A LETTERS JOURNAL EXPLORING
THE FRONTIERS OF PHYSICS

OFFPRINT

Combinatorial detection of determinism in noisy time series

J. M. Amigó, S. Zambrano and M. A. F. Sanjuán

EPL, **83** (2008) 60005

epl A LETTERS JOURNAL EXPLORING THE FRONTIERS OF PHYSICS

September 2008

EPL, **83** (2008) 60005
doi: 10.1209/0295-5075/83/60005

www.epljournal.org

Combinatorial detection of determinism in noisy time series

J. M. AMIGÓ[1(a)], S. ZAMBRANO[2] and M. A. F. SANJUÁN[2]

[1] *Centro de Investigación Operativa, Universidad Miguel Hernández - Avda. de la Universidad s/n., 03202 Elche, Spain, EU*
[2] *Departamento de Física, Universidad Rey Juan Carlos - Tulipán s/n, 28933 Madrid, Spain, EU*

received 7 February 2008; accepted in final form 31 July 2008
published online 9 September 2008

PACS 05.45.Tp – Time series analysis
PACS 05.40.Ca – Noise
PACS 02.50.-r – Probability theory, stochastic processes, and statistics

Abstract – This paper deals with the distinction between white noise and deterministic chaos in multivariate noisy time series. Our method is combinatorial in the sense that it is based on the properties of topological permutation entropy, and it becomes especially interesting when the noise is so high that the standard denoising techniques fail, so a detection of determinism is the most one can hope for. It proceeds by i) counting the number of the so-called ordinal patterns in independent samples of length L from the data sequence and ii) performing a χ^2 test based on the results of i), the null hypothesis being that the data are white noise. Holds the null hypothesis, so should all possible ordinal patterns of a given length be visible and evenly distributed over sufficiently many samples, contrarily to what happens in the case of noisy deterministic data. We present numerical evidence in two dimensions for the efficiency of this method. A brief comparison with two common tests for independence, namely, the calculation of the autocorrelation function and the BDS algorithm, is also performed.

Introduction. – Roughly speaking, an ordinal pattern of length L is a digest of the up-and-down in a length-L segment of a sequence whose elements can be linearly ordered. Ordinal patterns have been used in different ways to study time series since Bandt and Pompe discussed them in ref. [1]; see, *e.g.*, [2–4]. This paper deals with the application of ordinal patterns to the detection of determinism in multivariate time series corrupted by observational white noise; the univariate case was considered in ref. [5] from a more qualitative point of view. By detection of determinism we mean as usually in physics, that the data of a random-looking sequence are actually not independent with a high degree of confidence. Our method is based on some recent results on the topological permutation entropy of expansive maps of q-dimensional intervals endowed with lexicographical order [6], although we conjecture that these results hold also true for more general maps. Specifically, the orbits of expansive maps cannot materialize all possible ordinal patterns of sufficient length, contrarily to randomly generated orbits, in which any ordinal pattern is allowed with finite probability. Here and henceforth, "random" means generated by an unconstrained, stochastic process taking on arbitrary values. This result applies also to piecewise monotone maps from a one-dimensional interval into itself [7], which encompass virtually all one-dimensional interval maps encountered in practice. Missing ordinal patterns in a deterministic sequence are called (true) forbidden patterns and have two basic properties: a) robustness against observational noise and b) super-exponential growth with the length. Notice, however, that real time series are finite and thus real random sequences may also have missing ordinal patterns (called false forbidden patterns) just because of their finite length.

Noisy univariate and multivariate time series have been intensively studied in the last few decades [8,9]. Depending on the noise level of the data, one can expect to recover the full deterministic dynamics, to reconstruct the geometry of the denoised signal in some appropriate space or just to ascertain the existence of an underlying determinism. The method presented in this paper falls in the third category but, unlike other proposals, it does not rely on analytical leverage but on discrete tools: the number of forbidden patterns in a finite, random sequence and in a noisy, deterministic sequence of the same length, is different and, thanks to the properties a) and b) mentioned above, this difference persists into very high levels of observational noise. Moreover, the lexicographical

(a) E-mail: jm.amigo@umh.es

order reduces in practice to ordering vectors according to their "leading" coordinate, usually the first, so that the analysis of multivariate and univariate time series is formally the same. The bottom line is that determinism in noisy vector time series can be decided by observing a single component, a possibility that can come in handy in experimental situations. This decision requires to contrast the randomness hypothesis of the data by means of a χ^2 test based on the number of visible ordinal patterns in sliding, non-overlapping windows.

Forbidden ordinal patterns. – For definiteness, we will consider $\mathbb{R}^q$, $q \geqslant 1$, endowed with a product or lexicographical order $<$ defined as follows: $(x^{(1)}, x^{(2)}, \ldots, x^{(q)}) < (y^{(1)}, y^{(2)}, \ldots, y^{(q)})$ iff $(x^{(1)}, x^{(2)}, \ldots, x^{(q)}) \neq (y^{(1)}, y^{(2)}, \ldots, y^{(q)})$ and i) $x^{(1)} < y^{(1)}$ or ii) $x^{(k)} = y^{(k)}$ for $1 \leqslant k < q$ and $x^{(k+1)} < y^{(k+1)}$. This particular convention singles out the first coordinate as the leading coordinate, *i.e.*, the one that "decides" most of the time the order of the points. Other conventions are of course possible and may be sometimes more convenient. We say that a set $D \subset \mathbb{R}^q$ is a (q-dimensional) *simple domain* if it is obtained by a continuous and invertible deformation of a closed and bounded q-dimensional interval (think, *e.g.*, of the invariant region of the Hénon map or, more generally, of the domains considered in integration theory on $\mathbb{R}^q$); in dimension $q = 1$, simple domains are always closed intervals. As a subset of $\mathbb{R}^q$, D is also lexicographically ordered. Given the map $f: D \to D$, we say that $x \in D$ defines an *ordinal L-pattern* $\pi = \langle \pi_0, \pi_1, \ldots, \pi_{L-1} \rangle$, if

$$f^{\pi_0}(x) < f^{\pi_1}(x) < \ldots < f^{\pi_{L-1}}(x), \qquad (1)$$

where $f^0(x) \equiv x$ and $f^n(x) \equiv f(f^{n-1}(x))$. Alternatively we say that x *is of type* π, and also that π *is realized by* x. Thus, an ordinal L-pattern (or an *ordinal pattern of length* L) is just a permutation of $\{0, 1, \ldots, L-1\}$ written between angular brackets, that encapsulates the order of the points $x_k = f^k(x)$, $0 \leqslant k \leqslant L-1$, in D. For example, if $x_1 = f(x_0)$, $x_2 = f^2(x_0) = f(x_1)$ and $x_2 < x_0 < x_1$, then x_0 defines the ordinal 3-pattern (or is of type) $\pi = \langle 2, 0, 1 \rangle$.

The set of ordinal patterns of length L will be denoted by $\mathcal{S}_L$. Ordinal patterns that are missing in the orbits of f are called *forbidden patterns* for f; otherwise, they are *allowed* or *admissible*. Hence

$$\begin{aligned} |\mathcal{S}_L| &= |\{\pi \in \mathcal{S}_L : \pi \,\text{admissible for} f\}| \\ &\quad + |\{\pi \in \mathcal{S}_L : \pi \,\text{forbidden for} f\}|, \end{aligned} \qquad (2)$$

where $|\cdot|$ denotes cardinality.

Let now X be a compact metric space with metric d. A continuous map $f: X \to X$ is said to be (positively) *expansive* if there exists a $\delta > 0$ such that $d(f^n(x), f^n(y)) \leqslant \delta$ for all $n \geqslant 0$ implies $x = y$. In particular, expansive maps are sensitive to initial conditions. Intuitively, the orbits of an expansive map f can be resolved by taking a sufficiently high precision. Standard examples of expansive maps include expanding maps on the circle, topological Markov chains, and hyperbolic toral automorphisms.

It can be shown that if

(C1) f is a piecewise monotone map of a one-dimensional closed interval (*i.e.*, there is a finite partition of the interval into subintervals such that f is continuous and strictly monotone on each of those subintervals),

or

(C2) f is an expansive map of a q-dimensional simple domain, $q > 1$,

then,

$$|\{\pi \in \mathcal{S}_L : \pi \,\text{admissible for} f\}| \sim e^{L h_{top}(f)}, \qquad (3)$$

where $\sim$ means "asymptotically" as $L \to \infty$, and $h_{top}(f)$ is the topological entropy of f [1,10].

On the other hand, Stirling's formula,

$$|\mathcal{S}_L| = L! \sim e^{L(\ln L - 1) + (1/2)\ln 2\pi L},$$

spells out that the number of ordinal L-patterns defined by the orbits of f grows super-exponentially with L. We conclude from (2) and (3) that maps complying with (C1) or (C2) have necessarily forbidden ordinal patterns and, moreover, these grow super-exponentially with L. In particular, the orbits of such maps cannot realize all possible ordinal patterns for sufficiently long patterns. Details on the super-exponential growth of forbidden patterns can be found in [5,11].

Combinatorial detection of determinism. – In the literature there is a wealth of different methods to detect [4,12] and recover [13,14] the deterministic dynamics from time series contaminated with different degrees of observational noise (see also [9] and references therein). Our objective is modest in that we only seek to discriminate deterministic, noisy time series from white noise. As a compensation, the method we will describe shortly is much simpler, and it has a remarkable success even when powerful denoising techniques, like the one described in [13], fail to deliver due to the high level of noise. A comparison with two independence tests is made also below. Since the method is based on the existence of forbidden patterns, we should in principle always assume that the condition (C1) or (C2) is fulfilled, although we conjecture, based on numerical evidence with chaotic maps, that forbidden patterns exist under more general conditions. Furthermore, we can dispense with a set of Lebesgue measure zero in numerical simulations, since points of such a set have probability zero to contribute to the quantity being calculated. This means that the conditions guaranteeing the existence of forbidden patterns can be weakened in practice to hold "almost everywhere".

Thus, a basic difference between deterministic and random sequences is that the latter have no forbidden patterns on account of all ordinal patterns being allowed with finite probability. Unfortunately, real sequences are finite and this entails *false forbidden patterns*, *i.e.*, ordinal patterns missing in a finite, random or deterministic sequence just by chance. On the opposite end stand the *true forbidden patterns*, that is, those ordinal patterns that cannot appear in (finite nor infinite) deterministic sequences and whose "existence" is warranted at least under the hypothesis (C1) or (C2) —and shown numerically for more general maps. In other words, true forbidden patterns never occur in deterministic (finite or infinite) sequences, but allowed patterns need not occur in finite (random or deterministic) sequences either. In the second case, we say sometimes that a pattern is *visible* to stress that it is realized in the finite time series under analysis. Missing patterns in finite sequences can be true or false forbidden patterns and will be generically referred to as forbidden (or just *missing*).

If a data sequence has length N and we are sorting, say, K ordinal patterns down the sequence, a necessary condition for all possible patterns to be visible is certainly $L! \leqslant K$. Thus, given the length of the sequence in question, we may expect forbidden patterns to make a difference between deterministic and random generation only as long as $L! \ll K$ to allow for admissible patterns with low probability to become visible. In the numerical simulations we will refer to below, where a χ^2 goodness-of-fit hypothesis test is used, it is enough to have

$$5L! \lesssim K. \tag{4}$$

In particular, if the sequence is *white noise*, *i.e.* generated by N independent and identically distributed (i.i.d.) random variables, then the probability that any fixed ordinal pattern is missing goes to 0 exponentially as N grows.

At this point we need to consider observational noise. Two properties of forbidden patterns come to the rescue now. First of all, admissible (and thus forbidden) ordinal patterns are robust against noise because they are defined by inequalities (see eq. (1)), although the greater the sample size K, the higher the chance that a given missing pattern becomes visible. Moreover, we know already that the number of forbidden patterns grows super-exponentially with their length, so the odds are high that some of them will survive even in very noisy environments. Indeed, the simulations made in [5] with time series of the form

$$\xi_n = f^n(x_0) + w_n, \tag{5}$$

where $f(x)$ is the logistic map and w_n is white noise, confirm this expectation.

In order to distinguish white noise from noisy deterministic (uni- or multivariate) time series of the form (5), we propose a χ^2 test based on the count of visible ordinal patterns, to accept or reject the *null hypothesis*:

$$H_0\text{: the } \xi_n\text{'s are independent and identically distributed.} \tag{6}$$

Numerical simulations. – First of all, notice that the determination of visible ordinal patterns in multivariate time series can be speeded up in practice by looking at a single and the same component. Indeed, when calculating the orbit points $x_n = f^n(x_0)$, $n \geqslant 0$, where $x_n = (x_n^{(1)}, x_n^{(2)}, \ldots, x_n^{(q)}) \in D \subset \mathbb{R}^q$, we will find that, in general, $x_i^{(1)} \neq x_j^{(1)}$ for $i \neq j$, hence the ordinal patterns of f will be actually determined by the first components of the orbit points, or by any other fixed components for that matter, in correspondence to the different definitions of lexicographical order. In sum, it suffices in numerical simulations to consider the projection of the orbit of a generic initial condition on any dimension, to sort and count the ordinal patterns of the map.

Let $N_{\max}$ denote the length of the data time series under scrutiny, and let $\mathbf{n}(L,N)$ be the number of forbidden L-patterns in the initial segment $\xi_0, \xi_1, \ldots, \xi_{N-1}$ of variable length $N \leqslant N_{\max}$. In the simulations below, we count visible patterns in sliding, overlapping windows of length L. Since, in this case, the sample size is $K = N - L + 1$, we recommend to take

$$5L! \lesssim N \leqslant N_{\max} \tag{7}$$

to comply with (4).

Next, we will analyze numerically the forbidden patterns of lengths $4 \leqslant L \leqslant 7$ for self-maps in two and three dimensions. In order to estimate an average number $\langle \mathbf{n}(L,N) \rangle$ in sequences of the form (5) with $0 \leqslant n \leqslant N-1$ and N complying with the condition (7), we generate 100 samples of length $N_{\max} = 8000$ and normalize the corresponding count of missing L-patterns. In our simulations we have taken f to be the Hénon map (2D),

$$x_{n+1} = 1 - 1.4x_n^2 + 0.3y_n, \qquad y_{n+1} = x_n, \tag{8}$$

see ref. [15]. As justified above, it suffices to consider, say, the first component. As for the additive noise w_k, we have used Gaussian white noise,

$$\langle w_n \cdot w_m \rangle = \sigma^2 \delta_{nm},$$

with different standard deviations σ. The results with white noise uniformly distributed and different supports (not shown) are very much the same.

Figure 1 shows the return map ξ_{n+1} *vs.* ξ_n for a typical orbit of the Hénon map on its attractor (fractal dimension $D_0 = 1.28 \pm 0.01$ [15]) for Gaussian white noise with $\sigma = 0.25$ ($SNR \simeq 9.2\,\text{dB}$). Notwithstanding the fact that Hénon's attractor has been completely masked by the noise, the count of forbidden patterns (in logarithmic scales) in fig. 2(a) for signal plus noise, and in fig. 2(b) for only noise, are different, making feasible to distinguish between both situations by means of forbidden patterns.

Statistical analysis. – Consider now sliding windows of length L, overlapping at a single point (*i.e.*, the last point of a window is the first point of the next one) down a sequence of N entries. The number of such windows is

$$K = \left\lfloor \frac{N-1}{L-1} \right\rfloor,$$

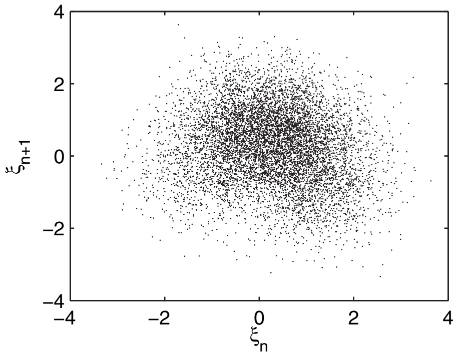


Fig. 1: Return map for a time series of the Hénon map contaminated with Gaussian white noise with $\sigma = 0.25$ ($SNR \simeq 9.2$ dB). The high noise level does not allow to recognize the underlying deterministic dynamics. However, the number of forbidden patterns is sensibly higher than in the purely random case.

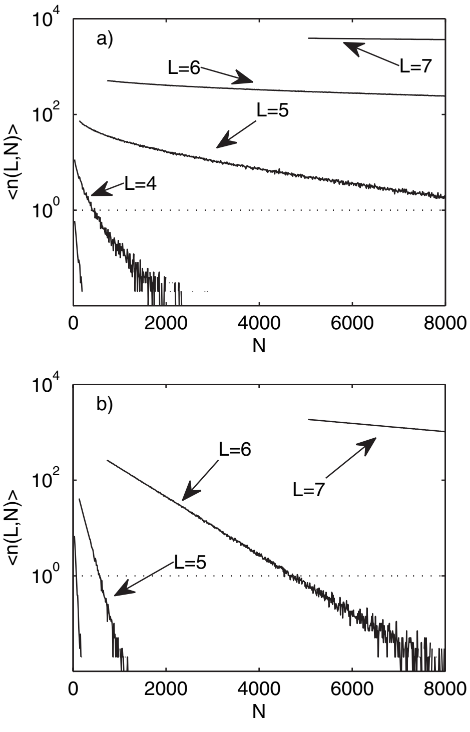


Fig. 2: Average number of forbidden patterns of length L found in a time series of length N, $\langle n(L,N)\rangle$ (in logarithmic scale), for time series of the Hénon map with Gaussian white noise with $\sigma = 0.25$ ($SNR \simeq 9.2$ dB) (a), and for a time series of Gaussian white noise (b).

each comprising the elements

$$\eta_k := \xi_{kL-k}, \ldots, \xi_{(k+1)L-(k+1)}, \quad 0 \leqslant k \leqslant K-1.$$

Notice that if the values $\xi_0, \xi_1, \ldots, \xi_{N-1}$ are independently drawn from a uniform probability distribution, then the same will happen to the ordinal patterns of length L defined by the components of $\eta_k \in \mathbb{R}^L$, which we denote by $\pi(\eta_k) \in \mathcal{S}_L$. Therefore, if one or several ordinal patterns of length L are missing in a sample obtained using non-overlapping windows, this might be a statistically significant signal that independence and/or equiprobability are/is not fulfilled. The level of significance will depend on the specifics of the sample.

Consider a time series $\{\xi_n : n \geqslant 0\}$ and the corresponding realization of the multivariate random process $\{\eta_k : k \geqslant 0\}$, and suppose that some ordinal patterns of length L are missing in the initial segment $\xi_0, \xi_1, \ldots, \xi_{N-1}$. Let ν_j be the number of η_k's such that η_k is of type π_j (*i.e.*, $\pi(\eta_k) = \pi_j \in \mathcal{S}_L$), $1 \leqslant j \leqslant L!$. Thus, $\nu_j = 0$ means that the pattern π_j has not been observed and, hence, it could be a true forbidden pattern of the time series.

In order to accept or reject the null hypothesis H_0 of (6) based on our observations (sampling, in statistical terms), we apply a χ^2 test with statistic

$$\chi^2 = \sum_{j=1}^{L!} \frac{(\nu_j - K/L!)^2}{K/L!}.$$

As said before, the rationale behind this procedure is that if the ξ_n's are i.i.d., then the ordinal patterns $\pi(\eta_k) \in \mathcal{S}_L$ are independent and uniformly distributed. If H_0 is true, then χ^2 converges in distribution (as $K \to \infty$) to a χ^2 distribution with $L! - 1$ degrees of freedom. Thus, for large K, a test with approximate level α is obtained by rejecting H_0 if $\chi^2 > \chi^2_{L!-1,1-\alpha}$, where $\chi^2_{L!-1,1-\alpha}$ is the upper $1-\alpha$ critical point for the χ^2 distribution with $L! - 1$ degrees of freedom [16]. Notice that since this test is based on distributions, it can happen that a deterministic map has no forbidden L-patterns, thus $\nu_j \neq 0$ for all j, and, however, the null hypothesis be rejected because the ν_j's are not evenly distributed.

Although in the tests below only the thresholds $\chi^2_{L!-1,1-\alpha}$ for $L = 4, 5$ and $\alpha = 0.05$ will be needed, we give next some typical values in this range. For $L = 4$ we have

$$\chi^2_{23,0.90} = 32.007;\ \chi^2_{23,0.95} = 35.172.$$

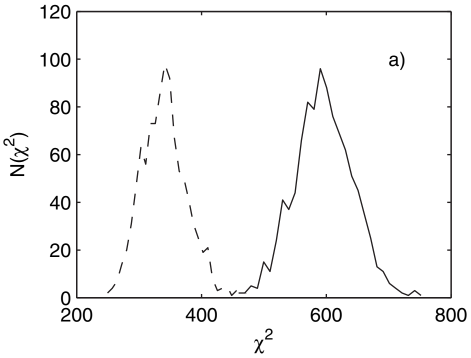
For $L \geqslant 5$, corresponding to degrees of freedom numbering more than 100, the following approximation is used [16]:

$$\chi^2_{L!-1,1-\alpha} \approx (L!-1)\left(1 - \frac{2}{9(L!-1)} + z_{1-\alpha}\sqrt{\frac{2}{9(L!-1)}}\right)^3,$$

where $z_{1-\alpha}$ is the upper $1-\alpha$ critical point for the standard normal distribution, $\mathcal{N}(0,1)$; in particular, $z_{0.90} = 1.282$ and $z_{0.95} = 1.645$. Table 1 shows $\chi^2_{L!-1,1-\alpha}$ for $4 \leqslant L \leqslant 6$ and $\alpha = 0.1, 0.05$.

Table 1: Some values of $\chi^2_{L!-1,1-\alpha}$.

$\chi^2_{L!-1,1-\alpha}$	$\alpha = 0.10$	$\alpha = 0.05$
$L = 4$	32.01	35.17
$L = 5$	139.15	145.46
$L = 6$	768.02	782.50



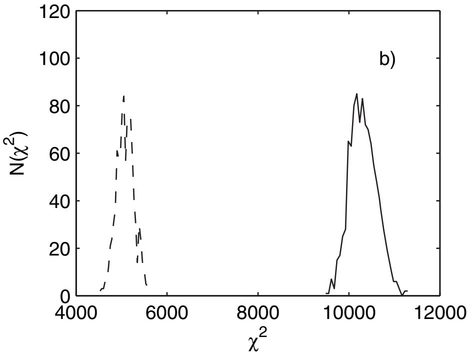


Fig. 3: Distribution $N(\chi^2)$ of χ^2 for 10000 noisy sequences generated with the Hénon map, for $L = 4$, $N = 1000$, and $\sigma = 0.10$ (continuous line), $\sigma = 0.25$ (dashed line) ($SNR \simeq 17, 9.2$ dB, respectively) (a), and for $L = 5$, $N = 8000$, $\sigma = 0.25$ ($SNR \simeq 9.2, 3.2$ dB, respectively) (continuous line), $\sigma = 0.50$ (dashed line) (b).

Furthermore,

$$\chi^2 = \frac{L!}{K}\left(\sum_{j=1}^{L!}\nu_j^2 - 2\frac{K}{L!}\sum_{j=1}^{L!}\nu_j + \left(\frac{K}{L!}\right)^2\sum_{j=1}^{L!}1\right)$$
$$= \frac{L!}{K}\sum_{j=1}^{L!}\nu_j^2 - 2K + K$$
$$= \frac{L!}{K}\sum_{j:\pi_j\,\mathrm{visible}}\nu_j^2 - K,$$

since (i) $\sum_{j=1}^{L!}\nu_j = K$ and (ii) $\nu_j = 0$ if π_j is not visible.

Figure 3 shows the distribution of the statistic χ^2 obtained for 10000 sequences generated by the *Hénon map* (8) contaminated with additive Gaussian noise. In (a) we used sequences of length $N = 1000$, non-overlapping windows of length $L = 4$ (thus, the sample comprises $K = 333$ ordinal 4-patterns) and noise with standard deviations $\sigma = 0.1, 0.25$ ($SNR \simeq 17, 9.2$ dB, respectively); the rejection threshold of the null hypothesis H_0 at level $\alpha = 0.05$ is $\chi^2_{23,0.95} = 35.17$, see table 1. In (b) we used sequences of length $N = 8000$, non-overlapping windows of length $L = 5$ (thus, the sample comprises $K = 1999$ ordinal 5-patterns) and noise with standard deviations $\sigma = 0.25, 0.50$ ($SNR \simeq 9.2, 3.2$ dB, respectively); the rejection threshold of the null hypothesis H_0 at level $\alpha = 0.05$ is $\chi^2_{119,0.95} = 145.46$, see table 1. In both cases, the χ^2 test clearly rejects H_0 based on the counts of visible ordinal patterns, with a high degree of confidence.

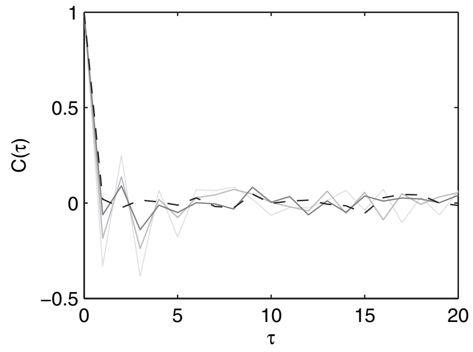


Fig. 4: Autocorrelation calculated for a time series of length $N = 1000$ of the Hénon map, in absence of noise (light gray), with $\sigma = 0.5$ ($SNR \simeq 3.2$ dB) (gray), with $\sigma = 1$ ($SNR \simeq -6.4$ dB) (dark gray) and for a random time series (dashed line). As σ is increased, the fluctuations around $C(\tau) = 0$ of the deterministic time series become closer to those of a random time series.

Comparison with other methods. – In order to evaluate the usefulness of our method, we can compare its performance with that of other approaches. A first qualitative approach to detect determinism in time series would be the calculation of the autocorrelation $C(\tau) = \langle \xi_n \cdot \xi_{n+\tau}\rangle$. In fig. 4 we can observe the computed $C(\tau)$ for a time series of length $N = 1000$ of the Hénon map contaminated with different noise levels. As the noise increases, the fluctuations of $C(\tau)$ around zero become comparable to those observed for random time series of the same length, making it difficult to distinguish between them. Thus, this approach does not lead to any significant advantage compared to counting the number of forbidden patterns appearing in the time series.

A more quantitative (statistical) method to determine whether a time series is random, is the Brock-Dechert-Scheickman (BDS) test described in [17]. This test exploits the fact that for random time series, finding two pairs of vectors of m consecutive elements whose coordinates differ in no more than a fixed ϵ should be statistically independent events. The parameter ϵ is chosen as a fraction of the standard deviation of the time series considered. This idea is substantiated in the form of a test whose statistics for random time series is perfectly characterized and, hence, can be used to accept/reject the null hypothesis H_0 (the time series is i.i.d.) for a given time series with different levels α.

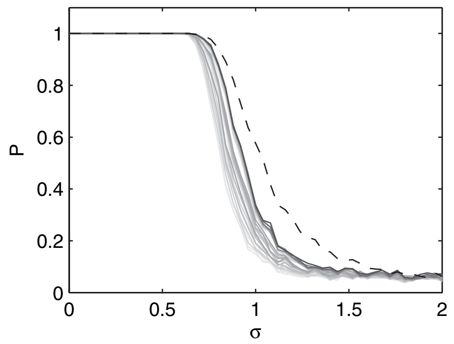


Fig. 5: Solid lines: probability P of rejecting the null hypothesis H_0 (the time series is i.i.d.) for 27 different BDS tests applied on a time series of the Hénon map of length 1000 contaminated with white noise with different σ values (SNRs down to approximately -8.9 dB). A lighter gray color indicates a bigger value of ϵ. Dashed line: probability P of rejecting H_0 using our chi-square test. All the tests have level $\alpha = 0.05$. Our test rejects the null hypothesis more often.

We have used the algorithm provided in ref. [18]. An important issue when implementing the BDS test is how to choose m and ϵ. Following the procedure given in [19] we have taken ϵ values of the form $\epsilon = 0.9^j$ and selected all the combinations of j and m such that the BDS test rejects H_0 as often as expected from the level α chosen.

This comparison is shown in fig. 5. There we show the probability P of rejecting the null hypothesis H_0 for the 27 different adequate BDS tests that can be performed for a time series of length $N = 1000$ of the Hénon map contaminated with white noise with σ between 0 and 2 (and thus with SNRs down to approximately -8.9 dB). In the same figure, we have also plotted the probability P of rejecting H_0 using our χ^2 test with the same level α. Notice that using our test the null hypothesis is rejected more often. In other words, the probability of a false positive is higher with the BDS test.

We have obtained analogous results when applying this test to time series under other conditions, and even for different chaotic systems. Furthermore, the BDS algorithm is $O(N^2)$ [18], whereas a simple estimation shows that our χ^2 test is approximately $O(N)$. This, together with the fact that our method does not require to adjust parameters like m and ϵ (we just have to investigate the distribution of the patterns of length L satisfying eq. (4)), allows us to conclude that our test compares favorably to standard methods to test for independence in time series.

Conclusion. – We proposed an ordinal pattern-based test to detect determinism in univariate and multivariate time series contaminated with observational white noise. This test exploits the fact that, under some mild mathematical assumptions, deterministic sequences exhibit forbidden patterns, while this is not the case in the random ones. Numerical simulations in two and three dimensions show that the number of forbidden patterns in noisy deterministic sequences is noticeably higher than in random sequences, even when the level of the i.i.d. uniform or Gaussian noise used in simulations, as measured by the amplitude or the standard deviation, respectively, is so high that the return map gives no hint about a hypothetical underlying dynamics. In the case of multivariate sequences, the method works also when applied to single components and orbits moving in two-dimensional attractors. Lastly, a χ^2 test based on the counts of visible ordinal patterns in non-overlapping windows, clearly discriminates between white noise and noisy deterministic time series. We have shown that our technique compares favorably to two well-established methods to test for independence in time series.

* * *

We thank our referees for their valuable comments and suggestions. This research has been supported by the Spanish Ministry of Education and Science under Projects No. MTM2005-04948 and FIS2006-08525.

REFERENCES

[1] Bandt C. and Pompe B., *Phys. Rev. Lett.*, **88** (2002) 174102.
[2] Keller K. and Lauffer H., *Int. J. Bifurcat. Chaos*, **13** (2003) 2657.
[3] Keller K. and Sinn M., *Physica A*, **356** (2005) 114.
[4] Rosso O. A., Larrondo H. A., Martin M. T., Plastino A. and Fuentes M. A., *Phys. Rev. Lett.*, **99** (2007) 154102.
[5] Amigó J. M., Zambrano S. and Sanjuán M. A. F., *Europhys. Lett.*, **79** (2007) 50001.
[6] Amigó J. M. and Kennel M. B., *Physica D*, **231** (2007) 137.
[7] Bandt C., Keller G. and Pompe B., *Nonlinearity*, **15** (2002) 1595.
[8] Abarbanel H. D. I., *Analysis of Observed Chaotic Data* (Springer, New York) 1996.
[9] Kantz H. and Schreiber T., *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, UK) 1997.
[10] Amigó J. M. and Kennel M. B., to be published in *Physica D* (DOI: 10.1016/j.physd.2008.05.003).
[11] Amigó J. M., Elizalde S. and Kennel M. B., *J. Comb. Theory, Ser. A*, **115** (2008) 485.
[12] Zhang J. and Small M., *Phys. Rev. Lett.*, **96** (2006) 238701.
[13] Mera M. E. and Morán M., *Chaos*, **16** (2006) 013116.
[14] Ortega G. J. and Louis E., *Phys. Rev. Lett.*, **81** (1998) 4345.
[15] Sprott J. C., *Chaos and Time-Series Analysis* (Oxford University Press, Oxford, UK) 2003.
[16] Law A. M. and Kelton W. D., *Simulation, Modeling Analysis*, 3rd edition (McGraw-Hill, Boston) 2000.
[17] Brock W. A., Dechert W. D., Scheinkman J. A. and LeBaron B., *Econom. Rev.*, **15** (1996) 197.
[18] LeBaron B., *Stud. Nonlinear Dyn. Econom.*, **2** (1997) 52.
[19] Liu T., Granger C. W. and Heller W. P., *J. Appl. Econom.*, **7** (1992) 25.