Complexity of switching chaotic maps

M. Antonelli[1], L. De Micco[1,2], O. A. Rosso[3,4] and H. A. Larrondo[1,2]
[1] Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Mar del Plata, Argentina.
[2] CONICET.
[3] LaCCAN/CPMAT Instituto de Computaçao, Universidade Federal de Alagoas, Maceió,Alagoas, Brazil.
[4] Laboratorio de Sistemas Complejos, Facultad de Ingeniería, Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina.

Abstract

In the last years, digital systems such as Digital Signal Processors (DSP), Field Programmable Gate Arrays (FPGA) and Application-Specific Integrated Circuits (ASIC), became the standard in all experimental sciences.

In these systems digital implementations have a custom made numerical representations, therefore finite arithmetic needs to be investigated. Binary floating- and fixed-point are the numerical representations available. Fixed-point representation is prefered over floating-point when speed, low power and/or small circuit area is necessary.

Chaotic systems implemented in finite precision will always become periodic with period T, each of them produces specific period and statistical characteristics that need to be evaluated. It has been recently shown that it is convenient to describe the statistical characteristic using both, a non causal and a causal probability distribution function (PDF). The corresponding entropies, must be evaluated to quantify these PDF's.

In previous works an interesting study about period was carried out using as an example two well known chaotic maps: the tent map and the logistic map. After that, switching techniques between these two maps were tested. In this paper we complement that analysis by characterizing the behaviour of these maps from an statistical point of view using causal and non causal quantifiers.

VERSION: May 4, 2017

1 Introduction

In the last years digital systems became the standard in all experimental sciences. By using virtual instruments, new programmable electronic devices such as Digital Signal Processors (DSP) and Reconfigurable electronics such as Field Programmable Gate Arrays (FPGA) or Application-Specific Integrated Circuits (ASIC), allow experimenters to design and modify their own signal generators, measuring systems, simulation models, etc.

The effect of finite precision in these new devices needs to be investigated. Floating-point is not always available when speed, low power and/or small circuit area are is required, a fixed-point solution is better in these cases. Fixed-point representation is critical if chaotic systems must be implemented, because due to roundoff errors digital implementations will always become periodic with period T and unstable orbits with a low periods may become stable destroying completely the chaotic behavior.

Grebogi and coworkers [1] studied this subject and they saw that the period T scales with roundoff ϵ as $T \sim \epsilon^{-d/2}$ where d is the correlation dimension of the chaotic attractor. To have a large period T is an important property of chaotic maps, in [2] Nagaraj et. als studied the effect of switching over the average period lengths of chaotic maps in finite precision. They saw that the period T of the compound map obtained by switching between two chaotic maps is higher than the period of each map. Liu et. als [3] studied different switching rules applied to linear systems to generate chaos. Switching issue was also addressed in [4], author considers some mathematical, physical and engineering aspects related to singular, mainly switching systems. Switching systems naturally arise in power electronics and many other areas in digital electronics. They have also interest in transport problems in deterministic ratchets [5] and it is known that synchronization of the switching procedure affects the output of the controlled system.

Stochasticity and mixing are also relevant, to characterize these properties several quantifiers were studied [6]. Among them the use of an entropy-complexity representation ($H-C$ plane) and causal-noncausal entropy ($H_{BP}-H_{Val}$ plane) deserves special consideration [7,8,9,6,10,11]. A fundamental issue is the criterium to select the probability distribution function (PDF) assigned to the time series. Causal and non causal options are possible. Here we consider the non-causal traditional PDF obtained by normalizing the histogram of the time series. Its statistical quantifier is the normalized entropy H_{Val} that is a measure of equiprobability among all allowed values. We also consider a causal PDF that is obtained by assigning ordering patterns to segments of trajectory of length D. This PDF was first proposed by Bandt & Pompe in a seminal paper [12]. The corresponding entropy H_{BP} was also proposed as a quantifier

by Bandt & Pompe. Amigó and coworkers proposed the number of forbidden patterns as a quantifier of chaos [13]. Essentialy they reported the presence of forbidden patterns as an indicator of chaos. Recently it was shown that the name forbidden patterns is not convenient and it was replaced by missing patterns (MP) [14]. PONER ALGO DE BPW

In this paper we study the statistical characteristics of five maps, two well known maps: (1) the tent map (TENT) and (2) logistic map (LOG), and three additional maps generated from them: (3) SWITCH, generated by switching between TENT and LOG; (4) EVEN, generated by skipping all the elements in odd position in SWITCH time series and (5) ODD, generated by discarding all the elements in an even position in SWITCH time series. Binary floating- and fixed-point numbers are used, these specific numerical systems may be implemented in modern programmable logic devices.

The main contributions of this paper are:

(1) the definition of different statistical quantifiers and their relationship with the properties of the series generated by the studied chaotic maps.
(2) the study of how these quantifiers detect the evolution of stochasticity and mixing of the chaotic maps according as the numerical precision varies.
(3) the effect on the period and the statistical properties of the time series of switching between two different maps.
(4) the effect on the period and the statistical properties of the time series of skipping values in the switched maps.

Organization of the paper is as follows: section 2 describes the statistical quantifiers used in the paper and the relationship between their value and characteristics of the causal and non causal PDF considered; section 3 shows and discusses the results obtained for all the numerical representations. Finally section 4 deals with final remarks and future work.

2 Information theory quantifiers

Dynamical systems are systems that evolve in time. In practice, one may only be able to measure a scalar time series $\mathcal{X}(t)$ which may be a function of variables $\mathcal{V} = \{v_1, v_2, \cdots, v_k\}$ describing the underlying dynamics (i.e. $d\mathcal{V}/dt = f(\mathcal{V})$. Given a time series or other observational data, we try to infer properties of an unfamiliar system from the analysis of measured time record of it behavior (time series). how much information are these data revealing about the dynamics of the underlying system or processes? The information content of a system is typically evaluated via a probability distribution function (PDF) P describing the apportionment of some measurable or observable quantity,

generally a time series $\mathcal{X}(t)$. We can define Information Theory quantifiers as measures able to characterize relevant properties of the PDF associated with these time series, and in this way we should judiciously extract information on the dynamical system under study. These quantifiers represent metrics on the space of PDFs for data sets, allowing to compare different sets and classifying them according to the properties of underlying processes - broadly, stochastic vs. deterministic.

In our case, we are interested in chaotic dynamics. Thus we are interested in metrics which take the temporal both order of observations explicitly into account; i.e. the approach is fundamentally *causal* and *statistical* in nature. In a purely statistical approach, correlations between successive values from the time series are ignored or simply destroyed via construction of the PDF; while a causal approach focuses on the PDFs of data sequences.

The quantifiers selected are based on symbolic counting and ordinal pattern statistics. For an application of alternative quantifiers based on Symbolic Dynamics to environmental data, we refer to [**?**]. The metrics to be used can be broadly classified along two categories: those which quantify the *information content* of data versus those related to their *complexity*. Note that we are referring to the space of probability density functions here, not physical space. For the sake of clarity and simplicity, we only introduce Information Theory quantifiers that are defined on discrete PDFs in this section, since we are only dealing with discrete data (time series). However, all the quantifiers also have definitions for the continuous case [15,**?**] .

2.1 Shannon entropy and statistical complexity

Entropy is a basic quantity that can be regarded to as a measure of the uncertainty associated (information) to the physical process described by P. When dealing with information content, the Shannon entropy is often considered as the foundational and most natural one [15,**?**]. Regarded as a measure of uncertainty, is the most paradigmatic example of these information quantifiers.

Let a $P = \{p_i; i = 1, \ldots, N\}$ with $\sum_{i=1}^{N} p_i = 1$, be a discrete probability distribution, with N the number of possible states of the system under study. The Shannon's logarithmic information measure reads

$$\mathrm{S}[P] \;=\; -\sum_{i=1}^{N} p_i \ln [p_i] \;. \tag{1}$$

if $\mathrm{S}[P] = \mathrm{S}_{\min} = 0$, we are in position to predict with complete certainty which of the possible outcomes i, whose probabilities are given by p_i, will

actually take place. Our knowledge of the underlying process described by the probability distribution is maximal in this instance. In contrast, our knowledge is minimal for a uniform distribution $P_e = \{p_i = 1/N, \forall i = 1, \ldots, N\}$ since every outcome exhibits the same probability of occurrence, and the uncertainty is maximal, i.e., $\mathrm{S}[P_e] = \mathrm{S}_{\max} = \ln N$. These two situations are extreme cases, therefore we focus on the 'normalized' Shannon entropy, $0 \leq \mathcal{H} \leq 1$, given as

$$\mathcal{H}[P] \;=\; \mathrm{S}[P]/\mathrm{S}_{\max} \;. \tag{2}$$

AAA

Contrary to information content, there is no universally accepted definition of complexity. Here, we focus on describing the *complexity of time series* and do not refer to the complexity of the underlying *systems*. In fact, "simple" models might generate complex data, while "complicated" systems might produce output data of low complexity [**?**].

An intuitive notion of a quantitative complexity attributes low values both to perfectly ordered data (i.e. with vanishing Shannon entropy) as well as to uncorrelated random data (with maximal Shannon entropy), as both cases can be well described in a compact manner. For example, the statistical complexity of a simple oscillation or trend (ordered), but also of uncorrelated white noise (unordered) would be classified as low. Hence, linear transient trends and measurement noise of some geophysical variable would exhibit small complexity values, but the two processes would differ considerably in their Shannon entropy (Fig. 1). Between the two cases of minimal and maximal entropy, data are more difficult to characterize and hence the complexity should be higher. We seek some functional $\mathcal{C}[P]$ quantifying structures present in the data which deviate from these two cases. These structures relate to organization, correlational structure, memory, regularity, symmetry, patterns, and other properties [**?**].

One suitable way to guarantee the desired properties for a complexity measure is to build the product of a measure of information and a measure of disequilibrium, i.e. some kind of distance from the uniform ('equilibrium') distribution of the accessible states of a system. In this respect, [16] an effective *Statistical Complexity Measure* (SCM) $\mathcal{C}$ was introduced, that is able to detect and discern basic dynamical properties of datasets.

Based on the seminal notion advanced by Lopez-Ruiz *et al.* [**?**], this statistical complexity measure[17,16] is defined through the functional product form

$$\mathcal{C}[P] \;=\; \mathcal{Q}_J[P, P_e] \cdot \mathcal{H}[P] \tag{3}$$

of the normalized Shannon entropy $\mathcal{H}$, see Eq. (8), and the disequilibrium $\mathcal{Q}_J[P, P_e]$ defined in terms of the Jensen-Shannon divergence $\mathcal{J}[P, P_e]$. The latter is given by

$$\mathcal{Q}_J[P, P_e] \;=\; Q_0 \mathcal{J}[P, P_e] \;=\; Q_0\{\mathrm{S}[(P + P_e)/2] - \mathrm{S}[P]/2 - \mathrm{S}[P_e]/2\}, \quad (4)$$

where Q_0 is a normalization constant chosen such that $0 \leq \mathcal{Q}_J \leq 1$:

$$Q_0 \;=\; -2\left\{\frac{N+1}{N}\ln(N+1) - \ln(2N) + \ln N\right\}^{-1} \quad , \quad (5)$$

The inverse of this value is obtained for the non-normalized Jensen-Shannon divergence when one of the components of P, say p_m, is equal to one and the remaining p_j's are zero.

The Jensen-Shannon divergence quantifies the difference between probability distributions and is especially useful to compare the symbolic composition of different sequences[**?**]. Note that the above introduced Information Theory quantifier depends on two different probability distributions: one associated with the system under analysis, P, and the other the uniform distribution P_e. For the latter, other reference distributions can be chosen to test whether an observed distribution is close to a target distribution. It has been shown that there are *limit curves* for complexity: for a given value of $\mathcal{H}$ and any data set, the possible $\mathcal{C}$ values vary between a minimum $\mathcal{C}_{min}(\mathcal{H})$ and a maximum $\mathcal{C}_{max}(\mathcal{H})$, restricting the possible values of the complexity measure [18].

BBB

Complexity denotes a state of affairs that one can easily appreciate when confronted with it; however, is rather difficult to define it quantitatively, probably due to the fact that there is no universal definition of complexity. In between the two special instances of perfect order and complete randomness, a wide range of possible degrees of physical structure exists that should be reflected in the features of the underlying probability distribution P. One would like to assume that the degree of correlational structures would be adequately captured by some functional $\mathcal{C}[P]$ in the same way that Shannon's entropy $\mathrm{S}[P]$ [15] "captures" randomness.

Clearly, the ordinal structures present in a process is not quantified by randomness measures, and consequently, measures of statistical or structural complexity are necessary for a better understanding (characterization) of the system dynamics represented by their time series [**?**]. The opposite extremes of perfect order and maximal randomness are very simple to describe, because they do

not have any structure. The complexity should be zero in these cases. At a given distance from these extremes, a wide range of possible ordinal structures exists. Complexity can be characterized by a certain degree of organization, structure, memory, regularity, symmetry, and patterns [**?**]. The complexity measure does much more than satisfy the boundary conditions of vanishing in the high- and low-entropy limits. In particular the maximum complexity occurs in the region between the system's perfectly ordered state and the perfectly disordered one. Complexity is allows us to detect essential details of the dynamics, and more importantly to characterize the correlational structure of the orderings present in the time series.

The perfect crystal and the isolated ideal gas are two typical examples of systems with minimum and maximum entropy, respectively. However, they are also examples of simple models and therefore of systems with zero complexity, as the structure of the perfect crystal is completely described by minimal information (i.e., distances and symmetries that define the elementary cell) and the probability distribution for the accessible states is centered around a prevailing state of perfect symmetry. On the other hand, all the accessible states of the ideal gas occur with the same probability and can be described by a "simple" uniform distribution.

Statistical complexity is often characterized by the paradoxical situation of a complicated dynamics generated from relatively simple systems. Obviously, if the system itself is already involved enough and is constituted by many different parts, it clearly may support a rather intricate dynamics, but perhaps without the emergence of typical characteristic patterns [**?**]. Therefore, a complex system does not necessarily generate a complex output. Statistical complexity is therefore related to patterned structures hidden in the dynamics, emerging from a system which itself can be much simpler than the dynamics it generates [**?**].

According to López-Ruiz, Mancini and Calbet [**?**], and using an oxymoron, an object, a procedure, or system is said to be complex when it does not exhibit patterns regarded as simple. It follows that a suitable complexity measure should vanish both for completely ordered and for completely random systems and cannot only rely on the concept of information (which is maximal and minimal for the above mentioned systems). A suitable measure of complexity can be defined as the product of a measure of information and a measure of disequilibrium, i.e. some kind of distance from the equiprobable distribution of the accessible states of a system. In this respect, Rosso and coworkers [16] introduced an effective *Statistical Complexity Measure* (SCM) $\mathcal{C}$, that is able to detect essential details of the dynamical processes underlying the dataset.

Based on the seminal notion advanced by López-Ruiz *et al.* [**?**], this statistical

complexity measure[17,16] is defined through the functional product form

$$\mathcal{C}[P] \;=\; \mathcal{Q}_J[P, P_e] \cdot \mathcal{H}[P] \tag{6}$$

of the normalized Shannon entropy $\mathcal{H}$, see Eq. (8), and the disequilibrium $\mathcal{Q}_J$ defined in terms of the Jensen-Shannon divergence $\mathcal{J}[P, P_e]$. That is,

$$\mathcal{Q}_J[P, P_e] \;=\; Q_0 \mathcal{J}[P, P_e] \;=\; Q_0\{\mathrm{S}[(P + P_e)/2] - \mathrm{S}[P]/2 - \mathrm{S}[P_e]/2\}, \tag{7}$$

the above-mentioned Jensen-Shannon divergence and Q_0, a normalization constant such that $0 \leq \mathcal{Q}_J \leq 1$:

$$Q_0 \;=\; -2\left\{\frac{N+1}{N}\ln(N+1) - \ln(2N) + \ln N\right\}^{-1}, \tag{8}$$

are equal to the inverse of the maximum possible value of $\mathcal{J}[P, P_e]$. This value is obtained when one of the components of P, say p_m, is equal to one and the remaining p_j are zero.

The Jensen-Shannon divergence, which quantifies the difference between probability distributions, is especially useful to compare the symbolic composition between different sequences[**?**,**?**,**?**]. Note that the above introduced SCM depends on two different probability distributions: one associated with the system under analysis, P, and the other the uniform distribution, P_e. Furthermore, it was shown that for a given value of $\mathcal{H}$, the range of possible $\mathcal{C}$ values varies between a minimum $\mathcal{C}_{min}$ and a maximum $\mathcal{C}_{max}$, restricting the possible values of the SCM [18]. Thus, it is clear that important additional information related to the correlational structure between the components of the physical system is provided by evaluating the statistical complexity measure.

If our system, with associated discrete PDF, lies in a very ordered state, will be represented by an extremely narrow PDF, that is almost all the p_i–values are almost zero except for a particular state $k \neq i$ with $p_k \cong 1$,then both the normalized Shannon entropy and statistical complexity are close to zero ($\mathcal{H} \approx 0$ and $\mathcal{C} \approx 0$)

On the other hand, when the system under study is represented by a very disordered state, that is when all the p_i–values oscillate around the same value, we have $\mathcal{H} \approx 1$ while $\mathcal{C} \approx 0$

This point was extensively discussed by Rosso and co-workers [**?**,**?**]. The summands can be regarded to as a kind of "distance" between two contiguous probabilities. Thus, a different ordering of the pertinent summands would lead to a different FIM-value, hereby its local nature. In the present work, we follow the Lehmer lexicographic order [**?**] in the generation of Bandt and Pompe PDF (see next section).

YOOOOO

The entropy $H[P]$ is the normalized version of the Entropy proposed by Shannon [15]:

$$H[P] = S[P]/S_{max}, \tag{9}$$

where $S[P] = -\sum_{j=1}^{M} \; p_j \; \ln(p_j)$ and S_{max} is the normalizing constant:

$$S_{max} = S[P_e] = \ln M, \tag{10}$$

and $P_e = \{1/M, \ldots, 1/M\}$ is the uniform distribution.

The statistical complexity $C[P]$ is given by:

$$C[P] = Q_J[P, P_e] \cdot H[P] \; , \tag{11}$$

, and Q_J is named disequilibrium and it is the distance between P and P_e in the probability space. The metric used in this paper is based on the Jensen-Shannon divergence [16]:

$$Q_J[P, P_e] = Q_0 \cdot \{S[\frac{P + P_e}{2}] - S[P]/2 - S[P_e]/2\} \; . \tag{12}$$

The normalization constant Q_0 is:

$$Q_0 = -2 \left\{ \left(\frac{N+1}{N} \right) \ln(N+1) - 2\ln(2N) + \ln N \right\}^{-1} . \tag{13}$$

From the statistical point of view the disequilibrium Q_J is an intensive magnitude, and it is 0 if and only if $P = P_e$. It has been proved that the $C[P]$ quantifies the presence of nonlinear correlations typical of chaotic systems [17,16]. The complexity $C[P]$ is independent from the entropy $H[P]$, as far as different P's share the same entropy $H[P]$ but they have different complexity $C[P]$.

2.2 Determination of a probability distribution

AAA

The quantifiers from Information Theory rely on a probability distribution associated to the time series. The determination of the most adequate PDF is a fundamental problem because the PDF P and its sample space Ω are inextricably linked. The usual histogram technique is inadequate since the data are treated purely stochastic and the temporal information is completely

lost. However, Bandt and Pompe (BP)[12] introduced a simple and robust symbolic methodology that takes into account time ordering of the time series by comparing neighboring values in a time series. The symbolic data are: *(i)* created by ranking the values of the series; and *(ii)* defined by reordering the embedded data in ascending order, which is tantamount to a phase space reconstruction with embedding dimension (pattern length) D and time lag τ (see Information for a more detailed description). In this way, it is possible to quantify the diversity of the ordering symbols (patterns) derived from a scalar time series. Note that the appropriate symbol sequence arises naturally from the time series, and no model-based assumptions are needed. As such, it allows to uncover important details concerning the ordinal structure of the time series [7] and can also yield information about temporal correlation [**?**].

This type of analysis of a time series entails losing details of the original series' amplitude information. However, the symbolic representation of time series by recourse to a comparison of consecutive ($\tau = 1$) or nonconsecutive ($\tau > 1$) values allows for an accurate empirical reconstruction of the underlying phase-space, even in the presence of weak (observational and dynamic) noise [12]. Furthermore, the ordinal patterns associated with the PDF are invariant with respect to nonlinear monotonous transformations; nonlinear drifts or scaling artificially introduced by a measurement device will not modify the estimation of quantifiers, a nice property if one deals with experimental data (see, e.g., [**?**]. The only condition for the applicability of the BP method is a very weak stationary assumption: for $k \leq D$, the probability for $x_t < x_{t+k}$ should not depend on t. For a review of BP's methodology and its applications to physics, biomedical and econophysics signals, see [**?**].

Regarding the selection of the parameters, Bandt and Pompe suggested working with $4 \leq D \leq 6$ for typical time series lengths, and specifically considered a time lag $\tau = 1$ in their cornerstone paper [12].For the artificially generated time series shown below(Figs 1 and 2),we chose D= 6 and follow the Lehmer-permutation scheme [35] to calculate the Fisher Information. For themeasured andmodelled time series analyzed here, the embedding dimension is chosen as D= 4 throughout due to time series length requirements (S1 File), in particular at coarse tem- poral resolution, and to achieve comparability across the different analyses. In all cases, the delay parameter has been set to tau = 1.

Incorporating amplitude information: Weighted ordinal pattern distribution: Recently, the permutation entropy was extended to incorporate also amplitude information [40].Hence, a potential disadvantage of ordinal pattern statistics, namely the loss of amplitude information, can be addressed by introducing-weights in order to obtain a 'weighted permuta- tion entropy (WPE)'. In the context of environmental time series,which typically exhibit a pro- nounced seasonal cycle and thus seasonally varying signal to noise ratios, this ideamight be particularly useful to address (noisy) low-variance patterns (e.g. during

dormancy periods in winter). Weighting the probabilities of individual patterns according to their variance alleviates potential issues regarding to 'high noise, low signal' patterns, because low-variance patterns that are stronlgy affected by noise are down-weighted in the resulting 'weighted ordinal pattern distributions'. For example, [40] show that a weighted entropymeasure is sensitive to sudden changes in the variance of the time series.Here, we extend the idea ofWPE following [40] to derive a weighted permutation entropy (Hw), weighted statistical complexity (Cw), and weighted Fisher Information (Fw). Non-normalizedweights are computed for each temporalwindowfor the time seriesX, such that X D wj Here, the embedding dimension is denoted by D and XD j denotes the arithmeticmean of the time series in the currentwindowwith index j. Thus, theweight of each windowof lengthDis given by its variance in the equation above. The weights wj are then used tomodify the relative frequencies of each ordinal pattern pj

The denominator of this equation provides the normalization, and the Kronecker delta fasd in the enumerator serves to indicatewhich pattern occurs in each windowj: (1; if piasdf pj ; dpi;pj asdf 0; if pi 6fasd pj : After having calculated the appropriate asdfi for each i = 1, . . .,D!, the same Eqs (2), (3) and (6) are applied to obtain the weighted versionsHw , Cw and Fw of the ITQ. The weighting of the ITQcan be considered as noise filter, provided that noise is character- ized by relatively low variance, and thus enhances the signal contained in the time series. In any case, rare patterns are suppressed in favour ofmore frequent ones. Obviously, the window-basedvariance is but one out ofmany weighting recipes, others are easily conceivable. There is also a connection to the celebrated Rènyi entropy ([41]) ! X D! Hq asdfXSS asdf 1 1 ? q log2 pq i ifds1 where the Rènyi exponent q suppresses low-frequency patterns for q¿1; for q = 1, the Shan- non entropy is obtained. The resulting 'Rènyi ordinal pattern distribution' could be considered as a special case of the weighted pattern distribution, characterized by a single exponent; this has not been investigated so far, however. Summarizing, ITQcomputed based on an appropriately weighted formof the ordinal pat- tern distribution are suitable to analyse data setswith considerable amplitude information (e.g. seasonal variation) froman information-theoretic viewpoint. Since this is an issue for the time series investigated here, we willmostly performour ITQanalysis using theweighted versions describedhere.

BBB

The evaluation of the Information Theory derived quantifiers, like those previously introduced (Shannon entropy, Fisher information and statistical complexity), suppose some prior knowledge about the system; specifically, a probability distribution associated to the time series under analysis should be provided beforehand. The determination of the most adequate PDF is a fundamental problem because the PDF P and the sample space Ω are inextricably

linked.

Usual methodologies assign to each time point of the series $\mathcal{X}(t)$ a symbol from a finite alphabet $\mathcal{A}$, thus creating a *symbolic sequence* that can be regarded to as a *non causal coarse grained* description of the time series under consideration. As a consequence, order relations and the time scales of the dynamics are lost. The usual histogram technique corresponds to this kind of assignment. *Causal information* may be duly incorporated if information about the past dynamics of the system is included in the symbolic sequence, i.e., symbols of alphabet $\mathcal{A}$ are assigned to a portion of the phase-space or trajectory.

Many methods have been proposed for a proper selection of the probability space (Ω, P). Among others, of type non causal coarse grained, we can mention frequency counting [**?**], procedures based on amplitude statistics [8], binary symbolic dynamics [**?**], Fourier analysis [**?**], or wavelet transform [**?**]. The suitability of each of the proposed methodologies depends on the peculiarity of data, such as stationarity, length of the series, the variation of the parameters, the level of noise contamination, etc. In all these cases, global aspects of the dynamics can be somehow captured, but the different approaches are not equivalent in their ability to discern all relevant physical details. Bandt and Pompe (BP)[12] introduced a simple and robust symbolic methodology that takes into account time causality of the time series (causal coarse grained methodology) by comparing neighboring values in a time series. The symbolic data are: *(i)* created by ranking the values of the series; and *(ii)* defined by reordering the embedded data in ascending order, which is tantamount to a phase space reconstruction with embedding dimension (pattern length) D and time lag τ. In this way, it is possible to quantify the diversity of the ordering symbols (patterns) derived from a scalar time series.

Note that the appropriate symbol sequence arises naturally from the time series, and no model-based assumptions are needed. In fact, the necessary "partitions" are devised by comparing the order of neighboring relative values rather than by apportioning amplitudes according to different levels. This technique, as opposed to most of those in current practice, takes into account the temporal structure of the time series generated by the physical process under study. As such, it allows us to uncover important details concerning the ordinal structure of the time series [7,19,**?**,**?**] and can also yield information about temporal correlation [**?**,**?**].

It is clear that this type of analysis of a time series entails losing details of the original series' amplitude information. Nevertheless, by just referring to the series' intrinsic structure, a meaningful difficulty reduction has indeed been achieved by BP with regard to the description of complex systems. The symbolic representation of time series by recourse to a comparison of consecutive

($\tau = 1$) or nonconsecutive ($\tau > 1$) values allows for an accurate empirical reconstruction of the underlying phase-space, even in the presence of weak (observational and dynamic) noise [12]. Furthermore, the ordinal patterns associated with the PDF are invariant with respect to nonlinear monotonous transformations. Accordingly, nonlinear drifts or scaling artificially introduced by a measurement device will not modify the estimation of quantifiers, a nice property if one deals with experimental data (see, e.g., [**?**]). These advantages make the BP methodology more convenient than conventional methods based on range partitioning, i.e., a PDF based on histograms.

To use the BP methodology[12] for evaluating the PDF, P, associated with the time series (dynamical system) under study, one starts by considering partitions of the D-dimensional space that will hopefully "reveal" relevant details of the ordinal structure of a given one-dimensional time series $\mathcal{X}(t) = \{x_t; t = 1, \ldots, M\}$ with embedding dimension $D > 1$ ($D \in \mathcal{N}$) and time lag τ ($\tau \in \mathcal{N}$). We are interested in "ordinal patterns" of order (length) D generated by

$$(s) \mapsto \left(x_{s-(D-1)\tau}, x_{s-(D-2)\tau}, \ldots, x_{s-\tau}, x_s\right) \ , \tag{14}$$

which assign to each time s the D-dimensional vector of values at times $s-(D-1)\tau, \ldots, s-\tau, s$. Clearly, the greater D, the more information on the past is incorporated into our vectors. By "ordinal pattern" related to the time (s), we mean the permutation $\pi = (r_0, r_1, \ldots, r_{D-1})$ of $[0, 1, \ldots, D-1]$ defined by

$$x_{s-r_{D-1}\tau} \leq \ x_{s-r_{D-2}\tau} \leq \cdots \leq \ x_{s-r_1\tau} \leq x_{s-r_0\tau}. \tag{15}$$

In this way the vector defined by Eq. (17) is converted into a unique symbol π. We set $r_i < r_{i-1}$ if $x_{s-r_i} = x_{s-r_{i-1}}$ for uniqueness, although ties in samples from continuous distributions have null probability.

In order to illustrate BP method, we will consider a simple example: a time series with seven ($M = 7$) values $\mathcal{X} = \{4, 7, 9, 10, 6, 11, 3\}$ and we evaluate the BP-PDF for $D = 3$ and $\tau = 1$. In this case the state space is divided into 3! partitions and 6 mutually exclusive permutation symbols are considered. The triplet $(4, 7, 9)$ and $(7, 9, 10)$ represent the permutation pattern [012] since they are in increasing order. On the other hand, $(9, 10, 6)$ and $(6, 11, 3)$ correspond to the permutation pattern [201] since $x_{s+2} < x_s < x_{s+1}$, while $(10, 6, 11)$ has the permutation pattern [102] with $x_{s+1} < x_s < x_{s+2}$. Then, the associated probabilities to the 6 patterns are: $p([012]) = p([201]) = 2/5$; $p([102]) = 1/5$; $p([021]) = p([120]) = p([210]) = 0$.

Figure **??** illustrates the construction principle of the ordinal patterns of length $D = 2$, 3 and 4 with $\tau = 1$[**?**]. Consider the sequence of observations $\{x_0, x_1, x_2, x_3\}$. For $D = 2$, there are only two possible directions from x_0 to x_1: up and down. For $D = 3$, starting from x_1 (up) the third part of the pattern can be above x_1, below x_0, or between x_0 and x_1. A similar situation

can be found starting from x_1 (down). For $D = 4$, for each one of the six possible positions for x_2, there are four possible localizations for x_3, yielding $D! = 4! = 24$ different possible ordinal patterns. In Fig. **??**, full circles and continuous lines represent the sequence values $x_0 < x_1 > x_2 > x_3$, which leads to the pattern $\pi = [0321]$. A graphical representation of all possible patterns corresponding to $D = 3, 4$ and 5 can be found in Fig. 2 of Parlitz *et al.*[**?**].

For all the $D!$ possible orderings (permutations) π_i when embedding dimension is D, and time-lag τ, their relative frequencies can be naturally computed according to the number of times this particular order sequence is found in the time series, divided by the total number of sequences,

$$p(\pi_i) = \frac{\#\{s|s \leq N - (D-1)\tau; (s) \text{ is of type } \pi_i\}}{N - (D-1)\tau}, \tag{16}$$

where # denotes cardinality. Thus, an ordinal pattern probability distribution $P = \{p(\pi_i), i = 1, \ldots, D!\}$ is obtained from the time series.

The embedding dimension D plays an important role in the evaluation of the appropriate probability distribution, because D determines the number of accessible states $D!$ and also conditions the minimum acceptable length $M \gg D!$ of the time series that one needs in order to work with reliable statistics [**?**]. In the present work, we follow the Lehmer lexicographic order [**?**] in the generation of Bandt and Pompe PDF.

Regarding the selection of the parameters, Bandt and Pompe suggested working with $4 \leq D \leq 6$, and specifically considered a time lag $\tau = 1$ in their cornerstone paper [12]. Nevertheless, it is clear that other values of τ could provide additional information. It has been recently shown that this parameter is strongly related, if it is relevant, to the intrinsic time scales of the system under analysis [**?**,**?**,**?**].

Additional advantages of the method reside in *i)* its simplicity (it requires few parameters: the pattern length/embedding dimension D and the time lag τ), and *ii)* the extremely fast nature of the calculation process. The BP methodology can be applied not only to time series representative of low dimensional dynamical systems, but also to any type of time series (regular, chaotic, noisy, or reality based). In fact, the existence of an attractor in the D-dimensional phase space in not assumed. The only condition for the applicability of the BP method is a very weak stationary assumption: for $k \leq D$, the probability for $x_t < x_{t+k}$ should not depend on t. For a review of BP's methodology and its applications to physics, biomedical and econophysics signals see Zanin *et al.* [**?**]. Moreover, Rosso *et al.* [7] show that the above mentioned quantifiers produce better descriptions of the process associated dynamics when the PDF is computed using BP rather than using the usual histogram methodology.

The BP proposal for associating probability distributions to time series (of an underlying symbolic nature) constitutes a significant advance in the study of nonlinear dynamical systems [12]. The method provides univocal prescription for ordinary, global entropic quantifiers of the Shannon-kind. However, as was shown by Rosso and coworkers [**?**,**?**], ambiguities arise in applying the BP technique with reference to the permutation of ordinal patterns. This happens if one wishes to employ the BP-probability density to construct local entropic quantifiers, like the Fisher information measure, which would characterize time series generated by nonlinear dynamical systems.

The local sensitivity of the Fisher information measure for discrete PDFs is reflected in the fact that the specific "i-ordering" of the discrete values p_i must be seriously taken into account in evaluating Eq. (**??**). The numerator can be regarded to as a kind of "distance" between two contiguous probabilities. Thus, a different ordering of the summands will lead, in most cases, to a different Fisher information value. In fact, if we have a discrete PDF given by $P = \{p_i, i = 1, \ldots, N\}$, we will have $N!$ possibilities for the i-ordering.

The question is, which is the arrangement that one could regard as the "proper" ordering? The answer is straightforward in some cases, like the one that pertains to histogram-based PDFs. For extracting a time-series PDF P via an histogram procedure, one first divides the interval $[a, b]$ (with a and b the minimum and maximum amplitude values in the time series) into a finite number N_{bin} ($N \equiv N_{bin}$ in Eqs. (7) and (**??**)) of non overlapping equal sized consecutive subintervals $A_k : [a, b] = \bigcup_{k=1}^{N_{bin}} A_k$ and $A_n \bigcap A_m = \emptyset \ \forall n \neq m$. Then, recourse to the usual histogram method, based on counting the relative frequencies of the time series' values within each subinterval, is made. Of course, in this approach the temporal order in which the time-series values emerge plays no role at all. The only pieces of information we have here are the x_t-values that allow one to assign inclusion within a given bin, ignoring just where they are located (this is, the subindex i). Note that the division procedure of the interval $[a, b]$ provides the natural order sequence for the evaluation of the PDF gradient involved in Fisher's information measure.

From now on, we assume that the all Information Theory quantifiers will be evaluated with BP-PDF's, by this reason usually they are called permutation quantifiers or time causal Information Theory quantifiers. In our current paper, we chose the lexicographic ordering given by the algorithm of Lehmer [**?**], among other possibilities, due to its better distinction of different dynamics in the Shannon–Fisher plane, $\mathcal{H} \times \mathcal{F}$ (see [**?**,**?**]).

YOOOO

All the statistical quantifiers studied here are functionals of the PDF associated to the time series. The first step to quantify the statistical properties of

the values (amplitude statistics) of a time series $\{x_i, \forall\ i = 1, \ldots, N\}$ using information theory is to determine the concomitant PDF. This is an issue studied in detail in previous papers [**?**]. Let us summarize the procedure:

(1) a finite alphabet with M symbols $\mathcal{A} = \{a_1, \ldots, a_M\}$ is chosen.
(2) each one of these symbols is assigned to each: (a) value of the time series (or non-overlapped set of consecutive values) or (b) portion of length D of the trajectory.
(3) the normalized histogram of the symbols is the desired PDF.
(4) a randomness quantifier is calculated over the PDF. In our case we calculate (a) normalyzed Shannon entropy H, (b) normalyzed statistical complexity C and (d) MP.

Note that if option (a) is chosen in step 2 then the PDF is non causal, because all the information about the time evolution of the system generating $\{x_i\}$ is completely lost. On the contrary if option (b) is chosen in step 2 then the PDF is causal, in the sense it includes information about the temporal evolution.

NOMBRAR LA BPW

NOMBRAR LOS PLANOS QUE TERMINEMOS ELIGIENDO

Of course there are infinite possibilities to choose the alphabet $\mathcal{A}$ as well as the length D. Bandt & Pompe made a proposal for a causal PDF that has been shown to be easy to implement and useful in a great variety of applications REFERENCIAS A APLICACIONES. The procedure is the following [12,20,21]: ESTA PARTE ESTÁ EN EL TIEMPO Y DEBERÍA ESTAR EN LAS MUESTRAS

- Given a series $\{x_t, \forall\ t = 0, \Delta t, \cdots, N\Delta t\}$, a sequence of vectors of length D is generated.

$$(s) \longmapsto \left(x_{t-(d-1)\Delta t}, x_{t-(d-2)\Delta t}, \ldots, x_{t-\Delta t}, x_t\right) \tag{17}$$

  Each vector turns out to be the "history" of the value x_t. Clearly, the longer the length of the vectors D, the more information about the history would the vectors have but a higher value of N is required to have an adequate statistics.

- The permutations $\pi = (r_0, r_1, \cdots, r_{D-1})$ of $(0, 1, \cdots, D-1)$ are called "order of patterns" of time t, defined by:

$$x_{t-r_{D-1}\Delta t} \leq x_{t-r_{D-2}\Delta t} \leq \cdots \leq x_{t-r_1\Delta t} \leq x_{t-r_0\Delta t} \tag{18}$$

  In order to obtain an unique result it is considered $r_i < r_{i-1}$ if $x_{t-r_i\Delta t} =$

$x_{t-r_{i-1}\Delta t}$.

In this way, all the $D!$ possible permutations π of order D, and the PDF $P = \{p(\pi)\}$ is defined as:

$$p(\pi) = \frac{\sharp\{s|s \leq N - D + 1; (s) \quad \texttt{has type } \pi\}}{N - D + 1} \tag{19}$$

In the last expression the $\sharp$ symbol means "number".

This procedure has the advantages of being *i*) simple, *ii*) fast to calculate, *iii*) robust in presence of noise, and *iv*) invariant to lineal monotonous transformations. DICE QUE ES ROBUSTO FRENTE A LA PRESENCIA DE RUIDO PERO NO, REFERENCIAS?

It is applicable to weak stationarity processes (for $k = D$, the probability that $x_t < x_{t+k}$ doesn't depend on the particulary t [12]).The causality property of the PDF allows the quantifiers (based on this PDFs) to discriminate between deterministic and stochastic systems [22].

According to this point Bandt and Pompe suggested $3 \leq D \leq 7$. $D = 6$ has been adopted in this work.

2.3 Information Planes

AAA

A particularly useful visualization of the quantifiers from Information Theory is their juxtaposition in two-dimensional graphs. Three such *causal information planes* can be defined: *a)* The *causality entropy–complexity plane*, $\mathcal{H} \times \mathcal{C}$, is based only on global characteristics of the associated time series PDF (both quantities are defined in terms of Shannon entropies); while *b)* the *causality Shannon-Fisher plane*, $\mathcal{H} \times \mathcal{F}$, is based on global and local characteristics of the PDF; finally, *c)* the plane spanned by $\mathcal{C} \times \mathcal{F}$ is of course fully redundant to the other two. In the case of $\mathcal{H} \times \mathcal{C}$ the value range is $[0, 1] \times [\mathcal{C}_{min}(\mathcal{H}), \mathcal{C}_{max}(\mathcal{H})]$, while in the causality plane $\mathcal{H} \times \mathcal{F}$ the range is presumably $[0, 1] \times [0, 1]$; no limit curves have been shown to exist so far.

These diagnostic tools were shown to be particularly efficient to distinguish between the deterministic chaotic and stochastic nature of a time series since the permutation quantifiers have distinct behaviours for different types of processes, see Figs. **??** and **??**, respectively.

Chaotic maps have intermediate entropy $\mathcal{H}$ and Fisher $\mathcal{F}$ values, while their complexity $\mathcal{C}$ reaches larger values, very close to those of the upper complexity limit [7,**?**]. For regular processes, entropy and complexity have small

values, close to zero, while the Fisher information is close to one. Uncorrelated stochastic processes are located in the planar location associated with $\mathcal{H}$ near one and $\mathcal{C}$, $\mathcal{F}$ near zero, respectively. It has also been found that correlated stochastic noise processes with a power spectrum proportional to $1/f^k$, where $1 \leq k \leq 3$, are characterized by intermediate permutation entropy and intermediate statistical complexity values [7], as well as intermediate to low Fisher information [**?**]. In both causal information planes ($\mathcal{H} \times \mathcal{C}$, see Fig. **??** and $\mathcal{H} \times \mathcal{F}$, see Fig. **??**), stochastic data are clearly localized at different planar positions than deterministic chaotic ones. These two causal information planes have been profitably used to visualize and characterize different dynamical regimes when the system parameters vary [**?**,**?**,10,**?**,9,**?**,**?**,**?**,**?**,**?**,**?**,**?**]; to study temporal dynamic evolution [**?**,**?**,**?**]; to identify periodicities in natural time series [**?**]; to identify deterministic dynamics contaminated with noise [19,**?**]; to estimate intrinsic time scales and delayed systems [**?**,**?**,**?**]; for the characterization of pseudo-random number generators [**?**,**?**]; to quantify the complexity of two-dimensional patterns [**?**], and for other biomedical and econophysics applications (see [**?**] and references therein).

BBB

In statistical mechanics one is often interested in isolated systems characterized by an initial, arbitrary, and discrete probability distribution. Evolution towards equilibrium is to be described, as the overriding goal. At equilibrium, we can suppose, without loss of generality, that this state is given by the equiprobable distribution $P_e = \{p_i = 1/N, \forall i = 1, \ldots, N\}$. The temporal evolution of the above introduced Information Theory quantifiers, Shannon entropy $\mathcal{H}$, statistical complexity $\mathcal{C}$ and Fisher information measure $\mathcal{F}$, can be analyzed using a two-dimensional (2D) diagrams of the corresponding quantifiers versus time t. However, the second law of thermodynamics states that, for isolated systems, entropy grows monotonically with time ($d\mathcal{H}/dt \geq 0$) [**?**]. This implies that entropy $\mathcal{H}$ can be regarded as an arrow of time, so that an equivalent way to study the temporal evolution of these quantifiers is using the normalized entropy $\mathcal{H}$ as substitute for the time-axis.

Two causal information planes are defined (the term causality remembers the fact that temporal correlations between successive samples are taken into account through the Bandt and Pompe PDF recipe used to estimate both Information Theory quantifiers): *a)* The *causality entropy–complexity plane*, $\mathcal{H} \times \mathcal{C}$, is based only on global characteristics of the associated time series PDF (both quantities are defined in terms of Shannon entropies); while *b)* the *causality Shannon-Fisher plane*, $\mathcal{H} \times \mathcal{F}$, is based on global and local characteristics of the PDF. In the case of $\mathcal{H} \times \mathcal{C}$ the variation range is $[0, 1] \times [\mathcal{C}_{min}, \mathcal{C}_{max}]$ (with $\mathcal{C}_{min}$ and $\mathcal{C}_{max}$ the minimum and maximum statistical complexity values, respectively, for a given $\mathcal{H}$ value [18]), while in the causality plane $\mathcal{H} \times \mathcal{F}$ the range is $[0, 1] \times [0, 1]$.

These two diagnostic tools were shown to be particularly efficient to distinguish between the deterministic chaotic and stochastic nature of a time series since the permutation quantifiers have distinctive behaviors for different types of motion. According to the findings obtained by Rosso *et al.* [7,?,?], chaotic maps have intermediate entropy $\mathcal{H}$ and Fisher $\mathcal{F}$ values, while complexity $\mathcal{C}$ reaches larger values, very close to those of the limit. For regular processes, entropy and complexity have small values, close to zero, while the Fisher is close to one. Finally, totally uncorrelated stochastic processes are located in the planar location associated with $\mathcal{H}$ near one and, $\mathcal{C}$, $\mathcal{F}$ near to zero, respectively. It has also been found that $1/f^{\alpha}$ correlated stochastic processes with $1 \leq \alpha \leq 3$ are characterized by intermediate permutation entropy and intermediate statistical complexity values [7], as well as, intermediate low Fisher information [7,?,?]. Moreover, note that in both causal information planes the localization of these stochastic behavior look like a separation border with respect to chaotic, which are localized above of it. In addition, these two causal information planes have been profitably used to visualization and characterization of different dynamical regimes when the system parameters vary [?,?,10,?,9,?,?,?,?,?,?]; to study time dynamic evolution [?,?,?]; identifying periodicities in natural time series [?]; identification of deterministic dynamics contaminated with noise [19,?]; estimating intrinsic time scales and delayed systems [?,?,?,?]; characterization of pseudo-random number generators [?,?] measure of complexity of two-dimensional patterns [?] among other biomedical and econophysics applications (see [?] and references therein).

YOOOOO

Two representation planes are considered: H_{BP} vs H_{hist} [8] and H_{BP} vs C_{BP} [7]. In the first plane a higher value in any of the entropies, H_{BP} and H_{hist}, implies an increase in the uniformity of the involved PDF. The point $(1, 1)$ represents the ideal case with uniform histogram and uniform distribution of ordering patterns. In the second plane not the entire region $0 < H_{BP} < 1$, $0 < C_{BP} < 1$ is achievable. In fact for any PDF the pairs (H, C) of possible values fall between two extreme curves in the plane H-C [23]. Fig. 1 shows two regions labeled as deterministic and stochastic. In fact transition from one region to the other are smooth and the division is a bit arbitrary. A more detailed discussion can be seen in [7]. Ideal random systems having uniform Bandt & Pompe PDF, are represented by the point $(1, 0)$ [24] and a delta-like PDF corresponds with the point $(0, 0)$.

We also used the number of MP as a quantifier[14]. As shown recently by Amigó *et al.* [25,26,27,28], in the case of deterministic one-dimensional maps, not all the possible ordinal patterns can be effectively materialized into orbits, which in a sense makes these patterns forbidden. Indeed, the existence of these forbidden ordinal patterns becomes a persistent fact that can be regarded as a new dynamical property. Thus, for a fixed pattern-length (embedding

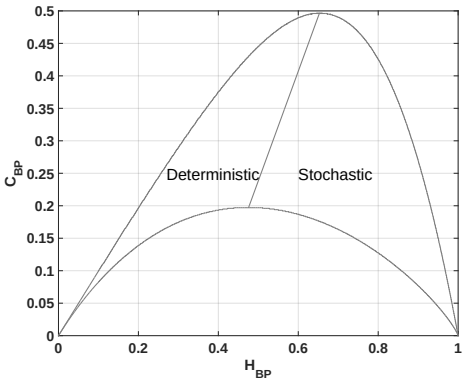

Figure 1. Entropy-Complexity plane.

dimension D) the number of forbidden patterns of a time series (unobserved patterns) is independent of the series length N. Remark that this independence does not characterize other properties of the series such as proximity and correlation, which die out with time [26,28].

A full discussion about the convenience of using these quantifiers is out of the scope of this work. Nevertheless reliable bibliographic sources do exist [29,30,31,8,10,18,14].

3 Results

Five pseudo chaotic maps were studied, two simple maps and three combination of them. For each one floating-point representation and fixed-point numbers with $1 \leq B \leq 53$ representation are considered, where B is the number of bits that represents the fractional part. For each one of these 54 representations 100 time series were generated using randomly chosen initial conditions within the interval $[0, 1]$, that is the attraction domain.

The studied maps are logistic (LOG), tent (TENT), sequential switching between TENT and LOG (SWITCH) and skipping discarding the values in the odd positions (EVEN) or the values in the even positions (ODD) respectively.DEFINIR ACRONIMOS Y EXPLICAR MEJOR.

Logistic map is interesting because is representative of the very large family of quadratic maps. Its expression is:

$$x_{n+1} = 4\, x_n(1 - x_n) \tag{20}$$

with $x_n \in \mathcal{R}$.

Note that to effectively work in a given representation it is necessary to change the expression of the map in order to make all the operations in the chosen representation numbers. For example, in the case of LOG the expression in binary fixed-point numbers is:

$$x_{n+1} = 4\epsilon \,\text{floor}\left\{\frac{x_n(1-x_n)}{\epsilon}\right\} \tag{21}$$

with $\epsilon = 2^B$ where B is the number of bits that represents the fractional part.

The Tent map has been extensively studied in the literature because theoretically it has nice statistical properties that can be analytically obtained. For example it is easy to proof that it has a uniform histogram and consequently an ideal $H_{val} = 1$. The Perron-Frobenius operator and its corresponding eigenvalues and eigenfunctions may be also be analytically obtained for this map [**?**]. PONER ALGUNA REFERNCIA A PERRON Y AGREGAR ALGO EN ITQS

This map is represented with the equation:

$$x_{n+1} \;=\; \begin{cases} 2\;x_n & ,\text{if}\;\; 0 \le x_n \le 1/2 \\ 2\;(1-x_n) & ,\text{if}\;\; 1/2 < x_n \le 1 \end{cases} \tag{22}$$

with $x_n \in \mathcal{R}$.

In base-2 fractional numbers rounding, equation 25 becomes:

$$x_{n+1} \;=\; \begin{cases} 2\;x_n & ,\text{if } 0 \le x_n \le 1/2 \\ \epsilon\;\text{floor}\{\frac{2\;-\;2\;x_n}{\epsilon}\} & ,\text{if } 1/2 < x_n \le 1 \end{cases} \tag{23}$$

with $\epsilon = 2^{-B}$.

Switching, even skipping and odd skipping procedures are shown in Fig. 2.

SWITCH map is expressed as:

$$\begin{cases} x_{n+1} = \begin{cases} 2\,x_n, & ,\text{if } 0 \le x_n \le 1/2 \\ 2\,(1-x_n) & ,\text{if } 1/2 < x_n \le 1 \end{cases} \\ x_{n+2} = 4\,x_{n+1}\,(1-x_{n+1}) \end{cases} \tag{24}$$

with $x_n \in \mathcal{R}$ and n an even number.

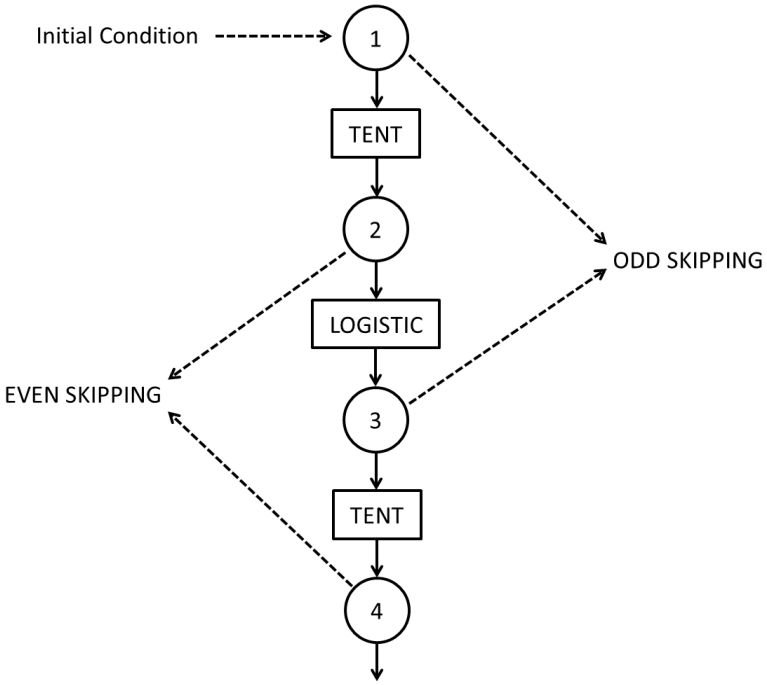

Figure 2. Sequential switching between Tent and Logistic maps. In the figure are also shown even and odd skipping strategies.

Skipping is a usual randomizing technique that increases the mixing quality of a single map and correspondingly increases the value of H_{BP} and decreases C_{BP} of the time series. Skipping does not change the values of H_{val} and C_{val} because their are evaluated using the non causal PDF (normalized histogram of values)[8].

In the case under consideration we study even and odd skipping of the sequential switching of Tent and Logistic maps:

(1) Even skipping of the sequential switching of Tent and Logistic maps (EVEN).
If $\{x_n, \forall\ n = 1, \ldots, \infty\}$ is the time series generated by 27, discard all the values in odd positions and retain the values in even positions.
(2) Odd skipping of the sequential switching of Tent and Logistic maps. If $\{x_n, \forall\ n = 1, \ldots, \infty\}$ is the time series generated by 27, discard all the values in even positions and retain all the values in odd positions.

The reason for studying even and odd skipping cases is the sequential switching map SWITCH is the composition of two different maps. Even skipping may be expressed as the composition function TENT$\circ$LOG while odd skipping may be expressed as LOG$\circ$TENT. The evolution of period as function of precision was reported in [2].

Let us detail our results for each of these maps.

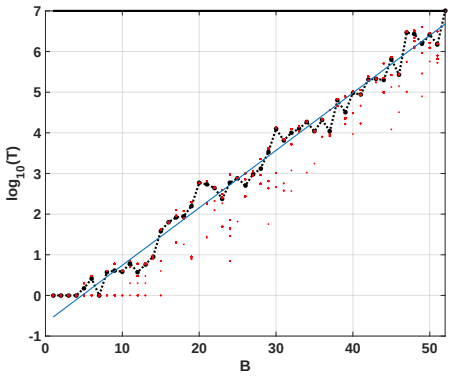

Figure 3. Period as function of precision in binary digits

Table 1
Period T as a function of B for the maps considered

map	m	b
TENT	0	0
LOG	0.139	-0.6188
SWITCH	0.1462	-0.5115
EVEN	0.1447	-0.7783
ODD	0.1444	-0.7683

3.1 Period T as a function of B

Grebogi and coworkers [1] have studied how the period T is related with the precision. There they saw that the period T scales with roundoff ϵ as $T \sim \epsilon^{-d/2}$ where d is the correlation dimension of the chaotic attractor.

Nagaraj et als. [2] studied the case of switching between two maps. They saw that the period T of the compound map obtained by switching between two chaotic maps is higher than the period of each map and they found that a random switching improves the results. Here we have considered sequential switching to avoid the use of another random variable, because it can include its own statistical properties in the time series.

Fig. 3 shows T vs B in semi logarithmic scale. The experimental averaged points can be fitted by a straight line expressed as $log_2 T = mB + b$ where m is the slope and b is the y-intercept. Results for all considered maps are summarized in Table 1.

Results are compatible for those obtained in [2]. Switching between maps

increases de period T but skipping procedure decreases by almost half.

3.2 Quantifiers of simple maps

Here we report our results for both simple maps, LOG and TENT

3.2.1 Logistic map (LOG)

Figs. 4 (a) to (f) show the statistical properties of LOG map in floating-point and fixed-point representation. All these figures show: 100 red points for each fixed-point precision (B) and in black their average (dashed black line connecting black dots), 100 horizontal dashed blue lines that are the results of each run in floating-point and a black solid line their average. In this case, all the lines of the floating-point are overlapped.

According as B grows, statistical properties vary until they stabilize. For $B \geq 30$ the value of H_{val} remains almost identical to the values for the floating-point representation whereas H_{BP} and C_{BP} stabilizes at $B > 21$. Their values are: $\langle H_{val} \rangle = 0.9669$; $\langle H_{BP} \rangle = 0.6269$; $\langle C_{BP} \rangle = 0.4843$. Note that the stable value of missing patterns $MP = 645$ makes the optimum $H_{BP} \leq ln(75)/ln(720) \simeq 0.65$. Then, $B = 30$ is the most convenient choice because an increase in the number of fractional digits does not improve the statistical properties.

Some conclusions can be drawn regarding BP and BPW quantifiers. For $B = 1, 2, 3$ and 4, the averaged BP quantifiers are almost 0 while the averaged BPW quantifiers can not be calculated (see in Figs. 4 c and e the missing black dashed line). For those sequences where the initial condition where 0 all iterations result to be a sequence of zeros (the fixed point of the map), this happens when using small precisions because of the roundoffs.

When B increases, $B = 7, 9$ and 12, the initial conditions are rounded to zero less frequently. So the generated sequences start from some value but many of them fall to zero with a short transitory. This can be seen in Figs.4 c and e, BPW quantifiers show a high dispersion unlike BP quantifiers. This is because BPW procedure takes into account only the transient discarding fixed points, unlike BP procedure considers all the values of the sequence.

The same results are shown in double entropy planes with the precision as parameter (Fig. 5). These figures show: 100 red points for each fixed-point precision (B) and in black their average (dashed black line connecting black dots), 100 blue dots that are the results of each run in floating-point and the black star is their average. In Fig. 5 the 100 blue points and their average are overlapped.

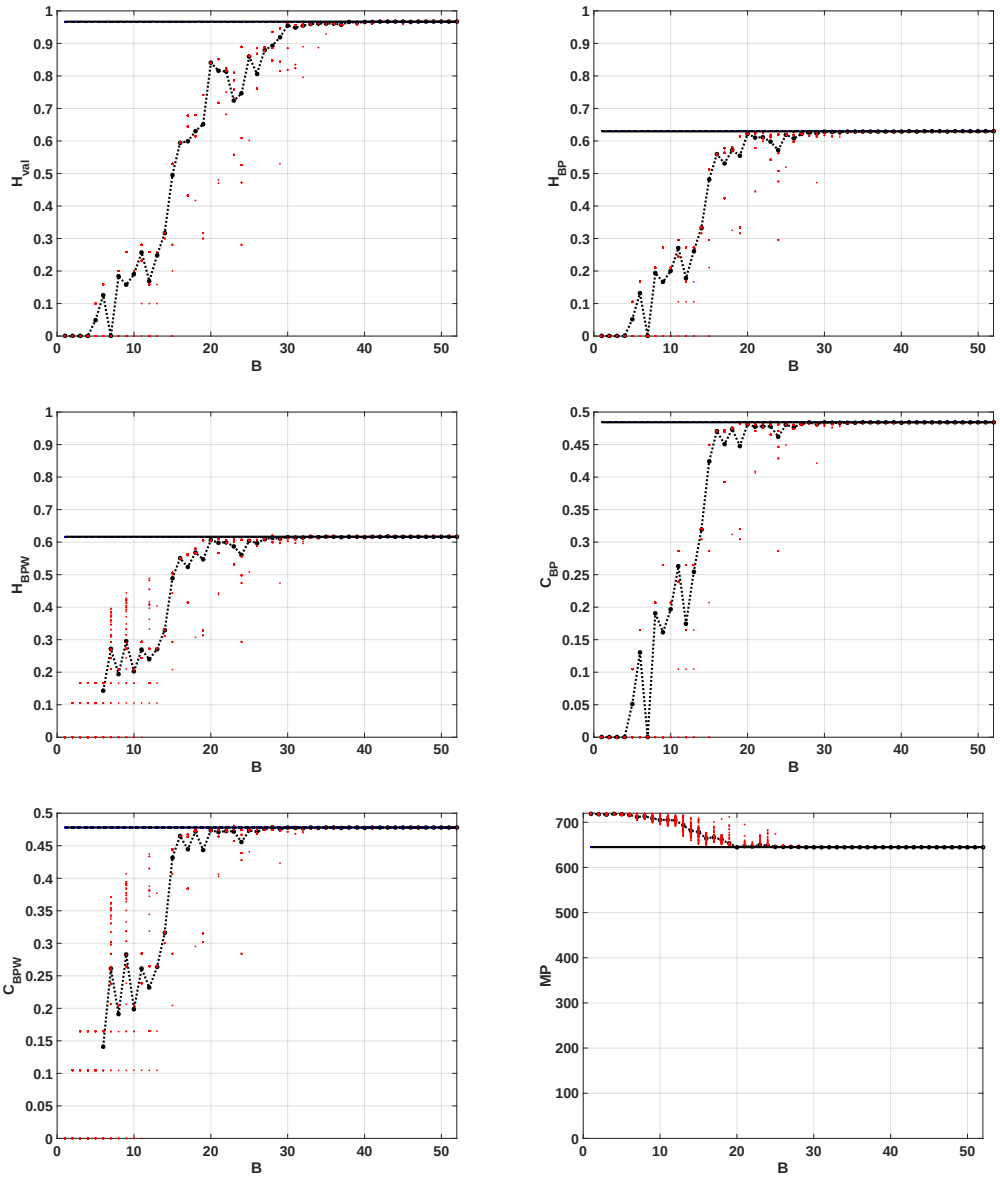

Figure 4. Statistical properties of the LOG map: (a) H_{val} vs B (b) H_{BP} vs B (c) C_{BP} vs B (d) MP vs B.

As expected, the fixed-point architecture implementation converges to the floating-point value as B increases. For both, Hbp-Hval and Hbpw-Hval, from $B = 20$, H_{val} improves but H_{BP} remains constant. It can be seen that the distribution of values reaches high values ($\langle H_{val} \rangle = 0.9669$) but their mixing is poor ($\langle H_{BP} \rangle = 0.6269$).

In Fig. 6 we show the entropy-complexity planes. Dotted gray lines are the upper and lower margins, it is expected that a chaotic system remains near the upper margin. This results characterize a chaotic behaviour, in $H_{BP} - C_{BP}$ plane we can see a low entropy and high complexity.

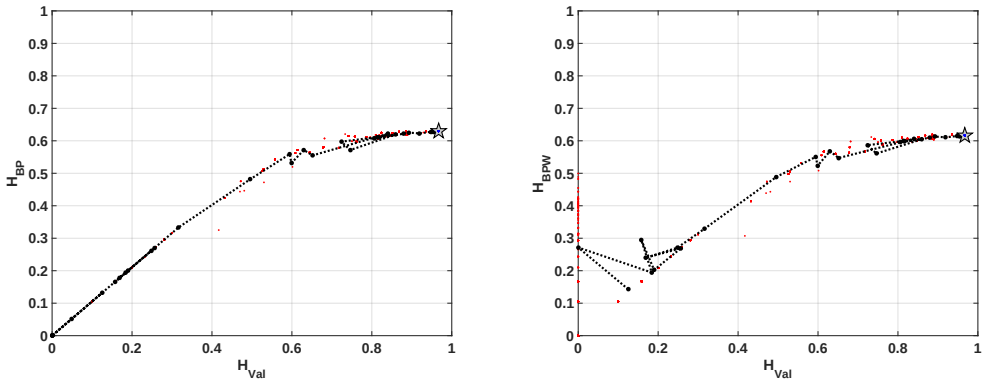

Figure 5. Evolution of statistical properties in double entropy plane of LOG map: (a) H_{val} vs H_{BP} (b) H_{val} vs H_{BPW}.

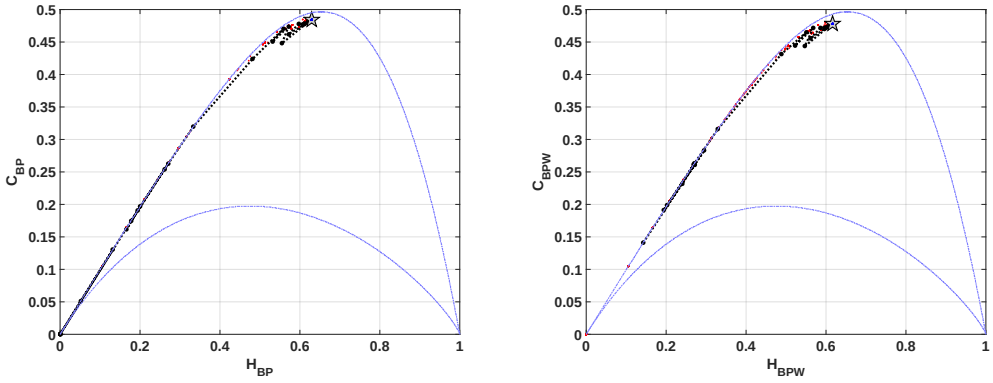

Figure 6. Evolution of statistical properties in entropy-complexity plane of LOG map: (a) C_{BP} vs H_{BP} (b) C_{BPW} vs H_{BPW}.

3.2.2 Tent map (TENT)

When this map is implemented in a computer using any numerical representation system (even floating-point!) truncation errors rapidly increase and make unstable fixed point in $x^* = 0$ to become stable. The sequences within the attractor domain of this fixed point will have a short transitory of length between 0 and B followed by an infinite number of 0's [32,33]. This issue is easily explained in [34], the problem appears because all iterations have a left-shift operation that carries the 0's from the right side of the number to the most significant positions. Some authors [**?**] have proposed to add random perturbations to avoid this drawback of the Tent map. This procedure improves the statistical properties of the time series, but what really happens is that the statistical properties of the random perturbations are mixed with those of the Tent map. Here we study the Tent map "as it is" without any artifact to evaluate its real behavior, instead of theoretical statistical properties.

Figs. 7 (a) to (e) show the quantifiers for floating- and fixed-point numerical representations. Quantifiers H_{val}, H_{BP} and C_{BP} are equal to zero for all

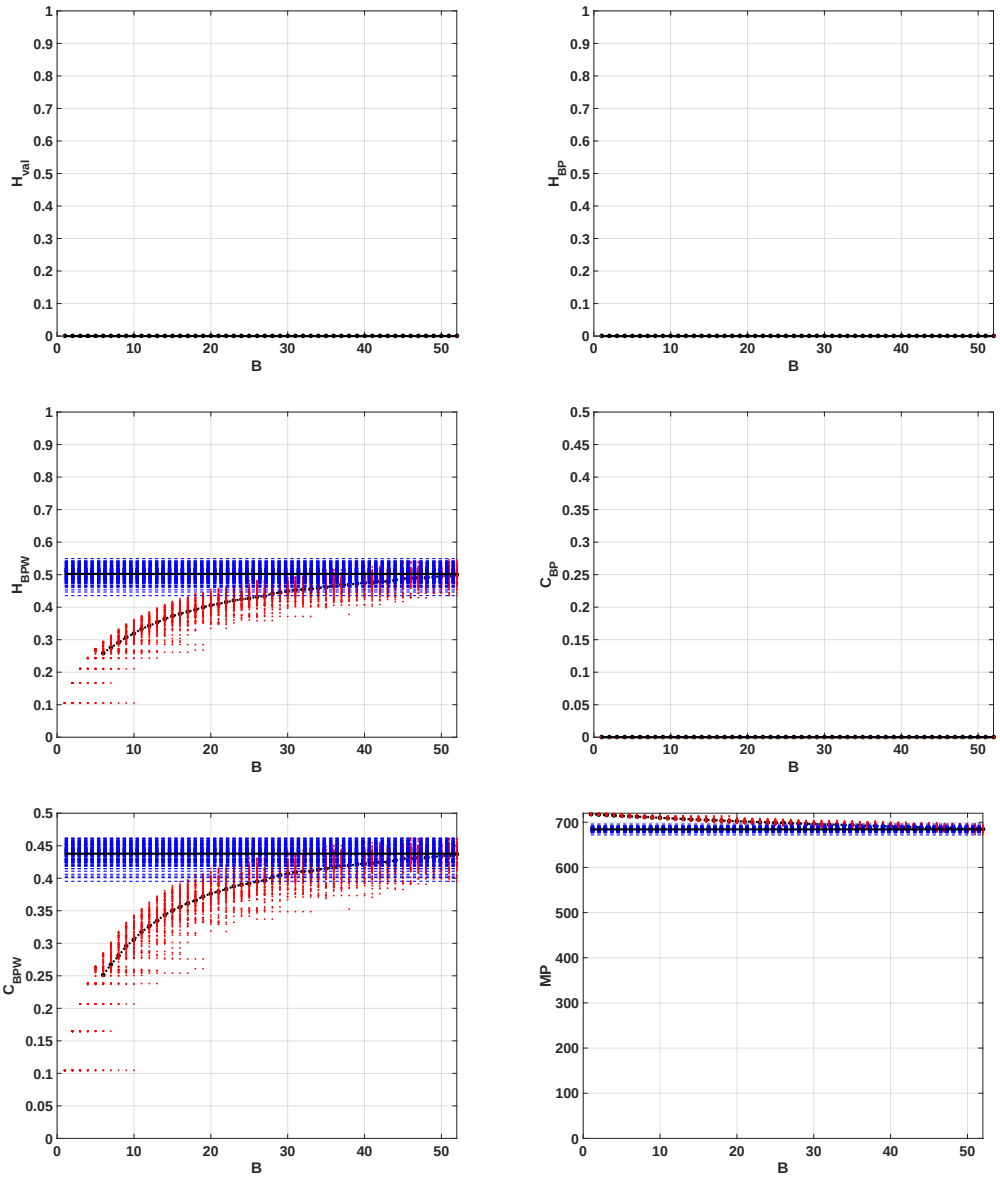

Figure 7. Statistical properties of the TENT map: (a) H_{val} vs B (b) H_{BP} vs B (c) C_{BP} vs B (d) MP vs B.

precisions, this reflects that the series quickly converge toward a fixed point for almost all sequences. In the case of H_{BPW} and C_{BPW} quantifiers they are different from zero because BPW procedure discards the elements once they reach the fixed point. The high dispersions in H_{BPW}, C_{BPW} and MP are related to the short length of transient. These transient that converges to a fixed point has a maximum length of B iterations for fixed-point arithmetic and 54 for floating-point (double precision).

In summary in spite of using a high number of bits (with any 2-based numerical representation) to represent the digitalized TENT map it always loses the

chaotic behaviour.

3.3 Quantifiers of combined Maps

Here we report our results for the three combinations of the simple maps, SWITCH, EVEN and ODD.

3.3.1 Sequential switching between Tent and Logistic maps (SWITCH)

Results with sequential switching are shown in Figs. 8 (a) to (f). The entropy value calculated in floating-point is $\langle H_{val} \rangle = 0.9722$, this value is slightly higher than the one obtained for the LOG map. For fixed-point arithmetic this value is reached in $B = 24$, but it stabilizes from $B = 28$. Regarding the ordering patterns the number of MP decreases to 586, this value is lower than the one obtained for LOG map. It means the entropy H_{BP} may increase up to $ln(134)/ln(720) \simeq 0.74$. BP and BPW quantifiers reach their maximum of $\langle H_{BP} \rangle = 0.6546$ and $\langle H_{BPW} \rangle = 0.6313$ at $B = 16$, but they stabilize from $B = 24$. Complexities are lower than for LOG, $\langle C_{BP} \rangle = 0.4580$ and $\langle C_{BPW} \rangle = 0.4578$, these values are reached for $B \geq 15$ but they are stable from $B \geq 23$. Compared with LOG, statistical properties are better with less amount of bits, for $B \geq 24$ this map reaches optimal characteristics in the sense of random source.

Furthermore, we encountred one initial condition in floating-point long double precision with an anomalous behavior. The quantifiers based on BPW procedure cant detect an anomaly, 8 (a), (b) and (d) an horizontal blue dashed line is far from the average value but this can not be seen in the other figures 8 (c) and (e). We detected a falling to a fixed point after a long transitory, the BPW procedure discards the values corresponding with a fixed point and calculates only the transitory.

Double entropy plane H_{val} vs H_{BP} is showed in Fig. 9. The point reached in this plane for SWITCH map is similar to that reached for LOG map. The mixing is slight better in this case.

Entropy-complexity plane C_{BP} vs H_{BP} is showed in Fig. 10. If we compare with the same plane in the case of LOG (Fig. 6. a.), C_{BP} is lower for SWITCH, this fact shows a more random behaviour.

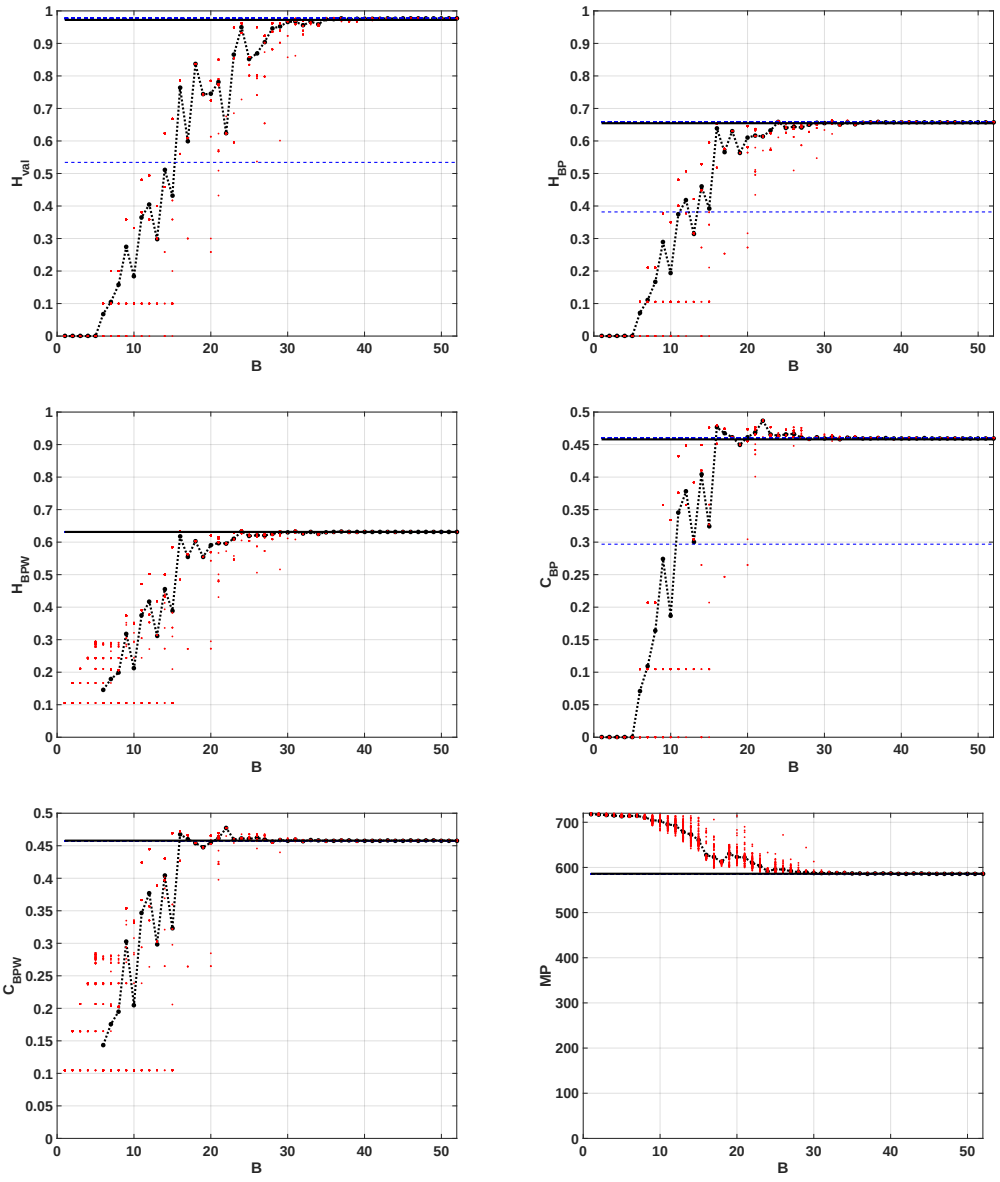

Figure 8. Statistical properties of the SWITCH map: (a) H_{val} vs B (b) H_{BP} vs B (c) C_{BP} vs B (d) MP vs B.

3.3.2 Skipping on sequential switching between Tent and Logistic maps (EVEN and ODD)

In figures 11.a and 12.a, quantifiers related to the normalized histogram of values slightly degrades with the skipping procedure due to finite data length. For example $\langle H_{val} \rangle$ reduces from 0.9722 without skipping to 0.9459 for EVEN and 0.9706 for ODD. This difference between EVEN and ODD in floating point is because a high dispersion was obtained for H_{val}, H_{BP} and C_{BP} but not for H_{BPW} or C_{BPW}.

Figs. 11.b to 11.f. and Figs. 12.b to 12.f show the results of BP and BPW

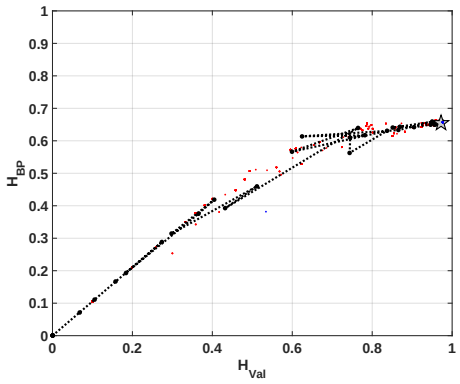

Figure 9. Evolution of statistical properties in double entropy plane of SWITCH map H_{val} vs H_{BP}.

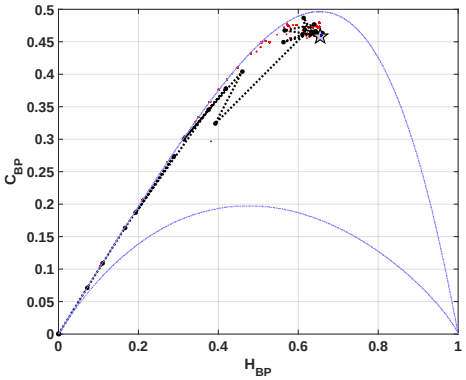

Figure 10. Evolution of statistical properties in entropy-complexity plane of SWITCH map C_{BP}.

quantifiers for EVEN and ODD respectively. Higher accuracy is required to achieve lower complexity than without using skipping. From the point of view of MP a great improvement is obtained using any of the skipping strategies but ODD is slightly better than EVEN. Missing patterns are reduced to $MP = 118$ for EVEN and ODD, increasing the maximum allowed Bandt & Pompe entropy that reaches the mean value $\langle H_{BP} \rangle = 0.8381$ for EVEN, and $\langle H_{BP} \rangle = 0.9094$. The complexity is reduced to $\langle C_{BP} \rangle = 0.224$ for EVEN and $\langle C_{BP} \rangle = 0.282$ for ODD. The amount of bits necessary to converge to this value is $B > 40$ for both EVEN and ODD maps.

FALTA VER PLANOS DOBLE ENTROPÍA

FALTA VER PLANOS ENTROPÍA-COMPLEJIDAD

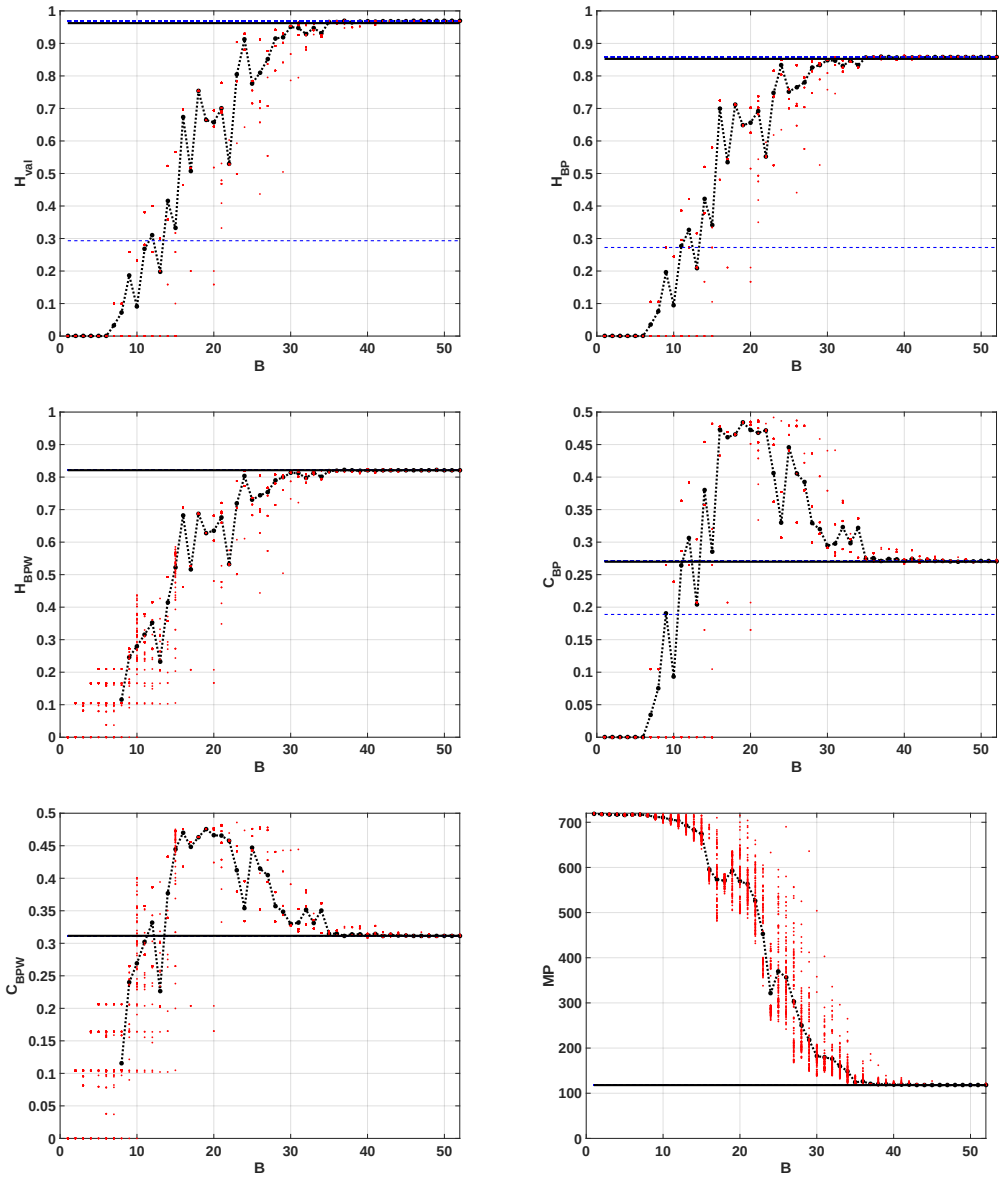

Figure 11. Statistical properties of the EVEN map: (a) H_{val} vs B (b) H_{BP} vs B (c) C_{BP} vs B (d) MP vs B.

4 Conclusions

In summary:

- Not every number base can be represented by a device with a certain base. For example, a base ten number can not be exactly represented in a conventional computer, it will always have an error inherent of the system..No todas las bases numéricas son representables con una máquina de base distinta. Por ejemplo, no se puede representar la base 10 con base 2.
- En una máquina de cálculo "a medida", como la que puede implementarse

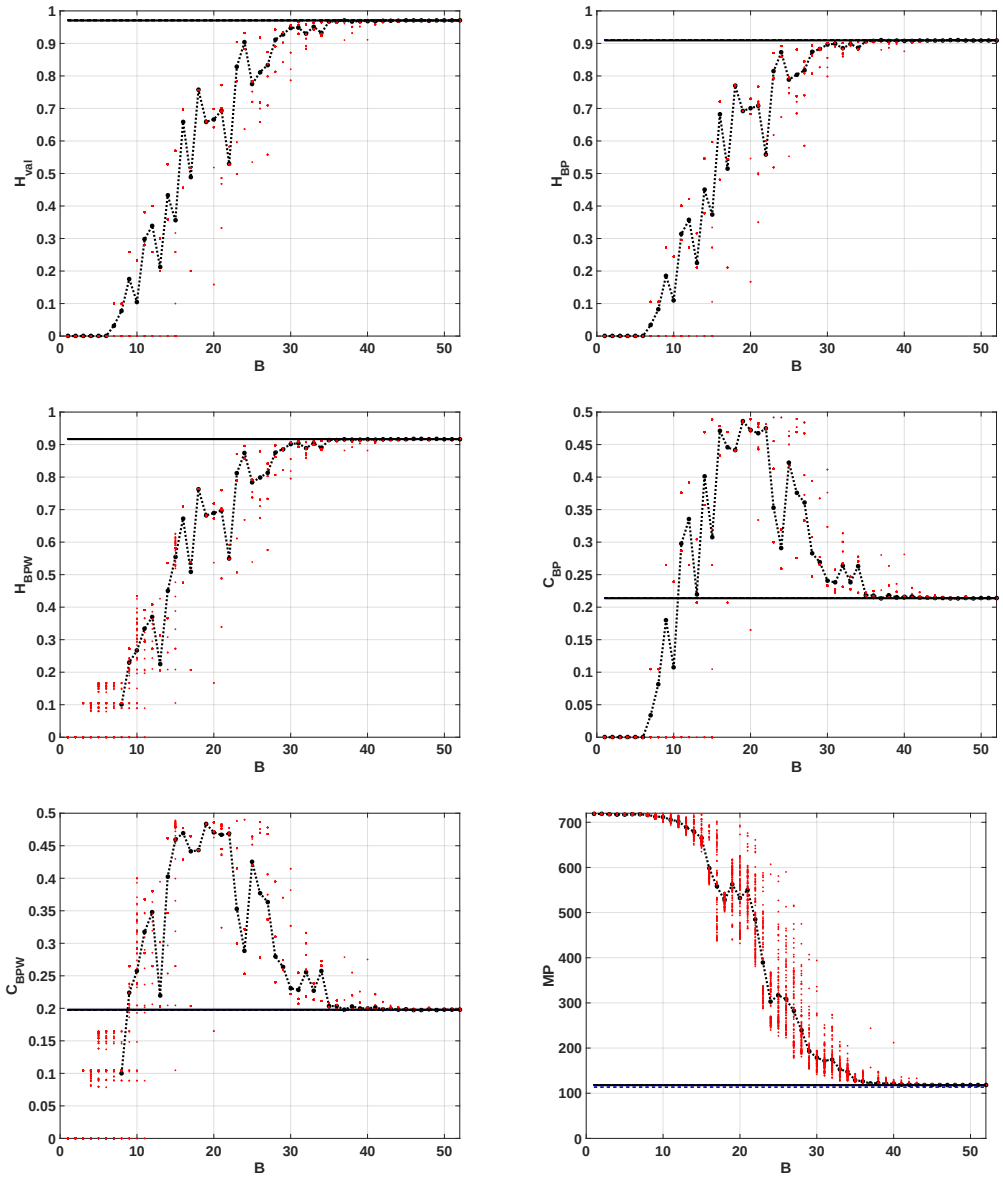

Figure 12. Statistical properties of the ODD map: (a) H_{val} vs B (b) H_{BP} vs B (c) C_{BP} vs B (d) MP vs B.

en ASICs o FPGAs existen limitaciones en el bus de datos y en la electrónica de cálculo. Si la electrónica de cálculo debe ser reducida se recomienda usar mapas que puedan ser calculados sólo con sumas y restas de la variable pseudoaleatoria.

- Los mapas que sólo tienen operaciones de shifteo en la base de la máquina de cálculo inevitablemente caerán a cero en tantas iteraciones como el largo de la mantisa de representación. Por ejemplo el tent en base 2.
- La comparación entre BP y BPW permite detectar el comportamiento del sistema. Puede detectarse si el atractor cae a un punto fijo y diferenciar si el transitorio es corto o largo, respecto de la cantidad de itaraciones del mapa.

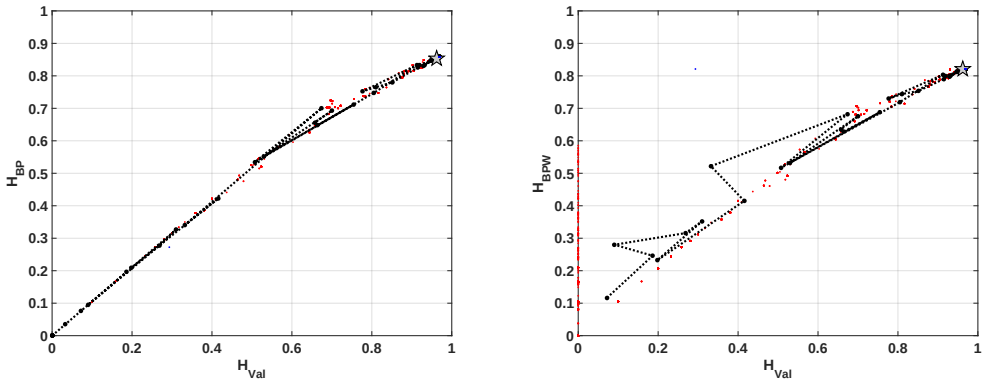

Figure 13. Evolution of statistical properties in double entropy plane of EVEN map: (a) H_{val} vs H_{BP} (b) H_{val} vs H_{BPW}.

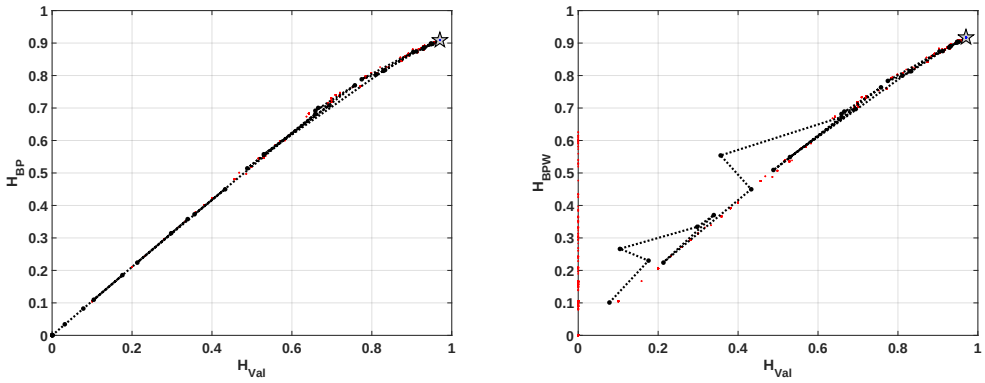

Figure 14. Evolution of statistical properties in double entropy plane of ODD map: (a) H_{val} vs H_{BP} (b) H_{val} vs H_{BPW}.

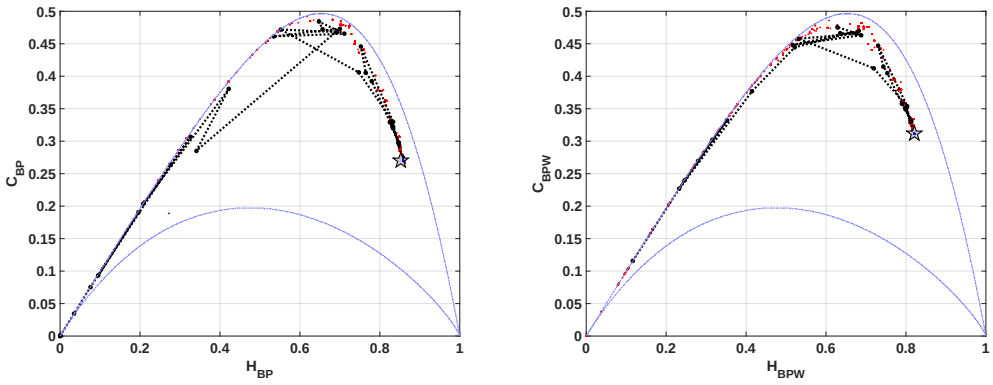

Figure 15. Evolution of statistical properties in entropy-complexity plane of EVEN map: (a) C_{BP} vs H_{BP} (b) C_{BPW} vs H_{BPW}.

- Como se menciona en el paper de referencia, el período del mapa SWITCH aumenta respecto del simple. También se nota una mejora marginal en la mezcla de la secuencia. La distribución de valores es buena en todos los casos.

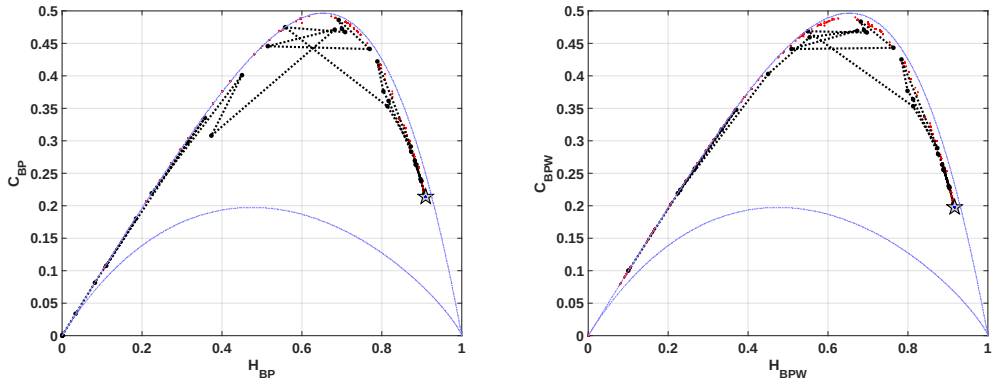

Figure 16. Evolution of statistical properties in entropy-complexity plane of ODD map: (a) C_{BP} vs H_{BP} (b) C_{BPW} vs H_{BPW}.

- el skipping empeora el período pero mejora sustancialmente la mezcla de los valores. Esto puede verse en BP, BPW y MP.

produces a non-monotonous evolution toward the floating point result. This result is relevant because it shows that increasing the precision is not always recommended. It is specially interesting to note that some systems (TENT) with very nice statistical properties in the world of the real numbers, become "pathological" when binary numerical representations are used.

Acknowledgment

This work was partially financed by CONICET (PIP2008), and UNMDP.

References

[1] C. Grebogi, E. Ott, J. A. Yorke, Roundoff-induced periodicity and the correlation dimension of chaotic attractors, Phys. Rev. A 38 (7) (1988) 3688–3692. doi:10.1103/PhysRevA.38.3688.

[2] N. Nagaraj, M. C. Shastry, P. G. Vaidya, Increasing average period lengths by switching of robust chaos maps in finite precision, Eur. Phys. J. Spec. Top. 165 (1) (2008) 73–83. doi:10.1140/epjst/e2008-00850-4.

[3] X. Liu, K.-L. Teo, H. Zhang, G. Chen, Switching control of linear systems for generating chaos, Fractals An Interdiscip. J. Complex Geom. Nat. 30 (3) (2006) 725–733. doi:10.1016/j.chaos.2005.03.020.

[4] E. Gluskin, The nonlinear-by-switching systems (a heuristic discussion of some basic singular systems)arXiv:0801.3652.

[5] D. G. Zarlenga, H. A. Larrondo, C. M. Arizmendi, F. Family, Complex synchronization structure of an overdamped ratchet with discontinuous periodic forcing, Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys. 80 (1) (2009) 011127. doi:10.1103/PhysRevE.80.011127.

[6] L. De Micco, H. A. Larrondo, A. Plastino, O. A. Rosso, Quantifiers for randomness of chaotic pseudo-random number generators., Philos. Trans. A. Math. Phys. Eng. Sci. 367 (1901) (2009) 3281–3296. arXiv:0812.2250, doi:10.1098/rsta.2009.0075.

[7] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, M. A. Fuentes, Distinguishing Noise from Chaos, Phys. Rev. Lett. 99 (15) (2007) 154102. doi:10.1103/PhysRevLett.99.154102.

[8] L. De Micco, C. M. González, H. A. Larrondo, M. T. Martin, A. Plastino, O. A. Rosso, Randomizing nonlinear maps via symbolic dynamics, Phys. A Stat. Mech. its Appl. 387 (14) (2008) 3373–3383. doi:10.1016/j.physa.2008.02.037.

[9] L. De Micco, J. G. Fernández, H. A. Larrondo, A. Plastino, O. A. Rosso, Sampling period, statistical complexity, and chaotic attractors, Phys. A Stat. Mech. its Appl. 391 (8) (2012) 2564–2575. doi:10.1016/j.physa.2011.12.042.

[10] O. A. Rosso, L. De Micco, H. A. Larrondo, M. T. Martin, A. Plastino, Generalized Statistical Complexity Measure, Int. J. Bifurc. Chaos 20 (03) (2010) 775. doi:10.1142/S021812741002606X.

[11] M. Antonelli, L. De Micco, H. A. Larrondo, Measuring the jitter of ring oscillators by means of information theory quantifiers, Commun. Nonlinear Sci. Numer. Simul. 43 (2017) 139–150. doi:10.1016/j.cnsns.2016.05.002.

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

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

[14] O. A. Rosso, L. C. Carpi, P. M. Saco, M. G. Ravetti, H. A. Larrondo, A. Plastino, The Amigo paradigm of forbidden/missing patterns: A detailed analysis, Eur. Phys. J. B 85 (12). doi:10.1140/epjb/e2012-30307-8.

[15] C. E. Shannon, W. Weaver, The mathematical theory of communication, Bell Syst. Tech. J. 27 (3) (1948) 379–423. doi:10.1145/584091.584093.

[16] P. W. Lamberti, M. T. Martin, A. Plastino, O. A. Rosso, Intensive entropic non-triviality measure, Phys. A Stat. Mech. its Appl. 334 (1-2) (2004) 119–131. doi:10.1016/j.physa.2003.11.005.

[17] M. T. Martin, A. Plastino, O. A. Rosso, Statistical complexity and disequilibrium, Phys. Lett. Sect. A Gen. At. Solid State Phys. 311 (2-3) (2003) 126–132. doi:10.1016/S0375-9601(03)00491-2.

[18] M. T. Martin, A. Plastino, O. A. Rosso, Generalized statistical complexity measures: Geometrical and analytical properties, Phys. A Stat. Mech. its Appl. 369 (2) (2006) 439–462. doi:10.1016/j.physa.2005.11.053.

[19] O. A. Rosso, L. C. Carpi, P. M. Saco, M. G??mez Ravetti, A. Plastino, H. A. Larrondo, Causality and the entropy-complexity plane: Robustness and missing ordinal patterns, Phys. A Stat. Mech. its Appl. 391 (1-2) (2012) 42–55. arXiv:1105.4550, doi:10.1016/j.physa.2011.07.030.

[20] K. Keller, H. Lauffer, Symbolic Analysis of High-Dimensional Time Series, Int. J. Bifurc. Chaos 13 (9) (2003) 2657–2668. doi:10.1142/S0218127403008168.

[21] K. Keller, M. Sinn, Ordinal analysis of time series, in: Phys. A Stat. Mech. its Appl., Vol. 356, 2005, pp. 114–120. doi:10.1016/j.physa.2005.05.022.

[22] O. A. Rosso, L. Zunino, D. G. Pérez, A. Figliola, H. A. Larrondo, M. Garavaglia, M. T. Martin, A. Plastino, Extracting features of Gaussian self-similar stochastic processes via the Bandt-Pompe approach, Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys. 76 (6) (2007) 061114. doi:10.1103/PhysRevE.76.061114.

[23] C. Anteneodo, A. R. Plastino, Some features of the López-Ruiz-Mancini-Calbet (LMC) statistical measure of complexity, Phys. Lett. Sect. A Gen. At. Solid State Phys. 223 (5) (1996) 348–354. doi:10.1016/S0375-9601(96)00756-6.

[24] C. M. González, H. A. Larrondo, O. A. Rosso, Statistical complexity measure of pseudorandom bit generators, Phys. A Stat. Mech. its Appl. 354 (1-4) (2005) 281–300. doi:10.1016/j.physa.2005.02.054.

[25] J. M. Amigó, L. Kocarev, J. Szczepanski, Order patterns and chaos, Phys. Lett. Sect. A Gen. At. Solid State Phys. 355 (1) (2006) 27–31. doi:10.1016/j.physleta.2006.01.093.

[26] J. M. Amigó, L. Kocarev, I. Tomovski, Discrete entropy, Phys. D Nonlinear Phenom. 228 (1) (2007) 77–85. doi:10.1016/j.physd.2007.03.001.

[27] J. M. Amigó, S. Zambrano, M. A. F. Sanjuán, Combinatorial detection of determinism in noisy time series, Europhys. Lett. 83 (6) (2008) 60005. doi:10.1209/0295-5075/83/60005.

[28] J. Amigó, Permutation Complexity in Dynamical Systems, Springer Series in Synergetics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. arXiv:arXiv:1011.1669v3, doi:10.1007/978-3-642-04084-9.

[29] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, H. Scheingraber, A comparative classification of complexity measures, Chaos, Solitons and Fractals 4 (1) (1994) 133–173. doi:10.1016/0960-0779(94)90023-X.

[30] R. López-Ruiz, H. L. Mancini, X. Calbet, R. L6pez-Ruiz, H. L. Mancini, X. Calbet ', A statistical measure of complexity, Phys. Lett. A 209 (5-6) (1995) 321–326. doi:10.1016/0375-9601(95)00867-5.

[31] O. A. Rosso, H. A. Larrondo, M. T. Martin, A. Plastino, M. A. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (15) (2007) 154102. arXiv:1401.2139, doi:10.1103/PhysRevLett.99.154102.

[32] M. Jessa, The period of sequences generated by tent-like maps, IEEE Trans. Circuits Syst. I Fundam. Theory Appl. 49 (1) (2002) 84–89. doi:10.1109/81.974880.

[33] S. Callegari, G. Setti, P. Langlois, A CMOS tailed tent map for the generation of uniformly distributed chaotic sequences, in: Proc. 1997 IEEE Int. Symp. Circuits Syst. Circuits Syst. Inf. Age ISCAS '97, Vol. 2, IEEE, pp. 781–784. doi:10.1109/ISCAS.1997.621829.

[34] S. Li, When Chaos Meets Computers (2004). arXiv:0405038.