

Distinguishing Noise from Chaos

O. A. Rosso,^{1,2} H. A. Larrondo,³ M. T. Martin,⁴ A. Plastino,⁴ and M. A. Fuentes^{5,6}

¹*Centre for Bioinformatics, Biomarker Discovery and Information-Based Medicine,
School of Electrical Engineering and Computer Science, The University of Newcastle,
University Drive, Callaghan NSW 2308, Australia*

²*Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,
1428 Ciudad Universitaria, Buenos Aires, Argentina*

³*Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Juan B. Justo 4302, 7600 Mar del Plata, Argentina*

⁴*Instituto de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 727, 1900 La Plata, Argentina*

⁵*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

⁶*Instituto Balseiro, Centro Atómico Bariloche and CONICET, 8400 Bariloche, Argentina*
(Received 15 June 2007; published 12 October 2007)

Chaotic systems share with stochastic processes several properties that make them almost undistinguishable. In this communication we introduce a representation space, to be called the complexity-entropy causality plane. Its horizontal and vertical axis are suitable functionals of the pertinent probability distribution, namely, the entropy of the system and an appropriate statistical complexity measure, respectively. These two functionals are evaluated using the Bandt-Pompe recipe to assign a probability distribution function to the time series generated by the system. Several well-known model-generated time series, usually regarded as being of either stochastic or chaotic nature, are analyzed so as to illustrate the approach. The main achievement of this communication is the possibility of clearly distinguishing between them in our representation space, something that is rather difficult otherwise.

DOI: 10.1103/PhysRevLett.99.154102

PACS numbers: 05.45.-a, 02.70.Rr, 05.40.-a, 89.70.+c

Although being of a quite different physical origin, time series arising from chaotic systems (CS) share with those generated by stochastic processes (SP) several properties that make them almost undistinguishable: (1) a wide-band power spectrum (PS), (2) a delta-like autocorrelation function, (3) an irregular behavior of the measured signals, etc. In fact this similitude has made it possible to replace SP by CS in many practical applications. We attempt here to distinguish between SP and CS by recourse to an appropriate representation whose starring role is played by a so-called complexity measure. We deal with well-known *models* that generate time series according to prespecified rules. This is to be contrasted with the situation posed by *real data*, however, that always possess a stochastic component due to omnipresent dynamical noise [1,2]. Indeed, Wold proved [1] that any (stationary) time series can be decomposed into two different parts. The first (deterministic) part can be exactly described by a linear combination of its own past; the second part is a moving average component of a finite order. Hence it may seem superfluous to ask whether a time series generated by “natural processes” is either deterministic, chaotic, or stochastic. However, having in mind Wold’s theorem [2] it makes sense to ask, with respect to the deterministic part (predictable from the past), whether (i) it is dominant vis-à-vis the unpredictable stochastic part or (ii) it is of a regular or chaotic nature. CS always produce time series with a physical structure. Looking for this physical structure is our leitmotif. In order to do this several statistical complexity measures have been recently introduced in the literature, based on the notion of the so-called “disequilibrium”

[3–5]. In Ref. [5] it was advanced a statistical complexity measure version that is (i) able to grasp essential details of the dynamics, (ii) an intensive quantity, and (iii) capable of discerning among different degrees of periodicity and chaos. This measure, to be referred to as the intensive statistical complexity measure $C_{JS}[P]$, is a functional of the probability distribution P associated with the time series. C_{JS} writes

$$C_{JS}[P] = Q_J[P, P_e]H_S[P], \quad (1)$$

associating, to the probability distribution $P = \{p_j; j = 1, \dots, N\}$, the entropic measure

$$H_S[P] = S[P]/S_{\max} \quad (2)$$

with $S_{\max} = S[P_e] = \ln N$, ($0 \leq H_S \leq 1$). We denote by $P_e = \{1/N, \dots, 1/N\}$ the uniform distribution while $S[P] = -\sum_{j=1}^N p_j \ln(p_j)$ stands for the Shannon’s entropy. Q_J , the above referred disequilibrium, is defined in terms of the extensive Jensen-Shannon divergence (it induces a squared metric, in contrast to the Kullback-Leiber divergence [5]) as follows:

$$Q_J[P, P_e] = Q_0\{S[(P + P_e)/2] - S[P]/2 - S[P_e]/2\}, \quad (3)$$

with Q_0 a normalization constant ($0 \leq Q_J \leq 1$). The disequilibrium Q_J is an intensive quantity being different from zero only if there exist “privileged,” or “more likely” states among the accessible ones [3–5]. A critical point is that of using the methodology proposed by Bandt and Pompe [6] for evaluating the probability distribution P associated to the time series under study. This requires

suitable partitions of a D -dimensional embedding space that will, it is hoped, reveal important details concerning the ordinal structure of a given one-dimensional time series. For the time series $\{x_t: t = 1, \dots, M\}$ and an embedding dimension $D > 1$, one is interested in “ordinal patterns” of order D [6,7] generated by

$$(s) \mapsto (x_{s-(D-1)}, x_{s-(D-2)}, \dots, x_{s-1}, x_s), \quad (4)$$

which assign to each time s a D -dimensional vector of values pertaining to the times $s, s-1, \dots, s-(D-1)$. By the ordinal pattern related to the time (s) we mean the permutation $\pi = (r_0, r_1, \dots, r_{D-1})$ of $(0, 1, \dots, D-1)$ defined by $x_{s-r_{D-1}} \leq x_{s-r_{D-2}} \leq \dots \leq x_{s-r_1} \leq x_{s-r_0}$. In order to get a unique result we consider that $r_i < r_{i-1}$ if $x_{s-r_i} = x_{s-r_{i-1}}$. Thus, for all the $D!$ possible permutations π of order D , the probability distribution $P = \{p(\pi)\}$ is defined by ($\mathcal{Y} = M - D + 1$)

$$p(\pi) = \# \{s | s \leq \mathcal{Y}; \quad (s) \text{ has type } \pi\} / \mathcal{Y}. \quad (5)$$

In this expression, the symbol $\#$ stands for “number.” The Bandt-Pompe method [6] for evaluating the probability distribution P is based on the details of the attractor-reconstruction procedure and causal information is duly incorporated in the construction process that yields $P \in \Omega$ (with Ω the probability space) [8]. A notable Bandt-Pompe result is a clear improvement in the performance of the information quantifiers obtained using their P -generating algorithm. One must assume with them that the system fulfills a very weak stationary condition (for $k < D$, the probability for $x_t < x_{t+k}$ should not depend on t) [6] and that enough data are available for a correct attractor reconstruction. The advantages of the Bandt-Pompe method reside in (a) its simplicity, (b) the concomitant extremely fast calculation process, (c) its robustness, and (d) its invariance with respect to nonlinear monotonous transformations. Also, it can be applied to any type of time series (regular, chaotic, noisy, or reality based) [6]. Finally, it is important to remark that calculations made with the Bandt-Pompe prescription are robust in the presence of observational and dynamical noise [6]. Of course, the embedding dimension D plays an important role for the evaluation of the appropriate probability distribution, since D determines the number of accessible states $D!$ and tells us about the necessary length M of the time series needed in order to work with a reliable statistics. Concerning this last point in all calculations reported here the condition $M \gg D!$ is satisfied. It is essential for our present purposes to consider *rather small* D values. In particular, Bandt and Pompe suggest for practical purposes working with $3 \leq D \leq 7$, and this is what we do here (in the present work we used $D = 6$). If we allow D to grow without bounds, significant (and important from a theoretic viewpoint) consequences ensue [10]. We remark that the above statistical complexity measure quantifies not only randomness but also the degree of correlational structures, and consequently it is not a trivial function of the entropy in the sense that, for a given

H_S value, there exists a range of possible C_{JS} ’s values between a minimum C_{\min} and a maximum C_{\max} [3]. Thus, evaluating C_{JS} provides one important additional piece of information regarding the peculiarities of a probability distribution, not already carried by the entropy. The additional information disappears if, for example, a probability density function (PDF) based on histograms is used. A general procedure for obtaining the bounds C_{\min} and C_{\max} corresponding to the statistical complexity measures’ family is given by Martín, Plastino, and Rosso in Refs. [4]. One may content oneself with using just Q_J , instead of C_{JS} . If so doing, the advantage of having a quantifier guaranteeing (as C_{JS} does) (1) a zero value for both regular and completely random series and (2) maximum values for systems with “immersed” (or hidden) structures [4], would be lost. In order to study the time evolution of C_{JS} , a diagram of C_{JS} versus H_S can be used, the CH plane (in this case, H_S can be regarded as an arrow of time [12]). Also, this kind of diagram has been used to study changes in a system’s dynamics originated by modifications of some characteristic parameters [3,5,13–16]. Processes here studied were selected as illustrative examples of (a) CS and (b) SP, two different classes of processes in the sense indicated in the introduction.

We dealt with the following five kinds of CS: (1) *The logistic map*: [17] defined by

$$x_{n+1} = rx_n(1 - x_n). \quad (6)$$

Note that for $r = 4$ this map has a nonuniform natural invariant PDF.

(2) *The skew tent map*: one has [17]

$$\begin{cases} x/\omega & \text{for } x \in [0, \omega] \\ (1-x)/(1-\omega) & \text{for } x \in [\omega, 1] \end{cases}. \quad (7)$$

For any ω value this map has a uniform natural invariant PDF ($\omega = 0.1847$ is here considered).

(3) *Henon’s map*: it is a 2D extension of the logistic map [17] given by

$$\begin{cases} x_{n+1} = 1 - ax_n^2 + y_n \\ y_{n+1} = bx_n \end{cases}. \quad (8)$$

The values used here, $a = 1.4$ and $b = 0.3$, correspond to a chaotic attractor with a nonsmooth PDF.

(4) *The Lorenz map of Rossler’s oscillator*: for the 3D continuous Rossler oscillator [17] one has

$$\begin{cases} \dot{x} = -y - z \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c) \end{cases}, \quad (9)$$

where $a = 0.2$, $b = 0.2$, and $c = 5.7$ correspond to a chaotic attractor. The Lorenz map is obtained by storing only x -minimal values [17].

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

$$x_{n+1} = x_n + x_n^z, \quad \text{Mod } 1. \quad (10)$$

In particular, results for $z = 5/2, 2$, and $3/2$ are reported.

We considered the following two kinds of SP here: (6) *Noises with f^{-k} PS* generated as follows: (a) The MATLAB®RAND function is used to produce pseudo random numbers in the interval $(-0.5, 0.5)$ with an (i) almost flat PS, (ii) uniform PDF, and (iii) zero mean value. (b) Then, the fast Fourier transform (FFT) y_k^1 is obtained and multiplied by $f^{-k/2}$, yielding y_k^2 ; (c) Now, y_k^2 is symmetrized so as to obtain a real function and then the pertinent inverse FFT x_i is obtained, after discarding the small imaginary components produced by our numerical approximations. The ensuing time series x_i has the desired PS and, by construction, is representative of non-Gaussian noises.

(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 Ref. [16] and references therein). The normalized family of these Gaussian processes, $\{B^{\mathcal{H}}(t), t > 0\}$, is endowed with these properties: (i) $B^{\mathcal{H}}(0) = 0$ almost surely, i.e., with probability 1, (ii) $\mathbb{E}[B^{\mathcal{H}}(t)] = 0$ (zero mean), and (iii) covariance given by

$$\mathbb{E}[B^{\mathcal{H}}(t)B^{\mathcal{H}}(s)] = (t^{2\mathcal{H}} + s^{2\mathcal{H}} - |t - s|^{2\mathcal{H}})/2 \quad (11)$$

for $s, t \in \mathbb{R}$. Here $\mathbb{E}[\cdot]$ refers to the average computed with a Gaussian PDF. The power exponent $0 < \mathcal{H} < 1$ is commonly known as the Hurst parameter (exponent). These processes exhibit “memory” for any Hurst parameter except for $\mathcal{H} = 1/2$, as one realizes from Eq. (11). 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 antipersistent. Let us introduce the quantity $\{W^{\mathcal{H}}(t), t > 0\}$ (FBM “increments”)

$$W^{\mathcal{H}}(t) = B^{\mathcal{H}}(t + 1) - B^{\mathcal{H}}(t), \quad (12)$$

so as to express our Gaussian noise in the fashion

$$\begin{aligned} \rho(k) &= \mathbb{E}[W^{\mathcal{H}}(t)W^{\mathcal{H}}(t+k)] \\ &= \frac{1}{2}[(k+1)^{2\mathcal{H}} - 2k^{2\mathcal{H}} + |k-1|^{2\mathcal{H}}], \quad k > 0. \end{aligned} \quad (13)$$

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 processes 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. Because of their Gaussian nature, and other characteristics above enumerated, the Bandt-Pompe ideas are applicable to the FBM and FGN dynamical process [18]. For evaluating the FBM and FGN time series we adopt the Davies-Harte algorithm [19], as recently improved by Wood and Chan [20], which is both exact and fast.

For all the cases we studied here 10 time series of 2^{15} data each were analyzed, each series starting at a different initial condition. The concomitant mean values of both H_S and C_{JS} are plotted in Fig. 1.

All the CS under scrutiny have entropies that, in our causality plane, are seen to be (i) in the entropy region lying between 0.45 and 0.7, (ii) located near to the maximum C_{JS} . This entails that high C_{JS} values are produced by structures immersed in chaotic time series. Higher H_S values may be obtained using randomizing techniques to

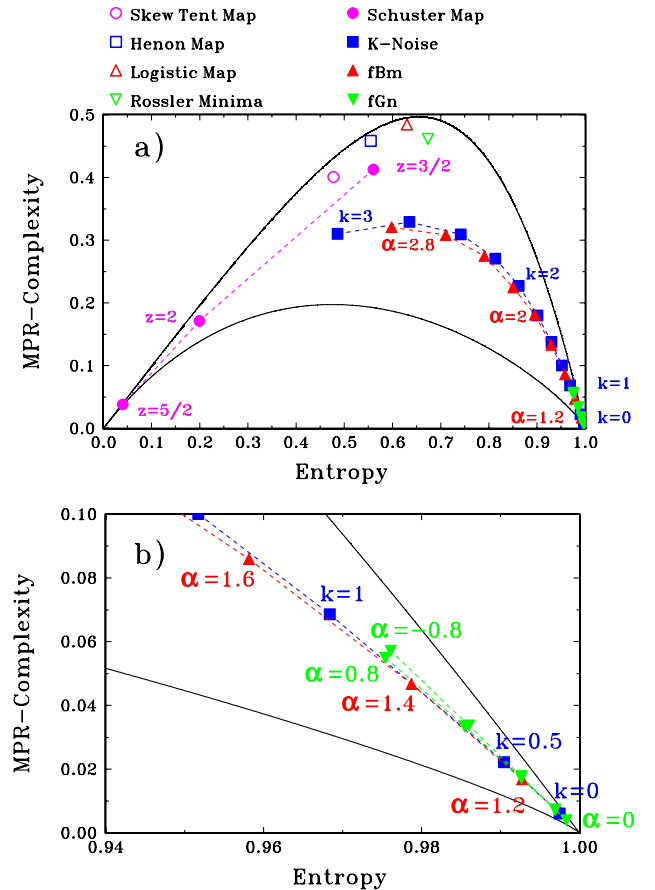


FIG. 1 (color online). Continuous lines represent minimum C_{\min} and maximum C_{\max} complexities. The area enclosed by them is the CH plane. (a) Localization of different CS and SP in the CH plane. (b) Enlargement near the ideal point $H_S = 1$, $C_{JS} = 0$. $D = 6$ is used. The graph illustrates the fact that, in the case of textbook models usually regarded as being of either stochastic or deterministic nature [17], our numerical results place them at clearly different planar locations.

increase the mixing and destroying correlational structures, as will be reported elsewhere. For the logistic map the effect of a control parameter $r < 4$ was also studied. We found that, in our causality plane, the pertinent points lie in a low entropy region, although they always remain near the maximum complexity curve, in agreement with what happens in the case of a binary PDF [5]. For the Henon map both the X and Y coordinates have the same ordinal structure and they have the same point in the CH plane. Schuster maps have also low entropy values $H_S < 0.6$. In this case the reason is that these maps exhibit laminar regions separated by chaotic bursts. Their complexity is lower than that of the chaotic case. When the parameter z decreases, H_S increases and the representative trajectory always remains below (but close) to that for chaotic maps. This is so because the size of laminar regions diminishes, entailing that the system becomes more similar to a fully chaotic one.

Noises with f^{-k} power spectrum, and $0 \leq k \leq 3$, exhibit medium-high entropy values ($0.45 < H_S < 1$) and C_{JS} values almost equidistant between the curves of maximum and minimum complexity. In particular, their C_{JS} are much lower than those for deterministic noises (Schuster and chaotic). For the small k values $k = 0$ and $k = 1$ they become almost ideal noises with $H_S \approx 1$ and $C_{JS} \approx 0$. As k increases, correlations among different values become apparent and, consequently, H_S decreases. FBM ($1 < \alpha < 3$) exhibits entropies near those of f^{-k} PS, but with a lower C_{JS} in relation to that of a non-Gaussian process. The associated FGN ($-1 < \alpha < 1$) has higher entropic values ($0.97 < H_S < 1$) and complexity values between 0 and 0.1. In addition these two kinds of value are higher in comparison to those for a f^{-k} PS [see Fig. 1(b)]. We associate this behavior with either the Gaussian or non-Gaussian nature of the respective processes. Ordinary Brownian motion ($\alpha = 2$) is characterized by a relative high entropy and low C_{JS} ($H_S \approx 0.9$ and $C_{JS} \approx 0.18$). Also, persistent FBM ($2 < \alpha < 3$)—long memory processes—are more complex than FBM antipersistent ($1 < \alpha < 2$)—short memory ones—in agreement with the intuitive idea for this kind of behavior. Complexity values for FGN are higher than those corresponding to a f^{-k} PS. In particular, persistent and antipersistent ($0 < |\alpha| < 1$) FGN display quite similar values [see Fig. 1(b)]. Maximum entropy and minimum complexity values are observed for $\alpha = 0$, which corresponds to white Gaussian noise. Note that, in the causality plane, this situation can be located in a position that lies below that corresponding to the case $k = 0$ of f^{-k} PS.

Summing up, our representation plane is here shown, with regards to signals generated by well-known models, to accommodate noise and chaos at different planar locations. Such representational property could be useful when dealing with real data (that always have a stochastic component due to omnipresent dynamical noise) so as to classify different degrees of “stochasticness.”

Our representation also distinguishes (a) Gaussian from non-Gaussian process and, (b) among different degrees of correlations (colored noises). Consequently, this representation plane is an effective tool for revealing the sometimes subtle difference between noise and chaos.

This work was partially supported by the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina (Nos. PIP 5687/05, PIP 6036/05) and ANPCyT, Argentina (No. PICT 11-21409/04). O. A. R. gratefully acknowledge support from Australian Research Council (ARC) Centre of Excellence in Bioinformatics, Australia.

-
- [1] H. Wold, *A Study in the Analysis of Stationary Time Series* (Almqvist and Wiksell, Upsala, Sweden, 1938).
 - [2] J. Kurths and H. Herzel, *Physica (Amsterdam)* **25D**, 165 (1987); S. Cambanis, C. D. Hardin, and A. Weron, *Probab. Theory Relat. Fields* **79**, 1 (1988).
 - [3] R. López-Ruiz, H. L. Mancini, and X. Calbet, *Phys. Lett. A* **209**, 321 (1995).
 - [4] M. T. Martín, A. Plastino, and O. A. Rosso, *Phys. Lett. A* **311**, 126 (2003); *Physica (Amsterdam)* **369A**, 439 (2006); O. A. Rosso, M. T. Martín, and A. Plastino, *ibid.* **347A**, 444 (2005).
 - [5] P. W. Lamberti, M. T. Martín, A. Plastino, and O. A. Rosso, *Physica (Amsterdam)* **334A**, 119 (2004).
 - [6] C. Bandt and B. Pompe, *Phys. Rev. Lett.* **88**, 174102 (2002).
 - [7] K. Keller and H. Lauffer, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **13**, 2657 (2003).
 - [8] The BP probability distribution is the only one that takes into account the temporal structure of the time series generated by the physical process under study [9].
 - [9] H. A. Larrondo, M. T. Martín, C. M. González, A. Plastino, and O. A. Rosso, *Phys. Lett. A* **352**, 421 (2006).
 - [10] The BP entropy [6] yields, for any chaotic system, the Kolmogorov-Sinai entropy [11] and one easily proves that $\lim_{D \rightarrow \infty} H_S \rightarrow 0$ and $\lim_{D \rightarrow \infty} C_{JS} \rightarrow 0$.
 - [11] Bandt *et al.*, *Nonlinearity* **15**, 1595 (2002); Amigó *et al.*, *Physica (Amsterdam)* **210D**, 77 (2005).
 - [12] A. R. Plastino and A. Plastino, *Phys. Rev. E* **54**, 4423 (1996).
 - [13] X. Calbet and R. López-Ruiz, *Phys. Rev. E* **63**, 066116 (2001).
 - [14] A. M. Kowalski, M. T. Martín, A. Plastino, and O. A. Rosso, *Int. J. Mod. Phys. B* **19**, 2273 (2005); *Physica (Amsterdam)* **233D**, 21 (2007).
 - [15] O. A. Rosso, M. T. Martín, A. Figliola, K. Keller, and A. Plastino, *J. Neurosci. Methods* **153**, 163 (2006).
 - [16] L. Zunino, D. G. Pérez, M. T. Martín, A. Plastino, M. Garavaglia, and O. A. Rosso, *Phys. Rev. E* **75**, 021115 (2007).
 - [17] H. G. Schuster, *Deterministic Chaos* (VCH, Weinheim, 1988) 2nd ed.
 - [18] C. Bandt and F. Shiha, *J. Time Ser. Anal.* **28**, 646 (2007).
 - [19] R. B. Davies and D. S. Harte, *Biometrika* **74**, 95 (1987).
 - [20] A. T. A. Wood and G. Chan, *J. Comput. Graph. Stat.* **3**, 409 (1994).