# Influence of Gaussian noise on the correlation exponent

Thomas Schreiber*
*Max Planck Institute for Physics of Complex Systems, Bayreuther Strasse 40, D-01187 Dresden, Germany*
(Received 20 March 1997)

Self-similarity of fractal point sets is broken if the points are contaminated by measurement noise. Different approaches have been taken to calculate the influence of Gaussian noise on the scaling of interpoint distances. We bring the results into a form which makes a comparison possible. Although the approaches use different norms and ways to quantify the scaling behavior, we find that the influence of noise leads to similar deviations from the pure scaling behavior. [S1063-651X(97)07307-8]

PACS number(s): 05.45.+b

## I. INTRODUCTION

The geometry of fractals has been subject to extensive mathematical studies, and there is ample empirical evidence for the relevance of fractal geometry in nature [1]. Mathematical fractals show self-similarity in the limit of small length scales. Such a limit cannot be taken in experimental observations due to the finite resolution of any measurement device. However, approximate self-symmetry often extended to small but finite length scales. It is therefore interesting to study how self-similarity is broken if the resolution is limited by measurement noise. In Refs. [2–6], this question has been addressed for the particular case of the scaling exponent of the distribution of interpoint distances, known as the Grassberger-Procaccia correlation integral, under the assumption that an infinite number of measurements is available, subject to Gaussian independent measurement noise. While Refs. [2,4–6] consider the Euclidean, or $L^2$, norm, Ref. [3] uses the maximum, or $L^\infty$, norm and makes the further assumption that the fractal measure arises as an attractor of a dynamical system and is given by a time delay reconstruction from a time series. In this paper we will comment on and compare these works. In particular, we will quote the different formulas in a form which makes it possible to compare the curves obtained. For the usual correlation integral in the Euclidean case, Refs. [2,6] make the most specific statement by giving the functional form of the contaminated scaling function of a noisy fractal. Although the derivation uses different arguments, both results coincide. The approach of Ref. [5] is only valid for length scales larger than the noise level. Thus for the theoretical understanding of the influence of noise on pair distributions, the more general result given in Refs. [2,6] seems preferable.

## II. SCALING OF PAIR DISTRIBUTIONS

One of the most popular quantities to characterize observed fractal distributions is the Grassberger-Procaccia correlation integral [7]. A continuum definition of the correlation integral $C(r)$ of a distribution $\mu(\vec{x})$ at length scale $r$ is

$$C(r)=\int d\vec{x}\int d\vec{y}\,\mu(\vec{x})\mu(\vec{y})K(\|\vec{y}-\vec{x}\|/r), \tag{1}$$

where usually the kernel function $K(x)$ is taken to be the Heaviside step function $\Theta(1-x)$. Then, $C(r)$ is simply the fraction of pairs of points with a distance (in some norm) smaller than $r$. For a fractal measure and in the limit that $r\to 0$, $C(r)$ scales as a power law, $C(r)\propto r^\alpha$. The scaling exponent $\alpha$ is also called the correlation dimension. $\alpha$ is a lower estimator of the information dimension, which is theoretically more interesting, and of the Hausdorff dimension of the supprt of $\mu$.

The probability density $n(r)$ for the interpoint distance can be obtained by taking the derivative of $C(r)$, yielding $n(r)=dC(r)/dr\propto r^{\alpha-1}$. Another way to express the same scaling behavior is by giving the local slope of a double logarithmic plot of $C(r)$:

$$D(r)=\frac{d}{d\ln r}\ln C(r)=\frac{rn(r)}{C(r)}. \tag{2}$$

$D(r)$ can be seen as a scale dependent effective *dimension*. The actual correlation dimension is obtained as the limit $\lim_{r\to 0}D(r)$.

In Ref. [4], a version of the correlation integral is used in which the hard kernel function is replaced by a Gaussian, $\exp(-x^2/4)$:

$$C^g(r)=\int d\vec{x}\int d\vec{y}\,\mu(\vec{x})\mu(\vec{y})e^{-\|\vec{y}-\vec{x}\|^2/4r^2}. \tag{3}$$

Smooth kernel correlation integrals are discussed in Ref. [8], where it is also shown that if $C$ does indeed scale as a power law, then so does $C^g$, with the same power. In the case of the Euclidean norm, the Gaussian kernel allows for the convolution of the kernel function and the Gaussian noise distribution.

## III. SCALING FUNCTIONS AND NOISE

If in an experiment the pair distribution has to be estimated from a large collection of points which are known only up to Gaussian measurement noise, we will have to account for the fact that $\mu$ itself is not available but only the convolution of $\mu$ and the noise distribution. Therefore we

*Permanent address: Physics Department, University of Wuppertal, D-42097 Wuppertal, Germany.

 

expect that at small length scales the scaling is dominated by the noise, $\widetilde{C}(r) \propto r^d$, where $d$ is the dimension of the space the set is embedded in. For larger $r$ there may be a region where the original scaling $\widetilde{C}(r) \propto r^\alpha$ is approximately valid. The modified scaling functions $\widetilde{n}(r)$, $\widetilde{C}(r)$, or $\widetilde{D}(r)$ are computed under different conditions in Refs. [2–6]. Due to the different assumptions, the results are not supposed to be identical but the differences should not be dramatic. Unfortunately, these references state their results in different forms and cannot be immediately compared. Reference [3] becomes most useful when $\widetilde{D}(r)$ is evaluated for different embedding dimensions. $\widetilde{C}(r)$ can also be given, $\widetilde{n}(r)$, however, only for integer $\alpha$. Reference [2] gives $\widetilde{C}(r)$ and Ref. [6] $\widetilde{n}(r)$ as expressions involving a transcendental function. Both expressions can indeed be related by taking a derivative. Given both $\widetilde{C}(r)$ and $\widetilde{n}(r)$, $\widetilde{D}(r)$ is also known. Reference [5] gives a formula for a corrected length $r'$ such that scaling of $\widetilde{C}(r)$ is restored. This is a very convenient way to express the result for practical applications. For the purpose of comparison, we will here use it to express $r$, $n$, and $\widetilde{D}$ as functions of $\widetilde{C}$. Reference [4] gives the Gaussian kernel equivalent to $\widetilde{C}(r)$ which is easily differentiated to yield $\widetilde{D}(r)$ and $\widetilde{n}(r)$. While the expressions for the Gaussian kernel correlation integral are by far the easiest to handle, the estimation of $\widetilde{C}^g(r)$ from a time series is somewhat more tedious.

## IV. MAXIMUM NORM

Here and in the following, the $\square$ superscript indicates the use of the maximum norm.

If the interpoint distances are calculated using the maximum norm and the point set is reconstructed from a time series, Ref. [3] derives an expression for the noise contribution to the correlation integral. If the embedding dimension is sufficiently much above that necessary for a proper embedding, the correlation integral can be split multiplicatively into one contribution containing signal and noise, and one containing noise only. The latter can be calculated analytically when the noise distribution is given. The result for Gussian noise of standard deviation $\sigma$ is

$$\widetilde{C}^{\square}(r) = \widetilde{C}^{\square}_m(r)[\sqrt{2}\,\mathrm{erf}(r/2\sigma)]^{d-m}. \tag{4}$$

Here and in the following, the $\square$ superscript indicates the use of the maximum norm.

The result is valid for the case where a reconstruction in $m$ dimensions yields a faithful embedding. Formally, this requires $m > 2\alpha$ but typically, $m > \alpha$ is sufficient. For the case of the maximum norm we have not been able to further evaluate $\widetilde{C}^{\square}_m(r)$ which contains mixed signal and noise contributions. The derivative of $\widetilde{C}^{\square}(r)$, $\widetilde{n}^{\square}(r)$ does not allow for the isolation of a pure noise component, except for the unusual case that $\alpha = m$ is integer and therefore $\widetilde{C}^{\square}_m(r) \propto r^m$. On the other hand, the decomposition becomes additive and thus more useful in the representation as $\widetilde{D}^{\square}(r)$. The result then reads

$$\widetilde{D}^{\square}(r) = \widetilde{D}^{\square}_m(r) + \frac{d-m}{\sigma\sqrt{\pi}} \frac{r e^{-r^2/4\sigma^2}}{\mathrm{erf}(r/2\sigma)}. \tag{5}$$

This formula provides a convenient means to evaluate the effect of noise on the scaling of a distribution. It has been used in particular to determine the noise level in time series data [3,10].

## V. EUCLIDEAN NORM

If the Euclidean norm is used, the contaminated scaling function can be derived assuming power law scaling of the unperturbed distances. The derivation is quite elegant if the Gaussian kernel correlation integral $C^g(r)$, Eq. (3), is used. For the usual hard kernel the algebra is more complicated and the result involves a special function.

### A. Gaussian kernel

Diks [4] observes that the Gaussian kernel correlation integral can be seen as a convolution of the probability distribution $\widetilde{\mu}(\vec{x})$ with a Gaussian of width $r$. The noisy distribution $\widetilde{\mu}(\vec{x})$ itself, however, depends on the true measure $\mu(\vec{x})$ through a convolution with the noise distribution, a Gaussian of width $\sigma$. Under the assumption that $C^g(r) \propto r^\alpha$ for the noise-free distribution, it is found in Ref. [4] that

$$\widetilde{C}^g(r) \propto \frac{r^d}{(r^2+\sigma^2)^{(d-\alpha)/2}}. \tag{6}$$

In order to compare to the other approaches let us quote this result in two alternative forms. Differentiation with respect to $r$ yields

$$\widetilde{n}^g(r) \propto \frac{r^{d-1}}{(r^2+\sigma^2)^{(d-\alpha)/2}} \frac{\alpha r^2 + d\sigma^2}{r^2+\sigma^2}. \tag{7}$$

The logarithmic derivative $\widetilde{D}^g(r) = [d/(d\ln r)]\ln \widetilde{C}^g(r)$ takes a particularly simple form:

$$\widetilde{D}^g(r) = \frac{\alpha r^2 + d\sigma^2}{r^2+\sigma^2}. \tag{8}$$

The crossover from $\widetilde{D}^g(r) \approx d$ for $r \ll \sigma$ to $\widetilde{D}^g(r) \approx \alpha$ for $r \gg \sigma$ is evident.

### B. Hard kernel

For the usual correlation integral, the convolution of the kernel and the noise distribution cannot be carried out in such an elegant way. However, Smith [2] was able to derive a formula for the noise contaminated correlation integral using local linearizations of the system equations. We will not repeat here the derivation but quote the final result. Assuming $C(r) = r^\alpha$ it reads

$$\widetilde{C}^{\circ}(r) = \frac{\Gamma((\alpha+1)/2)}{\Gamma((d+1)/2)} (2\sigma)^{\alpha-d} r^d M\left(\frac{d-\alpha}{2}, \frac{d+2}{2}; -\frac{r^2}{4\sigma^2}\right), \tag{9}$$

where $M(a,b;z)$ is the confluent hypergeometric, or Kummer's, function. Here and in the following, the $\bigcirc$ superscript indicates the use of the Euclidean norm.

Oltmans and Verheijen [6] analyze the influence of Gaussian isotropic noise on the interpoint distances, rather than the integrated distribution. The expression for the scal-

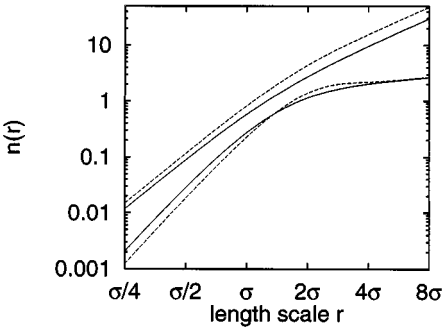


FIG. 1. Double logarithmic plot of $\tilde{n}^g(r)$ (solid) and $\tilde{n}^\circ(r)$ (dashed) for $d=4,\alpha=2.6$ (upper two curves) and for $d=5,\alpha=1.4$ (lower two curves).

ing function $\tilde{n}^\circ(r)$ of the noisy interpoint distances is obtained by combining Eqs. (18) and (19) of Ref. [6]:

$$\tilde{n}^\circ(r)=\frac{\Gamma(\alpha/2)}{\Gamma(d/2)}\alpha(2\sigma)^{\alpha-d}r^{d-1}M\left(\frac{d-\alpha}{2},\frac{d}{2};-\frac{r^2}{4\sigma^2}\right). \tag{10}$$

Using the recursion relations for Kummer's function [Eqs. (13.4.1)–(13.4.7) in Ref. [9]], one finds that $\tilde{n}^\circ(r)$ in Eq. (10) is indeed the derivative of $\tilde{C}^\circ(r)$ in Eq. (9). This is very satisfactory since Eq. (9) and Eq. (10) have been derived by very different means.

Combining Eqs. (9) and (10) and using Kummer's transformation [Eq. (13.1.27) in Ref. [9] ], we get quite a compact expression for $\tilde{D}^\circ(r)$:

$$\tilde{D}^\circ(r)=d\frac{M(\alpha/2,d/2;r^2/4\sigma^2)}{M((\alpha+2)/2,(d+2)/2;r^2/4\sigma^2)}. \tag{11}$$

Kugiumtzis [5] gives a correction formula for the length scale $r$ to be used in the presence of Gaussian noise. $\tilde{C}(r)$ is then supposed to show the proper scaling with the corrected $r'$. The formula is derived by accounting for those points which escape from a neighborhood due to the noise. Certain approximations (namely, exchanging the expectation operator with point counting) are made which assume that $r$ is at least of the order of $\sigma$. Therefore the implication of the result for the present work is less general than Refs. [2,6]. Let us assume that power law scaling $C(r)\propto r^\alpha$ is restored by the correction, such that for the measured correlation sum, $\tilde{C}(r)\propto r'^\alpha$ holds. Then we can rewrite Eq. (5) in Ref. [5] and obtain

$$r=\frac{d\tilde{C}^{4/\alpha}+(4d-1)(\sigma^2\tilde{C}^{2/\alpha}+\sigma^4)}{d(\tilde{C}^{2/\alpha}+2\sigma^2)^{3/2}}. \tag{12}$$

Note that Ref. [5] uses the $d$-scaled Euclidean norm, which will be accounted for when the results are compared. From Eq. (12) we can compute $\tilde{n}^>$ and $\tilde{D}^>$ as functions of $\tilde{C}$. The expressions are lengthy and not very instructive. They are therefore not given here, except for a particular choice of $d$ and $\alpha$, see Eq. (14) below.

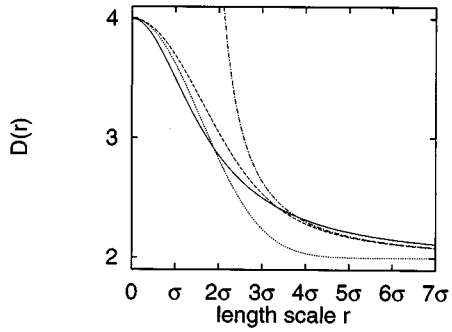


FIG. 2. Plot in linear scale of $\tilde{D}^\square(r)$ (dotted), $\tilde{D}^g(r)$ (solid), and $\tilde{D}^\circ(r)$ (dashed). Further, $\tilde{D}^>(\tilde{C})$ is plotted versus $r^>(\tilde{C})$ as a parametric function of $\tilde{C}$ (dashed-dotted). All curves are computed for $d=4,\alpha=2$. The result derived from Ref. [5], $\tilde{D}^>$, coincides with $\tilde{D}^\circ(r)$ for large enough $r$.

## VI. COMPARISON

When comparing the different results, we should observe that the length scales accessed by the different approaches at a given value of $r$ are not quite the same and cannot be made exactly equal. For a better comparison, one may want to use kernel functions of the same root mean squared variance. The Gaussian kernel used above has width $\sqrt{2}$ while the hard kernel has width $1/\sqrt{3}$. Further, the hard kernel in the maximum norm covers a hypercube whereas in the Euclidean norm it only covers a sphere. The corresponding rescaling would depend on the attractor dimension.

Let us first show exemplary curves $\tilde{n}^g(r)$ and $\tilde{n}^\circ(r)$ for the Euclidean norm, which are available analytically for all $d$ and $\alpha$. Figure 1 contains a double logarithmic plot of $\tilde{n}(r)$ for $d=4,\alpha=2.6$ and for $d=5,\alpha=1.4$. In order to correct for the different kernel widths, $\tilde{n}^g(r)$ has been shifted left by $\sqrt{2}$ and $\tilde{n}^\circ(r)$ right by $\sqrt{3}$. As expected, the asymptotics are the same but the crossover region around the noise level is different.

For certain particular (nonfractal) values of $d$ and $\alpha$, we can calculate analytical curves for all three cases for $\tilde{C}(r)$ and $\tilde{D}(r)$. For $d=4,\alpha=2$, we obtain the following formulas for $\tilde{D}(r)$:

$$\tilde{D}^\square(r)=2+\frac{2r}{\sigma\sqrt{\pi}}\frac{e^{-r^2/4\sigma^2}}{\mathrm{erf}(r/2\sigma)},$$

$$\tilde{D}^g(r)=\frac{2r^2+4\sigma^2}{r^2+\sigma^2},\quad \tilde{D}^\circ(r)=\frac{2r^2(1-e^{-r^2/4\sigma^2})}{r^2-4\sigma^2(1-e^{-r^2/4\sigma^2})}, \tag{13}$$

Further, the result of Ref. [5] can be used to obtain a parametric function $(r^>(\tilde{C}),\tilde{D}^>(\tilde{C}))$:

$$r^>(\tilde{C})=\frac{4\tilde{C}^2+15\sigma^2\tilde{C}+15\sigma^4}{4(\tilde{C}+2\sigma^2)^{3/2}}, \tag{14}$$

$$\widetilde{D}^{>}(\widetilde{C}) = \frac{4\widetilde{C}^3 + 23\sigma^2\widetilde{C}^2 + 45\sigma^4\widetilde{C} + 30\sigma^6}{4\widetilde{C}^3 + 17\sigma^2\widetilde{C}^2 + 15\sigma^4\widetilde{C}}.$$

The curves are shown in Fig. 2. Again, the different kernel widths are accounted for. On a two dimensional point set, the ratio of the volumes covered by a box of radius $r$ and a circle of the same radius is $4/\pi$. Therefore $\widetilde{D}^{\square}(r)$ is shifted left by $4\sqrt{2}/\pi$. Since $\widetilde{D}^{>}(\widetilde{C})$ has been calculated with the $d$-scaled Euclidean norm, the curve has been shifted right by an additional factor of $\sqrt{d}=2$ for comparison. Note that $\widetilde{D}^{>}(\widetilde{C})$ is only supposed to be valid for $r>\sigma$, and, in fact, diverges for $\widetilde{C}\to 0$ (at finite $r$). For large enough $r$, however, the agreement with the general result $\widetilde{D}^{\circ}(r)$ is very good, as expected.

## VII. CONCLUSION

We have brought together the results of different approaches which compute the influence of Gaussian noise on the correlation integral. Although different norms and different definitions of the correlation integral were used, the resulting curves are very similar. This is quite obvious for the asymptotic behavior for length scales much larger and much smaller than the noise level but it is also found for the crossover region. For the Euclidean case, the results of Refs. [2,6] coincide and the formula which can be derived from Ref. [5] indeed converges to the more general result of Refs. [2,6] for large enough $r$.

As we have mentioned before, the length scales cannot be compared directly for the different approaches. In particular, the reasoning in this paper does not take into account that scaling is both limited by noise from below and by the finite extent of the set from above. Thus neither the formulas nor the figures allow for a definitive statement as to how strongly the different ways of estimating the correlation integral are affected by noise.

All approaches allow for the estimation of the noise level, Refs. [2,4,6] allow for the simultaneous determination of the noise level and the scaling exponent by a nonlinear function fit. While in the case of the usual hard kernel the function to be fitted is rather complicated, the computation of the Gaussian kernel correlation integral from a time series is more tedious and computer time consuming. Reference [5] requires an estimate of the noise variance but then yields a convenient correction scheme for $C(r)$. The result [3] for the maximum norm allows for the determination of the noise level but not for a direct correction of a dimension estimate. An attractive alternative is to perform a nonlinear noise reduction step [11] in order to recover the correct scaling behavior.

## ACKNOWLEDGMENTS

We thank Cees Diks, Dimitris Kugiumtzis, and Peter Grassberger for useful discussions. This work was supported by the SFB 237 of the Deutsche Forschungsgemeinschaft.

[1] B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1985).
[2] R. L. Smith, J. R. Statist. Soc. B **54**, 329 (1992).
[3] T. Schreiber, Phys. Rev. E **48**, R13 (1993).
[4] C. Diks, Phys. Rev. E **53**, 4263 (1996).
[5] D. Kugiumtzis, Int. J. Bifurcation Chaos (to be published).
[6] H. Oltmans and P. J. T. Verheijen, Phys. Rev. E **56,** 1160 (1997).
[7] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[8] J. M. Ghez and S. Vaienti, Nonlinearity **5**, 777 (1992).
[9] *Handbook of Mathematical Functions*, edited by M. Abramowitz and I. A. Stegun (Dover, New York, 1965).
[10] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, in press).
[11] E. J. Kostelich and T. Schreiber, Phys. Rev. E **48**, 1752 (1993).