

Article

Evaluating Temporal Correlations in Time Series Using Permutation Entropy, Ordinal Probabilities and Machine Learning

Bruno R. R. Boaretto ¹, Roberto C. Budzinski ^{2,3}, Kalel L. Rossi ⁴, Thiago L. Prado ¹, Sergio R. Lopes ¹
and Cristina Masoller ^{5,*}

- ¹ Department of Physics, Universidade Federal do Paraná, Curitiba 81531-980, Brazil; brrb12@fisica.ufpr.br (B.R.R.B.); prado@fisica.ufpr.br (T.L.P.); lopes@fisica.ufpr.br (S.R.L.)
² Department of Mathematics, Western University, London, ON N6A 3K7, Canada; rbudzins@uwo.ca
³ Brain and Mind Institute, Western University, London, ON N6A 3K7, Canada
⁴ Theoretical Physics/Complex Systems, ICBM, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany; kalel.rossi@uni-oldenburg.de
⁵ Department of Physics, Universitat Politecnica de Catalunya, 08034 Barcelona, Spain
* Correspondence: cristina.masoller@upc.edu

Abstract: Time series analysis comprises a wide repertoire of methods for extracting information from data sets. Despite great advances in time series analysis, identifying and quantifying the strength of nonlinear temporal correlations remain a challenge. We have recently proposed a new method based on training a machine learning algorithm to predict the temporal correlation parameter, α , of flicker noise (FN) time series. The algorithm is trained using as input features the probabilities of ordinal patterns computed from FN time series, $x_{\alpha}^{FN}(t)$, generated with different values of α . Then, the ordinal probabilities computed from the time series of interest, $x(t)$, are used as input features to the trained algorithm and that returns a value, α_e , that contains meaningful information about the temporal correlations present in $x(t)$. We have also shown that the difference, Ω , of the permutation entropy (PE) of the time series of interest, $x(t)$, and the PE of a FN time series generated with $\alpha = \alpha_e$, $x_{\alpha_e}^{FN}(t)$, allows the identification of the underlying determinism in $x(t)$. Here, we apply our methodology to different datasets and analyze how α_e and Ω correlate with well-known quantifiers of chaos and complexity. We also discuss the limitations for identifying determinism in highly chaotic time series and in periodic time series contaminated by noise. The open source algorithm is available on Github.

Keywords: ordinal analysis; symbolic analysis; machine learning; time series analysis; permutation entropy; complexity; chaos; noise



Citation: Boaretto, B.R.R.; Budzinski, R.C.; Rossi, K.L.; Prado, T.L.; Lopes, S.R.; Masoller, C. Evaluating Temporal Correlations in Time Series Using Permutation Entropy, Ordinal Probabilities and Machine Learning. *Entropy* **2021**, *23*, 1025. <https://doi.org/10.3390/e23081025>

Academic Editors: Miguel C. Soriano and Luciano Zunino

Received: 8 July 2021

Accepted: 3 August 2021

Published: 9 August 2021

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



Copyright: © 2021 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

Thanks to huge advances in data science and computing power, a wide repertoire of time series analysis methods [1–7] are available for the quantitative characterization of time series and are routinely used in all fields of science and technology, social sciences, economy and finance, etc. Since different methods have different requirements and involve different approximations, no single method can be expected to perform well over all types of data. Therefore, despite huge advances, extracting reliable information from stochastic or high dimensional signals remains a challenge. As any algorithm will return, at least, a number (i.e., a “feature” that encapsulates some property of the time series), in order to interpret the information in the obtained features and to assess the performance of different algorithms, appropriate surrogates [8] or a “reference model” (where the systems that generates the data are known) need to be used. The comparison of the features obtained from the time series of interest with those obtained from surrogate time series or from

reference time series allows testing (and even quantifying) some particular property of the time series of interest.

We have recently proposed a new method for estimating the strength of the temporal correlations in a given time series, which uses flicker noise (FN), a fully stochastic process, as the reference model [9]. A FN time series, $x_{\alpha}^{FN}(t)$, is characterized by a power spectrum $P(f) \propto 1/f^{\alpha}$, with α being a parameter that quantifies the temporal correlations present in the signal [10]. The method proposed in [9] combines the use of symbolic ordinal analysis [11,12] and machine learning (ML): We utilize the ordinal probabilities computed from FN time series generated with different α values as input features to a ML algorithm. The algorithm is trained to return the value of α_e , which estimates the real α value of a FN time series from the $D!$ probabilities of the ordinal patterns of length D , calculated from the same time series. Then, after the training stage, the ordinal probabilities now computed from a time series of interest, $x(t)$, are provided as input features to the ML algorithm that returns a value, α_e , which encapsulates reliable information about the strength of the temporal correlations present in $x(t)$. By calculating the difference, Ω , between the permutation entropy (PE) of $x(t)$ and that of a FN time series generated with $\alpha = \alpha_e$, $x_{\alpha_e}^{FN}(t)$, we were also able to identify determinism in $x(t)$.

Our approach is, thus, based on reducing a large number of features (with $D = 6$, we have $D! = 720$ ordinal probabilities) to only two: the α_e value returned by the ML algorithm; and the permutation entropy, \bar{S} , computed with the ordinal probabilities. Dimensionality reduction is a well known technique [13–15] that has been used to tackle a variety of problems. With ordinal probabilities, for instance, it is possible to distinguish between noise and chaos by reducing the set of probabilities to only two features—the permutation entropy and the complexity; or the permutation entropy and the Fisher information—as demonstrated in [16–19]. In our methodology, we not only apply dimensionality reduction but also use a fully stochastic “reference” FN time series: we compare the value of \bar{S} of the time series of interest with that of a FN time series generated with $\alpha = \alpha_e$. We have shown that the entropy difference, Ω , may provide good contrast for distinguishing fully stochastic time series from a time series with a degree of determinism [9].

The method we proposed has in fact a large degree of flexibility because, instead of ordinal analysis and the permutation entropy, different symbolization rules [20,21] and different entropies [22–24] could be tested. In addition, while we use a simple artificial neural network, other algorithms could be evaluated. Different combinations may provide different results and particular combinations may result in optimized performance for the analysis of particular types of time series.

We have shown that the algorithm returns meaningful information even from time series that are very short: for the synthetic examples considered in [9], we could distinguish whether the dynamics is mainly chaotic or stochastic with only 100 data points. However, an open question is as follows: What are the limitations in terms of the level of chaos, the level of noise and the length of the time series? Here, we address this issue by using as examples the time series generated with the Logistic map, the βx map and the Schuster map. We also address the following questions: Can we distinguish a highly chaotic time series from a stochastic one? Can we identify a periodic signal hidden by noise? In addition, to gain insight into the information encapsulated by α_e and Ω , we contrast them with well known quantifiers of chaos and complexity: the maximum Lyapunov exponent and the ordinal-based statistical complexity measure [25].

The organization of the paper is as follows: Section 2 describes the methodology, Section 3 presents the datasets and systems analyzed, Section 4 presents the results and Section 5 presents the discussion and our conclusions.

2. Methodology

The methodology proposed in [9] can be described in a few steps:

1. Calculate the ordinal probabilities (OPs) of a large set of FN time series generated with different values of α and use them as features to train a ML algorithm to return the (known) value of α_e ;
2. Calculate the OPs of the time series of interest, $x(t)$, and use them as features to the trained ML algorithm, which returns a value α_e (see Section 2.1);
3. Generate a FN time series with $\alpha = \alpha_e$ and calculate its permutation entropy (PE), \bar{S}_{FN} (see Section 2.2);
4. Calculate the relative difference, Ω , between the PE of $x(t)$, \bar{S} and \bar{S}_{FN} :

$$\Omega = \frac{|\bar{S}_{\text{FN}} - \bar{S}|}{\bar{S}_{\text{FN}}}; \quad (1)$$

5. Use the value of α_e to quantify the strength of the temporal correlations in the time series of interest and use the value of Ω to identify underlying determinism: if $\Omega \approx 0$, $x(t)$ is mainly stochastic, otherwise there is some determinism.

In the implementation proposed in [9], the probabilities of the 720 ordinal patterns of length $D = 6$ (described in the next section) were used as input features to the ML algorithm. Then, these features were reduced to only two: the scalar value returned by the ML algorithm, α_e , which quantifies the temporal correlations presented in the time series of interest $x(t)$; and the permutation entropy, \bar{S} (Equation (4)). Then, the value of \bar{S} was compared with the PE of a FN time series generated with $\alpha = \alpha_e$ and the relative difference, Ω (Equation (1)), was found to provide contrast for identifying determinism in $x(t)$. The value of Ω can be used to organize a set of time series according to their level of stochasticity (the lower the value of Ω , the larger the stochasticity level) and, by appropriately selecting a threshold value, Ω can be used to classify the time series into two categories: mainly stochastic and mainly deterministic.

2.1. Machine Learning Algorithm

A wide range of ML algorithms are available nowadays. Since we want to regress the information of the features ($D!$ probabilities) into one real value, α_e (a classical scalar regression problem), an appropriate simple option is a feed forward artificial neural network (ANN). Mathematically, the ANN can be described as follows. Considering a set of inputs \mathbf{X} of $D!$ features with output α_e , the ANN can be sketched by the following:

$$\mathbf{X}' = f(\mathbf{W} * \mathbf{X} + \mathbf{b}), \quad (2)$$

$$\alpha_e = f'(\mathbf{W}' * \mathbf{X}' + \mathbf{b}'), \quad (3)$$

where \mathbf{W} and \mathbf{W}' are matrices of *weights*, \mathbf{b} and \mathbf{b}' are *biases* column vectors, f and f' correspond to activation functions and the “*” symbols corresponds to a tensor product. In this sense, \mathbf{X}' is the result of the transformation $f(\mathbf{W} * \mathbf{X} + \mathbf{b})$, which can be understood as a new representation of the inputs \mathbf{X} . The elements of the tensors \mathbf{W} , \mathbf{W}' , \mathbf{b} , \mathbf{b}' are the parameters of the ANN, which are calibrated in the training state. ANNs are well known, and we refer the reader to our previous work [9] for details about the network structure and the training procedure. We remark that our ANN is a fast and automatic tool and it performed well in all the cases we tested, with a computational time and in a standard notebook, of a few seconds for the analysis of time series with 10^6 data points. However, we do not claim that an ANN is an optimal choice, since different ML algorithms may be even more efficient. It is important to notice also that the optimal choice will likely depend on the characteristics of the time series (length, frequency content, level of noise, etc.).

2.2. Ordinal Analysis and Permutation Entropy

Ordinal analysis and the permutation entropy were proposed by Bandt and Pompe [11] almost 20 years ago and they are now well-known. For their interdisciplinary applications, we refer the reader to a recent *Focus Issue* [12].

Here, we compute the ordinal patterns of length $D = 6$ with the algorithm proposed in [26]. For $D = 6$, there are $6! = 720$ possible patterns. The patterns are calculated with an overlap of $D - 1$ data points, i.e., for a time series with N data points, $N - D$ ordinal patterns are obtained, which are then used to evaluate the ordinal probabilities $\mathcal{P} = \{\mathcal{P}(i); i = 1, \dots, D!\}$. Then, the normalized permutation entropy is calculated as follows.

$$\bar{S} = -\frac{\sum_{i=1}^{D!} \mathcal{P}(i) \ln [\mathcal{P}(i)]}{\ln [D!]}. \quad (4)$$

2.3. Quantifiers of Chaos and Complexity

The Lyapunov exponents measure the local divergence of infinitesimally close trajectories, and positive exponents indicate chaos [27]. While there are important challenges when calculating them from high-dimensional and/or noisy data [3,28], in the case of a one-dimensional dynamical system for which its governing equation is known, the Lyapunov exponent, $x_{n+1} = f(x_n)$, can be straightforwardly calculated as follows.

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_n \ln [|f'(x_n)|]. \quad (5)$$

A popular measure for characterizing complex systems is the statistical complexity measure [25] that takes into account the distance to the most regular state and the most random states of the system. It is defined as follows:

$$C = \bar{S}(\mathcal{P})Q(\mathcal{P}, \mathcal{P}_e), \quad (6)$$

where \bar{S} is the entropy and Q is the distance between the distribution that describes the state of the system, \mathcal{P} , and the equilibrium distribution, \mathcal{P}_e , that maximizes the entropy.

As we use the probabilities of the ordinal patterns [16,29], \bar{S} is the normalized permutation entropy and $\mathcal{P}_e = \{1/D!, 1/D!, \dots, 1/D!\}$. The distance between \mathcal{P} and \mathcal{P}_e is calculated with the Jensen–Shannon divergence:

$$Q(\mathcal{P}, \mathcal{P}_e) = Q_0 \left\{ S\left(\frac{\mathcal{P} + \mathcal{P}_e}{2}\right) - \frac{S(\mathcal{P})}{2} - \frac{S(\mathcal{P}_e)}{2} \right\}, \quad (7)$$

where Q_0 a normalization factor $Q_0 = -2 \left\{ \left(\frac{D!+1}{D!} \right) \ln [D! + 1] - 2 \ln [2D!] + \ln [D!] \right\}^{-1}$.

3. Datasets

We analyze time series generated by the following stochastic or deterministic dynamical systems. Typical time series examples are presented in Figure 1.

3.1. Flicker Noise

Flicker noise (FN), also called as colored noise, is used for training the ML algorithm. FN time series are stochastic and characterized by a temporal correlation coefficient α . The power spectrum of this signal is given by $1/f^\alpha$, and different α values result in different “colors”. We used the open Python library *colorednoise.py* [30,31] to generate FN time series with different α values. An example is shown in panel (e) of Figure 1.

3.2. Uniform Noise

Uniform noise is a stochastic process with no memory and uniform distribution in $[0, 1]$. Panels (a) and (b) of Figure 1 depict an example of a uniformly distributed white noise and its PDF, respectively.

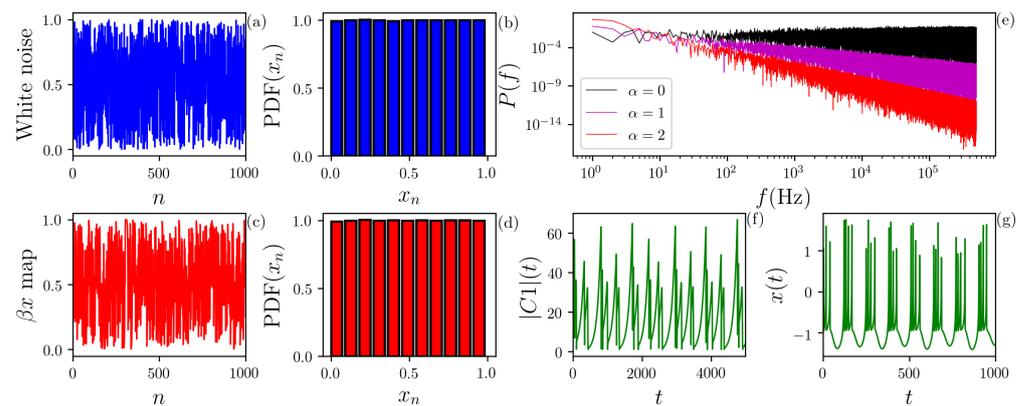


Figure 1. (a,b) Uniformly distributed white noise and its probability distribution function (PDF). (c,d) Time series generated by iteration of the βx map with $\beta = 2$ and its PDF. We observe that the deterministic map depicts a very similar PDF of the white noise. (e) Power spectral density (PSD) of flicker noise (FN) $\alpha = 0$ (black), $\alpha = 1$ (magenta) and $\alpha = 2$ (red). By definition, the PSD of the FN decays as $1/f^\alpha$. (f,g) Time evolution of two continuous chaotic systems: (f) $|C1|$ of the three-waves system and (g) x of the Hindmarsh–Rose model.

3.3. Random Walk

A one-dimensional random walk time series is defined by $x_{n+1} = x_n + \epsilon_n$, where ϵ is a memory-less stochastic Gaussian process with mean 0 and standard deviation 1.

3.4. Periodic or Chaotic Signals Contaminated by Noise

In order to generate time series that represent periodic or chaotic signals contaminated by noise, we used the map equation $Z_n = (1 - \eta)X_n + \eta Y_n$. For a periodic signal, X_n is the sin map, $X_{n+1} = \sin(2\pi n/\tau)$ with period τ ; for a chaotic signal, X_n is the βx map (see below). In both cases, X_t is normalized $[0, 1]$, Y_t is a uniform white noise and $\eta \in [0, 1]$ controls the stochastic component of Z_t ; for $\eta = 0$, the signal is fully deterministic (periodic or chaotic depending on the map used), while for $\eta = 1$ the signal is fully stochastic and memory-less.

3.5. Logistic Map

The Logistic map is a popular nonlinear dynamical system defined by $x_{n+1} = rx_n(1 - x_n)$, where r is the control parameter that allows us to obtain periodic or chaotic signals.

3.6. βx Map

The generalized Bernoulli chaotic map, also known as βx map, is defined by $x_{n+1} = \beta x_n \pmod{1}$, where β controls the dynamical characteristic of the map. Panel (c) and (d) of Figure 1 depict the evolution of a βx map with $\beta = 2$ and its PDF, respectively. Here, we observe that it cannot be visually distinguished from uniform noise; however, the level of chaos can be estimated with the Lyapunov exponent, which has an exact solution in this case: $\lambda = \ln(\beta)$ [27]. Therefore, the higher the β parameter, the more chaotic the signal is.

3.7. Schuster Map

The Schuster map [32] generates intermittent signals with a power spectrum $P(f) \sim 1/f^z$ and is defined as $x_{n+1} = x_n + x_n^z \pmod{1}$, where we used z as a parameter control.

3.8. Lorenz System

The Lorenz system is also a well-known dynamical system [33], defined by three rate equations: $dx/dt = \sigma(y - x)$, $dy/dt = x(R - z) - y$ and $dz/dt = xy - bz$. Here we use typical parameters $\sigma = 16$, $R = 45.92$ and $b = 4$ that generate chaotic trajectories.

3.9. Rossler System

The Rossler system [34] is also a well-known dynamical system, defined by the equations $dx/dt = -y - z$, $dy/dt = x + ay$ and $dz/dt = b + z(x - c)$. Here, we used typical parameters for generating chaotic trajectories: $a = b = 0.2$ and $c = 5.7$.

3.10. Three Waves System

This system is composed of three first-order autonomous differential equations written in terms of the complex slowly varying wave amplitude [35] $dC_1/dt = \gamma_1 C_1 + C_2 C_3$, $dC_2/dt = \gamma_2 C_2 - C_1 C_3^* + i\delta C_2$ and $dC_3/dt = \gamma_3 C_3 - C_1 C_2^*$, where $C_j \in \mathbb{C}$ and we used $\gamma_1 = 1$, $\gamma_2 = \gamma_3 = -17$ and $\delta = 2$. The chaotic dynamics can be observed at the amplitude variations of $|C_1|$, which is depicted in panel (f) of Figure 1.

3.11. Hindmarsh–Rose Model

The Hindmarsh–Rose model [36] is a set of three rate equations that models the neural activity. An example is shown in panel (g) of Figure 1. The equations are $dx/dt = y - ax^3 + bx^2 - z + I$, $dy/dt = c - dx^2 - y$ and $dz/dt = r[s(x - x_r) - z]$ and the parameters used are $a = c = 1$, $b = 3$, $s = 4$, $x_r = -8/5$ and $r = 0.006$.

4. Results

A demonstration of the methodology is presented in Figure 2 that shows, in the plane (\bar{S}, α_e) , the values obtained from time series generated with the dynamical systems described in the previous section. All the time series analyzed possesses $N = 2^{14}$ points and the error bars represent the standard deviation over 1000 time series generated with different initial conditions or noise seeds.

Figure 2a presents results for discrete systems, while Figure 2b, for continuous systems. In both panels the black line represents FN signals generated with different values of α , which are perfectly recovered by the ANN (that returns a value α_e equal to α). As expected, for $\alpha_e = 0$, $\bar{S} \approx 1$ since the FN signal is white noise [37]. For $\alpha_e > 0$, some ordinal patterns occur in the time series more frequently than others, and the value of \bar{S} decreases.

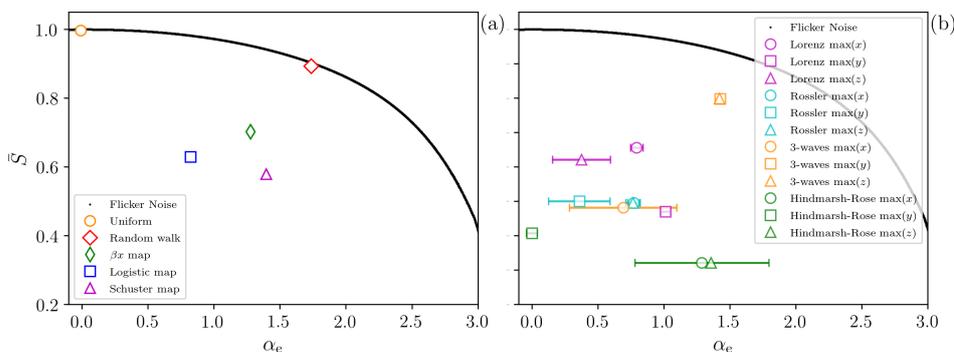


Figure 2. Characterization of stochastic and chaotic time series in the plane (normalized permutation entropy \bar{S} ; correlation coefficient α returned by the ML algorithm). Panel (a) presents results for discrete systems. The black line represents a set of 10,000 FN signals and \bar{S} decreases as the temporal correlation increases. The stochastic cases (uniform noise and random walk) are very close to the FN signals. Otherwise, the chaotic signal (βx , logistic and Schuster maps signals) depict a lower \bar{S} compared to the FN signals. Panel (b) presents results for continuous systems. Here, we analyze the sequence of maxima of each variable and, again, we observe that the vertical distance to the FN curve reveals that the time series are not fully stochastic. All time series possesses $N = 2^{14}$ points and the error bars represent the standard deviation over 1000 time series generated with different initial conditions or noise seeds. Small error bars are not shown.

In panel (a), the orange circle represents time series of uniform white noise, and the ANN returns the correct value $\alpha_e = 0$ (There is almost no dispersion in the returned value of $\alpha_e = 0$, therefore, the error bar is not shown). The red diamond represents

random walk signals; for them, the ANN returns $\alpha_e \approx 1.75$ (Again, there is almost no dispersion in the value of α_e). We observe that the red diamond is very close to the black line, providing a clear indication of the highly stochastic nature of a time series generated by a one-dimensional random walk.

On the other hand, when we analyze chaotic signals (time series from βx map with $\beta = 2$, from logistic map with $r = 4$ and from Schuster map with $z = 0.5$), we observe that the distance between the symbols and the FN noise curve (black line) allows identifying the signals as not fully stochastic, i.e., the distance to the FN curve uncovers determinism in the signals.

For continuous dynamical systems, the results obtained with ordinal analysis strongly depend on the lag time between the data points, which can be performed in multiple ways, for instance using the maxima of each variable or even the first minimum of the mutual information [38]. While this dependence with the lag allows identifying different time scales in a complex signal [39,40], it renders it difficult to compare different signals. Therefore, for the continuous dynamical systems considered (Lorenz, Rossler, 3-waves and Hindmarsh–Rose), the time series analyzed here are the sequence of maxima of each variable of the system. Due to the fact that the variables obey different equations and have oscillations with different properties, we can expect obtaining different values of α_e . Despite a large dispersion in the values obtained, it can be observed in Figure 2b that all the systems depict a substantial distance to the FN curve, clearly revealing that they are not fully stochastic. The dispersion in the permutation entropy values is of the order of 1%; therefore, vertical error bars are not shown.

4.1. Comparison with Standard Quantifiers

Next, we compare, for the three chaotic maps, the quantifiers obtained with our approach, α_e and Ω , with well-known quantifiers of chaos and complexity including the Lyapunov exponent, λ and the ordinal statistical complexity, C (described in Section 2.3).

The results are depicted in Figure 3. Panels (a), (b) and (c) show results for the βx map. For $\beta < 1$, the map is not chaotic since $\lambda = \ln(\beta) < 0$. In this range, Ω decreases while α_e remains constant. The βx map is chaotic for $\beta > 1$, since $\lambda > 0$. At $\beta = 1$, Ω and α_e vary abruptly and both decrease as β increases. There is negative correlation between \bar{S} and Ω and also between λ and Ω . If β is too large, $\Omega \approx 0$ and $\bar{S} \approx 1$ determinism can no longer be identified. We note that the small oscillations of α_e capture the changes in dynamics when β approaches an integer number since the PDF of βx is homogeneous for integer β but becomes inhomogeneous for non-integer values [41], which is not clearly observed in the other quantifiers.

In panels (d), (e) and (f) for the logistic map, periodic windows embedded in chaotic regions are detected in the Lyapunov exponent, and the quantifiers \bar{S} , C and α_e also show abrupt variations. We note that Ω identifies the deterministic nature of the signals since, for the entire interval of r , Ω remains well above zero and \bar{S} remains well below one.

Similar results are found for the Schuster map, panels (g), (h) and (i): Ω is anti-correlated with λ and confirms that the signal is always deterministic since relatively large Ω values are observed for the entire interval of z . Even though a signal generated by the Schuster map has the same power spectrum as Flicker noise, its α_e value varies non-monotonically with z [32], contrary to the line $\alpha = \alpha_e$ that is obtained from FN signals. This is due to the fact that signals generated by the Schuster map and FN signals have different sets of ordinal probabilities.

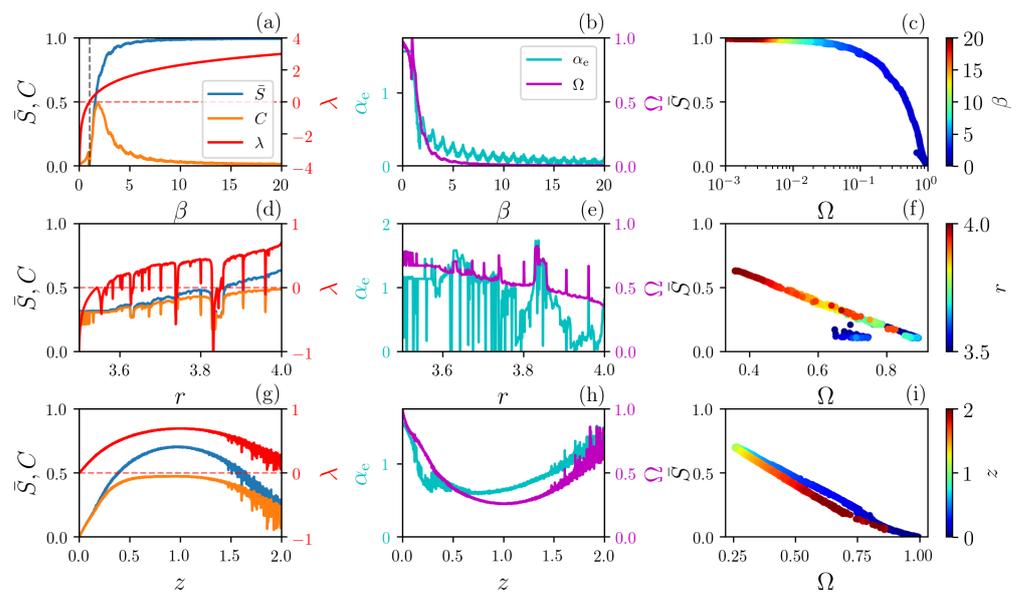


Figure 3. Analysis of deterministic signals generated by chaotic maps. First column depicts the normalized permutation entropy \bar{S} , the complexity C (left vertical scale) and the Lyapunov exponent λ (right scale). Second column depicts the new quantifiers α_e (left scale) and Ω (right scale). The third column presents \bar{S} vs. Ω ; the color codes represent the control parameter of each map. Panels (a–c) illustrate the case of βx map as a function of β ; (d–f) show the Logistic map as a function of r ; and (g–i) depict the Schuster map as a function of z .

4.2. Influence of the Length of the Time Series

The results so far indicate that our methodology is very precise for characterizing chaotic and stochastic signals when long time series are analyzed (in Figures 2 and 3 $N = 2^{14}$). However, an important question is the following: What is the role of the time series size in the analysis? In order to address this issue, we investigate both stochastic and chaotic time series with different lengths N . Figure 4 displays the role of time series length N in \bar{S} (a), in the output of the ANN α_e (b) and in Ω (c) for chaotic and for stochastic signals. Here, chaotic time series are represented by signals generated with the βx map, while stochastic ones are given by uniform white noise and a random walk.

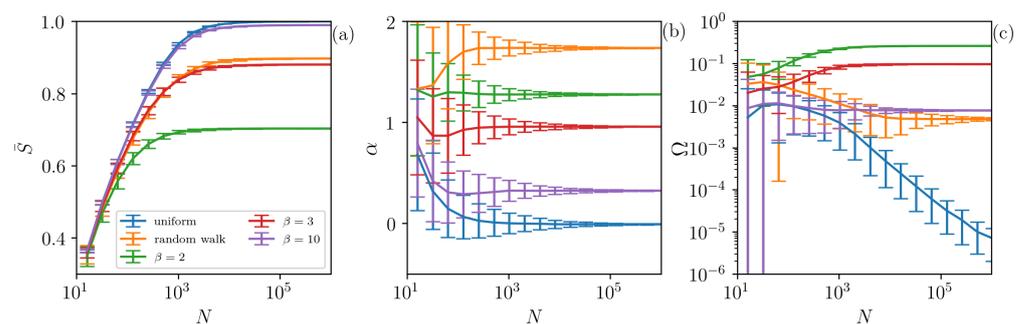


Figure 4. The role of time series length N is studied using \bar{S} (a), the output of the ANN α_e (b) and Ω (c) for both chaotic and stochastic signals. The results are stable for $N \geq 10^4$, indicating the robustness of our methodology. Here, we analyze βx time series and also uniform noise and random walk signals. Despite having similar values of \bar{S} , our method is able to distinguish between the chaotic and stochastic cases even for $N \leq D! = 720$ (number of ordinal patterns) for $\beta = 2$ and $\beta = 3$. However, for $\beta = 10$ the chaoticity is too high ($\lambda = \ln 10$) therefore existing a great local divergence of the trajectory. In this case, the deterministic nature cannot be detected since the ordinal probability distribution is as uniform as a distribution of a white noise signal with $\bar{S} \approx 1$.

One can observe that, in general, for a small number of data points ($N < 100$), our method is unable to distinguish the signals as all of them are overlapping. Interestingly,

for $N \approx 100$, our method is able to distinguish the chaotic signal generated by the β map with $\beta = 2$ (green lines), since Ω depicts values that are higher than those of the other signals. As N is increased, the signal of $\beta = 3$ is also separated and characterized as chaotic. Surprisingly, for $N < D! = 720$ (the number of features), we are already able to distinguish the chaotic signals of $\beta = 2$ and $\beta = 3$ from the stochastic ones (uniform and random walk) by using the quantifier Ω . The results improve as the number of points increases and, for $N \geq 10^4$, the analysis stabilizes. Here, an important point is that even if the chaotic and stochastic signals have similar permutation entropy values (\bar{S}), Ω is able to capture their difference. On the other hand, if the chaoticity of the signal is too high, as observed for $\beta = 10$ (purple lines), $\bar{S} \approx 1$. In this case the ordinal probabilities of the signal are as uniform as those of white noise, which means that we are unable to identify determinism.

Figure 5 shows that the above observations remain robust when noise is added to the chaotic signal. As explained in Section 3.4, we control the amount of noise in the analyzed time series by varying the parameter η . If the time series is too long and has too much noise, Ω does not identify determinism (blue region). On the other hand, if the time series is not long, Ω identifies determinism even when there is no determinism for $\eta = 1$ (orange region). Therefore, for a correct interpretation of the information contained in the value of Ω , we need to compare it with the value of \bar{S} obtained from “reference” time series of the same length as the time series of interest, generated by a known stochastic process such as Flicker noise.

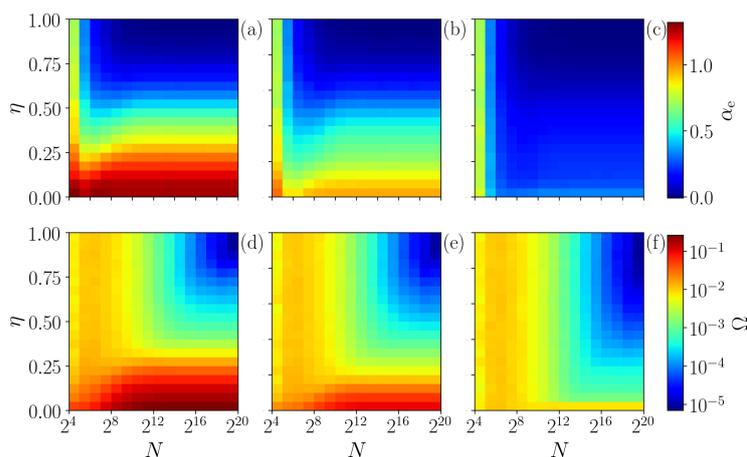


Figure 5. Robustness of the method to the application of additive noise in the βx map. The upper row (a–c) depicts the α_e values in color-code as a function of the percentage of noise η and the series’ size N for $\beta = 2, 3, 10$, respectively. The lower row (d–f) shows Ω for the same simulations.

4.3. Analysis of Periodic Signals Contaminated by Noise

Another challenge to our methodology is the identification of periodic signals with added noise. The results obtained when varying the parameter η (see Section 3.4) are depicted in Figure 6, where panel (a) shows the temporal correlation α_e returned by the ANN and panel (b) shows the quantifier Ω , both as a function of η . Here, the periodic signals with $N = 2^{14}$ points and different frequencies ω are analyzed.

First, one can observe that, for $\eta = 0.0$, all signals are characterized as deterministic (high value of Ω) with a high temporal correlation (high value of α_e), which is expected from a periodic signal. Secondly, as we can also expect for $\eta = 1.0$, all signals are identified as stochastic and memory-less (zero temporal correlation), since the added noise is white. However, for intermediate values of η , an interesting behavior is observed. The frequency of the signal is indeed important, because periodic signals with low frequencies are characterized as stochastic time series for lower values of η (panel (b)). For instance, for $\omega < 10^{-2}$, the signal is characterized as noisy even for very small values of η , while for signals with $\omega \approx 10^{-1}$, the deterministic nature of the original time series is kept for $\eta < 0.35$. As expected, the temporal correlation of the signals α_e (panel (a)) also depends

on the frequency $\omega = 2\pi/\tau$. As η is increased, the higher the frequency of the signal, the slower the way $\alpha_e \rightarrow 0$ (i.e., high-frequency signals loss memory slower than low-frequency signals). Since the noise being added is white and thus, is memory-less, one can expect that $\alpha_e \rightarrow 0$ and $\Omega \rightarrow 0$ for $\eta \rightarrow 1$. However, we see that the temporal correlation α_e can be high even when Ω is very small. Therefore, these signals are first identified as noisy with non-zero time correlation, and, subsequently, they are identified as noisy with no time correlation.

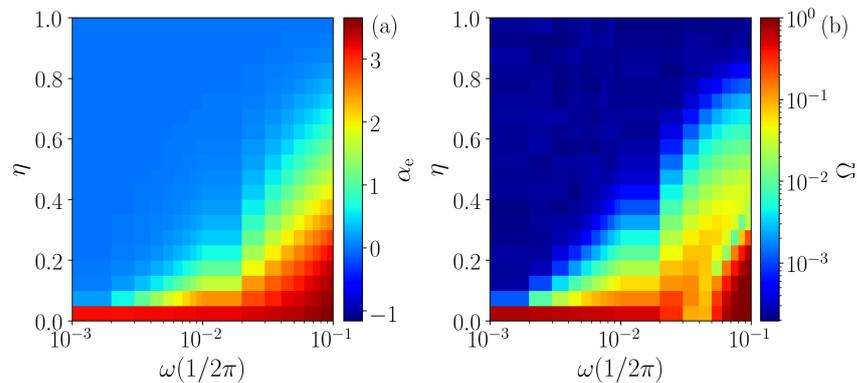


Figure 6. Analysis of periodic signals contaminated with white noise using the temporal correlation α_e (a) and the quantifier Ω (b). The results show that, when the noise strength, η , increases, the deterministic nature of the periodic signal gradually vanishes and, for large enough η , only stochastic dynamics is identified. However, the frequency $\omega = 2\pi/\tau$ of the signal is important, because low frequency signals are identified as stochastic at lower noise levels. Moreover, even when the signal is characterized as stochastic, a nonzero temporal correlation can be estimated.

These results and conclusions are robust for different realizations of the noise, as is shown in Figure 7, which depicts curves with fixed values of ω (upper row) as a function of η . Panel (a) shows the temporal correlation coefficient α_e , while panel (b) shows the quantifier Ω . Following the same idea, the lower row shows three examples with fixed η values as a function of the periodic signal frequency ω . In all cases, the error bars represent the dispersion over 1000 different noise realizations, and we can observe that the trends discussed previously persist over different noise realizations.

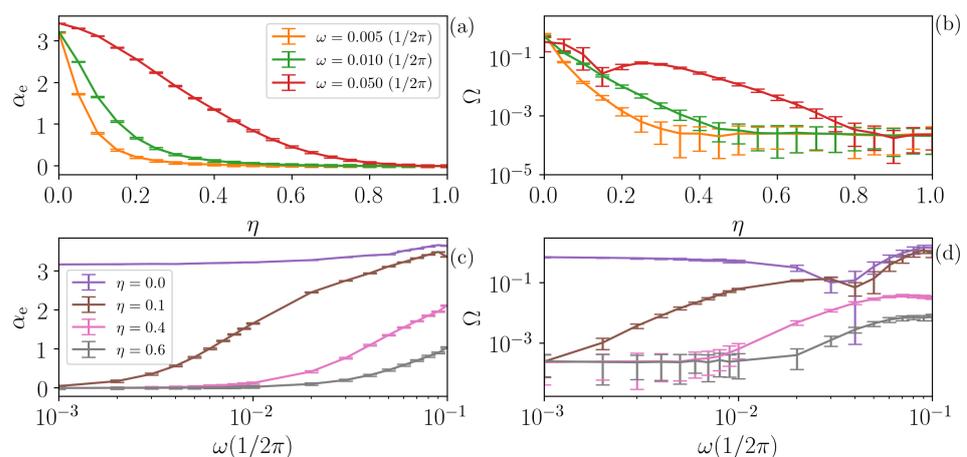


Figure 7. The effects of noise on a periodic signal do not depend on the noise’s specific realization. The error bars indicate the dispersion over 1000 analyses with different realization of the added noise. Here, the upper row represents three examples of α_e (a) and Ω (b) as a function of η for a fixed value of ω . The lower row (c,d) shows three examples as a function of ω for fixed values of η following the same representation. The dispersion in all cases is sufficiently low that the trends discussed in Figure 6 remain.

4.4. Analysis of Two Stochastic Processes

In experimental stochastic systems, two (or more) stochastic dynamics are often present. In order to test the performance of the algorithm in this situation, we consider the same approach as in the previous section, $Z_n = (1 - \eta)X_n + \eta Y_n$, but now X_n is not a periodic signal but a Flicker noise. The results are depicted in Figure 8, where panels (a) and (b) present α_e and Ω , respectively, and are calculated from 1000 FN time series with $\alpha = 0$ (purple), $\alpha = 1$ (brown) and $\alpha = 2$ (pink). For $\eta = 0$ (FN noise) $\alpha_e = \alpha$, as η increases due to the influence of the uniform white noise, α_e decreases to 0 and $\alpha = 2$ (pink line) decays faster due to the slower dynamics, while $\alpha = 0$ (purple) does not change since there is no time correlation in both X_n and Y_n . Ω is extremely low for all cases (<0.01) identifying the full stochasticity of the time series.

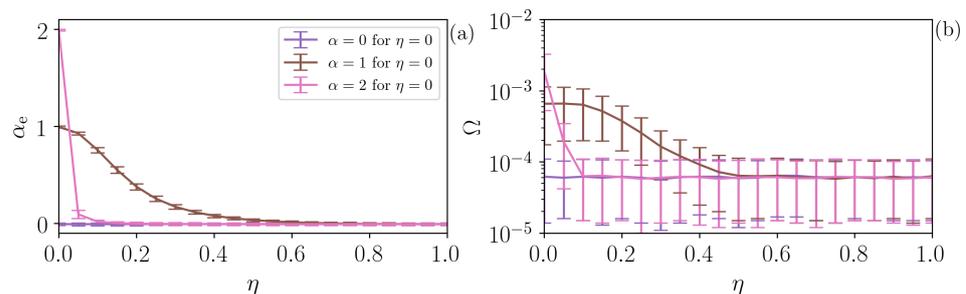


Figure 8. Analysis of FN time series contaminated with uniform noise. (a) The addition of uniform noise decreases the α_e predicted by the ML algorithm, leading from the value α of the FN time series to value to $\alpha_e = 0$ for uniform noise. η also decreases the value of Ω (b), revealing the increase in the degree of stochasticity of the time series. The addition of uniform noise is, thus, similar in FN and in deterministic time series.

5. Discussions and Conclusions

We have analyzed stochastic and deterministic time series using an algorithm [9] that automatically reduces the dimensionality of the feature space from 720 probabilities of ordinal patterns to 2 features: the degree of stochasticity (Ω) and the strength of the temporal correlations (α_e). We have analyzed the performance and limitations of the algorithm, presenting applications to different datasets, including highly chaotic and periodic signals with added white noise.

For the analysis of chaotic time series, we have shown that α_e and Ω are able to capture the rich dynamics that chaotic systems can depict, where the transitions between periodic windows and chaos are evident. In general, negative correlation between the Lyapunov exponent and Ω and between the permutation entropy and Ω were found. For highly chaotic signals, when the time evolution of the system is very similar to a stochastic process, $\bar{S} \approx 1$ and $\Omega \approx 0$ and our methodology characterizes highly chaotic signals as stochastic ones.

In addition, we have studied periodic signals contaminated with noise. In this case, our method captures the transition from deterministic time series to stochastic ones. However, we have shown that the period of the signal is indeed important, with fast signals being identified as deterministic even with large noise. We have found that when the noise contamination increases, periodic signals lose their deterministic feature but preserve a nonzero temporal correlation.

For future work, it will be interesting to analyze whether the performance of our methodology can be improved by using different lengths of the ordinal pattern, D , or different lags between the data points that define the ordinal patterns. It is well known that the ordinal patterns distribution varies with the time scale of the analysis [19] and it will also be interesting, for future work, to address the relevant question of whether our method is able to estimate, from the same time series, different values of α_e by using different lags.

We remark that the automatic and easy-to-use time series analysis tool that we propose is here is freely available at [42]. We believe that it will be a valuable contribution to the

wide repertoire of time series analysis tools that are available nowadays. Many applications are foreseen. As an example, for ultra-fast optical random number generation [43–45], it is crucial to generate optical chaotic signals that are as uncorrelated and as “pseudorandom” as possible, for which its deterministic nature is hidden by noise-like properties. Many different setups have been proposed to generate such broad-band, high-entropy optical signals [46–50]. Our algorithm allows an automatic comparison of the strength of the correlations and the level of randomness of signals generated by different setups. Moreover, the algorithm may be used to identify, in a given experimental setup, the optimal operation conditions and parameters that produce optical signals with the lowest temporal correlations (lowest α_c) or with highest level of randomness (lowest Ω).

Author Contributions: Conceptualization: B.R.R.B., K.L.R., R.C.B. and C.M.; methodology and Software, B.R.R.B.; investigation, B.R.R.B.; validation, B.R.R.B., K.L.R. and R.C.B.; discussion, B.R.R.B., K.L.R., R.C.B., T.L.P., S.R.L. and C.M.; visualization: B.R.R.B.; manuscript preparation: B.R.R.B., K.L.R., R.C.B. and C.M. All authors have read and agreed to the published version of the manuscript.

Funding: B.R.R.B., T.L.P. and S.R.L. acknowledge partial support of Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brazil, Grant No. 302785/2017-5, 308621/2019-0 and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 001. R.C.B. was supported by BrainsCAN at Western University through the Canada First Research Excellence Fund (CFREF) and by the NSF through a NeuroNex award (#2015276). C. M. was funded by the Spanish Ministerio de Ciencia, Innovación y Universidades (PGC2018-099443-B-I00) and the ICREA ACADEMIA program of Generalitat de Catalunya. K.L.R. was supported by a scholarship from the Deutsche Akademischer Austauschdienst (DAAD).

Data Availability Statement: All codes and datasets relevant to our methodology are freely available at [42].

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

References

- Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.L.C.; Shih, H.H.; Zheng, Q.N.; 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. A* **1998**, *454*, 903–995. [CrossRef]
- Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: Cambridge, MA, USA, 1998.
- Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*; Cambridge University Press: Cambridge, UK, 2005.
- Brunton, S.L.; Kutz, J.N. *Data-driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*; Cambridge University Press: Cambridge, UK, 2019.
- Fulcher, B.D.; Jones, N.S. A computational framework for automated time-series phenotyping using massive feature extraction. *Cell Syst.* **2017**, *5*, 527–531. [CrossRef]
- Zou, Y.; Donner, R.V.; Marwan, N.; Donges, J.F.; Kurths, J. Complex network approaches to nonlinear time series analysis. *Phys. Rep.* **2019**, *787*, 1–97. [CrossRef]
- Carlsson, G. Topological methods for data modelling. *Nat. Rev. Phys.* **2020**, *2*, 697–708. [CrossRef]
- Lancaster, G.; Iatsenko, D.; Pidge, A.; Ticcinelli, V.; Stefanovska, A. Surrogate data for hypothesis testing of physical systems. *Phys. Rep.* **2018**, *748*, 1–60. [CrossRef]
- Boaretto, B.R.R.; Budzinski, R.C.; Rossi, K.L.; Prado, T.L.; Lopes, S.R.; Masoller, C. Discriminating chaotic and stochastic time series using permutation entropy and artificial neural networks. *Sci. Rep.* **2021**, *11*, 15789. [CrossRef]
- Beran, J.; Feng, Y.; Ghosh, S.; Kulik, R. *Long-Memory Processes*; Springer: Berlin/Heidelberg, Germany, 2016.
- Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [CrossRef]
- Rosso, O.A. Permutation Entropy & Its Interdisciplinary Applications. Available online: https://www.mdpi.com/journal/entropy/special_issues/Permutation_Entropy (accessed on 4 August 2021).
- Tenenbaum, J.B.; De Silva, V.; Langford, J.C. A global geometric framework for nonlinear dimensionality reduction. *Science* **2000**, *290*, 2319–2323. [CrossRef] [PubMed]
- Maaten, L.V.D.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
- Van Der Maaten, L.; Postma, E.; Van den Herik, J. Dimensionality reduction: A comparative review. *J. Mach. Learn. Res.* **2009**, *10*, 66–71.
- Rosso, O.A.; Larrondo, H.A.; Martin, M.T.; Plastino, A.; Fuentes, M.A. Distinguishing noise from chaos. *Phys. Rev. Lett.* **2007**, *99*, 154102. [CrossRef]

17. Olivares, F.; Plastino, A.; Rosso, O. Contrasting chaos with noise via local versus global information quantifiers. *Phys. Lett. A* **2012**, *376*, 1577–1583. [CrossRef]
18. Ravetti, M.G.; Carpi, L.C.; Gonçalves, B.A.; Frery, A.C.; Rosso, O.A. Distinguishing noise from chaos: Objective versus subjective criteria using horizontal visibility graph. *PLoS ONE* **2014**, *9*, e108004. [CrossRef]
19. Spichak, D.; Kupetsky, A.; Aragoneses, A. Characterizing complexity of non-invertible chaotic maps in the Shannon-Fisher information plane with ordinal patterns. *Chaos Solitons Fractals* **2021**, *142*, 110492. [CrossRef]
20. Daw, C.S.; Finney, C.E.A.; Tracy, E.R. A review of symbolic analysis of experimental data. *Rev. Sci. Inst.* **2003**, *74*, 913–930. [CrossRef]
21. Corso, G.; Prado, T.d.L.; Lima, G.Z.S.; Kurths, J.; Lopes, S.R. Quantifying entropy using recurrence matrix microstates. *Chaos Interdiscip. Nonlinear Sci.* **2018**, *28*, 083108. [CrossRef] [PubMed]
22. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
23. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef]
24. Prado, T.L.; Corso, G.; dos Santos Lima, G.Z.; Budzinski, R.C.; Boaretto, B.R.R.; Ferrari, F.A.S.; Macau, E.E.N.; Lopes, S.R. Maximum entropy principle in recurrence plot analysis on stochastic and chaotic systems. *Chaos Interdiscip. Nonlinear Sci.* **2020**, *30*, 043123. [CrossRef] [PubMed]
25. Martin, M.T.; Plastino, A.; Rosso, O.A. Generalized statistical complexity measures: Geometrical and analytical properties. *Physica A* **2006**, *369*, 439–462. [CrossRef]
26. Parlitz, U.; Berg, S.; Luther, S.; Schirdewan, A.; Kurths, J.; Wessel, N. Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics. *Comput. Biol. Med.* **2012**, *42*, 319–327. [CrossRef]
27. Ott, E. *Chaos in Dynamical Systems*; Cambridge University Press: Cambridge, UK, 2002.
28. Mitschke, F.; Dammig, M. Chaos vs. noise in experimental data. *Int. J. Bifurc. Chaos* **1993**, *3*, 693–702. [CrossRef]
29. Rosso, O.A.; Masoller, C. Detecting and quantifying stochastic and coherence resonances via information-theory complexity measurements. *Phys. Rev. E* **2009**, *79*, 040106. [CrossRef]
30. Library to Generate a Flicker Noise. Available online: <https://github.com/felixpatzelt/colorednoise> (accessed on 4 August 2021).
31. Timmer, J.; Koenig, M. On generating power law noise. *Astron. Astrophys.* **1995**, *300*, 707.
32. Schuster, H.G.; Just, W. *Deterministic Chaos: An Introduction*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
33. Lorenz, E.N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
34. Rössler, O.E. An equation for continuous chaos. *Phys. Lett. A* **1976**, *57*, 397–398. [CrossRef]
35. Lopes, S.; Chian, A.L. Controlling chaos in nonlinear three-wave coupling. *Phys. Rev. E* **1996**, *54*, 170. [CrossRef]
36. Hindmarsh, J.L.; Rose, R.M. A model of neuronal bursting using three coupled first order differential equations. *Proc. R. Soc. Lond. B Biol. Sci.* **1984**, *221*, 87–102.
37. Little, D.J.; Kane, D.M. Permutation entropy of finite-length white-noise time series. *Phys. Rev. E* **2016**, *94*, 022118. [CrossRef]
38. De Micco, L.; Fernández, J.G.; Larrondo, H.A.; Plastino, A.; Rosso, O.A. Sampling period, statistical complexity, and chaotic attractors. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 2564–2575. [CrossRef]
39. Zunino, L.; Soriano, M.C.; Fischer, I.; Rosso, O.A.; Mirasso, C.R. Permutation-information-theory approach to unveil delay dynamics from time-series analysis. *Phys. Rev. E* **2010**, *82*, 046212. [CrossRef] [PubMed]
40. Aragoneses, A.; Carpi, L.; Tarasov, N.; Churkin, D.V.; Torrent, M.C.; Masoller, C.; Turitsyn, S.K. Unveiling Temporal Correlations Characteristic of a Phase Transition in the Output Intensity of a Fiber Laser. *Phys. Rev. Lett.* **2016**, *116*, 033902. [CrossRef]
41. Lopes, S.R.; Prado, T.d.L.; Corso, G.; Lima, G.Z.d.S.; Kurths, J. Parameter-free quantification of stochastic and chaotic signals. *Chaos Solitons Fractals* **2020**, *133*, 109616. [CrossRef]
42. Repository with the ANN. Available online: https://github.com/brunorrboaretto/chaos_detection_ANN/ (accessed on 4 August 2021).
43. Uchida, A.; Amano, K.; Inoue, M.; Hirano, K.; Naito, S.; Someya, H.; Oowada, I.; Kurashige, T.; Shiki, M.; Yoshimori, S. Fast physical random bit generation with chaotic semiconductor lasers. *Nat. Photonics* **2008**, *2*, 728–732. [CrossRef]
44. Sakuraba, R.; Iwakawa, K.; Kanno, K.; Uchida, A. Tb/s physical random bit generation with bandwidth-enhanced chaos in three-cascaded semiconductor lasers. *Opt. Express* **2015**, *23*, 1470–1490. [CrossRef]
45. Zhang, L.; Pan, B.; Chen, G.; Guo, L.; Lu, D.; Zhao, L.; Wang, W. 640-Gbit/s fast physical random number generation using a broadband chaotic semiconductor laser. *Sci. Rep.* **2017**, *7*, 45900. [CrossRef]
46. Oliver, N.; Soriano, M.C.; Sukow, D.W.; Fischer, I. Dynamics of a semiconductor laser with polarization-rotated feedback and its utilization for random bit generation. *Opt. Lett.* **2011**, *36*, 4632–4634. [CrossRef] [PubMed]
47. Zhang, J.; Feng, C.; Zhang, M.; Liu, Y.; Zhang, Y. Suppression of Time Delay Signature Based on Brillouin Backscattering of Chaotic Laser. *IEEE Photon J.* **2017**, *9*, 1502408. [CrossRef]
48. Li, S.S.; Chan, S.C. Chaotic Time-Delay Signature Suppression in a Semiconductor Laser With Frequency-Detuned Grating Feedback. *IEEE J. Sel. Top. Quantum Electron.* **2015**, *21*, 541–552.

-
49. Correa-Mena, A.G.; Lee, M.W.; Zaldívar-Huerta, I.E.; Hong, Y.; Boudrioua, A. Investigation of the Dynamical Behavior of a High-Power Laser Diode Subject to Stimulated Brillouin Scattering Optical Feedback. *IEEE J. Quantum Electron.* **2020**, *56*, 1–6. [[CrossRef](#)]
 50. Bouchez, G.; Malica, T.; Wolfersberger, D.; Sciamanna, M. Optimized properties of chaos from a laser diode. *Phys. Rev. E* **2021**, *103*, 042207. [[CrossRef](#)] [[PubMed](#)]