Test to determine the Markov order of a time series

E. Racca,* F. Laio, D. Poggi, and L. Ridolfi
Dipartimento di Idraulica, Trasporti e Infrastrutture Civili, Politecnico di Torino, Torino, Italy
(Received 20 July 2006; revised manuscript received 6 October 2006; published 25 January 2007)

The Markov order of a time series is an important measure of the "memory" of a process, and its knowledge is fundamental for the correct simulation of the characteristics of the process. For this reason, several techniques have been proposed in the past for its estimation. However, most of this methods are rather complex, and often can be applied only in the case of Markov chains. Here we propose a simple and robust test to evaluate the Markov order of a time series. Only the first-order moment of the conditional probability density function characterizing the process is used to evaluate the memory of the process itself. This measure is called the "expected value Markov (EVM) order." We show that there is good agreement between the EVM order and the known Markov order of some synthetic time series.

DOI: 10.1103/PhysRevE.75.011126 PACS number(s): 02.50.Ga, 05.45.Tp

I. INTRODUCTION

A process is Markovian of order n if

$$p(x_t|x_{t-1},x_{t-2}, \ldots ,x_{t-n}, \ldots ,x_{t-m}) = p(x_t|x_{t-1}, \ldots ,x_{t-n}), \quad (1)$$

where x is a time series of observations of the process and $p(x_t|\ldots)$ represents the conditional probability density function (PDF) of x_t, the value of x at time t. The definition in Eq. (1) corresponds to setting the state of the process at time t as a function of the previous n states of the process, from x_{t-1} to x_{t-n}, and does not depend on what happened before x_{t-n}. Note that in the literature x_t is often considered Markovian just in the case $n=1$, namely,

$$p(x_t|x_{t-1},x_{t-2}, \ldots ,x_{t-m}) = p(x_t|x_{t-1}). \quad (2)$$

Here, we consider Eq. (1) as the definition of a Markov process and we define as "first-order Markovian" a process satisfying Eq. (2). The estimation of the Markov order of a process has received constant attention over the past few decades. In fact, it is the first step in the direction of building models able to reproduce the main features of the process. Several techniques have been developed, and applications can be found in DNA-model building [1], in the design of wireless networks [2], in the analysis of chemical reactions [3], in the analysis of high-frequency financial series [4], and in the study of turbulence flow fields [5,6]. Most of these works have been devoted to analyzing the case of Markov chains, a particular class of Markov processes in which the domain of x_t is a discrete set of states (e.g., [1,7–9]). As a consequence, if one deals with continuous processes, it is necessary to discretize the state space (for example in [7,8,10]). This operation does not always give good results, particularly when the signal amplitude is wide. In fact, these methods require a low number of possible discrete states, demanding then a coarse description of the time series. Moreover, the use of these techniques has some problematic aspects. First, they were proposed to test low-order processes, whereas in many cases one is interested in higher-order Markov processes. Second, these techniques often require one to make assumptions on the Markov order of the time series, which is unknown, and to check which hypothesis gives the best results (e.g., [11–13]). This requires one to assume *a priori* the set of possible values of the Markov order of the process, which is often a difficult task when dealing with real-valued time series.

To overcome these problems, in recent years a new method has been proposed in turbulence research. In particular, to measure the Markov order of a turbulent time series, Renner *et al.* [6] used a statistical test, named the "Wilcoxon test" after its developer. This test can be applied to continuous time series, without being limited to low Markov orders and without making assumptions on its value. However, it is rather complex to implement and its performances are affected by the presence of a large number of parameters.

In this paper we introduce a further test to assess the Markov order of a process, that allows one to overcome some of the above difficulties. This test is different from the previous techniques in that it does not consider directly the conditional PDF in Eq. (1), but only its moments. In particular, we focus on the expected value of the conditional PDF. This test evaluates a measure of the memory of the process that is called the expected value Markov (EVM) order. In order to verify the qualities and drawbacks of the test, we compare the EVM order with the known Markov order n of some synthetic time series.

II. THE PROPOSED METHOD

A. The expected value Markov order

The Markov order of a process, as stated before, can be expressed by means of Eq. (1). Note that the information contained in the conditional PDF of Eq. (1) can also be obtained through the analysis of all of its moments. Therefore, an alternative way to estimate the Markov order of a process is to consider the moments instead of the conditional PDF. However, this perspective alone does not allow one to simplify the problem, because each conditional PDF has an infinite number of moments and it is clearly impossible to handle all of them. A possible simplification consists in dealing with just some of the PDF's moments. Since we are

*Electronic address: enrico.racca@polito.it

interested in proposing a simple and robust procedure, here we choose to consider only the first moment, the expected value. By applying the expectation operator E to Eq. (1), the EVM hypothesis is then expressed as

$$E(x_t|x_{t-1},x_{t-2},\ldots,x_{t-n},\ldots,x_{t-m}) = E(x_t|x_{t-1},\ldots,x_{t-n}). \tag{3}$$

We are aware that Eq. (3) is a necessary but not sufficient condition for Eq. (1) to hold. Nevertheless, if it is tenable that the two approaches provide equivalent results, then a useful simplification of the problem is possible. The aim of this paper is to show that this assumption works well and that the EVM order can be used to infer the Markov order n of the process.

A property of conditional expectation is that the expected value of x_t is not a function of x_t itself, but in general it depends on the states of the process from x_{t-1} to x_{t-m}. The problem is then to determine which is the greatest value of i, i.e. n, whose corresponding x_{t-i} contributes significantly to $E(x_t|x_{t-1},\ldots,x_{t-m})$. Toward this goal, a second simplification consists in expanding $E(x_t)$ in a Taylor series

$$\begin{aligned}E(x_t|x_{t-1},x_{t-2},\ldots,x_{t-m}) = b_0 &+ b_{11}x_{t-1} + \cdots + b_{1l}x_{t-1}^l + b_{21}x_{t-2} \\ &+ \cdots + b_{2l}x_{t-2}^l + b_1^* x_{t-1}x_{t-2} + \cdots \\ &+ b_{m1}x_{t-m} + \cdots + b_{ml}x_{t-m}^l + \cdots \end{aligned} \tag{4}$$

where the b_{ik} are the coefficients of the powers of x_{t-i}. Note that, if the functional dependence of $E(x_t)$ on x_{t-i} is not smooth, the Taylor series expansion does not exist and Eq. (4) represents only a polynomial interpolation of this functional dependence.

To determine the Markov order we search for the greatest i for which at least one of the coefficients b_{ik} is statistically different from zero. When all the coefficients are equal to zero for $i > n$, one can state that

$$E(x_t|x_{t-1},x_{t-2},\ldots,x_{t-n},\ldots,x_{t-m}) = E(x_t|x_{t-1},\ldots,x_{t-n}), \tag{5}$$

and the process is said to be nth-order Markovian in the average, i.e., the EVM order is equal to n.

Proceeding in a similar manner, it would be possible to formulate expressions analogous to Eq. (3) also for the higher-order moments. We could then define a Markov order "in the variance," one "in the skewness," and so on. However, this would correspond to losing the simplicity and robustness of the proposed approach. Therefore, we will consider only the first-order moment.

B. The evaluation of the expansion's terms

The EVM order can be found by evaluating which is the greatest i for which at least one of the terms containing x_{t-i} in Eq. (4) has a b_{ik} coefficient statistically different from zero. To do that, one should test the partial dependence of $E(x_t)$ on the regressors x_{t-i}. However, simple statistical tests for the partial dependence are not available. The available tests of independence account for the indirect dependencies too. For example, consider an auto-regressive process of the first order (AR1). If one tests the dependence of $E(x_t)$ on x_{t-2}, one will find that $E(x_t)$ depends on x_{t-2}; however, an AR1 is first-order Markovian. The dependence one finds is not a direct one, but it is due to the dependence on x_{t-1}. Owing to these difficulties in testing the partial dependence, we choose to test the partial correlation via the multiple linear regression scheme. Again, we are aware that the possible lack of partial correlation is a necessary but not sufficient condition for the partial independence of $E(x_t)$ on the x_{t-i} terms. However, since we are interested in a simple and robust test, we accept this simplification, provided that it produces acceptable results, as will be shown in the rest of the paper.

Being interested in the robustness of the test, we are also interested in avoiding the proliferation of the number of terms in Eq. (4). With this aim, we introduce a further simplification, consisting in omitting the mixed products. We will return later on to the implication of this simplification. Without the mixed terms Eq. (4) becomes

$$\begin{aligned}E(x_t|x_{t-1},x_{t-2},\ldots,x_{t-m}) = b_0 &+ b_{11}x_{t-1} + \cdots + b_{1l}x_{t-1}^l + b_{21}x_{t-2} \\ &+ \cdots + b_{2l}x_{t-2}^l + \cdots + b_{m1}x_{t-m} \\ &+ \cdots + b_{ml}x_{t-m}^l + \cdots . \end{aligned} \tag{6}$$

In the multiple linear regression scheme, the dependent variable y is expressed as a linear function of $p-1$ independent terms, z_i,

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_{p-1}z_{p-1} + \varepsilon, \tag{7}$$

where ε is the error term. In the case under analysis, $y=x_t$ and the z_i terms are the powers of x_{t-i},

$$\begin{aligned} &z_{11} = x_{t-1}, \quad z_{12} = x_{t-1}^2, \quad \cdots, \quad z_{1l} = x_{t-1}^l \\ &z_{21} = x_{t-2}, \quad z_{22} = x_{t-2}^2, \quad \cdots, \quad z_{2l} = x_{t-2}^l \\ &\vdots \\ &z_{m1} = x_{t-m}, \quad z_{m2} = x_{t-m}^2, \quad \cdots, \quad z_{ml} = x_{t-m}^l, \end{aligned}$$

where l is the maximum considered power for each x_{t-i} and m is the number of considered orders. Equation (7) can be written for each value of t. In this way we get a linear system of equations

$$\mathbf{y} = \mathbf{Z}\beta + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ is the column vector of the x_t values for each t, β is the vector of the regression coefficients, $\mathbf{Z}$ is a matrix whose columns correspond to the powers of the various x_{t-i}, whereas the rows correspond to the values they assume when t varies. Finally, $\boldsymbol{\epsilon}$ represents the vector of the differences between the x_t values and their corresponding linear estimators, for each t.

This system can be solved by means of the multiple linear regression technique (e.g., see [14]), giving the estimators $\hat{\beta}$ of the coefficients β,

$$\hat{\beta} = (\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{y}. \tag{8}$$

A test variable T_{ik} for the hypothesis $\beta_{ik}=0$ is then defined as

$$T_{ik} = \frac{\hat{\beta}_{ik}}{\sqrt{\hat{\sigma}^2 c'_{ii}}}, \tag{9}$$

where c'_{ii} is the ith diagonal element of the $(\mathbf{Z}^T\mathbf{Z})^{-1}$ matrix and the subscript ik means that we are considering the significance of the kth power of x_{t-i}. In the case that the errors ϵ have a normal distribution, the test variable T_{ik} has a Student's t distribution with $j-p$ degrees of freedom, $t(T_{ik}, j-p)$, j being the number of elements of the time series (the size of $\mathbf{y}$) and $p-1$ the number of parameters to be estimated (the size of β). The significance S_{ik} of the coefficient β_{ik}, with respect to the hypothesis $\beta_{ik}=0$, is then

$$S_{ik} = \int_{-T_{ik}}^{T_{ik}} t(T', j-p)dT'. \tag{10}$$

We consider significantly different from zero the powers x_{t-i}^k for which $S_{ik} > 1-\alpha$, where α is the selected significance level.

III. CASE STUDIES

In order to show the reliability of the introduced simplifications, we apply the test to some synthetic time series. These time series are chosen to be the observations of processes of known Markov order n. We consider the following processes: (i) the Hénon map, which is a deterministic chaotic process, (ii) two autoregressive (AR) processes, for which x_t linearly depends on the powers of x_{t-i}, and (iii) a nonlinear autoregressive (NAR) process, that instead is characterized by a nonlinear dependence of x_t on the powers of the past states of the process. This last case study is chosen to show that the omission of the mixed terms in Eq. (4) does not worsen the test's results. For each of the considered processes, we evaluate the EVM order and check if this is equal to n.

A. The Hénon map

The Hénon map is a deterministic chaotic process, generated by the equation

$$x_t = 1 - 1.4x_{t-1}^2 + 0.3x_{t-2}. \tag{11}$$

Since it is a deterministic process, it is certainly a Markov process, even if of a particular type (see [15], p. 74). In particular, the Hénon map is a Markov process of order 2, since every x_t depends only on x_{t-1} and x_{t-2}. To avoid numerical problems when handling the matrix $\mathbf{Z}$ [see Eq. (8)], we add a small dynamical Gaussian noise $0.001\xi_t$, where ξ_t has zero mean and unitary standard deviation. Figure 1(a) shows the significance of the coefficients β_{ik} obtained by applying the EVM method. As for the other case studies, we choose to represent $\log_{10}(1-S_{ik})$ instead of S_{ik}, which is the significance of x_{t-i}^k. This allows one to stretch the part of the plot closer to 1. We also fix $S_{ik}=1-10^{-6}$ when $S_{ik} \geq 1-10^{-6}$, for graphical convenience. The abscissas correspond to the powers x_{t-i}^k, with i varying from 1 to m and k increasing from 1 to l, and the ordinates are the corresponding significance values. The test correctly recognizes that only the

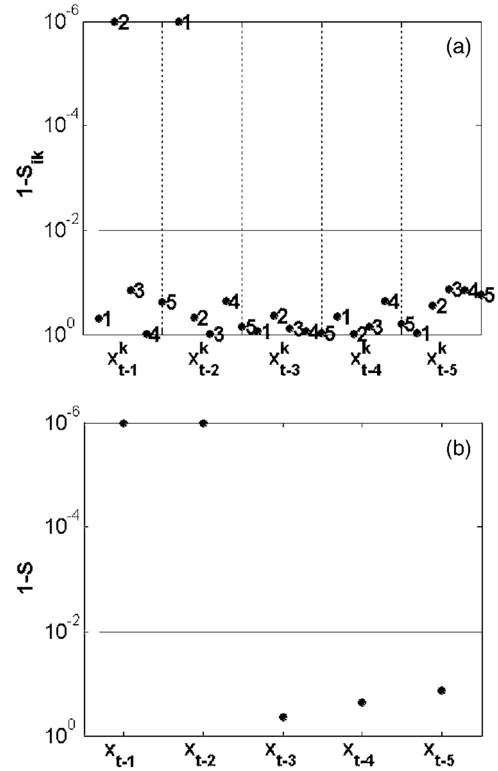

FIG. 1. Significance of the first five orders of the Hénon map (a). The horizontal line is the threshold of significance, $\alpha=0.01$: x_{t-i}^k is statistically significant when the corresponding full circle is over the threshold. Five powers are considered for each order, each corresponding to the number reported near the full circle. Only x_{t-1}^2 and x_{t-2} are statistically significant. In (b) we report only the maximum value of S_{ik}, where k varies from 1 to 5 (full circles).

coefficients multiplying x_{t-1} and x_{t-2} are statistically different from zero.

Since we are not interested in finding the exact dependence of $E(x_t)$ on the previous states $x_{t-1}, \ldots, x_{t-m}$, but to estimate the memory of the process, in what follows only the maximum S_{ik} (hereafter S) for each x_{t-i} is shown. Therefore, in the present case we obtain the diagram in Fig. 1(b), where the abscissas are the orders from x_{t-1} to x_{t-m}, whereas the ordinates correspond to the maximum significance among the powers of the corresponding order. The process is recognized from this figure to have EVM order is equal to 2, which coincides with the actual Markov order of the process.

Some methodological details deserve further comments. Here and in the following cases, the significance level α is fixed at 0.01. Being $\alpha=0.01$, one has a 1% probability to consider statistically significant the contribution of a nonsignificant term. Since we choose to represent just the maximum among the S_{ik} values, with fixed i and k varying from 1 to l, the probability is greater than α and approximately equal to $l\alpha$. However, this expression for the significance would be valid just in the case the powers x_{t-i}^k were independent. Since they are not, the actual significance is between α and $l\alpha$. The number l of powers for each order is chosen to be equal to 5 in all of the case studies, whereas the number of orders m is initially fixed at 5, and successively increased if necessary, following the indications reported in the next subsection. The choice of the number of orders and powers in the Taylor expansion of Eq. (4) is particularly important and has effects on the stability, robustness, and power of the test. The following section is devoted to this topic. In this section we explain also the reason for the choice $m=l=5$.

B. The choice of the number of orders and powers in the expansion

Before proceeding with the other case studies, it is necessary to establish the behavior of the test when l and m are changed. Moreover, the stability, robustness, and power of the test need to be evaluated when the series length N varies. This allows one to understand which criteria have to be considered in the choice of l and m in the Taylor expansion of Eq. (6). In building the test we are interested in increasing m and l and to include the mixed products of powers of the terms from x_{t-1} to x_{t-m}, in order to obtain the best possible approximation of E in Eq. (4). On the other hand, when the number of the expansion's terms is increased, the stability, robustness, and power of the test are reduced as a consequence. The final choice should be a balance between these two contrasting necessities.

Consider first the case of the mixed products: if one keeps them in the expansion, the number of considered powers should be limited to 2 or 3, to avoid an excessive increase of the number of β coefficients. A possible alternative is to omit the mixed terms and increase the number of considered powers. This second alternative is found to give better results. As for the choice of m and l, we explore the behavior of the test for the Hénon map when a different number of powers l is chosen for a fixed number of orders m [see Fig. 2(a)], and when different values of m are investigated with a varying l [Fig. 2(b)]. We find that the power of the test decreases when the total number of parameters grows. In both Figs. 2(a) and 2(b), this can be seen by observing that the power of the test becomes lower for a fixed length N of the series, when the number of powers or the number of orders is increased. This behavior is reasonable, because one has to evaluate a greater number of parameters with the same number N of elements; as a consequence the uncertainty grows. On the other hand, one cannot excessively decrease l and m because otherwise the Taylor expansion of E in Eq. (4) would lose its validity, when more complex processes are considered. As mentioned, a compromise should be searched for, which we find by taking $l=5$ and $m=5$. It is important to notice that for $m=3$ the power of the test is almost constant and close to 100 [see Fig. 2(b); the exact value is ~97]. This is due to the fact that the Hénon map is a second-order Markov process, so 3 is the most efficient choice for m in this case. It could, however, happen that the chosen m is lower than the Markov order of the process. In this case we find that x_{t-m} results to be significant. When the last considered order x_{t-m} is significant, one must increase m in a progressive manner, until x_{t-m} turns out not to be significant. We show an example of this procedure in the following, by analyzing the case of the AR10 process.

The results reported in Figs. 1(a) and 1(b) refer to an analysis carried out on a time series of 10 000 elements. The result of the test is stable when other series of the same length are generated and analyzed. In fact, for 93 out of 100 generated series the test correctly recognizes the Hénon map as a second-order Markov process. Note that the power of the test is not simply equal to $1-\alpha$ because we are testing the maximum of the S_{ik} values rather than each of them separately. We also find that the power of the test ceases to increase when N is greater than 250, for every value of l and m. Similar results can be observed also for the other case studies, i.e., the method does not need a great amount of data to be applied with success.

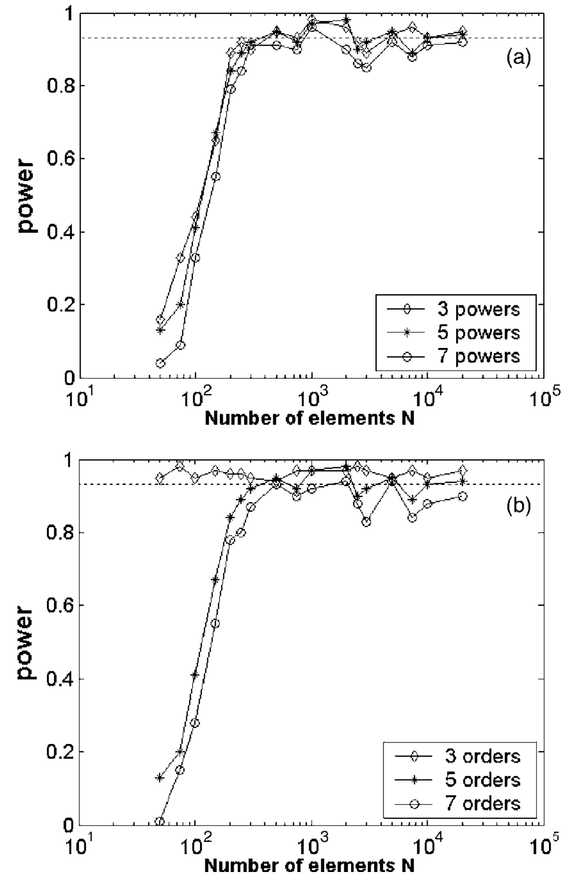

FIG. 2. (a) Percentage of successes of the proposed test when it is applied with five orders and a number of powers increasing from 3 to 7. (b) Percentage of successes of the proposed test when it is applied with five powers and a number of orders increasing from 3 to 7. In both cases the series' length N varies from 50 to 20 000 elements and the horizontal line represents the power of the test for $m=l=5$ and $N=10\,000$.

C. An AR1 process

The AR1 process is an autoregressive model of the first order, governed by the equation

$$x_t = \rho x_{t-1} + \sigma \xi_t, \tag{12}$$

where $\rho=0.8$, $\sigma=0.5$, and ξ_t is a Gaussian white noise term with zero mean and unitary standard deviation. The AR1 is obviously a Markov process of the first order, i.e., $n=1$. The test works well also in this case (see Fig. 3): in fact, the test

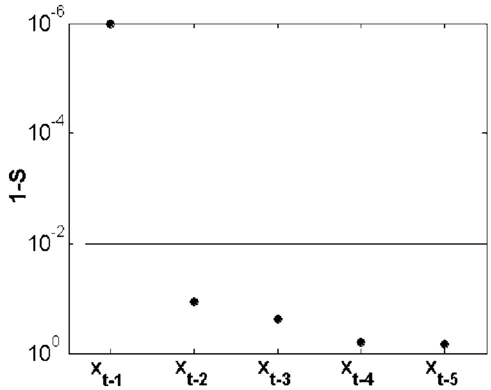

FIG. 3. Significance of the first five orders of the AR1 process. The first order is significant, so that the EVM order is equal to 1.

correctly determines an EVM order equal to 1. This happens for 89 out of 100 series of 10 000 elements each. The results do not change when the series length varies, provided that at least 150 elements are available.

D. An AR10 process

The AR10 process is an autoregressive process of the tenth order, whose present state x_t depends on x_{t-1} and x_{t-10} according to the expression

$$x_t = a_1 x_{t-1} + a_{10} x_{t-10} + \xi_t, \tag{13}$$

where $a_1 = -0.5$, $a_{10} = -0.5$, and ξ_t has the same properties mentioned in the preceding sections. We first apply the test with five orders and find that x_{t-5} is significative [Fig. 4(a)]. Since the last order is significant, we then increase the number of orders, still obtaining the same results for all $m \leq 10$ [Figs. 4(b) and 4(c) report the results for seven and nine orders]. Finally, the test is applied with m greater than 10, with good results [see Fig. 4(d) for the case $m = 15$]. In fact, the test establishes that the EVM order is equal to 10. The results are again stable when one considers different series of the same length, with powers slightly lower than before (82 successes out of 100, for a 10 000-element series). This is an expected consequence of the fact that the number of considered orders has to be increased. The series length above which the power of the test ceases to increase is 500 elements in this case.

E. A nonlinear autoregressive process

To demonstrate that the results do not change when neglecting the mixed products in Eq. (4), we apply the test with and without the mixed products to a time series that is the observable of a nonlinear autoregressive model, generated by the equation

$$\begin{aligned} x_t = & -0.00113 + 0.0613 x_{t-1}^2 - 0.0628 x_{t-2} - 0.053 x_{t-2} x_{t-1} \\ & + 0.0573 x_{t-2} x_{t-1}^2 - 0.0234 x_{t-2}^2 x_{t-1} - 0.0675 x_{t-3} - 0.02 x_{t-3}^2 \\ & - 0.5071 x_{t-3} x_{t-2} - 0.044 x_{t-3}^2 x_{t-1} + 0.001 \xi_t, \end{aligned}$$

where ξ_t is the same as in the previous cases. This NAR process is third-order Markovian, since x_t depends on x_{t-1}, x_{t-2}, and x_{t-3} (i.e., $n=3$). Depending in a complex way on the mixed terms, the NAR process is an ideal candidate to check if the test works also in this case. In this case study, we apply the test with $l=3$, because it is necessary to evaluate a greater number of parameters to take into account the mixed

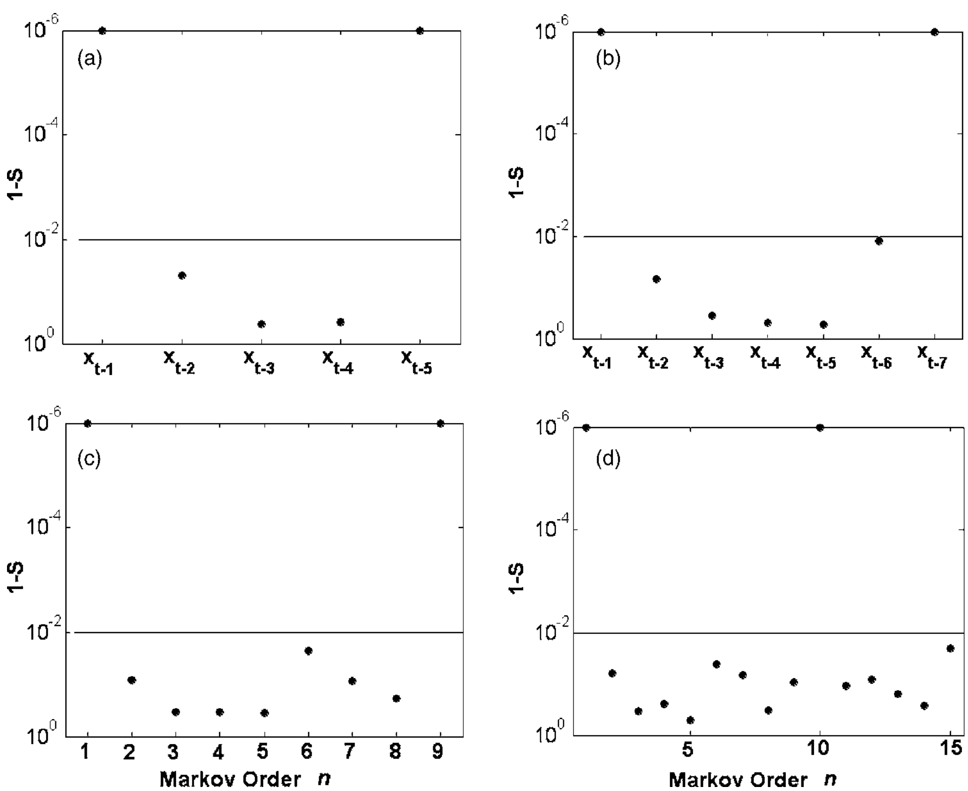

FIG. 4. (a) Significance of the first five orders of the AR10 process; x_{t-5} is significant, so the number of orders is increased to (b) 7 (c) 9, obtaining the same behavior; (d) finally the number m of considered orders is increased to 15 (d): the algorithm correctly recognizes the AR10 as a tenth-order process.

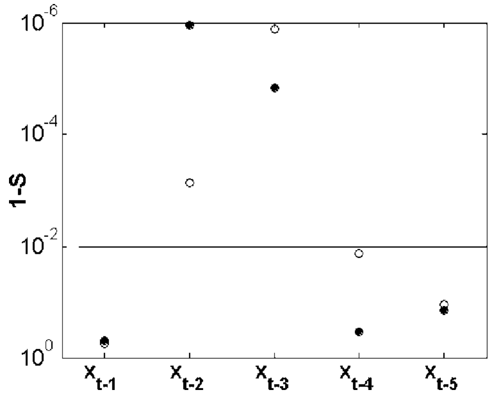

FIG. 5. Significance of the first five orders of the considered NAR process. The empty circles correspond to the algorithm with the mixed products, whereas the full ones correspond to the Taylor expansion of E without these terms. In both cases the test recognizes the considered process as third-order Markovian (EVM order is equal to $n=3$).

products in the Taylor expansion (i.e., $l=5$ would make the test unstable). There is not a great difference in the results the test gives with and without the mixed products (see Fig. 5). We find that the test, also when one neglects the mixed terms, allows one to obtain good results, since it estimates EVM order equal to 3, with high power (91 out of 100 positive answers, when the series has 10 000 elements). However, the minimum series length that guarantees a positive result is greater than before. The series must consist of, at least, 10 000 elements, both when one takes into account the mixed products and when one neglects them. It is interesting to mention that x_{t-1} in both cases is not significant, whereas it appears in the equation of the process. This behavior is probably due to the fact that there is just one term that depends only on x_{t-1} and it has a very low multiplicative coefficient. The other terms containing x_{t-1} depend also on other x_{t-i} terms, with $i>1$, and they are attributed to the greater ith orders: i.e., the term $-0.0234x_{t-2}^2x_{t-1}$ is considered contributing to order 2 in the multiple linear regression technique. Thus, the dependence on x_{t-1} is so weak that it is "covered" by the stronger dependencies on x_{t-2}, x_{t-3}, and their products.

IV. CONCLUDING REMARKS

In the past few decades the problem of determining the Markov order of a time series has received constant attention, being the first step in the direction of building a model able to reproduce the characteristics of the process. Many works have been devoted to this topic, but a simple, flexible technique that can be applied to a continuous time series of unknown Markov order is still lacking. To overcome some of the above difficulties, a simple and robust test to determine the Markov order of a process has been proposed. The test measures the memory of the process by means of the conditional PDF's first-order moment alone. To reach this aim a Taylor expansion of the expected value is performed and the statistical significance of the terms in the expansion is evaluated using a multiple linear regression scheme. The maximum i for which the terms containing x_{t-i} give a statistically nonzero contribution to the expected value of x_t represents a measure of the memory of the expected value. This measure, used to evaluate the Markov order of the process, has been called the expected value Markov order. The use of the proposed test has some advantages. First, the test has a high power. In fact, applying the test to some synthetic time series of known Markov order n, the EVM order coincides with n in a high percentage of realizations. Second, the test is robust, i.e., it maintains a high power also when applied to short time series. Finally, the test is simple to use and can be applied to both continuous and discrete time series, without making *a priori* assumptions on the Markov order of the process.

ACKNOWLEDGMENT

This work has been partially supported by Fondazione Cassa di Risparmio di Torino—Progetto Lagrange.

[1] C. Kelly, Biometrics **50**, 653 (1994).
[2] M. R. Hueda and C. E. Rodriguez, IEEE Trans. Veh. Technol. **54**, 425 (2005).
[3] A. Fuliński, Z. Grzywna, I. Mellor, Z. Siwy, and P. N. R. Usherwood, Phys. Rev. E **58**, 919 (1998).
[4] Ch. Renner, J. Peinke, and R. Friedrich, Physica A **298**, 499 (2001).
[5] R. Friedrich and J. Peinke, Phys. Rev. Lett. **78**, 863 (1997).
[6] Ch. Renner, R. Friedrich, and J. Peinke, J. Fluid Mech. **433**, 383 (2001).
[7] I. Nagy, Z. Gingl, L. B. Kiss, and J. Vinkó, Physica B **216**, 79 (1995).
[8] M. J. van der Heyden, C. G. C. Diks, B. P. T. Hoekstra, and J. DeGoede, Physica D **117**, 299 (1998).
[9] H. Tong, J. Appl. Probab. **12**, 488 (1975).
[10] S. M. Pincus, Proc. Natl. Acad. Sci. U.S.A. **89**, 4432 (1992).
[11] R. W. Katz, Technometrics **23**, 243 (1981).
[12] M. Harrison and P. Waylen, Int. J. Climatol. **20**, 1861 (2000).
[13] F. G. Karssenberg, C. Piel, A. Hopf, V. B. F. Mathot, and W. Kaminsky, Macromol. Theory Simul. **14**, 295 (2005).
[14] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, 3 Vols. (Charles Griffin & Co., London, 1967, 1968, 1969).
[15] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).