Physics Letters A 376 (2012) 1577–1583

Contents lists available at SciVerse ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Contrasting chaos with noise via local versus global information 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, IFLP-CCT, Universidad Nacional de La Plata (UNLP), C.C. 727, 1900 La Plata, Argentina
[c] LaCCAN/CPMAT – Instituto de Computação, Universidade Federal de Alagoas, BR 104 Norte km 97, 57072-970 Maceió, Alagoas, Brazil
[d] Laboratorio de Sistemas Complejos, Facultad de Ingeniería, Universidad de Buenos Aires, 1063 Av. Paseo Colón 840, Ciudad Autónoma de Buenos Aires, Argentina
[e] Fellow of CONICET, Argentina

ARTICLE INFO

Article history:
Received 22 December 2011
Received in revised form 2 March 2012
Accepted 19 March 2012
Available online 21 March 2012
Communicated by C.R. Doering

ABSTRACT

At issue here is the distinction between noise and chaos. They are different phenomena but sometimes produce results that resemble each other. From a numerical viewpoint, in particular, subtle differences that exist between them are often difficult to discern. We present here a conceptual scheme, based on Information Theory, that successfully distinguishes between these two regimes. The idea is to look for the location of the pertinent signal on a special plane, called the information-one, whose axes are entropic-like measures. Using these quantifiers (one local, the other global), the contrast between the two dynamical regimes becomes apparent.

1. Introduction

Temporal sequences of measurements (or observations) are, of course, the basic elements for the study of natural phenomena. In particular, from these sequences, commonly called time series (TS), one should judiciously extract information on dynamical systems. The TS-analysis and characterization is of relevance to a broad range of research domains, as indicated by the variety of time series studied in different areas of science. Irregular and apparently unpredictable behavior is often observed in natural time series. Consequently, it is of great interest to establish whether the underlying dynamical process is of either deterministic or stochastic character in order to i) model the associated phenomenon and ii) determine which are the relevant quantifiers. Many procedures have been proposed to such effect. In the present work we focus attention on Information Theory quantifiers and try to use them in order to distinguish between time series generated by pure deterministic (chaos) and pure stochastic (noise) processes.

Given the TS $\mathcal{S}(t) \equiv \{x_t\}$, the starting point is extracting from it a probability distribution function (PDF) $P \equiv \{p_i\}$. The determination of the most adequate PDF is a fundamental issue. P and the pertinent sample space Ω are inextricably linked. Indeed, many methods have been proposed for a proper selection of the probability space (Ω, P), but their applicability depends on specific features of the data, such as stationarity, TS-length, variation of parameters, level of noise contamination, etc. Global aspects of the dynamics can be "captured", but the different approaches are not equivalent in their ability to discern the relevant physical details.

The permutation of ordinal patterns proposed by Bandt–Pompe (BP) for generating a TS-PDF [1] is one of the most simple symbolization techniques available and takes into account time-causality in the concomitant process. In particular, Rosso et al. [2] showed that the BP methodology may be profitably used in the so-called entropy-complexity plane $(H \times C)$ so as to separate and differentiate amongst chaotic and deterministic systems.

Another relevant information plane is the Shannon–Fisher one, introduced by Vignat and Bercher in [3]. These authors showed that the simultaneous examination of both Shannon's entropy and Fisher's information measure (FIM) is required to characterize the non-stationary behavior of a complex signal. They also demonstrated that scaling and uncertainly properties, together, highlight the fact that FIM and Shannon's entropy are intrinsically linked (see [4] for explicit relationships), so that the characterization of complex signals should improve when considering their localization in the Shannon–Fisher plane. It is worth noting that some ambiguity arises in trying to apply the Bandt–Pompe methodology in a Fisher environment, a point exhaustively examined in [5], using the logistic map as illustrative example.

In this communication we introduce a representation space, to be called the causality Shannon–Fisher information plane $(H \times F)$. Its horizontal and vertical axis are suitable functionals of the pertinent probability distribution, namely, the normalized Shannon

* Corresponding author at: Laboratorio de Sistemas Complejos, Facultad de Ingeniería, Universidad de Buenos Aires, 1063 Av. Paseo Colón 840, Ciudad Autónoma de Buenos Aires, Argentina.
E-mail addresses: folivares@fisica.unlp.edu.ar (F. Olivares), plastino@fisica.unlp.edu.ar (A. Plastino), oarosso@fibertel.com.ar, oarosso@gmail.com (O.A. Rosso).

0375-9601/$ – see front matter
http://dx.doi.org/10.1016/j.physleta.2012.03.039

entropy (H) and the normalized FIM (F). It is of the essence for us to evaluate these quantifiers as functions of a TS-PDF obtained via the Bandt and Pompe recipe [1]. In the present work, the same dynamical systems studied in [2] are revisited. We will show that the causality Shannon–Fisher information plane helps one to get additional insights with respect to those gathered via the entropy-complexity plane [2]. The distinction between stochastic noise (true noise) and deterministic dynamics (chaos) gets now more clearly delineated with relation to differences in planar location.

2. Shannon entropy and Fisher information measure

Given a continuous probability distribution function (PDF) $f(x)$ (with $x \in [x_{\min}, x_{\max}]$ and $\int_{x_{\min}}^{x_{\max}} f(x)\,dx = 1$), its associated *Shannon Entropy* S [6] is

$$S[f] = -\int_{x_{\min}}^{x_{\max}} f \ln(f)\,dx, \tag{1}$$

a measure of "global character" [7] 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 [8,9], which constitutes a measure of the gradient content of the distribution f, thus being quite sensitive even to tiny localized perturbations. It reads

$$F[f] = \int_{x_{\min}}^{x_{\max}} \frac{|\vec{\nabla} f|^2}{f}\,dx. \tag{2}$$

The above is not the most general Fisher-definition, but in physical applications it is the one often employed [9]. 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 [9].

In the previous definition of FIM (Eq. (2)) the division by $f(x)$ is not convenient if $f(x) \to 0$ at certain x-values. Frieden [9] has shown that such points can be avoided if the integral in Eq. (2) is understood in terms of Cauchy's principal value. Another way out is work with a real probability amplitudes $f = \psi^2$ [8,9], then

$$F[\psi] = 4\int_{x_{\min}}^{x_{\max}} |\vec{\nabla}\psi|^2\,dx, \tag{3}$$

which is a simpler form than Eq. (2) (no divisors) and shows that F simply measures the gradient content in $\psi(x)$. The gradient operator significantly influences the contribution of minute local f-variations to FIM's value. Accordingly, this quantifier is called a "local" one [7,9]. Note that Shannon's entropy decreases for skewed distributions, while FIM increases in such a case. Local sensitivity is useful in scenarios whose description necessitates appeal to a notion of "order" (see below).

Let now $P = \{p_i;\ i = 1, \ldots, N\}$ be a discrete probability distribution, with N the number of possible states of the system under study. The concomitant problem of information-loss due to discretization has been thoroughly studied (see, for instance, [10–12], 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, we define a "normalized" Shannon entropy as

$$H[P] = S[P]/S_{\max} = \left\{-\sum_{i=1}^{N} p_i \ln(p_i)\right\}/S_{\max}, \tag{4}$$

where the denominator $S_{\max} = S[P_e] = \ln N$ is that attained by a uniform probability distribution $P_e = \{p_i = 1/N,\ \forall i = 1, \ldots, N\}$.

Eqs. (2) or (3) can be used as a FIM-starting point in the discrete case. The proposal of Frieden [9] and of Ferri et al. [13] take, as starting point, Eq. (2). However, the concomitant discretizations can be considered just approximations to what results when using Eq. (3) as the starting point. We emphasize that a discrete normalized FIM, convenient for our present purposes, is given by

$$F[P] = F_0 \sum_{i=1}^{N-1} \left[(p_{i+1})^{1/2} - (p_i)^{1/2}\right]^2, \tag{5}$$

which follow from Eq. (3). Here the normalization constant F_0 reads

$$F_0 = \begin{cases} 1 & \text{if } p_{i^*} = 1 \text{ for } i^* = 1 \text{ or } i^* = N \text{ and } p_i = 0\ \forall i \neq i^*, \\ 1/2 & \text{otherwise.} \end{cases} \tag{6}$$

Important note: some ambiguity arises in trying to apply the Bandt–Pompe methodology to the construction of local entropic quantifiers, an issue discussed and clarified by us in [5], using the logistic map as illustrative example. An important inequality has to be obeyed by FIM, that was advanced by Vignat–Bercher [3]. Let $Y = \frac{1}{2\pi e}\exp[2S]$ and J equal the un-normalized FIM-quantifier. Then, the Bercher and Vignat inequality reads

$$Y \cdot J \geqslant 1. \tag{7}$$

We have numerically verified that Eq. (7) is duly satisfied in all our calculations. The concomitant results are not displayed because of space-saving reasons.

If our system lies in a very ordered state, being thus represented by an extremely narrow PDF, we have a normalized Shannon entropy $H \sim 0$ and a normalized Fisher's Information Measure $F \sim 1$. On the other hand, when the system under study is represented by a very disordered state, one gets an almost flat PDF and $H \sim 1$ while $F \sim 0$. One can state that the general FIM-behavior is opposite to that of the Shannon entropy [13–15]. 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 seriously taken into account in evaluating the sum in Eq. (5). The summands 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, hereby its local nature.

3. Description of our chaotic and stochastic systems

Here we study both chaotic and stochastic systems, selected as illustrative examples of different classes of signals, namely, (a) chaotic dynamic maps and (b) truly stochastic processes. Of the first kind we deal with:

(1) The Logistic Map: the classical one-dimensional quadratic map [16] defined by

$$x_{n+1} = r x_n (1 - x_n). \tag{8}$$

For the fully chaotic case $r = 4$ this map has a non-uniform natural invariant histogram PDF.

(2) The Skew Tent Map: one has [16]

$$\begin{cases} x/\omega & \text{for } x \in [0, \omega], \\ (1-x)/(1-\omega) & \text{for } x \in [\omega, 1]. \end{cases} \tag{9}$$

The case $\omega = 0.1847$ is here considered. For any ω-value this map has a uniform, invariant-histogram PDF.

(3) Henon's Map: This is a $2D$ extension of the Logistic Map [16] given by:

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n, \\ y_{n+1} = b - x_n. \end{cases} \tag{10}$$

The values used here, $a = 1.4$ and $b = 0.3$, correspond to a chaotic dynamics.

(4) The Lorenz Map of Rossler's oscillator: for the $3D$ continuous Rossler oscillator [16] one has

$$\begin{cases} \dot{x} = -y - z, \\ \dot{y} = x + ay, \\ \dot{z} = b + z(x - c), \end{cases} \tag{11}$$

where $a = 0.2$, $b = 0.2$, and $c = 5.7$ correspond to a chaotic dynamics. The evolution of the system was determined by recourse to a variable-step Runge–Kutta–Fehlberg approach with a sampling time of $\Delta t = 10^{-5}$. The Lorenz map is obtained by storing just x-minimal values.

(5) Schuster Maps: Schuster and coworkers [17] introduced a class of maps which generate intermittent signals with chaotic bursts that also display $1/f^z$ noise

$$x_{n+1} = x_n + x_n^z, \quad \text{Mod } 1. \tag{12}$$

In particular, results for $1.25 \leqslant z \leqslant 2$ with $\Delta z = 0.25$ are reported.

Additionally, the following *stochastic* processes are considered in the present study. They are chosen because of their popularity in characterizing and modeling the behavior of natural time series.

(6) Noises with f^{-k} power spectrum: The corresponding time series are generated as follows [18]:

1) Using the Mersenne twister generator [19] through the MATLAB© *RAND* function we generate pseudo random numbers y_i^0 in the interval $(-0.5, 0.5)$ with an
   (a) almost flat power spectra (PS),
   (b) uniform PDF, and
   (c) zero mean value.
2) Then, the Fast Fourier Transform (FFT) y_i^1 is first obtained and then multiplied by $f^{-k/2}$, yielding y_i^2.
3) Now, y_i^2 is symmetrized so as to obtain a real function.

The pertinent inverse FFT is now at our disposal, after discarding the small imaginary components produced by the numerical approximations. The resulting time series $\eta^{(k)}$ exhibits the desired power spectra and, by construction, is representative of non-Gaussian noises. In the present communication, noises with $0 \leqslant k \leqslant 5$ and $\Delta k = 0.25$ are considered.

(7) Fractional Brownian motion (fBm) and fractional Gaussian noise (fGn): fBm is the only family of processes which is (a) Gaussian, (b) self-similar, and (c) endowed with stationary increments (see [20] and references therein). The normalized families of these Gaussian processes, $\{B^{\mathcal{H}}(t), t > 0\}$, are endowed with the following properties:

i) $B^{\mathcal{H}}(0) = 0$ almost surely (a.s.), i.e., with probability 1,
ii) $\mathbb{E}[B^{\mathcal{H}}(t)] = 0$ (zero mean), and
iii) covariance given by

$$\mathbb{E}\big[B^{\mathcal{H}}(t)B^{\mathcal{H}}(s)\big] = \big(t^{2\mathcal{H}} + s^{2\mathcal{H}} - |t-s|^{2\mathcal{H}}\big)/2, \tag{13}$$

for $s, t \in \mathbb{R}$. Here $\mathbb{E}[\cdot]$ refers to the average computed with a Gaussian probability density. The power exponent $0 < \mathcal{H} < 1$ is commonly known as the Hurst parameter or Hurst exponent. These processes exhibit "*memory*" for any Hurst parameter *except for* $\mathcal{H} = 1/2$, as one realizes from Eq. (13). The $\mathcal{H} = 1/2$-case corresponds to classical Brownian motion and successive "motion-increments" are as likely to have the same sign as the opposite (there is no correlation among them). Thus, Hurst's parameter defines two distinct regions in the interval $(0, 1)$. When $\mathcal{H} > 1/2$, consecutive increments tend to have the same sign so that these processes are *persistent*. For $\mathcal{H} < 1/2$, on the other hand, consecutive increments are more likely to have opposite signs, and we say that they are *anti-persistent*.

Let us introduce the quantity $W^{\mathcal{H}}(t) = B^{\mathcal{H}}(t+1) - B^{\mathcal{H}}(t)$, $t > 0$, (fBm-"increments") so as to express our Gaussian noise in the fashion

$$\begin{aligned} \rho(k) &= \mathbb{E}\big[W^{\mathcal{H}}(t)W^{\mathcal{H}}(t+k)\big] \\ &= \frac{1}{2}\big[(k+1)^{2\mathcal{H}} - 2k^{2\mathcal{H}} + |k-1|^{2\mathcal{H}}\big], \quad k > 0. \end{aligned} \tag{14}$$

Note that for $\mathcal{H} = 1/2$ all correlations at nonzero lags vanish and $\{W^{1/2}(t), t > 0\}$ thus represents *white noise*. The fBm and fGn are continuous but non-differentiable processes (in the classical sense). As a non-stationary process, they do not possess a spectrum defined in the usual sense. However, it is possible to define a *generalized power spectrum* of the form $\Phi \propto |f|^{-\alpha}$, with $\alpha = 2\mathcal{H} + 1$, $1 < \alpha < 3$ for fBm and, $\alpha = 2\mathcal{H} - 1$, $-1 < \alpha < 1$, for fGn. Due to their Gaussian nature, and other characteristics enumerated above, the Bandt–Pompe ideas are applicable to the fBn and fGn dynamical processes [20,21]. To simulate the fBm and fGn time series we adopt the Davies–Harte algorithm [22], as recently improved by Wood and Chan [23], which is both exact and fast.

4. Results and discussion

4.1. Generalities

As we mentioned in Section 2, the "i-ordering" for the discrete-PDFs must be specified so as to compute the FIM. The question is, *which is the arrangement that one could regard as the "proper" ordering?* The answer is straightforward in some cases, histogram-based PDF constituting a conspicuous example. For such a 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 on non-overlapping sub-intervals (bins). Thus, 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, for example, the time series generated by the logistic map with $r = 4$, whose associated dynamics is totally developed chaos [16]. It is well known that in this situation the logistic map exhibits an almost flat PDF-histogram (referred to the amplitude values in the interval $[0, 1]$), with peaks at $x = 0$ and $x = 1$. This histogram-PDF constitutes an invariant measure of the system [16]. Thus, if we use this PDF we obtain $H[P] \cong 1$ 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*. Accordingly, if one wants to use quantifiers based on Information Theory with the purpose of distinguishing deterministic signals (chaos) from noise (random process), one should demand for improvements in the methodology used for associating a PDF to a time series generated by a dynamical system. Such goal can be achieved if the time-causality (in the series' values) is duly taken into account when extracting from the TS the associated PDF, something that one gets automatically from the Bandt–Pompe (BP) methodology.

BP introduced their successful approach for such extraction using a symbolization technique [1]. Their symbolic data are created by appropriately ranking the series' values. This rank is defined (this is the essential feature) by reordering the suitably embedded data in ascending order with an embedding dimension D. *"Causal"*

information is, in such a manner, properly incorporated into the "building-up" process that yields (Ω, P).

4.2. The Bandt and Pompe approach to the PDF determination

The Bandt–Pompe methodology is not restricted to time series representative of low-dimensional dynamical systems but can be applied to any type of time series $\{x_t\}$ (regular, chaotic, noisy, or reality based). No attractor reconstruction is assumed for our D-dimensional space. The only limitation for its applicability is that the underlying dynamical process (time series) fulfills a very weak stationarity condition: for $k \leqslant D$, the probability for $x_t < x_{t+k}$ should not depend on t [1]. It is clear that with this ordinal time series analysis details of the original amplitude information are lost. However, a meaningful reduction of the complex systems to their basic intrinsic structure is provided. This way of symbolizing time series, based on a comparison of consecutive points, allows for a more accurate empirical reconstruction of the underlying phase space in the case of chaotic time series affected by weak (observational and dynamical) noise. Furthermore, ordinal pattern distributions are invariant with respect to nonlinear monotonous transformations. Thus, nonlinear drifts or scalings artificially introduced by a measurement device do not modify the quantifiers' estimations, a highly desirable property for the analysis of experimental data. We see that many BP-advantages recommend it as an alternative to more conventional methods based on range partitioning.

To use the Bandt and Pompe [1] methodology for evaluating the PDF P associated to the time series (dynamical system) under study, one starts by considering partitions of the pertinent D-dimensional space that will hopefully "reveal" relevant details of the ordinal-structure of a given one-dimensional time series $\mathcal{S}(t) = \{x_t; t = 1, \ldots, M\}$ with embedding dimension $D > 1$ ($D \in \mathbb{N}$) and embedding time delay τ ($\tau \in \mathbb{N}$). We are interested in "ordinal patterns" of order (length) D generated by $(s) \mapsto (x_{s-(D-1)\tau}, x_{s-(D-2)\tau}, \ldots, x_{s-\tau}, x_s)$, which assigns to each time s the D-dimensional vector of values at times $s, s-\tau, \ldots, s-(D-1)\tau$. Clearly, the greater the D-value, 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 $x_{s-r_{D-1}\tau} \leqslant x_{s-r_{D-2}\tau} \leqslant \cdots \leqslant x_{s-r_1\tau} \leqslant x_{s-r_0\tau}$. In order to get a unique result we set $r_i < r_{i-1}$ if $x_{s-r_i} = x_{s-r_{i-1}}$. This is justified if the values of x_t have a continuous distribution so that equal values are very unusual. Thus, for all the $D!$ possible permutations π of order D, their associated relative frequencies can be naturally computed by the number of times this particular order sequence is found in the time series divided by the total number of sequences.

In order to illustrate Bandt–Pompe method (BP), we will consider a simple example: a time series with seven ($M = 7$) values $x = \{4, 7, 9, 10, 6, 11, 3\}$ and we evaluate the BP-PDF for $D = 3$ and $\tau = 1$. 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: $p(\{012\}) = p(\{201\}) = 2/5$; $p(\{102\}) = 1/5$; $p(\{021\}) = p(\{120\}) = p(\{210\}) = 0$.

Consequently, it is possible to quantify the diversity of the ordering symbols (patterns of length D) derived from a scalar time-series, by evaluating the so-called permutation entropy (that Shannon's entropy special version corresponding to the Bandt–Pompe PDF). Of course, 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 a reliable statistics [24]. With respect to the selection of the other parameter, Bandt and Pompe suggest to work with $4 \leqslant D \leqslant 6$ and specifically considered an embedding delay $\tau = 1$ in their cornerstone paper [1]. 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, with the intrinsic time scales of the system under analysis [25,26].

Let us insist: the Bandt and Pompe technique is computationally fast and constitutes the only procedure, among those currently in use, that takes into account the temporal structure of the original time series (that is in turn generated by the physical process under study). In addition, the BP technique allows one to i) uncover important details concerning the TS-ordinal structure [2] and ii) to disclose information about temporal correlations [27,28]. For the BP-computed PDF we follow the lexicographic ordering [29,30]. Such is the order sequence used in the evaluation of FIM's PDF gradient.

4.3. The Bandt–Pompe PDF for deterministic and stochastic process

A central issue is to be now brought up. For deterministic one-dimensional maps, Amigó et al. [31,32] have conclusively demonstrated that *not all possible ordinal patterns* (as defined using Bandt–Pompe's methodology) can effectively materialize into orbits, which in a sense makes these patterns "forbidden". Remark that *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 [31,32]. For example, in the time series generated by the logistic map with $r = 4$, if we consider patterns of length $D = 3$ and $\tau = 1$, the pattern $\{210\}$ is forbidden. That is, the pattern $x_{s+2} < x_{s+1} < x_s$ never appears [31]. In consequence, the BP-PDF will not be finite 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 the normalized Shannon's entropy obey $0 < H[P] < 1$ while we have for FIM $0 < F[P] < 1$.

Stochastic processes can also exhibit forbidden patterns [33,34]. However, in the case of either uncorrelated (white noise) or certain correlated stochastic processes (noise with power law spectrum f^{-k} with $k \geqslant 0$, i.e., fBm, fGn), 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 [31,32]. If the time series is large 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–Pompe technique to an unconstrained stochastic process with enough data, the PDF is $p_i = 1/N \neq 0 \ \forall i$ and one has $H[P] = 1$ and $F[P] = 0$.

For correlated stochastic processes (like noise with power spectrum f^{-k} with $k > 0$, fBm and fGn) the probability of observing individual patterns depends not only on the time series length M but also on the correlations' structure [33]. The existence of a non-observed ordinal pattern does not qualify them as forbidden but only as missing, since the effect could be attributed due finitude of the time series. A similar observation also holds for the case of real data, that always possess a stochastic component due to the

omnipresence of dynamical noise [35]. Thus, the existence of missing ordinal patterns could be either related to stochastic processes (correlated or uncorrelated) or to deterministic, noisy processes, which is the case for observational time series. In consequence, the Bandt–Pompe PDF for constrained stochastic process with enough data will be a PDF different from the uniform one and characterized by $0 < H[P] < 1$ and $0 < F[P] < 1$.

4.4. Our findings

For each system (see Section 3) 10 different TS of $M = 2^{15}$ data each were analyzed. In all cases each series possesses a different initial condition. For all of them, the corresponding Bandt–Pompe PDF was evaluated with a pattern-length (embedding dimension) $D = 6$ and embedding delay $\tau = 1$, using the lexicographic pattern-ordering proposed by Lehmer [30]. Similar results are obtained (but not shown) when a small variation of the lexicographic pattern-ordering [29] is instead implemented. For additional discussion on BP-PDF and the i-ordering see [5]. The concomitant mean values of both the normalized Shannon entropy, $H[P]$, and normalized Fisher Information Measure, $F[P]$ are plotted in Fig. 1. Qualitatively same results are obtained when the evaluations were made with embedding dimensions $D = 4$ and 5.

As mentioned in Section 2, different FIM-discretization procedures are available. We have compared the present FIM-numerical results with those obtained via de discretization approach of Ferri et al. [13]. One can in this way discover that, even if the ensuing pertinent pairs of FIM-values are not identical, their position and distribution in the Shannon–Fisher plane are not qualitatively different. The associated causality's global aspects are invariant against discretization-technique-change, entailing that our planar representation is robust. Let us reiterate that we have numerically verified fulfillment of the Vignat–Bercher inequality, Eq. (7), a sort of uncertainty property which relates Shannon's entropy with FIM and that Vignat–Bercher advance as their Cramer–Rao version. The inequality is obeyed by all the dynamical systems (time series) analyzed in the present work.

Let us consider localization in the $H \times F$-plane. From Fig. 1 we clearly see that all chaotic dynamic maps under study (Logistic map, Henon's map, Lorenz Map of Rossler's oscillator, Schuster's map) are located at the entropic region lying between 0.45 and 0.7 and reach high FIM values. This entails that characteristics structures in the chaotic time series lead to a planar position near to the middle-top of the entropy-causality plane ($H \times C$) (see [2]). For the Henon chaotic map both the X and Y time series coordinates have been considered, sharing the same point-localization in the $H \times F$ plane. Schuster's maps exhibit intermediate entropic and FIM values ($0.2 < H < 0.65$, $0.4 < F < 0.6$). The reason for this behavior is that these maps exhibit laminar regions separated by chaotic bursts. When the parameter z decreases, FIM increases approaching a location in the vicinity of the chaotic maps. This is so because the size of the laminar regions is small, entailing that the system becomes more similar to a fully chaotic one.

Noises with f^{-k} power spectrum (with $0 \leqslant k \leqslant 5$) exhibit a wide range of entropic values ($0.1 \leqslant H \leqslant 1$) and FIM values lying between $0 \leqslant F \leqslant 0.5$. A smooth transition in the planar location is observed in the passage from uncorrelated noise ($k = 0$ with $H \sim 1$ and $F \sim 0$) to correlated one ($k > 0$). The correlation degree grows as the k value increases. From Fig. 1 we gather that, for stochastic time series with increasing correlation-degree, the associated entropic values H decrease, while Fisher's values F grow in a fashion that converges to the planar location $(0.105, 0.492)$. Note that, for the present time series' length $M = 2^{15}$, the planar location for those time series with $5 < k \leqslant 8$ always remains in the vicinity of the "convergence point" previously mentioned. Thus, this behavior can be associated to the specific TS-length chosen by us.

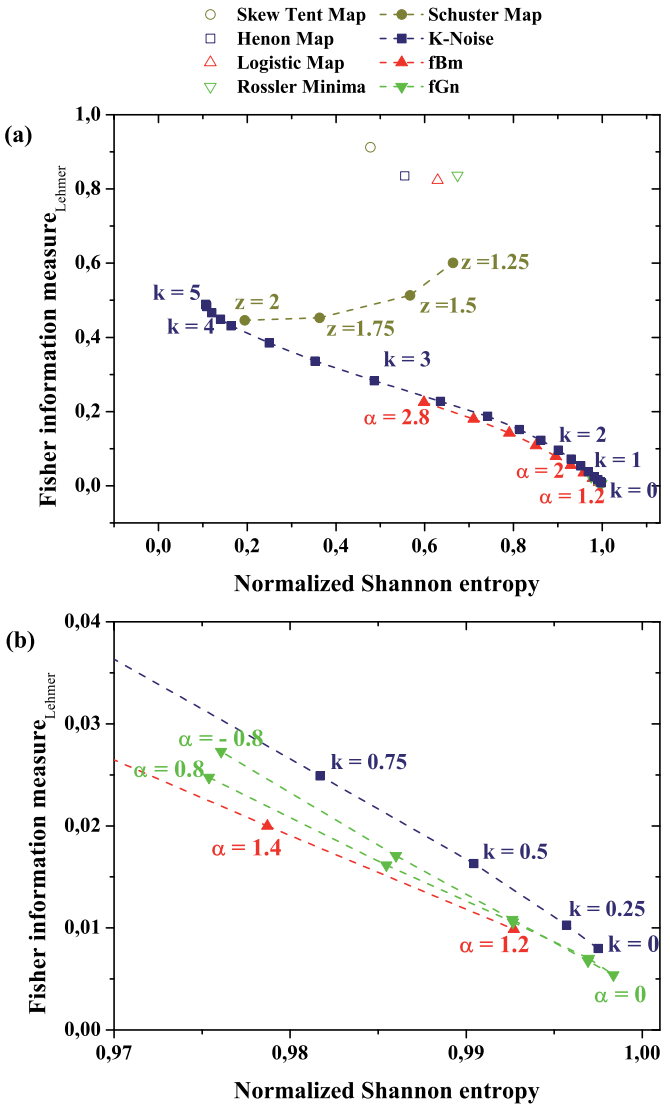

Fig. 1. (a) Localization of the different chaotic and stochastic (noises) systems in the information Shannon–Fisher causality plane, for $D = 6$ and $\tau = 1$. (b) Enlargement near the point $H = 1$, $F = 0$.

As for the planar location for fractional Brownian motion ($1 < \alpha < 3$), one finds entropy values lying between $0.5 \leqslant H \leqslant 1$ with low Fisher information $F < 0.3$. The persistent fBm behavior ($2 < \alpha < 3$) – *long memory processes* – displays higher FIM values than its counterpart fBm anti-persistent one ($1 < \alpha < 2$) – *short memory ones* – which places the FIM as a memory-quantifier for such kind of processes. On the other hand, fractional Gaussian noise ($-1 < \alpha < 1$) exhibits higher entropy values ($0.96 < H < 1$) while its Fisher information ranges between $0 \leqslant F \leqslant 0.03$. Note that these Gaussian noises ($0 < |\alpha| < 1$) exhibit similar values for both informational quantifiers. In the causality $H \times F$ information plane (see Fig. 1(b)) both fBm and fGn are encountered lying below the position attained by noises with f^{-k} PS. We associate this fact to the Gaussian nature of the respective processes. The white Gaussian noise for $\alpha = 0$ reaches maximum entropy and minimum Fisher information, as expected. Note that its location is lower than that pertaining to the case $k = 0$ of f^{-k}-noise.

The causality Shannon complexity plane, $H \times C$, is based only on global characteristics of the associated TS-PDF (both quantities are defined in terms of Shannon entropies [2]), while the causality Shannon–Fisher plane, $H \times F$, is based on global *and* local charac-

teristics of the TS-PDF. In the case of $H \times C$ the variation range is $[0,1] \times [C_{\min}, C_{\max}]$ (with $C_{\min}$ and $C_{\max}$ the minimum and maximum statistical complexity values, respectively, for a given H-value [36]), while in the causality $H \times F$ the range is $[0,1] \times [0,1]$. This range-modification clearly facilitate the distinction between deterministic (chaos) and stochastic dynamics (in a sense, there is "more room"). At first sight, when one compare the present $H \times F$-results (Fig. 1) with those obtained for the $H \times C$-plane (see Fig. 1 in [2]) a qualitatively agreement is discerned. Both planes afford a good distinction between deterministic and stochastic dynamics. However, important improvements are obtained when the FIM-associated local characteristics are incorporated to the picture, especially when system parameters vary and thus give rise to different dynamical regimes.

Look at the planar behavior of the Schuster map when the parameter z is changed. In the $H \times C$-plane the corresponding points are localized very close to the curve of maximum statistical complexity, $C_{\max}$ independently of the value of z. In fact, this kind of behavior in this plane, is typical for all chaotic dynamics studied (even for chaotic maps [2,37], as well as, in continuous chaotic systems [38]). That is, chaotic dynamics present near maximum statistical complexity. When local aspect of the TS-PDF are included, clear variations in FIM values are observed, facilitating in this way the identification of the different dynamical behavior of the system (see Fig. 1).

As second example, consider the logistic map's planar representation (control parameter range $3.8 \leqslant r \leqslant 3.87$) which includes a mixture of order and chaos (the corresponding bifurcation diagram is presented in Fig. 2(a) [5,13]). 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 end up 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 denominated *Chaos* 3, which "lives" in a wider region that includes the three sectors of the previous attractor. In Figs. 2(b) and 2(c) the two causality planes, $H \times C$ and $H \times F$ are depicted. From these plots one can conclude that a more clear distinction between the different regimes associated to the control parameter variation is provided by the causality $H \times F$-plane that we are advancing in this communication.

5. Conclusions

In this work we have presented an extensive series of numerical simulations/computations. On that basis, we suggest that the following conclusions may reasonably be gathered. In the present communication we have contrasted the characterizations of deterministic chaotic and noisy-stochastic dynamics, as represented by time series of finite length. The pertinent characterizations can be successfully achieved with reference to an information plane (its two coordinate axis being different information quantifiers). One just has to look at the different planar locations of our two dynamical regimes. In point of fact, the plane is defined by i) the Shannon entropy, responsible for global features and ii) the Fisher information measure, accountable for local attributes. The two quantifiers are evaluated using Bandt–Pompe's ordinal patterns-probability distribution function (PDF). Such PDF can adequately take into account the causal time ordering present in the time series. Our present results are in agreement with those obtained some time ago with the causality Shannon complexity plane and complement the pertinent results by adding a different perspective. When is the $H \times F$ more recommendable than other related planar representations? This is so for those cases in which the characterization of different dynamical regimes *associated to system's parameter-changes* are the focus of interest. This allows us to conclude that causality information planes do con-

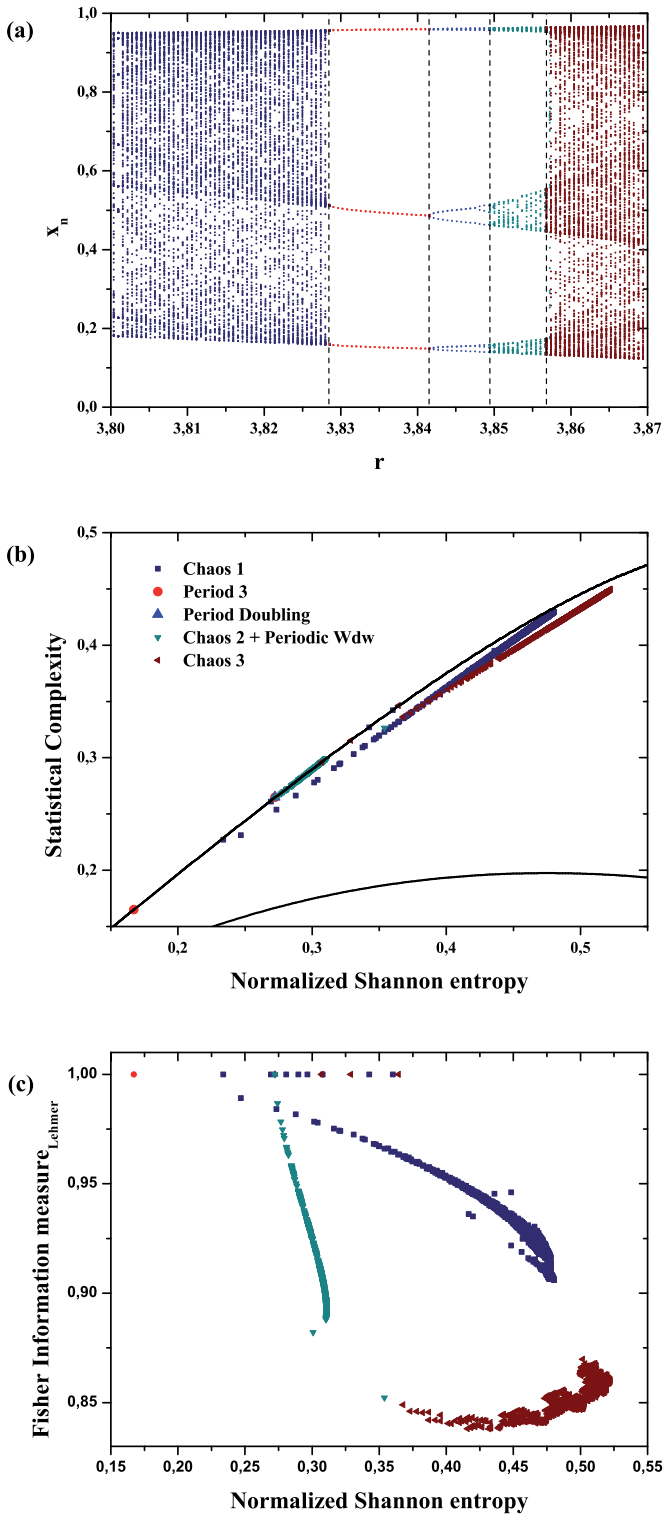

Fig. 2. Results for the logistic map, as a function of the parameter $3.8 \leqslant r \leqslant 3.87$. (a) Bifurcation diagram ($\Delta r = 0.0006$). The vertical lines represent the different dynamical windows described in the text. (b) Causality $H \times C$-map with $D = 6$ and $\tau = 1$, $\Delta r = 1 \times 10^{-5}$. The continue lines correspond to the minimum and maximum complexity. (c) Causality $H \times F$-map with $D = 6$ and $\tau = 1$, $\Delta r = 1 \times 10^{-5}$.

stitute a useful new tool for the analysis and characterization of time series. Moreover, these planes are able to reveal subtle differences between noise and chaos, that are related but different phenomena.

Acknowledgements

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

References

[1] C. Bandt, B. Pompe, Phys. Rev. Lett. 88 (2002) 174102.
[2] O.A. Rosso, H.A. Larrondo, M.T. Martín, A. Plastino, M.A. Fuentes, Phys. Rev. Lett. 99 (2007) 154102.
[3] C. Vignat, J.F. Bercher, Phys. Lett. A 312 (2003) 27.
[4] A.R. Plastino, A. Plastino, Phys. Rev. E 52 (1995) 4580.
[5] F. Olivares, A. Plastino, O.A. Rosso, Physica A 391 (2012) 2518.
[6] C. Shannon, W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Champaign, IL, 1949.
[7] F. Pennini, A. Plastino, Phys. Lett. A 365 (2007) 263.
[8] R.A. Fisher, Philos. Trans. R. Soc. Lond. Ser. A 222 (1922) 309.
[9] B. Roy Frieden, Science from Fisher Information: A Unification, Cambridge University Press, Cambridge, 2004.
[10] K. Zografos, K. Ferentinos, T. Papaioannou, Canad. J. Stat. 14 (1986) 355.
[11] L. Pardo, D. Morales, K. Ferentinos, K. Zografos, Kybernetika 30 (1994) 445.
[12] M. Madiman, O. Johnson, I. Kontoyiannis, in: IEEE Int. Symp. Inform. Theory, Nice, June 2007.
[13] G.I. Ferri, F. Pennini, A. Plastino, Phys. Lett. A 373 (2009) 2210.
[14] O.A. Rosso, L. De Micco, A. Plastino, H.A. Larrondo, Physica A 389 (2010) 4604.
[15] F. Pennini, A. Plastino, Phys. Rev. E 71 (2005) 047102.
[16] H.O. Peitgen, H. Jürgens, D. Saupe, Chaos and Fractals, New Frontiers of Science, 2nd ed., Springer-Verlag, New York, 1992.
[17] H.G. Schuster, Deterministic Chaos, 2nd ed., VCH, Weinheim, 1988.
[18] H.A. Larrondo, Matab program: *noisefk.m*, http://www.mathworks.com/matlabcentral/fileexchange/35381, 2012.
[19] M. Matsumoto, T. Nishimura, ACM Transactions on Modeling and Computer Simulation 8 (1998) 3.
[20] L. Zunino, D.G. Pérez, M.T. Martín, A. Plastino, M. Garavaglia, O.A. Rosso, Phys. Rev. E 75 (2007) 021115.
[21] C. Bandt, F. Shiha, J. Time Series Analysis 28 (2007) 646.
[22] R.B. Davies, D.S. Harte, Biometrika 74 (1987) 95.
[23] A.T.A. Wood, G. Chan, J. Comput. Graph. Stat. 3 (1994) 409.
[24] A. Kowalski, M.T. Martín, A. Plastino, O.A. Rosso, Physica D 233 (2007) 21.
[25] L. Zunino, M.C. Soriano, I. Fischer, O.A. Rosso, C.R. Mirasso, Phys. Rev. E 82 (2010) 046212.
[26] M.C. Soriano, L. Zunino, O.A. Rosso, I. Fischer, C.R. Mirasso, IEEE J. Quantum Electron. 47 (2011) 252.
[27] O.A. Rosso, C. Masoller, Phys. Rev. E 79 (2009) 040106(R).
[28] O.A. Rosso, C. Masoller, European Phys. J. B 69 (2009) 37.
[29] K. Keller, M. Sinn, Physica A 356 (2005) 114.
[30] http://www.keithschwarz.com/interesting/code/factoradic-permutation/FactoradicPermutation.hh.html.
[31] J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, Europhys. Lett. 79 (2007) 50001.
[32] J.M. Amigó, Permutation Complexity in Dynamical Systems, Springer-Verlag, Berlin, Germany, 2010.
[33] L.C. Carpi, P.M. Saco, O.A. Rosso, Physica A 389 (2010) 2020.
[34] O.A. Rosso, L.C. Carpi, P.M. Saco, M. Gómez Ravettti, A. Plastino, H. Larrondo, Physica A 391 (2012) 42.
[35] H. Wold, A Study in the Analysis of Stationary Time Series, Almqvist and Wiksell, Upsala, Sweden, 1938.
[36] M.T. Martín, A. Plastino, O.A. Rosso, Physica A 369 (2006) 439.
[37] O.A. Rosso, L. De Micco, H. Larrondo, M.T. Martín, A. Plastino, Int. J. Bif. and Chaos 20 (2010) 775.
[38] L. De Micco, J.G. Fernndez, H.A. Larrondo, A. Plastino, O.A. Rosso, Physica A 391 (2012) 2564.