Send Orders for Reprints to reprints@benthamscience.net
Concepts and Recent Advances in Generalized Information Measures and Statistics, **2013, 283-308** **283**

CHAPTER 12

Statistical Complexity of Chaotic Pseudorandom Number Generators

Hilda A. Larrondo [a,e,*], Luciana De Micco [a,e], Claudio M. González [a], Angelo Plastino [b,e], Osvaldo A. Rosso [c,d,e]

[a] Facultad de Ingeniería, Universidad Nacional de Mar del Plata, Juan B. Justo 4302, 7600 Mar del Plata, Argentina

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

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

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

[e] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Abstract

This chapter deals with the use of the Statistical Complexity Measure, as defined by López Ruiz, Mancini and Calbet [Phys. Lett. A 209 (1995) 321–326] and modified by Rosso and coworkers [P. W. Lamberti, M. T. Martín, A. Plastino, O. A. Rosso; Physica A 334 (2004) 119–131] to characterize pseudo random number generators (*PRNG's*) obtained from chaotic dynamical systems. It is shown that two probability distribution functions are required for a proper characterization: the first one is based on the histogram and is used to characterize the uniformity of the values in the time series; the second one is based on the permutation procedure proposed by Bandt and Pompe [Phys. Rev. Lett. 88 (2002) 174102] and characterize the uniformity of patterns of several consecutive values of the time series.

Keywords: Chaos; Random number generators; Entropy; Statistical Complexity.

* Corresponding author. E-mail: larrondo@fi.mdp.edu.ar

1 Introduction

Let us briefly discuss at the very beginning of this chapter five controversial questions one wants to answer.

(1) Is the world deterministic or stochastic?

In his famous lectures Feynman describes a physicist as an observer watching a chess game, making models and experiments to discover the laws behind the players' moves. This vision implies that those rules really exist and our limited knowledge precludes us to know them in a complete way.

Statistical mechanics is founded on the opposite view. The physicist observes the game but they do not care about the existence of detailed rules. Predictions are made as a list of possible events and the probabilities for each one. A lot of theoretical and practical knowledge came from these ideas, and Statistical Mechanics appears today as one of the most fruitful branches of physics.

Two theories put this controversy between determinism and stochasticity on its top: quantum mechanics and deterministic chaos. The last one will be the subject of this chapter. Even if being from a quite different physical origin, time series arising from chaotic systems share with those generated by stochastic processes several properties that make them almost indistinguishable: *a)* a wide-band power spectrum, *b)* a delta-link linear autocorrelation function, *c)* an irregular behavior of the measures signals. In fact this similitude has made it possible to replace stochastic process by chaotic systems in many practical applications.

Wold proved 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 (random) is a moving average component of a finite order [1]. Chaotic systems always produce time series with a physical structure. This structure may be discovered and measured by statistical complexity measures [2] (see Chapter 8). In Ref. [3], time series generated by different chaotic and stochastic systems were studied. It was shown that using the Bandt and Pompe procedure to assign a probability distribution function (PDF) to the time series [4], it is possible to distinguish the chaotic systems from stochastic ones (*i.e.*, correlated and uncorrelated noises) when they are localized in the entropy-complexity plane. In point of fact, chaos is representative of deterministic processes in which time-causality constitutes an important fact that must be taken into account for a proper characterization.

(2) What are random number generators (RNG) and pseudo random number generators

(PRNG)?

Randomness is the opposite of deterministic. An ideal random number generator must produce numbers from a set of possible values with two basic properties:

- the values must be equally probable, *i.e.*, the sequence of numbers must have a uniform histogram;
- each number of the sequence must be statistically independent of the others, *i.e.*, it is impossible to predict the next value on the basis of the previous ones.

Mathematical algorithms can only produce pseudo random numbers. Some researches speculate that physical sources can produce truly random numbers. Some remarkable websites to get *physical* random numbers are [5], where random numbers are produced by measuring atmospheric noise, [6] that uses a radioactive decay, and [7] where a charge-coupled device (CCD) inside a dark can. But do physical noise sources produce true random numbers? If we accept a deterministic model of the universe the answer is no. The only difference between physical or algorithmic time series is that the putative underlying model may be unknown in the physical case, but in fact algorithmic time series may also be underlaid by a model that those who analyze them are unaware of.

The relevant issue in our opinion is to analyze the time series and quantify its statistical characteristics. The existence or not of a deterministic model is not relevant at all. This point of view was proposed by Marsaglia [8]. Quoting him, his DIEHARD famous pseudorandom number generator was *produced by combining two or more of the most promising deterministic generators with sources of random noise from three physical devices for those who feel that physical sources of randomness are better than deterministic sources.* But let us stress that the statistical quality of DIEHARD is governed by the deterministic part!

Another important issue is the alphabet of the source of random numbers. In this review we will consider the case of a finite alphabet with N different simbols (numbers), because that is the case in all the applications, as far as a computer has a finite precision and it is also impossible to measure a noise source with infinite precision.

(3) Can a chaotic dynamical system behave like and acceptable PRNG?

Here we are interested in chaotic PRNG. The interesting thing with deterministic chaos is that it faces us with a new situation: in spite the model is deterministic and simple, an unpredictable behavior of the variables may arise due to the sensitivity to initial conditions.

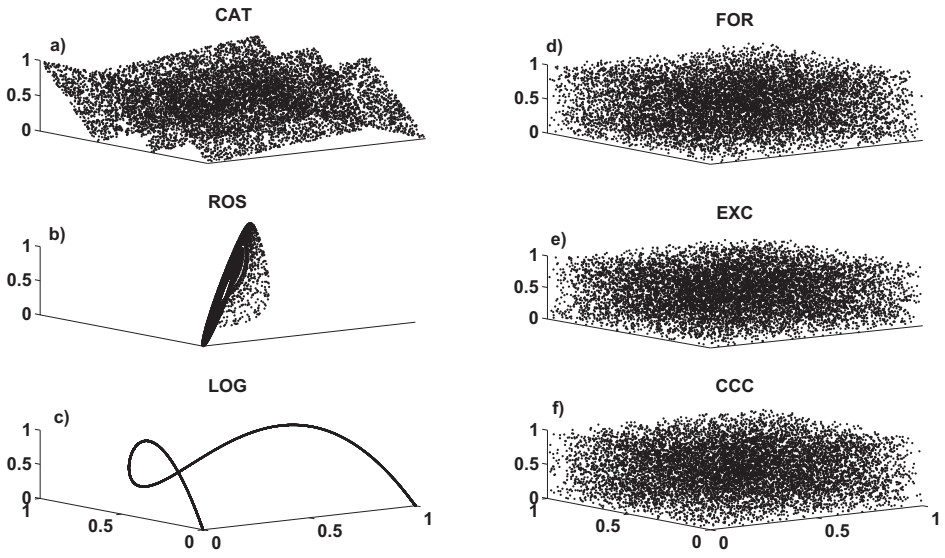

Figure 1: 3D distribution for for three chaotic (see (a) to (c)) and, three non-chaotic ((d) to (f)) PRNG's.

Then statistical mechanics is the best theory to analyze their long time behavior.

In other words: if we could use real numbers with an infinite number of significance figures chaotic systems are strictly deterministic, but real numbers are a mental idealization created by human minds, as the rigid body, the Galileo's inertia law, and many other very fruitful ideas. When evaluations are made, numbers must be represented by a discrete number of significance figures and the sensitivity to initial conditions or the "butterfly effect" Lorenz described in his famous paper [9] produces unpredictable behavior. The sensitivity to initial conditions and the broadband power spectrum of chaotic time series make them good candidates to generate PRNG's with a behavior very similar to physical noise sources. In this review we will focus on discrete chaotic dynamical systems (or maps).

(4) How can we measure the quality of a PRNG?

The first requirement pointed in *(2)*, an uniform histogram, is easy to be tested and the Shannon Entropy is the obvious quantifier. The second requirement, refers to statistical independence. This is hard to theoretically substantiate. Numerically, it is impossible. Instead, several basic properties are tested: long cycle length, uniformity of the power spectrum, mixing constant, *etc.* Also the speed of generation, reproducibility and portability

are essential requirements. In engineering applications, for example, it is usual to replace this second requirement for a flat power spectrum. The reason is, a uniform spectrum means that no periodicities exist and the next value may be any of the allowed ones. For discrete sources the power spectrum is replaced by the square of the absolute value of the fast Fourier transform (FFT) of the time series.

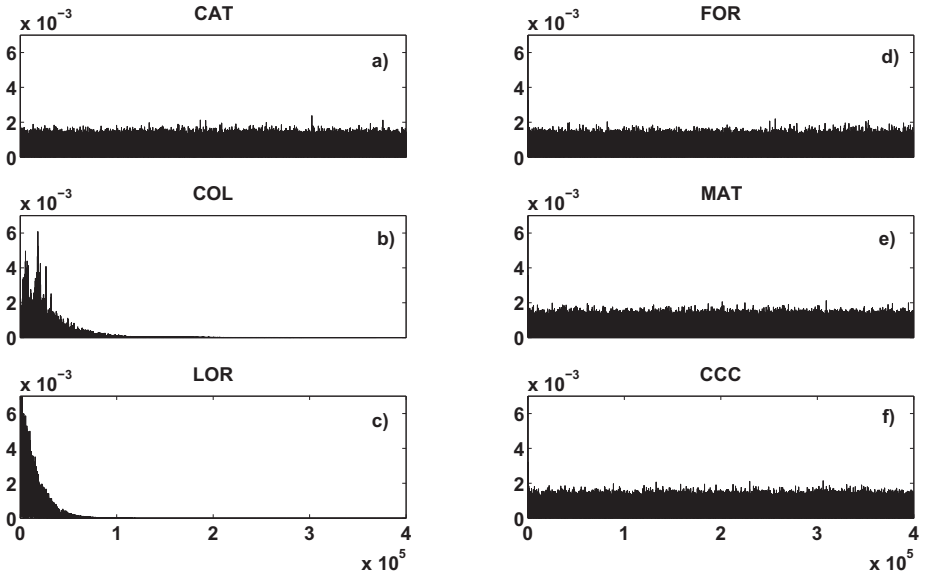

Figure 2: Power spectrum for for three non-chaotic (a) to (c) and three chaotic (d) to (f) PRNG's.

Some general purpose test suites, readily available to researchers in academia and industry who wish to analyze their newly developed PRNG are: tests by George Marsaglia [8], *Crypt-XS* by Helen Gustafson of the Queensland University of Technology [10], the National Institute of Standards and Technology Statistical Test Suite [11] and Rukhin suite of statistical procedures [12]. Additional requirements may be imposed in a specific application.

Of course a statistical test can never prove that a sequence generated by a PRNG is random (*because it is not !*), but it helps to detect certain kinds of weaknesses a generator may have. Furthermore none of these tests can prove that a given generator is reliable in all applications. Vattulainen *et al.* [13], for example, proposed three additional physical tests to detect deficiencies of several PRNG's used in Monte Carlo simulations.

In previous works we studied the ability of different quantifiers to characterize the uni-

formity of the histogram and the statistical independence in the case of chaotic PRNG's [14]. There we studied quantifiers based on Information Theory, recurrence plots [15] and intrinsic computation [16]. In this chapter we will focus on quantifiers based on Information Theory only.

In [14] we concluded that a very good classification scheme of PRNG's may be done using Information Theory quantifiers but it is mandatory to use two different PDF's: one to measure the uniformity of the outputs and the other one to measure correlation between consecutive samples. In this chapter we will focus on this approach where the Shannon-Jensen Statistical Complexity plays a fundamental role.

(5) What is the strategy to improve the quality of a chaotic PRNG?

As we pointed out several times any PRNG must have a uniform histogram and statistical independence of consecutive values. For a large family of chaotic maps, the histogram is characteristic of the map and it is more precisely defined as the *invariant measure*. The invariant measure is one of the eigenfunctions of the Perron-Frobenius operator of the map. The statistical independence of consecutive values of the time series may be estimated by means of the study of the mixing properties of the map. A mixing constant r_{mix} given by one of the eigenvalues of the Perron-Frobenius operator of the map [17, 18, 19], may be used to compare mixing maps.

When a chaotic time series has not a constant invariant measure, or a null mixing constant, or both, randomizing techniques may be designed to improve it. In this chapter we only consider one of this techniques called skipping, that improves the mixing but does not change the histogram. It works fine when the chaotic time series has a uniform histogram as is the case in piecewise linear maps. More details about more general randomization processes can be seen in [20].

This chapter is organized as follows: in Section 2 we briefly describe fourteen algorithmic PRNG. Half of them are non chaotic and half of them are chaotic. Results for these 14 PRNG's are used to exemplify the analysis tools. In Section 3 we briefly describe how the statistical description of a chaotic map may be done by studying the Perron-Frobenius operator associated with the map. Section 4 deals with classical tests to characterize chaotic PRNG's and the quantifiers based on Information Theory are introduced. In section 5 we describe one randomization technique called skipping, used to improve the mixing and consequently the statistical independence between consecutive values, for chaotic maps.

The characterization of the PRNG considered in the entropy-complexity plane is given in Section 6. Finally conclusions are presented in Section 7.

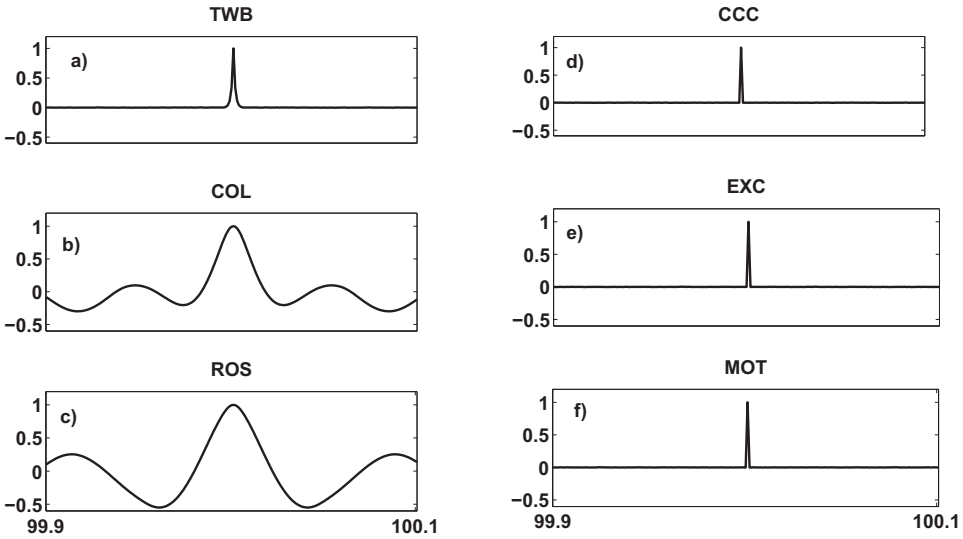

Figure 3: Linear autocorrelation for three chaotic (see (a) to (c)) and three non-chaotic ((d) to (f)) PRNG's.

2 Algorithmic Pseudo Random Number Generators

Most non chaotic algorithmic PRNGs are based on three methods [8]:

- *Congruential:* These generators use a linear transformation on the ring of reduced residues of some modulus m, to produce a sequence of integers, $x_1,\ x_2,\ x_3,\ \cdots$, with $x_n \equiv a\ x_{n-1} + b \mod m$.

- *Shift-register:* These generators iterates a binary $1 \times n$ vector β by means of a linear transformation with a matrix T. β is the seed and T must be a non singular $n \times n$ matrix having order $2^n - 1$ to produce a sequence $\beta,\ \beta T,\ \beta T^2,\ \cdots$, with period $2^n - 1$.

- *Lagged-Fibonacci:* These generators use an initial set of elements $x_1, x_2, \cdots, x_r$, and two lags r and s with $r > s$. Successive elements are generated by the recursion

$x_n = x_{n-r} \bullet x_{n-s}$, where $\bullet$ is some binary operation, $n > r$, the initial seed elements are computer words and the binary operation might be $+$, $-$, $*$ or $\oplus$ (exclusive or). For operations $+$ or $-$ the x's might be integers mod 2^k or single, or double precision reals mod 1. For the operations $*$ the x's might be odd integers mod 2^k. Each lagged-Fibonacci generator depends on details of the particular binary operation and the set of elements it operates on.

Best results are obtained by combining two or more simple generators through a convenient computer operation such as $+$, $-$, $*$ or $\oplus$.

In this review we will consider seven non chaotic algorithmic generators widely used:

- Excel© RNG (EXC).

- RNG available in Intel© fortran compiler (FOR).

- RNG available in Borland © C++ compiler (CCC).

- Matlab© RAND function (MAT).

- *Mother* RNG, available in Marsaglia website [8] (MOT), given by

$$\begin{cases} u_{n+1} = 2111111111\ x_{n-3} + 1492\ x_{n-2} + 1776\ x_{n-1} + 5115\ x_n + c_n\ , \\ x_{n+1} = u_{n+1} \bmod 2^{32}\ , \\ c_{n+1} = floor(u_{n+1}/2^{32})\ . \end{cases} \tag{1}$$

- *Multiple with carry* RNG (MWC) [8], given by

$$\begin{cases} u_{n+1} = 2131995753\ x_n + c_n\ , \\ x_{n+1} = u_{n+1} \mod 2^{32}\ , \\ c_{n+1} = floor(u_{n+1}/2^{32})\ . \end{cases} \tag{2}$$

- *Lehmer* RNG (LEH), given by

$$\begin{cases} u_{n+1} = 16807\ x_n\ , \\ x_{n+1} = u_{n+1}\ \mod 2^{31} - 1\ . \end{cases} \tag{3}$$

All these algorithmic non-chaotic PRNGs are widely used and they pass most tests by Marsaglia [8].

We will also consider seven chaotic generators; the first three are continuous dynamical systems discretized by the Euler procedure and the last four are chaotic maps. They are:

- *Lorenz discrete dynamical system* (LOR):

$$\begin{cases} x_{n+1} = x_n + k\ [-\delta\ (x_n - y_n)]\ , \\ y_{n+1} = y_n + k\ [-x_n\ z_n + r\ x_n - y_n]\ , \\ z_{n+1} = z_n + k\ [x_n\ y_n - b\ z_n]\ . \end{cases} \tag{4}$$

  with $k = 1/64$, $\delta = 8$, $r = 24$, $b = 2$.

- *Collpits discrete dynamical system* (COL):

$$\begin{cases} x_{n+1} = x_n + k\ [a\ z_n - b\ u_n]\ , \\ y_{n+1} = y_n + k\ [a\ z_n + c\ u_n - y_n - d]\ , \\ z_{n+1} = z_n + k\ [e\ d - z_n - x_n - y_n]\ , \\ u_{n+1} = \begin{cases} 0 \quad \text{for } y_n\ \geq\ -1\ , \\ -y_n - 1 \text{ for any other case.} \end{cases} \end{cases} \tag{5}$$

  with $k = 1/128$, $a = 12$, $b = 800$, $c = 4$, $d = 8$, $e = 8$.

- *Rössler discrete dynamical system* (ROS):

$$\begin{cases} x_{n+1} = x_n + k\ [-y_n - z_n]\ , \\ y_{n+1} = y_n + k\ [x_n + a\ y_n]\ , \\ z_{n+1} = z_n + k\ [b + x_n\ z_n - \mu\ z_n]\ . \end{cases} \tag{6}$$

  with $k = 1/16$, $a = 1/4$, $b = 1/2$, $\mu = 8$.

- *Cat map* (CAT):

$$\begin{cases} x_{n+1} = x_n + y_n - floor(x_n + y_n) \ , \\ y_{n+1} = x_n + 2\ y_n - floor(x_n + 2\ y_n) \ . \end{cases} \tag{7}$$

- *Tent map* (TEN):

$$x_{n+1} = \begin{cases} (2 * x_n + 1 - a)/(1 + a) & \text{if } -1 \leq x_n \leq a \ , \\ (-2 * x_n + 1 + a)/(1 - a) & \text{if } a < x_n \leq 1 \ . \end{cases} \tag{8}$$

  with $a = 0.26$.

- *Logistic map* (LOG):

$$x_{n+1} = r\ x_n\ (1 - x_n) \ . \tag{9}$$

  with $r = 4$.

- *Three way Bernoulli map* (TWB):

$$x_{n+1} \ = \ \begin{cases} 3x_n & \text{if } 0 \leq x_n \leq 1/3 \ , \\ 3x_n - 1 & \text{if } 1/3 < x_n \leq 2/3 \ , \\ 3x_n - 2 & \text{if } 2/3 < x_n \leq 1 \ . \end{cases} \tag{10}$$

Some of these chaotic PRNG's do not pass the tests by Marsaglia [8] but they may be randomized as will be explained in Section 5.

3 Perron-Frobenius approach for chaotic maps

The statistical properties of chaotic maps have been characterized by means of the Perron-Frobenius operator associated with the map [21, 22, 23]. This issue will not be detailed in this chapter because there exists an extensive bibliography about it (see for example [17, 18, 19]).

We only give here some general aspects. Let f be a chaotic map on the interval $[0, 1]$. Suppose the map has an invariant measure $\mu(x)$. Then the map is *ergodic* if for any

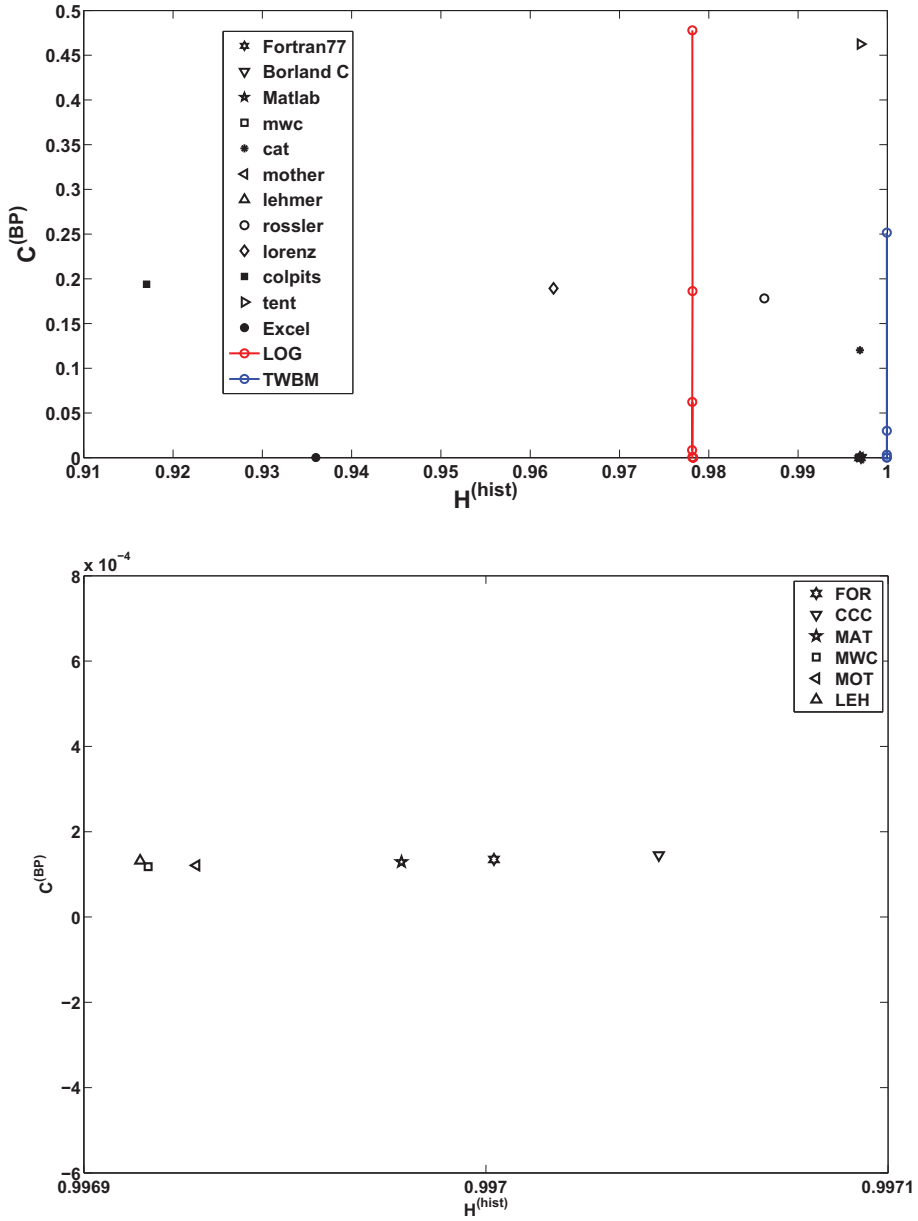

Figure 4: (a) Entropy complexity plane $H^{(hist)} \times C^{(BP)}$ for all the PRNGs of Section 2. (b) Zoom of Fig. (a) around the point (0.997,0), where several PRNGs are very close to each other.

integrable test function $Q(x)$, and for an arbitrary initial condition x_0 (up to a set of zero

μ-measure), the time average is equal to the ensemble average:

$$\overline{Q} = \langle Q \rangle \ . \tag{11}$$

Equation (11) is a consequence of the famous Birkhoff ergodic theorem [24]. *Mixing* is an even stronger requirement than ergodicity. A map is called "*mixing*" if any smooth initial probability density $\rho(x)$ converges to the invariant measure $\mu(x)$ after enough successive iterations. Mixing implies ergodicity. The reverse, however, is not true [25].

There exists an equivalent definition of mixing *via correlation functions*. Let $\phi_1(x)$ and $\phi_2(x)$ be two integrable test-functions and define the generalized correlation function of the map f by

$$C(\phi_1, \phi_2, n) = \lim_{J\to\infty} \frac{1}{J} \sum_{j=0}^{J-1} \phi_1(x_{j+n})\phi_2(x_j) - \langle \phi_1 \rangle \langle \phi_2 \rangle \ , \tag{12}$$

where

$$\langle \phi_i \rangle = \lim_{J\to\infty} \frac{1}{J} \sum_{j=0}^{J-1} \phi_i(x_j) \ . \tag{13}$$

The map is mixing if, for arbitrary ϕ_1 and ϕ_2,

$$\lim_{n\to\infty} C(\phi_1, \phi_2, n) = 0 \ . \tag{14}$$

Let us stress that it is not easy to prove that f is a *mixing map* because the mixing condition given in equation (14) must be fulfilled for *arbitrary* test functions. Formally, every mixing map f has an associated determines the time evolution of any initial density $\rho_0(x)$ towards the invariant measure $\mu(x)$:

$$\rho_{n+1} = \mathcal{L}[\rho_n] \ . \tag{15}$$

The explicit formal expression for the Perron-Frobenius operator for a one-dimensional map f is given in [25].

$$\mathcal{L}[\rho_y] = \sum_{x \epsilon f^{-1}(y)} \frac{[\rho_0(x)]}{|f'(x)|} \ . \tag{16}$$

This operator $\mathcal{L}$ has a set of eigenfunctions $\psi_\alpha(x)$ and eigenvalues η_α. The invariant measure $\mu(x)$ is the eigenfunction corresponding to the largest eigenvalue $\eta_0 = 1$. The full set of eigenfunctions and eigenvalues may be used as a basis to express any density:

$$\begin{aligned} \rho_0(x) &= \sum_\alpha c_\alpha \psi_\alpha(x) \ , \\ \rho_n(x) &= \mathcal{L}^n \rho_0(x) = \sum_\alpha \eta_\alpha^n c_\alpha = c_0\psi_0(x) + R_n \ . \end{aligned} \tag{17}$$

The eigenvalue with the second largest absolute value, η_1, has a "distinguished" physical meaning: it is related with the *mixing constant* r_{mix} that governs the relaxation of *exponentially mixing* maps:

$$|R_n| \sim |\eta_1|^n \sim \exp(-n/r_{mix}) \; . \tag{18}$$

In the case of mixing maps the natural invariant measure $\mu(x)$ is identical to the normalized histogram of the time series. Consequently the first main requirement of a RPNG is $\mu(x) = const$. The mixing constant r_{mix} gives the characteristic transient time [18, 26, 27, 28, 29, 30, 31]. A mixing constant r_{mix} pretends to be an indirect measure of statistical independence: if $r_{mix} = 0$ the maps mixes the consecutive values making them *look as* statistically independent random variables. For a PRNG the ideal value is $r_{mix} = 0$.

Usually, the analytical expression of the invariant measure $\mu(x)$ of a given map f is not known. Exceptions are the logistic map in full chaos, and piecewise-linear maps. The mixing constant r_{mix} has been analytically obtained only for piecewise-linear maps. For other maps it must be numerically obtained by means of a piecewise linear approximation of the map.

4 Classical Tests and quantifiers based on Information Theory

4.1 Classical Tests

Let us start describing four classical tests usually applied to PRNGs, that are also useful for the study of chaotic PRNGs:

- $2D$ *and* $3D$ *distribution:*

  These are excellent graphic or visual tools to discover hidden structures produced by nonlinear correlations in the time series. They are obtained by plotting vectors $\overrightarrow{x}_i$ with components equal to successive values of the time series, *i.e.*, $\overrightarrow{x}_i = (x_i,\, x_{i+1})$ in the case of $2D$ and $\overrightarrow{x}_i = (x_i,\, x_{i+1},\, x_{i+2})$ in the case of $3D$. The sequence $\{\overrightarrow{x}_i\}$ must fill the complete $2D$ or $3D$ space in an uniform way. It only means that any combination of 2 or 3 successive values appear the same number of times. Note that to assure randomness based on this test it is necessary to extend the procedure to infinite dimensions.

In Figs. 1 it is shown the $3D$ distribution for some of the non-chaotic and chaotic systems reported in this chapter. Several chaotic systems do not fill the complete $2D$ or $3D$ space in an uniform way. This is an evidence that the system does not constitute a good PRNG and must be randomized.

- *Power Spectrum:*

  The Fast Fourier Transform (FFT) $\{X_k\}$ of the time series $\{x_n\}$ is performed and the corresponding discrete power spectrum $P_s = X_k \; X_k^*$ is computed. A flat power spectrum, with equal frequency contribution for all frequencies is indicative of statistical independence between successive values. Consequently a good PRNG must have a uniform power spectrum.

  In Figs. 2 it is shown the power spectrum for several non-chaotic and chaotic systems reported in this chapter.

- *Autocorrelation Function:*

  The discrete linear autocorrelation function given by

$$Corr(x,x)_j = \sum_k \; x_{nk} \; x_{nk+j} \; , \tag{19}$$

  where j is the lag. An almost constant value of $Corr(x,x)_j$, independent of the lag j, will be indicative of an uncorrelated random series. But this test only considers linear correlations and it is not a good test for chaotic PRNGs.

  In Fig. 3 we show the autocorrelation function for several PRNG's studied in this chapter. Note that for example LOG has a delta like autocorrelation in spite consecutive values are not statistically independent and are strongly correlated by the quadratic map.

- *Marsaglias's tests:*

  Marsaglia proposed a set of tests he called *stringent tests*. They are based on the fact that a random number generator is supposed to produce a sequence of statistical independent and identically distributed random variables x_1, x_2, x_3, $\cdots$. Any function

of elements of that sequence may serve as a test. Marsaglia's tests provide a matrix of 127 values to characterize the PRNG. The idea is that a good PRNG must pass all the tests but it is not easy to compare widely different PRNG's by the values obtained. Quoting Marsaglia: "If the RNG is to be used for a particular problem one should try to create a test based on a similar problem for which the underlying distributions are known or lacking that at least compared with results produced by widely different RNG's".

For those interested in test suites we refer to [8, 10, 11, 32].

.

4.2 Information Theory quantifiers for chaotic maps

As we need to characterize respectively the uniformity of the histogram and the statistical independence, our approach is to use two quantifiers and represent the PRNG's on a plane with one axis for each quantifier. The goal of our approach is to get an easy way to compare PRNG's of different origin. The uniformity of the histogram is obviously measured by the Shannon Entropy. To quantify the statistical independence between consecutive values we studied several quantifiers for chaotic systems in previous works the FFT, the mixing constant, recurrence plots quantifiers, intrinsic computation quantifiers, *etc.* We will focus in this chapter to the use of quantifiers based on Information Theory [14, 20, 33, 34, 35] (see also Chapter 2 and 8).

Information Theory quantifiers are appropriate functionals of the probability distribution function (PDF). Let $\{x_i\}$ be the time series under analysis, with length M. There are infinite possibilities to assign a PDF to a given time series, a subject that will be given due consideration below. In the meantime, suppose that the PDF is discrete and is given by $P = \{p_i; i = \cdots, N\}$. One defines then various quantities, namely,

- *Normalized Shannon Entropy $H[P]$:*

  Let $S[P]$ be the Shannon Entropy

$$S[P] = -\sum_{i=1}^{N} p_i \ \ln(p_i) \ . \tag{20}$$

Is is well known that the maximum $S_{max} = \ln(N)$ is obtained for $P_e = \{1/N, \cdots, 1/N\}$, that is, the uniform PDF. A "normalized" entropy $H[P]$ can also be defined in the fashion

$$H[P] = S[P]/S_{max} \ . \tag{21}$$

- *Shannon-Jensen Statistical Complexity Measure $C[P]$:*
  For the statistical complexity measure, in this chapter we adopt the functional form definition given by López Ruiz-Mancini-Calbet [36] (see Chapter 7) with the modifications advanced by Lamberti and co-workers in [2] (see Chapter 8), so as to ensure that the concomitant SCM-version becomes *(i)* able to grasp essential details of the dynamics, *(ii)* an intensive quantity (in the thermodynamical sense and, *(iii)* capable of discerning both among different degrees of periodicity and chaos [3]. The ensuing measure, to be referred to as the Shannon-Jensen statistical complexity (see Chapter 8), is a functional $C[P]$ that reads

$$C[P] = Q_J[P, P_e] \cdot H[P] \ , \tag{22}$$

where Q_J is the "disequilibrium", defined in terms of the so-called extensive Jensen-Shannon divergence (which induces a squared metric) [2]. One has

$$Q_J[P, P_e] = Q_0 \cdot \{S[(P + P_e)/2] - S[P]/2 - S[P_e]/2\} \ , \tag{23}$$

with Q_0 a normalization constant ($0 \leq Q_J \leq 1$) that reads

$$Q_0 = -2 \left\{ \left(\frac{N+1}{N} \right) \ln(N+1) - 2\ln(2N) + \ln N \right\}^{-1} \ . \tag{24}$$

We see that the disequilibrium Q_J is an intensive quantity that reflects on the systems's "architecture", being different from zero only if there exist "privileged", or "more likely" states among the accessible ones. $C[P]$ quantifies the presence of correlational structures as well [2, 37]

The opposite extremes of perfect order and maximal randomness possess no structure to speak of and, as a consequence, $C[P] = 0$. In between these two special instances a wide range of possible degrees of physical structure exist, degrees that should be reflected in the features of the underlying probability distribution. In the case of a PRNG the "ideal" values are $H[P] = 1$ and $C[P] = 0$.

As pointed out above, P itself is not a uniquely defined object and several approaches have been employed in the literature so as to "extract" P from the given time series. Just to mention some frequently used extraction procedures: *a)* frequency count [38] *b)* procedures based on amplitude statistics (histograms) [20], *c)* binary symbolic-dynamics [39], *d)* Fourier analysis [40], *e)* wavelet transform [41, 42], *f)* partition entropies [43], *g)* permutation entropy [4, 44], *h)* discrete entropies [45], *etc.* There is ample liberty to choose among them. De Micco *et al.* [20] proposed two probability distributions as relevant for testing respectively the uniformity of $\mu(x)$ and the mixing constant: *(a)* a P based on time series' histograms and *(b)* a P based on ordinal patterns permutation ordering that derives from using the Bandt-Pompe method [4].

For extracting P *via* the histogram divide the interval $[0,1]$ into a finite number N_{bin} of non overlapping subintervals A_i: $[0,1] = \bigcup_{i=1}^{N_{bin}} A_i$ and $A_i \bigcap A_j = \emptyset \ \forall i \neq j$. Note that N in equation (20) is equal to N_{bin}. Of course, in this approach the temporal order of the time-series plays no role at all. The quantifiers obtained *via* the ensuing PDF are called in this paper $H^{(hist)}$ and $C^{(hist)}$. Let us stress that for time series within a finite alphabet it is relevant to consider an optimal value of N_{bin} (see *i.e.*, [20]).

In extracting P by recourse to the Bandt-Pompe method the resulting probability distribution P is based on the details of the attractor-reconstruction procedure. *Causal information* is, consequently, duly incorporated into the construction-process that yields P. The quantifiers obtained *via* the ensuing PDF are called in this paper $H^{(BP)}$ and $C^{(BP)}$. A notable Bandt-Pompe result consists in getting a clear improvement in the quality of Information Theory-based quantifiers [3, 34, 35, 46, 47, 48, 49, 50].

The extracting procedure is as follows. For the time-series $\{x_t : t = 1, \cdots, M\}$ and an embedding dimension $D > 1$, one looks for "ordinal patterns" of order D [4, 44, 51] generated by

$$(s) \mapsto \left(x_{s-(D-1)}, x_{s-(D-2)}, \cdots, x_{s-1}, x_s\right) \ , \tag{25}$$

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

$$x_{s-r_{D-1}} \leq x_{s-r_{D-2}} \leq \cdots \leq x_{s-r_1} \leq x_{s-r_0}. \tag{26}$$

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

$$p(\pi) = \frac{\sharp\{s|s \leq M - D + 1; (s) \quad \texttt{has type } \pi\}}{M - D + 1}. \tag{27}$$

In the last expression the symbol $\sharp$ stands for "number".

The advantages of the Bandt-Pompe method reside in *a)* its simplicity, *b)* the associated extremely fast calculation-process, *c)* its robustness in presence of observational and dynamical noise, and *d)* its invariance with respect to nonlinear monotonous transformations. The Bandt-Pompes methodology is not restricted to time series representative of low dimensional dynamical systems but can be applied to any type of time series (regular, chaotic, noisy, or reality based), with a very weak stationary assumption (for $k = D$, the probability for $x_t < x_{t+k}$ should not depend on t [4]. One also assumes that enough data are available for a correct phase space reconstruction. 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 $D!$. Also, it conditions the minimum acceptable length $M \gg D!$ of the time series that one needs in order to work with a reliable statistics. In relation to this last point Bandt and Pompe suggest, for practical purposes, to work with $3 \leq D \leq 7$ with a time lag $\tau = 1$. This is what we do here (in the present work $D = 6$ is used).

d	LOG	TWBM
1	0.56789	0.333333333
2	0.31848	0.111111111
3	0.13290	0.037037037
4	0.05788	0.012345679
5	0.03646	0.004115226
6	0.01791	0.001371742
7	0.01152	0.000457247
8	0.00515	0.000152416

Table 1: r_{mix} as a function of the iteration-order d for the TWBM and LOG chaotic maps, respectively.

$\mathcal{S}_{IN}$	$\mathcal{S}_{OUT} \equiv f^d$				
$(\mathbb{R})$	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$
0.010559404					
0.041791613	0.041791613				
0.160180296		0.160180296			
0.538090276	0.538090276		0.538090276		
0.994196523				0.994196523	
0.023079185	0.023079185	0.023079185			0.023079185
0.090186145					
0.328210418	0.328210418		0.328210418		
0.881953358		0.881953358			
0.416446528	0.416446528			0.416446528	
0.972075269					
0.10857976	0.10857976	0.10857976	0.10857976		0.10857976
0.387160782					
0.949069244	0.949069244				
0.193347257		0.193347257		0.193347257	
0.62385638	0.62385638		0.62385638		

Table 2: Illustrating the **Skipping** procedure

5 Skipping randomization procedure

It is possible to show that the invariant measure of the iterated map f^d is identical to the invariant measure of the original map f. Also, the mixing constant r_{mix} for f^d is lower that the mixing constant for f. As an example, in Table 1 the value or r_{mix} of f and f^d is shown for two of the chaotic maps studied in this chapter.

The iteration of a map is then the simplest randomization procedure proposed in the literature, used to diminish r_{mix} or equivalently to increase the statistical independence of consecutive values. This procedure is also known as **Skipping** because iterating a map is tantamount to skipping values in the original time series.

The procedure is as follows: let f be a chaotic map. Starting from a randomly chosen initial condition, the map is iterated generating the chaotic time-series (CHTS) $\mathcal{S}_{IN} =$

$\{x_0, x_1, \cdots\}$. This CHTS is to be regarded as the input for the randomizing process (see Table 2). Let us assume that x_i is a floating-point number (in the IEEE normalized representation). Without loss of generality, we consider values restricted to the interval $[0, 1]$. As previously stated, the central idea here is that of employing the well-known symbolic dynamics-randomization process called Skipping to obtain the new time-series (STS) $\mathcal{S}_{OUT}$, obtained by discarding $d-1$ values of $\mathcal{S}_{IN}$, are "skipped" to get the STS $\mathcal{S}_{OUT}$ which originates the name **Skipping** for this technique. In other words, one employs, instead of the original map f, its d-times iterated one f^d. This randomization technique is routinely (and successfully) used with piecewise linear maps in many applications [18]. In Table 2, an example with different values of d is displayed.

6 Characterization of PRNG's by means of the Entropy-Complexity plane

In this section we present the main result reported in this chapter. That is a representation of all the PRNG's on a plane with one axis used to characterize the uniformity of the histogram and the other axis used to characterize the statistical independence.

The uniformity of the histogram is measured by $H^{(hist)}$, the normalized Shannon entropy of the histogram. The statistical independence is measured by $C^{(BP)}$, the Shannon-Jensen statistical complexity measure using the PDF obtained by the Bandt-Pompe procedure. Figs. 4 shows our results. We represent on the plane $H^{(hist)} \times C^{(BP)}$ all the PRNG's presented in Section 2. The ideal PRNG is one with $H^{(hist)} = 1$ and $C^{(BP)} = 0$ and Fig. 4 shows that the best PRNG is TWB, after the skipping procedure is applied. Near the ideal there are several PRNGs: MOT, FOR, CCC and MAT. LOG does not produce a good result even after the skipping procedure. The reason is the skipping procedure does not change the histogram of the map and consequently it is impossible to improve the value of $H^{(hist)}$ by this randomization procedure. For other strategies that can be used for maps with nonuniform histogram like LOG see [20].

It is interesting to see that EXC and also LOR, ROS and COL are not good choices. All these generators have low entropy and high complexity. For LOR, ROS and COL it is possible to use a bigger step in the Euler procedure to get better results [52].

7 Conclusions

In summary two main characteristics define the quality of a PRNG: equiprobability of all the values and statistical independence between consecutive outputs. Both characteristics may be quantified by means of an Information Theory approach but two different PDF's are required. Other quantifiers have been also studied in previous works: $3D$ distributions, power spectrum uniformity, r_{mix} constant, recurrence plots, intrinsic computation quantifiers [14]. The advantage of a procedure based only on Information Theory quantifiers is to use very well understood concepts as Entropy and Statistical Complexity. In fact the first above mentioned requirement - the equiprobability of all the values - is equivalent to a uniform histogram and implies a normalized Shannon entropy $H^{(hist)} = 1$. Consequently one PRNG will be better than other if it has $H^{(hist)}$ closer to the ideal value 1.

To quantify the statistical independence it is required to use a causal PDF, it means a PDF for trajectories and not for individual values, in order to capture correlations between consecutive outputs. Our approach is to define a probability distribution function by means of the Bandt-Pompe procedure and evaluate the Shannon-Jensen statistical complexity $C^{(BP)}$ of that distribution. The ideal PRNG will have $C^{(BP)} = 0$. The representation of different PRNG's on a plane with axis $H^{(hist)}$ and $C^{(BP)}$ allows one to easily compare them with each other. The ideal PRNG must be represented by the point $(1, 0)$ in this plane.

In the present work we used the skipping randomization procedure in order to improve the PRNG performance. Its viability was checked using the entropy-complexity plane $H^{(hist)} \times C^{(BP)}$. We have shown that the TWB, MOT, FOR, CCC and MAT are very near the ideal point $(1, 0)$ after skipping procedure. However, the proposed procedure for improvement the PRNG performance do not work for the case of LOG due to the characteristics that its PDF-histogram present. Something similar happen with EXC, LOR, ROS and COL due to they present low entropy and high complexity values.

It is interesting to note that the representation points of piecewise chaotic maps can almost reach this ideal point by means of the skipping procedure. The reason is skipping diminishes correlation between consecutive values. For other chaotic maps more involved randomization techniques are required [20] Let us finally stress that statistical quality is the main requirement for a PRNG used in many applications such as Monte Carlo simulations, electromagnetic compatibility improvement [53], filtering by random sampling [29], *etc.* But it is not enough to guarantee good cryptographic properties [54]. Cryptographic properties

are not the subject of this chapter.

Conflict of interest: None.

Acknowledgments

This work was partially supported by Universidad Nacional de Mar del Plata, ANPCyT (PICT 2010-2335) and CONICET (PIP 112-200801-1420 CCTMDP). OAR acknowledges partial support from CONICET, Argentina, and CAPES, PVE fellowship, Brazil.

References

[1] Wold H. *A Study in the Analysis of Stationary Time Series.* Almqvist and Wiksell Book Co., Upsala, Sweden, 1938.

[2] Lamberti, P.W.; Martín, M.T.; Plastino, A.; Rosso, O.A. Intensive entropic non-triviality measure. *Physica A* **334** 119–131 (2004).

[3] Rosso, O.A.; Larrondo, H.A.; Martín, M.T.; Plastino, A.; Fuentes, M.A. Distinguishing noise from chaos. *Phys. Rev. Lett.* **99** 154102 (2007).

[4] Bandt, C.; Pompe, B. Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* **88** 174102 (2002).

[5] http://www.random.org

[6] http://www.fourmilab.ch/hotbits/

[7] http://www.lavarnd.org

[8] Marsaglia, G. *The marsaglia random number cdrom including the diehard battery of tests of randomness.* http://www.stat.fsu.edu/pub/diehard/, 1995.

[9] Lorenz, E.N. Deterministic non periodic flow. *Journal of the Atmospheric Sciences* **20** 130–141 (1963).

[10] Gustafson, H.M.; Dawson, E.P.; Nielsen, L.; Caelli, W.J. A computer package for measuring the strength of encryption algorithms. *J. Comput. Security* **13** 687–697 (1994).

[11] Soto, J. *Statistical testing of random number generators, available online.* http://www.itl.nist.gov/div893/staff/soto/jshome.html.

[12] Rukhin, A.L. Testing randomness: a suite of statistical procedures. *Theory Probab. Appl.* **45** 111–132 (2000).

[13] Vattulainen, I.; Ala-Nissila, T.; Kankaala, K. Physical tests for random numbers in simulations. *Phys. Rev. Lett.* **73** 2513–2516 (1994).

[14] De Micco, L.; Larrondo, H.A.; Plastino, A.; Rosso, O.A. Quantifiers for randomness of chaotic pseudo random number generators. *Philosophical Transactions of the Royal Society A* **367** 3281–3296 (2009).

[15] Marwan, N.; Romano, M.C.; Thiel; M., Kurths, J. Recurrence plots for the analysis of complex systems. *Physics Reports* **438** 237–329 (2007).

[16] Feldman, D.P.; McTague, C.S.; Crutchfield, P. The organization of intrinsic omputation: complexity-entropy diagrams and the diversity of natural information processing. *Chaos* **18** 043106 (2008).

[17] Beck, C.; Schlögl, F. *Thermodynamics of chaotic systems: an introduction.* Cambridge University Press, 1997.

[18] Setti, G.; Mazzini, G.; Rovatti, R.; Callegari, S. Statistical modeling of discrete-time chaotic processes: Basic finite-dimensional tools and applications. *Proceedings of the IEEE* **90** 662–690 (2002).

[19] Lasota, A.; Mackey, M.C. Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics. *Applied Mathematical Sciences* **97**. Springer Verlag, 2nd. edition, 1994.

[20] De Micco, L.; González, C.M.; Larrondo, H.A.; Martín, M.T.; Plastino, A.; Rosso, O.A. Randomizing nonlinear maps via symbolic dynamics. *Physica A* **387** 3373–3383 (2008).

[21] Dellnitz, M.; Froyland, G.; Sertl, S. On the isolated spectrum of the Perron-Frobenius operator. *Nonlinearity* **13** 1171–1188 (2000) .

[22] Ding, J.; Zhou, A. Finite approximations of Frobenius-Perron operators. A solution to Ulam's conjecture to multi-dimensional transformations. *Physica D* **92** 61–68 (1996).

[23] Pingel, D.; Schmelcher, P. Theory and examples of the inverse Frobenius-Perron problem for complete chaotic maps. *Chaos* **9** 357–66 (1999).

[24] Cornfeld, P.; Fomin, S.V.; Sinai, Ya. G. *Ergodic Theory.* Springer, New York, 1982.

[25] Beck, C. Ergodic properties of a kicked damped particle. *Commun. Math. Phys.* **130** 51–60 (1990).

[26] De Micco, L.; Petrocelli, R.A.; Larrondo, H.A. Constant envelope wideband signals using arbitrary chaotic maps. *Proceedings of XII RPIC*, 2007.

[27] De Micco, L.; Arizmendi, C.M.; Larrondo, H.A. Zipping characterization of chaotic sequences used in spread spectrum communication systems. *Institute of Physics Conference Proceedings* **913** 139–144 (2007).

[28] De Micco, L.; Petrocelli, R.A.; Carrica, D.O.; Larrondo, H.A. Muestreo caótico para la adquisición de señales de baja frecuencia con ruido de alta frecuencia. *Proceedings of XII RPIC*, 2007.

[29] Petrocelli, R.A.; De Micco, L.; Carrica, D.O.; Larrondo, H.A. Acquisition of low frequency signals immersed in noise by chaotic sampling and fir filters. *IEEE International Symposium on Intelligent Signal Processing 2007. WISP2007* 1–6 (2007).

[30] Rovatti, R.; Mazzini, G.; Setti, G. On the ultimate limits of chaos-based asynchronous DS-CDMA-I: Basic definitions and results. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **51** 1336–1347 (2004).

[31] Rovatti, R.; Mazzini, G.; Setti, G. On the ultimate limits of chao-sbased asynchronous DS-CDMA-II: Analytical results and asymptotics. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **51** 1348–1364 (2004).

[32] Rukhin, A.L. *Random number generation.* Available from http://www.iro.umontreal.ca/lecuyer/myftp/papers/handstat.pdf, 2004.

[33] González, C.M.; Larrondo, H.A.; Rosso, O.A. Statistical complexity measure of pseudorandom bit generators. *Physica A* **354** 281–300 (2005).

[34] Larrondo, H.A.; González, C.M.; Martín, M.T.; Plastino, A.; Rosso, O.A. Intensive statistical complexity measure of pseudorandom number generators. *Physica A* **356** 133–138 (2005).

[35] Larrondo, H.A.; Martín, M.T.; González, C.M.; Plastino, A.; Rosso, O.A. Random number generators and causality. *Phys. Lett. A* **352** 421–425 (2006).

[36] López-Ruiz, R.; Mancini, H.L.; Calbet, X. A statistical measure of complexity. *Phys. Lett. A* **209** 321–326 (1995).

[37] Martín, M.T.; Plastino, A.; Rosso, O.A. Statistical complexity and disequilibrium. *Phys. Lett. A* **311** 126–132 (2003).

[38] Rosso, O.A.; Craig, H.; Moscato, P. Shakespeare and other english renaissance authors as characterized by Information Theory complexity quantifiers. *Physica A* **388** 916–926 (2009).

[39] Mischaikow, K.; Mrozek, M.; Reiss, J.; Szymczak, A. Construction of symbolic dynamics from experimental time series. *Phys. Rev. Lett.* **82** 1114–1147 (1999).

[40] Powell, G.E.; Percival, I.C. A spectral entropy method for distinguishing regular and irregular motion of hamiltonian systems. *J. Phys. A: Math. Gen.* **12** 2053–2071 (1979).

[41] Blanco, S.; Figliola, A.; Quian Quiroga, R.; Rosso, O.A.; Serrano, E. Time-frequency analysis of electroencephalogram series (III): Wavelet packets and Information Cost Function. *Phys. Rev. E* **57** 932–940 (1998).

[42] Rosso, O.A.; Blanco, S.; Jordanova, J.; Kolev, V.; Figliola, A.; Schürmann, M.; Başar, E. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *Journal of Neuroscience Methods* **105** 65–75 (2001).

[43] Ebeling, E.; Steuer, R. Partition-based entropies of deterministic and stochastic maps. *Stochastics and Dynamics* **1** 1–17 (2001).

[44] Keller, K.; Sinn, M. Ordinal analysis of time series. *Physica A* **356** 114–120 (2005).

[45] Amigó, J.M.; Kocarev, L.; Tomovski, I. Discrete entropy. *Physica D* **228** 77–85 (2007).

[46] Kowalski, A.M.; Martín, M.T.; Plastino, A.; Rosso, O.A. Bandt-Pompe approach to the classical-quantum transition. *Physica D* **233** 21–31 (2007).

[47] Rosso, O.A.; Zunino, L.; Pérez, D.G.; Figliola, A.; Larrondo, H.A.; Garavaglia, M.; Martín, M.T.; Plastino, A. Extracting features of gaussian selfsimilar stochastic processes via the Bandt and Pompe approach. *Phys. Rev. E* **76** 061114 (2007).

[48] Rosso, O.A.; Vicente, R.; Mirasso, C.R. Encryption test of pseudo-aleatory messages embedded on chaotic laser signals: an Information Theory approach. *Phys. Lett. A* **372** 1018–1023 (2008).

[49] Zunino, L.; Pérez, D.G.; Martín, M.T.; Plastino, A.; Garavaglia, M.; Rosso, O.A. Characterization of gaussian self-similar stochastic processes using wavelet-based informational tools. *Phys. Rev. E* **75** 021115 (2007).

[50] Zunino, L.; Pérez, D. G.; Martín, M.T.; Garavaglia, M.; Plastino, A.; Rosso, O.A. Permutation entropy of fractional Brownian motion and fractional Gaussian noise. *Physics Letters A* **372** 4768–4774 (2008).

[51] Keller, K.; Lauffer, H. Symbolic analysis of high-dimensional time series. *Int. J. Bifurcation and Chaos* **13** 2657–2668 (2003).

[52] De Micco, L.; Zabaleta, O.G.; González, C.M.; Arizmendi, C.M.; Larrondo, H.A. Estocasticidad de un atractor caótico determinista implementado en FPGA. *Proceedings IBERCHIP 2010.*

[53] De Micco, L.; Petrocelli, R.A.; Rosso, O.A.; Plastino, A.; Larrondo, H.A. Mixing chaotic maps and electromagnetic interference reduction. *Int. J. App. Math. Stat*, **26**, 106–120 (2012).

[54] Liberatori, M.C.; Castieira Moreira, J.; Petruzzi, D.M.; Honary, B. Trellis-hopping turbo coding. *IEE Proc.-Commun.* **153** 966–975 (2006).