Physica A 391 (2012) 2518–2526

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Ambiguities in Bandt–Pompe's methodology for local entropic quantifiers

Felipe Olivares [a], Angelo Plastino [b,e], Osvaldo A. Rosso [c,d,e,*]

[a] Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), C.C. 67, 1900 La Plata, Argentina

[b] Instituto de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), C.C. 727, 1900 La Plata, Argentina

[c] Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais (UFMG), Av. Antônio Carlos, 6627 - Campus Pampulha, 31270-901 Belo Horizonte - MG, Brazil

[d] Chaos & Biology Group, Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (UBA), Pabellón II, Ciudad Universitaria, 1428 Ciudad de Buenos Aires, Argentina

[e] Fellow of CONICET, Argentina

ARTICLE INFO

Article history:
Received 20 October 2011
Received in revised form 17 December 2011
Available online 27 December 2011

Keywords:
Shannon entropy
Fisher information measure
Bandt–Pompe probability distribution
Nonlinear time series analysis

ABSTRACT

The Bandt–Pompe (BP) prescription for building up probability densities [C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102] constituted a significant advance in the treatment of time-series. However, as we show here, 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 that would characterize time-series generated by nonlinear dynamical systems. Explicit evidence of this fact is presented by comparing two different procedures, frequently found in the literature, that generate sequences of ordinal patterns. In opposition to the case of *global* quantifiers in the orthodox Shannon fashion, the proper order of the pertinent symbols turns out to be not uniquely predetermined for local entropic indicators. We advance the idea of employing the Fisher–Shannon information plane as a tool to resolve the ambiguity and give illustrative examples.

1. Introduction

The Bandt–Pompe (BP) proposal for associating probability distributions to time series (of an underlying symbolic nature), constitutes a significant advance in the study of non linear dynamical systems [1]. The method provides univocal prescription for ordinary, global entropic quantifiers of the Shannon-kind.

We show here that for local quantifiers [2] some ambiguity arises that deserves careful analysis. The issue is of some importance from the information-theoretic viewpoint because the formalisms of (a) electromagnetism, (b) classical physics, (c) quantum mechanics, (d) general relativity, and many others, can be re-derived using local information quantifiers, of which Fisher's information measure is the paradigmatic example (see, for instance, Ref. [3] and references therein). Consequently, there is ample motivation for re-discussing in this light the BP approach: the opening of doors to its utilization in a Fisher-treatment of non linear physics time series.

* Corresponding author at: Departamento de Física, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais (UFMG), Av. Antônio Carlos, 6627 - Campus Pampulha, 31270-901 Belo Horizonte - MG, Brazil.

E-mail addresses: folivares@fisica.unlp.edu.ar (F. Olivares), plastino@fisica.unlp.edu.ar (A. Plastino), oarosso@fibertel.com.ar, orosso@gmail.com (O.A. Rosso).

0378-4371/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2011.12.033

1.1. Time series and their associated probability distributions

For time series $\mathcal{S}(t) \equiv \{x_t\}$ characterization using information theory quantifiers, a probability distribution function (PDF) $P \equiv \{p_i\}$ associated to the time series under analysis should be provided beforehand. The determination of the most adequate PDF is a fundamental problem because P and the sample space Ω are inextricably linked. Many methods have been proposed for a proper selection of the probability space (Ω, P). We can mention: (a) frequency counting [4], (b) procedures based on amplitude statistics [5], (c) binary symbolic dynamics [6], (d) Fourier analysis [7], (e) wavelet transform [8], and (f) permutation ordinal patterns [1], among others.

Their applicability depends on particular characteristics of the data, such as stationarity, time series length, variation of the parameters, level of noise contamination, etc. In all these cases the dynamics' global aspects can be somehow captured, but the different approaches *are not equivalent in their ability to discern all the relevant physical details*. One must also acknowledge the fact that the above techniques, (a)–(e), are introduced in a rather "ad hoc fashion". They *are not directly derived from the dynamical properties themselves of the system under study*.

This is not the case of the permutation ordinal patterns proposed by Bandt and Pompe for generating PDF [1] (for details of the Bandt and Pome approach, see Appendix). It is one of the most simple symbolization techniques and takes into account time-causality in the evaluation of the PDF associated to a time series. In particular, Rosso et al. [9] showed that the Bandt and Pompe methodology may be profitably used in the causality entropy-complexity plane $(H \times C)$ so as to separate and differentiate amongst stochastic, chaotic, and deterministic systems. It was shown in Refs. [10,11] that temporal correlations are nicely displayed by the Bandt and Pompe PDF.

2. Shannon entropy & Fisher information measure

Given a continuous probability distribution function (PDF) $f(x)$, its *Shannon entropy* S is [12]

$$S[f] = -\int f \ln(f)\, \mathrm{d}x, \tag{1}$$

a measure of "global character" that it is not too sensitive to strong changes in the distribution taking place on a small-sized region.

Such is not the case with *Fisher's Information Measure* (FIM) F [3,13], which constitutes a measure of the gradient content of the distribution f, thus being quite sensitive even to tiny localized perturbations. It reads [3]

$$F[f] = \int \frac{|\vec{\nabla} f|^2}{f}\, \mathrm{d}x. \tag{2}$$

The above is not the most general Fisher-definition, but in physical applications it is the one often employed [3]. FIM can be variously interpreted as a measure of the ability to estimate a parameter, as the amount of information that can be extracted from a set of measurements, and also as a measure of the state of disorder of a system or phenomenon [3,14].

We emphasize that the gradient operator significantly influences the contribution of minute local f-variations to FIM's value, so that the quantifier is called a "local" one. Note that Shannon's entropy decreases with skewed distributions, while Fisher's information increases in such a case. Local sensitivity is useful in scenarios whose description necessitates appeal to a notion of "order" (see below).

Now, let $P = \{p_i; i = 1, \ldots, N\}$ be a discrete probability distribution set, with N the number of possible states of the system under study. The concomitant problem of loss of information due to the discretization has been thoroughly studied (see, for instance, Refs. [15–17] and references therein) and, in particular, it entails the loss of FIM's shift-invariance, which is of no importance for our present purposes. In the discrete case, Shannon's quantifier is evaluated via

$$S[P] = -\sum_{i=1}^{N} p_i \ln(p_i), \tag{3}$$

and we define a "normalized" Shannon entropy as $H[P] = S[P]/S_{\max}$, where the denominator is obtained for a uniform probability distribution.

For the FIM-computation, we follow the proposal of Ferri and coworkers [18,19] (among others)

$$F[P] = \frac{1}{4} \sum_{i=1}^{N-1} 2 \frac{(p_{i+1} - p_i)^2}{(p_{i+1} + p_i)}. \tag{4}$$

Shannon's entropy $S[P]$, as well as, FIM's $F[P]$, can be thought of as a measure of the degree of order/disorder of a signal, so it can provide useful information about the underlying dynamical process associated with a signal. One can state that the general behavior of the Fisher information measure is opposite to that of the Shannon entropy [20].

The local sensitivity of FIM for discrete-PDFs is reflected in the fact that the specific i-"ordering" of the discrete values p_i must be carefully examined in evaluating the sum in Eq. (4). The pertinent numerator can be regarded as a kind of "distance" between two contiguous probabilities. Thus, a different ordering of the pertinent summands would lead to a different FIM-value. In fact, if we have a discrete PDF given by $P = \{p_i, i = 1, \ldots, N\}$ we will have $N!$ possibilities.

The question is, which is the one that could be regarded 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 values in the time series) into a finite number N_{bin} ($N \equiv N_{bin}$ in Eqs. (3) and (4)) of non overlapping equal sized consecutive subintervals $A_i : [a, b] = \bigcup_{i=1}^{N_{bin}} A_i$ and $A_i \bigcap A_j = \emptyset \forall i \neq j$. 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.

Consider, in example, the time series generated by the logistic map with $r = 4$, whose associated dynamics is the totally developed chaos [21]. It is well known that in this situation the logistic map exhibits an almost flat PDF-histogram (histogram on the amplitude values in the interval [0, 1]), with peaks at $x = 0$ and $x = 1$. This PDF constitutes an invariant measure of the system [21]. In consequence, if we use this PDF we obtain $S[P] \cong S_{\max}$ and $F[P] \cong 0$, which makes the logistic map almost indistinguishable from a pure-noise signal (uncorrelated random process), although one certainly knows that *chaos is not noise*. Thus, if one want to use quantifiers based on information theory with the purpose of characterizing and distinguishing deterministic signals (chaos) from noise (random process) one should demand for some improvements in the methodology used for associating a PDF to a time series (extracted from a dynamical system). This can be achieved if the time causality (in the series' values) is duly taken into account for generating the pertinent PDF, something that one gets automatically from the Bandt–Pompe methodology.

2.1. The problem to be addressed here

The Bandt–Pompe procedure does not specify which order for generating ordinal patterns should be privileged. That is, how to specify the ordering of index series $\{i\}$ and, consequently, for a pattern length D, we have $D!$ possible i-sequences. Therefore, we face an ambiguity when local information measures are evaluated. The issue does not affect at all the evaluation of global entropic quantifiers (like Shannon entropy). Different orders would lead, though, to different "local" information contents. *This is the crux of the problem to be addressed in this communication.*

As in the case of the PDF-histogram, we can reduce drastically the number of available $D!$ possibilities if we proceed to form patterns of length D starting from those of length $D - 1$ (see Appendix). However, some lack of precise definition remains in the assignation of the pattern indices $\{i\}$. We will analyze serious consequences that arise out of such ambiguity by considering two different manners of (i) sequentially generating ordinal patterns and (ii) deciding just how to assign our pattern of indices $\{i\}$. The two procedures we use are frequently found in the literature:

- the algorithm described by Keller and Sinn [22].
- a different type of implementation, called the Lehmer one [23].

The differences that arise out of following one or the other algorithm will be illustrated with reference to the pattern of indices resulting in a scenario for which the embedding dimensions are $D = 2$, 3, or 4, as displayed in Table 1. We see that the sequences of all possible ordinal patterns are different. This affects the information-content FIM measured, as evidenced by its application to the dynamics of the logistic map and its variation with the parameter r.

2.2. Using the Fisher–Shannon information plane

The Fisher–Shannon information plane was introduced by Vignat and Bercher in Ref. [24]. These authors showed that the simultaneous examination of both Shannon's entropy and Fisher's information is required to characterize the non-stationary behavior of a complex signal. Also, Vignat and Bercher demonstrated that scaling and uncertainly properties, together, highlight the fact that FIM and Shannon entropy are intrinsically linked (see Ref. [25] for explicit relationships), so that the characterization of signals should be improved when considering their localization in the Fisher–Shannon plane. Accordingly, it is plausible to expect that the Fisher–Shannon information plane may help to get insights concerning the evaluation of PDFs based on Bandt and Pompe's time-series procedure. The best sequence for ordinal-patterns' generation will be the one that, as a function of the dynamical system's parameters, produces the best planar-separation between the different dynamics regimes.

3. Logistic map application

The logistic map constitutes a paradigmatic example, often employed as a testing-ground in order to illustrate new concepts in the treatment of dynamical systems. Thus, it is almost obligatory to test our BP ideas in such scenario. It is well-known that the logistic map is a polynomial mapping of degree 2, $\mathcal{F} : x_n \to x_{n+1}$ [21], described by the ecologically motivated, dissipative system described by the first order difference equation

$$x_{n+1} = r \cdot x_n \cdot (1 - x_n), \tag{5}$$

Table 1
Sequence-patterns for both Keller's and Lehmer's generating algorithms. The pattern lengths (embedding dimensions) are, respectively, $D = 2$, 3, and 4.

D = 2			**D** = 3			**D** = 4		
{*i*}	*Keller*	*Lehmer*	{*i*}	*Keller*	*Lehmer*	{*i*}	*Keller*	*Lehmer*
1	{01}	{01}	1	{012}	{012}	1	{0123}	{0123}
						2	{0132}	{0132}
						3	{0312}	{0213}
						4	{3012}	{0231}
			2	{021}	{021}	5	{0213}	{0312}
						6	{0231}	{0321}
						7	{0321}	{1023}
						8	{3021}	{1032}
			3	{201}	{102}	9	{2013}	{1203}
						10	{2031}	{1230}
						11	{2301}	{1302}
						12	{3201}	{1320}
2	{10}	{10}	4	{102}	{120}	13	{1023}	{2013}
						14	{1032}	{2031}
						15	{1302}	{2103}
						16	{3102}	{2130}
			5	{120}	{201}	17	{1203}	{2301}
						18	{1230}	{2310}
						19	{1320}	{3012}
						20	{3120}	{3021}
			6	{210}	{210}	21	{2103}	{3102}
						22	{2130}	{3120}
						23	{2310}	{3201}
						24	{3210}	{3210}

with $0 \leq x_n \leq 1$ and $0 \leq r \leq 4$. We scrutinize the logistic map dynamics for the parameter range $3.8 \leq r \leq 3.87$, which includes a mixture of order and chaos (the corresponding bifurcation diagram is presented in Fig. 1(a)) [18,19].

A period-three attractor arises through a saddle–node bifurcation at $r_1 \cong 3.82842$ (tangent bifurcation) till $r_2 \cong 3.8415$ (flip bifurcation). The chaotic dynamics that prevails before reaching r_1 is called *Chaos* 1. As r grows beyond r_2, the period-three solutions experience a new sequence of period-doubling bifurcations that ends in a totally chaotic dynamics at $r_3 \cong 3.84943$. The chaotic attractor consists of three narrow disjoint segments with several periodic windows, referred to as *Chaos* 2 *with periodic window*. At $r_4 \cong 3.85681$ (interior crisis) this chaotic attractor is again replaced by another one designed as *Chaos* 3 which "lives" in a wider region that includes the three sectors of the previous attractor.

The FIM-behavior, evaluated with both Keller and Lehmer's sequential order, in the parameter range $3.8 \leq r \leq 3.87$, is depicted in Fig. 1(b). Globally, one finds a good characterization of the different dynamics. Differences arise, though, in the numerical values that emerge in the distinct chaotic zones. For $r < r_1$ we see a smooth increasing behavior until $r \sim r_1$, where one finds for either F_{Keller} and F_{Lehmer} a smooth transition between Chaos 1 and period 3 regimes. For $r_2 \leq r \leq r_3$ a period-doubling region occurs, where $F_{Keller} = F_{Lehmer} = 1$. At the parameter range $r_3 \leq r \leq r_4$ another remarkable feature emerges, a "band splitting" phenomenon. The iterates alternate between the three bands in periodic fashion. However, in the interior of each band motion becomes chaotic. The behavior of both measure is similar. At $r > r_4$ the Chaos 3-zone arises and we encounter a clear difference between the results from the two ways of computing the information measures. While F_{keller} tends to reach values similar to those of Chaos 1, F_{Lehmer} descends to considerably lower values than those of Chaos 1.

In Fig. 2, we display the plane $F_{Keller} \times F_{Lehmer}$ for the parameter range $3.8 \leq r \leq 3.87$ (the control parameter does not explicitly appear in the graph, of course). The continuous line in the graph represents $F_{Keller} = F_{Lehmer}$. We see that both measures' results coincide for the periodic regime and for the Chaos 2 with periodic window zone, as above. They do differ in the remaining chaotic zones.

The Fisher–Shannon information plane [24], in which the Normalized Shannon entropy is used ($H[P] = S[P]/S_{\max}$) is depicted in Fig. 3. From this plot we see that, in the parameter range variation $3.8 \leq r \leq 3.87$, the results based on the Lehmer-sequential algorithm for the generation of ordinal patterns ($D = 6$ and $T = 1$) produce the better differentiation (planar localization without overlapping) between the different dynamical behaviors associated to the values of r.

The Fisher–Shannon planar behavior-characterization based on Lehmer's sequence (see Fig. 3(b)) reveals fine details concerning the intermittency-regime in the chaotic zones named *Chaos* 1 and *Chaos* 3. A feature of the intermittency is the *convergent* trajectory-clustering to the left of a tangent bifurcation, which gives rise to a period-3 window at r_1. After the interior crisis at r_4 one detects least 8 *divergent* clusters of trajectories leading to a totally developed chaos. We thus appreciate intermittency-differences in a chaotic dynamics that the planar representation recognizes via their respective locations, signaling the existence of two dynamical regimes without overlapping, a perhaps surprising result.

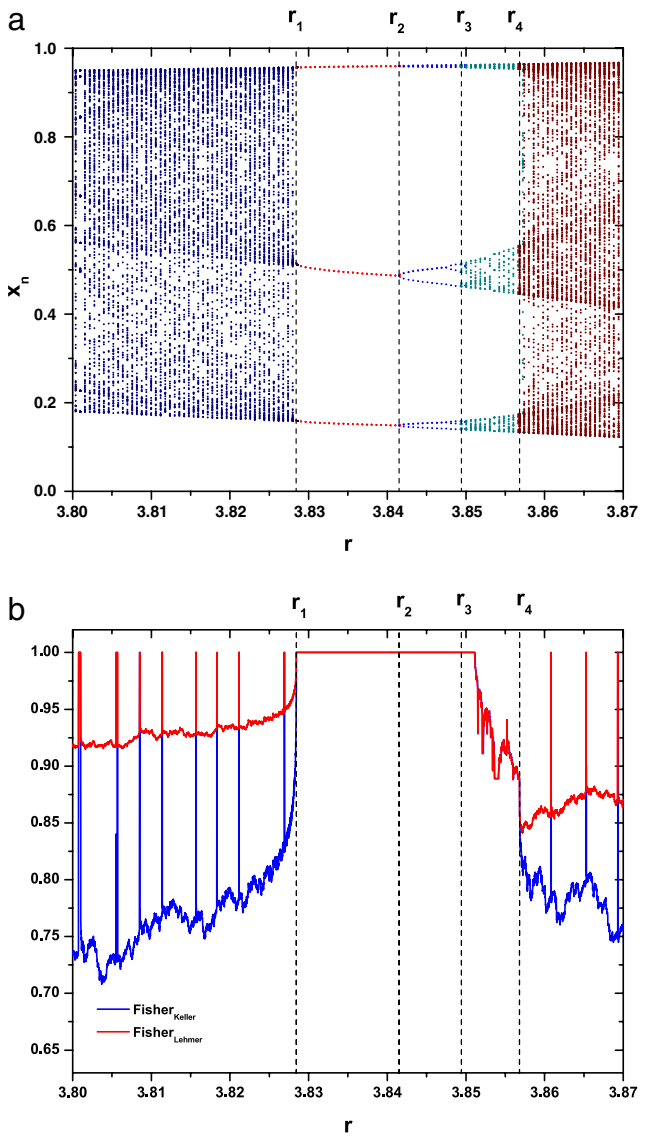

Fig. 1. Results for the logistic map, as a function of the parameter $3.8 \leq r \leq 3.87$. (a) Bifurcation diagram ($\Delta r = 0.0006$). (b) Fisher information measure evaluated with the Keller and Lehmer pattern-sequence ($\Delta r = 1 \times 10^{-5}$). In the logistic time series-generation we discard the first 10^5 iterations and then $N = 2 \times 10^6$ data time series are generated. In building up Bandt and Pompe' PDF we consider $D = 6$ and $T = 1$. The vertical lines represent the different dynamical windows described in the text.

4. Conclusions

We have conclusively shown that some ambiguity arises in trying to apply the Bandt–Pompe methodology to the construction of local entropic quantifiers. This happens when these indicators are associated to time series generated by a nonlinear dynamical systems. We have given explicit evidence of this fact with reference to Fisher's information measure. The order of the pertinent symbols turns out to be not uniquely predetermined.

We analyzed the results obtained when two usual algorithms for sequential, ordinal pattern-generation are employed. In order to decide which one is the best, we advanced using the localization-behavior of the local quantifiers in the Fisher–Shannon information plane. The best way of ordering our symbols should be that which produces the most clear distinction between different nonlinear dynamic regimes. We illustrated this assertion with reference to the logistic map.

Acknowledgments

F. Olivares is supported by a Fellowship of the Chilean Government, CONICYT. O.A. Rosso gratefully acknowledge support from CAPES, PVE fellowship, Brazil.

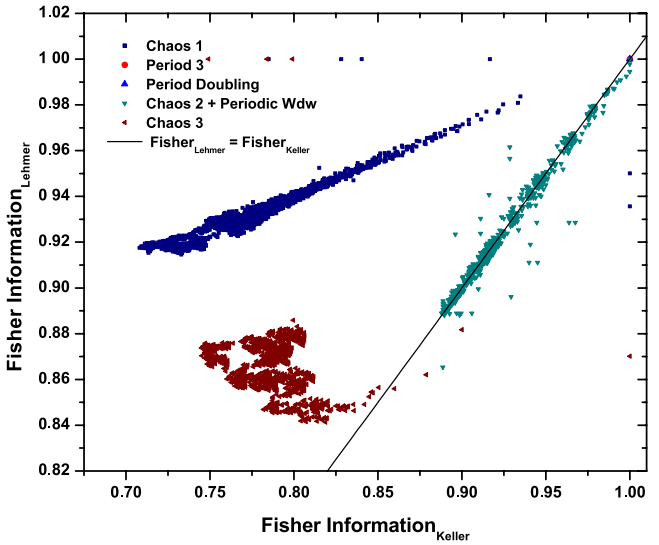

Fig. 2. Fisher Keller–Fisher Lehmer plane for the logistic map. The parameter range is $3.8 \leq r \leq 3.87$. For the Bandt and Pompe's PDF we consider $D = 6$ and $T = 1$.

Appendix

We present here some technical details that may help understanding the text but can be omitted at a first reading of the manuscript.

A.1. Dealing with symbolic sequences

Different *symbolic sequences* may be assigned to a given time series [26]. In this respect, an issue of some importance is that of ascertaining whether the order in which the distinct time-series' values emerge is taken into account or not. In the first case one says that causal information has been taken into account [10,11]. This is not the case, obviously, for the second instance above. If one merely assigns a symbol a of the finite alphabet $\mathfrak{A}$ to each x_t of the time series, the ensuing *symbolic sequence* can be regarded as a *non causal coarse grained* description of the time series under consideration. The PDF "extracted" from the time-series will not have here any *causal information*. The usual histogram-technique corresponds to this kind of assignment.

Causal information may be duly incorporated into the construction-process that yields P if one symbol of a finite alphabet $\mathfrak{A}$ is assigned instead *to a (phase-space) trajectory's portion*, i.e., we assign "words" to each trajectory-portion. The Bandt and Pompe (BP) methodology for extracting a PDF from a time series corresponds to the causal type of assignment and the resulting probability distribution P constitutes thus a *causal coarse grained* description of the system (for methodological details see below).

The advantage of the Bandt and Pompe approach lies in the fact that it solves the problem of finding time-series' generating partitions. Thus one expect that for increasing patterns' length (embedding dimension) the BP approach retains all relevant essentials of the original continuous (in phase-space) dynamics. We consider that the Bandt and Pompe approach for generating PDF's is one of the simplest symbolization techniques which incorporates causality in the evaluation of the PDF associated to a time series. Its use has been shown to yield a clear improvement on the quality of information theory-based quantifiers (see i.e. Refs. [9,27–29] and reference therein).

A.2. The Bandt and Pompe approach to building up PDF

Which is the appropriate probability distribution function (PDF) that one is to associate to a given time-series? Bandt and Pompe (BP) [1] answered this question by introducing a simple and robust symbolic method that takes notice of time causality in dealing with a given system's dynamics. The appropriate symbol sequence arises naturally from the time series. No model-based assumptions are needed. "Partitions" are devised by comparing the order of neighboring relative values rather than by apportioning amplitudes according to different levels.

Given a time series $\mathcal{S}(t) = \{x_t; t = 1, \ldots, M\}$, an embedding dimension $D > 1$ ($D \in \mathbb{N}$), and an embedding delay T ($T \in \mathbb{N}$), the BP-pattern of order D generated by

$$s \mapsto \left(x_{s-(D-1)T}, x_{s-(D-2)T}, \ldots, x_{s-T}, x_s\right), \tag{6}$$

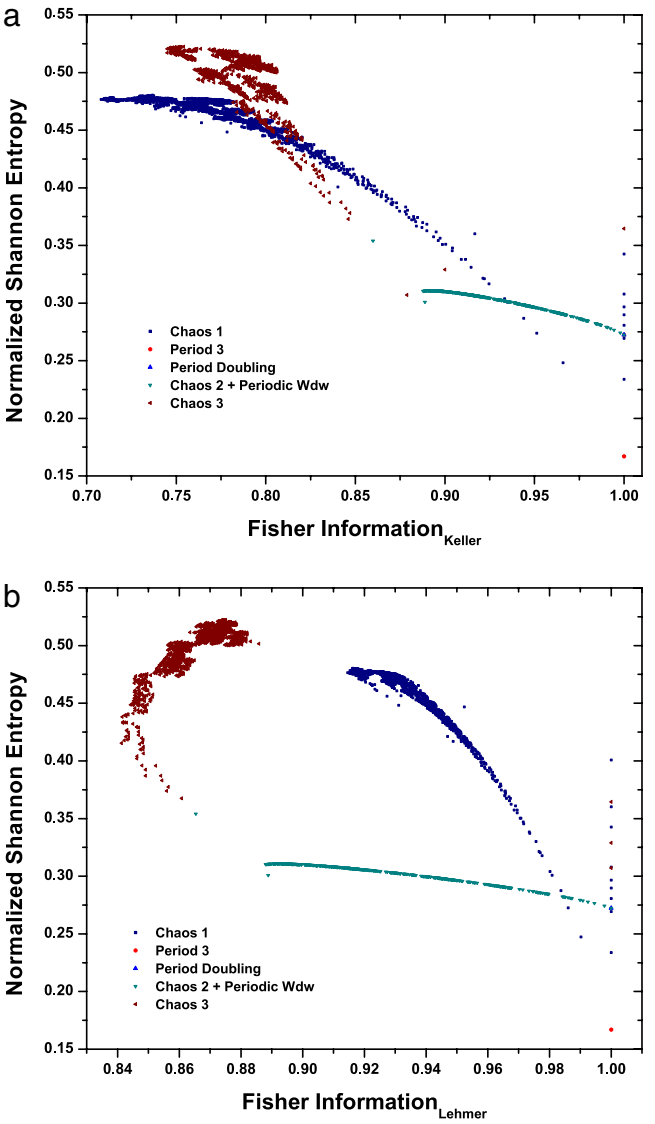

Fig. 3. Fisher–Shannon information plane in the parameter range $3.8 \leq r \leq 3.87$. (a) Evaluating the Fisher information measure with Keller's algorithm for the generation of sequence ordinal patterns and (b) Lehmer's algorithm.

is the one to be considered. To each time s BP assign a D-dimensional vector that results from the evaluation of the time series at times $s-(D-1)T, \ldots, s-T, s$. Clearly, the higher the value of D, the more information about "the past" is incorporated into the ensuing vectors. By the ordinal pattern of order D related to the time s, BP mean the permutation $\pi = (r_0, r_1, \ldots, r_{D-1})$ of $(0, 1, \ldots, D-1)$ defined by

$$x_{s-r_{D-1}T} \leq x_{s-r_{D-2}T} \leq \cdots \leq x_{s-r_1T} \leq x_{s-r_0T}. \tag{7}$$

In this way, the vector defined by Eq. (6) is converted into a definite symbol π. So as to get a unique result BP consider that $r_i < r_{i-1}$ if $x_{s-r_iT} = x_{s-r_{i-1}T}$. This is justified if the values of x_t have a continuous distribution so that equal values are very unusual.

For all the $D!$ possible orderings (permutations) π_i when the embedding dimension is D, their associated 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,

$$p(\pi_i) = \frac{\sharp\{s | s \leq M-(D-1)T;\, (s)\ \mathtt{has\ type}\ \pi_i\}}{M-(D-1)T}. \tag{8}$$

In the last expression the symbol $\sharp$ stands for "number". Thus, an ordinal pattern probability distribution $P = \{p(\pi_i), i = 1, \ldots, D!\}$ is obtained from the time series.

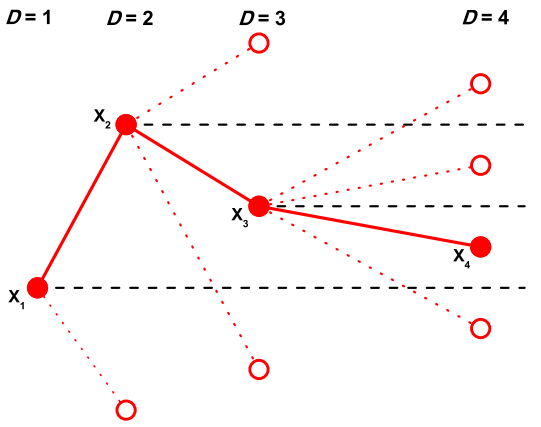

Fig. 4. Illustration of the construction-principle for ordinal patterns of length D [30]. If $D = 4$, full circles and continuous lines represent the sequence values $x_1 < x_2 > x_3 < x_4$ which lead to the pattern {0321}.

In Fig. 4, we illustrate the construction principle of the ordinal patterns of length $D = 2$, 3 and 4 [30]. Consider the values sequence $\{x_1, x_2, x_3, x_4\}$. For $D = 2$, there are only two possible directions from x_1 to x_2, up and down. For $D = 3$, starting from x_2 (up) the third part of the pattern can be above x_2, below x_1 or between x_1 and x_2 as illustrated in the Fig. 4. We can have similar situation starting from x_2 (down). For $D = 4$, for each one of the 6 possible positions for x_3 we will have 4 possible localizations for x_4, leading in this way finally to the $D! = 4! = 24$ different ordinal patterns. In the diagram of Fig. 4, with full circles and continuous line we represent the sequence values $x_1 < x_2 > x_3 < x_4$ which lead to the pattern {0321}. A graphical representation of all possible patterns corresponding to $D = 3$, 4 and 5 can be found in Fig. 2 of Ref. [30].

A.3. Bandt–Pompe and time series

It is clear that this ordinal time-series' analysis entails losing some details of the original time-series' amplitude-information. Nevertheless, a meaningful difficulty-reduction of the problem posed by the description of a given complex system is indeed achieved by just referring to its intrinsic structure via BP. The symbolic representation of time-series by recourse to a comparison of consecutive points ($T = 1$) or non consecutive ($T > 1$) points allows for an accurate empirical reconstruction of the underlying phase-space, even in the presence of weak (observational and dynamical) noise [1]. Furthermore, the ordinal-pattern's associated PDF is invariant with respect to nonlinear monotonous transformations. Accordingly, nonlinear drifts or scalings artificially introduced by a measurement device will not modify the quantifiers' estimation, a nice property if one deals with experimental data (see Ref. [31]). These advantages make the BP approach more convenient than conventional methods based on range partitioning. Additional advantages of the method reside in (i) its simplicity (we need few parameters: the pattern length/embedding dimension D and the embedding delay T) and (ii) the extremely fast nature of the pertinent 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), with a very weak stationary assumption (that is, for $k = D$, the probability for $x_t < x_{t+k}$ should not depend on t [1]).

The BP-generated probability distribution P is obtained once we fix the embedding dimension D and the embedding delay T. The former parameter plays an important role in the evaluation of the appropriate probability distribution, since D determines the number of accessible states, given by $D!$. Moreover, it has been established that the length M of the time series must satisfy the condition $M \gg D!$ in order to achieve a reliable statistics and proper distinction between stochastic and deterministic dynamics [9]. With respect to the selection of the parameters, BP suggest in their cornerstone paper [1] to work with $3 \leq D \leq 7$ with a time lag $T = 1$. Nevertheless, other values of T might provide additional information. Soriano et al. [32,33] and Zunino et al. [34] have recently showed that this parameter is strongly related, when it is relevant, to the intrinsic time scales of the system under analysis. In the present work, $D = 6$ and $T = 1$ are used. Of course it is also assumed that enough data are available for a correct embedding-time delay procedure (attractor-reconstruction).

A.4. Forbidden patterns

For deterministic one-dimensional maps, Amigó et al. [35–38] have conclusively demonstrated that *not all possible ordinal patterns* (as defined using Bandt and Pompe's methodology) can effectively materialize into orbits, which in a sense makes these patterns "forbidden". We insist: *this is an established fact, not a conjecture*. The existence of these *forbidden ordinal patterns* becomes indeed a persistent feature, a "new" dynamical property. For a fixed pattern-length (embedding dimension D) the number of forbidden patterns of a time series (unobserved patterns) is independent of the series length M. Such

independence does not characterize other properties of the series like proximity and correlation, which die out with time [36,38]. For example, in the time series generated by the logistic map with $r = 4$, if we consider patterns of length $D = 3$, the pattern $\{2, 1, 0\}$ is forbidden. That is, the pattern $x_{k+2} < x_{k+1} < x_k$ never appears [36]. Moreover, one can numerically verify that the number of forbidden patterns of length D in the case of logistic map decreases as the parameter r of the logistic increase, becomes maximal at $r = 4$ (see Fig. 4 in Ref. [29]). In consequence, the PDF of Bandt and Pompe's will not everywhere $p_i \neq 0$ since for the forbidden patterns we necessarily have $p_i = 0$ (independently of the time series length M). Thus, for such a PDF, its logistic map version makes Shannon's entropy obey $0 < S[P] < S_{\max}$ (see [9,29]) and FIM $0 < F[P] < F_{\max}$.

Stochastic processes could also exhibit forbidden patterns [27,28]. However, in the case of either uncorrelated (white noise) or certain correlated stochastic processes (noise with power low spectrum f^{-k} with $k \geq 0$, ordinal Brownian motion, fractional Brownian motion, and fractional Gaussian noise), it can be numerically shown that *no* forbidden patterns emerge. In the case of time series generated by an *unconstrained stochastic process* (uncorrelated process) every ordinal pattern has the same probability of appearance [35–38]. If the time series is long enough, all the ordinal patterns should eventually appear. If the number of time-series' observations is sufficiently large, the associated probability distribution function is the uniform distribution, and the number of observed patterns should depend only on the length M of the time series under study. Accordingly, in applying the Bandt and Pompe technique to an unconstrained stochastic process with enough data, the PDF is $p_i = 1/N \neq 0\ \forall i$ and, one has $S[P] = S_{\max}$ and $F[P] = 0$.

References

[1] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102.
[2] F. Pennini, A. Plastino, Localization estimation and global vs. local information measures, Phys. Lett. A 365 (2007) 263–267.
[3] B. Roy Frieden, Science From Fisher Information: A Unification, Cambridge University Press, Cambridge, 2004.
[4] O.A. Rosso, H. Craig, P. Moscato, Shakespeare and other English renaissance authors as characterized by information theory complexity quantifiers, Phys. A 388 (2009) 916–926.
[5] L. De Micco, C.M. González, H.A. Larrondo, M.T. Martín, A. Plastino, O.A. Rosso, Randomizing nonlinear maps via symbolic dynamics, Phys. A 87 (2008) 3373–3383.
[6] K. Mischaikow, M. Mrozek, J. Reiss, A. Szymczak, Construction of symbolic dynamics from experimental time series, Phys. Rev. Lett. 82 (1999) 1144–1147.
[7] G.E. Powell, I.C. Percival, A spectral entropy method for distinguishing regular and irregular motion of hamiltonian systems, J. Phys. A: Math. Gen. 12 (1979) 2053–2071.
[8] O.A. Rosso, S. Blanco, J. Jordanova, V. Kolev, A. Figliola, M. Schürmann, E. Başar, Wavelet entropy: a new tool for analysis of short duration brain electrical signals, J. Neurosci. Meth. 105 (2001) 65–75.
[9] O.A. Rosso, H.A. Larrondo, M.T. Martín, A. Plastino, M.A. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (2007) 154102.
[10] O.A. Rosso, C. Masoller, Detecting and quantifying stochastic and coherence resonances via information theory complexity measurements, Phys. Rev. E 79 (2009) 040106(R).
[11] O.A. Rosso, C. Masoller, Detecting and quantifying temporal correlations in stochastic resonance via information theory measures, Eur. Phys. J. B 69 (2009) 37–43.
[12] C. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Champaign, IL, 1949.
[13] R.A. Fisher, On the mathematical foundations of theoretical statistics, Philos. Trans. R. Soc. Lond. Ser. A 222 (1922) 309–368.
[14] A.L. Mayer, C.W. Pawlowski, H. Cabezas, Fisher information and dynamic regime changes in ecological systems, Ecol. Modeling 195 (2006) 72–82.
[15] K. Zografos, K. Ferentinos, T. Papaioannou, Discrete approximations to the Csiszár, Renyi, and Fisher measures of information, Canad. J. Statist. 14 (1986) 355–366.
[16] L. Pardo, D. Morales, K. Ferentinos, K. Zografos, Discretization problems on generalized entropies and *R*-divergences, Kybernetika 30 (1994) 445–460.
[17] M. Madiman, O. Johnson, I. Kontoyiannis, Fisher information, compound Poisson approximation, and the Poisson channel, IEEE Int. Symp. Inform. Th., Nice, June 2007.
[18] G.I. Ferri, F. Pennini, A. Plastino, LMC-complexity and various chaotic regimes, Phys. Lett. A 373 (2009) 2210–2214.
[19] O.A. Rosso, L. De Micco, A. Plastino, H.A. Larrondo, Info-quantifiers' map-characterization revisited, Phys. A 389 (2010) 4604–4612.
[20] F. Pennini, A. Plastino, Reciprocity relations between ordinary temperature and the Frieden–Soffer Fisher temperature, Phys. Rev. E 71 (2005) 047102.
[21] J.C. Sprott, Chaos and Time Series Analysis, Oxford University Press, Oxford, 2004.
[22] K. Keller, M. Sinn, Ordinal analysis of time series, Phys. A 356 (2005) 114–120.
[23] http://www.keithschwarz.com/interesting/code/factoradic-permutation/FactoradicPermutation.hh.html.
[24] C. Vignat, J.F. Bercher, Analysis of signals in the Fisher–Shannon information plane, Phys. Lett. A 312 (27) (2003) 27–33.
[25] A.R. Plastino, A. Plastino, Fisher information and bounds to the entropy increase, Phys. Rev. E 52 (1995) 4580–4582.
[26] C.S. Daw, C.E.A. Finney, E.R. Tracy, A review of symbolic analysis of experimental data, Review of Scientific Instruments 74 (2003) 915–930.
[27] O.A. Rosso, L.C. Carpi, P.M. Saco, M. Gómez Ravetti, A. Plastino, H.A. Larrondo, Causality and the entropy-complexity plane: robustness and missing ordinal patterns, Phys. A 391 (2012) 42–55.
[28] O.A. Rosso, L.C. Carpi, P.M. Saco, M. Gómez Ravetti, H.A. Larrondo, A. Plastino, Noisy-chaotic time series and the forbidden/missing patterns paradigm, (unpublished), 2011, http://arxiv.org/abs/1110.0776v1.
[29] O.A. Rosso, L. de Micco, H.A. Larrondo, M.T. Martín, A. Plastino, Generalized statistical complexity measure, Internat. J. Bifur. Chaos 20 (2010) 775–785.
[30] U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, N. Wessel, Classifying cardiac biosignals using pattern statistics and symbolic dynamics, Comput. Biol. Med. (2011) in press (doi:10.1016/j.compbiomed.2011.03.017).
[31] P.M. Saco, L.C. Carpi, A. Figliola, E. Serrano, O.A. Rosso, Entropy analysis of the dynamics of El Niño/Southern oscillation during the holocene, Phys. A 389 (2010) 5022–5027.
[32] M.C. Soriano, L. Zunino, L. Larger, I. Fischer, C.R. Mirasso, Distinguishing fingerprints of hyperchaotic and stochastic dynamics in optical chaos from a delayed opto-electronic oscillator, Optics Letters 36 (2010) 2212–2214.
[33] M.C. Soriano, M.C., L. Zunino, O.A. Rosso, I. Fischer, C.R. Mirasso, Time scales of a chaotic semiconductor laser with optical feedback under the lens of a permutation information analysis, IEEE J. Quantum Electron. 47 (2010) 252–261.
[34] L. Zunino, M.C. Soriano, I. Fischer, O.A. Rosso, C.R. Mirasso, Permutation information theory approach to unveil delay dynamics from time series analysis, Phys. Rev. E 82 (2010) 046212.
[35] J.M. Amigó, L. Kocarev, J. Szczepanski, Order patterns and chaos, Phys. Lett. A 355 (2006) 27–31.
[36] J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, True and false forbidden patterns in deterministic and random dynamics, Europhys. Lett. 79 (2007) 50001.
[37] J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, Combinatorial detection of determinism in noisy time series, Europhys. Lett. 83 (2008) 60005.
[38] J.M. Amigó, Permutation Complexity in Dynamical Systems, Springer-Verlag, Berlin, Germany, 2010.