Quantifying entropy using recurrence matrix microstates

Gilberto Corso, Thiago de Lima Prado, Gustavo Zampier dos Santos Lima, Jürgen Kurths, and Sergio Roberto Lopes

Citation: Chaos **28**, 083108 (2018); doi: 10.1063/1.5042026
View online: https://doi.org/10.1063/1.5042026
View Table of Contents: http://aip.scitation.org/toc/cha/28/8
Published by the American Institute of Physics

Quantifying entropy using recurrence matrix microstates

Gilberto Corso,[1] Thiago de Lima Prado,[2] Gustavo Zampier dos Santos Lima,[3] Jürgen Kurths,[4,5] and Sergio Roberto Lopes[4,5,6,a)]

[1]*Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, Natal 59078-970, Brazil*
[2]*Instituto de Engenharia, Ciência e Tecnologia, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Janaúba 39440-000, Brazil*
[3]*Escola de Ciências e Tecnologia, Universidade Federal do Rio Grande do Norte, Natal 59078-970, Brazil*
[4]*Potsdam Institute for Climate Impact Research, Telegraphenberg A 31, 14473 Potsdam, Germany*
[5]*Department of Physics, Humboldt University Berlin, Berlin 12489, Germany*
[6]*Departamento de Física, Universidade Federal do Paraná, Curitiba 81531-980, Brazil*

(Received 29 May 2018; accepted 16 July 2018; published online 24 August 2018)

We conceive a new recurrence quantifier for time series based on the concept of information entropy, in which the probabilities are associated with the presence of microstates defined on the recurrence matrix as small binary submatrices. The new methodology to compute the entropy of a time series has advantages compared to the traditional entropies defined in the literature, namely, a good correlation with the maximum Lyapunov exponent of the system and a weak dependence on the vicinity threshold parameter. Furthermore, the new method works adequately even for small segments of data, bringing consistent results for short and long time series. In a case where long time series are available, the new methodology can be employed to obtain high precision results since it does not demand large computational times related to the analysis of the entire time series or recurrence matrices, as is the case of other traditional entropy quantifiers. The method is applied to discrete and continuous systems. *Published by AIP Publishing.* https://doi.org/10.1063/1.5042026

A large number of quantifiers of complexity based on data can be found in the literature. In general, all these quantities were developed to distinguish regular, chaotic, and random properties of a time series. The knowledge of these properties is fundamental since it has been reported that complexity is an important characteristic of data of heart and brain signals, stock market, climatology, seismology, etc. The main types of complexity measures are entropies, Lyapunov's exponents, and fractal dimensions. Here, we develop a new entropy based on recurrence properties of a time series. The new recurrence entropy has a good correlation with the maximum Lyapunov exponent of the system and a weak dependence on the vicinity threshold parameter, a critical parameter on recurrence analysis. Furthermore, the new method works adequately even for small segments of data, bringing consistent results for short and long time series.

I. INTRODUCTION

Nonlinear time series analyses try to extract information from the underlying dynamics of the data. In this way, nonlinear techniques supply new tools for data diagnostics using a whole set of quantities such as divergence rates, predictability, scaling exponents, and entropies in symbolic representation.[1,2] All these methods are based on more general phase space properties like trajectory recurrences. Nonlinear time series analysis is a practical by-product from complex dynamical systems theory since nonlinear concepts allow extracting information that cannot be resolved using classical linear techniques, e.g., power spectrum or spectral coherence.[2,3] These analyses have influenced many areas of science, from physics, chemistry, and engineering to life science and ecology; from economics to linguistics, and more recently, neuroscience.[3–8] Time series analysis has turned out to be a key issue providing the most direct link between nonlinear dynamics and the real world.[2]

In this work, we explore a new entropy quantifier of nonlinear time series based on more general properties of the recurrence space than just its periodicities in time or space. It intends to be useful in the analysis of nonlinear properties of the signal, particularly its (dis)order and/or chaoticity levels, concepts closely related to the Lyapunov exponent of the system, a more familiar nonlinear quantifier of time series, but harder to estimate.[2]

The concept of recurrence dates back to the work of Poincaré[9] and it is a fundamental attribute of dynamical systems. A visualization method known as recurrence plot (**RP**) was introduced later on by Eckmann *et al.*[10] The **RP** is a graphical tool to identify recurrence of a trajectory $\mathbf{x}_i \in \mathbb{R}^d$ phase space, $i \in 1, 2, \ldots, K$, and it is based on the recurrence matrix[10]

$$\mathbf{R}_{ij} = \begin{cases} 1 & \text{if } ||\mathbf{x}_i - \mathbf{x}_j|| \leq \varepsilon, \\ 0 & \text{if } ||\mathbf{x}_i - \mathbf{x}_j|| > \varepsilon, \end{cases} \quad i,j \in 1, 2, \ldots, K, \qquad (1)$$

where $||\cdot||$ is an appropriate norm. ε is the vicinity threshold parameter consisting of a maximum distance among two points in a trajectory such that both can be considered recurrent to each other, and K is the length of the analyzed time series. The **RP** is a matrix of "ones" and "zeros," where one (zero) indicates a recurrent (non-recurrent) pair of points in phase space.

a)Electronic mail: lopes@fisica.ufpr.br

An example of **RP** is depicted in Fig. 1, where a "one" ("zero") is represented by a black (white) pixel. The recurrence analysis technique is conceptually simpler than spectral linear (Fourier) or nonlinear (Wavelets) analyses and numerically easier to perform since it does not have to decompose the signal within a basis.[11] Instead, the **RP** is computed using repetitions (or recurrences) of segments of the signal which produce a time mosaic of the recurrence signal, an imprinting of signal time patterns.

Based on the statistical properties of the **RP**, a large number of quantifiers have been developed to analyze its details.[12] Many of them deal with statistical properties like mean lengths and frequency of occurrence of diagonal/vertical/horizontal recurrence lines. An important class of recurrence quantifiers are those that try to capture the level of disorder of a signal. As an example, we mention the entropy based on diagonal line statistics.[13,14] This quantity has been correlated with other dynamical quantifiers such as the largest Lyapunov exponent since both capture properties of the disorder level of the dynamics.[12]

Nevertheless, at this point, it is important to emphasize that sometimes the diagonal entropy (ENTR)[12] as originally defined behaves in an unexpected way; indeed, a well known problem occurs in the recurrence quantification of the dynamics of the logistic map, namely, ENTR decreases despite the increase of nonlinearity and chaoticity (evaluated by the increase of the Lyapunov exponent).[15] In fact, to deal with that, an adaptive method has been presented to compute recurrences and to conciliate the behavior of a decreasing ENTR with the increasing disorder of the logistic map.[15]

Another important method for recurrence-based entropies has been built based on the concept of weighted recurrence plots.[16] The restriction of unweighted **RP**s imposes a strong dependency on the vicinity threshold parameter, a free parameter in recurrence analyses. A weighted variant of the **RP** allows one to relax the condition of defining the vicinity threshold parameter and its sensitive dependence on the results. Using the concept weighted recurrence, the associated Shannon entropy also correlates with the largest Lyapunov exponent enabling the computation of the Shannon entropy of weighted recurrence plots based on the distances between the points in the phase space.

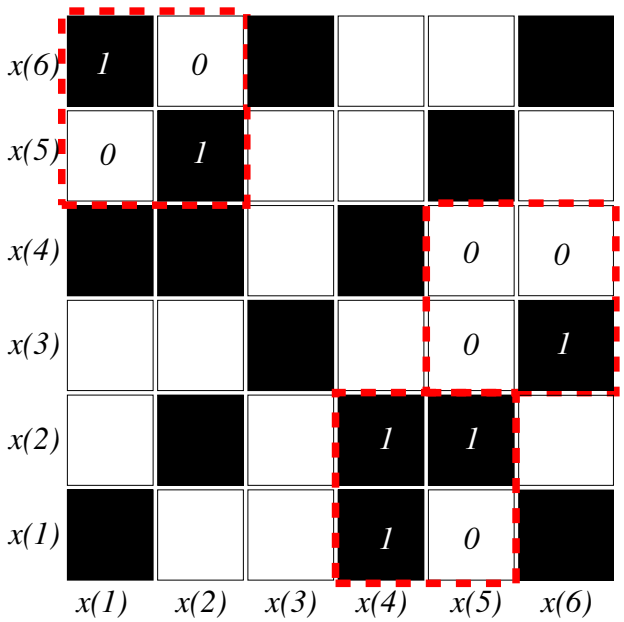

FIG. 1. An example of **RP** of a short 6 elements time series. Black (white) pixels denote a pair of recurrent (non-recurrent) points. Red dashed squares are representative of recurrence microstates randomly selected in the **RP** and obtained for $N = 2$.

Out of the recurrence based methods to compute entropy and evaluate the complexity of a time series, other promising methods can be found. One of the most qualified methods is the permutation entropy,[17] a method based on comparisons of neighboring values in real time series. Permutation entropy transforms raw time series into a corresponding sequence of symbols. The diversity observed in the set of symbols can be used to quantify data complexity. In fact, a large number of other methods to measure complexity and correlated topics such as fractal or correlation dimensions, Lyapunov's exponents, and mutual information have been widely developed to compare time series and distinguish regular, chaotic, and random behaviors as well as serving as a tool to quantify complexity.[2,18]

Here, we develop a new entropy recurrence quantifier based on the formal definition of system entropies and making use of small square matrices of $N \times N$ elements (we show results for matrix sizes up to 4×4) defined as *recurrence microstates*. In order to introduce this new quantifier, we consider a **RP** selecting on it random samples of microstates. The extraction of microstate samples of the **RP** allows us to define an ensemble of microstates. As we shall see, a sufficiently large ensemble of microstates reflects well the dynamical recurrent patterns of the time series. Indeed, the definition of an entropy quantifier based on the frequency that each microstate appears on **RP** provides a quantifier that correctly captures the relation between the entropy and the level of disorder and/or chaoticity of the time series, fixing the behavior displayed by the former diagonal entropy quantifier ENTR.[12] As we will show, making use of just one time series, e.g., one variable of a dynamical system, it also provides substantial results for the identification of major dynamical changes that occur in dynamical systems, including a close correlation with the maximal Lyapunov exponent of the system. The new methodology can also be faster compared to other computed entropies such as permutation and recurrence matrix based methods.

The rest of this paper is organized as follows: in Sec. II, we briefly show the recurrence technique and present our methodology. In Sec. III, we apply the new methodology to three well-known time series, namely, a white noise signal, a discrete logistic map time series, and a time series of the well known nonlinear flux described by the Lorenz equations. Finally, in Sec. IV, we discuss our results and point out future perspectives.

II. METHODOLOGY

The methodology section is divided into two parts. First, we introduce the recurrence analysis technique based on the recurrence matrix $\mathbf{R}_{ij}$, as defined in Eq. (1), and in Sec. II B,

we define a way to represent microstates extracted from the recurrence matrix, which is used to calculate our proposed entropy.

A. Recurrence plots and recurrence quantification analysis

A complete compilation in the literature of this issue is given in Ref. 12, and a particular use of recurrence plots for spatial profiles is developed in Refs. 19 and 20. An example of a recurrence plot for a short six element time series is depicted in Fig. 1, where a (white) black pixel denotes a (non) recurrence between two points (i,j) of the time series. An accurate extraction of specific features of a time series based on the **RP** can be obtained by using a set of tools developed initially by Zbilut and Webber.[13,14] These tools are called recurrence quantification analysis (**RQA**) or just recurrence quantifiers and measure dynamical properties of a signal based on the recurrence matrix.

The **RQA** studies different aspects of the **RP**, from the density of recurrent (non-recurrent) points to the statistics of vertical, horizontal, or diagonal lines.[12] An important question in recurrence analysis is the measure of diagonal lines that represent recurrence segments of trajectories as the one composed by elements $x(4), x(5)$, and $x(6)$ in Fig. 1. Diagonal lines are $\mathbf{R}_{ij}$ structures parallel to the line of identity [the main diagonal of the $\mathbb{RP}$ defined as $\mathbf{R}_{i+k,j+k} = 1$ $(i,j \in 1,2,\ldots,K-\ell; k \in 1,2,\ldots,\ell)$ and $\mathbf{R}_{i,j} = \mathbf{R}_{i+\ell+1,j+\ell+1} = 0$, where ℓ is the length of the diagonal line]. Two pieces of a trajectory following a diagonal line undergo for a certain time (the length of the diagonal) a similar evolution once they have visited the same region of phase space at different times. This is the key idea behind recurrence and thus a clearcut signature of a deterministic behavior in the time series. Accordingly, $P(\ell)$, $\ell \in 1,2,\ldots,K$, can be defined as the distribution of the lengths of diagonal lines, ℓ_i, and it is used to compose a recurrence quantifier based on the Shannon entropy[12]

$$\mathcal{S} = -\sum_{i=1}^{Q} p(i)\,\ln p(i), \tag{2}$$

where $p(i)$ measures the probability of occurrence of a specific state i. Q is the number of accessible states such that $\mathcal{S}$ captures, in this sense, how much information resides on a collection of states. The Shannon entropy is an open tool that can be adapted to any probability space $p(i)$. The **RP** space of probability offers many different possibilities, e.g., the probability distribution of diagonal lines $p(\ell) = P(\ell)/\sum_{\ell=1}^{K} P(\ell)$. In that context, Eq. (2) assumes the form: $\mathrm{ENTR} = -\sum_{\ell=\ell_{\min}}^{\ell_{\max}} p(\ell)\,\ln p(\ell)$. Despite the use of this tool in the literature,[14] it presents some inconsistency as reported by Letellier.[15] While the entropy was primarily conceived as a quantification of disorder, this first approach based on entropy when applied to the chaotic logistic map provides an unsatisfactory result, sometimes indicating a more organized regime for an arising chaoticity level.

B. A new recurrence entropy

In this paper, we developed a novel methodology to extract information from the recurrence matrix to properly define an entropy. Particularly, we introduce a new concept of microstates for a **RP** that is associated with features of the dynamics of the time series. These microstates are defined as small matrices of dimension $N \times N$ that are randomly sampled on the **RP**. These so called *recurrence microstates* allow several configurations, as exemplified in Fig. 1 as dashed red squares, for the simplest situation of $N = 2$. Since the **RP** is a binary matrix, where recurrences are graphically depicted by black pixels assuming a value 1, while non-recurrencies are displayed by white pixels assuming a value 0, the three microstates shown in Fig. 1 are **1001**, **0001**, and **1110** for the upper left, middle right, and down right squares, respectively. In general, the number of microstates for a given N is $N^* = 2^{N^2}$, such that $N^* = 16$ for $N = 2$.

Defining n_i as the number of times that a microstate i is observed in $\bar{N}$ samples, so $P_i = n_i/\bar{N}$ is the probability related to the microstate i and we define an entropy of the **RP** associated with the probabilities of occurrence of microstates as

$$S(N^*) = -\sum_{i=1}^{i=N^*} P_i\,\ln P_i. \tag{3}$$

Although N^* grows quickly as a function of N, for all examples explored here and probably for many other examples, just a small number of microstates are effectively populated. We have observed that for all deterministic systems treated here, for $N = 3$ just $\sim$30 (5% of N^*) different microstates effectively exist in the distribution of randomly selected microstates. For $N = 4$ only 0.3% of N^* are populated. For a stochastic signal where we expect the largest number of populated microstates, numerical simulations show that this number does not grow significantly, reaching $\sim$20% of N^* for $N = 3$ and just $\sim$1.5% of N^* for $N = 4$. So, the effective set of microstates needed to compute adequately the entropy can be populated by just $\bar{N}$ random samples obtained from the recurrence matrix, and a fast convergence of Eq. (3) is expected. In general, we found that $\bar{N} \ll N^*$ for $N > 3$ such that $\bar{N} \sim 10\,000$ is enough for all cases treated here, and probably it is enough for many other situations, turning the method extremely fast even for moderate values of microstate sizes N. This observation also points out that a microstate size $N = 4$ is sufficient for many dynamical and stochastic systems.

As defined, the microstates are representative of all possible short time recurrence patterns of the system. Nevertheless, a particular time series is characterized by a reduced set of populated microstates reflecting the diversity of possible sequence of values (trajectories or projections of trajectories) departing from a given value of the time series.

A clear advantage of the new methodology to compute the entropy using Eq. (3) over the former diagonal entropy is the possibility of computing information over all possible microstates. In addition, it is possible to estimate analytically the maximum value of the entropy $S(N^*)$ corresponding to the case in which all populated microstates are equally probable.

For this case $P_i = 1/N^*$ and we have

$$S(N^*) = \ln N^* = N^2 \ln 2. \tag{4}$$

Analogously, the minimum value of the entropy corresponds to the situation in which all sampling matrices are at the same microstate and $S(N^*) = 0$. We will show next that our approach offers advantages when compared to the more traditional recurrence entropy quantifier, ENTR of non-recurrent points as defined by Letellier.[15] In particular, we show that S computed using Eq. (3) leads to a similar behavior of the entropy of non-recurrent points, but our method is faster for time series length $K > 500$ and more precise results can be acquired when applied to short length time series ($K < 100$). On the other hand, the new methodology also facilitates an analysis using longer time series lengths, since the computational effort does not depend on the length of the time series analyzed but only on the number of samples ($\bar{N}$) extracted from the time series.

III. RESULTS

To explore in detail the results obtained by the novel methodology to compute the recurrence entropy, we apply this tool to three illustrative data: white noise, logistic map, and the Lorenz system time series. We test the robustness of entropy against (i) the vicinity parameter ε and the microstate sizes N, (ii) the length K of the time series, and (iii) the number of randomly selected samples $\bar{N}$ used to populate the set of possible microstates. Moreover, we compare the novel methodology against other well studied recurrence quantification entropy methods.

A. White noise

First, let us use a white noise time series case to construct a model for the maximal entropy in the recurrence plot methodology. The lack of correlation of the white noise is implicated in a theoretical maximum entropy. This result is straightforward, but we have to take into account the border effects of the recurrence space in the methodology to compute a correct result of S. In this simple model, it is possible to extract the exact value of ε for which the entropy is a maximum. Consider a random data signal x, being $0 \leqslant x \leqslant 1$. It is clear that for the vicinity parameter $\varepsilon = 1$, the recurrence rate (RR) is maximum and equal to 1, independently of the point where ε is centered on the time series. In this case, all points in phase space are recurrent. Nevertheless, when the vicinity parameter is less than 1, $\varepsilon = 0.5$, for example, the particular position in the time series where the vicinity is computed is important due to the boundaries of the recurrence space. Suppose that we are computing the recurrence points of the first value in the time series, in this case there are no recurrence points on the left side and just half of the phase space is recurrent. In another circumstance, for the same value of $\varepsilon = 0.5$ but considering the central point of the time series, the same methodology results that the entire phase space will be recurrent to this point.

In the case of white noise, we expect that the entropy will be maximal for RR $= 0.5$ for which all populated states are equally populated. However, the computation of RR must take into account that the threshold ε needs to be considered for all points of the data, even those close to the first value of the time series where it captures less recurrent points. We explore in more detail this situation in Fig. 2, where we take into consideration the position for which ε is computed in the time series. Figure 2 depicts a graphical view of the recurrence space border effect. The figure illustrates the percentage of the recurrence phase space that will be recurrent versus the position for which the threshold of the vicinity parameter ε is computed, for different values of ε. We emphasize that this uniformity can be assumed only for white noise. For simplicity, we use a normalized phase space. In Fig. 2, RR is computed as the area below each trapezoidal curve. Figure 2 also highlights a particular case for which $\varepsilon = 0.1$ (dashed line), showing explicitly that it does not correspond to 10% or 20% of the phase space as recurrent but an intermediate amount. The general expression for the value of RR as a function of ε can be written as

$$\mathrm{RR} = 2\varepsilon - \varepsilon^2. \tag{5}$$

Finally, since we expect RR$= 0.5$ for a maximal entropy of the white noise case, Eq. (5) gives us the optimum value of the threshold, namely, $\varepsilon \approx \mathbf{0.293}$. In fact, despite some numerical inaccuracy, Fig. 3 confirms this maximum for three different microstate sizes, N.

To understand better the behavior of the entropy S, Fig. 3 depicts the behavior of S normalized by its maximum possible value $S_{\max}(N) = N^2 \ln 2$ as a function of the vicinity threshold ε. We employ three values of $N = 2, 3, 4$, as indicated in the legend. Considering an infinite set of data, a uniform distribution of all extracted microstates and a situation where $RR = 0.5$, the entropy should present a maximum $S_{\max} = N^2 \ln 2$ when applied to random data. It occurs for a unique ε value. Nevertheless, Fig. 3 shows that for finite time series, we observe that starting in a vanishing value, an

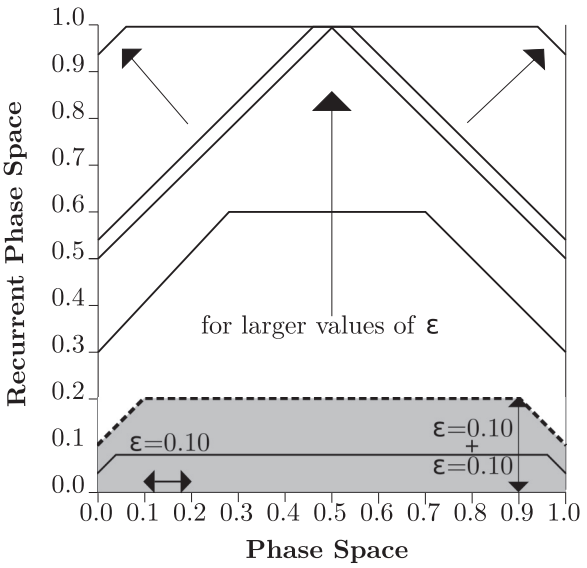

FIG. 2. Graphical representation of the recurrence rate as trapezoidal areas in the (normalized) recurrence phase space for white noise. The trapezoidal shape is expected since points on the left (right) borders of the phase space do not have recurrences on the left (right) sides. Here, the y axis gives the recurrence percentage of the phase space for different values of ε.

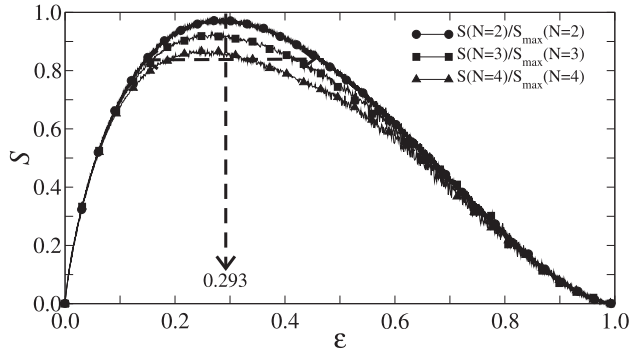

FIG. 3. Entropy S normalized by the maximum entropy of random data, $N^2 \ln(2)$, as a function of the recurrent threshold (ε) for 3 different microstate sizes, and applied to white pseudo-random data series. Note that the curve shape is insensible on the amount of possible microstates. Considering the entropy should reach its maximum for random data, we assume a range around max(S) as the optimum interval to be used for the vicinity threshold. As observed, a large interval of values of ε exists, spanning from approximately $0.14 \lesssim \varepsilon \lesssim 0.45$, independently of the microstate size, N, used.

increasing ε leads the entropy S to reach a maximum value decreasing for larger ε. In this general case, we define a conceivable interval of validity of the analyses using S, namely, an interval around its maximum. This result puts in evidence the resilience of the methodology against the vicinity threshold parameter. Figure 3 shows that the method allows the use of a large interval for ε, spreading from values around 0.14 until 0.45 for all microstate sizes tested when applied to white noise. For deterministic signals, the shape of the curve observed in Fig. 3 is still present. Nevertheless, the maximum occurs for smaller values of ε since in these cases, the distinction between structures (diagonal, vertical, or horizontal lines) in the **RP** is more relevant. In deterministic cases, we have used $\varepsilon = 0.13$. The behaviors for $\varepsilon \to 0$ ($\varepsilon \to 1$) in Fig. 3 are expected since in these limits of ε there will be just non-recurrent (recurrent) points in the **RP**. In these cases, for microstate size $N = 2$ as exemplified in Fig. 1, the microstates **0000** or **1111** dominate the distribution and the entropy level results to be small. In this way, intermediary values of ε will produce a richer distribution of microstates among classes and a more confident entropy output. The length of the optimum ε interval depends on the number of samples we take from the recurrence matrix, $\bar{N}$, but has a minor dependency on the size of the microstates, as observed in Fig. 3.

B. The logistic map

To explore further the new entropy methodology, we test the behavior of S as a function of system parameters considering a case of chaotic dynamics. We also compare our results against other more traditional entropies, namely, the non-recurrence points entropy,[15] the entropy of weighted recurrence plots,[16] and the permutation entropy.[17] To do so, we employ the paradigmatic logistic map[21,22] defined by

$$x_{n+1} = r x_n (1 - x_n). \tag{6}$$

The parameter r controls the non-linearity of the system and is responsible for the bifurcation cascade route to chaos and windows of periodic behavior as shown in the well known bifurcation diagram [Fig. 4(f)].

Figure 4(a) shows the normalized (by its maximum) entropy S plotted against r for a microstate size $N = 4$ and for time series of length $K = 1000$, $\bar{N} = 10\,000$ random samples, and $\varepsilon = 0.13$. Figures 4(b)–4(d) depict results for normalized entropies computed as in Refs. 15–17. For panels (c) and (d), both entropies are evaluated in the interval $(3.5 < r < 4.0)$ since the methods do not give acceptable results for periodic orbits. Visually, all methodologies to compute the system entropy lead to similar results. To explore further the behavior of the entropy S, Fig. 4(e) plots the Lyapunov exponent of the logistic map as a function of r and Fig. 4(f) depicts its bifurcation diagram. As it is the case of the entropies computed in Refs. 15–17, one of the main results of the entropy S defined here is its strong correlation to the Lyapunov exponent for all values of the nonlinear parameter r. Notwithstanding, we would like to call attention for the fact that S captures good results even for extreme short time series. To show that, Table I summarizes results of the maximum Pearson correlation between the signals S as well as results for the non-recurrence points entropy,[15] the permutation entropy,[17] and the entropy of weighted recurrence plots,[16] against the Lyapunov exponent λ, for time series lengths varying from $K = 20$ to $K = 2000$. Table I also depicts values for the standard deviation σ of the Pearson correlation of each entropy versus the Lyapunov exponent. To compute σ, we consider 10 time series of length K and 1000 values of the nonlinear parameter of the logistic map r, in the interval $(0, 1)$. To compute S, we have used $\bar{N} = 10\,000$ random samples for each time series. It is important to mention that the entropy S correlates with the Lyapunov exponent almost independently of the time series length, for time series lengths as low as $K = 20$. For large K length time series, a larger value of sample $\bar{N}$

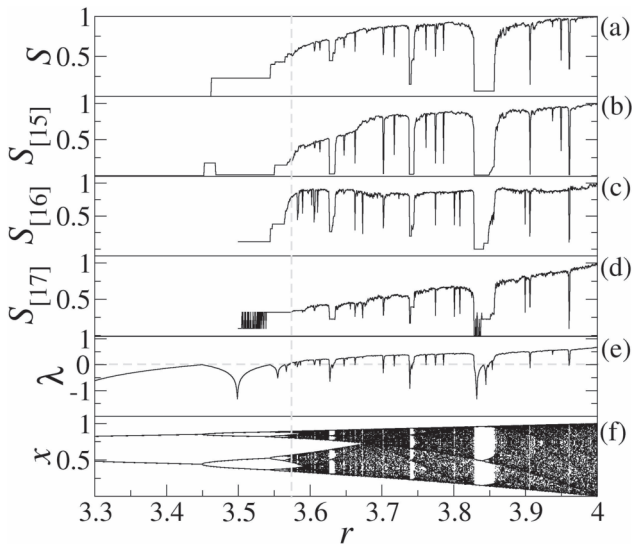

FIG. 4. (a) Normalized entropy quantifier S. [(b)–(d)] Normalized entropy as defined in Ref. 15 (non-recurrence entropy), Ref. 16 (weighted recurrence plot based entropy), and Ref. 17 (permutation entropy). (e) Lyapunov's exponent and (f) bifurcation diagram for the logistic map versus the parameter r. In panel (a), we use a microstate size $N = 4$, $\varepsilon = 0.13$, and time series length $K = 1000$. As observed, both entropy correlates well with the Lyapunov exponent for the entire range of r.

TABLE I. Maximum Pearson's correlation of S, S (Ref. 15), S (Ref. 16), and S (Ref. 17) against the Lyapunov exponent $\lambda(r)$ for different time series lengths, K.

Entropy	$K = 20$	$K = 60$	$K = 100$	$K = 300$	$K = 500$	$K = 1000$	$K = 2000$
S	0.8701	0.8939	0.8973	0.8998	0.9004	0.9006	0.9007
$\sigma(S)$	$\pm$ 0.0026	$\pm$ 0.0012	$\pm$ 0.0009	$\pm$ 0.0003	$\pm$ 0.0002	$\pm$ 0.0001	$\pm$ 0.0001
S (Ref. 15)	0.8208	0.8836	0.8983	0.9086	0.9094	0.9096	0.9098
$\sigma(S)$ (Ref. 15)	$\pm$ 0.0074	$\pm$ 0.0017	$\pm$ 0.0015	$\pm$ 0.0003	$\pm$ 0.0004	$\pm$ 0.0003	$\pm$ 0.0002
S (Ref. 16)	0.7426	0.8266	0.8411	0.8536	0.8558	0.8571	0.8571
$\sigma(S)$ (Ref. 16)	$\pm$ 0.0049	$\pm$ 0.0011	$\pm$ 0.0008	$\pm$ 0.0003	$\pm$ 0.0002	$\pm$ 0.0002	$\pm$ 0.0003
S (Ref. 17)	0.7196	0.8855	0.8939	0.8960	0.8960	0.8970	0.8970
$\sigma(S)$ (Ref. 17)	$\pm$ 0.0101	$\pm$ 0.0012	$\pm$ 0.0002	$\pm$ 0.0003	$\pm$ 0.0003	$\pm$ 0.0003	$\pm$ 0.0003

can be used. In these cases, a better correlation between S and the Lyapunov exponent can be expected. Although the results of Table I show that $\bar{N} = 10\,000$ is enough to acquire a high value for the correlation. Besides, S is much faster than the traditional methods when longer time series are necessary. For example, for $K = 2000$, S is 10 times faster than the entropy defined in Ref. 15. If we consider $K = 16\,000$, the entropy S is $\approx 10^3$ times faster. When compared with the methods proposed in Refs. 16 and 17, S is also faster and turn to be much faster for larger time series since S is dependent only on the number of random samples $\bar{N}$ in contrast with the evaluation of large **RP** matrices or large number of possible permutations.

To explore further the robustness of the methodology, we compute S against the length of the time series. In Figs. 5(a)–5(f), we plot the normalized entropy S of the logistic map, for 6 different time series lengths using 3 values of microstate sizes, $N = 2$ (green lines), $N = 3$ (blue lines), and $N = 4$ (black lines). For each considered size of microstates and for the value of r in the interval $3.3 < \varepsilon < 4.0$, we consider time series of length K and compute its entropy using $\bar{N} = 10\,000$ randomly sampled microstates on the recurrence matrix. We have set $\varepsilon = 0.13$. Error bars of different simulations are not shown since they are smaller than the natural dispersion of the data. Observe that, despite the larger dispersion of computed values obtained for smaller K, all curves are similar. In panel (f), we explicitly plot the Lyapunov exponent (red curve) for visual correlation between $S(r)$ and $\lambda(r)$. We call attention for Figs. 5(e) and 5(f) that show results for time series lengths larger than usually used in recurrence analysis, explicitly showing that the new methodology to compute the entropy S can be applied to longer time series. This particular characteristic of the methodology can be useful in extreme cases, where even a minimal dispersion on computed values has to be avoided. In fact, dispersions observed for smaller values of K, mainly in Figs. 5(a) and 5(b) are exclusively due to the length of the time series. Short time series result on a, relatively, poorer distribution of randomly selected microstates (including the possibility of repeated samples) leading to larger dispersion of computed values of S. The smallest value of the microstate size ($N = 2$) does not capture all features of the dynamics, for example, the gradual increases of the chaoticity level near the onset of chaos, $r \approx 3.57$, and for $r > 3.9$. Results for $N = 3$ (blue lines) and $N = 4$ (black lines) show a clear convergence to the correct behavior. Moreover, the convergence to the correct behavior is very quick when we consider larger microstate sizes, as can be observed for the $N = 3$ and $N = 4$ results.

Another important issue is the quantity of randomly chosen microstate sample ($\bar{N}$) needed to obtain reliable results. Since already mentioned in Sec. II B, the number of microstates that is, effectively, populated by the dynamics of the systems is much smaller than N^*, so the number of samples necessary to obtain a good distribution of microstates can be drastically reduced. To test this point, Figs. 6(a)–6(f) display results for the logistic map considering 6 different sampling sizes and 3 values of microstate sizes, $N = 2$ (green lines), $N = 3$ (blue lines), and $N = 4$ (black lines). Observe that for a number of samples as small as $\bar{N} = 1000$, the new methodology to compute the entropy is already in good agreement with the logistic dynamics, even using the smallest microstate size $N = 2$, for almost the entire interval of r

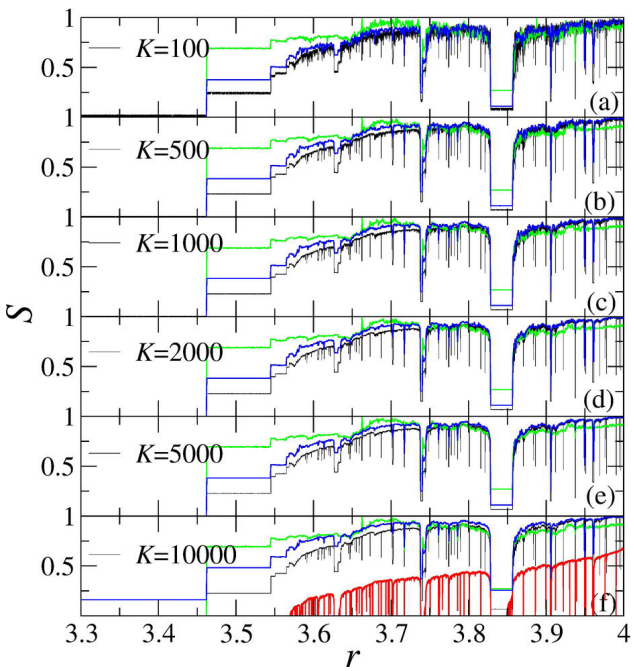

FIG. 5. Normalized entropy quantifier S for the logistic map versus the parameter r. Each panel from (a) to (f) corresponds to a different time series length, K, used to compute S. For all time series lengths, S extract similar information about the system. Notice that, for longer time series, the dispersion of S turn to be small as we should expect. All simulations have used $N = 2$ (green lines), $N = 3$ (blue lines), $N = 4$ (black lines), $\varepsilon = 0.13$, and $\bar{N} = 10\,000$. In panel (f), we also plot the Lyapunov exponent (red thick line). As observed the visual correlation between both quantifiers is expressive.

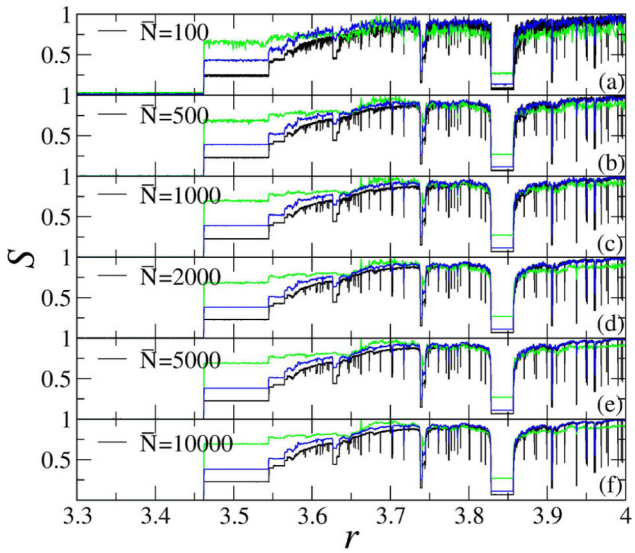

FIG. 6. Normalized entropy quantifier S for the logistic map versus the parameter r. Each panel from (a) to (f) corresponds to a different number of random microstate samples, $\bar{N}$, taken in the recurrence matrix and used to compute S. For $\bar{N} \geq 1000$, S has almost no dependence on the number of random samples taken into the recurrence matrix. All simulations have used $N = 2$ (green lines), $N = 3$ (blue lines), $N = 4$ (black lines), $\varepsilon = 0.13$, and $K = 1000$.

excluding a small interval around the onset of chaos occurring for $r > 3.57$. For smaller time series, $\bar{N} = 100$ [panel (a)] and $\bar{N} = 500$ [panel (b)], for example, the results for the microstate size $N = 2$ start to degrade, mainly near the onset of the chaotic behavior. For larger microstate sizes ($N = 3$ and $N = 4$), all results are in good agreement with the logistic dynamics.

We call attention to other properties of the new methodology to compute entropy of time series. Figure 7(a) depicts the computational effort to obtain S as a function of the number of collected microstate samples in a time series of size $K = 1000$. As expected, the time to compute S scales linearly with the number of sampling $\bar{N}$ for all microstate sizes N, contrarily to the traditional recurrence analysis that scales as N^2. Finally, and perhaps one of the most important results of the new methodology to compute the entropy is depicted in Fig. 7(b), which shows that more precise results, obtained from the analysis of longer time series, can be acquired without any additional computational time, as far as $\bar{N}$ is kept fixed (data computed using $\bar{N} = 10\,000$ microstate samples in this example). This characteristic results directly from the methodology to compute S since the computational time is not sensitive to the size of time series, but only the number of collected microstate samples, contrarily of traditional methods that scale quadratically with the size of the analyzed time series.

Nevertheless, we emphasize that the size of the time series is important. We can say that if the time series size K is large enough, it is expected that it captures all important properties of the phenomenon. So, $\bar{N}$ microstate samples collected, randomly, into the entire time series acquire a good characterization of all these properties. For the case, the time series is not large enough to characterize all properties of the system underlying the data, the same $\bar{N}$ microstate samples still bring useful information about the distribution of microstates. In

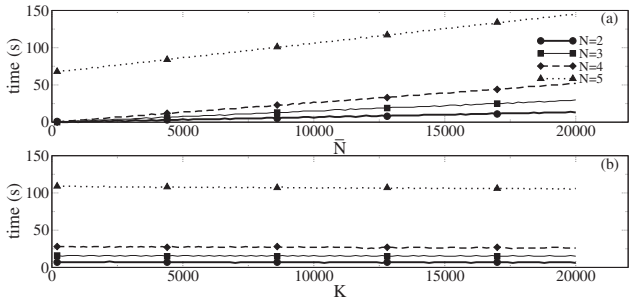

FIG. 7. Computational times as a function of the number of samples, $\bar{N}$, taken into account to compute S, for 4 different microstate sizes and time series of size $K = 1000$ [panel (a)], and as a function of the analyzed time series lengths (K), using $\bar{N} = 10\,000$ samples [panel (b)]. Observe the linear growth of the computational time as a function of the number of samples used to compute S and the invariance of computational time even if longer time series are used.

these cases, due to the short time series length used to collect the ensemble of microstates, the results obtained from the collected microstates can be noisier. A good example of this relation is observed in Fig. 5(a) where an extreme short time series $K = 100$ is analyzed using a large number of microstates $\bar{N} = 10\,000$. For such a short time series, the dispersion in S values is relatively larger, reflecting the bad statistical distribution obtained when a large number of randomly selected microstates is captured in a few among of data. Despite this, the result is still in good agreement with the time series properties. As stated before, the computational enforce to perform the characterization will be the same, independently of the time series since it depends only on the number of collected microstates, but the quality of the results is dependent on the time series size.

C. The Lorenz equations

To illustrate the results of the new methodology to obtain the system entropy applied to a continuous chaotic system, we present results of S using 3 different microstate sizes, $N = 2$, $N = 3$, and $N = 4$ applied to the classical Lorenz system.[21] The Lorenz model is a well studied continuous dynamical system that displays nonlinear behavior and chaos.[21,22] These equations are a reduction from seven to three differential equations, originally developed to model a convection motion in the atmosphere.[21] The three equations that describe the model are written as follows:

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - bz, \end{aligned} \tag{7}$$

with three free parameters: the Rayleigh number ρ, the Prandtl number σ, and the quantity b, a geometric factor. Here, we fix $\sigma = 10$ and $b = 8/3$, varying ρ in the interval $20 < \rho < 240$ for which the system behaves chaotic, displaying also some periodic windows.

An example of the application of our methodology to compute the recurrence entropy S to the Lorenz system is depicted in Fig. 8. In panels (a)–(c), we plot the entropy $S(\varepsilon)$ for 3 microstate sizes, $N = 2, N = 3$, and $N = 4$, respectively.

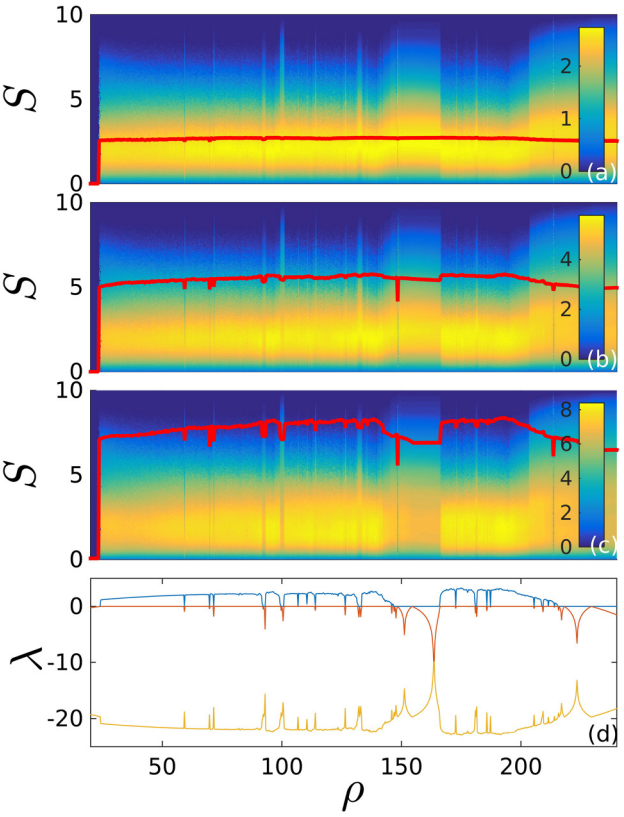

FIG. 8. Entropy and Lyapunov's spectrum for the Lorenz equations considering 3 microstate sizes, $N = 2$ (a), $N = 3$ (b), and $N = 4$ (c). Panels (a)–(c) depict, in color coded representation, the entropy $S(\varepsilon)$ computed for the interval $0.005 < \varepsilon < 1.0$, $K = 1000$, and $\bar{N} = 10\,000$ random chosen microstate samples, superimposed to the color maps of $S(\varepsilon)$, we plot the maximum of S for each value of ρ for the Lorenz equations (red lines). Panel (d) depicts the Lyapunov spectrum for the Lorenz equations. Results show that for $N = 2$, the 2^4 possible microstates are not sufficient to capture all subtle changes occurring in the dynamics as ρ varies [panel (a)]. Panel (b) shows that $N = 3$, or 2^9 possible microstates are marginally able to capture all changes in the dynamics. Panel (c) depicts results for $N = 4$, or 2^{16} possible microstates. In this S captures well the subtle dynamical changes, displaying a good correlation between $\max[S(\varepsilon)]$ and the maximum Lyapunov exponent, depicted in panel (d).

We consider a large interval of the vicinity threshold parameter $0.005 < \varepsilon < 1.0$ (color coded for a step size of 0.005) against ρ for the interval $20 < \rho < 240$. The results were computed using $K = 1000$ and $\bar{N} = 10\,000$ random chosen microstate samples. The continuous x time series for the system was sampled using a discrete time interval of $\tau = 0.2$ time units. y and z dynamics produce similar results. We highlight the maximum value of $S(\varepsilon)$ for each ρ as a red line superimposed to the color coded values of $S(\varepsilon)$ in panels (a)–(c). Panel (d) displays the Lyapunov spectrum for the Lorenz system computed using the time evolution of the linear space.[23] Considering the results for $N = 2$ depicted in Fig. 8(a), it is possible to conclude that for this microstate size, the few 2^4 possible microstates are not able to capture the subtle changes occurring in the dynamics of the Lorenz system as the parameter ρ is varied, resulting in an almost flat curve. Figure 8(b) depicts the same analyses, but using microstates of size $N = 3$. In this case, the bigger and more diverse ensemble of 2^9 microstates allows the detection of almost all small variability in the chaoticity level displayed by the Lorenz system in the considered interval of ρ. Nevertheless, the slow increase of chaoticity in the interval $25 < \rho < 70$ is just marginally captured by this set of microstates. A further increase of the microstate size to $N = 4$ results in 2^{16} possible microstates. In this case, S allows a better characterization of all subtle changes occurring in the dynamics, as observed in Fig. 8(c). For this case, the entropy S correlates qualitatively well with the maximum Lyapunov exponent. The correlations are valid for a large range of ε, as suggests the color coded representation of $S(\varepsilon)$. Important details of the dynamics are revealed by $S(\varepsilon)$, using microstates of size $N = 4$, for example, the smooth growing of the chaoticity level for the interval $24 < \rho < 70$, its almost constant value for $70 < \rho < 195$ (excluding periodic windows), and its decrease occurring after $\rho = 195$, all features also detected by the maximum Lyapunov exponent. All small windows of regular behavior are also captured. For the entire interval, the maximum Lyapunov exponent and $\max[S(\varepsilon)]$ oscillate in the intervals $1.24 < \lambda < 2.46$ and $7.02 < S < 8.33$, respectively, as shown in Figs. 8(c) and 8(d).

The maximum values obtained for S for each microstate sizes, $N = 2$, $N = 3$, and $N = 4$, are 2.74, 5.77, and 8.33, respectively. These values are smaller than the maximum possible values of the entropy $S = 4\ln(2) \approx 2.77$ for $N = 2$, $S = 9\ln(2) \approx 6.24$ for $N = 3$, and $S = 16\ln(2) \approx 11.09$ for $N = 4$, since the maximum entropy should exist only for the condition of a random system and supposing an infinite number of microstate samples. The use of the smallest microstate size $N = 2$ almost reaches the maximum allowed value. This high value of S, compared to the maximum allowed, reflects the fact that in this case the small number of possible microstates is not enough to distinguish among chaotic and random states. For larger values of the microstate sizes, the obtained maxima differ consistently from the values of random data. Such observation, occurring even for the smallest microstate size $N = 2$, puts in evidence the possibility of S to distinguish between random and deterministic (chaotic) signals.

IV. DISCUSSION AND FINAL REMARKS

This work has explored a new tool to study recurrence patterns in time series. The method brings a novel quantifier that analyzes microstates obtained from sampled matrices extracted from the **RP**. In a broad sense, the paper has studied the diversity of the computed microstates in the **RP**.

To quantify the diversity of accessed microstates, we have computed a proper Shannon entropy for the system. To demonstrate the validity of our method, we have applied the new methodology to calculate the time series entropy to a random signal (white noise), a discrete chaotic system (the logistic map), and a continuous system exemplified by the Lorenz model. Moreover, we have tested the methodology for diverse microstate sizes, recurrence vicinity size, ε, microstate sizes, N, length of the analyzed time series, K, as well as the number of random samples taken into the recurrence matrix, $\bar{N}$.

The main advantage of employing the Shannon entropy based on recurrence microstates, as proposed here, is the fact that it is intrinsic to the meaning of an entropic quantifier. The new methodology to compute the recurrence entropy

increases naturally with the chaoticity (concept closely associated to disorder and complexity) of the system, similarly to any other entropies, like the Boltzmann entropy, the Kolmogorov-Smirnov entropy, or the Shannon entropy. In particular, it correlates well with the Lyapunov exponent, a traditional way to compute the chaoticity level. At this point, we should emphasize that the more traditional diagonal entropy ENTR[12] computed as a recurrence quantification does not have this propriety. In fact, this particular point has been considered in the literature.[15–17] When compared with these distinct methods to compute system entropy, the new concept of recurrence entropy described in this paper brings similar and/or better results. We should mention that our methodology results to be faster, more tolerant to small time series segments, as well as allows the evaluation of time series lengths as large as necessary to obtain suitable small dispersions of the computed values. The former characteristic is an important point in critical situations where a precise result and an as low as possible dispersion are required.

The entropy S proved to be useful to distinguish random from non-random signals or even to distinguish a signal displaying different properties along the time. At this point, we would like to mention that such property of S can find applications in many areas such as the analysis of signals from unknown sources since it opens a possibility to distinguish random from non-random signals. The sensitiveness of the recurrence entropy to capture the characteristics of a signal results to be important to unveil correlation of different systems based only on time signals and can find applications in many research areas like climatology or economy as well as complex network analysis.

An important issue is the computational effort to estimate the entropy quantifier as defined here. Using a number of random microstate samplings as low as $\bar{N} = 1000$, we have shown that, in most cases, reliable results can be obtained even when just small segments of data are available. More than that, good results can be obtained for a large range of time series sizes bringing reliable results for time series as short as $K = 100$ data points. On the other hand, the new methodology can extract precise results for the entropy of a signal for the case where longer time series are available, thanks to the fact that longer time series can be analyzed without additional computation time. In this particular point, the new methodology to compute the entropy differs from all other traditional recurrence methods found in the literature where the totality of the **RP** must be analyzed. In such cases, a traditional **RP** analysis of a standard 10^3 points in a time series leads to a recurrence matrix of 10^6 points that should be evaluated to compute a proper quantification of diagonals, verticals, or simple density of recurrences.

In addition, it is well known within recurrence techniques that the value of the vicinity size ε and also the minimal sizes of diagonal, $l_{\min}$, or vertical lines, $v_{\min}$, affect the value of the quantifiers.[10,12] The method proposed in this work shows great robustness against changes of ε and does not need to define minimal sizes of diagonal or vertical lines. The new methodology for the entropy obtains stable results using microstates with sizes as small as $N \geq 2$ and shows to be adequate to diverse discrete and continuous systems.

Finally, it is important to mention that the evaluation of the entropic level of a time series is an important question nowadays, finding applications in many areas of science. In this scenario, some other important methods have been developed to quantify changes in the entropic level of a system.[15–17,24] One of the main uses of these quantifiers is the construction of "2-D state maps" where two quantifiers or momenta of these quantifiers are plotted to distinguish different system states. In this scenario, the development of a new method to evaluate the entropy shows a great potential since it can capture different properties of the systems due to its different fundamentals properties, contributing to a better understanding of some particular phenomena. Considering its simplicity to be computed and its good results when applied to the evaluation of the entropic level of the system, the methodology presented here remains to be applied to numerous other possible systems, as experimental data, intermittent systems, complex systems dynamics phenomenology, transitions chaos-hyperchaos, and many others.

ACKNOWLEDGMENTS

The authors acknowledge the support of Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brazil (Grant No. 302785/2017-5), Coordenação de Aperfeiçoamento de pessoal de Nível Superior, CAPES, through Project Nos. 88881.119252/2016-01 and BEX: 11264/13-6, and Financiadora de Estudos e Projetos (FINEP).

[1]E. Bradley and H. Kantz, Chaos **25**, 097610 (2015).
[2]H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, 2004), Vol. 7.
[3]S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Westview Press, 2014).
[4]J. M. Cushing, R. F. Costantino, B. Dennis, R. Desharnais, and S. M. Henson, *Chaos in Ecology: Experimental Nonlinear Dynamics* (Elsevier, 2002), Vol. 1.
[5]D. A. Hsieh, J. Fin. **46**, 1839 (1991).
[6]H. Herzel, W. Ebeling, and A. O. Schmitt, Phys. Rev. E **50**, 5061 (1994).
[7]E. M. Izhikevich, *Dynamical Systems in Neuroscience* (MIT Press, 2007).
[8]T. L. Prado, S. R. Lopes, C. A. S. Batista, J. Kurths, and R. L. Viana, Phys. Rev. E **90**, 032818 (2014).
[9]H. Poincaré, Acta Math. **13**, A3 (1890).
[10]J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, Europhys. Lett. **4**, 973 (1987).
[11]S. Mallat, *A Wavelet Tour of Signal Processing* (Academic Press, 1999).
[12]N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, Phys. Rep. **438**, 237 (2007).
[13]J. P. Zbilut and C. L. Webber, Phys. Lett. A **171**, 199 (1992).
[14]C. L. Webber and J. P. Zbilut, J. Appl. Physiol. **76**, 965 (1994).
[15]C. Letellier, Phys. Rev. Lett. **96**, 254102 (2006).
[16]D. Eroglu, T. K. D. Peron, N. Marwan, F. A. Rodrigues, L. d. F. Costa, M. Sebek, I. Z. Kiss, and J. Kurths, Phys. Rev. E **90**, 042919 (2014).
[17]C. Bandt and B. Pompe, Phys. Rev. Lett. **88**, 174102 (2002).
[18]N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, and J. Kurths, Phys. Lett. A **373**, 4246 (2009).
[19]D. B. Vasconcelos, S. R. Lopes, R. L. Viana, and J. Kurths, Phys. Rev. E **73**, 056207 (2006).
[20]T. L. Prado, P. P. Galúzio, S. Lopes, and R. L. Viana, Chaos **24**, 013106 (2014).
[21]K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos* (Springer, 1996).
[22]A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics*, 2nd ed. (Springer-Verlag, 1992), Vol. 1.
[23]A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).
[24]L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno, Proc. Natl. Acad. Sci. U.S.A. **105**, 4972 (2008).