

Noise reduction in chaotic time series data

A BHOWAL* and T K ROY†

*Institute for Theoretical Physics, University of Cologne, D-50923, Cologne, Germany

†Saha Institute of Nuclear Physics, 1/AF, Bidhan Nagar, Calcutta 700 064, India

Email: *ajanta@thp.Uni-Duisberg.De; †tarun@tnp.saha.ernet.in

MS received 30 July 1998

Abstract. With a prescription for an ‘effective noise’ in the given noisy chaotic time series and searching out ‘similar’ strings of data, we give a simple method of ‘cleaning’ the data by an iterative process converging with decreasing length of the string. This has been found efficient even for small amount of data.

Keywords. Chaotic time series data; noise reduction; Lyapounov exponent; correlation dimension.

PACS Nos 47.52; 05.45

1. Introduction

Even with the most sophisticated setup, noise is often encountered in time series data obtained from the experiments. Even if the recordings of the observation are accurate, there may be fluctuations due to the environment which are difficult to avoid. For example, the readings of the luminosity of a variable star cannot exclude the stray lights from the atmosphere or due to other objects. These fluctuations can be looked as noise, the removal of which may reveal a system with a few degrees of freedom. If the system under observation is in a steady state and the underlying dynamics is described by a low dimensional chaos, then noise will obscure this simple behaviour particularly when some of its frequencies coincide with those of the pure signal. This is a reason why for a noisy chaotic signal, filtering of frequencies to clean the data is not recommended because it damps out some parts of the signal.

For a chaotic signal one needs to know the different characteristics of the system, like the rate of divergence of nearby trajectories or the Lyapounov exponent, the embedding and correlation dimensions of the invariant set to which the data belongs (attractors in case of dissipative systems), etc., to get a preliminary idea of the system under observation. There exists a number of methods to determine these characteristics. For example, the local divergence plot of Gao and Zheng [1] produces efficiently the minimum embedding dimension and the Lyapounov exponent from a time series data at equal intervals of time. The method even distinguishes the noise in the series. The divergence rate of nearby trajectories becomes different when the distances in phase space are of the order of noise level. If

$$\vec{V}_i \equiv \{x_i, x_{i+L}, \dots, x_{i+(m-1)L}\} \quad (1)$$

are m -dimensional vectors constructed from the time series $\{x_i\}$ with a properly chosen time delay L , the initial difference of vectors \vec{V}_i and \vec{V}_j with

$$\Delta v_0 = |\vec{V}_i - \vec{V}_j| \leq \epsilon \quad (2)$$

where ϵ is small, grows exponentially in k (k large) time steps as

$$\Delta v_k = |\vec{V}_{i+k} - \vec{V}_{j+k}| \approx \lambda^k \Delta v_0 \quad (3)$$

Therefore the slope of $\log(\Delta v_k / \Delta v_0)$ vs. k gives an estimate of the rate of divergence λ . ϵ is small to ensure local linearity of the dynamics. But if it is of the order of noise, then the description of the dynamics fails and the divergence plot shows a different rate of growth. Above the noise level, the different curves follow an envelope both with the change of ϵ and the embedding dimension m of the vectors \vec{V}_i . This is illustrated in figure 1, where the time series is obtained from a Henon mapping [2], polluted by a 1% Gaussian noise. When ϵ is less than 2% of the extent of the data, the divergence curve gets detached from the envelope of curves obtained for different ϵ 's above 2%. For 1% noise the m -dimensional phase points within a radius of $2\sqrt{m}$ % become unreliable for the calculation of divergent rates.

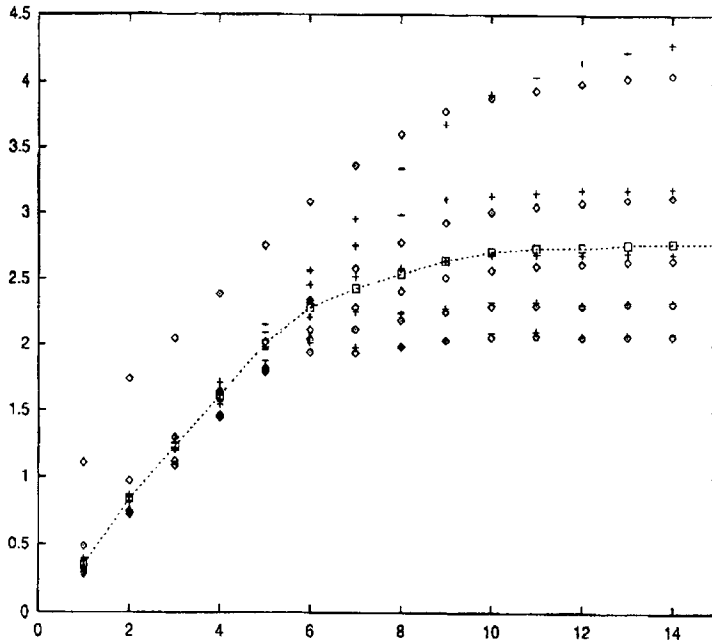


Figure 1. Divergence plot $\log(\Delta v_k / \Delta v_0)$ vs. k for Henon data, polluted by 1% Gaussian noise (\diamond) for different ranges of initial separation $\epsilon = \Delta v_0$, compared with that for pure data (+). The range for the topmost curve is $0 < \epsilon < 1\%$, the next $1\% < \epsilon < 2\%$, and so on, with that for the bottom most curve, $4\% < \epsilon < 5\%$. Also shown the curve for pure data with $\epsilon < 5\%$ (\square). Pure data normalised to $[0, 1]$.

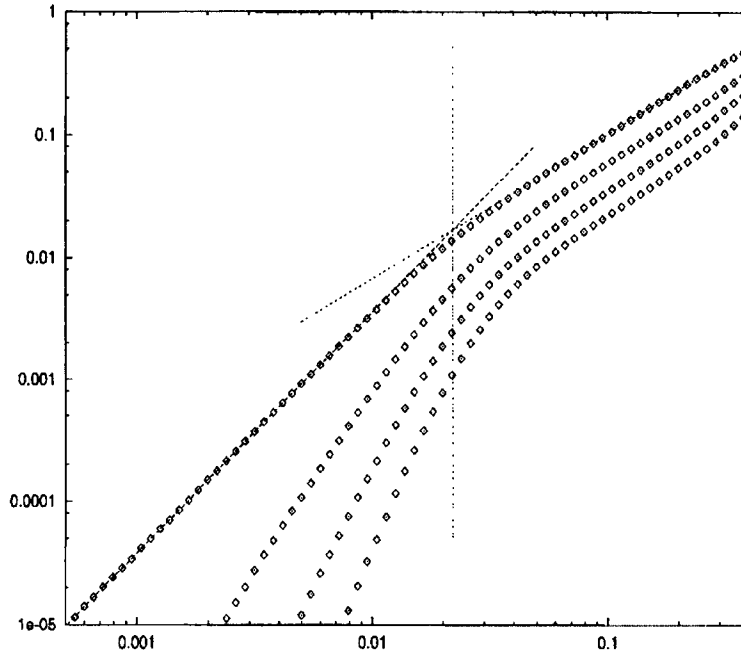


Figure 2. Plot of $N(\epsilon)$, the number of pairs of phase points vs. ϵ , the cell size in embedding dimensions 2 to 4 for the noisy data as in figure 1. The slopes above and below the noise level meet at around $\epsilon = 2\%$.

The correlation dimension is another popular measure to characterise chaotic data. The number of pairs of points in the phase space, $N(\epsilon)$ in a cell of dimension ϵ , is proportional to ϵ^ν , where ν is the correlation dimension, estimated from the slope of $\log P(\epsilon)$ vs. ϵ , $P(\epsilon)$ being the probability of observing a pair of phase points within a distance ϵ :

$$\nu = \lim_{\epsilon \rightarrow 0} \frac{\log P(\epsilon)}{\log \epsilon} \quad (4)$$

There are standard methods [3] to determine this. The estimate of ν is correct for ϵ at lower scales only, but with noise it becomes unreliable. The $\log N(\epsilon)$ vs. $\log \epsilon$ plot (figure 2) becomes different at the noise level and the slope tends to the dimension of the phase space, indicating that the distribution is homogeneous at these scales. One can just read out the noise level from this plot [4]. Kostelich and Yorke [5] point out that for a noise level of 1% in the time series it is impossible to determine the correlation dimension for distances below 3% of the extent of data. We shall see later, that this characteristic can be used to estimate what we term as the ‘effective noise’ present in the data.

Therefore it is necessary that the data be processed such that the noise is reduced with the nature of the signal being preserved as far as possible. Although the different methods for noise reduction, as pointed out by Kostelich and Schreiber [6], are similar in nature still one has to look for simpler and easy to implement methods and which are convergent. One such method as proposed by Schreiber and Grassberger [7], which takes into account the ‘past’ and ‘future’ values from a point in the phase space is especially attractive due to its simplicity and efficiency compared to other methods. In this paper we present a similar

convergent process to clean the data once an 'effective' noise in the data is estimated. Our method is comparable to that of Schreiber and Grassberger, but with input parameters that can be determined from the given noisy series. It gives a relatively good 'cleanliness', especially for small volume of data.

In the following, in §2 we define 'similar sequences' that are parts of trajectories and which will be required in the method of noise reduction described in §3. Section 4 presents the results and also establishes the convergence of the procedure. A concluding summary is given in §5.

2. Local divergence and similar sequences

The system under observation is supposed to follow the dynamics

$$\dot{\vec{V}} = \vec{F}(\vec{V}) \quad (5)$$

where \vec{F} is vectorial and $\vec{V} \equiv \{x, y, z, \dots\}$, of which a particular component, say x , is recorded at equal time intervals resulting in the series x_1, x_2, x_3, \dots . With a proper choice of the phase space dimension M , the above dynamics can be described by a mapping in M -dimension:

$$\vec{V}_{i+1} = \vec{f}(\vec{V}_i) \quad (6)$$

where \vec{V}_i is the sequence $\{x_{i-M+1}, x_{i-M+2}, \dots, x_{i-1}, x_i\}$ and also a part of the trajectory, and denotes a vector embedded in M -dimension. Gao *et al* [1] determines the minimum embedding dimension by constructing vectors with a time delay L , for which the components are $\{x_{i-(m-1)L}, x_{i-(m-2)L}, \dots, x_{i-L}, x_i\}$, and calculating the average divergence rate, which does not change after a certain m .

For a deterministic system, the initial condition $\{x_1, x_2, \dots, x_M\}$, specifies the whole of the trajectory as well as any V_i , since both the past and future are determined by the dynamics. Therefore, the sequence V_i can be termed as a signature of the trajectory. This signature has a minimum length, in this case equal to M , the number of initial values. A signature of length less than M will not produce the dynamics.

Alternative to eq. (2), the dynamics can also be represented by a M -th order difference mapping, where the trajectory is determined by the initial values $\{x_1, x_2, \dots, x_M\}$:

$$x_{i+1} = f(x_{i-M+1}, x_{i-M+2}, \dots, x_i). \quad (7)$$

The deviation Δx_0 at a point x_i in the time series will in general result in a growth of deviation at x_{i+k} ,

$$\Delta x_k \approx \lambda^k \Delta x_0 \quad (8)$$

where λ is the local rate of divergence depending on $\{x_{i-M+1}, x_{i-M+2}, \dots, x_i\}$. When the dynamics is described in terms of past and future values, λ depends on $\{x_{i-D}, x_{i-D+1}, \dots, x_i, \dots, x_{i+D-1}, x_{i+D}\}$, where $D = M/2$, since the trajectory is deterministic.

In an experiment, the time series polluted by noise will have distorted signatures, the distortion increasing with the noise amplitude. The discussions following will be for a class of noise, with zero mean and having equal probabilities of positive and negative values of

noise (symmetric with respect to mean). It also belongs to the type 'measurement' noise or fluctuations that do not affect the system. The common occurrence in the laboratories is that with a Gaussian distribution, but in our numerical calculations an illustration will also be given for a flat distribution over twice the noise amplitude.

A distorted signature actually represents a bundle of trajectories denoting different states of the system, in which one of them is a pure one. If the length of the signature is increased then some of these trajectories will be rejected. The longer the length the more closely we are able to identify the pure trajectory. The accuracy increases with the length of the signature but suffers with increasing noise for a given length. In a noise reduction scheme, it is therefore advantageous to use signatures of 'similar strings' (see below) with their lengths as large as possible, but then their number will be limited by the amount of data available. Use of smaller signatures, which are less accurate, but more available, requires a pre-processing. We shall see how this is taken into account in our noise reduction scheme.

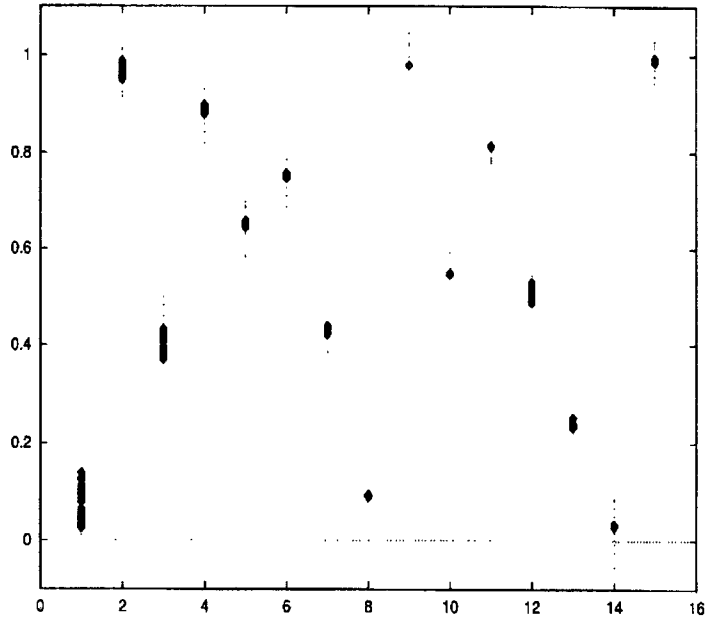
In a phase space of dimension $m = 2k + 1$ (say), if the system is chaotic, two nearby points V_i and V_j separate from each other in either or both of the 'future' or 'past' directions according to the local rate of divergence. Therefore, if there are two strings of length k , with their initial ends Δx_0 apart, the other ends will diverge to a separation of Δx_k . Thus, if two equal parts of a trajectory lie within some separation, which we denote by 'similar sequences', then it is most likely that the middle parts are close to each other, since towards either or both of the ends the deviations increase. This holds true even if all the eigenvalues (of separation growth) are nearly equal to 1, in which case a longer part in the middle has very little variation. An illustration in figure 3 shows that for equal length sequences of a time series from a Henon mapping, within a given separation, the middle of the sequences are closer the longer similar sequences are.

Therefore, in a noisy time series, if different equal sequences of length $2k + 1$, with components $\{x_{-k}, x_{-k+1}, \dots, x_0, \dots, x_{k-1}, x_k\}$, searched from the data, lie within some separation from a reference string, which we need to modify, then the components x_{-r} to x_r where r depends on the length of the sequences, lie around the actual values of the pure string, the extent of variation being the noise amplitude plus a quantity decreasing with the length of the sequence (from eq. (8)). Averaging will reduce the variation due to noise, the reduction being proportional to the number of sequences available.

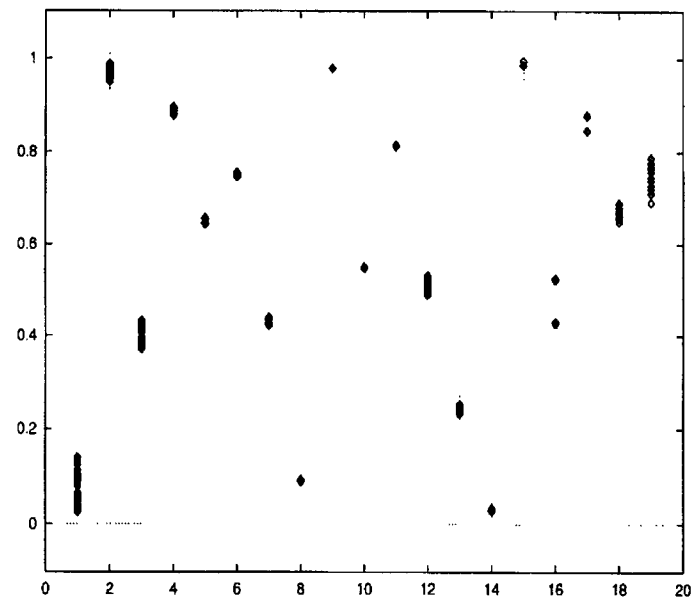
3. Reconstruction from noisy data

To reconstruct the data, the standard practice [6] is to find a number of similar sequences from the data, where by similarity is meant that they lie within a separation δ , depending on the noise amplitude, and estimate the midcomponent by averaging or the least square method. The result is again treated like a raw data and the process iterated, where this time δ is taken as the root mean squared deviation of the modified series from the previous series.

The above procedure has the disadvantage that after the very first iteration some 'information' is lost, especially when the data is short which gives rise to a distortion. Rather, the original noisy time series is expected to have more information. Hence it is better that comparison be made each time with the original time series. Further, averaging tends to



(a)



(b)

Figure 3. Similar strings of data, of (a) length 15, from pure (\diamond) Henon time series (same as of figure 1) and that with 5% Gaussian noise (\bullet), within a separation of $\delta = 10\%$ and (b) same as in (a) for string length 19. Note how the points in the middle of pure data strings (\diamond) are spread by noise (\bullet).

shift the results towards higher phase space densities because of overlap of phase points due to noise. This effect is more prominent for low k values. Besides, although the standard method is convergent, there is no reason to take the subsequent δ 's as the root mean squared deviation from the previous series, where instead it should depend on the noise in the modified series.

The method we propose avoids these problems in an iterative process. The results of successive iterations do not depend on the phase space densities. Instead of the processed data from an iteration to be treated like raw data, we always compare the modified data with the original noisy series and process further.

Before we start cleaning we estimate some of the parameters that will be required in the process. These are

i) the separation δ within which two strings of data of equal length are called similar. This can be determined from the divergence plot, where the divergence rate curve just separates from the envelope of correct curves. For 1% Gaussian noise, the error in the divergence plot is seen as ϵ , the maximum initial separation of vectors $|\vec{V}_i - \vec{V}_j|$ is less than 2%. Also, from the correlation dimension plot we can prescribe a δ . At the noise level the correlation dimension tends to m , the dimension of the phase space itself and the slope is a constant for smaller values of ϵ . Well above the noise level the curve shows the true correlation dimension and is a straight line here also. The point where these two straight lines meet gives a measure of the 'effective noise' amount which we prescribe for δ . This is also seen to be consistent as with that obtained from the divergence plot. For a Gaussian noise δ is about twice the width of the distribution while for uniformly distributed noise it is almost same as the amplitude. We have tested that a small error in the estimation does not change the results for reconstruction significantly.

ii) the value of k_{\min} up to which the iteration process will continue. From the two curves we discussed above one can determine the minimum embedding dimension M . Then

$$k_{\min} = M/2. \quad (9)$$

The method we suggest does not depend on the time delay L . Although Takens' theorem [4,8] points out the importance of a time delay for the determination of the characteristics, it turns out that for reconstruction it is the local rate of divergence and the similarity of the sequences available in the data that helps in noise reduction.

Denoting the cleaned data from an iteration by the y 's and the original time series (raw data) by x 's, we proceed as follows:

- i) Initialise y with x . Start with a high value of k so that there is very little modification after the first iteration. In case of oversampled data, it is sufficient if the initial $k \geq k_{\min} + 10$.
- ii) From the sequences $\{x_{-k}, x_{-k+1}, \dots, x_{k-1}, x_k\}$ in the raw data similar to $\{y_{-k}, y_{-k+1}, \dots, y_{k-1}, y_k\}$, replace the midstring $\{y_{-r}, y_{-r+1}, \dots, y_{r-1}, y_r\}$, where $r = \text{integer}(k/4)$, by the mean of maximum and minimum of the variations. If say, z_0 is the pure datum corresponding to y_0 , then the different x_0 's available from the searched sequences are spread around z_0 symmetrically.
- iii) Calculate the root mean squared deviation E_{xy} , of the y 's from the x 's:

$$E_{xy} = \sqrt{\frac{1}{N} \sum (x_i - y_i)^2} \quad (10)$$

- iv) If E_{xy} changes very little, decrease k by 1 until $k = k_{\min}$, before going to step (ii).

v) If the change in E_{xy} is small compared to that obtained by previous k or if for some k with converging E_{xy} , the minimum number n_{\min} of similar sequences is at least a certain preassigned value (say, n_{\min} is 100), go to step (vi), otherwise step (ii).

vi) When E_{xy} converges, find the centroid with $k = k_{\min}$ and then $k = k_{\min} + 1$ (to remove false points due to noise), i.e., from the sequences $\{x_{-k}, x_{-k+1}, \dots, x_{k-1}, x_k\}$ in the raw data similar to $\{y_{-k}, y_{-k+1}, \dots, y_{k-1}, y_k\}$, replace the midstring $\{y_{-r}, y_{-r+1}, \dots, y_{r-1}, y_r\}$, where $r = \text{integer}(k/4)$, by the average of the variations of $\{x_{-r}, x_{-r+1}, \dots, x_{r-1}, x_r\}$ (similar to the 'constant fit' method as in [6]). If the minimum number of matching strings n_{\min} is high one can also calculate the centroid at a higher k .

vii) If $E_{xy} < \sigma_{\text{noise}}$ and $n_{\min} > 1$, where σ is the standard deviation of the noise ($\delta/2$), repeat steps i) to vi) replacing x by y and δ by $\delta/\sqrt{n_{\min}}$, because now the standard deviation of the error is expected to be reduced by $\sqrt{n_{\min}}$.

Due to the second step in the above procedure the points do not move towards higher phase space densities. The retrieved data keeps on improving as compared with a known case (a set of noisy data from Henon mapping) to be presented below.

As used by Schreiber and Grassberger [7], the 'cleanliness' of the processed data, of size N , is usually expressed by the quantities r_0 and r_{dyn} defined below:

$$r_0 = E_{xz}/E_{yz}, \tag{11}$$

where E_{xz} and E_{yz} are the root mean squared deviations of the noisy (x) and the cleaned (y) data from the true values (z), viz.

$$E_{xz} = \sqrt{\frac{1}{N} \sum (x_i - z_i)^2} \tag{11a}$$

and

$$E_{yz} = \sqrt{\frac{1}{N} \sum (y_i - z_i)^2}. \tag{11b}$$

$$r_{\text{dyn}} = E_{xf}/E_{yf} \tag{12}$$

where E_{xf} and E_{yf} are the estimates of the errors due to functional representations, viz.

$$E_{xf} = \sqrt{\frac{1}{N} \sum [x_i - f(x_{i-1}, x_{i-2}, \dots)]^2} \tag{12a}$$

and

$$E_{yf} = \sqrt{\frac{1}{N} \sum [y_i - f(y_{i-1}, y_{i-2}, \dots)]^2}. \tag{12b}$$

r_{dyn} can be calculated only if the evolution function f is known. We present these values for the Henon mapping, given by

$$x_{i+1} = 1 - ax_i^2 + bx_{i-1} \tag{13}$$

with $a = 1.4$ and $b = 0.3$, which gives rise to an attractor.

Tables 1 and 2 show the continuous improvement of data with iterations and the convergence with decreasing k for two types of noise both with symmetric distribution. In table 3 we see how the ratios r_0 and r_{dyn} become better with the available number of data. This is

Table 1. Results for Henon data, $N = 64000$ with 1% Gaussian noise.

Iterations	k	n_{\min}	δ	E_{xy}	r_{dyn}	r_0
1	9	1	0.02000	0.00148	1.0117	1.0108
10	8	1	0.02000	0.00248	1.0332	1.0304
20	6	1	0.02000	0.00525	1.1884	1.1722
30	5	1	0.02000	0.00741	1.5041	1.4510
40	4	1	0.02000	0.00872	2.2184	2.0355
50	1	2	0.02000	0.00918	3.2793	2.7384
51	1	28	0.02000	0.00960	6.8688	3.4241
52	2	14	0.02000	0.00987	9.2523	3.9815
1	9	1	0.00535	0.00988	9.2740	3.9936
2	9	1	0.00535	0.00988	9.2740	3.9931
3	8	1	0.00535	0.00989	9.3022	4.0050
4	7	1	0.00535	0.00991	9.3393	4.0209
5	7	1	0.00535	0.00991	9.3411	4.0212
6	6	1	0.00535	0.00993	9.3952	4.0555
7	6	1	0.00535	0.00994	9.4075	4.0565
8	5	1	0.00535	0.00997	9.5316	4.1066
9	5	1	0.00535	0.00997	9.5347	4.1118
10	4	1	0.00535	0.01001	9.7874	4.1718

Table 2. Results for Henon data, $N = 16000$ with 10% uniform noise, ($\sigma = 0.058$).

Iterations	k	n_{\min}	δ	E_{xy}	r_{dyn}	r_0
1	10	1	0.11000	0.00461	1.0039	1.0030
11	9	1	0.11000	0.00820	1.0113	1.0093
18	8	1	0.11000	0.01537	1.0375	1.0330
25	7	1	0.11000	0.02948	1.1690	1.1439
32	6	1	0.11000	0.04662	1.7252	1.5248
39	5	1	0.11000	0.05465	3.1874	2.2119
44	4	5	0.11000	0.05760	7.4278	2.8715
47	3	16	0.11000	0.05777	8.1396	2.9072
48	2	99	0.11000	0.05779	8.3464	2.9081
49	1	220	0.11000	0.05780	8.3687	2.9079
50	1	214	0.11000	0.05824	9.5793	2.8503
51	2	98	0.11000	0.05871	11.2091	2.8542

also evident from figure 4 for a series with 64000 data from the Henon mapping polluted by 1% Gaussian noise. The finer structures of the attractor are retrieved implying that for such a noise this amount of data is sufficient for reconstruction. Because the mapping function is known for this case it was possible to know r_{dyn} also. A comparison of the ratios obtained by the standard [7] methods with that of the present shows that with half the amount of data as in [7] (where in most cases k is not mentioned) the ratios obtained by this simple method, are comparable to and sometimes better than those obtained with a least square fit. It is therefore obvious that still better ratios could be achieved by applying the least square method as in [7]. In table 4 we present the results for a noisy Lorenz system [9] with its data oversampled and input *time delay* $L = 1$. With oversampling there is definitely an advantage in reconstruction. The convergence was obtained earlier at $k = 3$. In this case note that since the functional dependence of the future values is not known in terms of the past and present we could calculate r_0 only.

Table 3. Results for different N 's for Henon data, with Gaussian noise.

%noise	N	k	r_{dyn}	r_0	r_{dyn} [7]	r_0 [7]
5	500				1.7	1.4
5	500	2	2.52	1.71		
5	1000				2.4	1.8
5	1000	2	3.79	1.88		
5	20000				7.8	3.1
5	20000	2	10.19	2.67		
5	10000	2	8.45	2.65		
5	64000				9.9	3.3
5	64000	2	11.80	2.69		
5	32000	2	10.63	2.70		
1	20000				3.8	2.4
1	20000	2	6.14	3.12		
1	10000	2	3.89	2.44		
1	64000				6.4	3.0
1	64000	2			9.65*	3.55*
1	64000	3			8.77*	4.05*
1	64000	2	9.79	4.17		
1	32000	2	7.41	3.63		

* with least square fit

Table 4. Results for oversampled Lorenz data, with 5% Gaussian noise
 $\dot{x} = -\sigma(y - x)$, $\dot{y} = x(r - z) - y$, $\dot{z} = xy - bz$,
 $\sigma = 10$, $b = 8/3$, $r = 45.92$, $\delta t = 2^{-5}$, $N = 32000$, $L = 1$.

Iterations	k	n_{min}	δ	E_{xy}	r_0
1	17	1	0.10000	0.01934	1.0689
14	16	1	0.10000	0.03420	1.2739
21	15	1	0.10000	0.03940	1.4294
28	14	1	0.10000	0.04356	1.6621
35	13	1	0.10000	0.04604	1.8793
40	12	1	0.10000	0.04760	2.0939
44	11	1	0.10000	0.04808	2.2027
47	10	1	0.10000	0.04838	2.2672
49	9	1	0.10000	0.04847	2.2926
52	8	1	0.10000	0.04865	2.3521
54	7	1	0.10000	0.04874	2.3684
56	6	2	0.10000	0.04882	2.4005
57	5	28	0.10000	0.04883	2.4010
58	4	64	0.10000	0.04883	2.4040
59	3	120	0.10000	0.04884	2.4069
60	3	175	0.10000	0.05079	2.5258

Having reduced the noise we now see how the characteristics of the system are dependent on the cleaned data. Figure 5 shows the divergence rate obtained from the cleaned Henon data and figure 6 the correlation dimension plot. Although the tables presented above show a reduction of noise by approximately 3 times (given by r_0 when compared point to point), it is seen from the above figures that the effective reduction seems to be more than indicated.

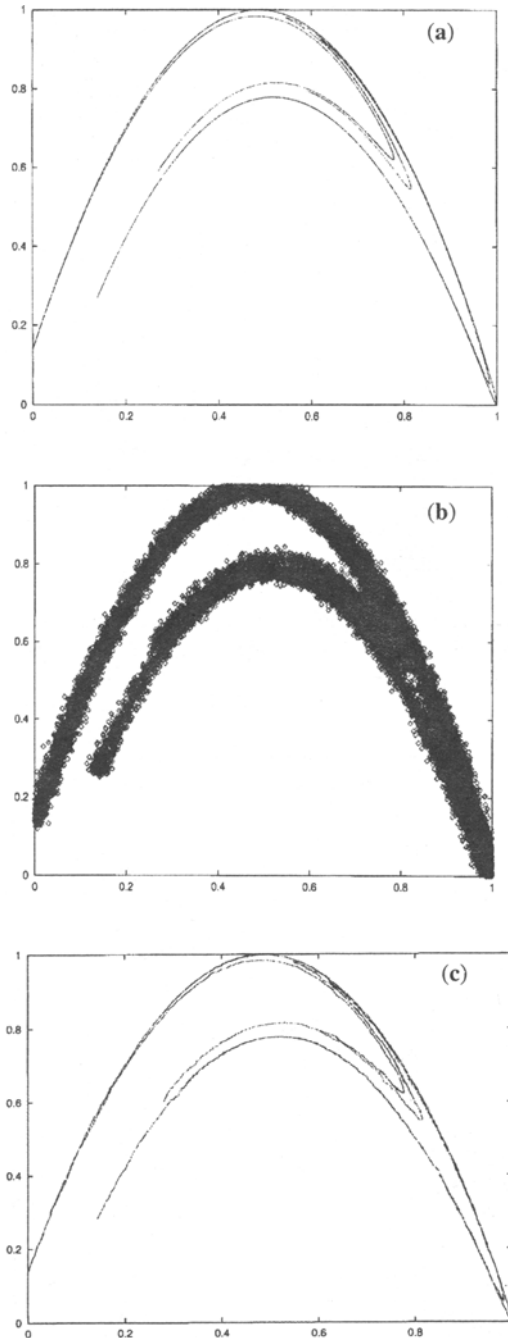


Figure 4. (a) Henon data from $x_{i+1} = 1 - ax_i^2 + bx_{i-1}$ with $a = 1.4$ and $b = 0.3$, (b) the same with 1% Gaussian noise (for comparison points with noise outside the range $[0,1]$ are not shown) and (c) the reconstruction from 64000 data.

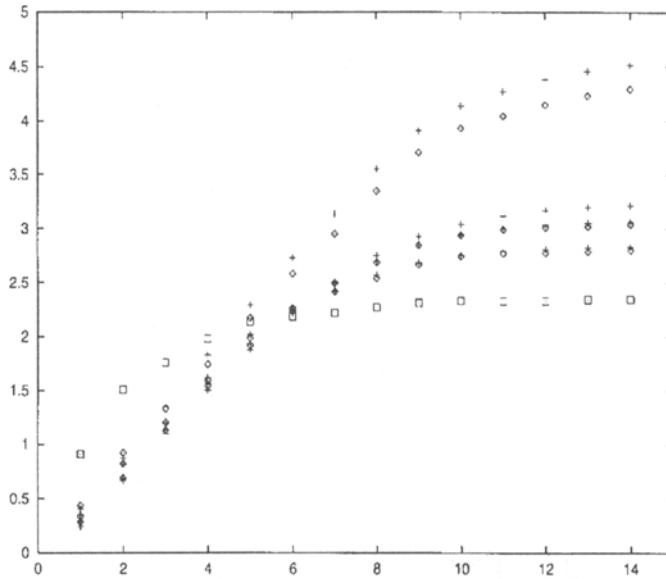


Figure 5. Divergence plot as in figure 1 for reconstructed data (+) compared with that for pure (\diamond) and noisy data (\square) for a time series of 16000 Henon data with 5% Gaussian noise.

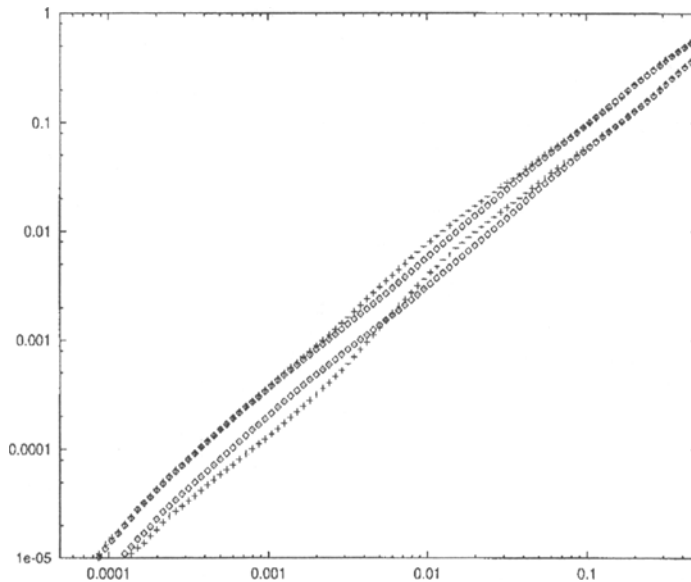


Figure 6. A plot of $\log P(\epsilon)$, the probability of finding a pair of phase points within a distance ϵ , vs. $\log \epsilon$ for the same reconstructed data (dashed curves) as used in figure 5, for $m = 2$ (upper curve) and $m = 3$ (lower curve) compared with that for pure data (\diamond).

5. Conclusions

We have prescribed a method of noise reduction in chaotic time series with input parameters that can be determined from the data. Given the data we first sort out the parameters needed as input for the method. First an effective noise is determined from the divergence and correlation dimension plots. These also give an estimate of the minimum embedding dimension. Then comparing similar strings of data, with the lengths $(2k + 1)$ initially large we modify the midportions of the strings and then repeat the process with convergence at each k up to the length not less than the minimum embedding dimension as estimated from the earlier plots. Finally we average at the next higher k , similar to the 'constant fit' method as applied by Schreiber *et al* [6]. The method is found to work well even for small amount of data and also tested for two different types of symmetric noise distribution. It is important to note that in each iteration the cleaned data is remodified by the original raw data only. This has the advantage of improving the data in each iteration because the original noisy set is expected to give 'more correct information'.

Acknowledgement

We thank Professors Avijit Lahiri and Gautam Ghosh for many useful discussions and suggestions. One of us (AB) is grateful to the Department of Theoretical Physics, University of Cologne, for the computational facilities.

References

- [1] J Gao and Z Zheng, *Phys. Rev.* **E49**, 3807 (1994)
- [2] M Henon, *Commun. Math. Phys.* **50**, 69 (1979)
- [3] J Theiler, *Phys. Rev.* **A34**, 2427 (1986);
P Grassberger, *Phys. Lett.* **A148**, 63 (1990)
- [4] H G Schuster, *Deterministic chaos: An introduction* (Physik-Verlag GmbH, 1984)
- [5] E J Kostelich and J A Yorke, *Physica* **D41**, 183 (1990)
- [6] E J Kostelich and T Schreiber, *Phys. Rev.* **E48**, 1752 (1993)
- [7] T Schreiber and P Grassberger, *Phys. Lett.* **A160**, 411 (1991);
T Schreiber, *Phys. Rev.* **47**, 2401 (1993)
- [8] F Takens, *Lecture Notes in Math.* **898** (Springer, Heidelberg-New York, 1981)
- [9] E N Lorenz, *J. Atmos. Sci.* **30**, 130 (1963)