Chaotic Measures and Real-World Systems:

Does the Lyapunov exponent always measure chaos?

Milan Paluš

Institute of Computer Science, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic; and
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
E-mail: `mp@uivt.cas.cz, mp@santafe.edu`

Abstract. Direct estimation of the largest Lyapunov exponent as a measure of exponential divergence of nearby trajectories is well established in the case of deterministic dynamical systems. Questions are naturally raised about applicability of Lyapunov exponents and other "chaotic measures" when analyzing data from real-world systems, which are either stochastic or affected by numerous external influences, which cannot be described in any other way than a stochastic component in system dynamics. In a series of numerical experiments, Gaussian random deviates were added to a set of chaotic time series with different Lyapunov exponents. It is demonstrated that the estimated Lyapunov exponents fail to distinguish different noisy chaotic time series when relatively small scales are used. The distinction can be reestablished by using larger scales. Using larger scales, however, the estimated Lyapunov exponent is determined by macroscopic statistical properties of the series and provides the same information as the autocorrelation function and/or coarse-grained mutual information.

1 Introductory Remark

This paper is meant as an informal but, it is hoped, informative contribution to a discussion of problems related to nonlinear techniques used in analysis of physiological or other experimental real-world time series. It is addressed to a broad audience of readers with different educational backgrounds, to both theorists and practitioners, therefore I limit the use of mathematical formulae to the necessary minimum, and try to explain discussed facts verbally and by presenting graphical material. Nevertheless, I suppose that a reader is familiar with basic notions related to time series analysis, both linear theory (spectrum, autocorrelation) and chaos approach (dimensions, Kolmogorov-Sinai entropy, Lyapunov exponents).

2 Introduction

Distinction and classification of different dynamical phenomena or systems, or, distinction and classification of different dynamical states of a system is a common problem in many areas of natural and social sciences. In many cases this problem can be translated into a task of quantitative characterization of observable signals, i.e., into an estimation of a quantitative measure from registered time series.

In physiology and medicine time series are recorded which are related to different physiological and/or pathological states of an organism or its parts. Classification of such data is a challenging task with significance for understanding underlying physiological processes and for medical diagnostics. Traditional – linear techniques for time-series analysis have been applied successfully in various areas of physiology and medicine, however, they may have serious limitations due to the fact that physiological processes are usually nonlinear. Therefore it is not surprising that recently developed methods for nonlinear time series analysis have immediately found their way into physiology and biomedical research.

Many of the current methods in nonlinear signal processing have not arisen as an extension of linear analysis, but have been conceived due to the entirely new idea of deterministic chaos. These innovative techniques have provided experimentalists with new ways of understanding the implications of their data, though the limitations of these new techniques have not always been understood and necessary precautions fully appreciated.

Estimation from time series of descriptive measures such as dimensions, Lyapunov exponents or Kolmogorov entropy, derived from theory of deterministic chaos ("chaotic measures") is well established in the case of data generated by low-dimensional deterministic dynamical systems in numerical and laboratory experiments. Questions are naturally raised about applicability of the chaotic measures when analyzing data from real-world systems, which are either stochastic or affected by numerous external influences, which cannot be described in any other way than a stochastic component in system dynamics. Analyzing time series from physiological systems, many authors have realized that low-dimensional chaos in such systems is improbable, however, they have demonstrated that formal estimates of the chaotic measures may possess some discriminating power with respect to data recorded in different experimental conditions (Layne et al., 1986), (Mayer-Kress and Layne, 1987), (Koukkou et al., 1993), (Wackermann et al., 1993). This "relative characterization" of different datasets may surely have its importance in diagnostics, however, the question should be asked: *What do the quantities for measuring chaos actually measure, when in processed data there is no chaos, or a chaotic phenomenon is obscured by noise?* The term "measuring chaos" should be deciphered considering a particular chaos-based method used: What do the small numbers, obtained from dimensional algorithms, actually mean, when the underlying system is high-dimensional or stochastic? What do the estimates of Lyapunov exponents, designed to measure exponential divergence of nearby trajectories, actually characterize, when there is no exponential divergence of trajectories, or even there are no trajectories in the data, or the trajectories are obscured by noise? These questions are important from both practical and theoretical points of view. When the chaotic measures, designed for characterization of low-dimensional dynamics, are applied to analysis of high-dimensional or stochastic systems, precision of their estimates, their robustness with respect to noise, or their sensitivity to changes in underlying dynamics can hardly be established. In theoretical aspect, correct interpretation of obtained results is unclear, while using the original meaning and interpretations of the chaotic measures, i.e., using a "low-dimensional language" for high-dimensional or stochastic systems can be misleading.

This paper does not have the ambition to answer generally these important questions. It is just a demonstration of a particular situation of applying the direct method for estimating the largest Lyapunov exponent (LLE) (Wolf et al., 1985) to noisy chaotic and linear stochastic data. A set of chaotic time series with different positive Lyapunov exponents was generated. It is shown that the LLE algorithm correctly distinguishes and orders the series according to their positive LE's. Then Gaussian noises with zero means and different standard deviations (SD's) are added to the series. The distinction of the series is lost when the LLE algorithm uses scales comparable to, or smaller than the SD of the noise. The noisy chaotic series can be correctly distinguished and ordered, when the scales used in the LLE estimation are larger than the noise's SD.

The requirement to use relatively large scales in practical estimation of the chaotic measures is very typical due to finite precision measurements, limited amounts of data and/or noise in the data, as in this case. Using large scales, however, do the chaotic measures indeed "measure chaos", i.e., does the LLE algorithm measure the exponential divergence of nearby trajectories, or something else? Using isospectral surrogate data approach we show that the LLE in fact distinguishes time series with different autocorrelation functions. This property is probably typical also for dimensional estimates and other measures which explore distributions of distances between points. Thus the chaotic measures, estimations of which usually possess high computational cost and vulnerability to various experimental and numerical factors, in many cases provide the same information as the results obtained by standard linear time series tools such as the spectral or autocovariance analysis. In highly nonlinear systems, when the linear analysis is inadequate, the autocovariance/autocorrelation function should be substituted by (coarse-grained) mutual information and related measures, which can provide reliable relative classification of different dynamical states of nonlinear systems.

3 The Largest Lyapunov Exponent and the Baker Map

Given a scalar time series $x(t)$, an m-dimensional trajectory is reconstructed using the time-delay method (Takens, 1981) as $\mathbf{x}(t) = \{x(t), x(t+\tau), \ldots, x(t+[m-1]\tau)\}$, where τ is the *delay time* and m is the *embedding* *dimension.* A neighbour point $\mathbf{x}(t')$ is located so that the initial distance δ_I, $\delta_I = ||\mathbf{x}(t) - \mathbf{x}(t')||$, is $s_{min} \leq \delta_I \leq s_{max}$. $||.||$ means the Euclidean distance. The *minimum and maximum scales* s_{min} and s_{max}, respectively, are chosen so that the points $\mathbf{x}(t)$ and $\mathbf{x}(t')$ are considered to be in a common "infinitesimal" neighborhood. After an *evolution time* $T \in \{1, 2, 3, \ldots\}$, the resulting final distance δ_F is calculated: $\delta_F = ||\mathbf{x}(t+T) - \mathbf{x}(t'+T)||$. Then the local exponential growth rate per time unit is:

$$\lambda_1^{local} = \frac{1}{T} \log(\delta_F/\delta_I). \quad (1)$$

To estimate the overall growth rate, in the case of deterministic dynamical systems the largest Lyapunov exponent (LLE) λ_1, the local growth rates are averaged along the trajectory:

$$\lambda_1 = <\lambda_1^{local}> = \frac{1}{T}[< log(\delta_F) > - < log(\delta_I) >], \quad (2)$$

where $< . >$ denotes averaging over all initial point pairs fulfilling the condition $s_{min} \leq \delta_I \leq s_{max}$.

These ideas are applied in the fixed evolution time program for estimating LLE as proposed by Wolf et al. (1985). More details, as well as the code of the program FET1, used in this study, can be found in (Wolf et al., 1985).

The set P of numerical parameters:

$$P = \{m, \tau, T, s_{min}, s_{max}\} \quad (3)$$

is chosen by a user.

The data for this study were generated using the well-known chaotic baker transformation:

$$(x_{n+1}, y_{n+1}) = (\beta x_n, \frac{1}{\alpha} y_n)$$

for $y_n \leq \alpha$, or:

$$(x_{n+1}, y_{n+1}) = (0.5 + \beta x_n, \frac{1}{1-\alpha}(y_n - \alpha)) \quad (4)$$

for $y_n > \alpha$;
$0 \leq x_n, y_n \leq 1$, $0 < \alpha, \beta < 1$, β was set to $\beta = 0.25$. For this system the positive Lyapunov exponent λ_1 can be expressed analytically as the function of the parameter α (Hentschel and Procaccia, 1983), (Farmer et al., 1983):

$$\lambda_1(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}. \quad (5)$$

The second LE of this system is negative and is given as (Schuster, 1988):

$$\lambda_2 = \log \beta, \quad (6)$$

in this case $\lambda_2 = -2\log 2$.

Varying the parameter α from 0.01 to 0.49 with step 0.005, ninety-seven time series with different positive Lyapunov exponents λ_1 were generated. The component y was recorded[1], the series length $N = 1024$ samples in each case of this study was used. In addition to the original strictly deterministic series, also noisy data were prepared. The noise considered in this study is the additive "measurement" noise, i.e., the strictly deterministic series y_n, $n = 1, \ldots, N$, were generated according to Eq. (4). Then a noisy series ξ_n, $n = 1, \ldots, N$, of Gaussian random deviates with zero mean and unit variance, generated using the GASDEV procedure from Ref. (Press et al., 1986), were added to the deterministic series:

$$z_n \;=\; y_n + c\xi_n, \tag{7}$$

and the noisy series z_n were analyzed. The coefficients c were defined so that the standard deviation (SD) of the noise was equal to a defined portion of the SD of the original noise-free data. That is, the term "10% of noise" means that the SD of the added noise is equal to 0.1SD of the original data.

The set of 97 baker series with different $\lambda_1(\alpha)$ is an ideal material for simulating the task of relative characterization, i.e., the task of distinguishing and ordering the series according to their "chaoticity", i.e., according to their λ_1. The exact dependence of $\lambda_1(\alpha)$ on the parameter α, based on the analytic formula (5), is displayed in Fig. 1a. Figure 1b presents estimates of λ_1 from noise-free data using the following numerical parameters: $m = 2$, $\tau = 2$, $T = 1$, s_{min} =0.01SD, i.e., 1% of SD of a particular series, s_{max} is always defined as $s_{max} = 10 s_{min}$ in this study. The λ_1 estimates in Fig. 1b agree with the correct $\lambda_1(\alpha)$ values only for small α, while the majority of the results in Fig. 1b are overestimated. It is possible to "tune" the results by changing some parameters from P (Eq. 3), e.g., the estimates would decrease using larger evolution time T. Trying to simulate a real problem of classifying experimental time series, where the correct values of λ_1 are unknown (or, in strict mathematical sense they do not exist), it may be dangerous to tune the parameters P for each estimate individually[2]. As the methodologically correct approach we consider using the same parameters P for the whole set of time series, i.e., in each plot of the type of Fig. 1b the estimated LLE's were obtained using the same numerical parameters. The only varying parameter is the parameter α from Eq. (4), used in generating the series. Then, we are not interested in absolute values of estimated LLE's, but in relative quantification of different series. In this case, the results can be considered as

[1] Thus we concentrate to the chaotic dynamics in the y-direction, which is equivalent to a one-dimensional system known as the tilted tent map (Hilborn, 1994).

[2] This may lead to a subjective bias and false positive results. Even from white-noise data any positive value of the λ_1 estimate may be obtained by tuning the parameters P (Dämmig and Mitschke, 1993).

successful, if a similar curve as that in Fig. 1a was obtained, irrespectively of a scale on the ordinate. The principal shape of the theoretical curve $\lambda_1(\alpha)$ is reproduced by the λ_1 estimates in Fig. 1b, however, the curve is not smooth due to numerical instability[3] of the estimates. Fluctuations of the estimates occur due to a relatively short time series length (1k=1024 samples) used. For a significant decrease of the fluctuations and obtaining smooth curves resembling the theoretical one (Fig. 1a) the series length must be increased by one or two orders of magnitude. We will, however, continue the study using 1k series and consider the results in Fig. 1b as a "good" classification considering "available" amount of data. It should also be noted that the results

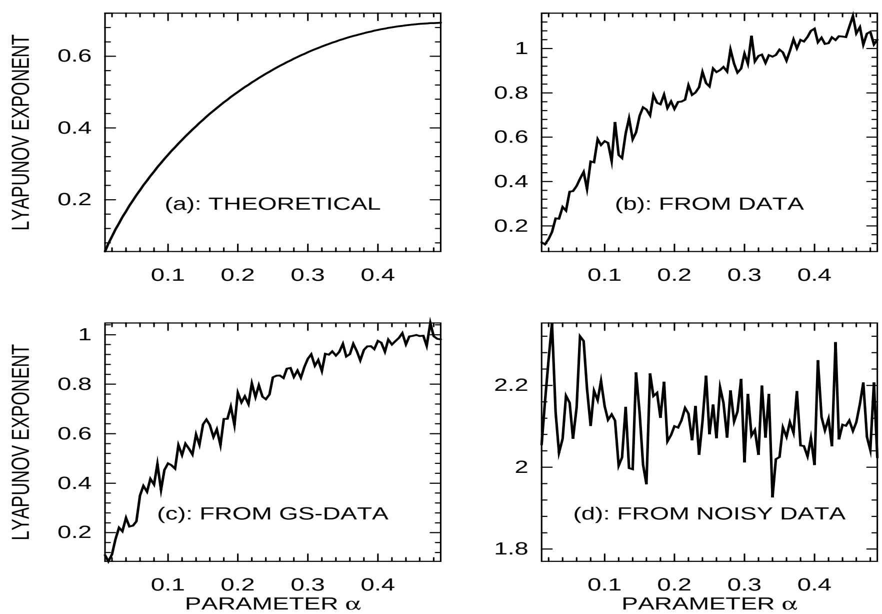

Fig. 1. *Positive Lyapunov exponents of the baker system as functions of the parameter α. (a) Theoretical dependence $\lambda_1(\alpha)$, (b) estimates from noise-free data, (c) estimates from noise-free gaussianized data, (d) estimates from noised data (10% of noise). The parameters, used in estimations (b, c, d) are: m=2, $\tau = 2$, $T = 1$, s_{min} =0.01SD, s_{max} is always $s_{max} = 10 s_{min}$.*

presented below were obtained from (noisy) baker series which underwent

[3] To decrease fluctuations of the λ_1 estimates due to local properties of time series, as a final estimate we used the mean value from the last third of all iterates – see Fig. 5.

so called gaussianization (Paluš, 1995) – a nonlinear transformation which transformed the marginal distribution of the data into a normal distribution. The reason for this transformation is comparison of the results from the "experimental" data with the results from surrogate data, described below. Although the gaussianization has some influence on the estimated λ_1 values, principal dynamical properties and related classification of the series were not changed – cf. Figs. 1b and 1c, the former was obtained from the original baker series, the latter from the same series after the gaussianization, using the same parameters P.

The situation dramatically changed when 10% of noise was added to the baker series and λ_1 estimations were repeated using the same parameters P as in the case of Figs. 1b,c. In this case the LLE algorithm failed to distinguish the series, it yielded random values irrelevant to the actual dynamical properties of the data (Fig. 1d).

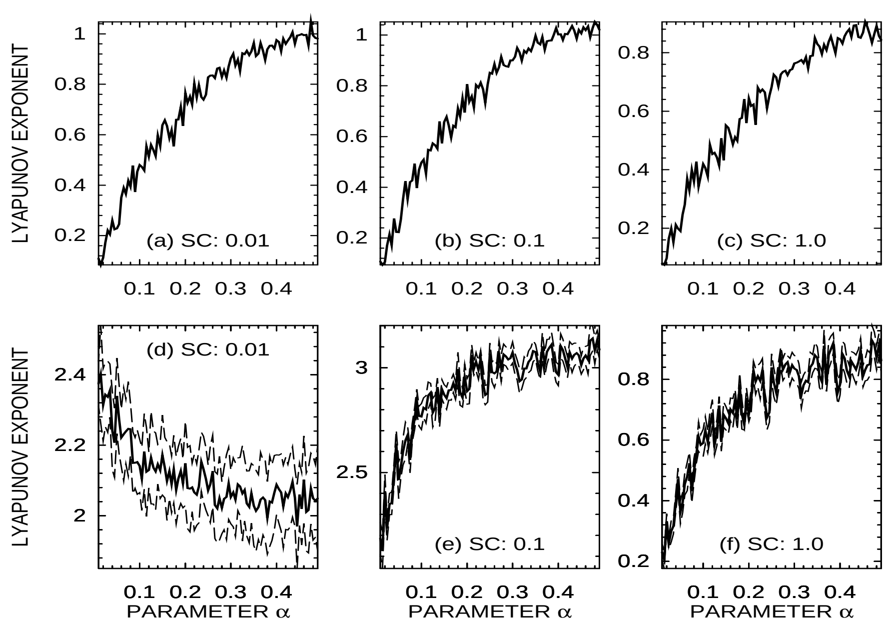

Fig. 2. *Estimates of the positive Lyapunov exponent from noise-free baker series (a, b, c) and their surrogate data (d, e, f), plotted as functions of the parameter α. In plots d-e-f solid lines and dashed lines depict mean λ_1 and mean$\pm$SD, respectively, of 15 realizations of the surrogates for each value of α. The scales s_{min} =0.01SD (a, d), s_{min} =0.1SD (b, e), and s_{min} =1.0SD (c, f) were used. The parameters m=2, $\tau = 2$, $T = 1$ were used in all estimations.*

In the following we compare the λ_1 estimates obtained from the baker series with different portions of noise, using different scales s_{min} ($s_{max} = 10s_{min}$). The values of the parameters m, τ, T are the same as above. The results for the noise-free data are presented in Fig. 2, the minimum scales are s_{min}=0.01SD (Fig. 2a), 0.1SD (Fig. 2b) and 1.0SD (Fig. 2c). The largest Lyapunov exponents λ_1, estimated from the noise-free low-dimensional chaotic series, are stable with respect to different scales (cf. Figs. 2a and 2b), only in the case of the largest scales (Fig. 2c) the estimates have lower values and the curve $\lambda_1(\alpha)$ is partially distorted, but still able to classify the series in the relative sense. With 5% of noise in the data the classification is practically impossible for s_{min}=0.01SD (Fig. 3a), possible, though with a higher error rate for s_{min}=0.1SD (Fig. 3b), while for s_{min}=1.0SD (Fig. 3c) the results are almost as good as for the noise-free data. For the data with 10% of noise

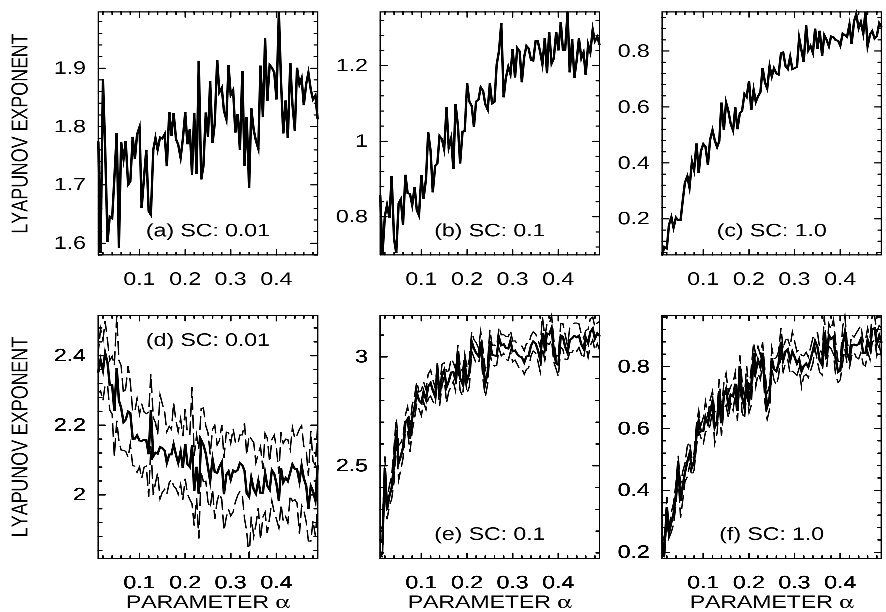

Fig. 3. *Estimates of the positive Lyapunov exponent from noisy (5% of noise) baker series (a, b, c) and their surrogate data (d, e, f), plotted as functions of the parameter α. In plots d-e-f solid lines and dashed lines depict mean λ_1 and mean±SD, respectively, of 15 realizations of the surrogates for each value of α. The scales s_{min} =0.01SD (a, d), s_{min} =0.1SD (b, e), and s_{min} =1.0SD (c, f) were used. The parameters m=2, $\tau = 2$, $T = 1$ were used in all estimations.*

(Fig. 4), the classification is impossible for both s_{min}=0.01SD (Fig. 4a) and

s_{min}=0.1SD (Fig. 4b), while the classification ability of the algorithm is restored using s_{min}=1.0SD (Fig. 4c). Using s_{min}=1.0SD we obtained the same results as in Fig. 4c also for 30% of noise in the data. (Data with higher portions of noise were not tested.) Thus, the generally known advice that the scales, used in estimating the chaotic measures, should be above the noise level, seems to be valid. Considering, however, that the chaotic measures are

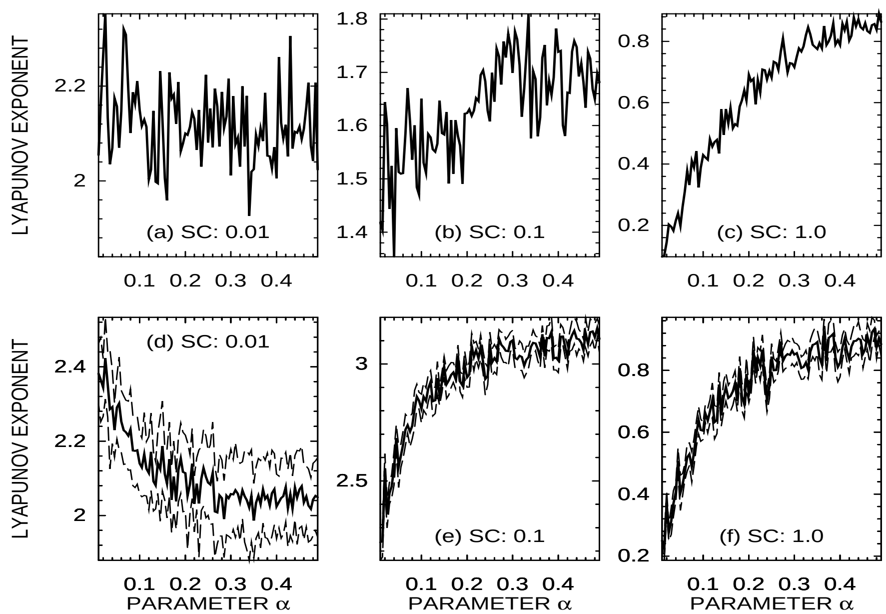

Fig. 4. *Estimates of the positive Lyapunov exponent from noisy (10% of noise) baker series (a, b, c) and their surrogate data (d, e, f), plotted as functions of the parameter α. In plots d-e-f solid lines and dashed lines depict mean λ_1 and mean±SD, respectively, of 15 realizations of the surrogates for each value of α. The scales s_{min} =0.01SD (a, d), s_{min} =0.1SD (b, e), and s_{min} =1.0SD (c, f) were used. The parameters m=2, $\tau = 2$, $T = 1$ were used in all estimations.*

defined in terms of vanishing distances between points, one could doubt what is actually measured using the large, macroscopic scales. In this study, is it really the exponential divergence of nearby trajectories, which is reflected in the results in plots b) and c) of Figs. 2–4? Searching for an answer, the technique of surrogate data (Theiler et al., 1992), (Paluš, 1995) was used. The surrogate data to an "observed" series are, in this case, realizations of a Gaussian linear stochastic process with the same spectrum as the "observed" series.

For each time series analyzed above, a set of 15 realizations of the surrogates were constructed and the largest Lyapunov exponents λ_1 were estimated using the same parameters P as for the λ_1 of the relevant "observed" series. The results from surrogates are presented in plots d, e, f of Figs. 2–4. Solid lines are used for mean λ_1, dashed lines depict mean±SD of λ_1 estimated from the set of 15 realizations of the surrogates.

Estimating LLE λ_1 from linear stochastic data one could ask whether such estimates converge. The positive answer is illustrated in Fig. 5, where the convergence of λ_1 estimates is presented for a noise-free baker series, generated with $\alpha = 0.3$ (upper panel in Fig. 5) and a realization of its surrogates (lower panel in Fig. 5).

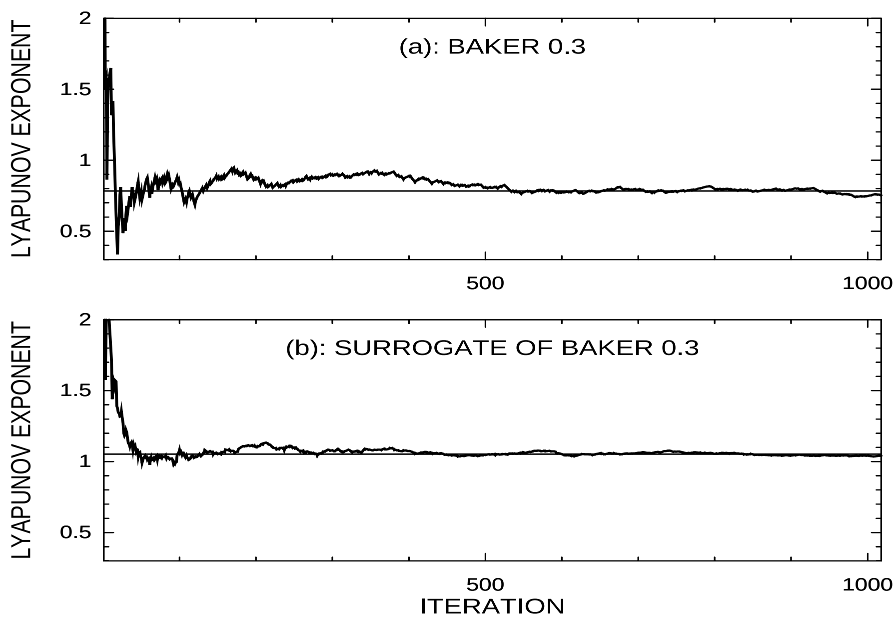

Fig. 5. *Convergence of estimates of the positive Lyapunov exponent in the course of averaging along the trajectory ("iteration") for a baker series generated with* $\alpha = 0.3$ *(a) and a realization of its surrogate data (b). Estimation parameters:* s_{min} *=0.1SD,* m*=2,* $\tau = 2$, $T = 1$. *The horizontal line presents the final estimate, defined as the average value of the last third of all iterations.*

Exploring relatively small scales (s_{min}=0.01SD, plots d in Figs. 2–4), LLE's λ_1 estimated from the surrogates do not reflect the "chaoticity", i.e., the dependence $\lambda_1(\alpha)$ of the original data. Such a result could be expected as far as the chaotic dynamics and nonlinear properties of the original data were destroyed by phase randomization in the surrogates. Using larger scales s_{min}=0.1SD and 1.0SD (plots e and f in Figs. 2–4), however, a relative classification, similar to the ordering of the baker series according to their λ_1, is again observed, though, in the surrogate data, there is no exponential divergence of trajectories, or even no trajectories in the deterministic sense! These time series are realizations of *Gaussian linear stochastic* processes, thus their dynamics are fully characterized by their power spectra or, equivalently, by their autocovariance functions. In this situation one can infer that the algorithm for the largest Lyapunov exponent distinguishes time series with different autocorrelation functions.

4 Discussion: From Chaotic to Stochastic Measures and Back

For understanding the results from the previous section we will briefly review relations between two kinds of dynamical measures of chaos — Lyapunov exponents and Kolmogorov-Sinai entropy (KSE), between KSE and mutual information (MI) and between MI and a standard linear statistical measure – the autocorrelation function.

Consider that a time series $x(t)$ started at time t_0 with an initial value $x(t_0)$. If the process underlying the series is not regular, but either stochastic, or chaotic and our knowledge about $x(t)$ is limited by finite precision measurement, after an evolution time $\tau = t_1 - t_0$ it is impossible to find an exact relation between $x(t_0)$ and $x(t_1)$. That is, knowing $x(t_0)$ one cannot exactly[4] predict $x(t_1)$, or, knowing $x(t_1)$ one cannot exactly compute backward the value of $x(t_0)$. In its evolution the underlying system is forgetting the information about its initial condition or, in other words, the system is creating new information, which has to be obtained by new measurement to know the state $x(t)$ of the system. The rate, how quickly the new information is created, is characterized by the *entropy rate* h of the system. When the underlying process is described by a dynamical system, the special case of the entropy rate h can be defined, known as the Kolmogorov-Sinai entropy (KSE). The famous theorem of Pesin (1977) says that the KSE h of a dynamical system is equal to the sum of its positive Lyapunov exponents. In the case of the baker map there is only one positive λ_1 and thus

$$h(\alpha) = \lambda_1(\alpha). \tag{8}$$

The mutual information $I(\tau) \equiv I[x(t); x(t+\tau)]$ (Shannon and Weaver, 1964), (Cover and Thomas, 1991), (Paluš, 1995), (Paluš, 1996a), (Pompe, this volume) quantifies the average amount of information about $x(t+\tau)$ that is contained in $x(t)$, and vice-versa. The rate of decrease of $I(\tau)$ with

[4] "Exactly" should be understood as "with precision comparable to the precision of measurement".

increasing τ should be related to the rate of the information creation of the underlying system, thus to the system's entropy rate h[5]. In the case of the baker system, where $h = \lambda_1$ increases with the parameter α according to Fig. 1a, the larger α is used for generating the series, the faster should be the decrease of the mutual information $I(\tau)$, estimated from that series. This behaviour is demonstrated in Fig. 6, where $I(\tau)$ is presented, extracted

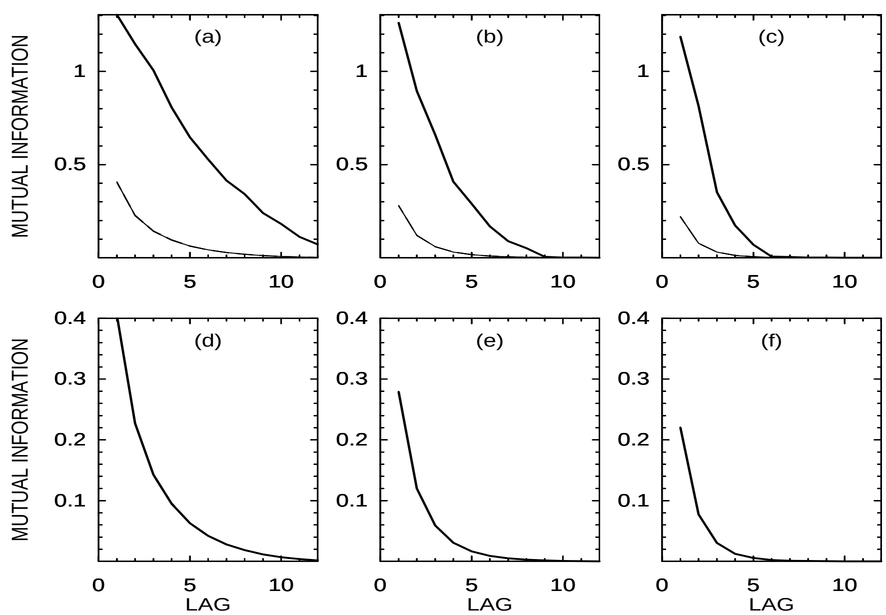

Fig. 6. *Mutual information $I(\tau)$ as a function of the lag τ for baker series and their surrogates, generated with $\alpha = 0.1$ (a, d), $\alpha = 0.2$ (b, e), and $\alpha = 0.3$ (c, f). In the plots a-b-c $I(\tau)$ from the baker data is plotted using thick lines, thin lines are used for $I(\tau)$ from the surrogates, which are plotted once more in the plots d-e-f, using the appropriate scale.*

from the baker series generated using $\alpha = 0.1$ (Fig. 6a), $\alpha = 0.2$ (Fig. 6b), and $\alpha = 0.3$ (Fig. 6c). The chaoticity of different baker series is reflected in the character of the τ-dependence of the mutual information $I(\tau)$ estimated from the baker series (thick lines in Figs. 6a-c), but surprisingly also in $I(\tau)$ estimated from related surrogate data (thinner, lower curves in Figs. 6a-c

[5] More details about the relation between higher-order mutual information (called marginal redundancy) and the entropy rates can be found in Refs. (Fraser, 1989), (Paluš, 1993), (Paluš, 1996b) and references therein.

illustrate mean values for the sets of 15 realizations of the surrogates). Thus the chaoticity of the baker series is not only encoded into their "non-linear properties", characterized by their $I(\tau)$, but also reflected into their "linear properties", which are preserved in the surrogates. $I(\tau)$ from the surrogates are once more displayed using an appropriate scale in Figs. 6d,e,f.

The surrogates are realizations of Gaussian linear stochastic processes, thus their mutual information $I(\tau)$ can be expressed as a function of their autocorrelation functions $C(\tau)$ (Morgera, 1985), (Paluš, 1995) as

$$I = -\frac{1}{2}\log(1 - C^2). \tag{9}$$

Then also the autocorrelation functions $C(\tau)$ (and spectra) of the baker series and their surrogates contain the information about the baker series' chaoticity (dependence $\lambda_1(\alpha)$). This explains why the Gaussian linear stochastic surrogates, related to different baker series, can be distinguished and ordered in the same way (in the relative sense) as the original baker series are classified according to their positive Lyapunov exponents. The question is, however, why this classification was possible to perform by using the Lyapunov exponent algorithm, designed to quantify the exponential divergence of nearby trajectories of chaotic systems.

The LLE algorithm explores changes of initial distances δ_I of pairs of points into final distances δ_F after an evolution time T. Consider a time series generated by white noise (independent identically distributed — IID process). For any initial distance δ_I, the final distance δ_F is a random number independent of δ_I. The averaged $< \delta_F >$ is then equal to the overall average distance of the data points. The averaged initial distance $< \delta_I >$ is influenced by the choice of the scales s_{min}, s_{max}. Then, choosing the scales so that $< \delta_I >$ is smaller than $< \delta_F >$, a positive estimate of λ_1 is obtained. When considered noise is not white but "coloured", i.e., there is some correlation $C(T)$ between $x(t)$ and $x(t+T)$, the increase of distance after the time T is smaller for series with stronger correlations, i.e., the larger $C(T)$, the smaller is the estimated λ_1, and vice versa.

Dämmig and Mitschke (1993) have derived analytic formulae for the λ_1 estimates when applying the considered LLE algorithm to white noise and a very special kind of coloured noise (white noise filtered by a "brickwall filter", the filter function is equal to one for a defined spectral bandwidth, and to zero otherwise). As one could expect, λ_1 estimated from white noise depends exclusively on the parameters P, in the case of the coloured noises λ_1 depends on P and on the spectral bandwidth. Thus for fixed P the estimated Lyapunov exponent λ_1 classifies the series according to their spectra, or, equivalently, according to their autocorrelation functions.

In the case studied here, where the coloured (autocorrelated) noises – the surrogate data – were generated according to given nontrivial spectra, derivation of an analytic formula is probably impossible, but the dependence of λ_1 on autocorrelations has been demonstrated in the presented numerical study.

Consider next the results in Figs. 2b and 2e, where λ_1 was estimated from the baker series and their surrogates using the same scale s_{min} =0.1SD. We can see that the stronger is the dependence between $x(t)$ and $x(t+T)$, the smaller is the estimate of λ_1. The differences between the strength of the dependence between $x(t)$ and $x(t+T)$, measured by $I(T)$, in the original baker series and their surrogates can be observed in Figs. 6a-c. The estimates of λ_1 for the baker series reach values between 0.1 and 1.1 (Fig. 2b), while in the case of surrogates they are between 2.1 and 3.2 (Fig. 2e). Above it has been found that in linear stochastic series the estimates of λ_1 are determined by the series' autocorrelation function, here it can be inferred that in general nonlinear series the estimates of λ_1 are determined by general (linear + nonlinear) temporal dependences in the series, which can be measured, e.g., by the mutual information $I(\tau)$.

Using the largest scales $s_{min} = 1.0SD$, however, the λ_1 estimates obtained from the baker data and the surrogates are not significantly different (plots c and f in Figs. 2-4). In these scales dynamical properties of the series and the λ_1 estimates are dominated by linear properties of the series. Considering this result as a formal surrogate-based test for nonlinearity (Theiler et al., 1992), (Paluš, 1995), the Lyapunov exponent as a statistic fails to distinguish the nonlinear chaotic baker series from their isospectral linear stochastic surrogates. Using smaller scales ($s_{min} = 0.1SD$, Figs. 2b and 2e), however, the differences between the baker series and their surrogates are statistically significant, though in both cases (the data and the surrogates), equivalent relative classifications of the series were observed.

The direct comparison of the values of λ_1 estimated from the baker series and from the surrogates was possible due to using the methodology of the surrogate-based tests for nonlinearity. The surrogate data had the same linear properties (spectra, autocorrelations) as the original baker series, and also the same marginal histograms, which also influence the estimates of chaotic and other measures. The latter was achieved by the "gaussianization" – a histogram transformation which transformed the marginal distributions of the baker series into the Gaussian distribution. The surrogates were Gaussian by the construction.

An equivalent approach is using the original baker data without transformation, but transforming the surrogates from Gaussian into the distribution of the original data. Using this approach a shift in scales was observed: For s_{min} =1.0SD, the estimates of λ_1 were negative and irrelevant to the actual chaoticity of the baker series (i.e., the average initial distance, given by s_{min}, was already larger than the overall average distance). Then, the results of this approach for scales s_{min} =0.1SD and 0.01SD were equivalent to the results from the former approach using s_{min} =1.0SD and 0.1SD, respectively, while the scale $s_{min} = 0.001$SD was the first "non-macroscopic" scale, in which the surrogates were not classified according to the $\lambda_1(\alpha)$ scheme, i.e., this result is equivalent to the result from the scale $s_{min} = 0.01$SD in the former approach. Using the first or the other approach, the differences in the results are of a technical level, but the main messages of this study, formulated in the Conclusion (items 1 and 2), are not changed.

5 Conclusion: From Stochastic to Chaotic Measures and Back

The findings of this study can be summarized as follows:

1. The surrogates of the baker series can be relatively classified according to the $\lambda_1(\alpha)$ scheme (Fig. 1a). Because of the linear stochastic nature of these processes, this classification must be accessible using linear techniques and, consequently, the original chaotic baker series can be distinguished and ordered equivalently to the $\lambda_1(\alpha)$ ordering also by using linear statistical techniques. This result may hold also for other chaotic systems, but NOT generally for all nonlinear systems.

2. The classification of the linear stochastic surrogates was performed by using the algorithm for estimating the largest Lyapunov exponent. This finding may probably be generalized:[6] The chaotic measures may provide meaningful classification (relative characterization) even for linear stochastic data.

The relation between existence of chaos in a system underlying data and the ability of chaotic measures to classify different systems states is not straightforward. Linear techniques may be used successfully for some chaotic systems, while chaotic measures may give meaningful results for linear stochastic data. A successful application of a chaotic measure in relative characterization of system states does not necessarily imply chaos in the system.

About a decade ago the chaotic measures became frequently used in analysis of complex time series as an alternative to stochastic, mostly linear techniques. Deterministic chaos has been usually considered as an opposite alternative to random effects in attempts to explain complicated dynamics. Recent results indicate, however, that low-dimensional chaos may be

[6] Esp. for those chaotic measures which explore distributions of distances between points like the correlation dimension (Grassberger and Procaccia, 1983a), (Grassberger and Procaccia, 1983b). For example, in an EEG study it was observed that classification of EEG signals, obtained by using the correlation dimension, had been possible to reproduce by linear measures (Paluš et al., 1992).

rather a rare than ubiquitous phenomenon[7], or, the strict separation between deterministic-chaotic and stochastic dynamics may be impossible (Ellner and Turchin, 1995). And even in data generated by a low-dimensional chaotic system, microscopic properties, which are characterized by the chaotic measures, may be unaccessible due to finite precision and measurement noise, as demonstrated in this study. A more comprehensive approach to study real-world systems is emerging, based on mathematical theory of nonlinear stochastic systems. This approach offers data analysis methods that explicitly consider randomness and have a firm basis in statistical theory.

The entropy rate (Cover and Thomas, 1991), (Paluš, 1996b), i.e., the rate of information creation by a system, was the property which made possible to classify the above studied time series. The entropy rate can be defined for both chaotic and stochastic systems. Although the exact entropy rate of a continuous system may be unaccessible from data, there is always a possibility to estimate its "coarse-grained" versions, suitable for classification of system states. An example of such measures, applied to the same baker series, as considered here, is presented in Ref. (Paluš, 1996b). A comprehensive review of "complexity" measures, related to entropies and entropy rates, can be found in Ref. (Wackerbauer et al., 1994).

Creation of information by a system, characterized by its entropy rate, may be caused either by intrinsic dynamical noise, or by a system's "chaoticity" — sensitivity to initial conditions; or by a combination of the two. Detection and characterization of the "stochastic chaos", i.e., of the initial-condition sensitivity of nonlinear stochastic systems, is a task of immense importance in nonlinear time-series analysis. The sources of positive entropy rates can be distinguished neither by the measures of entropy rate, nor by the chaotic measures such as the Lyapunov exponents. Specific techniques, designed for nonlinear stochastic systems, should be used, such as the conditional mean/variance or conditional probability approaches, advocated by Yao & Tong (1994). An interesting overview of nonlinear time series analysis from a chaos perspective has been recently published by Tong (1995).

Acknowledgements

The author was supported by the Grant Agency of the Czech Republic (grant No. 201/94/1327) and by the Academy of Sciences of the Czech Republic (grant No. 230404). Majority of numerical experiments related to this study were performed using computational facilities of the Santa Fe Institute, supported by the core funding from the John D. and Catherine T. MacArthur Foundation, the National Science Foundation (PHY-8714918), and the U.S. Department of Energy (ER-FG05-88ER25054).

[7] Especially when considering open, real-world systems, such as those studied in physiology and medicine. For recent results on nonlinearity/chaos in EEG see Refs. (Theiler and Rapp, 1996), (Paluš, 1996c).

References

Cover T.M., Thomas J.A. (1991): *Elements of Information Theory.* (J. Wiley & Sons, New York)

Dämmig M., Mitschke F. (1993): Estimation of Lyapunov exponents from time series: the stochastic case. *Phys. Lett. A* **178** 385-394

Ellner S., Turchin P. (1995): Chaos in noisy world: new methods and evidence from time series analysis. *The American Naturalist* **145(3)** 343-375

Farmer J.D., Ott E., Yorke J.A. (1983): The dimension of chaotic attractors. *Physica D* **7** 153-180

Fraser A.M. (1989): Information and entropy in strange attractors. *IEEE Transactions on Information Theory* **35** 245-262

Grassberger P., Procaccia I. (1983a): On the characterization of strange attractors. *Phys.Rev.Lett.* **50** 346-349

Grassberger P., Procaccia I. (1983b): Measuring the strangeness of strange attractors. *Physica D* **9** 189-208

Hentschel H.G.E., Procaccia I. (1983): The infinite number of generalized dimensions of fractals and strange attractors. *Physica D* **8** 435-444

Hilborn R.C. (1994): *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers.* (Oxford University Press, Oxford, New York)

Koukkou M., Lehmann D., Wackermann J., Dvořák I., Henggeler B. (1993): Dimensional complexity of EEG brain mechanisms in untreated schizophrenia. *Biol. Psychiatry* **33** 397-407

Mayer-Kress G. (ed.) (1986): *Dimensions and Entropies in Chaotic Systems* (Springer, Berlin)

Mayer-Kress G., Layne S.P. (1987): Dimensionality of the human electroencephalogram, in: Mandell A.S., Koslow S. (eds.): *Perspectives in Biological Dynamics and Theoretical Medicine, Ann. N.Y. Acad. Sci.* **504** 62-87

Morgera S.D. (1985): Information theoretic covariance complexity and its relation to pattern recognition. *IEEE Trans. on Systems, Man, and Cybernetics* **SMC-15, 5** 608-619

Layne S.P. , Mayer-Kress G., Holzfuss J. (1986): Problems associated with dimensional analysis of electroencephalogram data In: Mayer-Kress G. (ed.): *Dimensions and Entropies in Chaotic Systems* (Springer, Berlin) p. 246.

Paluš M., Dvořák I., David I. (1992): Spatio-temporal dynamics of human EEG. *Physica A* **185** 433-438

Paluš M., Albrecht V., Dvořák I. (1993): Information-theoretic test for nonlinearity in time series. *Phys. Lett. A* **175** 203-209

Paluš M. (1993): Identifying and quantifying chaos by using information-theoretic functionals. In: Weigend A.S., Gershenfeld N.A. (eds.): *Time Series Prediction: Forecasting the Future and Understanding the Past.* Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XV (Addison–Wesley, Reading, Mass.) 387-413

Paluš M. (1995): Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D* **80** 186-205

Paluš M. (1996a): Detecting nonlinearity in multivariate time series. *Phys. Lett. A* **213** 138-147

Paluš M. (1996b): Coarse-grained entropy rates for characterization of complex time series, *Physica D* **93** 64-77

Paluš M. (1996c): Nonlinearity in Normal Human EEG: Cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.,* in press. Electronic preprint: `ftp://ftp.santafe.edu/pub/mp/eeg2/bcnpr.ps`.

Pesin Ya.B. (1977): Characteristic Lyapunov exponents and smooth ergodic theory, *Russian Math. Surveys* **32** (1977) 55-114.

Pompe B.: Ranking and entropy estimation in nonlinear time series analysis, this volume.

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T. (1986): *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge)

Schuster H.G. (1988): *Deterministic Chaos, An Introduction.* (Physik, Weinheim)

Shannon C.E., Weaver W. (1964): *The Mathematical Theory of Communication.* (University of Illinois Press, Urbana)

Takens F. (1981): Detecting strange attractors in turbulence. In: Rand D.A., Young L.-S. (eds.): *Dynamical systems and turbulence. Springer Lecture Notes in Mathematics* **898** 366-81

Theiler J., Eubank S., Longtin A., Galdrikian B., Farmer J.D. (1992): Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58** 77-94

Theiler J., Rapp P.E. (1996): Re-examination of the evidence for low-dimensional nonlinear structure in the human electroencephalogram. *EEG Clin. Neurophysiol.* **98** 213-222

Tong H. (1995): A personal overview of nonlinear time series analysis from a chaos perspective, *Scandinavian J. Statistics* **22(4)** 399-445

Wackerbauer R., Witt A., Atmanspacher H., Kurths J., Scheingraber H. (1994): A comparative classification of complexity measures. *Chaos, Solitons and Fractals* **4(1)** 133-173

Wackermann J., Lehmann D., Dvořák I., Michel C.M. (1993): Global dimensional complexity of multi-channel EEG indicates change of human brain functional state after a single dose of a nootropic drug. *EEG Clin. Neurophysiol.* **86** 193-198

Weigend A.S., Gershenfeld N.A. (eds.) (1993): *Time Series Prediction: Forecasting the Future and Understanding the Past.* Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XV (Addison–Wesley, Reading, Mass.)

Wolf A., Swift J.B., Swinney H.L., Vastano J.A. (1985): Determining Lyapunov exponents from a time series. *Physica D* **16** 285-317

Yao Q., Tong H. (1994): On prediction and chaos in stochastic systems. *Philos. Trans. Roy. Soc. (Lond.)* **A 348** 357-369