Physics Letters A 169 (1992) 313–322
North-Holland

PHYSICS LETTERS A

A scaling test for correlation dimensions

Heinz Isliker
Institut für Astronomie, ETH Zentrum, CH-8092 Zurich, Switzerland

Received 28 October 1991; revised manuscript received 27 May 1992; accepted for publication 5 August 1992
Communicated by A.P. Fordy

A quantitative test is presented to check scaling and convergence of the correlation integral and consistency of two different algorithms. An application to two known attractors demonstrates that it allows one to judge fast and reliably the quality of a conjectured scaling behaviour above all in the case of short or noisy data. Results concerning minimum data amount and maximum noise level confirm earlier work, the crucial parameter concerning data length turns out, however, to be not the number of points, but the number of cycles in phase space (peaks in the time series).

1. Introduction

Correlation dimensions have been used to investigate whether a process is deterministic or not. However, the last years showed that there is a need for criteria of reliability of dimension estimates. Erroneous dimensions can have two causes:

On the one hand, there is a mathematical reason. A class of stochastic processes shows a finite correlation dimension, such as fractional Brownian motion [1]. This motion is approximately a self-similar random path. Self-similarity locally leads to a finite fractal dimension of which correlation dimension is a representative. The dimensions are "erroneous" just in the sense that they are not indicative of deterministic behaviour. Theiler [2] systematically investigates these processes and proposes methods to identify them.

On the other hand, erroneous dimensions can result from statistical effects, in three ways: (i) The number of available points might not be sufficient. (ii) The influence of noise is too severe. (iii) There is an uncertainty in identifying a linear scaling region of the logarithm of the correlation integral.

The present investigation deals with point (iii) in this second group of error sources. We will present a method to determine the reliability of the linear scaling region in section 3. Applying it in sections 4 and 5 to two well-known examples of attractors of intermediate dimension (about 2 and 3.5), its efficiency is demonstrated mainly for the case of short or noisy data sets, where the convergence gets worse and worse. As an aside we give reference plots of typical behaviour of attractors if the parameter noise and data length are varied. They show that the crucial measure of data length is not the number of points, but the number of peaks in a time series.

We start with reviewing two ways of calculating correlation dimensions.

2. Two common ways to determine the correlation dimension

According to the time delay method of Takens [3], the phase space is reconstructed from a given a time series $\{X(t_i)\}_{i=1}^{N}$, yielding the d-dimensional vectors $\xi(t_i)$. Different kinds of dynamical systems yield different kinds of limit sets in state space, so-called attractors, distinguishable by the notion of dimension. For a quasi-periodic motion it is a torus. For a deterministic chaotic process it generally is a complicated invariant set of finite dimension. Finally, a stochastic process is an erratic movement and spans a subset of phase space whose dimension equals the one of the reconstructed state space – at least in many cases, compare however the fractional Brownian motion mentioned in the introduction. We resume two popular and easily workable definitions of dimension:

(A) *The Grassberger–Procaccia (GP) method.* Grassberger and Procaccia [4] define the correlation integral

$$C_d^{(2)}(\epsilon) := \lim_{N\to\infty} \frac{2}{N(N-1)} \sum_{i<j}^{N} \Theta(\epsilon - |\boldsymbol{\xi}_i - \boldsymbol{\xi}_j|) , \tag{1}$$

with the Heaviside function $\Theta(\ )$ and any vector norm $|\ \ |$. The correlation dimension $D^{(2)}$ is defined via $C_d^{(2)}(\epsilon) \sim \epsilon^{D^{(2)}}$, for a large enough embedding dimension d.

(B) *Maximum likelihood (ML) estimate of correlation dimensions* [5,6]. The distances $|\boldsymbol{\xi}(t_i) - \boldsymbol{\xi}(t_j)|$ are expected to have a probability distribution $\epsilon^{D^{(2)}}$, with parameter $D^{(2)}$. This parameter can be calculated by means of a maximum likelihood estimate: Assume the scaling law $\epsilon^{D^{(2)}}$ to hold for $\gamma_1 \leqslant \epsilon \leqslant \gamma_2$. The maximum likelihood formalism yields

$$D^{(2)} = -\frac{n-K}{\sum_{ij=1}^{n} \ln(r_{ij}/\gamma_2)} , \tag{2}$$

where the n distances $r_{ij} := \max(|\boldsymbol{\xi}(t_i) - \boldsymbol{\xi}(t_j)|, \gamma_1)$ are chosen at random, and K denotes the number of distances r_{ij} equal to γ_1. Ellner [6] shows that n chosen as half of the length of the original time series yields reasonably accurate results, $n = \frac{1}{2}N$.

(C) *Intrinsic error estimate of correlation dimensions.* Often, the error of a regression in the $\log C_d^{(2)}(\epsilon)$–$\log \epsilon$ representation is taken to be the error of the estimated correlation dimension. However, this error is not an *intrinsic* one. An adequate error estimate, based entirely on the probability interpretation of the notions, is the following [6]: calculating the mean value and the standard deviation of the distribution $\epsilon^{D^{(2)}}$ and propagating the error the Gauss way into $D^{(2)}$ yields an error Δ of $D^{(2)}$ as

$$\Delta = \frac{1.96 D^{(2)}}{\sqrt{n}} \frac{\sqrt{1 + 2\ln(r_0^{D^{(2)}}) r_0^{D^{(2)}} - r_0^{2D^{(2)}}}}{1 + \ln(r_0^{D^{(2)}}) r_0^{D^{(2)}} - r_0^{D^{(2)}}} , \tag{3}$$

with $r_0 := \gamma_1/\gamma_2$, and the factor 1.96 stemming from adopting a 5% confidence interval, as proposed by Ellner.

3. A measure of reliability of the "plateau"

In the GP algorithm, $D^{(2)}$ is usually determined as the linear scaling region in a $\log C_d^{(2)}(\epsilon)$–$\log \epsilon$ representation, the so-called plateau. How to judge the quality of a plateau? The idea is to use the probabilistic character of the notions: $C_d^{(2)}(\epsilon)$ is a distribution of distances $|\boldsymbol{\xi}(t_i) - \boldsymbol{\xi}(t_j)|$, expected to be of the form $\epsilon^{D^{(2)}}$. By means of a χ^2-test it is quantitatively possible to compare the empirical distribution $C_d^{(2)}(\epsilon)$ with the conjectured one $\epsilon^{D^{(2)}}$, checking by that whether it is justified to assume a particular scaling of the correlation integral.

Scaling and convergence test. Assume a plateau in the range $\gamma_1 \leqslant \epsilon \leqslant \gamma_2$, suggesting a correlation dimension $D^{(2)}$. Then the test proceeds as follows:

(a) Divide the ϵ-space (space of distances $|\boldsymbol{\xi}(t_i) - \boldsymbol{\xi}(t_j)|$) into classes,

$$[0, \gamma_1], (\gamma_1, \epsilon^{(2)}], (\epsilon^{(2)}, \epsilon^{(3)}], \ldots, (\epsilon^{(m_{\rm cl}-1)}, \gamma_2] . \tag{4}$$

The first class includes the distances smaller than γ_1.

(b) Calculate the frequencies of distances $|\boldsymbol{\xi}(t_i) - \boldsymbol{\xi}(t_j)|$ in these classes. Let $n_k^{(\rm e)}$ denote the "empirical" frequency of distances in the kth class, as it is given by the distribution $C_d^{(2)}(\epsilon)$, and $n_k^{(\rm th)}$ the "theoretical" one in the same class, calculated from $\epsilon^{D^{(2)}}$:

$$n_k^{(\rm e)} := Z[C_d^{(2)}(\epsilon^{(k)}) - C_d^{(2)}(\epsilon^{(k-1)})] ,$$
$$n_k^{(\rm th)} := Z\left[\left(\frac{\epsilon^{(k)}}{\gamma_2}\right)^{D(2)} - \left(\frac{\epsilon^{(k-1)}}{\gamma_2}\right)^{D(2)}\right] , \tag{5}$$

where the normalization Z equals $\frac{1}{2}N(N-1) \times C_d^{(2)}(\gamma_2)$.

(c) The test quantity χ^2 is defined as

$$\chi^2 := \sum_{k=1}^{m_{\rm cl}} \frac{(n_k^{(\rm e)} - n_k^{(\rm th)})^2}{n_k^{(\rm th)}} \tag{6}$$

and expected to have a χ^2 distribution with $m_{\rm cl} - 2$ degrees of freedom. To accept a scaling we demand that χ^2 satisfies $\text{prob}[x^2 \leqslant \chi^2] \leqslant 95\%$, plateaus with a too low significance are rejected.

Comments. The used value of $D^{(2)}$ is estimated from a $\log C_d^{(2)}(\epsilon)$–$\log \epsilon$ representation. The values of $\epsilon^{(k)}$ in step (a) can be the values for which the distribution $C_d^{(2)}(\epsilon)$ has anyway been evaluated. We generally used $5 \lesssim m_{\rm cl} \lesssim 10$.

This answers the question of goodness of the scaling fit and, by comparing different embedding dimensions, of convergence of the GP algorithm.

Consistency. We further demand that the GP value of $D^{(2)}$ is reproduced by the ML algorithm, within the – meaningful – error Δ of the latter (eq. (3)). This is an intuitive and qualitative criterion. To quantify it we use throughout the upper χ^2-test the value $D^{(2)}$ calculated the ML way. As $D^{(2)}$ is estimated from the same data as the correlation integral, the usual degree of freedom of the χ^2-test is still lowered by one, yielding again $m_{\text{cl}}-2$ degrees of freedom.

Figure 1 shows a well-behaved case of convergence of the GP algorithm for a time series generated with the Mackey–Glass equations (see eq. (11)). It yields a significance of 95% for the marked range, the plateau is accepted. Going to smaller data sets, or increasing the noise level, the quality of the plateau gets worse, and extended numerical experiments with time series generated by different equations (sine wave, Lorenz equations, Rössler equations, Mackey–Glass equations) show that even in these well behaved cases the χ^2-test often fails, it does not approve the – distinct – plateaus. This may partly have numerical reasons, however more important are two theoretical findings:

Badii and Politi [7] and Smith et al. [8] show that for any lacunar set, i.e. a set with sparse regions, the scaling law $\epsilon^{D^{(2)}}$ cannot hold anymore in a strict sense. It must be replaced by the more general form

$$\epsilon^{D^{(2)}} \to \epsilon^{D^{(2)}} \psi(\ln \epsilon / P) , \tag{7}$$

with ψ being an unknown periodic function with period 1: There will generally be superimposed on the plateau an oscillation with completely unknown period and amplitude, both depending on the individual structure of lacunarity.

Furthermore, Smith [9] showed that a plateau will be skew, decreasing from its true value at small radii r towards larger radii. A phenomenon caused by the finite size of the attractor: Near the edge, a point is surrounded by other points only on one side. The

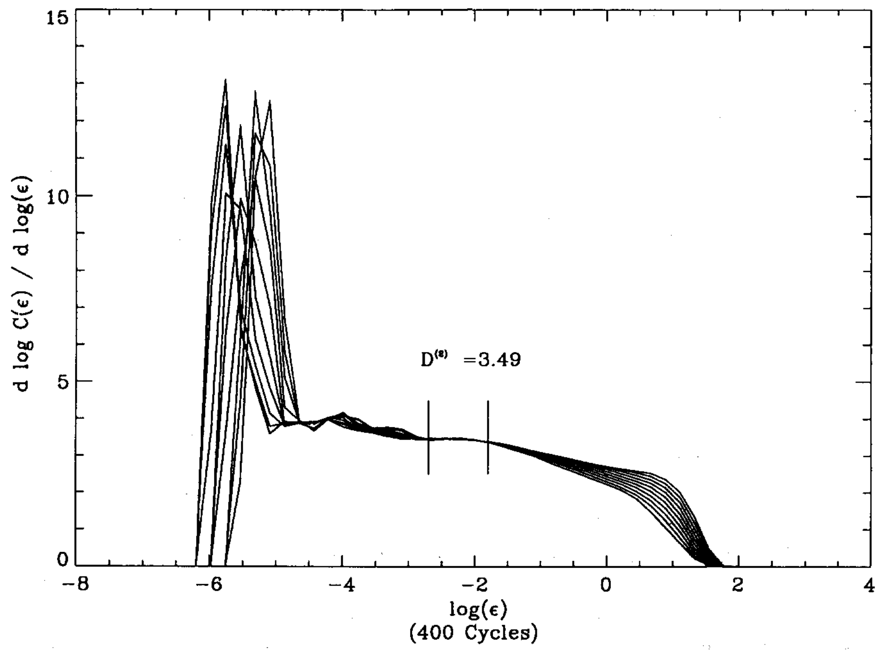

Fig. 1. According to the Grassberger–Procaccia method the slope of the correlation integral $\log C_d^{(2)}(\epsilon)$ against $\log \epsilon$ is plotted (see eq. (1)) for a time series calculated from the Mackey–Glass equation (eq. (11)). We used 16000 points, the time resolution τ was chosen so that $t_{\text{corr}}/\tau \approx 10$, hence the number of cycles n_{S} equals 1600 (see eq. (12)). There is a distinctly flat scaling region for intermediate radii ϵ, modulated by an oscillation towards smaller radii, and by skewness towards larger radii.

probability distribution for distances, $\epsilon^{D^{(2)}}$, is therefore biased.

The typical plateau in fig. 1 for the Mackey–Glass system clearly shows both the above deviation features. The χ^2-test is obviously sensitive enough to detect these oscillating and skew deviations from a wrongly assumed flat scaling region. We are wrong if we expect a behaviour $\epsilon^{D^{(2)}}$. We do not have access to the true values of the deviations, however we have at least the intrinsic error Δ for the correlation dimension (eq. (3)). It seems therefore reasonable to adjust the χ^2-test, allowing for deviations within the error of the correlation dimension. We propose to use

$$\chi^2_{\text{adapted}} := \sum_{k=1}^{m_{\text{cl}}} \frac{\max^2[\,|n_k^{(\text{e})} - n_k^{(\text{th})}| - \eta_k;\, 0]}{n_k^{(\text{th})}} \tag{8}$$

instead of χ^2 of eq. (6). η_k is the error of frequency in the kth class for the distribution $\epsilon^{D^{(2)}}$, due to the error Δ in $D^{(2)}$ (eq. (3)):

$$\eta_k = Z\{[(\epsilon^{(k)})^{D^{(2)}} \ln \epsilon^{(k)}]^2 + [(\epsilon^{(k-1)})^{D^{(2)}} \ln \epsilon^{(k-1)}]^2\}^{1/2} \Delta\,. \tag{9}$$

This allows a plateau to fluctuate between $D^{(2)} - \Delta$ and $D^{(2)} + \Delta$ and to be still acknowledged. Care is to be taken that the frequencies $n_k^{(\text{th})}$ and $n_k^{(\text{e})}$ are well above the theoretically allowed minimal value of 5. If they are not, the test has a tendency to become too tolerant, the deviations from a plateau it allows for becoming too large. For small numbers of distances imply a large error Δ (eq. (3)), and by that large tolerances η_k. In numerical experiments, this has proven to be a very adequate tool: Slightly oscillating or skew plateaus are still approved. This is illustrated in the next two sections.

4. Application I: the number of points

We investigate two intermediately-high-dimensional systems and, with the method of the previous section, we address two problems: How many points are necessary to detect a correlation dimension, and what is the dependence on the number of points per cycle in phase space?

– The Rössler attractor, which is the limit set of the three-dimensional Rössler system:

$$\dot{X} = -(Y+Z), \quad \dot{Y} = X + 0.2Y, \quad \dot{Z} = 0.2 + Z(X - 5.7) \tag{10}$$

(see ref. [10]). We solved it by a Runge–Kutta algorithm and took the X-coordinate as a time series to analyze – fig. 2 shows a part of this time series.

– The Mackey–Glass system, whose equation is of the delay type and therefore existing in an infinite-dimensional space (the actually used degrees of freedom can be lower, e.g. if an attractor exists):

$$\dot{X}(t) = \frac{aX(t-\tau)}{1 + [X(t-\tau)]^{10}} - bX(t)\,. \tag{11}$$

We choose $a = 0.2$, $b = 0.1$. For a study of this system see ref. [4], whose method of solving the equation is used here, converting the equation into a system of 600 difference equations (for a time profile see fig. 3).

Both systems were analyzed in a chaotic regime, using solutions just after a certain amount of time to avoid transient features. The references values of the corresponding correlation dimensions were taken from long time series with ten points per cycle in phase space: 16000 points yielded a dimension of 3.49 ± 0.17 for the Mackey–Glass attractor; 8000 points in the case of the Rössler attractor showed a dimension of 1.86 ± 0.06 (the errors are calculated via eq. (3), with r_0 estimated from the GP procedure and $D^{(2)}$ from the ML procedure).

The results are visualized in figs. 4a and 4b for the Rössler and the Mackey–Glass attractor, respectively. The calculations were done in an automatic way, the decision whether there is a plateau or not can be left to the computer – a question else becoming awkward for short time series. The user has only to propose ranges $\gamma_1 < \epsilon < \gamma_2$ of conjectured plateaus to the procedure. We implemented a loop in order to check always several possible regions. It turned out, however, that the procedure is just weakly dependent on the finer details of the choice of γ_1 and γ_2. The plateau widths are found to be $3 \lesssim \gamma_2/\gamma_1 \lesssim 4$ for the Rössler system, $2 \lesssim \gamma_2/\gamma_1 \lesssim 2.5$ for the Mackey–Glass system. Figure 4 shows that the method reproduces the results of other practical inquiries showing that $N \gtrsim 1000$ can deliver reliable, though not very precise results – for dimensions between about 1 and 7 [11–14]. The scaling test may there-

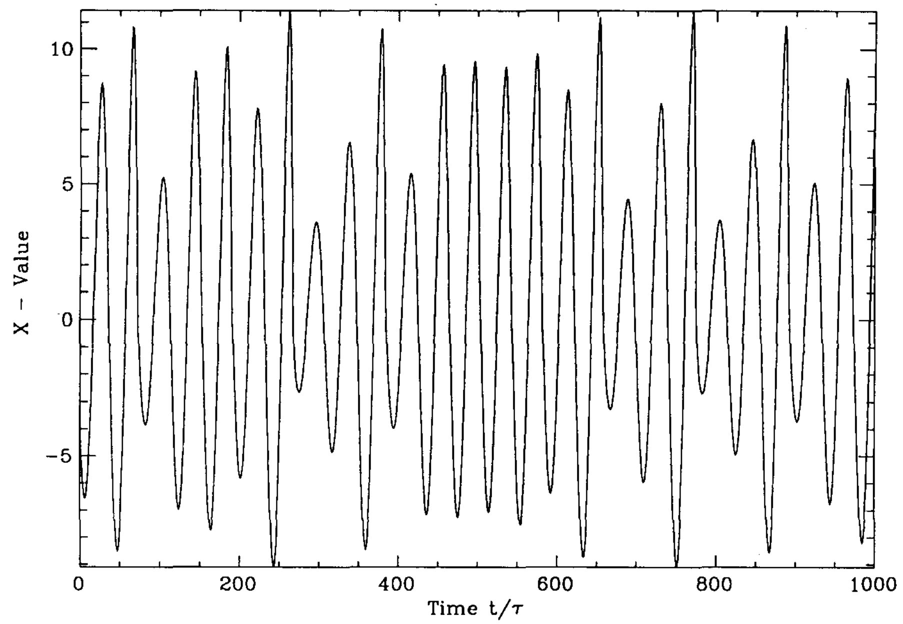

Fig. 2. Time profile of the X-coordinate of the Rössler system. The time step τ is chosen so that $t_{corr}/\tau \approx 20$.

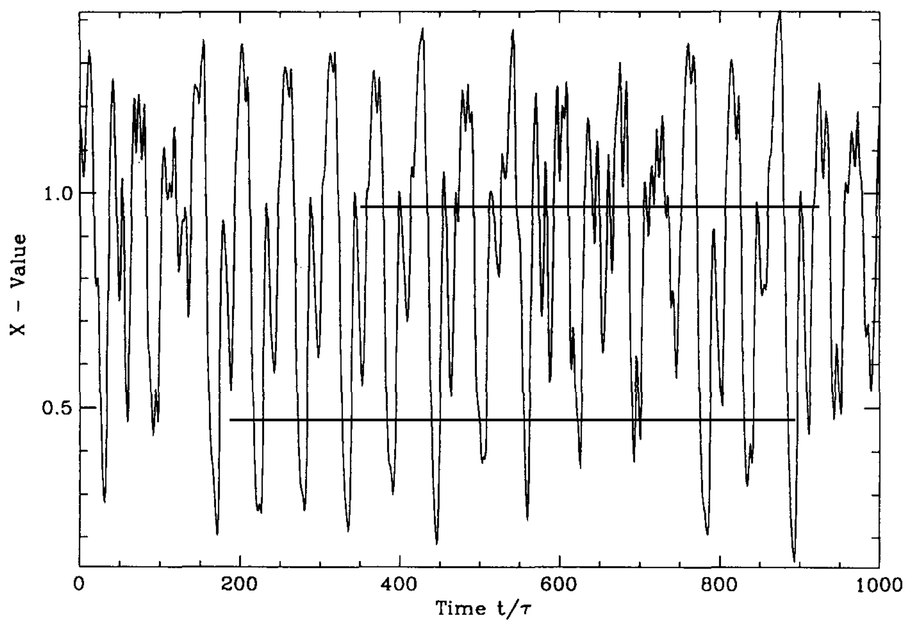

Fig. 3. Time profile of a solution $X(t)$ of the Mackey–Glass system. The time step τ is chosen so that $t_{corr}/\tau \approx 20$.

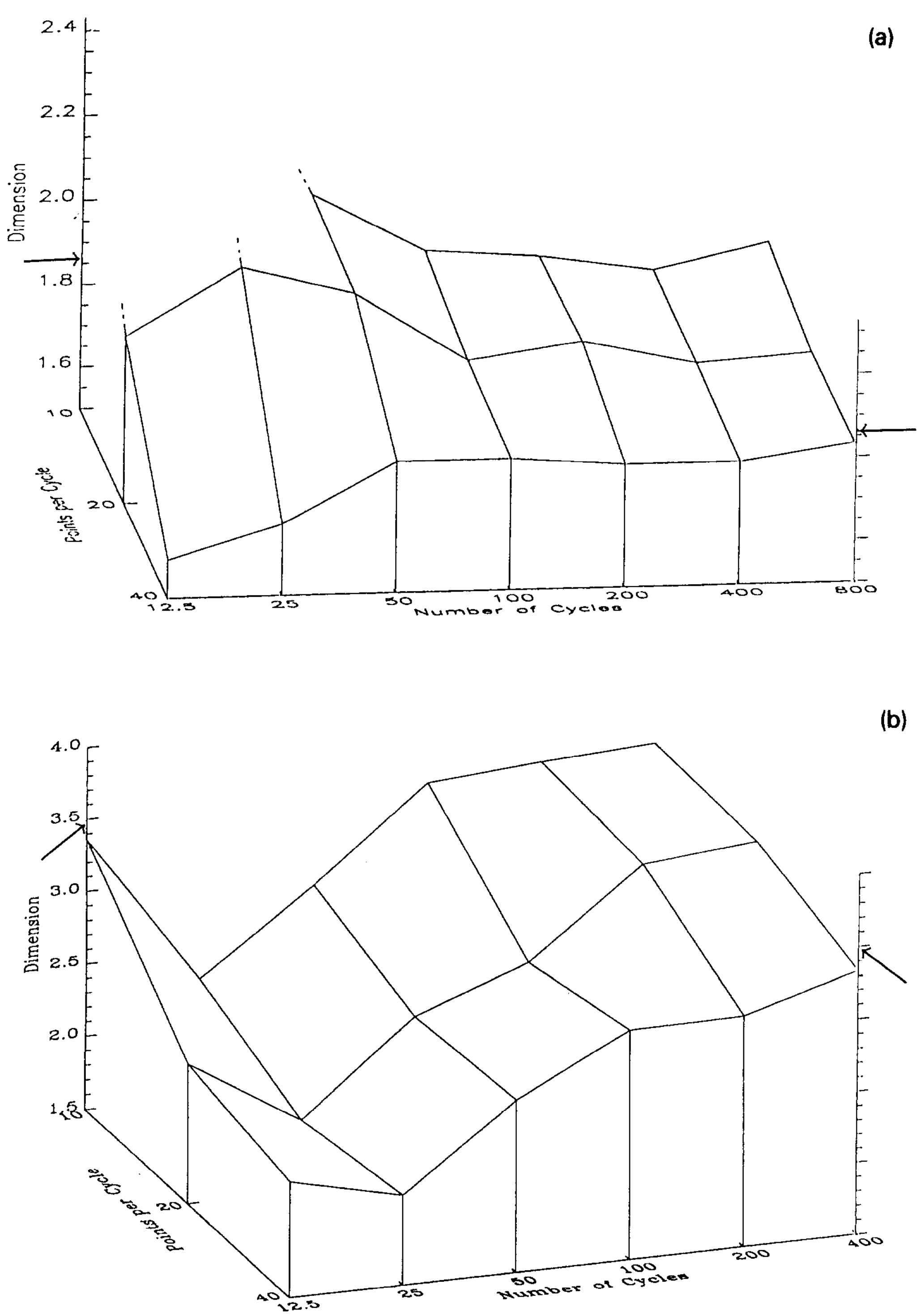

Fig. 4. Correlation dimension as a function of n_S, the number of cycles, and of t_{corr}/τ, the number of points per least cycle, (a) for the Rössler system, (b) for the Mackey–Glass system. The true values are marked by arrows.

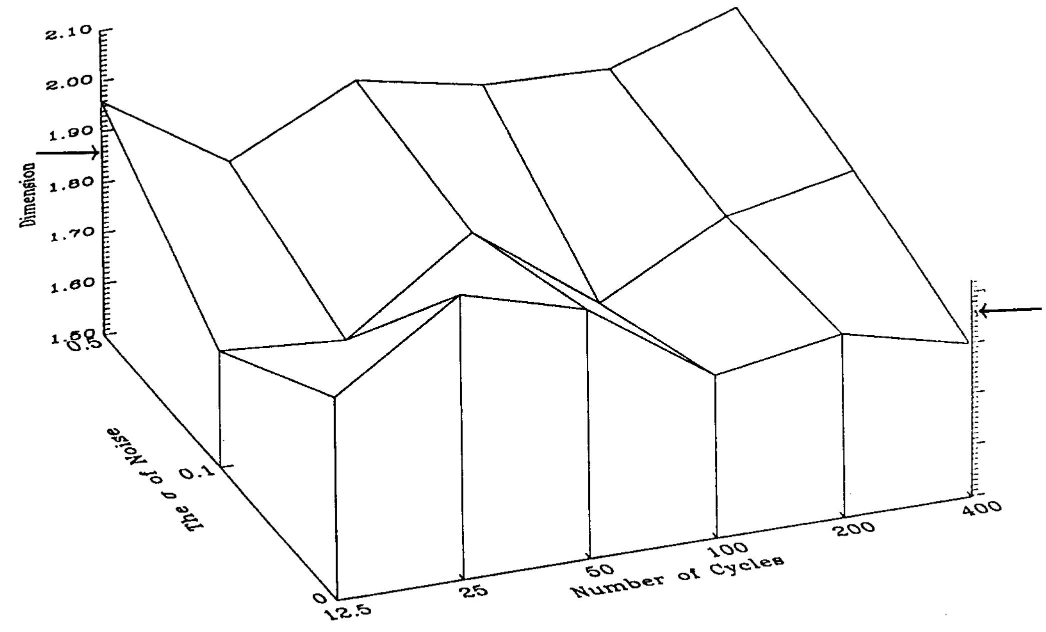

Fig. 5. Correlation dimension as a function of n_S, the number of cycles, and the variance σ of Gaussian noise added to the time series, for the X-coordinate of the Rössler system. The true value is marked by arrows.

fore be considered as a fast, easy to use, reliable and automatic aid in estimating correlation dimensions. Note that this limit $N \gtrsim 1000$ is well above the one derived from theoretical arguments in order to observe a plateau of width $2: N \gtrsim 2^{D^{(2)}/2}$ [15,16].

Two remarks aside. (a) Many authors agree that the number of points per cycle is well chosen whenever $10 \lesssim t_{corr}/\tau \lesssim 20$, with t_{corr} being the first minimum of the autocorrelation function. This is nicely seen in figure 4, where the dimension for $t_{corr}/\tau \approx 20$ fluctuates the least (cf. also ref. [17]). (b) Regarding the choice of the time delay Δt in the reconstruction procedure: the choice $\frac{1}{2} t_{corr}/\tau \leqslant d\Delta t/\tau \leqslant t_{corr}/\tau$ yields a set which is maximally spread, without the coordinates of the single vectors being decorrelated. The experience in our inquiry proves it to be adequate. Albano et al. [18] made a systematic inquiry (cf. also ref. [17]).

An interesting new finding is the following: The number of points clearly shows to be not the decisive quantity. The quality of the plateau depends much more on the *number of structures*, defined in the following way: We take again the first minimum of the autocorrelation as a definition of the autocorrelation time t_{corr}. This provides a relatively precise measure of the average length of the shortest existing temporal structure (a peak-like feature in our cases), leading to the fastest cycle in the reconstructed phase space. If a time series consists of N points, the number of structures n_S can be defined

$$n_S := N\tau / t_{corr} . \qquad (12)$$

This quantity and t_{corr}/τ, the number of points per structure – cycle in phase space – are the two ordinates in fig. 4. Clearly, n_S governs the behaviour. At $n_S = 50$, the error is smaller than 10% for the Rössler system and the dimension of the Mackey–Glass system is recognizable. For both attractors, the dimensions are within 10% of the respective true values as soon as

$$n_S \gtrsim 100 . \qquad (13)$$

Havstad and Ehlers [14] find in their numerical experiment an even lower limit for the number of points

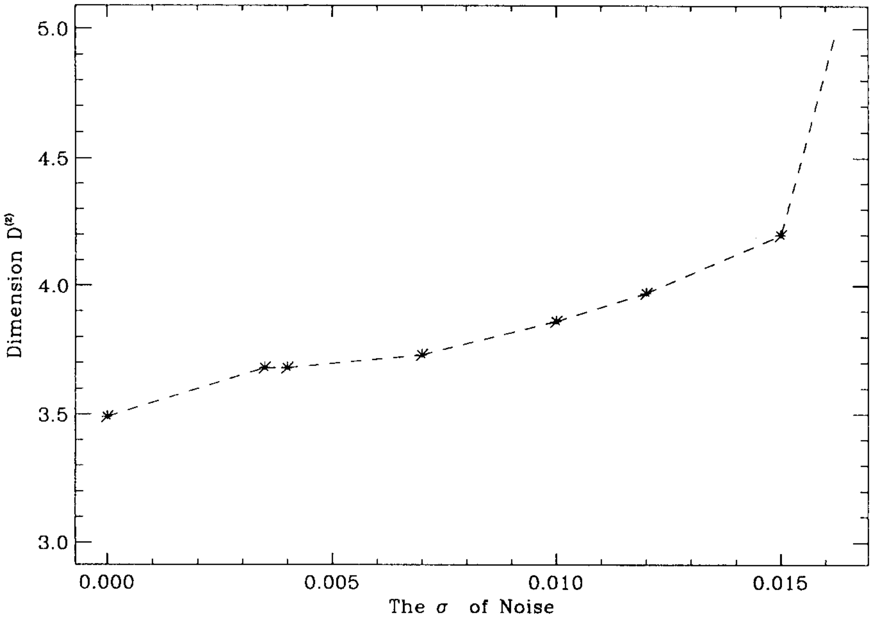

Fig. 6. Correlation dimension as a function of σ, the variance of Gaussian noise added to a time series which is generated by the Mackey–Glass system. The same units are used as in fig. 3. Its length is 4000 points, the time resolution τ is chosen so that $t_{corr}/\tau \approx 20$ and the number of cycles $n_S \approx 200$. For $\sigma > 0.015$ the correlation integral did not converge any more.

necessary, using a lot of care in selecting the points on the attractor, however.

The quantity n_S is also a rough estimate of the number of independent points in phase space, since t_{corr} is a measure for the duration of dynamical correlation. From theoretical arguments, Smith [9] claims that this number of independent points should fulfill $n_S \gtrsim (5.5\gamma_2/\gamma_1)^{D^{(2)}}$ in order that the dimension estimate has an accuracy of 10%. For the Rössler system ($\gamma_2/\gamma_1 \approx 3.5$) this yields $n_S \gtrsim 250$, and for the Mackey–Glass system $n_S \gtrsim 6500$ (where $\gamma_2/\gamma_1 \approx 2.25$). Both numbers are much higher than our empirically found limit of about 100.

5. Application II: the influence of noise

We show how the method is able to recognize plateaus spoilt this time not by shortness, but by noise. Gaussian noise of different variances σ was added to the generated time series, in two situations:

With the Rössler attractor we looked at the variation of the dimension with the number of structures n_S if an intermediate noise level is present (noise to signal ratio between 1% and 10%). Figure 5 presents the result for this attempt, where the added noise level is $\sigma=0$, $\sigma=0.1$ and $\sigma=0.5$. The time series has a quite well-defined average amplitude of about 7 (see fig. 2), so that σ corresponds to a noise to signal ratio of about 0%, 1.5% and 7%, respectively. The dimension estimates are spread around the true value, with no systematic deviation to be seen, and the fluctuations smaller than 10%.

With the Mackey–Glass attractor we increased the noise level for a fixed data length to follow the way how the estimated dimension changes and finally disappears. In fig. 6 is shown how sensitive to noise this attractor is. There is a transition zone for the σ's, where the dimension increases in a smooth way from its reasonable value to higher values. At $\sigma=0.0175$, the correlation integral did not converge anymore, noise dominates. The noise to signal ratio has no typical values, since the time series shows amplitudes on many quite different scales. For an average amplitude of 0.1, the $\sigma=0.0175$ corresponds to a noise level of 17.5%. At a noise level of about 10%, the deviation from the true value of the dimension gets larger than 10%.

6. Conclusions

A test of significance of a power law scaling region of the correlation integral is desired to overcome the somewhat subjective part of the method: identifying a plateau. A regression into the $\log C_d^{(2)}(\epsilon)$–$\log \epsilon$ relation would not help for it is always possible and the errors are not meaningful. A test is proposed here that uses the probabilistic character of the correlation integral and determines the significance of a scaling behaviour by a χ^2-test. That the test has to be adapted has strong theoretical reasons. The test is a completion of dimension estimate methods towards being an algorithm, i.e. an automatic proceeding. It is easy to implement, several conjectured scaling ranges can be tested, and their quality can be compared. The given applications prove that it is delivering, in the adjusted form, reliable and fast decisions. It efficiently reproduces dimensions with 10% accuracy under the condition that $n_S \gtrsim 50$ or 100, and that the noise level is below 10%. The plots of the applications are also practical references, illustrating typical behaviour of chaotic systems. They lead to the following finding: For small amounts of data the crucial parameter is not the data length but the number of cycles in phase space (peaks in the time series). This becomes plausible from the following two considerations:

First: A strange attractor is fractal just perpendicular to the trajectories $\xi(t)$: Locally, at the point $\xi(t_0)$, an attractor A looks like the product of a line segment (a piece $\xi(t)$ of the trajectory) and a fractal set $F_{\xi(t_0)}$ which is perpendicular to the trajectory, $F_{\xi(t_0)} \perp \xi(t)$: $A = F_{\xi(t_0)} \times \xi(t)$ [19]. The dimension of $F_{\xi(t_0)}$ can be expected to be the attractor dimension minus one, $D^{(2)} - 1$ [20]. Each cycle on the attractor will contribute a point to $F_{\xi(t_0)}$, so that intuitively the fractal structure of this set is the easier detectable, the more cycles there are given; and by that the fractal structure of the attractor as well.

Second: The number of structures roughly estimates the number of independent vectors in phase space which can be built from a time series (section 4). Both the two applied dimension estimate methods demand that the used vectors are chosen independently. Correlated vectors bias the estimate. Therefore, the number of structures will naturally influence the quality of a dimension estimates more than the number of points in the time series does. Smith [9] investigates conditions on the number of independent vectors. We find the limits he claims to be too high in the investigated cases (section 4).

The importance of the number of structures, which is related to the number of independent points in phase space, should be included in theoretical inquiries on the minimum number of necessary data points.

Acknowledgement

I would like to thank A.O. Benz and J. Kurths for helpful discussions and criticisms, and H. Scheingraber for passing on to me his experience. This work was supported by the Swiss National Science Foundation (grant No. 2000-5.499).

References

[1] A.R. Osborne and A. Provenzale, Physica D 35 (1989) 357.
[2] J. Theiler, Phys. Lett. A 155 (1991) 480.
[3] F. Takens, in: Lecture notes in mathematics, Vol. 898. Dynamical systems and turbulence (Springer, Berlin, 1981) p. 366.
[4] P. Grassberger and I. Procaccia, Physica D 9 (1983) 189.
[5] F. Takens, in: Lecture notes in mathematics, Vol. 1125. Dynamical systems and turbulence (Springer, Berlin, 1984) p. 99.
[6] S. Ellner, Phys. Lett. A 133 (1988) 128.
[7] R. Badii and A. Politi, Phys. Lett. A 104 (1984) 303.
[8] L.A. Smith, J.-D. Fournier and E.A. Spiegel, Phys. Lett. A 114 (1986) 465.
[9] L.A. Smith, Phys. Lett. A 133 (1988) 283.
[10] O.E. Rössler, Phys. Lett. A 57 (1976) 397.
[11] J. Kurths and H. Herzel, Physica D 25 (1987) 165.
[12] H. Atmanspacher, H. Scheingraber and W. Voges, Phys. Rev. A 37 (1988) 1314.
[13] N.B. Abraham, A.M. Albano, B. Das, G. de Guzman, S. Yong, R.S. Gioggia, G.P. Puccioni and J.R. Tredicce, Phys. Lett. A 114 (1986) 217.
[14] J.W. Havstad and C.L. Ehlers, Phys. Rev. A 39 (1989) 845.
[15] A. Brandstater and H.L. Swinney, Phys. Rev. A 35 (1987) 2207.
[16] D. Ruelle, Proc. R. Soc. A 427 (1990) 241.

[17] J. Kurths, A.O. Benz and M.J. Aschwanden, Astron. Astrophys. 248 (1991) 270.
[18] A.M. Albano, J. Muench, C. Schwartz, A.I. Mees and P.E. Rapp, Phys. Rev. A 38 (1988) 3017.
[19] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. 57 (1985) 617.
[20] K.J. Falconer, Fractal geometry, mathematical foundations and applications (Wiley, New York, 1990).