation (4) with  $\log_2$  instead of  $\ln$ .

### 2.2.4. Wavelet Entropy

Wavelet entropy (WaEn) was introduced by Rosso and Blanco [52] in 2001. The WaEn is an indicator of the disorder degree associated with the multi-frequency signal response [11]. The definition of WaEn is given as follows:

$$WaEn = - \sum_{i < 0} p_i \cdot \log(p_i) \quad (12)$$

where  $p_i$  denotes the probability distribution of time-series, and  $i$  represents different resolution levels.

### 2.2.5. Empirical Mode Decomposition Energy Entropy

In 2006, Yu et al. [53] proposed the Empirical Mode Decomposition Energy entropy (EMDEnergyEn) that quantifies the regularity of time-series with the help of the intrinsic mode functions (IMFs) [12] obtained by the empirical mode decomposition (EMD).

Assume that we have obtained  $n$  IMFs, three steps are required to obtain the EMDEnergyEn as follows:

1. Calculate  $E_i$  energy for each  $i$ th IMFs  $c_i$ :

$$E_i = \sum_{j=1}^m |c_i(j)|^2 \tag{13}$$

where  $m$  represents the length of IMF.

2. Calculate the total energy of these  $n$  efficient IMFs:

$$E = \sum_{i=1}^n E_i. \tag{14}$$

3. Calculate the energy entropy of IMF:

$$H_{en} = - \sum_{j=1}^n p_i \cdot \log(p_i) \tag{15}$$

where  $H_{en}$  denotes the EMDEnergyEn in the whole of the original signal, and  $p_i = \frac{E_i}{E}$  denotes the percentage of the energy of the IMF number  $i$  relative to the total energy entropy.

### 2.2.6. $\Delta$ –Entropy

In 2011, the  $\Delta$ –entropy was introduced by Chen et al. [54]. This measure is sensitive to the dynamic range of the time-series. The  $\Delta$ –entropy contains two terms, in which the first term measures the probabilistic uncertainty obtained with Shannon entropy. The second term measures the dispersion in the error variable:

$$H_{\Delta}(X) = - \sum_i p_i \cdot \log(p_i) + \log \Delta(X) \tag{16}$$

where  $X = \{x_i, i = 1, \dots, N\}$  is the time-series and  $\Delta(X) = \frac{1}{N-1} \sum_{i=1}^{N-1} |x_{i+1} - x_i|$ .

The  $\Delta$ –entropy converges to DE when the scale ( $\Delta(X)$ ) tends to zero. When the scale defaults to the natural numbers ( $\Delta(X) = 1$  and therefore  $\log(\Delta(X)) = 0$ ) the  $\Delta$ –entropy is indistinguishable from SE.

### 2.3. Kolmogorov, Topological and Geometric Entropies

In 1958, Kolmogorov [55] introduced the concept of entropy in the dynamical system as a measure-preserving transformation and studied the attendant property of completely positive entropy ( $K$ -property). In 1959, his student Sinai [56,57] formulated Kolmogorov-Sinai entropy (or metric entropy or Kolmogorov entropy [58]) that is suitable for arbitrary automorphisms of Lebesgue spaces. The Kolmogorov-Sinai entropy is equivalent to a generalized version of SE under certain plausible assumptions [59].

For the case of a state space, it can be partitioned into hypercubes of content  $\epsilon^m$  in an  $m$ -dimensional dynamical system and observed at time intervals  $\delta$ , defining the Kolmogorov-Sinai entropy as [60]:

$$H_{KS} = - \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n\delta} \sum_{K_1, \dots, K_n} p(K_1, \dots, K_n) \cdot \ln(p(K_1, \dots, K_n)) = \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n\delta} H_n \tag{17}$$

where  $p(K_1, \dots, K_n)$  denote the joint probability that the state of the system is in the hypercube  $k_1$  at the time  $t = \delta$ ,  $k_2$  at  $t = 2\delta, \dots$ , in the hypercube  $k_n$  at  $t = n\delta$  and  $H_n = H(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1)$ . For stationary processes it can be shown that

$$H_{KS} = - \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} (H_{n+1} - H_n). \tag{18}$$

It is impossible to calculate Equation (17) for  $n \rightarrow \infty$  because different estimation methods have been proposed, such as approximate entropy [61] and Rényi entropy [62], and compression [63].

The Kolmogorov-Sinai entropy and conditional entropy coincide for a stochastic process  $X$ , where  $X_n$  is the random variable obtained by sampling the process  $X$  at the present time  $n$ , and  $X_{n-1} = \{X_1, \dots, X_{n-1}\}$  [64].

In this case, conditional entropy quantifies the amount of information in the current process that cannot be explained by its history. If the process is random, the system produces information at the maximum rate, producing the maximum conditional entropy. If, on the contrary, the process is completely predictable, the system does not produce new information, and conditional entropy is zero. When the process is stationary, the system produces new information at a constant rate, meaning that the conditional entropy does not change over time [65]. Note that conditional entropy is, more broadly, the entropy of a random variable conditioned to the knowledge of another random variable [44].

In 1965, inspired by Kolmogorov-Sinai entropy, the concept of the topological entropy (TopEn) was introduced by Adler et al. [66] to describe the complexity of a single map acting on a compact metric space. The notion of TopEn is similar to the notion of metric entropy: instead of a probability measure space, we have a metric space [67].

After 1965, many researchers proposed other notions of TopEn [68–72]. Most of the new notions extended the concept to more general functions or spaces, but the idea of measuring the complexity of the systems was preserved among all these new notions. In [56], the authors review the notions of topological entropy, give an overview of the relation between the notions and fundamental properties of topological entropy.

In 2018, Rong and Shang [10] introduced TopEn based on time-series to characterizes the total exponential complexity of a quantified system with a single number. The authors began by choosing a symbolic method to transform the time-series  $X = \{x_i, i = 1, \dots, N\}$  into a symbolic sequence  $Y = \{y_i, i = 1, \dots, N\}$ . At the same time, consider  $k$  the number of different alphabets of  $Y$  and  $P(n)$  represents the number of different sets of words with length  $n$ . Note that  $P(N) = 1$  and the TopEn of the time-series was defined as:

$$h_n = \frac{\log_k P(n)}{n} \quad (1 \leq n \leq N). \tag{19}$$

The maximum of  $P(n)$  is  $k^n$ , then TopEn can reach a maximum of 1, while TopEn’s minimum value is 0, which is reached when  $n = N$ . Addabbo and Blackmore [73] showed that metric entropy, with the imposition of some additional properties, is a special case of topological entropy and Shannon entropy is a particular form of metric entropy.

In 1988, Walczak et al. [74] introduced the geometric entropy (or foliation entropy) to study a foliation dynamics, which can be considered as a generalization of TopEn of a single group [75–77]. Recently, Rong and Shang [10] proposed geometric entropy for time-series based on the multiscale method (see Section 2.10) and the original definition of geometric entropy provided by Walczak. To calculate the value of TopEn and geometric entropy in their work, the authors used horizontal visibility graphs [78] to transform the time-series into a symbolic series. More details about horizontal visibility graphs in Section 2.8.

In Figure 4, we show the relationships between the entropies described in this section.

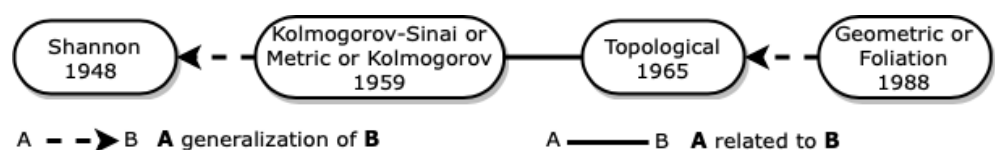


Figure 4. Relationship between Kolmogorov, topological and geometric entropies.

### 2.4. Rényi Entropy

In 1961, Rényi entropy (RE) or q-entropy was introduced by Alfréd Rényi [79] and played a significant role in information theory. Rényi entropy, a generalization of Shannon



entropy, is a family of functions of order  $q$  ( $R_q$ ) for quantifying the diversity, uncertainty, or randomness of a given system defined as:

$$R_q(X) = \frac{1}{1-q} \log_2 \sum_i p_i^q. \tag{20}$$

In the following sections, we describe particular cases of Rényi entropy,  $\epsilon$ -smooth Rényi entropy, Rényi entropy for a continuous random variable, and discuss issues with the different definitions of quadratic entropy.

Figure 5 illustrates the relations between Shannon entropy, Rényi entropy, particular cases of Rényi entropy, and  $\epsilon$ -smooth Rényi entropy.

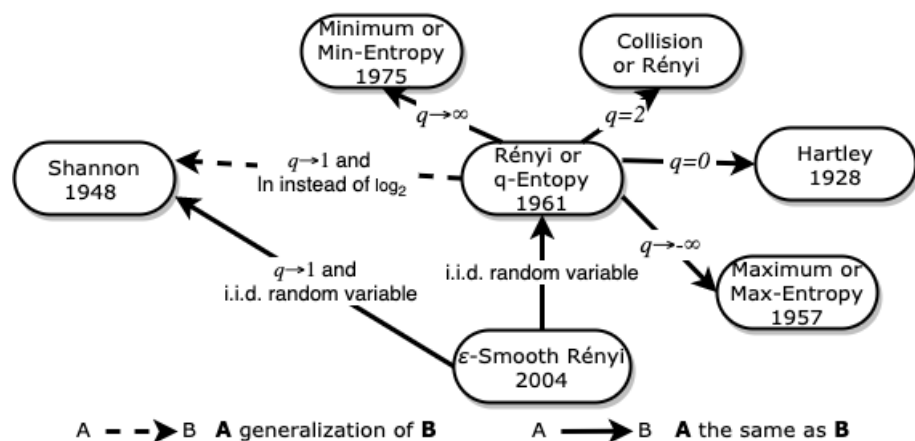


Figure 5. Relationship between Rényi entropy and its particular cases.

### 2.4.1. Particular Cases of Rényi Entropy

There are some particular cases of Rényi entropy [80–82], for example, if  $q = 0$ ,  $R_0 = \frac{1}{1-0} \log_2 \sum_{i=1}^N p_i^0 = \log_2 N$  is the Hartley entropy (see Equation (3)).

The Rényi entropy converges to the well-known Shannon entropy, eventually, with multiplication by a constant resulting from the base change property, i.e., when  $q \rightarrow 1$  and  $\log_2$  is replaced by  $\ln$ .

When  $q = 2$ ,  $R_2 = -\log_2 \sum_{i=1}^N p_i^2$  is called collision entropy [80] and is the negative logarithm of the likelihood of two independent random variables with the same probability distribution to have the same value. The collision entropy measures the probability for two elements drawn according to this distribution to collide. Many papers refers to Rényi entropy [81] when using  $q = 2$ , even when that choice it is not explicit.

The concept of maximum entropy (or Max-Entropy) arose in statistical mechanics in the nineteenth century and has been advocated for use in a broader context by Edwin Thompson Jaynes [83] in 1957. The Max-Entropy can be obtained from Rényi entropy when  $q \rightarrow -\infty$ , and the limit exists, as:

$$R_{-\infty} = -\log_2(\min_{i=1,\dots,N} p_i), \tag{21}$$

It is the largest value of  $R_q$ , which justifies the name maximum entropy. The particular cases of Rényi entropy when  $q \rightarrow \infty$ , and the limit exists as

$$R_{\infty} = -\log_2(\max_{i=1,\dots,N} p_i) \tag{22}$$

is called minimum entropy (or Min-Entropy) because it is the smallest value of  $R_q$ . The Min-Entropy was proposed by Edward Posner [84] in 1975. According to reference [85], the collision entropy and Min-Entropy are related by Equation by the following:

$$R_{\infty} \leq R_2 \leq 2R_{\infty}. \tag{23}$$

#### 2.4.2. $\epsilon$ -Smooth Rényi Entropy

In 2004, Renner and Wolf [86] proposed the  $\epsilon$ -smooth Rényi entropy for characterizing the fundamental properties of a random variable, such as the amount of uniform randomness that may be extracted from the random variable:

$$R_q^\epsilon(X) = \frac{1}{1-q} \inf_{Q \in B^\epsilon(P)} \log_2 \left( \sum_{z \in Z} Q(z)^q \right) \quad (24)$$

where  $B^\epsilon(P) = \{Q : (\sum_z |P(z) - Q(z)|)/2 \leq \epsilon\}$  is the set of probability distributions which are  $\epsilon$ -close to  $P$ ;  $P$  is the probability distribution with range  $Z$ ;  $q \in [0, \infty]$  and  $\epsilon \geq 0$ . For the particular case of a significant number of independent and identically distributed (i.i.d.) random variables,  $\epsilon$ -smooth Rényi entropy approaches the Rényi entropy. In this special case, if  $q \rightarrow 1$  the  $\epsilon$ -smooth Rényi entropy approaches the Rényi entropy.

#### 2.4.3. Rényi entropy for Continuous Random Variables and the Different Definition of Quadratic Entropy

According to Lake [87], if  $X$  is an absolutely continuous random variable with density  $f$ , the Rényi entropy of order  $q$  (or  $q$ -entropy) is defined as:

$$R_q(X) = \frac{1}{1-q} \log_2 \left( \int_{-\infty}^{\infty} f(x)^q \right) \quad (25)$$

where letting  $q$  tend to 1, and using L'Hospital's rule results in differential entropy, i.e.,  $DE(x) = R_1(X)$ . Lake [87] uses the term quadratic entropy when  $q = 2$ .

In 1982, Rao [88,89] gave a different definition for quadratic entropy. He introduced quadratic entropy as a new measure of diversity in biological populations, which considers differences between categories of species. For the discrete and finite case, the quadratic entropy is defined as:

$$\sum_{i=1}^s \sum_{j=1}^s d_{i,j} p_i p_j \quad (26)$$

where  $d_{i,j}$  is the difference between the  $i$ -th and the  $j$ -th category and  $p_1, \dots, p_s$  ( $\sum p_i = 1$ ) are the theoretical probabilities corresponding to the  $s$  species in the multinomial model.

#### 2.5. Havrda-Charvát Structural $\alpha$ -Entropy and Tsallis Entropy

The Havrda-Charvát Structural  $\alpha$ -entropy was introduced in 1967 within information theory [90]. It may be considered as a new generalization of the Shannon entropy, different from the generalization given by Rényi. The Havrda-Charvát Structural  $\alpha$ -entropy,  $S_\alpha$ , was defined as:

$$S_\alpha = \frac{2^{\alpha-1}}{2^{\alpha-1} - 1} \left( 1 - \sum_i p_i^\alpha \right) \quad (27)$$

for  $\alpha > 0$  and  $\alpha \neq 1$ . When  $\alpha \rightarrow 1$ , the Havrda-Charvát Structural  $\alpha$ -entropy converges for Shannon entropy minus multiplication by a constant ( $1/\ln 2$ ).

In 1988, the concept of Tsallis entropy (TE) was introduced by Constantino Tsallis as a possible generalization of Boltzmann-Gibbs entropy to nonextensive physical systems and prove that the Boltzmann-Gibbs entropy is recovered as  $q \rightarrow 1$  [7,91]. TE is identical in form to Havrda-Charvát Structural  $\alpha$ -entropy and Constantino Tsallis defined it as:

$$S_q = \frac{k}{q-1} \left( 1 - \sum_i p_i^q \right), \quad q \in \mathbb{R} \quad (28)$$

where  $k$  is a positive constant [91]. In particular, when  $q \rightarrow 1$  implies that

$$S_1 = \lim_{q \rightarrow 1} S_q = -k \sum_i p_i \cdot \ln(p_i) \tag{29}$$

When  $k = 1$  in Equation (29) we recover Shannon entropy [92]. In Equation (29) if we consider  $k$  to be the Boltzmann constant and  $p_i = \frac{1}{N}, \forall i$ , we recover the Boltzmann entropy to  $W = N$ .

Figure 6 summarizes the relationships between the entropies that we describe in this section.

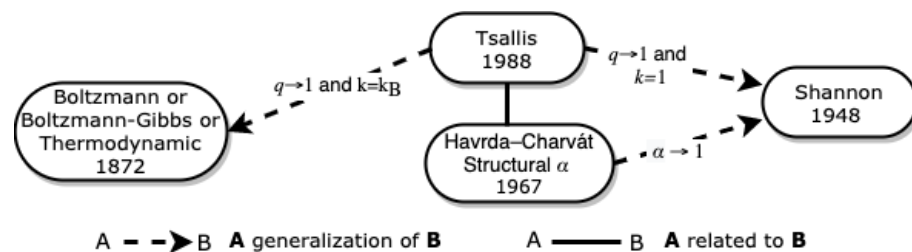


Figure 6. Relationship between Havrda–Charvát structural  $\alpha$ -entropy Tsallis and others entropies.

### 2.6. Permutation Entropy and Related Entropies

Bandt and Pompe [13] introduced permutation entropy (PE) in 2002. PE is an ordinal analysis method, in which a given time-series is divided into a series of ordinal patterns for describing the order relations between the present and a fixed number of equidistant past values [93].

A time-series, as a set of  $N$  consecutive data points, is defined as  $X = \{x_i, i = 1, \dots, N\}$ . PE can be obtained by the following steps. From the original time-series  $X$ , let us define the vectors  $X_m^\tau(i)$  as:

$$X_i = X_m^\tau(i) = (x_i, x_{i+\tau}, x_{i+2\tau}, \dots, x_{i+(m-1)\tau}) \tag{30}$$

with  $i = 1, \dots, N - (m - 1) * \tau$ , where  $m$  is the embedding dimension, and  $\tau$  is the embedding lag or time delay. Reconstruct the phase space so that the time-series maps to a trajectory matrix  $X_m^\tau$ ,

$$X_m^\tau = (X_1, X_2, \dots, X_{N-(m-1)\tau})^\Gamma \tag{31}$$

Each state vector  $X_i$  has an ordinal pattern. It is determined by a comparison of neighboring values. Sort  $X_i$  in ascending order, and the ordinal pattern  $\pi_i$  is the ranking of  $X_i$ . The trajectory matrix  $X_m^\tau$  is thus transformed into ordinal pattern matrix,

$$\pi = (\pi_1, \pi_2, \dots, \pi_{N-(m-1)\tau})^\Gamma \tag{32}$$

It is called permutation because there is a transformation from  $X_i$  to  $\pi_i$  and for a given embedding dimension  $m$ , at most  $m!$  permutations exist in total.

Select all distinguishable permutations and number them  $\pi_j, j = 1, 2, \dots, m!$ . For all the  $m!$  possible permutations  $\pi_j$ , the relative frequency is calculated by

$$p(\pi_j) = \frac{\#\{X_i | X_i \text{ has ordinal pattern } \pi_j\}}{N - (m - 1)\tau} \tag{33}$$

where  $1 \leq i \leq N - (m - 1)\tau$  and  $\#$  represents the number of elements.

According to information theory, the information contained in  $\pi_j$  is measured as  $-\ln(p(\pi_j))$  and PE is finally computed as:

$$H(m) = - \sum_{j=1}^{m!} p(\pi_j) \cdot \ln(p(\pi_j)). \tag{34}$$

Since  $H(m)$  can maximally reach  $\ln(m!)$ , PE is generally normalized as:

$$PE(m) = - \frac{\sum_{j=1}^{m!} p(\pi_j) \cdot \ln(p(\pi_j))}{\ln(m!)} \tag{35}$$

This is the information contained in comparing  $m$  consecutive values of the time-series. Bandt and Pompe [13] defined the permutation entropy per symbol of order  $m$  ( $pe(m)$ ), dividing by  $m - 1$  since comparisons start with the second value:

$$pe(m) = \frac{PE(m)}{m - 1} \tag{36}$$

In reference [13], “Permutation entropy — a natural complexity measure for time-series” the authors also proposed the sorting entropy (SortEn) defined as:  $d_m = H(m) - H(m - 1)$ ,  $d_2 = H(2)$ . This entropy determines the information contained in sorting the  $m$ th value among the previous  $m - 1$  when their order is already known.

Based on the probability of the permutation  $p_j = p(\pi_j)$ , other entropy can be defined, e.g., Zhao et al. [14], in 2013, introduced Rényi permutation entropy (RPE) as:

$$RPE_q = \frac{1}{1 - q} \log_2 \sum_{j=1}^{m!} p_j^q \tag{37}$$

Zunino et al. [92], based on the definition of TE, proposed the normalized Tsallis permutation entropy (NTPE) and defined it as:

$$NTPE_q = \frac{\sum_{j=1}^{m!} (p_j - p_j^q)}{1 - m^{1-q}} \tag{38}$$

In Figure 7, we showed how Tsallis, Rényi, Shannon and the entropies described in this section are related to each other. The rank-based and bubble entropies, described the Section 2.7, are related to permutation entropy and are also represented in the Figure 7.

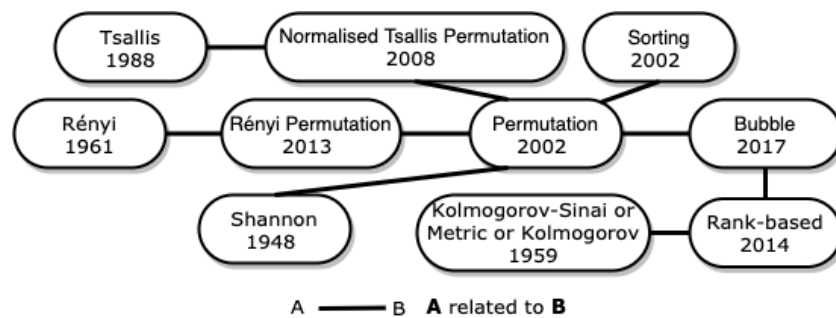


Figure 7. Entropies related to permutation entropy.

2.7. Rank-Based Entropy and Bubble Entropy

Citi et al. [94] proposed the rank-based entropy (RbE), in 2014. The RbE consists of an alternative entropy metric based on the amount of shuffling required for ordering the mutual distances between  $m$ -long vectors when the next observation is considered, that is, when the corresponding  $m + 1$ -long vectors are considered [94]. Operationally, RbE can be defined by the following steps:

1. Compute, for  $1 \leq i < j \leq N - m$ , the mutual distances vectors:  $d_{k(i,j)} = \|v_{m,i} - v_{m,j}\|_\infty$  and  $d'_{k(i,j)} = |x_{i+m} - x_{j+m}|$  where  $\|v_{m,i}\|_\infty$  is the infinity norm of vector  $v_{m,i} = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$  that is,  $\|v_{m,i}\|_\infty = \max_{1 \leq l \leq i+m-1} |x_l|$  and  $k = k(i, j)$  is the index assigned for each  $(i, j)$  pair, with  $1 \leq k \leq K = (N - m - 1)(N - m)/2$ .
2. Consider vector  $d_k$  and find the permutation  $\pi(k)$  such that the vector  $S_k = d_{\pi(k)}$  is sorted in ascending order. Now, if the system is deterministic, we expect that if the

vectors  $v_{m,i}$  and  $v_{m,j}$  are close, then the new observation from each vector  $x_{i+m}, x_{j+m}$  should be close too. In other words,  $S'_k = d'_{\pi(k)}$  should be almost sorted too. Compute inversion count which is a measure of a vector's disorder.

3. Determine the largest index  $k_\rho$  satisfying  $S_{k_\rho} < \rho$  and compute the number  $I$  of inversion pairs  $(k_1, k_2)$  such that  $k_1 < k_\rho, k_1 < k_2$  and  $S'_{k_1} > S'_{k_2}$ .
4. Compute the RbE as:

$$RbE = -\ln\left(1 - \frac{I}{(2K - k_\rho - 1)k_\rho/2}\right). \tag{39}$$

The concept of bubble entropy (BEn) was introduced by Manis et al. [95] in 2017 and it is based on permutation entropy (see Section 2.6), where the vectors in the embedding space are ranked and this rank is inspired by rank-based entropy.

The computation of BEn is as follows:

1. Sort each vector  $X_i$ , defined in Equation (30), of  $m - 1$  elements in ascending order, counting the number of swaps  $n_i$  necessary. The  $n_i$  number is obtained by bubble sort [96]. For more details about bubble sort see paper [97].
2. Compute an histogram from  $n_i$  values and normalize it by  $N - m$ , to obtain the probabilities  $p_i$  (describing how likely a given number of swaps  $n_i$  is).
3. Consider  $RPE_2^{m-1}$  according to Equation (37) for  $q = 2$  and from  $p_i$  compute the  $RPE_{swaps}^{m-1}$  [98]:

$$RPE_{swaps}^{m-1} = -\log_2 \sum_{i=1}^n p_i^2. \tag{40}$$

4. Repeat steps 1 to 3 to compute  $RPE_{swaps}^m$ .
5. Compute BEn by:

$$BEn = \frac{(RPE_{swaps}^m - RPE_{swaps}^{m-1})}{\log_2\left(\frac{m}{m-2}\right)}. \tag{41}$$

### 2.8. Topological Information Content, Graph Entropy and Horizontal Visibility Graph Entropy

In the literature, there are variations in the definition of graph entropy [99]. Inspired by Shannon's entropy, the first researchers to define and investigate the entropy of graphs were Rashevsky [100], Trucco [101], and Mowshowitz [102–105]. In 1955, Rashevsky introduced the information measures for graphs  $G = (V, E)$  called topological information content and defined as:

$$V_I(G) = -\sum_{i=1}^k \frac{|N_i|}{|V|} \log\left(\frac{|N_i|}{|V|}\right) \tag{42}$$

where  $|N_i|$  denotes the number of topologically equivalent vertices in the  $i$ th vertex orbit of  $G$  and  $k$  is the number of different orbits. Vertices are considered as topologically equivalent if they belong to the same orbit of a graph  $G$ . In 1956, Trucco [101] introduced similar entropy measures applying the same principle to the edge automorphism group.

In 1968, Mowshowitz [102–105] introduced an information measure using chromatic decompositions of graphs and explored the properties of structural information measures relative to two different equivalence relations defined on the vertices of a graph. Mowshowitz [102] defined graph entropy as:

$$I(G) = \min_{\hat{V}} \left\{ -\sum_{i=1}^h \frac{n_i(\hat{V})}{|V|} \log\left(\frac{n_i(\hat{V})}{|V|}\right) \right\} \tag{43}$$

where  $\hat{V} = \{V_i | 1 \leq i \leq h\}$ ,  $|V_i| = n_i(V)$  denotes an arbitrary chromatic decomposition of a graph  $G$ ,  $h = \chi(G)$  is the chromatic number of  $G$ .

In 1973, János Körne [106] introduced a different definition of graph entropy linked to problems in information and coding theory. The context was to determine the performance of the best possible encoding of messages emitted by an information source where the symbols belong to a finite vertex set  $V$ . According to Körne, the graph entropy of  $G$ , denoted by  $H(G)$  is defined as:

$$H(G) = \min_{X,Y} I(X;Y) \tag{44}$$

where  $X$  is chosen uniformly from  $V$ ,  $Y$  ranges over independent sets of  $G$ , the joint distribution of  $X$  and  $Y$  is such that  $X \in Y$  with probability one, and  $I(X;Y)$  is the mutual information of  $X$  and  $Y$ . We will not go deeper into the analysis of the graph entropy defined by Körner because it involves the concept of mutual information.

There is no unanimity in the research community regarding the first author that defined graph entropy. Some researchers consider that graph entropy was defined by Mowshowitz in 1968 [50,99] but other researchers consider that it was introduced by János Körne in 1973 [107,108]. New entropies have emerged based on both concepts of graph entropy. When the term graph entropy is mentioned, care must be taken and understand the underlying definition.

The concept of horizontal visibility graph entropy, a measure based on one definition of graph entropy and in the concept of horizontal visibility graph, was proposed by Zhu et al. [109], in 2014. The horizontal visibility graph is a type of complex network, based on a visibility algorithm [78], and was introduced by Luque et al. [110], in 2009. Let a time-series  $X = \{x_i, i = 1, \dots, N\}$  be mapped into a graph  $G(V, E)$  where a time point  $x_i$  is mapped into a node  $v_i$ . The relationship between any two points  $(x_i, x_j)$  is represented by an edge  $e_{ij}$ , which are connected if and only if the maximal values between  $x_i$  and  $x_j$  are less than both of them. Therefore, each edge can be defined as [111]:

$$e_{i,j} = \begin{cases} 1 & x_j > \max(x[(j+1) \dots (i-1)]) \\ 1 & j+1 = i \\ 0 & \text{other cases} \end{cases}$$

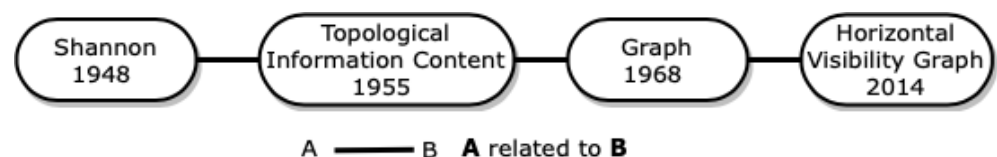
where  $e_{ij} = 0$  implies that the edge does not exist, otherwise it does.

The graph entropy, based on either vertices or edges, can be computed as follows:

$$GE = - \sum_{k=1}^n p(k) \cdot \log(p(k)) \tag{45}$$

where  $p(k)$  is a probability degree  $k$  over a degree sequence of the graph  $G$ . It is obtained by counting the number of nodes having degree  $k$  divided by the size of the degree sequence. Notice that, the more fluctuating the degree sequence is, the larger the graph entropy is. There are others graph entropy calculation methods based on either vertices or edges [99].

We summarize how the entropies described in this section are related to each other, in Figure 8.



**Figure 8.** Relations between topological information content, graph entropy and horizontal visibility graph entropy.

### 2.9. Approximate and Sample Entropies

Pincus [112] introduced the approximate entropy (ApEn), in 1991. ApEn is derived from Kolmogorov entropy [113] and it is a technique used to quantify the amount of

regularity and the unpredictability of fluctuations over time-series data [114]. In order to calculate the ApEn the new series of a vector of length  $m$  (embedding dimension) are constructed,  $X_m^1$ , based on Equation (30). For each vector  $X_m^1(i)$ , the value  $C_m^r(i)$ , where  $r$  is referred as a tolerance value, is computed as:

$$C_m^r(i) = \frac{\text{number of } d[X_m^1(i), X_m^1(j)] \leq r}{N - m + 1}, \quad \forall j. \quad (46)$$

Here the distance between the vector  $X_m^1(i)$  and its neighbor  $X_m^1(j)$  is:

$$d[X_m^1(i), X_m^1(j)] = \max_{k=1, \dots, m} |x(i+k-1) - x(j+k-1)|. \quad (47)$$

The value  $C_m^r(i)$  can also be defined as:

$$C_m^r(i) = \frac{1}{N - m + 1} \sum_{j=1}^{N-m+1} \theta(\max_{k=1, \dots, m} |x(i+k-1) - x(j+k-1)| - r) \quad (48)$$

where  $\theta$  is the Heaviside function

$$\theta(t) = \begin{cases} 1 & \text{if } t \leq 0 \\ 0 & \text{if } t > 0 \end{cases}$$

Next, the average of the natural logarithm of  $C_m^r(i)$  is computed for all  $i$ :

$$\Phi_m^r = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln(C_m^r(i)). \quad (49)$$

Since in practice  $N$  is a finite number, the statistical estimate is computed as:

$$ApEn(m, r) = \begin{cases} \Phi_m^r - \Phi_{m+1}^r & \text{for } m > 0 \\ -\Phi_1^r & \text{for } m = 0 \end{cases}$$

The disadvantages of ApEn are: it lacks relative consistency, its strong dependence on data length and is often lower than expected for short records [15]. To overcome the disadvantages of ApEn, in 2000 Richman and Moorman [15] proposed the sample entropy (SampEn) to replace ApEn by excluding self-matches and thereby reducing the computing time by one-half in comparison with ApEn. For the SampEn [15] calculation, the same parameters defined for the ApEn,  $m$ , and  $r$  are required. Considering  $A$  as the number of vector pairs of length  $m + 1$  having  $d[X_m^1(i), X_m^1(j)] \leq r$ , with  $i \neq j$  and  $B$  as the total number of template matches of length  $m$  also with  $i \neq j$ , the SampEn is defined as:

$$SampEn = -\ln \frac{A}{B}. \quad (50)$$

Many entropies related to ApEn and SampEn have been created. In Figure 9, some entropies related to ApEn and SampEn are represented, which are described in the following sections.

### 2.9.1. Quadratic Sample Entropy, Coefficient of Sample Entropy and Intrinsic Mode Entropy

The SampEn has a strong dependency on the size of the tolerance  $r$ . Normally, smaller  $r$  values lead to higher and less confident entropy estimates because of falling numbers of matches of length  $m$  and, to an even greater extent, matches of length  $m + 1$  [16].

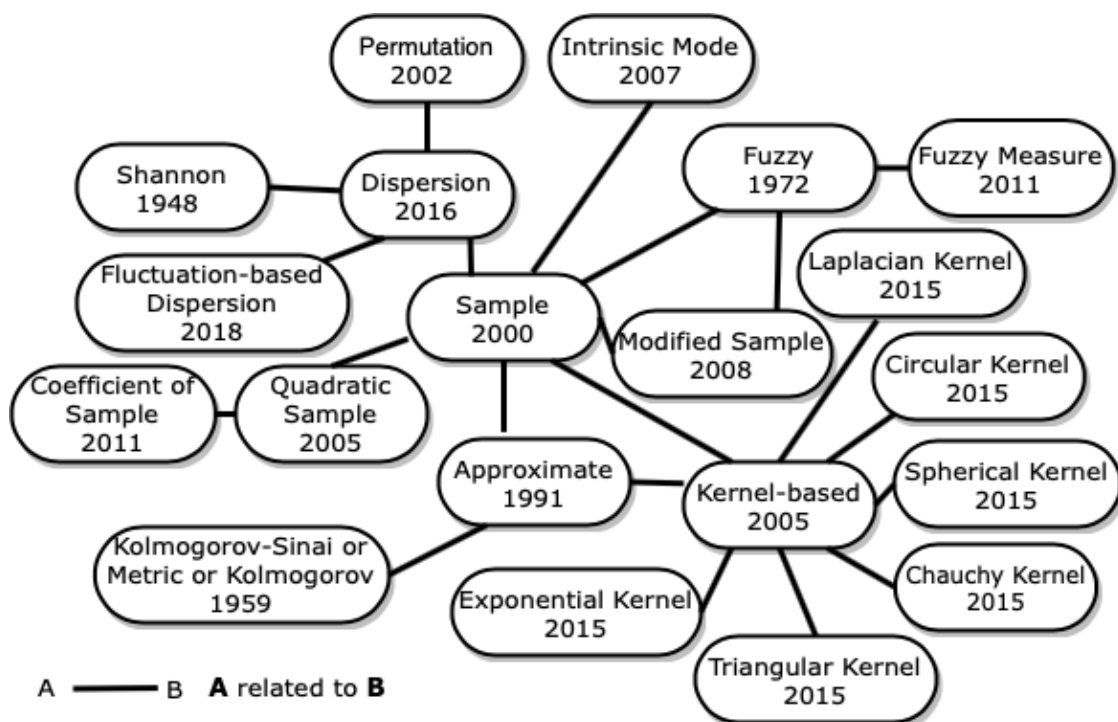


Figure 9. Entropies related to sample entropy.

In 2005, Lake [87] introduced the concept of quadratic sample entropy (QSE) (Lake called it quadratic differential entropy rate) to solve the aforementioned problem. QSE normalizes the value of  $r$  and allows any  $r$  for any time-series and the results to be compared with any other estimate. The QSE is defined as follows:

$$QSE = -\ln\left(\frac{A}{\frac{B}{2r}}\right) = SampEn + \ln(2r). \tag{51}$$

where  $A$  is the number of vector pairs of length  $m + 1$  having  $d[X_m^1(i), X_m^1(j)] \leq r$ , with  $i \neq j$  and  $B$  is the total number of template matches of length  $m$  also with  $i \neq j$ , as in SampEn calculation.

In 2010, derived from QSE it was introduced the coefficient of sample entropy (COSEn) [16]. This measure was first devised and applied to the detection of atrial fibrillation through the heart rate. CosEn is computed similarly as QSE:

$$COSEn = -\ln\left(\frac{A}{\frac{B}{2r}}\right) = SampEn - \ln(2r) - \ln(\mu) \tag{52}$$

where  $\mu$  is the mean value of the time-series.

In 2007, Amoud et al. [115] introduced the intrinsic mode entropy (IME). The IME is essentially the SampEn computed on the cumulative sums of the IMF [12] obtained by the EMD.

### 2.9.2. Dispersion Entropy and Fluctuation-Based Dispersion Entropy

The SampEn is not fast enough, especially for long signals, and PE, as a broadly used irregularity indicator, considers only the order of the amplitude values and hence some information regarding the amplitudes may be discarded. In 2016, to solve these problems, Rostaghi and Azami [18] proposed the dispersion entropy (DispEn) applied to a univariate signal  $X = x_1, x_2, \dots, x_N$  whose algorithm is as follows:



1. First,  $x_j (j = 1, 2, \dots, N)$  are mapped to  $c$  classes, labeled from 1 to  $c$ . The classified signal is  $u_j (j = 1, 2, \dots, N)$ . To do so, there are a number of linear and nonlinear mapping techniques. For more details see [116].
2. Each embedding vector  $U_m^{\tau,c}(i)$  with  $m$  embedding dimension and  $\tau$  time delay is created according to  $U_m^{\tau,c}(i) = (u_i^c, u_{i+\tau}^c, u_{i+2\tau}^c, \dots, u_{i+(m-1)\tau}^c)$  with  $i = 1, \dots, N - (m - 1)\tau$ . Each time-series  $U_m^{\tau,c}(i)$  is mapped to a dispersion pattern  $\pi_{v_0 v_1 \dots v_{m-1}}$ , where  $u_i^c = v_0, u_{i+\tau}^c = v_1, \dots, u_{i+(m-1)\tau}^c = v_{m-1}$ . The number of possible dispersion patterns that can be assigned to each time-series  $U_m^{\tau,c}(i)$  is equal to  $c^m$ , since the signal has  $m$  members and each member can be one of the integers from 1 to  $c$  [18].
3. For each  $c^m$  potential dispersion patterns  $\pi_{v_0 v_1 \dots v_{m-1}}$ , their relative frequency is obtained as follows:

$$p(\pi_{v_0 v_1 \dots v_{m-1}}) = \frac{\#\{i | i \leq N - (m - 1)\tau, \text{ has type } \pi_{v_0 v_1 \dots v_{m-1}}\}}{N - (m - 1)\tau} \tag{53}$$

where # represents their cardinality.

4. Finally, the DispEn value is calculated, based on the SE definition of entropy, as follows:

$$DispEn(X, m, c, \tau) = - \sum_{\pi=1}^{c^m} p(\pi_{v_0 v_1 \dots v_{m-1}}) \cdot \ln(p(\pi_{v_0 v_1 \dots v_{m-1}})). \tag{54}$$

In 2018, Azami and Escudero [116] proposed the fluctuation-based dispersion entropy (FDispEn) as a measure to deal with time-series fluctuations. FDispEn considers the differences between adjacent elements of dispersion patterns. According to the authors, this forms vectors with length  $m - 1$ , which each of their elements changes from  $-c + 1$  to  $c - 1$ , soon, becoming  $(2c - 1)^{m-1}$  potential fluctuation-based dispersion patterns. The only difference between DispEn and FDispEn algorithms is the potential patterns used in these two approaches [116].

### 2.9.3. Fuzzy Entropy

The uncertainty resulting from randomness is best described by probability theory, while the aspects of uncertainty resulting from imprecision are best described by the fuzzy sets introduced by Zadeh [117], in 1965. In 1972, De Luca and Termini [118] used the concept of fuzzy sets and introduced a measure of fuzzy entropy that corresponds to Shannon’s probabilistic measure of entropy. Over the years, other concepts of fuzzy entropy have been proposed [119]. In 2007, Chen et al. [120] introduced fuzzy entropy (FuzzyEn), a measure of time-series regularity, for the characterization of surface electromyography signals. In this case, FuzzyEn is the negative natural logarithm of the probability that two similar vectors for  $m$  points remain similar for the next  $m + 1$  points. This measure of FuzzyEn is similar to ApEn and SampEn, replaces the 0-1 judgment of Heaviside function associated with ApEn and SampEn by a fuzzy relationship function [121], the family of exponential functions  $\exp(-d_{ij}^m/r)$ , to get a fuzzy measurement of the two vectors similarity based on their shapes. This measure also comprises the removal of the local baseline which may allow for minimizing the effect of non-stationarity in the time-series. Besides possessing the good properties of SampEn superior to ApEn, FuzzyEn also succeeds in giving an entropy definition for the case of small parameters. The method can also be applied to noisy physiological signals with relatively short databases [50]. Consider a time-series  $\{x_i, i = 1, \dots, N\}$  with embedding dimension  $m$  to calculate FuzzyEn form vector sequence as follows:

$$X_m(i) = \{x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1}\} - x_0(i) \tag{55}$$

with  $i = 1, \dots, N - m + 1$ . In Equation (55),  $X_m(i)$  represents  $m$  consecutive  $x$  values, starting with the  $i$ th point and generalized by removing a baseline:

$$x_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x_{i+j}. \tag{56}$$

For computing FuzzyEn, consider a certain vector  $X_m(i)$ , define the distance function  $d_{m,ij}$  between  $X_m(i)$  and  $X_m(j)$  as the maximum absolute difference of the corresponding scalar components:

$$d_{m,ij} = d[X_m(i), X_m(j)] = \max_{k=0, \dots, m-1} |x_{i+k} - x_0(i) - x_{j+k} - x_0(j)| \tag{57}$$

with  $i, j = 1, \dots, N - m, j \neq i$ .

Given  $n$  and  $r$ , calculate the similarity degree  $D_{m,ij}$  between  $X_m(i)$  and  $X_m(j)$  through a fuzzy function:

$$D_{m,ij} = \mu(d_{m,ij}, n, r) \tag{58}$$

where the fuzzy function  $\mu(d_{m,ij}, n, r)$  is the exponential function

$$\mu(d_{m,ij}, n, r) = \exp\left(-\frac{d_{ij}^n}{r}\right). \tag{59}$$

For each vector  $X_m(i)$ , averaging all the similarity degree of its neighboring vectors  $X_m(j)$ , we get:

$$B_m(n, r, i) = \frac{1}{N - m} \sum_{j=1, j \neq i}^{N-m} D_{m,ij}. \tag{60}$$

Determine the function  $B_m(n, r)$  as:

$$B_m(n, r) = \frac{1}{N - m} \sum_{i=1}^{N-m} B_m(n, r, i). \tag{61}$$

Similarly, form the vector sequence  $\{X_{m+1}(i)\}$  and get the function  $A_m(n, r)$ :

$$A_m(n, r, i) = \frac{1}{N - m} \sum_{j=1, j \neq i}^{N-m} D_{m+1,ij} \tag{62}$$

$$A_m(n, r) = \frac{1}{N - m} \sum_{i=1}^{N-m} A_m(n, r, i). \tag{63}$$

Finally, we can define the parameter  $FuzzyEn(m, n, r)$  of the sequence as the negative natural logarithm of the deviation of  $B_m(n, r)$  from  $A_m(n, r)$ :

$$FuzzyEn(m, n, r) = \lim_{N \rightarrow \infty} (\ln B_m(n, r) - \ln A_m(n, r)) \tag{64}$$

which, for finite databases, can be estimated from the statistic:

$$FuzzyEn(m, n, r, N) = \ln B_m(n, r) - \ln A_m(n, r) = -\ln\left(\frac{A_m(n, r)}{B_m(n, r)}\right). \tag{65}$$

There are three parameters that must be fixed for each calculation of FuzzyEn. The first parameter  $m$ , as in ApEn and SampEn, is the length of sequences to be compared. The other two parameters,  $r$  and  $n$ , determine the width and the gradient of the boundary of the exponential function respectively.

### 2.9.4. Modified Sample Entropy

The measure SampEn may have some problems in the validity and accuracy because the similarity definition of vectors is based on the Heaviside function, of which the boundary is discontinuous and hard. The sigmoid function is a smoothed and continuous version of the Heaviside function.

In 2008, a modified sample entropy (mSampEn), based on the nonlinear Sigmoid function, was proposed to overcome the limitations of SampEn [17]. The mSampEn is similar to FuzzyEn, the only differences is that instead of Equation (59), mSampEn uses the fuzzy membership function:

$$D_{m,ij} = \mu(d_{m,ij}, r) = \frac{1}{1 + \exp[(d_{m,ij} - 0.5)/r]} \tag{66}$$

### 2.9.5. Fuzzy Measure Entropy

In 2011, based on FuzzyEn definition, the fuzzy measure entropy (FuzzyMEn) was proposed by Liu and Zhao [122]. FuzzyMEn combines local and global similarity in a time-series and allows a has discrimination for time-series with different inherent complexities. It is defined as:

$$FuzzyMEn(m, r_L, r_G, n_L, n_G, N) = FuzzyEn(m, n_L, r_L, N) + FuzzyEn(m, n_G, r_G, N) \tag{67}$$

where  $FuzzyEn(m, n_L, r_L, N)$  and  $FuzzyEn(m, n_G, r_G, N)$  are obtained by Equation (65) considering in the Equation (55) the local vector sequence  $XL_m(i)$  and global vector sequence  $XG_m(i)$ :

$$XL_m(i) = X_m(i) = \{x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1}\} - x_0(i) \tag{68}$$

$$XG_m(i) = \{x_i, x_{i+1}, x_{i+2}, \dots, x_{i+m-1}\} - \bar{x} \tag{69}$$

The vector  $XL_m(i)$  represents  $m$  consecutive  $x(i)$  values but removing the local baseline  $x_0(i)$ , which is defined in Equation (56). The vector  $XG_m(i)$  also represents  $m$  consecutive  $x(i)$  values but removing the global mean value  $\bar{x}$  which is defined as:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \tag{70}$$

### 2.9.6. Kernel Entropies

In 2005, Xu et al. proposed another modification of ApEn, the approximate entropy with Gaussian kernel [19]. It exploits the fact that the Gaussian kernel function can be used to give greater weight to nearby points by replacing the Heaviside function, in Equation (48), by  $k(i, j, r)$ . The first kernel proposed  $k(i, j, r)$  is defined as:

$$k(i, j, r) = \exp\left(-\frac{\|x(i) - x(j)\|^2}{10r^2}\right) \tag{71}$$

and the kernel-based entropy (KbEn) [123] is given as:

$$KbEn(m, r) = \Phi_m^r - \Phi_{m+1}^r \tag{72}$$

where  $\Phi_m^r$  was defined in Equation (49).

Therefore, if  $k(i, j, r) = \theta(\max_{k=1, \dots, m} |x(i+k-1) - x(j+k-1)| - r)$  and  $\theta$  is the Heaviside function then the resulting entropy value is the ApEn. The same procedure of changing the distance measure can be applied to define the sample entropy with Gaussian kernel [61].

In 2015, Mekyska et al. [124] proposed 6 other kernels based in approximate and sample entropies: exponential kernel entropy (EKE):

$$k(i, j, r) = \exp\left(-\frac{\|x(i) - x(j)\|}{2r^2}\right); \quad (73)$$

Laplacian kernel entropy (LKE):

$$k(i, j, r) = \exp\left(-\frac{\|x(i) - x(j)\|}{r}\right); \quad (74)$$

circular kernel entropy (CKE):

$$k(i, j, r) = \frac{2}{\pi} \arccos\left(-\frac{\|x(i) - x(j)\|}{r}\right) - \frac{2}{\pi} \frac{\|x(i) - x(j)\|}{r} \sqrt{1 - \left(\frac{\|x(i) - x(j)\|}{2r^2}\right)^2} \quad (75)$$

for  $\|x(i) - x(j)\| < r$ , and zero otherwise; spherical kernel entropy (SKE):

$$k(i, j, r) = 1 - \frac{3}{2} \frac{\|x(i) - x(j)\|}{r} + \frac{1}{2} \left(\frac{\|x(i) - x(j)\|}{2r^2}\right)^3 \quad (76)$$

for  $\|x(i) - x(j)\| < r$ , and zero otherwise; Cauchy kernel entropy (CauchyKE):

$$k(i, j, r) = \frac{1}{1 + \frac{\|x(i) - x(j)\|^2}{r}} \quad (77)$$

for  $\|x(i) - x(j)\| < r$ , and zero otherwise; triangular kernel entropy (TKE):

$$k(i, j, r) = 1 - \frac{|x(i) - x(j)|}{r} \quad (78)$$

for  $|x(i) - x(j)| < r$ , zero otherwise.

Zaylaa et al. [123], called Gaussian entropy, exponential entropy, circular entropy, spherical entropy, Cauchy entropy and triangular entropy to kernels entropies based in ApEn.

### 2.10. Multiscale Entropy

The multiscale entropy approach (MSE) [60,125] was inspired by Zhang's proposal [126], and considers the information of a system's dynamics on different time scales. Multiscale entropy algorithms are composed of two main steps. The first one is the construction of the time-series scales: using the original signal, a scale,  $s$ , is created from the original time-series, through a coarse-graining procedure, i.e., replacing  $s$  non-overlapping points by their average. In the second step, the entropy (sample entropy, permutation entropy, fuzzy entropy, dispersion entropy, and others) is computed for the original signal and for the coarse-grained time-series to evaluate the irregularity for each scale. For this reason, methods of multiple scale entropy (such as entropy of entropy [127], composite multiscale entropy [128], refined multiscale entropy [129], modified multiscale entropy [130], generalized multiscale entropy [131], multivariate multiScale entropy [132], and others) were not explored in this paper.

### 3. The Entropy Universe Discussion

The Entropy Universe is presented in Figure 10.

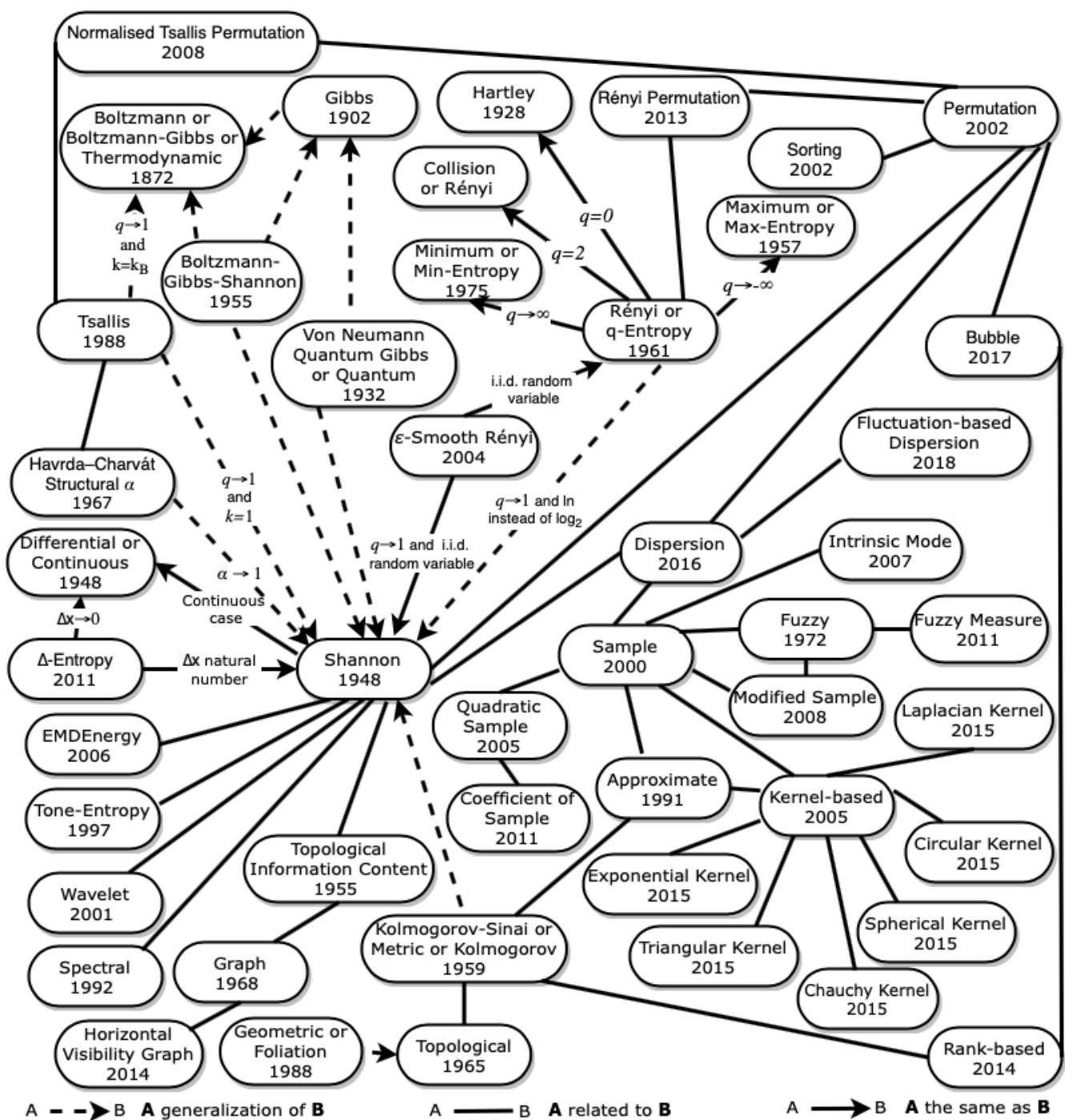


Figure 10. The Entropy Universe.

Entropy is the uncertainty of a single random variable. We can define conditional entropy  $H(X|Y)$ , which is the entropy of a random variable conditional on another random variable’s knowledge. The reduction in uncertainty due to the knowledge of another random variable is called mutual information.

The notion of Kolmogorov entropy as a particular case of conditional entropy includes various entropy measures, and estimates proposed to quantify the complexity of a time-series aimed at the degree of predictability of the underlying process. Measures such as approximate entropy, sample entropy, fuzzy entropy, and permutation entropy are prevalent for estimating Kolmogorov entropy in several fields [65].

Other measures of complexity are also related to entropy measures such as the Kolmogorov complexity. The original ideas of Kolmogorov complexity were put forth independently and almost simultaneously by Kolmogorov [133], Solomonoff [134], and Chaitin [135]. Teixeira et al. [136] studied relationships between Kolmogorov complexity and entropy measures. Consequently, in Figure 10 a constellation stands out that relates these entropies.

The concept of entropy has a complicated history. It has been the subject of diverse reconstructions and interpretations, making it difficult to understand, implement, and interpret. The concept of entropy can be extended to continuous distributions. However, as we mentioned, this extension can be very problematic.

Another problem we encountered was the use of the same name for different measures of entropy. The quadratic entropy was defined by Lake [87] and by Rao [89], but despite having the same name, the entropies are very different. The same is valid with graph entropy, which was defined differently by Mowshowitz [102] and by János Körne [106]. Different fuzzy entropies have been proposed over time, as mentioned in Section 2.9.3, but always using the same name. On the other hand, we find different names for the same entropy. It is common in the literature to find thermodynamic entropy, differential entropy, metric entropy, Rényi entropy, geometric entropy, and quantum entropy with different names (see Figure 10). Moreover, the same name is used for a family of entropies and a particular case. For example, the term Rényi entropy is used for all Rényi entropy families and to the particular case  $q = 2$ .

It is also common to have slightly different definitions of some entropies. The Tsallis entropy formula in some papers appears with  $k$ , but in other papers, the authors consider  $k = 1$ , which is not mentioned in the text.

Depending on the application area, different log bases are used for the same entropy. In this paper, in the Shannon entropy formula, we use the natural logarithm, and in the Rényi entropy formula, we use base 2 logarithm. However, it is common to find these entropies defined with other logarithmic bases in the scientific community, which needs to be considered when relating entropies.

In this paper, we study only single variables, but many other entropies have not been addressed. Nonetheless, some entropies were created to measure the information between two or more time-series. The relative entropy (also called Kullback–Leibler divergence) is a measure of how one probability distribution is different from a second, reference probability distribution. The cross-entropy is an index for determining the divergence between two sets or distributions. Many other like, conditional entropy, mutual information, information storage [65], relative entropy, will not be covered in this paper. While belonging to the same universe, they should be considered in a different galaxy.

Furthermore, in the literature, we find connections between the various entropies and other dispersion and uncertainty measures. Ronald Fisher defined a quantity called Fisher information as a strict limit on the ability to estimate the parameters that define a system [137,138]. Several researchers have shown that there is a connection between Fisher's information and Shannon's differential entropy [46,139–141]. The relationship between the variance and entropies was also explored in many papers such as [141–143]. Shannon defined what he called the derived quantity of entropy power (also called entropy rate power) to be the power in a Gaussian white noise limited to the same band as the original ensemble and having the same entropy. The entropy power defined by Shannon is the minimum variance that can be associated with the Gaussian differential entropy [144]. However, the entropy power can be calculated associated with any given entropy. The in-depth study of the links between *The Entropy Universe* and other measures of dispersion and uncertainty remains as future work. As well as the exploration of the possible log-concavity of these measures along with the heat flow, for example, as recently established for Fisher information [145].

Like our universe, the universe of entropies can be considered infinite because it is continuously expanding.

#### 4. Entropy Impact in the Scientific Community

The discussion and the reflection on the definitions of entropy and how entropies are related are fundamental. However, it is also essential to understand each entropy's impact in the scientific community and in which areas each entropy applies. The main goals of this section are two-fold. The first is to analyze the number of citations for each entropy in recent years. The second is to understand the application areas of each entropy discussed in the scientific community.

For the impact analysis, we have used the Web of Science (WoS) and Scopus databases. Those databases provide access to the analysis of citations of scientific works from all scientific branches and have become an important tool in bibliometric analysis[146]. On both websites, researchers and the scientific community can access databases, analysis, and insights. Several researchers have studied the advantages and limitations of using the WoS or Scopus databases [147–151].

The methodology used was the following in both databases: we searched for the paper's title in which entropy was proposed and collected the year publication, the number of citations total and by year, for the last 16 years. In WoS, we selected in *Research Areas* the ten areas with a higher record count, and in Scopus, we selected the ten *Documents by subject area* with a larger record count for each entropy.

##### 4.1. Number of Citations

In Table 1, we list the paper in which each entropy was proposed. We also present the number of citations of the paper in Scopus and WoS since the paper's publication.

**Table 1.** Reference and number of citations in Scopus and WoS of the paper that presented each entropy.

Name of Entropy	Reference	Year	Scopus	Web of Science
Boltzmann entropy	[25]	1900	5	-
Gibbs entropy	[26]	1902	1343	-
Hartley entropy	[29]	1928	902	-
Von Neumann entropy	[30]	1932	1887	-
Shannon/differential entropies	[33]	1948	34,751	32,040
Boltzmann-Gibbs-Shannon	[42]	1955	8	7
Topological information content	[100]	1955	204	-
Maximum entropy	[83]	1957	6661	6283
Kolmogorov entropy	[55]	1958	693	662
Rényi entropy	[79]	1961	3149	-
Topological entropy	[66]	1965	728	682
Havrda-Charvát structural $\alpha$ -entropy	[90]	1967	744	-
Graph entropy	[102]	1968	207	195
Fuzzy entropy	[118]	1972	1395	-
Minimum entropy	[84]	1975	22	17
Geometric entropy	[74]	1988	71	-
Tsallis entropy	[91]	1988	5745	5467
Approximate entropy	[112]	1991	3562	3323
Spectral entropy	[49]	1992	915	26
Tone-entropy	[51]	1997	85	76
Sample entropy	[15]	2000	3770	3172
Wavelet entropy	[52]	2001	582	465
Permutation/sorting entropies	[13]	2002	1900	1708
Smooth Rényi entropy	[86]	2004	112	67
Kernel-based entropy	[19]	2005	15	13
Quadratic sample entropy	[87]	2005	65	68
Empirical mode decomposition energy entropy	[53]	2006	391	359

Table 1. Cont.

Name of Entropy	Reference	Year	Scopus	Web of Science
Intrinsic mode dispersion entropy	[115]	2007	59	55
Tsallis permutation entropy	[92]	2008	35	37
Modified sample entropy	[17]	2008	58	51
Coefficient of sample entropy	[16]	2011	159	136
$\Delta$ -entropy	[54]	2011	13	10
Fuzzy entropy	[122]	2011	23	18
Rényi permutation entropy	[14]	2013	28	26
Horizontal visibility graph entropy	[109]	2014	22	-
Rank-based entropy	[94]	2014	6	6
Kernels entropies	[124]	2015	46	39
Dispersion entropy	[18]	2016	98	84
Buble entropy	[95]	2017	25	21
Fluctuation-based dispersion entropy	[116]	2018	16	10

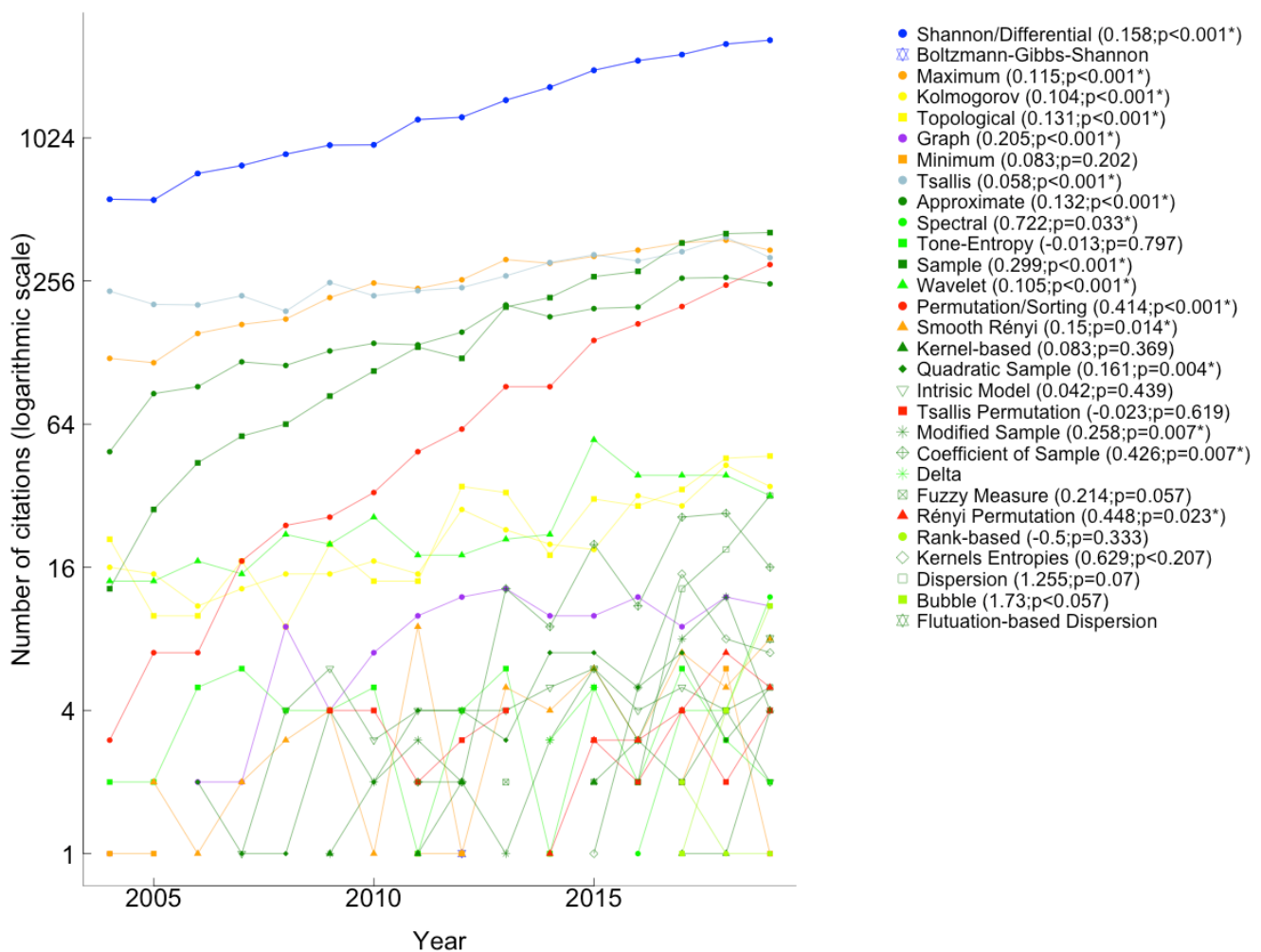
Legend: -paper not found in database.

In general, as expected, the paper that firstly defined each entropy of the universe has more citations in Scopus than in WoS. Based on the number of citations in the two databases, the five most popular entropies in the scientific community are the Shannon/differential, maximum, Tsallis, SampEn, and ApEn entropies. Considering the papers published until 2000, only the papers that proposed the Boltzmann, the Boltzmann-Gibbs-Shannon, the minimum, the geometric, and the tone-entropy entropies have less than one hundred citations in each database. The papers that proposed the empirical mode decomposition energy entropy and the coefficient of sample entropy are the most cited ones among the entropies introduced in the last sixteen years (see Table 1). Of the entropies proposed in the last five years, the dispersion entropy paper was the most cited one, followed by the one introducing the kernels entropies.

Currently, the WoS platform only covers the registration and analysis of papers after 1900. Therefore, we have not found all the documents needed for the analysis to be developed in this section. Ten papers were not found in the WoS, as shown in Table 1. In particular, entropies' papers widely cited in the scientific community were not found, such as Gibbs, Quantum, Rényi, and Fuzzy entropies. Currently, Scopus has more than 76.8 million main records: 51.3 million records after 1995 with references and 25.3 million records before 1996, the oldest from 1788 [152]. We found the relevant information regarding all papers that proposed each entropy in Scopus.

Next, we analyzed the impact of each entropy on the scientific community in recent years. Figures 11 and 12 show the number of citations of the paper proposing each entropy in the last 16 years (from 2004 to 2019), respectively, in the Scopus and WoS databases. Note that the colors used in the Figures 11 and 12 are in accordance with the colors in Figure 1 and we did not use the 2020 information once we considered that those values were incomplete at the research time. The range of citations of the papers that introduced each entropy is extensive; therefore, we chose to use a logarithmic scale in Figures 11 and 12. For entropies whose paper that introduced it has been cited more than three times in the past 16 years, we have drawn the linear regression of the number of citations as a function of the year. In the legends of Figures 11 and 12, we present the slope of the regression line,  $b$ , and the respective  $p$ -value. For  $p < 0.05$  the slope was considered significant. Shannon/Differential paper is the most cited in the last sixteen years, as shown in Figures and Table 1.





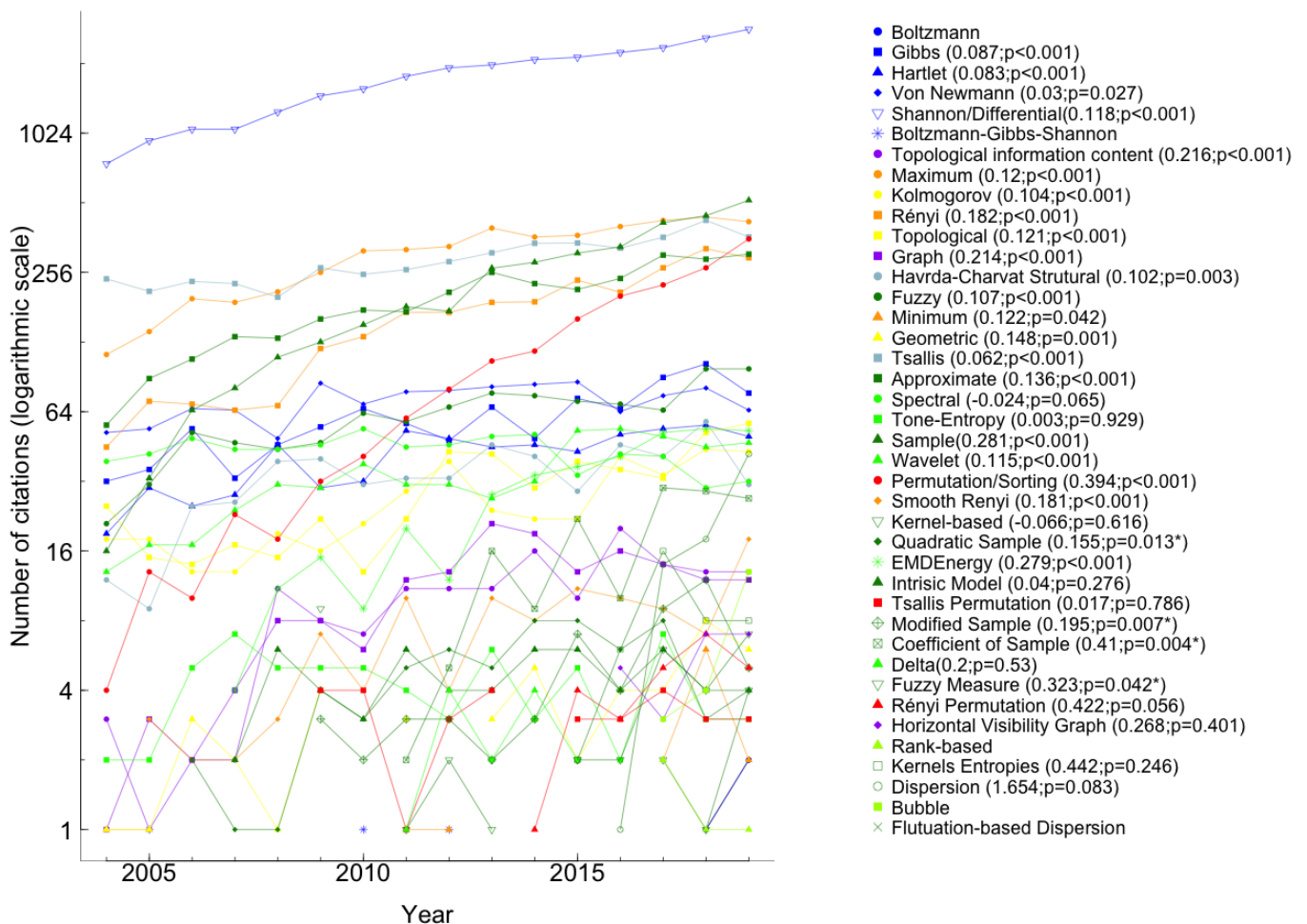
**Figure 11.** Number of citations by year in the WoS between 2004 and 2019 of the papers proposing each measure of entropy, in logarithmic scale ( $\log_2(\text{Number of citations})$ ). In the legend, the ordered pair ( $\beta$ ,  $p$ -value), in papers cited in more than three years, corresponds to the slope of the regression line,  $\beta$ , and the respective  $p$ -value. Statistically significant slopes ( $p < 0.05$ ) are marked with \*.

In recent years, Shannon/differential, sample, maximum, Tsallis, permutation/sorting, and approximate were the entropies most cited, as we can see in both Figures. However, Rényi entropy joins that group in Figure 12. Fluctuation-based dispersion entropy was proposed in 2018, so we only have citations in 2019 that correspond to the respective values in Table 1.

In the last sixteen years, there were several years in which Boltzmann-Gibbs-Shannon, graph, minimum, spectral, kernel-based, Tsallis permutation entropy,  $\Delta$ -entropy, and fuzzy measure entropies had not been cited, according to WoS information (Figure 11). While, Boltzmann, Boltzmann-Gibbs-Shannon, minimum, geometric,  $\Delta$ -entropy, kernel-based, Tsallis permutation, and fuzzy measure entropies papers have not been cited in all years as reported in Scopus (Figure 12). In particular, the paper that introduced Boltzmann-Gibbs-Shannon entropy, with the least number of citations in the last sixteen years, had only one quote in 2012 on the WoS platform, and in Scopus, it had three citations (2005, 2010, and 2012).

Figure 11 shows that the number of citations of the paper proposing sample entropy has increased significantly and is the fastest: the number of citations per year went from 13 in 2004 to 409 in 2019. The permutation/sorting paper had an increasing number of citations as well; the number of citations per year increased significantly from 3 in 2004 to 300 in 2019. From this information, we infer that these entropies have increased their

importance in the scientific community. Based on the WoS database, the two entropies proposed in the last eight years, whose paper has strictly increased the number of citations, are dispersion and bubble entropies, respectively, from 2 in 2016 to 32 in 2019 and 1 in 2017 to 11 in 2019. However, from the ones proposed in the last eight years, the two most cited are dispersion and kernels entropies' papers. Regarding the paper that introduced the wavelet entropy, and according to Figure 11, the year with the most citations was 2015. Since then, the number of citations per year has been decreasing.

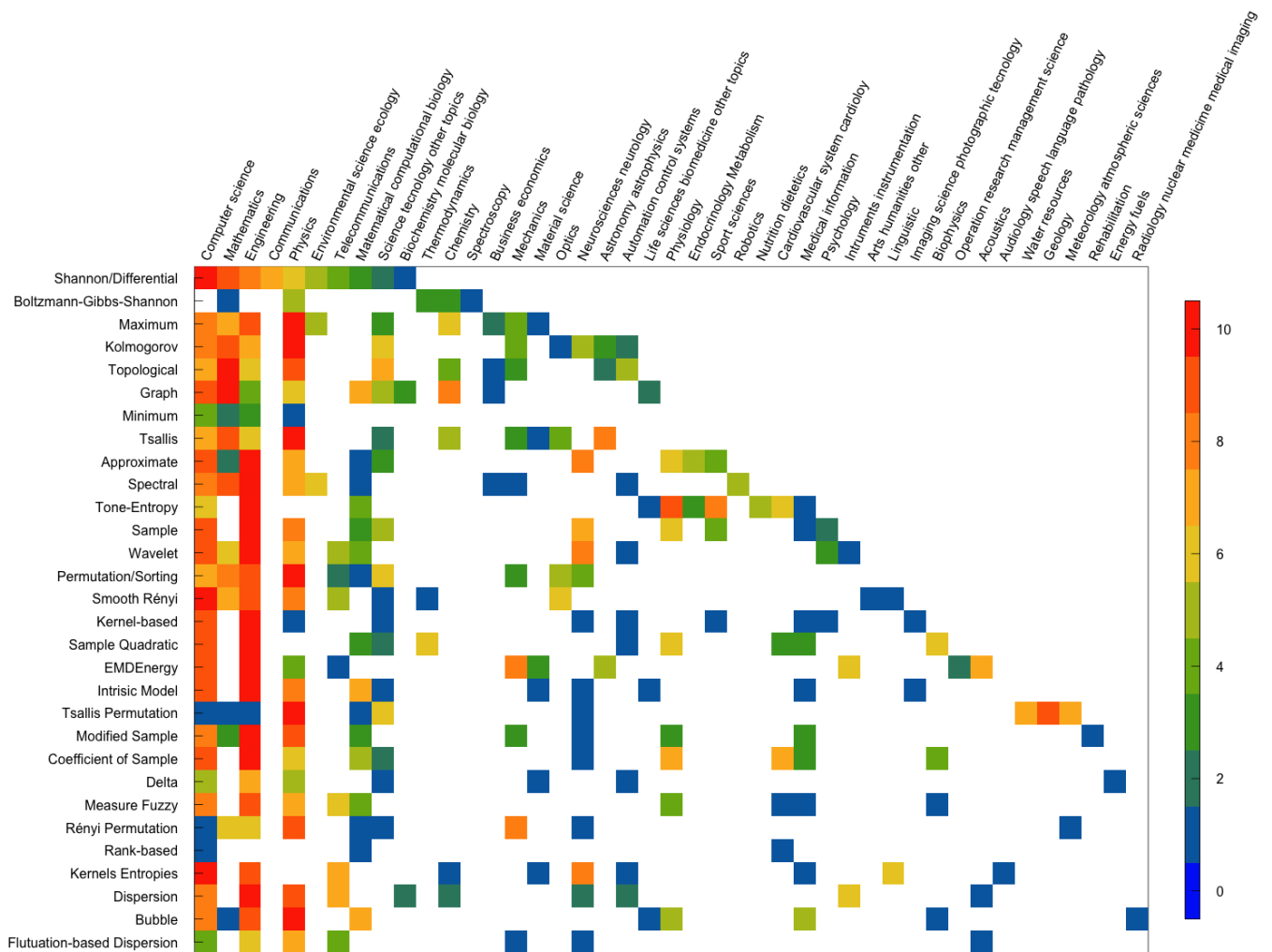


**Figure 12.** Number of citations by year in Scopus between 2004 and 2019 of the papers proposing each measure of entropy, in logarithmic scale ( $\log_2(\text{Number of citations})$ ). In the legend, the ordered pair ( $\beta$ ,  $p$ -value), in papers cited in more than three years, corresponds to the slope of the regression line,  $\beta$ , and the respective  $p$ -value. Statistically significant slopes ( $p < 0.05$ ) are marked with \*.

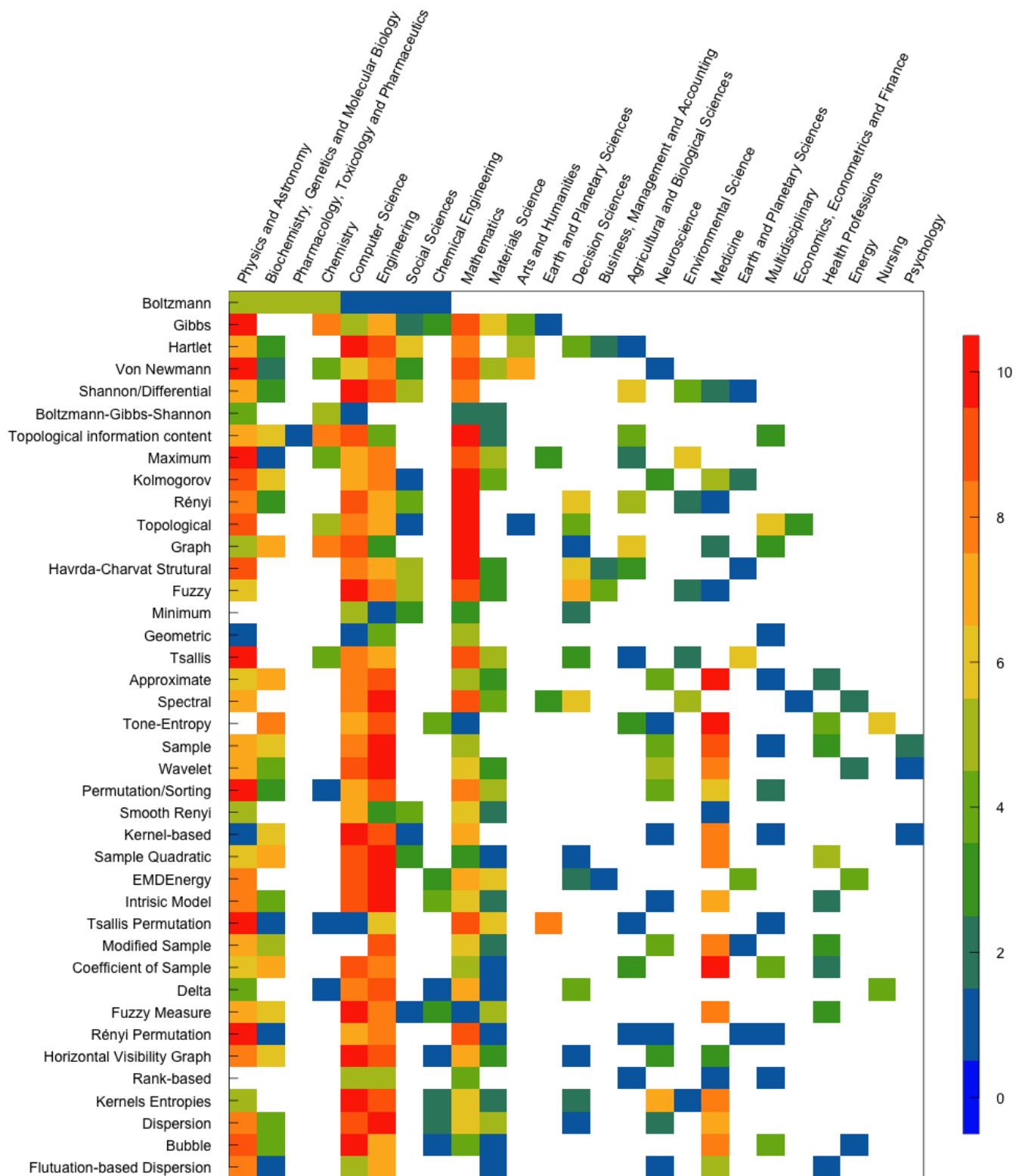
The results from Scopus, displayed in Figure 12, show that among the entropy articles where the number of citations increased significantly are the Shannon/differential (from 754 in 2004 to 2875 in 2019) and the sample (from 16 in 2004 to 525 in 2019). More recently, the number of citations of the papers that introduced dispersion and bubble entropies continuously increased from 1 in 2016 to 42 in 2019 and from 3 in 2017 to 13 in 2019, respectively. In 2014, there were 12 entropy papers more cited than the sample entropy paper, but in 2019 the sample entropy paper was the second most cited. This progress implies that the impact of SampEn on the scientific community has increased significantly. On the other hand, the spectral entropy paper lost positions in 2004, it was the 6th most cited, and moved to position 16 in 2019 (see Figure 12). The results obtained on the WoS and Scopus platform on the impact of entropies on scientific communities complement and reinforce each other.

### 4.2. Areas of Application

To understand each entropy’s areas of application in the scientific community, we consider the ten areas that most cited the paper that introduced the entropy, according to the largest number of records obtained in Web of Science and the Scopus databases. The results are displayed in Figures 13 and 14, respectively. The entropies are ordered in the chronological order of the work that proposed them, meaning the first line corresponds to the oldest entropy paper. There are ten color tones in the figures representing the range from 0 (blue) to 10 (red), where 0 is assigned to the research area that least mentioned the paper that introduced the entropy. In contrast, 10 is assigned to the research area that most cited the paper.



**Figure 13.** The ten areas that most cited each paper introducing entropies according to the *Research Areas* of the WoS. Legend: range 0 (research area least cited)-10 (research area most cited).



**Figure 14.** The ten areas of most cited papers that introduced entropies according to the *Documents by subject area* of the Scopus. Legend: range 0 (research area least cited)-10 (research area most cited).

Note that there are papers in which the entropies were proposed that do not have ten application areas in the databases used. For example, the Rényi permutation entropy has only nine areas in the WoS, and the ranked-based entropy has only six areas of application in Scopus.

Figures 13 and 14 show that, over the years, the set of areas that apply entropies is increasingly numerous. However, we did not obtain the same results for the two figures.

Based on the ten *Research Areas* with the largest number of records obtained on the WoS, we obtained 45 *Research Areas* (Figure 13).

Initially, the research areas that most cited the entropy papers were Physics, Mathematics, and Computer Science. However, according to data from the WoS, there is a time interval between 1991 (approximate entropy) and 2007 (intrinsic model entropy) in which the emerging entropies have more citations in the area of Engineering. Of the thirty works covered by the WoS database, sixteen were most cited in the Engineering area, nine in the Physics area, five in the Computer Science area, and two in the Mathematics area. The second area of research in which the paper that introduced tone-entropy was most cited is Psychology. The second area in which the paper that proposes Tsallis permutation entropy was most cited is Geology. According to the WoS database, the papers that introduced entropies were little cited in medical journals.

Figure 14, based on the top 10 *Documents by subject area* of Scopus, covers all papers that proposed each entropy described in this paper. However, this top 10 contains about half of the *Research Areas* provided on the WoS (25 *Documents by area subject*). Also, we observe that the set of research areas that most cite the works is broader. Of the forty works covered by the Scopus database, the twelve most cited papers are in the area of Computer Science, eight in the area of Physics and Astronomy, seven in the area of Engineering, seven in the area of Mathematics, three in the area of Medicine, one in the area of Chemistry and two papers with the same number of citations in different areas. Over the years, the papers that introduced entropies have been most cited in medical journals by the Scopus database.

According to the Scopus platform, areas such as Business, Management and Accounting, and Agricultural and Biological Sciences are part of the ten main areas that cite the papers that proposed the universe's entropies. When we did the data collection on the WoS, the paper that introduced minimum entropy was mentioned by papers from only four areas, but it is complete in the top 10 of Scopus. On the WoS, the papers that proposed approximate entropy, tone-entropy, and coefficient of sample entropy are the most cited in the Engineering area. In contrast, in the Scopus database, they are most cited in the medical field.

There are several differences between the results obtained in the application areas of the two platforms. However, we also find similarities. Shannon/differential, maximum, topological, graph, minimum, tone-entropy, wavelet, EMDEnergyEn, intrinsic mode, Rényi permutation, kernels, dispersion and fluctuation-based dispersion entropies' introductory papers show the same application areas with more citations in Scopus and WoS database.

We believe that the differences found may be due to the number of citations in the papers and the list of areas for each platform being different. Therefore, it is important to show the results obtained based on the WoS and Scopus since, when the results are the same, it strengthens, while when the results are different, the two analyzes complement each other.

## 5. Conclusions

In this paper, we introduce *The Entropy Universe*. We built this universe by describing in-depth the relationship between the most applied entropies in time-series for different scientific fields, establishing bases for researchers to properly choose the variant of entropy most suitable for their data and intended objectives. We have described the obstacles surrounding the concept of entropy in scientific research. We aim that this work help researchers choosing the best entropy for their time-series analysis. Among the problems, we discussed and reflected on the extension of the concept of entropy from a discrete variable to a continuous variable. Also, we point out that different entropies have the same name and the same entropies have different names.

The papers that proposed entropies have been increasing in the number of citations, and the Shannon/differential, Tsallis, sample, permutation, and approximate had been

the most cited entropies. Permutation/sorting entropy were the ones that most increased the impact on scientific works, in the last sixteen years. Of the entropies proposed in the past five years, kernels and dispersion entropies are the ones that have the greatest impact on scientific research. Based on the ten areas, with the largest number of records of the paper that introduced each new entropy, obtained from WoS and Scopus, the areas that most applied the entropies are Computer Science, Physics, Mathematics, and Engineering. However, there are differences between the results achieved by the two databases. According to the WoS database, the papers that introduced the entropies were rarely cited in medical journals. However, we did not obtain the same result from the Scopus database.

*The Entropy Universe* is an ongoing work since the number of entropies is continually expanding.

**Author Contributions:** Conceptualization, M.R., L.A., A.S. and A.T.; methodology and investigation, M.R., L.C., T.H., A.T., and A.S.; writing, editing and review of manuscript, M.R., L.C., T.H., A.T. and A.S.; review and supervision, L.A. and C.C.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** M.R. acknowledges Fundação para a Ciência e a Tecnologia (FCT) under scholarship SFRH/BD/138302/2018. A.S. acknowledges funds of Laboratório de Sistemas Informáticos de Grande Escala (LASIGE) Research Unit, Ref. UIDB/00408/2020, funds of Instituto de Telecomunicações (IT) Research Unit, Ref. UIDB/EEA/50008/2020, granted by FCT/MCTES, and the FCT projects Confident PTDC/EEI-CTP/4503/2014, QuantumMining POCI-01-0145-FEDER-031826, and Predict PTDC/CCI-CIF/29877/2017 supported by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), and by the Regional Operational Program of Lisboa. This article was supported by National Funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., within CINTESIS, R&D Unit (reference UIDB42552020).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ApEn	approximate entropy
BEn	bubble entropy
CauchyKE	Cauchy kernel entropy
CKE	circular kernel entropy
CosEn	coefficient of sample entropy
DE	differential entropy
DispEn	dispersion entropy
EKE	exponential kernel entropy
EMD	empirical mode decomposition
EMDEnergyEn	empirical mode decomposition energy entropy
FDispEn	fluctuation-based dispersion entropy
FuzzyEn	fuzzy entropy
FuzzyMEn	fuzzy measure entropy
i.i.d.	independent and identically distributed
IME	intrinsic mode entropy
IMF	intrinsic mode functions
InMDen	intrinsic mode dispersion entropy
KbEn	kernel-based entropy
LKE	Laplacian kernel entropy
mSampEn	modified sample entropy
NTPE	normalized Tsallis permutation entropy

PE	permutation entropy
QSE	quadratic sample entropy
RbE	rank-based entropy
RE	Rényi entropy
RPE	Rényi permutation entropy
SampEn	sample entropy
SE	shannon entropy
SKE	spherical kernel entropy
SortEn	sorting entropy
SpEn	spectral entropy
TE	Tsallis entropy
T-E	tone-entropy
TKE	triangular kernel entropy
TopEn	topological entropy
WaEn	wavelet entropy
WoS	web of science

## References

- Flores Camacho, F.; Ulloa Lugo, N.; Covarrubias Martínez, H. The concept of entropy, from its origins to teachers. *Rev. Mex. Física* **2015**, *61*, 69–80.
- Harris, H.H. Review of Entropy and the Second Law: Interpretation and Misinterpretations. *J. Chem. Educ.* **2014**, *91*, 310–311. [[CrossRef](#)]
- Shaw, D.; Davis, C.H. Entropy and information: A multidisciplinary overview. *J. Am. Soc. Inf. Sci.* **1983**, *34*, 67–74. [[CrossRef](#)]
- Kostic, M.M. The elusive nature of entropy and its physical meaning. *Entropy* **2014**, *16*, 953–967. [[CrossRef](#)]
- Popovic, M. Researchers in an entropy wonderland: A review of the entropy concept. *arXiv* **2017**, arXiv:1711.07326.
- Batten, D.F. A review of entropy and information theory. In *Spatial Analysis of Interacting Economies*; Springer: Berlin/Heidelberg, Germany, 1983; pp. 15–52.
- Amigó, J.M.; Balogh, S.G.; Hernández, S. A brief review of generalized entropies. *Entropy* **2018**, *20*, 813. [[CrossRef](#)] [[PubMed](#)]
- Tempesta, P. Beyond the Shannon–Khinchin formulation: The composability axiom and the universal-group entropy. *Ann. Phys.* **2016**, *365*, 180–197. [[CrossRef](#)]
- Namdari, A.; Li, Z. A review of entropy measures for uncertainty quantification of stochastic processes. *Adv. Mech. Eng.* **2019**, *11*. [[CrossRef](#)]
- Rong, L.; Shang, P. Topological entropy and geometric entropy and their application to the horizontal visibility graph for financial time series. *Nonlinear Dyn.* **2018**, *92*, 41–58. [[CrossRef](#)]
- Blanco, S.; Figliola, A.; Quiroga, R.Q.; Rosso, O.; Serrano, E. Time-frequency analysis of electroencephalogram series. III. Wavelet packets and information cost function. *Phys. Rev. E* **1998**, *57*, 932. [[CrossRef](#)]
- Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
- Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)] [[PubMed](#)]
- Zhao, X.; Shang, P.; Huang, J. Permutation complexity and dependence measures of time series. *EPL Europhys. Lett.* **2013**, *102*, 40005. [[CrossRef](#)]
- Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [[CrossRef](#)] [[PubMed](#)]
- Lake, D.E.; Moorman, J.R. Accurate estimation of entropy in very short physiological time series: The problem of atrial fibrillation detection in implanted ventricular devices. *Am. J. Physiol. Heart Circ. Physiol.* **2011**, *300*, H319–H325. [[CrossRef](#)]
- Xie, H.B.; He, W.X.; Liu, H. Measuring time series regularity using nonlinear similarity-based sample entropy. *Phys. Lett. A* **2008**, *372*, 7140–7146. [[CrossRef](#)]
- Rostaghi, M.; Azami, H. Dispersion entropy: A measure for time-series analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [[CrossRef](#)]
- Xu, L.S.; Wang, K.Q.; Wang, L. Gaussian kernel approximate entropy algorithm for analyzing irregularity of time-series. In Proceedings of the 2005 international conference on machine learning and cybernetics, Guangzhou, China, 15–21 August 2005; Volume 9, pp. 5605–5608.
- Martin, J.S.; Smith, N.A.; Francis, C.D. Removing the entropy from the definition of entropy: Clarifying the relationship between evolution, entropy, and the second law of thermodynamics. *Evol. Educ. Outreach* **2013**, *6*, 30. [[CrossRef](#)]
- Chakrabarti, C.; De, K. Boltzmann–Gibbs entropy: Axiomatic characterization and application. *Int. J. Math. Math. Sci.* **2000**, *23*, 243–251. [[CrossRef](#)]

22. Haubold, H.; Mathai, A.; Saxena, R. Boltzmann-Gibbs entropy versus Tsallis entropy: Recent contributions to resolving the argument of Einstein concerning “Neither Herr Boltzmann nor Herr Planck has given a definition of  $W$ ”? *Astrophys. Space Sci.* **2004**, *290*, 241–245. [[CrossRef](#)]
23. Cariolaro, G. Classical and Quantum Information Theory. In *Quantum Communications*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 573–637.
24. Lindley, D.; O’Connell, J. Boltzmann’s atom: The great debate that launched a revolution in physics. *Am. J. Phys.* **2001**, *69*, 1020. [[CrossRef](#)]
25. Planck, M. On the theory of the energy distribution law of the normal spectrum. *Verh. Deut. Phys. Ges.* **1900**, *2*, 237–245.
26. Gibbs, J.W. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*; C. Scribner’s Sons: Farmington Hills, MI, USA, 1902.
27. Rondoni, L.; Cohen, E. Gibbs entropy and irreversible thermodynamics. *Nonlinearity* **2000**, *13*, 1905. [[CrossRef](#)]
28. Goldstein, S.; Lebowitz, J.L.; Tumulka, R.; Zanghi, N. Gibbs and Boltzmann entropy in classical and quantum mechanics. *arXiv* **2019**, arXiv:1903.11870.
29. Hartley, R.V. Transmission of information 1. *Bell Syst. Tech. J.* **1928**, *7*, 535–563. [[CrossRef](#)]
30. Von Neumann, J. *Mathematische Grundlagen der Quantenmechanik*; Springer: Berlin/Heidelberg, Germany, 1932.
31. Legeza, Ö.; Sólyom, J. Optimizing the density-matrix renormalization group method using quantum information entropy. *Phys. Rev. B* **2003**, *68*, 195116. [[CrossRef](#)]
32. Coles, P.J.; Berta, M.; Tomamichel, M.; Wehner, S. Entropic uncertainty relations and their applications. *Rev. Mod. Phys.* **2017**, *89*, 015002. [[CrossRef](#)]
33. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
34. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1949; pp. 1–117.
35. Weaver, W. Recent contributions to the mathematical theory of communication. *ETC Rev. Gen. Semant.* **1953**, *10*, 261–281.
36. Rioul, O.; LTCl, T.P. This is it: A primer on Shannon’s entropy and information. *L’Information, Semin. Poincare* **2018**, *23*, 43–77.
37. Kline, R.R. *The Cybernetics Moment: Or Why We Call Our Age the Information Age*; JHU Press: Baltimore, MA, USA, 2015.
38. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; Number pt. 11 in Illini books; University of Illinois Press: Urbana, IL, USA, 1963.
39. Smith, J.D. Some observations on the concepts of information-theoretic entropy and randomness. *Entropy* **2001**, *3*, 1–11. [[CrossRef](#)]
40. Ochs, W. Basic properties of the generalized Boltzmann-Gibbs-Shannon entropy. *Rep. Math. Phys.* **1976**, *9*, 135–155. [[CrossRef](#)]
41. Plastino, A.; Plastino, A.; Vucetich, H. A quantitative test of Gibbs’ statistical mechanics. *Phys. Lett. A* **1995**, *207*, 42–46. [[CrossRef](#)]
42. Stratonovich, R. The entropy of systems with a random number of particles. *Sov. Phys. JETP-USSR* **1955**, *1*, 254–261.
43. Khinchin, A.Y. *Mathematical Foundations of Information Theory*; Courier Corporation: Dover, NY, USA, 2013.
44. Cover, T.M.; Thomas, J.A. *Elem. Inf. Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
45. Aczél, J.; Daróczy, Z. *On Measures of Information and Their Characterizations*; Academic Press: New York, NY, USA, 1975; p. 168.
46. Kullback, S. *Information Theory and Statistics*; Courier Corporation: Dover, New York, USA, 1997.
47. Chakrabarti, C.; Chakrabarty, I. Shannon entropy: Axiomatic characterization and application. *Int. J. Math. Math. Sci.* **2005**, *2005*, 2847–2854. [[CrossRef](#)]
48. Marsh, C. *Introduction to Continuous Entropy*; Department of Computer Science, Princeton University: Princeton, NJ, USA, 2013.
49. Kapur, J.N.; Kesavan, H.K. Entropy optimization principles and their applications. In *Entropy and Energy Dissipation in Water Resources*; Springer: Berlin/Heidelberg, Germany, 1992; pp. 3–20.
50. Borowska, M. Entropy-based algorithms in the analysis of biomedical signals. *Stud. Logic Gramm. Rhetor.* **2015**, *43*, 21–32. [[CrossRef](#)]
51. Oida, E.; Moritani, T.; Yamori, Y. Tone-entropy analysis on cardiac recovery after dynamic exercise. *J. Appl. Physiol.* **1997**, *82*, 1794–1801. [[CrossRef](#)] [[PubMed](#)]
52. Rosso, O.A.; Blanco, S.; Yordanova, J.; Kolev, V.; Figliola, A.; Schürmann, M.; Başar, E. Wavelet entropy: A new tool for analysis of short duration brain electrical signals. *J. Neurosci. Methods* **2001**, *105*, 65–75. [[CrossRef](#)]
53. Yu, Y.; Junsheng, C. A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *J. Sound Vib.* **2006**, *294*, 269–277. [[CrossRef](#)]
54. Chen, B.; Zhu, Y.; Hu, J.; Pri, J.C.  $\Delta$ -Entropy: Definition, properties and applications in system identification with quantized data. *Inf. Sci.* **2011**, *181*, 1384–1402. [[CrossRef](#)]
55. Kolmogorov, A.N. *A New Metric Invariant of Transient Dynamical Systems and Automorphisms in Lebesgue Spaces*; Doklady Akademii Nauk; Russian Academy of Sciences: Moscow, Russia, 1958; Volume 119, pp. 861–864.
56. Wong, K.S.; Salleh, Z. A note on the notions of topological entropy. *Earthline J. Math. Sci.* **2018**, 1–16. [[CrossRef](#)]
57. Sinai, I. On the concept of entropy for a dynamic system. *Dokl. Akad. Nauk. SSSR* **1959**, *124*, 768–771.
58. Farmer, J.D. Information dimension and the probabilistic structure of chaos. *Z. Naturforschung A* **1982**, *37*, 1304–1326. [[CrossRef](#)]
59. Frigg, R. In what sense is the Kolmogorov-Sinai entropy a measure for chaotic behaviour?—bridging the gap between dynamical systems theory and communication theory. *Br. J. Philos. Sci.* **2004**, *55*, 411–434. [[CrossRef](#)]
60. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906. [[CrossRef](#)] [[PubMed](#)]



61. Orozco-Arroyave, J.R.; Arias-Londono, J.D.; Vargas-Bonilla, J.F.; Nöth, E. Analysis of speech from people with Parkinson's disease through nonlinear dynamics. In *International Conference on Nonlinear Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 112–119.
62. Zmeskal, O.; Dzik, P.; Vesely, M. Entropy of fractal systems. *Comput. Math. Appl.* **2013**, *66*, 135–146. [[CrossRef](#)]
63. Henriques, T.; Gonçalves, H.; Antunes, L.; Matias, M.; Bernardes, J.; Costa-Santos, C. Entropy and compression: Two measures of complexity. *J. Eval. Clin. Pract.* **2013**, *19*, 1101–1106. [[CrossRef](#)]
64. Eckmann, J.P.; Ruelle, D. Ergodic theory of chaos and strange attractors. In *The Theory of Chaotic Attractors*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 273–312.
65. Xiong, W.; Faes, L.; Ivanov, P.C. Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations. *Phys. Rev. E* **2017**, *95*, 062114. [[CrossRef](#)]
66. Adler, R.L.; Konheim, A.G.; McAndrew, M.H. Topological entropy. *Trans. Am. Math. Soc.* **1965**, *114*, 309–319. [[CrossRef](#)]
67. Feng, D.J.; Huang, W. Variational principles for topological entropies of subsets. *J. Funct. Anal.* **2012**, *263*, 2228–2254. [[CrossRef](#)]
68. Nilsson, J. On the entropy of a family of random substitutions. *Monatshefte Math.* **2012**, *168*, 563–577. [[CrossRef](#)]
69. Bowen, R. Entropy for group endomorphisms and homogeneous spaces. *Trans. Am. Math. Soc.* **1971**, *153*, 401–414. [[CrossRef](#)]
70. Cánovas, J.; Rodríguez, J. Topological entropy of maps on the real line. *Topol. Appl.* **2005**, *153*, 735–746. [[CrossRef](#)]
71. Bowen, R. Topological entropy for noncompact sets. *Trans. Am. Math. Soc.* **1973**, *184*, 125–136. [[CrossRef](#)]
72. Handel, M.; Kitchens, B.; Rudolph, D.J. Metrics and entropy for non-compact spaces. *Isr. J. Math.* **1995**, *91*, 253–271. [[CrossRef](#)]
73. Addabbo, R.; Blackmore, D. A dynamical systems-based hierarchy for Shannon, metric and topological entropy. *Entropy* **2019**, *21*, 938. [[CrossRef](#)]
74. Ghys, E.; Langevin, R.; Walczak, P. Entropie géométrique des feuilletages. *Acta Math.* **1988**, *160*, 105–142. [[CrossRef](#)]
75. Hurder, S. *Entropy and Dynamics of C1 Foliations*; University of Illinois: Chicago, IL, USA, 2020.
76. Biś, A. Entropy of distributions. *Topol. Appl.* **2005**, *152*, 2–10. [[CrossRef](#)]
77. Hurder, S. Lectures on foliation dynamics: Barcelona 2010. *arXiv* **2011**, arXiv:1104.4852.
78. Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuno, J.C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4972–4975. [[CrossRef](#)] [[PubMed](#)]
79. Rényi, A. 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: California, CA, USA, 1961.
80. Bosyk, G.; Portesi, M.; Plastino, A. Collision entropy and optimal uncertainty. *Phys. Rev. A* **2012**, *85*, 012108. [[CrossRef](#)]
81. Easwaramoorthy, D.; Uthayakumar, R. Improved generalized fractal dimensions in the discrimination between healthy and epileptic EEG signals. *J. Comput. Sci.* **2011**, *2*, 31–38. [[CrossRef](#)]
82. Müller, M.P.; Pastena, M. A generalization of majorization that characterizes Shannon entropy. *IEEE Trans. Inf. Theory* **2016**, *62*, 1711–1720. [[CrossRef](#)]
83. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620. [[CrossRef](#)]
84. Posner, E. Random coding strategies for minimum entropy. *IEEE Trans. Inf. Theory* **1975**, *21*, 388–391. [[CrossRef](#)]
85. Chevalier, C.; Fouque, P.A.; Pointcheval, D.; Zimmer, S. Optimal randomness extraction from a Diffie-Hellman element. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 572–589.
86. Renner, R.; Wolf, S. Smooth Rényi entropy and applications. In *Proceedings of the International Symposium on Information Theory*, Chicago, IL, USA, 27 June–2 July 2004; p. 233.
87. Lake, D.E. Renyi entropy measures of heart rate Gaussianity. *IEEE Trans. Biomed. Eng.* **2005**, *53*, 21–27. [[CrossRef](#)]
88. Botta-Dukát, Z. Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *J. Veg. Sci.* **2005**, *16*, 533–540. [[CrossRef](#)]
89. Rao, C.R. Diversity and dissimilarity coefficients: A unified approach. *Theor. Popul. Biol.* **1982**, *21*, 24–43. [[CrossRef](#)]
90. Havrda, J.; Charvát, F. Quantification method of classification processes. Concept of structural  $\alpha$ -entropy. *Kybernetika* **1967**, *3*, 30–35.
91. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [[CrossRef](#)]
92. Zunino, L.; Pérez, D.; Kowalski, A.; Martín, M.; Garavaglia, M.; Plastino, A.; Rosso, O. Fractional Brownian motion, fractional Gaussian noise, and Tsallis permutation entropy. *Phys. A Stat. Mech. Its Appl.* **2008**, *387*, 6057–6068. [[CrossRef](#)]
93. Bandt, C. Ordinal time series analysis. *Ecol. Model.* **2005**, *182*, 229–238. [[CrossRef](#)]
94. Citi, L.; Guffanti, G.; Mainardi, L. Rank-based multi-scale entropy analysis of heart rate variability. In *Proceedings of the Computing in Cardiology 2014*, Cambridge, MA, USA, 7–10 September 2014; pp. 597–600.
95. Manis, G.; Aktaruzzaman, M.; Sassi, R. Bubble entropy: An entropy almost free of parameters. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2711–2718.
96. Friend, E.H. Sorting on electronic computer systems. *J. ACM (JACM)* **1956**, *3*, 134–168. [[CrossRef](#)]
97. Astrachan, O. *Bubble Sort: An Archaeological Algorithmic Analysis*; ACM SIGCSE Bulletin; ACM: New York, NY, USA, 2003; Volume 35, pp. 1–5.
98. Bodini, M.; Rivolta, M.W.; Manis, G.; Sassi, R. Analytical Formulation of Bubble Entropy for Autoregressive Processes. In *Proceedings of the 2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)*, Pisa, Italy, 15 July 2020; pp. 1–2.

99. Dehmer, M.; Mowshowitz, A. A history of graph entropy measures. *Inf. Sci.* **2011**, *181*, 57–78. [[CrossRef](#)]
100. Rashevsky, N. Life, information theory, and topology. *Bull. Math. Biophys.* **1955**, *17*, 229–235. [[CrossRef](#)]
101. Trucco, E. A note on the information content of graphs. *Bull. Math. Biophys.* **1956**, *18*, 129–135. [[CrossRef](#)]
102. Mowshowitz, A. Entropy and the complexity of graphs: I. An index of the relative complexity of a graph. *Bull. Math. Biophys.* **1968**, *30*, 175–204. [[CrossRef](#)]
103. Mowshowitz, A. Entropy and the complexity of graphs: II. The information content of digraphs and infinite graphs. *Bull. Math. Biophys.* **1968**, *30*, 225–240. [[CrossRef](#)]
104. Mowshowitz, A. Entropy and the complexity of graphs: III. Graphs with prescribed information content. *Bull. Math. Biophys.* **1968**, *30*, 387–414. [[CrossRef](#)]
105. Mowshowitz, A. Entropy and the complexity of graphs: IV. Entropy measures and graphical structure. *Bull. Math. Biophys.* **1968**, *30*, 533–546. [[CrossRef](#)]
106. Körner, J. Coding of an information source having ambiguous alphabet and the entropy of graphs. In Proceedings of the 6th Prague Conference on Information Theory, Prague, Czech Republic, 18–23 August 1973; pp. 411–425.
107. Csiszár, I.; Körner, J.; Lovász, L.; Marton, K.; Simonyi, G. Entropy splitting for antiblocking corners and perfect graphs. *Combinatorica* **1990**, *10*, 27–40. [[CrossRef](#)]
108. Simonyi, G. Graph entropy: A survey. *Comb. Optim.* **1995**, *20*, 399–441.
109. Zhu, G.; Li, Y.; Wen, P.P.; Wang, S. Analysis of alcoholic EEG signals based on horizontal visibility graph entropy. *Brain Informatics* **2014**, *1*, 19–25. [[CrossRef](#)]
110. Luque, B.; Lacasa, L.; Ballesteros, F.; Luque, J. Horizontal visibility graphs: Exact results for random time series. *Phys. Rev. E* **2009**, *80*, 046103. [[CrossRef](#)] [[PubMed](#)]
111. Zhu, G.; Li, Y.; Wen, P.P. Epileptic seizure detection in EEGs signals using a fast weighted horizontal visibility algorithm. *Comput. Methods Programs Biomed.* **2014**, *115*, 64–75. [[CrossRef](#)]
112. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)] [[PubMed](#)]
113. Liang, Z.; Wang, Y.; Sun, X.; Li, D.; Voss, L.J.; Sleigh, J.W.; Hagihira, S.; Li, X. EEG entropy measures in anesthesia. *Front. Comput. Neurosci.* **2015**, *9*, 16. [[CrossRef](#)]
114. Pincus, S.M.; Gladstone, I.M.; Ehrenkranz, R.A. A regularity statistic for medical data analysis. *J. Clin. Monit.* **1991**, *7*, 335–345. [[CrossRef](#)]
115. Amoud, H.; Snoussi, H.; Hewson, D.; Doussot, M.; Duchene, J. Intrinsic mode entropy for nonlinear discriminant analysis. *IEEE Signal Process. Lett.* **2007**, *14*, 297–300. [[CrossRef](#)]
116. Azami, H.; Escudero, J. Amplitude-and fluctuation-based dispersion entropy. *Entropy* **2018**, *20*, 210. [[CrossRef](#)] [[PubMed](#)]
117. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [[CrossRef](#)]
118. De Luca, A.; Termini, S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Inf. Control* **1972**, *20*, 301–312. [[CrossRef](#)]
119. Parkash, C. *Fuzzy and Non Fuzzy Measures of Information and Their Applications to Queueing Theory*; Guru Nanak Dev University: Punjab, India, 2014.
120. Chen, W.; Wang, Z.; Xie, H.; Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [[CrossRef](#)]
121. Yeniyayla, Y. Fuzzy Entropy and Its Application. Ph.D. Thesis, Dokuz Eylül University, Fen Bilimleri Enstitüsü, İzmir, Turkey, 2011.
122. Liu, C.; Zhao, L. Using fuzzy measure entropy to improve the stability of traditional entropy measures. In Proceedings of the 2011 Computing in Cardiology, Hangzhou, China, 18–21 September 2011; pp. 681–684.
123. Zaylala, A.; Saleh, S.; Karamah, F.; Nahas, Z.; Bouakaz, A. Cascade of nonlinear entropy and statistics to discriminate fetal heart rates. In Proceedings of the 2016 3rd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Beirut, Lebanon, 13–15 July 2016; pp. 152–157.
124. Mekyska, J.; Janousova, E.; Gomez-Vilda, P.; Smekal, Z.; Rektorova, I.; Eliasova, I.; Kostalova, M.; Mrackova, M.; Alonso-Hernandez, J.B.; Faundez-Zanuy, M.; et al. Robust and complex approach of pathological speech signal analysis. *Neurocomputing* **2015**, *167*, 94–111. [[CrossRef](#)]
125. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 068102. [[CrossRef](#)]
126. Zhang, Y.C. Complexity and  $1/f$  noise. A phase space approach. *J. Phys. I* **1991**, *1*, 971–977. [[CrossRef](#)]
127. Hsu, C.F.; Wei, S.Y.; Huang, H.P.; Hsu, L.; Chi, S.; Peng, C.K. Entropy of entropy: Measurement of dynamical complexity for biological systems. *Entropy* **2017**, *19*, 550. [[CrossRef](#)]
128. Wu, S.D.; Wu, C.W.; Lin, S.G.; Wang, C.C.; Lee, K.Y. Time series analysis using composite multiscale entropy. *Entropy* **2013**, *15*, 1069–1084. [[CrossRef](#)]
129. Valencia, J.F.; Porta, A.; Vallverdú, M.; Claria, F.; Baranowski, R.; Orłowska-Baranowska, E.; Caminal, P. Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects. *IEEE Trans. Biomed. Eng.* **2009**, *56*, 2202–2213. [[CrossRef](#)] [[PubMed](#)]

130. Wu, S.D.; Wu, C.W.; Lee, K.Y.; Lin, S.G. Modified multiscale entropy for short-term time series analysis. *Phys. A Stat. Mech. Appl.* **2013**, *392*, 5865–5873. [[CrossRef](#)]
131. Costa, M.D.; Goldberger, A.L. Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. *Entropy* **2015**, *17*, 1197–1203. [[CrossRef](#)] [[PubMed](#)]
132. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918. [[CrossRef](#)]
133. Kolmogorov, A.N. Three approaches to the quantitative definition of information. *Probl. Inf. Transm.* **1965**, *1*, 1–7.
134. Solomonoff, R.J. A formal theory of inductive inference. Part I. *Inf. Control* **1964**, *7*, 1–22. [[CrossRef](#)]
135. Chaitin, G.J. On the length of programs for computing finite binary sequences. *J. ACM (JACM)* **1966**, *13*, 547–569. [[CrossRef](#)]
136. Teixeira, A.; Matos, A.; Souto, A.; Antunes, L. Entropy measures vs. Kolmogorov complexity. *Entropy* **2011**, *13*, 595–611. [[CrossRef](#)]
137. Zegers, P. Fisher information properties. *Entropy* **2015**, *17*, 4918–4939. [[CrossRef](#)]
138. Fisher, R.A. Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*; Cambridge University Press: Cambridge, UK, 1925; Volume 22, pp. 700–725.
139. Blahut, R.E. *Principles and Practice of Information Theory*; Addison-Wesley Longman Publishing Co., Inc.: Cambridge, MA, USA, 1987.
140. Stam, A.J. Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control* **1959**, *2*, 101–112. [[CrossRef](#)]
141. Borzadaran, G.M. Relationship between entropies, variance and Fisher information. In Proceedings of the AIP Conference Proceedings; American Institute of Physics: Melville, NY, USA, 2001; Volume 568, pp. 139–144.
142. Mukherjee, D.; Ratnaparkhi, M.V. On the functional relationship between entropy and variance with related applications. *Commun. Stat. Theory Methods* **1986**, *15*, 291–311.
143. Toomaj, A.; Di Crescenzo, A. Connections between weighted generalized cumulative residual entropy and variance. *Mathematics* **2020**, *8*, 1072. [[CrossRef](#)]
144. Gibson, J. Entropy power, autoregressive models, and mutual information. *Entropy* **2018**, *20*, 750. [[CrossRef](#)]
145. Ledoux, M.; Nair, C.; Wang, Y.N. *Log-Convexity of Fisher Information along Heat Flow*; University of Toulouse: Toulouse, France, 2021.
146. Vieira, E.; Gomes, J. A comparison of Scopus and Web of Science for a typical university. *Scientometrics* **2009**, *81*, 587–600. [[CrossRef](#)]
147. Liu, W.; Tang, L.; Hu, G. Funding information in Web of Science: An updated overview. *arXiv* **2020**, arXiv:2001.04697.
148. Franceschini, F.; Maisano, D.; Mastrogiacomo, L. Empirical analysis and classification of database errors in Scopus and Web of Science. *J. Inf.* **2016**, *10*, 933–953. [[CrossRef](#)]
149. Meho, L.I.; Yang, K. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *J. Am. Soc. Inf. Sci. Technol.* **2007**, *58*, 2105–2125. [[CrossRef](#)]
150. Mugnaini, R.; Strehl, L. Recuperação e impacto da produção científica na era Google: Uma análise comparativa entre o Google Acadêmico e a Web of Science. In *Revista Eletrônica de Biblioteconomia e ciência da Informação*; Encontros Bibli: Florianópolis, Brazil, 2008; n. esp., pp. 92–105.
151. Falagas, M.E.; Pitsouni, E.I.; Malietzis, G.A.; Pappas, G. Comparison of PubMed, Scopus, Web of Science, and Google scholar: strengths and weaknesses. *FASEB J.* **2008**, *22*, 338–342. [[CrossRef](#)]
152. Scopus. Content Selection and Advisory Board (CSAB). Available online: <https://www.elsevier.com/solutions/scopus/how-scopus-works/content> (accessed on 18 June 2020).