Markov chain Monte Carlo method in Bayesian reconstruction of dynamical systems from noisy chaotic time series

E. M. Loskutov,* Ya. I. Molkov, D. N. Mukhin,† and A. M. Feigin
Institute of Applied Physics, Russian Academy of Sciences, 46, Ulianov Street, Nizhny Novgorod 603950, Russia
(Received 30 March 2007; revised manuscript received 14 November 2007; published 23 June 2008)

The impossibility to use the MCMC (Markov chain Monte Carlo) methods for long noisy chaotic time series (TS) (due to high computational complexity) is a serious limitation for reconstruction of dynamical systems (DSs). In particular, it does not allow one to use the universal Bayesian approach for reconstruction of a DS in the most interesting case of the unknown evolution operator of the system. We propose a technique that makes it possible to use the MCMC methods for Bayesian reconstruction of a DS from noisy chaotic TS of arbitrary length.

DOI: 10.1103/PhysRevE.77.066214 PACS number(s): 05.45.-a, 05.10.-a, 95.75.Wx

I. INTRODUCTION

The presence of a noise component in the time series generated by a dynamical system (DS) results in finite accuracy of reconstruction of the evolution operator (EO): Any model that can be constructed will, generally speaking, differ from the original system. One of the ways to take this fact into consideration is to describe parameters of the model EO as random quantities [1]. Two limiting cases may be distinguished [2]: The “perfect model” scenario when the form of the operator describing the DS evolution is reliably known, and the “imperfect model” scenario when EO is unknown.

The perfect model scenario is realized, for instance, when the problem of hidden information transmission is solved using chaotic time series [3]. In this case we are interested in the characteristics of the system (“perfect model”) parameters (the most probable value, mean, dispersion, distribution). In investigations of many “natural” (atmospheric-oceanic, tectonic, biological) systems the EO is unknown, which corresponds to the imperfect model scenario. Then, information about model parameters is of no value as the physical meaning of the parameters is unknown, and we are interested in the properties of the model defined by these parameters and reflecting the properties of the reconstructed DS. Nevertheless, in this case too evaluation of statistical characteristics of model parameters is the key element without which the problem of reconstructing a DS generating the original TS cannot be solved.

The mathematical body used to reconstruct the DS is determined by a specific application. For example, for finding the most probable set of model (perfect or imperfect) parameters it suffices to determine the maximum of posterior probability density of model parameters [4–6]. However, it should be remembered that in the presence of noise the inverse problem of reconstruction becomes ill-posed, i.e., it admits an infinite set of solutions. Selection of the “most correct” solutions demands regularization that is a physically justified constraint on admissible values of parameters based solely on a priori information [7]. This approach is usually called Bayesian. Within the Bayesian approach, there are methods which make it possible to estimate dispersions and mathematical expectations of the model parameters [8]. A significant advantage of such methods is their relatively low computational resource requirements. However, due to the model nonlinearity, the form of parameter distribution may differ strongly from the normal one. In this case, estimation of expectation and dispersion is insufficient to construct a correct model, and one must use the Bayesian approach in full. In other words, we must construct models which include distributions of parameters and take a priori information about the system into account correctly. In this formulation the problem can be solved using the Markov chain Monte Carlo (MCMC) algorithms [9]. However, in the case of reconstruction of DSs by noisy chaotic time series, computational resources required for these algorithms strongly depend on distribution dimensionality. This limits applicability of the algorithms to a narrow class of problems, e.g., the case of short TS [10]. The method proposed in this paper makes it possible to expand the applicability domain of the MCMC algorithms to the case of arbitrarily long TS.

II. FORMULATION OF THE PROBLEM

In what follows, we will assume that the available time series can be used to state the fact of dynamism of the system that has generated it, to determine minimum embedding dimension \( d \) [11], and to reconstruct the phase trajectory [12] \( \{x_i\}_{i=0}^T, \ x_i \in R^d \), and \( t \) enumerates moments of the discrete time. Let us assume that the system experimented upon has a set of properties (parameters) \( \mu \) which cannot be measured directly. Let us consider, for the sake of certainty, a DS with discrete time

\[
x_i = U_i + \xi_i, \quad U_i = f(U_{i-1}, \mu) + \eta_i.
\] (1)

Here, the vector \( x = \{x_i\}_{i=1}^d \), as mentioned above, is an observable quantity; \( U = \{U_i\}_{i=1}^d \) is a latent variable characterizing the true state of the dynamical system in the \( d \)-dimensional phase space; \( f(U, \mu) \) is the vector function describing the DS evolution operator; \( \mu \) are unobserved parameters; and \( \xi_i \) and \( \eta_i \) are random quantities (“noises”) with distributions \( w_\xi \) and \( w_\eta \). Then, in accordance with the Bayesian theorem [13], the