Autoregressive moving average model for analyzing edge localized mode time series on Axially Symmetric Divertor Experiment (ASDEX) Upgrade tokamak

G. Zvejnieks[a] and V. N. Kuzovkov
Institute of Solid State Physics University of Latvia, Euratom-University of Latvia Association, 8 Kengaraga str., Riga LV-1063, Latvia

O. Dumbrajs
Helsinki University of Technology, Euratom-Tekes Association, FIN-02015 HUT Helsinki, Finland

A. W. Degeling
Centre de Recherches en Physique des Plasmas, École Polytechnique Fédérale de Lausanne, Association EURATOM, Confédération Suisse, Lausanne CH-1015, Switzerland

W. Suttrop, H. Urano, and H. Zohm
Max-Planck-Institut für Plasmaphysik, Euratom Association, Garching D-85740, Germany

(Received 6 April 2004; accepted 16 September 2004; published online 12 November 2004)

An approach to analysis of time series of edge localized modes (ELMs) is proposed. It is based on the use of the autoregressive moving average model, which decomposes time series into deterministic and noise components. Despite the inclusion of nonlinearity in the model, the resulting deterministic equations for the ELM time series measured on Axially Symmetric Divertor Experiment Upgrade tokamak turn out to be linear. This contrasts with the findings on JAERI tokamak (JT-60U) and tokamak à configuration variable that ELMs exhibit features of chaotic dynamics, namely, the presence of unstable periodic orbits. This methodology for distinguishing chaotic behavior is examined, and found to be susceptible to misinterpretation. © *2004 American Institute of Physics*. [DOI: 10.1063/1.1814368]

I. INTRODUCTION

It was first observed at the Axially Symmetric Divertor Experiment (ASDEX) tokamak in 1982, see Ref. 1, that externally heated tokamak plasmas can suddenly reach an operating regime of improved confinement. The regimes of low and high confinement are referred to as *L* mode and *H* mode, respectively.[2] The transition from *L* mode to *H* mode is normally accompanied by appearance of recurrent magnetohydrodynamic instabilities known as edge localized modes (ELMs).[3,4] These manifest themselves as short bursts of energy and particles as the outer layer of plasma is suddenly pealed off and then flows along the magnetic field lines to the divertor plates.

The loss of energy and particles deteriorates confinement, which would suggest that ELMs are not very desirable. Additionally, the short but intense particle and power loads on divertor plates cause erosion of the plates, which might become a serious concern in large, future machines, such as international thermonuclear experimental reactor.[5] On the other hand, ELMs exhaust impurities and helium ash that otherwise would accumulate in the plasma and eventually terminate the fusion burn, and they provide a means for density control. With few exceptions, see Ref. 6, stationary *H* mode has been observed only in the presence of ELMs (ELMy *H* mode). Hence the understanding and control of ELMs is vital for the development of fusion power.

Although a number of ELM types have been classified and there are several models for ELMs,[7] the physics of ELMs is still to a large extent unresolved. An alternative approach to the analysis of ELMs is to determine whether the fluctuations in the time interval between ELMs (ELM period) are the result of a deterministic process in chaotic state or whether they are random. This approach was initiated in Ref. 8 on JAERI Tokamak (JT-60U) (Ref. 9) in 1999 and continued in Ref. 10 on tokamak à configuration variable (TCV).[11] Recently these studies have been extended in Refs. 12 and 13 by analyzing data from ASDEX Upgrade tokamak.

The motivation for this approach is that chaos can be controlled even without any prior analytical knowledge of the system dynamics. All chaotic dynamical systems have an infinite number of periodic solutions that are unstable. Such solutions are called unstable periodic orbits (UPOs) and they are the signature of chaotic systems. Chaos can be controlled by stabilizing UPOs, as proposed in Ref. 14. This control scheme uses very small perturbations of available control parameters of the system and has found a wide utility in many branches of natural sciences and medicine (see, e.g., Ref. 15 for a review).

In Refs. 8, 10, 12, and 13 UPOs have been searched by means of the recurrence method. In Refs. 12 and 13 also the fixed point transform method was employed for this purpose. In a number of ELM time-series UPOs indeed have been found and on this basis a claim has been raised that at least traces of deterministic chaos have been detected in some ELM time series.

[a]Electronic mail: guntars@latnet.lv

The goal of the present work is to show that this claim might be premature. We use a method based on the financial engineering principles, which allow the decomposition of the ELM time series into deterministic and noise components. One can expect that a sufficiently complicated nonlinear deterministic part would indicate the presence of deterministic chaos, while a simple linear deterministic part would deny this hypothesis. On the basis of this analysis it is also argued that presence of UPOs in time series is a necessary, but not a sufficient condition for ubiquity of deterministic chaos.

The structure of the paper is as follows. Section II briefly describes the two known methods—the recurrence method and the fixed point transform method—for detecting deterministic chaos in time series. Section III introduces both the autoregressive moving average (ARMA) and the nonlinear autoregressive (NAR) methods. Section IV contains the results of the ARMA and NAR applications including analysis of ELM time series measured at ASDEX Upgrade. Conclusions are presented in Sec. V and the mathematics is summarized in the Appendix.

II. DETECTION OF DETERMINISTIC CHAOS

A. Recurrence method

Information about ELMs is stored in the form of discrete time series originating from the D_α signal. This information is processed in order to simplify the further analysis. For example, for type I ELMs maxima of D_α signals are used for determination of ELM times t_n, as given in Ref. 16, and subsequently only ELM periods, i.e., differences between successive maxima $T_n = t_n - t_{n-1}$ are considered. Just these new series constitute the subject of our analysis and are denoted by $y = \{y_1, y_2, \ldots, y_N\}$ in what follows.

A dynamical system is in chaotic state if all its periodic solutions are unstable. If the phase space is bounded, as it must be for physical solutions, then recurrent nonperiodic behavior ensues. Occasionally, when the solution is close to an unstable periodic solution (or orbit), its displacement from the periodic solution is briefly described by the linearized system.

A useful method for finding UPOs in scalar, discrete data is the so-called *k*th return map, which is a scatter plot of all y_{i+k} versus y_i. On the return map, the consecutive points of the departure phase with exponential divergence follow straight lines as illustrated in Fig. 1.

In Fig. 2 is illustrated how the straight lines on the return map look like in the actual time series. The exponential divergence away from the unstable fixed point is clearly seen. The return map suggests that period-one UPOs can be found by searching for sets of consecutive points in the data that follow a straight line on the return map. The unstable fixed point is found at the intersection of this line with the diagonal.

B. Fixed point transform

The fixed point transform[17] concentrates on the fixed point instead of sequences of collinear points. It is based on the transformation $\hat{y}$ of a given time series y, by

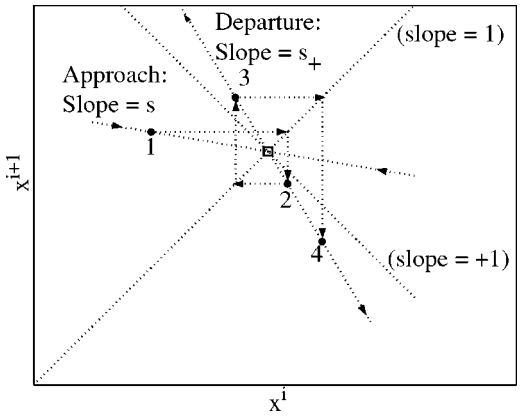

FIG. 1. Schematic diagram of an UPO on the return map. The square denotes the fixed point. Reproduced with permission from Ref. 12. (Here x^i corresponds to y_i in our notations.)

$$\hat{y}_i = \frac{y_{i+1} - s_i(k) y_i}{1 - s_i(k)}, \tag{1}$$

where

$$s_i(k) = \frac{y_{i+2} - y_{i+1}}{y_{i+1} - y_i} + k(y_{i+1} - y_i). \tag{2}$$

The idea of the transformation is that the transformed points cluster near the fixed points of the original series.

C. Statistical tests

As noted in Ref. 10, the finding of any individual signal resembling an UPO carries no significance in itself because it is also likely for such signals to occur by chance in a nonchaotic noise driven system. The key question is whether the statistical properties of such signals, called UPO candidates,

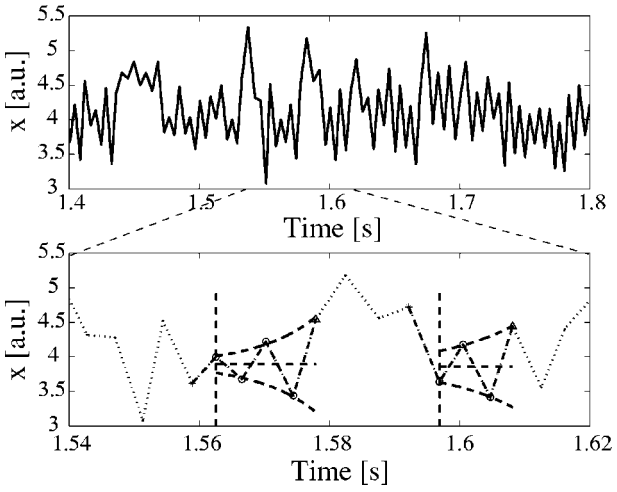

FIG. 2. Illustration of the exponential divergence in the departure phase. The beginning of the approach phase is denoted by an asterisk, the points in the departure phase are denoted by circles except the last point, which is denoted by a triangle. The position of the fixed point is denoted by the horizontal dashed line and the exponential divergence is shown by the curved dashed lines. The vertical dashed line denotes the point of closest approach. Reproduced with permission from Ref. 12. (Here *x* corresponds to *y* in our notations.)

are significantly different from those that would be expected if they all occurred by chance. In order to answer this question, the statistics of such chance occurrences needs to be defined. This is done by creating surrogate data that have the same statistical properties as the original data and comparing the results of a statistical test on the original and surrogate data. Arguably the easiest way to create surrogates is to shuffle, i.e., randomly reorder, the original time series. This preserves the probability distribution but effectively destroys any possible determinism in the data.

The null hypothesis in the statistical analysis for both methods is that all detected UPOs or fixed points are chance occurrences governed by the probability distribution function of the data. In such a case no information is carried in the ordering of the time series, and hence any reordering should be equivalent to the original series and give rise to a similar number of UPOs or similar peaks in the histogram of the fixed point transform.

Quantitatively it is described in the recurrence method by the UPO number deviation from the mean

$$K = \frac{N - \langle N_s \rangle}{\sigma_s}, \tag{3}$$

where N is the number of detected UPOs in the original time series, $\langle N_s \rangle$ is the mean of fake UPOs in surrogate sets, and σ_s is the standard deviation of fake UPOs in surrogates. One can reject the null hypothesis if $K \geqslant 3$.

The statistical tests of the fixed point transform are very similar to those of the recurrence method. Many different realizations of surrogate data are generated by shuffling the original time series, and the same procedure is applied for calculating the histogram approximation of the probability density function.

III. STATISTICAL DESCRIPTION OF TIME SERIES: THE ARMA METHOD

A. General aspects

The UPO approach assumes that in the time series the cause-effect relationship is hidden. We express this assumption as existence of some unknown nonlinear function f

$$y_n = f(y_R), \tag{4}$$

where y_R denotes past observations $\{y_{n-1}, \ldots, y_{n-R}\}$ (the parameter R determines the maximum order of past observations taken into account). Certain relationships between few successive observation[18] (UPOs) indicate that the function f must be sufficiently complicated to be able to generate the deterministic chaos. Controlling chaos does not require the knowledge of the complete nonlinear function f. It is sufficient to know its linearized form in the vicinity of UPOs.[18]

In the case of ELM time series, the initial assumption Eq. (4) is considered too optimistic. These time series are projections of complicated processes occurring in the plasma and one has to assume the existence of "hidden" degrees of freedom in the series which manifest themselves as noise. We proceed differently from the UPO method and replace Eq. (4) by its more general stochastic recursive form

$$y_n = F(y_R, \varepsilon_M) + \varepsilon_n, \tag{5}$$

where ε_n represents an innovation. Innovations are defined as a sequence or unanticipated disturbances. The simplest assumption is that innovations are random numbers obeying the normal distribution. A more general concept assumes that the variance (squared standard deviation) of innovations can be time dependent.

Thus F is an unknown function of past observations y_R and past innovations $\varepsilon_M = \{\varepsilon_{n-1}, \ldots, \varepsilon_{n-M}\}$, where M is maximum order of past innovations to be taken into account. Hence, time series can be decomposed into deterministic and random components. The deterministic component can be associated with signal forecasting, while the random component describes the error or uncertainty of prediction. In what follows we will show that the number of needed past observations R and past innovations M can be obtained by analyzing the initial time series. On the basis of correlation analysis we will restrict ourselves to the class of functions which depend on past observations. This form resembles the simple Langevine equation

$$y_n = f(y_R) + \varepsilon_n. \tag{6}$$

If decomposition of the signal into deterministic f and random ε components is possible (this issue will be discussed later), the deterministic approach could be justified if Eq. (6) would be complex enough for generating the deterministic chaos in the limit $\varepsilon \to 0$. In such a case analysis of Eqs. (4) and (6) lead to the same results. However, in a general case there are no reasons to assume that the deterministic part in Eq. (6) has a sufficient complexity.

B. The ARMA model

Let us begin with the following assumptions. We consider ELM time series as a statistical phenomenon. Events y can be described by means of some statistical distribution. Generally speaking, this distribution is non-Gaussian.[10] In comparison with the normal distribution, the histograms show thin waists and heavy tails. These characteristics are quite common for data obtained in such areas as hydrology, reliability engineering, telecommunications, (re)insurance, and finance.[19] We note that these quite distinct research fields face mathematically identical problems. Our interest in *statistical models* describing ELM events is strongly motivated by the potential to control ELMs.

We proceed by introducing discrete models for ELM time-series y, since their structure is discrete as well. One of the simplest models which can be used for their description is the general ARMA model, which encompasses autoregressive (AR) and moving average (MA) parts in any combination. The general ARMA (R,M) model is obtained from Eq. (5) with the function F,[20,21]

$$F(y_R, \varepsilon_M) = C + \sum_{i=1}^{R} (\mathrm{AR})_i y_{n-i} + \sum_{j=1}^{M} (\mathrm{MA})_j \varepsilon_{n-j}, \tag{7}$$

where C, $(\mathrm{AR})_i$, and $(\mathrm{MA})_j$ are coefficients. It is easy to notice, that the dependence on past events y_R and past innovations ε_M is introduced linearly in the ARMA model.

The optimal order R of (AR) terms and M of (MA) terms can be estimated by qualitative analysis of correlation structures of events. For example, the cutoff lag of the autocorrelation function (ACF) indicates the order of (MA) terms, while the cutoff lag of the partial autocorrelation function (PACF) indicates the order of (AR) terms. A quantitative estimate of the order of the model can be found from the extremum principle, e.g., Bayesian information criterion (BIC). Similarly, values of the coefficients C, $(\mathrm{AR})_i$, and $(\mathrm{MA})_j$ are found from the extremum condition for the log-likelihood function (LLF), which determine the quality of decomposition of time series into the deterministic and random parts. More detailed discussion of these methods is given in the Appendix. Thus, the model order is fixed by means of two criterions. First, the LLF part ensures that for the given model order, R and M innovations of the decomposed signal resemble the normal (see, the Appendix) distribution as closely as possible. Second, the BIC criterion is used to select the best model order by finding the optimal balance between the data description quality for the given model order, and the number of model parameters. This plays the role of the penalizing factor (see, e.g., the Appendix).

The necessity of such a penalizing factor can be demonstrated by the following simple example. The recursive equation depending on the first past event $y_n = c_1 y_{n-1}$ can be written as the recursive equation depending on the second past event $y_n = c_1^2 y_{n-2}$. However, the same equation can be rewritten as, e.g., $y_n = a c_1 y_{n-1} + (1-a) c_1^2 y_{n-2}$, where a is some arbitrary constant. Thus, equations of different order result in identical LLF, but with different information content. BIC must discard all phantom equations containing higher order terms. As was noted above, the correlation structure of real ELM time series is such that (MA) terms can be omitted in the ARMA model ($M=0$). In such a case one can say that the autoregressive form (AR) of the type of Eq. (6) (Langevine type equation) can be used as a linear approximation, describing those time series where the function $f(y_R)$ is given by

$$f(y_R) = C + \sum_{i=1}^{R} (\mathrm{AR})_i y_{n-i}. \tag{8}$$

C. The NAR form

The analysis of the structure of the function $f(y_R)$ can serve as an additional criterion for distinguishing between noise and the deterministic chaos. If the decomposition procedure is able to distinguish between the deterministic and random components of the initial signal with the sufficient accuracy, we can interpret the signal as the deterministic chaos only in the case, when Eq. (6) predicts chaos in the limit $\varepsilon_n \to 0$. We must generalize the approach used in the linear AR form to the NAR form.[22–24] This generalization step is discussed in details in the Appendix. However, the NAR method has the following natural drawback. The BIC procedure is based on the selection of the best parameter order among all combinatorially possible cases. Complementing the AR form by nonlinear terms considerably increases the number of such cases.

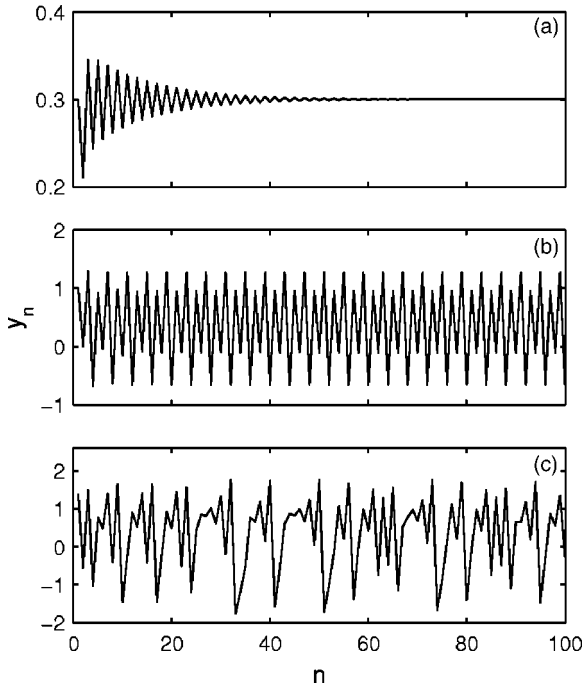

FIG. 3. The Henon map in the absence of noise, Var=0. Here $b=0.3$ and $a=0.3$ (a), $a=1.0$ (b), $a=1.4$ (c).

IV. APPLICATIONS OF THE ARMA AND NAR METHOD

A. Henon map examples

Before practical application let us test the NAR method on time series with good statistical properties (the length of the signal has to be sufficient) and known deterministic component. As a test model we select the well known Henon map.[25] This corresponds to the following choice of the function $f(y_R)$:

$$f(y_R) = a - y_{n-1}^2 + b y_{n-2} \tag{9}$$

in Eq. (4). Depending on the values of the two parameters a and b, Eq. (9) leads to qualitatively different types of behavior. In the absence of noise one can observe relaxation, periodic, and chaotic oscillations, see Fig. 3, respectively. We generate the test time-series y by means of Eq. (6) with the function Eq. (9) and the given statistical properties for innovations ε (innovations are quantified by the Var, variance which is the standard deviation squared). These time series are analyzed by means of the NAR method, i.e., by decomposing the series into the deterministic and random components. The quality of the decomposition is determined by agreement between statistical properties of the initial random component and the calculated innovations, on the one hand, and agreement between the initial function Eq. (9) and the calculated deterministic part, on the other hand.

The NAR method is a statistical method of decomposing the test signal into the deterministic and random components. The best results obtained with the NAR correspond to cases

TABLE I. Calculated NAR form for the Henon map model, Eq. (9). Here $b=0.3$.

a	Var	NAR form	Err	Var	BIC
1.4	0	$1.4-1.0y_{n-1}^2+0.3y_{n-2}$	$\pm 10^{-5}$	$(2\pm 1)\times 10^{-9}$	−16 351
0.3	0.05	$0.27-0.96y_{n-1}^2+0.31y_{n-2}$	±0.04	0.053±0.002	−70
0.3	0.005	$0.3-0.9y_{n-1}^2+0.3y^{n-2}$	±0.1	0.0052±0.0002	−2402

when the spectrum of y values is of the order of the signal length. Neither relaxation, nor periodic oscillations fulfill this condition in the absence of noise. Indeed, after a short instability region the relaxation leads to some constant value $y=\text{const}$. It is impossible to recover the generating equation and the corresponding constants from a single value, because of insufficient statistics. On the other hand, in considering only the instability region of relaxation with a sufficient number of events, we can rely on the NAR method. The same can be said in the case of periodic oscillations with a limited set of events y. At first the NAR method gives linear AR forms which describe time series but depend on initial conditions. However, by omitting linear cases, we can still recover the generating polynomial. In the presence of noise these limitations disappear.

As the first test example for the NAR method, we choose time series generated by the Henon map in a chaotic regime without noise with parameters $a=1.4$, $b=0.3$, and Var=0.0. The shift in time R (the number of relevant previous innovations which have to be taken into account) in the function $f(y_R)$ is estimated on the basis of the cutoff lag of the PACF. This indicates that $R=2$ in accordance with the function Eq. (9) which memorize two previous events.

The results obtained by means of the NAR method are shown in Table I. It is seen that in this case the NAR method leads to the exact equation (the needed polynomial terms are distinguished) and to the corresponding coefficients (governing parameters). It should be emphasized that the coefficients are determined with high accuracy: the error (Err) is of the order 10^{-5}. Moreover, the method indicates that the noise level is below 10^{-9} (see Table I). These small errors can be attributed to the errors inherent in the numerical methods used for the determination of the extremal BIC criterion.

It should be noted that the given test signal can be analyzed also by means of the linear AR form (see Table II). The AR model can be obtained from the NAR model in the case when the basis functions are restricted to the first-order polynomials. However, comparison of the corresponding BIC values (2594 and −16 351 for the AR and NAR forms, respectively) allows us easily to conclude that the linear form of description is in this case unacceptable and has to be abandoned, because the difference between the BIC values is too large. This example indicates the procedure to be followed in the case when *a priori* there are no reasons to assume a polynomial dependence. In such a case one can try to decompose the test signal into various sets of basis vectors and to compare the corresponding BIC values. The smallest BIC value corresponds to the best set of basis vectors. It should be emphasized that one can compare the BIC values for different models, but only for the same time series. Comparison of BIC values of different time series is meaningless.

It was already noted that in the absence of noise the NAR method does not work. Let us test the noise level which is still sufficient for obtaining a reliable estimate of the deterministic component. As the next test signal, we will consider the Henon map with dynamically added noise. We choose the parameter values $a=0.3$ and $b=0.3$ which correspond to relaxation described by the deterministic component, see Fig. 3(a). Figure 4 shows Henon map time series with different dynamical noise levels. The results of the NAR analysis are shown in Table I. As before, the decomposition reproduces the correct form of the equation with the terms present in the Henon map for the two variance values 0.005 and 0.05. However, we have to compare the results obtained by the NAR and AR methods to be sure that the nonlinear form is preferable. Here in one case the variance value is 0.05 and the two BIC values considerably differ. In another case the variance value is 0.005. It is obvious that lowering the noise level we obtain time series which resemble stationary behavior [see Fig. 3(a)]. Since stationarity can be described by a constant (i.e., by the linear AR model), the BIC values for both the NAR and the AR model are of the same order (see the results for the Var=0.005 case in Tables I and II). In this case our method fails to find the nonlinear form. The method would select the linear form, because it is simpler and even the BIC value is slightly larger (within a few percent error of this value). One should note that in the cases presented in Table I the variance estimates coincide with the input values of Var chosen for the time-series generation.

We have demonstrated that the NAR method of decomposition is capable of distinguishing the deterministic and

TABLE II. Calculated AR form for the Henon map model, Eq. (9). Here $b=0.3$.

a	Var	AR form	Err	Var	BIC
1.4	0	$0.8-0.3y_{n-1}-0.4y_{n-3}-0.2y_{n-4}-0.2y_{n-5}-0.1y_{n-6}$	±0.1	0.75±0.04	2594
0.3	0.05	$0.26-0.37y_{n-1}+0.24y_{n-2}$	±0.03	0.067±0.002	159
0.3	0.005	$0.33-0.49y_{n-1}+0.35y_{n-2}$	±0.02	$(5.5\pm 0.2)\times 10^{-3}$	−2335

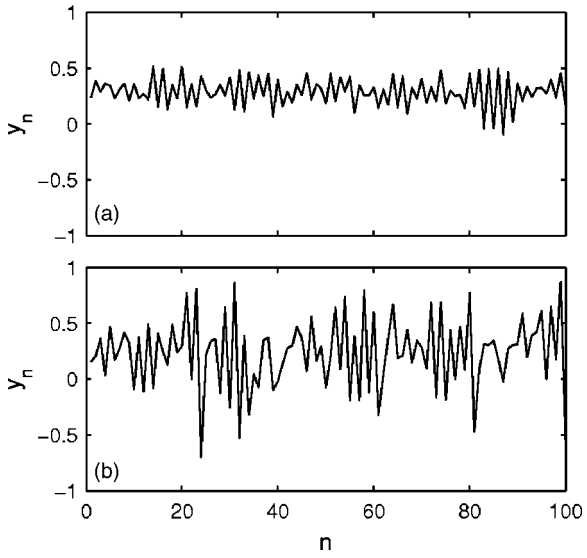

FIG. 4. The Henon map in the presence of noise. Here $a=0.3$ and $b=0.3$; Var=0.005 (a) and Var=0.05 (b).

noise components in the initial time series. It should be emphasized that in the case of a known (or guessed) class of basis functions used for decomposing the time series, the NAR method reproduces the correct type of the initial equation. The method makes it possible to correctly estimate the level of noise present in the series. Hence, a qualitatively new way of behavior classification from time series emerges.

On the basis of the first example shown in (Fig. 4), it could be argued that in the limit of $\varepsilon \to 0$ the deterministic component without noise describes a trivial relaxation. This, however, is true only for particular coefficients in the polynomial form. If we assume that the values of the coefficients of the system can change, then we can extrapolate the obtained analytical results to a different region in the parameter space. Thus, we can claim that the nonlinear equation for the deterministic component can in general describe other types of behavior in addition to relaxation. For example, in the considered case of the Henon map a change of the value of the parameter a leads from relaxation to increasingly complex periodic and finally chaotic types of behavior, see Fig. 3.

B. The detection of fake UPOs in ARMA generated time series

It has been found that the time series generated by the ARMA model can give rise to a false detection of chaotic behavior using the standard UPO method. For given values of coefficients $(\mathrm{AR})_i$ Eq. (6) can be used as a signal generator. Here as a test case we consider the generating equation

$$y_n = -1.1y_{n-1} - 1.0y_{n-2} - 0.8y_{n-3} - 0.5y_{n-4} - 0.2y_{n-5} \tag{10}$$

with Var=0.009 and length of 1000 points.

The time series were generated using the first five, four, and three terms of Eq. (10), respectively. Then they were searched for UPOs using the recurrence method, see Table III. Significantly more UPO candidates were found in ARMA generated time series with both five and three generating terms than in the surrogate sets produced by reshuffling the original data. According to Ref. 10 this indicates that the generated signal is chaotic, which is of course false. This reveals a deficiency in the method used to generate the surrogate sets, and also an easily made misinterpretation of the outcome of the null hypothesis test. Evidently, the surrogate generation by means of shuffling the data only results in a subset of all possible nonchaotic sets that could produce an irregular time series. It does not lead to a conclusive null hypothesis test. Moreover, it seems that the properties of fake UPOs found in nonchaotic noisy time series are system dependent. For this reason caution is needed in interpreting the null hypothesis test using this method. Strictly speaking, such a test only can distinguish the examined dynamical system from the system modeled by the surrogates. The more general the model used to generate the surrogates, the more powerful the test. However it is unlikely that every conceivable irregular nonchaotic system can be represented by a single surrogate set generator. Therefore the null hypothesis test is incapable of proving *conclusively* that a given time series is chaotic, as this would require testing of all possible null hypotheses. In any event it is clear that the results reported in Ref. 10 need to be reexamined in the light of these arguments. The reconsidered UPO method and the new results of the analysis are presented in Ref. 26.

C. Analysis of the ASDEX Upgrade ELM time series

From the ASDEX Upgrade database the shots with type I ELMs were selected for analysis as being the most complete. In these shots time intervals were determined where plasma parameters characterizing the plasma current, heating power, and geometry could be considered as stationary. Since we are interested in statistical properties, the ELMs were chosen for analysis within these time intervals. If the number of the observed ELMs was larger than 100, the corresponding shot was selected for the analysis. The beginning t_{beg} and end t_{end} times of stability intervals as well as mean ELM periods $\langle T_n \rangle$ for these shots can be found in Table IV.

TABLE III. Test of the UPO method using ARMA [see Eq. (10)] generated time series.

Number of AR terms	N	$\langle N_s \rangle$	σ_s	K	Comment
5	24	11.7	3.4	3.6	Chaotic
4	21	11.8	3.3	2.8	Nonchaotic
3	107	12.5	3.4	27.8	Chaotic

TABLE IV. Analysis of ASDEX Upgrade type I ELMs.

Shot	t_{beg} [s]	t_{end} [s]	$\langle T_n \rangle$ [s]	$f(y_R)$	Var
13 513	2.00	4.70	0.0087	$0.004-0.1y_{n-3}+0.3y_{n-8}+0.4y_{n-9}$	1×10^{-6}
13 582	2.80	4.80	0.0130	$0.009+0.3y_{n-1}$	6×10^{-6}
13 881	1.90	3.25	0.0096	$0.006+0.4y_{n-1}$	1×10^{-6}
13 946	4.50	5.70	0.0041	0.0041	2×10^{-7}
14 440	3.11	4.50	0.0096	$0.008+0.2y_{n-1}$	2×10^{-6}
14 441	3.40	4.70	0.0066	0.0066	1×10^{-6}
14 830	5.02	6.07	0.0055	$0.004+0.1y_{n-1}$	1×10^{-6}
14 870	2.80	4.30	0.0040	$0.003+0.2y_{n-1}$	1×10^{-6}
14 908	4.60	5.90	0.0049	$0.003+0.3y_{n-1}$	1×10^{-6}
14 985	2.50	5.00	0.0112	$0.009+0.2y_{n-1}$	1×10^{-5}
15 113	2.70	3.30	0.0059	$0.004+0.3y_{n-1}$	3×10^{-6}
15 145	4.60	6.00	0.0127	0.0127	4×10^{-5}
15 400	2.80	4.00	0.0042	$0.005-0.2y_{n-1}$	1×10^{-7}
15 687	3.32	6.44	0.0132	0.0132	4×10^{-5}
15 733	2.80	4.10	0.0043	0.0043	1×10^{-6}
15 741	2.70	5.50	0.0273	0.0273	1×10^{-4}
15 743	3.00	4.50	0.0072	0.0072	1×10^{-6}
15 840	4.80	5.80	0.0054	0.0054	4×10^{-7}
15 844	4.80	5.80	0.0058	0.0058	1×10^{-6}
15 859	5.00	6.80	0.0024	$0.002+0.1y_{n-1}$	2×10^{-7}
15 873	3.80	5.20	0.0053	$0.004+0.2y_{n-1}$	1×10^{-6}
15 875	4.00	5.50	0.0054	$0.004+0.2y_{n-1}$	1×10^{-6}
15 895	4.00	6.00	0.0070	0.0070	1×10^{-6}
15 905	5.50	7.00	0.0124	$0.015-0.2y_{n-2}$	2×10^{-6}
16 204	4.00	5.00	0.0080	$0.005+0.3y_{n-1}$	1×10^{-5}

As an example, we consider the D_α signal of the shot No. 14 908, Fig. 5. For this shot the plasma parameters can be regarded as constant within the time interval 4.6–5.9 s. The corresponding ELM periods are shown in Fig. 6.

The correlation function (see Fig. 7) can be used for qualitative analysis of time series. Both the autocorrelation and the partial correlation function show correlations with lag one, while for lags larger than unity the correlations are within the confidence limits. In terms of the ARMA model this means that a single AR term or a single MA term can be used to describe the data (see Sec. III B). Since we introduced the Langevine condition, we will use a qualitative model with a single AR term.

The quantitative analysis of periods y_n was performed with the NAR method. It was found that the best form describing the signal corresponds to a linear AR equation

$$y_n = 0.0035 + 0.3y_{n-1} + \varepsilon_n \tag{11}$$

with BIC equal to −2905 and Var$=10^{-6}$, see Table IV. It can be seen that the NAR method leads to the equation with a

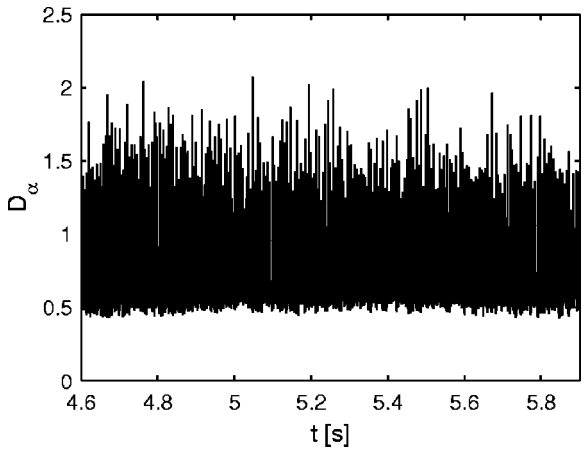

FIG. 5. D_α signal in a stationary plasma parameter region. ASDEX Upgrade shot No. 14 908.

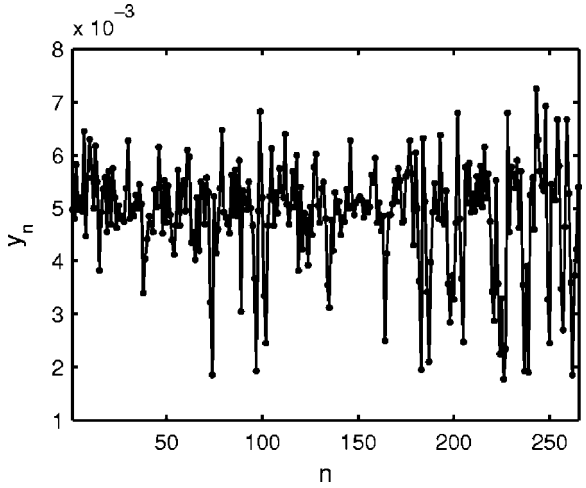

FIG. 6. Periods y_n between ELM events. ASDEX Upgrade shot No. 14 908.

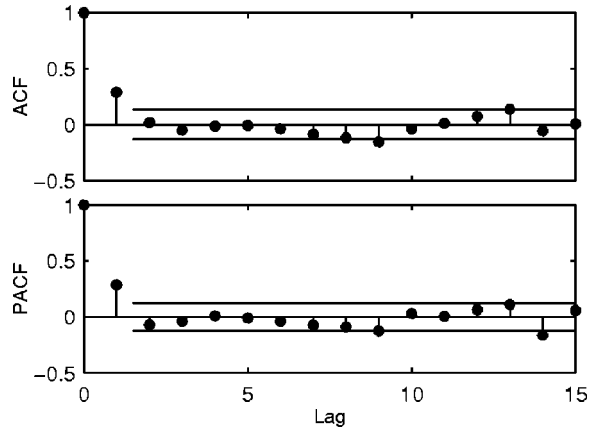

FIG. 7. ACF and PACF analysis of y_n. Confidence levels are marked by horizontal lines. ASDEX Upgrade shot No. 14 908.

single linear term in accordance with the qualitative estimates. The linear form of Eq. (11) is in conflict with the hypothesis obtained by the standard UPO method that the periods of these ELMs (shot No. 14 908) are chaotic.[12] One should note, that the improved UPO method[26] finds no chaos in this shot.

It was shown above that once the form of the equation is known, it is possible to generate new time series with the same statistical properties as the initial time series. We have generated two time series with and without the dynamical noise (innovations ε_n were added in each step n before calculating the next $n+1$ value) (see Fig. 8). The noiseless case relaxes to the unconditional mean value of the stationary y_n time series. By contrast, the case with the dynamically added noise demonstrates the stochastic relaxation to the unconditional mean value.

In the same manner we have analyzed with the NAR method the ELM intervals for the rest of the available shots. The resulting functions $f(y_R)$ exhibit similar behavior as the function corresponding to the shot No. 14 908 (see Table IV). We have detected only linear AR forms in the ASDEX Upgrade ELM time series which indicates the relaxation type

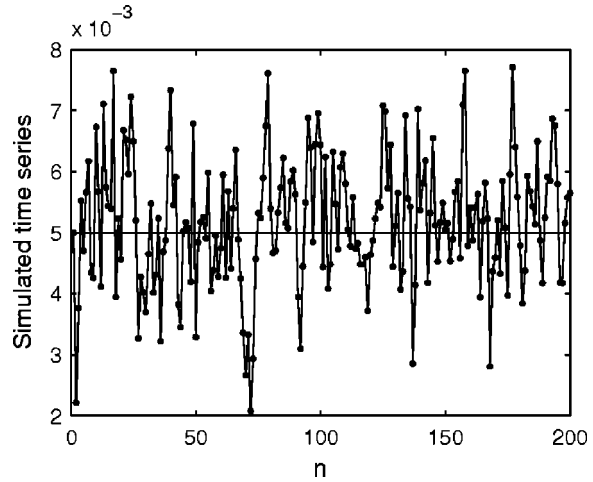

FIG. 8. Time series simulated with Eq. (11): (a) noiseless case, Var=0 (line) and (b) with noise, Var=10^{-6} (line with dots).

of behavior of ELM periods. In the cases when it was sufficient to use a constant form $f=\text{const}$, it was found that this constant coincided with the mean value of ELM periods, as expected. It should be emphasized that the linear form is not sufficient for generating a complex deterministic chaoslike behavior. In turn it implies that the idea of ELM control using standard chaos control techniques is not applicable for the ASDEX Upgrade ELM shots.

On the other hand, we should remember that we have analyzed very short and simplified time series containing only ELM periods. All information contained in the D_α signal is reduced to location of maximums in time. More advanced technique would be to consider both periods and amplitudes of ELMs. This step requires more elaborated methods of analysis treating vector data type instead of scalar. However, the information reduction to, e.g., periods, is well defined for type I ELMs. The type II ELM time series contain no explicit maximums of ELMs and finding the way to analyze such signals remains a challenge for the future.

V. CONCLUSIONS

An approach to analysis of ELM time series has been proposed. It is based on the ARMA model which decomposes time series into deterministic and noise components. Here in contrast to the traditional UPO method it is assumed that a noise is playing an important role. For example, a complicated form of a time series can be a result of an interplay between the relaxation and noise rather than a manifestation of the deterministic chaos. Here the analysis is performed by means of a deterministic equation in the recursive form with added noise. A deterministic solution is searched for in the space of polynomials. Generalizations to other types of basis functions are possible.

The ARMA model and its extension, the nonlinear NAR model, have been validated using artificially generated time series. As a generating function the Henon type equations with and without dynamically added noise have been used. Particularly, the NAR model successfully recovered both the correct Henon map equation and the parameter values of the time series generated in the chaotic regime. Moreover, the NAR method was able to detect the correct deterministic part as well as the noise component used in generating the time series.

A test of the standard UPO method was performed by applying this method to time series generated by means of the linear ARMA model. It was found that the UPO method detected in some of these time-series unstable periodic orbits. This is inconsistency, because the linear recursive equation with noise cannot generate chaotic behavior. The detection of ghost UPOs can be attributed to deficiencies in statistical tests used in the standard UPO method.

The ARMA and NAR analysis of selected ELM time series measured at ASDEX Upgrade fail to show presence of the deterministic chaos in the series. All detected deterministic components can be described by simple linear functions which cannot generate complex chaoslike behavior. These findings imply that the existing ASDEX Upgrade ELMs can-

not be controlled using the standard chaos control techniques.

The traditional UPO methodology recently has been reconsidered,[26] since it detects unstable periodic orbits in time series generated by a linear model. The modified UPO method finds no chaos in the ASDEX Upgrade time series, which now is in agreement with the results of our analysis.

New experimental possibly long and stationary ELM time series at fixed values of some preselected plasma parameters would be needed for further studies of this fascinating subject.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to V. Hynönen for useful discussions and for cross checking some of the results obtained. One of the authors (G.Z.) wishes to thank S. Lu for many stimulating discussions.

APPENDIX: MATHEMATICAL BACKGROUND

Let us consider the NAR Eq. (6)

$$y_n = f(y_R) + \varepsilon_n, \quad n = 1, \ldots, N, \tag{A1}$$

where $f(y_R)$ is a nonlinear function of past events y and ε_n is an innovation value. The unknown function f which describes the given series in the best possible way has to be found. The best description assumes that the entire deterministic information is contained in the f function, while ε values are normally distributed.

In general there are two possibilities for solving this problem. The first is minimization of the mean-squared error between the observed and approximated data,[27] which is equivalent to minimization of the norm of innovations.

The second is maximization of the innovation's likelihood function if the probability distribution function p is known. It can be shown, see Ref. 28, that in the case of the normal distribution p these two approaches are equal. It should be noted, however, that the least-squares method does not require the normal distribution of ε values.

In the present work we use a combination of the two methods. First, a pool of basis functions is formed for the given observation vector y. It consists of both the basis functions y_{n-i} of the linear AR model with the time lag $i = 1 \ldots R$, and the nonlinear basis functions of y_{n-i}. In general, nonlinearity can be introduced in a number of ways,[28] e.g., the monomial/polynomial,[22,27] rational,[29] radial basis,[30] functions and wavelets.[24] We use for simplicity polynomials of y_{n-i}.

Second, the basis functions in the pool are ordered by means of the standard procedure based on the fast orthogonal search (FOS) method.[27] Here the function with the longest observation vector projection is used. Alternatively the approach proposed in Ref. 31 could be used in which the FOS method is extended to include the optimal parameter search algorithm based on the nonorthogonal projection search criterion. After ordering the cutoff number of the basis functions is specified and further only the functions with the largest projection values are considered.

Third, all possible combinations (models) of the resulting basis functions are considered and specific parameter values for each model are found by maximizing the innovation likelihood function.

Let us denote all coefficients prior to basis functions as a parameter vector a. For example, in the case of the linear AR model the vector a consists of $(\mathrm{AR})_i$ coefficients and the constant C. In order to find the values of the parameter vector a of this model, we rewrite Eq. (A1) emphasizing the dependence of innovations on the value of the parameter vector a and on the corresponding choice of the basis functions,

$$\varepsilon_n(a) = y_n - f(y_R|a), \tag{A2}$$

where y_n is the observed value and $f(y_R|a)$ is the value obtained for a particular set of basis functions and the parameter vector a. We assume that there exists some unknown probability distribution function $p(\varepsilon|a, N, \theta)$ which provides observed probabilities of innovations ε_n for the given sample size N and the set of distribution parameters θ. Since innovations by definition contain no deterministic information, it is assumed that they obey the normal distribution

$$p(\varepsilon|a, N, \theta) = \sum_{n=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_n} \exp - \frac{[\varepsilon_n(a) - \mu]^2}{2\sigma_n^2}, \tag{A3}$$

where μ is the mean, σ is the standard deviation of innovations, and $\theta = \{\mu, \sigma\}$ is the distribution parameter. Next for observations y the set of parameters a is found which determines θ and results in innovations obeying the normal distribution. To this end, the likelihood function $L(a|\varepsilon, N, \theta)$ is defined which returns relative likelihoods for different values of parameters a for the given sample size N and innovations ε.

In the case of the normal distribution the likelihood function can be expressed as[32]

$$L(a|\varepsilon, N, \theta) = p(\varepsilon|a, N, \theta). \tag{A4}$$

The best set of parameters a can be obtained by maximizing the likelihood function

$$\hat{a} = \arg\max_a L(a|\varepsilon, N, \theta). \tag{A5}$$

Since the log-likelihood function (LLF) is more convenient in calculations and has the same extremum properties as Eq. (A5), the function

$$\hat{a} = \arg\max_a \ln L(a|\varepsilon, N, \theta) \tag{A6}$$

is used. Here

$$\ln L(a|\varepsilon, N, \theta) = -\frac{N}{2}\ln(2\pi) - \sum_{n=1}^{N} \ln(\sigma_n) - \sum_{n=1}^{N} \frac{[\varepsilon_n(a) - \mu]^2}{2\sigma_n^2}. \tag{A7}$$

In our studies we assume that the variance (the squared standard deviation) of innovations is a time-independent constant

$$\sigma_n^2 = K. \tag{A8}$$

The values of the coefficient $\hat{a}$ for the model (a particular set of the basis functions) are obtained by finding the extremum of the LLF function, Eq. (A6). The negative value of the LLF Eq. (A7) is minimized by means of the method of constrained nonlinear optimization. Finally, different models are compared using the Ockham's Razor criterion. On the one hand, large number of parameters (and of corresponding basis functions) ensures good quality of data description but, on the other hand, it complicates the model excessively. Various methods used in selecting the order of the model are described in Ref. 31, and references therein. Here the balance principle based on the Bayesian information criterium is used (BIC) (Ref. 33)

$$\mathrm{BIC} = -2 \ln L(\hat{a}|\varepsilon, N, \theta) + N_c \ln(N), \tag{A9}$$

where N_c is the total number of parameters in Eqs. (A2) and (A8). The first term in Eq. (A9) is responsible for the quality of data description and the second term plays the role of the penalizing factor reducing complexity of the model.

[1] F. Wagner, G. Becker, K. Behringer *et al.*, Phys. Rev. Lett. **49**, 1408 (1982).
[2] ASDEX team, Nucl. Fusion **29**, 1959 (1989).
[3] H. Zohm, Plasma Phys. Controlled Fusion **38**, 105 (1996).
[4] J. W. Connor, Plasma Phys. Controlled Fusion **40**, 531 (1998).
[5] ITER Physics Basis, Nucl. Fusion **39**, 2137 (1999).
[6] ASDEX Upgrade Team, W. Suttrop, M. Maraschek, G. D. Conway *et al.*, Plasma Phys. Controlled Fusion **45**, 1399 (2003).
[7] J. W. Connor, Plasma Phys. Controlled Fusion **40**, 191 (1998).
[8] P. E. Bak, R. Yoshino, N. Asakura, and T. Nakano, Phys. Rev. Lett. **83**, 1339 (1999).
[9] http://www-jt60.naka.jaeri.go.jp
[10] A. W. Degeling, Y. R. Martin, P. E. Bak, J. B. Lister, and X. Llobet, Plasma Phys. Controlled Fusion **43**, 1671 (2001).
[11] http://crppwww.epfl.ch/crpp_tcv.html
[12] V. Hynönen, Master's thesis, Helsinki University of Technology, 2003.
[13] V. Hynönen, O. Dumbrajs, A. W. Degeling, T. Kurki-Suonio, and H. Urano, Plasma Phys. Controlled Fusion **46**, 1409 (2004).
[14] E. Ott, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **64**, 1196 (1990).
[15] J. F. Lindner and W. L. Ditto, Appl. Mech. Rev. **48**, 795 (1995).
[16] H. Urano, Plasma Phys. Controlled Fusion **45**, 1571 (2003).
[17] P. So, E. Ott, S. J. Schiff, D. T. Kaplan, T. Sauer, and C. Grebogi, Phys. Rev. Lett. **76**, 4705 (1996).
[18] G. Nitsche and U. Dressler, Physica D **58**, 153 (1992).
[19] *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, edited by R. Adler, R. Feldman, and M. S. Taggu (Birkhäuser, Boston, 1998).
[20] G. U. Yule, J. R. Stat. Soc. A **84**, 497 (1921).
[21] G. U. Yule, Philos. Trans. R. Soc. London, Ser. A **226**, 267 (1927).
[22] S. A. Billings, S. Chen, and M. J. Korenberg, Int. J. Control **49**, 2157 (1989).
[23] L. A. Aguirre and S. A. Billings, Int. J. Bifurcation Chaos Appl. Sci. Eng. **5**, 449 (1995).
[24] S. A. Billings and D. Coca, Int. J. Bifurcation Chaos Appl. Sci. Eng. **9**, 1263 (1999).
[25] M. Henon, Commun. Math. Phys. **50**, 69 (1976).
[26] A. W. Degeling, J. B. Lister, Y. R. Martin, and G. Zvejnieks, Plasma Phys. Controlled Fusion **46**, L15 (2004).
[27] M. Korenberg, S. A. Billings, Y. P. Liu, and P. J. McIlroy, Int. J. Control **48**, 193 (1988).
[28] A. Hornstein, Diplomarbei, Drittes Physikalisches Institut der Georg-August-Universität zu Göttingen, 2001.
[29] M. V. Corrêa, L. A. Aguirre, and E. M. A. M. Mendes, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 1019 (2000).
[30] Q. M. Zhu and S. A. Billings, Int. J. Control **64**, 871 (1996).
[31] S. Lu, K. H. Ju, and K. H. Chon, IEEE Trans. Biomed. Eng. **48**, 1116 (2001).
[32] G. D'Agostini, Rep. Prog. Phys. **66**, 1383 (2003).
[33] *Time Series Analysis: Forecasting and Control*, 3rd ed., edited by G. E. P. Box, G. M. Jenkins, and G. C. Reinsel (Prentice Hall, Englewood Cliffs, NJ, 1994).