HTML ABSTRACT + LINKS



# Estimation of dynamical invariants without embedding by recurrence plots

M. Thiel and M. C. Romano
*University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany*

P. L. Read
*Atmospheric, Oceanic and Planetary Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, United Kingdom*

J. Kurths
*University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany*

(Received 18 September 2003; accepted 16 January 2004; published online 24 March 2004)

In this paper we show that two dynamical invariants, the second order Rényi entropy and the correlation dimension, can be estimated from recurrence plots (RPs) with arbitrary embedding dimension and delay. This fact is interesting as these quantities are even invariant if no embedding is used. This is an important advantage of RPs compared to other techniques of nonlinear data analysis. These estimates for the correlation dimension and entropy are robust and, moreover, can be obtained at a low numerical cost. We exemplify our results for the Rössler system, the funnel attractor and the Mackey–Glass system. In the last part of the paper we estimate dynamical invariants for data from some fluid dynamical experiments and confirm previous evidence for low dimensional chaos in this experimental system. © *2004 American Institute of Physics.*
[DOI: 10.1063/1.1667633]

**We investigate a putatively intuitive method to represent graphically the structures of the phase space of dynamical systems called recurrence plot (RP). We present evidence that there are dynamical invariants (such as the Rényi entropy of second order $K_2$ and the correlation dimension $D_2$) which can be estimated from the plot, and that they are independent of the embedding parameters. These invariants can even be computed if no embedding is used. The application to prototypical systems (Rössler, funnel attractor, Mackey–Glass), and to experimental flow data in the last part of the paper, shows that the estimation of invariants by means of recurrence plots is very robust and confirms previous evidence for low dimensional chaos for the flow.**

## I. INTRODUCTION

Recurrence plots (RPs) visualize the behavior of trajectories in phase space.[1,2] They are a graphical representation of the matrix

$$\mathbf{R}_{i,j}=\Theta(\varepsilon-\|\mathbf{x}_i-\mathbf{x}_j\|), \quad i,j=1,\ldots,N, \tag{1}$$

where $\mathbf{x}_i \in \mathcal{R}^d$ stands for the point in phase space at which the system is situated at time $i$, $\varepsilon$ is a predefined threshold, and $\Theta(\cdot)$ is the Heaviside function. One assigns a ''black'' dot to the value one and a ''white'' dot to the value zero. The two dimensional graphical representation of $R_{i,j}$ then is called a RP.

To quantify the structures that occur in RPs, Webber and Zbilut have proposed several measures in their seminal paper,[3] that constitute the recurrence quantification analysis (RQA). Basing on these measures, RPs have become very popular and have been applied to various experimental data, especially in physiology and earth science,[3–8] A further potential advantage of RPs is that they enable the computation of well-known dynamical invariants, such as the correlation entropy $K_2$ and correlation dimension $D_2$.[9,10] One crucial point is whether these estimates depend on the parameters used in the computations. To calculate a recurrence plot, one has to fix three parameters in advance. One of them is the threshold $\varepsilon$. Recently a lower bound for $\varepsilon$ in the presence of noise has been calculated.[11] It has been shown that, in order to resolve fine structures, $\varepsilon$ should not be chosen too large either.[9,5] So upper and lower bounds for $\varepsilon$ are known, at least theoretically.

In the case of experimental data there is often only one component (i.e., a univariate time series) available. Hence, the embedding dimension $d$ and the delay $\tau$ needed for the embedding of the time series

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

have additionally to be fixed.[12] To estimate the ''optimal'' embedding dimension $d$, the method of false nearest neighbors or single value decomposition has usually been applied.[13,14] The choice of the delay $\tau$ on the other hand is still under debate. The most frequently applied methods for the estimation of $\tau$ are based on the autocorrelation function or on the mutual information. The choice of the method is rather arbitrary.[15] There are further more sophisticated methods for the reconstruction of the phase space (e.g., differential embedding, integral embedding, etc.), cf. Ref. 16 and further references therein.

Figure 1 presents RPs for three different systems (uniformly distributed and independent noise, a sine function and the chaotic Rössler system with standard parameters) and for different embedding parameters. The left panel shows the plots for embedding dimension $d=1$, i.e., no embedding is

 © 2004 American Institute of Physics

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

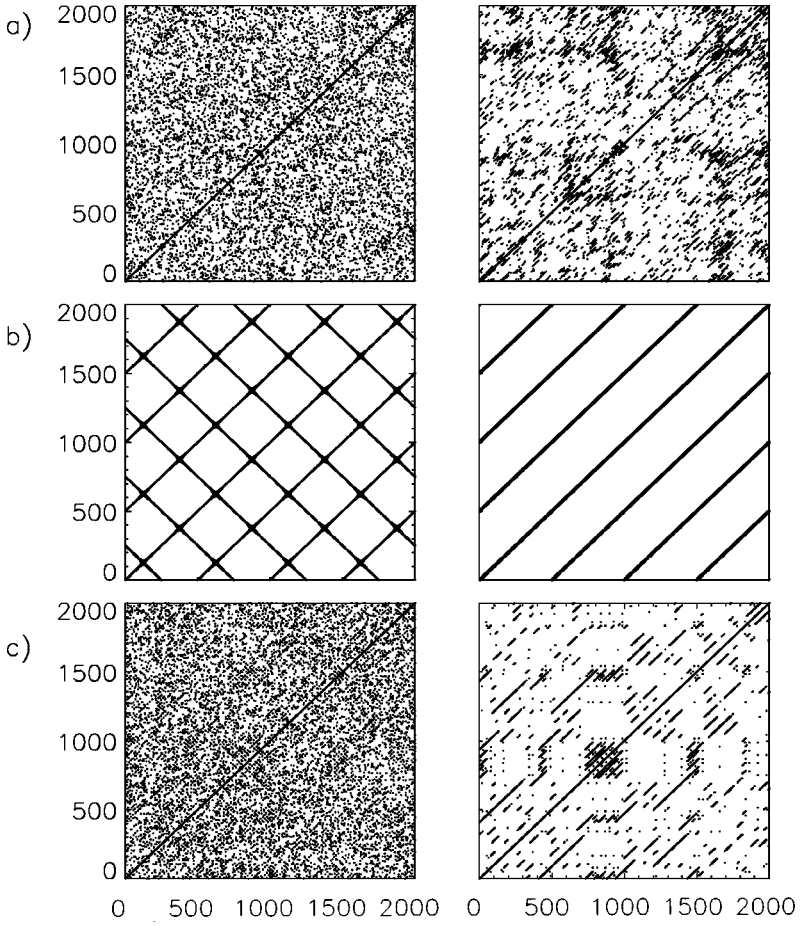


FIG. 1. RPs for uniformly distributed noise (a), the sine function (b), and the Rössler system (c). The left panel shows the plots for $d=1$. The right panel represents the plots for $d=14$, $d=2$, and $d=3$, respectively, from top to bottom. $\varepsilon$ is chosen so that the recurrence rate (percentage of black point in the plot) is the same for the embedded and nonembedded time series.

used. The right panel illustrates the same graphics for $d>1$. The plots show that the embedding via Eq. (2) influences the structure in the RPs drastically. We will study in this paper whether embedding changes the estimates of the measures that are used to quantify the dynamics of the underlying system.

The outline of the paper is as follows: In Sec. II we show that the correlation entropy and dimension can be estimated from RPs and that the estimation can be performed without embedding. The results hold for discrete and continuous systems. In Sec. III A we estimate the entropy and dimension for the Rössler system for standard parameters and in the funnel regime. We then present (Sec. III B) results for the (infinite dimensional) Mackey–Glass system. For all these different attractors we obtain, at a low numerical cost, estimates that are in accordance with literature. Finally, we analyze data from fluid dynamical experiments and confirm previous results that indicate low dimensional chaos in the flow system (Sec. IV).

## II. INDEPENDENCE OF THE EMBEDDING PARAMETERS

We first summarize some results concerning the cumulative distribution of the diagonals $P_\varepsilon^c(l)$, i.e., the lines of at least length $l$, in a RP. First we recall the definition of recurrence rate RATE, i.e., the probability to find a black or recurrence point in the RP

$$\text{RATE}=\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{i,j}. \tag{3}$$

$P_\varepsilon^c(l)$ is the probability to find in the RP a diagonal line of at least length $l$.[3] Considering in Eq. (1) the maximum norm ($L_\infty$ norm), we get

$$P_\varepsilon^c(l)=\frac{1}{N^2}\sum_{i,j=1}^{N}\prod_{m=0}^{l-1}\Theta(\varepsilon-\|\mathbf{x}_{i+m}-\mathbf{x}_{j+m}\|_\infty). \tag{4}$$

To link $P_\varepsilon^c(l)$ to known dynamical measures, we start with the correlation integral[17]

$$C(\varepsilon)=\lim_{N\to\infty}\frac{1}{N^2}\times\{\text{number of pairs } (i,j) \text{ with } \|\mathbf{x}_i-\mathbf{x}_j\|<\varepsilon\}. \tag{5}$$

There the Euclidean norm was used, but in principle the choice is arbitrary. It is important to emphasize that the definition of the recurrence rate of a RP of length $N$ coincides with the definition of the correlation integral if $N$ tends to infinity

$$C(\varepsilon)=\lim_{N\to\infty}\frac{1}{N^2}\sum_{i=1}^{N}\Theta(\varepsilon-\|\mathbf{x}_i-\mathbf{x}_j\|)\overset{\text{Eq. (1)}}{=}\lim_{N\to\infty}\frac{1}{N^2}\sum_{i,j=1}^{N}\mathbf{R}_{i,j}=\lim_{N\to\infty}\text{RATE}. \tag{6}$$

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

This fact allows us to link the known results about the correlation integral to the structures in RPs.

Suppose we have a trajectory $\mathbf{x}(t)$ in the basin of attraction of an attractor in phase space and the state of the continuous system is measured at time intervals $\tau$. Let $\{1,2,\ldots,M(\varepsilon)\}$ be a partition of the attractor in boxes of size $\varepsilon$. Then $p(i_1,\ldots,i_l)$ is the joint probability that $\mathbf{x}(t=\tau)$ is in box $i_1$, $\mathbf{x}(t=2\tau)$ is in box $i_2,\ldots,$ and $\mathbf{x}(t=l\tau)$ is in box $i_l$. The order-2 Rényi entropy $K_2$ (Refs. 18, 19) is then defined as

$$K_2=-\lim_{\tau\to 0}\lim_{\varepsilon\to 0}\lim_{l\to\infty}\frac{1}{l\tau}\ln\sum_{i_1,\ldots,i_l}p^2(i_1,\ldots,i_l). \tag{7}$$

In order to estimate $\Sigma_{i_1,\ldots,i_l}p^2(i_1,\ldots,i_l)$, one considers the number of pairs $(\mathbf{x}_i,\mathbf{x}_j)$ such that

$$\|\mathbf{x}_{i+k}-\mathbf{x}_{j+k}\|<\varepsilon \quad \text{for } k=1,\ldots,l. \tag{8}$$

Grassberger and Procaccia claim[20] that this is roughly equivalent to demanding that the sets $\{\mathbf{x}_{i+1},\mathbf{x}_{i+2},\ldots,\mathbf{x}_{i+l}\}$ and $\{\mathbf{x}_{j+1},\mathbf{x}_{j+2},\ldots,\mathbf{x}_{j+l}\}$ are pairwise in the same box of the space–time mesh. They define

$$C_l(\varepsilon)=\lim_{N\to\infty}\frac{1}{N^2}\times\{\text{number of pairs } (i,j) \text{ with } |\mathbf{x}_{i+1}-\mathbf{x}_{j+1}|^2+\cdots+|\mathbf{x}_{i+l}-\mathbf{x}_{j+l}|^2\leqslant\varepsilon^2\} \tag{9}$$

and state that due to the exponential divergence of the trajectories, requiring $\Sigma_{k=1}^{l}|\mathbf{x}_{i+k}-\mathbf{x}_{j+k}|^2\leqslant\varepsilon^2$ is essentially equivalent to the condition given by Eq. (8). Therefore, up to a factor of order unity one finds

$$C_l(\varepsilon)\simeq\sum_{i_1,\ldots,i_l}p^2(i_1,\ldots,i_l) \tag{10}$$

and the relationship

$$C_l(\varepsilon)\sim\varepsilon^{D_2}\exp(-l\tau K_2) \tag{11}$$

holds, where $D_2$ is the correlation dimension.[21]

Note that the condition $|\mathbf{x}_{i+k}-\mathbf{x}_{j+k}|<\varepsilon$ for $k=1,\ldots,l$ coincides exactly with the definition of the cumulative distribution of diagonals in the RP of $\{\mathbf{x}_i\}_{i=1}^{N}$. Hence, we can infer also an interrelation of Eq. (11) with RPs

$$P_\varepsilon^c(l)\simeq\sum_{i_1,\ldots,i_l}p^2(i_1,\ldots,i_l)\simeq C_l(\varepsilon)\sim\varepsilon^{D_2}\exp(-l\tau K_2). \tag{12}$$

Therefore, if we represent $\ln(P_\varepsilon^c(l))$ versus $l$ we obtain a straight line with slope $-\hat{K}_2(\varepsilon)\tau$, where $\hat{K}_2(\varepsilon)$ is an estimator for $K_2$.

Next, we treat the case of a reconstructed phase space. As we have shown, embedding can change the structure of the RPs considerably (Fig. 1). The question we want to tackle is how the estimates of $K_2$ and also of the correlation dimension $D_2$ depend on the embedding parameters.

Again we consider the cumulative distribution of diagonal lines $P_\varepsilon^c(l)$ in the RP. To make clear that we use the delay reconstructed phase space—since we have only observed one component—we write for the cumulative distribution $P_{\varepsilon,\tau}^d$ from now on, where $d$ is again the embedding dimension and $\tau$ the delay used for the reconstruction. As we have decided to choose the maximum norm in Eq. (1), $P_{\varepsilon,\tau}^d(l)$ is given by

$$P_{\varepsilon,\tau}^d(l)=\frac{1}{N^2}\sum_{i,j=1}^{N}\prod_{m=0}^{l-1}\Theta(\varepsilon-\max_{k=0,\ldots,d-1}|x_{i+m+k\tau}-x_{j+m+k\tau}|). \tag{13}$$

Obviously,

$$\prod_{m=0}^{l-1}\Theta(\varepsilon-\max_{k=0,\ldots,d-1}|x_{i+m+k\tau}-x_{j+m+k\tau}|)=\Theta\left(\varepsilon-\max_{\substack{m=0,\ldots,l-1\\k=0,\ldots,d-1}}|x_{i+m+k\tau}-x_{j+m+k\tau}|\right) \tag{14}$$

holds. Equation (14) can be interpreted as testing whether the conditions

$$|x_{i+m+k\tau}-x_{j+m+k\tau}|<\varepsilon \quad \forall \begin{matrix} m=0,\ldots,l-1\\ k=0,\ldots,d-1\end{matrix} \tag{15}$$

are met. The terms on both sides of Eq. (14) are equal to one if all conditions are simultaneously met and zero otherwise. Hence, $P_{\varepsilon,\tau}^d(l)$ in Eq. (13) can be interpreted as an estimate of the probability that all the conditions in Eq. (15) are simultaneously met.

Note that for the estimation of the correlation dimension and entropy the Grassberger–Procaccia integral is frequently used in the form

$$C_\varepsilon^{d,\tau}=P_{\varepsilon,\tau}^d(1)=\frac{1}{N^2}\sum_{i,j=1}^{N}\Theta(\varepsilon-\max_{k=0,\ldots,d-1}|x_{i+k\tau}-x_{j+k\tau}|), \tag{16}$$

which differs from Eq. (13). The indices of subsequent entries $x_n$ of the time series that enter analog conditions, as given by Eq. (15), are now separated by $\tau$ steps.

Equations (15) are a set of $l\cdot d$ conditions that are in general not independent. If for example $m+k\tau=m'+k'\tau$, one of the two conditions

$$\varepsilon>|x_{i+m+k\tau}-x_{j+m+k\tau}|$$

or

$$\varepsilon>|x_{i+m'+k'\tau}-x_{j+m'+k'\tau}|$$

is redundant. If $l$ is sufficiently large, i.e., $l>\tau$, we can condense the conditions Eqs. (15) and find the $l+(d-1)\tau$ relations

$$|x_{i+m}-x_{j+m}|<\varepsilon \quad \forall \quad m=0,\ldots,l-1+(d-1)\tau. \tag{17}$$

These conditions have to be met in order to find a line of at least length $l-1+(d-1)\tau$ if the time series $x_i$ is not embedded, i.e.,

$$P_\varepsilon^1(l-1+(d-1)\tau)=\frac{1}{N^2}\sum_{i,j=1}^{N}\prod_{m=0}^{l-1+(d-1)\tau}\Theta(\varepsilon-|x_{i+m}-x_{j+m}|). \tag{18}$$

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

Note, that the further condition $l>\tau$ has to be met. More generally one finds

$$\begin{aligned} P^{d}_{\varepsilon,\tau}(l) &= P^{1}_{\varepsilon,\tau}(l-1+(d-1)\tau) \\ &= P^{1}_{\varepsilon,\tau}\Big(\underbrace{[l\mp\Delta d\cdot\tau]}_{l'}-1+(\underbrace{[d\pm\Delta d]}_{d'}-1)\tau\Big) \\ &= P^{d'}_{\varepsilon,\tau}(l') \end{aligned} \tag{19}$$

provided that $l,l'>\tau$ and $d,d'\geqslant 1$. Equation (19) shows that the decay of $P^{d}_{\varepsilon,\tau}(l)$ is essentially the same for different embedding dimensions and delays. The curve is only shifted to larger $l$'s if the dimension is decreased. The condition for $P^{d}_{\varepsilon,\tau}(l)=P^{d'}_{\varepsilon,\tau}(l')$ is

$$l+(d-1)\tau=l'+(d'-1)\tau'. \tag{20}$$

This equality only holds (strictly) due to the special choice of the maximum norm. However, for other choices of the norm and other embeddings it still holds approximately. As only the slope of $P^{d}_{\varepsilon,\tau}(l)$ for large $l$ is relevant for the estimation of $K_2$ by means of RPs, the estimate does not depend on $d$ and $\tau$. The main consequence of this independence is that the correlation entropy (and—as we will show—also the correlation dimension) can be estimated even if no embedding is used.

An estimator of the correlation dimension $\hat{D}_2$ (see Ref. 9) from the RP has also been derived

$$\hat{D}_2(\varepsilon)=\ln\left(\frac{P^{c}_{\varepsilon}(l)}{P^{c}_{\varepsilon+\Delta\varepsilon}(l)}\right)\left(\ln\left(\frac{\varepsilon}{\varepsilon+\Delta\varepsilon}\right)\right)^{-1}, \tag{21}$$

where $P^{c}_{\varepsilon}(l)$ is the cumulative distribution of the real coordinates. Using delay embedding and substituting Eq. (19) in Eq. (21), we find

$$\begin{aligned} &\ln(P^{d}_{\varepsilon_1,\tau}(l))-\ln(P^{d}_{\varepsilon_2,\tau}(l)) \\ &\quad=\ln(P^{d'}_{\varepsilon_1,\tau}(l'))-\ln(P^{d'}_{\varepsilon_2,\tau}(l')), \end{aligned}$$

i.e.,

$$\ln\left(\frac{P^{d}_{\varepsilon_1,\tau}(l)}{P^{d}_{\varepsilon_2,\tau}(l)}\right)=\ln\left(\frac{P^{d'}_{\varepsilon_1,\tau}(l')}{P^{d'}_{\varepsilon_2,\tau}(l')}\right),$$

i.e., the estimate $D_2$ is independent of the choice of the embedding parameters and also applies if no embedding is used.

## III. MODEL SYSTEMS

In this section we estimate the correlation entropy and dimension from RPs for three different topological situations. First, we analyze the phase coherent Rössler system with standard parameters. Second, we perform the analysis for the nonphase coherent Rössler system in a funnel regime. Then we consider the chaotic Mackey–Glass system. For each of these cases we obtain robust estimates.

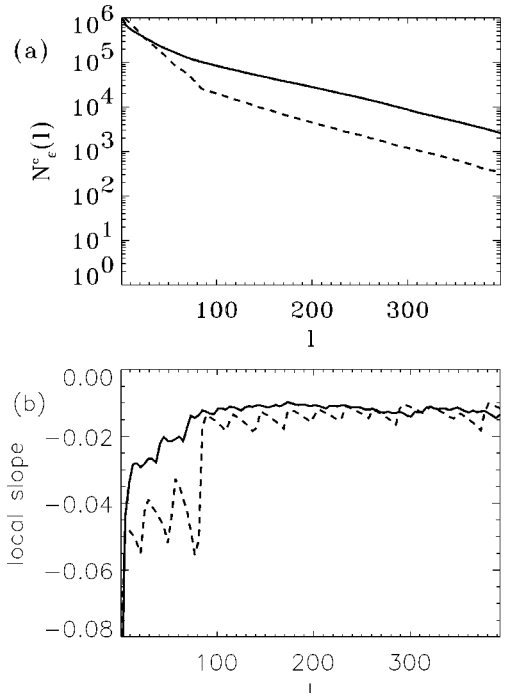


FIG. 2. (a) Comparison of the cumulative number of diagonal lines of length $l$ ($N^{c}_{\varepsilon}=N^2\cdot P^{d}_{\varepsilon,\tau}(l)$) for the Rössler system with real coordinates (dashed line) and embedded using $d=3,\tau=8$ (bold line) using the $x$-component. (b) Local slope of the curves represented in (a), for the real coordinates (dashed) and for the embedded coordinates (bold).

### A. The Rössler system

We first exemplify our results for the case of the chaotic, phase-coherent Rössler system[22]

$$\begin{aligned} \dot{x}&=-y-z, \\ \dot{y}&=x+ay, \\ \dot{z}&=b+(x-c)z, \end{aligned} \tag{22}$$

setting $a=b=0.2$ and $c=5.7$. The integration step is $h=0.01$ and the sampling rate $\Delta t=20$, i.e., in absolute time units two subsequent points are separated by a lag of 0.2. The length of the time series is $N=10\,000$ (after neglecting the first 5 000 points).

As we have shown in Sec. II, it is possible to define measures to quantify the structures in RPs which are independent of the embedding parameters. These quantities are invariants known from the theory of dynamical systems and can be estimated by the cumulative distribution of diagonals $P^{c}_{\varepsilon}(l)$. Figure 2(a) represents $N^{c}_{\varepsilon}(l)=N^2\cdot P^{c}_{\varepsilon}(l)$ resp. $N^{c}_{\varepsilon}(l)=N^2\cdot P^{d}_{\varepsilon,\tau}(l)$ for both real and embedded coordinates of the

TABLE I. Estimates of $K_2$ and $D_2$ estimated by RPs based on the original coordinates of the Rössler system, on the embedded ones for arbitrary (!) embedding parameters, and the same values estimated by the Grassberger–Procaccia algorithm.

| System | $K_2$ | $D_2$ |
|---|---|---|
| "Original coordinates" | $0.0675\pm0.004$ | $2.07\pm0.01$ |
| Embedding, $x$-component, arbitrary $d,\tau$ | $0.067\pm0.007$ | $2.06\pm0.06$ |
| G–P algorithm | $0.070\pm0.003$ | $1.81\pm0.02$ |

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

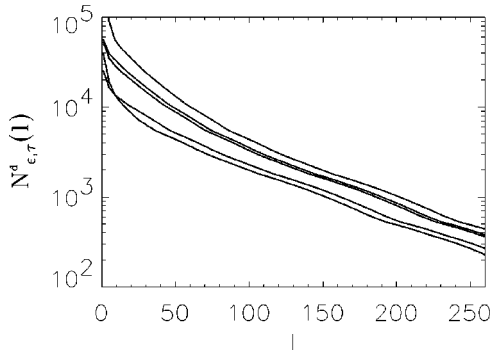


FIG. 3. Comparison of the cumulative number of diagonal lines of length $l$ ($N^d_{\varepsilon,\tau}=N^2\cdot P^d_{\varepsilon,\tau}(l)$) for the Rössler system ($a=b=0.2$, $c=5.7$) for different embedding parameters ($\{d=1\}$, $\{d=3,\tau=6\}$, $\{d=3,\tau=8\}$, $\{d=6,\tau=8\}$, and $\{d=3,\tau=25\}$ from top to bottom). For all cases the $x$-component was used for the embedding.

Rössler system. For small $l$ both curves behave quite differently, but for larger $l$ their slopes become almost identical [Fig. 2(b)], as expected by Takens' theorem. The slope in the latter part corresponds, as we have shown in Sec. II, to the second order Rényi entropy, $K_2$, of the Rössler system (see also Ref. 9 for a more detailed analysis of the Rössler system).

A remarkable point is the occurrence of an additional first slope of $N^c_\varepsilon(l)$ (Fig. 2) that is pronounced in the case of the original coordinates. Thiel *et al.* have recently reported that the first slope may due to the nonhyperbolicity of the attractor.[9] They find that the first scaling region extends from $l=1$ to $l\approx 84$ (i.e., $84\cdot 0.2=16.8$ in absolute time units). Comparing this with the distribution for the embedded coordinates, we obtain three regions which are not so well pronounced. The first one is found between $l=1$ and $l\approx 7$ (i.e., 1.4 in absolute time units), and the second one between $l=8$ and $l\approx 68$ (i.e., 13.6 in absolute time units). The third and last scaling region, which is parallel to the second part of the distribution for the original coordinates, sets in at $l>68$.

Also, using the $y$- and $z$-components of the Rössler system, the slopes of the second part of the logarithmic plot of $P^c(l)$, i.e., the estimate of $K_2$, are the same (to within numerical errors). Table I summarizes the results of the estimates of $D_2$ and $K_2$ based on the RP method for original and embedded coordinates for the Rössler system (parameters see Sec. III A), and estimated by the Grassberger–Procaccia algorithm. The values for the correlation entropy are identical within the error bounds, whereas the estimates of the correlation dimension are slightly higher for the RP method than for the G–P algorithm. Sprott and Rowlands report a correlation dimension of $D_2=1.988\pm 0.078$ and a Kaplan–Yorke dimension of $D_{\mathrm{KY}}=2.031$.[23] The largest Lyapunov exponent is reported to be $\lambda=0.0714$.[24] The independence of the embedding dimension $d$ and the delay $\tau$ is a very important point for the analysis of observed time series. Even though the embedding dimension $d$ and the delay $\tau$ may be difficult to determine, the slope of $P^d_{\varepsilon,\tau}(l)$ for large $l$ will be independent of the special choice of the parameters. Hence, the estimates for $K_2$ and $D_2$ do not depend on these parameters. Figure 3 shows the number of diagonals $N^c_\varepsilon(l)=N^2$

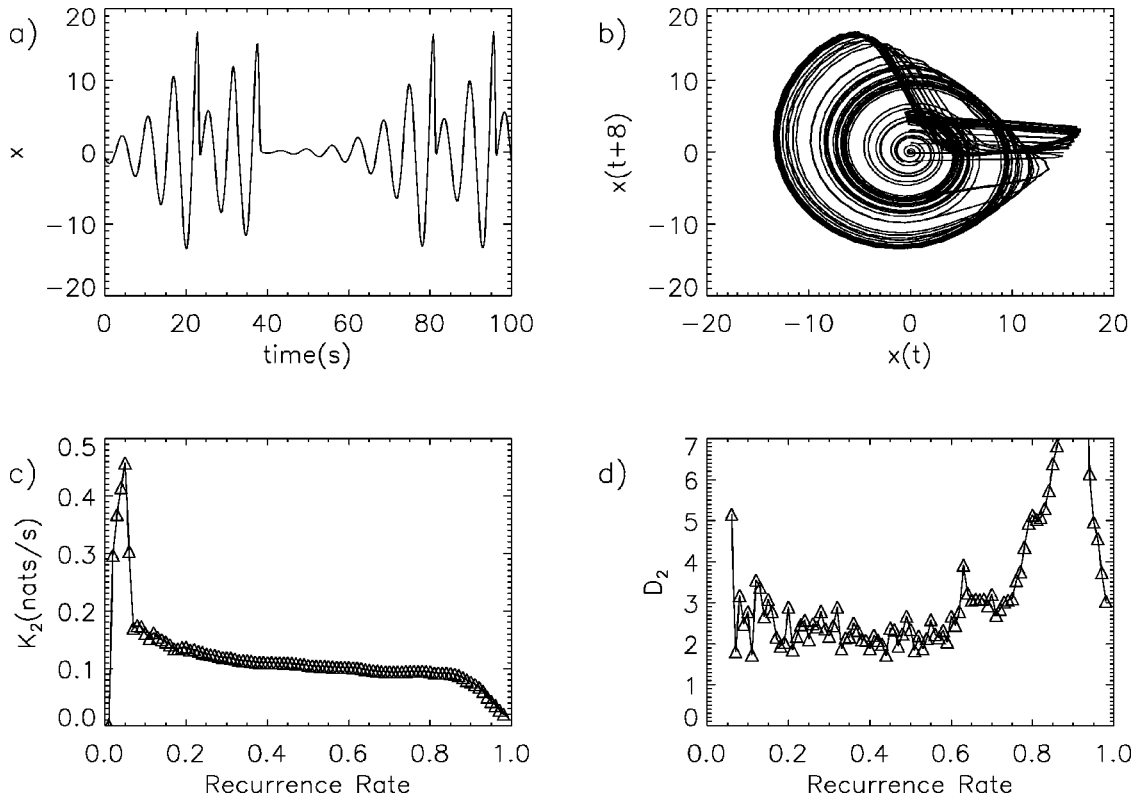


FIG. 4. (a) Sample of the time series of the Rössler attractor in the funnel regime ($a=0.25$, $b=0.4$, and $c=8.5$); (b) reconstructed attractor; (c) dependence of correlation entropy on the recurrence rate; (d) dependence of correlation dimension on the recurrence rate.

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

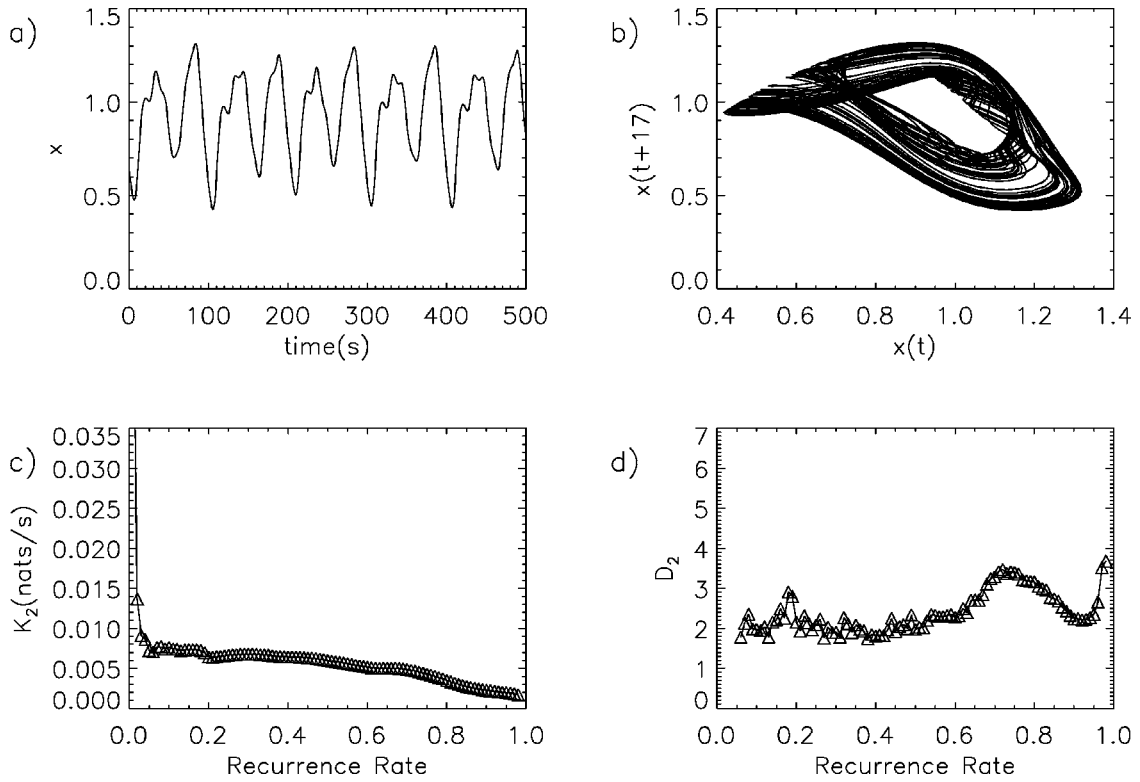


FIG. 5. (a) Sample of the time series of the Mackey–Glass system for $\tau=17$ (chaotic regime); (b) reconstructed attractor; (c) correlation entropy in dependence on the recurrence rate; (d) correlation dimension in dependence on the recurrence rate.

$\cdot P_\varepsilon^c(l)$ on a logarithmic scale for different embedding parameters (see caption). $K_2$ is given by the slope of $N_\varepsilon^c(l)$ for large $l$. For $l>100$ the graphs are (approximately) parallel and so the estimate of $K_2$ is independent of the special choice of the embedding parameters in accordance with Eq. (19). Upon changing the parameters of the Rössler system to $a=0.25$, $b=0.4$, and $c=8.5$, the structure of the Rössler attractor changes into a different, more complicated topology; a funnel attractor as shown in Figs. 4(a) and 4(b). Due to the change of the parameters, the system loses its phase coherency. Performing the upper analysis one finds $K_2=0.112\pm0.050$ nats/s and $D_2=2.25\pm0.22$. Figures 4(c) and 4(d) also show that, without using any embedding, we obtain a broad plateau, that allows the estimation of both invariants. Therefore, the extent of the plateau has to be estimated. This estimation was—in this and the following examples—performed by several independent scientists. The resulting variations were taken into account for the estimation of the errors. However, the error bars are reasonably small. Based on the program "lyap_k" of the TISEAN program package,[28] one obtains an estimate for the largest Lyapunov exponent of $\lambda=0.113\pm0.002$ nats/s, again in very good agreement with our results and, using the program D2 of the same package for the estimate of the correlation dimension, one obtains $D_2=1.93\pm0.28$. Note, that we represent in Figs. 4(c) and 4(d) the entropy and the dimension versus the recurrence rate, not versus $\varepsilon$. We obtain a plateau, as the recurrence rate depends strictly monotonously on $\varepsilon$. However, this representation makes the plateau easier to determine.

### B. The Mackey–Glass system

The independence of the estimates of the correlation dimension and entropy are of special advantage for the analysis of infinite dimensional systems. A dynamical system is infinite dimensional if an infinite set of independent numbers is required to specify an initial condition, i.e., its phase space dimension is infinite. The Mackey–Glass equation,

$$\frac{d}{dt}x(t)=\frac{ax(t-\tau)}{1+x^c(t-\tau)}-bx(t), \tag{23}$$

is a prototypical example of such a system.[25,27] It is used as a model for the investigation of blood production. In this study we set $a=0.2$, $b=0.1$, $c=10$, and $\tau=17$. The integration step is $h=0.01$. Figure 5(a) shows a section of the time series. Figure 5(b) represents a phase portrait of the attractor. To estimate the correlation dimension and entropy one would usually have to embed the time series $x(t)$. Therefore, one has to determine the embedding dimension. The embedding dimension is not necessarily infinite. An embedding is a smooth map $f:X\to Y$ that is, a diffeomorphism from a smooth manifold $X$ to a smooth submanifold $Y$. The embedding dimension $d$ is then defined as the minimum dimension of a subset of Euclidean space into which a smooth manifold containing the attractor can be "embedded," i.e., $d$ variables are sufficient to uniquely specify a point on the attractor. As mentioned in the preceding section, the determination of the embedding dimension is problematic for observed time se-

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

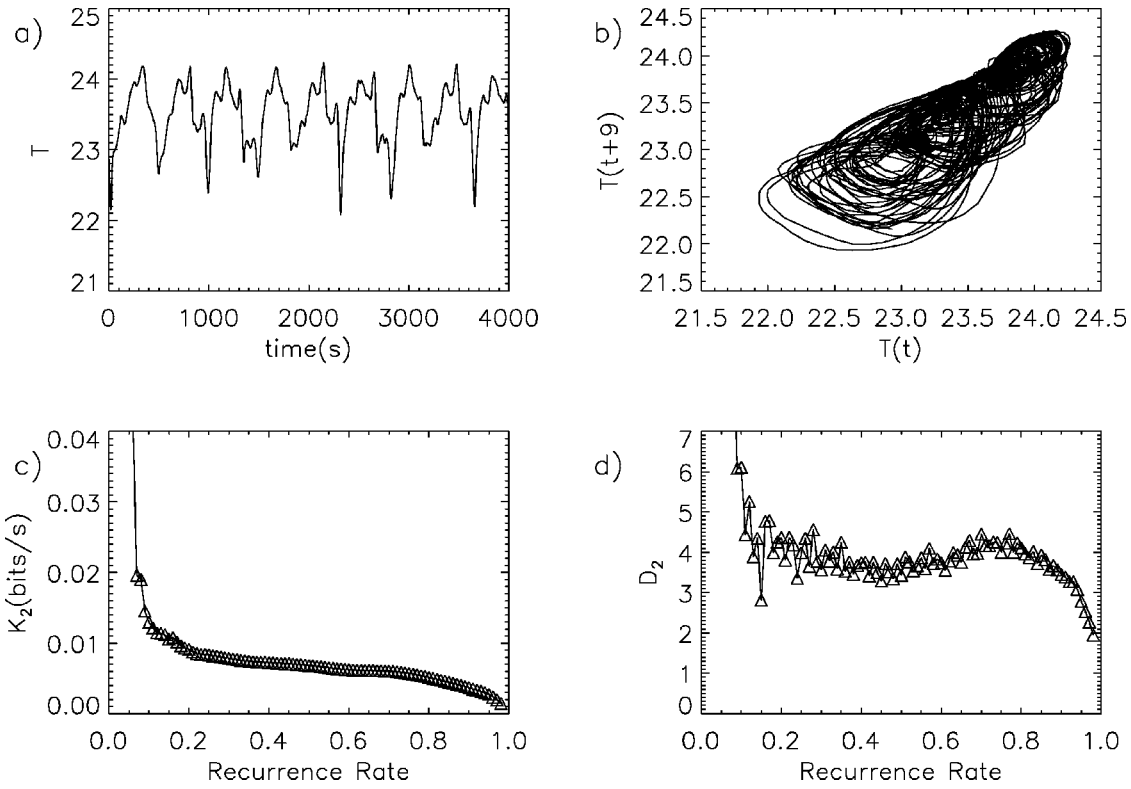


FIG. 6. (a) Sample of the time series for the flow in chaotic regime; (b) reconstructed attractor; (c) dependence of correlation entropy on the recurrence rate; (d) dependence of correlation dimension on the recurrence rate.

ries. However, since the RP based method for the estimation of $K_2$ and $D_2$ is independent of the embedding parameters, we do not use any embedding in this study.

Figure 5(c) shows the estimates of $K_2$ for different recurrence rates (we have obtained similar results for larger $\tau$, e.g., $\tau=30$). The plateau allows us to estimate the correlation entropy for the Mackey–Glass system to be $K_2=6.66\cdot 10^{-3}\pm 10^{-4}$ nats/s. This result is in good accordance with values reported in the literature for its Lyapunov exponents $[0.007,0,-0.071,-0.15,\ldots]$,[25] since the correlation entropy is numerically close to the sum of the positive Lyapunov exponents. The metric entropy for nonhyperbolic systems is approximately the sum of the positive Lyapunov exponents (slightly less) and the metric entropy is an upper bound for the correlation entropy.[26]

The RP based estimate for the correlation dimension is $D_2=2.13\pm 0.03$. Farmer computes $D_F=2.13\pm 0.03$ for the fractal dimension and $D_{KY}=2.10\pm 0.02$ for the Kaplan–Yorke dimension, again in very good accordance with our results.[27]

Our estimates were obtained from the scalar time series without any embedding. However, if we had used embedding we would have obtained the same estimates for all choices of the embedding parameters.

Note that we also computed successfully estimates of $K_2$ and $D_2$ for larger $\tau$, e.g., $\tau=30$ and $\tau=300$.

## IV. APPLICATION TO FLOW DATA

In this section we estimate invariants from the RP obtained from an analysis of fluid flow data. The experiment consists of a rotating, differentially-heated cylindrical annulus, in which a fluid (a water–glycerol mixture) is contained within the annular gap between two coaxial, brass cylinders and horizontal, thermally-insulating base and lid. The apparatus is rotated uniformly about its vertical axis of symmetry, and motion is driven by differential heating of the cylindrical sidewalls. Further details may be found in Ref. 29. The time series consisted of temperatures measured in the fluid at intervals of 2 s for periods of up to $8\times 10^4$ s, and were obtained from copper–constantan thermocouples on fine-wire probes located at mid-height and mid-radius in the convection chamber. The flows measured were in the baroclinically unstable regime, and took the form of azimuthally-propagating travelling waves with various quasiperiodic or chaotic modulations. The particular time series investigated here were taken from a single thermocouple probe for cases (ii) and (iii) of Ref. 29; case (iii) was identified as a quasiperiodic amplitude-modulated wavenumber $m=3$ flow, while case (ii) was identified as a low-dimensional chaotically-modulated wave with both $m=3$ and $m=2$ present.

Figure 6(a) shows a section of the time series from case (ii) and Fig. 6(b) the reconstructed attractor. To estimate $K_2$ we compute the distribution of diagonals $P_\varepsilon^c(l)$ and determine the slope for different thresholds $\varepsilon$ (respectively versus the recurrence rate) in the second scaling region. We also included the Theiler correction[30,31] based on the first zero of the autocorrelation function which was determined to be at $t=118$ s. Figure 6(c) shows the result. The plateau allows us to estimate $K_2$ reasonably unambiguously.

Based on Eq. (21) we then estimate $D_2$. Figure 6(d)

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

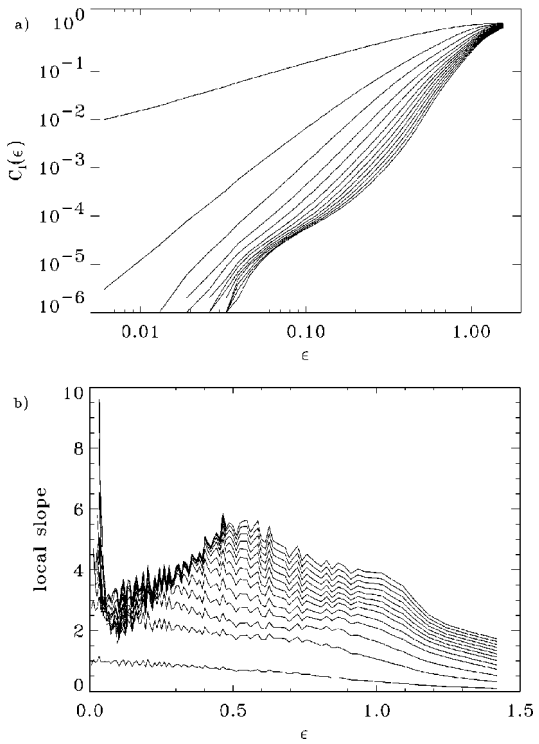


FIG. 7. (a) Correlation integral for the flow data. The estimation of the invariants is problematic in this representation as there is no clear straight line for large embedding dimensions. (b) Local slope of the curves represented in (a).

represents the outcome for different recurrence rates. We obtain for the correlation entropy $K_2 = 6.8 \cdot 10^{-3}$ bits/s $\pm 2.5 \cdot 10^{-4}$ bits/s and for the correlation dimension $D_2 = 3.6 \pm 0.4$. Our results are in accordance with previous results[29] and indicate low dimensional chaos. Note that our results were obtained from 40 000 data points without filtering and without embedding.

We have also performed the analogous computation based on the Grassberger–Procaccia algorithm (Fig. 7) and found that the CPU time needed is about one or two orders of magnitude higher. The values we have obtained for the correlation entropy are $K_2 = 1.1 \cdot 10^{-3} \pm 5.0 \cdot 10^{-4}$ bits/s and for the correlation dimension $D_2 = 4.1 \pm 0.4$.

The estimation of the correlation dimension and the entropy is rather problematic as the scaling is not very well pronounced for large embedding dimensions. The estimation based on the RP method is more robust.

In the quasiperiodic case (iii) we obtain for the correlation entropy $K_2 = 4.63 \cdot 10^{-4} \pm 2.4 \cdot 10^{-5}$ bits/s and for the correlation dimension $D_2 = 2.39 \pm 0.29$ also in very good agreement with previously reported results.[29]

## V. EFFECTS OF NOISE

In this section we discuss the effects of noise on the results of the analysis presented in this paper. A more detailed discussion can be found in Ref. 11.

Naively speaking, both observational and dynamical noise have two effects on the analysis. First, due to noise, two points that are neighoring can be separated, and second, two points that are not neighboring can become neighbors.

The first effect breaks long lines in smaller fractions and hence increases the slope of $-\ln(P(l))$ and therefore the entropy. Furthermore, it can be shown,[11] that the slope of $\ln(P(l))$ versus $l$ for a stochastic time series is not constant with respect to $\varepsilon$ resp. the recurrence rate. Hence, especially for small recurrence rates, the scaling region breaks down if $\varepsilon$ is smaller than about 5 times the standard deviation of the noise.[11] However, this effect is at least partially corrected by an appropriate choice of the scaling region in the ''$K_2$ versus recurrence rate'' diagram.

The second effect increases the number of short lines, mainly single points. This effect is—under normal conditions—nearly completely corrected when the scaling region in the $\ln(P(l))$ versus $l$ diagram is fixed, as short lines are ignored.

For practical purposes, a noise level of about 5%–10% can be treated without denoising the data. Further, Ref. 11 also allows for a mathematical correction if the noise level can be estimated.

Next, we exemplify the effects of noise on the analysis for the Rössler system in the funnel regime (see Fig. 4) which we have already studied in Sec. III A. We use the same parameters for the integration and add 10% of observational noise. Then, we perform the estimation based on the RP method. The results are summarized in Fig. 8. We find for the correlation entropy $K_2 = 0.130 \pm 0.050$, i.e., the value is about 16% higher than in the case without observational noise. This increase of the entropy is expected. Also the plateau breaks down for small recurrence rates. However, the scaling region is still reasonably broad.

For the estimation of the correlation dimension we obtain a rather small plateau. Despite of this problem, we obtain $D_2 = 2.27 \pm 0.28$ which is in a very good agreement with the result for the case without noise. The deviation is less than 1%. Note, that the error bars for both, the entropy and the dimension, are larger than the deviations due to the observational noise.

## VI. CONCLUSIONS

In this paper we have shown that it is possible to estimate $K_2$ and $D_2$ from RPs and that these estimates are independent of the choice of the embedding parameters. Even without an explicit embedding, it is possible to estimate these dynamical invariants directly. This fact is of considerable advantage, since means of making the optimal choice for the embedding parameters is still under debate. Hence, the method of RPs may be advantageous with respect to other methods, as presented in Ref. 20.

These results are also hinged to a train of arguments by Ivanski and Bradley on the one hand and Gao and Cai on the other. Ivanski and Bradley reported that some measures do not depend on the embedding.[32] Gao and Cai then stated that there the measures computed by Ivanski and Bradley are rather independent from the embedding, but that some structures in RPs are linked to the Lyapunov exponents and entropies, i.e., dynamical invariants. Then, in general, one has

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

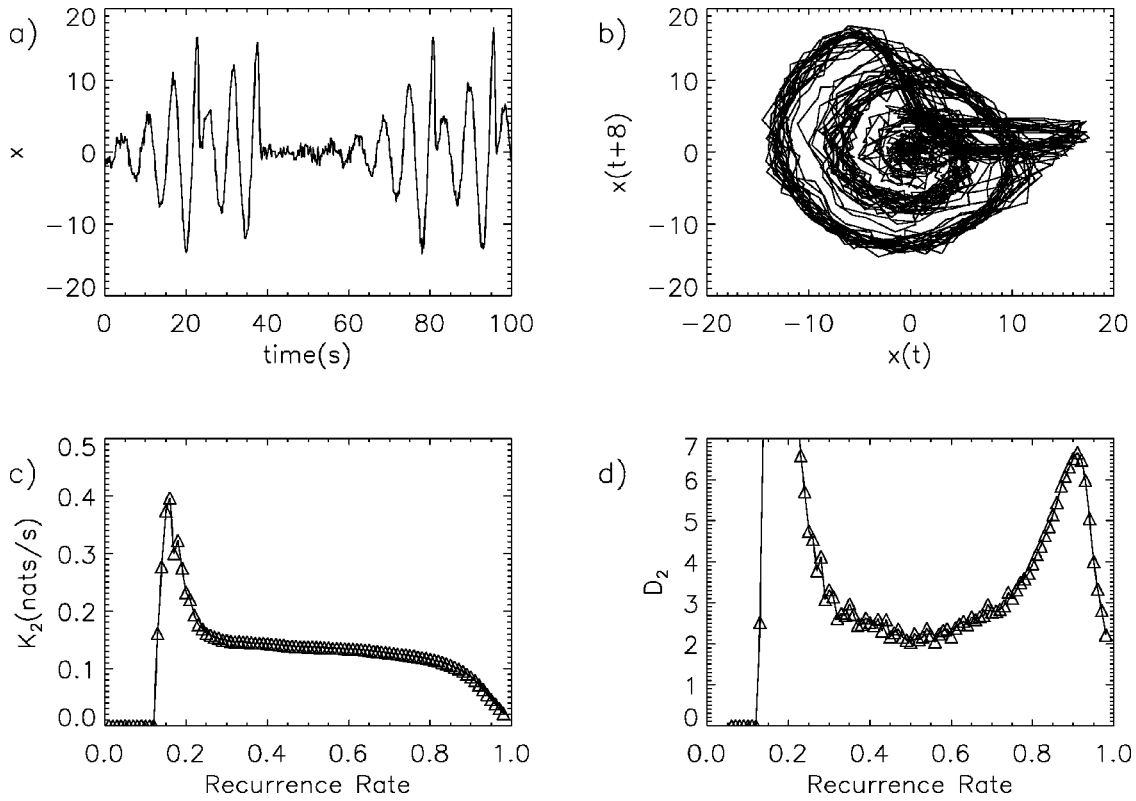


FIG. 8. (a) Sample of the time series of the Rössler attractor in the funnel regime ($a=0.25$, $b=0.4$, and $c=8.5$) with observational noise (10% of the standard deviation of the signal); (b) reconstructed attractor; (c) dependence of correlation entropy on the recurrence rate; (d) dependence of correlation dimension on the recurrence rate.

to embed the time series.[33] We think that both groups are in some sense right. However, we have shown analytically that the correlation entropy and the correlation dimension are independent from the embedding. Some other invariants, such as the mutual information, are not.

We also have estimated $K_2$ and $D_2$ for prototypical systems with different types of attractors, such as the chaotic and phase-coherent Rössler system with standard parameters, the Rössler funnel attractor, which is nonphase-coherent, and the Mackey–Glass system, which is infinite dimensional. For all systems the estimates are in very good agreement with results that are reported elsewhere in the literature.

For experimental flow data we were able to confirm the previous results that the time series comes from a low-dimensional chaotic system. The CPU time needed to obtain these results is by a factor 2 or 3 lower than that needed to apply the Grassberger–Procaccia algorithm. The benchmark was performed with the program D2 from the TISEAN package. We find numerically that the estimates are very robust, which may be due to the fact that the functions $\ln P(l)$ and $K_2(RR)$ which we consider have rather clear scaling regions. Furthermore, we exclude a first scaling region in $\ln P(l)$ for short lines, which improves the robustness of our estimates.

## ACKNOWLEDGMENTS

This work was supported by the ''DFG-Schwerpunkt 1114,'' the EU HPRN-CT-2000-00158, and the ''Internationales Promotionskolleg computational neuroscience of behavioral cognitive dynamics.''

[1] J.-P. Eckmann, S.O. Kamphorst, and D. Ruelle, Europhys. Lett. **4**, 973 (1987).
[2] J.P. Zbilut and C.L. Webber, Phys. Lett. A **171**, 199 (1992).
[3] C.L. Webber and J.P. Zbilut, J. Appl. Physiol. **76**, 965 (1994).
[4] N. Thomasson, C.L. Webber, and J.P. Zbilut, Int. J. Comp. Appl. **9**, 1 (2002).
[5] J.P. Zbilut, N. Thomasson, and C.L. Webber, Med. Eng. Phys. **24**, 53 (2002).
[6] J. Kurths, U. Schwarz, C.P. Sonett, and U. Parlitz, Nonlinear Processes in Geophysics **1**, 72 (1994).
[7] N. Marwan, M. Thiel, and N.R. Nowaczyk, Nonlinear Processes in Geophysics **9**, 325 (2002).
[8] N. Marwan and J. Kurths, Phys. Lett. A **302**, 299 (2002).
[9] M. Thiel, M.C. Romano, and J. Kurths, Jzv. VUZov—Applied Nonlinear Dynamics (in press).
[10] P. Faure and H. Korn, Physica D **122**, 265 (1998).
[11] M. Thiel, M.C. Romano, J. Kurths, R. Meucci, E. Allaria, and T. Arecchi, Physica D **171**, 138 (2002).
[12] F. Takens, *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics, edited by D.A. Rand and L.-S. Young (Springer, Berlin, 1980), Vol. 898.
[13] M.B. Kennel, R. Brown, and H.D.I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).
[14] D. Broomhead and K. King, Physica D **20**, 217 (1986).
[15] P. Grassberger, T. Schreiber, and C. Schaffrath, Int. J. Bifurcation Chaos Appl. Sci. Eng. **1**, 521 (1991).
[16] R. Gilmore and M. Lefranc, *The Topology of Chaos* (Wiley, New York, 2002).
[17] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[18] A. Rényi, *Probability Theory* (North–Holland, Amsterdam, 1970), Appendix.
[19] P. Grassberger, Phys. Lett. **97A**, 227 (1983).
[20] P. Grassberger and I. Procaccia, Physica D **13**, 34 (1984).
[21] P. Grassberger and I. Procaccia, Phys. Rev. A **28**, 2591 (1983).
[22] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, Cambridge, 1993).

Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp

[23] J.C. Sprott and G. Rowlands, Int. J. Bifurcation Chaos Appl. Sci. Eng. **11**, 1865 (2001).
[24] O. Rössler, Phys. Lett. **57A**, 397 (1976).
[25] J.D. Farmer, Physica D **4**, 366 (1982).
[26] C. Beck and F. Schlögl, *Thermodynamics of Chaotic Systems* (Cambridge University Press, Cambridge, 1993).
[27] M.C. Mackey and L. Glass, Science **197**, 287 (1977).
[28] R. Hegger, H. Kantz, and T. Schreiber, Chaos **9**, 413 (1999).
[29] P.L. Read, M.J. Bell, D. Johnson, and R.M. Small, J. Fluid Mech. **238**, 599 (1992).
[30] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[31] J. Kurths and H. Herzel, Physica D **25**, 165 (1987).
[32] J. Ivanski and E. Bradley, Chaos **8**, 861 (1998).
[33] J. Gao and H. Cai, Phys. Lett. A **270**, 75 (1999).

**Downloaded 25 Mar 2004 to 141.89.176.8. Redistribution subject to AIP license or copyright, see http://chaos.aip.org/chaos/copyright.jsp**