Physica D 55 (1992) 221–234
North-Holland

PHYSICA D

Singular-value decomposition in attractor reconstruction: pitfalls and precautions

Milan Paluš and Ivan Dvořák

Laboratory for Applied Mathematics and Bioengineering, Prague Psychiatric Centre, Ústavní 91, CS-181 03 Prague 8, Czechoslovakia

Received 23 April 1991
Revised manuscript received 27 August 1991
Accepted 8 September 1991
Communicated by A.V. Holden

Applicability of singular-value decomposition for reconstructing the strange attractor from one-dimensional chaotic time series, proposed by Broomhead and King, is extensively tested and discussed. Previously published doubts about its reliability are confirmed: singular-value decomposition, by nature a linear method, can bring distorted and misleading results when nonlinear structures are studied.

1. Introduction

Interpretation of irregular dynamics of various systems as a deterministic chaotic process is increasingly popular and widely used in almost all fields of science [1–4]. It is based on an idea that structurally complex systems can perform dynamics with only a few degrees of freedom reflecting a strange attractor in the system phase space. Takens' embedding theorem [5] offers the possibility of reconstructing n-dimensional dynamics from one system observable – measurable one-dimensional signal and thus to estimate dynamical invariants (e.g. dimensions, entropies, Lyapunov exponents [1, 6, 7]) from experimental data. Results of these analyses, giving finite and even low dimensions, are presented as evidence for deterministic chaotic nature of the examined dynamics and obtained values of dynamical invariants are employed for characterization of the system under study.

Algorithms used for estimations of dynamical invariants from experimental data [8–11] are extremely time-consuming and usually must be repeated many times. Broomhead and King [12] have proposed to use so-called singular system analysis – very efficient and at the first sight promising method able to give basic characterization of the studied dynamics and to make further analyses simpler and more purposive. Numerical experience, however, led several authors [13–15] to express some doubts about reliability of singular system analysis in the attractor reconstruction.

In this paper we explain why singular-value decomposition, the heart of the singular system analysis and by nature a linear method, may become misleading technique when used in nonlinear dynamics studies. Our considerations are illustrated by numerical examples in which chaotic data from Hénon [16, 1] and Lorenz [17, 1] attractors are subjected to singular-value decomposition and by simultaneous use [13] of

0167-2789/92/$05.00

singular-value decomposition and Grassberger–Procaccia [9, 10] algorithm for estimation of the correlation exponent (dimension).

In section 2 the basic theory of analysis of experimental data based on nonlinear dynamical systems and deterministic chaos theory ("nonlinear analysis") is surveyed. This is followed by description of the singular-value decomposition method and its supposed application in nonlinear analysis (section 3). Confusing numerical results of the latter and their origin – an attempt to install linear order in nonlinear world – are then presented and discussed in sections 4 and 5.

2. Dynamical systems with one observable

Let the system under study be described as the ergodic dynamical system:

$$\frac{\mathrm{d}y}{\mathrm{d}t} = F(y), \tag{1}$$

in an n-dimensional continuous phase space $\mathscr{S}$. If smooth vector field $F(y)$: $\mathscr{S} \to \mathscr{S}$ satisfies well-known conditions of the existence and uniqueness theorem [18, 19], then (1) defines an initial value problem: for any $y_0 \in \mathscr{S}$ there is the unique solution curve $y(t, y_0)$ passing through y_0 which can be formally written as $y(t, y_0) = \phi_t y_0$. Here ϕ_t represents one-parameter map family ϕ_t: $\mathscr{S} \to \mathscr{S}$, also called flow of vector field F.

If the system (1) is known, its dynamics can be characterized qualitatively by analysis of topological (or dynamical) invariants of the system phase portrait (such as singular trajectories or limit sets). In practice, however, the system under study gives usually one observable, i.e. the only information about the system is noisy one-dimensional signal sampled with a finite precision. We have knowledge neither about system equations nor about geometry of its phase portrait.

Suppose there is a smooth compact m-dimensional manifold $\mathscr{M} \subset \mathscr{S}$ such that:

(i) $\mathscr{M}$ is invariant in the sense: $\forall x \in \mathscr{M}$ and $\forall t > 0 \Rightarrow \phi_t x \in \mathscr{M}$;

(ii) $\mathscr{M}$ is attracting in the sense that the evolution curves $\phi_t x$ starting in almost all points outside $\mathscr{M}$ tend to $\mathscr{M}$ for $t \to \infty$;

(iii) smooth flow ϕ_t has an attractor within $\mathscr{M}$; and there is smooth function v: $\mathscr{M} \to \mathbb{R}^1$.

The according to Takens' theorem [5] it is a generic property that the map $\Phi_{F,v}(y)$: $\mathscr{M} \to \mathbb{R}^{2m+1}$ defined by

$$\Phi_{F,v}(y) = \big(v(y), v(\phi_1(y)), \ldots, v(\phi_{2m}(y))\big)^{\mathrm{T}} \tag{2}$$

is an embedding. ϕ_i denotes $\phi_{i\tau}$ – flow of the continuous vector field F, τ is a so-called time delay; for discrete dynamical system ϕ on $\mathscr{M}$ ϕ_i denotes ϕ^i; the upper index T denotes transposition. (Embedding is a smooth map Ψ from the manifold $\mathscr{M}$ to some metric space $\mathscr{U}$ such that its image $\Psi(\mathscr{M}) \subset \mathscr{U}$ is a smooth submanifold of $\mathscr{U}$ and that Ψ is a diffeomorphism between $\mathscr{M}$ and $\Psi(\mathscr{M})$.) Dimension $2m+1$ is, according to Whitney's theorem [20, 21], a sufficient embedding dimension for m-dimensional compact smooth manifold. Often some dimension k of $\mathbb{R}^k$, $m \leq k \leq 2m+1$, is a sufficient embedding dimension, depending on geometric complexity of the manifold.

Using the above embeddings, geometry and qualitative dynamical (topological) invariants can be estimated from one-dimensional experimental data because a diffeomorphism – topological equivalence – preserves all the topological invariants.

3. Singular-value decomposition

Let the experimental signal (the observable of the system under study) be recorded with sampling frequency f_s. Then time series $Y(I)$, $I = 1, 2, \ldots, N_0$ is obtained. The map into $\mathbb{R}^n$ – n-dimensional series X_J^i (superscript $i = 1, \ldots, n$ is the index determining the component of

n-dimensional space, subscript $J = 1, \ldots, N$ is the time index) is constructed according to Takens' time-delay method [5]:

$$X_1 = (Y(1), Y(1+d), \ldots, Y(1+(n-1)d))^{\mathrm{T}},$$
$$\ldots$$
$$X_L = (Y(L), Y(L+d), \ldots, Y(L+(n-1)d))^{\mathrm{T}},$$
$$\ldots$$
$$X_n = (Y(N), Y(N+d), \ldots, Y(N+(n-1)d))^{\mathrm{T}}, \quad (3)$$

where $d = \tau f_s$, τ is the chosen time delay, $N + (n-1)d = N_0$. The obtained $n \times \mathrm{N}$ matrix $\mathbf{X} = \{X_J^i\}$ (reconstructed n-dimensional trajectory) is called the *trajectory matrix*.

Let k be the smallest sufficient embedding dimension for the studied problem, i.e. the attractive manifold of the system under study can be embedded into $\mathbb{R}^k$ but not into $\mathbb{R}^{k-1}$. As we have stressed earlier, the dimension of the manifold is not known a priori, so that k is not known, either. In experimental practice – let us consider, for concreteness, estimation of the correlation exponent (CE) [9, 10] (see also remark at the end of this section) – maps (3) for increasing sequence of n are constructed and CE is estimated from them. For $n < k$ CE is underestimated and increasing with n, for $n \geq k$ CE saturates on the correct value. (Let us disregard practical numerical problems now.)

In order to avoid these extensive computations, Broomhead and King [12] proposed method for determination of k, or, more precisely, for estimation of the upper bound of sufficient embedding dimension, or the upper bound of dimension of the system under study. (For example, for the three-dimensional Lorenz system Broomhead and King obtained $k \leq 4$ [12].)

The basic idea of Broomhead and King is that the sufficient embedding dimension k is equal to the number of linearly independent vectors that can be obtained from the columns of trajectory matrix $\mathbf{X}$, hence to the rank of $\mathbf{X}$. Instead of determining the rank of $n \times N$ trajectory matrix $\mathbf{X}$ it is more convenient to determine the rank of the symmetric $n \times n$ matrix $\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}$, because $\mathrm{rank}(\mathbf{X}) = \mathrm{rank}(\mathbf{C})$ [12]. The elements of $\mathbf{C}$ are

$$C_{ij} = \frac{1}{N} \sum_{L=1}^{N} X_L^i X_L^j, \quad (4)$$

where $1/N$ is the proper normalization. The rank of a matrix can be determined by singular-value decomposition (SVD) [22, 12, 13 and references therein]: Let $\mathbf{U}$ and $\mathbf{V}$ be unitary matrices such that

$$\mathbf{C} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}},$$

where $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, σ_i are non-negative singular values (SVs) of $\mathbf{C}$ by convention given in descending order:

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0.$$

If $\mathrm{rank}(\mathbf{C}) = k < n$, then

$$\sigma_1 \geq \ldots \geq \sigma_k > \sigma_{k+1} = \ldots = \sigma_n = 0. \quad (5)$$

Conversely, if $\sigma_k > 0$ and $\sigma_{k+1} = 0$ then $\mathrm{rank}(\mathbf{C}) = k$. In case of a symmetric matrix $\mathbf{C}$ equality $\mathbf{V} = \mathbf{U}$ holds.

SVD provides a set of n singular vectors of the matrix $\mathbf{C}$ which conform the orthonormal bases of both the range (the singular vectors corresponding to the non-zero singular values) and the null space (the singular vectors corresponding to the zero singular values) of matrix $\mathbf{C}$.

In experimental practice with noised and finite-precision data, all singular values are shifted and are non-zero. Relation (5) takes the form [12, 13]

$$\sigma_1 \geq \ldots \geq \sigma_k \gg \sigma_{k+1} \geq \ldots \geq \sigma_n > 0 \quad (6)$$

and the image of the map (3) in $\mathbb{R}^n$ can be decomposed into two orthogonal parts: a deterministic subspace corresponding to σ_i for $i = 1, \ldots, k$ and a subspace dominated by noise cor-

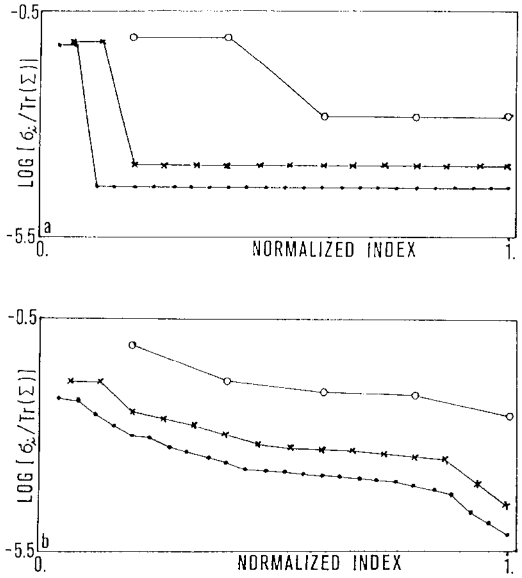

Fig. 1. Normalized singular values, plotted against the normalized index (the jth value is plotted at position j/n on the abscissa, n is the embedding dimension), obtained from 4096 data points of the noised sine signal (a) and 4096 data points generated by the Hénon map in the chaotic state (b). Embedding dimensions were 5 (open circles), 15 (crosses) and 25 (full circles).

responding to σ_i for $i = k+1, \ldots, n$; with relevant orthonormal bases of singular vectors [12, 13]. Values of σ_i, $i = k+1, \ldots, n$, in the so-called noise floor should *reflect the noise level in the data* [12, 13]. Hence one can estimate the rank of $\mathbf{C}$ as the number k of singular values greater than the noise level and sufficient embedding dimension should be equal to k or k should be the upper bound for sufficient embedding dimension.

Practically, maps (3) are constructed and their correlation matrices $\mathbf{C}$ decomposed for increasing dimension n. When sufficient embedding dimension is reached, number of singular values greater than the noise level should remain constant and all the residual singular values should fall onto the noise floor. For illustration the method was applied to the noised sine signal and results are presented in fig. 1a. The correct dimension two for embedding of a cycle was obtained as expected.

In summary, singular-value decomposition of the matrix $\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}$ should provide the finite upper bound k for sufficient embedding dimension and information about noise level, that is about a portion of the noise in the studied data. Moreover, it should offer the possibility of noise reduction: reconstructed n-dimensional trajectory (trajectory matrix $\mathbf{X}$) can be transformed into basis of singular vectors of $\mathbf{C}$. As this transformation is only a rotation it will not change the topological properties of the image of map (3). If relation (6) holds, elements W_j^i of trajectory in the basis of singular vectors for $i = k+1, \ldots, n$ could be discarded so that the embedding dimension is reduced and noise effects should be suppressed [13].

Remarks. In almost all papers concerning reconstruction of embeddings into $\mathbb{R}^n$ from one-dimensional series any map of type (3) is called embedding without considering whether its image is diffeomorphic to the original attractive manifold or not. There are no differences considered between the attractive manifold and the attractor. For simplicity we shall follow this convention as well.

In this paper we use the term "correlation exponent" instead of broadly used "correlation dimension". Determination of the correlation dimension, according to its definition [1, 7], requires a limit transition to "infinitely fine" space partition, while the correlation exponent characterizes the scaling, that is the macroscopic property of the signal probability distribution. What is really estimated by the Grassberger–Procaccia algorithm [9, 10] or its modification [24] is the correlation exponent not the correlation dimension.

4. Numerical results and their discussion

In our experiments we used chaotic series of the x-components generated by the Hénon map for parameters $a = 1.4$ and $b = 0.3$ (eq. (4) in ref. [16]) and by numerical integration of the Lorenz

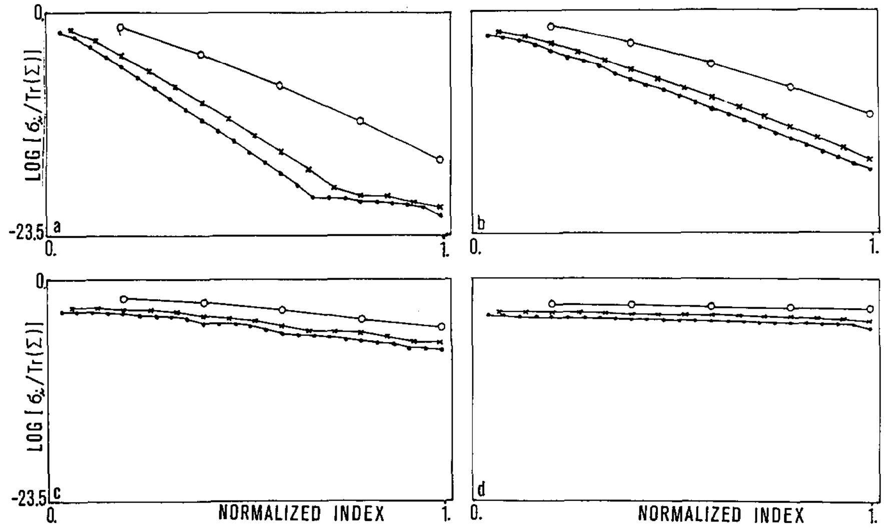

Fig. 2. The normalized singular spectra obtained from 4096 data points generated by the x component of the Lorenz system in chaotic state with integration step 0.02. Embedding dimensions used were 5 (open circles), 15 (crosses) and 25 (full circles). Time delays, used in embedding reconstructions, were $\tau = 0.02$ (a), 0.04 (b), 0.12 (c) and 0.4 (d).

equations for $\sigma = 10$, $r = 28$ and $b = 8/3$ (eqs. (25), (26), (27) in ref. [17]). The Bulirsh–Stoer algorithm [23, p. 563] for numerical solution of systems of ordinary differential equations was used. Integration steps in the Lorenz model were 0.0025 (results in figs. 3, 6 and 7b) and 0.02 (other results). Before calculating matrix **C** from matrix **X**, the series X_J^i were centered and scaled. i.e. instead of series X_J^i series $(X_J^i - a^i)/\delta^i$, where

$$a^i = \frac{1}{N} \sum_{J=1}^{N} X_J^i,$$

$$\delta^i = \left(\frac{1}{N} \sum_{J=1}^{N} \left(X_J^i - a^i \right)^2 \right)^{1/2},$$

were used. The topological structure is not changed by this transformation and **C** is now the correlation matrix of the reconstructed trajectory. Matrix **C** was decomposed using the SVDCMP routine according to ref. [23, p. 52]. Obtained singular values were plotted in conventional way: on ordinate, there are logarithms of normalized SVs: $\log[\sigma^i/\mathrm{Tr}(\Sigma)]$, on abscissa, there is a so-called normalized index: the first SV is plotted at position $1/n$ on abscissa, the second at $2/n, \ldots$, and the nth singular value at position 1. Finally, original reconstructed trajectories and those transformed into (a part of) basis of singular vectors of **C** were used for estimation of the correlation exponent using Grassberger–Procaccia algorithm [7, 9, 10], slightly modified according to Dvořák and Klaschka [24].

Computations of the presented singular spectra were performed with 4096 data points. Numerical experiments with data of 12288 points brought no significant differences in relevant singular spectra. Correlation exponents were estimated using time series of primary length 12288 points.

4.1. Upper bound for the embedding dimension

4.1.1. Numerical results

In fig. 1 data from two various two-dimensional systems are subjected to SVD. Results for noised sine signal (fig. 1a) are as expected: two singular values are greater than the noise level independently of the embedding dimension ($n = 5, 15, 25$). It does not hold for the discrete chaotic Hénon

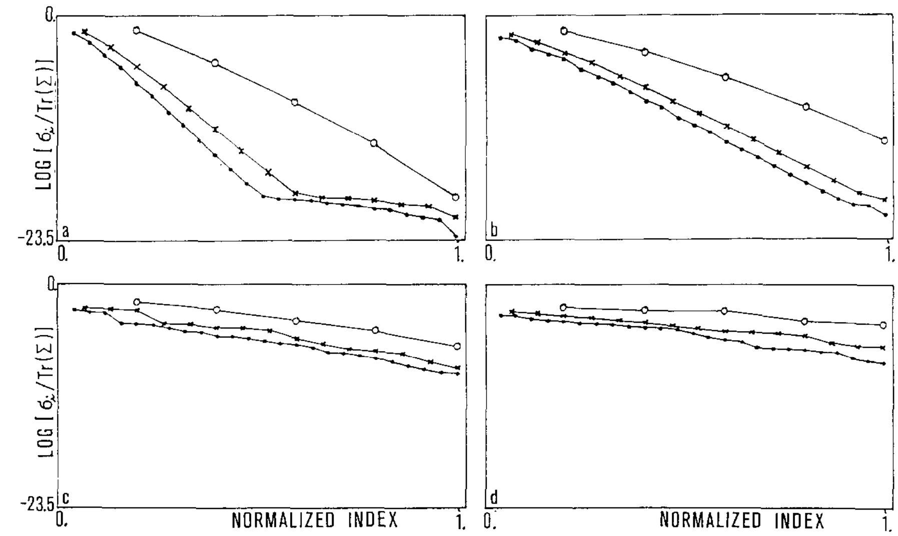

Fig. 3. The normalized singular spectra obtained from 4096 data points generated by the x-component of the Lorenz system in chaotic state with integration step 0.0025. Embedding dimensions used were 5 (open circles), 15 (crosses) and 25 (full circles). Time delays, used in embedding reconstructions, were $\tau = 0.02$ (a), 0.04 (b), 0.12 (c) and 0.4 (d).

map (fig. 1b): no noise floor is detected here; from the viewpoint of SVD it is classified either as a pure noise or as a deterministic system of high dimension and not as the two-dimensional system which it is actually.

Figs. 2 and 3 present results of SVD application to data from the continuous Lorenz system. The number of singular values greater than the noise level increases with increasing embedding dimension, $n = 5, 15, 25$, or, more frequently, no noise floor is detected. (Influence of time delay and integration step – "sampling frequency" – on singular spectra, depicted on the above figures, will be discussed further.)

Influences of data precision and of a defined amount of noise in the data on the singular spectra were studied next. The same data from the x-component of the Lorenz system were generated with various precisions: 6, 10, 14 and 18 bits, and 18-bit data were jammed by Gaussian noise so that contents of about 10% and 30% of noise in the data were obtained. Fig. 4a confirms the report of Mees et al. [13] that the higher the precision the more singular values are greater than noise level (noise floor is sinking with increasing precision of data). Also an increase of the amount of noise brings more singular values to fall onto the noise floor – see fig. 5a. It seems that the number of singular values over the noise floor depends on numerical accuracy, precision of the data and amount of noise in the data rather than on intrinsic dynamical structure of the system under study. (Influence of time delay τ on the singular spectra will be discussed further.) We should note, however, that each of the "upper bounds" (numbers of singular values above the noise floor) obtained is sufficient to embed faithfully the attractor. Since experimentalists need saturation of the value in question on well-defined and possibly the lowest upper bound in case the dimension of the studied system is unknown, SVD is not too reliable for this purpose.

4.1.2. Discussion of results

Let us try to explain the obtained results. Let k be the sufficient embedding dimension in previously defined sense, e.g. for data from the Lorenz system $k = 3$ (for proper time delay τ, see further). Let embedding dimension n be $n \geq 2k$. It can possibly occur that there are two subspaces of

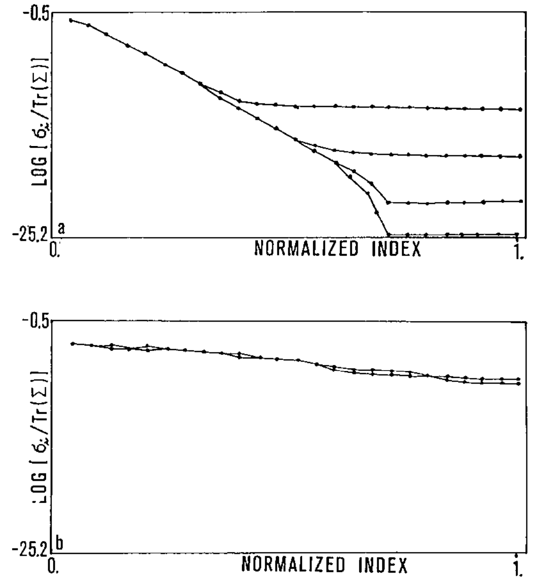

Fig. 4. The normalized singular spectra computed from 4096 data points generated by the x component of the Lorenz system in chaotic regime with integration step 0.02. In each case the embedding dimension was 25. The four different spectra were obtained from data of various precisions: 6, 10, 14 and 18 bits (reading from top to bottom). Time delays, used in embedding reconstructions, were $\tau = 0.02$ (a) and 0.12 (b). In case of $\tau = 0.12$ (b) the singular spectra are practically undistinguishable whatever the data precision is.

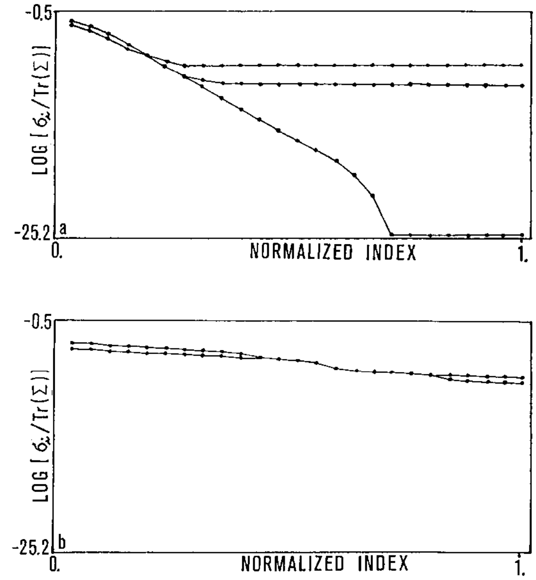

Fig. 5. The normalized singular spectra computed from 4096 data points generated by the x component of the Lorenz system in chaotic state with integration step 0.02 and recorded with precision 18 bits. In each case the embedding dimension was 25. The three different spectra were obtained from data jammed by different portions of the Gaussian noise: 30%, 10% and no noise (reading from top to bottom). Time delays, used in embedding reconstructions, were $\tau = 0.02$ (a) and 0.12 (b). In case of $\tau = 0.12$ (b) the singular spectra are practically undistinguishable whatever the portion of noise in data is.

$\mathbb{R}^n$ containing diffeomorphic images of the original attractive manifold. As a diffeomorphism is a relation of equivalence, these two images are mutually diffeomorphic, i.e. relations:

$$X^i = f^i(X^1, X^2, \ldots, X^k),$$
$$\text{for } i = k+1, \ldots, n \qquad (7)$$

and their inverses hold. (The superscript denotes again the component index, time index or argument is not used for simplification of formulas.) Diffeomorphic relations f^i need not be linear so that in embedding (3) more than k, possibly even n, linearly independent components can be found. As correlation is measure of a linear dependence [14, 22, 25] and correlation (or covariance) matrix **C** reflects the structure of this linear dependence, it is clear that the singular-value decomposition of **C**, extracting the number of linearly independent components in the embedding, cannot discover the true dimensionality of the system under study. Especially, as a one-dimensional chaotic time series has a finite correlation length τ_c (its autocorrelation function is usually exponentially decreasing [1]), any number of linearly independent delayed series can be constructed (up to some practical bound given by precision, sampling frequency etc.) using time delay $\tau > \tau_c$. As a consequence the SVD procedure gives unboundedly increasing number of singular values greater than the noise level or no noise level is obtained in singular spectrum at all.

4.1.3. *The role of the "window length"*

In all our experiments with SVD, except of those depicted in fig. 6, when we increased embedding dimension n, we kept *constant time delay* τ. Alternatively, Broomhead and King [12] pro-

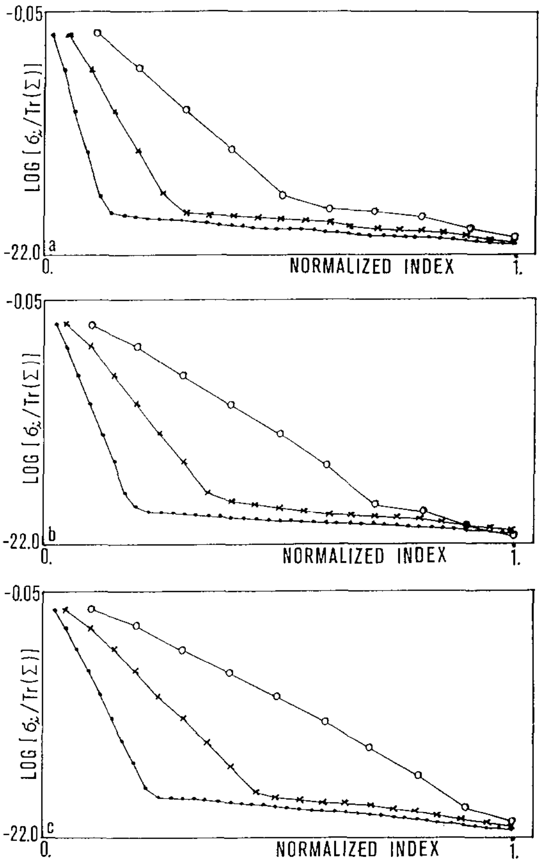

Fig. 6. The normalized singular spectra computed from 4096 data points generated by the x component of the Lorenz system in chaotic state with integration step 0.0025 and recorded with precision 20 bits. Embedding dimensions were $n = 10$ (open circles), 20 (crosses) and 40 (full circles). The constant "window lengths" $\tau_w = 0.1$ (a), 0.2 (b) and 0.3 (c) were used, i.e. time delay τ, used in particular embedding reconstruction, was $\tau = \tau_w / n$.

posed to keep *constant* a so-called *window length* $\tau_w = n\tau$, i.e. when we increase dimension n, delay τ must proportionally decrease. According to our experience this is exactly the way which can bring "false positive" results when SVD is applied: As a time series obtained from a continuous chaotic system has bounded but usually nonzero correlation length $0 < \tau_c \ll \infty$ [1], for given window length a limited number of linearly independent delayed series can be constructed and any other series from the window is correlated with previous ones. As a consequence, we can state the following conjecture: For any $k \in \mathbb{N}$ there is a window length $\tau_{w,k}$ so that keeping constant $\tau_{w,k}$ the number k of singular values greater than the noise level will remain constant for $n \geq k$. (There are again some numerical restrictions.) This phenomenon is illustrated in fig. 6. Using the same x series of the Lorenz system and keeping constant window lengths $\tau_{w,4} = 0.1$, $\tau_{w,6} = 0.2$ and $\tau_{w,8} = 0.3$, constant numbers of "deterministic" singular values $k = 4 - 5$, 6 and 8, respectively, were obtained.

There are examples quoted in ref. [12] where a proper choice of window length (around the correlation time of the signal) brings satisfactory results. These examples document that the constant window length – if ever used – should be chosen and used with utmost care.

4.2. *SVD and noise control*

4.2.1. *Numerical results*

Discovered dependence of the noise level in singular spectra on data precision (fig. 4a) and on the amount of noise in the data (fig. 5a) led some authors [13] to assume that singular-value decomposition could be used as an efficient tool for detecting the amount of noise in the studied data. Comparing parts (a) and (b) of figs. 4 and 5 we can see that the detected noise level and even its occurrence in a singular spectrum depends on time delay τ used in an embedding construction rather than on precision of the data or on the actual amount of noise in the data: Results presented in figs. 4b and 5b were obtained by computations in the same conditions as those in figs. 4a and 5a, respectively; except of time delay τ which is $\tau = 0.02$ for spectra given in figs. 4a and 5a and $\tau = 0.12$ for spectra in figs. 4b and 5b. The singular spectra in figs. 4b and 5b are practically the same whatever the precision of data or the amount of noise in the data are.

4.2.2. *Discussion of results*

Occurrence of the noise level in the singular spectrum of the correlation matrix of an embedding means that there are linearly dependent components in the embedding, or, in other words, that components of the embedding are correlated. Correlation of two identical but lagged series (which is the case in the Takens' embedding method) depends on lag τ and on the sampling rate of the signal (on the integration step in the generated data). Occurrence of a noise floor implies the possibility that a signal is oversampled and certainty that the lag τ is less than a "proper" time delay τ_p (see further) and/or the correlation length τ_c of the series. (Condition $\tau_p > \tau_c$ is not always necessary because even actual components of the system can be slightly correlated.)

The theory [5] gives broad possibility for setting the value of time delay τ in the embedding construction. When the embedding dimension is sufficient [20, 21] the maps of type (2) and (3) are (diffeomorphic) embeddings *generically*. This holds for function v, flow ϕ in (2) and lag τ in (3). In general, an object from some set or space of objects is *generic* if it is the element of the intersection of countably many open sets, dense in this space. In experimental practice, however, the situation is different and map (3) is (diffeomorphic) embedding usually for drastically reduced subset of theoretically accepted values of τ.

Based on numerical experience we could state that the time delay τ_p is "proper" when the image of the map (3) constructed using τ_p is diffeomorphic with the original geometry of the system attractive manifold and the embedding dimension is not greater than the smallest sufficient embedding dimension. Embeddings constructed with $\tau < \tau_p$ are wrong or improper in the following way: a diffeomorphic image of the system attractor is reached in $\mathbb{R}^n$ with dimension n higher than the embedding dimension, which is sufficient for τ_p and computations of e.g. correlation exponent are superfluously complicated. In the worst case, whatever the dimension n is, the components of map (3) are so strongly correlated that estimations of CE are heavily biased downwards or even tend to unity (fig. 7a). On the other

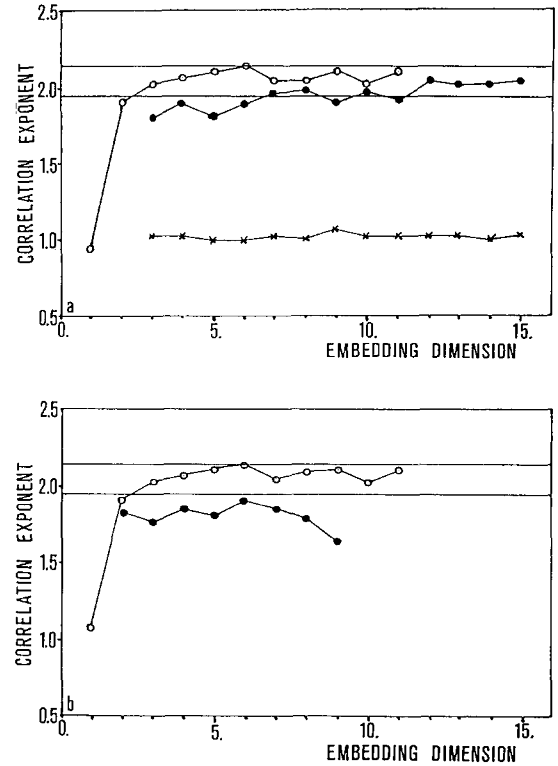

Fig. 7. Estimates of correlation exponent from 12288 data points of the x components of the Lorenz attractor dynamics plotted against the embedding dimension. (a) Dependence of the CE estimation on time delay τ, used in the embedding reconstruction, for $\tau = 0.12$ (open circles), 0.04 (full circles) and 0.02 (crosses). Data were generated with integration step 0.02 and recorded with precision 20 bits. (b) Dependence of the CE estimation on the integration step τ_i of the data generation ("sampling frequency") for $\tau_i = 0.02$ (open circles) and 0.0025 (full circles). Time delay $\tau = 0.12$. Horizontal lines at positions 1.95 and 2.15 on the ordinate are frontiers of 5% error of estimation.

hand, for τ greater than some "greatest proper τ_p" the estimates of CE are overestimated.

There are several methods for determining the proper time delay τ_p. The most promising, according to our experience, is the method of the first minimum of mutual information [26] giving for our Lorenz series $\tau_p = 0.12$.

Further study of the dependence of singular spectra on lag τ is depicted in figs. 2 and 3. Using $\tau_p = 0.12$ a noise floor disappears (figs. 2c and 3c). Comparing the numbers of singular values on the noise levels in figs. 2a and 3a an influence of the integration step (sampling frequency) on the singular spectra can be seen.

4.3. Test of the embeddings by CE computations

"Quality" of the above embeddings of the Lorenz attractor for various τ was studied using the Dvořák–Klaschka version [24] of the Grassberger–Procaccia algorithm [9, 10] for estimation of the correlation exponent (fig. 7). Using data with integration step 0.02 and $\tau_p = 0.12$ determined by the first minimum of mutual information [26], according to our expectation, embedding dimension $n = 3$ was found as sufficient and unbiased value CE = 2.03 [11] (with standard deviation less than 0.1) was obtained. For $n > 3$ CE saturates (with slight random oscillations) on this value. Using $\tau = 0.04$ CE is underestimated up to $n = 10$–12, when the correct value is reached. Delay $\tau = 0.02$, which seemed in previous computations so promising for the noise control, gives so strongly correlated components of the embedding that the value CE = 1 is estimated whatever the dimension n of the embedding is (fig. 7a). Using series with integration step 0.0025 and the same series length as previously ($N_0 = 12288$), even for $\tau_p = 0.12$ CE is underestimated, probably due to correlations caused by "oversampling" of the signal (fig. 7b).

4.4. SVD and noise reduction

4.4.1. Numerical experiment

Before further testing of the SVD applications, let us summarize the results of the above discussion: The aim of the attractor reconstruction from a one-dimensional time series is to reconstruct an uncorrelated embedding with the lowest possible embedding dimension. (Slightly correlated components are acceptable; a more general formulation could be "to obtain the highest information content in the lowest possible number of components".) Using the method of the first minimum of the mutual information [26] and its generalization [27] such embedding can be in principle constructed. It seems that using the singular-value decomposition loses its sense: with increasing embedding dimension n we find no upper bound for the number of linearly independent components. No noise floor need to be detected and a frontier between "deterministic" and "noise" components in a singular spectrum need not be found. So that *using SVD as a noise reduction technique* by rejecting the "noise" components of the embedding in the basis of singular vectors *is dubious*. But, in order to lead the discussion to the whole end, let us suppose that one can know the sufficient embedding dimension and wants to use SVD only for the noise reduction.

Let us estimate CE for the Lorenz data jammed with 10% of the Gaussian noise (with time delay $\tau_p = 0.12$). While for the same signal without noise starting with $n = 3$ CE saturates approximately on the expected value 2.03–2.08, presence of noise causes overestimation of CE: for $n = 3$ CE = 2.77 was obtained. (For $n = 4$ CE = 3.36, etc.) If we insist on strict application of the algorithm for CE estimation based on saturation of its values with increasing n, addition of noise blurs the saturation and CE – strictly speaking – is not defined. We can estimate, however, the value of CE for each n and with the aforementioned signals we can test the noise reduction ability of the singular-value decomposition.

Embedding n/k in the following means that an n-dimensional embedding was constructed and its $n \times n$ correlation matrix $\mathbf{C}$ was decomposed. The reconstructed trajectory was transformed into the basis of singular vectors of $\mathbf{C}$ ordered according to decreasing singular values. The first k components of the transformed trajectory (corresponding to the k largest singular values of $\mathbf{C}$) were used in CE estimations as proposed when using SVD as the noise-reduction technique [13]. With the noise level in the singular spectrum absent (fig. 2c) we tried to make use of the previous experience: using $\tau_p = 0.12$ for these data the three-dimensional embedding could be sufficient. Embedding $5/3$ gave CE = 2.59, for $7/3$

CE = 2.61 and for 9/3 CE = 2.43 were obtained. Overestimation was attenuated. Was it the effect of noise reduction or not? We repeated these computations for pure, not-noised data, which gave for $n=3$ CE = 2.03, $n=4$ CE = 2.07, etc., as discussed in section 4.3 (fig. 7, $\tau_p = 0.12$). Embedding 5/3 gave CE = 2.07, for 7/3 CE = 1.96 and for 9/3 CE = 1.77 were obtained. Results from the embedding sequence $n/3$ are depicted in fig. 8a. It can be seen that for $n > 6$ CE for embeddings $n/3$ are underestimated. It seems that SVD can reduce dynamical information rather than noise.

4.4.2. Discussion of results

In order to find a possible explanation of the above phenomenon let us recall relations (7) and propose the special form of them holds for some indexes $j \in [k+1, \ldots, n]$:

$$X^j = f^j(X^1, X^2).$$

After SVD the first three components in the singular-vector basis related to the three largest singular values can be written as:

$$W^1 = g^1(X^1, X^2),$$
$$W^2 = g^2(X^1, X^2),$$
$$W^3 = g^3(X^1, X^2), \qquad (8)$$

because singular values are given by the variance of the transformed components [22, p. 501] and do not reflect the nonlinear dynamical structure of the data. Functions g^i, $i = 1,2,3$, can be nonlinear and the three components (8) can be linearly independent. But, on the other hand, these three components can obtain only some two-dimensional projection of the Lorenz attractor. (Estimation of CE is 1.8.) It is clear that using all the transformed components full embedding (contained several times in the original embedded trajectory) must be obtained. In order to

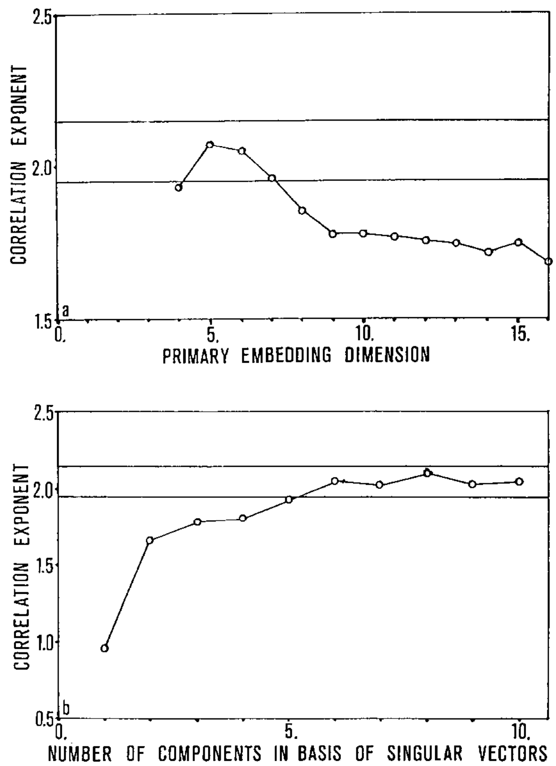

Fig. 8. Estimates of correlation exponent from 12288 data points of the x component of the Lorentz attractor dynamics, obtained from: (a) the first three components (related to the three largest singular values) of the embedding in the basis of singular vectors of the correlation matrix of primary (Takens' time delay) embedding, depicted against relevant primary embedding dimension; (b) the first k components (sorted according to descending singular values) of the embedding in the basis of singular vectors of correlation matrix of the primary 11-dimensional embedding, depicted against the number k of components used in the estimation. Each primary embedding was reconstructed from 20-bit data generated with integration step 0.02, time delay $\tau = 0.12$. Horizontal lines at positions 1.95 and 2.15 on the ordinate are frontiers of 5% error of estimation.

verify this, CE was computed for the sequence of embeddings $11/k$, $k = 1,2,\ldots,11$ (fig. 8b). Fig. 8b illustrates that "full" - diffeomorphic embedding is reached earlier - for $k = 6$.

As we mentioned above, this particular situation (no noise floor in the singular spectrum) is not the case for noise reduction as described in previous papers [13]. But this "enlarged" analysis is a good illustration of the "nonlinear phantom" haunting people accustomed to the linear world only. Even in case of noise floor presence the supra-noise (linearly independent) components could be created only from one true dynamical component which is mapped nonlinearly from one linearly independent component to another. And other true dynamical components (together with their nonlinear "ghosts") could be suppressed onto the noise floor due to lower variance.

In order to bring more support for these considerations, we studied dependence structure of components of the embedding of the Lorenz attractor in the basis of singular vectors obtained by SVD of the correlation matrix of primary 11-dimensional embedding ($\tau_p = 0.12$, as above). The correlation function and mutual information [26, 28, 29] were calculated. The components of the singular basis were ordered, as usual, according to descending singular values.

The mutual information $I(x, y)$ is, in general, a measure of the stochastic dependence of two random variables x, y. Roughly speaking, it can be taken as a generalization of the correlation function to nonlinear dependence. (For more details see refs. [26, 28, 29].) Mutual information (MI) is non-negative, zero is equivalent to independent variables x, y. For a more vivid representation of the results correlation functions in fig. 9 are depicted in absolute values.

Correlation functions (CF) and mutual information (MI) of couples of the series are displayed as functions of lag τ, $0 \leq \tau \leq 2.0$, albeit values of CF and MI for $\tau = 0$ are of main interest here.

Figs. 9a and 9b reflect the dependence of the first and the second, and the first and the third singular basis components, relevant to the three largest singular values of the correlation matrix of the original embedding. Correlation functions for $\tau = 0$ are due to orthogonality of the basis close to zero. But it does not hold for mutual information between the components (figs. 9a and 9b), which are significantly greater than zero, i.e. the components are *not independent*. The dependence level between the first and the sixth components (see fig. 9c) is much lower. This corresponds to the above findings that the first six components provide again the diffeomorphic embedding of the attractor (see fig. 8b). (Here we must remark that neither for the ideal embedding one can expect fully independent components, because even the original components of the system under study are tied by the particular trajectory.)

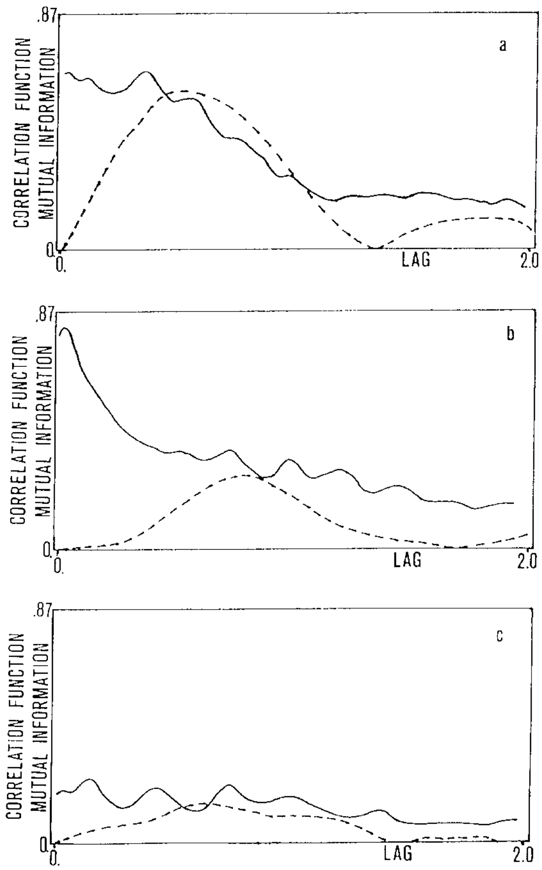

Fig. 9. Mutual information (full lines) and absolute values of correlation functions (dashed lines) of the first and the second (a), of the first and the third (b) and of the first and the sixth components of the embedding in the basis of singular vectors (sorted according to descending singular values) of the correlation matrix of the primary 11-dimensional embedding of the Lorenz attractor, reconstructed by the same way as the caption in fig. 8 describes. The scales of CF and MI are identical accidentally.

5. Concluding discussion

Reliability of singular-value decomposition in reconstructing a strange attractor from a one-dimensional chaotic time series was extensively tested. SVD, an effective method in discovering the linear structures in the data under study, can be misleading when nonlinear structures are of interest. Hence it can be hardly useful for determining an upper bound for the embedding dimension of the system attractor, or even dimensionality of the studied nonlinear system. Its role in noise control is questionable, too, because the so-called noise level and even its occurrence in singular spectra depends on the sampling rate of the series and on the time delay used in the embedding reconstruction drastically more than on the precision of the measurement and the actual amount of noise in the data. Components of the orthonormal basis given by SVD are linearly independent but can be strongly related in a nonlinear sense. As a consequence, the "most important" components, related to the largest singular values, can contain less dynamical information than the same number of components given by simple time delay method.

Presented criticism of SVD is not meant to refuse SVD technique as a whole or even denounce authors recommending application of SVD in nonlinear analysis. Though we in principle agree with the majority of statements in ref. [13], we have no intention to evoke a discussion of the type "you criticize claims we never made" like in ref. [30, 31]. But the results presented in our detailed analysis are warning to experimentalists, trying to find dynamical structures or even to detect strange attractor in their data, to be precautious in SVD applications and especially in interpreting the results it yields.

While application of SVD in reconstructing an attractor from a one-dimensional time series by the time-delay method is questionable, it may be of great importance when multi-dimensional time series are registered [32]. Neither in this case must we forget that nonlinearities in the data can distort the results.

Nonlinear relations in dynamical data are usually equivalent to the curvature of the relevant attractive manifold. However, a manifold with nonzero curvature is locally flat. This fact was the inspiration for applying SVD locally, i.e. on a subset of trajectory confined to a neighborhood of selected points [33, 34]. At the first sight this approach, in general not a new one (using local decomposition for estimating intrinsic dimensionality of data was proposed by Fukunaga and Olsen [35] twenty years ago for general data without any concept of dynamical systems), seems to be promising even for determining the actual dimension of the attractive manifold. But, analyzing experimental data, many questions appear, like "is the local subset local enough?", or "what is the noise effect in local case?", etc. Therefore several authors try to use local SVD together with some statistical or informational criteria for estimating the number of deterministic components in the singular spectra [36, 37].

SVD is surely an established method of signal analysis, but it should be applied with caution. Also results it provides must be interpreted with care. Improper use of singular-value decomposition (like that of Grassberger–Procaccia or any other algorithm) can bring misleading results interpreted like a (false) discovery of a strange attractor in any data. We are afraid that at least a part of recently published papers, alleging detection of chaos, are based on misinterpreted results, which cause "inflation" of the scientific value of this approach and jeopardize the prestige of deterministic chaos theory for experimental data analysis.

Acknowledgement

The authors thank V. Albrecht and L. Pecen for stimulating discussions and M. Pipíška for assistance with figures.

References

[1] H.G. Schuster, Deterministic Chaos: An Introduction, (Physik, Weinheim, 1984).

[2] P. Schuster, ed., Stochastic Phenomena and Chaotic Behaviour in Complex Systems, Springer Series in Synergetics, Vol. 21 (Springer, Berlin, 1984).

[3] P.E. Kloeden and A.I. Mees, Chaotic phenomena, Bull. Math. Biol. 47 (1985) 697.

[4] P. Cvitanovic, ed., Universality in Chaos (Hilger, Bristol, 1984).

[5] F. Takens, Detecting strange attractors in turbulence, in: Dynamical systems and turbulence, Warwick, 1980, Lecture Notes in Mathematics No. 898, eds. D.A. Rand and D.S. Young (Springer, Berlin, 1981) p. 366.

[6] G. Mayer-Kress, ed., Dimensions and Entropies in Chaotic Systems (Springer, Berlin, 1986).

[7] P. Grassberger and I. Procaccia, Dimensions and entropies of strange attractors from a fluctuating dynamics approach, Physica D 13 (1984) 34.

[8] P. Grassberger and I. Procaccia, Estimation of the Kolmogorov entropy from a chaotic signal, Phys. Rev. A 28 (1983) 2591.

[9] P. Grassberger and I. Procaccia, On the characterization of strange attractors, Phys. Rev. Lett. 50 (1983) 346.

[10] P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors, Physica D 9 (1983) 189.

[11] P. Grassberger, Estimating the fractal dimensions and entropies of strange attractors, in: Chaos, ed. A.V. Holden (Manchester Univ. Press, Manchester, 1986) p. 291.

[12] D.S. Broomhead and G.P. King, Qualitative dynamics from experimental data, Physica D 20 (1986) 217.

[13] A.I. Mees, P.E. Rapp and L.S. Jennings, Singular-value decomposition and embedding dimension, Phys. Rev. A 36 (1987) 340.

[14] A.M. Fraser, Reconstructing attractors from scalar time series: A comparison of singular system and redundancy criteria, Physica D 34 (1989) 391.

[15] A. Brandstater, H.L. Swinney and G.T. Chapman, Characterizing turbulent channel flow, in: Dimensions and Entropies in Chaotic Systems, ed. G. Mayer-Kress (Springer, Berlin, 1986) p. 150.

[16] M. Henon, A two-dimensional mapping with strange attractor, Comm. Math. Phys. 50 (1976). 69.

[17] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (1963) 130.

[18] E. Kamke, Differentialgleichungen Lösungsmethoden und Lösungen (Leipzig, 1959) (in Russian: Nauka, Moscow, 1971).

[19] V.I. Arnold, Ordinary Differential Equations (MIT Press, Cambridge, 1973).

[20] H. Whitney, Differentiable manifolds, Ann. Math. 37 (1936) 645.

[21] J. Palis Jr. and W. de Melo, Geometric Theory of Dynamical Systems, An Introduction (Springer, New York, 1982).

[22] C.R. Rao, Linear Statistical Inference and Its Applications, (Wiley, New York, 1965).

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

[24] I. Dvořák and J. Klaschka, Modification of the Grassberger–Procaccia algorithm for estimating the correlation exponent of chaotic systems with high embedding dimension, Phys. Lett. A 145 (1990) 225.

[25] M.M. Rao, Harmonizable, Cramer and Karhunen classes of processes, Handbook of Statistics, Vol. 5, eds. E.J. Hannan, P.R. Krishnaiah and M.M. Rao (Elsevier, Amsterdam, 1985) p. 279.

[26] A.M. Fraser and H.L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33 (1986) 1134.

[27] A.M. Fraser, Information and entropy in strange attractors, IEEE Trans. Information Theory 35 (1989) 245.

[28] R.G. Gallager, Information Theory and Reliable Communication (Wiley, New York, 1968).

[29] S. Kullback, Information Theory and Statistics (Wiley, New York, 1959).

[30] D.S. Broomhead, R. Jones and G.P. King, Comment on "Singular-value decomposition and embedding dimension", Phys. Rev. A 37 (1988) 5004.

[31] A.I. Mees and P.E. Rapp, Reply to "Comment on 'Singular-value decomposition and embedding dimension'", Phys. Rev. A 37 (1988) 5006.

[32] M. Paluš, I. Dvořák and I. David, Remarks on spatial and temporal dynamics of EEG, in: Mathematical Approaches to Brain Functioning Diagnostics, eds. I. Dvořák and A.V. Holden (Manchester Univ. Press, Manchester, 1991) p. 369.

[33] D.S. Broomhead, R. Jones and G.P. King, Topological dimension and local coordinates from time series data, J. Phys. A 20 (1987) L563.

[34] E.R. Pike, Singular system analysis of time series data, in: Computational Systems – Natural and Artificial, ed. H. Haken (Springer, Berlin, 1987) p. 86.

[35] K. Fukunaga and D.R. Olsen, An algorithm for finding intrinsic dimensionality of data, IEEE Trans. Computers C-20 (1971) 176.

[36] A. Goel, S.S. Rao and A. Passamante, Estimating local intrinsic dimensionality using thresholding techniques, in: Measures of Complexity and Chaos, eds. N.B. Abraham, A.M. Albano, A. Passamante and P.E. Rapp (Plenum Press, New York, 1989) p. 125.

[37] A. Passamante, T. Hediger and M.E. Farrell, Analysis of local space/time statistics and dimensions of attractors using singular value decomposition and information theoretic criteria, in: Measures of Complexity and Chaos, eds. N.B. Abraham, A.M. Albano, A. Passamante and P.E. Rapp (Plenum Press, New York, 1989) p. 173.