Characterizing time series via complexity-entropy curves

Haroldo V. Ribeiro,1,* Max Jauregui,1 Luciano Zunino,2,3 and Ervin K. Lenzi4
1Departamento de Física, Universidade Estadual de Maringá, Maringá, PR 87020-900, Brazil
2Centro de Investigaciones Ópticas (CONICET La Plata - CIC), C.C. 3, 1897 Gnonet, Argentina
3Departamento de Ciencias Básicas, Facultad de Ingeniería, Universidad Nacional de La Plata (UNLP), 1900 La Plata, Argentina
4Departamento de Física, Universidade Estadual de Ponta Grossa, Ponta Grossa, PR 84030-900, Brazil

(Received 11 February 2017; revised manuscript received 6 April 2017; published 5 June 2017)

The search for patterns in time series is a very common task when dealing with complex systems. This is usually accomplished by employing a complexity measure such as entropies and fractal dimensions. However, such measures usually only capture a single aspect of the system dynamics. Here, we propose a family of complexity measures for time series based on a generalization of the complexity-entropy causality plane. By replacing the Shannon entropy by a monoparametric entropy (Tsallis q entropy) and after considering the proper generalization of the statistical complexity (q complexity), we build up a parametric curve (the q-complexity-entropy curve) that is used for characterizing and classifying time series. Based on simple exact results and numerical simulations of stochastic processes, we show that these curves can distinguish among different long-range, short-range, and oscillating correlated behaviors. Also, we verify that simulated chaotic and stochastic time series can be distinguished based on whether these curves are open or closed. We further test this technique in experimental scenarios related to chaotic laser intensity, stock price, sunspot, and geomagnetic dynamics, confirming its usefulness. Finally, we prove that these curves enhance the automatic classification of time series with long-range correlations and interbeat intervals of healthy subjects and patients with heart disease.

DOI: 10.1103/PhysRevE.95.062106

I. INTRODUCTION

The study of complex systems often shares the goal of analyzing empirical time series aiming to extract patterns or laws that rule the system dynamics. In order to perform this task, it is very common to employ a complexity measure such as algorithmic complexity [1], entropies [2], relative entropies [3], fractal dimensions [4], and Lyapunov exponents [5]. Researchers have actually defined several complexity measures, a fact directly related to the difficulty of accurately defining the meaning of complexity. However, the majority of the available measures depends on specific algorithms and tuning parameters, which usually creates great difficulties for research reproducibility.

To overcome this problem, Bandt and Pompe [6] have introduced a complexity measure based on the comparison of neighboring values that can be easily applied to any time series. The application of this technique (the permutation entropy) is widely spread over the scientific community [7–13] mainly because of its simplicity and ability to distinguish among regular, chaotic, and random time series. Regarding the particular issue of distinguishing between chaotic and stochastic processes, Rosso et al. [14] have shown that the permutation entropy alone is not enough for accomplishing this task. They observed, for instance, that the value of the permutation entropy calculated for the logistic map at fully developed chaos is very close to the value obtained for long-range correlated noises. Because of that, Rosso et al. [14] have employed the ideas of Bandt and Pompe together with a diagram proposed by López-Ruiz et al. [15]. This diagram is composed of the values of a relative entropic measure (the statistical complexity) versus the Shannon entropy, both calculated within the framework of Bandt and Pompe. Rosso et al. have named this diagram as complexity-entropy causality plane, and by using it they were able to distinguish between several time series of stochastic and chaotic nature. The causality plane has also proved its usefulness in several applications [9,10,12,16–18] and has been generalized for considering higher dimensional data [19], and different time [20] and spatial resolutions [21].

Here, we propose to extend the causality plane for considering a monoparametric entropy in replacement of the Shannon entropy. In particular, we have considered the Tsallis q entropy [22,23] (that recovers the Shannon entropy for q = 1) together with the proper generalization of the statistical complexity [24] (q complexity). The values of the parameter q in the Tsallis entropy give different weights to the underlying probabilities of the system, accessing different dynamical scales and producing a family of complexity measures that capture some of the different meanings of complexity. Moreover, the Tsallis q entropy has already proved to be useful for enhancing the performance of computational techniques such as in optimization problems [25–27] and image thresholding [28,29] as well as has been previously implemented for characterizing fractal stochastic processes [30].

Thus, for a given time series, we build up a parametric curve composed of the values of the q complexity versus the q entropy, which we will call the q-complexity-entropy curve. Based on simple exact results, we discuss some general properties of these curves, and next we present an exhaustive list of applications based on numerical simulations and empirical data. These applications show that the q-complexity-entropy curve can capture dynamical aspects of time series that are not properly identified only by the point for q = 1, which corresponds to the complexity-entropy causality plane of Rosso et al. The rest of this article is organized as follows. Section II is devoted to reviewing the Bandt and Pompe
approach and the complexity-entropy causality plane of Rosso et al. Also in this section, we present our generalization for considering the Tsallis q entropy as well as some general properties of the q-complexity-entropy curve. Section III presents our numerical experiments with time series from the fractional Brownian motion, harmonic noise, and chaotic maps. In Sec. IV, we discuss some real world applications involving time series from laser dynamics, sunspot numbers, stock prices, human heart rate, and Earth’s magnetic activity. Section V ends this paper with some concluding remarks.

II. GENERALIZED ENTROPY AND COMPLEXITY MEASURES WITHIN THE BANDT AND POMPE FRAMEWORK

We start by reviewing the approach of Bandt and Pompe [6] for extracting the probabilities related to the ordinal dynamics of the elements of a time series. For a given time series \( \{x_i\}_{i=1,...,n} \), we construct \( (n-d+1) \) overlapping partitions of length \( d > 1 \) represented by

\[
s \rightarrow \{x_{s-(d-1)},x_{s-(d-2)},...,x_s\},
\]

where \( s = d, d+1, ..., n \). For each \( s \), we evaluate the permutations \( \pi_j = \{\pi_0,\pi_1,\cdots,\pi_{d-1}\} \) of \( \{0,1,\cdots,d-1\} \) defined by the ordering \( x_{\pi_0} \leq x_{\pi_1} \leq \cdots \leq x_{\pi_{d-1}} \), and we associate to each permutation \( \pi_j \) (with \( j = 1, \cdots, d! \)) the probability

\[
p_j(\pi_j) = \frac{\text{the number of } s \text{ that has type } \pi_j}{n-d+1}.
\]

The components of the probability distribution \( P = \{p_j(\pi_j)\}_{j=1,\cdots,d!} \) represent the odds of finding a segment of length \( d > 1 \) within the time series in a given order. For instance, for the time series \( \{2,4,3,5\} \), we can create tree partitions of size \( d = 2 \): \( (s = 2) \rightarrow \{2,4\} \), \( (s = 3) \rightarrow \{4,3\} \), and \( (s = 4) \rightarrow \{3,5\} \). For each one, we associate the permutations \( \{0,1\} \), \( \{1,0\} \), and \( \{0,1\} \), respectively; consequently, the probability distribution is \( P = \{2/3,1/3\} \). Thus, the probability distribution \( P = \{p_j(\pi_j)\}_{j=1,\cdots,d!} \) provides information about the ordering dynamics for a given time scale defined by the value of \( d \), often called the embedding dimension. In the Bandt and Pompe framework, \( d \) is a parameter whose value must satisfy the condition \( n \gg d! \) in order to obtain reliable statistics for all \( d! \) possible permutations occurring in the time series.

Given the probability distribution \( P = \{p_j(\pi_j)\}_{j=1,\cdots,d!} \), Bandt and Pompe proposed to employ the normalized Shannon entropy

\[
H_1(P) = \frac{S_1(P)}{S_1(U)},
\]

where \( S_1(P) = \sum_{j=1}^{d!} p_j \ln \frac{1}{p_j} \) is the Shannon entropy and \( U = \{1/d!\}_{j=1,\cdots,d!} \) is the uniform distribution [so \( S_1(U) = \ln d! \)], as a natural measure of complexity. By following Bandt and Pompe’s idea together with the diagram of López-Ruiz et al. [15], Rosso et al. [14] have proposed to further calculate a second complexity measure defined by

\[
C_1(P) = \frac{D_1(P,U)H_1(P)}{D^*_1},
\]

where \( D_1(P,U) \) is a relative entropic measure (the Jensen-Shannon divergence) between the empirical distribution \( P = \{p_j(\pi_j)\}_{j=1,\cdots,d!} \) and the uniform distribution \( U = \{1/d!\}_{j=1,\cdots,d!} \). This relative measure can be defined in terms of the symmetrized Kullback-Leibler divergence \( [K_1(P|R) = -\sum p_j \ln p_j/p_i, \text{ with } P \text{ and } R \text{ probability distributions}] \) and is written as

\[
D_1(P,U) = \frac{1}{2}K_1\left(\left|\frac{P + U}{2}\right|\right) + \frac{1}{2}K_1\left(\left|\frac{U - P}{2}\right|\right).
\]

with \( P_{x+U} = \left\{ \frac{p_j(\pi_j)+(1/d!)}{2} \right\}_{j=1,\cdots,d!} \); while

\[
D^*_1(P,U) = \max_p D_1(P,U)
\]

is a normalization constant [obtained by calculating \( D_1(P,U) \) when one component of \( P \) is one and all others are zero].

In spite of the fact that the statistical complexity \( C_1(P) \) is defined by the product of \( D_1(P,U) \) and \( H_1(P) \), \( C_1(P) \) is not a trivial function of \( H_1(P) \) in the sense that, for a given value of \( H_1(P) \), there is a range of possible values for \( C_1(P) \) [24].

Because of that, Rosso et al. [14] proposed a representation space composed of the values of \( C_1(P) \) versus \( H_1(P) \), building up the complexity-entropy causality plane and finding that chaotic and stochastic time series occupy different regions of this diagram.

Despite being successfully applied for studying several systems, the values of \( C_1(P) \) and \( H_1(P) \) are not enough for capturing different scales of the system dynamics as well as different meanings for complexity. Because of that, we propose to replace the normalized Shannon entropy [Eq. (3)] and the statistical complexity [Eq. (4)] by monoparametric generalizations based on the Tsallis q entropy. This entropic form is a generalization of the Shannon entropy and can be defined as [22,23]

\[
S_q(P) = \sum_{j=1}^{d!} p_j \ln_q \frac{1}{p_j},
\]

where \( q \) is a real parameter and \( \ln_q x = \int_1^x t^{-q} \, dt \) is the \( q \) logarithm (\( \ln_1 x = \frac{x^q-1}{q} \) if \( q \neq 1 \) and \( \ln_1 x = \ln x \) for any \( x > 0 \) [23]). We will use the convention \( 0 \ln_q (1/0) = 0 \) whenever \( q > 0 \). It is worth noting that \( S_1 \) is the Shannon entropy.

Once defined the \( q \) entropy, we further consider its normalized version [analogously to the Eq. (3)]

\[
H_q(P) = \frac{S_q(P)}{S_q(U)},
\]

where \( S_q(U) = \ln d! \) is the maximum value of the \( q \) entropy [23]. Furthermore, by following the developments of Martin, Plastino, and Rosso [24], we assume the generalized version of the statistical complexity (4), the \( q \) complexity, to be

\[
C_q(P) = \frac{D_1(P,U)H_q(P)}{D^*_q},
\]

where \( D^*_q \) is a normalization constant [obtained by calculating \( D_1(P,U) \) when one component of \( P \) is one and all others are zero].
where
\[ D_q(P, U) = \frac{1}{2} K_q \left( P \left| P + \frac{U}{2} \right| \right) + \frac{1}{2} K_q \left( U \left| P + \frac{U}{2} \right| \right) \]
\[ = -\frac{1}{2} \sum_{i=1}^{d!} p_i \ln \frac{p_i + 1/d!}{2p_i} - \frac{1}{2} \sum_{i=1}^{d!} \ln \frac{p_i + 1/d!}{2d!/d!} \]  

(10)
is a distance between \( P \) and \( U \). \( K_q(P|R) = -\sum p_i \ln p_i / p_i \) is a generalization of Kullback-Leibler divergence within the Tsallis formalism [24], and
\[ D_q^* = \max_P D_q(P, U) \]
\[ = \frac{2^{2-q}d! - (1 + d!)^{1-q} - d!(1 + 1/d!)^{1-q} - d! + 1}{(1-q)2^{2-q}d!} \]  

(11)
is a normalization constant.

Thus, the quantities \( H_q \) and \( C_q \) are generalizations of the normalized entropy and complexity within the symbolic approach of Bandt and Pompe [6], introduced by Rosso et al. [14], which are included as the particular case \( q = 1 \). Here, we are interested in the parametric representation of the ordered pairs \( (H_q(P), C_q(P)) \) on \( q > 0 \) for a fixed distribution \( P \).

We call this curve the \( q \)-complexity-entropy curve, and we shall see that this representation has superior capabilities of distinguishing time series when compared with the point \( (H_1(P), C_1(P)) \) in the complexity-entropy causality plane.

Before we proceed to the applications, let us enumerate some general properties of the \( q \)-complexity-entropy curves. For a given probability distribution \( P = \{ p_j(\pi_j) \}_j=1,...,d! \), let \( r \) be the number of nonzero components of \( P \) (that is, the number of permutations \( \pi_j \) that actually occur in the time series) and \( \gamma = r/d! \) (essentially the fraction of occurring permutations among all \( d! \) possible). From the definitions of \( H_q \) and \( C_q \), it is not difficult to prove the following statements (see Appendix A):

1. If \( r = 1 \), then \( H_q(P) = 0 \) and \( C_q(P) = 0 \) for any \( q > 0 \).
2. \( H_q(P) \rightarrow \gamma \) and \( C_q(P) \rightarrow \gamma(1 - \gamma) \) as \( q \rightarrow 0^+ \).
3. If \( r > 1 \), then \( H_q(P) \rightarrow 1 \) and \( C_q(P) \rightarrow 1 - \gamma \) as \( q \rightarrow \infty \).

These general properties of \( H_q \) and \( C_q \) have the following consequences for the \( q \)-complexity-entropy curves:

1. The \( q \)-complexity-entropy curve of a time series that only displays one permutation \( \pi_j \) (that is, for \( r = 1 \)) collapses onto the point \((0,0)\).
2. For a time series that has all possible permutations \( \pi_j \) (that is, \( r = d! \) and \( \gamma = 1 \)), the \( q \)-complexity-entropy curve is a loop that starts at the point \((1,0)\) for \( q = 0^+ \) and ends at the same point for \( q \rightarrow \infty \).
3. For a time series that does not display all permutations \( \pi_j \), the \( q \)-complexity-entropy curve starts at the point \((\gamma, \gamma(1 - \gamma))\) for \( q = 0^+ \) and ends at the point \((1,1-\gamma)\) for \( q \rightarrow \infty \). Here, \( 0 < \gamma < 1 \), and the number of occurring permutations \( r \) can be obtained from \( \gamma \) via \( r = (d! - 1)\gamma + 1 \).

We shall see that noisy time series are usually characterized by closed \( q \)-complexity-entropy curves, whereas chaotic time series have open curves (especially for large embedding dimensions). This last feature is related to the existence of forbidden ordinal patterns in the chaotic dynamics that is common in several chaotic maps [31–34], but that can also appear in stochastic processes depending on the time series length [35–37].

III. NUMERICAL APPLICATIONS

In this section, we present several applications of the \( q \)-complexity-entropy curve for numerically generated time series of stochastic and chaotic nature.

A. Fractional Brownian motion

As a first application, we study time series generated from the fractional Brownian motion [4]. The fractional Brownian motion is a stochastic process that has stationary, long-range correlated, and Gaussian increments. It is usually defined in terms of a parameter \( h \) (the so-called Hurst exponent): for \( h < \frac{1}{2} \), the fractional Brownian motion is antipersistent, meaning that positive increments are followed by negative increments (or vice versa) more frequently than by chance; while for \( h > \frac{1}{2} \), it is persistent, meaning that positive increments are followed by positive increments and negative increments are followed by negative increments more frequently than by chance. Also, we have fully persistent motion in the limit of \( h \rightarrow 1 \), whereas the usual Brownian motion is recovered in the limit of \( h \rightarrow \frac{1}{2} \).

In order to calculate the \( q \)-complexity-entropy curves associated with the fractional Brownian motion, we numerically generate time series of length \( 2^{17} \) following the procedure of Hosking [39] for different values of the Hurst exponent \( h \). Figures 1(a) and 1(b) show these curves for the embedding dimension \( d = 3 \) and \( h \) in \((0.2,0.3, \ldots, 0.9)\), while Figs. 1(c) and 1(d) are the same for \( d = 4 \). These plots show the average values (over 100 realizations) of the ordered pairs \( (H_q(P), C_q(P)) \), with \( q \) from \( 10^{-4} \) (assumed to be \( q = 0^+ \)) to 1000 in steps of \( 10^{-4} \). We note that all \( q \)-complexity-entropy curves are closed, indicating that time series of length \( 2^{17} \) of the fractional Brownian motion displays all possible permutations \( \pi_j \) for \( d = 3 \) and 4. The presence of forbidden ordinal patterns in the fractional Brownian motion was studied by Rosso et al. [35,36] and Carpi et al. [37], where they observed that the number of forbidden ordinal patterns decreases with the time series length with a rate that depends on the Hurst exponent \( h \). In particular, Carpi et al. [37] showed that for \( d = 4 \) and very small time series (around 100 steps), the fractional Brownian motion may have a few number of forbidden patterns; however, this number vanishes for series of length larger than 500 terms, which agrees with our findings. Also, for the fractional Brownian motion is possible to obtain the exact expression for the \( q \)-complexity-entropy curves, because Bandt and Shiha [38] have calculated the exact form of the probability distribution \( P = \{ p_j(\pi_j) \}_j=1,...,d! \) for \( d = 3 \) and 4 [all values of \( p_j(\pi_j) \) are provided in Appendix B]. By using these distributions, the expressions (8) and (9) lead to the exact values for the ordered pairs \( (H_q(P), C_q(P)) \).
Fig. 1. Dependence of the entropy $H_q$ and complexity $C_q$ on the parameter $q$ for the fractional Brownian motion. Panels (a) and (b) show the $q$-complexity-entropy curves with $d = 3$ and several values of the Hurst exponent $h$ [in $(0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$, as indicated in the plots]. The values of $q$ are increasing (from $q = 0^+$ to $q = 1000$ in size steps of $10^{-4}$) in the clockwise direction. Panels (c) and (d) show the same, but with embedding dimension $d = 4$. All colored curves represent the average value of $H_q$ and $C_q$ over 100 realizations of a fractional Brownian walker with $2^{17}$ steps. The dashed lines stand for 95% confidence intervals estimated via bootstrapping (over 100 independent realizations). The dashed lines in Fig. 1 show the exact $q$-complexity-entropy curves with $H_q$ and $C_q$ obtained from the simulations and the exact results for the fractional Brownian motion. Panel (a) shows the values of $q = q^*_h$ for which $H_q$ reaches a minimum as a function of the Hurst exponent $h$ and $d = 3$. Each dot corresponds to the average value of $q^*_h$, obtained from 100 realizations of a fractional Brownian walker with $2^{17}$ steps. Error bars stand for 95% confidence intervals estimated via bootstrapping (over 100 independent realizations). Panel (b) shows the values of $q = q^*_c$ for which $C_q$ reaches a maximum as a function of the Hurst exponent $h$ and $d = 3$. Again, the dots are the average values calculated from 100 independent realizations of a fractional Brownian walker with $2^{17}$ steps, and the error bars are 95% confidence intervals. In both plots, the continuous lines are the exact results. Panels (c) and (d) are the analogous of (a) and (b) when considering the embedding dimension $d = 4$. In all cases, we note an excellent agreement between the simulations and the exact results.

The results of Fig. 1 also reveal that the $q$-complexity-entropy curves distinguish the different values of the Hurst exponent $h$. We observe that the larger the value of $h$, the broader the loop formed by the $q$-complexity-entropy curve. We also find that the normalized entropy $H_q$ as a function of $q$ has a minimum value at $q = q^*_h$ and that the complexity $C_q$ as a function of $q$ has a maximum value at $q = q^*_c$, both extreme values of $q$ depend on the Hurst exponent $h$ and also on the embedding dimension $d$. This dependence is shown in Fig. 2 for $d = 3$ and 4, where a good agreement between the exact and the numerical values of these extreme values is observed. We further notice that $q^*_h$ increases with $h$ up to a maximum and then starts to decrease [Figs. 2(a) and 2(c)], whereas $q^*_c$ is a monotonically increasing function of the Hurst exponent $h$ [Figs. 2(b) and 2(d)].

The extreme values of the normalized $q$ entropy and the $q$ complexity-entropy ($H_{q^*_h}$ and $C_{q^*_c}$) represent the largest contrast between $H_q$ (as well as $C_q$) calculated for the system distribution $P = \{ p_j (\pi_j) \}_{j=1,...,d!}$ and the uniform distribution $U = \{1/d! \}_{j=1,...,d!}$. In the context of ecological diversity, the values of $H_{q^*_h}$ were found to enhance contrast among ecological communities when compared with usual diversity indexes [40]. Similarly, the values of $H_{q^*_h}$ and $C_{q^*_c}$ may enhance the differentiation among time series of the fractional Brownian motion with different Hurst exponents. In order to verify this hypothesis, we test the performance of the values $H_{q^*_h}$ and $C_{q^*_c}$ (in comparison with $H_1$ and $C_1$) for classifying time series of the fractional Brownian motion with different Hurst exponents. For this, we generate an ensemble with 100 time series for each value of the Hurst exponent $h$ in $(0.03, 0.05, \ldots, 0.97)$. Next, we train a $k$ nearest neighbors algorithm [41] in a threefold cross-validation strategy, considering these 48 different values of $h$ as possible classes for the algorithm. This machine learning classifier is one of the simplest algorithms for supervised learning [41]; it basically assigns a class (here the value of the Hurst exponent) to an unlabeled point based on the class of the majority of the nearby points.

Figure 3(a) shows the confusion matrices (that is, the fractions of time series with a particular Hurst exponent that are classified with a given Hurst exponent, accuracy) when considering $H_1$ and $C_1$ (usual causality plane, $q = 1$ (left panel)) and the $H_{q^*_h}$ and $C_{q^*_c}$ (optimized causality plane (right panel)) for time series with $n = 2^{10}$ terms. We observe that both matrices have nonzero elements only around the diagonal, indicating that the misclassified Hurst exponents are labeled with values close to actual values. We also note that the “diagonal stripe” of these matrices is narrow for the optimized causality plane (specially for $h > 0.5$), showing that the optimized values enhance the performance of the
we still observe that the accuracy is larger when employing the optimized values $H_q^\ast$ and $C_q^\ast$. In fact, Fig. 3(c) shows that overall accuracy is always enhanced (regardless of $n$) when considering the optimized causality plane in comparison with usual causality plane.

B. Harmonic noise

For another numerical application, we consider times series generated from the harmonic noise [42]. This stochastic process is a generalization of the Ornstein-Uhlenbeck process [43] and can be defined by the following system of Langevin equations [42]:

$$\frac{dy}{dt} = s,$$

$$\frac{ds}{dt} = -\Gamma s - \Omega^2 y + \sqrt{2\varepsilon} \Omega^2 \xi(t),$$

where $\xi(t)$ is a Gaussian noise with zero mean, $(\xi(t)) = 0$ (here $\langle \ldots \rangle$ stands for ensemble average) and uncorrelated, $(\xi(t)\xi(t')) = \delta(t-t')$. That is, a harmonic oscillator driven by a white noise. This noise has an oscillating correlation function given by [42]

$$\langle y(t)y(t+\tau) \rangle = \frac{\varepsilon\Omega^2}{\Gamma} \exp\left(-\frac{\Gamma}{2}\tau\right) \times \left[\cos(\omega\tau) + \frac{\Gamma}{2\omega} \sin(\omega\tau)\right],$$

where

$$\omega = \sqrt{\Omega^2 - (\Gamma/2)^2}$$

is the frequency of oscillation. In practical terms, this noise is a mixture of random and periodic behaviors. Notice that the Ornstein-Uhlenbeck process is recovered in the limit $\Omega \rightarrow \infty$ and $\Gamma \rightarrow \infty$, while the ratio $\Gamma/\Omega$ remains fixed [for this case, $\langle y(t)y(t+\tau) \rangle \sim \exp(-\tau)$].

In order to produce time series from the harmonic noise, we integrate the system of equations (12) by using the Euler method with a step size $dt = 10^{-3}$ [an approach that produces a good agreement between the exact correlation function of Eq. (13) and the numerical results] up to maximum integration time of 1320 and for particular values of parameters $\Gamma$, $\Omega$, and $\varepsilon$. In particular, we first investigate the role of the frequency $\omega$ on the form of the $q$-complexity-entropy curve. To do so, we choose $\varepsilon = 1$, $\Gamma = 0.05$, and several values of $\omega$ ranging from 1 to 60 [$\Omega$ is obtained from Eq. (14)]. For these parameters, the shape of the correlation function is similar to an underdamped simple harmonic motion. Figure 4 shows some $q$-complexity-entropy curves for the embedding dimension $d = 3$. We note that all curves form loops and the broader the loop, the smaller the value of $\omega$. Also, the largest contrasts among the values of $\omega$ occur around the regions of minimum entropy and maximum complexity (as indicated by the insets). The curves are strongly overlapped for very small or very large values of $q$. We further observe that values of the complexity for $q = 1$ (black dots in the first inset) practically do not change with $\omega$.

Figures 4(b) and 4(c) depict the individual behavior of $H_q$ and $C_q$ versus $q$ (now for more values of $\omega$), where the insets
show the form of these curves around their extreme values (that are indicated by small dots). Finally, in Figs. 4(d) and 4(e) we study the dependence of the extreme values $q_H^*$ and $q_C^*$ on the parameter $\omega$. We find that $q_H^*$ monotonically increases with $\omega$ in a relationship that can be approximated by an exponential approach to the value $q_H^* = 0.867$. The shape of $q_C^*$ is more intriguing because it suddenly changes around the value $\omega \approx 16.5$, a behavior that is similar to a phase transition in a bistable system.

We further investigate the shape of the $q$-complexity-entropy curves in a situation that is closer to a pure Ornstein-Uhlenbeck process, that is, a process with a correlation function that decays exponentially. For this, we fix $\omega = 10^{-4}$ and choose different values for $\Gamma$, ranging from close to zero up to $10^4$ [again $\varepsilon = 1$ and $\Omega$ is obtained from Eq. (14)]. The small value of $\omega$ ensures that the oscillation period of the correlation function [Eq. (13)] is much larger than the integration time. Figure 5(a) shows some $q$-complexity-entropy curves for the embedding dimension $d = 3$ and different values of $\Gamma$ ranging from 0.55 to 100. These curves form loops whose broadness decreases as $\Gamma$ increases; in fact, the form of these curves is approaching a limit loop that is similar to the one observed for the fractional Brownian motion with $h = \frac{1}{2}$.

Thus, the $q$-complexity-entropy curve can also distinguish among different time series with short-range correlations. Figures 5(b) and 5(c) show the individual behavior of $H_q$ and $C_q$ versus $q$, where we find that the extreme values $q_H^*$ and $q_C^*$ depend on $\Gamma$, as illustrated in Figs. 5(d) and 5(e). For small values of $\Gamma$, $q_H^*$ logarithmically increases with $\Gamma$ up to $\Gamma \approx 300$, where it saturates around $q_H^* \approx 1.12$. Similarly, $q_C^*$ logarthimically decreases with $\Gamma$ up to $\Gamma \approx 1000$, where it saturates around $q_C^* \approx 2.53$. These limit values for $q_H^*$ and $q_C^*$ are very close to those obtained for the fractional Brownian motion with $h = 0.5$ [a random walk, see Figs. 2(a) and 2(b)].

C. Chaotic maps at fully developed chaos

We now focus on analyzing the shape of the $q$-complexity-entropy curves for time series associated with chaotic processes. To do so, we generate time series by iterating eight chaotic maps: Burgers, cubic, Gingerbreadman, Henon, logistic, Ricker, sine, and Tinkerbell. We set their parameters to ensure a fully chaotic regime and iterate over $2^{17} + 10^5$ steps, dropping the initial $10^5$ steps for avoiding transient behaviors. Also, for the two-dimensional maps (Burgers, Gingerbreadman, Henon, and Tinkerbell), we have considered the squared sum of the two coordinates. The definition of each map and the parameters employed are given in Appendix C. Figure 6 shows the $q$-complexity-entropy curves for each map and for the embedding dimensions between $d = 3$ and 6. Differently from our previous results for noises, these curves do not form loops for all embedding dimensions, showing that there are permutations of $\pi$, that never appear in these time series.

For comparison, we also show in Fig. 6 some $q$-complexity-entropy curves for the fractional Brownian motion and harmonic noise, calculated from time series of the same length used for the chaotic maps. For the fractional Brownian motion,
of length $2^L$. As reported by Rosso et al. [35,36] and Carpi et al. [37] for the fractional Brownian motion, we expect the number of forbidden permutations to vanish with the length of the time series, and loops should appear for longer time series. For instance, we find that the curves for the harmonic noise shown in Fig. 6 become loops for time series 1000 times longer, which does not happen for the chaotic maps. Thus, the shape of the $q$-complexity-entropy curve (closed or open) can be used as an indicative of chaos (open curves) or stochasticity (closed curves). Another characteristic that can distinguish between chaotic and stochastic time series is the existence of a minimum value for the normalized entropy $H_q$. We note that a minimum value exists in all time series from harmonic and fractional noise for $3 \leq d \leq 6$, which does not happen for the chaotic maps ($H_q \rightarrow 1$ monotonically for most $d$ values).

D. Logistic map

Still on chaotic processes, we investigate the logistic map in more detail. This map is a quadratic recurrence equation defined as

$$y_{k+1} = a y_k (1 - y_k), \quad (15)$$

where $a$ is a parameter whose values of interest are in the interval $0 \leq a \leq 4$ (for which $0 \leq y_k \leq 1$). Depending on $a$, this map can exhibit simple periodic behavior (e.g., $a = 3.05$), stable cycles of period $m$ (e.g., $m = 4$ for $a = 3.5$ and $m = 8$ for $a = 3.55$), and chaos (most values of $a > 3.56994567 \ldots$ and $a = 4$).

This map is particularly interesting for our study because we can find the exact expression of the $q$-complexity-entropy curve when $d = 3$ and $a = 4$. Amigó et al. [31–34] have shown that the list $\{y_1, y_{15}, y_{34}\}$ always corresponds to the ordinal pattern $[0,1,2]$ when $0 < y_k < \frac{1}{2}$. Similarly, the ordinal pattern $[0,2,1]$ occurs for $\frac{1}{4} < y_k < \frac{5 - \sqrt{5}}{8}$, $[2,0,1]$ for $\frac{5 - \sqrt{5}}{8} < y_k < \frac{3}{4}$, $[1,0,2]$ for $\frac{3}{4} < y_k < \frac{5 + \sqrt{5}}{8}$, and $[1,2,0]$ for $\frac{5 + \sqrt{5}}{8} < y_k < 1$, and the ordinal pattern $[2,1,0]$ never appears. Combining these results with the fact that the probability distribution of $y_k$ is a beta distribution [45], $\rho(y) = \frac{\Gamma(1 + y) \Gamma(1 - y)}{\Gamma(2)}$, we can find the probability distribution $\mathcal{P}(\pi_j(\pi_{j+1})) = \frac{\Gamma(1 + y_k)}{\Gamma(1 + y_k)} \frac{\Gamma(1 - y_k)}{\Gamma(1 - y_k)}$ by integrating the beta distribution over each one of the previous intervals of $y_k$ (for instance, the probability associated with the pattern $[0,1,2]$ is $\frac{1}{3}$). These integrals yield $\mathcal{P}(\pi_j(\pi_{j+1})) = \frac{1}{3}$, which is the exact form of the $q$-complexity-entropy curve, where an excellent agreement is observed.

We further estimate the $q$-complexity-entropy curve for other values of $a$, as shown in Fig. 7(b) for $d = 4$. In these plots, we choose values of $a$ for which the map oscillates between two values ($a = 3.05$), four values ($a = 3.50$), and for two chaotic regimes: one ($a = 3.593$) close to the onset of chaos and another at fully developed chaos ($a = 4$). We note that these different regimes of the logistic map correspond to different curves. However, the values of $H_1$ and $C_1$ alone are not enough for a complete discrimination; for instance, these loops are observed for all values of $d$ and $h$; however, for the harmonic noise there are some open $q$-complexity-entropy curves (when $\omega = 10^{-3}$ and $\Gamma = 0.55$ or 1.30), indicating that this noise presents forbidden permutations, even for time series.

FIG. 5. Dependence of the entropy $H_q$ and complexity $C_q$ on the parameter $q$ for the harmonic noise: changes with the damping coefficient $\Gamma$. Panel (a) shows $q$-complexity-entropy curves with embedding dimension $d = 3$, $\omega = 10^{-4}$, $\epsilon = 1$, and for some values of the parameter $\Gamma$ (shown in the plot). The values of $q$ are increasing (from $q = 0^+$ to 1000 in size steps of $10^{-3}$) in the clockwise direction. The black dots indicate the points $(H_q, C_q)$ for $q = 0$, cross markers for $q = q_0^0$, and asterisk markers for $q = q_0^c$. Panel (b) shows the dependence of $H_q$ on $q$ for several values of $\Gamma$ (indicated by the color code), where the black dots indicate the values of $q = q_0^0$ that minimize $H_q$. Panel (c) shows the dependence of $C_q$ on $q$ for several values of $\Gamma$ (the same of panel (b)), where the black dots indicate the values of $q = q_0^c$ that maximize $C_q$. Panels (d) and (e) show the dependence of the extreme values of $q$ ($q_0^0$ and $q_0^c$) on the damping parameter $\Gamma$ (the markers are the average values over 100 realizations of the harmonic noise with maximum integration time of 1320 and step size of $10^{-3}$, and the error bars stand for 95% bootstrap confidence intervals). We note that $q_0^0$ logistically increases with $\Gamma$ up to $\Gamma \approx 300$ [$q_0^0 = 0.73 + 0.07 \ln(\Gamma)$, as indicated by the continuous line], where it saturates around $q_0^0 \approx 1.12$ (continuous line). We further notice that $q_0^c$ logistically decreases with $\Gamma$ up to $\Gamma \approx 1000$ [$q_0^c \approx 7.4 - 0.47 \ln(\Gamma)$, as indicated by the continuous line], where it saturates around $q_0^c \approx 2.53$ (continuous line).
values are practically the same for $a = 3.50$ and $3.593$, while the values for $q \sim 0$ are very different for these two regimes.

**IV. EMPIRICAL APPLICATIONS**

Another important test for the $q$-complexity-entropy curve is related to empirical time series. These time series usually have some degree of randomness only associated with the experimental technique employed to study a system, a feature that is well known to hinder the discrimination between experimental chaotic and stochastic signals [20]. Thus, in order to test the $q$-complexity-entropy curve in an experimental scenario, we first consider two empirical time series of the West Texas Intermediate from January 2, 1986 to July 10, 2012 (freely available in Ref. [48]). The results are shown in Figs. 8(a) and 8(b). We observe that the shape of the curve for the laser intensity is similar to those reported for chaotic maps, that is, it forms a loop only for $d = 3$ (such as the Burgers map), while for higher embedding dimensions the curve is open. On the other hand, the curves for price time series form loops with a shape that resembles those of the fractional Brownian motion. We further study a time series of the monthly smoothed sunspot index, whose stochastic or chaotic nature is still debated [49–55]. By analyzing the 13-month smoothed monthly sunspot number from 1749 to 2016 ($n = 3202$, freely available in Ref. [56]), we have built the $q$-complexity-entropy curves shown in Fig. 8(c). We note that the curve is closed for $d = 3$ and open for $d = 4$ and 5, showing a minimum value for $H_q$ for the three values of $d$; moreover, the shapes of the curves are similar to those of the fractional Brownian motion.
obtained from the simulations and the exact results for the logistic
dependence of $H_q$ on $q$. Panel (b) shows the $q$-complexity-entropy curve and the dependence of $H_q$ and $C_q$ on $q$ for $d = 4$ and four values of the parameter $a$: $a = 3.05$ (oscillating behavior between two values), $a = 3.50$ (oscillating behavior among four values), $a = 3.593$ (chaotic behavior), and $a = 4$ (fully developed chaos). We note that the complete differentiation among these regimes of the logistic map is only possible when considering different values of $q$. In particular, we observe that the points $(H_q, C_q)$ for $q = 1$ (indicated by the black dots) are in about the same location for $a = 3.50$ and 3.593.

![Fig. 7](image)

**FIG. 7.** Dependence of the entropy $H_q$ and complexity $C_q$ on the parameter $q$ for the logistic map. Panel (a) shows a comparison between the values $H_q$ and $C_q$ (as well as their dependence on $q$) obtained from the simulations and the exact results for the logistic map with $a = 4$ and $d = 3$. Notice that a practically perfect agreement is found. Panel (b) shows the $q$-complexity-entropy curve and the dependence of $H_q$ and $C_q$ on $q$ for $d = 4$ and four values of the parameter $a$: $a = 3.05$ (oscillating behavior between two values), $a = 3.50$ (oscillating behavior among four values), $a = 3.593$ (chaotic behavior), and $a = 4$ (fully developed chaos). We note that the complete differentiation among these regimes of the logistic map is only possible when considering different values of $q$. In particular, we observe that the points $(H_q, C_q)$ for $q = 1$ (indicated by the black dots) are in about the same location for $a = 3.50$ and 3.593.

the harmonic noise. Thus, our results suggest that the sunspot index can be described by an oscillatory behavior combined with irregularities of stochastic nature. A similar description was proposed by Mininni et al. [53,54], where a Van der Pol oscillator with a noise term was found to reproduce several features of the sunspot index. Figure 8(d) shows the values of $q$ that optimize $H_q$ and $C_q$ ($q^*_H$ and $q^*_C$) for each system. For $d = 4$ and 5, we note that $q^*_H$ is substantially smaller for the laser intensities than the values observed for the two other systems (which are very similar). This agrees with the results observed in Fig. 6, where we verified that $H_q$ does not have a minimum value for all embedding dimensions in the case of chaotic maps (which corresponds to $q^*_H \rightarrow 0$). The price dynamics presents the smallest values of $q^*_C$, followed by the laser intensities and the sunspots index (respectively), indicating that $q^*_H$ and $q^*_C$ are associated to different dynamical scales of these systems.

Next, we test if the $q$-complexity-entropy curve can improve the discrimination of physiological signals of healthy subjects and patients with congestive heart failure. In particular, we investigate time series of the interbeat intervals from 46 healthy subjects (age $= 65.9 \pm 4.0$, $n = 106 235 \pm 10 900$) and 15 patients (age $= 69.7 \pm 6.4$, $n = 109 031 \pm 12 826$) with severe congestive heart failure (NYHA class III). All time series are made freely available by the PhysioNet web page [57,58]. Figure 9(a) shows the average curves of all healthy subjects and patients, where loops are found for both conditions. However, the loop is broader for patients than for healthy subjects, which is compatible with the fact that the Hurst exponents of these time series are usually larger for patients than for healthy subjects [59]. We also verify whether the values of $H_q$ and $C_q$, in comparison with $H_1$ and $C_1$, can enhance the differentiation among time series from healthy subjects and patients in a classification task. To do so, we proceed as in the fractional Brownian motion case, that is, we train a $k$ nearest neighbors algorithm in a threefold cross-validation strategy. Our results show that optimized values provide a greater accuracy when compared with the usual values for $q = 1$ ($\approx 80\%$ against $\approx 76\%$), as shown in Fig. 9.

Finally, as a last application, we consider a time series related to the Earth’s magnetic activity: the disturbance storm time index (DST). This index reflects the average change
in the Earth’s magnetic field based on measurements of the equatorial ring current from a station network located along the equator on the Earth’s surface. The injection of energetic ions from the solar wind into the ring current produces a magnetic field that (at the equator) is opposite to the Earth’s field, often resulting in a sharp decreasing of the DST index and defining a geomagnetic storm [60]. Figure 10(a) shows the evolution of the DST index (hourly resolution) from January 1, 1989 to May 24, 1989 (144 days) based on data freely available by the World Data Center for Geomagnetism [61]. During this period, a great geomagnetic storm occurred (the March 13, 1989 geomagnetic storm [62]), making the DST as lower as −600 nT. In order to verify if the \( q \)-complexity-entropy curve can distinguish between the different regimes present in the DST index, we segment the data of Fig. 10(a) into time series of 18 days (\( n = 432 \)) and calculate the curves for each segment with \( d = 3 \). Figure 10(b) shows that all curves are characterized by loops of different broadness. In particular, the period just after the beginning of the storm is characterized by the broadest loop, whereas the next data segment has the narrowest loop. We further note that, after the storm, the curve width is progressively restored to a shape similar to the one observed before the storm, reflecting the recovering dynamics of a geomagnetic storm [60]. We find that the values of \( q \) that optimize \( H_q(q^*_n) \) are close to 1 [Fig. 10(c), upper panel] and that they are not efficient for identifying the geomagnetic storm. However, the values of \( q \) that optimize \( C_q(q^*_n) \) are very different from 1 [Fig. 10(c), bottom panel] and capable of identifying the geomagnetic storm (notice that \( q^*_n \approx 2.85 \) during the geomagnetic storm, and \( q^*_n < 2.75 \) in all other periods). We also study the \( q \)-complexity-entropy curves for shuffled versions of each time series segment, as shown in Fig. 10(d). After shuffling, the loops are very narrow, and no significant differences among the curves are observed. It is worth noting that the fluctuations in the values of \( C_q \) are not much larger than \( 10^{-3} \). Assuming that the fluctuations in the original time series would have the same magnitude, this result suggests that the difference observed in Fig. 10(b) is statistically significant.

V. SUMMARY AND CONCLUSIONS

We have proposed an extension to the complexity-entropy causality plane of Rosso et al. [14] by considering a monoparametric generalization of the Shannon entropy (Tsallis \( q \) entropy, \( H_q \)) and of the statistical complexity (\( q \) complexity, \( C_q \)). Our approach for characterizing time series is based on the parametric representation of the ordered pairs \((H_q(P), C_q(P))\) on \( q > 0 \), which we have called the \( q \)-complexity-entropy curve. In a series of applications involving numerically generated and empirical time series, we have shown that the \( q \)-complexity-entropy curves can be very useful for characterizing and classifying time series, outperforming the
possible to show that the normalized Rényi\textsuperscript{α} analog of Eq. (8) is a monotonically decreasing function of the support of the CNPq under Grant No. 303642/2014-9. Support of CAPES and CNPq. L.Z. acknowledges Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, for financial support. E.K.L. thanks the financial support of the CNPq under Grant No. 303642/2014-9.

APPENDIX A: LIMITING EXPRESSIONS FOR $H_q$ AND $C_q$ WHEN $q \to 0^+$ AND $q \to \infty$

**Theorem.** Let $P = \{p_j\}_{j=1,...,d^!}$ be a probability distribution, $r$ be the number of nonzero components of $P$, and $\gamma = \frac{r}{d^!}$. The following statements are true:

1. If $r = 1$, then $H_q(P) = 0$ and $C_q(P) = 0$ for any $q > 0$.
2. $H_q(P) \to \gamma$ as $q \to 0^+$;
3. $C_q(P) \to (1 - \gamma) \gamma$ as $q \to 0^+$;
4. If $r > 1$, then $H_q(P) \to 1$ as $q \to \infty$;
5. If $r > 1$, then $C_q(P) \to 1 - \gamma$ as $q \to \infty$.

**Proof.**

1. This follows immediately from the definition of $H_q$ and $C_q$, given in Eqs. (8) and (9).

2. For any $x > 0$, it is clear that $\ln_q x \to x - 1$ as $q \to 0^+$. Using this fact in Eq. (8), we obtain

$$\lim_{q \to 0^+} H_q(P) = \frac{1}{d^! - 1} \sum_{i=1}^{d^!} p_i \left( \frac{1}{p_i} - 1 \right) = \frac{r - 1}{d^! - 1}. \quad (A1)$$

3. It follows immediately from Eq. (11) that $D_q^* \to \frac{d^! - 1}{d^!}$ as $q \to 0^+$. From Eq. (10) we have

$$\lim_{q \to 0^+} D_q(P,U) = -\frac{1}{2} \sum_{i=1}^{d^!} p_i \left( \frac{p_i + 1/d^!}{2 p_i} - 1 \right) - \frac{1}{2} \sum_{i=1}^{d^!} \left( \frac{p_i + 1/d^!}{2/d^!} - 1 \right) = \frac{1}{4} \left( \sum_{i=1}^{d^!} \left( d^! - p_i \right) - \frac{1}{4} \sum_{i=1}^{d^!} \left( p_i - 1/d^! \right) \right) = \frac{1}{4} \left( 1 - \frac{r}{d^!} \right). \quad (A2)$$

Using these results and item (2), we obtain from Eq. (9) that

$$\lim_{q \to 0^+} C_q(P) = \left( \frac{4d^!}{d^! - 1} \right) \left( \frac{d^! - r}{4d^!} \right) \left( \frac{r - 1}{d^! - 1} \right) = \left( 1 - \frac{r}{d^! - 1} \right) \left( \frac{r - 1}{d^! - 1} \right). \quad (A3)$$

4. For $q > 1$, Eq. (8) can be written as

$$H_q(P) = \sum_{i=1}^{d^!} \frac{1}{p_i - p_i^q} \frac{p_i}{1 - (d^!)^{1-q}}. \quad (A4)$$

Then, if $r > 1$, $H_q(P) \to \sum_{i=1}^{d^!} p_i = 1$ as $q \to \infty$.

5. We have from Eq. (11) that

$$D_q^* = \frac{K_q}{(1 - q)^{2^{q-1}}}, \quad (A5)$$

where

$$K_q = \frac{2^{q-1}d^! - (1 + d^!)^{1-q} - d^!(1 + 1/d^!)^{1-q} - d^! + 1}{d^!}. \quad (A6)$$

We note immediately that $K_q \to (1 - d^!)/d^!$ as $q \to \infty$. We have from Eq. (10) that, for $q > 1$,

$$D_q(P,U) = -\frac{2^{1-q}}{K_q} \left[ \sum_{i=1}^{d^!} p_i \left( \frac{1}{2} + \frac{1}{2p_i d^!} \right)^{1-q} \right] + \sum_{i=1}^{d^!} \frac{1}{d^!} \left( p_i d^! + \frac{1}{2} \right)^{1-q} - 2.$$
\[ = - \frac{1}{K_d} \left[ \sum_{i=1}^{d_t} p_i \left( \frac{1}{p_i d_1^{1-q}} \right) \right] + \sum_{i=1}^{d_t} \frac{1}{d_i!} (p_i d_i! + 1)^{1-q} - 2^{-q} \]  
\text{(A7)}

Hence, for \( r > 1 \),
\[ \lim_{q \to \infty} \frac{D_q(P,U)}{D_q^*} = \left( \frac{d_t!}{d_t! - 1} \right) \left( d_t! - r \right) \]
\[ = 1 - \frac{r - 1}{d_t! - 1}. \]  
\text{(A8)}

Therefore, using these results and item (4) on Eq. (9), \( C_q(P) \to 1 - \gamma \) as \( q \to \infty \) whenever \( r > 1 \). \( \blacksquare \)

APPENDIX B: ORDINAL PROBABILITIES FOR THE FRACTIONAL BROWNIAN MOTION

By following the results of Bandt and Shiha [38], we can write the ordinal probabilities (for the embedding dimension \( d = 3 \)) of the fractional Brownian motion with Hurst exponent \( h \) as

\[ p((0,1,2)) = \frac{\alpha}{2}, \]
\[ p((0,2,1)) = \frac{1 - \alpha}{4}, \]
\[ p((1,0,2)) = \frac{1 - \alpha}{4}, \]
\[ p((2,0,1)) = \frac{1 - \alpha}{4}, \]
\[ p((1,2,0)) = \frac{1 - \alpha}{4}, \]
\[ p((2,1,0)) = \frac{\alpha}{2}, \]

where
\[ \alpha = \frac{2}{\pi} \arcsin(2^{h-1}). \]  
\text{(B1)}

Similarly, for the embedding dimension \( d = 4 \), we have [38]
\[ p((0,1,2,3)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_1 + 2 \arcsin \alpha_2 \right), \]
\[ p((0,1,3,2)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_3 - \arcsin \alpha_1 - \arcsin \alpha_2 \right), \]
\[ p((0,2,1,3)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_4 - 2 \arcsin \alpha_3 \right), \]
\[ p((0,2,3,1)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_3 + \arcsin \alpha_8 - \arcsin \alpha_5 \right), \]
\[ p((0,2,3,1)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_3 + \arcsin \alpha_8 - \arcsin \alpha_5 \right), \]
\[ p((0,3,1,2)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_7 - \arcsin \alpha_4 - \arcsin \alpha_5 \right), \]
\[ p((0,3,2,1)) = \frac{1}{8} + \frac{1}{4\pi} \left( \arcsin \alpha_6 - \arcsin \alpha_8 + \arcsin \alpha_2 \right). \]

\[ p((1,0,2,3)) = p((0,1,3,2)), \]
\[ p((1,0,3,2)) = \frac{1}{8} + \frac{1}{4\pi} \left( 2 \arcsin \alpha_6 + \arcsin \alpha_1 \right), \]
\[ p((1,2,0,3)) = p((0,3,1,2)), \]
\[ p((1,2,3,0)) = p((0,3,2,1)), \]
\[ p((1,3,0,2)) = p((0,2,3,1)), \]
\[ p((1,3,2,0)) = p((0,2,3,1)), \]
\[ p((2,0,1,3)) = p((0,2,3,1)), \]
\[ p((2,0,3,1)) = p((0,3,2,1)), \]
\[ p((2,1,0,3)) = p((0,3,2,1)), \]
\[ p((2,1,3,0)) = p((0,3,1,2)), \]
\[ p((2,3,0,1)) = p((1,0,3,2)), \]
\[ p((2,3,1,0)) = p((0,1,3,2)), \]
\[ p((3,0,1,2)) = p((0,3,2,1)), \]
\[ p((3,0,2,1)) = p((0,3,1,2)), \]
\[ p((3,1,0,2)) = p((0,2,3,1)), \]
\[ p((3,1,2,0)) = p((0,2,1,3)), \]
\[ p((3,2,0,1)) = p((0,1,3,2)), \]
\[ p((3,2,1,0)) = p((0,1,2,3)), \]

where
\[ \alpha_1 = \frac{1 + 3^{2h} - 2^{2h+1}}{2}, \]
\[ \alpha_2 = 2^{2h-1} - 1, \]
\[ \alpha_3 = \frac{1 - 3^{2h} - 2^h}{2 \times 6^h}, \]
\[ \alpha_4 = \frac{3^{2h} - 1}{2^{2h+1}}, \]
\[ \alpha_5 = 2^{h-1}, \]
\[ \alpha_6 = \frac{2^{2h} - 3^{2h} - 1}{2 \times 3^h}, \]
\[ \alpha_7 = \frac{3^{2h} - 2^{2h} - 1}{2^{h+1}}, \]
\[ \alpha_8 = \frac{2^{2h} - 1}{3^h}. \]

By using these values, we find the exact form of the distribution \( P = \{p(x_j)\}_{j=1,...,d} \) and the \( q \)-complexity-entropy curve \( (H_q(P),C_q(P)) \).

APPENDIX C: DEFINITION OF THE EIGHT CHAOTIC MAPS EMPLOYED IN OUR STUDY

The Burgers map is defined as
\[ x_{k+1} = ax_k - y_k^2, \]
\[ y_{k+1} = by_k + x_k y_k, \]
and we have chosen \( a = 0.75 \) and \( b = 1.75 \). The time series that we have analyzed is \((x_k + y_k)^2\) with \( x_0 = -0.1 \) and \( y_0 = 0.1 \).
The cubic map is defined as
\[ x_{k+1} = ax_k(1 - x_k^2), \]
and we have chosen \( a = 3 \) and \( x_0 = 0.1 \).

The Gingerbreadman map is defined as
\[ x_{k+1} = 1 - y_k + |x_k|, \]
\[ y_{k+1} = y_k, \]
and we have chosen \( x_0 = 0.5 \) and \( y_0 = 3.7 \). The time series that we have analyzed is \((x_k + y_k)^2\).

The logistic map is defined as
\[ x_{k+1} = ax_k(1 - x_k), \]
and we have chosen \( a = 4 \) and \( x_0 = 0.1 \).

The Hénon map is defined as
\[ x_{k+1} = 1 - ax_k^2 + y_k, \]
\[ y_{k+1} = b x_k, \]
and we have chosen \( a = 1.4 \) and \( b = 0.3 \). The time series that we have analyzed is \((x_k + y_k)^2\) with \( x_0 = 0 \) and \( y_0 = 0.9 \).

The Ricker map is defined as
\[ x_{k+1} = ax_k e^{-y_k}, \]
and we have chosen \( a = 20 \) and \( x_0 = 0.1 \).

The sine map is defined as
\[ x_{k+1} = a \sin(\pi x_k), \]
and we have chosen \( a = 1 \) and \( x_0 = 0.1 \).

The Tinkerbell map is defined as
\[ x_{k+1} = x_k^2 - y_k^2 + ax_k + by_k, \]
\[ y_{k+1} = 2x_k y_k + cx_k + d y_k, \]
and we have chosen \( a = 0.9, b = -0.6, c = 2.0, \) and \( d = 0.5 \). The time series that we have analyzed is \((x_k + y_k)^2\) with \( x_0 = -0.1 \) and \( y_0 = 0.1 \).