

Filtered noise can mimic low-dimensional chaotic attractors

P. E. Rapp

Department of Physiology, The Medical College of Pennsylvania, 3300 Henry Avenue, Philadelphia, Pennsylvania 19129

A. M. Albano

Department of Physics, Bryn Mawr College, Bryn Mawr, Pennsylvania 19010

T. I. Schmah

Department of Mathematics, Bryn Mawr College, Bryn Mawr, Pennsylvania 19010

L. A. Farwell

Human Brain Research Laboratory, 12944C Travilah Road, Potomac, Maryland 20854

(Received 14 May 1992; revised manuscript received 5 November 1992)

This contribution presents four results. First, calculations indicate that when examined by the Grassberger-Procaccia algorithm alone, filtered noise can mimic low-dimensional chaotic attractors. Given the ubiquity of signal filtering in experimental investigations, this is potentially important. Second, a criterion is derived which provides an estimate of the minimum data accuracy needed to resolve the dimension of an attractor. Third, it is shown that a criterion derived by Eckmann and Ruelle [*Physica D* **56**, 185 (1992)] to estimate the minimum number of data points required in a Grassberger-Procaccia calculation can be used to provide a further check on these dimension estimates. Fourth, it is shown that surrogate data techniques recently published by Theiler and his colleagues [in *Nonlinear Modeling and Forecasting*, edited by M. Casdagli and S. Eubanks (Addison Wesley, Reading, MA, 1992)] can successfully distinguish between linearly correlated noise and nonlinear structure. These results, and most particularly the first, indicate that Grassberger-Procaccia results must be interpreted with far greater circumspection than has previously been the case, and that the algorithm should be used in combination with additional procedures such as calculations with surrogate data. When filtered signals are examined by this algorithm alone, a finite noninteger value of D_2 is consistent with low-dimensional chaotic behavior, but it is certainly not a definitive diagnostic of chaos.

PACS number(s): 05.45.+b, 02.60.-x

I. INTRODUCTION

Experimentally acquired signals are typically filtered. Additionally, analytical methods in which interevent interval data are reduced to continuous functions are, in effect, filters. Given the widespread use of filtering, it is important to determine its effects on dynamical analysis. This contribution focuses on a specific issue: what effect does filtering have on the estimation of the correlation dimension using the Grassberger-Procaccia algorithm? This question has already received attention from a number of investigators. One paper explicitly considering the effect of filtering on dynamical analysis that has come to our attention is that of Badii and Politi [1]. They considered a restricted class of filter, the single-pole low-pass filter, and concluded that "a dimension increase can be theoretically predicted, contrary to the idea that filtering reduces the dimensionality of the signal." These results were generalized by Badii *et al.* [2] and Mitschke, Möller, and Lange [3], who have argued that whenever the effects of a filter can be described by a differential equation, then filtering should increase dimension estimates. Conversely, empirical experience with very noisy signals produced the opposite effect [4]. Also, digital filters themselves can be a source of chaos [5].

Mitschke's more recent results [6] indicate that acausal filters do not necessarily cause an artifactual increase in dimension estimates. This is consistent with results of Yip, Holstein-Rathlou, and Marsh [7], who demonstrated that an appropriately tuned Kaiser-Bessel filter can accurately resolve a noise corrupted low-dimensional attractor. In a recent paper Broomhead, Huke, and Muldoon [8] have shown theoretically that finite-order nonrecursive filters leave invariant quantities that can be estimated using embedding techniques, including, for example, the correlation dimension. However, they are careful to make the important distinction between that which is theoretically true and practical issues of numerical estimation using finite data sets.

In this paper we construct an extremely simple filter than can produce the spurious identification of chaotic behavior by the Grassberger-Procaccia algorithm. Additionally, we show that calculations using a surrogate-data method described by Theiler and his colleagues [9] can successfully distinguish between linearly correlated noise and nonlinear structure. As Theiler *et al.* stress, surrogate-based hypothesis testing has a long history. These papers are reviewed in Sec. V. Given the ubiquity of filtered signals in experimental science, these results suggest that the estimates obtained from the

Grassberger-Procaccia algorithm must be interpreted with far greater circumspection than has typically been the case. If filtered signals are being examined, this method should only be used in combination with other procedures, such as surrogate data calculations.

II. SPECIFICATION OF THE RANDOM DATA SET AND THE FILTER

A set of random numbers, uniformly distributed on the unit interval, was generated using a procedure by Press *et al.* [10]. This procedure follows a design specified by Knuth [11]. In order to define a time scale for subsequent calculations, it is assumed that the random-number generator was sampled at a frequency of 1 Hz; 8192 data points are used in all calculations unless otherwise specified.

This data set was filtered by a procedure modified from the smoothing algorithm published by Press *et al.* [10]. The Fourier transform of the original data set is calculated and the j th element of the resulting Fourier series is multiplied by factor F_j ,

$$F_j = \max[0, 1 - kj^2], \quad k = 0.37 \times 10^{-6}.$$

Since the Fourier transform is linear, the resulting filter is linear; that is, if $x(t)$ and $y(t)$ are functions of time, a and b are constant multipliers, and $F(\cdot)$ represents the action of the filter, then $F(ax(t) + by(t)) = aF(x(t)) + bF(y(t))$. Filters are also classified as causal or acausal. If $F(x(t_j))$ depends only on $x(t_1), \dots, x(t_j)$, then the filter is said to be causal. In the filter described here, the value of $F(x(t_j))$ depends on the entire time series. Filters of this type are acausal.

It is appropriate to ask what signal properties remain invariant under the filtering process, and most particularly what happens to the dimension. In the Appendix we show that it is possible to present our filter in a form that corresponds to the finite-order, nonrecursive filters described by Broomhead, Huke, and Muldoon [8]. Filters that satisfy the conditions of this theorem do not change properties that are determined after embedding by the method of delays. This would include Lyapunov exponents, Kolmogorov entropy and the dimension. From the equation for F_j given above, we see that our filtered signal has 10^3 independent frequencies. Therefore, theoretically, the dimension of a signal will be unaltered by the filter if its dimension is less than 499. As Broomhead, Huke, and Muldoon stress, "the situation is not so clear, however, even for a finite-order nonrecursive filter, when constructing numerical estimates of dimension." This is the issue of focus for this paper. We are concerned here with the practical, numerical difficulties that are encountered when we try to estimate dimension using finite, filtered data sets. Numerical estimates of dimension are not invariant under the filtering process.

Figure 1 shows a section of the original set of random numbers and the corresponding filtered output. The spectrum of the original data set is uniform, and the spectrum of the filtered data set follows the profile given by F_j as defined above. All spectra were calculated four times using square, Welch, Hanning, and Parzen windows. No

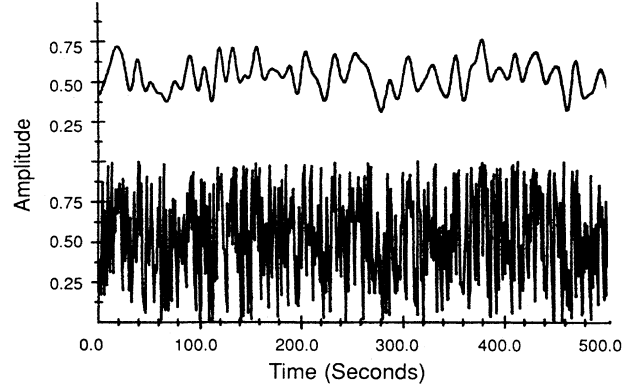


FIG. 1. The effect of filtering random numbers with the filter specified in the text. The input signal (lower trace) is the set of random numbers. The filter's output is shown in the upper trace.

significant window-dependent differences were observed. These two data sets constitute the experimental material for the next set of calculations.

III. CONVERGENCE CRITERIA FOR THE GRASSBERGER-PROCACCIA ALGORITHM

In order to establish several definitions precisely, we begin the discussion with an abbreviated didactic statement of the Grassberger-Procaccia algorithm [12,13]. The investigation begins with a data set $\{V_j\}$ measured at a time interval T_S . These data are used to construct points $\{X_k\}$ in an N -dimensional embedding space,

$$X_K = (V_{1+(K-1)J}, V_{1+(K-1)J+L}, \\ V_{1+(K-1)J+2L}, \dots, V_{1+(K-1)J+(N-1)L}).$$

lag L , together with embedding dimension N and sampling interval determine the window, $W = (N-1)LT_S$, of the embedding. The window is the length of time required to measure the data used to construct any point X_K . The important parameter in an embedding is not LT_S or N separately, but the value of window. Albano *et al.* [14] have shown over a wide range of values of window that if the window is the same, even though the embedding dimension and lag are different, the resulting correlation integrals nearly coincide and produce nearly equal estimates of dimension. They have also shown that the selection of an appropriate value of window is crucial to the successful application of the Grassberger-Procaccia algorithm. An inappropriate window can result in the unnecessary failure to resolve a finite-dimensional attractor. Several candidate measures for predicting a successful window are summarized by Albano *et al.* [14,15], Liebert and Schuster [16], and Gibson *et al.* [17].

The correlation integral corresponding to this embedding is then constructed,

$$C_N(r) = (1/N_p) \sum_{i=1}^{K-1} \sum_{j=i+1}^K \Theta(r - |X_i - X_j|),$$

where Θ is the Heaviside function, K is the number of points in the embedding space, and N_p is the number of distinct pairs of points. In all of the calculations presented here, the Euclidean metric is used. There are two reasons for making this choice of metric. First, the Euclidean metric is typically used in the Grassberger-Procaccia literature that we wish to address in this paper. Second, we have suggested [14] using the orthogonal transformation specified by the singular-value decomposition prior to calculations of the Grassberger-Procaccia algorithm as a noise-reduction technique. The Euclidean metric is invariant under a rotation. Other metrics, such as the max norm, are not.

Grassberger and Procaccia demonstrate that if a series of conditions are satisfied, then the double-logarithmic plot of the correlation integral versus the distance in the embedding space constructed using data obtained from a finite-dimensional system will display a linear region called the scaling region. The slope of the scaling region is the correlation dimension of the attractor. The conditions of the theorem require using a sufficient number of data points of high quality (low noise) and a sufficiently large value of embedding dimension. In the calculations presented here, we have estimated the derivative of the correlation integral by determining the slope of 11 consecutive points. The partition width of the $\ln(r)$ axis was 0.05 in all calculations. It should be stressed that the dimension estimate itself is determined directly from the correlation integral. The derivative calculation is used only to determine boundaries r_L and r_U of the scaling region. Though its limitations are recognized, this method results in an objective, quantitative procedure that allows systematic comparisons between cases. Following a recommendation of Caswell and Yorke [18], we include the specification of r_L and r_U in the report of the dimension estimates summarized in Tables I and II. r_{\max} is the largest interpoint distance in the embedding space.

Three convergence criteria should be applied to dimension estimation based on the Grassberger-Procaccia algorithm. First, the scaling region must be linear. The degree of variation of the derivative in the scaling region should not exceed some specified standard. This factor

TABLE I. Results from random numbers.

N	D_2	$\ln r_L$	$\ln r_U$	$\ln r_U - \ln r_L$	$\ln r_{\max}$	D_{\max}
1	1.00±0.15	-11.00	-1.90	9.10	0.00	6.95
2	2.02±0.30	-7.75	-1.45	6.30	0.30	7.38
3	2.94±0.44	-4.85	-0.95	3.90	0.45	8.69
4	3.85±0.58	-3.80	-0.75	3.05	0.55	9.26
5	4.73±0.71	-3.25	-0.90	2.35	0.60	8.57
6	5.71±0.86	-2.90	-0.80	2.10	0.65	8.89
7	6.46±0.97	-2.20	-0.80	1.40	0.70	9.14
8	7.16±1.07	-1.90	-0.40	1.50	0.75	10.98
9	7.98±1.20	-1.50	-0.50	1.00	0.75	11.06
10	8.79±1.32	-1.35	-0.35	1.00	0.80	11.78
11	9.63±1.44	-1.40	-0.20	1.20	0.85	12.18
12	10.62±1.59	-1.15	-0.20	0.95	0.85	12.73
13	12.01±1.80	-1.00	-0.40	0.60	0.90	11.58
14	12.44±1.87	-0.95	-0.15	0.80	0.90	13.14
15	12.72±1.91	-0.80	-0.10	0.70	0.95	13.45

TABLE II. Results from filtered random numbers.

N	D_2	$\ln r_L$	$\ln r_U$	$\ln r_U - \ln r_L$	$\ln r_{\max}$	D_{\max}
1	1.00±0.15	-14.80	-2.25	12.55	-0.55	7.53
2	1.98±0.30	-8.50	-3.55	4.95	-0.20	4.47
3	2.97±0.45	-7.70	-5.05	2.65	0.00	3.18
4	3.02±0.45	-6.35	-4.60	1.75	0.15	3.41
5	2.95±0.44	-5.15	-3.85	1.30	0.20	4.00
6	3.70±0.56	-6.00	-4.50	1.50	0.30	3.41
7	3.85±0.58	-6.05	-4.10	1.95	0.35	3.60
8	4.07±0.61	-5.70	-4.25	1.45	0.40	3.51
9	4.00±0.60	-5.15	-3.70	1.45	0.40	3.93
10	4.18±0.63	-4.95	-3.50	1.45	0.45	4.07
11	4.36±0.65	-4.70	-3.30	1.40	0.45	4.27
12	4.57±0.69	-4.80	-3.35	1.45	0.50	4.16
13	4.35±0.65	-4.10	-2.70	1.40	0.50	4.91
14	4.40±0.66	-4.05	-2.45	1.60	0.50	5.21
15	4.45±0.67	-4.00	-2.20	1.80	0.50	5.56

establishes an optimistic estimate of uncertainty in the dimension estimate. Second, the scaling region must be of significant length. In Tables I and II, we report $\ln(r_U) - \ln(r_L)$. A length of $\Delta \ln(r) = 1.6$ corresponds to a fivefold variation in r . We regard this as being the minimally acceptable scaling length. Third, the estimate of dimension should be robust against variations in the embedding window. It should be noted that the estimate should be robust against changes in window, not against changes in embedding dimension alone.

A crucial issue to any experiment is the accuracy of the experimental measurement. Möller *et al.* [19] have shown that inadequate resolution causes an underestimate of the dimension. Let Δr be the uncertainty in the measurement of a distance in embedding space. For the dimension estimate to be meaningful we require $\Delta r < r_L$, where r_L is the lower bound of the scaling region:

$$r \sim [N(x-y)^2]^{1/2},$$

where N is the embedding dimension, and x and y are representative values of the measured variable. Let Δx and Δy be the associated uncertainties in x and y . Δx and Δy are determined by several factors including arithmetic precision in the case of theoretical calculations and digitizer resolution in experimental investigations,

$$\Delta r^2 = \left[\frac{\partial r}{\partial x} \right]^2 \Delta x^2 + \left[\frac{\partial r}{\partial y} \right]^2 \Delta y^2,$$

$$\Delta r^2 \sim 2 \left[\frac{\partial r}{\partial x} \right]^2 \Delta x^2 = 2N^2(x-y)^2 \Delta x^2 / r^2.$$

Since r is of order $N^{1/2}(x-y)$,

$$\Delta r \sim (2N)^{1/2} \Delta x.$$

The resolution requirement $\Delta r < r_L$ therefore becomes

$$\Delta x < r_L / (2N)^{1/2}.$$

The value of r_L is determined from the dimension calculation. If Δx does not satisfy this condition, reduced confidence must be ascribed to the dimension estimate.

Conversely, existing implementations of the algorithm can be modified to begin the search for r_L at $r_L = (2N)^{1/2} \Delta x$. A modest acceleration in the computation can result. This criterion is satisfied in all of the dimension calculations presented here.

A final consideration is N_{data} , the number of data points required to resolve the attractor. As previously noted [20] there is no universal response to this question since the answer will depend on the distribution of points on the attractor. Smith's analysis [21] indicates that data requirements increase exponentially with the dimension of the attractor, 42^D . However, Grassberger, Schreiber, and Schaffrath [22] have argued that this pessimistic estimate is based on assumptions that are not applicable to the correlation dimension. Eckmann and Ruelle [23] have produced a somewhat more optimistic estimate. According to their criterion, the maximum dimension estimate that can be supported by N_{data} data points is

$$D_{\text{max}} = 2 \ln N_{\text{data}} / \ln(1/\rho),$$

where $\rho = r/r_{\text{max}}$. Here r_{max} is the maximum distance vector and r is a representative value of r in the scaling region. In our calculations we use $r_{\text{mid}} = (r_L + r_U)/2$ as the representative value. D_{max} should be calculated as part of the dimension-estimation procedure. If the estimated value of dimension exceeds D_{max} , that estimate is questionable.

IV. DIMENSION CALCULATIONS USING RANDOM NUMBERS AND FILTERED RANDOM NUMBERS

The results of the dimension calculations using our set of random numbers are summarized in Fig. 2 and in Table I. The behavior expected for random numbers is observed, namely the estimated dimension should be equal to the embedding dimension. The failure to fill the embedding space at higher dimensions is a consequence of the finite number of data points used. This failure indicates that care will have to be exercised when examin-

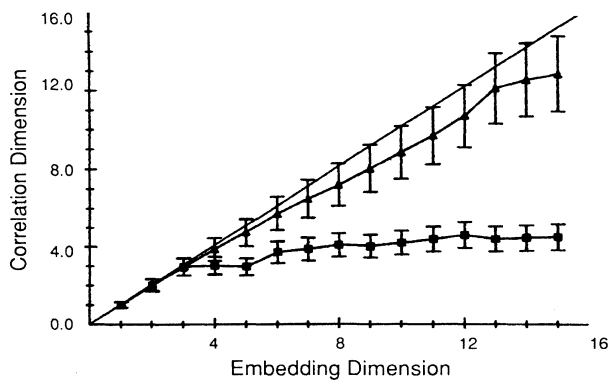


FIG. 2. Estimated correlation dimension as determined by the Grassberger-Procaccia algorithm as a function of embedding dimension. The input data sets were 8192 random numbers (upper trace) and 8192 filtered random numbers (lower trace). In each embedding $L=J=1$. Numerical details of these calculations are presented in Tables I and II.

ing dimension calculations in higher-dimensional spaces (roughly when $N \geq 10$) if only 8192 points are used. In summary, these calculations with random numbers show an estimated value of dimension that increases with increasing embedding dimension. This data set would not be mistaken for a system with dimension less than 16.

The results of the dimension calculations using filtered random numbers are shown in Fig. 2 and in Table II. The filter described in Sec. II effectively reduces the number of independent points in the data set. Thus, the estimated dimension of filtered noise should be less than the dimension of the original set of random numbers. These results were obtained when the maximum permitted excursion of the derivative in the scaling region was limited to 15%. The plateau length is stable and the value of the dimension estimate is comparatively robust, certainty within 15%, with changes in the embedding window. Also, D_{max} is greater than the dimension estimate in the critical embeddings that display a stable value of dimension. It should be noted that these calculations are based on a naive embedding. In each case $L=J=1$. Preliminary calculations suggest that if the dimension calculations were preceded by a systematic effort to establish a better window, the scaling-region duration would be significantly improved. These calculations suggest that inappropriately filtered random signals can mimic finite-dimensional chaotic attractors when they are analyzed by the Grassberger-Procaccia algorithm alone.

Theiler [24] has demonstrated that erroneous results can be obtained from the Grassberger-Procaccia algorithm when correlations from consecutively sampled points produce spuriously low values of the correlation dimension. These correlations are to be contrasted from geometrically significant correlations which result when the trajectory in embedding space returns to the vicinity of an earlier point on the trajectory. Theiler argues that a simple modification of the correlation integral can correct for this bias,

$$C_N(r) = (1/N_P) \sum_{i=1}^{K-B} \sum_{j=i+B}^K \Theta(r - |X_i - X_j|).$$

The case $B=1$ reduces to the conventionally defined correlation integral. The choice of B must be considered. A natural time scale for any time series is its autocorrelation time (the time required for the autocorrelation function to drop to $1/e$ of its original value). For the filtered random numbers the autocorrelation time is 6 s. We recalculated the dimension of this data set for $B=6, 60,$ and 600 ($M=8192, N=1-15, L=1, J=1$). No significant differences were observed. This insensitivity to B should be expected since for this time series all of the correlations result from time-correlation pairs. The correlation time introduced by the filter is very small compared to the total time span, and thus deleting some of the time-correlated pairs does not significantly change the final result.

V. CALCULATIONS WITH ALGORITHM 1 SURROGATE DATA

Scheinkman and LeBaron [25] have proposed a simple and easily implemented shuffle test that can be used to ex-

clude at least some classes of spurious Grassberger-Procaccia results. In this test the correlation integral is calculated with a data set obtained by shuffling the original data. A representative result obtained by using filtered random numbers and a data set produced by randomly shuffling the data is shown in Fig. 3. (Following Albano *et al.* [14], the x axis in this diagram is normalized against r_{\max} , the largest interpoint distance in the embedding space.) The differences between the correlation integral obtained from the original data and its shuffled variant are readily observed and indicate that there is a nonrandom structure in the filtered random numbers. However, the shuffle test detects any correlation. In the case of filtered noise the existence of a correlation in the data is not in dispute. Rather, our objective is to determine if a meaningful structure underlies this observed correlation.

Several investigators have demonstrated the importance of bringing more rigor to measurements of dimension [26–29]. Notable in this regard is the work of Brock and his colleagues [30–32]. A recent review of their work is given in Brock and Potter [33]. As a result of the recognized need for improved procedures, a number of investigators have independently proposed using a class of methods that, collectively, have come to be called surrogate data techniques. The surrogate data method is only one of many techniques that can be used to avoid spurious estimates of dimension. Early variants of the method include contributions by Grassberger [34], Brock [35], Brock, Dechert, and Scheinkman [36], Osborne *et al.* [37], Theiler [38], Elgar and Mayer-Kress [39], Theiler [40], Kaplan and Cohen [41], and Pijn [42]. Theiler and his colleagues [9] have expanded on this work in a recent paper. Following the nomenclature in their paper, the most elementary form of the method is referred to as the Algorithm 1 test. This procedure directly examines the null hypothesis that the signal in question is linearly correlated noise. This hypothesis is tested by constructing a surrogate data set in a three-step process.

(1) Determine the Fourier transform of the original data set.

- (2) Randomize the phases of this Fourier transform.
- (3) Produce a surrogate data set by taking the inverse transform.

The surrogate and original data will have the same power spectrum and autocorrelation function. A measure is then applied to the original data set and to its surrogate. In this paper we are specifically interested in identifying possible sources of errors that have appeared in published estimates of dimension. Therefore, the correlation dimension as determined by the Grassberger-Procaccia algorithm is the calculated measure. However, other measures of dynamical behavior can be employed. In addition to dimension, forecasting error [43,44], Lyapunov exponents [45–47], and direct measurements of local vector fields [48] have been used with surrogate data sets. Our present understanding of the numerical difficulties associated with estimation of different measures is expanding rapidly. For example, Eckmann and Ruelle [23] have recently argued that measuring Lyapunov exponents may require much more data than has previously been supposed. Also, it is now recognized that the Grassberger-Procaccia algorithm may not be the best approach to estimating dimension. Recent work by Judd [29] has resulted in an alternative procedure that produces dimension estimates based not on the correlation integral, but on the distribution of interpoint intervals that is used to calculate the integral.

If the measures obtained from the original and its surrogate are significantly different, the null hypothesis fails, and it is concluded that the original data set is not simply linearly correlated noise. A crucial word in the preceding sentence is “significantly.” In order to determine the confidence with which the null hypothesis can be rejected, several distinct surrogate data sets should be generated by the three-step procedure. Let D_{data} be the correlation dimension of the original data. Let $\langle D_{\text{surr}} \rangle$ be the mean correlation dimension of the surrogates, and σ_{surr} be the standard deviation of the correlation-dimension values obtained from the surrogates. The significance S is defined as

$$S = |D_{\text{data}} - \langle D_{\text{surr}} \rangle| / \sigma_{\text{surr}} .$$

Theiler [9], suggest that a value of $S \sim 2$ is not significant but $S \sim 10$ is highly significant. In that paper they also derive an analytic expression for ΔS , the uncertainty in S . As an example of the method, we began by applying it to data where deterministic, nonlinear structure is known to exist. Test data were generated by the Hénon attractor

$$x_{n+1} = 1 - ax_n^2 + y_n, \quad y_{n+1} = bx_n, \quad a = 1.4, \quad b = 0.3 .$$

In calculations using these data and the Grassberger-Procaccia algorithm, satisfactory scaling regions can be obtained when the derivative criterion is reduced to 10%. The requirement $D_2 \leq D_{\max}$ was easily met in each embedding [49]. When a correlation integral obtained from the Hénon data is compared with that obtained from one of its surrogates, dramatic differences are observed. Table III (case 1) shows the results of systematic comparisons of the Hénon data and its Algorithm 1 sur-

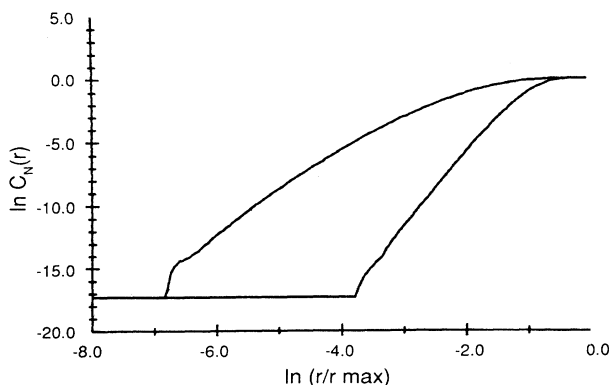


FIG. 3. Correlation integrals obtained from filtered random numbers and its random shuffle. The random shuffle has the steeper slope. In both calculations $N_{\text{data}} = 8192$, $N = 6$, $L = 1$, $J = 1$.

TABLE III. Significance as function of embedding dimension. Case 1: Hénon data, $N=8192$, $L=J=1$, ten Algorithm 1 surrogates. Case 2: Linearly filtered random numbers, $N=8192$, $L=J=1$, ten Algorithm 1 surrogates. Case 3: Nonlinearly transformed, linearly filtered random numbers, $N=8192$, $L=J=1$, ten Algorithm 1 surrogates. Case 4: Nonlinearly transformed, linearly filtered random numbers, $N=8192$, $L=J=1$, ten Algorithm 2 surrogates.

N_{emb}	Case 1	Case 2	Case 3	Case 4
2	54.00±12.08	0.13±0.32	8.25±1.86	1.14±0.41
3	98.64±22.06	1.53±0.47	10.00±2.26	0.10±0.32
4	63.29±14.16	3.04±0.75	6.25±1.43	0.71±0.35
5	46.09±10.31	2.53±0.65	3.84±0.92	0.97±0.38
6	48.75±10.91	0.10±0.32	4.57±1.07	0.34±0.33
7	50.22±11.23	0.09±0.32	3.07±0.76	0.46±0.33
8	25.09±5.62	0.13±0.32	3.63±0.87	0.19±0.32
9	21.74±4.87	0.23±0.32	2.79±0.70	0.31±0.32
10	35.23±7.88	0.22±0.32	1.66±0.49	2.69±0.68
11	22.59±5.06	0.02±0.32	1.88±0.53	1.95±0.53
12	19.41±4.35	0.22±0.32	1.92±0.53	1.51±0.46
13	16.10±3.61	0.20±0.32	2.82±0.71	1.15±0.41
14	21.13±4.74	0.32±0.32	1.98±0.54	1.97±0.54
15	13.27±2.98	0.15±0.32	1.91±0.53	1.59±0.48

rogates. Given the very high S values, it is possible to reject the null hypothesis, that the Hénon data are linearly correlated noise, with a high level of confidence. Algorithm 1 was then applied to filtered random numbers. When correlation integrals obtained from these data and one of their surrogates are compared, they very nearly superimpose. Results from calculations using ten surrogates are also summarized in Table III (case 2). Using this test it is possible to conclude that the data set of filtered random numbers is, indeed, indistinguishable from linearly correlated noise.

VI. CALCULATIONS WITH ALGORITHM 2 SURROGATES

The set of linearly filtered random numbers is clearly an artificial data set. In actual laboratory applications some degree of nonlinear distortion of the signal often occurs. Suppose that the filtered noise was subjected to a monotonic nonlinear transformation h . Would Algorithm 1 correctly identify the random origin of the observed signal? This question is addressed in the next set of calculations. The filtered random numbers were transformed by the function

$$h(x) = \frac{\left[\frac{x - x_{\min} - 0.0001}{x_{\max} - x + 0.0001} \right]^{\rho}}{1 + \left[\frac{x - x_{\min} - 0.0001}{x_{\max} - x + 0.0001} \right]^{\rho}},$$

where x_{\min} and x_{\max} are the minimum and maximum values of the set of filtered random numbers and ρ , in this calculation, is equal to three. The results of dimension calculations using filtered random numbers transformed

by this nonlinear function are given in Fig. 4. Grassberger-Procaccia dimension estimates obtained with these data are much less robust than those obtained using the original set of filtered random numbers. In order to satisfy the $D_2 < D_{\max}$ condition and to get scaling regions of even marginal acceptability, it is necessary to increase the derivative criterion to 20%. Even without the benefit of surrogate data calculations, one would conclude that the identification of a low-dimensional structure in this data set was very tentative.

When correlation integrals calculated using the transformed data and one of its Algorithm 1 surrogates are compared, differences are observed. However, these differences are not as dramatic as those seen with Hénon data. These calculations are also summarized in Table III (case 3). These results are intermediate to case 1 (Hénon data) and case 2 (linearly filtered random numbers). The average significance for embedding dimensions 2 to 15 is $S=3.89$. However, it was previously observed that because only 8192 points were used, results obtained for $N \geq 10$ are suspect. The average significance for $N=2$ to $N=9$ is $S=5.29$. When examined with Algorithm 1 surrogates, the nonlinearly transformed filtered random numbers might be deemed to have low-dimensional behavior.

The problem of nonlinearly transformed signals has also been considered by Theiler *et al.* [9]. Algorithm 1 addresses the null hypothesis: the signal is linearly correlated noise. Algorithm 2 addresses the null hypothesis: the signal is linearly correlated noise that has been transformed by a static, monotonic, nonlinear function. The surrogates that are used to investigate this hypothesis are constructed by the following procedure.

(A) let $\{x_i\}$ denote the original times series. A time series $\{y_i\}$ of equal length is constructed with two properties: the elements $\{y_i\}$ are drawn from a random Gaussian distribution, and the time series $\{y_i\}$ has the same rank ordering as $\{x_i\}$; that is, if x_{27} is the third largest value of $\{x_i\}$, then y_{27} is the third largest value of

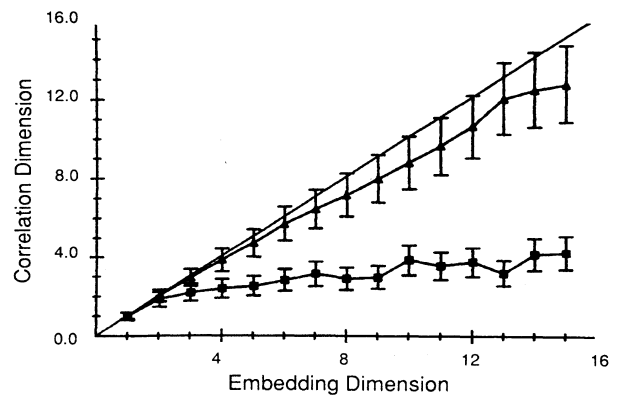


FIG. 4. Estimated correlation dimension as determined by the Grassberger-Procaccia algorithm as a function of embedding dimension. The input data sets were 8192 random numbers (upper trace) and 8192 numbers generated by a nonlinear transformation of filtered random numbers. In each embedding $L=J=1$.

$\{y_i\}$. Given the hypothesis that the underlying process is random, the object is to construct in $\{y_i\}$ the time series that would have been observed prior to the action of the nonlinear transformation h ; $y_i = h^{-1}(x_i)$.

(B) Given $\{y_i\}$, an Algorithm 1 surrogate, denoted $\{y'_i\}$, is constructed. If the null hypothesis is correct (the underlying signal is random), then randomizing the phase of $\{y_i\}$ to create $\{y'_i\}$ will not result in a loss of structure. The measured properties of the two time series should be the same.

(C) A surrogate, denoted $\{x'_i\}$, is constructed by shuffling the original time series $\{x_i\}$ so that the rank structure of $\{y'_i\}$ determines the rank structure of $\{x'_i\}$. The surrogate time series $\{x'_i\}$ is used in subsequent calculations, in this instance, in calculations of dimension.

Algorithm 2 is thus seen to be a shuffle test, but unlike the previously described shuffle it is very specific. The essential step in Theiler's analysis is the numerical approximation of h^{-1} . For this construction to be successful and for the subsequent shuffle of $\{x_i\}$ to be other than just a random shuffle, the data set $\{x_i\}$ has to be large and well distributed over its domain. This raises questions concerning data requirements for Algorithm 2 which require systematic investigation.

When correlation integrals from the nonlinearly transformed filtered random numbers and from one of its Algorithm 2 surrogates are compared they are found to be nearly identical. The significance values from ten surrogates are summarized in Table III (case 4). Calculations with Algorithm 2 correctly identify the random nature of the underlying signal. As Theiler *et al.* [9] observe, these algorithms consider only two of many possible null hypotheses. It is in principle possible to construct a sequence of hypotheses and corresponding surrogates to conduct dynamical investigations of increasing specificity and rigor.

VII. CONCLUSIONS

In this investigation random numbers, filtered by a linear, acausal filter were examined by the Grassberger-Procaccia algorithm. These calculations alone failed to identify the underlying random nature of the time series. These results suggest that considerable care must be exercised when filtered signals are examined. Given the ubiquity of filters in experimental equipment, this observation is applicable to a broad range of experimental studies. A finite, noninteger value of D_2 is consistent with low-dimensional chaotic behavior, but it is not a definitive diagnostic of chaos.

Many experimental investigations begin with the measurement of interevent time intervals. Examples include the measurement of the time intervals between neural action potentials and between heart beats. In some studies, event-interval data are used to construct a continuous function in time that gives an approximation of the frequency at any given moment. This is done in order to use analysis procedures that presuppose the existence of a continuous function defined at uniformly spaced time intervals. The construction of a continuous wave form is sometimes used by these investigators as a preliminary

step prior to application of the Grassberger-Procaccia algorithm. The process by which interval data is converted to a wave form is itself a filtering procedure. The results presented here suggest that this filtering process could have artifactually produced the appearance of low-dimensional structures that do not, in fact, exist in the original data. A comparison with surrogate data should be used to determine if pretreatment of the original data in this way has resulted in the spurious identification of low-dimensional attractors in biological data. The Grassberger-Procaccia algorithm can be applied directly to interval data. Unnecessary filtering should certainly be avoided. Calculations with original interval data should be combined with calculations using appropriately constructed surrogates.

While the numerical results presented here sound several cautionary notes, our overall conclusions are optimistic. The surrogate data technology is broadly applicable and may prove to be particularly helpful in the analysis of noisy biological data. Several factors combine to make the dynamical analysis of biological signals exceptionally difficult [50], but preliminary results by Theiler *et al.* [9] are encouraging. They present an investigation of an electroencephalographic (EEG) signal using the Algorithm 2 surrogates. The estimated dimension increases with embedding dimension. If this EEG record had been examined by the Grassberger-Procaccia algorithm alone, it would have been concluded that low-dimensional behavior was absent, and the search for statistically significant evidence of low-dimensional structure would have been abandoned. However, the values of S and ΔS obtained from this data set and its Algorithm 2 surrogates indicate that significant higher-dimensional nonlinear structure is present.

Note added. We recently became aware of unpublished work by James Theiler, who also found that filtered noise can mimic low-dimensional chaos.

ACKNOWLEDGMENTS

P.E.R. would like to acknowledge support from NIH Grant No. NS19716. We wish to thank the College Computer Center of the Medical College of Pennsylvania for its support. Professor Donald Marsh of the University of Southern California and Professor Alistair Mees of the University of Western Australia provided advice about filtering procedures. Additionally, we would like to acknowledge with gratitude Dr. Theiler's valuable insights and recommendations.

APPENDIX

In this appendix we show that the filter described in Sec. II is an acausal, nonrecursive filter of finite order. Therefore, according to a theorem proved by Broomhead, Huke, and Muldoon [8] (see below), the filter in principle leaves invariant those quantities that can be obtained using embedding techniques.

Let $V = \{v_0, v_1, \dots, v_{N-1}\}$ be a series of real-valued measurements made on a dynamical system. We assume that the system evolves on an attractor \mathcal{A} , which is a sub-

set of a differentiable manifold \mathcal{M} of dimension, m . A nonrecursive filter applied to V produces a new time series, $U = \{u_0, u_1, \dots\}$, with

$$u_j = \sum_{k=0}^{n-1} a_k v_{j-k} \equiv \mathbf{a} \cdot \mathbf{v}_j. \quad (\text{A1})$$

This filter is *causal* if $k \geq 0$ and is said to be of *finite order* if n is finite.

The filter described in Sec. II uses the fast-Fourier-transform algorithm which calculates the discrete-Fourier-transform,

$$\mathcal{V}_n = \mathcal{D}[v_k] = \sum_{k=0}^{N-1} v_k e^{2\pi i k n / N}, \quad (\text{A2})$$

of the sequence, $V = \{v_0, v_1, \dots, v_{N-1}\}$. The inverse of Eq. (A2) is

$$v_k = \mathcal{D}^{-1}[\mathcal{V}_n] = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{V}_n e^{-2\pi i k n / N} \quad (\text{A3})$$

(see, e.g., Ref. 10). Note that the sequence $\{\mathcal{V}_n\}$ is periodic of period N .

Using Eqs. (A2) and (A3), the action of the filter of Sec. II may be written as

$$u_k = \mathcal{D}^{-1}[F_n \mathcal{V}_n], \quad (\text{A4})$$

where F_n is the filter function. Again using Eqs. (A2) and (A3), Eq. (A4) can be rewritten as

$$u_k = \sum_{j=0}^{N-1} a_{k-j} v_j \quad (\text{A5})$$

$$\text{with } a_m = \mathcal{D}^{-1}[F_k] = \frac{1}{N} \sum_{k=0}^{N-1} F_k e^{-2\pi i k m / N}.$$

The sequence $\{a_m\}$ is the (inverse) discrete Fourier transform of the filter function F_n . We shift the summation index in Eq. (A5) to get

$$u_k = \sum_{m=k}^{k-N+1} a_m v_{k-m} \equiv \mathbf{a} \cdot \mathbf{v}_k. \quad (\text{A6})$$

Comparison of Eqs. (A1) and (A6) shows that the filter defined in Sec. II is an acausal, nonrecursive filter of order N , which is finite, satisfying the hypotheses of the theorem of Broomhead, Huke, and Muldoon [8] which we restate here for completeness.

Let \mathcal{M} be a compact manifold of dimension m ; $\phi: \mathcal{M} \rightarrow \mathcal{M}$ a smooth diffeomorphism; $\nu: \mathcal{M} \rightarrow \mathcal{R}$ a smooth function; and the map, $\Phi_{(\phi, \nu)}: \mathcal{M} \rightarrow \mathcal{R}^{2m+1}$ defined by

$$\Phi_{(\phi, \nu)}(x) = (v_0 = \nu(x), v_{-1} = \nu(\phi^{-1}(x)), \dots, v_{-2m} = \nu(\phi^{-2m}(x)))$$

an embedding.

Theorem: Let V be a time series of measurements made on a dynamical system, (ϕ, \mathcal{M}) , satisfying the above conditions. Then, for triples (\mathbf{a}, ϕ, ν) it is a generic property that the method of delays, which constructs, from the time series U , vectors of the form

$$(u_j, u_{j-1}, \dots, u_{j-l+1})$$

$$\text{where } l \geq 2m + 1 \text{ and } u_j = \mathbf{a} \cdot \mathbf{v}_j,$$

where \mathbf{a} describes a nonrecursive filter of finite order n , gives an embedding of \mathcal{M} .

The proof of this theorem given in Ref. [8] requires n to be at least equal to $2m + 1$ so that the filter acts as a map of the original embedding.

-
- [1] R. Badii and A. Politi, in *Dimensions and Entropies in Chaotic Systems*, edited by G. Mayer-Kress (Springer-Verlag, Berlin, 1992).
- [2] R. Badii, G. Broggi, B. Derighetti, M. Ravani, S. Ciliberti, A. Politi, and M. A. Rubio, *Phys. Rev. Lett.* **60**, 979 (1988).
- [3] F. Mitschke, M. Möller, and W. Lange, *Phys. Rev. A* **37**, 4518 (1988).
- [4] P.-C. Lo and J. Principe, *Proc. IEEE Eng. Med. Biol. Soc.* **89**, 638 (1989).
- [5] L. O. Chua and T. Lin, *IEEE Trans. Circuits Syst.* **35**, 648 (1988).
- [6] F. Mitschke, *Phys. Rev. A* **41**, 1169 (1990).
- [7] K.-P. Yip, N.-H. Holstein Rathlou, and D. J. Marsh, *Am. J. Physiol.* **261**, F400 (1991).
- [8] D. S. Broomhead, J. P. Huke, and M. R. Muldoon, *J. R. Stat. Soc. B* **54**, 373 (1992).
- [9] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J. D. Farmer, in *Nonlinear Modeling and Forecasting*, edited by M. Casdagli and S. Eubanks (Addison-Wesley, Reading, MA, 1992).
- [10] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
- [11] D. E. Knuth, *Seminumerical Algorithms* (Addison-Wesley, Reading, MA, 1981).
- [12] P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983).
- [13] P. Grassberger and I. Procaccia, *Phys. Rev. D* **13**, 34 (1983).
- [14] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp, *Phys. Rev. A* **38**, 3017 (1988).
- [15] A. M. Albano, A. Passamante, and M. E. Farrell, *Physica D* **54**, 85 (1991).
- [16] W. Liebert and H. G. Schuster, *Phys. Lett. A* **142**, 107 (1988).
- [17] J. F. Gibson, J. D. Farmer, M. Casdagli, and S. Eubank, *Physica D* **57**, 1 (1992).
- [18] W. E. Caswell and J. A. York, in *Dimensions and Entropies in Chaotic Systems*, edited by G. Mayer-Kress (Springer, Berlin, 1986).
- [19] M. Möller, W. Lange, F. Mitschke, N. B. Abraham, and U. Hübner, *Phys. Lett. A* **138**, 176 (1989).
- [20] A. M. Albano, A. I. Mees, G. C. deGuzman, and P. E. Rapp, in *Chaos in Biological Systems*, edited by H. Degn, A. V. Holden, and L. F. Olsen (Plenum, New York, 1987).
- [21] L. A. Smith, *Phys. Lett. A* **133**, 283 (1988).
- [22] P. Grassberger, T. Schreiber, and C. Schaffrath, *Int. J. Bi-*

- furcation Chaos **1**, 521 (1991).
- [23] J.-P. Eckmann and D. Ruelle, *Physica D* **56**, 185 (1992).
- [24] J. Theiler, *Phys. Rev. A* **34**, 2427 (1986).
- [25] J. A. Scheinkman and B. LeBaron, *J. Business* **62**, 311 (1989).
- [26] S. Ellner, *Phys. Lett. A* **133**, 128 (1988).
- [27] R. L. Somorjai and M. K. Ali, *Can. J. Chem.* **66**, 979 (1988).
- [28] J. B. Ramsey and H.-J. Yuan, *Phys. Lett. A* **134**, 287 (1989).
- [29] K. Judd, *Physica D* **56**, 216 (1992).
- [30] W. A. Brock, *J. Econ. Theory* **40**, 168 (1986).
- [31] W. A. Brock and W. E. Dechert, in *Measures of Complexity and Chaos*, edited by N. B. Abraham, A. M. Albano, T. Passamante, and P. E. Rapp (Plenum, New York, 1989).
- [32] W. Dechert (unpublished).
- [33] W. A. Brock and S. M. Potter, in *Nonlinear Modeling and Forecasting*, edited by M. Casdagli and S. Eubank (Addison-Wesley, Reading, MA, 1992).
- [34] P. Grassberger, *Nature (London)* **323**, 609 (1986).
- [35] W. A. Brock, *J. Econ. Theory* **40**, 168 (1986).
- [36] W. A. Brock, W. D. Dechert, and J. A. Scheinkman (unpublished).
- [37] A. R. Osborne, A. D. Kirwin, A. Provenzale, and L. Ber-gamasco, *Physica D* **23**, 75 (1986).
- [38] J. Theiler, Ph.D. thesis, California Institute of Technology, 1988.
- [39] S. Elgar and G. Mayer-Kress, *Physica D* **37**, 104 (1989).
- [40] J. Theiler, *Phys. Rev. A* **41**, 3038 (1990).
- [41] D. T. Kaplan and R. J. Cohen, *Circ. Res.* **67**, 886 (1990).
- [42] J. P. M. Pijn, Ph.D. thesis, University of Amsterdam, 1990.
- [43] J. D. Farmer and J. J. Sidorowich, in *Evolution, Learning and Cognition*, edited by Y. C. Lee (World Scientific, Singapore, 1988).
- [44] M. Casdagli, *J. R. Stat. Soc. B* **54**, 303 (1991).
- [45] M. Sano and Y. Sawada, *Phys. Rev. Lett.* **55**, 1082 (1985).
- [46] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, *Physica D* **16**, 285 (1985).
- [47] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, and S. Ciliberto, *Phys. Rev. A* **34**, 4971 (1986).
- [48] D. T. Kaplan and L. Glass, *Phys. Rev. Lett.* **68**, 427 (1992).
- [49] In the interest of producing a brief publication, the details of dimension calculations reported in Tables I and II are not presented for other data sets. These tables have been prepared and can be obtained from the authors.
- [50] P. E. Rapp, *The Biologist* (to be published).