Extremely simple nonlinear noise-reduction method

Thomas Schreiber*
Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark
(Received 2 March 1992)

A very simple method to reduce noise in experimental data with nonlinear time evolution is presented. Locally constant fits are used to obtain a less noisy trajectory consistent with the dynamics as well as with the measured data. Neighborhoods are defined by coordinates both from the past and from the future. The method is applied to the Hénon map and to a discretized form of the Mackey-Glass equation.

PACS number(s): 05.45.+b, 06.50.Dc

I. INTRODUCTION

In the past decades progress has been made in the analysis of time series from phenomena with nonlinear behavior (for a review see, e.g., [1]). Every realistic measurement is to some extent contaminated by noise, which limits the performance of many techniques of modeling, prediction, and control. Thus, an important goal of time-series analysis is noise reduction. Traditional linear filters are based on the assumption that signal and noise components can be distinguished in the spectrum. For coarsely sampled signals from nonlinear systems this poses a problem since the signal itself can have a broadband spectrum. In this case one has to apply nonlinear methods in order not to distort the signal.

A number of nonlinear noise reduction algorithms have been developed by different authors [2–7] that take into account the nonlinear nature of the data. A comparison of the more recent approaches will be included in [8], where also a compromise is proposed which is considered optimal at this time. Different as they are, all of these methods have in common that their implementation is nontrivial, in particular if one is interested in optimal results.

The purpose of this paper is to present a nonlinear method of noise reduction which is very easy to implement and which needs only little resources (CPU time and memory). Except for small data sets, the simplicity has to be bought by slightly weaker results.

II. ALGORITHM

Suppose we have a scalar time series $\{x_i\}$, $i=1,\ldots,T$, where the x_i are composed of a clean signal y_i with some noise η_i added, $x_i=y_i+\eta_i$. Then $\sigma^2=\langle\eta^2\rangle$ is called the absolute noise level.

The main idea of the method presented in this paper is to replace each measurement x_i by the average value of this coordinate over points in a suitably chosen neighborhood. The neighborhoods are defined in a phase space reconstructed by delay coordinates. To define the neighborhoods, first fix positive integers k and l and construct embedding vectors

$$\mathbf{x}_i=(x_{i-k},\ldots,x_{i+l})$$

as usual [9]. Note that past and future coordinates are involved.

Further, choose a radius ϵ for the neighborhoods. For each value x_i find the set $\mathcal{U}_i^\epsilon$ of all neighbors x_j for which

$$\sup\{|x_{j-k}-x_{i-k}|,\ldots,|x_{j+l}-x_{i+l}|\}\equiv\|\mathbf{x}_j-\mathbf{x}_i\|_{\sup}<\epsilon\ ,$$

i.e., all segments of the trajectory which are close during

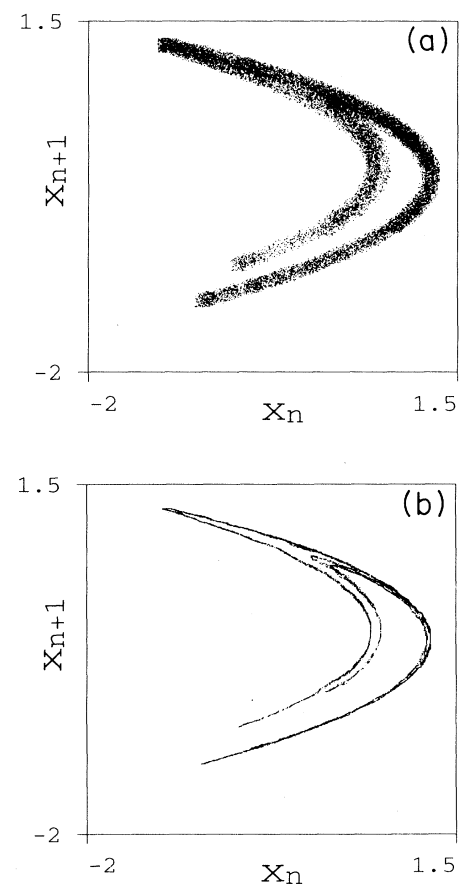

FIG. 1. Phase plots of iterates of the Hénon map. (a) A sample with 5% noise and (b) the same after noise reduction. Each panel contains 20 000 points.

a time lasting from k iterations in the past to l iterations in the future.

Replace the "present" coordinate x_i by its mean value [10] in $\mathcal{U}_i^\epsilon$,

$$x_i^{\text{corr}} = \frac{1}{|\mathcal{U}_i^\epsilon|} \sum_{\mathcal{U}_i^\epsilon} x_j \ . \tag{1}$$

The reason why only the central coordinate in the delay window is corrected is that only this coordinate is optimally controlled from past and future, i.e., its value is fixed along both the unstable and the stable manifolds [11]. The errors induced by this replacement are of statistical and geometrical nature. If the points in $\mathcal{U}_i^\epsilon$ are regarded as a random sample distributed according to the natural measure, the statistical uncertainty of the center of mass will be damped out like $|\mathcal{U}_i^\epsilon|^{-1/2}$ whereas the error introduced by replacing the geometrical center of the neighborhood by the center of mass depends on the nonuniformity of the distribution within $\mathcal{U}_i^\epsilon$ and will in general grow with the size of the neighborhood. We can expect the method to work when these errors are smaller than the individual errors of the coordinates. Figure 1 shows the effect of this procedure on a noisy sample of the Hénon map.

III. IMPLEMENTATION AND RESULTS

The implementation of the algorithm is straightforward. To obtain optimal results it is essential to choose ϵ, the size of the neighborhoods, appropriately. For the examples studied in this paper a value ϵ of about three times the amplitude of the noise σ was found to be optimal. In Fig. 2 the dependence of the amount of noise reduction on the parameter ϵ is shown. In an experimental situation one would have chosen ϵ at the point where the correction starts increasing more slowly. Above this value the geometrical error leads to a "correction" in the wrong direction.

If this behavior is not pronounced enough to estimate the optimal value of ϵ one can estimate the noise level from the correlation integral. An enlarged scaling region of the correlation integral after applying the algorithm can be taken as an indicator of noise reduction. In cases of doubt smaller values of ϵ are preferable since there is little risk of doing harm to the data by choosing small ϵ. In the worst case no neighbors are found, in which case the center of mass coincides with the original coordinate and no correction is made. This method does not have the problem of singular fits encountered in higher-order approximations [3–8].

With experimental data care should be taken and a statistical analysis of the corrections made. When only additive noise is removed, the corrections and the remaining signal should be uncorrelated. A problem common to all noise reduction methods is how to judge the effect of the procedure on a given experimental data set. A discussion of possible diagnostic tools is beyond the scope of this note, but material in [12] is applicable for the present algorithm as well.

The procedure can be iterated. If one takes the rms of the correction made as a new value for ϵ, ϵ will decrease exponentially with the number of iterations until eventually (typically after 2–6 iterations) no neighbors are found for any point and no further correction is made.

The only CPU time-consuming step is the neighbor search. Therefore it is essential to apply a fast searching algorithm. Recommended are box-assisted methods using linked lists as, e.g., described in [1,13,14]. Using this technique, the complete noise reduction algorithm can be coded in a few lines [15]. Even to handle long data sets (say 65 000 points), only a few minutes are needed on a DEC station 5000, depending on the noise level. Sets of about 1000 points can be processed in a few seconds.

To quantify the results, define r_{dyn} as, e.g., in [1]:

$$r_{\text{dyn}} = \left[\frac{\sum_i (x_i - f(x_{i-1}, \ldots))^2}{\sum_i (x_i^{\text{corr}} - f(x_{i-1}^{\text{corr}}, \ldots))^2} \right]^{1/2} .$$

This is the improvement of the deviation from the exact dynamical evolution. The improvement of the distance from the original noise-free trajectory is given by

$$r_0 = \left[\frac{\sum_i (x_i - y_i)^2}{\sum_i (x_i^{\text{corr}} - y_i)^2} \right]^{1/2} .$$

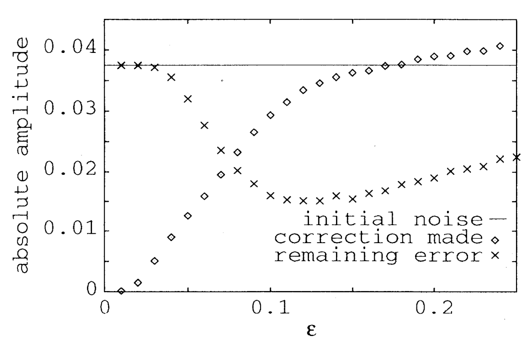

FIG. 2. Dependence of the amount of noise reduction in the first step on the size ϵ of the neighborhoods chosen. For 10 000 points of the Hénon map the initial amount of noise (5%), the rms of the correction made, and the deviation of the resulting signal from the noise-free one are shown.

In this paper we apply the algorithm to the Hénon map and a discretized version of the Mackey-Glass equation. Results are given in Figs. 1 and 3 and in Tables I and II. For the Hénon map [16] $y_{n+1} = 1 - a y_n^2 + b y_{n-1}$ we choose the canonical parameter values $a = 1.4$ and $b = 0.3$.

The Mackey-Glass delay differential equations [17] is integrated with a discrete time step so that it can be written as a $(M+1)$-dimensional map [18]:

$$y_{n+1}=\frac{1}{2M+b\tau}\left[(2M-b\tau)y_n + \tau a\left[\frac{y_{n-M}}{1+y_{n-M}^{10}}+\frac{y_{n-M+1}}{1+y_{n-M+1}^{10}}\right]\right]. \quad (2)$$

We choose parameter values $a=0.2$, $b=0.1$, and $\tau=30$. To obtain a reasonably sampled signal we choose $M=40$ corresponding to a step size $\Delta t=\tau/M=0.75$. That means that the integration is quite rough and maybe deviates considerably from the original differential equation. Indeed the attractor dimension is found to be slightly below 3 (as opposed to 3.1 for the differential equations).

The only reason to deal with the whole—still oversampled—signal is to be able to compute r_{dyn}. The high values of r_{dyn} (see Table II) are partly due to the fact that it is always easier to clean an oversampled signal. Even a simple smoothing procedure would reduce the noise to some extent. In fact, the present method degenerates to a smoothing scheme in the limit of high sampling and small ϵ since the neighbors found are points nearby in time.

The results in Tables I and II show that the method is surprisingly efficient, in particular on short data sets with a visible amount of noise. This is the type of data one often has to deal with in field experiments. A different question is what can be deduced from such small samples even if they were noise free.

In order to test the effect of noise reduction on dimension estimates we computed the correlation sum C_2 for three samples obtained from Eq. (2). Sample (a) consisted of 20 000 noise-free data points, sample (b) of the same points with 5% Gaussian noise added, and (c) of sample

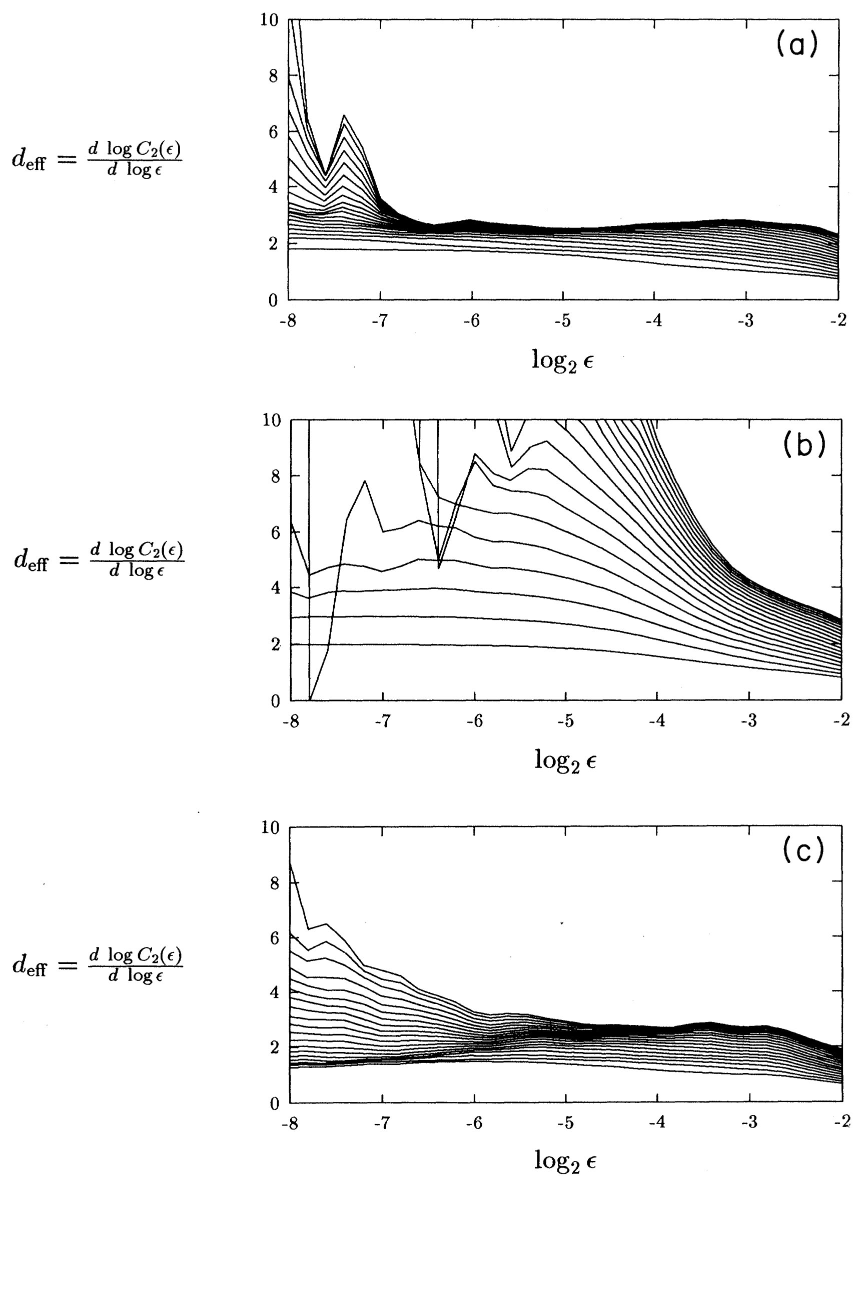

FIG. 3. Local slopes in a $\log_2 C_2(\epsilon)$ vs $\log_2\epsilon$ plot for 20 000 iterates of the discretized Mackay-Glass equation (2). Obtained from (a) the noise-free sample, (b) the same data with 5% Gaussian noise added, and (c) the noisy data of panel (b) after six steps of noise reduction. All three panels were obtained with a delay time of two steps and embedding dimensions 2–20.

TABLE I. Noise reduction on Hénon map.

T	% noise	r_{dyn}	r_0
500	5	1.7	1.4
	10	2.1	1.7
1 000	5	2.4	1.8
	10	3.6	2.2
5 000	5	4.4	2.5
	10	5.6	2.8
20 000	1	3.8	2.4
	5	7.8	3.1
	10	7.5	2.9
65 000	1	6.4	3.0
	5	9.9	3.3
	10	7.8	2.9
	Ref. [4]	4.76	
	Ref. [5]	9.65	4.05
	Ref. [8]	12.5	4.4

TABLE II. Noise reduction on Mackay-Glass equation.

T	% noise	r_{dyn}	r_0
500	5	5.1	2.0
	10	7.1	2.4
1 000	5	7.5	2.2
	10	9.0	2.1
5 000	5	11.5	2.4
	10	13.3	2.2
20 000	1	4.6	1.6
	5	11.9	2.3
	10	15.2	2.3
65 000	1	6.0	1.7
	5	12.8	2.4
	10	15.8	2.3

(b) after six iterations of noise reduction. The correlation sum [18] is defined as

$$C_2(\epsilon)=\frac{1}{T^2}\sum_{i,j}\Theta(\epsilon-\|\mathbf{x}_i-\mathbf{x}_j\|) , \tag{3}$$

where $\mathbf{x}_n$ are the embedding vectors. Figure 3 clearly shows that the scaling region is improved towards what can be obtained from the noise free signal.

The present method works best on moderate amounts of data with noise levels above 1%. If longer series with only a small amount of noise are available, e.g., in a laboratory experiment, it will be worth while to apply a more sophisticated procedure [8] to suppress noise well below 1%, which is indeed desirable to observe scaling properties at small length scales.

Finally it should be mentioned that the method can be easily generalized to multivariate time series. Zeroth-order fits also give a simple and very robust forecasting method. Their use for this purpose was proposed in [19].

In conclusion an extremely simple and quite efficient noise reduction scheme based on ideas from nonlinear dynamics has been presented.

ACKNOWLEDGMENTS

I want to thank my colleagues in Wuppertal, namely Peter Grassberger and Arkady Pikovsky, for encouraging discussions. Rainer Hegger provided assistance by reproducing the results. The author received a grant within the framework of the SCIENCE program of the Commission of the European Communities under Contract No. B/SC1*-900557.

*Permanent address: Physics Department, University of Wuppertal, Gauss-Strasse 20, D-5600 Wuppertal 1, Germany.

[1] P. Grassberger, T. Schreiber, and C. Schaffrath, J. Bifurcat. Chaos, **1**, 521 (1991).

[2] S. M. Hammel, Phys. Lett. A **148**, 421 (1990).

[3] J. D. Farmer and J. Sidorowich, Physica D **47**, 373 (1991).

[4] E. J. Kostelich and J. A. Yorke, Phys. Rev. A **38**, 1649 (1988).

[5] T. Schreiber and P. Grassberger, Phys. Lett. A **160**, 411 (1991).

[6] T. Sauer, Physica D **58**, 193 (1992).

[7] R. Cawley and G.-H. Hsu, Phys. Rev. A **46**, 3057 (1992).

[8] P. Grassberger, R. Hegger, H. Kantz, C. Schaffrath, and T. Schreiber, Chaos (to be published).

[9] T. Sauer, J. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).

[10] With small data sets some points happen to have the same neighborhoods and are therefore corrected to exactly the same value. This can be avoided [A. Pikovsky (private communication)] by introducing distance dependent weights in Eq. (1). In the examples studied in this paper this leads to a slight improvement of the results.

[11] If all coordinates would be corrected one would obtain the trivial case of the methods of [6] and [7] in which *all* directions are projected out.

[12] H. Kantz, in *Predicting the Future and Understanding the Past: A Comparison of Approaches,* Proceedings of the NATO ARW on Time Series Analysis and Forecasting held in Santa Fe, NM, 1992, edited by A. S. Weigend and N. A. Gershenfeld Santa Fe Institute Studies in the Science of Complexity (Addison-Wesley, Reading, MA, in press, 1993).

[13] P. Grassberger, Phys. Lett. A **148**, 63 (1990).

[14] T. Schreiber, in *Predicting the Future and Understanding the Past: A Comparison of Approaches* (Ref. [12]).

[15] The source code in the form of a FORTRAN subroutine is available from the author; electronic address: schreib @ complex.nbi.dk

[16] M. Hénon, Commun. Math. Phys. **50**, 169 (1976).

[17] M. C. Mackey and L. Glass, Science **197**, 287 (1977).

[18] P. Grassberger and I. Procaccia, Phys. Rev. A **28**, 2591 (1983).

[19] A. Pikovsky, Sov. J. Commun. Technol. Electron. **31**, 81 (1986).