Chaotic Time Series Analysis: Identification and Quantification Using Information-Theoretic Functionals

Milan Palus

SFI WORKING PAPER: 1993-03-017

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
We present a technique for analyzing experimental time series that provides detection of nonlinearity in data dynamics and identification and quantification of underlying chaotic dynamics. It is based on evaluation of redundancies, information-theoretic functionals which have a special form for linear processes, and on the other hand, when estimated from chaotic data, they have specific properties reflecting positive information production rate. This rate is measured by metric (Kolmogorov-Sinai) entropy that can be estimated directly from the redundancies.

1. INTRODUCTION

The inverse problem for dynamical systems, i.e., determination of the underlying dynamical process based on processing experimental dynamical data (Abraham et al., 1989; Mayer-Kress, 1986), can be considered as a recent alternative to generally used stochastic methods in the area of time series analysis. Algorithms have
been developed which can in principle serve for identification and quantification of underlying chaotic dynamics (Cohen & Procaccia, 1985; Dvořák & Klaschka, 1990; Grassberger & Procaccia, 1983a, Grassberger & Procaccia, 1983c; Wolf et al., 1985). Analyzing experimental and usually short and noisy data, however, ordinary estimators of dimensions or Lyapunov exponents can be fooled, e.g., by autocorrelation of the series under study and can consider as chaotic the process which is in fact linear and stochastic (Osborne & Provenzale, 1989). These complications evoked the necessity of developing methods testing the basic properties of chaotic systems, i.e., nonlinearity, independently of the dimensional or Lyapunov exponents algorithms.

In this paper, we propose an original method suitable for two-step assessment of the character of a time series under study. The first step is testing for nonlinearity in the data dynamics in a general sense, and the second step is identification of chaotic dynamics with the possibility of its subsequent quantification. The whole method is based on evaluation of so-called redundancies, information-theoretic functionals which have a special form for linear processes, and thus a comparison of their linear and general versions can demonstrate the linear or nonlinear nature of the data. On the other hand, when redundancies are estimated from chaotic data, they have specific properties reflecting a positive information production rate. This rate is measured by metric (Kolmogorov-Sinai) entropy that can be estimated directly from the redundancies. In this part of the method, i.e., in estimations of the metric entropy, we follow the original work of Fraser (1989a).

Section 2 introduces the information-theoretic functionals—mutual information and redundancies. The redundancy-based test for nonlinearity is explained and illustrated by numerical examples in Section 3. Its comparison with other methods, testing for nonlinearity or determinism, is given in Section 4. The relation between redundancies computed from time series and metric entropy of underlying dynamical system, giving the possibility of identification and quantification of chaotic dynamics, is explained in Section 5. Section 6 presents the application of the presented methodology on competition data. The conclusion is given in Section 7. In Appendix 1, remarks concerning the algorithm for estimating redundancies are presented. Details on numerically generated data, used in testing the properties of the presented method, can be found in Appendix 2.

2. MUTUAL INFORMATION AND REDUNDANCIES

In this section we will define basic functionals introduced in information theory. More details can be found in any book on information theory (Billingsley, 1965; Gallager, 1968; Khinchin, 1957; Kullback, 1959; Shannon & Weaver, 1964).
Let $x, y$ be random variables with probability distribution densities $p_x(x)$ and $p_y(y)$. The entropy of the distribution of a single variable, say $x$, is defined as:

$$H(x) = -\int p_x(x) \log(p_x(x)) \, dx.$$  

(1)

For the joint distribution $p_{x,y}(x,y)$ of $x$ and $y$, the joint entropy is defined as:

$$H(x,y) = -\int \int p_{x,y}(x,y) \log(p_{x,y}(x,y)) \, dx \, dy.$$  

(2)

The conditional entropy $H(x|y)$ of $x$ given $y$ is defined as:

$$H(x|y) = -\int \int p_{x,y}(x,y) \log \left( \frac{p_{x,y}(x,y)}{p_y(y)} \right) \, dx \, dy.$$  

(3)

The entropy of the distribution of the discrete random variable $z_i$ with probability distribution $p_z(z_i)$, $i = 1, \ldots, k$, is defined as:

$$H(z) = -\sum_{i=1}^{k} p_z(z_i) \log(p_z(z_i)).$$  

(4)

Definitions (2) and (3) for discrete variables can be derived straightforwardly. The average amount of information about the variable $y$ that the variable $x$ contains is quantified by the mutual information $I(x;y)$:

$$I(x;y) = H(x) + H(y) - H(x,y).$$  

(5)

Clearly, $I(x;y) = 0$ iff $p_{x,y}(x,y) = p_x(x)p_y(y)$, i.e., iff $x$ and $y$ are statistically independent.

Generalization of the definition of the joint entropy for $n$ variables $x_1, \ldots, x_n$ is straightforward: The joint entropy of distribution $p(x_1, \ldots, x_n)$ of $n$ variables $x_1, \ldots, x_n$ is

$$H(x_1, \ldots, x_n) =$$

$$-\int \ldots \int p(x_1, \ldots, x_n) \log(p(x_1, \ldots, x_n)) \, dx_1 \ldots dx_n.$$  

(6)

Generalization of the mutual information for $n$ variables can be constructed in two ways: In analogy with Eq. (5) we define

$$R(x_1; \ldots; x_n) = H(x_1) + \ldots + H(x_n) - H(x_1, \ldots, x_n).$$  

(7)
This difference between the sum of the individual entropies and the entropy of the $n$-tuple $x_1, \ldots, x_n$ vanishes iff there is no dependence among these variables. Quantity (7), which quantifies the average amount of common information contained in the variables $x_1, \ldots, x_n$, is called the redundancy of $x_1, \ldots, x_n$.

Besides Eq. (7) we define the marginal redundancy $g(x_1, \ldots, x_{n-1}; x_n)$ quantifying the average amount of information about the variable $x_n$ contained in the variables $x_1, \ldots, x_{n-1}$:

$$g(x_1, \ldots, x_{n-1}; x_n) = H(x_1, \ldots, x_{n-1}) + H(x_n) - H(x_1, \ldots, x_n).$$

Clearly, the marginal redundancy $g(x_1, \ldots, x_{n-1}; x_n)$ vanishes iff $x_n$ is independent of all $x_1, \ldots, x_{n-1}$.

The following relations between redundancies and entropies can be obtained by a simple manipulation:

$$g(x_1, \ldots, x_{n-1}; x_n) = R(x_1; \ldots; x_n) - R(x_1; \ldots; x_{n-1})$$

(9)

and

$$g(x_1, \ldots, x_{n-1}; x_n) = H(x_n) - H(x_n | x_1, \ldots, x_{n-1}).$$

(10)

Now, let $x_1, \ldots, x_n$ be an $n$-dimensional normally distributed random variable with zero mean and covariance matrix $C$. In this special case, redundancy (7) can be computed straightforwardly from the definition:

$$R(x_1; \ldots; x_n) = \frac{1}{2} \sum_{i=1}^{n} \log(c_{ii}) - \frac{1}{2} \sum_{i=1}^{n} \log(\sigma_i).$$

(11)

where $c_{ii}$ are diagonal elements (variances) and $\sigma_i$ are eigenvalues of the $n \times n$ covariance matrix $C$. (See, e.g., the work of Morgera (1985), two-dimensional case, i.e., mutual information, was derived also by Fraser (1999b).)

Formula (11) obviously may be associated with any positive definite covariance matrix. Thus we use formula (11) to define the linear redundancy $L(x_1; \ldots; x_n)$ of an arbitrary $n$-dimensional random variable $x_1, \ldots, x_n$, whose mutual linear dependencies are described by the corresponding covariance matrix $C$:

$$L(x_1; \ldots; x_n) = \frac{1}{2} \sum_{i=1}^{n} \log(c_{ii}) - \frac{1}{2} \sum_{i=1}^{n} \log(\sigma_i).$$

(12)

If formula (12) is evaluated using the correlation matrix instead of the covariance matrix, then particularly $c_{ii} = 1$ for every $i$, and we obtain

$$L(x_1; \ldots; x_n) = -\frac{1}{2} \sum_{i=1}^{n} \log(\sigma_i).$$

(13)

Furthermore, in analogy with Eq. (9), we can define the marginal linear redundancy of $x_1, \ldots, x_{n-1}$ and $x_n$ as:

$$\lambda(x_1, \ldots, x_{n-1}; x_n) = L(x_1; \ldots; x_n) - L(x_1; \ldots; x_{n-1}).$$

(14)
3. TESTING FOR NONLINEARITY

Nonlinearity is the necessary condition for deterministic chaos (Schuster, 1984). Thus, in searching for chaotic dynamics in time series, the first step should be assessing its nonlinearity. We can do it using the above-defined redundancies.

In a typical experimental situation, one deals with a time series $Y(t)$. It is usually considered as a realization of a stochastic process $\{x_i\}$ which is stationary and ergodic.

We will study redundancies for variables

$$x_i(t) = Y(t + (i-1)\tau), \quad i = 1, \ldots, n,$$

where $\tau$ is a time delay and $n$ is the so-called embedding dimension (Takens, 1981). Redundancies of the type

$$R(Y(t); Y(t+\tau); \ldots; Y(t+(n-1)\tau))$$

are, due to stationarity of $Y(t)$, independent of $t$. We introduce the notation:

$$R^0(\tau) = R(Y(t); Y(t+\tau); \ldots; Y(t+(n-1)\tau))$$

for the redundancy and

$$L^0(\tau) = L(Y(t); Y(t+\tau); \ldots; Y(t+(n-1)\tau))$$

for the linear redundancy of the $n$ variables $Y(t), Y(t+\tau), \ldots, Y(t+(n-1)\tau)$; and

$$g^0(\tau) = g(Y(t), Y(t+\tau), \ldots, Y(t+(n-2)\tau); Y(t+(n-1)\tau))$$

for the marginal redundancy and

$$\lambda^0(\tau) = \lambda(Y(t), Y(t+\tau), \ldots, Y(t+(n-2)\tau); Y(t+(n-1)\tau))$$

for the marginal linear redundancy of the variables $Y(t), Y(t+\tau), \ldots, Y(t+(n-2)\tau)$ and the variable $Y(t+(n-1)\tau)$.

Relations (9) and (14) can be rewritten as

$$g^n(\tau) = R^n(\tau) - R^{n-1}(\tau)$$

and

$$\lambda^n(\tau) = L^n(\tau) - L^{n-1}(\tau),$$

respectively.

The linear redundancy, according to its definition (13), reflects dependence structures contained in the correlation matrix $C$ of the variables under study. In
the special case considered here, when all the variables are, according to Eq. (15), lagged versions of the series \( Y(t) \), each element of \( C \) is given by the value of the autocorrelation function of the series \( Y(t) \) for a particular lag. As the correlation is the measure of linear dependence, the linear redundancy characterizes linear structures in the data under study.

We propose to compare the linear redundancy \( L^n(\tau) \) with the redundancy \( R^n(\tau) \) (or the marginal linear redundancy \( \lambda^n(\tau) \) with the marginal redundancy \( \varrho^n(\tau) \)) considered as the functions of the time lag \( \tau \). If their shapes are the same or very similar, a linear description of the process under study should be considered sufficient. Large discrepancies suggest important nonlinearities in links among the variables, or, recalling Eq. (15), among the studied time series and its lagged versions, i.e., in the dynamics of the process under study.

We would like to emphasize that we compare shapes of redundancies as functions of lag \( \tau \), not particular values of the redundancies. Estimated values of \( R^n(\tau) \) and \( \varrho^n(\tau) \) depend on a numerical procedure used ("quantization"—see Appendix 1), while the shapes of their \( \tau \)-plots are usually consistent for a large extent of numerical parameters used in the redundancy estimations. Therefore each figure, depicting redundancies against time lag \( \tau \), is drawn in its individual scale. Redundancies \( R^n(\tau) \) and \( L^n(\tau) \) are plotted as \( R^n(\tau)/(n-1) \) and \( L^n(\tau)/(n-1) \). All the redundancies are in bits and time lags in number of samples. In the case of the Rössler and Lorenz systems, time lags are in relevant time units; in the case of EEG, in milliseconds. Different curves in each figure correspond to redundancies of different numbers \( n \) of variables (embedding dimension); \( n \) is from two usually to five, reading from bottom to top.

Let us recall that the equivalence of redundancy \( R^n(\tau) \) and linear redundancy \( L^n(\tau) \) can be proved only for a special type of linear processes—the processes with the multivariate Gaussian distribution. In a general case, however, we cannot neglect the possibility that differences between redundancies and linear redundancies are not due nonlinearity, but due a non-Gaussian distribution of the studied data. Nevertheless, after extensive numerical study we can conjecture that the shapes of \( \tau \)-dependence of redundancy \( R^n(\tau) \) (marginal redundancy \( \varrho^n(\tau) \)) and linear redundancy \( L^n(\tau) \) (marginal linear redundancy \( \lambda^n(\tau) \)) are approximately the same or similar also for different kinds of linear processes. Only for nonlinear processes the differences seem to be of qualitative level. This conclusion is illustrated by the following examples.

We start with a simple torus (two-periodic) time series. (For details about the numerically generated data, see Appendix 2.) Figures 1(a) and (b) illustrate the linear redundancy \( L^n(\tau) \) and the redundancy \( R^n(\tau) \), respectively, as functions of the time lag \( \tau \), computed from the torus series 51,200 samples long containing 50% of uniformly distributed noise. The multivariate distribution of these data is not exactly Gaussian; however the shapes of both the \( \tau \)-plots of the redundancies are almost the same. All the dependence structures in the data dynamics are detectable on the linear level; i.e., linear description of this data is sufficient.
The same data as above, but without noise, i.e., the “pure” torus, has a multivariate distribution very different from the Gaussian one. This is the cause of differences between $L^n(\tau)$ and $R^n(\tau)$, and Figures 1(c) and (d) are different. Detailed study of redundancies on Figures 1(c) and (d), however, shows that both $L^n(\tau)$ and $R^n(\tau)$ detect the same dependence structures: the maxima and minima of the redundancies are at the same values of the lag $\tau$, and only the relations of their magnitudes are different. These differences we do not consider qualitative and, therefore, in this case, too, we conclude that the linear description of the data should be sufficient.

What we mean by “qualitative differences” of $L^n(\tau)$ and $R^n(\tau)$ is illustrated by the following three data sets, generated by nonlinear dynamical systems.

Figure 2 illustrates redundancies for the series generated by the system with a “strange nonchaotic” attractor, introduced by Grebogi et al. (1984) (51,200 samples of the series length; for details on all numerically generated data, see Appendix 2). Comparing $L^n(\tau)$ and $R^n(\tau)$ on Figures 2(a) and (b), respectively, we can see qualitatively different structures: there are not consistent $\tau$-positions of minima or maxima of the redundancies, there are even different numbers of extrema of $L^n(\tau)$ and $R^n(\tau)$. $R^n(\tau)$ clearly reflects a periodic structure that, unlike that in the above
FIGURE 2 (a) $L^n(\tau)$, (b) $R^n(\tau)$, (c) $\lambda^n(\tau)$, and (d) $\phi^n(\tau)$ of the series generated by the Grebogi strange nonchaotic system.

FIGURE 3 (a) $L^n(\tau)$, (b) $R^n(\tau)$, (c) $\lambda^n(\tau)$, and (d) $\phi^n(\tau)$ of the series generated by the Lorenz system in a chaotic state.
torus data, is not detectable on the linear level. This difference indicates that a linear description of the data is not sufficient and a nonlinear model should be considered.

$L^n(\tau)$ and $R^n(\tau)$ for time series generated by the Lorenz system (Figures 3(a) and (b), respectively) are again different qualitatively: the linear redundancy decreases quickly to values close to zero and detects no dependence for $\tau > 0.4$. On the other hand, the redundancy $R^n(\tau)$ detects nonlinear dependence, the level of which is oscillating with $\tau$ and is characterized by the long-term decreasing trend.

In case of time series generated by the Rössler system, both $L^n(\tau)$ (Figure 4(a)) and $R^n(\tau)$ (Figure 4(b)) are of a similar oscillating nature, but the linear redundancies are not able to detect the long-term decreasing trend, clearly reflected in the redundancy $R^n(\tau)$. This difference we again consider as important or qualitative. The importance and nature of this decreasing trend will be discussed in Section 5.

The series length in both the chaotic cases was 1,024,000 samples. These extensive computations were performed in the study of estimation of the metric entropy (Paluš, submitted)—see Section 5—and are not necessary in testing for nonlinearity itself when the nonlinear character of the data can be identified using time series lengths of several thousands or even hundreds of samples.

FIGURE 4 (a) $L^n(\tau)$, (b) $R^n(\tau)$, (c) $\lambda^n(\tau)$, and (d) $\rho^n(\tau)$ of the series generated by the Rössler system in a chaotic state.
In this section we demonstrated how we can distinguish linear time series from nonlinear ones. At this level we can use either redundancies $L^n(\tau)$ and $R^n(\tau)$ or marginal redundancies $\lambda^n(\tau)$ and $\varphi^n(\tau)$. In Section 5 we will demonstrate the importance of the marginal redundancy for specific detection of chaotic dynamics.

Considering rigorous mathematical results (Eq. 11), we would like to stress that differences between linear redundancy and redundancy cannot be considered as evidence for nonlinearity. We have demonstrated, however, that this difference can be understand as a serious "signature" of nonlinearity. On the other hand, the opposite result, i.e., equality of $L^n(\tau)$ and $R^n(\tau)$ (in the above-considered sense) presents strong support for the hypothesis that the data under study were generated by a linear (stochastic) process. Such a result is valuable— one knows that there is no need to continue analysis of the data by methods of nonlinear dynamics and deterministic chaos; in some cases, it can be also surprising: Figures 5(a) and 5(b) present linear redundancy $L^n(\tau)$ and redundancy $R^n(\tau)$, respectively, computed from a normal human EEG signal (15,360 samples). In spite of many published papers, declaring that "EEG is chaotic" ( Başar, 1990), based on results presented in Figure 5, we can conclude that explanation by a linear (stochastic) process is consistent with these data. The same results we obtained from EEG data of ten other subjects recorded in three independent laboratories (Paluš, in press).

4. COMPARISON WITH OTHER TESTS FOR NONLINEARITY AND DETERMINISM

Several authors were led to developing methodology for testing the basic necessary conditions for chaos—i.e., nonlinearity and/or determinism by never-ending discussions on the relevance of "evidences of chaos" in experimental time series by estimating correlation dimension or other dynamical invariants. Our original approach to this problem was introduced in the previous section. Here we present its...
comparison with two other methods. Both are described by their authors in original papers (Theiler et al., 1992a; Kaplan & Glass, 1992) and also in this volume (Theiler et al., this volume; Kaplan, this volume), so we keep the maximum brevity.

Theiler et al. (1992a) proposed to test for nonlinearity in time series using so called surrogate data. The surrogate data are numerically prepared data with the same statistical properties (mean, variance, spectrum) as the studied experimental data, but the surrogate data are usually created as a linear stochastic process. Then one estimates dimensions and/or other dynamical invariants from both the original and surrogate data. If there are no significant differences between these estimates, the explanation by a linear stochastic process ("null hypothesis") is consistent with the experimental data. Significant differences enable us to reject the null hypothesis of linear stochastic origin of the data and are proposed as signatures of nonlinearity.

We generated surrogate data for all the studied data sets by computing the forward fast Fourier transform (FFT), randomizing the phases and computing the backward FFT. Thus we obtained linear stochastic time series with the same spectra as the original data and computed redundancies from them. We have found that linear redundancies \( L^n(\tau)s \) of surrogate data are equal to redundancies \( R^n(\tau)s \) of surrogate data, and both are equal to linear redundancies \( L^n(\tau)o \) of original data. (Here we again compare time lag plots of redundancies, not their absolute values. \( o \) means original, \( s \) surrogate data.) These results are not surprising: The equality \( L^n(\tau)s = R^n(\tau)s \) follows directly from generating the surrogate data as a linear stochastic process. If two series have the same spectra, then they also have the same autocorrelation functions. Recalling the relation between the autocorrelation function and the linear redundancy, the equality \( L^n(\tau)o = R^n(\tau)o \) also can be understood. Then, providing that both the tests, Theiler's and ours, give consistent results, one need not generate surrogate data; the comparison of \( L^n(\tau) \) and \( R^n(\tau) \) of the studied data is sufficient. But are the tests consistent?

Let us realize that surrogate data are linear and stochastic. When the null hypothesis is rejected, then negation of "linear AND stochastic" is "nonlinear OR deterministic"; i.e., the following hypotheses are possible:

1. nonlinear and deterministic,
2. nonlinear and stochastic, or
3. linear and deterministic.

On the other hand, in our redundancy test our null hypothesis is linearity. After the results presented in previous sections, we can specify our understanding of linearity: we classify a time series as linear if all the dependence structures can be detected by (auto)correlations and, consequently, by linear redundancy; and

\[\text{[1]}\text{The method of surrogate data was used, independently of Theiler et al., by other authors, e.g., Elgar \\& Mayer-Kress (1989) and others. For simplicity we use the term "Theiler's test."} \]
application of (general) redundancy $R^n(\tau)$ brings no new information. The equality $L^n(\tau)[o] = R^n(\tau)[o]$ also implies the equality

$$L^n(\tau)[s] = R^n(\tau)[s] = L^n(\tau)[o] = R^n(\tau)[o]$$

which means that our test cannot distinguish deterministic linear oscillations from a linear stochastic process. Rejecting the null hypothesis (linearity), the following hypotheses are available:

1. nonlinear and deterministic, or
2. nonlinear and stochastic.

That means, in general, these two tests are consistent in detection of processes (1) and (2). The example for case (3) is our "pure" torus series which we classified as linear (Section 3). Estimations of correlation dimension from this data saturate on 2, but estimations from its surrogate data do not; i.e., the Theiler's null hypothesis is rejected.

Detecting nonlinearity in a time series, we are interested in intrinsic nonlinearity in its dynamics. Influence of a "static" nonlinearity on the nonlinearity tests should be considered. Let us suppose that the underlying dynamics of the studied system is linear—i.e., there is an original (stochastic) linear process $\{x_i\}$, but we can measure series $\{y_i\}$, $y_i = f(x_i)$, where $f$ is a nonlinear function. Such nonlinearity, which is not intrinsic for the dynamics of the system under study but can be caused e.g., by a measurement apparatus, we call static nonlinearity.

A static nonlinearity can influence the results of Theiler's test using linear stochastic surrogate data. Therefore, Theiler et al. (1992a) proposed a more complicated algorithm for generating the surrogate data tailored to this specific null hypothesis including static nonlinearity.

We studied the influence of several types of static nonlinearities on our redundancy test. Figures 6(a) and (b) present linear redundancy $L^n(\tau)$ and redundancy $R^n(\tau)$, respectively, for the noisy torus series (51,200 samples) passed through the quadratic nonlinearity. We can see that $L^n(\tau)$ and $R^n(\tau)$ are not exactly the same, but there is no significant (qualitative) difference in the sense discussed in Section 3, i.e., they reflect the same dependence structures. Redundancies $L^n(\tau)$ and $R^n(\tau)$ for the noisy torus data passed through function tangent hyperbolicus are presented in Figures 6(c) and (d), respectively. They are practically the same, which is the nice result demonstrating the robustness of our test: The function tanh transforms the data to the interval $[-1,1]$, and the shape of the distribution of the transformed data is approximately like the character M. Two-dimensional distribution of the series and its lagged twin is like a cube with a conic hole. It is very different from the Gaussian "bell."
The static nonlinearity effectively means that the distribution of the studied data is not Gaussian. Then the above-demonstrated robustness of our test against the static nonlinearity (i.e., that it indicates only dynamic nonlinearity) is compatible with our original conjecture in Section 3 that the equality of $L^n(\tau)$ and $R^n(\tau)$ can hold not only for Gaussian, but also for different types of (dynamically) linear processes.

We have demonstrated that our redundancy test for nonlinearity has slightly better properties than Theiler's surrogate data test in its basic version. The latter, however, can be adapted for more specific types of surrogate data, depending on the particular problem under study. Such specialized tests can improve one's insight to the origin of the particular data, but lose their general applicability.

Kaplan and Glass (1992) proposed a test for determinism based on computation of average directional vectors\footnote{A similar idea was applied in the paper of Cremers & Hübler (1987). We will refer, however, to Kaplan & Glass because we follow their particular implementation.} in a coarse-grained $n$-dimensional embedding of a time series. They showed that in the case of random-walk data magnitudes of these vectors decrease with the number $N$ of passes of the trajectory through the box (i.e., region of coarse-grained state space) as $1/N^{1/2}$. On the other hand, in the...
ideal case of deterministic dynamics, these magnitudes are independent of $N$ and, using normalized directional vectors, are equal to one. In an experimental practice, however, using short and noisy time series, magnitudes of averaged directional vectors of deterministic systems can also decrease with $N$. On the other hand, there are random processes—autocorrelated noises, exhibiting in this test "deterministic" behavior for some range of time lags: Figure 7 illustrates the dependence of the magnitudes of averaged directional vectors for the Rössler, i.e., deterministic series (Figures 7(a) and (c)), and for the surrogate data (i.e., linear stochastic series) of the same series (Figures 7(b) and (d)). We can see that both the series can be considered as clearly deterministic using a short lag $\tau=0.62$ (Figures 7(a) and (b)). Using a larger lag $\tau=3.14$ (Figures 7(c) and (d)), the magnitudes of averaged directional vectors of both the series decrease with $N$, but are above the theoretical dependence derived for random-walk data.

In order to prevent such false detection of determinism, Kaplan & Glass (1992) proposed the new quantity $A$—the weighted average of magnitudes of averaged
directional vectors, averaged through all the occupied boxes—and compared its
dependence on time lag \( \tau \) with (a) the autocorrelation function of the series under
study and (b) \( \Lambda(\tau) \) (\( \tau \)-trace of \( \Lambda \)) computed from the (linear stochastic) surrogate
of the studied data.

They showed that in the case of surrogate (i.e., random) data, the \( \tau \)-dependence
of \( \Lambda \) reflects the shape of the autocorrelation function (of both the original and sur-
rogate data—see discussion above). \( \Lambda(\tau) \) of a deterministic (and nonlinear, we add)
time series is different from its autocorrelation function (and \( \Lambda(\tau) \) of its surrogate
data).

At this step the Kaplan-Glass test becomes similar to our redundancy one
and we demonstrate that it also gives very similar results: Figures 8(a) and (b)
present the \( \Lambda(\tau) \) dependence for the time series generated by the Rössler system
and its surrogate, respectively; Figures 8(c) and (d) illustrate the same for the
series generated by the Lorenz system and its surrogate. Compare this figures with
Figures 3(a) and (b) and 4(a) and (b) illustrating the linear redundancy \( L^m(\tau) \) and
the redundancy \( R^m(\tau) \) for these data. (Remember that \( R^m(\tau) \) of the surrogate data
is the same as \( L^m(\tau) \) of the original data.)
5. DETECTION AND QUANTIFICATION OF CHAOS

In the previous two sections, we discussed the possibility of detecting nonlinearity in a time series. If the series is identified as nonlinear, there are still several possible types of underlying processes: stochastic or deterministic or, more specifically, chaotic processes as a special type of nonlinear deterministic processes. In this section we will demonstrate how the latter can be identified. The tools for identification and also for subsequent quantification of chaotic dynamics are marginal redundancies $\rho^m(\tau)$, and the basic theoretical concept is the concept of classification of dynamical systems by information rates. The latter was introduced by Kolmogorov (1959) who, inspired by information theory, generalized the notion of the entropy of an information source. In ergodic theory of dynamical systems, it is known as the metric or Kolmogorov-Sinai entropy (Kolmogorov, 1959; Martin & England, 1981; Petersen, 1983; Sinai, 1959, 1976; Paluš, submitted; Walters, 1982).

The rigorous definition of the metric (Kolmogorov-Sinai) entropy requires introduction of basic notions of ergodic theory, which is beyond the content of this chapter. Therefore we refer to any book on ergodic theory (Billingsley, 1965; Cornfeld, Fomin, & Sinai, 1982; Martin & England, 1981; Petersen, 1983; Sinai, 1976, Walters, 1982) as well as to the author’s recent paper (Paluš, submitted), and here we will return to an experimental time series considered as a realization of a stationary ergodic stochastic process $\{x_i\}$. We will point out correspondence between a process $\{x_i\}$ and a measure-preserving dynamical system (Cornfeld, Fomin, & Sinai, 1982; Sinai, 1959, 1976; Paluš, submitted; Walters, 1982) for which we can define the metric entropy in terms of entropies introduced in Section 2.

A measure-preserving dynamical system can correspond to a stochastic process $\{x_i\}$ in the sense that an orbit, presenting a single evolution of the system in its state space, is mapped by a measurable map from this space to the set of real numbers. The resulting series of real numbers represents values of physically observable variable measured in successive instants of time—experimental time series.

Conversely, any stationary stochastic process corresponds to a measure-preserving system in a standard way: One can construct a map $\Phi$ mapping variables of a stochastic process to a sequence of points $\{z_i\}$ of a measure space $Z$ and define the shift transformation $\sigma$ on the sequence $\{z_i\}$ as

$$\sigma z_n = z_{n+1}.$$  

Due to the stationarity of the original process, such a system is a measure-preserving transformation (dynamical system). (For more details see Petersen’s book (1983).)

Considering this correspondence we can express the metric entropy of an underlying dynamical system in terms of entropies of a sequence of random variables $x_i$. Let us remark first that in practice there is never a continuous probability distribution density, and we consider all the entropies defined using discrete probability distributions on a finite partition of the state space.
Under the above consideration we define the entropy $h(T, \xi)$ of the dynamical system $T$ with respect to the partition $\xi$ as

$$h(T, \xi) = \lim_{n \to \infty} H(x_n | x_1, \ldots, x_{n-1}).$$  \hfill (22)

The metric (Kolmogorov-Sinai) entropy of the dynamical system $T$ is then

$$h(T) = \sup_{\xi} h(T, \xi),$$  \hfill (23)

where the supremum is taken over all the finite partitions (Petersen, 1983; Sinai, 1976; Walters, 1982).

Among all the finite partitions, the key role is played by so-called generating partitions. The partition $\alpha$ is generating with respect to the transformation (dynamical system) $T$ if all the measurable sets of the system state space (more precisely $\sigma$-algebra) can be generated by countably-fold application of $T$ on $\alpha$ (Petersen, 1983; Sinai, 1976; Walters, 1982). Then the Kolmogorov-Sinai theorem (Petersen, 1983; Sinai, 1976; Walters, 1982) holds:

$$h(T) = h(T, \alpha).$$  \hfill (24)

The following theorem is important for further consideration here: Let $T$ be a discrete measure-preserving dynamical system. Then

$$h(T^k) = |k|h(T)$$  \hfill (25)

for every whole number $k$. For the continuous flow $T_t$ and any real $t$, the equality

$$h(T_t) = |t|h(T_1)$$  \hfill (26)

holds (Sinai, 1976).

Now we can proceed to the relation between the metric entropy and the marginal redundancy $\varrho^n(\tau)$: Comparing Eqs. (22) and (10) for $n \to \infty$, we have

$$\varrho^n(\tau) \approx A_\xi - h(T_\tau, \xi),$$

where $A_\xi$ is a parameter independent of $n$ and $\tau$ (and, clearly, dependent on $\xi$) and $h(T_\tau, \xi)$ is the entropy of (continuous) dynamical system $T_\tau$ with respect to the partition $\xi$.

Let $\xi$ be the generating partition with respect to $T$. Then, considering Eqs. (26) and (24), we have

$$\lim_{n \to \infty} \varrho^n(\tau) = A - |\tau|h(T_1).$$  \hfill (27)

This assertion was originally conjectured by Fraser (1989b). Application of ideas and methods of information theory in nonlinear dynamics was originally proposed by Shaw (1981).
Let us consider further that the studied time series was generated by an \( m \)-dimensional continuous-time dynamical system \( T \), fulfilling the conditions of the existence and uniqueness theorem (Arnold, 1973; Kamke, 1959) and the particular trajectory of \( T \) is mapped from the state space \( S \) to the set of real numbers. There is the unique trajectory passing through each point \( s \in S \) so that the evolution on the particular trajectory is fully determined by one \( m \)-dimensional point \( s \in S \). On the other hand, \( m \)-tuples of \( m \) successive points \( Y(t), \ldots, Y(t+(m-1)\tau) \) can, according to the theorem of Takens (1981), form a mapping of the process \( \{Y(t)\} \) to a space \( Z \), so that the sequence \( \{z_i\} \) of images of \( m \)-tuples \( Y(t), \ldots, Y(t+(m-1)\tau) \) is topologically equivalent to the original trajectory \( \{s_i\} \) in \( S \). Hence, a particular \( m \)-tuple \( Y(t), \ldots, Y(t+(m-1)\tau) \) is equivalent to a point from \( S \), and thus it determines the rest of the series \( \{Y(t)\} \). It means that only the redundancies \( q^n(\tau) \) for \( n \leq m \) should be finite and, for \( n > m \), redundancies \( q^n(\tau) \) should diverge. This is, however, a theoretic consideration providing infinite precision. In experimental and numerical practice, the measurement noise and finite precision emerge, and all the estimated redundancies \( q^n(\tau) \) are finite but increasing with \( n \). We can only suppose that the increase of \( q^n(\tau) \) for \( n > m \) is lower than for \( n \leq m \) and is independent of \( \tau \).

Intuitively we can explain this supposition by the fact that adding another variable to \( n \) variables, \( n < m \), the common information measured by \( q^n(\tau) \) is increased by specific dynamical information; i.e., the increase \( q^{n+1}(\tau) - q^n(\tau) \) depends on \( n \) and \( \tau \). Addition of another variable when \( n > m \) is, considering the increase \( q^{n+1}(\tau) - q^n(\tau) \) of the common information, (approximately) adequate to addition of a "noise term" contributing only nonspecific information relevant to noise and finite precision. Therefore, for \( n > m \), we expect the curves \( q^n(\tau) \) as functions of \( \tau \) have the same shape, only they are shifted; i.e., \( q^{n+1}(\tau) = q^n(\tau) + \text{const} \). Thus the limit behavior of \( q^n(\tau) \) for \( n \to \infty \), in the case of an \( m \)-dimensional dynamical system, is attained for very small \( n \), actually for \( n = m+1, m+2, \ldots \).

Let us consider that the probability distribution \( p(x_1, \ldots, x_n) \) used in the estimation of \( q^n(\tau) \) corresponds to the generating partition of the studied \( m \)-dimensional dynamical system for a certain extent of \( \tau \). Then the limit behavior (27) of the marginal redundancy, i.e., \( q^n(\tau) \approx A - \tau h(\lambda_1) \), is attained for \( n = m+1, m+2, \ldots \). And this is actually the behavior of \( q^n(\tau) \) for low-dimensional dynamical systems. The extent of \( \tau \) for which marginal redundancies approach the linearly decreasing function is usually bounded by some \( \tau_1 \) and \( \tau_2 \), as is discussed in the author's recent paper (Paluš, submitted).

This phenomenon is illustrated in Figures 3(d) and 4(d) presenting the marginal redundancies \( q^n(\tau) \) for the Lorenz and the Rössler system, respectively. \( q^n(\tau) \) for \( n = 4 \) and \( 5 \) in both the cases approach linearly decreasing function \( A - \tau h \), where the slope \( h \) is an estimation of the positive metric entropy of the underlying chaotic dynamical system.

Thus we can consider the linear decrease of the marginal redundancy \( q^n(\tau) \) of the studied data as the signature of chaos and the value of its slope—estimation
of the metric entropy can serve as the quantitative description of the data under study.

We use the term “signature” and not the popular term “evidence for chaos,” because this property, like all so-called evidences, is not a sufficient but necessary condition for chaos (e.g., the implication “chaos ⇒ finite dimension” holds, but the inverse does not).

The conditions for successful identification of chaos by \( g^n(\tau) \), as we explained above, are \( n \) exceeding the dimension of the studied system and a partition fine enough to be generating (redundancy is estimated by box-counting method—see Appendix 1). (Each refinement of a generating partition is generating; i.e., for systems with finite entropy, the boxlike generating partitions exist.) Fine and high-dimensional partitions require a large amount of data (otherwise, estimations of redundancies are heavily biased by the effect of “overquantization”—see Appendix 1); therefore, these conditions cannot always be fulfilled and chaos in time series identified. This is, however, a general property: Using an insufficient amount of data, any method can give biased results.

6. ANALYSIS OF THE COMPETITION DATA

We applied the above-described redundancy method on the sets A, D, and E of the competition data.

DATA SET A

The results, linear marginal redundancy \( \lambda^n(\tau) \) and marginal redundancy \( g^n(\tau) \), are presented in Figures 9(a) and 9(b), respectively. We can see that \( g^n(\tau) \) for \( n = 4 \) and 5 approach more or less linearly decreasing functions. This is

![Figure 9](image-url)
1. a qualitative difference from the shape of $\lambda^n(\tau)$, i.e., the signature of nonlinearity, and
2. also the signature of chaos—it detects a positive value of the metric entropy of the system underlying this series. The marginal redundancy $\theta^n(\tau)$ for $n = 5$ is, however, distorted due to “overquantization”; i.e., there is not enough data to obtain “better” results—consistent slopes for $n = 5$ or higher (see Appendix 1).

The result correctly identified the chaotic laser dynamics, which can be described by the system of three ordinary nonlinear differential equations (Hübner, this volume), having an attractor of correlation dimension about 2.05 and positive metric entropy. The amount of the data available (10,000 samples) is unfortunately not sufficient to obtain reliable estimation of the value of the metric entropy (Paluš, submitted).

The dynamical system describing this laser dynamics is equivalent to that of the Lorenz system (Hübner, this volume). The attractors of these systems have very similar topological properties, including dimension. Considering our results (Figure 9), we can state that the information production or “the levels of chaoticity” of these systems are different. In order to explain this point, let us remember Figures 3(c) and 4(c), depicting $\lambda^n(\tau)$ of the Lorenz and the Rössler systems, respectively. In the case of the Lorenz system, the (marginal) linear redundancy $\lambda^n(\tau)$ for $\tau > 0$ decreases immediately to a close-to-zero level, classifying the series as the sequence of independent random variables—white noise (i.e., $\lambda^n(\tau)$ is not able to detect any dependence at all due to the high rate of information production destroying correlations), while in the case of the Rössler system, $\lambda^n(\tau)$ stays on the significant nonzero level as in the case of regular oscillations (i.e., $\lambda^n(\tau)$ is not able to detect the production of information here). We explain this difference by different "levels of chaoticity" of these systems—the Lorenz system is "strongly chaotic" and the Rössler one is "weakly chaotic." To be more specific, the metric entropy or the positive Lyapunov exponent of the Lorenz system is about ten times greater than that of the Rössler system (Wolf et al., 1985).

According to Figure 9(a) the “chaoticity” of the laser dynamics under study is between those of the Lorenz and the Rössler systems. Examples of such “moderately chaotic systems” in which the linear redundancy is nonzero for a non-negligible extent of lag $\tau$, usually oscillating but decreasing in such a way that the envelope of oscillations of $\lambda^n(\tau)$ reflects the exponential or power-law decrease, are also known (Kot, Sayler, & Schultz, 1992). And this is the case of our chaotic laser (Figure 9(a)).
FIGURE 10 (a) $L^n(\tau)$ and (b) $R^n(\tau)$ for the data set E (13th continuous segment)—light curve of the white dwarf star.

FIGURE 11 (a) $L^n(\tau)$, (b) $R^n(\tau)$, (c) $\lambda^n(\tau)$, and (d) $\rho^n(\tau)$ for the data set D. The embedding dimensions are $n = 2$ to 8 reading from bottom to top. From $n = 6$ the overquantization effect emerges (d).
DATA SET E

Figures 10(a) and (b) present the linear redundancy $L^n(\tau)$ and the redundancy $R^n(\tau)$ computed from the 13th continuous segment (2,568 samples) of set E—the light curve of the variable white dwarf star (Clemens, this volume). We can see there is no significant difference between $L^n(\tau)$ and $R^n(\tau)$, and we conclude that explanation by a linear (stochastic) process is consistent with the data. Analysis of the other two longest segments gave the same results.

DATA SET D

We can see all the redundancies in Figure 11. (Redundancies $R^n(\tau)$ and $\rho^n(\tau)$ for $n = 6, 7, 8$ are distorted due to the overquantization effect—see Appendix 1.) Linear redundancies decrease quickly to near-zero values indicating noiselike dependence level for lags $\tau > 5$ samples, while redundancies $R^n(\tau)$ (or $\rho^n(\tau)$) are clearly above zero up to lags of approximately 50 samples, and reflect an oscillating phenomenon. This difference we consider as the signature for nonlinearity. There is, however, no specific “chaotic” behavior of marginal redundancies $\rho^n(\tau)$: They decrease, but a tendency to the linear decrease is not apparent. This long-term trend is closer to an exponential or power-law decrease. Therefore we should assume the following possibilities for the series dynamics:

1. nonlinear stochastic,
2. nonlinear deterministic, but not chaotic, and
3. chaotic, but for identification of its chaoticity, finer partition is necessary and/or its dimensionality is higher than used in analysis. To test these possibilities, longer time series are necessary than the given 100,000 samples.

Above (Figures 1 and 2) we have found that redundancies of deterministic nonchaotic series have no (i.e., zero) long-term trend. A decreasing trend of redundancies of the data set D, then, can be explained by the above hypotheses (1) and (3). Can we reject hypothesis (2)?

Looking at the data set D as a whole, we can see its nonstationary character—sudden changes of the extent of amplitudes. Using time-delay reconstruction we can observe that the dynamics is oscillating but switching among several areas in the phase space.

In the theory of Sections 2, 3, and 5, we required stationarity of the time series under study. Being mathematically strict we should reject application of our redundancy test to nonstationary data. But, let us make the exception for a while and study the effect of nonstationarity on time-lag dependence of the redundancies.

We added the Gaussian drift (generated as integration of a Gaussian random variable) to (a) the amplitude of our torus time series, and (b) parameter $\omega$ (see Appendix 2) of the system of Grebogi et al. (1984), analyzed above. The parameter was drifted in each iteration step.
The results are presented in Figure 12. Nonstationarity induces the long-term decreasing trend of all the redundancies. (Remember that these data in their original, stationary forms, did not exhibit any long-term decrease of redundancies—Figures 1 and 2.) It is also interesting that the character of these data, i.e., either linear or nonlinear, is not changed (cf. $L^n(\tau)$ with $R^n(\tau)$ in Figures 12(a)–(b) and 12(c)–(d)). We must remark, however, that the roots of this result lie not only in the different nature of the data, but also in different levels of the dynamics influenced by the drift. When we simply added the Gaussian drift to the amplitude of the series generated previously by the system of Grebogi et al. (1984), we found that while the small amplitude drifts do not affect the redundancies, the drifts with larger amplitude "Gaussianize" the series: redundancies $R^n(\tau)$ become similar to the linear redundancies $L^n(\tau)$; i.e., the nonlinear character of the series is hidden by the Gaussian noise added.

Our conclusion based on this experimentation is that hypothesis (2) cannot be rejected. The data set D could be generated by a system regularly oscillating in normal conditions, but disturbed by some drift or other nonstationarity in its structural parameters, and in fact this is the case. Using post-competition information, we know that the data set D reflects the dynamics of a driven particle in a

![Figure 12](image_url)

**FIGURE 12** (a) $L^n(\tau)$, (b) $R^n(\tau)$ for the torus series with the Gaussian drift in the amplitude, (c) $L^n(\tau)$, (d) $R^n(\tau)$ for the Grebogi system series iterated with the small Gaussian drift in the system parameter $\omega$. 
four-well potential (Gershenfeld, this volume). Normally the evolution after short transient period ends in regular nonlinear oscillations in one of the wells: Figures 13(a) and 13(b) present $L^n(\tau)$ and $R^n(\tau)$, respectively, for this case, i.e., for the series obtained by integrating the system without the drift in potential. There is no long-term decreasing trend in these redundancies, and looking closely at them (Figure 13(c) illustrates $L^n(\tau)$ and $R^n(\tau)$ for $n = 2$) we can see that there are several peaks in redundancy $R^n(\tau)$ missing in linear redundancy $L^n(\tau)$; i.e., specific nonlinear dependence structures which are not detectable on the linear level. Considering the phase portrait of these oscillations, we can see that it is not a simple cycle (Figure 13(d)).

The specific behavior of the data set $D$ was induced by a small Gaussian drift in one of the potential parameters. This caused prolonged transient dynamics, consisting of nonlinear oscillations within the wells and random switching among them. Thus the system exhibits a “dynamical loss of memory” (i.e., a long-term decrease of redundancies), but it is of a different origin than this phenomenon in the case of chaotic dynamical systems.

Without information about the origin of this data set, however, we are able to detect only its nonlinearity and none of the above nonlinear hypotheses (1)–(3) can be rejected.
7. CONCLUSION

We presented the method for analysis of experimental time series suitable for assessing the nonlinearity of a series and for identification and quantification of the underlying chaotic dynamics. It is based on examination of time-lag dependence of redundancies, the information-theoretic functionals computed from the studied time series and its lagged versions. It was demonstrated in the extensive numerical study that this method is able to discern linear from nonlinear processes and specifically detect low-dimensional chaotic dynamics. It is valuable also for quantitative characterization of chaotic systems by measuring their information production rates in terms of metric entropy.

ACKNOWLEDGMENT

The author would like to thank V. Albrecht, I. Dvořák, A. M. Fraser, A. Longtin, and J. Theiler for stimulating discussions and D. Bensen for careful reading of the manuscript. Special thanks are due to A. S. Weigend and N. A. Gershenfeld.

This study was performed as a part of the projects “Information and entropy properties of the spontaneous brain electrical activity,” supported within the framework of the “IBM Academic Initiative in the CSFR,” and “Application of EEG complexity measures for quantification of changes of the brain state,” supported by the Ministry of Health of the Czech Republic.

APPENDIX 1—REDUNDANCY ALGORITHM

Linear redundancies were computed according to Eq. (13). Eigenvalues of each correlation matrix were obtained by its decomposition using SVDCMP routine (Press et al., 1988).

Practical computations of mutual information and redundancies of continuous variables are always connected with the problem of quantization. By quantization we understand the definition of the finite-size boxes covering the state space. The probability distribution is then estimated as relative frequencies of occurrence of data samples in particular boxes. The naive approach to estimate redundancies of continuous variables should be the use of the finest possible quantization given, e.g., by a computer memory or measurement precision. We must remember, however, that we usually have a finite number $N$ of data samples. Hence, using too fine quantization can cause the estimation of entropies and redundancies to be heavily
biased: Estimating joint entropy of $n$ variables using $Q$ marginal quantization levels, one obtains $Q^n$ boxes covering the state space. If $Q^n$ approaches the number $N$ of data samples, or even $Q > N$, the estimate of $H(x_1, \ldots, x_n)$ can be equal to $\log(N)$ or, in any case, it can be determined mainly by the number of the data samples and/or by the number of the distinct data values and not by the data structure, i.e., by the properties of the system under study. We say, in such a case, that the data are overquantized. (Even the "natural" quantization of experimental data given by an A/D converter is usually too fine for reliable estimation of the redundancies.)

The emergence of overquantization is given by the number of boxes covering the state space; i.e., the higher the space dimension (the number of variables), the lower the number of marginal quantization levels that can cause overquantization. Recalling definition (7) of the redundancy of $n$ variables, one can see that while the estimate of the joint entropy can be overquantized, i.e., saturated on some value given by the number of the data samples and/or by the number of the distinct data values, the estimates of the individual entropies are not and they increase with finer quantization. Thus the overquantization causes overestimation of redundancy $R^n(T)$ and smearing of its dependence on $T$.

Recalling $g^n(T) = R^{n+1}(T) - R^n(T)$, one can see that overquantization causes overestimation of marginal redundancy and, moreover, attenuation of its decrease with increasing $T$. Further overquantization can lead to a paradoxical, unreal result of $g^n(T)$ increasing with $T$ which formally implies negative metric entropy (Paluš, submitted).

Therefore, one must be very careful in defining the quantization. Fraser & Swinney (1986) have proposed an algorithm for constructing the locally data-dependent quantization. We have found, however, that one need not develop such a complicated algorithm; a simple box-counting method is sufficient. The only "special prescriptions," based on our extensive numerical experience, concern the method of data quantization:

a. The type of quantization: We propose to use the marginal equiquantization method; i.e., the boxes for box counting are defined not equidistantly, but so that there is approximately the same number of samples in each marginal box.

b. Number of quantization levels (marginal boxes): We have found that the requirement for the effective\[3\] series length $N$ using $Q$ quantization levels in computation of $n$-dimensional redundancy is

$$N \geq Q^{n+1};$$

otherwise, the results can be heavily biased due to overquantization as stated above.

\[3\] The effective series length $N$ is $N = N_0 - (n - 1)\tau$, where $N_0$ is the total series length, $n$ is the embedding dimension, and $\tau$ is the time delay used in the reconstruction of the $n$-dimensional embedding.
APPENDIX 2

Two-periodic noisy data were generated according to the following formula:

\[ Y(t) = (R_1 + R_2 \sin(\omega_2 t + \phi)) \sin(\omega_1 t) + \xi, \]

where \( R_1 : R_2 = 5 : 4, \omega_1 : \omega_2 = 10 : 9, \phi = 1.3\pi, \) and \( \xi \) are random numbers uniformly distributed between \(-\Xi\) and \(\Xi\). The term "50% of noise" means that \( R_1 : R_2 : \Xi = 5 : 4 : 9 \). For the torus series without noise, the same formula with \( \xi = 0 \) holds.

Chaotic data were generated by numerical integration, based on the Bulirsch-Stoer method (Press et al., 1986), of the Rössler system (Rössler, 1976)

\[
\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} = (-z - y, x + 0.15y, 0.2 + z(x - 10)),
\]

with initial values \((11.120979, 17.496796, 51.023544)\), integration step 0.314, and accuracy 0.0001; and the Lorenz system (Lorenz, 1963)

\[
\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} = (10(y - x), 28x - y - xz, xy - 8z/3),
\]

with initial values \((15.34, 13.68, 37.91)\), integration step 0.04, and accuracy 0.0001. Component \(x\) was used in both the cases.

The time series from the "strange nonchaotic attractor" (Grebogi et al., 1984) was obtained by iterating the system:

\[
\begin{align*}
\Theta_{n+1} &= (\Theta_n + 2\pi\omega) \text{mod}(2\pi) \\
u_{n+1} &= \Lambda(u_n \cos(\Theta) + v_n \sin(\Theta)) \\
v_{n+1} &= -0.5\Lambda(u_n \cos(\Theta) - v_n \cos(\Theta))
\end{align*}
\]

where \( \omega = (5^{1/2} - 1)/2 \) and \( \Lambda = 2/(1 + u_n^2 + v_n^2) \). Component \( \Theta \) was recorded.

The Gaussian drift was generated by the same algorithm as in the program used for the generation of data set D (Gershenfeld, this volume).
REFERENCES

7. Clemens, C. “Whole Telescope Observations of the White Dwarf Star (PG1159-035).” This volume.
18. Gerchenfeld, N. This volume.


