Modified Bayesian approach for the reconstruction of dynamical systems from time series

D. N. Mukhin,* A. M. Feigin, E. M. Loskutov,† and Ya. I. Molkov
Institute of Applied Physics, Russian Academy of Sciences, Russia

(Received 5 May 2005; revised manuscript received 17 January 2006; published 16 March 2006)

Some recent papers were concerned with applicability of the Bayesian (*statistical*) approach to reconstruction of *dynamic* systems (DS) from experimental data. A significant merit of the approach is its universality. But, being correct in terms of meeting conditions of the underlying theorem, the Bayesian approach to reconstruction of DS is hard to realize in the most interesting case of noisy *chaotic* time series (TS). In this work we consider a modification of the Bayesian approach that can be used for reconstruction of DS from noisy TS. We demonstrate efficiency of the modified approach for solution of two types of problems: (1) finding values of parameters of a known DS by noisy TS; (2) classification of modes of behavior of such a DS by short TS with pronounced noise.

DOI: 10.1103/PhysRevE.73.036211 PACS number(s): 05.45.Tp, 02.70.Rr, 02.30.Zz

I. STATISTICAL APPROACH TO RECONSTRUCTION OF DYNAMIC SYSTEMS

Reconstruction of DS from TS generated by this system is usually understood as seeking its evolution operator. When a DS is known, it is necessary to find values of parameters that determined evolution of the system during TS generation. Such a formulation of the problem arises, for instance, when chaotic regimes of DS behavior are used for solution of the problem of coded transmission of information (see, e.g., Ref. [1]). In situations typical for most natural systems (atmospheric-oceanic, tectonic, biological), the DS that generated the observed TS is *unknown*. In this case, reconstruction of DS means construction on the basis of the information contained in the TS of a parametrized *model* of unknown evolution operator. Apparently, such a model cannot be ideal as, generally, there does not exist such a set of parameter values for any class of models that would make the model absolutely adequate to the modeled DS. Inevitable discrepancy between them is called "defect of the model." In this sense, the two above formulations of the problem of DS retrieval are sometimes classified as the "perfect model class scenario" and the "imperfect model class scenario" [2].

Assume that we have at our disposal the vector TS $\mathbf{x}$, which is formed by M observed quantities $\mathbf{x}=\{\mathbf{x}^{(m)}\}_{m=1}^{M}$ and connected with DS state $\mathbf{u}=\{\mathbf{u}^{(k)}\}_{k=1}^{d}$ via an observer h, $\mathbf{x}=h(\mathbf{u})$. Here d is DS dimensionality, $d \geqslant M \geqslant 1$. Under the perfect model class scenario d value is naturally known to us. When DS is unknown this quantity may be estimated by minimal embedding dimension of the attractor responsible for its observed evolution. Methods of obtaining this information and quite substantial restrictions were discussed by numerous authors and summarized in different review papers and books (see, e.g., Refs. [3,4]). In this work we restrict our consideration to the first situation. The second situation (the "imperfect model class scenario") will be considered elsewhere.

In what will follow we will use the following formulation of the Bayes theorem [5]. Suppose that the system under experiment possesses a set of properties (parameters) $\boldsymbol{\mu}$ that cannot be measured directly and let values of $\mathbf{x}$ be recorded in experiment. Then, *posterior* conditional probability density of *unobservable* parameters (frequently referred to as likelihood) $p(\boldsymbol{\mu}|\mathbf{x})$ is proportional to the product of their *prior* probability density $p(\boldsymbol{\mu})$ and *prior* conditional probability density of the obtained experimental results, $p(\mathbf{x}|\boldsymbol{\mu})$:

$$p(\boldsymbol{\mu}|\mathbf{x}) = C \times p(\boldsymbol{\mu}) \times p(\mathbf{x}|\boldsymbol{\mu}). \tag{1}$$

It will be clear from what will follow that conditional probability density $p(\mathbf{x}|\boldsymbol{\mu})$ depends wholly on the way TS becomes noisy and on probability densities of all noises present in the TS. Factor $p(\boldsymbol{\mu})$ takes into account *a priori* information about the system. If such information is not available, probability density $p(\boldsymbol{\mu})$ must be chosen to be constant, with the width allowing for all possible values of parameters $\boldsymbol{\mu}$. Constant C in (1) is determined by the normalization condition: $C=[\int p(\boldsymbol{\mu})p(\mathbf{x}|\boldsymbol{\mu})d\boldsymbol{\mu}]^{-1}$.

The presence in experimental data of noise component [6] justifies application of the *probability* Bayesian approach to construction of models of *dynamic* systems. Consider as an example a DS with discrete time and assume for definiteness that measurement error ("noise") $\boldsymbol{\xi}$ is additive:

$$\boldsymbol{\xi}_t = \mathbf{x}_t - h(\mathbf{u}_t), \quad \mathbf{u}_{t+1} = \mathbf{f}(\mathbf{u}_t, \boldsymbol{\mu}). \tag{2}$$

Here, the subscript numbers discrete time counts, vector $\mathbf{u}_t = (u_t^{(k)})_{k=1}^{d}$ specifies now "true" (*latent*) state of the DS at the time instant t ($t=0,\ldots,T-1$) in d-dimensional phase space (embedding space), the discrete time map $f(\mathbf{u}_t,\boldsymbol{\mu})$ describes evolution operator of the DS, and $\boldsymbol{\mu}=\{\mu_m\}_{m=1}^{M}$ is the vector of parameters.

As "true" states of the DS are unknown, the probability densities entering (1) depend not only on parameters $\boldsymbol{\mu}$, but also on latent variables $\mathbf{u}=\{\mathbf{u}_t\}_{t=0}^{T-1}: p(\boldsymbol{\mu}|\mathbf{x}) \Rightarrow p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})$; $p(\boldsymbol{\mu}) \Rightarrow p(\mathbf{u},\boldsymbol{\mu}), p(\mathbf{x}|\boldsymbol{\mu}) \Rightarrow p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu})$, with *prior* conditional probability density $p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu})$ determined wholly by the properties of random quantities $\boldsymbol{\xi}$. If they are mutually independent and their probability densities are described by the

*Electronic address: mukhin@appl.sci-nnov.ru
†Electronic address: loskutov@appl.sci-nnov.ru

same function w_ξ, then an expression for $p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu})$ has the following form:

$$p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu}) = \prod_{t=0}^{T-1} w_\xi\{\mathbf{x}_t - h[f^t(\mathbf{u}_0,\boldsymbol{\mu})]\}. \quad (3)$$

Here, $\mathbf{u_0}$ is the value of latent variable at the initial moment of time, $f^t(\cdot)$ is t-multiple (successively reiterated t times) discrete time map, and $f^0(\mathbf{u_0},\boldsymbol{\mu}) := \mathbf{u_0}$. Note that $\mathbf{u_0}$ is the only latent variable when TS is generated by DS and the noise is observational.

The relations (1) and (3) solve (in probabilistic formulation) the problem of seeking values of DS parameters under which the TS was generated in the experiment. Besides, they allow noise filtering, i.e., finding the most probable noise-free values of the measured dynamic variable.

Unfortunately, application of the Bayesian approach to DS reconstruction from noisy experimental data faces, as was discussed in Refs. [2,5], fundamental difficulties. An increase in the length of the TS that is desirable for reducing effective level of measurement noise makes the use of $p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})$ impracticable for calculation of needed probability characteristics. Note that even the fastest numerical algorithms based on the *Markov Chain Monte Carlo* method [7] require great computer power even for simple, low-dimensional DSs. It is obvious that expression (3) for the chaotic TS at sufficiently large T will be too complicated for both using the Monte Carlo method and finding the most probable values of parameters and initial conditions. This is attributed to the fact that, because of the fractal nature of the strange attractor, an increase in T leads to extremely complicated shape of the regions of values of the initial conditions and of model parameters that ensure existence of the model phase trajectory in the noise-specified neighborhood of the trajectory reconstructed by the initial noisy TS. Accordingly, the likelihood (3) as a function of its arguments takes on a multimodal ("jagged") form.

Inapplicability of the "classical" Bayesian approach to DS reconstruction from chaotic TS explicated above was demonstrated in Ref. [2] on an example of a logistic map, the evolution operator of which is a first-order discrete time map $u_{n+1} = 1 - au_n^2$. The system may demonstrate both, regular (periodic) and chaotic behavior, depending on the value of the only parameter a. The transition to dynamical chaos at varying a occurs through a cascade of period doublings. Reconstruction of the value of parameter a from chaotic TS with additive measurement noise ξ,

$$x_n = u_n + \xi_n,$$

$$u_n = 1 - au_{n-1}^2 \quad (4)$$

was considered in Ref. [2]. It was assumed that the noise is δ-correlated and has Gaussian distribution with known standard deviation σ_ξ:

$$w_\xi(\boldsymbol{\xi}) \propto \exp\left(-\sum_i \xi_i^2 \Big/ 2\sigma_\xi^2\right).$$

It was shown that, in the classical Bayesian formulation of the problem (3), the characteristic scale of irregularity of conditional probability density $p(u_0,a|\mathbf{x})$ along the latent variable u_0 becomes less than computer accurate even at moderate noise $\sigma_\xi = 0.1$ [17] and fairly short TS ($T=70$). For finding a correct value of parameter, the root-mean-square scatter σ_{u_0} characterizing the width of *prior* distribution $p(u_0,a)$ must be $\sigma_{u_0} \leqslant 10^{-17}$. In other words, the classical Bayesian approach demands unattainably exact information about the initial state of DS.

Meyer and Christensen [8] proposed to modify the classical Bayesian formulation so as to overcome the above problem (i.e., incompatibility of the statistical approach with dynamical nature of the studied system). Meyer and Christensen criticized the work by McSharry and Smith [9], in which the choice of the cost function was not justified, and proposed to assume within the framework of the Bayesian approach that small dynamic noise is present in the system. Then the second equation of system (2) takes on the form $\mathbf{u}_{t+1} = f(\mathbf{u}_t,\boldsymbol{\mu}) + \boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t$ is a Gaussian random quantity. Formally, as was noted in Ref. [2], such a formulation is incorrect: system (2) that is *known to be dynamic* is replaced by a *stochastic* system. However, such an assumption is quite correct in terms of the Bayesian approach and just means weakened *a priori* requirements to the model: deterministic relationship (2) of latent variables is replaced *a priori* by a "less strict," probabilistic relationship. In the work [8] it was demonstrated that the resulting probability density (that now includes, besides u_0, other latent variables of the system) allows statistical analysis of *posterior* distribution, for example, by the MCMC method using noninformative *prior* distribution for nonobservables.

The modification of the Bayesian approach proposed in this paper has a different underlying idea that is, in a certain sense, opposite to that in Ref. [8]. We suggest using as completely as possible *a priori* information about dynamic origin of the system. In Sec. III it will be shown that such a modification enables one to construct *posterior* distribution much more informative than that proposed in Ref. [8].

In Sec. II we elucidate the idea of the modification and present an expression for a modified *posterior* probability density $p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})$. Further, for comparison with the results obtained by Meyer and Christensen, the modified approach is used for seeking values of the parameters of a known DS by the noisy TS generated by this system. Application of the modified approach for classification of DS types of behavior by very short and noisy TS is described in Sec. IV. In the Conclusion, we summarize the considered issues and discuss problems that must be solved for effective application of the Bayesian approach in more complicated situations, in particular, for making prognosis of qualitative behavior of *unknown* DSs from noisy chaotic TS.

II. MODIFICATION OF THE BAYESIAN APPROACH: PIECEWISE-DYNAMIC RECONSTRUCTION

We propose a modification of the Bayesian approach that is based on *a priori* ideas about the properties of chaotic processes generated by dynamic systems. Suppose that the reconstructed system is known *a priori* to be dynamic, i.e., its latent states u_t at different moments of time are related by a certain evolution operator dependent on a set of parameters $\boldsymbol{\mu}:\mathbf{u}_{t+1}=f(\mathbf{u}_t,\boldsymbol{\mu})$. The joint probability density of $\mathbf{u}$ and $\boldsymbol{\mu}$, describing this relationship has the form

$$p(\mathbf{u},\boldsymbol{\mu}) \propto \prod_{j=1}^{T-1} \delta(\mathbf{u}_j - f(\mathbf{u}_{j-1},\boldsymbol{\mu})), \quad (5)$$

where $\delta(\cdot)$ is the delta function. If the dynamic system that generated the TS functions in the *chaotic* mode, then the information coupling between the TS counts is known to decrease with increasing time interval between the counts. In other words, the system starts to "forget" its initial state with time. Hence, assuming the states of the system separated by large time intervals to be independent we can regard the latent variables to be coupled only at finite time periods (segments) of length w. In this case, the function (5) is factorized as follows:

$$p(\mathbf{u},\boldsymbol{\mu}) \propto \prod_{k=0}^{Q}\prod_{j=1}^{w} \delta(\mathbf{u}_{k(w+1)+j} - f(\mathbf{u}_{k(w+1)+j-1},\boldsymbol{\mu})),$$

where $Q=T/(w+1)-1$. Further, the Bayes theorem gives *posterior* joint conditional probability density for $\mathbf{u}$ and $\boldsymbol{\mu}:p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})\propto p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu})p(\mathbf{u},\boldsymbol{\mu})$, where $p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu})$ is found from the distribution of measured noise $w_\xi(\boldsymbol{\xi})$ known *a priori*. In accordance with (2),

$$p(\mathbf{x}|\mathbf{u},\boldsymbol{\mu}) = p(\mathbf{x}|\mathbf{u}) = \prod_{l=0}^{T-1} w_\xi[\mathbf{x}_l - h(\mathbf{u}_l)].$$

Finally, the obtained *posterior* probability density

$$p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x}) \propto \prod_{l=0}^{T-1} w_\xi[\mathbf{x}_l - h(\mathbf{u}_l)] \times \prod_{k=0}^{Q}\prod_{j=1}^{w} \delta(\mathbf{u}_{k(w+1)+j} - f(\mathbf{u}_{k(w+1)+j-1},\boldsymbol{\mu})),$$

may be integrated with respect to latent variables $\mathbf{u}_{k(w+1)+j}$, $j=1,\ldots,w$; $k=0,\ldots,Q$,

$$p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x}) \propto \prod_{k=0}^{Q}\prod_{j=0}^{w} w_\xi\{\mathbf{x}_{k(w+1)+j} - h[f^j(\mathbf{u}_{k(w+1)},\boldsymbol{\mu})]\}. \quad (6)$$

Apparently, expression (6) is meaningful for $w\in[1,T/2]$; it is impossible to divide the TS on segments of equal length for $w\in(T/2,T]$ [18].

Note that when the observed series is generated by a system functioning in a regular (nonchaotic) mode, the assumptions described above that underlie the factorization function (5) become, generally speaking, incorrect. Therefore, reconstruction of a dynamic system using the proposed modification will be less accurate than in the case of chaotic time series (see Sec. IV for more detail).

The transition from (3)–(6) increases the number of latent variables from one to $[T/(w+1)]$, but a reasonable choice of segment length w eliminates extensive irregularity of the distribution $p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})$, when $T\to\infty$. Clearly, in the case of chaotic TS, the "utmost-reasonable" segment length can be estimated by the inverse value of the largest Lyapunov exponent calculated from the reconstructed phase trajectory. It is also readily understood that extensive decreasing of w will reduce accuracy of finding values of latent variables and, hence, of parameters of the model. It is worth mentioning that, for Gaussian distribution of measurement noise $w_\xi=N(0,\varepsilon^2)$, the *cost function* (CF) following from (6) corresponds exactly to the CF corresponding to the algorithm of "multiple shooting" [10]. Note that in the recent work [11] it was proposed to seek an optimal set of parameters $\boldsymbol{\mu}$ for each segment *independently*. So, the corresponding CF [for $w_\xi=N(0,\varepsilon^2)$] is obtained from (6) on the substitution $\boldsymbol{\mu}\to\boldsymbol{\mu}_k$. Then, the set $\{\boldsymbol{\mu}_k\}$ is interpreted as a statistical ensemble for further estimations. Such an approach is very simple to realize but it is hardly applicable for not very long data series.

A quite apparent drawback of expression (6) is its nonsymmetry with respect to latent variables whose noisy values form the TS. The value of the probability density (6) depends on $[T/(w+1)]$ (of T) latent variables only. Before we start symmetrization we want to make two remarks. First, the set of the latent variables $\{\mathbf{u}_{l(w+1)}\}_{l=0}^{Q}$ used in (6) is specified unambiguously by the choice of the (noisy) state of the DS, $\mathbf{x}_0$, as the initial one. Second, in the case of steady and rather long TS, the probability density (6) must not depend on which of the noisy states x_t $(t=0,\ldots,w)$ is initial. These considerations allow us to write down *posterior* probability density as a geometric mean of w expressions (6) that differ by the choice of the initial state. For segments having length $w\in[1,T/2]$, we obtain, to an accuracy of normalization, the following expression:

$$p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x}) \propto \left(\prod_{k=0}^{T-w-1}\prod_{j=0}^{w} w_\xi\{\mathbf{x}_{k+j} - h[f^j(\mathbf{u}_k,\boldsymbol{\mu})]\}\right)^{1/(w+1)}. \quad (7)$$

Note that, unlike (6), the symmetrized expression for the modified probability density may be written for segments of length $w\in(T/2,T]$ too. For this we need to change the exponent $(w+1)^{-1}\to(T-w-1)^{-1}$ in (7).

Expression (7) for *posterior* probability density is the key expression in the statistical approach to reconstruction of DS by noisy TS. *Posterior* probability density $p(\boldsymbol{\mu}|\mathbf{x})$ is expressed through (7) in a natural fashion as

$$p(\boldsymbol{\mu}|\mathbf{x}) = \int p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})d\mathbf{u}. \quad (8)$$

The *posterior* probability density (7) allows for dynamic features of the reconstructed system to the extent maximum possible within the framework of the Bayesian approach. A measure of reconstruction of dynamic features (as well as of filtering measurement noise) is segment length w: the greater w, the less (7) differs from the probability density (3) that formally includes all information about the DS contained in the initial TS. In the case of *the perfect* model class scenario, maximum possible value of w for reconstruction of DS from a specific TS is determined by the level of noise and, in addition, by the used approach to investigation and further application of conditional probability density of model parameters. Obviously, it is reasonable to extend the length of the segment unless accuracy of reconstruction ceases to grow

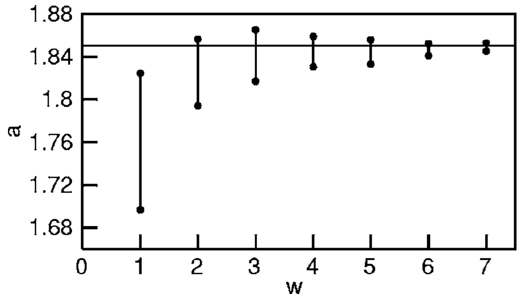

FIG. 1. Reconstruction of parameter a of logistic map: confidence intervals of reconstruction as a function of w (the horizontal line is for the "correct" value of a). The time series with the noise level of 0.4 is used for reconstruction.

(i.e., as long as informativity of *posterior* distribution is increasing). Examples of the dependence of w on the enumerated factors will be given in Secs. III and IV.

Note that the modified Bayesian approach is adapted to DS reconstruction at the cost of increased number of latent variables: they are $(T-w)>T/2$ in expression (7). Thus, when performing reconstruction by sufficiently extended TS containing *measurement* noise one can encounter the problem of an extremely large number of arguments for *posterior* probability density $p(\mathbf{u},\boldsymbol{\mu}|\mathbf{x})$. Solution of this problem needs special investigation. In the two sections to follow we will demonstrate efficiency of the modified Bayesian approach on examples of solution of problems in which this difficulty does not arise.

III. RECONSTRUCTION OF PARAMETER VALUES OF A KNOWN DS BY NOISY TS

The example from Ref. [2] given at the end of Sec. I clearly shows the inapplicability of the classical Bayesian approach, if initial conditions are not known precisely. We will show now that the modified approach allows estimation of parameter, even if there is no *a priori* information about the value of latent variable. We remind the reader that for solution of this problem Meyer and Christensen proposed a different modification of the Bayesian approach, namely, they suggested replacing the dynamic system by the stochastic one. For comparison with results of reconstruction obtained by Meyer and Christensen we took as observational data time series generated by a *logistic map* with the value of parameter a the same as in Ref. [6] ($a=1.85$) and the initial condition $x_0=0.3$. The time series is corrupted by white Gaussian noise with the $\sigma_{\text{noise}}/\sigma_{\text{signal}}$ ratio different for each series ranging from 0 to 2. An ensemble of values of parameter a, by which a 95% confidence interval was calculated, was generated by the MCMC method for each time series in accord with the *posterior* distribution (7) for different segment lengths w within the 1-to-7 interval.

The confidence interval of parameter a reconstructed by TS with the noise level of 0.4 is shown in Fig. 1 for different segment lengths $w=1,\ldots,7$. Clearly, in accord with the qualitative considerations given in Sec. I, an increase in w results in a decrease of both systematic bias relative to the correct parameter value and the distribution width determining error. Consequently, for the segment length $w=2$, the correct value of the parameter lies within the 95% confidence interval and at $w=7$ the bounds of the confidence interval differ from the correct value by less than 1%.

Variation of the confidence interval of parameter a with increasing noise level is shown in Fig. 2 for different segment lengths. Clearly, the use of the modified posterior probability density (7) with $w>3$ gives a better accuracy of reconstruction than the Meyer and Christensen approach (cf. an analogous dependence of the size of confidence interval on noise level in Ref. [8]). Thus, the modification of the Bayesian approach proposed in this paper seems to be more efficient.

To conclude this section we note that, for moderately noisy TS, it can be shown that the width σ_a of the distribution (8) is related to the largest Lyapunov exponent λ and segment length w by

$$\sigma_a \propto e^{-\lambda w}. \tag{9}$$

Figure 3 demonstrates that the theoretical dependence (9) σ_a vs w (the straight line in Fig. 3; the value of λ was calculated by the *logistic map* TS using the TISEAN software [12]) is in good agreement with results of reconstruction (the dots in Fig. 3).

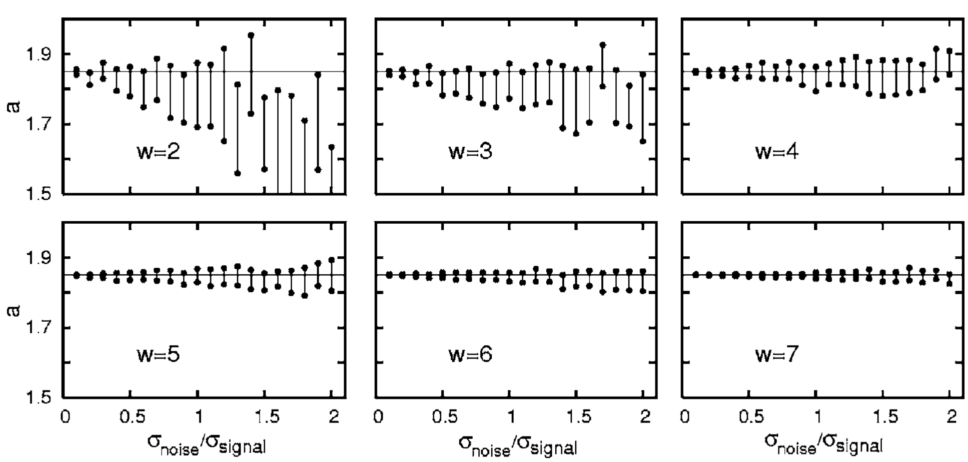

FIG. 2. Reconstruction of parameter a of *logistic map*: 95% confidence interval of parameter a as a function of noise level $\sigma_{\text{noise}}/\sigma_{\text{signal}}$ for different segment lengths w (the horizontal line is for the "correct" value of a).

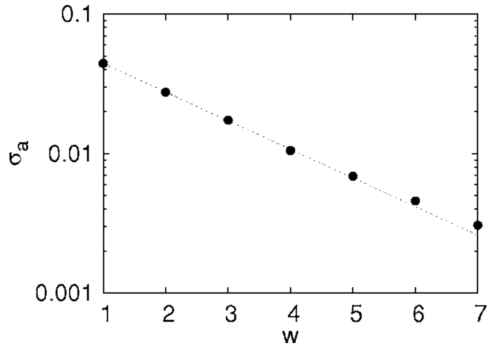

FIG. 3. Reconstruction of parameter a of logistic map: σ_a versus w: the dots show results of reconstruction ($\sigma_{\text{noise}}/\sigma_{\text{signal}}=0.2$), and the line is the theoretical dependence (9).

IV. CLASSIFICATION OF TYPES OF BEHAVIOR OF KNOWN DS BY SHORT NOISY TS

The results obtained in the preceding section enable us to formulate the problem of classification of types of DS behavior that is important from the practical viewpoint. Consider by way of example very short ($T=20$) TS generated by a more complex DS, namely, Henon map. Evolution operator of this DS is a second-order sequence function

$$u_{n+1} = v_n,$$

$$v_{n+1} = 1 - av_n^2 - bu_n. \tag{10}$$

System (10) may demonstrate diverse regular and chaotic modes of behavior, depending on values of parameters a and b. Regions of parameters corresponding to different modes of behavior are shown by graded gray colors in the bifurcation diagram (Fig. 4).

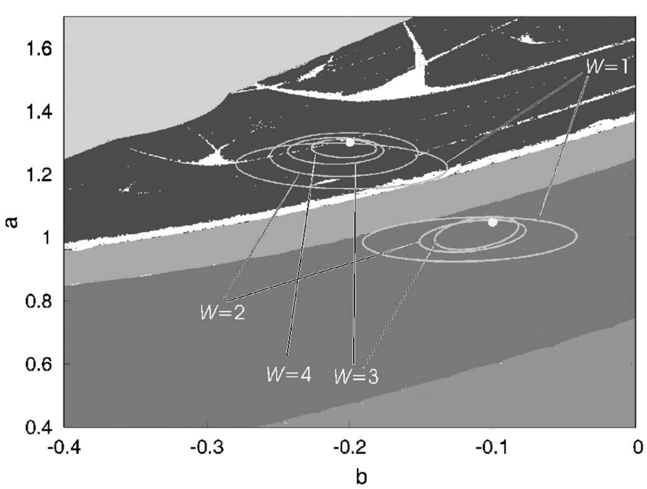

FIG. 4. Bifurcation diagram of Henon map. The graded gray colors show modes corresponding (from bottom to top) to stable equilibrium state, to two-period, four-period, and chaotic modes, and to the region of global instability of the model. The gray dots mark values of parameters at which the TS were generated. The ellipses are the boundaries of confidence intervals with 95% probability of parameter reconstruction for different w; $\sigma_\xi=0.06$.

We took x as a "measurable" quantity, that is latent variable u with additive noise: $x_n=u_n+\xi_n$. Measurement noise ξ was δ-correlated and had Gaussian distribution with dispersion σ_ξ^2. We analyzed the TS generated for different values of parameters a and b. In the absence of noise, the TS corresponding to different periodic and chaotic modes are readily distinguishable. By way of example we present in Figs. 5(a) and 5(b) the TS generated by the map (10) which correspond to two-period ($a=1.05$, $b=-0.1$) and chaotic ($a=1.3$, $b=-0.2$) modes. Addition of even relatively small noise [see Figs. 5(c) and 5(d), $\sigma_\xi=0.3$] makes the dynamic modes indistinguishable. We will demonstrate how this problem may be solved by means of the modified Bayesian approach.

The problem of classification of the mode of behavior of the DS generating the initial *noisy* TS is based on investigation of a statistical ensemble of *noiseless* TS corresponding to a statistical ensemble of *parameters* distributed according to the *posterior* probability density (8) [the ensemble is formed by means of the MCMC algorithm applied to the *posterior* distribution (7)]. First, the number of elements of the ensemble of noiseless TS corresponding to one or another dynamic regime is counted, and then probability of each recorded mode is calculated. Probability of the "correct" mode of behavior as a function of noise level is plotted in Fig. 6 for the chaotic and one of periodic noisy TS generated by the map (10).

Note that closeness to unity of the probability of the most probable type of behavior is a convenient quantitative characteristic of the quality of classification. Let us now consider dependence of this characteristic on segment length w. From the above mentioned it is clear that, in the case of the chaotic behavior, this probability will generally approach unity with increasing w. Probabilities of erroneous types will decrease almost always due to decreasing width of *posterior* parameter distribution density with increasing w.

Let us now try to understand whether this conclusion is true for the w dependence of the probability of "correct" regular mode. As was mentioned in Sec. II, factorization of the probability density (5) implying a decrease in information coupling between the TS counts with increasing time interval between them becomes incorrect for the system functioning in a regular (nonchaotic) mode. As a consequence of this incorrectness the dependence of parameter distribution width on w is somewhat different compared with the case of the chaotic mode. Namely, for regular regimes this dependence does not decay exponentially as in the case of chaotic mode [see (9)]. This is explained by the fact that there is no exponentially fast spreading of initially close phase trajectories on the attractor corresponding to a regular mode of a dynamic system; hence, there is no exponential growth of sensitivity of the current state of the system to initial conditions with increasing observation time. Nevertheless, reduction of the width of $p(\boldsymbol{\mu}|\mathbf{x})$ with increasing w ensures an increase of the probability of correct mode with increasing segment length for regular modes of behavior too. The above mentioned is confirmed by results of reconstruction of system (10) presented in Figs. 4 and 6. Figure 4 shows confidence regions of system parameters reconstruction by the corresponding noisy TS. One can see that for $w<4$ the quality of classification improves with increasing

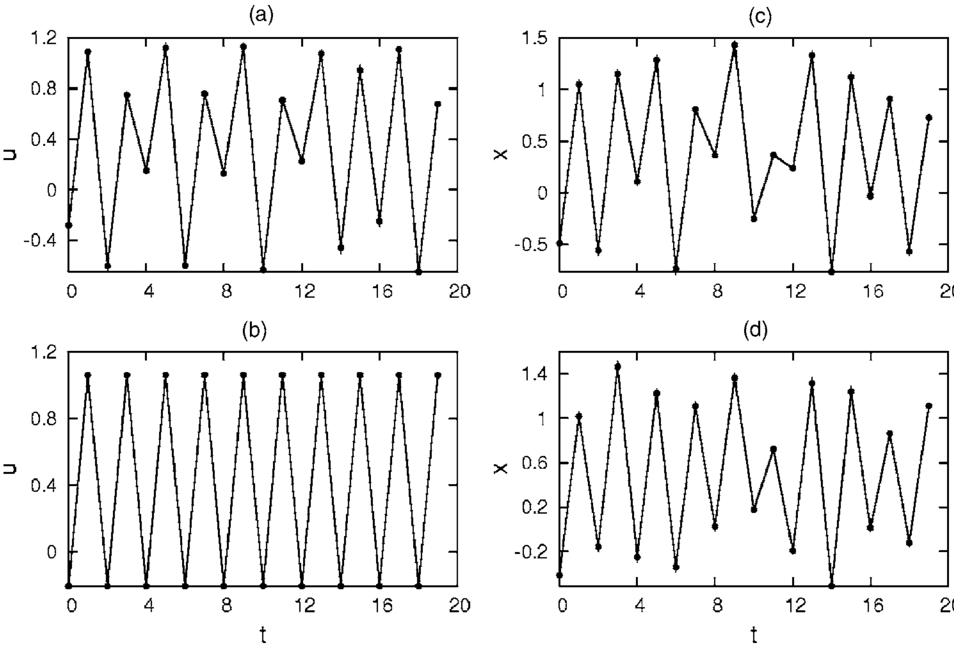

FIG. 5. (a) and (b) Chaotic and periodic TS corresponding to the values of parameters marked by dots in Fig. 2; (c) and (d) the same TS but with noise σ_ξ=0.3.

segment length for both chaotic and regular modes of behavior. The dependence, decaying exponentially with increasing w, of the value of the confidence region of the chaotic regime on segment length is plotted in Fig. 7 together with the same but nonmonotonic dependence for the regular, two-period regime.

V. CONCLUSION

We considered the statistical Bayesian approach to reconstruction of dynamic systems by time series generated by these very systems. We believe that this is the most adequate approach to reconstruction of "real" dynamic systems which, first, are subject to irregular external forcing in the course of signal generation and, second, the signal itself is randomly perturbed during propagation and recording. The statistical approach aims not only at reconstructing DS evolution operator but also at constructing probability density of parameters of this operator that are regarded to be random quantities. One of the main difficulties in realizing this approach arises during reconstruction by chaotic time series. The exponentially fast diverging of phase trajectories on a chaotic attractor results in an increase of irregularity of probability density of parameters as the duration of the TS grows. Hence, it is impossible to construct parameter probability density even for relatively short TS. We proposed a modification of the classical Bayesian approach that makes it possible to overcome this difficulty. The efficiency of the modified approach was demonstrated on an example of seeking unknown values of parameters of a known DS and classification of DS modes of behavior, both regular and chaotic, by short noisy TS.

The significant "technical" limitation for applying the modified approach is the absence of a "fast" algorithm of constructing an arbitrary distribution function (probability density) that depends on a very large number of arguments. For noisy TS, this number depends, primarily, on the TS length (the number of latent variables). In Sec. II we showed that, within the framework of the modified Bayesian approach, information about dynamic features of the system can be retained to a considerable extent by using almost maximum possible number of latent variables for the considered TS. Thus, the proposed solution to the first problem makes development of an effective algorithm for calculation of multidimensional distribution functions still more essential. We hope to complete this task in the near future.

To conclude we would like to emphasize universality of the proposed approach to reconstruction of DS. It may be used for any problems of information retrieval from TS generated by dynamic systems. A challenging problem of this kind is prognosis of qualitative behavior of *unknown* DS by chaotic TS. General formulation of this problem ("prognosis of bifurcations") was given in Ref. [5] where this problem was solved successfully by means of "transforming" the dynamic system to a formally stochastic one: the defect of the

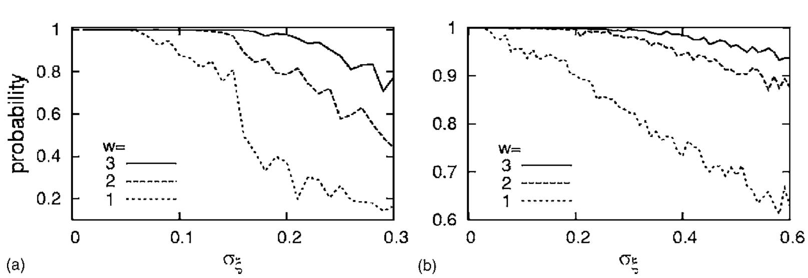

FIG. 6. Probability of recognition of "correct" dynamic mode as a function of noise level for different w: (a) chaotic TS; (b) periodic TS.

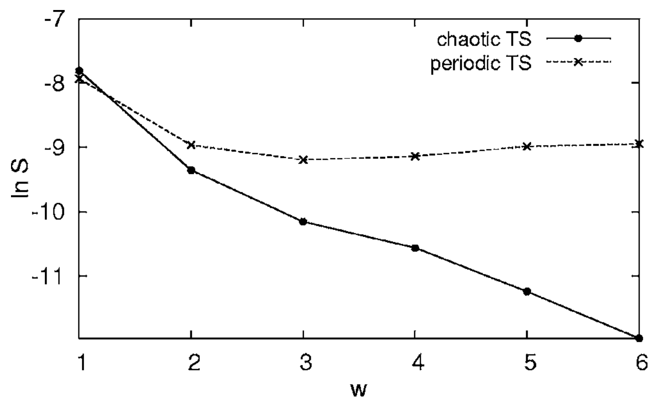

FIG. 7. Area of 95% confidence interval *versus* segment length for chaotic (the line with dots) and periodic (the line with crosses) TS.

model, inevitable by virtue of unknown DS, was described in Ref. [5] as additive dynamic "noise." The algorithm of bifurcation prognosis used in Ref. [5] is a particular case of the proposed modified Bayesian approach corresponding to the maximum short segment $w=1$. The results discussed in Secs. III and IV give us grounds to believe that the quality of bifurcation prognosis may be improved by increasing w up to a maximum possible value. Note that the example of constructing bifurcation prognosis given in Ref. [5], similarly to all the other examples available in the literature [13–16], refers to the case of "ideal," free of measurement noise TS generated by a low-dimensional DS. No advance has been made in this problem so far because of the absence of effective algorithm for constructing multidimensional probability density arbitrarily depending on its arguments.

ACKNOWLEDGMENTS

The work was done under support of the Russian Foundation for Basic Research (Contract No. 02-02-17080) and the Complex Program of Research "Fundamental Problems of Nonlinear Dynamics" under the umbrella of the Presidium of the Russian Academy of Sciences (project 3.3).

[1] V. S. Anishchenko and A. N. Pavlov, Phys. Rev. E **57**, 2455 (1998).
[2] K. Judd, Phys. Rev. E **67**, 026212 (2003).
[3] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data* (Springer-Verlag, New York, 1997).
[4] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, 1997).
[5] Y. I. Molkov and A. M. Feigin, *Nonlinear Waves'2002* (Institute of Applied Physics, Nizhny Novgorod, 2002), pp. 34–52 (in Russian).
[6] B. P. Carlin and T. A. Louis, *Bayes and Emperical Bayes Methods for Data Analysis* (Chapman and Hall, London, 1996).
[7] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice* (Chapman and Hall, London, 1996).
[8] R. Meyer and N. Christensen, Phys. Rev. E **62**, 3535 (2000).
[9] P. E. McSharry and L. A. Smith, Phys. Rev. Lett. **83**, 4285 (1999).
[10] H. Bock, *Recent Advances in Parameter Identification Techniques for O.D.E., Numerical Treatment of Inverse Problems in Differential and Integral Equations* (Birkhauser, Basel, 1983).
[11] V. F. Pisarenko and D. Sornette, Phys. Rev. E **69**, 036122 (2004).
[12] R. Hegger, H. Kantz, and T. Schreiber, Chaos **9**, 413 (1999).
[13] A. M. Feigin, Y. I. Molkov, D. N. Mukhin, and E. M. Loskutov, Report No. 508, Institute of Applied Physics, 1999 (in Russian).
[14] A. M. Feigin, Y. I. Molkov, D. N. Mukhin, and E. M. Loskutov, Radiophys. Quantum Electron. **44**, 348 (2001).
[15] A. M. Feigin, Y. I. Molkov, D. N. Mukhin, and E. M. Loskutov, Faraday Discuss. **120**, 105 (2002).
[16] A. M. Feigin, *Proceedings of the International Symposium "Topical Problems of Nonlinear Wave Physics"* (Nizhny Novgorod, Russia, 2003).
[17] For the TS considered in Ref. [2] [the TS generated by system (4) at $a=1.85$; $u_0=0.3$], $\sigma_\xi=0.1$ corresponds approximately to the 6:1 signal-to-noise ratio.
[18] Note that, formally, for $w=T$ expression (5) transforms to the "classical" expression (3).