The Amigó paradigm of forbidden/missing patterns: a detailed analysis

O.A. Rosso1,2,3, a, L.C. Carpi4, P.M. Saco4, M.G. Ravetti5, H.A. Larrondo3,6, and A. Plastino3,7,8

1 LaCCAN/CPMAT - Instituto de Computação Universidade Federal de Alagoas BR 104 Norte km 97, 57072-970 Maceió, Alagoas, Brazil
2 Laboratorio de Sistemas Complejos, Facultad de Ingeniería, Universidad de Buenos Aires, 1063 Av. Paseo Colón 840, Ciudad Autónoma de Buenos Aires, Argentina
3 Fellow CONICET, 1033 Av. Rivadavia 1917, Buenos Aires, Argentina
4 Civil, Surveying and Environmental Engineering. The University of Newcastle, University Drive, Callaghan NSW 2308, Australia
5 Departamento de Engenharia de Produção, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte, 31270-901 Belo Horizonte - MG, Brazil
6 Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Av. J.B. Justo 4302, 7600 Mar del Plata, Argentina
7 Universidad Nacional de La Plata, IFLP-CCT, C.C. 727, 1900 La Plata, Argentina
8 Universitat de les Illes Balears and IFISC-CSIC, 07122 Palma de Mallorca, Spain

Received 11 April 2012 / Received in final form 15 September 2012
Published online 24 December 2012 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2012

Abstract. We deal here with the issue of determinism versus randomness in time series (TS), with the goal of identifying their relative importance in a given TS. To this end we extend (i) the use of ordinal patterns based probability distribution functions associated to a TS [C. Bandt and B. Pompe, Phys. Rev. Lett. 88, 174102 (2002)] and (ii) the so-called Amigó paradigm of forbidden/missing patterns [J.M. Amigó et al., Europhys. Lett. 79, 50001 (2007)], to analyze deterministic finite TS contaminated with strong additive noises of different correlation-degree. Useful information on the deterministic component of the original time series is obtained with the help of the so-called causal entropy-complexity plane [O.A. Rosso et al., Phys. Rev. Lett. 99, 154102 (2007)].

1 Introduction

1.1 Preliminaries

The distinction between deterministic and random components in time series has attracted considerable attention. From previous research we may single out here work by Osborne and Provenzale [1], Sugihara and May [2], and Kaplan and Glass [3,4]. In particular, Kantz et al. [5] and Cencini et al. [6] analyzed recently the behavior of entropy quantifiers as a function of the coarse-graining resolution, and applied their ideas to the issue of trying to distinguish between chaos and noise. Why is this a relevant issue? Because the concept of deterministic chaos, derived from the modern theory of nonlinear dynamical systems, has profoundly changed our thinking regarding time-series’ analysis. Emphasis is now being placed on non-linear approaches, i.e., with nonlinear deterministic autonomous equations of motion, representative of chaotic systems, that give birth to irregular signals [7].

Clearly, signals emerging from chaotic time series occupy an intermediate position between (a) predictable regular or quasi-periodic signals and, (b) totally irregular stochastic signals (noise) that are completely unpredictable. However, they exhibit interesting phase-space structures. Chaotic systems display “sensitivity to initial conditions” which are the origin of instability everywhere in phase-space. This implies that instability itself uncovers information about the phase-space “population”, not available otherwise [8]. In turn this leads us to think of chaos as an information source, whose associated rate of generated-information is formulated in precise fashion via the Kolmogorov-Sinai’s entropy [9,10]. These considerations motivate our present interest in the computation of quantifiers based on Information Theory, like, “entropy”, “statistical complexity”, “entropy-complexity plane”, etc. These quantifiers can be used to detect determinism in time series [11]. Indeed, different information theory

a e-mail: oarosso@gmail.com
based measures (normalized Shannon entropy and statistical complexity) allow for a better distinction between deterministic chaotic and stochastic dynamics when ever “causal” information is incorporated via the Bandt and Pompe’s (BP) methodology [12]. For a review of BP’s methodology and its applications to physics, biomedical and econophysic signals, see [13].

1.2 The new paradigm of forbidden patterns

When nonlinear dynamics are involved, a deterministic system can generate “random-looking” results that nevertheless exhibit persistent trends, cycles (both periodic and non-periodic) and long-term correlations. Our main interest here, lies in the emergence of “forbidden/missing patterns” [14–18]. Why? Because they have the potential ability for distinguishing deterministic behavior (chaos) from randomness in finite time series contaminated with observational white noise [14,15,18]. A concomitant helpful feature, as will be further explained below, is the decay rate of “missing ordinal patterns” as a function of the time series length.

In fact, Zanin [19] and Zunino et al. [20] have recently studied the appearance of missing ordinal patterns in financial time series. The presence of missing ordinal patterns has also been recently construed as evidence of deterministic dynamics in epileptic states. Ouyang et al. [21] found that a missing patterns’ quantifier could be used as a predictor of epileptic absence-seizures. Other interesting applications in biology time series analysis are given at [22–24]. It is essential to point out those works have only considered the presence of uncorrelated noise (white noise), which makes the associated results somewhat incomplete, since the presence of colored noise might be of importance. A first step in such direction was recently trodden by us in [25]. We pursue matters here by investigating the robustness of an associated mathematical construct called the causal entropy-complexity plane [11], that plays a prominent role in some of the above cited discoveries. We intend to do this by analyzing the plane’s ability to distinguish between noiseless chaotic time series and the ones that are contaminated (weekly to very strongly) with additive correlated noise. The chaotic series studied here were generated by recourse to a logistic map to which noise with varying amplitudes was added.

1.3 Our goal

We intend here to re-analyze noise contamination for finite time series, i.e., noise with power spectrum $f^{-k}$ and amplitude $A > 1$, via a careful exploration of the whole plane $(A,k)$. Our goal is to understand the planar structural details with regards to their implications concerning the Amigó paradigm of forbidden/missing patterns.

Our basic tools are (i) the MPR-statistical complexity measure together with entropic quantifiers; (ii) the Bandt-Pompe approach to extract causal probability distributions functions (PDFs) from a given time-series; and (iii) the construction of an entropy-complexity plane. These subjects are covered in great detail in [18,25,26]. For detailed explanations, minutiae, and results for $A < 1$, the reader unfamiliar with these themes is strongly advised to look at these references. As for the organization of this communication, the forthcoming Section 2 details the problem to be discussed, while Section 3 revisits the Amigó paradigm. Our results are presented and discussed in Section 4, while some conclusions are drawn in Section 5.

2 Logistic map plus observational noise

The logistic map constitutes a canonic example, often employed to illustrate new concepts and/or methods for the analysis of dynamical systems. Here we will use the logistic map with additive correlated noise in order to exemplify the behavior of the normalized Shannon entropy $\mathcal{H}_S$ and the MPR-complexity $C_S$, both evaluated using a PDF based on the Bandt-Pompe’s procedure. We will also investigate the behavior of “missing ordinal patterns” in both (i) the logistic map with additive noise (observational noise) and (ii) a pure-noise series, where the number of time series data was fixed at $N$.

The logistic map is a polynomial mapping of degree 2, $F: x_n \rightarrow x_{n+1}$ [27], described by the ecologically motivated, dissipative system represented by the first-order difference equation

$$x_{n+1} = rx_n(1 - x_n),$$

with $0 \leq x_n \leq 1$ and $0 \leq r \leq 4$.

Let $\eta^{(k)}$ be a correlated noise with $f^{-k}$ power spectra generated as described in [11,25]. The steps to be followed are enumerated below.

1. Using the Mersenne twister generator [28] through the MATLAB® rand function we generate pseudo random numbers $y_i^0$ in the interval $(-0.5, 0.5)$ with an (i) almost flat power spectra (PS), (ii) uniform PDF, and (iii) zero mean value.

2. The Fast Fourier Transform (FFT) $y_i^1$ is first obtained and then multiplied by $f^{-k/2}$, yielding $y_i^2$. Then, $y_i^2$ is symmetrized so as to obtain a real function. Subsequently the pertinent inverse FFT is found, after discarding the small imaginary components produced by the numerical approximations. The resulting noisy time series is then re-scaled to the interval $[-1, 1]$, which produces a new time series $\eta^{(k)}$ that exhibits the desired power spectra and, by construction, is representative of non-Gaussian noises [29].

We consider time series of the form $S = \{S_n, n = 1, \ldots, N\}$ generated by the discrete system:

$$S_n = x_n + A \cdot \eta_n^{(k)},$$

in which, $x_n$ is given by the logistic map and $\eta_n^{(k)} \in [-1, 1]$ represents a noise with power spectrum $f^{-k}$ and amplitude $A$. 
In generating the logistic map’s component of our time-series we fix \( r = 4 \) and start the iteration procedure with a random initial condition. The first \( 5 \times 10^4 \) iterations are considered part of the transient behavior and discarded. After the transient part dies out, \( N = 10^5 \) values are generated. In the case of the stochastic component of our time-series we consider \( 0 \leq k \leq 2 \) with \( k \)-values changing by an amount \( \Delta k = 1 \) (other intermediate values have been considered in [25]). The noise amplitude was varied in order to cover different regimes from weak (the noise can be considered as perturbation to the logistic) to very strong (the logistic is consider a perturbation to the noise). Ten noisy time-series of length \( N = 10^5 \) values and using different seeds, are generated for each value of the \( k \)-exponent.

3 The Amigó-paradigm

We arrive at the central point of our discussion. For deterministic one-dimensional maps, Amigó et al. [14–17] have conclusively shown that not all the possible ordinal patterns (as defined using Bandt and Pompe’s methodology) can be effectively materialized 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 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 \( N \). It must be noted, that this independence does not characterize other properties of the series such as proximity and correlation, which die out with time [15,17]. For example, in the time series generated by the logistic map \( x_{k+1} = 4x_k(1-x_k) \), 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 [15].

Stochastic processes could also have forbidden patterns [25]. However, in the case of uncorrelated (white noise) or certain correlated stochastic processes (noise with power law 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 [14–17]. If the time series is long enough, all the ordinal patterns should eventually appear. If the number of time-series’ observations is sufficiently big, the associated probability distribution function should be the uniform distribution, and the number of observed patterns should depend only on the length \( N \) of the time series under study.

For correlated stochastic processes the probability of observing individual patterns depends not only on the correlation structure but also on the time series length \( N \) [18]. The existence of a non-observed ordinal pattern does not qualify it as “forbidden”, only as “missing”, and is due to the finite length of the time series. A similar observation also holds for the case of real data series, as they always possess a stochastic component due to the omnipresence of dynamical noise [30–32]. 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.

The Carpi-Amigó test

Amigó et al. [14,15] proposed a test that uses missing ordinal patterns to distinguish determinism (chaos) from pure randomness in finite time series contaminated with observational white noise (uncorrelated noise). The concomitant methodology [15] involves a graphic comparison between:

- the decay rate of the missing ordinal patterns (of length \( D \)) of the time series under analysis as a function of the series length \( N \), and
- the decay rate exhibited by white Gaussian noise.

This methodology was recently extended by Carpi et al. [18] for the analysis of missing ordinal patterns in stochastic processes with different degrees of correlation. We are speaking of fractional Brownian motion (fBm), fractional Gaussian noise (fGn), and noises with \( f^{-k} \) power spectrum (PS) and \( k \geq 0 \). Results show that for a fixed pattern length, the decay rate of missing ordinal patterns in stochastic processes depends not only on the series length but also on their correlation structures. In other words, missing ordinal patterns are more persistent in the time series with higher correlation structures. Carpi et al. [18] have also shown that the standard deviation of the estimated decay rate of missing ordinal patterns (\( \alpha \)) decreases with increasing \( D \). This is due to the fact that longer patterns contain more temporal information and are therefore more effective in capturing the dynamic of time series with correlation structures.

An important quantity for us, called \( \mathcal{M}(N,D) \), is the number of missing ordinal patterns of length \( D \) not observed in a time series with \( N \) values. As we mentioned before, for correlated stochastic processes the probability of observing an individual pattern of length \( D \) depends on the time series-length \( N \) and on the correlation structure (as determined by the type of noise \( k > 0 \)). In fact, as established for noises with an \( f^{-k} \) PS in [18], as the value of \( k > 0 \) augments – which implies that correlations grow – increasing values of \( N \) are needed in reaching the “ideal” condition \( \mathcal{M}(N,D) = 0 \). If the time series is chaotic but has an additive stochastic component, then one expects that as the time series’ length \( N \) increases, the number of “missing ordinal patterns” will decrease and eventually vanish. That this may happen or not is independent of the length \( N \), the underlying deterministic components of the time series, or the correlation-structure of the added noise. The number of missing ordinal patterns will be \( \mathcal{M}(N,D,k) \), where \( k \) is the noise’s characteristic parameter, representing its correlation degree.
4 Results

We wish to analyze and characterize the behavior of finite, noisy and chaotic time-series by recourse to patterns generated in the (causal) entropy-complexity plane. We intend to assess in particular the planar-geography of the forbidden patterns. The focus of the analysis is centered upon the roles of: (i) the noise amplitude $A$; (ii) the type of contaminating noise (degree of correlation). We consider noises with $f^{-k}$-power spectrum; (iii) the time series length $N$. Here we consider finite time series, with fixed length $N$.

We consider correlated noises as observational contaminations (additive noises). They introduce a dependence with $N$ in the pertinent results. As previously stated, if the noise correlations increase, we need $N$ to be larger for reaching the ideal condition $M(N,D,k) = 0$ with $k \neq 0$ [18].

For our present analysis, we fix the pattern-length at $D = 6$, the embedding time lag at $\tau = 1$, and the time series length at $N = 10^5$ values, for each one of the ten time series generated (see Eq. (2)) and for each $(A,k,N)$ value, with $N$ fixed, the normalized Shannon entropy $\mathcal{H}_S$ and the MPR-statistical complexity $C_{JS}$ were evaluated using the Bandt and Pompe PDFs. We deal with additive (observational) noise with $k = 0, 1, \text{and }2$, and analyze their position (each point results from an average over 10 different series) in the $\mathcal{H} \times C$-planar graphs. A previous study has extensively dealt with the subject, but only in a special scenario in which the noise is of perturbative character [25]. Here we deal with the behavior of $(A,k,N)$ in the planar-graphs, way beyond the perturbative $A$-zone, by considering also $A > 1$ values.

4.1 Case $0 \leq A \leq 1$

Figure 1 summarizes the planar behavior of the noisy chaotic time series considered (the logistic map with additive correlated noise with $f^{-k}$ PS), for a time series length of $N = 10^5$ data, considering noise amplitudes $0 \leq A \leq 1$ ($\Delta A = 0.1$) and $k = 0$ (uncorrelated), $k = 1, 2$ (correlated), see also Figure 8 in [25]. Clearly, given the noise-amplitude values here considered, the noise has a mere perturbative role vis-à-vis of the logistic map’s chaotic behavior.

For $A = 0$ we obtain the purely deterministic value corresponding to the logistic time-series, localized approximately at a medium-high entropic coordinate, near the highest possible complexity value. Note that such is the typical behavior observed for deterministic systems [11]. For a purely uncorrelated stochastic process ($k = 0$) we have $\mathcal{H}_S = 1$ and $C_{JS} = 0$. The correlated (colored) stochastic processes ($k \neq 0$) yield points located at intermediate values between the curves $C_{\text{min}}$ and $C_{\text{max}}$, with decreasing values of entropy and increasing values of complexity as $k$ grows [11] (see Fig. 1, open symbols). From this figure and, also from our previous work [25], we see that if $k \approx 0$, for increasing values of the amplitude $A$, entropy and complexity values change (starting from the value corresponding to the pure logistic series, i.e., $A = 0$) with a tendency to approach the values corresponding to pure noise, that is, $(\mathcal{H}_S \approx 1$ and $C_{JS} \approx 0$). Similar behavior is observed when correlated noises are considered, $k \neq 0$.

The main effect of the additive noise ($A \neq 0$) is to shift the point representative of zero noise ($A = 0$) towards increasing values of entropy $\mathcal{H}_S$ and decreasing values of complexity $C_{JS}$. This shift defines a kind of “trajectory” (curve) in the $\mathcal{H} \times C$-plane, that is located in the vicinity of the maximum complexity $C_{\text{max}}$-curve. Moreover, we found (by looking at other chaotic maps) that this trajectory is characteristic of the dynamical system under analysis. Such behavior can be linked to the persistence of forbidden patterns because they, in turn, imply that the deterministic logistic component is still operative and influencing the time series’ behavior [25]. As $k$ increases, the dependence on the noise-amplitude $A (0 \leq A \leq 1)$ tends to become attenuated, implying that $M(N,D,k) \neq 0$ (see Fig. 4 of [25]), while it almost disappears for $k \approx 2$. This fact is due to (i) the combined effect of the “coloring” correlations of the noise, that grow with increasing values of $k$ and, (ii) the finite time series’ length considered.

4.2 Case $A \geq 1$

Here we are interested in the physics associated to $A$’s growth, entailing an interchange between the roles of the logistic map and of the noise because a large $A$ makes noise the dominant feature. Several effects should be considered:

- noise contamination level (represented by $A$);
- noise correlation (represented by $k$);

![Fig. 1. (Color online) The causality entropy-complexity plane for all time series values (mean) corresponding to noise amplitude $0 \leq A \leq 1$ ($\Delta A = 0.1$) and $k = 0, 1, 2$. Values for strictly noisy time series are also shown as open symbols. The values shown were obtained from times series of length $N = 10^5$, and Bandt-Pompe parameters $D = 6$, $\tau = 1$. The lines represent the values of minimum and maximum statistical complexity $C_{\text{min}}$ and $C_{\text{max}}$ evaluated for the case of pattern length $D = 6$.](image-url)
the points corresponding to increasing values of $A$ continue to move in the plane towards the site representative of pure noise, with $\mathcal{M}(N, D, k) = 0$. The perturbative character of the noise is, of course, lost. This behavior is apparent in Figure 3, where the corresponding Bandt-Pompe PDF’s, for a typical noisy chaotic time-series, are displayed for increasing noise amplitudes $A$. The corresponding value of $\mathcal{M}(N, D, k)$ can also be observed in these graphs.

At the top-left corner of Figure 3 we depict the Bandt-Pompe’s PDF for the unperturbed logistic map ($A = 0$). The presence of forbidden patterns (that have probabilities $p_i = 0$) is clearly visible. In the same plot, at the bottom-right corner, we display the pure noise $k = 0$ (white noise) PDF. No forbidden patterns exist now, and we observe the characteristic flat distribution, $p_i \approx 1/M$ for all patterns $i = 1, \ldots, M$ ($M = D! = 720$). The noise-influence on the PDFs is clearly displayed in this graph (see also the plots for $k = 1$ and $k = 2$, Figs. 5 and 7). It can be gathered from this figure that the PDF evolves from that of the logistic map towards that of the pure-noise as one increases the value of $A$. The noise-effect also destroys the forbidden character of some patterns, whose associated probabilities grow slowly with the control parameter $A$, a feature that emphasizes the presence of a deterministic component in the noisy time series, even for $\mathcal{M}(N, D, k) = 0$, if $A$ is not too large. This effect disappears for very strong noise. Note that for $A = 5$ the PDF is practically equal to that of pure noise (see Fig. 3). The two PDFs are almost indistinguishable at $A = 10$, and their planar localizations coincide (see Fig. 2).

In Figures 4 and 6 one can appreciate details of the planar-localization of noisy chaotic time-series (mean values taken over ten realizations) that emerge out of the noise-contaminated logistic map. We display values for correlated noise with $k = 1$ and $k = 2$, with increasing noise amplitude $A \geq 1$. In these plots we also observe the planar locations of both the time-series for the unperturbed logistic map ($A = 0$) and the pure-correlated noise ($k = 1, 2$). As in the previous case of uncorrelated noise ($k = 0$), the planar locations move towards the pure-noise’s site (in the present cases, $k = 1$ and $k = 2$—pure noise, respectively) for increasing values of the noise-amplitude $A$. As noise correlation increases, large values of the noise amplitude ($A \geq 1$) need to be considered in order to see displacements in the $\mathcal{H} \times \mathcal{C}$-plane (compare the $A$-values for $k = 1$ and $k = 2$).

New behaviors emerge, as reflected by Figures 4 and 6. For varying noise-amplitude values we appreciate a trajectory formed by two curves: (i) in the first one, we see a displacement of the associated planar locations from left to right (starting form the unperturbed value $A = 0$), until reaching a point associated with a critical value $A_c$: $A_c \approx 7$ for $k = 1$ (see Fig. 4) and $A_c \approx 80$ for $k = 2$ (see Fig. 6); (ii) at these critical values the trajectory reverses its direction (now it moves from right to left, “below” the first curve) converging to the planar location of the pure-correlated noise. Identical behavior was observed for all correlated noises with $k$-values in the interval $0 < k < 2$.

---

**Fig. 2.** (Color online) (a) The causality entropy-complexity plane for all time series values (mean) corresponding to increasing values of noise amplitude $A$ and $k = 0$. Value for the case of strictly noisy time series is shown as open symbol. The values shown were obtained from times series of length $N = 10^5$, and Bandt-Pompe parameters $D = 6$, $\tau = 1$. The lines represent the values of minimum and maximum statistical complexity $C_{\text{min}}$ and $C_{\text{max}}$ evaluated for the case of pattern length $D = 6$. (b) Detail for high entropy values.

- persistence of the forbidden pattern;
- time series length $N$.

We should expect that as $A$ grows (with fixed values for $N$ and $D$), the missing-patterns numbers would diminish and eventually vanish. With strong or very strong noise we expect zero missing patterns and eventual predominance of stochastic features. This entails $\mathcal{M}(N, D, k) = 0$ independently of the $A$ value. However, taking into account our previous results [25], we expect that some new specific features will be revealed by the use of the causality entropy-complexity plane.

In the following discussion the time series length is fixed at $N = 10^5$ data. Consider our results for uncorrelated noise $k = 0$ and increasing values of noise-amplitude $A$ depicted in Figure 2. This graph reveals that for $A \geq 1$...
Fig. 3. (Color online) The Bandt-Pompe PDF ($M = 720$) for a typical noisy chaotic time series considered (uncorrelated noise $k = 0$), for different increasing values of the noise amplitude $A$. In the top-left corner the PDF corresponding to the unperturbed logistic map ($A = 0$) is presented. In the bottom-right corner, the PDF for a pure noise $k = 0$ time series is displayed. The values shown were obtained from times series of length $N = 10^5$, and Bandt-Pompe parameters $D = 6, \tau = 1$. 
that exhibit specific critical values of $A_c$. Clearly this noise amplitude critical value depends on the noise correlation, as well as on the time series’ length. We have $A_c(N, k)$.

Note that at $A_c$-value the normalized Shannon entropy, $\mathcal{H}$ and the MPR-complexity, $\mathcal{C} = \mathcal{H} \cdot \mathcal{Q}$, take their highest and lowest values, respectively. As the disequilibrium $Q[P, P_c] = Q_0 \cdot J[P, P_c]$ [11], defined in terms of Jensen-Shannon divergence is always positive ($J \geq 0$), the corresponding BP-PDF at $A_c$ is the one that most approximates to the uniform probability distribution ($P_c$). This behavior can clearly be seen in Figures 5 and 7, by looking at the PDF corresponding to $A_c = 7$ for $k = 1$ and $A_c = 80$ for $k = 2$.

The planar behavior of the noisy, chaotic time-series can be interpreted in terms of two main interacting effects, namely, those associated with (a) the persistence and robustness of the forbidden patterns in the deterministic chaotic component of the time-series and, (b) the correlations characterizing stochastic observational noise (which in the present cases do not display forbidden patterns, even in the presence of missing patterns, due to the finite length of the time series). These correlations grow with increasing values of $k > 0$. Looking to the associated Bandt-Pompe’s PDFs behavior, depicted in Figures 5 and 7 for different values of $A$, the observed planar trajectories can be succinctly described as itemized below:

- **amplitude range $0 < A \leq A_c$**: we deal with two scenarios, one (Case A) in which the correlated noise acts as a perturbation, with typical values $A \leq 1$ and $\mathcal{M}(N, D, k) \neq 0$, although the instance $\mathcal{M}(N, D, k) = 0$ can also be encountered. In Case B, the logistic and the correlated noise have the same hierarchy, and for this regime we have always $\mathcal{M}(N, D, k) = 0$. In terms of the trajectory in the $\mathcal{H} \times \mathcal{C}$-plane, the transition between these two regimes, perturbative (case A) and “same-hierarchy” (case B), is smooth. A better insight, for a specific $A$-value for the transition, can be gained by looking at the BP-PDF graphs. Figure 6b shows that for $k = 2$ and case A the entropic and complexities’ dispersion (given by the standard deviation) are both small and due to the perturbative noise. Contrariwise, for case B the dispersion values increase with the noise amplitude $A$.

- **amplitude range $A \geq A_c$**: now the correlated noise dominates. The deterministic component of the time series can be considered here as a perturbation. Again, for this range of noise amplitudes we have $\mathcal{M}(N, D, k) = 0$. We also observe that the dispersions of the entropy $\mathcal{H}_s$ and the complexity $\mathcal{C}_{JS}$ are small (see Fig. 6b).

In order to check the dependence of the behavior previously described with the time series length value ($N$), we re-evaluated our quantifiers but now considering time series with increasing length. We consider time series with $N = 5 \times 10^5$ and $10^6$ data. The corresponding values for the quantifiers (mean value and standard deviation) are compared with the obtained ones for $N = 10^5$ in Tables 1 and 2 for uncorrelated ($k = 0$) and correlated ($k = 1$) noises, respectively. Similar results were obtained for the case $k = 2$ (not shown).

By increasing the time series length, the quantifiers values for the logistic map without noise ($A = 0$) and for the pure noise (mean values) change slightly (see Tabs. 1 and 2). The observed main effect by increasing $N$ is that now $\mathcal{M}(N, D, k) = 0$ for all $A \geq 1$. A small variation on the noise amplitude critical values, $A_c$ is observed. From Tables 1 and 2 we can conclude that, even when small variations in the quantifiers values are observed, the behavior described previously (results for time series length $N = 10^5$ data) persists and it is not due to the finite number of data in the time series considered. The main reason for the lack of sensible dependence of the quantifiers on...
Fig. 5. (Color online) The Bandt-Pompe PDF ($M = 720$) for a typical noisy chaotic time series considered (correlated noise $k = 1$), for different increasing values of the noise amplitude $A$. In the top-left corner the PDF corresponding to the unperturbed logistic map ($A = 0$) is presented. In the bottom-right corner, the PDF for a pure noise $k = 1$ time series is displayed. The values shown were obtained from times series of length $N = 10^5$, and Bandt-Pompe parameters $D = 6$, $\tau = 1$. 
the values of minimum and maximum statistical complexity as, statistical complexity, can be achieved by adequately selecting the sample’s size.

Therefore, a desired degree of accuracy for quantifiers like entropies and divergences, as well as, statistical complexity, can be achieved by adequately selecting the sample’s size.

\[ \langle H \rangle_{\text{max}} - \langle C \rangle_{\text{min}} \]

\[ \langle H \rangle_{\text{max}} \times \langle C \rangle_{\text{min}} \]

Table 1. Variation of the Information Theory based quantifiers, normalized Shannon entropy and MPR-statistical complexity, (mean values and standard deviation) evaluated with PDF-Bandt and Pompe (parameters \( D = 6 \) and \( \tau = 1 \)), for time series length \( N = 10^5 \), \( 5 \times 10^5 \), \( 10^6 \) for the logistic map contaminated with uncorrelated noise, \( k = 0 \) and amplitude \( A \). See also Figure 2.

| Noise amplitude | \( \langle H \rangle \) | SD  | \( \langle C \rangle \) | SD  |
|-----------------|-----------------|-----|-----------------|-----|
| \( A = 0 \)     | 0.63009         | 0.48451 |                 |     |
| \( A = 1 \)     | 0.97081         | 0.00089 | 0.06219         | 0.00186 |
| \( A = 5 \)     | 0.99789         | 0.00012 | 0.00497         | 0.00028 |
| \( A = 10 \)    | 0.99856         | 0.73 \times 10^{-4} | 0.00344 | 0.00017 |
| pure noise      | 0.99858         | 0.00083 | 0.00340         | 0.00020 |

| Noise amplitude | \( \langle H \rangle \) | SD  | \( \langle C \rangle \) | SD  |
|-----------------|-----------------|-----|-----------------|-----|
| \( A = 0 \)     | 0.62986         | 0.48439 |                 |     |
| \( A = 1 \)     | 0.97013         | 9.42 \times 10^{-4} | 0.06351 | 0.00192 |
| \( A = 5 \)     | 0.99744         | 0.67 \times 10^{-4} | 0.00601 | 1.57 \times 10^{-4} |
| \( A = 10 \)    | 0.99802         | 0.30 \times 10^{-4} | 0.00474 | 0.73 \times 10^{-4} |
| pure noise      | 0.99810         | 0.46 \times 10^{-4} | 0.00457 | 1.13 \times 10^{-4} |

\[ k = 0 \); time series length \( N = 5 \times 10^5 \); data

| Noise amplitude | \( \langle H \rangle \) | SD  | \( \langle C \rangle \) | SD  |
|-----------------|-----------------|-----|-----------------|-----|
| \( A = 0 \)     | 0.62965         | 0.48430 |                 |     |
| \( A = 1 \)     | 0.96976         | 9.03 \times 10^{-4} | 0.06423 | 0.00184 |
| \( A = 5 \)     | 0.99829         | 0.51 \times 10^{-4} | 0.00400 | 1.8 \times 10^{-4} |
| \( A = 10 \)    | 0.99901         | 0.24 \times 10^{-4} | 0.00236 | 0.59 \times 10^{-4} |
| pure noise      | 0.99908         | 0.31 \times 10^{-4} | 0.00222 | 0.76 \times 10^{-4} |

5 Conclusions

We have revisited in some detail the paradigmatic concept of forbidden/missing ordinal patterns and used it as a tool for distinguishing between deterministic and stochastic behavior in empiric time-series. In the spirit of [14–16,19–21], we extended this kind of analysis by linking it to the physics described by the causality entropy-complexity plane [11]. Our considerations were made with regards to deterministic, finite time series contaminated with additive noises of different degree of correlation. Our analysis of the noise-determinism competition clearly demonstrates that forbidden patterns are a deterministic feature of nonlinear systems and also reaffirms the usefulness of the entropy-complexity plane as a powerful tool of the theoretical arsenal.

The planar localization in the causal entropy-complexity plane describing an uncontaminated, nonlinear deterministic system is displaced, by the addition of noise, towards a zone typical of pure stochasticity. This displacement generates a kind of trajectory in the \( H \times C \)-plane that is located in the vicinity of the curve corresponding to maximum statistical complexity.

Fig. 6. (Color online) (a) The causality entropy-complexity plane for all time series values (mean) corresponding to increasing values of noise amplitude \( A \) and \( k = 2 \). Value for the case of strictly noisy time series is shown as open symbol. The values shown were obtained from times series of length \( N = 10^5 \), and Bandt-Pompe parameters \( D = 6 \), \( \tau = 1 \). The lines represent the values of minimum and maximum statistical complexity \( C_{\text{min}} \) and \( C_{\text{max}} \) evaluated for the case of pattern length \( D = 6 \).

(b) Same, showing mean value and standard deviation for each \( A \) considered value.

\( N \) is due to the fact that the BP-PDF used here constitutes an invariant PDF for the maps contaminated with observational noise. In fact, let’s consider a sample \( \mathbf{x} \) of \( N \) observations. By the Law of Large Numbers, the sample’s probability for any quantity \( \Psi(\mathbf{x}) \), converges to its distributional counterpart as the sample size \( N \) increases (see, for instance [33,34]). Therefore, a desired degree of accuracy for quantifiers like entropies and divergences, as well as, statistical complexity, can be achieved by adequately selecting the sample’s size.
Fig. 7. (Color online) The Bandt-Pompe PDF \((M = 720)\) for a typical noisy chaotic time series considered (correlated noise \(k = 2\)), for different increasing values of the noise amplitude \(A\). In the top-left corner the PDF corresponding to the unperturbed logistic map \((A = 0)\) is presented. In the bottom-right corner, the PDF for a pure noise \(k = 2\) time series is displayed. The values shown were obtained from times series of length \(N = 10^5\), and Bandt-Pompe parameters \(D = 6, \tau = 1\).
Table 2. Variation of the Information Theory based quantifiers, normalized Shannon entropy and MPR-statistical complexity, (mean values and standard deviation) evaluated with PDF-Bandt and Pompe (parameters $D = 6$ and $\tau = 1$), for time series length $N = 10^5$, $5 \times 10^5$, $10^6$ for the logistic map contaminated with correlated noise, $k = 1$ and amplitude $A$. See also Figure 4.

$k = 1$; time series length $N = 10^5$ data

| Noise amplitude | $\langle H \rangle$ | $SD$ | $\langle C \rangle$ | $SD$ |
|-----------------|-------------------|-----|-----------------|-----|
| $A = 0$          | 0.63009           | 0.48451 |                  |     |
| $A = 1$          | 0.85883           | 0.01063 | 0.26581         | 0.01562 |
| $A = 5$          | 0.97969           | 0.00145 | 0.04582         | 0.00305 |
| $A = 7$          | 0.98188           | 0.00048 | 0.04161         | 0.00111 |
| $A = 10$         | 0.98059           | 0.00068 | 0.04362         | 0.00134 |
| $A = 30$         | 0.97458           | 0.00068 | 0.05544         | 0.00136 |
| $A = 50$         | 0.97273           | 0.00062 | 0.05919         | 0.00124 |
| $A = 100$        | 0.97123           | 0.00057 | 0.06223         | 0.00114 |
| $A = 300$        | 0.97014           | 0.00045 | 0.06444         | 0.00092 |
| pure noise       | 0.96902           | 0.00039 | 0.06554         | 0.00075 |

$k = 1$; time series length $N = 5 \times 10^5$ data

| Noise amplitude | $\langle H \rangle$ | $SD$ | $\langle C \rangle$ | $SD$ |
|-----------------|-------------------|-----|-----------------|-----|
| $A = 0$          | 0.62986           | 0.48439 |                  |     |
| $A = 1$          | 0.84657           | 0.00657 | 0.28729         | 0.00917 |
| $A = 5$          | 0.97792           | 0.00132 | 0.04936         | 0.00270 |
| $A = 7$          | 0.98144           | 2.13 $\times 10^{-4}$ | 0.04209       | 4.41 $\times 10^{-4}$ |
| $A = 10$         | 0.98099           | 3.14 $\times 10^{-4}$ | 0.04289       | 6.27 $\times 10^{-4}$ |
| $A = 30$         | 0.97475           | 3.54 $\times 10^{-4}$ | 0.05523       | 7.40 $\times 10^{-4}$ |
| $A = 50$         | 0.97273           | 2.32 $\times 10^{-4}$ | 0.05931       | 4.98 $\times 10^{-4}$ |
| $A = 100$        | 0.97105           | 2.06 $\times 10^{-4}$ | 0.06268       | 4.51 $\times 10^{-4}$ |
| $A = 300$        | 0.96987           | 1.51 $\times 10^{-4}$ | 0.06508       | 3.47 $\times 10^{-4}$ |
| pure noise       | 0.96924           | 1.33 $\times 10^{-4}$ | 0.06637       | 2.97 $\times 10^{-4}$ |

$k = 1$; time series length $N = 10^6$ data

| Noise amplitude | $\langle H \rangle$ | $SD$ | $\langle C \rangle$ | $SD$ |
|-----------------|-------------------|-----|-----------------|-----|
| $A = 0$          | 0.62965           | 0.48430 |                  |     |
| $A = 1$          | 0.84003           | 0.00403 | 0.29634         | 0.00549 |
| $A = 5$          | 0.97678           | 9.31 $\times 10^{-4}$ | 0.05171       | 0.00191 |
| $A = 7$          | 0.98153           | 1.91 $\times 10^{-4}$ | 0.04186       | 4.11 $\times 10^{-4}$ |
| $A = 10$         | 0.98180           | 1.76 $\times 10^{-4}$ | 0.04111       | 3.59 $\times 10^{-4}$ |
| $A = 30$         | 0.97607           | 2.11 $\times 10^{-4}$ | 0.05227       | 4.27 $\times 10^{-4}$ |
| $A = 50$         | 0.97399           | 1.48 $\times 10^{-4}$ | 0.05645       | 2.97 $\times 10^{-4}$ |
| $A = 100$        | 0.97223           | 1.44 $\times 10^{-4}$ | 0.06000       | 2.98 $\times 10^{-4}$ |
| $A = 300$        | 0.97007           | 1.44 $\times 10^{-4}$ | 0.06257       | 2.95 $\times 10^{-4}$ |
| pure noise       | 0.97030           | 1.53 $\times 10^{-4}$ | 0.06394       | 3.14 $\times 10^{-4}$ |

If the noise is uncorrelated (white noise), the displacement of the systems characteristic point in the $H \times C$-plane becomes more pronounced as the noise intensity (amplitude) increases. It exhibits a monotonic decreasing behavior (from pure determinism to pure stochasticity). The noise-effect tends to eliminate forbidden patterns. Since these, in turn, are features of a deterministic dynamics, their destruction signals deterministic-nature’s loss.

For contaminating correlated noise a new type of planar trajectory-behavior emerges as the noise intensity increases. Starting from a pure deterministic localization, the trajectory converges to a pure stochastic localization by following a loop-curve as the noise intensity increases. For a critical value of noise intensity ($A_c$), however, this trajectory reverses direction. One observes a planar movement of the system’s characteristic point that starts at the unperturbed value ($A = 0$) and closely approaches the curve of maximum complexity, from left to right, with increasing entropic values. At the critical noise intensity ($A_c$) this displacement reverses direction. It takes place now from right to left, below the original curve and converging to the planar location typical of pure-correlated noise. The value of the critical intensity $A_c$ depends on the correlation degree of the noise and on the deterministic component.

Three different scenarios have been studied here:

(a) the correlated noise acting as a perturbation (mostly $\mathcal{M}(N, D, k) \neq 0$). The net noise effect is to destroy the forbidden character of some of the patterns. However due to the low noise intensity and to its correlations, the number of affected patterns is relatively low. The dominance of the deterministic component over the noisy one is reflected by low dispersion values for both the entropy and the statistical complexity.

(b) the deterministic and the stochastic components have the same hierarchy and $\mathcal{M}(N, D, k) = 0$. However, the persistent character of the forbidden patterns of the deterministic component and their interplay with the correlations present in the noise are reflected in the system’s characteristic point-trajectory, that moves along the curve of maximum complexity, indicating that the pertinent patterns do not appear as frequently as the remaining ones. This behavior is indicative of a still active deterministic dynamic. The dispersion values for our two quantifiers increase with the noise intensity. Note that the two scenarios described above correspond to a range of noise intensities given by $0 \leq A \leq A_c(N, k)$.

(c) the noisy component is the dominating one and the deterministic component can be considered as a perturbation. This scenario corresponds to the noise intensity range $A > A_c(N, k)$ and we have $\mathcal{M}(N, D, k) = 0$ as well, with low dispersion values for entropy and statistical complexity.

This research has been partially supported by a scholarship from The University of Newcastle awarded to Laura C. Carpi. Osvaldo A. Rosso gratefully acknowledges support from CNPq fellowship, Brazil. Martín Gómez Ravetti acknowledges support from FAPEMIG and CNPq, Brazil. Angelo Plastino acknowledges support from (i) the project PIP1177 of CONICET (Argentina) and (ii) the project FIS2008-00781/FIS (MICINN) and FEDER (EU) (Spain, EU). H.A. Larrondo acknowledges support from CONICET project PIP112-200801-140 (Argentina).
References

1. A.R. Osborne, A. Provenzale, Physica D 35, 357 (1989)
2. G. Sugihara, R.M. May, Nature 344, 734 (1983)
3. D.T. Kaplan, L. Glass, Phys. Rev. Lett. 68, 427 (1992)
4. D.T. Kaplan, L. Glass, Physica D 64, 431 (1993)
5. H. Kantz, E. Olbrich, Physica A 280, 34 (2000)
6. M. Cencini, M. Falcioni, E. Olbrich, H. Kantz, A. Vulpiani, Phys. Rev. E 62, 427 (2000)
7. H. Kantz, T. Scheiber, Nonlinear Time Series Analysis (Cambridge University Press, Cambridge, 2002)
8. H.D.I. Abarbanel, Analysis of Observed Chaotic Data (Springer-Verlag, New York, 1996)
9. A.N. Kolmogorov, Dokl. Akad. Nauk. USSR 119, 861 (1959)
10. Y.G. Sinai, Dokl. Akad. Nauk. USSR 124, 768 (1959)
11. O.A. Rosso, H.A. Larrondo, M.T. Martín, A. Plastino, M.A. Fuentes, Phys. Rev. Lett. 99, 154102 (2007)
12. C. Bandt, B. Pompe, Phys. Rev. Lett. 88, 174102 (2002)
13. M. Zanin, L. Zunino, O.A. Rosso, D. Papo, Entropy 14, 1553 (2012)
14. J.M. Amigó, L. Kocarev, J. Szczepanski, Phys. Lett. A 355, 27 (2006)
15. J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, Europhys. Lett. 79, 50001 (2007)
16. J.M. Amigó, S. Zambrano, M.A.F. Sanjuán, Europhys. Lett. 83, 60005 (2008)
17. J.M. Amigó, Permutation Complexity in Dynamical Systems (Springer-Verlag, Berlin, 2010)
18. L.C. Carpi, P.M. Saco, O.A. Rosso, Physica A 389, 2020 (2010)
19. M. Zunino, Chaos 18, 013119 (2008)
20. L. Zunino, M. Zanin, B.M. Tabak, D. Pérez, O.A. Rosso, Physica A 388, 2854 (2009)
21. G. Ouyang, X. Li, C. Dang, D.A. Richards, Phys. Rev. E 79, 041146 (2009)
22. D. Li, X. Li, Z. Liang, L.J. Voss, J.W. Sleigh, J. Neural Eng. 7, 046010 (2010)
23. K. Schindler, H. Gast, L. Stieglitz, A. Stibal, M. Hauf, R. West, L. Mariani, C. Rummel, Epilepsia 52, 1771 (2011)
24. U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, N. Wessel, Comput. Biol. Med. 42, 319 (2012)
25. O.A. Rosso, L.C. Carpi, M.P. Saco, M. Gómez Ravetti, A. Plastino, H.A. Larrondo, Physica A 391, 42 (2012)
26. O.A. Rosso, L. De Micco, H. Larrondo, M.T. Martín, A. Plastino, Int. J. Bifurc and Chaos 20, 775 (2010)
27. J.C. Sprott, Chaos and Time Series Analysis (Oxford University Press, Oxford, 2004)
28. M. Matsumoto, T. Nishimura, ACM Trans. Model. Comput. Simul. 8, 3 (1998)
29. H.A. Larrondo, Matlab program: noisefk.m (2012), http://www.mathworks.com/matlabcentral/fileexchange/35381
30. H. Wold, A Study in the Analysis of Stationary Time Series (Almqvist and Wiksell, Upsala, 1938)
31. J. Kurths, H. Herzel, Physica D 25, 165 (1987)
32. S. Cambanis, C.D. Hardin, A. Weron, Probab. Theory Relat. Fields 79, 1 (1988)
33. S.I. Resnick, A Probability Path (Birkhäuser, Boston, 1999)
34. C.F. Manski, Analog Estimation Methods in Econometrics, in Monographs on Statistics and Applied Probability (Chapman & Hall, New York, 1988), Vol. 39