Physics Letters A 381 (2017) 3627–3635

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Detecting nonlinearity in short and noisy time series using the permutation entropy

Luciano Zunino [a,b,*], Christopher W. Kulp [c]

[a] Centro de Investigaciones Ópticas (CONICET La Plata – CIC), C.C. 3, 1897 Gonnet, Argentina
[b] Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
[c] The Department of Astronomy and Physics, Lycoming College, Williamsport, PA 17701, USA

ARTICLE INFO

Article history:
Received 3 July 2017
Received in revised form 17 September 2017
Accepted 18 September 2017
Available online 20 September 2017
Communicated by C.R. Doering

Keywords:
Time series analysis
Permutation entropy
Nonlinearity
Surrogate method

ABSTRACT

Permutation entropy contains the information about the temporal structure associated with the underlying dynamics of a time series. Its estimation is simple, and because it is based on the comparison of neighboring values, it becomes significantly robust to noise. It is also computationally efficient and invariant with respect to nonlinear monotonous transformations. For all these reasons, the permutation entropy seems to be particularly suitable as a discriminative measure for unveiling nonlinear dynamics in arbitrary real-world data. In this paper, we study the efficacy of a conventional surrogate method with a linear stochastic process as the null hypothesis but implementing the permutation entropy as a nonlinearity measure. Its discriminative power is tested by implementing several analyses on numerical signals whose dynamical properties are known *a priori* (linear discrete and continuous models, chaotic regimes of discrete and continuous systems). The performance of the proposed approach in real-world applications (chaotic laser data, monthly smoothed sunspot index and neuro-physiological recordings) is also demonstrated. The results obtained allow us to conclude that this symbolic tool is very useful for discriminating nonlinear characteristics in very short and noisy data.

1. Introduction

Determining whether a given time series comes from a deterministic chaotic or a stochastic system can be a big challenge [1–3]. It is well-known that nonlinearity is a necessary condition for chaos. Consequently, to determine if an arbitrary time series is compatible with chaotic dynamics for modeling and classification purposes, it is first necessary to demonstrate that the dynamics producing the time series is, in fact, nonlinear. Furthermore, the detection of nonlinearity is not a trivial task especially for experimental records that are often contaminated with unknown noise sources. Motivated by these facts, in the last decade, several techniques for identifying nonlinear processes in observational data have been introduced [4–8]. Despite the existing contributions, discriminating the nonlinear dynamics of a complex system from time series is still a challenging problem of current research [9].

In this paper, we implement and test the efficacy of the permutation entropy (PE) as a discriminating statistic in a standard surrogate framework [10] in order to detect nonlinearities in short and noisy time series data. The PE is the Shannon entropic measure evaluated using the successful encoding introduced by Bandt and Pompe (BP) [11] to extract the probability distribution associated with an input signal. Taking into account the widely recognized practical advantages of this symbolic information-theory quantifier, namely, *i)* simplicity, *ii)* low computational cost, *iii)* noise robustness, and *iv)* invariance with respect to nonlinear monotonous transformations, PE is demonstrated to be an alternative and/or complementary approach to more traditional techniques for unveiling nonlinear structures from complex systems. The proposed nonlinearity test relies on the well-established method of surrogate data [10] just as many other nonlinear discriminating approaches [4–6,12], and, obviously, the generation of proper surrogates is essential for the test's success. As will be shown below, linear and nonlinear short noisy scalar time series can be efficiently characterized supporting a remarkable reliability of PE as a discriminator in practical contexts.

Even though the permutation entropy has been used in countless applications, it has been rarely implemented within a surrogate framework for unveiling the nonlinear dynamics of complex systems from time series. The analysis developed by Tony et al. [13], to identify the deterministic nature of pressure measure-

* Corresponding author at: Centro de Investigaciones Ópticas (CONICET La Plata – CIC), C.C. 3, 1897 Gonnet, Argentina.
E-mail addresses: lucianoz@ciop.unlp.edu.ar (L. Zunino), kulp@lycoming.edu (C.W. Kulp).

http://dx.doi.org/10.1016/j.physleta.2017.09.032
0375-9601/

ments from a turbulent combustor, is one of these rare exceptions. Taking into account that, to the best of our knowledge, the performance of both techniques (permutation entropy and surrogate testing) together, as an unified approach, has not been analyzed in depth before, in this work we try to fill this gap through several numerical and real-world data tests. As it will be shown below, this approach is able to unveil the presence of nonlinear dynamics even in very short and noisy time series.

The remainder of the paper is organized as follows. In the next section, PE and the surrogate data analysis are briefly presented. Numerical and experimental analyses for testing the performance of the proposed nonlinearity test are detailed in Sections 3 and 4, respectively. Finally, Section 5 summarizes the results and contains concluding remarks.

2. Methodology

2.1. Permutation entropy

The symbolic encoding scheme due to BP [11], based on the ordinal relation between the amplitude of neighboring values of a given data sequence, has been implemented for estimating several information-theory quantifiers from time series. The BP ordinal method of symbolization naturally arises from the time series, inherits the causal information that stems from the temporal structure of the system dynamics, and also, avoids amplitude threshold dependencies that affect other more conventional symbolization recipes based on range partitioning [14,15]. These traits could be the main reasons behind notable success, as is evidenced by the enormous amount of applications in heterogeneous fields (see, *e.g.*, [16–26]). Furthermore, the ordinal pattern distribution is invariant with respect to nonlinear monotonous transformations. Thus, nonlinear drifts or scalings artificially introduced by a measurement device do not modify the quantifiers' estimations. It appears to be better suited to cope with usual problems (nonstationarities, nonlinearities, noise distortions) encountered when studying real time series compared to range-based encoding methods. Within this appealing encoding procedure, PE is undoubtedly the most widely-used descriptor. It should be stressed here that this entropic measure is applicable to noisy real time series from all class of systems, deterministic and stochastic, without the need to require any knowledge of the underlying mechanisms. As stated by Garland et al. [27], "It does not rely on generating partitions, and thus does not introduce bias into the results if one does not know the dynamics or cannot compute the partition. Permutation entropy makes no assumptions about, and requires no knowledge of, the underlying generating process: linear, nonlinear, the Lyapunov spectrum, etc." Furthermore, the relationship between the PE and the Kolmogorov–Sinai (KS) entropy has been discussed by several authors before. Basically, the growth rate of the PE is often used as a proxy for the KS entropy [28]. For more details, please see Refs. [29–31]. The KS entropy, probably the most appropriate indicator for distinguishing irregular deterministic from stochastic dynamics, requires specific knowledge of the generating process for its correct estimation. Finding the generating partition is not feasible for experimental data since they are inevitably contaminated with noise [32]. Consequently, the PE, which does not rely on generating partitions, emerges as a practical alternative to characterize data sets generated by an unknown dynamic process with unknown levels of noise [9].

Here, we will illustrate how to create ordinal patterns from the time series data with a simple example. Let us assume that we start with the time series $X = \{3, 2, 5, 8, 9, 6, 1\}$. In order to symbolize the series into ordinal patterns, first, two parameters, the order of the permutation symbols $D > 1$ ($D \in \mathbb{N}$, number of elements that form the ordinal pattern) and the lag τ ($\tau \in \mathbb{N}$, time separation between elements) are chosen. Next, the time series is partitioned into subsets of length D with lag τ similarly to phase space reconstruction by means of time-delay-embedding. The elements in each new partition (of length D) are replaced by their rank in the subset. For example, if we set $D = 3$ and $\tau = 1$, there are five different three-dimensional vectors associated with X. The first one $(x_0, x_1, x_2) = (3, 2, 5)$ is mapped to the ordinal pattern (102). The second three-dimensional vector is $(x_0, x_1, x_2) = (2, 5, 8)$, and (012) will be its related permutation. The procedure continues so on until the last sequence, $(9, 6, 1)$, is mapped to its corresponding motif, (210). In the case of two elements in the vector having the same value, the elements are ranked by index, for example, a vector $(7, 8, 7)$, which does not appear in X, would be mapped to (021). Afterward, an ordinal pattern probability distribution, $\boldsymbol{P} = \{p(\pi_i), i = 1, \ldots, D!\}$, can be obtained from the time series by computing the relative frequencies of the $D!$ possible permutations π_i. Continuing with the example: $p(\pi_1) = p(012) = 2/5$, $p(\pi_2) = p(021) = 0$, $p(\pi_3) = p(102) = 1/5$, $p(\pi_4) = p(120) = 1/5$, $p(\pi_5) = p(201) = 0$, and $p(\pi_6) = p(210) = 1/5$. PE is just the Shannon entropy estimated by using this ordinal pattern probability distribution, $S[\boldsymbol{P}] = -\sum_{i=1}^{D!} p(\pi_i)\log(p(\pi_i))$ ($0\log(0)$ is set to zero in accordance with its mathematical limit). Coming back to the example, $S[\boldsymbol{P}(X)] = -(2/5)\log(2/5) - 3(1/5)\log(1/5) = 1.3322$. PE quantifies the temporal structural diversity of a time series. Technically speaking, the ordinal pattern probability distribution $\boldsymbol{P}$ is obtained once we fix the order D and the lag τ. The PE estimation does not require the optimal reconstruction of the phase space that is necessary for estimating other quantifiers of chaotic signals. Consequently, D and τ are not usually selected following the methodologies often employed in a conventional phase space reconstruction (*e.g.*, the first zero of the autocorrelation function, the first minimum of the average mutual information, and the false nearest neighbor algorithm). Taking into account that there are $D!$ potential permutations for a D-dimensional vector, the condition $N \gg D!$, with N the length of the time series, must be satisfied in order to obtain a reliable estimation of $\boldsymbol{P}$ [33]. For practical purposes, BP suggest in their seminal paper to estimate the frequency of ordinal patterns with $3 \leq D \leq 7$ and $\tau = 1$ (consecutive points). However, it has been demonstrated that the analysis with lagged data points, *i.e.* $\tau \geq 2$, can be useful for reaching a better comprehension of the underlying dynamics [34–36]. Essentially, by changing the value of the lag τ, different time scales are being considered because this parameter physically corresponds to multiples of the sampling time of the signal under analysis. For further details about the BP methodology, we recommend [34,37,38]. It is common to normalize the PE, and therefore in this paper, a normalized PE is used and is given by

$$\mathcal{H}_S = S[\boldsymbol{P}]/S_{\max} = S[\boldsymbol{P}]/\log(D!) \tag{1}$$

with $S_{\max} = \log(D!)$ the value obtained from an equiprobable ordinal pattern probability distribution, *i.e.* $\boldsymbol{P} = \{p(\pi_i) = 1/D!, i = 1, \ldots, D!\}$.

2.2. Testing nonlinear dynamics in time series with surrogate methods

The method of surrogate data, introduced by Theiler et al. [10], represents a cornerstone in nonlinear time series analysis. Briefly, a statistic sensitive to nonlinearities is estimated for the original univariate time series $\{x_i\}_{i=1}^N$ and for an ensemble of M generated surrogate time series, $\{\hat{x}_i^j\}_{i=1}^N$ with $j = 1, \ldots, M$. Each surrogate (indexed with j) is a constrained realization of the original data that mimics its linear properties (autocorrelation function/power spectrum) while potential higher order correlations are randomized. When a statistically significant difference is found between

the computed measure for the original time series and the distribution of estimated values for the set of surrogates, a nonlinear dynamics can be concluded within a statistical confidence level. More specifically, to reach a significance level α, $M = 2/\alpha - 1$ surrogate realizations need to be generated in a two-sided test. The initial null hypothesis of a Gaussian linear stochastic process has been then generalized to cases where the data set is non-Gaussian in order to avoid spurious detection of nonlinearity. Consequently, surrogate realizations that preserve not only power spectrum but also amplitude probability distribution can be constructed through the iterative amplitude adjusted Fourier transform (IAAFT) scheme. We address the interested reader to [39] for further details about this method for generating surrogate data. Throughout this paper, we have implemented the IAAFT with its generalized null hypothesis as the surrogate testing method. The null hypothesis is rejected or not based on a standard rank-based nonparametric statistical test. Let us denote as $\mathcal{R}(\mathcal{H}_S)$ the rank assigned to the estimated normalized PE value for the original time series and $\mathcal{R}(\hat{\mathcal{H}}_S^j)$ those associated with surrogate realizations; the null hypothesis is rejected if $\mathcal{R}(\mathcal{H}_S) < \mathcal{R}(\hat{\mathcal{H}}_S^j)$ or $\mathcal{R}(\mathcal{H}_S) > \mathcal{R}(\hat{\mathcal{H}}_S^j)$ for $j = 1, \ldots, M$ [6]. Summarizing, calculated values of a discriminating statistic, normalized PE in our case, for the original record, $\{x_i\}_{i=1}^N$, and its IAAFT surrogates, $\{\hat{x}_i^j\}_{i=1}^N$, are compared through a two-sided rank order test of size α.

Taking into consideration that the analysis should be repeated on the collection of Monte Carlo realizations of the null hypothesis, the computational efficiency is another point in favor of PE and in detriment of other approaches that are more time demanding. It is also worth emphasizing here that, since our nonlinearity test depends directly on the appropriate generation of surrogate realizations, the Fourier phase correlations observed in IAAFT surrogates [40] may potentially be the source of false negative detection of weak nonlinearities.

3. Numerical results

In this section, synthetic time series from several models are generated and tested, to demonstrate the performance of PE as a discriminating nonlinear measure in controlled scenarios. Since the underlying nature is known, we can estimate the probabilities of *type-I errors* (incorrect rejections of true null hypothesis) and *type-II errors* (failures to reject a false null hypothesis). As mentioned previously, the null hypothesis here is that the series being studied is produced by a linear stochastic process. Therefore, in this case, a *type-I error* would occur if the test incorrectly rejects the null hypothesis when the system is, in fact, linear. In the case of a *type-I error*, the null hypothesis (linearity) is true but the test produced an incorrect result (the series is nonlinear). Furthermore, a *type-II error* would occur if the test fails to reject the null hypothesis when, in fact, the system is nonlinear. In this case, the null hypothesis (linearity) is false for the series being tested, but the test incorrectly supported the null hypothesis. Through these analyses, the discriminatory power of the proposed test can be quantified.

3.1. Tests on linear processes

As a first toy example, we have studied first-order autoregressive processes AR(1), *i.e.* $x_t = a x_{t-1} + \epsilon_t$ where ϵ_t are pseudorandom values from the standard normal distribution, for different values of the parameter a ($a \in \{0.05, 0.1, ..., 0.95\}$). One hundred independent realizations of length $N = 2^9$ data points were simulated for each value of the parameter a (potential transients were avoided by discarding the first 10^4 iterations). Each realization has been tested again $M = 199$ surrogate independent realizations generated from this realization, corresponding to a confidence level

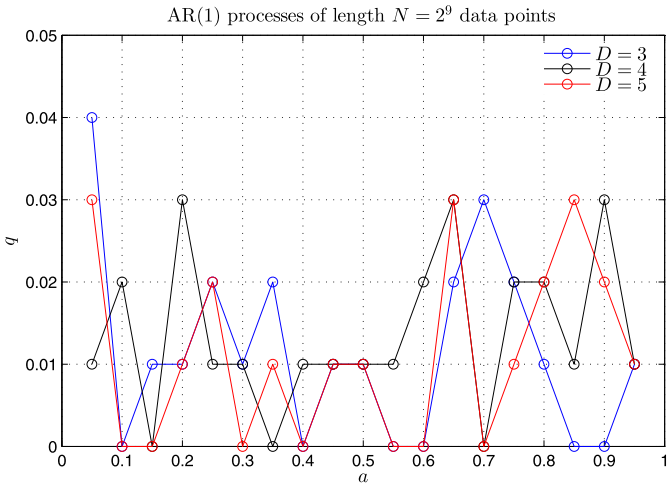

Fig. 1. Analysis of a first-order autoregressive processes AR(1) of length $N = 2^9$ data points with the PE as nonlinearity measure and the IAAFT as surrogate method. The fraction q of rejections of one hundred two-sided rank order tests of size $\alpha = 0.01$ as a function of parameter a is shown. The PE is estimated for different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of 99% ($\alpha = 0.01$). Different surrogate data sets have been used for each one of the one hundred realizations. The fraction q of numerical realizations for which the null hypothesis can be rejected as a function of the parameter a is plotted in Fig. 1. For this discrete null model, the rate of incorrect rejections of the true null hypothesis is very low ($q < 0.05$). Results are similar for different orders ($D = 3$, $D = 4$ & $D = 5$) and the PE is estimated with lag $\tau = 1$. We have also confirmed that similar probabilities of *type-I errors* are obtained for longer time series ($N = 2^{10}$ data points) generated for this model and for negative parameter values ($a \in \{-0.95, -0.9, ..., -0.05\}$).

As a second example, we have analyzed the following linear feedback system (second-order autoregressive process) proposed by Gunduz and Principe [4],

$$x_t = 1.5 x_{t-1} - 0.8 x_{t-2} + \epsilon_t \,, \tag{2}$$

where ϵ_t follows a white Gaussian distribution with zero mean and unit variance. One hundred independent realizations were generated for different lengths $N = 2^k$, $k \in \{8, 9, ..., 12\}$ (first 5,000 iterations are discarded to remove transients), and the null hypothesis was again tested through a two-sided rank order test of size $\alpha = 0.01$ ($M = 199$ surrogate time series). In this case, the fraction q of rejections has been analyzed as a function of the varying sample size N. Fig. 2a) shows results obtained for different orders $D = 3$, $D = 4$ and $D = 5$, and $\tau = 1$. A similar analysis has been developed with ϵ_t sampled from an exponential distribution also with zero mean (after being normalized by subtracting its mean value) and unit variance. False alarm rates (*type-I errors*) as a function of length N for this non-Gaussian linear process are plotted in Fig. 2b) for different orders. Overall, we have confirmed a very low rejection ($q \leq 0.06$) of the true null hypothesis in absence of nonlinearity for Gaussian white noise as well as for exponential noise for the different time series lengths and orders. Only in the non-Gaussian case and for $N = 2^{12}$ data points, was a false alarm rate around 0.1 found (for $D = 4$ and $D = 5$). This higher fraction of rejections of the true null hypothesis can be attributed to the presence of an exponential dynamical noise. Actually, much worse performances have been previously observed for non-Gaussian linear process with other discriminative statistics (bicorrelation [4], time reversibility [4], approximate entropy [41],

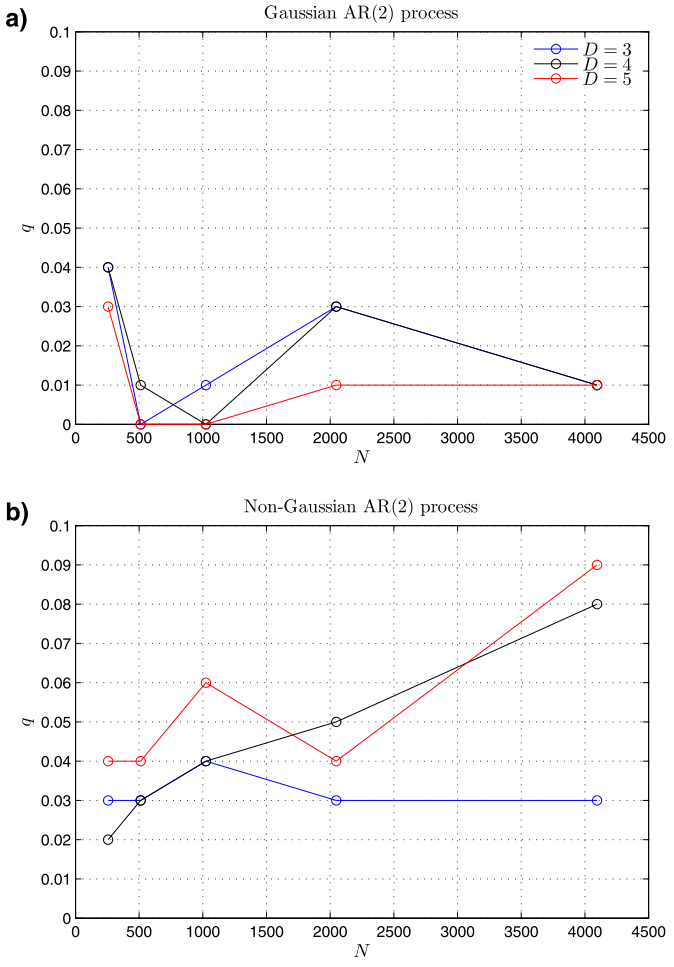

Fig. 2. a) Analysis of a Gaussian second-order autoregressive processes AR(2) (Eq. (2) with ϵ_t a Gaussian white noise) of different lengths ($N = 2^k$ data points, $k \in \{8, 9, ..., 12\}$) with the PE as nonlinearity measure and the IAAFT as surrogate method. The fraction q of rejections of one hundred two-sided rank order tests of size $\alpha = 0.01$ as a function of length N is depicted. b) Same analysis but for ϵ_t of Eq. (2) sampled from an exponential distribution with zero mean and unit variance. The PE is estimated for different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and Lempel–Ziv complexity [42]). For example, in the case of bi-correlation and time reversibility, false alarm rates higher than 0.3 are obtained for a signal-to-noise ratio equal to 20 dB and varying data lengths ($N \in \{500, 1000, \ldots, 5000\}$). Moreover, the *type-I error* increases with the data length, reaching a value near 0.9 for $N = 5000$ samples—please see Fig. 2 in Ref. [4].

Next, we addressed continuous null models. More precisely, colored noise with power-law spectra $1/f^{\beta}$ have been studied. Numerical realizations of noise with β ranging from -0.9 to 0.9 in 0.1 steps were generated with the widely implemented FFM algorithm. For more details about the FFM algorithm, we refer the interested reader to [43,44]. One hundred independent synthetic sequences of length $N = 2^9$ data points for each value of the β-exponent have been tested for nonlinearity using the proposed approach. The fraction q of (incorrect) rejections as a function of the power-law exponent β is plotted in Fig. 3. It is worth emphasizing here that $q < 0.05$ independently of β and the order D used for estimating PE. A similar finding is obtained for longer simulations ($N = 2^{10}$ data points). We have confirmed once again a relatively low fraction of rejections of a true null hypothesis.

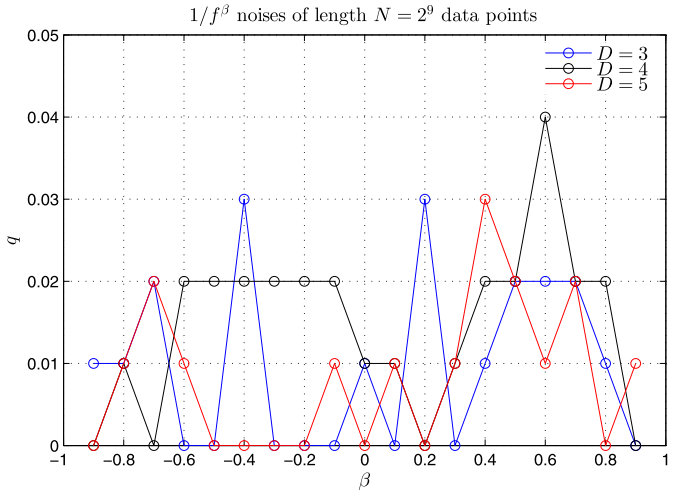

Fig. 3. The fraction q of rejections of true null hypothesis for colored noises of length $N = 2^9$ data points. Results are obtained from one hundred independent two-sided rank order tests of size $\alpha = 0.01$ for each value of the power-law exponent β. The PE is estimated with different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. Surrogates were generated with the IAAFT algorithm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Tests on nonlinear processes

In order to quantify the discriminative power of PE as a nonlinearity measure in a noisy environment, chaotic data sets generated from the logistic map given by

$$x_t = r x_{t-1}(1 - x_{t-1}) \tag{3}$$

with $r = 4$, and from the x-component of the Hénon map

$$x_t = 1 - A x_{t-1}^2 + y_{t-1}, \qquad y_t = B x_{t-1} \tag{4}$$

with $A = 1.4$ and $B = 0.3$, both contaminated with different amount of observational Gaussian white noise, have been analyzed for different lengths N. The noisy time series were produced from the logistic and Hénon maps in the following way. First, the initial 10^4 iterations were discarded to remove transient behavior. Second, a segment of the remaining series of length, N, was selected using the end-to-end mismatch criterion. This segment will be the series under study and will be used to produce surrogates. The end-to-end mismatch procedure is used because high frequency components are spuriously included in surrogate realizations when any mismatch between the beginning and the end of the original time series is present [5,45]. If the mismatch criterion is not satisfied, the resulting surrogates can lead to erroneous rejections of the null hypothesis. A criterion, that minimizes the average of the end-point mismatch and the first-derivative mismatch, has been introduced to reduce this effect. The aforementioned segment of the chaotic sequences have been chosen according to this criterion in order to minimize the chance of false rejections of the null hypothesis. Please see [45] for further details about the end-to-end mismatch problem and its solution. Once a segment of the chaotic series of length N has been chosen (using the end-to-end mismatch criterion), the series is then normalized to a unit standard deviation in order to easily quantify the noise that will be added to the series. Next, observational Gaussian white noise with a standard deviation, η, is added to the series. Twenty noise levels $\eta \in \{0.05, 0.1, ..., 1\}$ were used. The resulting noisy series is used to generate surrogates using the IAAFT method. For each noise level, one hundred independent realizations were tested for nonlinearity with the PE as a discriminative measure. Once again, a two-sided rank order test of size $\alpha = 0.01$ ($M = 199$ surrogates)

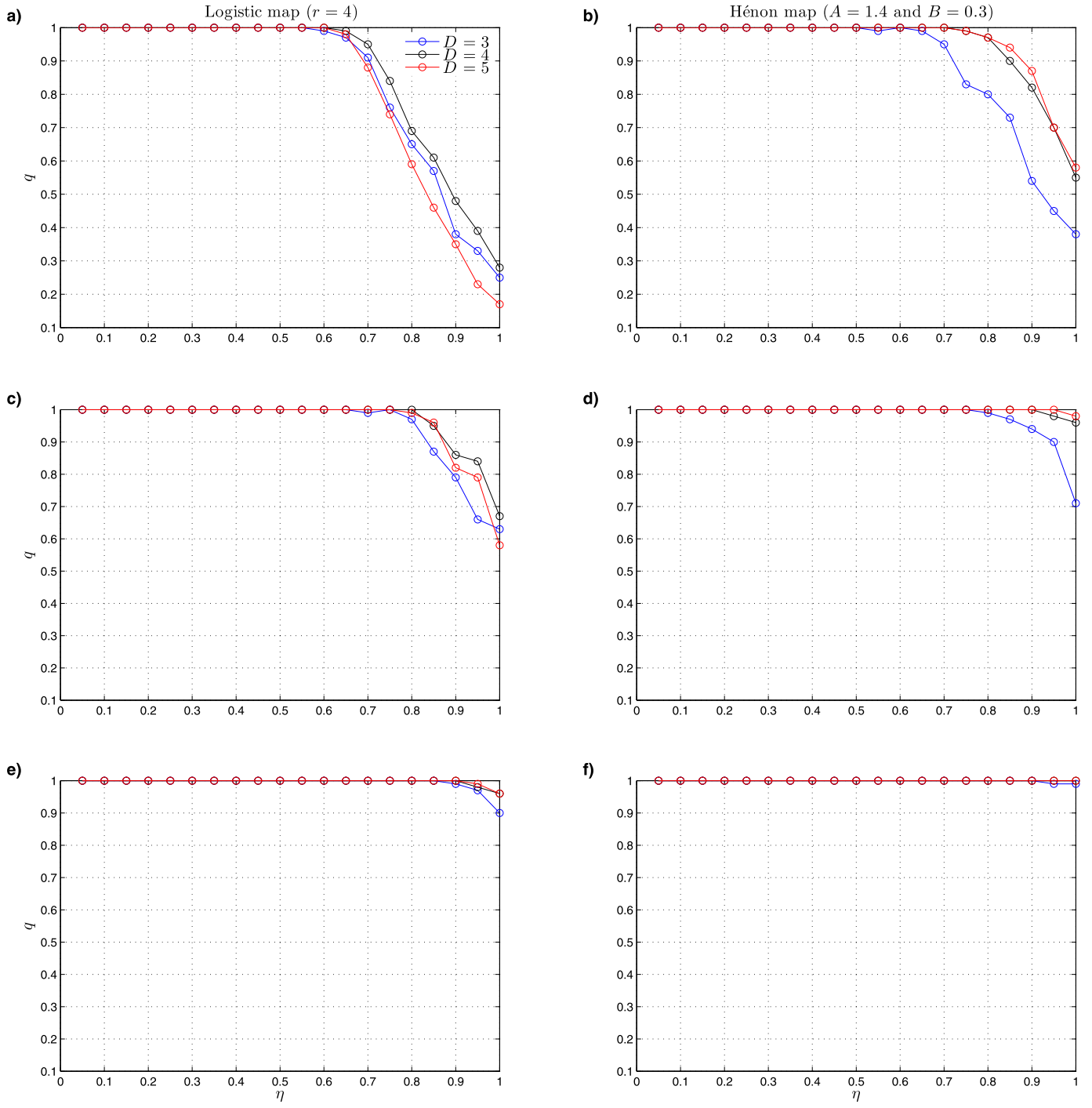

Fig. 4. The fractions of rejections obtained for chaotic generated sequences from the logistic map (left column) and the x-component of the Hénon map (right column) contaminated with different levels η of additive Gaussian white noise. Results for different time series lengths are included: a)–b) $N = 2^9$ samples; c)–d) $N = 2^{10}$ samples; e)–f) $N = 2^{11}$ samples. One hundred independent noise realizations have been tested for each noise level and for each length, with the PE as nonlinearity measure and the IAAFT as surrogate method. A two-sided rank order test of size $\alpha = 0.01$ ($M = 199$ surrogates) has been implemented for testing nonlinearity of each sequence. The PE is estimated for different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

has been implemented. In Fig. 4, fractions q of (correct) rejections of the false null hypothesis for the logistic map (left column) and for the x-component of the Hénon map (right column) are plotted as a function of the noise level η for three different lengths $N = 2^k$ with $k \in \{9, 10, 11\}$, and for different orders $D = 3$, $D = 4$ and $D = 5$, always with $\tau = 1$. As it was expected in the case of nonlinear dynamics, the fraction of rejections q is equal to one for low levels of observational noise. Likewise, q decreases as the amount of noise contaminating the original chaotic sequences increases. Since the practical usefulness of tests with fraction of (correct) rejections of false null hypothesis less than 0.7 is regarded as questionable [46], we conclude that suitable performances are obtained for noise levels around 80% for the chaotic logistic map and 95% for the chaotic Hénon map when considering the shortest sequences and order $D = 4$ (please see Figs. 4a–b). Obviously, better results are found for longer time series. Indeed, in the case of the

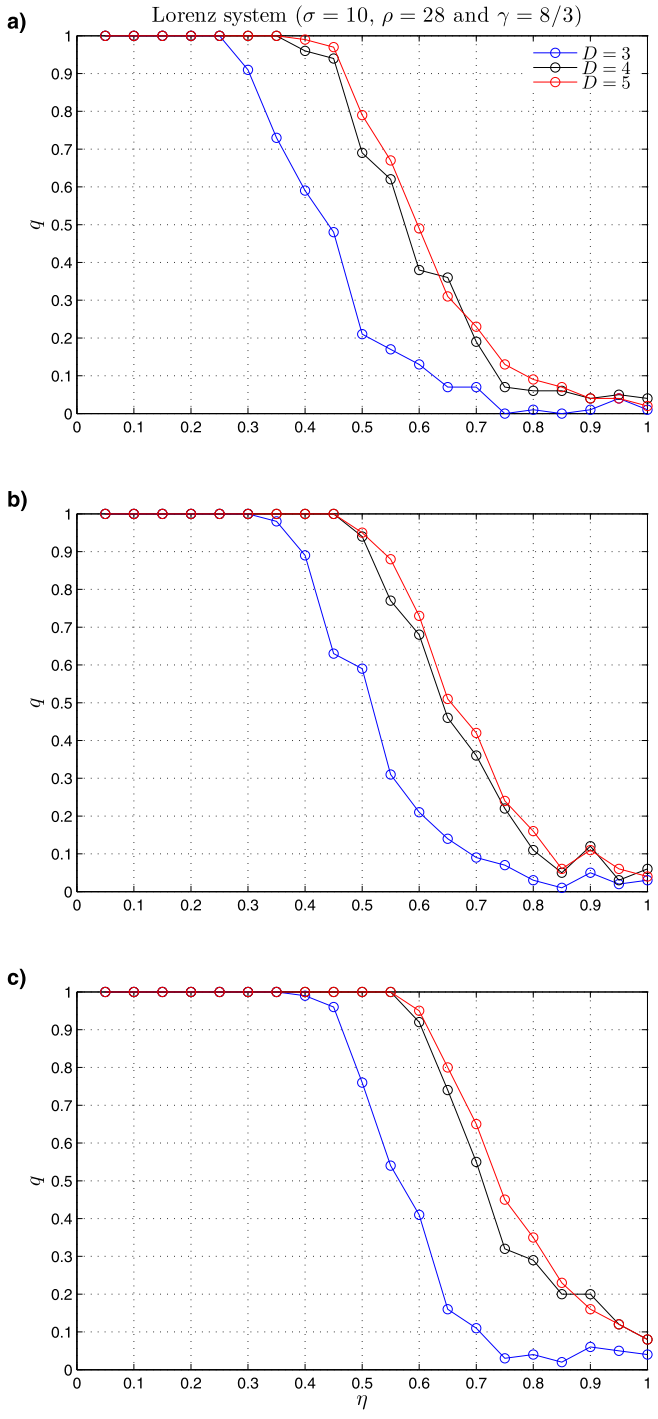

Fig. 5. Same as Fig. 4 but for the Lorenz system with parameters $\sigma = 10$, $\rho = 28$ and $\gamma = 8/3$. Results obtained for noisy chaotic sequences of lengths a) $N = 2^9$ samples, b) $N = 2^{10}$ samples and c) $N = 2^{11}$ samples are displayed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Hénon map and for the largest considered time series ($N = 2^{11}$ samples), a discrimination power very close to one is achieved for the three orders even for 100% of noise level (please see Fig. 4f)).

We have also developed a similar analysis for a continuous chaotic system. Being more precise, time series of length $N = 2^k$ data points with $k \in \{9, 10, 11\}$ of the x-variable of the Lorenz system:

$$\dot{x} = \sigma(y - x), \quad \dot{y} = x(\rho - z) - y, \quad \dot{z} = xy - \gamma z, \tag{5}$$

were generated with initial conditions $(x_0, y_0, z_0) = (1, 5, 10)$, and standard parameters $\sigma = 10$, $\rho = 28$ and $\gamma = 8/3$ for which the system exhibits chaotic behavior. Matlab's *ode45* function, that implements fourth and fifth order Runge–Kutta numerical integration algorithms, has been used to solve the system of differential equations with sampling period $\delta_t = 0.15$. The first 10^5 iterations were discarded to avoid possible transients. Again, observational Gaussian white noises of different levels have been added to characterize the noise effect. Results obtained for this continuous chaotic dynamics are shown in Fig. 5. In order to reach a discriminatory power of at least 0.7, the noise level can be as large as 50%, 60%, and around 70% for sequences of length $N = 2^9$, $N = 2^{10}$ and $N = 2^{11}$ data points, respectively. Better results are obtained for the larger orders ($D = 4$ and $D = 5$), similarly to the Hénon map analysis. Worse performances have been confirmed for other sampling periods ($\delta_t \in \{0.05, 0.1, 0.2, 0.25, 0.3, 0.35, 0.65\}$). As will be mentioned in Section 5, the study of the effects of sampling rate on the test is an open question and a planned avenue of future research.

4. Real-world applications

When dealing with real applications, the measured signals result from very complex dynamics and/or from coupled dynamics of many dimensional systems. Furthermore, these records are often contaminated by noise and other artifacts. Thus, extracting relevant features from them is usually a hard task to fulfill. Taking this into account, next we will test the performance of PE for detecting nonlinearity in these more challenging real-world situations.

4.1. Chaotic laser data

As a first nonlinear experimental example, we analyzed the chaotic intensity pulsations of a single-mode far-infrared NH_3 laser recorded by a LeCroy oscilloscope. This experimental time series was included in the Santa Fe Time Series Competition and it has been often used to check the performance of nonlinearity tests [4–6,12]. Further details of the recording procedure of this data set can be found in [47]. The longer data set ($N = 9,093$ data points) was considered. We have randomly selected one thousand subsequences from the original longer sequence of different sizes ($N \in \{50, 100, ..., 500\}$ data points) and, then, we have tested for nonlinearity by applying the proposed PE surrogate approach. As before, $M = 199$ IAAFT surrogates have been generated for each subsequence and a two-sided rank order test has been implemented for the rejection of the false null hypothesis. Fig. 6 shows the fraction q of successful rejections at significance level $\alpha = 0.01$ of the one thousand subsequences for each length N. It can be concluded that a tolerable performance, *i.e.* fraction of correct rejections of at least 0.7, is achieved with very short time series. Actually, for $D = 4$ and $D = 5$, one hundred samples are enough while around two hundred are needed for $D = 3$. Moreover, with $N = 200$ data points, an almost perfect discriminatory power is obtained for $D = 5$. Better performance is obtained for the larger values of D even when the condition $N \gg D!$ is not satisfied. We conjecture that this counter-intuitive result could be associated with the fact that, in order to detect nonlinearity, larger ordinal patterns provide more information about the underlying temporal structures, allowing an improved characterization of the intrinsic nonlinear dynamics.

4.2. Monthly smoothed total sunspot number

Following the analysis developed by De Domenico and Latora [6], we have tested the nonlinear dynamics for the monthly smoothed total sunspot number. These data were downloaded

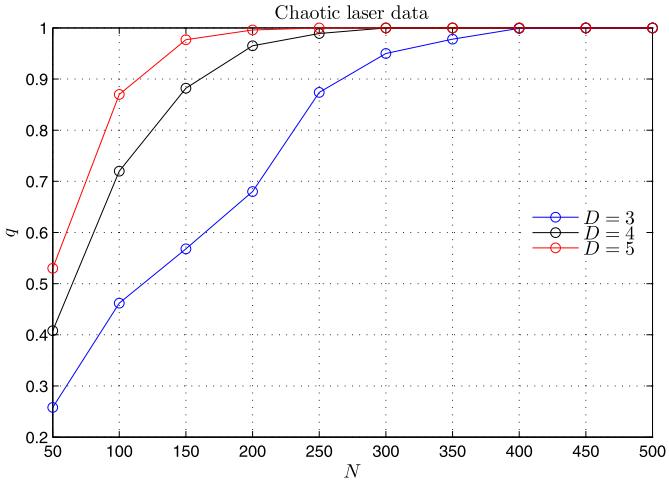

Fig. 6. The fraction q of correct rejections of a linear stochastic null hypothesis for experimental chaotic laser data. One thousand subsequences for each length N have been tested with the PE as nonlinearity measure and the IAAFT as surrogate method. A two-sided rank order test of size $\alpha = 0.01$ ($M = 199$ surrogates) has been implemented for rejecting linearity of each subsequence. The PE is estimated with different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

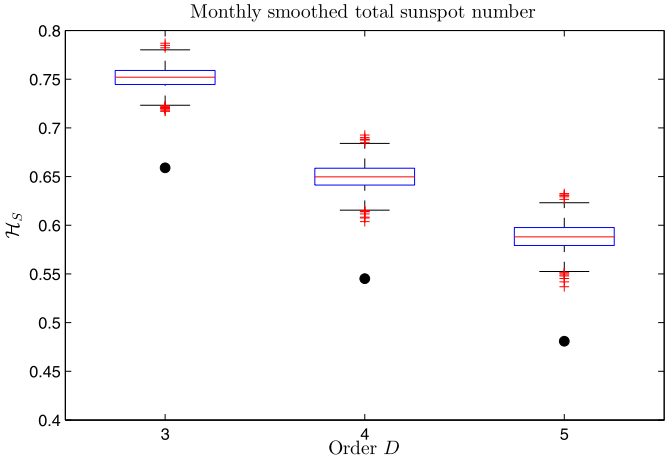

Fig. 7. Analysis of the monthly smoothed total sunspot number with the PE as nonlinear discriminators. $M = 999$ surrogates have been generated through the IAAFT algorithm. The PE has been estimated with different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. Black circles indicate the value estimated for the original time series while boxplots are used to display the distributions of estimated values for the surrogate realizations.

from the World Data Center SILSO, Royal Observatory of Belgium, Brussels. The 13-month smoothed monthly sunspot number from July 1749 to March 2016 ($N = 3{,}201$) has been considered. Further details about this database can be found at website: http://www.sidc.be/SILSO/. In this case, we have carried out a two-sided rank order test of size $\alpha = 0.002$ ($M = 999$ surrogates) for different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. As can be seen in Fig. 7, strong evidence in favor of a nonlinear underlying dynamics is clearly concluded. Being more precise, we can reject the null hypothesis of a linear stochastic process with 99.8% confidence level for the three orders. This finding is consistent with results obtained in [6], where rejection with 98% confidence level is achieved by applying the autoregressive-fit residuals kurtosis (ARK) as discriminator. As it was stated and discussed by De Domenico and Latora, this finding is significant for a better comprehension of the still debated mechanisms that govern the solar cycle dynamics. It is worth remarking here the higher confidence interval for the rejection obtained by implementing PE as nonlinearity measure.

4.3. Electroencephalograms from healthy and epileptic patients

Looking for signatures of nonlinear dynamical behavior in neuro-physiological data, we analyze five different sets of electroencephalogram (EEG) time series for different groups and recording regions: surface (scalp) EEG recordings from five healthy volunteers in an awake state with eyes open (Set A) and closed (Set B), intracranial EEG recordings from five epilepsy patients during the seizure free interval from outside (Set C) and from within (Set D) the seizure generating area, and intracranial EEG recordings of epileptic seizures (Set E). Each set contains one hundred single-channel EEG segments of 23.6 seconds of duration recorded at a sampling rate of 173.61 Hz ($N = 4{,}097$ data points). They were selected from continuous multichannel EEG recordings after visual inspection for artifacts. Additionally, only segments that satisfy a weak stationarity criterion were chosen. These artifact-free records are available at www.meb.unibonn.de/epileptologie/science/physik/eegdata.html. Further details about the recording technique of these EEG data can be found in the original paper by Andrzejak et al. [48]. We have tried to quantify the presence of nonlinearity in these five sets of EEGs through the PE discriminator estimated with different orders ($D = 3$, $D = 4$ & $D = 5$) and lag $\tau = 1$. For that purpose, $M = 999$ surrogates have been generated with the IAAFT algorithm for each one of these EEG original segments. Then, a two-sided rank order test was implemented for testing the null hypothesis of an underlying linear stochastic process. Fig. 8a) shows the fraction q of rejections for each set as a function of D. According to these results, a linear stochastic dynamics can be associated with EEG records of healthy subjects with eyes open (Set A) while nonlinear mechanisms are present during epileptic seizures (Set E). On the other hand, the number of rejections increases for the other data sets (B–D). This allows us to conjecture that the degree of nonlinearity is higher for the pathological cases. It should be stressed that these findings are in excellent accord with those obtained by Donges et al. [8] by implementing a complex network method on the same EEG database (please compare our results with those listed in Table 1 from [8]). We have also developed a multiscale analysis by testing the fraction q of rejection for the five sets as a function of the lag τ. Results obtained for the order $D = 5$ are displayed in Fig. 8b). Qualitative similar behaviors have been confirmed for $D = 3$ and $D = 4$. For $\tau = 4$ the fraction of rejections reaches a maximum for Sets E ($q = 0.95$) while Set A has a zero value. Likewise, at $\tau = 5$, the fraction of rejections is maximum for Set D ($q = 0.34$). Consequently, we can conclude that the nonlinear nature can be better discriminated at particular time scales. Finally, records from epilepsy patients during the seizure free interval from outside the seizure generating area (Set C) can not be distinguished from those obtained from healthy subjects (Sets A and B) through the proposed PE surrogate approach.

5. Conclusions

In this work, we have tested the permutation entropy as a nonlinear discriminative measure in a conventional surrogate framework. Numerically controlled analyses indicate that the proposed approach is efficient in recognizing the presence of nonlinear dynamics in several discrete and continuous chaotic systems, even when contaminated with a considerable amount of observational Gaussian white noise. Moreover, this discrimination can be significantly reached with time series of short length. The low fractions of rejections for true null hypothesis (linear Gaussian and non-Gaussian discrete and continuous models) confirm robustness

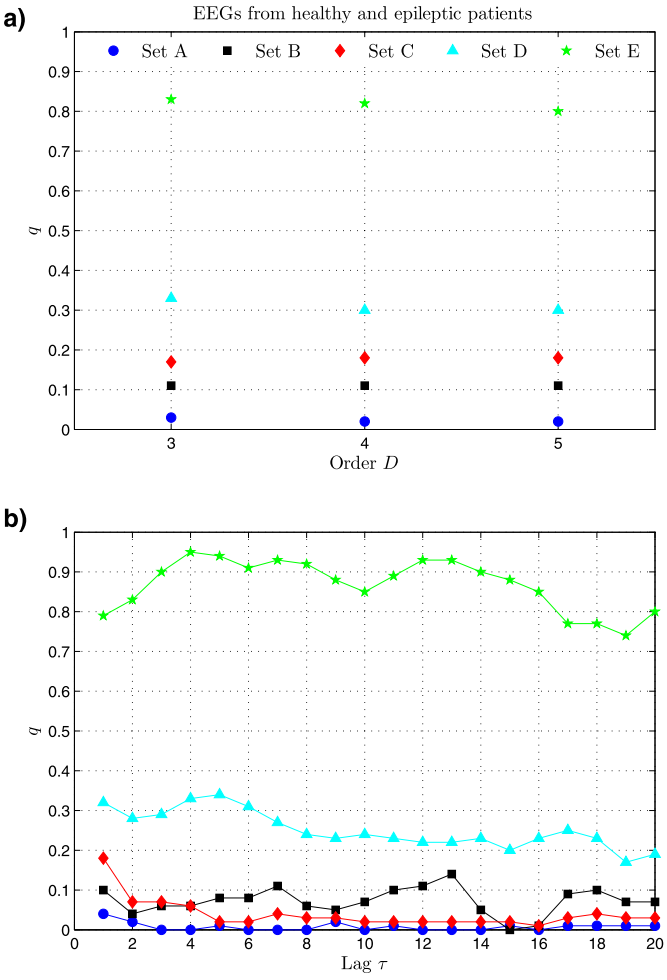

Fig. 8. a) The fraction q of rejections of a linear stochastic null hypothesis for the five different sets of EEGs as a function of the order D for lag $\tau = 1$. b) Analysis of q as a function of the lag τ for the order $D = 5$. Similar behaviors are obtained for $D = 3$ and $D = 4$. One hundred two-sided rank order tests of size $\alpha = 0.002$ ($M = 999$ surrogates) have been tested for each set. Set A (healthy, eyes open): blue circles; Set B (healthy, eyes closed): black squares; Set C (pathological, interictal, no epileptogenic zone): red diamonds; Set D (pathological, interictal, epileptogenic zone): cyan triangles; Set E (pathological, seizure activity): green stars. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

against *type-I errors* (false alarms). In addition, the high fractions of rejections obtained for false null hypothesis (chaotic regimes of discrete and continuous systems) evidence a significant discrimination power and, consequently, reliability against *type-II errors*. Finally, the PE has been implemented for testing the nonlinear dynamics in observational data from several natural and laboratory systems. Results obtained are in accord with those published in literature, confirming once again the versatility of this symbolic tool as nonlinear discriminating statistic.

According to the evidence gathered from the several analyses performed, we conclude that the surrogate method with the PE as a discriminating nonlinear measure can help shed new light on the detection of nonlinearities in very short and noisy time series, and ultimately, in a more appropriate physical interpretation of the underlying complex dynamics of real data. Several open methodological issues, such as the characterization of sampling effects and the analysis of irregularly sampled data for which the construction of surrogate data deserves special attention, will be considered in future research. We are also planning to assess the discriminative power of some generalized and improved ordinal descriptors, such as forbidden ordinal patterns [18,19], modified permutation entropy [49], weighted-permutation entropy [50], Tsallis permutation entropy [51], Rényi permutation entropy [52], and permutation min-entropy [53].

Acknowledgements

The authors thank two anonymous reviewers for their useful comments and suggestions that greatly helped to improve an earlier version of this paper. L.Z. gratefully acknowledges financial support from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. This work was supported by Lycoming College as part of C.W.K.'s sabbatical leave.

References

[1] M. Gómez Ravetti, L.C. Carpi, B.A. Gonçalves, A.C. Frery, O.A. Rosso, Distinguishing noise from chaos: objective versus subjective criteria using Horizontal Visibility Graph, PLoS ONE 9 (9) (2014) e108004, http://dx.doi.org/10.1371/journal.pone.0108004.

[2] Q. Li, Z. Fu, N. Yuan, Beyond Benford's law: distinguishing noise from chaos, PLoS ONE 10 (6) (2015) e0129161, http://dx.doi.org/10.1371/journal.pone.0129161.

[3] B. Ye, J. Chen, C. Ju, H. Li, X. Wang, Distinguishing chaotic time series from noise: a random matrix approach, Commun. Nonlinear Sci. Numer. Simul. 44 (2017) 284–291, http://dx.doi.org/10.1016/j.cnsns.2016.08.018.

[4] A. Gunduz, J.C. Principe, Correntropy as a novel measure for nonlinearity tests, Signal Process. 89 (1) (2009) 14–23, http://dx.doi.org/10.1016/j.sigpro.2008.07.005.

[5] S. Ramdani, F. Bouchara, J. Lagarde, Influence of noise on the sample entropy algorithm, Chaos 19 (1) (2009) 013123, http://dx.doi.org/10.1063/1.3081406.

[6] M.D. Domenico, V. Latora, Fast detection of nonlinearity and nonstationarity in short and noisy time series, Europhys. Lett. 91 (3) (2010) 30005, http://dx.doi.org/10.1209/0295-5075/91/30005.

[7] E. Roulin, U.S. Freitas, C. Letellier, Working conditions for safe detection of nonlinearity and noise titration, Phys. Rev. E 83 (4) (2011) 046225, http://dx.doi.org/10.1103/PhysRevE.83.046225.

[8] J.F. Donges, R.V. Donner, J. Kurths, Testing time series irreversibility using complex network methods, Europhys. Lett. 102 (1) (2013) 10004, http://dx.doi.org/10.1209/0295-5075/102/10004.

[9] E. Bradley, H. Kantz, Nonlinear time-series analysis revisited, Chaos 25 (9) (2015) 097610, http://dx.doi.org/10.1063/1.4917289.

[10] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J. Doyne Farmer, Testing for nonlinearity in time series: the method of surrogate data, Physica D 58 (1–4) (1992) 77–94, http://dx.doi.org/10.1016/0167-2789(92)90102-S.

[11] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (17) (2002) 174102, http://dx.doi.org/10.1103/PhysRevLett.88.174102.

[12] S. Basu, E. Foufoula-Georgiou, Detection of nonlinearity and chaoticity in time series using the transportation distance function, Phys. Lett. A 301 (5–6) (2002) 413–423, http://dx.doi.org/10.1016/S0375-9601(02)01083-6.

[13] J. Tony, E.A. Gopalakrishnan, E. Sreelekha, R.I. Sujith, Detecting deterministic nature of pressure measurements from a turbulent combustor, Phys. Rev. E 92 (6) (2015) 062902, http://dx.doi.org/10.1103/PhysRevE.92.062902.

[14] E.M. Bollt, T. Stanford, Y.-C. Lai, K. Życzkowski, Validity of threshold-crossing analysis of symbolic dynamics from chaotic time series, Phys. Rev. Lett. 85 (16) (2000) 3524–3527, http://dx.doi.org/10.1103/PhysRevLett.85.3524.

[15] E.M. Bollt, T. Stanford, Y.-C. Lai, K. Życzkowski, What symbolic dynamics do we get with a misplaced partition?: On the validity of threshold crossings analysis of chaotic time-series, Physica D 154 (3–4) (2001) 259–286, http://dx.doi.org/10.1016/S0167-2789(01)00242-1.

[16] Y. Cao, W.W. Tung, J.B. Gao, V.A. Protopopescu, L.M. Hively, Detecting dynamical changes in time series using the permutation entropy, Phys. Rev. E 70 (4) (2004) 046217, http://dx.doi.org/10.1103/PhysRevE.70.046217.

[17] O.A. Rosso, H.A. Larrondo, M.T. Martín, A. Plastino, M.A. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (15) (2007) 154102, http://dx.doi.org/10.1103/PhysRevLett.99.154102.

[18] J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, True and false forbidden patterns in deterministic and random dynamics, Europhys. Lett. 79 (5) (2007) 50001, http://dx.doi.org/10.1209/0295-5075/79/50001.

[19] M. Zanin, Forbidden patterns in financial time series, Chaos 18 (1) (2008) 013119, http://dx.doi.org/10.1063/1.2841197.

[20] A. Bahraminasab, F. Ghasemi, A. Stefanovska, P.V.E. McClintock, H. Kantz, Direction of coupling from phases of interacting oscillators: a permutation information approach, Phys. Rev. Lett. 100 (8) (2008) 084101, http://dx.doi.org/10.1103/PhysRevLett.100.084101.

[21] M. Staniek, K. Lehnertz, Symbolic transfer entropy, Phys. Rev. Lett. 100 (15) (2008) 158101, http://dx.doi.org/10.1103/PhysRevLett.100.158101.

[22] L. Zunino, M.C. Soriano, I. Fischer, O.A. Rosso, C.R. Mirasso, Permutation-information-theory approach to unveil delay dynamics from time-series analysis, Phys. Rev. E 82 (4) (2010) 046212, http://dx.doi.org/10.1103/PhysRevE.82.046212.
[23] Q. Li, Z. Fu, Permutation entropy and statistical complexity quantifier of nonstationarity effect in the vertical velocity records, Phys. Rev. E 89 (1) (2014) 012905, http://dx.doi.org/10.1103/PhysRevE.89.012905.
[24] C.W. Kulp, L. Zunino, Discriminating chaotic and stochastic dynamics through the permutation spectrum test, Chaos 24 (3) (2014) 033116, http://dx.doi.org/10.1063/1.4891179.
[25] P.J. Weck, D.A. Schaffner, M.R. Brown, R.T. Wicks, Permutation entropy and statistical complexity analysis of turbulence in laboratory plasmas and the solar wind, Phys. Rev. E 91 (2) (2015) 023101, http://dx.doi.org/10.1103/PhysRevE.91.023101.
[26] J.P. Toomey, A. Argyris, C. McMahon, D. Syvridis, D.M. Kane, Time-scale independent permutation entropy of a photonic integrated device, J. Lightwave Technol. 35 (1) (2017) 88–95, http://dx.doi.org/10.1109/JLT.2016.2626387.
[27] J. Garland, R. James, E. Bradley, Model-free quantification of time-series predictability, Phys. Rev. E 90 (5) (2014) 052910, http://dx.doi.org/10.1103/PhysRevE.90.052910.
[28] A. Politi, Quantifying the dynamical complexity of chaotic time series, Phys. Rev. Lett. 118 (14) (2017) 144101, http://dx.doi.org/10.1103/PhysRevLett.118.144101.
[29] C. Bandt, G. Keller, B. Pompe, Entropy of interval maps via permutations, Nonlinearity 15 (5) (2002) 1595–1602, http://dx.doi.org/10.1088/0951-7715/15/5/312.
[30] J.M. Amigó, The equality of Kolmogorov–Sinai entropy and metric permutation entropy generalized, Physica D 241 (7) (2012) 789–793, http://dx.doi.org/10.1016/j.physd.2012.01.004.
[31] K. Keller, A.M. Unakafov, V.A. Unakafova, On the relation of KS entropy and permutation entropy, Physica D 241 (18) (2012) 1477–1481, http://dx.doi.org/10.1016/j.physd.2012.05.010.
[32] C.S. Daw, C.E.A. Finney, E.R. Tracy, A review of symbolic analysis of experimental data, Rev. Sci. Instrum. 74 (2) (2003) 915–930, http://dx.doi.org/10.1063/1.1531823.
[33] M. Staniek, K. Lehnertz, Parameter selection for permutation entropy measurements, Int. J. Bifurc. Chaos 17 (10) (2007) 3729–3733, http://dx.doi.org/10.1142/S0218127407019652.
[34] U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, N. Wessel, Classifying cardiac biosignals using ordinal pattern statistics and symbolic dynamics, Comput. Biol. Med. 42 (3) (2012) 319–327, http://dx.doi.org/10.1016/j.compbiomed.2011.03.017.
[35] L. Zunino, M.C. Soriano, O.A. Rosso, Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach, Phys. Rev. E 86 (4) (2012) 046210, http://dx.doi.org/10.1103/PhysRevE.86.046210.
[36] A. Aragoneses, L. Carpi, N. Tarasov, D.V. Churkin, M.C. Torrent, C. Masoller, S.K. Turitsyn, Unveiling temporal correlations characteristic of a phase transition in the output intensity of a fiber laser, Phys. Rev. Lett. 116 (3) (2016) 033902, http://dx.doi.org/10.1103/PhysRevLett.116.033902.
[37] M. Zanin, L. Zunino, O.A. Rosso, D. Papo, Permutation entropy and its main biomedical and econophysics applications: a review, Entropy 14 (8) (2012) 1553–1577, http://dx.doi.org/10.3390/e14081553.
[38] M. Riedl, A. Müller, N. Wessel, Practical considerations of permutation entropy. A tutorial review, Eur. Phys. J. Spec. Top. 222 (2) (2013) 249–262, http://dx.doi.org/10.1140/epjst/e2013-01862-7.
[39] T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests, Phys. Rev. Lett. 77 (4) (1996) 635–638, http://dx.doi.org/10.1103/PhysRevLett.77.635.
[40] C. Räth, M. Gliozzi, I.E. Papadakis, W. Brinkmann, Revisiting algorithms for generating surrogate time series, Phys. Rev. Lett. 109 (14) (2012) 144101, http://dx.doi.org/10.1103/PhysRevLett.109.144101.
[41] R. Nagarajan, Surrogate testing of linear feedback processes with non-Gaussian innovations, Physica A 366 (2006) 530–538, http://dx.doi.org/10.1016/j.physa.2005.10.041.
[42] R. Nagarajan, J. Szczepanski, E. Wajnryb, Interpreting non-random signatures in biomedical signals with Lempel–Ziv complexity, Physica D 237 (3) (2008) 359–364, http://dx.doi.org/10.1016/j.physd.2007.09.007.
[43] H.A. Makse, S. Havlin, M. Schwartz, H.E. Stanley, Method for generating long-range correlations for large systems, Phys. Rev. E 53 (5) (1996) 5445–5449, http://dx.doi.org/10.1103/PhysRevE.53.5445.
[44] M. Gómez-Extremera, P. Carpena, P.C. Ivanov, P.A. Bernaola-Galván, Magnitude and sign of long-range correlated time series: decomposition and surrogate signal generation, Phys. Rev. E 93 (4) (2016) 042201, http://dx.doi.org/10.1103/PhysRevE.93.042201.
[45] T. Schreiber, A. Schmitz, Surrogate time series, Physica D 142 (3–4) (2000) 346–382, http://dx.doi.org/10.1016/S0167-2789(00)00043-9.
[46] T. Schreiber, A. Schmitz, Discrimination power of measures for nonlinearity in a time series, Phys. Rev. E 55 (5) (1997) 5443, http://dx.doi.org/10.1103/PhysRevE.55.5443.
[47] U. Hübner, N.B. Abraham, C.O. Weiss, Dimensions and entropies of chaotic intensity pulsations in a single-mode far-infrared NH3 laser, Phys. Rev. A 40 (11) (1989) 6354–6365, http://dx.doi.org/10.1103/PhysRevA.40.6354.
[48] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907, http://dx.doi.org/10.1103/PhysRevE.64.061907.
[49] C. Bian, C. Qin, Q.D.Y. Ma, Q. Shen, Modified permutation-entropy analysis of heartbeat dynamics, Phys. Rev. E 85 (2) (2012) 021906, http://dx.doi.org/10.1103/PhysRevE.85.021906.
[50] B. Fadlallah, B. Chen, A. Keil, J. Príncipe, Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information, Phys. Rev. E 87 (2) (2013) 022911, http://dx.doi.org/10.1103/PhysRevE.87.022911.
[51] L. Zunino, D.G. Pérez, A. Kowalski, M.T. Martín, M. Garavaglia, A. Plastino, O.A. Rosso, Fractional Brownian motion, fractional Gaussian noise, and Tsallis permutation entropy, Physica A 387 (24) (2008) 6057–6068, http://dx.doi.org/10.1016/j.physa.2008.07.004.
[52] X. Zhao, P. Shang, J. Huang, Permutation complexity and dependence measures of time series, Europhys. Lett. 102 (4) (2013) 40005, http://dx.doi.org/10.1209/0295-5075/102/40005.
[53] L. Zunino, F. Olivares, O.A. Rosso, Permutation min-entropy: an improved quantifier for unveiling subtle temporal correlations, Europhys. Lett. 109 (1) (2015) 10005, http://dx.doi.org/10.1209/0295-5075/109/10005.