# Influence of noise on power-law scaling functions and an algorithm for dimension estimations

Hans Oltmans* and Peter J. T. Verheijen
*Department of Chemical Process Technology, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands*
(Received 22 November 1996; revised manuscript received 19 February 1997)

The influence of Gaussian noise on power-law scaling functions of interpoint distances has been investigated. These functions appear in the estimation of the correlation dimension $\alpha$ of the attractor of a chaotic dynamical system, where the relative number of pairwise distances smaller than $r$ (correlation integral) theoretically scales as $r^{\alpha}$. Assuming the noise added to each measurement is independent and the distribution of the distances is governed completely by the power-law scaling rule in the noise-free case, the scaling functions of the perturbed distances have been calculated exactly. By considering the limiting cases for small and large distances, a method is presented to estimate the variance of the added noise and approximations of the scaling functions, which are suitable for data analysis, are derived. Dimension estimation can be improved by applying a nonlinear fit procedure to histograms of interpoint distances instead of the usual linear regression on log-log plots. [S1063-651X(97)04506-6]

PACS number(s): 02.70.−c, 05.45.+b, 82.40.Bj

## I. INTRODUCTION

### A. Spatial correlation of reconstructed attractors

In the analysis of chaotic behavior of dynamical systems, the correlation dimension of an attractor is an important characteristic. Of the various definitions of fractal dimensions (see, e.g., [1]), it is the most accessible one for an experimentalist [2,3]. The concept of correlation dimension is based on the fact that the pairwise distance $r$ between points of the attractor satisfies the scaling rule

$$n(r) \propto r^{\alpha-1}, \tag{1}$$

which holds for small values of $r$. Here $n(r)$ is the number of pairs of points with distance $r$ and $\alpha$ is the correlation dimension. The correlation dimension is a measure of the geometric complexity of the dynamics of the system [4,5].

In experiments, the attractor is reconstructed from a time series of measurements of a certain variable, being ''representative'' for the state of the system, using the method of delay-time embeddings [6,7]. The correlation dimension is then often calculated by determining the relative number of points $C(r)$ with a range smaller than $r$ for various values of $r$ and plotting $\ln C(r)$ against $\ln r$. This should yield a graph of a straight line with slope $\alpha$. If the points are reconstructed from a time series, the dimension of the reconstruction space $d$, called embedding dimension, should be big enough for a good representation of the state of the system. In theory, the slope of the graph will converge to $\alpha$ with increasing embedding dimension [8].

For small values of $r$, a deviation from the power-law behavior is observed because measurements are perturbed by noise [9]. The curves bend off to a larger slope with an associated dimension $d$, which can be explained by the fact that, on small scales, the dynamics is determined by noise, which has an infinite dimension if it is uncorrelated noise. Figure 1 shows an example of the influence of Gaussian noise, with standard deviation varying from zero (noise-free case) to 0.1, on correlation integral curves of the Hénon attractor (a well-known numerical example that we will use here, with $\alpha \approx 1.22$). It can be observed that the scaling behavior is almost entirely destroyed for the biggest noise level.

In this paper, equations are derived for the deviation from the power law caused by adding uncorrelated Gaussian noise to the measurements. In contrast with many numerical experiments to determine the statistical properties of dimension estimations (see, e.g., [10,11]), a statistical approach is used here, assuming the power law to hold exactly in the unperturbed case. A useful review of statistical methods in dimension estimations is given by Isham [12].

### B. Outline of the approach

In Sec. II, analytical expressions will be derived for the (non-normalized) probability density functions (PDFs) of the distances in the presence of independent Gaussian noise, added to the measurements. These expressions give a more correct description of the scaling behavior than Eq. (1) does

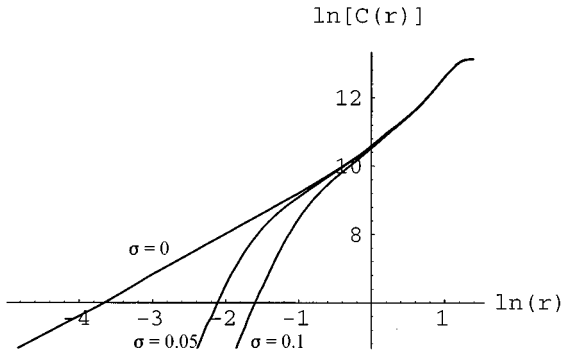


FIG. 1. Correlation integral curves for a time series of 10 000 points of the Hénon map with added Gaussian noise ($\sigma = 0, 0.05, 0.1$). The embedding dimension $d = 6$.

*Present address: Joh. Enschedé B.V., POB 464, 2000 AL Haarlem, The Netherlands.

 © 1997 The American Physical Society

and simulations can be made when the properties of the attractor and the standard deviation of the noise are known *a priori*. In the rest of this paper we will refer to the nonnormalized PDFs as ''scaling functions.'' We say, that $f_x(x)$ is a scaling function of the quantity $x$ if the PDF of the variable $x$ is proportional to $f_x(x)$ on a certain interval, say, $0<x<x_0$. Outside this interval, the PDF is unknown.

Asymptotic approximations for large and small values of $r$ are then derived, which can be useful in the analysis of measured data. As we shall see, the approximations are of the form

$$f_r(r)\approx\begin{cases} A(d,\alpha,\sigma)r^{d-1}P\left(\left(\frac{r}{\sigma}\right)^2\right) & (0<r<r_l) \\ r^{\alpha-1}Q\left(\left(\frac{\sigma}{r}\right)^2\right) & (r_h<r<r_0). \end{cases} \tag{2}$$

Here $\sigma$ is the standard deviation of the noise, $A$ is some constant, and $P(x)$ and $Q(x)$ are polynomials (truncated power series) of the form $1+p_1x+p_2x^2$, where the coefficients $p_k$ are dependent on $d$ and $\alpha$. The boundaries $r_l$ and $r_h$ demarcate the validity of the two approximations. All necessary input for the nonlinear regression procedure, adjusting Eq. (2) to fit the data, is given in Sec. III.

Section IV presents numerical experiments. The methods of estimating $\sigma$ and $\alpha$ are applied to a time series of the Hénon map and to a time series of pressure measurements in a fluidized bed column. Finally, in Sec. V conclusions are drawn about the usefulness of the analysis.

## II. THEORY

### A. Perturbation of fixed Euclidean distance by added noise

It has been illustrated above that the distribution of distances is governed by a cumulative distribution function (CDF) that is proportional to $r^\alpha$ on the interval $r<r_0$, the ''scaling region.'' In dimension estimations it is usual to compute the sample CDF (or ''correlation integral'') $C(r)$ from the data and obtain $\alpha$ as the slope of a log-log plot. Alternatively, a maximum likelihood estimate for $\alpha$ as a function of $r_0$ can be found [4].

In these procedures, all distances larger than $r_0$ are discarded. However, the introduction of this cutoff length is somewhat artificial. In this theoretical treatment the power law is assumed to hold exactly for all distances. We want to calculate the effect of the addition of noise on the scaling behavior. Therefore, $r_0$ is supposed to be infinite and we use the (unbounded) scaling functions

$$f_r(r)=r^{\alpha-1}. \tag{3}$$

Any fractal dimension is invariant to the choice of the norm. There are, for example, practical advantages in taking the $L^\infty$ norm [13,14]. However, the Euclidean norm (or $L^2$ norm) is preferred because its square value $\Sigma_{i=1}^d(x_i-y_i)^2$, considered as a stochastic variable, has a known distribution in many cases. Here $x_i, y_i$ are the Cartesian coordinates of the points $\vec{x},\vec{y}$ in $d$-dimensional space. In the case that the points are not perturbed by noise, the scaling function of the squared distance $s:=r^2$ equals

$$g(s)=f_r(\sqrt{s})\left|\frac{d\sqrt{s}}{ds}\right|=\frac{f_r(\sqrt{s})}{2\sqrt{s}}=\frac{1}{2}s^{\alpha/2-1}. \tag{4}$$

In the present study it is assumed that each coordinate of every point (each measurement) is perturbed by adding independent normally distributed noise with mean 0 and variance $\sigma^2$. Here the perturbed quantities (which are observed in an experiment) will be denoted by a tilde, for example,

$$\tilde{x_i}=x_i+\epsilon_{x_i}, \tag{5}$$

with $\epsilon_{x_i}\sim N(0,\sigma^2)$. The observed value of the squared distance is

$$\tilde{s}=\tilde{r}^2=\sum_{i=1}^d(\tilde{x_i}-\tilde{y_i})^2=\sum_{i=1}^d(x_i-y_i+\epsilon_i)^2$$
$$=s+2\sum_{i=1}^d(x_i-y_i)\epsilon_i+\sum_{i=1}^d\epsilon_i^2, \tag{6}$$

where $\epsilon_i:=\epsilon_{x_i}-\epsilon_{y_i}$, which has a $N(0,2\sigma^2)$ distribution. The second and third terms on the right-hand side (RHS) of this equation describe the perturbation of the square of the Euclidean distance $r$ due to the addition of the noise. We deduce directly from the distribution of $\epsilon_i$ that

$$E(\tilde{s})=s+2\sigma^2d,$$
$$\mathrm{Var}(\tilde{s})=8\sigma^2(s+\sigma^2d). \tag{7}$$

In this calculation we have used the fact that the odd moments of a standardized normal variable $\delta\sim N(0,1)$ vanish and the even moments are given by $E(\delta^{2n})=(2n-1)!!:=\Pi_{m=1}^n(2m-1)$.

So, on average, the square of the distance between two points increases due to the addition of noise. This can be checked easily for the case $d=1$. Each direction has a contribution of $2\sigma^2$.

### B. Derivation of perturbed scaling functions

Until now, the distance $r$ had been assumed to be fixed. We know, however, that it satisfies the scaling rule [Eq. (1) or (3)] and try to find the corresponding scaling law for $\tilde{r}$, the perturbed distance.

For convenience, a coordinate transformation is applied such that

$$(x_i-y_i)=\begin{cases} r & (i=d) \\ 0 & (i\neq d). \end{cases} \tag{8}$$

This means that the difference vector $\vec{x}-\vec{y}$ is taken in an arbitrary direction (the $x_d$ direction). This is allowed because there is rotation invariance on account of the independence of the noise. The expression for the observed squared distance of Eq. (6) can then be simplified to

$$\tilde{s}=(r+\epsilon)^2+\sum_{i=1}^{d-1}\epsilon_i^2, \tag{9}$$

where the index $d$ has been omitted for notational convenience. To investigate scaling properties of $\tilde{s}$, we will first calculate the scaling function of the quantity

$$u := |r+\epsilon| \tag{10}$$

under the assumption that $r$ and $\epsilon$ are independent stochastic variables. Subsequently, the second term on the RHS of Eq. (9) is considered. The PDF of $\epsilon$ is given by

$$p(\epsilon) = \frac{1}{2\sigma\sqrt{\pi}} e^{-\epsilon^2/4\sigma^2} \tag{11}$$

and the scaling function $f_r(r)$ follows from Eq. (3). Although this is no PDF, the influence of the addition of $\epsilon$ to $r$ can be studied by treating it as being a probability density function. Our basic assumption is that $r$ and $\epsilon$ are independent stochastic variables. Therefore, to obtain the scaling function $f_u(u)$ of $u$, the function $p(\epsilon)f_r(r)$ is integrated along curves of constant $u$ in the $(\epsilon,r)$ semiplane. A constant value of $u$ corresponds to two half lines in this plane, leaving the $\epsilon$ axis from the points $(-u,0)$ and $(u,0)$ at an angle of $3\pi/4$. The integral is

$$\begin{aligned} f_u(u) &= \int_{-\infty}^{-u} p(\epsilon) f_r(-u-\epsilon)\,d\epsilon + \int_{-\infty}^{u} p(\epsilon) f_r(u-\epsilon)\,d\epsilon \\ &= \int_0^{\infty} f_r(r) p(-u-r)\,dr + \int_0^{\infty} f_r(r) p(u-r)\,dr \\ &= \frac{1}{2\sigma\sqrt{\pi}} \int_0^{\infty} r^{\alpha-1} [e^{-(r+u)^2/4\sigma^2} + e^{-(r-u)^2/4\sigma^2}]\,dr. \end{aligned} \tag{12}$$

Although $f_r(r)$ is not bounded, these integrals do converge because for large values of $r$ the exponential functions decrease faster than the reciprocal of the power function $r^{\alpha-1}$ (for any $\alpha>0$). Note that since $u$ is non-negative, the second term on the RHS is much larger than the first because the exponential function attains its maximum at $r=u$. The integral has been solved using MATHEMATICA [15]:

$$\begin{aligned} f_u(u) = \frac{\sigma^{\alpha-1}}{2\sqrt{\pi}} e^{-u^2/4\sigma^2} \frac{u}{\sigma} \Bigg[ &\Gamma(\alpha) U\left(\frac{1+\alpha}{2}, \frac{3}{2}, \frac{u^2}{4\sigma^2}\right) \\ &+ 2^{\alpha}\Gamma\left(\frac{1+\alpha}{2}\right) M\left(\frac{1+\alpha}{2}, \frac{3}{2}, \frac{u^2}{4\sigma^2}\right)\Bigg]. \end{aligned} \tag{13}$$

Here $U(a,b,z)$ [sometimes denoted as $\Psi(a,b,z)$] is the confluent (or degenerate) hypergeometric function and $M(a,b,z)$ [sometimes denoted as $\Phi(a,b,z)$ or ${}_1F_1(a,b,z)$] is Kummer's confluent hypergeometric function (see [16], p. 504, and [17], pp. 337 and 1064). Using a relation between the functions $U(a,b,z)$ and $M(a,b,z)$ and the doubling formula of the gamma function (see [17], pp. 1058 and 938), the function $f_u(u)$ is written in a more compact form

$$f_u(u) = \frac{(2\sigma)^{\alpha-1}}{\sqrt{\pi}} e^{-u^2/4\sigma^2} \Gamma\left(\frac{\alpha}{2}\right) M\left(\frac{\alpha}{2}, \frac{1}{2}, \frac{u^2}{4\sigma^2}\right). \tag{14}$$

Though the RHSs of Eqs. (13) and (14) are identical, in some calculations it is more convenient to use the more complicated expression of Eq. (13), which avoided artificial singularities using MATHEMATICA, version 2.0.

Note that $u$ equals the observed distance in the case $d=1$; see Eqs. (9) and (10). For $d>1$, the influence of the remaining noise terms $\epsilon_1,\ldots,\epsilon_{d-1}$ will be taken into account. The scaling function of $u^2=(r+\epsilon)^2=:w$ follows from Eq. (14) by the transformation $f_w(w) = f_u(\sqrt{w})/(2\sqrt{w})$:

$$f_w(w) = \frac{(2\sigma)^{\alpha-1}}{2\sqrt{\pi w}} e^{-w/4\sigma^2} \Gamma\left(\frac{\alpha}{2}\right) M\left(\frac{\alpha}{2}, \frac{1}{2}, \frac{w}{4\sigma^2}\right). \tag{15}$$

The sum of the remaining noise terms in Eq. (9) is denoted by $v$. The quantity $v/2\sigma^2$ has a $\chi^2$ distribution with $d-1$ degrees of freedom. The PDF of $v$ is given by

$$f_v(v) = \frac{1}{2\sigma^2} f_{\chi^2_{d-1}}\left(\frac{v}{2\sigma^2}\right) = \frac{v^{(d-3)/2}}{(2\sigma)^{d-1}\Gamma\left(\frac{d-1}{2}\right)} e^{-v/4\sigma^2}. \tag{16}$$

This is called a gamma distribution with scale parameter $1/4\sigma^2$ and shape parameter $(d-1)/2$.

The observed squared distance is the sum of the non-negative quantities $w$ and $v$. They are independent because it has been assumed that the noise contributions are independent of each other and of $r$. Consequently, the scaling function is calculated from

$$f_{\tilde{s}}(\tilde{s}) = \int_0^{\tilde{s}} f_w(w) f_v(\tilde{s}-w)\,dw = \int_0^{\tilde{s}} f_w(\tilde{s}-v) f_v(v)\,dv. \tag{17}$$

Substitution of the functions of Eqs. (15) and (16) in Eq. (17) finally gives an expression for the scaling function of $\tilde{s}$, the observed squared distance:

$$\begin{aligned} f_{\tilde{s}}(\tilde{s}) &= \frac{1}{2(2\sigma)^{d-\alpha}} \frac{\Gamma\left(\frac{\alpha}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d-1}{2}\right)} e^{-\tilde{s}/4\sigma^2} \\ &\quad \times \int_0^{\tilde{s}} M\left(\frac{\alpha}{2}, \frac{1}{2}, \frac{w}{4\sigma^2}\right) \frac{(\tilde{s}-w)^{(d-3)/2}}{\sqrt{w}}\,dw \\ &= \frac{\tilde{s}^{d/2-1}}{2(2\sigma)^{d-\alpha}} \frac{\Gamma\left(\frac{\alpha}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} M\left(\frac{d-\alpha}{2}, \frac{d}{2}, -\frac{\tilde{s}}{4\sigma^2}\right). \end{aligned} \tag{18}$$

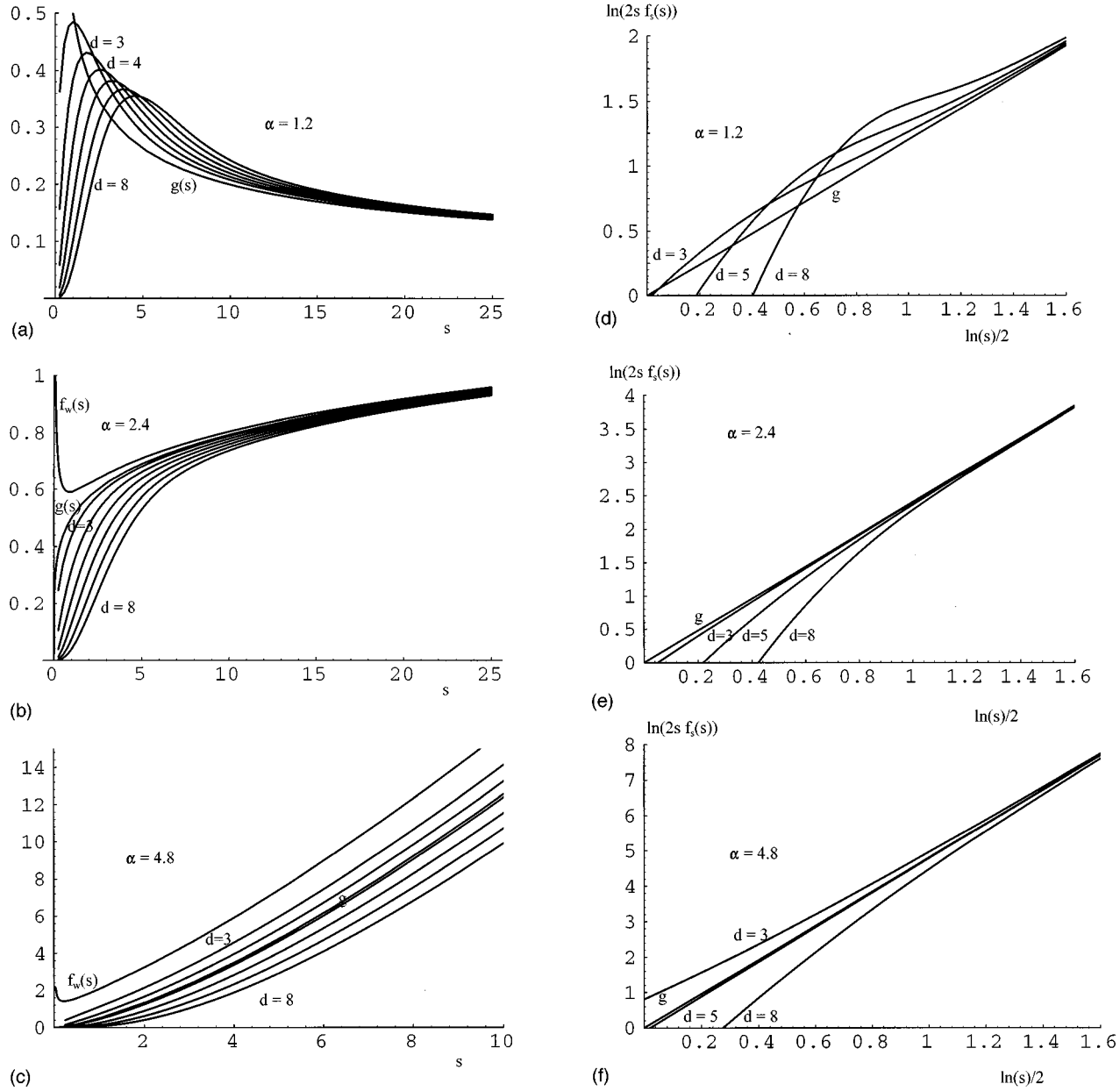


FIG. 2. (a), (b), and (c) Graphs of $f_s(s)$, the scaling function of the observed squared distance, for $\alpha=1.2,\ 2.4,\ 4.8$, and various values of $d$ with $\sigma=1/2$. Also $g(s)=1/2s^{\alpha/2-1}$, the scaling function of the exact squared distance [Eq. (4)], and $f_w(s)$, the scaling function of $(r+\epsilon)^2$ [Eq. (15)], are plotted. (For $\alpha=1.2$, the two functions are too close for both to be shown.) (d), (e), and (f) Graphs for $d=3,5,8$, plotted on a double logarithmic scale.

Figure 2 shows graphs of this function for various values of $\alpha$ and $d$ (with $d>\alpha$); $\sigma$ is kept constant at 1/2 because the only effect of changing $\sigma$ is a rescaling of the graphs: If $f_{\tilde{s}}(\tilde{s})$ for $\sigma=\frac{1}{2}$ is denoted by $\hat{f}(\tilde{s})$, then it follows that $f_{\tilde{s}}(\tilde{s})=(2\sigma)^{\alpha-2}\hat{f}(\tilde{s}/4\sigma^2)$. The distance has the same units as the measured variable $x$; in this text, however, for generality, the distance and the scaling functions are treated as dimensionless variables.

It is common to visualize the scaling function

$$f_{\tilde{r}}(\tilde{r})=2\tilde{r}f_{\tilde{s}}(\tilde{r}^2) \tag{19}$$

as a function of the distance $\tilde{r}$, with a double logarithmic scale. This should be a straight line with slope $\alpha$ for $\sigma\to 0$, as confirmed in the graphs of Fig. 2.

### C. Limiting behavior of scaling functions

The expression of Eq. (18) is rather complicated. Although some presently available fit programs can deal with almost any model function, it is useful to obtain some insight in the global properties of the scaling functions by considering their asymptotic behavior. In doing so we gain further understanding of the general shape of the curves in Fig. 2.

Here we will give approximations of the scaling function, which are valid for small ($r<r_l$) and large ($r>r_h$) values of the distance, compared to the noise level. For the moment,

we do not specify the values of $r_l$ and $r_h$ exactly in terms of $\sigma$ and $d$, but we assume that they are both ''of the order of $\sigma$.'' By using the Taylor-series expansion of Kummer's function (see [16], p. 504), we obtain the approximations of the scaling function $f_s(s)$ of Eq. (18) in both regions; $f_r(r)$ then follows from the transformation of Eq. (19). From now on, the tilde is omitted and we will only consider the observed quantities: So $r$ denotes the observed distance and $s$ its square. The resulting approximations are

$$f_r(r)=\frac{\Gamma\left(\frac{\alpha}{2}\right)}{(2\sigma)^{d-\alpha}\Gamma\left(\frac{d}{2}\right)}r^{d-1}\left[1-\frac{d-\alpha}{d}\left(\frac{r^2}{4\sigma^2}\right)+\frac{(d-\alpha)(d-\alpha+2)}{2d(d+2)}\left(\frac{r^2}{4\sigma^2}\right)^2+\cdots\right] \tag{20}$$

for small distances $(0<r<r_l)$ and

$$f_r(r)=r^{\alpha-1}\left[1-\frac{(\alpha-2)(d-\alpha)\sigma^2}{r^2}+\frac{(\alpha-2)(\alpha-4)(d-\alpha)(d-\alpha+2)\sigma^4}{2r^4}+\cdots\right] \tag{21}$$

for large distances $(r_h<r<r_0)$. Note that the lowest-order approximation is $r^{\alpha-1}$ for large distances and a constant times $r^{d-1}$ for small distances, as expected. The constant in Eq. (20) is the scale factor between the short-distance and the long-distance power relation. This factor is a function of $d$, $\alpha$, and $\sigma$.

For large distances, the first two terms of the expansion of the scaling function of the squared distance,

$$f_s(s)=\frac{s^{\alpha/2-1}}{2}\left[1-\frac{(\alpha-2)(d-\alpha)\sigma^2}{s}+\cdots\right], \tag{22}$$

are equal to the first two terms of the expansion of the function $\frac{1}{2}[s-(d-\alpha)2\sigma^2]^{\alpha/2-1}$. This is just $g(s-(d-\alpha)2\sigma^2)$, where $g(\,)$ is the unperturbed scaling function as introduced in Eq. (4):

$$g(s-(d-\alpha)2\sigma^2)=g(s)-g'(s)(d-\alpha)2\sigma^2+\cdots=\frac{s^{\alpha/2-1}}{2}\left[1-\frac{(\alpha-2)(d-\alpha)2\sigma^2}{2s}+\cdots\right]. \tag{23}$$

It follows that, in the first-order approximation for large distances, the scaling functions are shifted versions of the noise-free scaling function of Eq. (4), the shift being equal to $(d-\alpha)2\sigma^2$. This property can be verified by inspection of Fig. 2.

In an application, numerical values for the boundaries of the regions, in which our approximations are valid, have to be chosen. Because no general directives exist, we have used some more or less heuristic arguments.

The boundary $r_0$, the distance below which the power-law scaling rule is valid, is generally unknown. It is certainly much smaller than the total diameter of the attractor $\max_{m,n}|\vec{x}(m)-\vec{x}(n)|$, which is, of course, a function of the embedding dimension. We determine $r_0$ by inspection of the usual correlation integral curves.

If the approximated scaling functions are used instead of the exact model function, we also need values for $r_l$, the distance below which the distribution is ''governed'' by the noise, and $r_h$, the distance above which the influence of the noise becomes negligible. The boundaries $r_l$ and $r_h$ depend on $d$ and $\sigma$. The latter dependence makes things complicated because $\sigma$ is one of the parameters to be estimated.

In our analysis, the boundaries have been determined by using global properties of the $\chi^2$ distribution. If the distances were caused by noise alone, which can be assumed to be true for the distances that are originally very small, then the quantity $x:=r^2/2\sigma^2$ has a $\chi^2$ distribution with $d$ degrees of freedom. Then $E(x)=d$ and $\mathrm{Var}(x)=2d$. This means that the ''typical noise distance'' is $r_n:=\sigma\sqrt{2d}$ and its ''spread around the mean'' is approximated by $\Delta r_n=\sqrt{2}\sigma\Delta\sqrt{x}\approx\sqrt{2}\sigma\Delta x/2\sqrt{x}$, which just equals $\sigma$ if $x$ is replaced by its mean and $\Delta x$ by the standard deviation. The following boundary values now appear to be plausible:

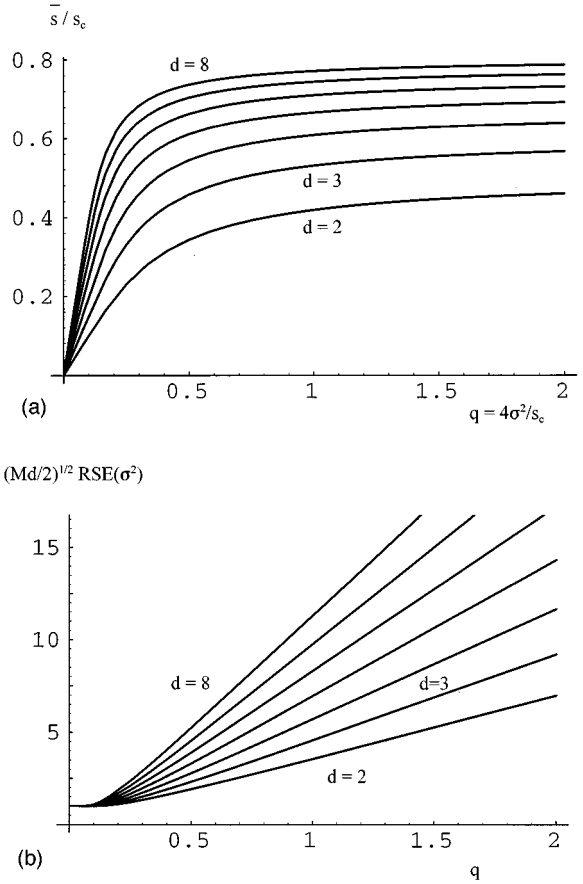


FIG. 3. (a) Plots of $\overline{s}/s_c$ against $q=4\widehat{\sigma^2}/s_c$ for various values of the embedding dimension $d$ [see Eq. (28)]. Using these curves, the maximum likelihood value of the variance can be determined from the cutoff value of the squared distance $s_c$ and the corresponding sample mean $\overline{s}$. (b) Plots of the RHS of Eq. (29) multiplied by $\sqrt{Md/2}$ as a function of $q=4\widehat{\sigma^2}/s_c$.

$$r_l = r_n F_l = \sigma\sqrt{2d}F_l,$$

$$r_h = r_n + \Delta r_n F_h = \sigma(\sqrt{2d} + F_h), \tag{24}$$

where $F_l < 1$ and $F_h > 1$ are constant factors. The choice of these factors is a trade-off between being safe concerning the validity of the Taylor-series approximation, on the one hand, and including enough data points, on the other hand. In Sec. IV we assume that $\sigma$ is much smaller than $r_0$ such that both regions of interest exist.

### D. Estimation of the variance

Consider the region of small distances $r \ll r_l$. In this region, the value of the observed distance is almost only determined by noise. Equation (9) then becomes

$$s \approx \sum_{i=1}^{d} \epsilon_i^2 . \tag{25}$$

As mentioned before, the RHS of this equation divided by $2\sigma^2$ has a $\chi^2$ distribution with $d$ degrees of freedom. This suggests that the smallest values of the squared distance can be used to estimate the noise variance $\sigma^2$.

We take a large sample $S = \{s_i | s_i < s_c\}$ of measured squared distances. The upper bound $s_c$ is the maximal value of $s$, for which Eq. (25) holds. The PDF of each observation $s$ is

$$f_s(s) = \begin{cases} \dfrac{\left(\dfrac{s}{4\sigma^2}\right)^{d/2-1}}{4\sigma^2 \gamma\left(\dfrac{d}{2}, \dfrac{s_c}{4\sigma^2}\right)} e^{-s/4\sigma^2} & (0 \leqslant s \leqslant s_c) \\ 0 & (s > s_c), \end{cases} \tag{26}$$

the cutoff version of the gamma distribution with scale parameter $1/4\sigma^2$ and shape parameter $d/2$. The function $\gamma$, appearing in the normalizing constant, is the incomplete gamma function (see [16], p. 260). The estimation of both parameters of a gamma distribution simultaneously from lowest-order statistics is described by Wilk *et al.* [18]. Here we have the special case that the shape parameter $d/2$ is known.

We choose the maximum likelihood estimator (MLE) (see, e.g., [19], p. 305), which is found by maximizing the log-likelihood function

$$L(\sigma^2) = \ln \prod_i f_s(s_i)$$

$$= C - M \ln \gamma\left(\frac{d}{2}, \frac{s_c}{4\sigma^2}\right) - M \frac{d}{2} \ln \sigma^2 - \frac{1}{4\sigma^2} \sum_i s_i \tag{27}$$

with respect to $\sigma^2$. Here $C$ is a constant independent of $\sigma^2$ and $M := \#S$, the number of samples smaller than $s_c$. The result is an implicit relation for the maximum likelihood estimator $\widehat{\sigma^2}$ for the variance $\sigma^2$,

$$\frac{\overline{s}}{s_c} = q\left[\frac{d}{2} - \frac{e^{-1/q}}{q^{d/2}\gamma\left(\dfrac{d}{2}, \dfrac{1}{q}\right)}\right], \tag{28}$$

where $\overline{s} := (1/M)\Sigma_i s_i$ is the sample mean and $q := 4\widehat{\sigma^2}/s_c$. The theory of maximum likelihood estimators also provides an expression for the minimum variance bound, an asymptotic approximation for the variance of a MLE (see [19], p. 308). The relative standard error ($R^{\mathrm{RSE}}$) of $\widehat{\sigma^2}$ is found to be

$$R^{\mathrm{RSE}}(\widehat{\sigma^2}) := \frac{\sqrt{\mathrm{Var}(\widehat{\sigma^2})}}{\widehat{\sigma^2}} \approx \frac{1}{\sqrt{M\left[\dfrac{e^{-1/q}}{q^{d/2}\gamma\left(\dfrac{d}{2}, \dfrac{1}{q}\right)}\left(\dfrac{d}{2} - 1 - \dfrac{1}{q} - \dfrac{e^{-1/q}}{q^{d/2}\gamma\left(\dfrac{d}{2}, \dfrac{1}{q}\right)}\right) + \dfrac{d}{2}\right]}}. \tag{29}$$

Figure 3 shows graphs of the function on the RHS of Eq. (28) and the relative standard error of Eq. (29) for several values of $d$. Given a value of $s_c$ and the corresponding sample mean $\overline{s}$, the MLE of $\sigma^2$ and the associated error can be read from these graphs.

The formulas above are valid if the only component in the signal is the noise. In Sec. II E it is shown that, in the first-order approximation, the value estimated here is really $\sigma^2 d/(d-\alpha)$ instead of $\sigma^2$ itself. This implies that from an observation of $\overline{s}$ and Eq. (28) an estimate for $q$ is found and

$$\widehat{\sigma^2} = \frac{s_c(d-\alpha)}{4d} q. \tag{30}$$

To check the result, this should be repeated for various values of $s_c$ and $d$. For increasing $d$, the estimation will improve because the influence of the noise is increased.

The graphs of Fig. 3(a) have a limit $d/(d+2)$ for large $q$. In this region they run almost horizontally and the estimation is very sensitive to small changes in the sample mean. The estimate is most reliable if it corresponds to a point in the steep part of the graph, say, $q < 1$. So $s_c$ cannot be taken too small, which means that $\sigma^2$ cannot be estimated from the leftmost part of the distribution. On the other hand, $s_c$ must be small enough for Eq. (25) to hold. If $s_c$ is taken too big, the $s$ values are not governed by the noise with dimension $d$, but by the scaling law of the underlying attractor. Then the limiting value of $\overline{s}/s_c$ will be $\alpha/(\alpha+2)$, independent of the embedding dimension and the noise level. Consequently, a

constant value of the MLE of the variance can only be expected in a limited region of $s_c$ values.

### E. Influence of scaling behavior on variance estimation

The determination of the variance of the noise by the maximum likelihood estimate is only correct in order of magnitude. There are two sources of error. First, the estimation is calculated from a limited number of observations, which is inevitable; this statistical error is given by Eq. (29). In addition, we have assumed that the distance on small scales is only determined by the noise; see Eq. (25). The scaling rule underlying the (unperturbed) values of the distance was neglected because the correlation dimension was unknown. However, it is possible to correct the result of the estimation by taking into account the influence of $\alpha$.

To derive a MLE, the scaling function has to be transformed into a PDF by cutting off at $s_c$ and normalizing. If only the zeroth- and first-order terms of the polynomial in Eq. (20) are taken into account and the truncated result is written as an exponential function having the same terms up to first order, only the first-order influence of $\alpha$ is calculated. The normalized PDF becomes

$$f_s(s)=\begin{cases}\dfrac{\left(\dfrac{s}{4\sigma^2}\right)^{d/2-1}}{4\sigma^2\left(\dfrac{d}{d-\alpha}\right)^{d/2}\gamma\left(\dfrac{d}{2},\dfrac{d-\alpha}{d}\dfrac{s_c}{4\sigma^2}\right)}e^{-[(d-\alpha)/d](s/4\sigma^2)} & (0\leqslant s\leqslant s_c)\\ 0 & (s>s_c).\end{cases} \tag{31}$$

This approximation is only valid if the first-order term is small compared to unity, so that $s/4\sigma^2 \ll d/(d-\alpha)$. It is easy to check, by comparing with the derivation following Eq. (26), that the MLE for $\sigma^2$ resulting from this PDF is just $(d-\alpha)/d$ times the value calculated before. So $(d-\alpha)/d$ is the ''first-order correction factor.'' This factor is close to unity for large $d$, as expected. For small $d$, it improves the estimate of the variance considerably, as can be seen from the examples.

## III. COMPUTATIONS TO ANALYZE EXPERIMENTAL DATA

A typical time series consists of $10^4-10^5$ data points. It is sometimes necessary to omit the first fraction of the data to avoid the influence of transients. The procedure is as follows:

(i) Compose a series of reconstruction vectors from the time series by the method of delay-time embeddings and generate a sample of distances.

(ii) Determine the MLE noise variance with Eq. (28) from only the short distances.

(iii) Use the short-distance data and long-distance data together to estimate the variance and the dimension with a nonlinear regression method.

In the second step, the actual distances are utilized. After that, the data are sorted in equidistant bins. The actually fitted observations are the number of occurrences in each bin and the error is its square root, implying conventional Poisson statistics. The fits can be made for each embedding dimension separately, but also for all embedding dimensions simultaneously. This section describes the computational implementation of these steps.

### A. Reconstructing the attractor

In the ideal case, for example, with known system equations, a multivariate time series with dimension $d$ is available that completely characterizes the state of the system. Then the state space vectors $\vec{x}(n)=[x_1(n),\ldots,x_d(n)]^T$ (with $n=1,\ldots,N$) can be used in the dimension calculations and $d$ is the proper embedding dimension. Here $N$ is the length of the recorded time series. For the practical case of a univariate time series $x(n)$, reconstruction vectors $\vec{x}(n)=[x(n),x(n+1),\ldots,x(n+d-1)]^T$ can be composed for various values of $d$. In the case of reconstruction vectors, the assumption that all noise contributions are independent is not strictly true because the $k$th coordinate of $\vec{x}(n)$ equals the $(k-1)$th coordinate of $\vec{x}(n+1)$ $(k=2,\ldots,d)$ and so on. Consequently, the corresponding noise contributions are equal. However, this will only affect $r=|\vec{x}(m)-\vec{x}(n)|$ if $|m-n|<d$. These pairs are excluded from the distance simulations to avoid this ''diagonal effect.''

### B. Inspecting the MLE for the noise variance

The MLE method above gives a pre-estimate of the noise variance. From a large sample of distances, $\hat{\sigma}^2$ is calculated for various values of $s_c$ by solving Eq. (28). If we plot $\hat{\sigma}^2$ as a function of $s_c$ (on a logarithmic scale), three regions with different behavior can be expected: (i) The left part of the graphs (''$s_c$ too small'') will be capricious as a consequence of the finite number of events simulated, (ii) in the center of the graphs we have the ''plateau'' we are looking for, and (iii) the right-hand part (''$s_c$ too big'') is an exponentially increasing function due to the underlying power-law scaling. The interpretation is done by visual inspection, which is sufficient here, as it is only used for an initial guess.

### C. Nonlinear fit procedure

In an experimental situation, a large sample of interpoint distances is generated. A histogram of these distances can be compared to the scaling function found above:

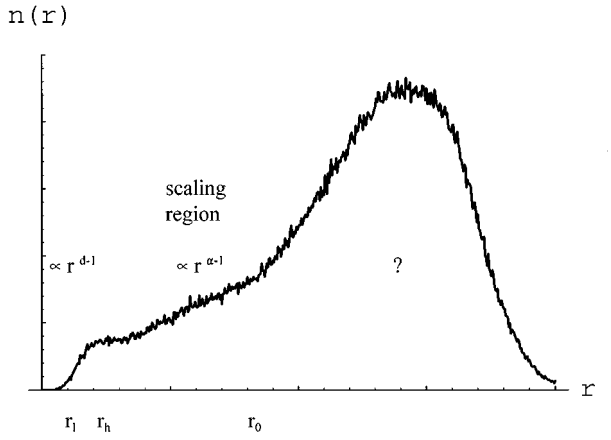


FIG. 4. Sketch of the general behavior of the PDF of the interpoint distances of the attractor of a chaotic system in the presence of noise.

$$n(r)=cf_r(r) \quad (r<r_0), \tag{32}$$

where $n(r)$ is the number of distances between $r-\delta r$ and $r$. In addition to the two parameters of the scaling function $\alpha$ and $\sigma$, also the factor of proportionality $c$ is unknown. It depends on the embedding dimension $d$ and the bin width $\delta r$.

Since the scaling function is essentially nonlinear in the parameters, a nonlinear fit procedure is required to determine the parameters from the sample histogram. We have used the nonlinear fit program PASTIFIT, which is based on the algorithm of Marquardt [20]. As mentioned in Sec. II C, it is possible to fit the data using the exact scaling function of Eq. (18) as the model. Because Kummer's function is not commonly available, it is replaced by its integral representation (see [16], p. 505) in the implementation,

$$cf_r(r)=\frac{c}{(2\sigma)^{d-\alpha}}\frac{r^{d-1}}{\Gamma\left(\frac{d-\alpha}{2}\right)} \times\int_0^1 e^{-(r^2/4\sigma^2)t}t^{(d-\alpha)/2-1}(1-t)^{\alpha/2-1}dt, \tag{33}$$

which involves numerical integration. This is a time-consuming process when many data points are concerned. To obtain a procedure that is easy to use on a routine basis, the approximated scaling functions given in Eqs. (20) and (21) are applied. Because they are valid in certain regions, we get a piecewise nonlinear model function [21]. The fit can be optimized by repeating it for various values of the region boundaries by adjusting the parameters $F_l$ and $F_h$ in Eq. (24).

The global behavior of the PDF of the distances is sketched in Fig. 4. For a nonlinear fit to succeed, preestimates (initial values) are needed, which are correct at least in order of magnitude. The initial value of the proportionality factor $c$ is not very critical, but those of $\alpha$ and $\sigma$ are. We use (a) the usual estimate of $\alpha$ from the slope of a log-log plot and Eq. (1) and (b) the MLE of $\sigma^2$ as described above.

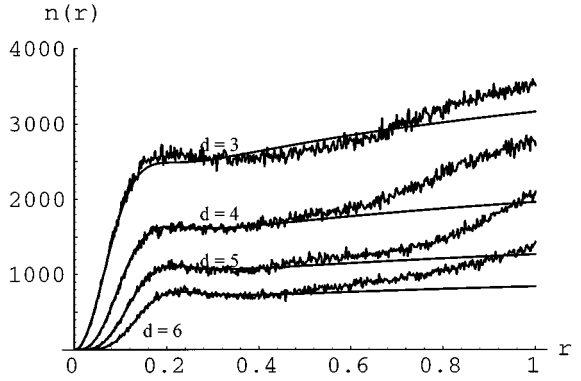


FIG. 5. Comparison of sample histograms with the model functions, for the Hénon map ($\alpha=1.22$). In this example $\sigma=0.05$, $d=3,4,5,6$, and $r_0=0.5$. The sample size $N_p=5\times10^6$. The bin width of the histograms is 0.002.

## IV. EXAMPLES

The following simulations and calculations have been done with the help of MATHEMATICA and MATLAB.

### A. The Hénon map

A well-known example of a dynamical system exhibiting chaotic behavior is the Hénon map [22]. The correlation dimension of the Hénon attractor is approximately 1.22. It has been calculated numerically by (among others) Grassberger and Procaccia [5].

The numerical experiment consisted of generating a time series of $10^4$ points by iterating $x(n)=1-1.4[x(n-1)]^2+0.3x(n-2)$. To each point, independent normally distributed noise has been added. Subsequently, a large sample of squared distances has been generated by repeatedly picking two points $m,n$ at random, taking into account the abovementioned diagonal effect, and calculating the Euclidean distance for various values of $d$.

In Fig. 5, plots of the sample histograms of the distances for $\sigma=0.05$ and $d=3,\ldots,6$ are shown (jagged curves). The bin width is 0.002. The histograms contain approximately 0.39 up to $1.31\times10^6$ events, smaller than 1, from a total sample size of $N_p=5\times10^6$. A comparison with the scaling functions requires appropriate normalization. The smooth curves in Fig. 5 are model functions, calculated from Eqs. (18) and (19), with $\alpha=1.22$ and $\sigma=0.05$, and normalized such that the area under the curve between 0 and $r_0$ is the same as that of the histogram:

$$c\int_0^{r_0} f_r(r)dr=C(r_0)\,\delta r, \tag{34}$$

which gives $c$ for each $d$. Here $C(r_0)$ is the number of distances smaller than $r_0$. The graphs exhibit a good match with the sample histograms.

As an illustration, Fig. 6 shows (non-normalized) model functions for $\alpha=1.22$ and various values of $d$ [with $\sigma=0.05$, Fig. 6(a)], and $\sigma$ [with $d=6$, Fig. 6(b)]. Of course, $\sigma=0$ gives the unperturbed scaling function $r^{0.22}$. All these functions have been calculated by evaluation of the RHS of Eq. (18) and the transformation of Eq. (19). The influence on the

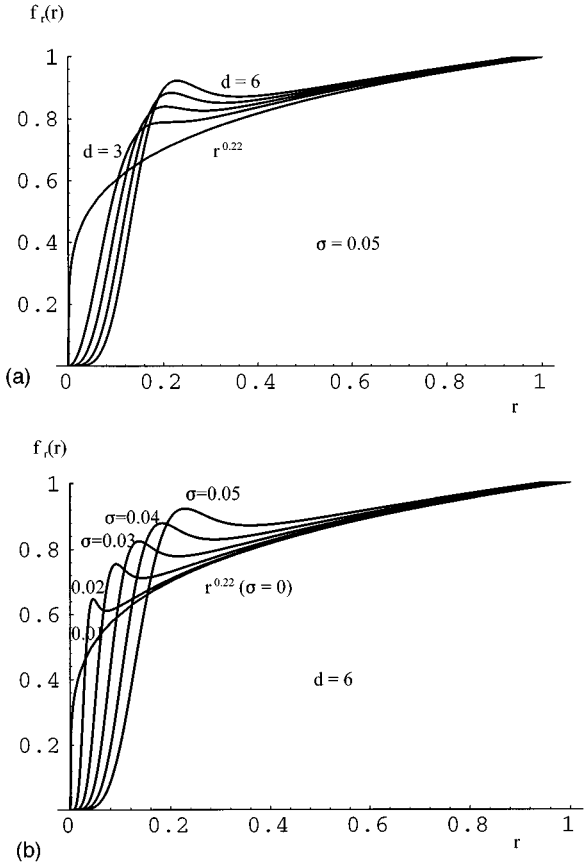


FIG. 6. Examples of numerically calculated model functions, with $\alpha=1.22$ for (a) $d=3,4,5,6$ and $\sigma=0.05$ and (b) $\sigma=0$ up to 0.05 with steps of 0.01 and $d=6$.

shape of the curves of changing $\sigma$ is big. This is also caused by the small value of $\alpha$ (compared to $d$).

In the case of an experimental time series, $\alpha$ and $\sigma$ are unknown. This case can be simulated by regarding the available Hénon time series as a series of successive measurements of some physical variable, whose generating mechanism is investigated. The unknown parameters can then be determined by the nonlinear fit procedure with the piecewise approximation function.

The pre-estimate of $\alpha$ is 1.3 from a plot as in Fig. 1. The pre-estimate of $\sigma^2$ is found from Fig. 7, which shows a plot of the MLE of $\sigma^2$ as a function of $s_c$ (on a logarithmic scale).

The behavior of these curves has been explained before. It is observed that the plateau decreases with increasing $d$, in agreement with the correction factor $d/(d-\alpha)$ of Eq. (30). The result is $\hat{\sigma}^2\approx 0.0024$, almost exactly the true value. The results of the piecewise nonlinear fit are listed in Table I. It is rewarding that the simulation parameters are recovered within a few percent of their actual values.

In this case, the data of Table I can be used for further analysis of the system. The factor of proportionality $c$ contains information on the correlation entropy $K$, which is a measure for the rate of divergence of initially nearby trajectories [5,14,23]. It quantifies the unpredictability of time evolutions. For small Euclidean distance $r$ and large $d$, we have $f_r(r)=cr^{\alpha-1}=c'e^{-Kd}d^{-\alpha/2}r^{\alpha-1}$. Applying this to the data gives $K\approx 0.30$, which agrees reasonably with the values found in [5,23].

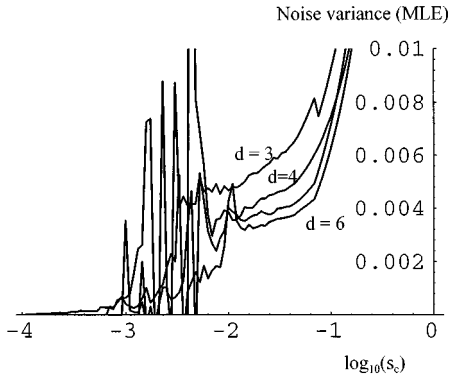


FIG. 7. Graphs of the ML estimate of the variance of the noise as a function of $s_c$, for various values of $d$, for a time series of $10^4$ points of the Hénon map. The estimates have been calculated from a random sample of $5\times 10^6$ interpoint distances (of which approximately $1.31\times 10^6$ are smaller than 1 for $d=3$ and $0.39\times 10^6$ for $d=6$).

### B. Fluidized bed data

The experimental time series that we have investigated consists of pressure measurements in a cylindrical fluidized bed (FB) column. Recent studies [24–26] have suggested that the hydrodynamics of a fluidized bed can be characterized by deterministic chaotic characteristics such as the (correlation) dimension and the entropy. In this context, the dimension can be used to characterize the hydrodynamic regime.

In the experiment, the FB column, with a diameter of 10 cm, was filled with spherical polystyrene particles with a diameter of 0.56 mm and a density of 1100 kg/m$^3$. Air was used as fluidization gas. The fluidization velocity was 20 cm/s. The sample frequency of the pressure sensor was 500 Hz.

In the dimension analysis, we have used a time series, to be denoted by ''FB,'' consisting of every tenth point of the measured time series. So the time step between the points $\delta t$ is 0.02 s. The length of the time series $N$ is again $10^4$ points (or 200 s). The first 500 points (10 s) of the time series are plotted in Fig. 8. The same analysis as in the previous example has been applied to the FB time series. Plots of the sample histograms of the distances and the MLE of the variance of the noise are shown in Fig. 9. The results of the piecewise nonlinear fit are summarized in Table I. To obtain a good fit for the FB time series, the data for different embedding dimensions had to be fitted separately; therefore, we have different values of $\alpha$ and $\sigma$. A simultaneous fit using the exact model function of Eq. (33) gives $\alpha\approx 1.777$ and $\sigma\approx 0.257$, which is in the same order. These results indicate that in this case single fits at higher embedding dimension tend to overestimate the correlation dimension and underestimate the noise variance.

## V. CONCLUDING REMARKS

It must be repeated that the above is an idealization. In practice, the scaling rule is only an approximation of a part of the real PDF of the distances. For large distance it is only known that the PDF will have a maximum, beyond which it

TABLE I. Fitted parameter values for Hénon time series with added noise ($\sigma=0.05$) and FB time series.

| Time series | $d$ | $r_0$ | $F_l$ | $F_h$ | $c$ | $\alpha$ | $\sigma$ |
|---|---|---|---|---|---|---|---|
| Hénon | 3 | 0.5 | 0.5 | 1.5 | 2991 | 1.1863 ± 0.0039 | 0.0522 ± 0.0003 |
| | 4 | | | | 1883 | | |
| | 5 | | | | 1204.2 | | |
| | 6 | | | | 799.0 | | |
| FB | 3 | 2.0 | 0.6 | 1.2 | 2174 | 1.6460 ± 0.0073 | 0.2046 ± 0.0034 |
| | 4 | | | | 1357 | 1.8946 ± 0.0089 | 0.1798 ± 0.0033 |
| | 5 | | | | 817.4 | 2.2466 ± 0.0111 | 0.1561 ± 0.0047 |
| | 6 | | | | 485.4 | 2.5595 ± 0.0153 | 0.1610 ± 0.0063 |

will decrease to zero, as in Fig. 4. For small distances there is always the problem of discretization (causing ''steps'' in the correlation integral curves; see [3]). Furthermore, in practice noise can be correlated or depending on the measurement.

By inspection of the graphs of Fig. 2, it can be observed that although $f_w$ deviates considerably from $g$, the function $f_{\tilde{s}}$ approximates $g$ very closely for $d\approx\alpha$. (Exact equality cannot be achieved because $\alpha$ is generally a noninteger.) This effect was explained in Eqs. (22) and (23). The consequence is that, if the embedding dimension is chosen as close to $\alpha$ as possible (i.e., $d=\lceil\alpha\rceil$), the effect of noise on the scaling function is minimized. Sauer and Yorke [27] proved that theoretically, in reconstructing an attractor from experimental data, $d>\alpha$ is indeed a sufficient condition for the power law to hold. Numerical limitations caused by lack of data are treated by Ding *et al.* [28].

Smith [29] has derived an approximation of the correlation integral based on a linearization of system equations. His approach is different, but he comes to an integral form of Eq. (18). This work shows that his approxiamtion is a special case of a more generally valid equation. A similar approach is followed by Schreiber [30], who derives an analytic expression for the perturbation of the correlation dimension as a function of $r$, using the $L^\infty$ norm. The parameter $\sigma$ is then determined by fitting this function to the data for various values of $d$. Another interesting algorithm is the one by Diks [23]. Here a modified definition of the correlation integral, containing a kernel function specially tailored to Gaussian noise and having the same power law property as the ''standard'' correlation integral, is used to fit $\sigma$, $\alpha$, and $K$.

From the numerical experiments, the following conclusions can be drawn.

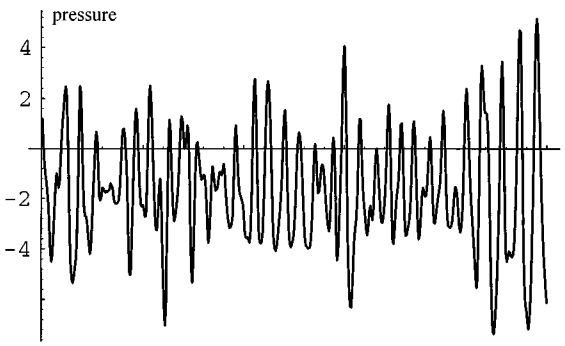


FIG. 8. Detail of the time series of pressure measurements in a cylindrical fluidized bed. The time interval is 0.02 s.

(i) The expression of Eq. (18) describes the reality in the idealized case of perfect scaling and normally distributed noise (if the standard deviation is a fraction of the scaling region boundary $r_0$).

(ii) In the Hénon example, it is observed (see Fig. 5) that the width of the scaling region increases with increasing embedding dimension, approximately with $\sqrt{d}$. This is understandable because Euclidean distances calculated from the same data are on average proportional to $\sqrt{d}$. Therefore, it could be sensible to use $r_0(d)=r_1\sqrt{d}$, with $r_1$ fixed.

(iii) In the fluidized bed example, the correlation dimension does not converge. For these data the model is not correct. This could mean that the FB system cannot be described

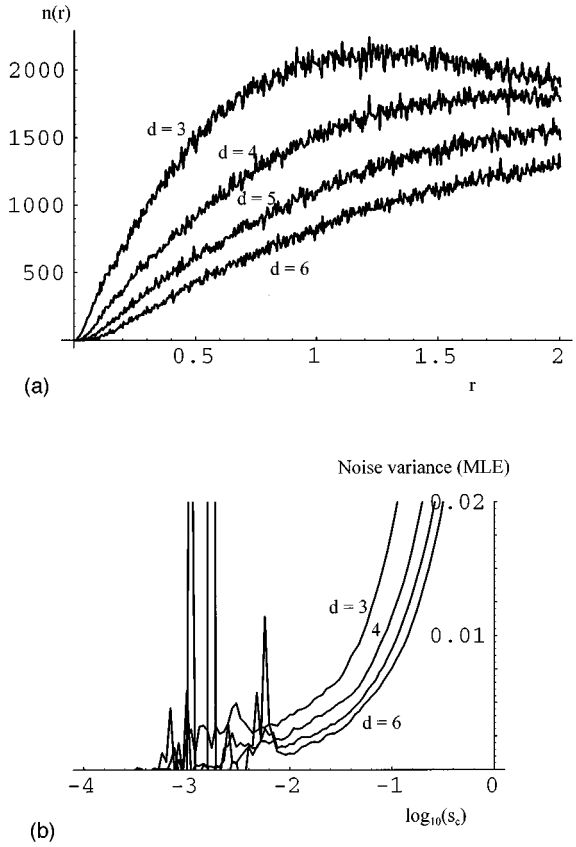


FIG. 9. (a) Graphs of sample data histograms for the FB time series. The bin width of the histograms is 0.004. (b) Graphs of the ML estimate of the variance of the noise as a function of $s_c$, for various values of $d$, for the same time series.

by a low-dimensional chaotic attractor or that the noise is of a different nature than assumed here.

(iv) The nonlinear fit procedure based on limiting behavior is a fast and useful method to recover the values of $\alpha$ and $\sigma$. However, the method can still be cumbersome in choosing the limits of the validity regions. Nonlinear regression with the complete function [Eq. (33)] is then an alternative.

The case of autocorrelated (or even dependent) noise could be investigated separately, but it is obvious from the foregoing that the expressions involved will be very complicated. More advanced statistical methods such as those used by Cheng and Tong [31] have to be applied.

## ACKNOWLEDGMENTS

The development of the data analysis method described here originates from research on the hydrodynamics in fluidized bed reactors, carried out in the Chemical Reaction Engineering Group of the Delft University of Technology and supervised by C.M. van den Bleek and J.C. Schouten.

[1] K.J. Falconer, *Fractal Geometry; Mathematical Foundations and Applications* (Wiley, Chichester, 1990).
[2] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[3] J. Theiler, J. Opt. Soc. Am. A **7**, 1055 (1990).
[4] F. Takens, in *Dynamical Systems and Bifurcations*, edited by B.L.J. Braaksma, H.W. Broer, and F. Takens, Lecture Notes in Mathematics Vol. 1125 (Springer-Verlag, Berlin, 1984), pp. 99–106.
[5] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
[6] F. Takens, in *Dynamical Systems and Turbulence*, edited by D.A. Rand and L.-S. Young, Lecture Notes in Mathematics Vol. 898 (Springer-Verlag, Berlin, 1981), p. 366.
[7] N.H. Packard, J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, Phys. Rev. Lett. **45**, 712 (1980).
[8] P. Grassberger and I. Procaccia, Physica D **13**, 34 (1984).
[9] A. Ben-Mizrachi, I. Procaccia, and P. Grassberger, Phys. Rev. A **29**, 975 (1984).
[10] J.B. Ramsey and H.-J. Yuan, Nonlinearity **3**, 155 (1990).
[11] J. Holzfuss and G. Mayer-Kress, in *Dimensions and Entropies in Chaotic Systems*, edited by G. Mayer-Kress (Springer-Verlag, Berlin, 1985), pp. 114–122.
[12] V. Isham, in *Networks and Chaos — Statistical and Probabilistic Aspects*, edited by O.E. Barndorff-Nielsen *et al.* (Chapman & Hall, London, 1993), pp. 124–200.
[13] P. Grassberger, in *Chaos, Nonlinear Science: Theory and Applications*, edited by A.V. Holden (Manchester University Press, Manchester, 1986), pp. 291–311.
[14] J.G. Caputo and P. Atten, Phys. Rev. A **35**, 1311 (1987).
[15] S. Wolfram, *Mathematica*, 2nd ed. (Addison-Wesley, Redwood City, CA, 1991).
[16] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I.A. Stegun (Dover, New York, 1968).
[17] I.S. Gradshteyn and I.M. Ryzhik, in *Table of Integrals, Series and Products*, edited by Alan Jeffrey, 4th ed. (Academic, New York, 1980).
[18] M.B. Wilk, R. Gnanadesikan, and M.J. Huyett, Biometrika **49**, 525 (1962).
[19] M.B. Priestley, *Spectral Analysis and Time Series* (Academic, London, 1981).
[20] D.W. Marquardt, J. Soc. Ind. Appl. Math. **11**, 431 (1963).
[21] One might remark that the scaling function for small distances is a polynomial [see Eq. (20)]: $n(r)=cf_r(r)=\Sigma_{k=0}^{m}a_k r^{d-1+2k}$, where $m\in\mathbb{N}$ is the order of approximation. This function is linear in the coefficients $a_k$, which are themselves functions of the three unkown parameters $c$, $\alpha$, and $\sigma$: $a_0=[c/(2\sigma)^{d-\alpha}][\Gamma(\alpha/2)/\Gamma(d/2)]$, $a_1/a_0=-[(d-\alpha/d)](1/4\sigma^2)$, and $a_2/a_1=-[(d-\alpha+2)/2(d+2)](1/4\sigma^2)$, for the first three coefficients. A linear regression directly calculates these coefficients. Then, by inversion of the formulas, we can solve for the required parameters: $\hat{\alpha}=(2a_0a_2-a_1^2)/[2a_0a_2/d-a_1^2/(d+2)]$, $\widehat{\sigma^2}=-(d-\hat{\alpha})a_0/4da_1$, and $\hat{c}=(2\hat{\sigma})^{d-\hat{\alpha}}[\Gamma(d/2)/\Gamma(\hat{\alpha}/2)]a_0$. This method does not work satisfactorily, because in the studied region the scaling behavior is only a small effect. As a consequence, the determination of $\alpha$ in this way is inaccurate.
[22] M. Hénon, Commun. Math. Phys. **50**, 69 (1976).
[23] C. Diks, Phys. Rev. E **53**, R4263 (1996).
[24] J.C. Schouten, M.L.M. vander Stappen, and C.M. van den Bleek, Chem. Eng. Sci. **51**, 1991 (1996).
[25] J.C. Schouten, F. Takens, and C.M. van den Bleek, Phys. Rev. E **50**, 1851 (1994).
[26] C.S. Daw, W.F. Lawkins, D.J. Downing, and N.E. Clapp, Jr., Phys. Rev. A **41**, 1179 (1990).
[27] T. Sauer and J.A. Yorke, Int. J. Bifurcation Chaos **3**, 737 (1993).
[28] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J.A. Yorke, Phys. Rev. Lett. **70**, 3872 (1993).
[29] R.L. Smith, J. R. Stat Soc. B **54**, 329 (1992).
[30] T. Schreiber, Phys. Rev. E **48**, R13 (1993).
[31] B. Cheng and H. Tong, J. R. Stat. Soc. B **54**, 427 (1992).