PHYSICAL REVIEW E VOLUME 50, NUMBER 3 SEPTEMBER 1994

Failure of dimension analysis in a simple five-dimensional system

Andreas Jedynak,[1] Michael Bach,[1,*] and Jens Timmer[2]
[1] *Universitäts-Augenklinik, Killianstrasse 5, D-79106 Freiburg, Germany*
[2] *Fakultät für Physik, Hermann-Herder-Strasse 3, D-79104 Freiburg, Germany*
(Received 18 October 1993; revised manuscript received 9 May 1994)

Dynamic systems may be characterized by their fractal dimension. The classical Grassberger-Procaccia algorithm is widely used to analyze time series. However, if this method is used beyond its intrinsic limitations it may cause incorrect classification of systems. We found that a simple deterministic five-dimensional system leads to erroneous dimension values around 5.5 if the following methods are used uncritically: The classical Grassberger-Procaccia algorithm, a pointwise correlation dimension algorithm, and an algorithm for calculation of the information dimension yielded this erroneous result for a wide range of numbers of data points ($N=30\,000-10^6$) and various delay times. Estimates of dimensions are only reliable if long plateaus of the local slope of the correlation integrals exist for small distances; these were not found in our example. This example suggests that a correlation dimension of 5 is too high to be recognized using even one million noise-free data points.

PACS number(s): 05.45.+b

I. INTRODUCTION

Since their introduction, correlation dimension values that indicate deterministic chaos have been found everywhere (EKG, climate, cries of babies, search behavior of ants, etc.). In human EEG's, correlation dimensions from 2, in cases of epilepsy, up to 11, for β EEG, have been published [1–4]. For α EEG's, values for the correlation dimension between 4 and 7 are typically reported. We failed to reproduce these findings in the α EEG of 14 subjects. In each case, we found insufficient scaling regions to estimate a valid correlation dimension, considering the improvements and avoiding the pitfalls of dimension analysis [5–8]. Possibly the typical sample sizes of EEG's ($\approx 4\,000-30\,000$) are too small to make dimension analysis feasible. Currently the minimal number of data points (N) necessary to estimate the correlation dimension is under debate. Smith [9] suggested that the minimum N is 42^D. This bound is far more demanding than has been realized in previous EEG analyses. Ruelle [10] proposed a minimum N to estimate the slope from correlation integrals of $D \le 2\log_{10}(N)$; this less demanding criterion is met in some EEG analyses.

To test whether the demands on N are justified and to investigate further possible pitfalls of dimension analysis, we studied a very simple five-dimensional system, the dimension of which is analytically known. We will show that data sizes of $N=1000-10^6$ used in dimension analysis are prone to yield erroneous correlation dimensions, which cannot even serve as approximations of the correct values. We will further present criteria to avoid such pitfalls.

II. RECONSTRUCTION PROCEDURE AND SCALING BEHAVIOR

We use the well-known time delay reconstruction from a scalar time series $x(t)$ [11] sampled with a sampling interval Δt. The vectors in the phase space $\mathbb{P}^d$ are constructed via the delay time τ:

$$\mathbf{x}_t := \{x(t), x(t+\tau), \ldots, x(t+(d-1)\tau\} \in \mathbb{P}^d . \quad (1)$$

The sampling time of the reconstructed trajectory $\mathbf{x}_t$ in the reconstructed phase space $\mathbb{P}^d$ is the same as in the original time series: $\ldots, \mathbf{x}_{t-\Delta t}, \mathbf{x}_t, \mathbf{x}_{t+\Delta t}, \ldots$.

In the case of a low-dimensional system the reconstructed trajectory $\mathbf{x}_t$ does not intersect itself with increasing embedding dimension d and occupies an invariant set, a subset of the embedding space. According to Grassberger and Procaccia [12] the correlation dimension D_2 is the number to characterize the scaling law of the number $\mathcal{N}$ of vector pairs $(\mathbf{x}_i;\mathbf{x}_j)$ within a distance r:

$$\mathcal{N}\{(\mathbf{x}_i;\mathbf{x}_j) \mid \|\mathbf{x}_i - \mathbf{x}_j\| \le r\} \propto r^{D_2} . \quad (2)$$

This number is calculated via the correlation integral $C_d(r)$:

$$C_d(r) := \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) ,$$

$$\text{with } \Theta(a) := \begin{cases} 1 & \text{for } a \ge 0 \\ 0 & \text{for } a < 0 \end{cases} , \quad (3)$$

and the correlation dimension is obtained by

$$D_2 := \lim_{r \to 0} \frac{d \ln[C_d(r)]}{d \ln(r)} . \quad (4)$$

As the reconstructed invariant set lies in a submanifold of the embedding space, it does not fill the whole embedding space. Therefore, dimension analysis does work reli-

*Author to whom correspondence should be addressed. Electronic address: bach@sun1.ruf.uni-freiburg.de

1063-651X/94/50(3)/1770(11)/$06.00

ably for low-dimensional systems embedded in very high-dimensional spaces.

In order to obtain a reliable reconstruction of the invariant set, it is necessary to include points from time intervals which are long compared to typical time scales of the process. Each region of the set should be passed sufficiently often.

In the case of high-dimensional systems the density of data points in the reconstructed phase space $\mathbb{P}^d$ is too low. This effect may be responsible for the failure of dimension analysis. Neighboring vector pairs are rare and the scaling behavior for $r \rightarrow 0$ is not observable. We stress the limit $r \rightarrow 0$ in the formula above. Experimental time series are sampled with discrete values in time and in amplitude. Typically 12-bit analog-to-digital converters (ADC's) are used to digitize continuous time series. Simulating a 12-bit ADC ($0 \leq r < 4096$) the limit $r \rightarrow 0$ means that the scaling region $[r_1; r_2]$ ought to start at small distances r in comparison to the diameter of the attractor; e.g., $r_1 < 0.05 r_{\max}$ and should exhibit one order of magnitude in length.

III. THREE METHODS TO CALCULATE DIMENSIONS

First we propose the basic concepts of dimension calculations and discuss computational details of the calculation procedures afterwards. There are several algorithms for dimension analysis. Basically, the algorithms may be subdivided in fixed-mass and fixed-state algorithms.

Fixed-size methods keeps the distances r fixed while counting the number of pairs $(\mathbf{x}_i; \mathbf{x}_j)$ closer than r. The frequency of distances r is calculated for a predefined set (for a 12-bit ADC, the set of possible distances using the maximum norm is $\{r | 0 \leq r \leq 4096$ and r integer$\}$). Estimates for the correlation integrals for these distances r are $C_d(r)$. The correlation integral is a function of the distance r and depends on the embedding dimension.

On the other hand, one can keep the mass k (number of points) constant and ask for the distance r, which is necessary to cover the nearest k points around a reference point $r_i(k)$. The average of the distances $r_i(k)$ of a set of L reference points yields $\langle r(k) \rangle$: $= \sum_{i=1}^{L} r_i(k) / -L$. The result is a function $\langle r(k) \rangle$ of the mass k. Following Badii and Politi [13], the information dimension D_1 is defined by

$$D_1 := -\lim_{k \rightarrow 1} \frac{\ln(N)}{\langle \ln[r(k)] \rangle} \quad \text{with } N \rightarrow \infty \ . \tag{5}$$

Here we use the method of Termonia and Alexandrowicz [14] to determine D_1 with a fixed number of data points N from the scaling behavior of the averaged logarithmic distances $\langle \ln[r(k)] \rangle$ of the k nearest neighbors.

We tested two fixed-size and one fixed-mass method to estimate dimension values. The pointwise correlation dimensions [15], calculated with a fixed-size method, might be of interest to analyse the local structure of an attractor and the whole spectrum of dimensions [16].

IV. COMPUTATIONAL DETAILS OF THE CALCULATION PROCEDURES

A. Classical Grassberger-Procaccia method

The classical Grassberger-Procaccia method uses all pairs $(\mathbf{x}_i; \mathbf{x}_j)$ with $i < j$ [12,17]. A modification of the correlation integral $C_d(r)$ by introducing a minimal time distance μ ($\mu \geq \tau$) as suggested by Theiler [18] avoids the influence of temporal neighbors of the vectors $\mathbf{x}_t$,

$$\tilde{C}_d(r) := \frac{2}{(N-\mu)(N-\mu-1)} \times \sum_{i=1}^{N-\mu-1} \sum_{j=i+\mu+1}^{N} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \ , \tag{6}$$

since the normalizing factor is of no interest:

$$C_d(r) := \sum_{i=1}^{N-\mu-1} \sum_{j=i+\mu+1}^{N} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \ . \tag{7}$$

As the values of the time series are integers, we used a standard technique to fill the histograms $H_d(r)$ where $0 \leq r \leq 4096$ for $d = 1, \ldots, 30$: for the maximum norm $\| \ \|_\infty$ of the distances of $(\mathbf{x}_i; \mathbf{x}_j)$ the maximal value is 4096. For each pair $(\mathbf{x}_i; \mathbf{x}_j)$ the distances for every d, using the result of $\| \ \|_{\infty, d'}$ for $\| \ \|_{\infty, d''}$, $d' < d''$, can be calculated for every $d = 1, \ldots, 30$ nearly as fast as calculating the distance for the greatest $d = 30$ only,

$$\begin{aligned} H_d(r) &:= \sum_{i=1}^{N-\mu-1} \sum_{j=i+\mu+1}^{N} \delta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \ , \\ C_d(r) &:= \sum_{r'=0}^{r} H_d(r') \ . \end{aligned} \tag{8}$$

The most time consuming task is filling the histograms $H_d(r)$. The most critical task is to derive the slope of the correlation integral in a double logarithmic plot of $\log_{10}[C_d(r)]$ vs $\log_{10}(r)$. Often scaling regions are selected manually. If the scaling regions are made very short, the final choice of the slopes may be influenced by the experimenter's expectation. Selecting the scaling region manually may be done in correlation integrals of low-dimensional systems, where it is no problem to find large scaling regions (e.g., $\log_{10}(r_2) - \log_{10}(r_1) \geq 1$ in our case). Only local slopes $s(r)$ calculated over small ranges of r show variations or plateaus of the slope. The local slope can be derived by linear regression from a fixed number of successive values of the correlation integral (linear regression $[(\{\log_{10}(r_j); \ \log_{10}[C_d(r_j)]\} | r_j = r + j$ and $j = 1, \ldots, 50)] \rightarrow s_d(r)$) or a small range of the distance r, linear regression

$$[(\{\log_{10}(r_j); \log_{10}[C_d(r_j)]\} | \ \|r_j - r\| \leq \Delta \log_{10}(r) = 0.25)] \rightarrow s_d(r) \ .$$

We applied both techniques here: manual selection of the scaling region and calculation of local slopes.

B. Pointwise correlation dimension

In general the correlation dimension depends on the position $\mathbf{x}_i$ in phase space. A small subset of reference points $\mathbf{x}_i$ ($\approx$200–1000 out of 10^4–10^6 possible vectors $\mathbf{x}_j$) was used to calculate the pointwise correlation dimensions. The correlation integral for a single reference point $\mathbf{x}_i$ is defined by

$$\tilde{C}_{i,d}(r):=\frac{2}{N-2\mu-1}\sum_{j\wedge|i-j|>\mu}^{N}\Theta(r-\|\mathbf{x}_i-\mathbf{x}_j\|)\ . \tag{9}$$

Actually, we have used the histogram techniques mentioned above and discarded the normalization factors

$$H_{i,d}(r):=\sum_{j\wedge|i-j|>\mu}^{N}\delta(r-\|\mathbf{x}_i-\mathbf{x}_j\|)$$

and get (10)

$$C_{i,d}(r):=\sum_{r'=0}^{r}H_{i,d}(r')\ .$$

For each reference point $\mathbf{x}_i$ slopes $s_{i,d}$ were automatically fitted to the d correlation integrals $C_{i,d}(r)$, as discussed in Holzfuss and Mayer-Kress [3]. To determine pointwise correlation dimensions $D_{2,i}:=D_2(\mathbf{x}_i)$,

$$D_{2,i}:=\lim_{r\to 0}\frac{d\ln[C_{i,d}(r)]}{d\ln(r)}\ ; \tag{11}$$

the slopes $s_{i,d}$ must saturate as d increases. Only in this case can one assign a pointwise correlation dimension value at the point $\mathbf{x}_i$. The values of the slopes $s_{i,d}$ may fluctuate around a mean value $\langle s_i\rangle$ for the $d\in[d_s;d_{\max}]$ in the saturation region,

$$\hat{D}_{2,i}=\langle s_i\rangle:=\frac{1}{d_{\max}-d_s+1}\sum_{d=d_s}^{d_{\max}}s_{i,d}\ . \tag{12}$$

If the reconstructed attractor of the trajectory $\mathbf{x}(t)$ is homogeneous in $\mathbf{P}^d$, all the reference points will show the same value for the correlation dimension theoretically. But due to the fitting procedure and the finite data set a histogram of the pointwise correlation dimension values is obtained. The reference points need not be chosen at random. Walking along the trajectory with equidistant time distances between reference points, the new "time series" of pointwise correlation dimension should be constant if the attractor is homogeneous. For high-dimensional systems embedded in relatively low embedding dimensions, intersections of the trajectories in phase space $\mathbf{P}^d$ might be seen. These intersections produce artifacts in the correlation dimension estimates. For high embedding dimensions they will disappear.

Only for homogeneous attractors the slopes $s_{i,d}$ may be averaged over the reference points i and the standard deviation of $s_{i,d}$ may give an idea of the variation of the slope estimates $\langle s_d\rangle$ for embedding dimension d:

$$\langle s_d\rangle:=\frac{1}{N_{\text{ref}}}\sum_{i=1}^{N_{\text{ref}}}s_{i,d},\quad \hat{D}_2=\frac{1}{d_{\max}-d_s+1}\sum_{d=d_s}^{d_{\max}}\langle s_d\rangle\ , \tag{13}$$

$$\sigma^2_{\hat{D}_2,d}=\text{var}[\hat{D}_{2,d}]:=\frac{1}{N_{\text{ref}}-1}\sum_{i=1}^{N_{\text{ref}}}(s_{i,d}-\langle s_d\rangle)^2\ . \tag{14}$$

The pointwise correlation dimension is a powerful tool to find interesting phase space points and the fine structure of the attractor in long time series and in homogeneous attractors. Another advantage of the pointwise calculation is avoiding possible averaging of correlation integrals from different (reference) points. This averaging might destroy the scaling region of the averaged correlation integral if the (reference) points start their scaling regions at different distances r, as discussed in Holzfuss and Mayer-Kress [3].

C. D_1 calculation with a fixed-mass method

The information dimension D_1 is defined according to

$$D_1:=-\lim_{k\to 1}\frac{\ln(k/N)}{\langle\ln[r(k)]\rangle}\quad\text{with } N\to\infty\ . \tag{15}$$

Again the histogram technique was applied to determine the distances of all points from a reference point $\mathbf{x}_i$. 100 reference points are sufficient to estimate the mean distance $\langle r(k)\rangle$. The maximal embedding dimension was $d_{\max}=50$. Based on the histogram $H_{i,d}(r)$, it is easy to derive $r_{i,d}(k)$, the distance of the kth nearest neighbor of reference point $\mathbf{x}_i$ embedded in d dimensions.

Although the values of interest are those of small k, the distances of k nearest neighbors with $k=1,2,\ldots,100,110,\ldots,1000,1100,\ldots,10^6$ have been stored to analyze the scaling behavior when getting close to the diameter of the attractor. Averaging the logarithms of the distances $\log_{10}[r_{i,d}(k)]$ over the reference points $\mathbf{x}_i$ for every embedding dimension d yields the mean distance $\langle\log_{10}[r_d(k)]\rangle$ of the kth nearest neighbor.

The problems of determining the slopes are the same as those of the classical Grassberger-Procaccia method. Calculating the local slopes $s_d(r)$ is the only way to demonstrate the existence of a scaling region.

V. SIMPLE FIVE-DIMENSIONAL SYSTEM

The following very simple five-dimensional system avoids numerical complications and shows that dimension analysis fails for high system dimensions. If one uses complicated chaotic systems, e.g., the Mackey-Glass system, the true dimension values are not known analytically and the argumentation may be circular.

The simple five-dimensional system is built by summing five sine functions with incommensurable frequencies. The invariant set in phase space of this quasiperiodic process is a 5 torus. We used various frequencies and amplitudes with similar results. Here we discuss

$$x(t):=\sum_{i=1}^{5}A_i\sin\left[\frac{2\pi}{T_i}t\right] \tag{16}$$

with the relative amplitudes $A_1:=10$, $A_2:=5$, $A_3:=3$, $A_4:=4$, $A_5=6$, and the frequencies $\omega_1:=\sqrt{2}$, $\omega_2:=\sqrt{3}$, $\omega_3:=\sqrt{5}$, $\omega_4:=\sqrt{7}$, $\omega_5:=\sqrt{11}$, with corresponding periods $T_1:=4.442\ldots$, $T_2:=3.627\ldots$, $T_3:=2.809\ldots$, $T_4:=2.374\ldots$, $T_5:=1.894\ldots$. We chose a sampling time $\Delta t=0.2$. This sampling frequency is high, as the Nyquist rule would allow a sampling rate at

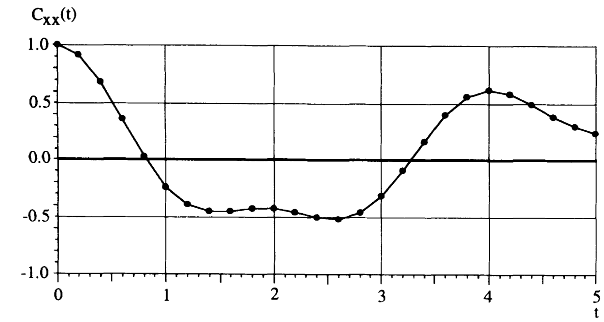

FIG. 1. Autocorrelation function of a simple five-dimensional system. The first zero crossing of the autocorrelation function is at $T=4\Delta t=0.8$. The autocorrelation function has no statistical meaning here, but is widely used to determine the delay time τ for dimension analysis.

$\Delta t \leq T_5/2 \approx 0.94$, but one should stay well below that. In the dimension analysis the data points can be selected in multiples of the sampling rate. Therefore, a fine sampling rate is useful as it allows various delay times τ. The minimal time distance μ was 10τ. To make these artificial time series comparable to experimental time series, expansion and discretization of the time series $x(t)$ according to a 12-bit analog-digital-conversion were applied. The new range of the integers $x(t)$ was $-2048 \leq x(t) \leq 2047$. There was no extra noise in the time series except discretization noise.

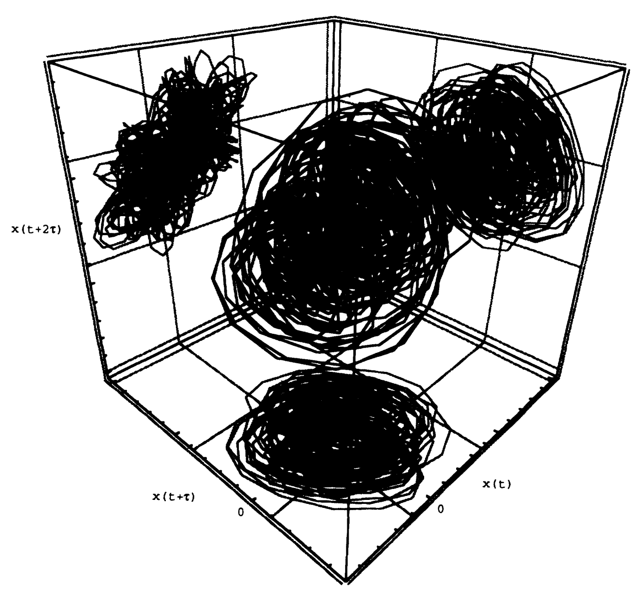

FIG. 2. Three-dimensional projection of a reconstructed trajectory of the five-dimensional system by means of delay coordinates $x(t)$, $x(t+\tau)$, $x(t+2\tau)$. The projection onto the $x(t)$-$x(t+\tau)$ plane looks complicated and the projection onto the $x(t)$-$x(t+2\tau)$ plane looks even "chaotic." The information content of those pictures is rather poor and one gets no idea of the dynamics of the process underlying the time series. Here 1000 data points of the time series of the five-dimensional system are plotted with $\tau=4\Delta t=0.8$.

VI. RESULTS

One of the difficulties in dimension analysis is choosing the delay time τ. The maximal volume expansion of the reconstructed attractor seems to be a good criterion for the choice of τ. In our case of oscillations, the first zero crossing of the autocorrelation function is a reasonable choice for τ. The autocorrelation function of our time series shows a first zero crossing at $T=0.8=4\Delta t$ (Fig. 1). The phase space portrait gives an impression of the complicated structure of the time series (Fig. 2). As the phase space portrait looks very complicated, it cannot be decided whether the time series is chaotic, regular, or stochastic.

Variation of the delay time τ $\{\tau=0.1;0.2;0.4;0.6;0.8;1.0\}$ ensured that our dimension analysis did not suffer from delay time artifacts. The local slopes are very similar for all delay times τ. As the results are not sensitive to variation of the delay time, we here present only the results for $\tau=0.8$.

Our simple five-dimensional system is stationary, which means that its dynamical properties do not change in time. Hence, an increase in N used for reconstruction of the invariant set in the phase space should improve the

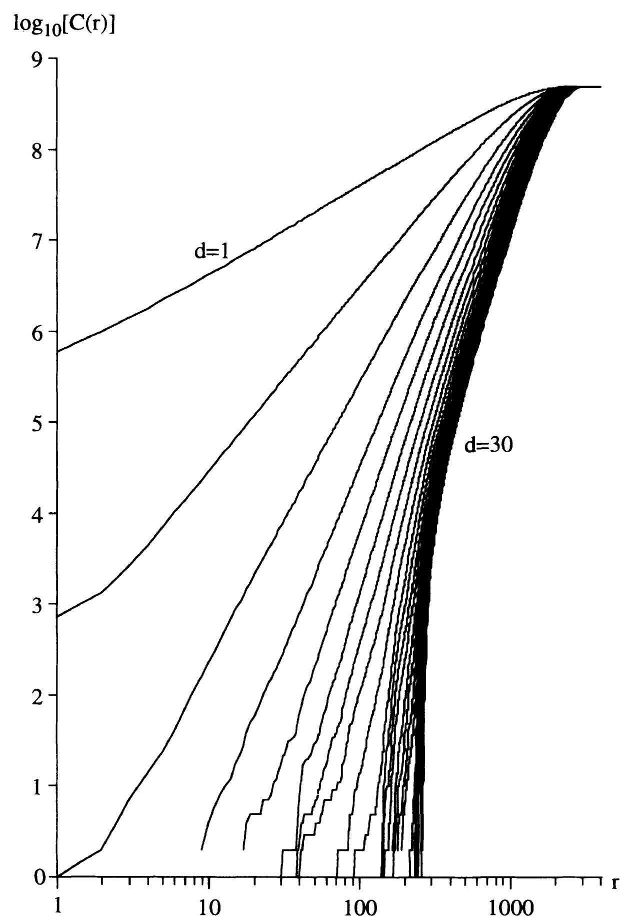

FIG. 3. Correlation integrals $C_d(r)$ of our five-dimensional system for embedding dimension $d=1,\ldots,30$. The number of data points was $N=32\,000$. For high embedding dimensions further increase of embedding dimensions does not change the correlation integrals $C_d(r)$ and saturation can be expected.

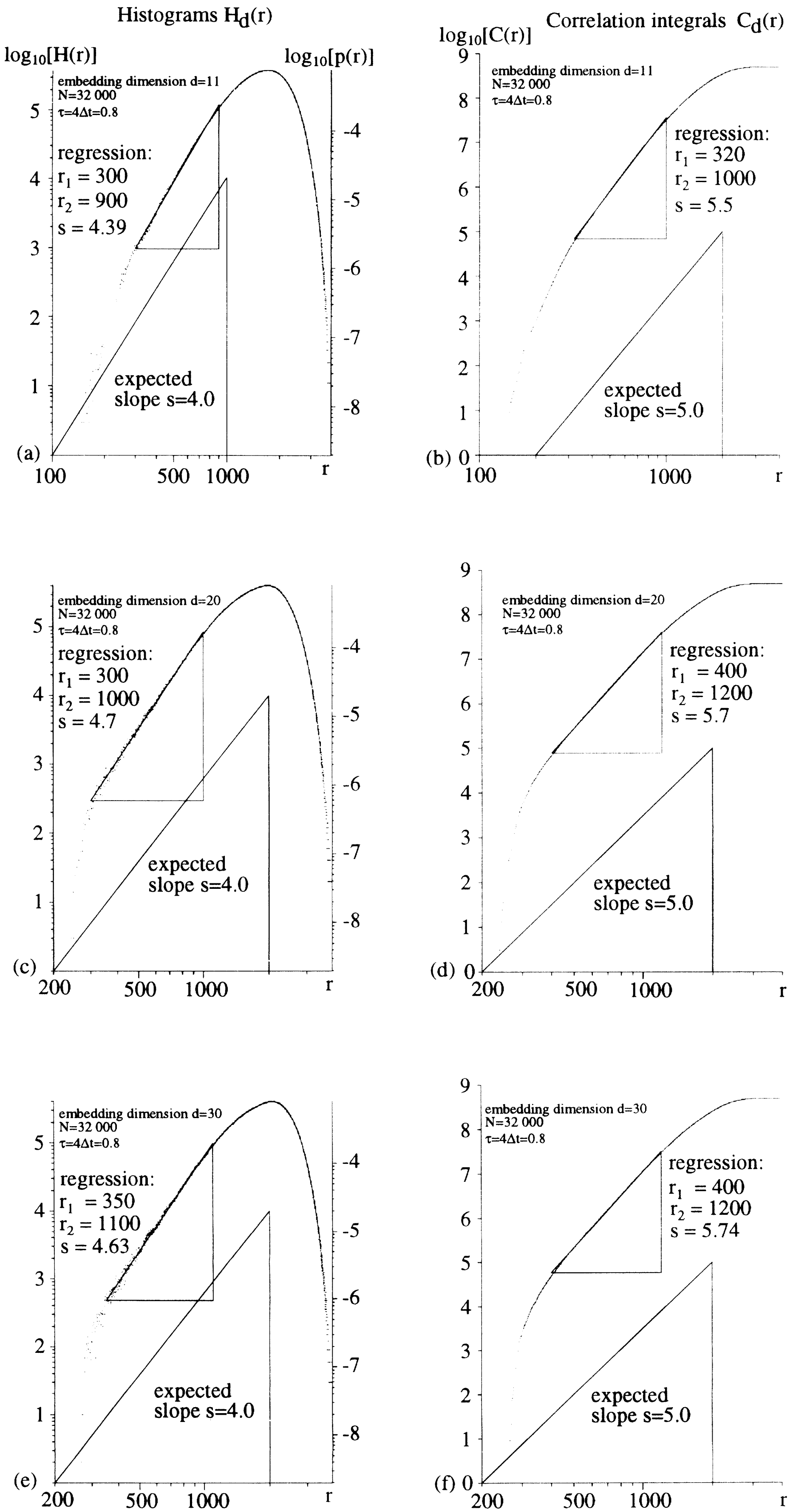

FIG. 4. Typical double logarithmic plots of correlation integrals and histograms for our five-dimensional system. The slopes of the manually selected scaling regions are all higher than $s=5.0$ and saturate around $s=5.7$.

estimates of the correlation integrals and therefore those of the correlation dimension. As the system is not chaotic but regular, the scaling regions of the correlation integrals start at $C_d(r) \approx O(N)$. This agrees with Theiler's results except for differing normalization [19].

A. Classical Grassberger-Procaccia method

For the classical Grassberger-Procaccia algorithm, we limited to $N = 32\,000$. To test the influence of N, it was varied: $N = 1000$, 2000, 4000, 8000, 16 000, and 32 000. We show the results for $N = 32\,000$ data points.

Figure 3 depicts the double logarithmic plot of the correlation integrals for embedding dimension $d = 1-30$. One might see a clear saturation of the slopes with increasing embedding dimension. A selection of several typical double logarithmic plots of correlation integrals [Figs. 4(a)–4(f)] shows the difficulties in selecting scaling regions for fitting a straight line in the double logarithmic plot of the correlation integral. In Fig. 4(b) it seems unproblematic for the embedding dimension $d = 11$. But the slope $s_{11} \approx 5.5$ is higher than the correct value of 5. For the embedding dimension $d = 20$ and $d = 30$ we found similar slopes near $s \approx 5.7$.

In addition to the correlation integrals $C_d(r)$, one can examine the histograms $H_d(r)$ directly. In double logarithmic plots of the histograms $\log_{10}[H_d(r)]$ vs $\log_{10}(r)$ the fluctuations of the density of the points of the reconstructed invariant set at the distance r [Fig. 4(a), 4(c), and 4(e)] are more pronounced. The manual selection of the scaling region is easier than in the $\log_{10}[C_d(r)]$ vs $\log_{10}(r)$ plots because fluctuations of $H_d(r)$ are larger than in the correlation integrals $C_d(r)$ and the maximum of $H_d(r)$ indicates that r is close to the diameter of the attractor. Here the results are very similar to those of the correlation integrals (Fig. 5).

To avoid manual selection of the scaling region, local slopes $s(r)$ can be calculated for short ranges of r and plotted against $\log_{10}(r)$. If they exhibit a plateau then there exists a scaling region. For the whole set of embedding dimensions $d = 1, \ldots, 30$ the plot of the local slope $s(r)$ vs $\log_{10}(r)$ indicates "evidence" of a chaotic system (Fig. 6). Obviously there is saturation but the "correlation dimension" is too high. The "scaling region" ($\approx [700;900]$), where the slopes $s_d(r)$ are nearly constant,

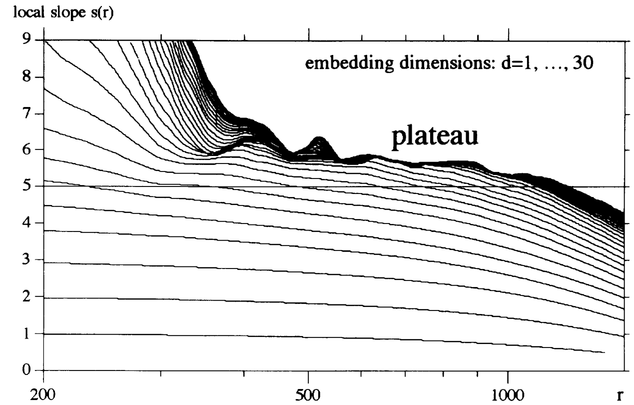

FIG. 6. The local slopes $s_d(r)$ show a plateau around 5.7. The length of the scaling region $\log_{10}(900) - \log_{10}(700) \approx 0.11 < 1$ is rather short and starts at large distances r. The local slopes $s_d(r)$ have been calculated from the correlation integrals in Fig. 3 by linear regression over 50 successive distance r_j: linear regression $[(\{\log_{10}(r_j);\ \log_{10}[C_d(r_j)]\}|r_j = r + j$ and $j = 1, \ldots, 50)] \rightarrow s_d(r)$.

is smaller than one order of magnitude: $\log_{10}(r_2) - \log_{10}(r_1) \approx 0.11$. The plateau is constant for 15 embedding dimensions ($d = 15, \ldots, 30$). Frequently, scaling regions comparable to ours are accepted [3,20–26].

B. Pointwise correlation dimension

For the estimation of pointwise correlation dimension only a few hundred reference points $\mathbf{x}_i$ (selected randomly or equitemporal on the reconstructed trajectory) with up to $N = 10^6$ other data points entered the calculations. For each reference point $\mathbf{x}_i$ the slopes $s_{i,d}$ for each embedding dimension d were fitted according to Holzfuss and Mayer-Kress [3]. If these slopes $s_{i,d}$ do show saturation (typical example in Fig. 7), the average of the slopes over the embedding dimensions d, where saturation is reached, is an estimate of pointwise correlation dimension. This average value $\langle s_i \rangle$ of the slopes $s_{i,d}$ is the estimate for the pointwise correlation dimension of the refer-

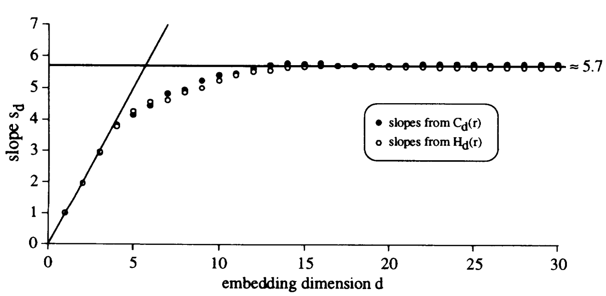

FIG. 5. The fitted slopes s_d of the manually selected scaling regions (several are shown in Fig. 4) saturate with increasing embedding dimension d. The estimate for the correlation dimension is $D_2 = 5.7$.

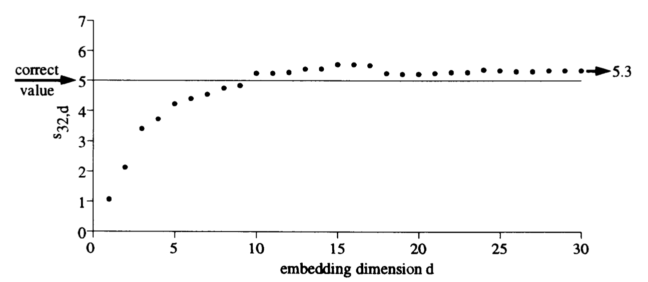

FIG. 7. Example of automatically fitted slopes $s_{32,d}$ vs embedding dimension d for a typical reference point ($i = 32$). From embedding dimension $d = 18-30$ the values of the slopes $s_{32,d}$ show saturation and the average of the slopes $s_{32,18} - s_{32,30}$ yields $\langle s_{32} \rangle = 5.3$ as an estimate for the pointwise correlation dimension.

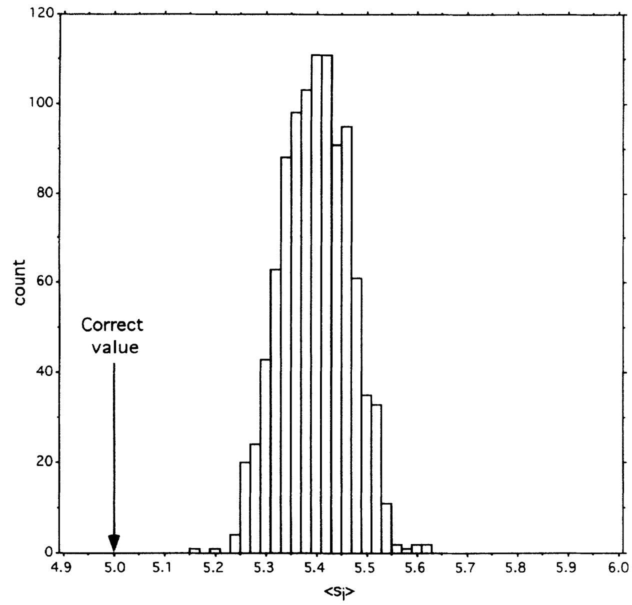

FIG. 8. Distribution of "pointwise correlation dimensions $D_2(\mathbf{x}_i)$." The $\langle s_i \rangle$ values are the averages of automatically fitted slopes $s_{i,d}$ of 1000 reference points $\mathbf{x}_i$ and 10^6 data points over embedding dimensions $d=15,\ldots,30$. None of the $\langle s_i \rangle$ values reach 5.0, the correct value of the correlation dimension of the five-dimensional system. The mean value is 5.4.

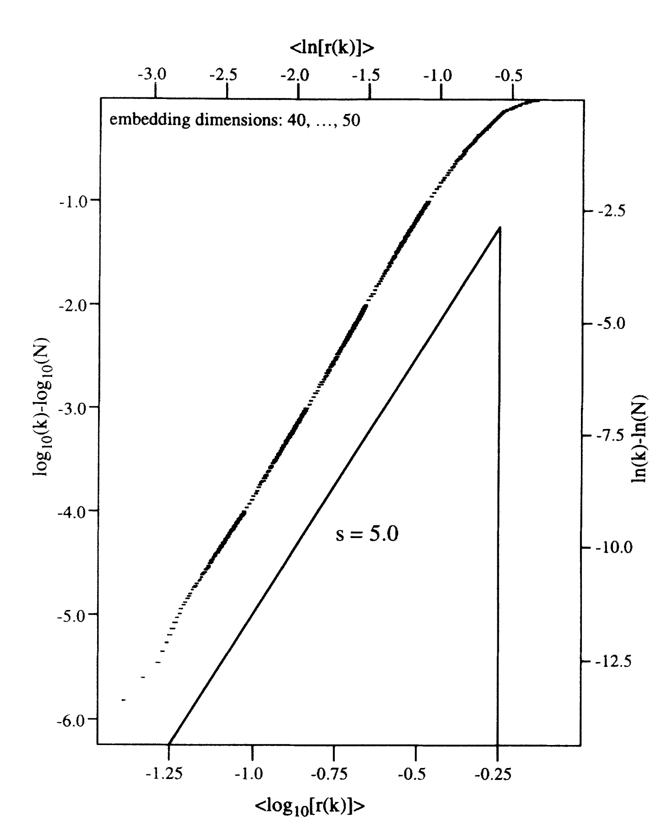

FIG. 9. The mean log distance $\langle \log_{10}[r(k)] \rangle$ is almost the same for the embedding dimensions $d=30,\ldots,50$. The scaling region is smaller than one order of magnitude and the difference between the slopes fitted by linear regression and the expected slope seems negligible but may tempt one to classify the system as a deterministic-chaotic one. Linear regression for the region $[-1;-0.5]$ of $\langle \log_{10}[r(k)] \rangle$ yields 5.48 for the slope s_d.

ence point $\mathbf{x}_i$. A histogram of these $\langle s_i \rangle$ values is shown in Fig. 8. None of the 1000 reference points shows a pointwise correlation dimension of 5 and all the values $\langle s_i \rangle$ are higher than 5.1. The mean value of the pointwise correlation dimension D_2 for our simple homogenous five-dimensional system is 5.4 and therefore too high.

C. Information dimension

Similar to the calculation of the pointwise-dimension values, $N=10^6$ data points and 100 reference points entered the calculation. The maximal embedding dimension was $d=50$.

Determining the linear scaling region poses the same problems as in the case of the classical Grassberger-Procaccia method. Figure 9 shows the mean distance $\langle r(k) \rangle$ for the mass k in the typical double logarithmic plot $\log_{10}(k)$ vs $\langle \log_{10}[r(k)] \rangle$. Although there is no problem fitting straight lines to determine the slopes s_d, the local slope $s_d(r)$ is depicted in Fig. 10. Each plateau is shorter than one order of magnitude of the distance r for the local slopes $s_d(r)$ in Fig. 10 and, especially, there is no plateau at the correct value 5.0. For embedding dimensions as high as $d=40,\ldots,50$ an information dimension of 5.56 may be claimed.

VII. DISCUSSION

Our very simple example of a five-dimensional system is easy to reproduce and the result is analytically known. Its dimension analysis can show dimension values higher than 5.0 if the methods are used uncritically. The numbers of data points are comparable to numbers in many published dimension analyses of experimental time series ($N=32\,000$) or even higher ($N=10^6$). The sampling rate for discretization (12-bit ADC) is comparable as well and the delay time τ is reasonable. We found a correlation dimension $D_2=5.7$ (classical Grassberger-Procaccia), a pointwise correlation dimensions $D_2\approx 5.4$, and an information dimension $D_1\approx 5.56$. The results look very consistent and would strongly suggest that the system under consideration is of a chaotic nature. Our findings demonstrate that dimension analysis can fail for high-dimensional systems. Therefore these methods are not suitable to distinguish reliably high-dimensional systems from stochastic processes.

A. Why the failure?

The failure might have been expected because our N was too low. If indeed $N_{\min}\geq 42^D$ data points [9] are necessary to calculate correlation dimensions, at least

$42^5 \approx 130 \times 10^6$ data points would have been necessary to adequately reconstruct the five-dimensional system. Our counterexample shows that the demands for dimension analysis of high-dimensional systems must be met by a sufficient N and sufficiently long scaling regions for small distances. Otherwise, even if the estimates of dimensions look very stable (independent of delay time τ and saturating for increasing embedding dimensions), the results may be wrong. This striking simple counterexample may be more convincing for many experimental research groups than the theoretical reasoning about N. If N is low, scaling regions at small distances cannot be found. Apparent scaling regions that depend on the global structure tend to be accepted. Instead of rejecting the existence of a scaling region and therefore considering the limitations of the method, the apparent scaling regions for large distances are erroneously taken as evidence that the methods are successful.

The scaling regions $[r_1;r_2]$ start for all three methods at distances $r_1 \approx 0.1 r_{\max}$, at 10% of the diameter of the attractor. 10% ($0.1 \times 4.096 \approx 410$) of the diameter of an attractor may look moderate, but it is larger than three of the amplitudes $\tilde{A}_i$ in the five-dimensional example:

$$\sum_{i=1}^{5} A_i = 28 \xrightarrow[\text{12-bit}]{} \tilde{A}_1 :\approx 731 : \tilde{A}_2 :\approx 366 : \tilde{A}_3 :\approx 219 ; \tilde{A}_4 :\approx 293 ; \tilde{A}_5 :\approx 439 \ . \tag{17}$$

This means that our estimated dimensions, ranging from 5.4–5.7, are not even approximations for the true dimension reflecting local properties of the invariant set, but result (quasi by accident) from the complicated global structure. The example of our five-dimensional system shows that the limit $r \to 0$ has to be taken seriously and that $N = 10^6$ is too low in the case of our high-dimensional system.

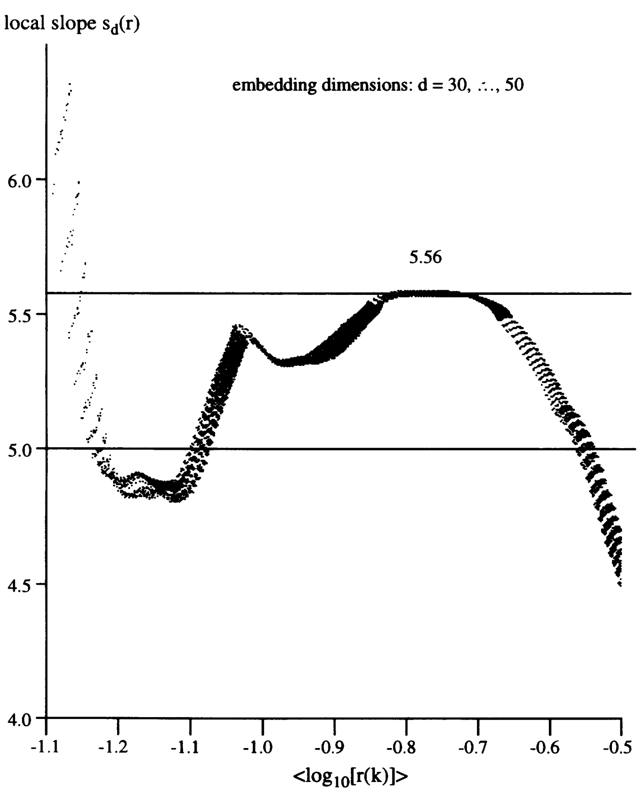

FIG. 10. The local slopes $s_d(r)$ of our five-dimensional system do not show plateaus with one order of magnitude in length. As the changes for embedding dimensions $d = 30, \ldots, 50$ are very small one might attribute this to saturation effects and accept the second line in the plot with a slope around $s = 5.56$. The local slopes were calculated by linear regression over a region of $\Delta \log_{10}(r) = 0.25$ in length: linear regression $[(\{\log_{10}(r_j);\ \log_{10}[C_d(r_j)]\} \mid \|r_j - r\| \leq 0.25)] \to s_d(r)$.

B. How to avoid erroneous dimension analysis

The evidence for deterministic chaos is often based on rather short experimental time series [1–4,27,28]. To check the validity of the algorithms for high-dimensional systems by simulations one may increase N until the result of the dimension analysis is in agreement with the expected values. The conclusion that N is sufficient to reconstruct a five-dimensional system is not correct. Increasing N from 1000, 2000, 4000, 8000, 16 000, up to 10^6 increases the correlation dimension for the pointwise correlation dimension for $N = 1000$: $D_2 = 4.6 \pm 0.4$; $N = 2000$: $D_2 = 4.8 \pm 0.3$; $N = 4000$: $D_2 = 5.0 \pm 0.2$ (Fig. 11). As the result for $N = 4000$ seems to be in the best agreement with the analytically expected value of 5.0,

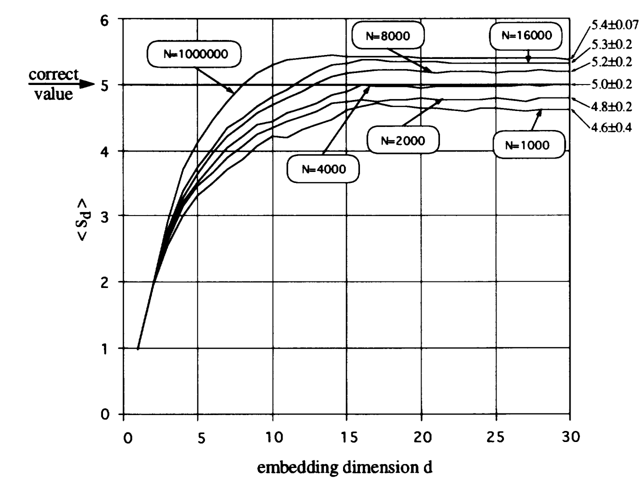

FIG. 11. We used several numbers of data points N ($N = 1000$; 2000; 4000; 8000; 16 000; 10^6) to calculate the pointwise dimensions $\hat{D}_{2,i} := \hat{D}_2(\mathbf{x}_i)$. Here we averaged the automatically fitted slopes $s_{i,d}$ over the reference points i and got $\langle s_d \rangle$ according to Eq. (12) for each N. For each N there is clear saturation of $\langle s_d \rangle$, but the value of saturation is depending on N. For small N the estimate of the correlation dimension is low ($N = 1000 \to D_2 = 4.6 \pm 0.4$), for $N = 4000$ it is apparently "best" ($D_2 = 5.0 \pm 0.2$), and for larger N it reaches $D_2 = 5.4 \pm 0.07$. The numbers to the right of the plot are the average values of the distribution of pointwise correlation dimension: $\langle \hat{D}_2 \rangle := \sum_{i=1}^{N_{\rm ref}} \langle s_i \rangle / N_{\rm ref}$ [Eq. (13)].

$N = 4000$ data points might seem to be sufficient to reconstruct the five-dimensional system and further increase of N might be stopped.

The results above show that this procedure is fallible. The usage of far too low numbers of data points N in dimension analysis forces one to select apparent scaling regions starting at large values of r. The pitfall of apparent scaling regions can be avoided if the limit $r \to 0$ in the definition of dimension values is considered. Plotting the local slopes of our simple five-dimensional example (Fig. 12) for the first 20 embedding dimensions yields nice plateaus for the first three embedding dimensions. If the system dimension would have been 3 or less, it could have been possible to identify the correct dimension ($N_{min} = 42^3 = 74\,088 \leq 10^6 = N$ data points used for dimension analysis). The length of the plateaus for an embedding dimension less than 3 is one order of magnitude and can be found for small distances r. In case of embedding dimension $d = 4$ the expected plateau does not appear, which is an indication of a lack of data points for analyzing four-dimensional systems.

For high embedding dimensions ($d > 10$), only pseudoplateaus at larger distances $r_1 > \langle \log_{10}[r(k)] \rangle \approx -1.0$ arise (see Fig. 10). The plateaus start at about 10% of the diameter of the reconstructed invariant set. If an algorithm under consideration is able to work for low-dimensional systems, the conclusion that it will work successfully for high-dimensional systems (e.g., ≥ 5) may not hold. To avoid the pitfalls mentioned above, simulations of very simple high-dimensional systems (quasiperiodic processes) with analytically known results ought to be recovered before high dimensions from experimental time series are acceptable. This is a first and minor requirement.

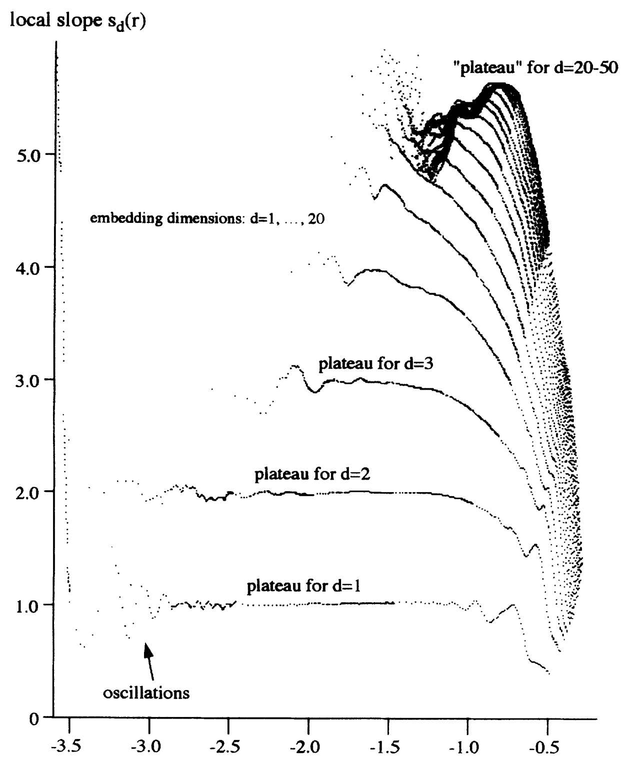

FIG. 12. The local slopes ($\Delta \log_{10}(r) = 0.25$) show only for low embedding dimensions plateaus of meaningful length. As the embedding dimensions increase, the number of vector pairs with distances smaller than $r = 0.1 r_{max}$ decreases and estimation of the scaling behavior of correlation integrals and mean distances in the limit $r \to 0$ is not possible. The plateau of the local slope at large distances r can hardly be taken seriously and is caused by the complicated structure of our five-dimensional set in high-dimensional spaces. As there are only plateaus for low embedding dimensions only low-dimensional systems (≤ 3) can be estimated accurately using 10^6 data points.

All three methods seem to fail in the case of our five-dimensional system. The main reason is that short scaling regions or short "plateaus" have been selected. If the slopes of the correlation integrals are estimated with only a single straight line, variations will not be detected. The manual selection or the corresponding automatic selection of scaling regions and the following fitting of a straight line may yield slopes for any correlation integral even if no scaling region exists. From this point of view the dimensions derived from Figs. 4, 6, 9, and 10 and the dimension derived by fitting the slopes automatically are victims of the pseudoplateaus resulting from the complicated structure of the embedded invariant set for large distances r. These slopes are artifactual and do not describe the local properties of our system. Thus, if the local slopes do not exhibit sufficient scaling regions (as for $d > 3$ in Fig. 12), respective data points in plots of slope vs embedding d (Figs. 5 and 7) are meaningless and misleading.

Gershenfeld [23] used up to 10^7 data points to reconstruct a 12-dimensional system. The scaling regions or plateaus of local slopes were rather small $[\log_{10}(r_2) - \log_{10}(r_1) \approx 0.3]$ but increased with N and the hope is that a further increase would build up plateaus for small distances r of sufficient length. He also showed that noisy data behave in a different way: Even for 10^7 data points, no plateau arises in a 20-dimensional embedding space for the slopes of the correlation integrals. So to make sure that a deterministic-chaotic system has been found, one should follow Theiler *et al.* [29] and produce "surrogate data" (same power spectrum but randomized phase) of the experimental time series in question. Then dimension analysis can be repeated for the surrogate data. If the local slopes of the surrogate data show plateaus up to embedding dimension $d_{surrogate}$ and the local slopes of experimental time series show plateaus at D_1 saturating with increasing embedding dimension and, furthermore, $D_1 < d_{surrogate}$, one is able to distinguish the results of dimension analysis of the experimental time series from corresponding noise data. If the plateaus of the local slopes are too short or if the slopes are calculated simply by linear regression from scaling regions of the correlation integrals, filtered noise can mimic low-dimensional chaotic attractors [30].

As an example of the proposed procedure, Fig. 13 depicts the results in the case of the Lorenz attractor with 10^6 data points. Because of its low dimensionality, the local slopes do not increase with the embedding dimension. On the other hand, the local slopes of the surrogate data

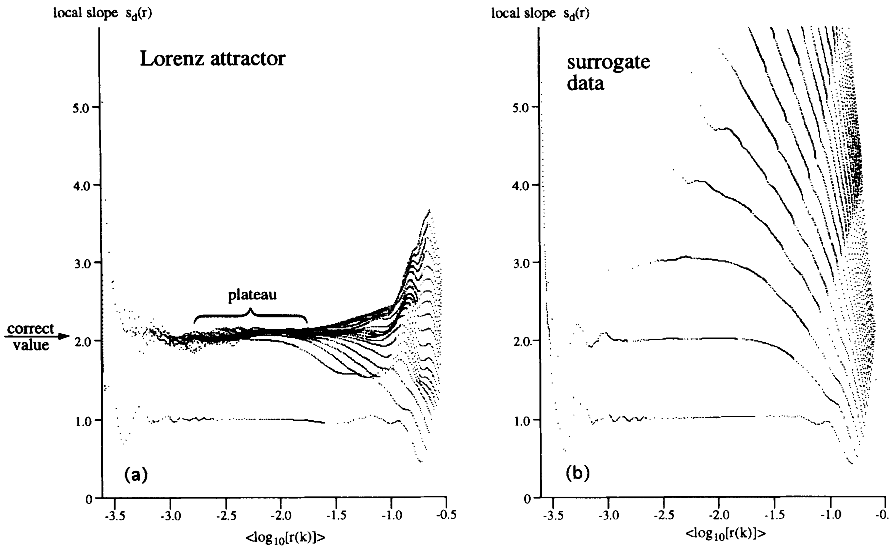

FIG. 13. (a) Local slopes of D_1 integrals for embedding dimension $d=1,\ldots,20$ of the Lorenz attractor [$\Delta \log_{10}(r)=0.2$]. Plateaus of one order of magnitude in length exist for small distances r even for high embedding dimensions. (b) Local slopes of the corresponding surrogate data show plateaus for embedding dimensions $d=1,2,3$. For embedding dimensions $d>3$ it is not possible to find plateaus of appropriate length and thus it is impossible to assign a dimension. As the height of the plateaus of the Lorenz time series saturate for high embedding dimensions near $D_I \approx 2$ and a plateau exists for the corresponding surrogate data at embedding dimension $d=3$, there is evidence of a finite information dimension.

increase with the embedding dimension and exhibit plateaus up to embedding dimension 3. Consequently, in this case, dimension analysis based on 10^6 data points is sufficient to detect the correct dimension of $d=2.06$ but would have been unable to detect dimensions of 3 or higher.

VIII. CONCLUSION

Straightforward uncritical application of dimension analysis may lead to incorrect conclusions. Proceeding this way, our example of $N=10^6$ data points taken from a very simple five-dimensional system yielded dimensions of 5.4–5.7.

Searching for the reasons of this failure we found that our results do not reflect, as required, the local structure determining the dimension of the system, rather they reflect some global structure of the reconstructed set. Thus these dimension values cannot even be taken as an approximation to the true value.

To obtain a meaningful estimation of the dimension one has to consider the local slopes of the logarithms of the correlation integrals in the limit $r \rightarrow 0$, which have to exhibit plateaus of sufficient length. This implies the requirement for huge values of N. In our example a dimension of 5 has been found to be "high" in the sense that it is not possible to detect it with 10^6 data points.

If data are derived from a low-dimensional system, plateaus have to appear even for high-dimensional embeddings. Thus embedding dimensions of at least 20 should be considered. If plateaus appear even for high-dimensional embedding dimensions, surrogate data can confirm the existence of a low-dimensional system. The example of the Lorenz attractor shows that low-dimensional systems are reliably detected by this procedure. We suggest that current results of dimension analysis concerning systems as complicated as human brains are far too optimistic and ought to be tested intensively before being applied as diagnostic tools [31–33].

ACKNOWLEDGMENTS

We would like to thank Wolfgang Jansen and Udo Schwarz (Nonlinear Dynamics Group, Max Planck Gesellschaft, University of Potsdam), for helpful discussions and corroboration of some of our computational results with independently implemented algorithms.

[1] E. Basar, *Chaos in Brain Function* (Springer, Berlin, 1990), p. 1.
[2] A. Babloyantz, in *Dimensions and Entropies in Chaotic Systems—Quantification of Complex Behavior,* edited by G. Mayer-Kress (Springer, Berlin, 1986), p. 241.
[3] J. Holzfuss and G. Mayer-Kress, in *Dimensions and Entropies in Chaotic Systems,* edited by G. Mayer-Kress (Springer, Berlin, 1986), p. 114.

[4] K. E. Graf and T. Elbert, in *Chaos in Brain Function,* edited by E. Basar (Springer, Berlin, 1990), p. 134.
[5] P. Grassberger, T. Schreiber, and C. Schaffrath, Int. J. Bifurc. Chaos **1**, 521 (1991).
[6] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. Rapp. Phys. Rev. A **38**, 3017 (1988).
[7] J. P. Caputo, B. Malrasion, and P. Atten in *Dimensions and Entropies in Chaotic Systems* (Ref. [3]), p. 57.
[8] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Phys. Rev. Lett. **25**, 3872 (1993).
[9] L. A. Smith, Phys. Lett. A **133**, 283 (1988).
[10] D. Ruelle, Proc. R. Soc. London Ser. A **427**, 241 (1990).
[11] F. Takens, in *Proceedings of the Symposium on Dynamical Systems and Turbulence, University of Warwick, 1979/1980,* edited by D. Rand and L.-S. Young, Lecture Notes in Mathematics Vol. 898 (Springer, Berlin, 1981), p. 366.
[12] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[13] R. Badii and A. Politi, Phys. Rev. Lett. **52**, 1661 (1984).
[14] Y. Termonia and Z. Alexandrowicz, Phys. Rev. Lett. **51**, 1265 (1983).
[15] J. D. Farmer, E. Ott, and J. Yorke, Physica D **7**, 153 (1983).
[16] R. Benzi, G. Paladin, and A. Vulpiani, J. Phys. A **17**, 3521 (1984).
[17] P. Grassberger, Phys. Lett. A **148**, 63 (1990).
[18] J. Theiler, Phys. Rev. A **34**, 3 (1986); **34**, 2427 (1986).
[19] J. Theiler, J. Opt. Soc. Am. A **7**, 1055 (1990).
[20] T. Buzug, T. Reimers, and G. Pfister, Europhys. Lett. **13**, 605 (1990).
[21] C. Essex and M. A. H. Nerenberg, Proc. R. Soc. London Ser. A **435**, 287 (1991).
[22] G. W. Frank, T. Lookman, M. A. H. Nerenberg, and C. Essex, Physica D **46**, 427 (1990).
[23] N. A. Gershenfeld, Physica D **55**, 135 (1992).
[24] J. W. Havstad and C. L. Ehlers, Phys. Rev. A **39**, 845 (1989).
[25] H. Herzel, J. Kurths, P. S. Landa, and M. G. Rosenblum, in *Irreversible Processes and Selforganization,* edited by W. Ebeling and H. Ulbricht (Teubner-Verlag, Leipzig, 1989).
[26] H. Isliker, Phys. Lett. A **169**, 313 (1992).
[27] C. Nicolis and G. Nicolis, Nature **311**, 529 (1984).
[28] J. Röschke and J. Aldenhoff, Bio. Cybern. **64**, 307 (1991).
[29] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Physica D **58**, 77 (1992).
[30] P. E. Rapp, A. M. Albano, T. I. Schmah, and L. A. Farwell, Phys. Rev. E **22**, 89 (1993).
[31] I. Dvorak, Phys. Lett. A **151**, 225 (1990).
[32] A. Babloyantz and A. Destexhe, in *Temporal Disorder in Human Oscillatory System,* edited by L. Rensing, U. van der Heiden, and M. C. Mackey (Springer, Heidelberg, 1987), p. 48.
[33] R. C. Watt, S. R. Hameroff, S. R. Alpha, and K. C. Mylrea, Anesthesiology **69**, A314 (1988).

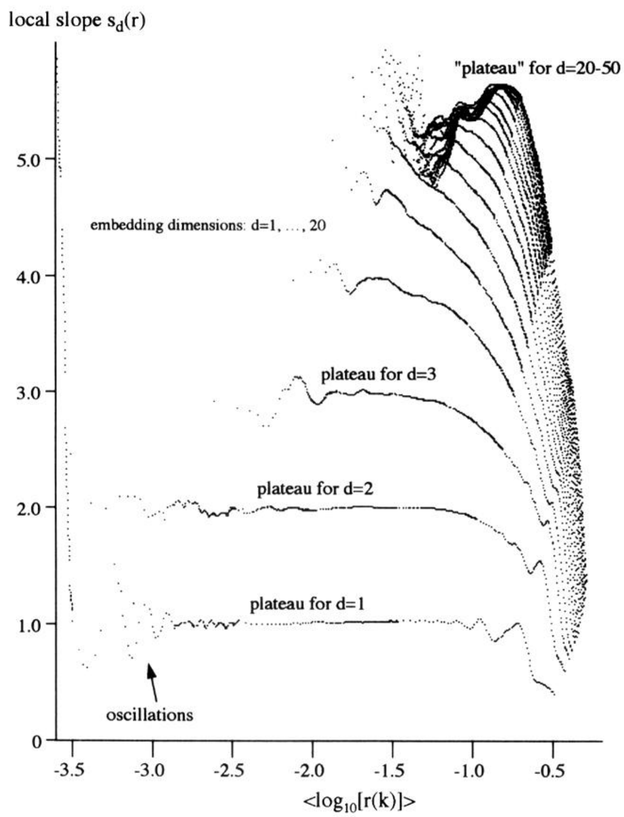

FIG. 12. The local slopes ($\Delta \log_{10}(r)=0.25$) show only for low embedding dimensions plateaus of meaningful length. As the embedding dimensions increase, the number of vector pairs with distances smaller than $r=0.1r_{\max}$ decreases and estimation of the scaling behavior of correlation integrals and mean distances in the limit $r \rightarrow 0$ is not possible. The plateau of the local slope at large distances r can hardly be taken seriously and is caused by the complicated structure of our five-dimensional set in high-dimensional spaces. As there are only plateaus for low embedding dimensions only low-dimensional systems (≤ 3) can be estimated accurately using 10^6 data points.

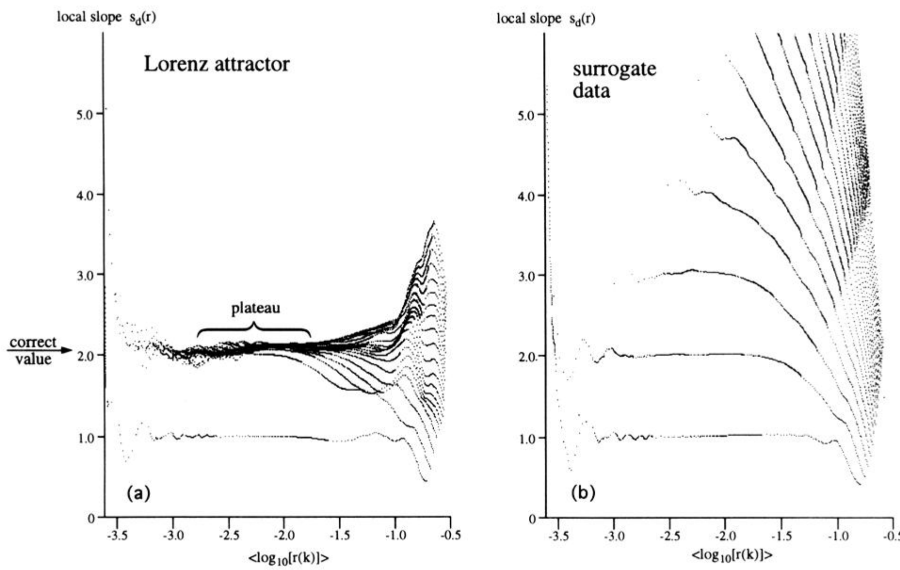

FIG. 13. (a) Local slopes of D_1 integrals for embedding dimension $d = 1, \ldots, 20$ of the Lorenz attractor $[\Delta \log_{10}(r) = 0.2]$. Plateaus of one order of magnitude in length exist for small distances r even for high embedding dimensions. (b) Local slopes of the corresponding surrogate data show plateaus for embedding dimensions $d = 1,2,3$. For embedding dimensions $d > 3$ it is not possible to find plateaus of appropriate length and thus it is impossible to assign a dimension. As the height of the plateaus of the Lorenz time series saturate for high embedding dimensions near $D_I \approx 2$ and a plateau exists for the corresponding surrogate data at embedding dimension $d = 3$, there is evidence of a finite information dimension.

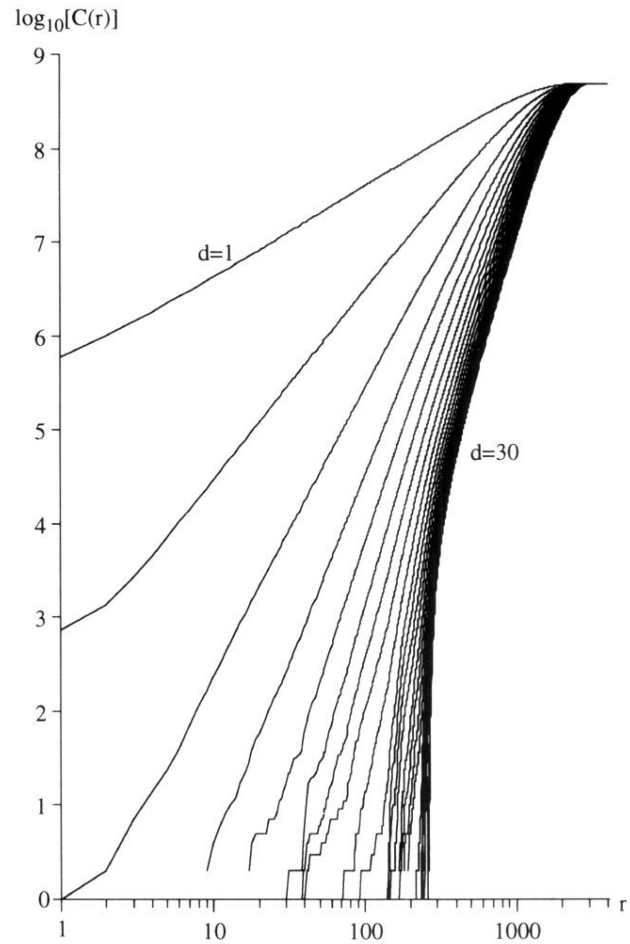

FIG. 3. Correlation integrals $C_d(r)$ of our five-dimensional system for embedding dimension $d=1,\ldots,30$. The number of data points was $N=32\,000$. For high embedding dimensions further increase of embedding dimensions does not change the correlation integrals $C_d(r)$ and saturation can be expected.

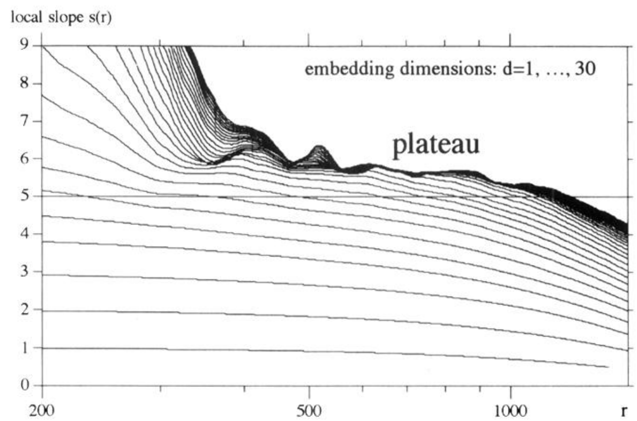

FIG. 6. The local slopes $s_d(r)$ show a plateau around 5.7. The length of the scaling region $\log_{10}(900)-\log_{10}(700) \approx 0.11<1$ is rather short and starts at large distances r. The local slopes $s_d(r)$ have been calculated from the correlation integrals in Fig. 3 by linear regression over 50 successive distance r_j: linear regression $[(\{\log_{10}(r_j);\ \log_{10}[C_d(r_j)]\} | r_j=r+j$ and $j=1,\ldots,50)] \rightarrow s_d(r)$.