PHYSICAL REVIEW E **71**, 015203(R) (2005)

Additive nonparametric reconstruction of dynamical systems from time series

Markus Abel,* Karsten Ahnert, Jürgen Kurths, and Simon Mandelj
Institute of Physics, University of Potsdam, 14415 Potsdam, Germany
(Received 22 July 2004; published 12 January 2005)

We present a nonparametric way to retrieve an additive system of differential equations in embedding space from a single time series. These equations can be treated with dynamical systems theory and allow for long-term predictions. We apply our method to a modified chaotic Chua oscillator in order to demonstrate its potential.

DOI: 10.1103/PhysRevE.71.015203 PACS number(s): 05.45.Tp, 47.52.+j

Casting physical observations into mathematical equations is one of the fundamental tasks to understand and predict dynamical systems. Basically, there are two complementary approaches to accomplish this task: theoretically, by convenient considerations, and empirically, by data analysis. Both approaches are essential for modern modeling strategies. If, for many systems, the dynamics is not directly accessible to theoretical considerations, then an appropriate data analysis is essential. This problem is very general; one can find it in classical fields of physics, e.g., classical mechanics, fluid dynamics, solid-state physics, statistical physics, as well as in more interdisciplinary fields, e.g., physiology, earth sciences, economics, or biological systems. In this paper the data analysis issue is addressed: we determine an analytically treatable set of additive equations in embedding space by the method of nonparametric embedding. This approach is *a priori* parameter-free; but *subsequent* parametrization can be helpful for analytical representation of the involved functions.

Often, the measurement of a complex system does not yield the whole set of state variables. The missing dynamics can be accessed by the *embedding* technique [1]. Given the measurement of a subset of variables, one can infer the missing information by an embedding map, e.g., by using the time-delayed variables or their derivatives. This has been proven rigorously for a wide class of systems [2]. It is, however, not known how the equations of the dynamical system in embedding space are structured. In this communication, we propose a technique to find a set of equations which allows a reproduction of the dynamics in phase space for the class of *additive* systems.

There are several excellent reviews about embedding [2–4]; therefore, we only repeat some basic facts. We consider a system governed by a set of ordinary differential equations,

$$\dot{\vec{x}} = F(\vec{x}), \tag{1}$$

where $\vec{x} \in \mathbb{R}^n$, $F:\mathbb{R}^n \mapsto \mathbb{R}^n$. This set of equations defines a flow, F_t, in phase space. We assume that there exists an attractor $\mathcal{A} \subset \mathbb{R}^n$ with the box-counting dimension $d \leqslant n$. In [2] it has been shown that almost every smooth map $\Psi:\mathbb{R}^n \mapsto \mathbb{R}^m$, $m > 2d$, is an embedding, i.e., a smooth diffeomorphism from $\mathcal{A}$ onto its image $\Psi(\mathcal{A})$. The condition $m > 2d$ is sufficient, therefore cases with $d < m < 2d$ can occur.

Due to differentiability, the dynamics of $\vec{\xi}(t) = \Psi(\vec{x}(t))$ obeys an ordinary differential equation in embedding space,

$$\dot{\vec{\xi}} = \Phi(\vec{\xi}), \tag{2}$$

with $\vec{\xi} \in \mathbb{R}^m$, $\Phi:\mathbb{R}^m \mapsto \mathbb{R}^m$. In this paper, we focus on *additive* models for the components Φ_i and show how to retrieve them from data.

One standard way of embedding is the use of the delay-coordinate map $H(f,\tau)$, with the smooth observation function, $f:\mathbb{R}^n \mapsto \mathbb{R}$, and τ, the time delay, some real number [2],

$$H(f,\tau) = (f, f[F_{-\tau}], \ldots, f[F_{-(m-1)\tau}]). \tag{3}$$

As an example, consider the particular case of $f(\vec{x}) = x_1$. Identifying the above embedding map Ψ with H, the coordinates in embedding space are $\xi_1(t) = f(\vec{x}) = x_1(t), \xi_2(t) = f(F_{-\tau}(x)) = x_1(t-\tau)$, etc.

In our analysis, we perform numerical simulations for some model systems to obtain the time series of various variables. We then discard all but one variable to embed the dynamical system (2) using the delay map. To avoid confusion, we will refer to dynamics from Eq. (1) as *original*. For the counterpart, Eq. (2), to be estimated by nonparametric regression, we will use the term *reconstructed*. If the embedding map Ψ is concerned, *embedded* will be used—the latter meaning that a time series from the original system is used, i.e., without knowing the dynamical system (2).

To find a dynamical system in embedding space, several approaches exist, e.g., local linear fits and parametric procedures as polynomial fits, radial basis functions, or neural networks (cf. [3]). Local fitting is a general concept, but the results are neither easy to access analytically nor to visualize due to the high dimensionality. Polynomial ansatzes tend to involve too many terms for a clear identification of a mathematical or physical structure; for neural networks a physical interpretation is very hard.

Now we describe our procedure in more detail: Considering each temporal measurement as a realization of the flow, one obtains as a best estimator of the components of Eq. (2) in the least-square sense [3]:

*Electronic address: markus@stat.physik.uni-potsdam.de

$$\Phi_i = E[\dot{\xi}_i | \xi_1, \ldots, \xi_m], \quad (4)$$

with $E[\cdot|\cdot]$ the conditional expectation value (CEV) operator. It is a very hard task to extract analytical models from Eq. (4); visualization is obviously impossible for $m > 2$.

To tackle this problem, we require the right-hand side (rhs) of Eq. (2) to be an additive model,

$$\Phi_i = \sum_{j=1}^{m} \phi_{ij}(\xi_j). \quad (5)$$

This is a subset of the class of models considered by Kolmogorov [5]: he showed rigorously that it is possible to represent any continuous function of a set of m variables as a $2m+1$-fold superposition of m functions of one argument. Below, we show that despite the less general formulation it is possible to reconstruct a chaotic dynamical system. Our model (5) is, however, still in a wider model class than in parametric methods, because we do not rely on a given set of basis functions. After having finally estimated the components ϕ_{ij}, we can easily visualize the functions and try analytical formulas.

It is worth noting the geometrical aspect of our approach: Equation (2) defines a differentiable manifold approximated by the sum of the functions ϕ_{ij}, cf. Eq. (5). This is possible within a certain scatter, which is quantified below by the correlation. If the manifold is found exactly by the model, the correlation is 1. Dynamical and topological properties of the original system are mirrored in embedding space. Long-term predictions of the dynamics are thus possible on the basis of the obtained model if the correlation is close to 1, which is a very strong advantage.

The optimal estimate for the ϕ_{ij} is calculated by the backfitting algorithm [6]. It works by alternately applying the CEV operator to projections of Φ_i on the coordinates: $\phi_{ij}(\xi_j) = E[\dot{\xi}_i - \sum_{k\neq j}\phi_{ik}|\xi_j]$, and is proven to converge to the global optimum in the least-square [Eq. (6)] or correlation [Eq. (7)] sense. For the application to spatiotemporal data analysis, see [7,8]. We calculate the CEV by smoothing splines, which are optimal for nonparametric regression [6], due to their smoothness and differentiability properties. It is important to note that the parameters used by splines or other estimators are method inherent and not prescribed by a preselected model; in this sense the model is parameter-free.

As an overall quality measure, the least-square error can be used,

$$\chi_i^2 = E\left[\left(\Phi_i - \sum_{k=1}^{m} \phi_{ik}\right)^2\right]. \quad (6)$$

The backfitting method, however, is formulated as optimal in the sense of correlation, i.e., the natural measure is the correlation coefficient C_{i0} between the rhs and lhs in Eq. (5). The correlation coefficient C_{ij}, given in Eq. (7), indicates its individual weight for the model,

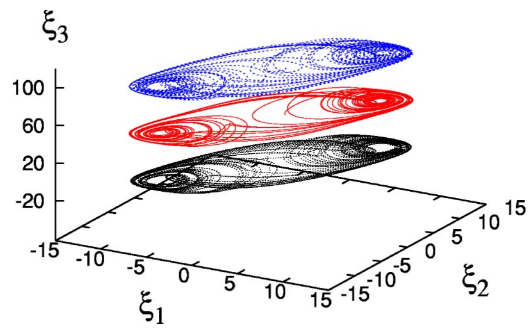

FIG. 1. Embedding of the first component of the system (8) with delay $\tau = 0.2$. The system, embedded with $m=3$ (bottom), is shown together with the reconstructed trajectory from the integration of the systems obtained by nonparametric regression for embedding dimensions 3 (middle) and 4 (top), respectively. An offset is added to avoid overlap of the attractors.

$$C_{i0} = C\left[\Phi_i; \sum_{k=1}^{m} \phi_{ik}\right], \quad C_{ij} = C\left[\phi_{ij}; \Phi_i - \sum_{k\neq j}^{m} \phi_{ik}\right]. \quad (7)$$

We will use C_{i0} as a quantitative measure, but not χ^2. Putting $\Phi_i = Y$, $\Sigma_{k=1}^{m}\phi_{ik} = X$, we give the relation between both measures: $2\times C_{i0}\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)} = \mathrm{Var}(X) + \mathrm{Var}(Y) + [E(X-Y)]^2 - \chi_i^2$, with $\mathrm{Var}(\cdot)$ the variance. A correlation close to 1 means the manifold described by Eq. (2) is approximated very well, lower correlations indicate scatter of data points around the manifold. In the case of experimental data, measurement noise can produce some additional scatter.

In the following, the procedure is illustrated by the example of a modified Chua circuit [9] with a third-order nonlinearity. The basic equations read

$$\dot{x}_1 = a(m_0 x_1 - 1/3 m_1 x_1^3 + x_2),$$

$$\dot{x}_2 = x_1 - x_2 + x_3, \quad \dot{x}_3 = -b x_2. \quad (8)$$

Written as an additive model (5) these equations read $\dot{x}_1 = f_{1,1}(x_1) + f_{1,2}(x_2)$, $\dot{x}_2 = f_{2,1}(x_1) + f_{2,2}(x_2) + f_{2,3}(x_3)$, $\dot{x}_3 = f_{3,2}(x_2)$, with the linear functions $f_{1,2}$, $f_{2,i}$, $f_{3,2}$, and $f_{1,1}$ a third-order polynomial.

We integrate the system (8) with $a=18$, $b=33$, $m_0=0.2$, $m_1=0.01$, numerically, by a Runge-Kutta algorithm of fourth order. The time series of the first component is used for embedding. Results do not change for other components. We first discuss embedding dimension $m=3$. From the time series $x(t)$, the points $\dot{\xi}_i = \dot{x}(t-\tau(i-1))$, $\xi_1 = x(t)$, $\xi_2 = x(t-\tau)$, $\xi_3 = x(t-2\tau)$ are used. The derivative is taken directly from the integration; this is more exact than the estimate by finite differences. The nonparametric regression yields the functions ϕ_{ij} for the resulting dynamical system (5).

First, we present results for the specific delay, $\tau = 0.2$; we study below the dependence of the results on the delay time. For $\tau = 0.2$, the embedded and the reconstructed attractor are shown together in Fig. 1. With respect to the data analysis, we want to quantify (i) the quality of the regression, (ii) the importance of the functions ϕ_{ij}, and (iii) the functions themselves.

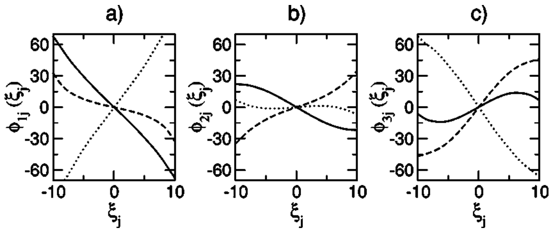

FIG. 2. Reconstructed functions ϕ_{ij}, $i=1,2,3$ in embedding space. (a) ϕ_{1j}, (b) ϕ_{2j}, (c) ϕ_{3j}, for $j=1$ (solid line), $j=2$ (dotted line), and $j=3$ (dashed line). All functions are important with $C_{11}=0.999$, $C_{12}=0.99$, $C_{13}=0.998$, $C_{21}=0.999$, $C_{22}=0.991$, $C_{23}=0.999$, $C_{31}=0.997$, $C_{32}=0.999$, $C_{33}=0.999$.

(i) *The quality of the regression* is given by the correlation C_{i0}, cf. (7). We find in our case $C_{10}=0.992$, $C_{20}=0.999$, $C_{30}=0.995$, such that the modeling error is very small.

(ii) *The importance of functions* is found by the coefficients C_{ij}, defined in Eq. (7) $(i,j=1,2,3)$. We find $C_{ij} > 0.99 \forall i,j$; consequently every function is substantial here (cf. Fig. 2). Given that we analyzed 50 000 data points, the C_{ij} refer to a very high correlation. Therefore we infer a property of the embedding transformation: each of the embedding space coordinates ξ_i contains information necessary for the dynamics.

(iii) *The nine functions* ϕ_{ij}, displayed in Fig. 2, are the most important result for an application. All functions are important and nonlinear, to a good approximation of cubic order; only ϕ_{13} appears to be a piecewise linear function. The quantitative comparison of the dynamics of the reconstructed and the original system is done by (i) calculation of the fixed points, (ii) their stability, and (iii) the Lyapunov exponents (LE's) of the reconstructed system. These quantities have to coincide with the ones of the original system.

(i) *Fixed points*. We solved $\sum_{j=1}^{3}\phi_{ij}=0$ $(i=1,2,3)$ numerically with the functions from the output of the analysis. The three fixed points of the embedded system are $(-7.75, -7.75, -7.75)$, $(0, 0, 0)$, $(7.75, 7.75, 7.75)$ with an accuracy of 10^{-3}. In the system, reconstructed with $\tau=0.2$, the fixed points are $\vec{\xi}_1^*=(-7.76,-7.55,-7.66)$, $\vec{\xi}_2^*=(7.75,7.52,7.68)$, and $\vec{\xi}_3^*=(0,0,0)$ with an error of less than 1%.

(ii) *Stability analysis*. The eigenvalues, corresponding to the above fixed points, are $\vec{\gamma}_1=(-7.68,0.47+i4.45,0.47-i4.45)$, $\vec{\gamma}_2=(-7.62,0.58+i4.55,0.59-i4.55)$, $\vec{\gamma}_3=(5.09,-1.16+i4.56,-1.16-i4.56)$ to be compared with the ones of the *original* Chua system: $\vec{\gamma}_{o,1}=(-8.76,0.28+i5.20,0.28-i5.20)$, $\vec{\gamma}_{o,2}=(-8.76,0.28+i5.20,0.28-i5.20)$, $\vec{\gamma}_{o,3}=(5.03,-1.21+i4.71,-1.21-i4.71)$. For the embedded attractor, there is nothing to calculate due to missing equations. Furthermore, the embedding conserves dynamical properties. The contraction rate from $\vec{\gamma}_{1,2}$ is found within 15%, the expansion rate from $\vec{\gamma}_3$ is found within 1%, the imaginary parts coincide within 15%.

(iii) *Lyapunov exponents and dependence on the delay*. We calculated the Lyapunov exponents of the reconstructed system for $0<\tau\leq 1$. For most of the delays no useful reconstruction is possible, however in the window $0.14<\tau<0.28$ the LE's are very close to the original ones [Fig. 3(a)]. By eye, it is hard to recognize which attractor is reconstructed or embedded (Fig. 1). With this study, we have determined the delay which is optimal in the sense of nonparametric embedding. Usually, the delay is chosen such that the information content in the delay coordinate vector is maximized. To do so one determines the minimum of the mutual information or the first zero of autocorrelation, or similar measures [3]. It turns out that these approaches do not yield a delay different from ours.

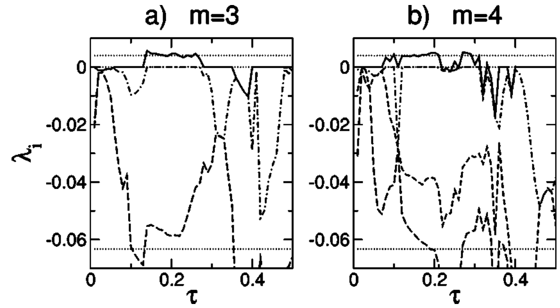

FIG. 3. Lyapunov exponents for original and reconstructed system for embedding dimension $m=3$ (a) and $m=4$ (b). Increasing m results in a larger window in the delay time for which the system is reconstructed, i.e., the LE's coincide well. The thin dotted lines indicate the LE's for the original system, $\lambda_1=0.432$, $\lambda_2=0$, $\lambda_3=-6.31$; the straight, dash-dotted, and dashed lines the correspondent ones for the reconstructed system.

If the embedding dimension is increased, one expects a good reconstruction in a larger delay-time window, because more information is used. This is confirmed by the calculation of the LE's with $m=4$ [Fig. 3(b)], where a good reconstruction is found for $0.08<\tau<0.36$. The attractor for $\tau=0.2$ is shown for comparison (Fig. 1, top).

At $m=4$, there is a breakdown of the reconstruction for $0.22<\tau<0.26$, whereas $m=3$ yields good results (Fig. 3). This is unexpected and a conclusive explanation requires further investigation.

A particularity of the modified Chua system is its additive structure. Next, we check whether a successful reconstruction can be found for dynamical systems with multiplicative terms, too, such as the Lorenz or the Rössler system. For both, we find a worse capability of our method to reconstruct the dynamics. For the Lorenz system $(\sigma=10,\rho=28,\beta=8/3)$, one of the best results appears for $\tau=0.09$; the corresponding reconstructed attractor is shown in Fig. 4. Clearly, some of the dynamics is lost, nevertheless a chaotic motion about the correct fixed points is found. The largest LE is found, $\lambda_{rec}=0.08$, to be compared with the original one, $\lambda_o=0.905$. Tests with $m=4$ and $m=5$ did not yield significant improvement. The reason lies probably in the topology of the attractor which cannot be produced by a purely additive model of reasonably low dimensionality. This is a limit of the additive modeling approach.

We have reconstructed a dynamical system by a set of ordinary differential equations in embedding space. We have considered additive models only, and have used as a typical chaotic system a modified Chua oscillator for illustration. The resulting equations can be analyzed by dynamical systems theory: we have investigated the fixed-point structure,

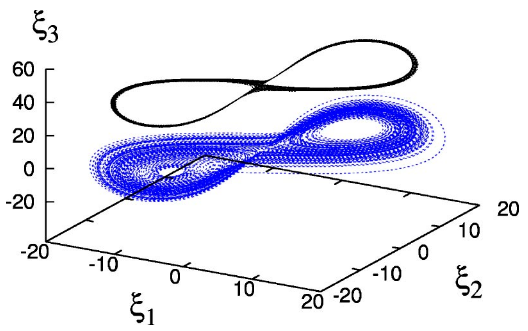

FIG. 4. Reconstruction for the Lorenz system ($m=3$). A part of the dynamics is not reconstructed, but may be recovered with higher dimensional embedding.

linear stability, and Lyapunov exponents, and have found that these dynamical characteristics quantitatively coincide with the ones of the original system. By studying the dependence of the results on the delay time, we could identify the window in which our method works very well. Higher embedding dimensions enlarge this window; overdetermination can, however, let the description break down. For nonadditive systems, our analysis works qualitatively. A quantitative comparison is in general not possible, although the result indicates which terms can be important in a more general model.

In the method, the statistical backfitting algorithm is used for an estimation of the CEV; the result is a set of optimal functions ϕ_{ij}. It is inherently insensitive against noise [6,8] and can be generalized in many ways. The results are functions of one variable and can be visualized and approximated by analytical formulas *after* the backfitting procedure. This yields an important advantage: when fitting polynomials or other basis systems one chooses these functions beforehand—this is not needed in the nonparametric approach—and the result is still interpretable. From a practical point of view, the input data are crucial for a good estimation on a connected region and an estimation of derivatives. Asymptotics and gaps due to missing data have to be treated with great care, or instead of derivatives one might rather use a mapping approach [10]. Decision on additivity of a model works by statistical measures (correlations or least-squares error), whereas dynamical measures, as Lyapunov exponents, indicate how well the dynamics is reproduced.

From a theoretical point of view, we formulate the following general question: Given a dynamical system (additive/multiplicative structure), which topology of a corresponding attractor is possible? Vice versa, given a topology and dynamics, which is the structure of the underlying dynamical system? We have treated a given topology (of the Chua, Lorenz, and Rössler system); taking into account the embedding theorem, the problem is transferred to embedding space. There, we have reconstructed a dynamical system of additive structure for the Chua system, less convincingly for the Lorenz and Rössler system. This suggests that the additive structure is kept. Mathematically, related questions have been treated in [5]. A key role is played by the nonlinear embedding transformation which can distort the system considerably. The above questions are open and touch the core of modern theories of dynamical systems.

Current and future activities focus on generalization to reconstruct mixed additive/multiplicative models following Kolmogorovs ideas, especially for real data. One goal is to follow the way from the general model (2) to a purely additive model (5). Finding the model which involves the least possible multiplicative and additive terms yields a considerable ease to analyze the systems. With an analytical expression a detailed analysis and long-term prediction of a (chaotic) orbit is possible; this is an unprecedented result. Applications for our method reach from geophysics and climatology, to biology and medicine, where the prediction of, e.g., climate change or illness detection are topics of great interest.

We thank B. Fiedler, P. Grassberger, M. Holschneider, and A. Pikovsky for helpful discussions. M. A. and K. A. are supported by the German Science Foundation, DFG; J. K and S. M. acknowledge a EU grant (HPRN-CT-2000-00158).

[1] F. Takens, *Detecting Strange Attractors in Turbulence*, Lecture Notes in Math Vol. 898 (Springer, Berlin, 1981).

[2] T. Sauer, J. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).

[3] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 1997).

[4] H. D. I. Abarbanel, M. E. Gilpin, and M. Rotenberg, *Analysis Of Observed Chaotic Data* (Springer, New York, 1997).

[5] A. Kolmogorov, Am. Math. Soc. Transl. **28**, 55 (1963); A. G. Vitushkin, Enseign. Math. **23**, 255 (1977).

[6] T. Hastie and R. Tibshirani, *Generalized Additive Models* (Chapman and Hall, London, 1990).

[7] H. Voss, M. J. Bünner, and M. Abel, Phys. Rev. E **57**, 2820 (1998); H. U. Voss, P. Kolodner, M. Abel, and J. Kurths, Phys. Rev. Lett. **83**, 3422 (1999).

[8] M. Abel, Int. J. Bifurcation Chaos Appl. Sci. Eng. **14**, 2027 (2004); H. U. Voss, J. Timmer, and J. Kurths, *ibid.* **14**, 1905 (2004).

[9] R. N. Madan, *Chua's Circuit: A Paradigm for Chaos* (World Scientific, London, 1993).

[10] A. J. Lichtenberg and M. A. Lieberman, *Regular and Chaotic Dynamics* (Springer, New York, 1992).