PRAMANA
—journal of
physics

© Printed in India

Vol. 49, No. 6,
December 1997
pp. 581–590

Detection of determinism and randomness in time series: A method based on phase space overlap of attractors

M S GOPINATHAN
Department of Chemistry, Indian Institute of Technology, Chennai 600 036, India
Email: msgopi@hotmail.com

MS received 19 March 1997; revised 25 August 1997

Abstract. The space overlap of an attractor reconstructed from a time series with a similarly reconstructed attractor from a random series is shown to be a sensitive measure of determinism. Results for the time series for Henon, Lorenz and Rössler systems as well as a linear stochastic signal and an experimental ECG signal are reported. The overlap increases with increasing levels of added noise, as shown in the case of Henon attractor. Further, the overlap is shown to decrease as noise is reduced in the case of the ECG signal when subjected to singular value decomposition. The scaling behaviour of the overlap with bin size affords a reliable estimate of the fractal dimension of the attractor even with limited data.

Keywords. Chaos; determinism; randomness; phase space overlap.

PACS No. 05.45

A basic problem in the analysis of a nonlinear dynamical system based on its measured time series is how to decide whether the dynamics is generated by a deterministic or a random process or by a combination of these two. The first step in the study of an experimental time series is usually the reconstruction of the attractor by the standard procedure using time delay vectors [1, 2] . Methods available to disentangle deterministic chaos from a stochastic process can be broadly divided into two categories: (a) Study of the temporal properties of the flows on the attractor. These include determination of Lyapunov exponents [3] measuring the divergence of trajectories, the average tangent of the flows on the attractor [4–6] and short term predictability of the time series [7]. (b) Methods based on the spatial structure of the attractor such as correlation dimension [8] and K_2 entropy [9]. Presence of even low levels of noise, which is unavoidable in a measured time series, can seriously affect the efficiency of these methods. For instance, determination of Lyapunov exponents and correlation dimension is unreliable [10–12] when noise level exceeds 1%.

Here we propose a simple geometric procedure to determine the overlap, η, in phase space of the reconstructed chaotic attractor under study and a surrogate random attractor reconstructed in the same manner, with identical reconstruction parameters, from a series of random numbers with a chosen distribution. It turns out that the magnitude of the overlap varies from $\eta \approx 1$ for a purely random process to $\eta \approx 0.1$ for deterministic chaos. η is found to be very sensitive especially to low levels of noise added to a chaotic time series. The basis for these results is that the extent of a chaotic attractor in phase space increases with additive noise and the attractor for a pure random series fills up the whole phase space.

In what follows, we first detail the procedure for the evaluation of the overlap η and then apply it to time series generated for standard chaotic systems like the Lorenz, Rössler and Henon systems [13] and to a linear stochastic signal [14]. An experimental ECG signal for a healthy human heart is also analysed. We then report the effect on η due to addition of different levels of noise to the Henon attractor as a test case. The scaling behaviour of η is also investigated and it is shown to give a fair estimate of the fractal dimension d_f of the attractor even with a limited number of data points. The sensitivity of η and d_f to noise is demonstrated.

The following procedure is used to determine the phase space overlap η of the attractor reconstructed from the time series with an attractor reconstructed similarly from a random series.

The time series $Y = (y_1, y_2, \ldots, y_M)$ of M data points is first linearly scaled so as to lie in a unit interval as

$$x_i = [(y_i - y_{\min})/(y_{\max} - y_{\min})] - 1/2. \tag{1}$$

Thus we have $-1/2 \leq x_i \leq 1/2$. This scaling fixes the volume of the state space in which the attractor is embedded to be unity, for convenience of comparison with other attractors. The attractor for this time series is then reconstructed by the method of delays [2] using delay vectors, $\mathbf{s}_i = \{x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(d-1)\tau}\}$, choosing appropriate values of reconstruction parameters namely delay time τ and embedding dimension d. This gives $N = M - (d-1)\tau$ points on the attractor. We call this attractor $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_N\}$.

The same reconstruction procedure with identical τ and d values is now applied to M computer generated random points of uniform distribution and a random attractor of N points is thus constructed within the same unit volume of phase space defined by the attractor $\mathbf{S}$. We designate the random attractor as $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\}$.

As an example, figure 1 shows the superposition of the reconstructed Rössler attractor and the corresponding random attractor for $N = 5000$ points, $\tau = 5$ and $d = 3$ as a typical case. The near coverage of the unit phase space by the random attractor and the limited region to which the Rössler attractor is confined may be noted.

We now define the overlap η of the two attractors $\mathbf{S}$ and $\mathbf{R}$ as

$$\eta(N, \varepsilon) = \frac{1}{N} \sum_{r_i \in R}^{N} [1 - f(g_{r_i})], \tag{2a}$$

where $f(z) = 0$ for $z \neq 0$ and $f(0) = 1$, and

$$g_{r_i} = \sum_{s_j \in S}^{N} \Theta(\varepsilon - |\mathbf{r}_i - \mathbf{s}_j|). \tag{2b}$$

Here Θ is the Heaviside step function. ε is the chosen bin size. The summation in (2a) goes over the N points of the random attractor $\mathbf{R}$ and that in (2b) goes over the N points of the attractor $\mathbf{S}$ under study. That η as defined is a measure of the overlap between two attractors in their common phase space can be seen as follows.

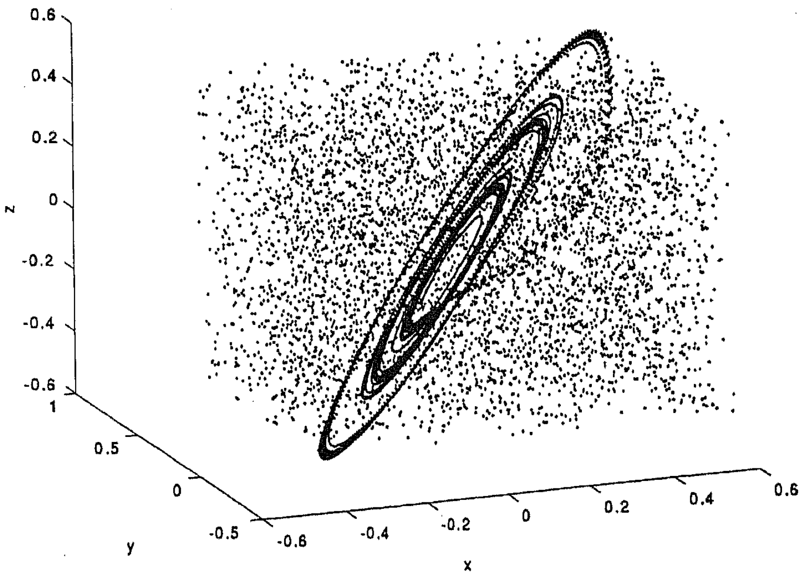

Figure 1. Rössler attractor reconstructed from x time series of the system superimposed on a random attractor of the same number of data points, $N = 5000$. x series is generated by 4th order Runge–Kutta integration of the Rössler system: $\{\dot{x}, \dot{y}, \dot{z} = -(y+z), x+ay, b+xz-cz\}$ with $a = 0.2$, $b = 0.2$ and $c = 5.7$, and the uniformly distributed random series is generated by the Matlab language command '*rand*'. Reconstruction parameters for both attractors are: dimension $d = 3$ and time delay $\tau = 5$.

Consider the two atrractors R and S to be identical. Thus for every point r_i, there is an identical point s_i so that $g_{r_i} > 0$ and hence $f(g_{r_i}) = 0$ for every r_i. Consequently, the sum in (2b) yields N and we therefore have $\eta = 1$ for this case of two identical attractors.

On the other hand, if the two attractors occupy non-overlapping regions of the unit phase space volume such that no point of R is within a distance ε of any point of S, we have $g_{r_i} = 0$ for every r_i. Therefore, $f(g_{r_i}) = 1$ for every r_i and we have $\eta = 0$ for this case. It is in this sense that we call η as the overlap between two attractors and its value can range from 1 to 0. η, as per eq. (2a) is the ratio of the number of points on the random attractor R that are within a distance ε of the points on the chaotic attractor S, to the total number of points N. Since the random attractor R is supposed to fill the whole of the unit phase space, η can be considered to be a measure of the spatial extent of the chaotic attractor S. This procedure is essentially a Monte Carlo [15] method of determining the volume of an object in a multidimensional space.

The overlap η should be distinguished from the correlation integral $C(\varepsilon)$ defined as [8]

$$C(\varepsilon) = \frac{1}{N^2} \sum_{i,j} \Theta(\varepsilon - |\mathbf{x}_i - \mathbf{x}_j|) \tag{3}$$

or the cross correlation sum [16]

$$C_{XY}(\varepsilon) = \frac{1}{|X|\,|Y|} \sum_{\mathbf{x}\in X} \sum_{\mathbf{y}\in Y} \Theta(\varepsilon - |\mathbf{x} - \mathbf{y}|). \tag{4}$$

These latter quantities represent the fraction of *pairs* of points on a given attractor or between two attractors that fall within a distance of ε.

Ideally $\eta(N, \varepsilon)$ should be studied as a function of N and ε. However we may make use of the assumption that the N random points uniformly cover the whole of the unit phase space volume so that we may fix ε for a given N as

$$\varepsilon^d = 1/N, \tag{5}$$

where d is the integer phase space dimension. From now on we denote η as $\eta(N)$ or as $\eta(\varepsilon)$ on account of the relation (5). This relation is prompted by the following reasoning. ε^d is the volume of a single bin and since we have scaled the phase space to unit volume, this should be given by $1/N_b$ where N_b is the number of bins. If N_b were chosen to be $\gg N$, corresponding to very small value of ε then $\eta \approx 0$ since the bin is so small that the probability of a given bin having two points, one each from the two attractors, would be very small. On the other hand, if $N_b \ll N$, corresponding to large ε, then $\eta \approx 1$ since every bin is likely to have a point from each of the two attractors. Hence setting $N_b = N$ is a reasonable choice, fixing the bin volume to the effective volume occupied by a point

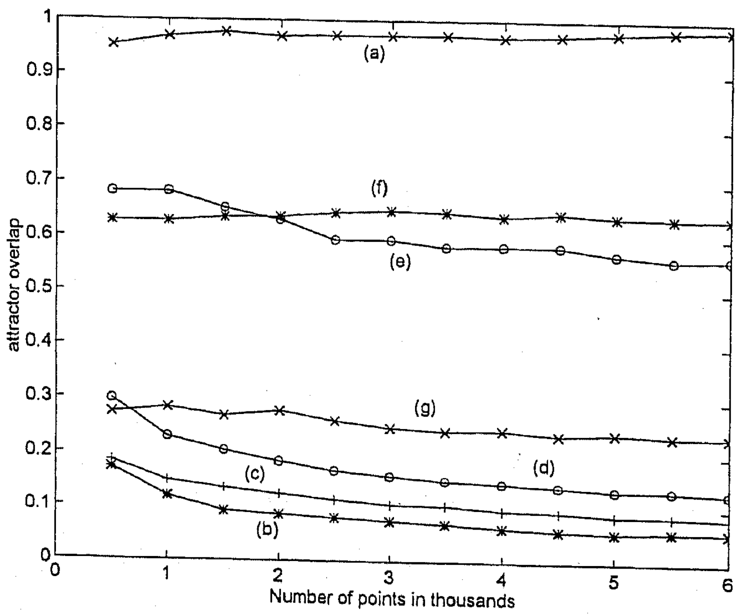

Figure 2. Overlap η (eq. 2) of reconstructed attractor with the reconstructed random attractor, as a function of data points used in reconstruction. For each curve, the same reconstruction parameters, namely phase space dimension d and time delay τ are used for the two attractors. **(a)** Another random attractor, $d = 3$, $\tau = 10$; **(b)** Rössler x series with parameters specified as in figure 1, $d = 3$, $\tau = 5$; **(c)** Lorenz x series of the system $\{\dot{x}, \dot{y}, \dot{z} = -\sigma x + \sigma y, rx - y - xz, -bz + xy\}$ with $\sigma = 10$, $b = 8/3$, $r = 28$, $d = 3$, $\tau = 5$; **(d)** Clean Henon map: $\{x_{k+1}, y_{k+1} = y_k + 1 - ax_k^2, bx_k\}$ with $a = 1.4$ and $b = 0.3$, $d = 2$, $\tau = 1$; **(e)** Stochastic time series of eq. (6), $d = 2$, $\tau = 5$; **(f)** x series of noisy Henon map, eq. (7), with $p = 50\%$, $d = 2$, $\tau = 5$; **(g)** Experimental ECG signal from a healthy human heart recorded at sampling frequency of 250 Hz with 12 bit resolution, $d = 2$, $\tau = 4$. In order to avoid minor fluctuations in η due to the random number generation algorithm, the random numbers are chosen from a very large set of random numbers (~ 15000) generated just once. The time delay τ used for the various attractors are rather arbitrary and no optimization of τ has been done in this work.

on the random attractor. This procedure is similar to the usual fixing of ε in the intermediate region where the logarithm of the correlation integral of (3) scales linearly with $\log(\varepsilon)$ [13].

$\eta(N)$ is plotted against N in figure 2 for a few dynamical systems, for N going from 500 to 6000 points. Curve (a) is for two independently realised random attractors in $3d$. It shows that $\eta(N) \approx 1$ for practically all values of N. Curves (b) and (c) are for the Rössler and Lorenz differential systems reconstructed in $3d$ and curve (d) is for the reconstructed discrete Henon attractor in $2d$. Remarkably, for all these attractors representing deterministic chaos, $\eta \approx 0.1$, for large N. Further, $\eta(N)$ levels off as $N \to 6000$ in all these cases. The dependence of η on N or equivalently on ε and its relation to the fractal dimension of the attractor is discussed later.

A time series generated by a linear stochastic process has been reported earlier [14] to correspond to a correlation dimension ≈ 2.5, a value normally considered indicative of low dimensional chaos. This series thus constitutes a counter-example to the trend of correlation dimension increasing with increasing embedding dimension for stochastic processes. The series is

$$\dot{x} = \theta x(t) + \omega(t), \tag{6}$$

where $\theta = -0.9$ and $\omega(t)$ is a standard Gaussian white noise process. Our estimate of η for the attractor for this series reconstructed in $2d$ is 0.56, clearly showing its stochastic origin (see curve (e) in figure 2).

Reported in figure (2), curve (f), is another noisy series generated by adding uniform random noise to the Henon map as

$$x'_n = x_n + (p/100)R_p l_x; \quad y'_n = y_n + (p/100)R_q l_y, \tag{7}$$

where (x_n, y_n) is the nth iterate of the clean Henon map and the primed coordinates denote the same with p per cent noise added. R_i is a random number with uniform distribution lying in the unit interval between -0.5 and $+0.5$ and l_x and l_y are the maximum extent of the clean Henon map in the x direction: $l_x = |x_{\max} - x_{\min}|$ and similarly for l_y. Figure 3, which is a superposition of the clean Henon attractor with one of 10% noise, clearly shows the spreading of the attractor with noise. Curve (f) in figure 2 is for the reconstructed Henon attractor with 50% noise and this gives an η value of 0.63 for large N, representing the high level of noise.

As a final example, the $\eta(N)$ values for an experimental time series, namely the ECG of a healthy human heart is displayed by curve (g) in figure 2. The large N value of η for this series is 0.23 which indicates deterministic chaos probably mixed with some noise. The presence of noise can be verified by applying one of the noise reduction techniques [10] to the time series and then determining η. Presently we applied the Broomhead and King's method [17] of singular value analysis to the ECG signal and reconstructed the attractor using the time series obtained by a projection of the trajectory matrix onto the most significant singular vector. This reconstructed attractor gives a considerably reduced η value of 0.16 indicating the presence of noise in the original data which corresponded to $\eta = 0.23$.

A study of η for the Henon map of (7), for a fixed value of $N = 5000$, at various values of the percentage of added noise, p, is reported in figure 4. The high sensitivity of η to noise, at low levels of noise is obvious. Note that η is proposed here as an index of

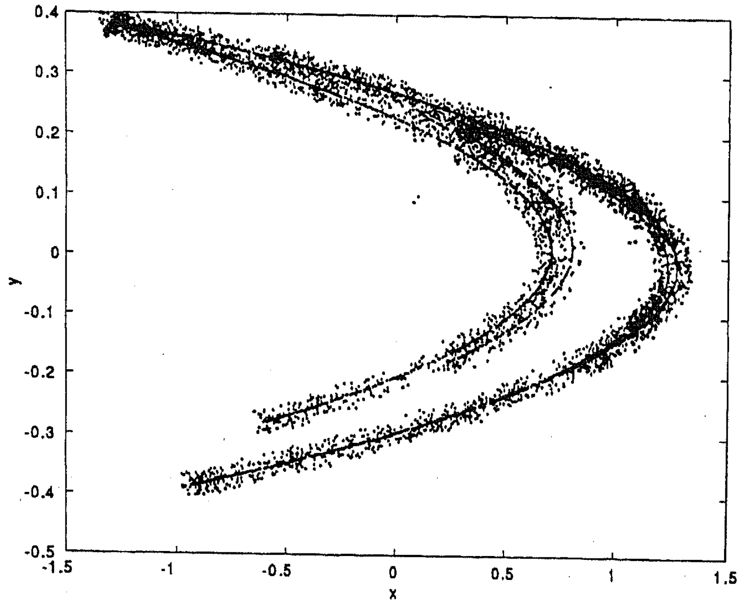

Figure 3. Clean Henon attractor as defined in figure 2, superimposed on another reconstructed noisy Henon attractor with $p = 50\%$ i.e. noise/signal ratio of 0.5. Number of data points = 5000, $d = 2$ and $\tau = 1$ for both attractors.

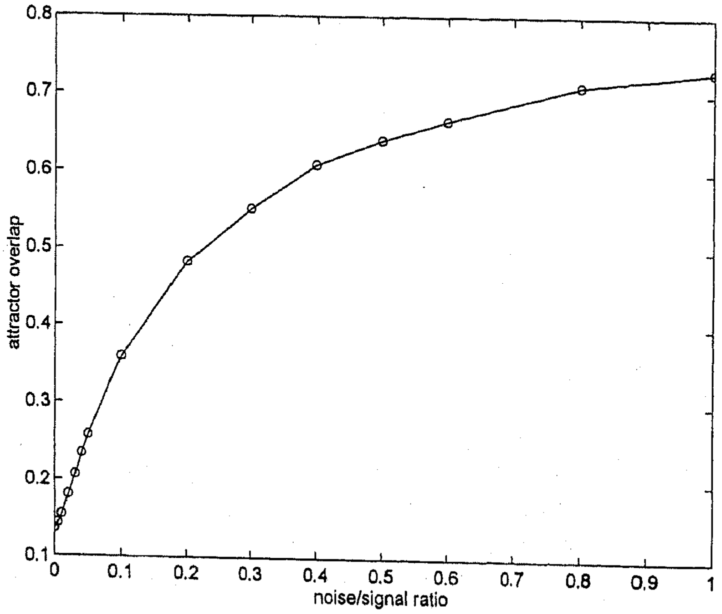

Figure 4. Overlap η of noisy Henon attractor with random attractor as a function of added noise. Noise to signal ratio corresponds to $p/100$ (eq. 7).

determinism in the time series and hence if noise is present in the observed signal, it should show up as increased value of η compared to that of a noise free signal. A major problem in the analysis of experimental signals is to discern if the observed nonlinearity is part of the dynamics or due to noise. Hence removal or addition of noise should alter the index of determinism (η in our case) sensitively, especially at low levels of noise which is the level normally present in a measured signal. At high levels of noise, η is anyway high, approaching 1, indicating the signal is dominantly noise. Figure 4 shows that η has this desirable property.

These numerical results indicate that the overlap η of a chaotic attractor with its surrogate random attractor in phase space can serve as one of the measures of stochasticity in a time series. η values of the order of ≈ 0.1 is indicative of deterministic chaos and significantly larger values point towards the presence of noise.

It is possible to determine, at least approximately, the fractal (box counting) dimension of the attractor from $\eta(\varepsilon)$ values determined as a function of ε. Denoting the number of points on the random attractor that fall on the chaotic attractor as N_{a}, we may set

$$N_{\mathrm{a}} = v/\varepsilon^{d_{\mathrm{f}}}, \tag{8}$$

where v is the constant volume occupied by the attractor in the total phase space of unit volume, after the transients have died down and d_{f} is the fractal dimension of the attractor. Combining (5) and (8) we have

$$\eta(\varepsilon) = \frac{N_{\mathrm{a}}}{N} \propto \varepsilon^{d-d_{\mathrm{f}}}. \tag{9}$$

Plots of $\log \eta$ vs. $\log \varepsilon$ for several attractors are given in figure 5. From the linear least squares fit (solid lines), it is clear that a scaling region exists for all these attractors, despite the different dimensions and time delay used, for moderate number of points in the range 500–6000. From the slope of these lines, $(d - d_{\mathrm{f}})$ and hence the fractal dimension d_{f} can be determined. This gives values of 2.04 ± 0.02 and 1.34 ± 0.01 for the Lorenz and Henon attractors for the chosen parameters, which compares well with the

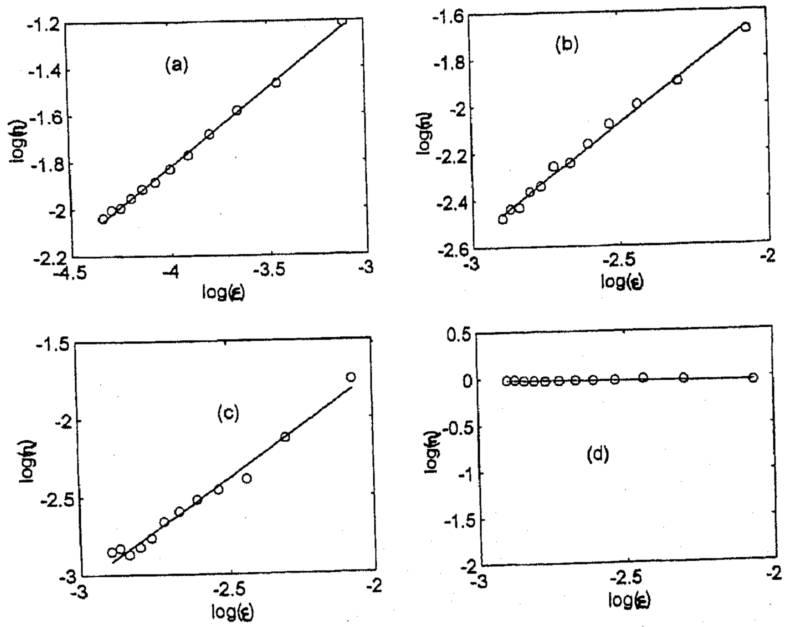

Figure 5. Log(η) vs. log(ε) for various attractors, for $N = 500$ to 6000 in steps of 500. **(a)** Henon attractor; **(b)** Lorenz attractor; **(c)** Rössler attractor; **(d)** random attractor. All time series and attractor reconstruction parameters are as defined in figure 2.

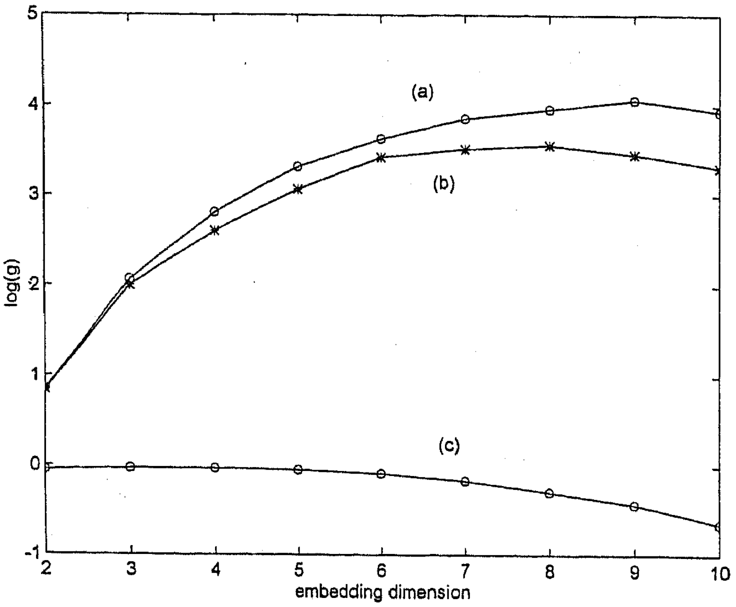

Figure 6. $\text{Log}(\eta/\varepsilon^{d-d_f})$ vs. embedding dimension d, (**a**) for Henon attractor with $N = 5000$; (**b**) Henon attractor with $N = 1000$; (**c**) random attractor with $N = 2000$. $d_f = 1.34$ for the Henon attractor and $d_f = d$ for the random attractor. Attractor and reconstruction parameters as in figure 2.

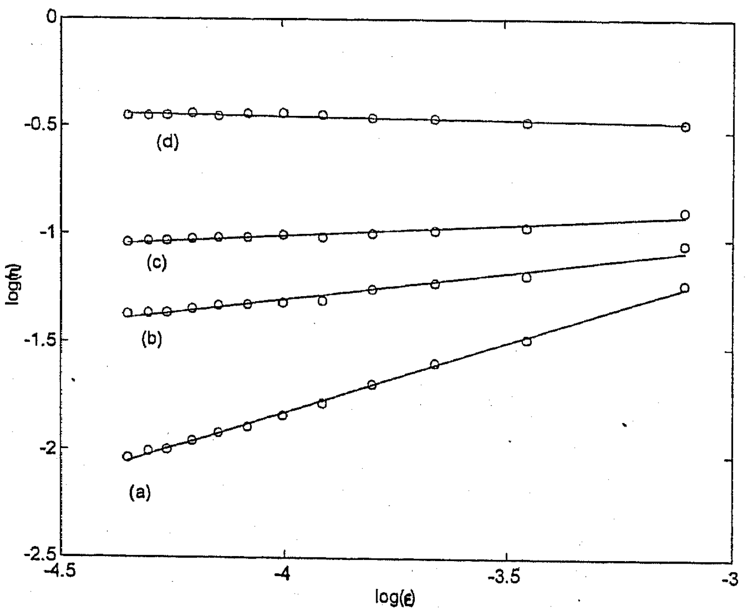

Figure 7. $\text{Log}(\eta)$ vs. $\log(\varepsilon)$ for Henon attractor with different noise levels, p, eq. (7). (**a**) Clean Henon, $p = 0\%$; (**b**) $p = 5\%$; (**c**) $p = 10\%$; (**d**) $p = 50\%$. Attractor and reconstruction parameters as in figure 2.

exact values 2.05 ± 0.01 and 1.21 ± 0.01 respectively [18]. For the Rössler attractor we get $d_f = 1.67 \pm 0.06$ for the parameters chosen here. For the random attractor, as expected, $(d - d_f) \approx 0$ showing that η is nearly independent of ε in this case (see figure 5(d)).

From (9), the ratio $g = \eta(\varepsilon)/\varepsilon^{d-d_f}$ should be a constant. A plot of $\log(g)$ as a function of the embedding dimension is shown in figure 6 for the Henon map, for fixed number of data points $N = 5000$ and $N = 1000$. Saturation of this ratio around $d \approx 6 - 7$ irrespective of N is clear. On the other hand, for a random attractor, as $d - d_f \approx 0$, g is nearly independent of the embedding dimension as seen from curve (c) in figure 7.

In order to study the effect of added noise on the attractor overlap, plots of $\log \eta$ vs. $\log \varepsilon$ for the Henon attractor with various levels of noise along with the best linear fit are displayed in figure 7. The systematic decrease in slope with increase in noise level is obvious. The slopes are 0.66, 0.26, 0.12 and −0.02 for levels of noise, $p = 0, 5, 10$ and 50% respectively corresponding to increasing fractal dimensions of 1.34, 1.74, 1.88 and 2.02 respectively.

It is also interesting to note that the quality of the linear fit as measured by the correlation coefficient deteriorates with increasing noise, these being 0.99, 0.98, 0.97 and −0.60 for the above p values. This demonstrates that the estimation of the dimension in the presence of very high noise levels of the order of 50% or more is unreliable.

In summary, we have proposed that the phase space overlap of reconstructed attractor with its random surrogate attractor can serve as an index of determinism. The low value of the order of ≈ 0.1 for the phase space overlap η of a deterministic chaotic attractor with its random attractor is found to be true for a variety attractors studied here, irrespective of the different phase space dimensions and time delay employed in the reconstruction of the attractors from time series. η is found to be extremely sensitive to noise, especially at low levels of noise. Reduction in η can also serve as a measure of the efficiency of noise-reduction technique applied to the time series. The existence of a scaling region for moderate number of points of a few thousands makes possible the economical determination of the fractal dimension of the attractor.

Further studies of the overlap method using surrogate data generated, for example by randomizing the phase of the Fourier transform of the time series [19], will be reported in future.

References

[1] F Takens, in *Dynamical Systems and Turbulence* of 'Springer Lecture Notes in Mathematics' edited by D A Rand and L-S Young (Springer-Verlag, New York, 1981) vol. 898, p. 366

[2] J-P Eckmann and D Ruelle, *Rev. Mod. Phys.* **57**, 617 (1985)

[3] A Wolf, J B Swift, H L Swinney and J A Vastano, *Physica* **D16**, 285 (1985)

[4] D T Kaplan and L Glass, *Phys. Rev. Lett.* **68**, 427 (1972)

[5] D T Kaplan and L Glass, *Physica* **D64**, 431 (1993)

[6] L W Salvino and R Cawley, *Phys. Rev. Lett.* **73**, 1091 (1994)

[7] D H Holton and R M May, in *The nature of Chaos* edited by T Mullin (Clarendon Press, Oxford, 1993) p. 149

[8] P Grassberger and I Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983)

[9] P Grassberger and I Procaccia, *Phys. Rev.* **A28**, 2591 (1983)

[10] E J Kostelich and T Schreiber, *Phys. Rev.* **E48**, 1752 (1993)

[11] T Schreiber, *Phys. Rev.* **E48**, R13 (1993)

[12] However, a recent report shows that for IID Gaussian noise, correlation dimension can be reliably estimated up to noise levels of 20%. See C Diks *Phys. Rev.* **E53**, R4263 (1996)

[13] S H Strogatz, *Nonlinear dynamics and chaos* (Addison-Wesley Publishing Co., Reading, 1994)

[14] A Provenzale, L A Smith, R Vio and G Murante, *Physica* **D58**, 31 (1992)
[15] W H Press, S A Teukolsky, W T Vetterling and B P Flannery, *Numerical recipes in Fortran* (Cambridge University Press, 1992) p. 295
[16] H Kantz, *Phys. Rev.* **E49**, 5091 (1994)
[17] D Broomhead and G P King, *Physica* **D20**, 217 (1986)
[18] P Grassberger and I Procaccia, *Physica* **D9**, 189 (1983)
[19] J Theiler, A Longlin, B Galdrikin and J D Farmer, *Physica* **D58**, 77 (1992)