Missing motor of on-off intermittency

J. Graf von Hardenberg,[1] F. Paparella,[1] N. Platt,[2] A. Provenzale,[1] E. A. Spiegel,[3] and C. Tresser[4]
[1]*Istituto di Cosmogeofisica, 10125 Torino, Italy*
[2]*Naval Surface Warfare Center, Code R44, Silver Spring, Maryland 20903-5000*
[3]*Department of Astronomy, Columbia University, New York, New York 10027*
[4]*IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598*
(Received 22 April 1996)

We illustrate the failure of standard algorithms to predict correctly the dimension of some composite maps that produce on-off intermittency. Examination of the phase portraits of some examples suggests that the culprit is a nearly singular density of points close to the origin and we describe ways to refine the algorithms so as to obtain correct results. [S1063-651X(96)03212-6]

PACS number(s): 05.45+b

I. INTRODUCTION

Given an irregular looking time series, it is now standard procedure to construct from it a phase portrait by the method of delay coordinates [1]. It is then possible to make this picture more quantitative by measuring some statistical properties of the trajectory such as the correlation dimension [1,2]. This may be done using the Grassberger-Procaccia (GP) algorithm [3], in which one determines the slope of a logarithmic plot of the correlation integral vs the separation for increasing values of the dimension D_e of the reconstructed phase space. If, for increasing values of D_e, the slopes converge to a finite value D_2, then this is a useful approximation to the desired dimension.

This procedure is normally quite successful when applied to laboratory experiments, though there may be problems when there are not sufficiently abundant data [4]. On the other hand, even for large and adequate data sets, situations do arise in which standard procedures do not give correct dimensional estimates [5–7]. What is disturbing is that in some of these cases the procedures themselves give no indication that all is not well, even when the so-called scaling curves look quite healthy. In such cases, more subtle approaches are needed and we discuss here an instance of this problem.

Our example is a form of intermittency that was first studied in fluid turbulence. The term intermittency was introduced to describe signals that alternated between flat portions and bursting ones, interpreted as the passage of laminar and turbulent states of the fluid passing by a probe [8]. In chaos theory, the term has been generally used to refer to any kind of alternation of behavior. Hence is has been useful to introduce the retronym *on-off intermittency* [9] when the original technical sense is intended, in which there is abrupt switching from periods of stasis to bursts of large variation and back again. Many simple models of this kind of behavior have been proposed and studied in the past fifteen years (e.g., [10–15]).

Most of the models of on-off intermittency have in common a single mechanism in which a nonlinear oscillator has a time-dependent control ''parameter'' $\mu(t)$ that determines its stability. The wealth of this time dependence determines the complexity of the consequent behavior. When the control parameter is a function of a variable generated by a chaotic or a stochastic system, we may encounter rich and intermittent oscillations.

To be more specific, suppose that the original oscillator operates in a phase space $\mathcal{S}$ with dimension D_X when the control parameter μ is constant in time. We assume that when μ is below a certain critical value μ_0, the oscillator is quiescent and the system is dominated by an attractor of low dimension, say a fixed point. For $\mu>\mu_0$, the oscillator becomes time dependent. Now let μ be time dependent and suppose that its time dependence is determined by a second system, the driver, which, when operating autonomously, occupies a phase space $\mathcal{D}$ with dimension D_Y. The total system in $\mathcal{S}\oplus\mathcal{D}$ has a phase space dimension $D_{\rm tot}=D_X+D_Y$ and it may be dominated by an attractor with correlation dimension D_2. In the cases of interest here, we observe only the dynamics of the forced oscillator, but wish to know something about the combined system, such as $D_{\rm tot}$ and/or D_2.

Our earlier studies of systems of this kind have suggested that the conventional dimension-finding algorithms give the same results, independently of the dimension of the driver. This outcome occasioned surprise on the part of those who quoted a theorem of Takens [16,17], suggesting that the standard procedures should give the dimension of the combined system correctly. However, in the fine print of this theorem there is the assumption that the coordinate that is measured is generic. When only one coordinate in $\mathcal{S}$ is measured, this assumption may not apply. On the other hand, there were those, especially Grassberger [private communication], who admitted the breakdown of the standard method, but felt that, *in principle*, it should be possible to tell something about the nature and dimension of the driver from a time series of the coordinate of the oscillator. The aim of this paper is to report on our efforts to learn the answer to the implied question by devising a sufficiently discriminating method to do the job. Such a method could also be useful in other kinds of problem and it seems worthwhile to describe the results of the search for it.

II. DETECTING THE DIMENSION OF ON-OFF INTERMITTENCY

A. Models

To illustrate the issue we are discussing we use the map X: $[0,1]\to[0,1]$ (described in [9]):

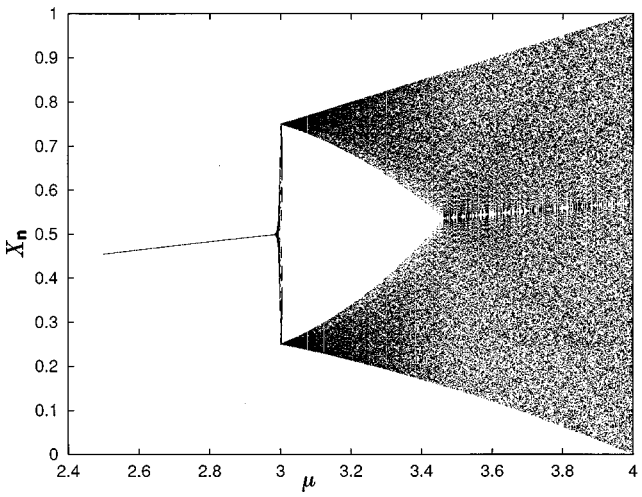

FIG. 1. Bifurcation plot of the map (1).

$$X_{n+1}=\begin{cases}\mu X_n & \text{if } 0\leqslant X_n\leqslant \frac{1}{4}\\ \frac{1}{3}\mu(1-X_n) & \text{if } \frac{1}{4}<X_n\leqslant 1.\end{cases} \tag{1}$$

This map always has the fixed point $X_0=0$, which is stable for $\mu<1$ and unstable for $\mu>1$. When $\mu>1$ there is a second fixed point $X_1=\mu/(\mu+3)$, which is stable for $\mu\leqslant 3$ and unstable otherwise. In Fig. 1, we show a bifurcation diagram for this system.

We are interested in seeing what happens when μ is allowed to depend on time; that is, we replace it in (1) by μ_n. We could, of course, begin with a simple time dependence for μ, and such matters have been widely studied elsewhere. Here we want to concentrate on the situation where μ_n is determined by a second dynamical system (the driver) Y: $[0,1]\rightarrow[0,1]$, which will produce an irregular sequence of values either chaotically or stochastically. Specifically, we shall work with the two drivers: $Y_{n+1}=2\,Y_n \bmod 1$ and Y_{n+1}, which is white noise uniformly distributed in the region $[0,1[$. The first map is sometimes called the *doubling map*, with $D_Y=1$; it is related to the so-called shift map. The noise algorithm has (in effect) $D_Y=\infty$.

We determine μ_n from the maps Y in this simple way:

$$\mu_n=aY_n, \tag{2}$$

where a is a constant to be specified (typically between 0 and 4).

In the present work we do not let Y experience any influence from X. This type of coupling gives the combined system that is called a *skew product structure*. This permits us to find interesting parameter regimes more readily than with feedback. We are interested in particular in parameter choices for which the output of X shows an alternation between quiescent and bursting behavior. Generally, we choose to fix the system Y and change a in order to obtain behavior for the signal $\{X_n\}$ ranging from no intermittent bursts at all ($X_n\simeq 0$ always), when a is below a critical value $a_c<2.7$, up to very dense bursts when a approaches the limiting value 4. (For $a>4$, X develops a repeller and we shall not investigate that behavior here.)

To give an early indication of the problem of interest we show in Fig. 2 the time series X_n obtained with two different values of a, for both the doubling map (DM) and white noise (WN) as drivers. What we find striking is that the low-dimensional doubling map and the noise give very similar looking results, which leads us to ask the question considered here: Can we from the time series X_n distinguish between the results produced by the two drivers without any further information about the system Y or the bare control parameter a? It has already been observed that this would be a formidable task as the system X is by its nature an amplifier of error and effectively masks the input from Y. In the rest of this section we describe how the standard procedure for determining the dimension of the system fails.

B. Correlation integrals

The two versions of $X\oplus Y$ evidently have completely different dimensions, so we naturally want to learn whether we could detect this by computing dimensions from a time series of the driven system X_n. For example, we have computed the correlation dimension of our data series using the familiar GP algorithm [3]. This we did in the usual manner by the time-delay procedure [16], constructing a set of D_e-dimensional vectors

$$\boldsymbol{X}_n=\{X_n, X_{n+1}, \ldots, X_{n+D_e-1}\}, \tag{3}$$

where D_e is the embedding dimension introduced earlier. Then we computed the correlation integral

$$C(r)=\frac{2}{N(N-w)}\sum_{i=1}^{N}\sum_{j=i+w}^{N}\Theta(r-|\mathbf{X}_i-\boldsymbol{X}_j|), \tag{4}$$

where Θ is the Heaviside step function and N is the length of our data series. The parameter w was introduced by Theiler [18] to make a distinction among close-lying vectors to separate those that are simply near each other because they are close in time from those that arise from close returns in the embedding space. We have varied w to see whether there is any significant change in going through the values one, ten, and a thousand and found none. This is to be expected for a map; Theiler's parameter has more of a role for flows. For all the results reported here, we use the value $w=10$.

As r tends to zero, we expect $C(r)$ to vanish also and we look for a scaling like $C(r)\propto r^{\nu}$ in the limit of small r. Grassberger and Procaccia suggested using ν to estimate the correlation dimension D_2, an approximation that works well within the range where scaling holds. Within that region, we expect ν to increase to a maximum value close to D_2 as D_e increases through D_{tot}.

What we are calling white noise in this discussion is really a signal from an algorithm that represents a system of very large dimension, too large to be of direct interest. Hence, for white noise, ν should not saturate as D_e is increased, whereas for a deterministic (that is, low order) system, the driver is of low dimension and can be expected to saturate at a limiting value D_2 as when D_e increases.

In practice, these determinations are carried out by plot-

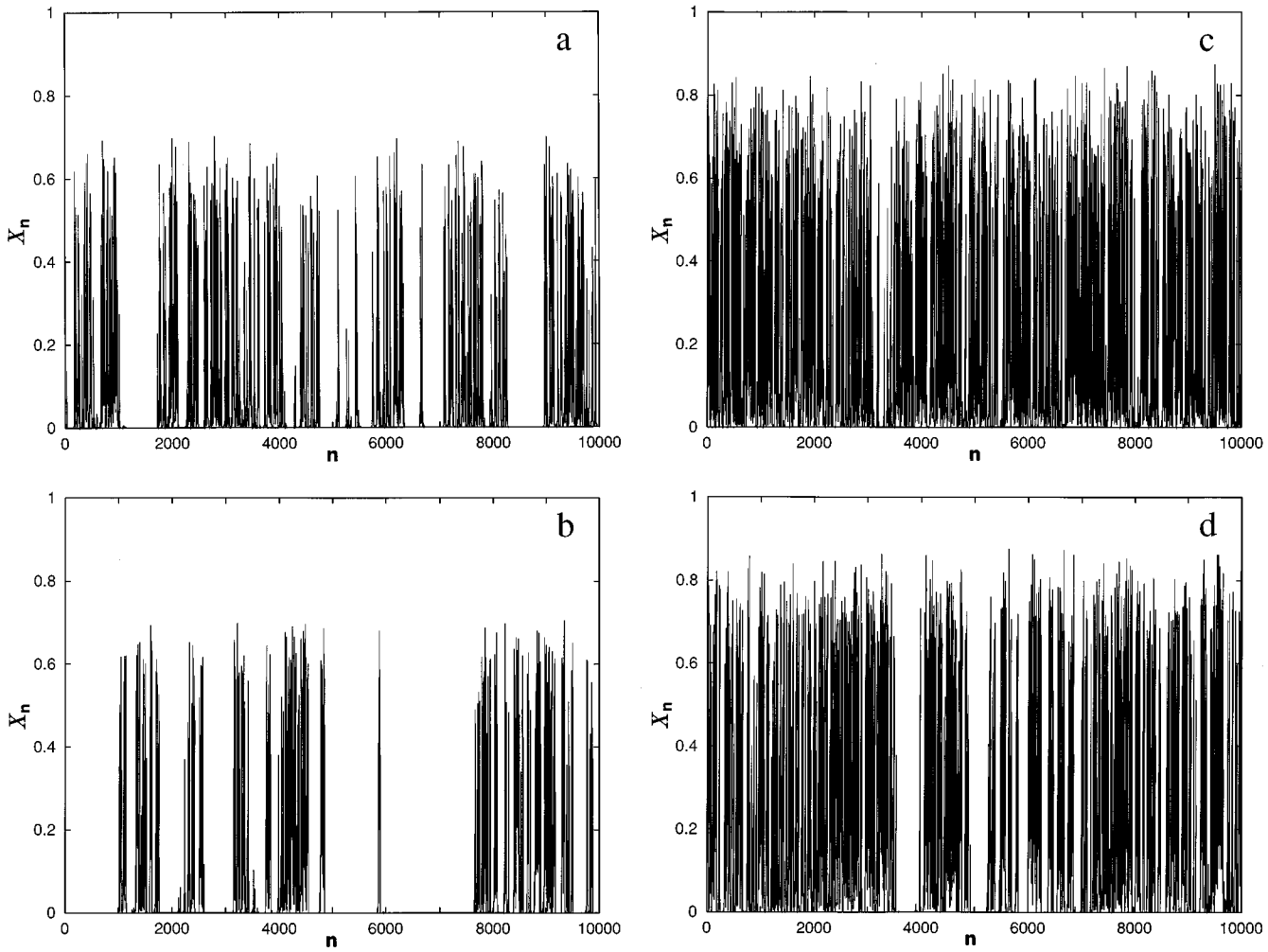

FIG. 2. Data series: (a) white-noise forcing with $a=2.88$, (b) doubling-map forcing with $a=2.88$, (c) white-noise forcing with $a=3.54$, and (d) doubling-map forcing with $a=3.54$.

ting $\log_{10}C(r)$ vs $\log_{10}r$ and determining the slope ν of this plot. In principle, we would determine this slope in the limit $r\to 0$, but, in fact, there is a lowest scale below which the scaling is not good and we must select some range of r over which to look at the fractality. When we proceed in this standard way, we encounter the problem that motivates the present work: ν saturates to a finite value for *both* the deterministic and the stochastic drivers. This is illustrated in Fig. 3, which shows the correlation integrals for the four time series shown in Fig. 2, where the maximum embedding dimension is $D_e=12$. Figure 4 shows the value of the correlation dimension as a function of the coupling strength a, for both the DM and the WN forcing. The dimensions have been computed as an average over ten realizations of the process; error bars are 1σ deviations.

We observe that D_2 is always very small ($D_2<1.5$) and that for both kinds of forcings there are some ranges of a that give the same correlation dimension. This means that given only the two signals, without any other information, one may not be able to distinguish the two cases and could, in any event, erroneously conclude that the signal forced by white noise comes from a total system of finite dimension. This failure of the standard procedure is not alleviated even when we use longer data series, as we have confirmed with a series of 10^5 points.

In making comparisons between runs using the deterministic and the stochastic driving we shall refer to the values of the coupling parameter a for the white noise as a_r and for the deterministic case as a_d. Aiming at some standardization, we typically choose a_r and a_d so as to give the same correlation dimension estimate in the naive application of the GP algorithm to the data series. From the plot in Fig. 4 we see that such a pair of values is $a_r=2.88$ and $a_d=3.54$.

In summary, we find that the standard approach cannot be relied on to give the correct dimension in the case of on-off intermittency. It does not even give an indication that there is something amiss. We have also investigated whether other techniques, such as calculating power spectra and autocorrelations, may help in this respect. While these showed interesting features in themselves, they did not prove helpful in this problem. Likewise, we tried the procedure of phase randomization [6,7]. For both the deterministic and the stochastic drivers, the dimension of the phase-randomized signal grows without bound, giving the false impression of a deterministic nature of both time series.

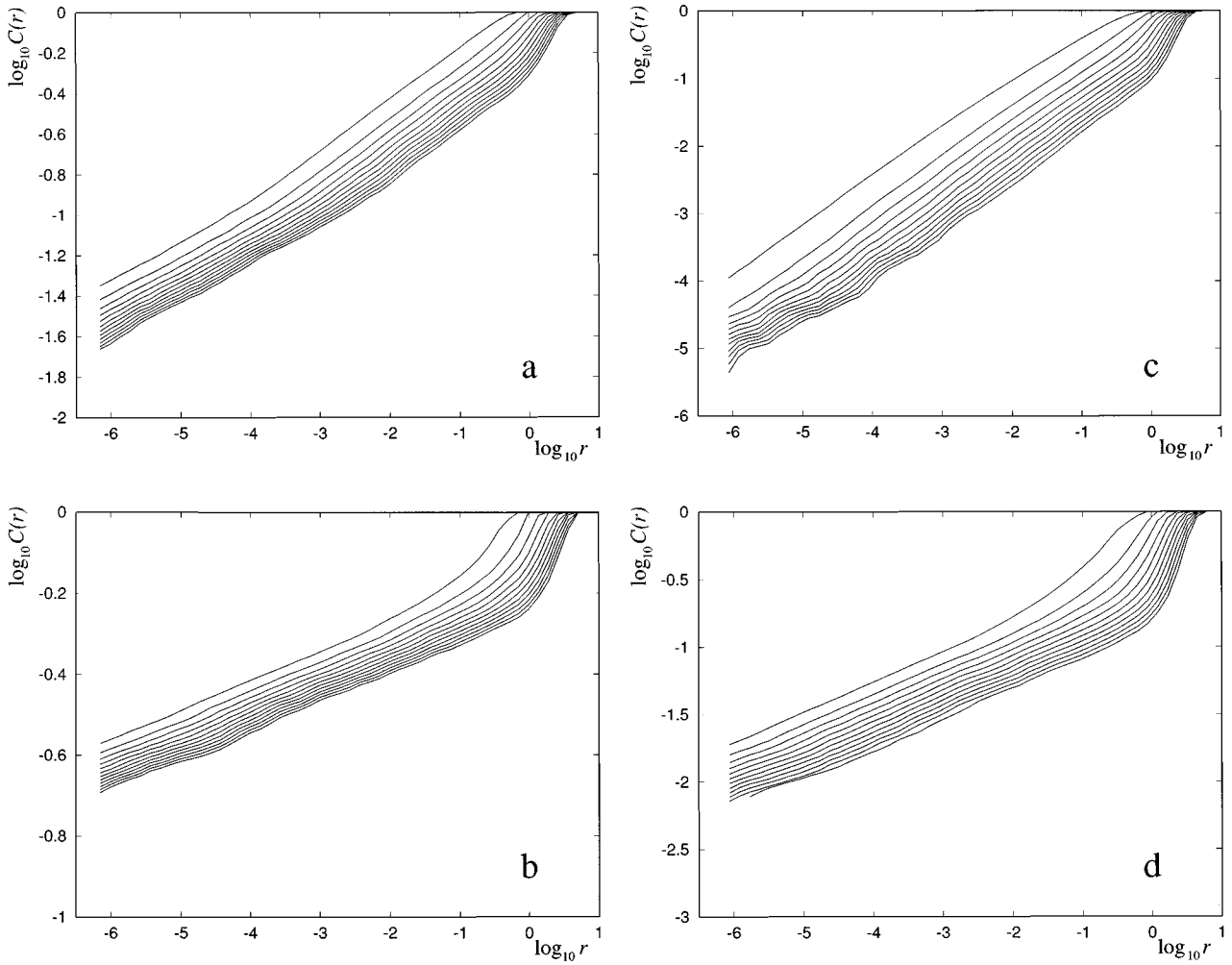

FIG. 3. Plots of $\log_{10}C$ vs $\log_{10}r$ for the various cases with the sequence of Fig. 2. Each curve refers to a different embedding dimension D_e; here $1 \leqslant D_e \leqslant 12$ and $w = 10$.

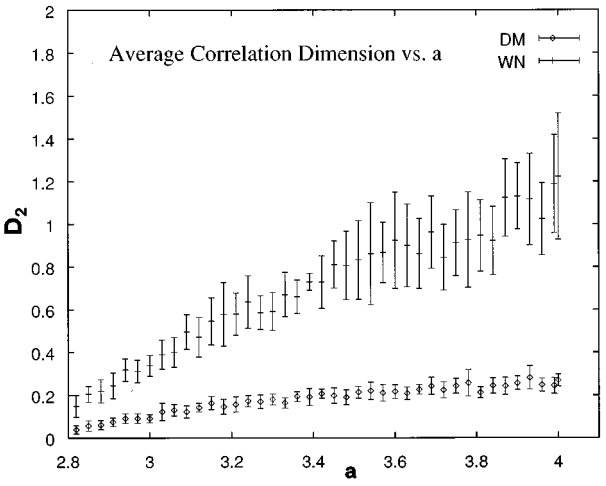

FIG. 4. D_2 vs a from the GP algorithm averaged over ten realizations for forcing by both noise and doubling map. Error bars are 1σ deviations.

III. MODIFIED APPROACHES

A. Phase portraits

In seeking a clue about how to proceed to discover the dimension of the system, we naturally examine the plots of X_{n+1} vs X_n. As we see from such plots in Figs. 5(a) and 5(b), the output that results when we use the doubling map as the driver has dense veils of points that do not appear when we drive with white noise. This distinction between the outputs produced when using the two drivers is even more striking on the plots of X_{n+2} vs X_n in Figs. 5(c) and 5(d). The output using the doubling map has very distinct structures that do not appear when we drive with white noise.

We naturally surmise that the appearance of distinct structures in the case of the doubling map is the signature of low-dimensional behavior and we want to take advantage of the differences in Fig. 5 to fashion a scheme of dimension detection. On the other hand, we are somewhat surprised that despite the absence of structures in the white-noise case, the GP algorithm produces low dimensions. A clue to the commonality of low-dimensional results in the two cases, which we see in looking at the plots, is the high density of points near the origin in both cases. These occur because of the ''off'' cycles in the data series, during which X_n remains

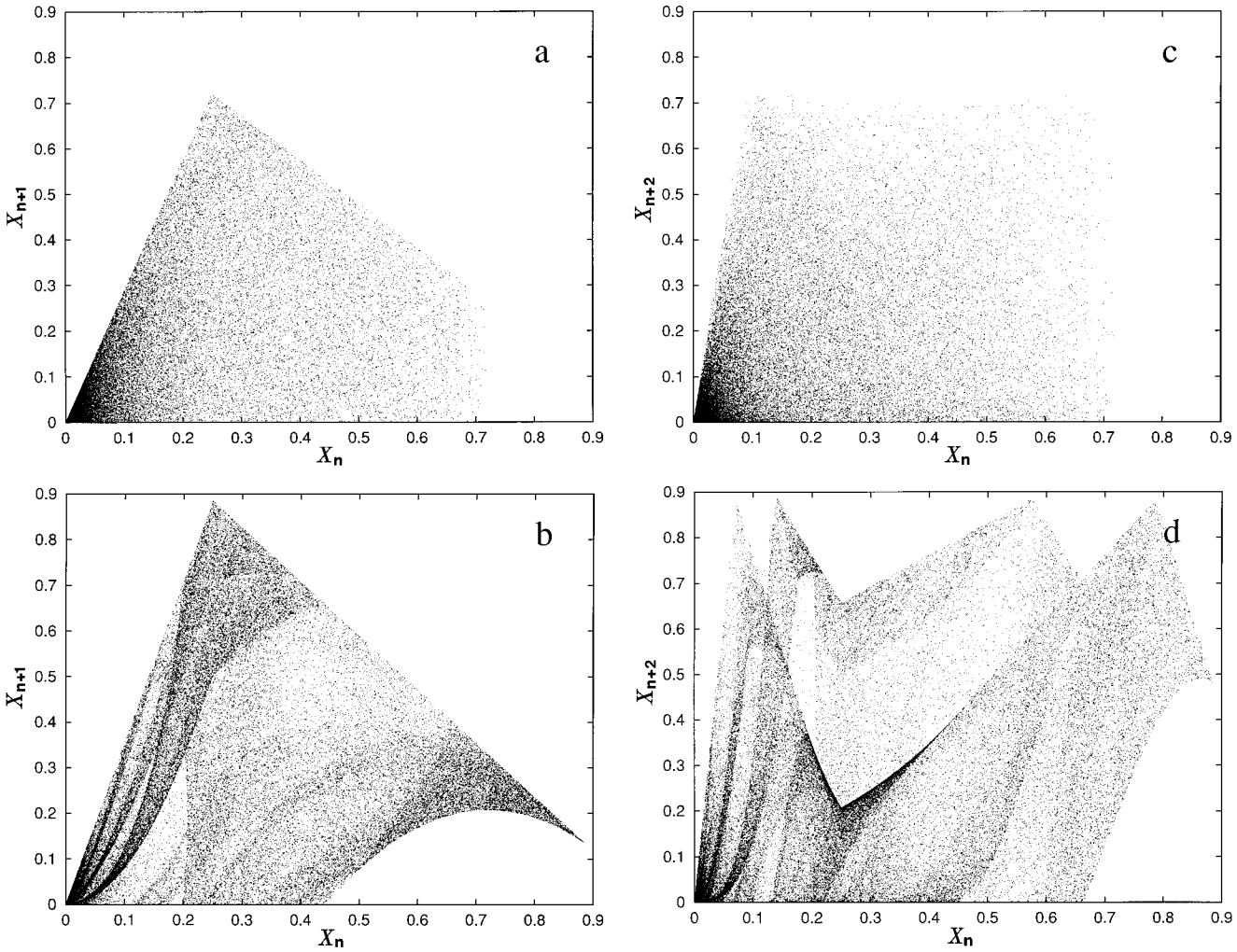

FIG. 5. Phase portraits: X_{n+1} vs X_n for (a) white noise and (b) doubling map and X_{n+2} vs X_n for (c) WN and (d) DM.

very close to zero for noticeable stretches of time. These weigh heavily in the dimension determinations. Such structures are reminiscent of self-similar excursions from the quiescent phase seen by Brooke and Moss [19] in a model of the solar dynamo and underlie the scaling analysis of Heagy, Platt, and Hammel [20] at the transition to on-off behavior.

The significance of all this is clearer when we recall that the plots in Fig. 5 may be thought of as delay reconstructions in two-dimensional embedding spaces with delays corresponding to $\Delta n=1$ or 2. Since the GP algorithm works by averaging on all the vector pairs in the embedding space, it may give misleading results when the scaling properties of the structures are as inhomogeneous as they appear in Fig. 5. In particular, we see in Fig. 5 that the typical scale of the structures produced by forcing with the doubling map becomes smaller as we approach the origin in the appropriate figure. This is a result of the alternations between bursting and quiescent periods.

The vectors are nonuniformly distributed over a range in lengths from about 10^{-30} up to 1. The underlying structure whose correlations might be important for the actual dimension determination is blurred because the lengths of the reconstructed vectors are not adequately resolved. A similar phenomenon may be expected to occur in higher embedding dimensions. The techniques we suggest next were devised with these features of the results in mind.

B. Approaches

1. Rescaling the data

The statistical properties of the phase portrait of our system are not changed by a smooth transformation of the coordinate X as the analyses of Takens [16] and of Sauer, Yorke, and Casdagli [17] show. Hence we choose to remap our data so as to bring out better the behavior near the origin. We introduce a function $M(X)$ designed to diminish the range of scales in the data. Two simple examples that work well are $M(X_n)=\log_{10}(X_n)$ and $M(X_n)=X_n^p$, the latter being useful for $0<p\ll 1$. The interest of using logarithmic variables for on-off intermittency has already been noted [20,21] and the case of p going to zero is related to that mapping. For the present purpose, it must be noted that these maps are not smooth at the origin, so we cannot cite the Takens theorem [16] to justify their use. Nevertheless, they do work quite well in practice as the singularity right at the origin does not seem to play any role.

The estimated correlation dimensions ν vs D_e for the two kinds of forcing and for the two scaling functions are shown

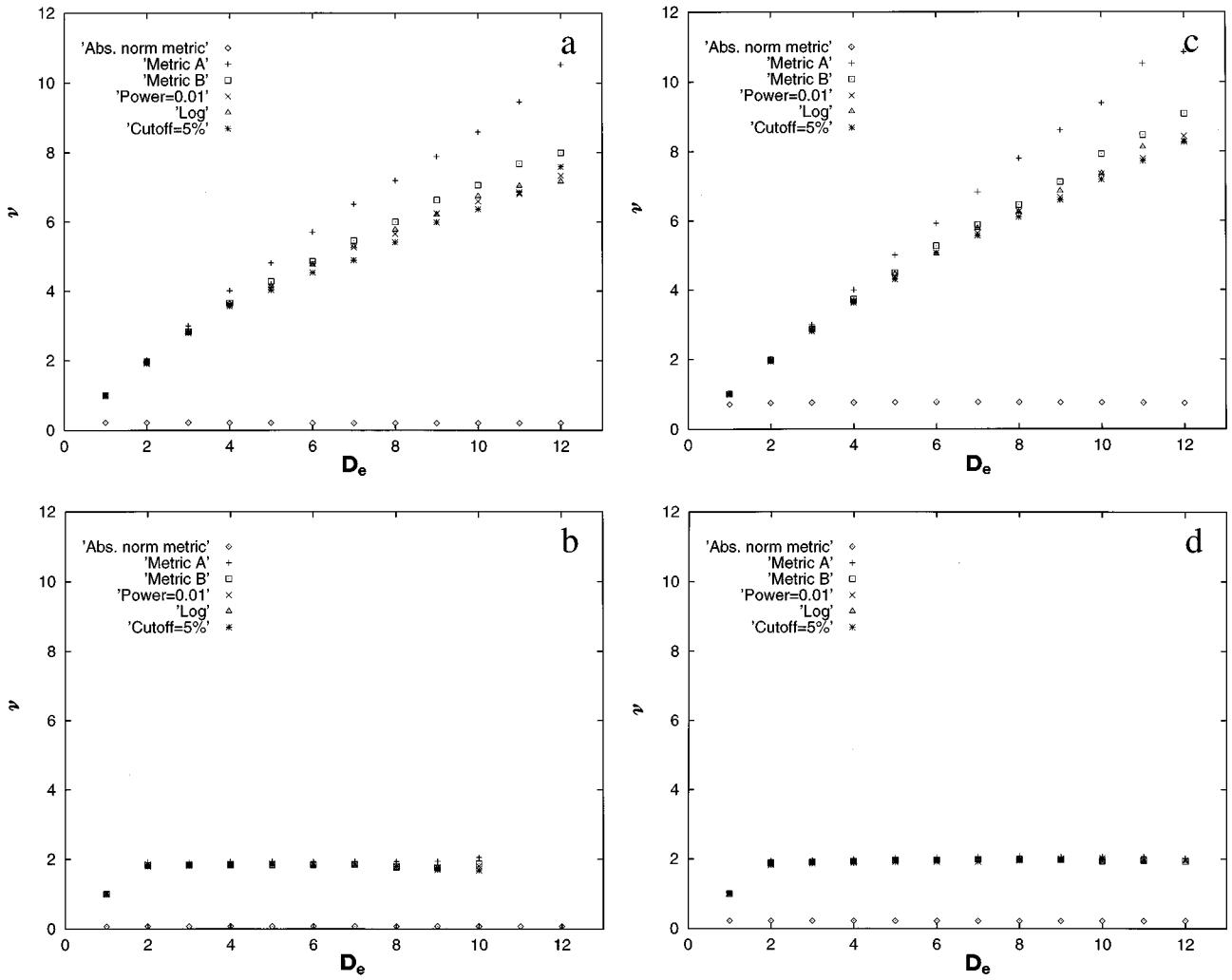

FIG. 6. ν vs D_e for the various cases with the sequence of Fig. 2. The different symbols refer to the various methods employed to calculate the correlation dimension. Open diamonds refer to the original results obtained when using the absolute distance. Open triangles and crosses refer to rescaling the data by, respectively, a logarithm and a power law with $p=0.01$; pluses and open squares refer to using the distances indicated, respectively, by Eqs. (6) and (7); stars refer to filtering the data by method (ii).

in Fig. 6. We can see how, using the small value $p=0.01$ or using the $\ln(X_n)$, we may readily distinguish the two kinds of forcing. We now get a convergence of ν to $D_2 \simeq 2.0$ for the doubling map and no convergence at all for the white noise.

2. Modifying the distance in the embedding space

An approach that is related to that of rescaling the data is that of changing the distance used in the embedding space. Heretofore, we have used the *absolute distance*

$$d(\mathbf{X},\mathbf{Y}) \equiv \sum_{i=1}^{D_e} |X^{(i)} - Y^{(i)}|, \tag{5}$$

where $X^{(i)}$ is the ith component of a vector $\mathbf{X}$ in the embedding space [see definition (3)]. To bring out the structures near the origin we have used several distances, including

$$d_A(\mathbf{X},\mathbf{Y}) \equiv \sum_{i=1}^{D_e} 2\frac{|X^{(i)} - Y^{(i)}|}{|X^{(i)}| + |Y^{(i)}|}, \tag{6}$$

$$d_B(\mathbf{X},\mathbf{Y}) \equiv \sum_{i=1}^{D_e} \frac{|X^{(i)} - Y^{(i)}|}{\max(|X^{(i)}|, |Y^{(i)}|)}. \tag{7}$$

With these two distances we obtained the results shown in Fig. 6. For the case forced with the doubling map, the correlation dimension converge to $D_2 \simeq 2.0$, while for forcing with white noise, convergence does not occur. This method works well, unless one wants to calculate D_0.

3 Filtering the data

As we have seen for the X_n vs X_{n+k} $(k=1, 2, \ldots)$ plots, the failure to resolve the structures near the origin results in the inability of some algorithms to discriminate among the various true dimensions in the system producing the signal. One way to remove this difficulty might be simply to filter out the points near the origin. We have tried various schemes of this kind and mention two here.

(i) Eliminate all the *data points* $X_n < C$ from our data series, which leaves out all the ''off'' segments.

(ii) Exclude all the reconstructed vectors **X** having *at least* one component $X^{(i)}<C$.

The first of these methods introduces some glitches where some welding of data segments is needed, but these do not seem to be very significant. This feature does not appear in the second method. Figure 6 shows the correlation dimension obtained for the second method using a cutoff value $C=5\%$ of the average signal amplitude.

It is interesting that there are forms of filtering that let us get at the dimensions although they force us to give up data. This is not a serious deficiency except in systems with very long off periods, in which case it would take prohibitively long to acquire enough data to make accurate dimension determinations.

IV. CONCLUSION

The case of on-off intermittency has revealed a situation where the standard algorithm for estimating dimension fails, while giving no warning of its failure. Since this form of intermittency is robust and fairly commonplace, we have felt that this issue needs to be addressed. There are already known situations where the present considerations may be helpful as in the case of the cosmic x-ray source known as the rapid burster and in certain examples of chemical chaos.

More importantly we see that a failure of resolution has led to the breakdown of the standard procedure. This has to do with the assumption that the inner (and possibly outer) scales of the fractal can be safely assigned. We learn from this that, when there are doubts of validity, the use of diverse definitions of distance in the dimension algorithms should routinely be introduced as a safeguard against false results. A typical sign of possible trouble is the hovering of the signal near a constant value. A slow drift through such a value, as in the Pomeau-Manneville intermittency, does not generally indicate trouble, as the drift is enough to lift the degeneracy. It has been suggested that a histogram of the data would reveal the singularity and that this should be adequate warning that dimension estimate might be problematic.

We see from the filtering example that the dimensions themselves do not tell as much of the story as one would like to know and other forms of information are needed. In turbulence theory, one uses a parameter called the intermittency factor measuring the fraction of time the system is bursting. We have found that this is not a reliable indicator of the nature of the data as its value is sensitive to the threshold of bursting that we use.

Another quantity that is much used in analyzing chaotic systems is the Lyapunov exponent [22,23]. As for the case of the dimensions, the distinction between the two kinds of drivers in on-off intermittency is hard to make with normal procedures. Signs of any difference between the two cases are not reliably found. However, the introduction of a suitable distance, as for the case of dimensions, or a suitable scaled variable makes the distinction in Lyapounov exponent readily detectable. For the present, it may suffice to conclude by repeating that in applying standard algorithms for dimension determination one should check to see whether the results are invariant under coordinate transformation and changes of the distance in the reconstructed phase space.

[1] *Coping with Chaos*, edited by E. Ott, T. Sauer, and J.A. Yorke (Wiley, New York, 1994).
[2] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, and L.Sh. Tsimring, Rev. Mod. Phys. **65**, 1331 (1993).
[3] P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).
[4] L.A. Smith, Phys. Lett. A **133**, 283 (1988).
[5] A.R. Osborne and A. Provenzale, Physica D **35**, 357 (1989).
[6] A. Provenzale, L.A. Smith, R. Vio, and G. Murante, Physica D **58**, 31 (1992).
[7] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J.D. Farmer, Physica D **58**, 77 (1992).
[8] A.A. Townsend, *The Structure of Turbulent Shear Flow* (Cambridge University Press, Cambridge, 1967).
[9] N. Platt, E.A. Spiegel, and C. Tresser, Phys. Rev. Lett. **70**, 279 (1993).
[10] E.A. Spiegel, Ann. N.Y. Acad. Sci. **617**, 305 (1980).
[11] H. Fujisaka and T. Yamada, Prog. Theor. Phys. **74**, 919 (1985).
[12] L. Yu, E. Ott, and Q. Chen, Phys. Rev. Lett. **65**, 2935 (1990).
[13] A.S. Pikovsky and P. Grassberger, J. Phys. A **24**, 4587 (1991).
[14] N. Platt, E.A. Spiegel, and C. Tresser, Geophys. Astrophys. Fluid Dyn. **73**, 147 (1993).
[15] E. Ott and J.C. Sommerer, Phys. Lett. A **188**, 39 (1994).
[16] F. Takens, in *Dynamical Systems and Turbulence, Warwick, 1980*, edited by D. Rand and L.-S. Young, Lecture Notes in Mathematics Vol. 898 (Springer, Berlin, 1981), p. 366.
[17] T. Sauer, J.A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).
[18] J. Theiler, Phys. Rev. A **34**, 2427 (1986).
[19] J.M. Brooke and D. Moss (unpublished).
[20] J.F. Heagy, N. Platt, and S.M. Hammel, Phys. Rev. E **49**, 1140 (1994).
[21] N. Platt, S.M. Hammel, and J.F. Heagy, Phys. Rev. Lett. **72**, 3198 (1994).
[22] J.P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).
[23] A. Wolf, in *Chaos*, edited by A.V. Holden (Manchester University Press, Manchester, 1986), p. 273.