Noise robust estimates of correlation dimension and K_2 entropy

Guido Nolte,[1,*] Andreas Ziehe,[2,†] and Klaus-Robert Müller[2,3,‡]
[1]*Department of Computer Science, University of New Mexico, Albuquerque, New Mexico 87131-1386*
[2]*GMD FIRST.IDA, Kekuléstrasse 7, 12489 Berlin, Germany*
[3]*University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany*
(Received 1 February 2001; published 15 June 2001)

Using Gaussian kernels to define the correlation sum we derive simple formulas that correct the noise bias in estimates of the correlation dimension and K_2 entropy of chaotic time series. The corrections are only based on the difference of correlation dimensions for adjacent embedding dimensions and hence preserve the full functional dependencies on both the scale parameter and embedding dimension. It is shown theoretically that the estimates, which are derived for additive white Gaussian noise, are also robust for moderately colored noise. Simulations underline the usefulness of the proposed correction schemes. It is demonstrated that the method gives satisfactory results also for non-Gaussian and dynamical noise.

DOI: 10.1103/PhysRevE.64.016112 PACS number(s): 02.50.−r, 05.45.Tp, 05.45.Ac

I. INTRODUCTION

Analysis of nonlinear dynamics plays an important role in science. Especially low-dimensional chaos has been found in various natural and technical systems, e.g., epileptic spike trains [1], EEG signals during sleep [2], or the cardiovascular system [3].

One of the important invariant measures to characterize a time series generated by nonlinear dynamics is the correlation dimension D [4–6]. Following in the lines of Grassberger and Procaccia [4] the fractal dimension of the attractor can be estimated from a time series (x_i) by using the power law behavior of the correlation sum

$$C(\epsilon,m) := \sum_{i<j} \Theta(\epsilon - |\mathbf{x}_i - \mathbf{x}_j|) \sim \epsilon^D, \tag{1}$$

where the vectors

$$\mathbf{x}_i = (x_i, x_{i+\tau}, \ldots, x_{i+\tau(m-1)})^T \tag{2}$$

are constructed from mapping the time series (x_i) into an m-dimensional embedding space (for choosing the value of τ see, e.g., [7]) and the Heaviside function Θ is used to count the number of points inside[1] a hypersphere of radius ϵ.

For Eq. (1) to hold, several requirements have to be fulfilled: (a) ϵ must be sufficiently small to avoid finite size effects, (b) ϵ must be sufficiently large to ensure sufficient statistics and to avoid discretization errors, (c) the embedding dimension m must be chosen large enough to unfold the attractor, and (d) the influence of noise—being itself infinite dimensional—must be negligible. Assuming that all these conditions hold, the correlation dimension can be estimated by

$$D \approx d(\epsilon,m) := \frac{d \ln C}{d \ln \epsilon}, \tag{3}$$

which should be independent of ϵ (''scaling'') and independent of m (''saturation'') in some ϵ range and for sufficiently large m.

In practice, the estimate of the correlation dimension will be biased if the number of available data points is insufficient [8,9]. Even more severe, the signals obtained from real-world systems are unavoidably contaminated with noise. This makes a reasonably accurate estimate of the correlation dimension extremely difficult if not impossible to achieve. Depending on the specific dynamics, already a noise level of 1–2 % can ruin scaling and saturation, which are necessary conditions to reliably assign a dimension to a time series [7].

So far, several methods have been developed to reduce the error in estimates of dimensions caused by additive noise (see [10,11] for overviews). One general idea is to account for the theoretically expected deviation from the simple scaling behavior of Eq. (1). This was done for the Grassberger-Procaccia type of correlation sum [12–16] and for the correlation sum with Gaussian kernels [17,18], which we will discuss in detail in the next section. Typically, the corrected estimate is found by a functional fit. Similar methods were applied to estimate the noise level [19,20]. A different approach is taken in [12] proposing a non-linear scale transformation to compensate for the noise bias. This method, however, is an approximation that becomes inaccurate for scales smaller than the noise level.

For practical applications, functional fitting of refined scaling laws to empirical correlation sums has two severe drawbacks.

(1) It is unclear in what range of ϵ this fit should be performed since the given functional form is only valid for ''sufficiently'' small ϵ and the estimated correlation dimension is statistically relevant only for ''sufficiently'' large ϵ. Since, especially, the former cannot be assessed without any knowledge of the underlying system, the estimation of D is likely to be based on wrong assumptions.

(2) In order to conclude that a time series originates from a (low-dimensional) deterministic dynamical system of a cer-

*Email address: nolte@cs.unm.edu
†Email address: ziehe@first.gmd.de
‡Email address: klaus@first.gmd.de

[1]Distances are measured by the Euclidian norm $|\cdot|$, which is appropriate for the purpose of this paper.

tain dimension, one has to verify that $d(\epsilon,m)$ does not depend on ϵ within a suitable range. Though the explicit evaluation of ''scaling regimes'' is to some extent subjective, we regard figures of an estimated dimension as a function of ϵ, not just as an intermediate step but as an important final result. However, when estimating D by a functional fit, the dependence on ϵ is lost by definition, and hence it is no longer possible to check the scaling behavior.

In order to avoid those functional fits in finite scaling ranges, we will use Gaussian kernels in the definition of the correlation sum as proposed in [17]. This will allow us to derive an explicit formula for the correlation dimension as a function of m and ϵ. We will show that for every scale, the necessary information needed to correct for the noise bias is completely contained in the difference of two uncorrected dimension estimates for adjacent embedding dimensions.

Though, to our opinion, the estimate of correlation dimension from fitting in scale space has significant drawbacks, the actual result may still be accurate. Since fitting, as proposed in [17], merely exploits ϵ dependence and we will here merely exploit m dependence, the results are based on mutually independent information. In practice, comparing both results may give additional insight into the dynamical system under study eventually leading to a stronger confirmation of the estimated dimension.

Section II is devoted to theoretical aspects of our approach. In Sec. II A we present the central idea of this paper by showing how one can remove the bias caused by white Gaussian noise in an extremely simple manner. Generalizations with respect to other invariants and nonwhite noise will be done in Sec. II B, II C, and II D. General remarks on the use of Gaussian kernels are given in Sec. II E. In Sec. III we demonstrate the usefulness of our method for various simulation examples and we finally give a conclusion in Sec. IV.

II. THEORY

A. Corrected dimensions for white Gaussian noise

In order to remove the bias caused by white Gaussian noise in the estimate of the correlation dimension we work with Gaussian kernels [17] and define

$$C^g(\epsilon,m)=\sum_{i<j-t_{min}} \exp\left(-\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4\epsilon^2}\right). \tag{4}$$

To avoid spurious effects arising from autocorrelation we have excluded pairs that are too close in time by introducing a minimal delay t_{min}. This was proposed in [21] where the recommendation is to choose t_{min} to be larger than the autocorrelation time.

In contrast to the formulation with the Heaviside step function (hard kernel) in Eq. (1), using a Gaussian function (soft kernel) has the effect that contributions from pairs with $|\mathbf{x}_i-\mathbf{x}_j|>\epsilon$ do not vanish but are exponentially suppressed. The power law scaling behavior, however, coincides in both cases as can be checked explicitly by applying the transformation law of Sec. II E 2.

In the noisy case the measured values are given by

$$y_i=x_i+\eta_i, \tag{5}$$

where we assume that η_i is white Gaussian noise with standard deviation σ.

Let us now calculate the expectation of the correlation sum in the presence of noise. Denoting an index pair (i,j) by α and $\mathbf{x}_\alpha=\mathbf{x}_i-\mathbf{x}_j$ (analogously for y and η), Eq. (4) reads for the noisy case

$$C^g(\epsilon,m)=\sum_\alpha \exp\left(-\frac{|\mathbf{x}_\alpha+\boldsymbol{\eta}_\alpha|^2}{4\epsilon^2}\right), \tag{6}$$

where $\boldsymbol{\eta}_\alpha$ is now a difference between two vectors of independent Gaussian random numbers with standard deviation σ, and hence corresponds to independent Gaussian noise with standard deviation $\sqrt{2}\sigma$, with probability density

$$p(\eta)=\prod_{k=1}^{m}\frac{1}{2\sigma\sqrt{\pi}}\exp\left(-\frac{\eta_k^2}{4\sigma^2}\right). \tag{7}$$

Accordingly, the expectation of the correlation sum with respect to the noise η reads

$$\begin{aligned}\langle C^g(\epsilon,m)\rangle &= \int D\eta\, C^g(\epsilon,m)p(\eta)\\ &=\sum_\alpha\prod_{k=1}^{m}\frac{1}{2\sigma\sqrt{\pi}}\int d\eta_k\\ &\quad\times\exp\left[-\frac{(x_{\alpha k}+\eta_k)^2}{4\epsilon^2}-\frac{\eta_k^2}{4\sigma^2}\right],\end{aligned} \tag{8}$$

where $x_{\alpha k}$ denotes the kth component of $\mathbf{x}_\alpha$ and $D\eta=\Pi_k d\eta_k$. Note, that we have omitted the irrelevant index α on η.

In order to perform the integration with respect to η the exponent of Eq. (8) is rewritten as

$$\begin{aligned}-\frac{(x_{\alpha k}+\eta_k)^2}{4\epsilon^2}-\frac{\eta_k^2}{4\sigma^2}&=-\frac{\sigma^2+\epsilon^2}{4\sigma^2\epsilon^2}\left(\eta_k+\frac{\sigma^2}{\sigma^2+\epsilon^2}x_{\alpha k}\right)^2\\ &\quad-\frac{x_{\alpha k}^2}{4(\sigma^2+\epsilon^2)}.\end{aligned} \tag{9}$$

With

$$\frac{1}{\sigma}\int d\eta_k \exp\left[-\frac{\sigma^2+\epsilon^2}{4\sigma^2\epsilon^2}\left(\eta_k+\frac{\sigma^2}{\sigma^2+\epsilon^2}x_{\alpha k}\right)^2\right]\sim\frac{\epsilon}{\sqrt{\sigma^2+\epsilon^2}} \tag{10}$$

we find

$$\langle C^g(\epsilon,m)\rangle\sim\left(\frac{\epsilon}{\sqrt{\sigma^2+\epsilon^2}}\right)^m\sum_\alpha \exp\left[-\frac{|\mathbf{x}_\alpha|^2}{4(\sigma^2+\epsilon^2)}\right] \tag{11}$$

$$\sim\left(\frac{\epsilon}{\sqrt{\sigma^2+\epsilon^2}}\right)^m (\sigma^2+\epsilon^2)^{D/2}. \tag{12}$$

This result was already found in [17] within a slightly different approach. It is proposed there to estimate D from a functional fit of Eq. (12) to the measured correlation sum.

Inserting Eq. (12) into Eq. (3) and using C^g as the correlation sum leads to the estimate (see also Eq. (8) in [19])

$$d(\epsilon,m)\approx D+(m-D)\frac{\sigma^2}{\sigma^2+\epsilon^2}. \tag{13}$$

The above-mentioned drawbacks of fitting the ϵ dependence can now be overcome by adding a *subtle but important* point: According to Eq. (13) one has

$$\frac{\sigma^2}{\sigma^2+\epsilon^2}\approx d(\epsilon,m+1)-d(\epsilon,m)=:\Delta(\epsilon,m) \tag{14}$$

and insertion into Eq. (13) and solving for D leads to the simple relation

$$D\approx d'(\epsilon,m):=\frac{d(\epsilon,m)-m\Delta(\epsilon,m)}{1-\Delta(\epsilon,m)}. \tag{15}$$

Therefore, it is no longer needed to determine the noise corrected dimension d' from the functional behavior over a finite range of ϵ. Instead, it is estimated independently for each value of ϵ by merely using the results of the ‘‘standard’’ correlation dimension estimates for two adjacent embedding dimensions.

We note that the corrected dimension estimate of Eq. (15) depends explicitly on m. For large m the correction becomes large and potentially inaccurate. However, from writing

$$D\approx d'(\epsilon,m)=d(\epsilon,m)-\frac{[m-d(\epsilon,m)]\Delta(\epsilon,m)}{1-\Delta(\epsilon,m)}, \tag{16}$$

we see that the correction is proportional to $m-d(\epsilon,m)$. Hence we can expect that the performance of the proposed scheme is essentially independent of $d(\epsilon,m)$ itself and can be also applied to systems with large correlation dimension as long as the embedding dimension does not exceed the correlation dimension by a large amount.

We want to emphasize that one should clearly distinguish between the ‘‘bare scale’’ ϵ and the ‘‘effective scale’’

$$\epsilon^{eff}=(\sigma^2+\epsilon^2)^{1/2} \tag{17}$$

of the exponent in Eq. (11). While ϵ can be set arbitrarily small, ϵ^{eff} is always larger than σ. Even with perfect bias correction, the correlation sum is blind to scales (of the noise-free signal) below the noise level. Note, however, that noise already severely distorts the correlation sum for bare scales ϵ far above noise level.

B. Robust estimates of the K_2 entropy

In order to correct an estimate of the K_2 entropy, let us first recall its definition (e.g., [6,22]). In the noise-free case and assuming proper scaling, one may write the m dependence of the correlation sum for large m as

$$C^g(\epsilon,m)=c\exp(-mK_2)\epsilon^D, \tag{18}$$

where c is a constant. K_2 can hence be estimated as the limit $m\to\infty$ of

$$K_2(\epsilon,m)=\ln[C^g(\epsilon,m)]-\ln[C^g(\epsilon,m+1)]. \tag{19}$$

The K_2 entropy measures the exponential increase of the uncertainty about future values given the past up to finite accuracy. More precisely, it is a lower bound on the sum of positive Lyapunov exponents and is hence a measure of ‘‘how chaotic’’ a system is (see, e.g., [24–26]). Linear systems, for example, must have $K_2=0$ while the correlation dimension can be arbitrarily large.

In the noisy case Eq. (19) has to be modified. Including m dependence of the noise-free correlation sum, the noisy correlation sum Eq. (12) can be written as

$$\langle C^g(\epsilon,m)\rangle\approx c\exp(-mK_2)\left(\frac{\epsilon}{\sqrt{\sigma^2+\epsilon^2}}\right)^m(\sigma^2+\epsilon^2)^{D/2}. \tag{20}$$

It follows that

$$\ln(\langle C^g(\epsilon,m)\rangle)-\ln[\langle C^g(\epsilon,m+1)\rangle]\approx K_2+\frac{1}{2}\ln\left(\frac{\sigma^2+\epsilon^2}{\epsilon^2}\right). \tag{21}$$

Noting that with Eq. (14)

$$1-\Delta(\epsilon,m)\approx\frac{\epsilon^2}{\sigma^2+\epsilon^2}, \tag{22}$$

we can calculate a bias-free estimator of the K_2 entropy from the limit $m\to\infty$ of

$$K_2'(\epsilon,m):=\ln[C^g(\epsilon,m)]-\ln[C^g(\epsilon,m+1)]+\frac{1}{2}\ln[1-\Delta(\epsilon,m)]. \tag{23}$$

It is well known that the estimation of the K_2 entropy requires in general a much larger embedding dimension to show a proper saturation than would be needed for the respective correlation dimension. However, in contrast to the correlation dimension, the correction here does not *explicitly* increase with m. Though this does not imply that implicit dependencies are present, we may expect that the entropy estimates are more robust than the dimension estimates for $m\gg D$.

C. Estimating the noise level

Similarly to the correction of the correlation dimension and K_2 entropy, the use of Gaussian kernels allows to obtain estimates of the noise level itself. Solving Eq. (14) for the noise level σ leads to

$$\sigma \approx \sigma(\epsilon,m) = \left(\frac{\epsilon^2 \Delta}{1-\Delta} \right)^{1/2}. \tag{24}$$

This estimate is now a function of ϵ and m allowing to check for scaling and saturation. Again, the estimate is based on dependencies of the correlation dimension on m alone. The dependence on ϵ can now serve as an independent consistency check, and comparisons with functional fits as done in [17–19] may mutually confirm the results obtained.

D. Colored noise

In the case of colored noise one finds approximately (see Appendix)

$$d(\epsilon,m) \approx D + (m-D)\Delta' \tag{25}$$

with

$$\Delta' = \frac{1}{m} \sum_{k=1}^{m} \frac{\sigma_k^2}{\sigma_k^2 + \epsilon^2}, \tag{26}$$

where σ_k^2 is the kth eigenvalue of the ''noise-autocorrelation matrix''

$$R_{ij} := \langle \eta_i \eta_j \rangle \tag{27}$$

and where $i,j = 1, \ldots, m$ denote the time points. For white noise $R_{ij} = \sigma^2 \delta_{ij}$, the eigenvalues are all equal to σ^2 and one arrives back at Eq. (13).

As explained in detail in the Appendix, colored noise leads to ellipsoidal Gaussian kernels. The approximation in Eqs. (25) and (26) consists of assuming that the correlation sum only depends on the volume of the ellipsoid and not on its shape, which also depends on the scale ϵ. This dependence becomes smaller for larger ϵ and hence Eqs. (25) and (26) are not only absolutely (because of the smaller bias) but also structurally more accurate for larger ϵ.

The correction formula Eq. (15) can be applied as long as $\Delta \approx \Delta'$, which is valid if Δ' is sufficiently independent of m. This is obvious for the ''useless'' case in the limit of $\epsilon \to 0$, since then $\Delta' = 1$ and the dependence on D drops out.

Less trivial, this is also the case for $\epsilon \gg \sigma_k$, since then

$$\Delta' \approx \frac{\sum_k \sigma_k^2}{m\epsilon^2} = \frac{\mathrm{tr}(R)}{m\epsilon^2} \tag{28}$$

and $\mathrm{tr}(R)$ is proportional to m if the diagonal elements are all equal as for the case of stationary noise. In general, the m dependence is a complicated function of the noise spectrum. Roughly speaking, Δ' corresponds to averages in the frequency domain: the larger the m, the better the resolution. If the resolution is sufficiently large to consider the power spectrum as locally constant, the m dependence will disappear.

In conclusion, in the case of colored noise the corrected dimension estimate of Eq. (15) becomes more accurate for larger ϵ not only because the noise bias is smaller but also because of two structural reasons: (a) the approximation in Eqs. (25) and (26) is not only absolutely but also relatively more accurate and (b) the dependence of Δ' on m decreases.

In place of $\Delta(\epsilon,m)$ from Eq. (14) one can also use, e.g.,

$$\tilde{\Delta}(\epsilon,m) := [d(\epsilon,m+1) - d(\epsilon,m-1)]/2, \tag{29}$$

which is statistically more robust, but a drawback is that a sufficient embedding is already required for $m-1$ instead of m. At first sight it seems that for colored noise this definition would be preferable to Eq. (14) because of its apparent symmetry.[2] However, after defining dimensions at half-integer embedding spaces $\hat{m} = m + 1/2$ by the mean

$$d(\epsilon,\hat{m}) \equiv [d(\epsilon,m+1) + d(\epsilon,m)]/2, \tag{30}$$

and correcting according to

$$d'(\epsilon,\hat{m}) = \frac{d(\epsilon,\hat{m}) - \hat{m}\Delta(\epsilon,m)}{1-\Delta(\epsilon,m)}, \tag{31}$$

it is readily seen that this definition is identical to Eq. (15), which can hence be regarded as a symmetric correction around half-integer embedding dimensions.

We finally note that for strictly Gaussian noise, the relation

$$\frac{d\langle \ln C^g(\epsilon,m) \rangle}{d \ln \epsilon} = m\Delta'(\epsilon,m) \tag{32}$$

holds exactly. This may be used to reveal spurious nonzero dimension estimates caused by correlations of the noise.

E. On the use of Gaussian kernels

1. Calculating the derivative

One might guess that a significant drawback of the use of Gaussian kernels is the apparent smearing of scales not present when using Heaviside kernels. However, the correlation sum calculated from step functions is not continuous: to differentiate it, one must use a finite difference over a considerable range of ϵ [7,27], which in fact also smears the scales. Since the correlation sum defined by Gaussian kernels is differentiable, the latter form of smearing can be avoided and we can directly calculate $d(\epsilon,m)$ as the derivative of the correlation sum

[2] Similar to the central difference for numerically approximating a derivative.

$$d(\epsilon,m)=\frac{d\ln C^g(\epsilon,m)}{d\ln(\epsilon)}=\frac{\sum_\alpha |\mathbf{y}_\alpha|^2\exp(-|\mathbf{y}_\alpha|^2/4\epsilon^2)}{2\epsilon^2\sum_\alpha \exp(-|\mathbf{y}_\alpha|^2/4\epsilon^2)}, \tag{33}$$

where again the double index α denotes all pairs $\{(i,j)|i<j\}$ that are included in the sum and $y_\alpha=y_i-y_j$. In all of our numerical simulations we use this formula to directly calculate the (uncorrected) correlation dimension.

2. Transformation of Heaviside kernels to Gaussian kernels

A correlation sum based on Gaussian kernels may be expressed by a correlation sum based on step functions

$$C^g(\epsilon,m)=\int d\epsilon' f(\epsilon,\epsilon')C(\epsilon',m) \tag{34}$$

with $C(\epsilon',m)$ from Eq. (4). In order to find the correct weighting function $f(\epsilon,\epsilon')$, it is sufficient to express the Gaussian kernel by the Heaviside kernel. From

$$\exp\left(-\frac{|y|^2}{4\epsilon^2}\right)=\int_0^\infty d\epsilon' f(\epsilon,\epsilon')\Theta(\epsilon'-|y|), \tag{35}$$

it follows by partial integration that

$$f(\epsilon,\epsilon')=\frac{\epsilon'}{2\epsilon^2}\exp\left(-\frac{\epsilon'^2}{4\epsilon^2}\right) \tag{36}$$

in agreement with [18]. The weight function[3] $f(\epsilon,\epsilon')$ can be used to transform any quantity calculated from hard kernels into the respective ones calculated from Gaussian kernels. Especially, it follows from

$$\int_0^\infty d\epsilon' f(\epsilon,\epsilon')\epsilon'^D\sim\epsilon^D \tag{38}$$

that the power law scaling—if it exists—will be the same for both correlation sums.

The relation between hard and Gaussian kernels could, in principle, be used to speed up the computation since the correlation sums according to Grassberger and Procaccia are much simpler to calculate. However, there is a tradeoff: for an accurate calculation a fine ϵ resolution is required that partly spoils the beneficial effect if the correlation sums are calculated directly by implementing Eq. (1).

An extremely efficient method to avoid this problem was presented in [18] where the authors proposed to first calculate the histogram of $|\mathbf{y}_i-\mathbf{y}_j|$ from which the correlation sum with respect to any kernel can readily be calculated in negligible computer time. We would like to suggest a slight modification of this method by calculating the histogram of the *logarithm* of the squared distances in order to ensure sufficient resolution also for small scales. Explicitly this means that one rewrites the relevant sums as

$$\begin{aligned}\sum_\alpha f(|\mathbf{y}_\alpha|^2)\exp(-|\mathbf{y}_\alpha|^2/4\epsilon^2)&=\int dz\, f(\exp[z])\exp[-\exp(z)/4\epsilon^2]\\&\quad\times\left\{\sum_\alpha \delta(z-\ln|\mathbf{y}_\alpha|^2)\right\}\end{aligned} \tag{39}$$

with $f(|\mathbf{y}_\alpha|^2)=1$ for the denominator and $f(|\mathbf{y}_\alpha|^2)=|\mathbf{y}_\alpha|^2$ for the numerator of Eq. (33). The term in curly brackets can now be approximated by the respective histogram and the integral is finally approximated by the respective sum. Taking, e.g., 100 values for a unit step of z results in essentially exact correlation sums.

III. SIMULATION RESULTS

A. Dimension estimates in the presence of white Gaussian noise

Numerical results will be given mainly for the Hénon map, which is defined by

$$x_{i+1}=1-ax_i^2+bx_{i-1}, \tag{40}$$

with $a=1.4$ and $b=0.3$ [4]. The time lag τ for embedding according to Eq. (2) is set to 1. For convenience, all time series considered in this paper were normalized according to

$$x_i\rightarrow\frac{x_i}{\sigma_x}, \tag{41}$$

where σ_x denotes the standard deviation of the time series (x_i). For additive noise, the normalization was done with respect to the noise-free data and for dynamical noise, with respect to the noisy data.

In order to evaluate our method we added white Gaussian noise with standard deviation $\sigma=0.1$ to the time series corresponding to a noise level of 10%.

From the noisy data we compute uncorrected dimension estimates with Eq. (33) for embedding dimensions $m=2,\ldots,8$ and subsequently correct these estimates according to Eq. (15). Apart from Fig. 3 (where we compare N=500 and N=20 000), in all simulations $N=5000$ time points are used for the estimation of the invariants.

In Fig. 1 we plot the results of uncorrected and corrected estimates of the correlation dimension. Indeed, a proper scaling and saturation behavior is completely ruined by the

[3]It should be noted that for fixed ϵ the maximum of $f(\epsilon,\epsilon')$ occurs at $\epsilon'=\sqrt{2}\epsilon$ implying a mismatch of scales. Replacing the definition of the Gaussian kernel according to

$$\exp\left(-\frac{y^2}{4\epsilon^2}\right)\rightarrow\exp\left(-\frac{y^2}{2\epsilon^2}\right) \tag{37}$$

ends up with a proper match of scales in the sense that $C^g(\epsilon,m)$ gets the largest contribution from $C(\epsilon',m)$ at the same scale. Of course, the specific choice is merely a convention, and in fact, the present definition is slightly more convenient for the analytical calculations.

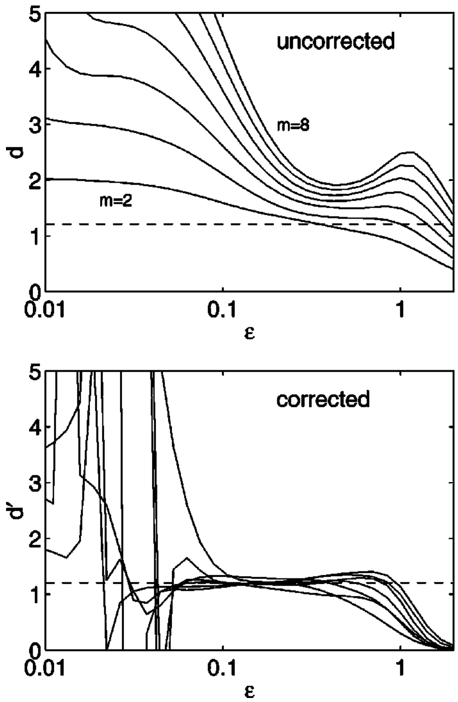

FIG. 1. Uncorrected and corrected estimates of the correlation dimension for embedding dimensions $m=2,\ldots,8$ for the Hénon map in the presence of 10% white Gaussian noise. The true dimension $D=1.21$ is indicated by a dashed line. For simulations with additive noise, the noise-free data, and with dynamical noise, the noisy data were always scaled to standard deviation 1. Hence, any quantity shown in this and in the following plots is dimensionless.

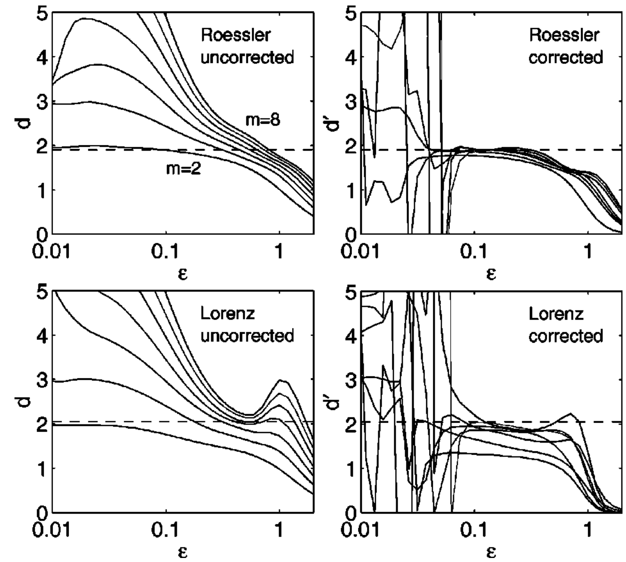

FIG. 2. Uncorrected and corrected estimates of the correlation dimension for embedding dimensions $m=2,\ldots,8$ for the Rössler and Lorenz system in the presence of 10% Gaussian noise. The ''true'' dimensions $D=2.05$ [4] for the Lorenz and $D=1.9$ for the Rössler system (estimated from [13]) are indicated by dashed lines.

Gaussian noise. In contrast, our corrected estimates show scaling and saturation at the correct dimension.

The correction breaks down if the scale ϵ becomes too small: the noisy correlation dimensions converge to the embedding dimensions and do not depend on the dimension of the noise-free signal, and hence, solving for the latter becomes ill defined. Note however, that for low embedding dimensions, ϵ may be considerably smaller than the noise level while still allowing for a reasonable dimension estimate.

The estimation becomes more and more difficult for higher embedding dimensions since then, the necessary relative correction increases strongly. We note again that the magnitude of the correction depends rather on the difference of embedding dimension and correlation dimension of the noise-free signal than on the embedding dimension itself. Thus, if the number of data points is correspondingly larger, one can expect to obtain similar results equally well also for higher-dimensional dynamics.

For other tests of the dimension estimation method we use the time series obtained from the Rössler and Lorenz systems [25] of differential equations that were then superimposed by white Gaussian noise with standard deviation $\sigma=0.1$. For continuous systems the delay time τ can take arbitrary values. Here, we set $\tau=1$ and $\tau=0.25$ for the Rössler and Lorenz systems, respectively.

The results are shown in Fig. 2. Again we find very satisfactory bias removal. While none of the uncorrected dimension estimates shows scaling, the Lorenz system at least approximately saturates at the correct value, at $\epsilon\approx0.5$. Remarkably, for the noisy Rössler system the correct dimension cannot even be anticipated before bias removal, but is nicely recovered by our correction scheme.

For very large noise or for very few data the correction still leads to qualitatively correct results. This can be seen in Fig. 3 where we show the dimension estimates for the Hénon map with 50% noise, now using N=20 000 data points and with 10% noise using N=500 data points. However, because of the relatively large fluctuations and the small scaling re-

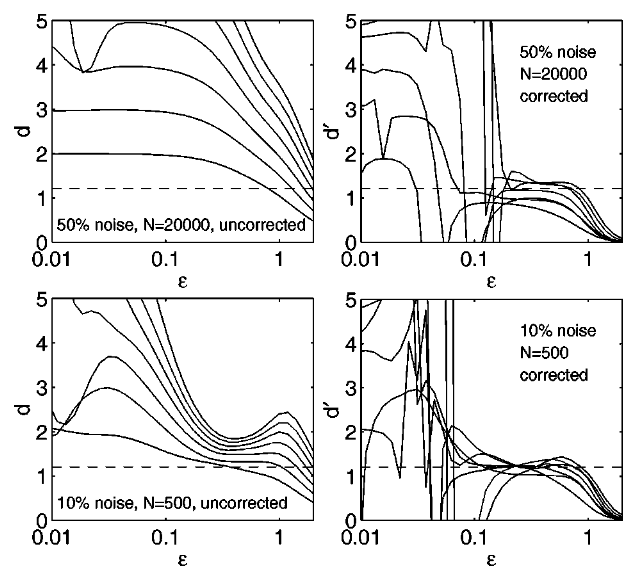

FIG. 3. Estimates of the correlation dimension for the Hénon map with 50% noise level using N=20 000 data points (upper panels) and with 10% noise using N=500 data points (lower panels).

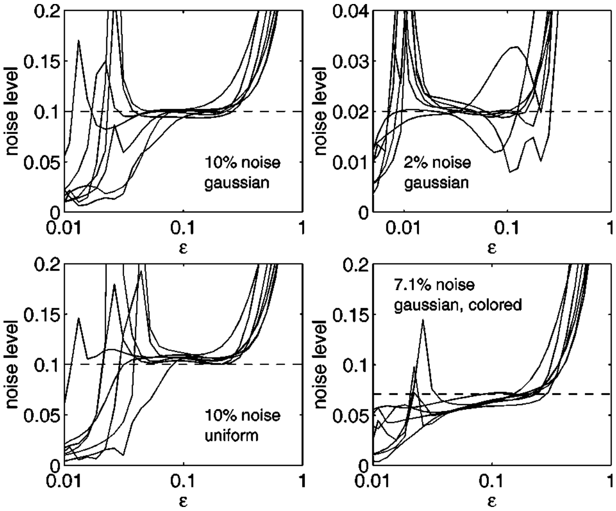

FIG. 4. Estimates of the respective noise levels (dashed line) for embedding dimensions $m=2,\ldots,8$ for the Hénon map data.

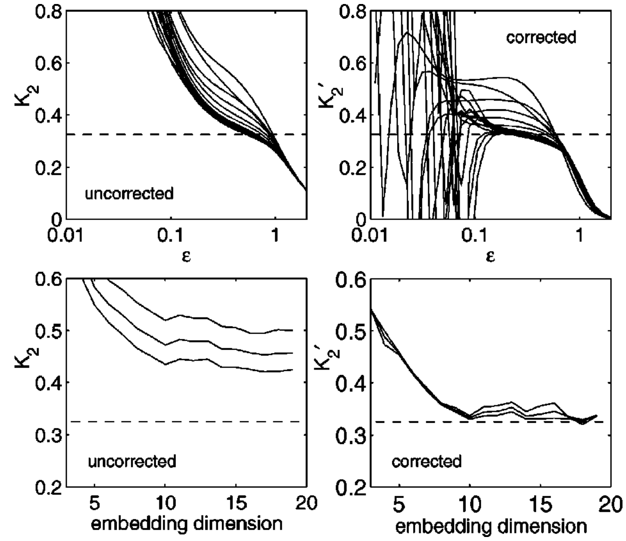

FIG. 5. Estimates of the correlation entropy $K_2 \approx 0.325$ (dashed line) for embedding dimensions $m=3,\ldots,19$ for the Hénon map in the presence of 10% Gaussian noise. Upper panels: K_2 as a function of ϵ. Lower panels: K_2 as a function of embedding dimension for $\epsilon=0.15,0.18,0.21$ chosen from the scaling region.

gime we regard these examples as being at the limit of a reasonable application of our method.

B. Noise level

In Sec. II C we derived an estimator $\sigma(m,\epsilon)$ for the standard deviation of the noise. Like the dimension and K_2 estimators, this construction has the advantage that in contrast to fitting procedures the consistency can be checked both in terms of scaling and saturation properties.

In the upper left panel of Fig. 4 we show the noise estimator for the Hénon map with 10% white Gaussian noise. A scaling region is well established at scales in the order of the noise level where the dependence of Δ on the noise level is maximal. Taking, e.g., the estimates in the middle of the scaling range at $\epsilon=0.1051$ results in a mean of $\sigma=0.0990$ with a standard deviation of 0.003 in excellent agreement with the true value.

The estimate of the noise level becomes less stable for smaller noise since estimates at smaller scales are based on fewer data. This can be seen in the upper right panel of Fig. 4 showing the result for 2% noise.

For comparison, we also plot the results for uniform and colored noise in the lower panels of Fig. 4. In both cases we find systematic but small deviations from the true noise level. Uniform noise typically results in nice scaling behavior with a small overestimation of the noise level, while in the case of colored noise we observe a systematic underestimation, which—analogous to the dimension estimates—is more pronounced for small ϵ.

C. K_2 entropy

The results for the K_2 entropy estimates are shown in Fig. 5. In contrast to the uncorrected estimates (left upper panel) the corrected ones scale properly (right upper panel). The lower boundary of the scaling range grows for increasing embedding dimensions, which is in fact a well-known property also for noise-free data [22]. A saturation behavior at the correct value is verifiable only for large embedding dimensions ($m \geqslant 10$) as it also would have been expected from the noise-free data [22].

The saturation is seen more clearly in the lower panels of Fig. 5 where we show the K_2 estimates as a function of the embedding dimension for three fixed scales within the scaling regime. In the literature different values are given for K_2 of the Hénon map. Although it is not the primary goal of this paper to settle this issue, our findings rather support $K_2 \approx 0.325$ as stated in [24] than $K_2 \approx 0.29$ from [23].

D. Other types of noise

For the derivation of the correction formulas we assumed additive white Gaussian noise. In real-world data this assumption will not hold exactly.

In order to test the validity of our method also for noise with other probability distributions, we added uniform white noise, again with standard deviation $\sigma=0.1$, to the data generated by the Hénon map and applied the same correction as in the previous section. The upper panels of Fig. 6 show the uncorrected and corrected dimension estimates. Though the estimates are slightly worse, if compared with the Gaussian noise case (cf. Fig. 1), both saturation and scaling are clearly visible after correction.

We now address the case of nonwhite noise that, in principle, could be overcome by choosing a large value of τ [see Eq. (2)] or by filtering the data appropriately before performing the actual analysis in order to ''whiten'' the noise. However, a too large τ also complicates the dimension estimation since due to the intrinsic chaotic nature of the dynamics, functional dependencies between consecutive data points are diminished, and the correct filter to whiten the noise without causing severe phase distortions of the system itself is usually unknown. Still, in order to get satisfactory results within

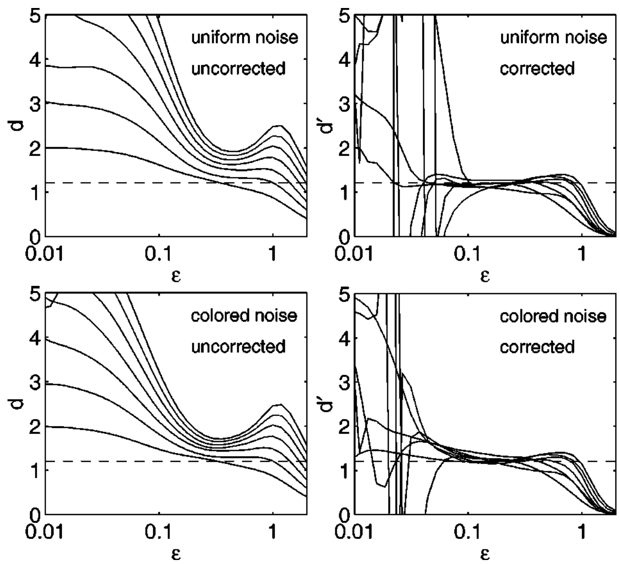

FIG. 6. Uncorrected (left panels) and corrected (right panels) estimates of the correlation dimension for embedding dimensions $m=2,\dots,8$ for the Hénon map in the presence of uniformly distributed white noise, $\sigma=0.1$ (upper panels) and Gaussian distributed colored noise, $\sigma=0.071$ (lower panels). The true dimension $D=1.21$ is always indicated by a dashed line.

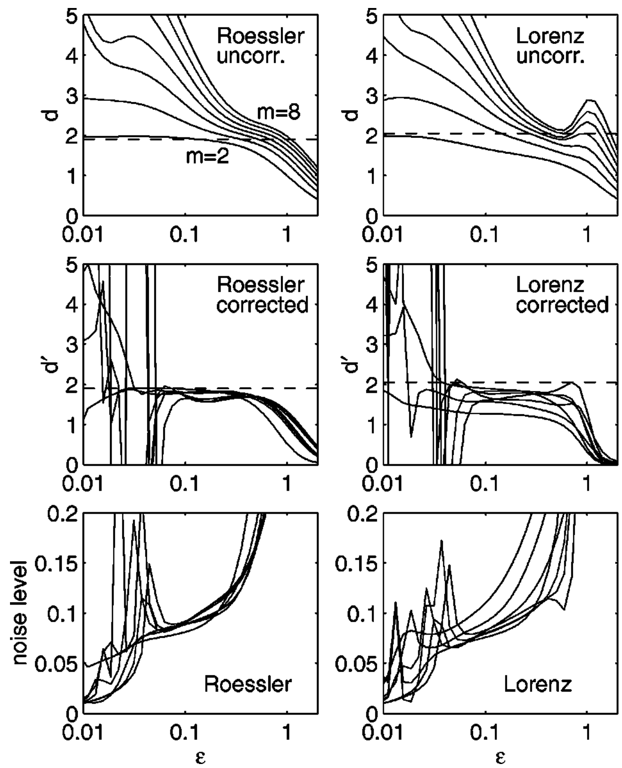

FIG. 7. Estimates of correlation dimension for the Rössler system (left panels) and for the Lorenz system (right panels) perturbed by dynamical noise. The σ estimates indicate a noise level of about 10%. Note, however, that no clear scaling regime can be observed for the noise level estimates.

the proposed approach it would be advisable to avoid extreme deviations from the white-noise case since the latter may ''mimic low-dimensional chaotic attractors'' [28].

In order to check the robustness of our method against non-iid noise, we added low pass filtered Gaussian noise to the time series. The low pass filter was implemented by applying a moving average of order two to white Gaussian noise with $\sigma=0.1$,

$$\eta_i \rightarrow (\eta_i + \eta_{i+1})/2, \tag{42}$$

$$\sigma \rightarrow \sigma/\sqrt{2} \approx 0.071. \tag{43}$$

The power spectrum of this colored noise then reads $P(\omega)=[1+\cos(\omega)]/2$, where $\omega \in [0,\pi]$ is the frequency and π is the Nyquist frequency.

The results are shown in the lower panels of Fig. 6. Again, we find a major improvement after applying the correction. However, if ϵ is smaller than the noise level we observe a systematic overestimate of the dimension in agreement with the theoretical considerations stating that the approximations are more accurate for larger ϵ.

We finally present two examples using dynamical noise that arises when the dynamical system itself and not merely the measurement is disturbed by noise. This was realized by a small distortion of one variable of the Rössler and Lorenz system in each step of the integration of the differential equations.[4] Though the estimates (see Fig. 7) are less stable than for additive white Gaussian noise, we approximately recover scaling and saturation at the correct values. Typically, the correction leads to small but systematic underestimates of the dimensions. Estimates of the noise level indicate that this dynamical noise roughly corresponds to 10% additive noise. However, the lack of a clear scaling regime can readily serve as an indicator that the assumption of additive white noise is inconsistent.

IV. CONCLUSION

We introduced a Gaussian kernel based method for reducing the noise bias in estimates of correlation dimension and K_2 entropy of dynamical system attractors.

In contrast to most proposed methods and to all existing exact methods, our approach is *local* in scale space and requires for each scale, only the knowledge of the function Δ: the difference of the uncorrected dimension estimates for *two adjacent* embedding dimensions. Hence, both scaling and saturation can still be checked after bias removal. For practical purposes this latter property is highly desirable since in most applications it is not clear, whether the time series under consideration is governed by deterministic chaotic dynamics or by an unstructured stochastic noise process.

We demonstrated the performance for various examples using data from the Hénon map and the Lorenz and Rössler system. In all cases the noise level was chosen to destroy scaling and saturation at the true correlation dimension for the uncorrected dimension estimate while these properties could be sufficiently recovered after bias removal. We could

[4]The Hénon map is unstable with respect to this perturbation.

show experimentally that the proposed approach is not sensitive to the distribution of the noise (Gaussian versus uniform). Experiments with dynamical noise led to qualitatively correct results with a small but systematic underestimate of the dimension.

Special emphasis was given to the problem of correlated noise. Also for this case we could derive an approximate refined scaling law that turns out to depend on the eigenvalues of the nontrivial noise covariance matrix in an m-dimensional embedding space. We could provide theoretical and experimental evidence that our method, which does not require knowledge of the noise characteristics itself, is practicable as long as the deviation from the white-noise case is not too large.

Estimation of the noise level and bias correction of K_2 entropy was achieved similarly. Again, Δ turned out to be the crucial quantity sufficient to define a ''noise function,'' which should scale and saturate at the correct noise level, and to construct a bias-free estimator of K_2. We could demonstrate a promising performance in case of the Hénon map even though the estimate of K_2 in the presence of noise is generally considered to be an exceptionally difficult task [18].

Future research will be devoted to applications of our estimation method to real-world data.

ACKNOWLEDGMENTS

K.-R.M. and A.Z. acknowledge funding by the EU project BLISS (Grant No. IST-1999-14190). This study was supported in part by the National Foundation for Functional Brain Imaging (USA).

APPENDIX

In case of colored noise, the probability distribution for η_α reads (henceforth omitting the index α on η)

$$p(\eta)\sim\frac{1}{\sqrt{\det(R)}}\exp\left(-\frac{\eta R^{-1}\eta}{4}\right) \tag{A1}$$

with R given by Eq. (27). In order to evaluate the expectation of the correlation sum

$$\langle C^g(\epsilon,m)\rangle\sim\sum_\alpha\frac{1}{\sqrt{\det(R)}}\int D\eta \times\exp\left(-\frac{|\mathbf{x}_\alpha+\eta|^2}{4\epsilon^2}-\frac{\eta R^{-1}\eta}{4}\right) \tag{A2}$$

with $D\eta:=\prod_{k=1}^m d\eta_k$, we reexpress the exponent as

$$-\frac{|\mathbf{x}_\alpha+\eta|^2}{4\epsilon^2}-\frac{\eta R^{-1}\eta}{4}=-\frac{\mathbf{x}_\alpha(1-A^{-1}A^{-T})\mathbf{x}_\alpha}{4\epsilon^2}-\frac{|\boldsymbol{\xi}+A^{-T}\mathbf{x}_\alpha|^2}{4\epsilon^2} \tag{A3}$$

with the definitions

$$A^TA:=1+\epsilon^2R^{-1} \tag{A4}$$

and

$$\boldsymbol{\xi}:=A\eta. \tag{A5}$$

Using $D\eta=D\xi/\det(A)$ we find

$$\langle C^g(\epsilon,m)\rangle\sim\frac{\epsilon^m}{\det A\sqrt{\det R}}\sum_\alpha\exp\left[-\frac{\mathbf{x}_\alpha(1-A^{-1}A^{-T})\mathbf{x}_\alpha}{4\epsilon^2}\right]. \tag{A6}$$

As we see, the presence of nonwhite noise has led to nonspherical, ellipsoidal Gaussian kernels with an m-dimensional volume V given by

$$V\sim\left[\det\left(\frac{1-A^{-1}A^{-T}}{4\epsilon^2}\right)\right]^{-1/2}. \tag{A7}$$

We now assume that the correlation sum scales in the noise-free case with the volume as $\sim V^{D/m}$. This is indeed an approximation because the exact scaling law can in general also depend on the specific shape of the ellipsoid that varies as a function of ϵ. The length l_k of the kth axis is given by the square root of the kth eigenvalue of the matrix in the exponent of Eq. (A6): $l_k=\sqrt{(\sigma_k^2+\epsilon^2)}$ with σ_k^2 being the kth eigenvalue of R. For large ϵ the ellipsoid becomes spherical; especially, its shape becomes independent of ϵ.

Ignoring ϵ dependence of the shape of the ellipsoid we arrive at

$$\langle C^g(\epsilon,m)\rangle\sim\frac{\epsilon^mV^{D/m}}{\det A\sqrt{\det R}}. \tag{A8}$$

Since A is merely a function of R we may express $\langle C^g(\epsilon,m)\rangle$ by the eigenvalues (σ_k^2), e.g.,

$$\det(A)=\det(\sqrt{1+\epsilon^2R^{-1}})=\prod_k\left(\frac{\sigma_k^2+\epsilon^2}{\sigma_k^2}\right)^{1/2}, \tag{A9}$$

leading finally to

$$\langle C^g(\epsilon,m)\rangle\sim\epsilon^m\prod_{k=1}^m(\sigma_k^2+\epsilon^2)^{(D-m)/2m}. \tag{A10}$$

The calculation of the correlation dimension as given by Eqs. (25) and (26) is now straightforward.

[1] A. Babloyantz and A. Destexhe, Proc. Natl. Acad. Sci. U.S.A. **83**, 3513 (1986).
[2] A. Babloyantz, J. M. Salazar, and C. Nicolis, Phys. Lett. **111A**, 152 (1985).
[3] V. K. Yeragani, K. Srinivasan, S. Vempati, R. Pohl, and R. Balon, J. Appl. Physiol. **75**, 2429 (1993).
[4] P. Grassberger and I. Procaccia, Phys. Rev. Lett. **50**, 346 (1983).
[5] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[6] K. Pawelzik and H. G. Schuster, Phys. Rev. A **35**, 481 (1987).
[7] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 1997).
[8] L. A. Smith, Phys. Lett. A **133**, 283 (1988).
[9] A. Stefanovska, S. Strle, and P. Kroselj, Phys. Lett. A **235**, 24 (1997).
[10] E. J. Kostelich and T. Schreiber, Phys. Rev. E **48**, 1752 (1993).
[11] T. Schreiber, Phys. Rev. E **56**, 274 (1997).
[12] D. Kugiumtzis, Int. J. Bifurcation Chaos Appl. Sci. Eng. **6**, 1283 (1997).
[13] H. Oltmans and P. J. T. Verheijen, Phys. Rev. E **56**, 1160 (1997).
[14] M. Möller, W. Lange, F. Mitschke, N. B. Abraham, and U. Hübner, Phys. Lett. A **138**, 176 (1989).
[15] E. Ott, E. D. Yorke, and J. A. Yorke, Physica D **16**, 62 (1985).
[16] J. C. Schouten, F. Takens, and C. M. van den Bleek, Phys. Rev. E **50**, 1851 (1994).
[17] C. Diks, Phys. Rev. E **53**, 4263 (1996).
[18] D. Yu, M. Small, R. G. Harrison, and C. Diks, Phys. Rev. E **61**, 3750 (2000).
[19] T. Schreiber, Phys. Rev. E **48**, R13 (1993).
[20] G. G. Szpiro, Physica D **65**, 289 (1993).
[21] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[22] M. Frank *et al.*, Physica D **65**, 359 (1993).
[23] P. Grassberger, R. Badii, and A. Politi, J. Stat. Phys. **51**, 135 (1988).
[24] P. Grassberger and I. Procaccia, Physica D **13**, 34 (1984).
[25] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).
[26] J. Kurths and H. Herzel, Physica D **25**, 165 (1987).
[27] J. Theiler and T. Lookman, Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 765 (1993).
[28] P. E. Rapp, A. M. Albano, T. I. Schmah, and L. A. Farwell, Phys. Rev. E **47**, 2289 (1993).