Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 369 (2006) 439–462
PHYSICA A
www.elsevier.com/locate/physa

Generalized statistical complexity measures: Geometrical and analytical properties

M.T. Martin[a], A. Plastino[a], O.A. Rosso[b,*]

[a]*Instituto de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata and Argentina's National Research Council (CONICET), C.C. 727, 1900 La Plata, Argentina*
[b]*Chaos & Biology Group, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina*

Received 18 April 2005; received in revised form 8 September 2005
Available online 23 February 2006

Abstract

We discuss bounds on the values adopted by the generalized statistical complexity measures [M.T. Martin et al., Phys. Lett. A 311 (2003) 126; P.W. Lamberti et al., Physica A 334 (2004) 119] introduced by López Ruiz et al. [Phys. Lett. A 209 (1995) 321] and Shiner et al. [Phys. Rev. E 59 (1999) 1459]. Several new theorems are proved and illustrated with reference to the celebrated logistic map.

Keywords: Statistical complexity; Entropy; Disequilibrium; Distances in probability space; Dynamical systems

1. Introduction

Complexity denotes a state of affairs that one easily appreciate when confronted with it. However, it is rather difficult to quantitatively define it in precise fashion. Perhaps the reason is that there is no universal definition of complexity. Different measures for complexity have been proposed in the literature. We can mention, among several ones, (a) Crutchfield and Young's complexity [1], which measures the amount of information about the past required to predict the future; (b) a measure of the self-organization "capacity" of a system [2]; (c) (for chaotic system) the dimension of the concomitant attractor is a measure of its complexity since it yields the number of active variables [2,3]; (d) the algorithmic-, information-, Kolmogorov-, or Chaitin-complexity [4–6] based on the size of the smallest computer program that can produce an observed pattern.

In trying to ascertain what is meant by the concept of statistical complexity (*not to be confused with that of algorithmic complexity*)one should start by excluding processes that are certainly not complex, such as those exhibiting periodic motion. A white-noise random process cannot be assumed to be complex either,

*Corresponding author. Tel./fax: +54 11 4576 3375.
E-mail addresses: mtmartin@venus.unlp.edu.ar (M.T. Martin), plastino@venus.unlp.edu.ar (A. Plastino), oarosso@fibetel.com.ar (O.A. Rosso).

0378-4371/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2005.11.053

notwithstanding its irregular and unpredictable character, since it does not contain any non-trivial structure. Statistical complexity has to do with intricate patterns hidden in the dynamics, emerging from a system which itself is much simpler than its dynamics [7]. Statistical complexity is characterized by the paradoxical situation of a complicated dynamics for simple systems. Of course, if the system itself is already involved enough and contains many different constituent parts, it clearly may support a rather intricate dynamics, but perhaps without the emergence of nitid and typical characteristic patterns [7].

A special generalized family of statistical complexity measures C will be the focus of attention here, whose members pretend to be measures of off-equilibrium "regularity", an "order" that is *not* the one associated with very regular structures (for instance, with crystals), for which the entropy is rather small. Biological life, like EEG signals [8–13], or the Classical Limit of Quantum Mechanics (CLQM) [14,15] are typical examples of the kind of regularities or "organization" one has in mind here, associated with relative large entropic values. In comparison with other definitions of complexity, the present family of statistical complexity is rather easy to compute due to it is evaluated in terms of common statistical mechanics' concepts. In fact, entropy and statistical complexity measures of tonic–clonic epileptic EEG time series support the conjecture that an epileptic focus, in this kind of seizure, triggers a self-organized brain state with both order and maximal complexity [8–13]. In the study of CLQM using the statistical complexity, dynamical features of quantum-classical transition zone can properly be appreciated and characterized [14,15]. Other field of application of present statistical complexity family is the quantification of the performance of Pseudorandom Numbers Generators [16,17].

The definitions of a statistical complexity measure can be partitioned into three categories [18]. They can either (a) grow with increasing disorder (decrease as order increases), (b) be quite small for large amounts of the degree of either order or disorder, with a maximum at some intermediate stage, or (c) grow with increasing order (decrease as disorder increases). The present contribution deals with a C-family of the second type.

Ascertaining the degree of unpredictability and randomness of a system is not automatically tantamount to adequately grasping the correlational structures that may be present, i.e. to be in a position to capture the relationship between the components of the physical system [1,19]. These structures strongly influence, of course, the character of the probability distribution that is able to describe the physics one is interested in. Randomness, on the one hand, and structural correlations on the other one, are not totally independent aspects of *this* physics [18,20–23]. Certainly, the opposite extremes of (i) perfect order and (ii) maximal randomness possess no structure to speak of Refs. [1,24,25]. In between these two special instances a wide range of possible degrees of physical structure exists that should be reflected in the features of the underlying probability distribution $P \equiv \{p_j\}$. One would like that they be adequately captured by some functional $\mathscr{F}_{stat}[P]$ in the same fashion that Shannon's information [26] "captures" randomness. A suitable candidate to this effect has come to be called the statistical complexity [1,18–23] (see the enlightening discussion of Ref. [19]). As mentioned above, $\mathscr{F}_{stat}[P]$ should vanish in the two special extreme instances mentioned above. In this work, we advance a rather general functional form for the statistical complexity measure and study its geometric and analytic properties. The celebrated logistic map is employed as an illustration-device.

2. Statistical complexity measures

An information measure $\mathscr{I}$ can primarily be viewed as a quantity that characterizes a given probability distribution. Shannon's logarithmic information measure [26], $\mathscr{I}[P] \equiv S[P] = -\sum_{j=1}^{N} p_j \ln(p_j)$, is regarded as the measure of the uncertainty associated to the physical processes described by the probability distribution $P = \{p_j, j = 1, \ldots, N\}$.

If $\mathscr{I}[P] = 0$ we are in a position to predict with certainty *which* of the possible outcomes j whose probabilities are given by the p_j will actually take place. Our *knowledge* of the underlying process described by the probability distribution is in this instance maximal. On the other hand, our *ignorance* is maximal for a uniform distribution. Then $\mathscr{I}[P] = \mathscr{I}_{max}$. This two extreme circumstances of (i) maximum foreknowledge ("perfect order") and (ii) maximum ignorance (or maximum "randomness") can, in a sense, be regarded as "trivial" ones.

We define for a given probability distribution P and its associate information measure $\mathscr{I}[P]$ [27], an amount of "disorder" H in the fashion

$$H[P] = \mathscr{I}[P]/\mathscr{I}_{max}, \tag{1}$$

where $\mathscr{I}_{max} = \mathscr{I}[P_e]$ and P_e is the probability distribution which maximizes the information measure, i.e. the uniform distribution. Then $0 \leqslant H \leqslant 1$. It follows that a definition of statistical complexity measure must not be made in terms of just "disorder" or "information". It might seen reasonable to propose a measure of "statistical complexity" by adopting some kind of distance $\mathscr{D}$ to the uniform distribution P_e [20,29,28].

For such a purpose we define the "disequilibrium" Q as

$$Q[P] = Q_0 \mathscr{D}[P, P_e], \tag{2}$$

where Q_0 is a normalization constant and $0 \leqslant Q \leqslant 1$. The disequilibrium Q would reflect on the system's "architecture", being different from zero if there are "privileged", or more likely states among the accessible ones. Consequently, we will adopt the following functional form for the statistical complexity measure:

$$C[P] = Q[P] \cdot H[P]. \tag{3}$$

This quantity reflects on the interplay between the amount of information stored in the system and its disequilibrium [20].

Here we will define $\mathscr{I}$ in terms of entropies (Shannon, Tsallis or Rényi) [30,31]. As for the choice of the distance $\mathscr{D}$ entering Q we face several possibilities, as for instance, (a) Euclidean or (b) Wootters's statistical one [32]. Moreover, since in analyzing complex signals by recourse to information theory tools, entropies, distances, and statistical divergences play a crucial role for prediction, estimation, detection and transmission processes [27,30,33], for the selection of $\mathscr{D}$ we consider also (c) relative entropies [30,31] and (d) Jensen divergences (Shannon, Tsallis, Rényi) [31,34].

2.1. Selection of the information measure $\mathscr{I}$ and generalized disorder

If $P \equiv \{p_1, \ldots, p_N\}$ is a discrete distribution, its associated Shannon entropy [26] reads

$$\mathrm{S}_1^{(S)}[P] = -\sum_{j=1}^{N} p_j \ln(p_j). \tag{4}$$

In 1998, Tsallis proposed a generalization of the celebrated Shannon–Boltzmann–Gibbs entropic measure [35]. The new entropy functional introduced by Tsallis along with its associated generalized thermostatistics is nowadays being hailed as the possible basis of a new theoretical framework appropriate to deal with non-extensive settings [36,37]. This entropy, for a discrete distribution, has the form

$$\mathrm{S}_q^{(T)}[P] = \frac{1}{(q-1)} \sum_{j=1}^{N} [p_j - (p_j)^q], \tag{5}$$

where the entropic index q is any real number.

Finally, the Rényi's entropy [38], for a discrete probability distribution P, is given by

$$\mathrm{S}_q^{(R)}[P] = \frac{1}{(1-q)} \ln\left\{ \sum_{j=1}^{N} (p_j)^q \right\}. \tag{6}$$

In what follows we denote with $\mathrm{S}_q^{(\kappa)}$ any of the above entropies (Shannon, Tsallis or Rényi) according to the κ assignment: S, T or R, so that $\mathscr{I} \equiv \mathrm{S}_q^{(\kappa)}$. Of course, in Shannon's instance one has $q = 1$. The generalized disorder $H_q^{(\kappa)}$ is based on these measures

$$H_q^{(\kappa)}[P] = \mathrm{S}_q^{(\kappa)}[P]/\mathrm{S}_{max}^{(\kappa)}, \tag{7}$$

where $S_{max}^{(\kappa)}$ is the maximum possible value for the information measure one is dealing with. Also,

$$\lim_{q\to 1} H_q^{(T)}[P] = \lim_{q\to 1} H_q^{(R)}[P] = H_1^{(S)}[P]. \tag{8}$$

In the following, and without loss of generality, in order to determine the maximum possible entropic value we just consider the case in which probability-normalization is the only constraint ($\sum_{j=1}^{N} p_j = 1$). The entropy is then maximized by the uniform probability $P_e = \{1/N, \ldots, 1/N\}$, so that $S_{max}^{(\kappa)} = S_q^{(\kappa)}[P_e]$, entailing

$$S_{max}^{(S)} = S_{max}^{(R)} = \ln N, \tag{9}$$

and

$$S_{max}^{(T)} = \frac{1 - N^{1-q}}{q-1} \tag{10}$$

for $q \in (0,1) \cup (1,\infty)$.

2.2. Choosing $\mathscr{D}$ and generalized disequilibrium

2.2.1. Euclidean and Wootters statistical distances

The "natural" election (the most simple one) for the distance $\mathscr{D}$ is the Euclidean one. If $\mathscr{D}$ is the Euclidean norm in $\mathbb{R}^N$, we get

$$\mathscr{D}_E[P, P_e] = \|P - P_e\|_E = \sum_{j=1}^{N} \left\{ p_j - \frac{1}{N} \right\}^2. \tag{11}$$

This straightforward definition of distance has been criticized by Wootters in an illuminating communication [32] because, in using the Euclidean norm, one is ignoring the fact that we are dealing with a space of probability-distributions and thus disregarding the stochastic nature of the distribution $P \equiv \{p_j\}$.

The concept of *statistical distance* originates in a quantum mechanical context. One uses it primarily to distinguish among different preparations of a given quantum state, and, more generally, to ascertain to what an extent two such states differ from one another. The concomitant considerations being of an intrinsic statistical nature, the concept can be applied to *any* probabilistic space [32]. The main idea underlying this notion of distance is that of adequately taking into account statistical fluctuations inherent to any finite sample. As a result of the associated statistical errors, the observed frequencies of occurrence of the various possible outcomes typically differ somewhat from the actual probabilities, with the result that, in a given fixed number of trials, two preparations are indistinguishable if the difference between the actual probabilities is smaller than the size of a typical fluctuation [32].

Given two probability distributions $P_i = \{p_j^{(i)}, j = 1, \ldots, N\}$ with $i = 1, 2$ the Wootters statistical distance is given by [28,32]

$$\mathscr{D}_W[P_1, P_2] = \cos^{-1}\left\{ \sum_{j=1}^{N} (p_j^{(1)})^{1/2} \cdot (p_j^{(2)})^{1/2} \right\}. \tag{12}$$

2.2.2. Relative entropies

Following Basseville [31] two divergence-classes can be built up starting with *functionals of the entropy* (see Appendix A). The first class includes divergences defined as relative entropies, while the second one deals with divergences defined as entropic differences.

Consider now the probability distribution P and the uniform distribution P_e. The distance between these two distributions, $\mathscr{D}_{KS}$, in Kullback–Shannon terms, will be

$$\mathscr{D}_{KS}[P, P_e] = K^{(S)}[P|P_e] = S_1^{(S)}[P_e] - S_1^{(S)}[P]. \tag{13}$$

We shall call $\mathscr{D}_{K_q^T}$ the accompanying distance to the uniform distribution in terms of Kullback–Tsallis entropy

$$\mathscr{D}_{K_q^T}[P, P_e] = \mathrm{K}_q^{(T)}[P|P_e] = N^{q-1}(\mathrm{S}_q^{(T)}[P_e] - \mathrm{S}_q^{(T)}[P]). \tag{14}$$

Similarly, $\mathscr{D}_{K_q^R}$ is the appropriate distance to the uniform distribution defined in terms of Kullback–Rényi's entropy

$$\mathscr{D}_{K_q^R}[P, P_e] = \mathrm{K}_q^{(R)}[P|P_e] = \mathrm{S}_q^{(R)}[P_e] - \mathrm{S}_q^{(R)}[P]. \tag{15}$$

2.2.3. Jensen divergence

In general, the entropic difference $E_D = \mathrm{S}[P_2] - \mathrm{S}[P_1]$ does not define an information gain (or divergence) because E_D is not necessarily positive definite. Something else is needed. An important example is provided by Jensen's divergence [31] (see Appendix A).

We will call $\mathscr{D}_J$ the distance to the uniform distribution associated to Jensen–Shannon' divergence for a special case ($\beta = \frac{1}{2}$, see the general definition in Appendix A),

$$\begin{aligned}\mathscr{D}_J[P, P_e] &= J_{\mathrm{S}_1^S}^{1/2}[P, P_e] \\ &= \mathrm{S}_1^{(S)}\left[\frac{P + P_e}{2}\right] - \frac{1}{2}\mathrm{S}_1^{(S)}[P] - \frac{1}{2}\ln N.\end{aligned} \tag{16}$$

Likewise, $\mathscr{D}_{\mathscr{J}_q^T}$ would be the distance to the uniform distribution associated to the Jensen–Tsallis divergence for $\beta = \frac{1}{2}$,

$$\begin{aligned}\mathscr{D}_{\mathscr{J}_q^T}[P, P_e] &= \mathscr{J}_{\mathrm{S}_q^T}^{1/2}[P, P_e] \\ &= \frac{1}{2}\mathrm{K}_q^{(T)}\left[P\middle|\frac{P + P_e}{2}\right] + \frac{1}{2}\mathrm{K}_q^{(T)}\left[P_e\middle|\frac{P + P_e}{2}\right].\end{aligned} \tag{17}$$

Finally, we have $\mathscr{D}_{\mathscr{J}_q^R}$, the distance to the uniform distribution associated to the Jensen–Rényi divergence for $\beta = \frac{1}{2}$,

$$\begin{aligned}\mathscr{D}_{\mathscr{J}_q^R}[P, P_e] &= \mathscr{J}_{\mathrm{S}_q^R}^{1/2}[P, P_e] \\ &= \frac{1}{2}\mathrm{K}_q^{(R)}\left[P\middle|\frac{P + P_e}{2}\right] + \frac{1}{2}\mathrm{K}_q^{(R)}\left[P_e\middle|\frac{P + P_e}{2}\right].\end{aligned} \tag{18}$$

2.2.4. Generalized disequilibrium

Let $Q_q^{(\nu)}$ be the generalized disequilibrium defined as

$$Q_q^{(\nu)}[P] = Q_0^{(\nu)}\mathscr{D}_\nu[P, P_e], \tag{19}$$

where ν stands for the different measures of distance $\mathscr{D}$ advanced above. Thus, $\nu = E, W, K, K_q, J, \mathscr{J}_q$. $Q_0^{(\nu)}$ is the concomitant normalization constant such that $0 \leqslant Q_q^{(\nu)} \leqslant 1$. We will have:

$$Q_0^{(E)} = \frac{N}{N-1}, \tag{20}$$

$$Q_0^{(W)} = 1/\cos^{-1}\left\{\left(\frac{1}{N}\right)^{1/2}\right\}, \tag{21}$$

$$Q_0^{(K^S)} = Q_0^{(K_q^R)} = \frac{1}{\ln N}, \tag{22}$$

$$Q_0^{(K_q^T)} = \frac{q-1}{N^{(q-1)} - 1}, \tag{23}$$

$$Q_0^{(J)} = -2\left\{\left(\frac{N+1}{N}\right)\ln(N+1) - 2\ln(2N) + \ln N\right\}^{-1}, \tag{24}$$

$$Q_0^{(\mathscr{J}_q^T)} = (1-q) \times\left\{1 - \left[\frac{(1+N^q)(1+N)^{(1-q)} + (N-1)}{2^{(2-q)}N}\right]\right\}^{-1}, \tag{25}$$

$$Q_0^{(\mathscr{J}_q^R)} = 2(1-q) \times\left\{\ln\left(\frac{\left[(N+1)^{(1-q)} + (N-1)\right]}{\left[2^{(1-q)}N\right]}\right) + (1-q)\ln\left(\frac{N+1}{2N}\right)\right\}^{-1}. \tag{26}$$

2.3. Generalized statistical complexity measures

On the basis of the functional product form $C = H \cdot Q$ we obtain then a family of statistical complexity measures $C_{v,q}^{(\kappa)}$ for the distinct disorder H- and disequilibrium Q-measures, namely,

$$C_{v,q}^{(\kappa)}[P] = Q_q^{(v)}[P] \cdot H_q^{(\kappa)}[P] \tag{27}$$

with $\kappa = S, R$ and T for fixed q. In Shannon's instance ($\kappa = S$) we have, of course, $q = 1$. The index $v = E, W, K, K_q, J$, and $\mathscr{J}_q$ tell us that the disequilibrium is to be evaluated with the appropriate distance measures: Euclidean, Wootters, Kullback, q-Kullback, Jensen, and q-Jensen, respectively.

For $v = K, K_q$ Eq. (27) becomes

$$C_{K,q}^{(\kappa)}[P] = (1 - H_q^{(\kappa)}[P]) \cdot H_q^{(\kappa)}[P]. \tag{28}$$

Eq. (28) is the generalized functional form advanced by Shiner, Davison and Landsberg (SDL) [18] for the statistical complexity measure. For $\kappa = S, q = 1$ we get the SDL statistical complexity $C_{SDL} \equiv C_{K,1}^{(S)}$. Note that the SDL complexity measure is calculated from the normalized information measure ($H_q^{(\kappa)}$), or entropy. One could rise the objection that C_{SDL} is just a simple function of the entropy. As a consequence, it might not contain new information vis-a-vis the measure of order. Such an objection is discussed at length in Refs. [23,39,40].

The remaining measures of the family $C_{v,q}^{(\kappa)}$ (with $v \neq K, K_q$) do not reduce, in general, to just functions of the entropy. On the contrary, for a given $H_q^{(\kappa)}$-value an ample range of statistical complexity measures is obtained, from a minimum one C_{min} up to a maximal value C_{max} [20–22,41]. Evaluation of $C_{v,q}^{(\kappa)}$ yields, in general, new information according to the peculiarities of the pertinent probability distribution. It is thus of interest to investigate particular properties of the $C = H \cdot Q$-family. In particular, we will obtain bounds to C_{min} and C_{max}.

In statistical mechanics one is often interested in isolated systems characterized by an initial, arbitrary, and discrete probability distribution and evolution towards equilibrium is to be described. It is the case for a very large class $\mathscr{K}_B$ of (''Boltzmannian'') systems (although NOT for all systems in general—denoted by $\mathscr{C}_B$) that, at equilibrium, the pertinent distribution is the uniform (equiprobability) one. In order to study the time evolution of the statistical complexity measure, a diagram of C versus time t can then be used. But, as we know, the second law of thermodynamics states that entropy grows monotonically with time ($\mathrm{d}H/\mathrm{d}t \geqslant 0$). This implies that H can be regarded as an arrow of time (see, for instance Ref. [42]). For systems $\mathscr{K}_B \subset \mathscr{C}_B$, an equivalent way to study the time evolution of the statistical complexity measure is to plot C versus H. In this way, for such systems, the normalized entropy substitutes for the time axis. If $\kappa = S$ ($q = 1$) and $v = E$ we recover the statistical complexity of López-Ruiz, Mancini and Calbet (LMC) [20], $C_{LMC} \equiv C_{E,1}^{(S)}$. LMC have shown that for an isolated system evolving in time its statistical complexity measure cannot attain arbitrary values in a C_{LMC} versus H map [20,21]. These values are restricted by certain bounds (C-maximum and C-minimum). A recipe for the evaluation of these two curves, for a given value of N, is given in Ref. [21]. The C_{LMC} measure has been exhaustively investigated by Anteneodo and Plastino [22], who, via a Monte Carlo exploration, determined that particular (discrete) probability distribution that maximizes C_{LMC}.

It has been pointed out by Crutchfield and co-workers [19] that the LMC measure is marred by some troublesome characteristics that we list below

- it is neither an intensive nor an extensive quantity,
- it vanishes exponentially in the thermodynamic limit for all one-dimensional, finite-range systems.

The authors of Ref. [19] forcefully argue that a reasonable statistical complexity measure should

- be able to distinguish among different degrees of periodicity and
- vanish only for periodicity unity.

Finally, and with reference to the ability of the LMC measure to adequately capture essential dynamical aspects, some difficulties have also been encountered in Ref. [22]. With the product functional form for the generalized statistical complexity it is impossible to overcome the second deficiency [1] mentioned above. In previous works [29,28] we have shown that, after performing some suitable changes, one is in a position to obtain a generalized statistical complexity measure that is (i) able to grasp essential details of the dynamics, (ii) an intensive quantity, and (iii) capable of discerning among different degrees of periodicity. For a short review see Section 5 below.

3. Analytical considerations concerning the $C^{(\kappa)}_{v,q}$-family

It is our purpose here that of determining the extremal values of the generalized statistical complexity measure $C^{(\kappa)}_{v,q}$ for $\Omega = \{P : \sum_{j=1}^{N} p_j = 1, p_j \geqslant 0\}$. Our procedure is independent of which is the particular measure under consideration, both for the disorder $H^{(\kappa)}_q$ and for the disequilibrium $Q^{(v)}_q$. We consider first a systematic procedure for obtaining $C^{(\kappa)}_{v,q}$-bounds in the case of constant $H^{(\kappa)}_q$. For the sake of a lighter notation we omit herefrom the indices κ, v, q save when they become unavoidable. $C[P] = Q[P] \cdot H[P]$ will now represent the generalized statistical complexity measure given by Eq. (27).

We generalize now the Anteneodo–Plastino formalism of Ref. [22] and show that, for $H^{(\kappa)}_q$ constant, the configurations $P = \{p_1, \ldots, p_N\}$ that optimize $C^{(\kappa)}_{v,q}$ admit just *two* distinct non-vanishing probabilities p_j [41]. Our considerations refer to the *statistical complexity measure* versus *normalized entropy* plane ($H \times C$).

Note that both $H \equiv H^{(\kappa)}_q$ and $Q \equiv Q^{(v)}_q$ (see Eq. (27)) are functions on the probability space Ω of the form

$$Q[P] = \phi\left(\sum_{j=1}^{N} f(p_j)\right), \quad H[P] = \psi\left(\sum_{j=1}^{N} g(p_j)\right) \tag{29}$$

for some particular functions ϕ, ψ, f and g. $\phi(u)$ and $\psi(u)$ are continuous functions of a strictly monotonous character. The second derivatives of $f(u)$ and $g(u)$ satisfy

$$f''(u) = A_1 u^{\alpha_1}, \quad g''(u) = A_2 u^{\alpha_2} \tag{30}$$

with $A_i, \alpha_i \in \mathbb{R}$ for $i = 1, 2$. Then, according with the Theorem 1 (see Appendix C) the probability distribution P that extremizes C for H constant, is of the form

$$P = \begin{cases} 0 & \text{for } 1 \leqslant j \leqslant m, \\ p & \text{for } m+1 \leqslant j \leqslant m+n, \\ (1-pn)/(N-m-n) & \text{for } m+n+1 \leqslant j \leqslant N \end{cases} \tag{31}$$

with $n, m \in \mathbb{N}$, $0 \leqslant m \leqslant N-1$ and $0 \leqslant n \leqslant N-m-1$.

In the particular instance $C = C^{(S)}_{E,q=1} \equiv C_{LMC}$, Anteneodo–Plastino [22] gave analytical arguments to the effect that the extremum of C_{LMC} is obtained for configurations P of the form of Eq. (31) and, afterwards, using numerical procedures, were able to ascertain that the global maximum corresponds to $n = 1$ in Eq. (31). *Our* interest in the following section is to investigate whether such a result $n = 1$ is of general validity for $C^{(\kappa)}_{v,q}$.

4. Geometrical considerations concerning the $C^{(\kappa)}_{v,q}$-family

Note that Ω is the $(N-1)$-simplex spanned by $\{e_1, e_2, \ldots, e_N\}$ (see Appendix B for geometrical definitions and additional considerations):

$$\Omega \equiv \Delta^{N_1} = [e_1, \ldots, e_N], \tag{32}$$

where $N_1 = N-1$ and $[e_1, \ldots, e_N]$ is the set of convex combinations of the N geometrically independent vectors $\{e_i\} \in \mathbb{R}^N$.

The statistical complexity measure C has identical values on each of the $N!$ barycentric subdivisions $\Delta^{N_1}_{\varphi}$ as proved in Theorem 2 (see Appendix B and C). One can thus study $C[P]$ by restricting attention just to the C-values on one of the $N!$ simplices arising out of the barycentric subdivision. In particular, we focus attention on the simplex corresponding to the *identity* permutation $\Delta^{N_1}_I = [b^1_I, \ldots, b^N_I]$.

In Fig. 1 we depict the 3-simplex Ω for $N = 4$. Ω belongs to an hyperplane of dimension 3 determined by the normalization restriction $\sum_{j=1}^{4} p_j = 1$ (Fig. 1a). The barycenter of Ω, $\mu_3 = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ is also shown. It is associated to the uniform probability distribution P_e. Fig. 1b displays Δ^3_I, the sub-simplex corresponding to the identity permutation I, with its $N = 4$ vertices b^j_I and its sides $L_{j,k}$. Of course, some of the sides join consecutive vertices. Whether this *is or is not* the case plays an important role in characterizing extrema.

For each $P \in \Omega$ we represent the pair $(H[P], C[P])$ in the plane $(H \times C)$, calling, respectively, C_{max} and C_{min} the curves determined by the extremal values of C a as a function of H. We are looking here for a procedure that will allow one to calculate these curves for any member of the family. The subindex I will herefrom be omitted for the sake of a lighter notation in speaking about the vertices of the simplex $\Delta^{N_1}_I = [b^1_I, \ldots, b^N_I]$.

From Theorem 3 (see Appendix C) it follows that vectors P of the form given by Eq. (31) are the sides of $\Delta^{N_1}_I$ and, in consequence, the C-extrema for H constant are reached just on the sides of $\Delta^{N_1}_I$. The *side* $L_{1,N}$ connects non-consecutive vertices (case $N = 2$ excluded) b^1 and b^N. b^1 is the barycenter of the entire simplex, associated to complete randomness, $H(b^1) = 1$. b^N is the barycenter of just a point, associated to a state of *certainty* $H(b^N) = 0$. Thus, $L_{1,N}$ joins two vertices associated to extremal entropic values, a fact that we interpret as suggesting the existence of an "order-relation" among the vertices. To probe into such a conjecture we consider entropic values H restricted to the side $L_{j,k}$. Keeping Eq. (59) in mind and assuming $j<k$, our values can be expressed as a function of ρ, in the fashion $H[\mathscr{L}_{j,k}(\rho)]$, for $0 \leqslant \rho \leqslant 1/(N-j+1)$, and then Theorems 4 and 5 do hold.

Note that $\mathscr{B}^{(1)} \equiv L_{1,N}$ represents the *side* joining extremal vertices while $\mathscr{B}^{(N_1)} \equiv \bigcup_{i=1}^{N_1} L_{i,i+1}$ is the *union of all sides* that connect consecutive vertices. Figs. 2a and b depict, for $N = 6$, the statistical complexity measures $C^{(S)}_E[P]$ and $C^{(S)}_W[P]$, respectively, as a function of $H^{(S)}[P]$, for P varying either on $\mathscr{B}^{(1)}$ or on $\mathscr{B}^{(N_1)}$. Eq. (60) tells

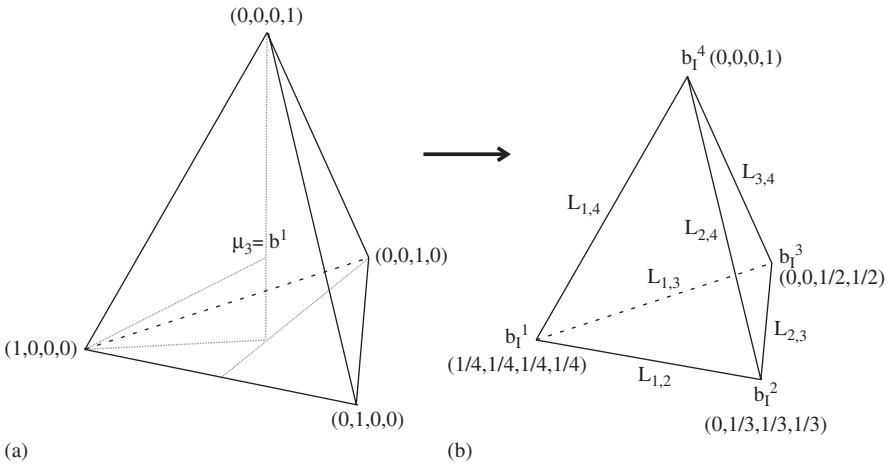

Fig. 1. Probability subspace Ω for $N = 4$: (a) $\Omega \equiv \Delta^2$ in an hyperplane of dimension 3. Dotted lines effect the barycentric subdivision with μ_3 the Ω-barycenter. (b) sub-simplex Δ^3_I.

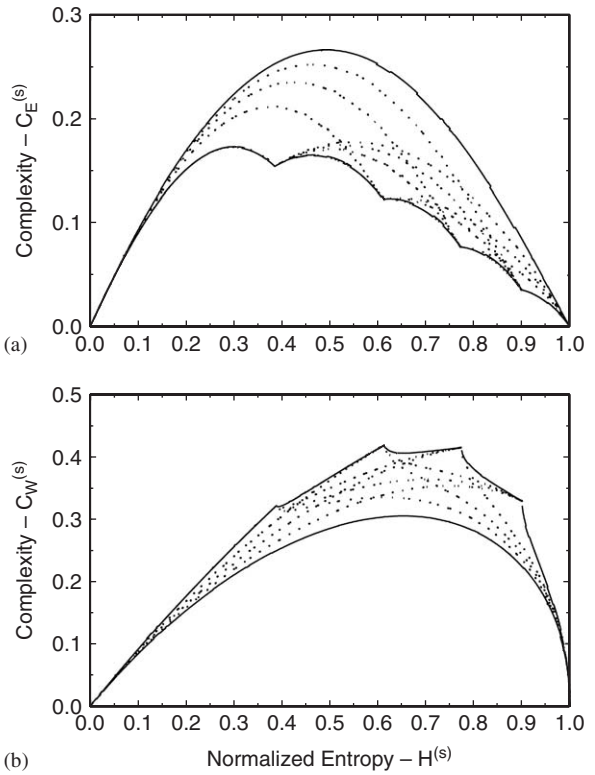

Fig. 2. Complexity as a function of $H^{(S)}[P]$ for P varying over $\mathscr{B}^{(1)}$, or over $\mathscr{B}^{(N_1)}$, for $N = 6$. (a) $C_E^{(S)}[P]$ and (b) $C_W^{(S)}[P]$.

us that $P \in \mathscr{B}^{(1)}$ fulfills Eq. (50) with $n = 1$. It is located in the interior of Ω ($m = 0$) and has just one coordinate of value p while the remaining ones adopt the value $(1-p)/(N-1)$. Now, if $P \in \mathscr{B}^{(N_1)}$ then P is on a side $L_{i,i+1}$ and, as a consequence, exhibits $m = i-1$ null coordinates, $n = 1$ coordinates of value p and $N-m-1$ ones whose value is $(1-p)/(N-1)$. In particular, if P belongs to $L_{1,2}$, it belongs to Ω's interior. Dotted lines in Fig. 2 represent the image of the sides joining non-consecutive vertices, who lie within the region delimited by the curves C_{max} and C_{min}.

Figs. 3a and b display values for $C_E^{(S)}[P]$ and $C_W^{(S)}[P]$, respectively, as a function of $H^{(S)}[P]$. The figures depicted are obtained via numerical simulation for $N = 6$ with 30 000 random probability distributions. Comparison with Figs. 2a and b, respectively, permits one to conclude that extremal values of the statistical complexity measure are in correspondence with values on $\mathscr{B}^{(1)}$ (or on $\mathscr{B}^{(N_1)}$). In particular, for $C_E^{(S)}[P]$ (Fig. 3a), the maximal statistical complexity measure C_{max} lies on $\mathscr{B}^{(1)}$, revalidating (cf. Eq. (50) for $n = 1$) the numerical results of Anteneodo–Plastino. We see in Fig. 2b, for $C_W^{(S)}[P]$), that C_{max} lies on $\mathscr{B}^{(N_1)}$ and C_{min} on $\mathscr{B}^{(1)}$.

It is also true that generalized statistical complexity measures $C_{v,q}^{(\kappa)}$ also adopt their extremal values in the sets $\mathscr{B}^{(1)}$ and $\mathscr{B}^{(N_1)}$. Using the notation of Theorem 5 we formulate now the following *regularity hypothesis*:

C satisfies *one* of the two conditions (see Fig. 4)

(I) For any face $[b^i, b^j, b^k]$ with $i<j<k$,
we have $C[P_h^{(ik)}] \geqslant C[P_h^{(ijk)}]$, $\forall h \in [H[b^k], H[b^i]]$.
(II) Given an arbitrary face $[b^i, b^j, b^k]$ with $i<j<k$,
one has $C[P_h^{(ik)}] \leqslant C[P_h^{(ijk)}]$ $\forall h \in [H[b^k], H[b^i]]$.

We introduce the notation, for each $H^* \in [0,1]$

$$P_h^{max} = \mathrm{Max}\{C[P], P \in \mathscr{E}_h\} \tag{33}$$

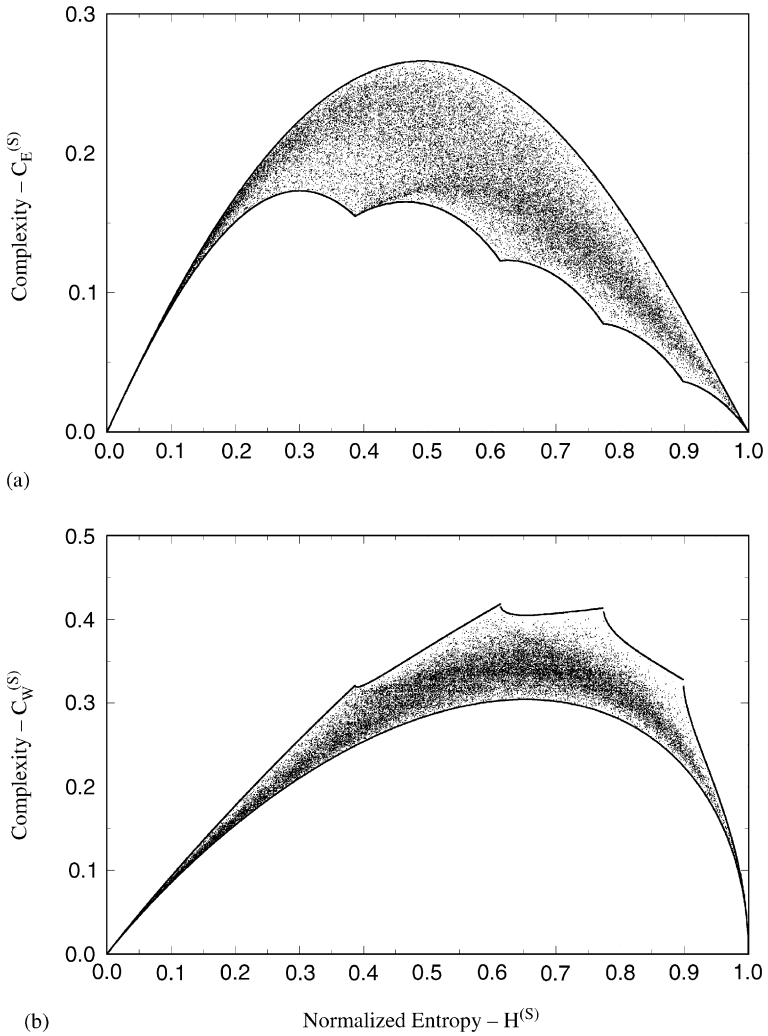

Fig. 3. Numerical simulation of the complexity as a function of $H^{(S)}[P]$, for P varying over Ω, for $N = 6$. (a) $C_E^{(S)}[P]$ and (b) $C_W^{(S)}[P]$.

and

$$P_h^{min} = \mathrm{Min}\{C[P], P \in \mathscr{E}_h\}. \tag{34}$$

Its follows that if C fulfills the regularity hypothesis (I) then $P_h^{max} = P_h^{(1)}$ and $P_h^{min} = P_h^{(N_1)}$. If instead, C fulfills hypothesis (II), then $P_h^{min} = P_h^{(1)}$ and $P_h^{max} = P_h^{(N_1)}$ (see Theorems 6, Appendix C).

Let

- $\mathscr{G}^{(1)}$ the map
  $\mathscr{G}^{(1)} : [0, 1] \to [0, 1]$, such that $\mathscr{G}^{(1)}(\mathrm{H}^*) = C[P_h^{(1)}]$,
- $\mathscr{G}^{(N_1)}$ the map
  $\mathscr{G}^{(N_1)} : [0, 1] \to [0, 1]$, such that $\mathscr{G}^{(N_1)}(\mathrm{H}^*) = C[P_h^{(N_1)}]$

with $P_h^{(1)}$ and $P_h^{(N_1)}$ defined above.

If C_{max} and C_{min} are upper and lower bounds for the statistical complexity measure in the plane $(H \times C)$ (see Theorem 7, Appendix C) then

- If C satisfies condition (I), $C_{max} \equiv \mathscr{G}^{(1)}$ and $C_{min} \equiv \mathscr{G}^{(N_1)}$.
- If C satisfies condition (II), $C_{max} \equiv \mathscr{G}^{(N_1)}$ and $C_{min} \equiv \mathscr{G}^{(1)}$.

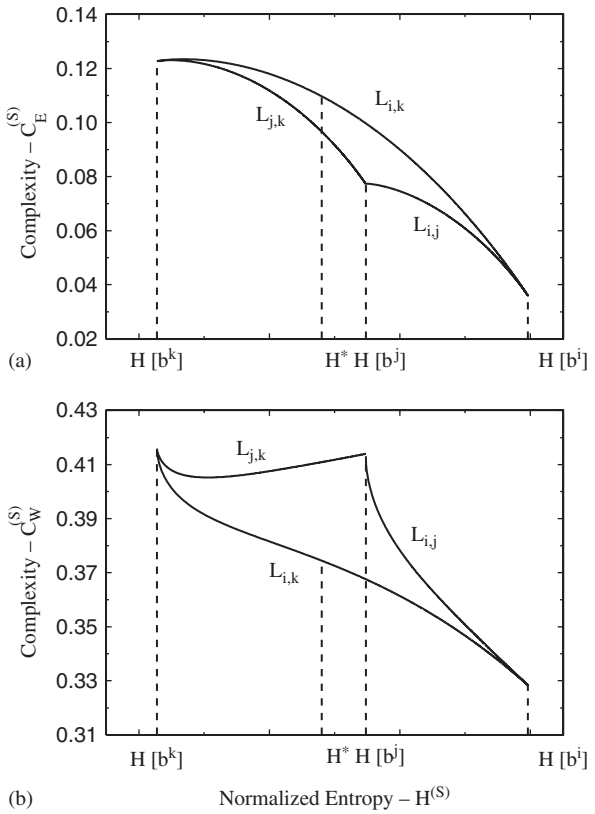

Fig. 4. Bi-dimensional face: plotting the complexity as a function of $H^{(S)}[P]$, for P varying over L_{ik} or over $L_{ij} \cup L_{jk}$. (a)$C_E^{(S)}[P]$ and (b)$C_W^{(S)}[P]$.

Note that, (a) $\forall \mathrm{H}^* \in [0,1]$ one has $P_h^{(1)} \in L_{1,N}$, the curve $\mathscr{G}^{(1)}$ arises after plotting the pairs $(H[P], C[P])$ for $P \in L_{1,N}$. Also, if (b) $P_h^{(N_1)} \in \bigcup_{i=1}^{N_1} L_{i,i+1}$, then the curve $\mathscr{G}^{(N_1)}$ arises after plotting $(H[P], C[P])$ for P traversing $L_{i,i+1}$, with $i = 1, \ldots, N_1$. This entails that $\mathscr{G}^{(N_1)}$ becomes defined piecewise. Each piece is associated to a side $L_{i,i+1}$ joining the consecutive vertices $b^{(i)}$ and $b^{(i+1)}$. It is easy to see from Eq. (60) that configurations corresponding to sides that join consecutive vertices have the form given by Eq. (50) with $n = 1$. Thus, $\mathscr{G}^{(N_1)}$ arises in configurations given by Eq. (50) with $n = 1$.

4.1. Evaluating C_{max} and C_{min}

One can assume, without loss of generality, that C obeys condition (I) and then $C_{max} \equiv \mathscr{G}^{(1)}$ and $C_{min} \equiv \mathscr{G}^{(N_1)}$. Since $\mathscr{G}^{(1)}(\mathrm{H}^*) = C[P_h^1]$ and $P_h^1 \in L_{1,N}$, in order to evaluate C_{max} we need that P varies over $L_{1,N}$. The pertinent curve is obtained plotting $(H[P], C[P])$ versus $P \in L_{1,N}$. The P-configurations are determined by the function $\mathscr{L}_{1,N} : [0, 1/N] \to L_{1,N}$

$$\mathscr{L}_{1,N}(\rho) = \rho \sum_{i=1}^{N-1} e_i + (1 - \rho(N-1))e_N. \tag{35}$$

P ranges from certainty to maximum ignorance via a set of configurations $\in \Omega$ $(m = 0)$. We deal here with a single probability $p_i = p$ $(n = 1)$, the remaining ones being equal to $(1-p)/(N-1)$.

Let us tackle now $C_{min} \equiv \mathscr{G}^{(N_1)}$. Since $\mathscr{G}^{(N_1)}(\mathrm{H}^*) = C[P_h^{N_1}]$ and $P_h^{N_1} \in \bigcup_{i=1}^{N_1} L_{i,i+1}$ one evaluates C_{min} by suitably varying P over $\bigcup_{i=1}^{N_1} L_{i,i+1}$ and plotting the curve $(H[P], C[P])$ with $P \in \bigcup_{i=1}^{N_1} L_{i,i+1}$. The

P-configurations are determined by the function $\mathscr{L}_{i,i+1} : [0, 1/(N-i+1)] \to L_{i,i+1}$

$$\mathscr{L}_{i,i+1}(\rho) = \rho e_i + \left(\frac{1-\rho}{N-i}\right) \sum_{l=i+1}^{N} e_l \tag{36}$$

with $i = 1, \ldots, N-1$. P ranges from certainty to maximum ignorance via a set of configurations belonging to consecutive sides. Minimal complexity is associated to configurations $P \in L_{1,2}$ within Ω ($m = 0$) or on its boundary. These configurations are of the following nature: if they belong to the side $L_{i,i+1}$, they possess $m = i-1$ vanishing coordinates. The non-vanishing ones are of the form $p_i = p$ ($n = 1$) for just one of them, the rest being equal to $(1-p)/(N-i)$.

5. Application

Information theory measures and probability spaces Ω are inextricably linked. In the evaluation of the statistical complexity, the determination of the probability distribution P associated to the dynamical system or time series under study is the basic element. Many procedures have been proposed for the election of $P \in \Omega$. We can mention procedures based on amplitude statistics (histograms) [41], symbolic dynamics [43], Fourier analysis [44], Wavelet transform [45], structure of the attractor reconstruction (permutation probability) [46], among others. The applicability of them depends on data-characteristics like, stationarity, length of the series, parameters variation, level of noise contamination, etc. In all these cases the global aspects of the dynamics might be "captured" but the different instances are not equivalent in their ability to discern relevant physical details.

The logistic map constitutes a paradigmatic example, often employed in order to illustrate new concepts in the treatment of dynamical systems. In such a vein we discuss here the application of the generalized statistical complexity measures (see Eq. (27)).

We deal with the map $F : x_n \to x_{n+1}$ [2,3], described by the ecologically motivated, dissipative system described by the first-order difference equation

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n) \tag{37}$$

with $0 \leqslant x_n \leqslant 1$ and $0 < r \leqslant 4$. Fig. 5a shows the well-known bifurcation diagram for the logistic map for $3.5 \leqslant r \leqslant 4.0$ while, in Fig. 5b, the corresponding Lyapunov exponent, Λ, is depicted.

Let us briefly review, with reference to Figs. 5a and b, some exceedingly well-known results for this map, that we need in order to put into an appropriate perspective the properties of our family of generalized statistical complexity measures. For values of the control parameter $1 < r < 3$ there exists only a single steady-state solution. Increasing the control parameter past $r = 3$ forces the system to undergo a period-doubling bifurcation. Cycles of period $8, 16, 32, \ldots$ occur and, if r_n denotes the value of r where a 2^n cycle first appears, the r_n converge to a limiting value $r_\infty \cong 3.57$ [3]. As r grows still more, a quite rich, and well-known structure arises. In order to be in a position to better appreciate at once the long-term behavior for all values of r lying between 3.5 and 4.0, we plot the pertinent orbit-diagram in Fig. 1a. We immediately note there the cascade of further period-doubling that occurs as r increases, until, at r_∞, the maps become chaotic and the attractors change from comprising a finite set of points to becoming an infinite set. For $r > r_\infty$ the orbit-diagram reveals an "strange" mixture of order and chaos. The large window beginning near $r = 3.83$ contains a stable period-3 cycle. In Fig. 5b we see that the non-zero Lyapunov characteristic exponent Λ remains negative for $r_\infty \cong 3.57$. We notice that Λ approaches zero at the period-doubling bifurcation. The onset of chaos is apparent near $r \cong 3.57$, where Λ first becomes positive. As stated above, for $r > 3.57$ the Lyapunov exponent increases globally, except for the dips one sees in the windows of periodic behavior. Note the large dip due to the period-3 window near $r = 3.83$.

Let us revisit now the binary treatment (symbolic dynamics) of the logistic map [20,22,28,29]. Following Ref. [20], for each parameter value, r, the dynamics of the logistic map was reduced to a binary sequence (0 if $x \leqslant \frac{1}{2}$; 1 if $x > \frac{1}{2}$) and binary strings of length 12 were considered as states of the system. The concomitant probabilities are assigned according to the frequency of occurrence after running over at least 2^{22} iterations.

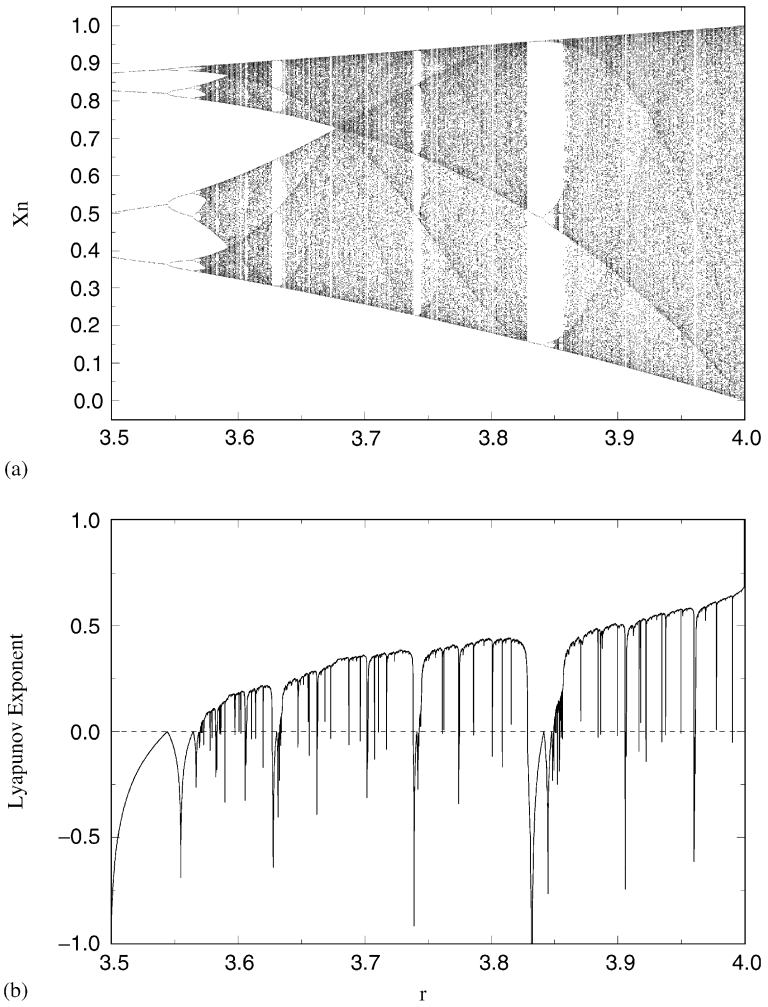

Fig. 5. (a) Orbit diagram, (b) Lyapunov exponent (Λ) for the logistic map as function of parameter r with step $\Delta r = 0.0001$.

Similar results were obtained when the number of iterations was increased to 2^{24}, and also when the binary strings' length was taken to be 15.

We depict in Figs. 6a and b the LMC, ($C_{LMC} \equiv C^{(S)}_{E,1}$) and SDL ($C_{SDL} \equiv C^{(S)}_{K,1}$) statistical complexity measures, evaluated for the logistic map (binary sequence) as functions of the parameter r. We observe, in both instances, an abrupt statistical complexity growth around $r > r_\infty \cong 3.57$. After we pass this point, the two measures behave in quite different fashions. Notwithstanding the fact that both measures almost vanish within the large stability windows, in the inter-windows region they noticeably differ. In the latter zone, the LMC measure globally decreases, reaching almost null values. Consider, for example, $r \in [3.58 < r < 3.62]$ in Fig. 6a. The many peaks indicate a local complexity growth. Comparison with the orbit-diagram (see Fig. 5a) indicates that the peaks coincide with the periodic windows. The original LMC measure regards this periodic motion as having a stronger statistical complexity-character than that pertaining to the neighboring chaotic zone. Note, instead, that the SDL measure does grow in the inter-windows region and rapidly falls within the periodic windows (see Fig. 6b).

The behavior ("degree of chaoticity") of the Lyapunov exponent Λ as a function of the parameter r is displayed in Fig. 5a. From this figure, we see that Λ and, as a result, the associated degree of chaoticity grows with r, reaching a maximum at $r = 4$. One would expect that a sensible statistical complexity measure should

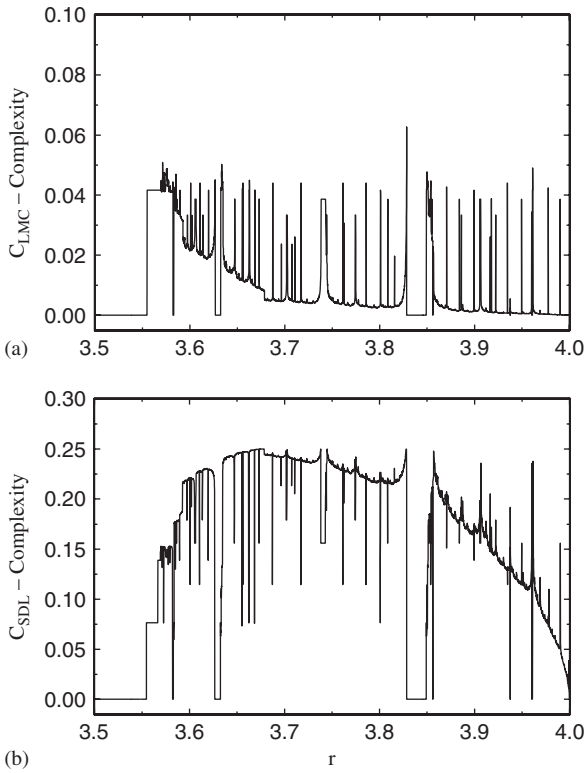

Fig. 6. (a) LMC statistical complexity ($C_{LMC} \equiv C^{(S)}_{E,1}$) and, (b) SDL statistical complexity ($C_{SDL} \equiv C^{(S)}_{K,1}$) measures evaluated for the logistic map (binary sequence) as function of parameter r.

accompany such a growth. In other words, a reasonable statistical complexity measure should take very small values for $r < r_\infty$ and then grow together with the degree of chaoticity till, the chaoticity becoming too large, the complexity falls again. When chaos prevails ($r = 4$) the statistical complexity should vanish. Clearly, this behavioral criteria is not satisfies by the LMC measure. In the case of the SDL measure (see Fig. 6b), even if it produces the expected behavior, it still exhibits a glaring flaw: it is a function of just the entropy. It cannot add information not previously contained in the entropy. This problem will not be remedied by recourse to non-extensive generalizations.

The $C_W \equiv C^{(S)}_{W,1}$ (Shannon–Wooters) and, $C_{JS} \equiv C^{(S)}_{J,1}$ (Jensen–Shannon) statistical complexity measures evaluated for the logistic map as functions of the parameter r are shown in Figs. 7a and b, respectively. The behavior of both measures, as the parameter r varies, is similar. Globally, the decay of C_{JS} is swifter the closer r becomes to the special value $r = 4$. Note that for the region $r_\infty \leqslant r \leqslant 4$ both statistical complexity measures do grow in the inter-windows region and rapidly fall within the periodic windows. More importantly, *we clearly appreciate* here the fact the both statistical complexity measures do distinguish among different periodicities, as rightly demanded in Ref. [19]. The ''Lyapunov''-based criterium mentioned above is satisfied by the two measures as well. For a physicist, $C_{JS} \equiv C^{(S)}_{J,1}$, the Jensen–Shannon statistical complexity measure, is the best one because it is an intensive quantity, which is not the case for $C^{(S)}_{W,1}$.

We present in Fig. 8 plots of a similar character to the interesting ones of Ref. [1] (see their Fig. 2), while in (our) Fig. 8a, we depict the Lyapunov exponent Λ versus the normalized entropy $H \equiv H^{(S)}_1$. Figs. 8b and c exhibit the $C \times H$-diagram for C_E and C_{JS}, respectively (we consider the r-range $3.5 \leqslant r \leqslant 4.0$, although the control parameter does not explicitly appear, of course). Two continuous curves represent C_{max} and C_{min}, evaluated as explained in the preceding section. Note that, for the case of periodic windows, if $H < \mathscr{H} \approx 0.3$, we can ascertain that $\Lambda < 0$, while for $H > \mathscr{H}$ we see that $\Lambda > 0$, which entails chaotic behavior. As evidenced by Fig. 8b, the LMC statistical complexity is larger for periodic than for chaotic motion, which is wrong!.

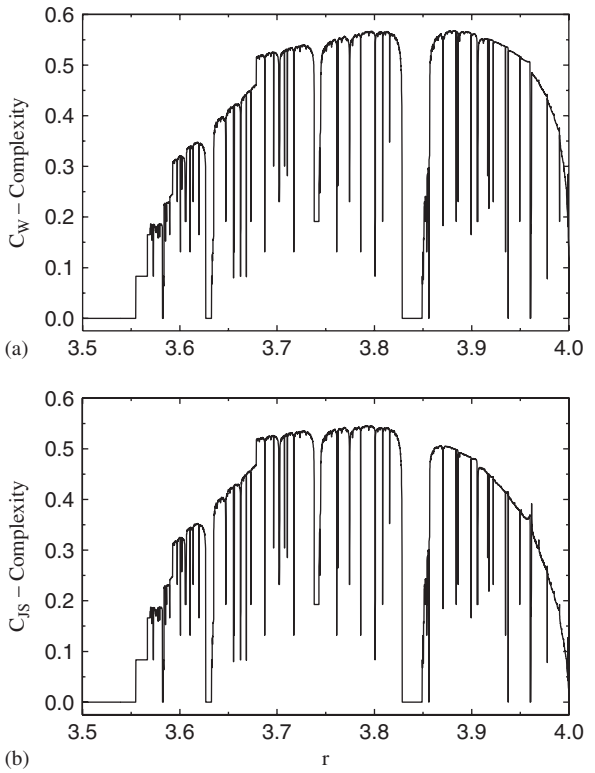

Fig. 7. (a) $C_W \equiv C^{(S)}_{W,1}$ statistical complexity (Shannon–Wooters) and, (b) $C_{JS} \equiv C^{(S)}_{J,1}$ statistical complexity (Jensen–Shannon) measures evaluated for the logistic map (binary sequence) as function of parameter r.

As illustrated by Fig. 8c, the Jensen–Shannon statistical complexity measure, C_{JS}, on the other hand, behaves in opposite manner, and is also different for distinct degrees of periodicity. Summing up: the Jensen–Shannon statistical complexity measure (i) becomes intensive, (ii) is able to distinguish among distinct degrees of periodicity, and (iii) yields a better description of dynamical features (a better grasp of dynamical details).

Extensions of the statistical complexity measures to non-extensive statistics are given by $C_{JT} \equiv C^{(T)}_{\mathscr{J}^T,q}$ (Jensen–Tsallis) and, $C_{JR} \equiv C^{(R)}_{\mathscr{J}^R,q}$ (Jensen–Rényi). Fig. 9 plots the C_{JT} and C_{JR} measures for $q = 0.4, 0.8, 1.0, 1.2, 1.6$ for the logistic map as a function of r. The global behavior of these statistical complexity measures is similar for the different q-values. Nevertheless, as q grows, a better (more detailed) picture ensues, that amplifies differences between different behaviors (both periodic and chaotic).

Changes in C_{min} and C_{max} versus q are also of interest. We consider then in the cases of the statistical complexity measures of the Jensen–Tsallis and Jensen–Rényi type in Figs. 10a and b, respectively. Let analyze these two curves in the plane $C^{(T)}_{\mathscr{J}^T,q} \times H^{(T)}_q$. As q increases, the disorder $H^{(T)}_q$ tends to 1 for all vertices of the "simplex", except for the vertex corresponding to certainty $(1, 0, \ldots, 0)$. The image of the sides and vertices of the simplex, i.e. the curve C_{max} (a piecewise one) reflects this property of the disorder. Notice the associated "rightwards motion" ($H^{(T)}_q \sim 1$). For an even better picture consider Figs. 10a. For large q the disorder $H^{(T)}_q$ is able to discriminate only 2 simplex' sides. For the remaining ones the normalized entropy ≈ 1. Instead, the Jensen–Tsallis divergence does distinguish among the sides, which gives rise to the triangular form apparent in the plane $C^{(T)}_{\mathscr{J}^T,q} \times H^{(T)}_q$.

Fig. 10b illustrates the fact that C_{min} and C_{max} (plane $C^{(R)}_{\mathscr{J}^R,q} \times H^{(R)}_q$) do not appreciably change with q. Images of the simplex' sides (the curve C_{max}) "move" in parallel fashion while vertices seem to "fall" along the same vertical line as q grows while keeping $H^{(R)}_q$ fixed. No triangular area ensues. The original picture obtained for $q = 1$ remains globally invariant as q changes.

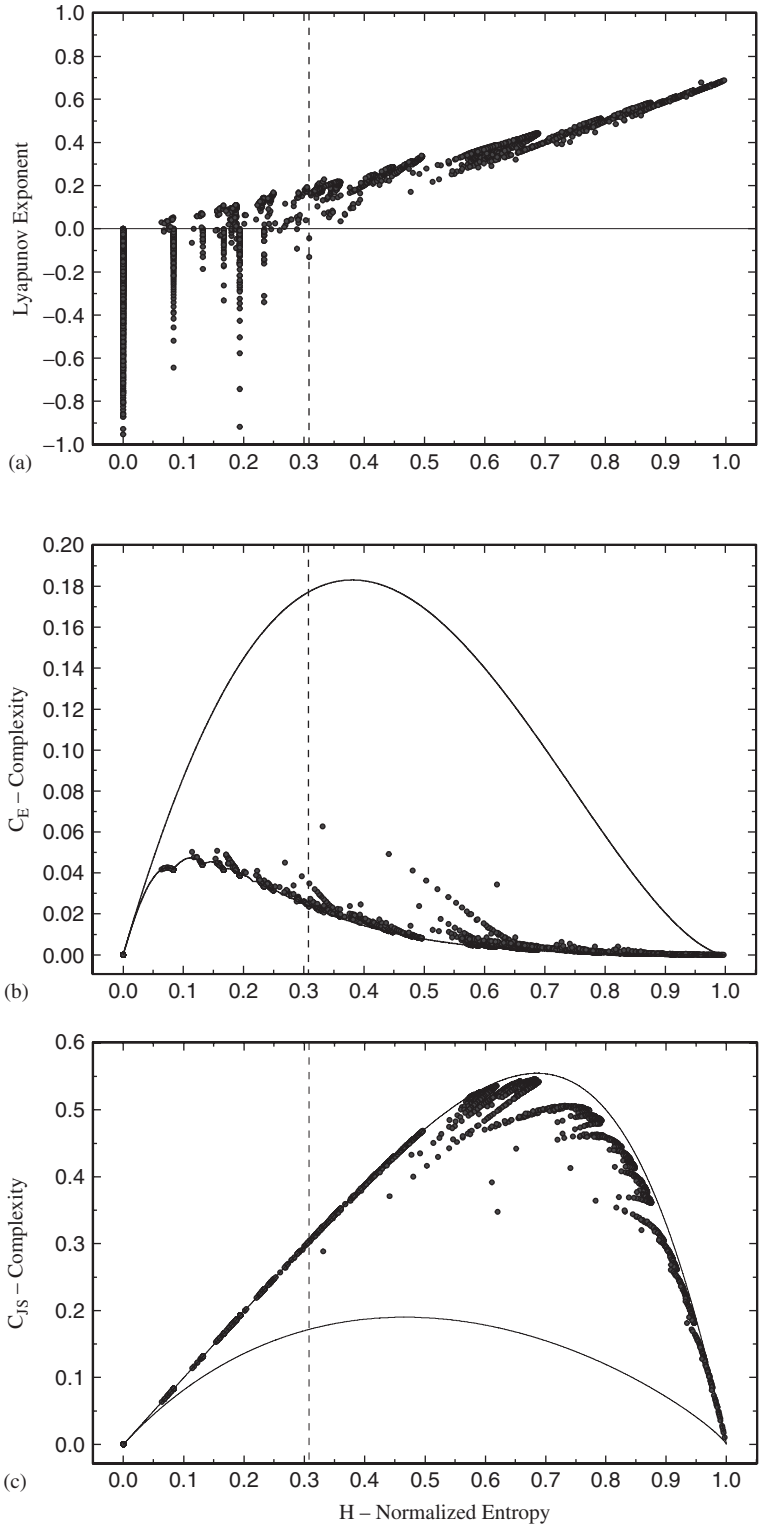

Fig. 8. Quantities of interest are plotted as a function of the normalized entropy $H^{(S)}$: (a) Lyapunov exponent (Λ). Note that the periodic windows are clearly distinguished for $\mathscr{H} \leqslant 0.3$. (b) $C_{LMC} \equiv C_{E,1}^{(S)}$ statistical complexity measure and (c) $C_{JS} \equiv C_{J,1}^{(S)}$ statistical complexity measure. In (b) and (c) we also display the maximum and minimum possible values of the statistical complexity (continuous lines). The vertical dashed line signals the value $\mathscr{H}$ for the end of the periodicity zone.

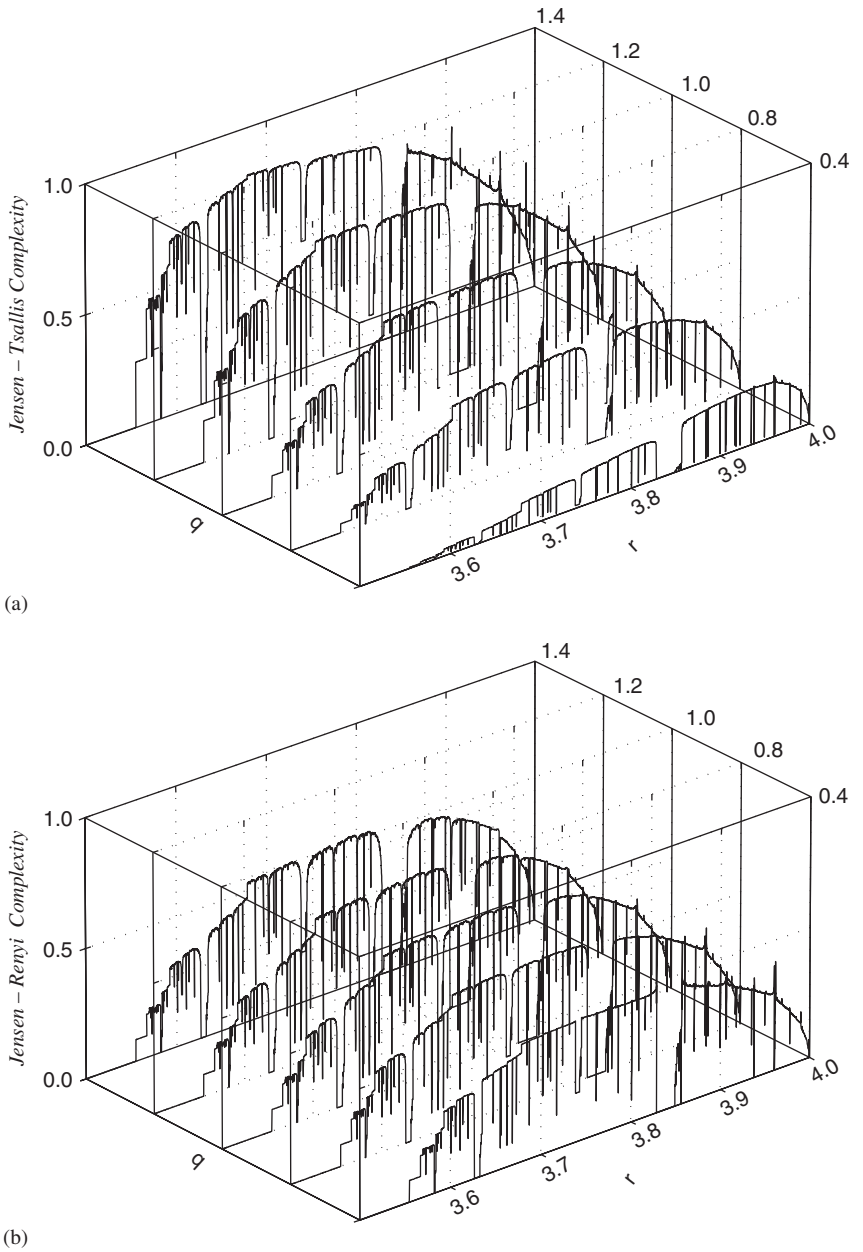

Fig. 9. (a) Jensen–Tsallis, $C_{JT} \equiv C^{(T)}_{\mathscr{J}^T,q}$, and, (b) Jensen–Rényi, $C_{JR} \equiv C^{(R)}_{\mathscr{J}^R,q}$, statistical complexity measures for non-extensive parameter $q = 0.4, 0.8, 1.0, 1.2, 1.6$ evaluated for the logistic map (binary sequence) as function of parameter r.

6. Conclusions

We have in this paper considered a family of statistical complexity measures of the generic form $C = H \cdot Q$, of which a particular instance is given by that of Lopez-Ruiz, Mancini, and Calbet. Non-extensive measures have also been incorporated within the family. A detailed analysis of the C-behavior in the C versus H plane demonstrates the existence of bounds to C that we called C_{max} and C_{min}. These bounds can be systematically evaluated by recourse to a careful geometric analysis performed in the space of probabilities. The scientific importance of this study resides in the fact that C versus H diagrams yield information of a system independently of the values that the different control parameters may adopt. The bounds yield also

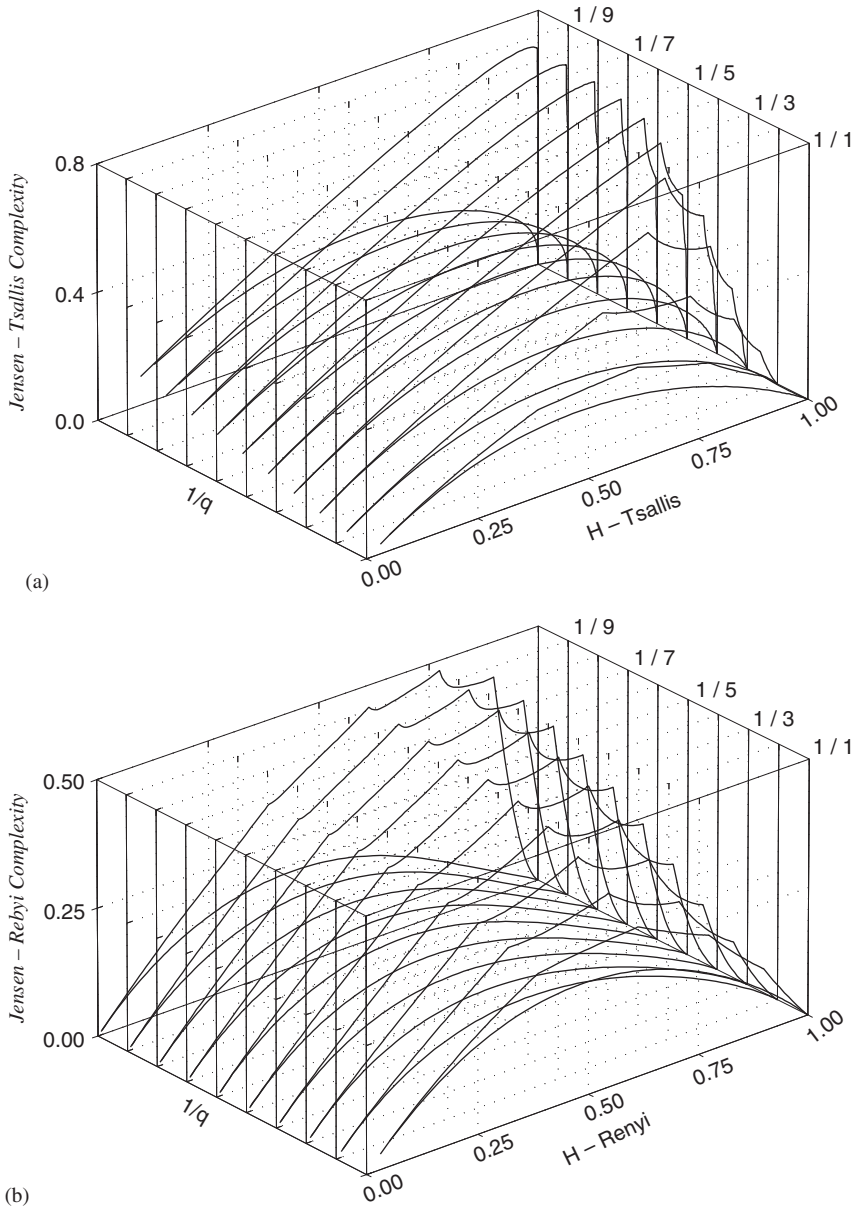

Fig. 10. Variation of the extreme values C_{min} and C_{max} for (a) $C^{(T)}_{\mathscr{F}^T,q} \times H^{(T)}_q$ and, (b) $C^{(R)}_{\mathscr{F}^R,q} \times H^{(R)}_q$ with the non-extensive parameter q. In present case, $N = 6$ and the parameter q was varied between 1 and 1 with step 1.

information that depends on the particular characteristics of a given system, as for instance, the existence of global extrema, or the peculiarities of the system's configuration for which such extrema obtain.

Acknowledgments

This work was partially supported by CONICET (PIP 5687/05, PIP 6036/05) Argentina, Agencia Nacional de Promoción Cientifica y Tecnólogica (PICT 20648/04) Argentina, and the International Office of BMBF (ARG-4-G0A-6A and ARG01-005), Germany.

Appendix A. Entropic definitions

Consider two discrete distributions $P_i = \{p_1^{(i)}, \ldots, p_N^{(i)}\}$, for $i = 1, 2$. The relative entropy of P_1 with respect to P_2 associated to Shannon's measure (see Eq. (4)) is the Kullback–Shannon entropy, that in the discrete case reads

$$\mathrm{K}^{(S)}[P_1|P_2] = \sum_{j=1}^{N} p_j^{(1)} \cdot \ln \frac{p_j^{(1)}}{p_j^{(2)}}. \tag{38}$$

The relative entropy associated to Tsallis' entropy (see Eq. (5)) is called the *relative Kullback– Tsallis entropy* (or q-Kullback measure). The q-Kullback entropy of P_1 with respect to P_2 (both discrete distributions) for $q \neq 1$ is

$$\mathrm{K}_q^{(T)}[P_1|P_2] = \frac{1}{(q-1)} \sum_{j=1}^{N} (p_j^{(1)})^q \left\{ (p_j^{(2)})^{1-q} - (p_j^{(1)})^{1-q} \right\}. \tag{39}$$

The relative entropy associated to Rényi's measure (see Eq. (6)) is the *relative Rényi-entropy*. The q-relative Rényi entropy (P_1 vis-a-vis P_2 discrete distributions) for $q \neq 1$ is

$$\mathrm{K}_q^{(R)}[P_1|P_2] = \frac{1}{(q-1)} \ln \left\{ \sum_{j=1}^{N} \left(p_j^{(1)} \right)^q (p_j^{(2)})^{1-q} \right\}. \tag{40}$$

Definition 1. Given a functional entropic form S and two distributions P_1 and P_2, the associate Jensen's divergence is defined by

$$J_{\mathrm{S}}^{\beta}[P_1, P_2] = \mathrm{S}[\beta P_1 + (1-\beta)P_2] - \beta \mathrm{S}[P_1] - (1-\beta)\mathrm{S}[P_2] \tag{41}$$

with $0 \leqslant \beta \leqslant 1$.

If $\mathrm{S} \equiv \mathrm{S}_1^{(S)}$ is Shannon's measure we have

$$\begin{aligned} J_{\mathrm{S}_1^S}^{\beta}[P_1, P_2] &= \mathrm{S}_1^{(S)}[\beta P_1 + (1-\beta)P_2] \\ &\quad - \beta \mathrm{S}_1^{(S)}[P_1] - (1-\beta)\mathrm{S}_1^{(S)}[P_2], \end{aligned} \tag{42}$$

that may also be expressed in terms of relative entropies

$$\begin{aligned} J_{\mathrm{S}_1^S}^{\beta}[P_1, P_2] &= \beta \mathrm{K}_1^{(S)}[P_1|\beta P_1 + (1-\beta)P_2] \\ &\quad + (1-\beta)\mathrm{K}_1^{(S)}[P_2|\beta P_1 + (1-\beta)P_2], \end{aligned} \tag{43}$$

with $\mathrm{K}_1^{(S)}$the relative Kullback–Shannon measure (Eq. (38)). Since $\mathrm{S}_1^{(S)}$ is a concave function, $J_{\mathrm{S}_1^S}^{\beta} \geqslant 0$ and $J_{\mathrm{S}_1^S}^{\beta}[P_1, P_2] = 0$ iff $P_1 = P_2$. Also, it is a symmetric function of its arguments P_1 and P_2.

If the Jensen divergence is associated to a measure for which definite concavity properties can not be established, the Jensen measure itself may lose its positive character. An example is provided by Rényi's instance $\mathrm{S} \equiv \mathrm{S}_q^{(R)}$ that is concave only for $0 < q < 1$. The relation between Eqs. (42) and (43) allows one to conceive of a generalization of Jensen's divergence, Eq. (41), by simply replacing the relative entropy in Eq. (43), using for the purpose, for instance, Tsallis $\mathrm{K}_q^{(T)}$ as given in Eq. (39). One finds

$$\begin{aligned} \mathscr{J}_{\mathrm{S}_q^T}^{\beta}[P_1, P_2] &= \beta \mathrm{K}_q^{(T)}[P_1|\beta P_1 + (1-\beta)P_2] \\ &\quad + (1-\beta)\mathrm{K}_q^{(T)}[P_2|\beta P_1 + (1-\beta)P_2]. \end{aligned} \tag{44}$$

The properties established for $J_{\mathrm{S}_1^S}^{\beta}$ do not change. For $q > 0$, $\mathscr{J}_{\mathrm{S}_q^T}^{\beta} \geqslant 0$ and $\mathscr{J}_{\mathrm{S}_q^T}^{\beta}[P_1, P_2] = 0$ iff $P_1 = P_2$. Note that Eq. (44) is not a measure obtained via a replacement of $\mathrm{S} = \mathrm{S}_q^{(T)}$ in Eq. (41). Using Jensen's inequality [31,34] one easily shows that, for $q > 1$

$$\mathscr{J}_{\mathrm{S}_q^T}^{\beta}[P_1, P_2] > J_{\mathrm{S}_q^T}^{\beta}[P_1, P_2], \tag{45}$$

entailing that the identity between Eqs. (43) and (42) is lost here. Thus, Eq. (44) defines a new measure based upon relative entropies that generalizes Eq. (41).

In a similar vein one could replace Kullback's relative entropy in Eq. (43) by the pertinent divergence associated to Rényi's entropy $\mathrm{K}_q^{(R)}$, given by Eq. (40), obtaining

$$\mathscr{J}^{\beta}_{\mathrm{S}_q^R}[P_1,P_2] = \beta \mathrm{K}_q^{(R)}[P_1|\beta P_1+(1-\beta)P_2] + (1-\beta)\mathrm{K}_q^{(R)}[P_2|\beta P_1+(1-\beta)P_2]. \tag{46}$$

Appendix B. Geometrical definitions

In order to discuss the geometry of $\Omega \subset \mathbb{R}^N$ we previously introduce some necessary definitions.

Definition 2. A set $\{a_1,a_2,\ldots,a_N\}$ of vectors of $\mathbb{R}^N$ is geometrically independent if, for any set of scalars $t_i \in \mathbb{R}$ we have (i) $\sum_{i=1}^N t_i = 0$ and (ii) $\sum_{i=1}^N t_i \cdots a_i = 0$, if and only if $\{t_i = 0, i = 1,\ldots,N\}$.

Definition 3. Let $\{a_1,a_2,\ldots,a_N\}$ be a set of geometrically independent vectors of $\mathbb{R}^N$. The (N − 1)-simplex spanned by $a_1,a_2,\ldots,a_N$ is the set

$$\Delta^{N-1} = \left\{ x \in \mathbb{R}^N : x = \sum_{i=1}^N t_i \cdot a_i \quad \text{with} \ \sum_{i=1}^N t_i = 1, t_i \geqslant 0, \forall i \right\}.$$

Intuitively, a $(N-1)$-simplex is a $(N-1)$-dimensional generalization of the triangle (in $\mathbb{R}^2$). Each a_i is a *vertex* and the *sides* are the segments joining pairs of vertices. A *face* of a simplex is the subsimplex spanned by a proper subset of $a_1,a_2,\ldots,a_N$, and *the centroid or barycenter* is the vector

$$\mu_{N-1} = \sum_{i=1}^N \frac{a_i}{N}. \tag{47}$$

Let φ be a bijective mapping $\varphi : \{1,\ldots,N\} \to \{1,\ldots,N\}$ and let $b_\varphi^j \in \Delta^{N_1}$ the point

$$b_\varphi^j \triangleq \textit{barycenter of } [e_{\varphi(j)},\ldots,e_{\varphi(N)}] = \sum_{i=\varphi(j)}^{\varphi(N)} e_{\varphi(i)} \Big/ (N-j+1) \tag{48}$$

for $j = 1,\ldots,N$. The point b_φ^j is a vertex of order j if it is the *centroid* of a subsimplex of dimension $(N-j)$. For example, b_φ^1 is the centroid of the N_1-simplex while b_φ^N is the centroid of just a *point*. We will call $\Delta_\varphi^{N_1} \subset \Delta^{N_1}$ the N_1-simplex

$$\Delta_\varphi^{N_1} = [b_\varphi^1,\ldots,b_\varphi^N]. \tag{49}$$

Definition 4. The barycentric subdivision of the N_1-simplex is the subset

$$\mathscr{SB} \triangleq \{\Delta_\varphi^{N_1} = [b_\varphi^1,\ldots,b_\varphi^N], \varphi \in \mathscr{P}_N\},$$

where $N_1 = N-1$ and $\mathscr{P}_N$ represents the set of permutations of $\{1,\ldots,N\}$.

Appendix C. Theorems

Theorem 1. *The probability distribution* (PD) *P that extremizes C for H constant, is of the form*

$$P = \begin{cases} 0 & \textit{for } 1 \leqslant j \leqslant m, \\ p & \textit{for } m+1 \leqslant j \leqslant m+n, \\ (1-pn)/(N-m-n) & \textit{for } m+n+1 \leqslant j \leqslant N \end{cases} \tag{50}$$

with $n,m \in \mathbb{N}$, $0 \leqslant m \leqslant N-1$ and $0 \leqslant n \leqslant N-m-1$.

Proof. Since $C = H \cdot Q$, finding extrema of C for H constant is tantamount to determining extrema for Q. These, in turn, are subjected to two constraints

(a) PD normalization $\sum_{j=1}^{N} p_j = 1$,
(b) constant $H[P] = \mathrm{H}^*$.

Lagrange's multipliers approach leads then to

$$\delta\left(Q[P] - \lambda_1 \left\{ \left(\sum_{j=1}^{N} p_j \right) - 1 \right\} - \lambda_2 \{H[P] - \mathrm{H}^*\} \right) = 0, \tag{51}$$

where λ_1 and λ_2 are Lagrange multipliers. Using Eqs. (29) and (30) we evaluate $\partial Q/\partial p_i$ and $\partial H/\partial p_i$ for $i = 1, \ldots, N$ and, after replacement in Eq. (51) one finds

$$\phi'\left(\sum_{j=1}^{N} f(p_j) \right) \cdot f'(p_i) - \lambda_1 - \lambda_2 \cdot \psi'\left(\sum_{j=1}^{N} g(p_j) \right) \cdot g'(p_i) = 0 \tag{52}$$

for $i = 1, \ldots, N$. Let us call $F(p_i)$ the left-hand side of Eq. (52) and rewrite it in the fashion

$$F(p_i) = f'(p_i) - a \cdot g'(p_i) - b = 0 \quad \text{for } i = 1, \ldots, N, \tag{53}$$

where

$$a = \lambda_2 \psi'\left(\sum_{j=1}^{N} g(p_j) \right) \Big/ \phi'\left(\sum_{j=1}^{N} f(p_j) \right),$$
$$b = \lambda_1 \Big/ \phi'\left(\sum_{j=1}^{N} f(p_j) \right). \tag{54}$$

Note that a and b are well defined and non-null due to the monoticity of ϕ and ψ. Also, since a and b are equal for all $F(p_i)$, $i = 1, \ldots, N$, we can regard them as constants in Eq. (53). After evaluation of the derivative of the left-hand side of Eq. (53) one writes

$$F'(u) = f''(u) - a \cdot g''(u) \tag{55}$$

and ascertains that $F'(u)$ has at the most one real positive solution. Indeed, from the functional form of $f''(u)$ and $g''(u)$ (see Eq. (30)), $F'(u) = 0$ leads to an equation of the form $u^{\alpha} = \beta$ with $\alpha, \beta \in \mathbb{R}$, that possesses at most one real positive solution. Thus, the mean value theorem allows us to infer that Eq. (53) cannot have more than two real positive roots, entailing that the non-vanishing p_i admit just two different values, as expressed by Eq. (50). Now, the condition ($p_i \geqslant 0$) determines a subset $\mathscr{A}$ of the $(N-2)$-dimensional manifold defined by the two constraints above. Lagrange's methodology yields the extrema within $\mathscr{A}(m = 0)$. Note that $\mathscr{A}$ is a bounded domain: the extrema might belong to its boundary as well. Assume then that the extremum P lies in the boundary of $\mathscr{A}$. In such a case, at least one of the $p_i = 0$.

Suppose now that P has m null components ($1 \leqslant m \leqslant N-1$). A similar line of reasoning on an $N-m$-dimensional space allows one to conclude that the $(N-m)$ non-vanishing components of P admit at the most two different values. In other words, the $P \in \Omega$ that extremizes C subject to H constant are of the form given by Eq. (50), which completes the proof. □

Theorem 2. *The statistical complexity measure C has identical values on each of the $N!$ barycentric subdivisions $\Delta_{\varphi}^{N_1}$.*

Proof. The C-value is independent of the order of the p_i, so that

$$C[\{p_{\varphi(1)}, \ldots, p_{\varphi(N)}\}] \equiv C[\{p_{\varphi'(1)}, \ldots, p_{\varphi'(N)}\}] \tag{56}$$

for any permutation φ and φ' defined on $\{1, \ldots, N\}$. □

Theorem 3. *Vectors P of the form given by Eq.* (50) *are the* sides *of $\Delta_I^{N_1}$.*

Proof. Let $L_{j,k}$ be the *side* joining b^j with b^k (see Fig. 1b). $L_{j,k}$ can be defined in parametric fashion via $\mathscr{L}^*_{j,k} : [0,1] \to L_{j,k} \subset \Omega$, where

$$\mathscr{L}^*_{j,k}(\lambda) = \lambda b^j + (1-\lambda) b^k. \tag{57}$$

Expressing vertices in the canonical basis and assuming $j<k$ we get

$$\mathscr{L}^*_{j,k}(\lambda) = \left(\frac{\lambda}{N-j+1}\right) \sum_{i=j}^{k-1} e_i + \left(\frac{\lambda}{N-j+1} + \frac{1-\lambda}{N-k+1}\right) \sum_{i=k}^{N} e_i. \tag{58}$$

After calling $\rho = \lambda/(N-j+1)$, the equation above defines now a function $\mathscr{L}_{j,k} : [0, 1/(N-j+1)] \to L_{j,k} \subset \Omega$

$$\begin{aligned} \mathscr{L}_{j,k}(\rho) &= \mathscr{L}^*_{j,k}(\rho(N-j+1)) \\ &= \rho \sum_{i=j}^{k-1} e_i + \left(\frac{1-\rho(k-j)}{N-k+1}\right) \sum_{i=k}^{N} e_i \end{aligned} \tag{59}$$

with $j<k$. Thus, if a vector P belongs to the *side* $L_{j,k}$, we have $P \equiv \mathscr{L}_{j,k}(\rho)$ for some value $0 \leqslant \rho \leqslant 1/(N-j+1)$ entailing that P will have the form of Eq. (50) with $m = j-1$ and $n = k-j$. Conversely, if P is of the form given by Eq. (50), it is easily seen that it belongs to an $L_{j,k}$. □

Corollary 1. *C-extrema for H constant are reached just on the sides of* $\Delta_I^{N_1}$.

Theorem 4. *For* $j<k$ *let* $\mathscr{L}_{j,k}(\rho)$ *be the parametric definition of Eq.* (59) *for the side* $L_{j,k}$. *Then* $H[\mathscr{L}_{j,k}(\rho)]$ *is a strictly increasing function of* ρ *for* $0 \leqslant \rho \leqslant 1/(N-j+1)$.

Proof. Because of H's definition, the derivative of $H\left[\mathscr{L}_{j,k}(\rho)\right]$ with respect to ρ is positive for $0 \leqslant \rho \leqslant 1/(N-j+1)$. □

Corollary 2. *The function* $H : \Omega \to [0,1]$ *restricted to a side* $L_{j,k}$, *with* $j \leqslant k$ *is an injective function.*

Corollary 3. *The H-values on the vertices of* $\Delta_I^{N_1}$ *constitute a partition of* $[0,1]$,

$$0 = H[b^N] < H[b^{(N-1)}] < \cdots < H[b^2] < H[b^1] = 1. \tag{60}$$

For any set $\mathscr{X} \subset \Omega$ we will use the notation $H[[\mathscr{X}]]$ to the image through H of $\mathscr{X}$. Thus, $H[[L_{1,N}]] = [0,1]$. Of $j<k$, $H[[L_{j,k}]] = [H[b^k], H[b^j]] \subset [0,1]$. Also, from the order relation of the preceding corollaries we deduce

Corollary 4. $\{H[[L_{i,i+1}]], i = N_1, \ldots, 1\}$ *is a series of disjoint sets whose union is the* $[0,1]$ *interval.*

For each value of $\mathrm{H}^* \in [0,1]$, $\mathscr{E}_h$ is the set $\mathscr{E}_h = \{P \in \Delta_I^{N_1} : H[P] = \mathrm{H}^*\}$.

Theorem 5.

(a) *For a bi-dimensional face* $[b^i, b^j, b^k]$, *with* $i<j<k$, *and for* $\mathrm{H}^* \in [H[b^k], H[b^i]]$, *there exists just one point* $P_h^{(ik)}$ *of* $L_{i,k}$ *and only one point* $P_h^{(ijk)}$ *of* $L_{i,j} \cup L_{j,k}$ *within the set* $\mathscr{E}_h$.
(b) *For the simplex* $\Delta_I^{N_1}$ *and for* $\mathrm{H}^* \in [0,1]$, *there exists* (i) *one and only one point* $P_h^{(1)}$ *of* $\mathscr{B}^{(1)} \equiv L_{1,N}$ *and* (ii) *one and only one point* $P_h^{(N_1)}$ of $\mathscr{B}^{(N_1)} \equiv \bigcup_{i=1}^{N_1} L_{i,i+1}$ *within the set* $\mathscr{E}_h$.

Proof. Both (a) and (b) follow trivially from Corollary 4 together with the injectivity mentioned in Corollary 2. □

Theorem 6. *If C fulfills the regularity hypothesis* (I) *then*

$$P_h^{max} = P_h^{(1)} \quad \text{and} \quad P_h^{min} = P_h^{(N_1)}. \tag{61}$$

If, instead, C fulfills the regularity hypothesis (II), *then*

$$P_h^{min} = P_h^{(1)} \quad \text{and} \quad P_h^{max} = P_h^{(N_1)}. \tag{62}$$

Proof. Assume that C satisfies the regularity hypothesis (I). From Corollary 3, H-values at the vertices yield a partition for $[0,1]$. Thus, $\{\exists j$ such that $H[b^{j+1}] \leqslant \mathrm{H}^* \leqslant H[b^j]\}$ (see 4). Let i, k be sub-indices such that $i<j<k$. Consider now $P_h^{(ik)}$, $P_h^{(ijk)}$, $P_h^{(1)}$, and $P_h^{(N_1)}$ as in Proposition 4. Applying twice the regularity condition (I), firstly to the bi-dimensional face $[b^1, b^j, b^k]$ and then to $[b^j, b^k, b^N]$, one easily sees that $C[P_h^{(1)}] \geqslant C[P_h^{(ik)}]$. Thus, since $H[b^{j+1}] \leqslant \mathrm{H}^* \leqslant H[b^j]$ and $P_h^{(N_1)} \in \mathscr{B}^{(N_1)}$, then $P_h^{(N_1)}$ belongs to $L_{j,j+1}$. Applying once again the regularity hypothesis (I) in twofold fashion, firstly to the bi-dimensional face $[b^i, b^j, b^{j+1}]$ and then to $[b^j, b^{j+1}, b^k]$, it is straightforwardly seen that $C[P_h^{(N_1)}] \leqslant C[P_{\mathrm{h}}^{(ik)}]$. Thus, for any i,k such that $i<j<j+1<k$ it holds that $C[P_h^{(N_1)}] \leqslant C[P_h^{(ik)}] \leqslant C[P_{\mathrm{h}}^{(1)}]$. Remembering thatwe know (Corollary 1) that extremal C-complexity values for constant H are reached on the sides of the simplex, it is then proved that

$$C[P_h^{(N_1)}] \leqslant C[P] \leqslant C[P_h^{(1)}] \quad \forall P \in \mathscr{E}_h. \tag{63}$$

In other words, if C fulfills the regularity hypothesis (I), then $P_h^{max} = P_h^{(1)}$ and $P_h^{min} = P_h^{(N_1)}$. One proceed in analogous fashion if C satisfies the regularity hypothesis (II). □

Theorem 7. *If C_{max} and C_{min} are upper and lower bounds for the statistical complexity measure in the plane* $(H \times C)$, *then*

- *If C satisfies the regularity hypothesis* (I), $C_{max} \equiv \mathscr{G}^{(1)}$ *and* $C_{min} \equiv \mathscr{G}^{(N_1)}$.
- *If C satisfies the regularity hypothesis* (II), $C_{max} \equiv \mathscr{G}^{(N_1)}$ *and* $C_{min} \equiv \mathscr{G}^{(1)}$.

Proof. It is straightforwardly obtained by application of Theorem 5 to each $\mathrm{H}^* \in [0,1]$. □

References

[1] J.P. Crutchfield, K. Young, Phys. Rev. Lett. 63 (1989) 105.
[2] J.C. Sprott, Chaos and Time Series Analyis, Oxford University Press, Oxford, 2004.
[3] E. Ott, T. Sauer, J.A. Yorke, Coping with Chaos, Wiley, New York, 1994.
[4] R.J. Solomonoff, Inf. Control 7 (1964) 224.
[5] A.N. Kolmogorov, Problems Inf. Transmission 1 (1965) 1.
[6] G.J. Chaitin, J. Assoc. Comput. Mach. 13 (1966) 547.
[7] H. Kantz, J. Kurths, G. Meyer-Kress, Nonlinear Analysis of Physiological Data, Springer, Berlin, 1998.
[8] O.A. Rosso, M.T. Martin, A. Plastino, Physica A 320 (2003) 497.
[9] M.T. Martin, A. Plastino, O.A. Rosso, in: M.P. Das, F. Green (Eds.), Condensed Matter Theories, vol. 17, Proceedings of Workshop on Condensed Matter Theories, Sydney, Australia, 2001, Nova Science Publishers, New York, 2003, pp. 279–291.
[10] O.A. Rosso, M.T. Martin, A. Plastino, in: O. Descalzi, J. Martinez, S. Ricca (Eds.), Instabilities and Non-equilibrium Structures, vol. IX, Kluwer Academic Publishes, Dordrech, 2004, pp. 281–290.
[11] A. Plastino, M.T. Martin, O.A. Rosso, in: M. Gell-Mann, C. Tsallis (Eds.), Nonextensive Entropy—Interdisciplinary Applications, Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press, Oxford, 2004, pp. 261–293.
[12] O.A. Rosso, M.T. Martin, A. Plastino, Physica A 347 (2005) 444.
[13] O.A. Rosso, M.T. Martin, A. Figliola, K. Keller, A. Plastino, J. Neurosci. Methods (2006) in press.
[14] A.M. Kowalski, M.T. Martin, A. Plastino, A.N. Proto, O.A. Rosso, Phys. Lett. A 311 (2003) 180.
[15] A.M. Kowalski, M.T. Martin, A. Plastino, O.A. Rosso, Int. J. Mod. Phys. B 19 (2005) 2273.
[16] C.M. González, H.A. Larrondo, O.A. Rosso, Physica A 354 (2005) 281.
[17] H.A. Larrondo, C.M. González, M.T. Martin, A. Plastino, O.A. Rosso, Physica A 356 (2005) 133.
[18] J.S. Shiner, M. Davison, P.T. Landsberg, Phys. Rev. E 59 (1999) 1459.
[19] D.P. Feldman, J.P. Crutchfield, Phys. Lett. A 238 (1998) 244.
[20] R. López-Ruiz, H.L. Mancini, X. Calbet, Phys. Lett. A 209 (1995) 321.
[21] X. Calbet, R. López-Ruiz, Phys. Rev. E 63 (2001) 066116.
[22] C. Anteneodo, A.R. Plastino, Phys. Lett. A 223 (1997) 348.
[23] J.S. Shiner, M. Davison, P.T. Landsberg, Phys. Rev. E 62 (2000) 3000.
[24] B.A. Huberman, T. Hogg, Physica D 22 (1986) 376.

[25] P. Grassberger, Int. J. Theor. Phys. 25 (1986) 907.
[26] C.E. Shannon, Bell System Technol. J. 27 (1948) 379;623.
[27] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[28] M.T. Martin, A. Plastino, O.A. Rosso, Phys. Lett. A 311 (2003) 126.
[29] P.W. Lamberti, M.T. Martin, A. Plastino, O.A. Rosso, Physica A 334 (2004) 119.
[30] J. Aczél, Z. Daròczy, On Measures of Information and their Characterization, Mathematics in Science and Engineering, vol. 115, Academic Press, New York, 1975.
[31] M. Basseville, Information: Entropies, Divergences et Mayennes, Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA), Publication Interne, vol. 1020, 1996. (Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France).
[32] W.K. Wootters, Phys. Rev. D 23 (1981) 357.
[33] W.T. Grandy Jr., P.W. Milonni (Eds.), Physics and probability: Essays in Honor of Edwin T. Jaynes, Cambridge University Press, New York, 1993.
[34] P.W. Lamberti, A.P. Majtley, Physica A 239 (2003) 81.
[35] C. Tsallis, J. Stat. Phys. 52 (1988) 479.
[36] M. Gell-Mann, C. Tsallis (Eds.), Nonextensive Entropy—Interdisciplinary Applications, Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press, Oxford, 2004.
[37] A Regularly updated bibliography on Tsallis' formalism and applications can be found at web page http://tsallis.cat.cbpf.br/TEMUCO.pdf.
[38] A. Rényi, On Measures of Entropy and Information, in: Fourth Berkeley Symposium on Mathematics, Statistics and Probability, vol. 1, 1970, p. 547.
[39] J.P. Crutchfield, D.P. Feldman, C.R. Shalizi, Phys. Rev. E 62 (2000) 2996.
[40] P.M. Binder, N. Perry, Phys. Rev. E 62 (2000) 2998.
[41] M.T. Martin, Ph.D. Thesis, Department of Mathematics, Faculty of Sciences, University of La Plata, La Plata, Argentina, 2004.
[42] A.R. Plastino, A. Plastino, Phys. Rev. E 54 (1996) 4423 and references therein.
[43] K. Mischaikow, M. Mrozek, J. Reiss, A. Szymczak, Phys. Rev. Lett. 82 (1999) 1114.
[44] G.E. Powell, I.C. Percival, J. Phys. A: Math. Gen. 12 (1979) 2053.
[45] O.A. Rosso, M.L. Mairal, Physica A 312 (2002) 469.
[46] C. Bandt, B. Pompe, Phys. Rev. Lett. 88 (2002) 174102.