Discontinuous and Nondifferentiable Functions and Dimension Increase Induced by Filtering Chaotic Data

Louis M. Pecora and Thomas L. Carroll

Code 6341

Naval Research Laboratory

Washington, DC 20375

Abstract

We show that one can use recently introduced statistics for continuity and differentiability to show the effect of filters of infinite extent in time on a chaotic time series. The statistics point to a discontinuous or nondifferentiable function between the unfiltered attractor and the filtered attractor as the origin of attractor dimension increase when the filtering is severe. The density of discontinuities as a function of resolution follows a scaling relation. We present direct visualization of this effect in the filtered Henon attractor where the origin of dimension increase becomes obvious.

PACS NO. 05.45.+b, 47.52.+j, 06.50.-x, 02.40.-k, 02.50.-r

Final version submitted to CHAOS 2 Apr 1996.

Reference for this paper:

L. Pecora and T. Carroll, CHAOS 6, 432-439 (1996).

I. Introduction

In this paper we show that recently devised statistics for testing continuity and differentiability [1] can be useful in testing for damage to a chaotic time series from the application of a filter. We first outline past explorations into the filtering of chaotic time series. We then introduce the statistics and show their application to filtered data sets. The results suggest a way to view filtering of a time series which explains why any filter of infinite extent in time has potential to increase the dimension of a chaotic attractor.

Filtering of time-series data is common-place in signal processing, but filtering of data from a chaotic system can lead to a time-series that does not convey the correct information about the original system. This was first demonstrated by Badii *et al* [2] using a low-pass dynamical filter, commonly called a linear, time-invariant filter (LTI). The LTI filter is given by

$$\frac{dz}{dt} = -\eta z + x_1 , \tag{1}$$

where z is the result of filtering and x_1 is a component of the dynamical system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, i.e. $x_1$generates the time-series that is filtered by weighted averaging into the past.

Badii *et al* showed that when the time series is highly filtered (small η values) the attractor reconstructed from the filtered time series z has a dimension that is larger than that of the original attractor. They went on to show that because of the dynamic nature of the filter the dimension increase could be correlated with an increase in the Kaplan-York dimension [3]. Later Badii *et al* showed that one can also view the dimension increase as a phase transition, using a thermodynamics approach to dimensions [4].

Chennaoui *et al* showed a mechanism for dimension increase for the particular case of an LTI filter acting on the logistic map [5]. For the logistic map parameter value of r=4, they obtained analytic results. Their work showed that in this particular case the filter induces self-similarity in the density of points in the two-dimensional space of time-series points and filtered values.

Other work by Mitschke [6] showed that acausal filters (which average symmetrically into the future as well as the past) may do little or no damage to some chaotic time series even for the highly filtered case. Mitschke suggested that the phase-preserving nature of the acausal filter kept the dimensions invariant under the filtering. This phase criteria remains unclear.

Isabelle [7], basing his conclusions on the Lipshitz condition, suggested that problems would ensue with LTI filters when the filter parameter η exceeded the minimum Lyapunov exponent of the system. This is essentially correct and is related to the Kaplan-Yorke dimension. We will show that this can occur for filters other than LTI filters.

Finally, Broomhead *et al* [8] showed that filters of finite extent in time, for certain cases called finite impulse response (FIR) filters, would, in principle, induce a diffeomorphism between the original attractor reconstructed from the signal (x_1) and that reconstructed from the filtered signal (z). The LTI filters like Eq. (1) and Mitschke's acausal filter are infinite in extent or infinite impulse response (IIR) filters and are not necessarily safe to use. Broomhead *et al* did point out that in practice FIR filters with a very long time window may act like IIR filters and for finite precision data cause problems, too.

In all the forgoing research what becomes clear is the need for a criterion to help determine when a filter is safe to apply to a data set or use during data collection in an experiment. The Kaplan-Yorke formula only works for filters that are dynamical systems. Recently Sauer and Yorke have given such a criteria [9]. They provide a theorem that shows that when the filtering induces a map from the unfiltered attractor to the filtered attractor which is continuously differentiable (C^1) the dimension quantities associated with the measure on the attractor (e.g. information or correlation dimensions) are preserved.

The theorem implies a need for a numerical algorithm to test for the C^1 condition in filtered data. Below we show how tests for continuity and differentiability we have devised fulfill this need. The tests also provide statistics that suggest a very interesting scaling behavior in highly filtered data that is not obvious from dimension measurements. This leads to a clear geometric picture of the action of a filter on a chaotic attractor and an explanation for dimension increase.

II. Statistics for Continuity and Differentiability

The action of any convolutional filter on a time series from a dynamical system can be written as

$$z(t) = \int R(t-t')\, x_1(t')\, dt' = \int R(t-t')\,(\Phi_{t'-t}[\mathbf{x}(t)])_1\, dt' \,, \qquad (2)$$

where Φ is the flow on the system's phase space and $R(t)$ is the filter. The action of the LTI filter (Eq. (1)) can also be written this way. It is not hard to show [1] that this induces a point-to-point mapping $\boldsymbol{f}$ from the attractor reconstructed from the original time series to that reconstructed from the filtered time series: $\mathbf{z}(t) = \boldsymbol{f}(\mathbf{x}(t))$, where $\mathbf{z}$ is the time delay vector generated from $z(t)$, $z(t+\Delta t)$,... and $\mathbf{x}$ is the time delay vector generated from $x_1(t)$, $x_1(t+\Delta t)$,... , etc. The Sauer-Yorke theorem states that we must show that $\boldsymbol{f}$ is a C^1 function. We next show how to do this with time-series data.

Our goal is to develop simple statistics that are (1) clearly connected to the one mathematical property they measure and (2) have a separate statistical hypothesis associated with them. Number (1) criteria guarantees that we are sure what property we are testing. Number (2) guarantees we know what we are using as a gauge of that property. These goals are, in a sense, reversed from what one usually finds in the literature where tests which are an amalgam of mathematical properties and numerical algorithms are used to test for attributes which are derived, for example, determinism [10,11], general synchronization [12], and correct embedding dimension [13,14]. Although these numerical tests can be useful, it is not always clear what mathematical property is being tested.

We are testing for continuity (C^0) and differentiability (C^1). We give two separate statistical tests, one for each property.

A test for continuity must start with the definition of continuity: the function $\boldsymbol{f}$ is continuous at a point $\mathbf{x}_0$ if $\varepsilon>0$ $\delta>0$ such that $\| \mathbf{x}-\mathbf{x}_0\|<\delta$ $\|\boldsymbol{f}(\mathbf{x})-\boldsymbol{f}(\mathbf{x}_0)\|<\varepsilon$.. In simpler terms if we restrict ourselves to some local region around $\boldsymbol{f}(\mathbf{x}_0)$, then there must exist a local region around $\mathbf{x}_0$ *all* of whose points are mapped into the $\boldsymbol{f}(\mathbf{x}_0)$ region. From this definition we can generate an algorithm to apply to time-series reconstructions.

Our reconstructions are matched in a "one-to-one" fashion (which we call $\boldsymbol{f}$) in that for every point $\mathbf{x}$ we have a corresponding point $\mathbf{z}$; however, we do not know how points in local neighborhoods are paired. We choose an ε -sized set around a fiducial point $\mathbf{z}_0$, we also choose a δ -sized set around it's preimage $\mathbf{x}_0$. We check whether all the points in the δ set map *into* the ε set. If not, we reduce δ and try again. We continue until we run out of points or all points from a small-enough δ set fall in the ε set. We count the number of points in the ε set (n_ε) and the δ set (n_δ). We do not include the fiducial points $\mathbf{z}_0$ or $\mathbf{x}_0$, since they are present by construction. Generally $n_\varepsilon \geq n_\delta$, since points other than those near $\mathbf{x}_0$ can also get mapped to the ε set, but this does not affect continuity.

We now choose the null hypothesis which helps us generate a probability that one should find n_ε and n_δ points in such an arrangement. Many null hypotheses are possible. We choose the simplest, namely, that placements of the points on the $\mathbf{x}$ and $\mathbf{z}$ attractors are independent of each other. This null hypothesis is not trivial. It is typical of what one would like to disprove early on in any attractor analysis, namely that the data have a relation to each other. We will see that it can show much more.

Given the null hypothesis we approximate the probability of a point from the δ set falling at random in the ε set as $p = n_\varepsilon /N$, where N is the total number of points on the attractor. Then the probability that n_δ points will fall in the ε set is p^{n_δ} . We obtain a likelihood that this will happen by taking the ratio of this probability to the probability for

the most likely event, p_{binmax}. The latter is just the maximum of the binomial distribution for n_δ points given probability p for each individual event. We see that p^{n_δ} is simply the "tail-end" of the binomial distribution. The maximum generally will occur for some intermediate number of δ points, say m ($<n_\delta$), falling in the ε set. If $p^{n_\delta} \ll p_{binmax}$, then the null hypothesis is not likely and can be rejected.

We define the continuity statistic as $\Theta_C 0 = 1 - p^{n_\delta} / p_{binmax}$. When $\Theta_C 0 \approx 1$ we can confidently reject the null hypothesis. The points in the ε set are behaving as though they are generated by a continuous function on the δ set. When $\Theta_C 0 \approx 0$ we cannot reject the null hypothesis and the points are behaving as though they are independent. Note that if we run out of points ($n_\delta = 0$), then we usually take the logical position that we cannot reject the null hypothesis and set $\Theta_C 0 = 0$. $\Theta_C 0$ will depend on ε, the resolution, and we will examine the statistic for a range of ε's. To get a global estimate of the continuity of $\boldsymbol{f}$ on the attractor we average $\Theta_C 0$ over the entire attractor or over a random sampling of points on it. We present those averages here.

The differentiability statistic is generated in the same vein as the continuity statistic. We start with the mathematical definition of a derivative and apply it locally to the two reconstructions. The generation of the linear map that approximates the derivative and the likelihood estimate associated with it are more complex than for continuity. We show the details in our longer paper [1]. We only outline the differentiability statistic here.

The definition of a derivative at a point $\mathbf{x}_0$ is that a linear operator A exists such that $\varepsilon>0 \quad \delta>0$ for which $\| \mathbf{x}-\mathbf{x}_0\|<\delta \quad \|\boldsymbol{f}(\mathbf{x}_0) + A(\mathbf{x}-\mathbf{x}_0) - \boldsymbol{f}(\mathbf{x})\| < \varepsilon\|\mathbf{x}-\mathbf{x}_0\|$. This means that there is a linear map that approximates the function at nearby points with an error in the approximation that is proportional to the distance between those points. Note that ε serves a purpose here different from continuity. Here it bounds the error in a linear mapping; it does not signify a distance.

The algorithm that we generate from this definition is to first choose an ε (error bound) and a δ. Then we find all the points in the local δ set $\{\mathbf{x}_i\}$ and their $\mathbf{z}$

counterparts $\{\mathbf{z}_i\}$. We approximate the linear operator A as the least squares solution of the linear equations $A(\mathbf{x}_i-\mathbf{x}_0) = (\mathbf{z}_i-\mathbf{z}_0)$. The solution is accomplished by singular value decomposition (SVD) [1]. The least-squares approach appears to be the best since it minimizes the errors in approximation which is just what ε is supposed to bound. We check if $\| \mathbf{z}_i-\mathbf{z}_0-A(\mathbf{x}_i-\mathbf{x}_0) \| < \varepsilon \ \| \mathbf{x}-\mathbf{x}_0\|$. If not, we decrease δ and try again with fewer, but nearer points. We continue this until we have success or we run out of points.

We choose the null hypothesis that the two sets of vectors $\{\mathbf{x}_i\}$ and $\{\mathbf{z}_i\}$ have zero correlation. We show [1] that this generates a likelihood that any two such sets will give the operator A "by accident" as $e^{-\frac{1}{2}(n_\delta - r_x)(n_\delta - r_z)r^2 d}$, where r^2 is the usual multivariate statistical correlation between $\{\mathbf{x}_i\}$ and $\{\mathbf{z}_i\}$, d=min(r_x, r_z), and r_x, r_z are the ranks of the $\mathbf{x}$ and $\mathbf{z}$ spaces that come out of the SVD [1]. This is an asymptotic formula. The differentiability statistic Θ_C1 is given by one minus this likelihood. When $\Theta_C1 \approx 1$ we can reject the possibility that the points are accidentally related by a linear operator, a derivative. When $\Theta_C1 \approx 0$, we cannot reject the null hypothesis. As before, when we shrink δ so small that no points other than $\mathbf{x}_0$ remain, we set $\Theta_C1 =0$. Analogous to Θ_C0, the statistic Θ_C1 depends on ε. We typically calculate Θ_C1 for a range of ε's and average over the attractor or over a random sampling of points on it.

We now have two statistics Θ_C0 and Θ_C1 that we can use to test for the Sauer-Yorke criteria for filtered chaotic data. We will mostly concentrate on the continuity statistic, since the differentiability statistic follows it very closely in this case. We give all ε and δ values in units of the standard deviation of the reconstructed attractor, so that, for example $\varepsilon = 1.0$ is an ε set with radius of one standard deviation. For all of our studies so far this seems to be a good normalization for set sizes.

III. Application to LTI and acausal filtered data

The two filters applied to the data were a LTI filter and an acausal filter. The LTI filter gives a new quantity z as in Eq. (1). This is the same as convolving x_1 with an exponential in time (with $-\eta$ as the exponent) or multiplying the Fourier transform of x_1

by $1/(i\omega+\eta)$. The acausal filter has a Fourier transform that is similar to the LTI filter, except it is real and causes no phase shifts: $1/\sqrt{\omega^2+\eta^2}$ [6].

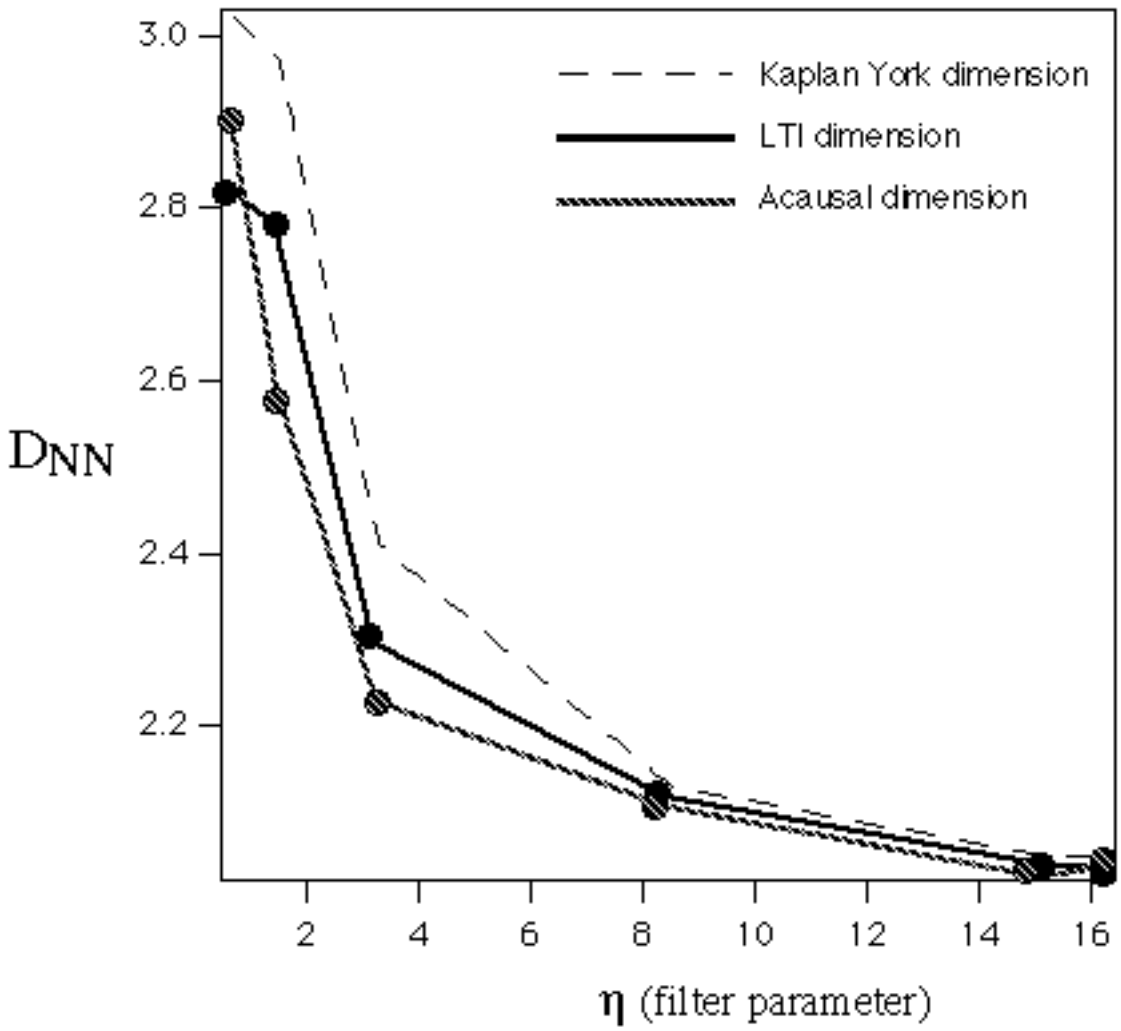

Figure 1. Nearest neighbor dimension of LTI and acausal filtered Lorenz attractors as a function of filter parameter η, along with the Kaplan-Yorke dimension estimate for the LTI filtered case.

We applied these filters to time series from two chaotic systems. The first was a Lorenz system with parameters $\sigma=10$, $b=8/3$, and $\rho=60$ which was integrated for 128,000 points using a 4th order Runge-Kutta routine with a 0.02 time step. The filtered x

component was used to reconstruct an attractor in six-dimensional space. Various time series lengths were used from 4000 to 128000 points. The second system was a chaotic Henon system with parameters $a = 1.4$ and $b = 0.3$. The filtered Henon x component was used in reconstructions in six dimensional space. We used six dimensions to insure complete unfolding of the attractors. Similar results were obtained for five-dimensional reconstructions and eight-dimensional reconstructions.

Fig. 1 shows the nearest-neighbor dimension [15] (which is the same as the information dimension) for the filtered Lorenz data as a function of filter parameter η.

We can analyze the LTI results using Kaplan Yorke formula for the fractal dimension D_L of an attractor : $D_L = j + \sum_{k=1}^{j} \lambda_k / |\lambda_{j+1}|$, where λ_k are the Lyapunov exponents and j is the largest index for which the sum is positive. According to this formula the dimension of the LTI filtered data should begin to increase when η goes below 15.085 (the smallest exponent for the Lorenz system). And at $\eta \leq 1.414$ we should see a leveling off in the increase in the dimension. The values of D_L for the Lorenz system are plotted in Fig. 1. They match the observed trend in the filtered attractor dimension nicely.

The problem is that the Kaplan-Yorke formula cannot explain the dimension increase observed from the use of the acausal filter. That filter (see Fig. 1) causes somewhat smaller increases in dimension.

We can begin to get some feel for how the filtering process affects the attractor by examining the continuity of the functional relation between the original, reconstructed attractor and the reconstructed, filtered attractor. Fig. 2 shows the results of calculating the continuity statistic $\Theta_C 0$ vs. the filtering parameter η for this function at a particular ε value. We choose $\varepsilon = 0.06$ which is a small set size on the attractor, about 0.1% of the attractor points are in such a set on average (see Fig. 2 inset for an example). We see from Fig. 2 that as we increase the amount of filtering, the continuity for the function between the original attractor and the filtered attractor drops off, with the acausal

continuity dropping more slowly in agreement with the smaller increase in dimension in the acausal, filtered attractors.

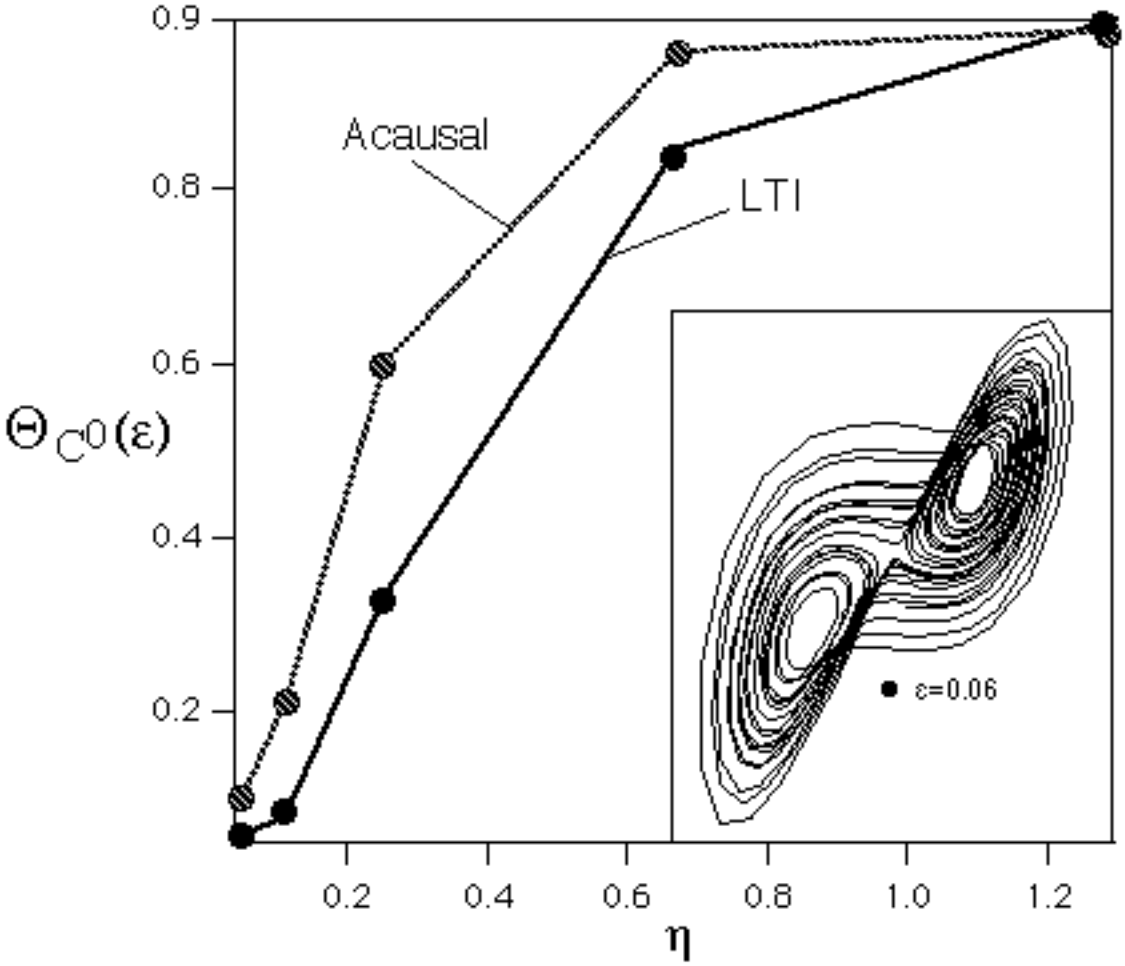

Figure 2. Continuity statistic vs. filter parameter for LTI and acausal filtered Lorenz attractors. The inset shows a two-dimensional projection of the attractor and the ε size used to calculate $\Theta_C 0$ in the figure.

Many times a drop-off in the $\Theta_C 0$ statistic is caused by the fact that as we decrease ε we decrease the number of points that can be found at smaller resolutions and the statistic becomes the victim of the finite amount of data. To test this we plotted $\Theta_C 0$ as a function of ε for acausal filtering for various time series lengths for two filtering parameters, one

η =16.0 where we are sure we should get good continuity and the other at η =0.113 where the filtering is extreme.

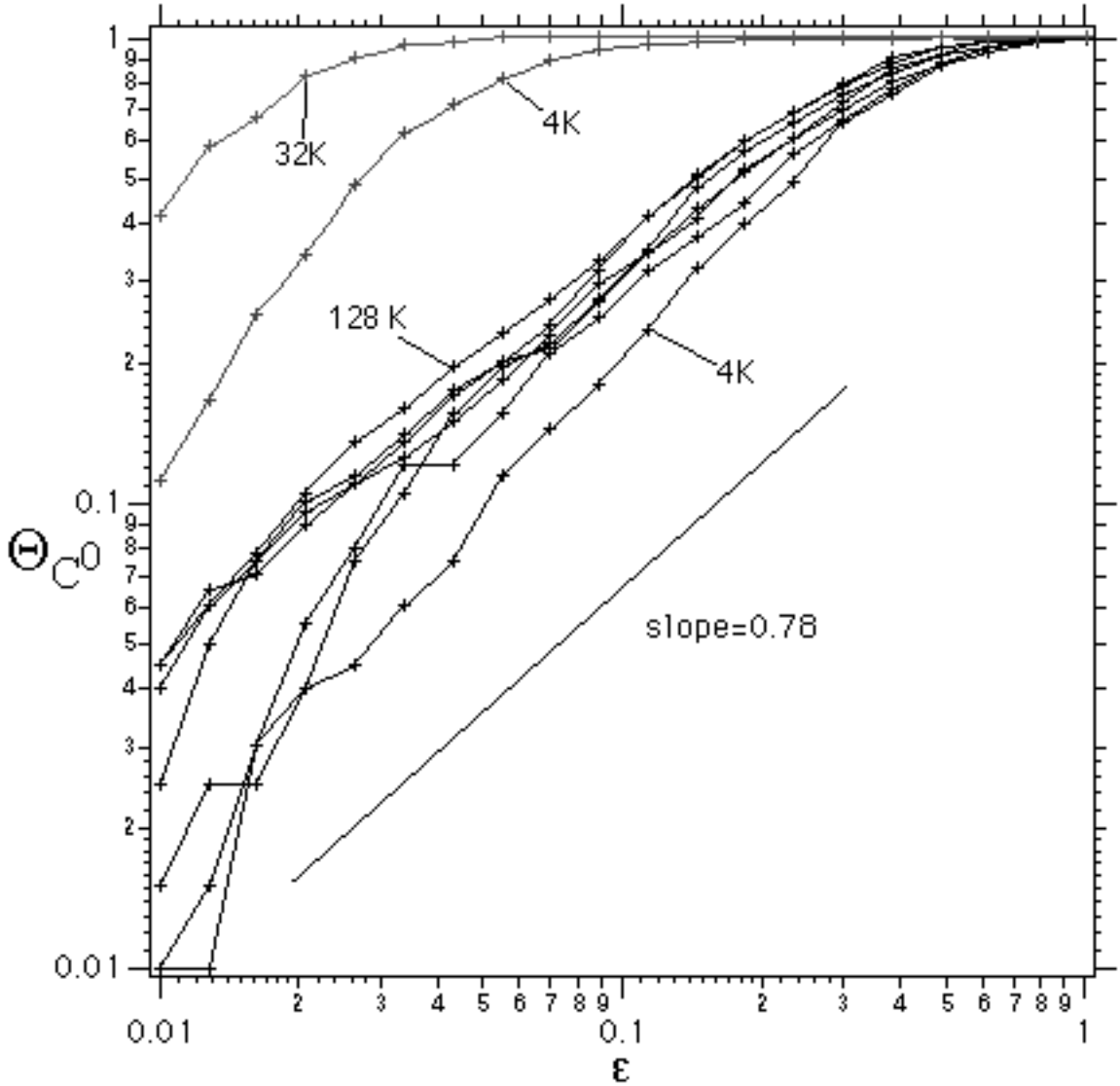

Figure 3. Continuity statistic vs. ε for various length filtered Lorenz time series for the acausal filter. The dotted lines are for mild filtering with no dimension increase (η = 1.6). The solid lines are for severe filtering with dimension increase (η = 0.113). The straight of slope 0.78 is from a least squares fit to the log-log plots for severe filtering using the time series with 32000, 64000, and 128000 points.

Fig. 3 shows these plots. For the mildly filtered case we see two features. One is the continuity statistic is nearly 1.0 down to very small ε sizes and the other is that the continuity drop-off at small ε can be reversed by simply taking longer time series.

For the heavily filtered time series we see that the features of the mildly filtered case do not show up. The continuity statistic is small even for larger ε sizes and it does not markedly improve with increasing time series length. In fact in going from 4000 to 128000 points we see very little improvement. From 32000 points and higher there is almost no improvement. The statistic saturates as a function of the amount of data; adding data does not improve it. This is a sign that there are true discontinuities in the function mapping the original attractor to the filtered attractor [1].

Fig. 3 gives us a clue into the nature of those discontinuities. In this log-log plot we see that the continuity statistic is nearly a straight line for a decade of ε sizes. The continuity null hypothesis is such that $\Theta_C 0$ is nearly 1 or 0 at each point on the reconstructed attractor [1]. This means that $\Theta_C 0\ (\varepsilon)$ is a good approximation to the fraction or density of points on the attractor that have discontinuities of size ε or greater. The linear behavior in the plot implies that the density of discontinuities $\rho_C 0$ should scale as $\rho_C 0 = \varepsilon^{\alpha}$, with $\alpha \approx 0.78$ in Fig. 3.

We can gain some more insight by plotting $\Theta_C 0$ as a function of the number of point in each ε set, n_ε. This gets rid of any effects associated with specific set-size scaling. The results are shown in Fig. 4 for the Lorenz system for the same filtering as in Fig. 3. We see we also get a scaling relation (this is a log-log plot) in which we can write $\Theta_C 0 \sim n_\varepsilon^{\beta}$. We can relate this to the $\Theta_C 0 \sim \varepsilon^{\alpha}$ scaling above as follows. We expect the number of points in each set, n_ε., to scale as ε^{D}, where D is the fractal dimension of the attractor. Then from our plots we would deduce $\Theta_C 0 \sim n_\varepsilon^{\beta} = \varepsilon^{\beta D} = \varepsilon^{\alpha}$, or that $\beta D = \alpha$. From the plots we have $\beta = 0.38$ and $\alpha = 0.78$. The fractal dimension D of the Lorenz attractor for our parameters is slightly greater than 2.0. Hence, the scaling

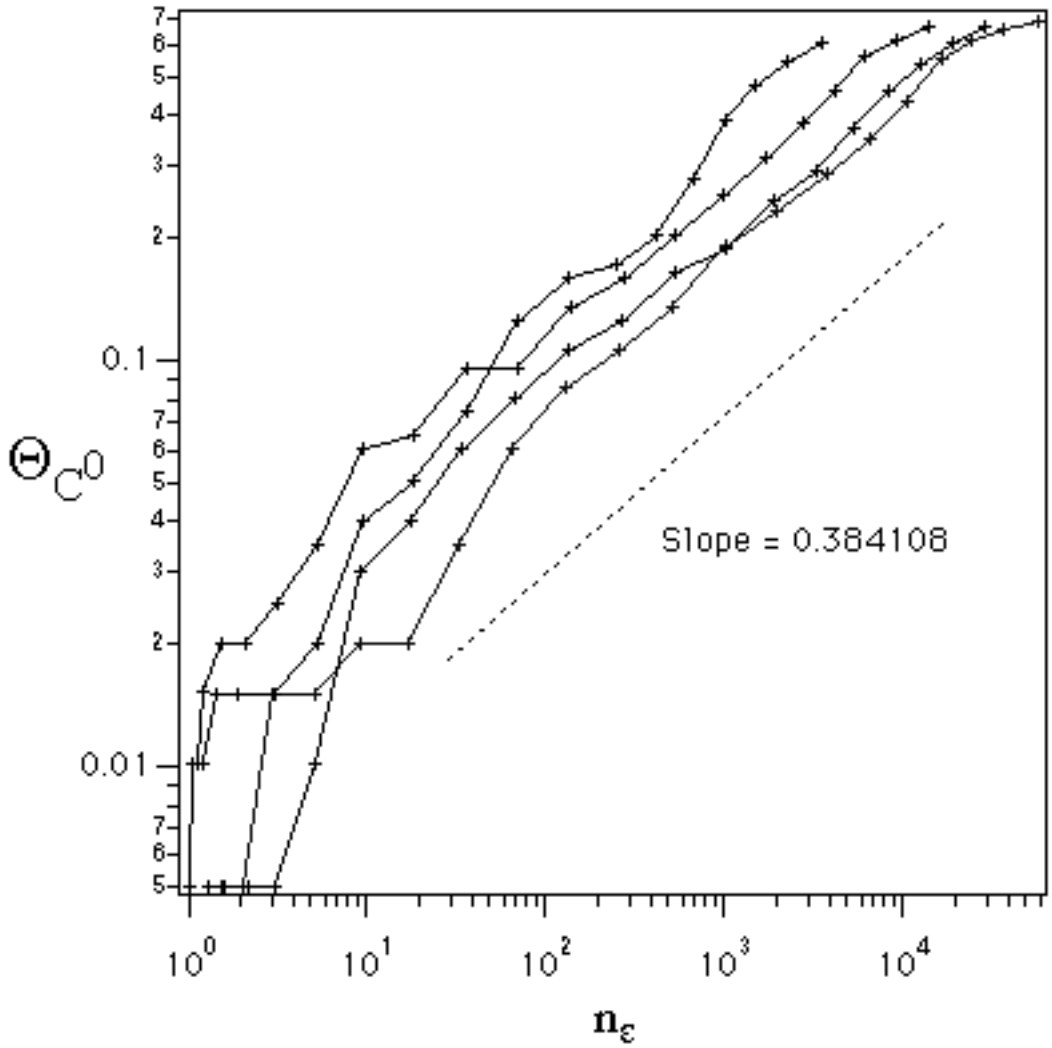

Figure 4. Plot of $\Theta_C 0$ vs. n_ε with average fit to all curves showing a scaling relation between $\Theta_C 0$ and n_ε.

exponents are consistently related. The scaling exponents' meanings are that if we choose a euclidean measure for spatial volume (the ε set), we will see one scaling of the density of discontinuities given by α. If we choose to use the fractal measure of the attractor as a measure of the volume on which the discontinuities appear we will see another scaling exponent β given by α / D. That is the change in scaling exponents simply reflects the change in volume measures as we would expect. With either measure the conclusion

would be that the set of discontinuities is dense on the attractor and the size of the discontinuities decreases as ε decreases.

We see similar behavior in LTI filtered time series. Typically scaling exponents for the LTI statistics are larger than for the acausal statistics.

In all our studies the differentiability statistic $\Theta_C 1$ follows the discontinuity statistic $\Theta_C 0$ rather closely implying the loss of differentiability is resulting from the loss of continuity, although this does not have to be the case. It may be that the mapping from unfiltered to filtered is actually becoming nondifferentiable with decreasing η , but remaining continuous. In this case the function would fluctuate wildly analogous to the Weirstrass function, for example. Therefore, if the nondifferentiability is "severe" the function will appear discontinuous at lower resolution (larger ε sizes). We may not have the data quantity and precision to go to small-enough ε 's to see this behavior. However, the mechanism for dimension increase given below will not, in the first approximation, be affected by this issue.

The behavior of filtered Henon time series is very similar to the Lorenz case. For the Henon map the LTI filter causes a larger increase in fractal dimension than the acausal filter [1]. The LTI filter dimension increases begin at "Kaplan-Yorke value" of $\eta = 1.58$ and proceeds to the "breakpoint" value $\eta = 0.417$ where it levels off as in the Lorenz case. Similarly, for the heavily filtered time series both the LTI and acausal attractors are related to the original attractor by a function that has a continuity statistic $\Theta_C 0$ that saturates with time series length and has a power-law scaling.

This scaling behavior is similar to that found for discontinuities in the devil's staircase [16]. This implies that the discontinuities (or apparent discontinuities due to nondifferentiability) caused by extreme filtering are dense, in some way, on the original attractor. That is, these are not isolated discontinuities, but at any ε we will find discontinuities "almost everywhere." A set of such discontinuities may increase the

fractal dimension of the original attractor. The questions now are, what in the filtering process causes the discontinuities and how can we view this process?

IV. A geometric interpretation of filtering

We answer the previous questions by starting with the second one. We view the action of a filter by augmenting the phase-space reconstruction (or the original attractor's phase space, if we have access to it) by a component z which will represent the filtered time series. Likewise, we augment the points on the attractor with another component, the values z_n resulting from application of the filter. For example, in the discrete case we can define

$$z_n = a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} + \ldots \tag{3}$$

as the action of an LTI filter on a time series (with properly chosen a_i values). In general, Eq. (2) shows the relationship of the filtered time series $z(t)$ to the phase space point $\mathbf{x}(t)$. The main point is that each point on the attractor is associated with a filtered value. Then we consider the augmented vectors $(x_n, x_{n-1}, \ldots, x_{n-(d-1)}, z_n)$ in d+1 dimensional space, where d was the original embedding dimension.

Fig. 5 shows this for the Henon map. In Fig. 5(a) we have LTI filtering below the level necessary to increase the dimension of the attractor. We see there a curvilinear version in three dimensions (x, y, z) of the original attractor as plotted in the (x, y) plane. The projection of 4(a) onto the (x, y) plane still gives the original Henon attractor, of course, but the object we are reconstructing from the filtered time series actually comes from the one in three-dimensional (3D) space. The f in the figure represents the mapping from the original unfiltered attractor to the new filtered attractor. Here f is continuous – points nearby in the (x, y) plane are also nearby in the three-dimensional object. The mapping is also smooth and therefore C^1. The dimension of the new object remains the same as the original Henon attractor. In this figure we can embed the filtered attractor in three dimensions because we have all three components. In our time-delay

reconstructions this is not the case and we must go to 5 or 6 dimensions to insure proper embedding.

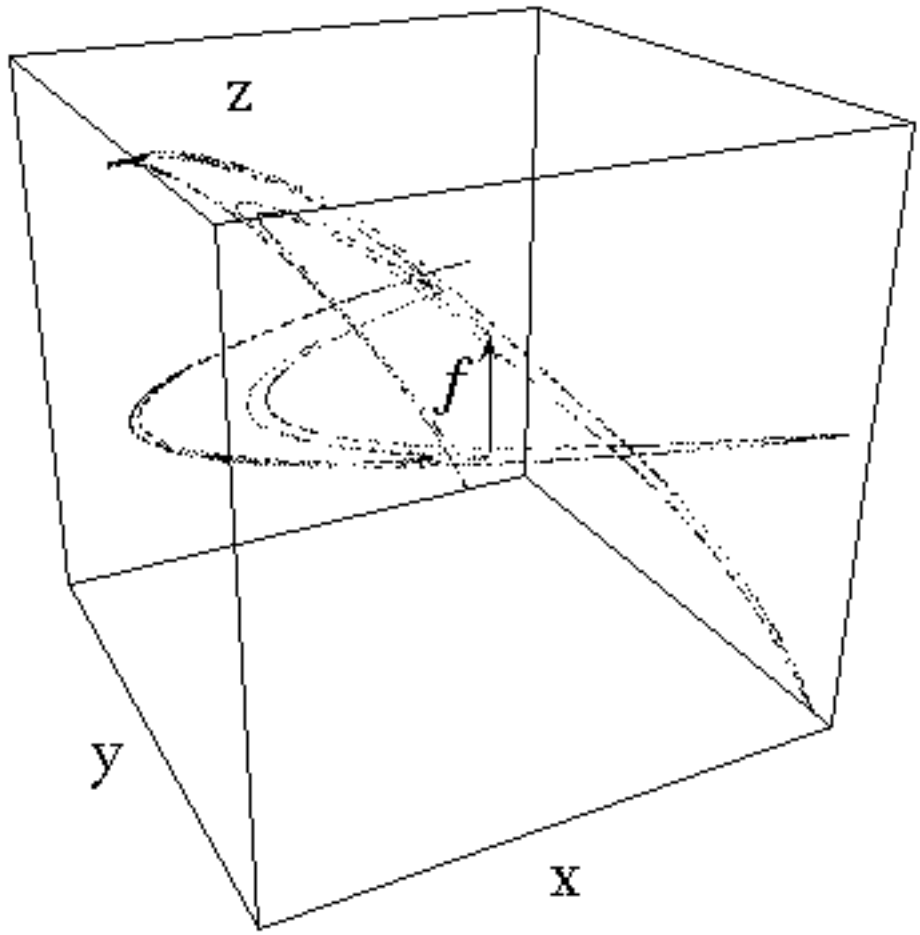

Figure 5. (a) Plot of the Henon attractor in the (x,y) plane and the filtered values z related to each Henon (x,y) point for acausal filtering below the threshold for dimension increase (η = 1.6).

In Fig. 5(b) we see a similar plot for the heavily filtered case. Now rather than a smooth object in 3D we see that the Henon attractor has been "smeared out" into the z-direction. In Fig. 5(c) we view the 3D object of Fig. 5(b) from above and see that it still projects down to the original Henon attractor as we would expect from the construction of our new augmented phase space, but no longer do points nearby in the (x, y) plane map to

points nearby in the 3D object. The mapping *f* is discontinuous – nearby points typically map to points that are distant in the 3D space. The "layers" of the Henon attractor which follow the folding unstable manifold [17] are smeared up into the *z*-direction like curved sheets. This is the discontinuity between the original attractor and the filtered attractor that our statistic $\Theta_C 0$ is detecting.

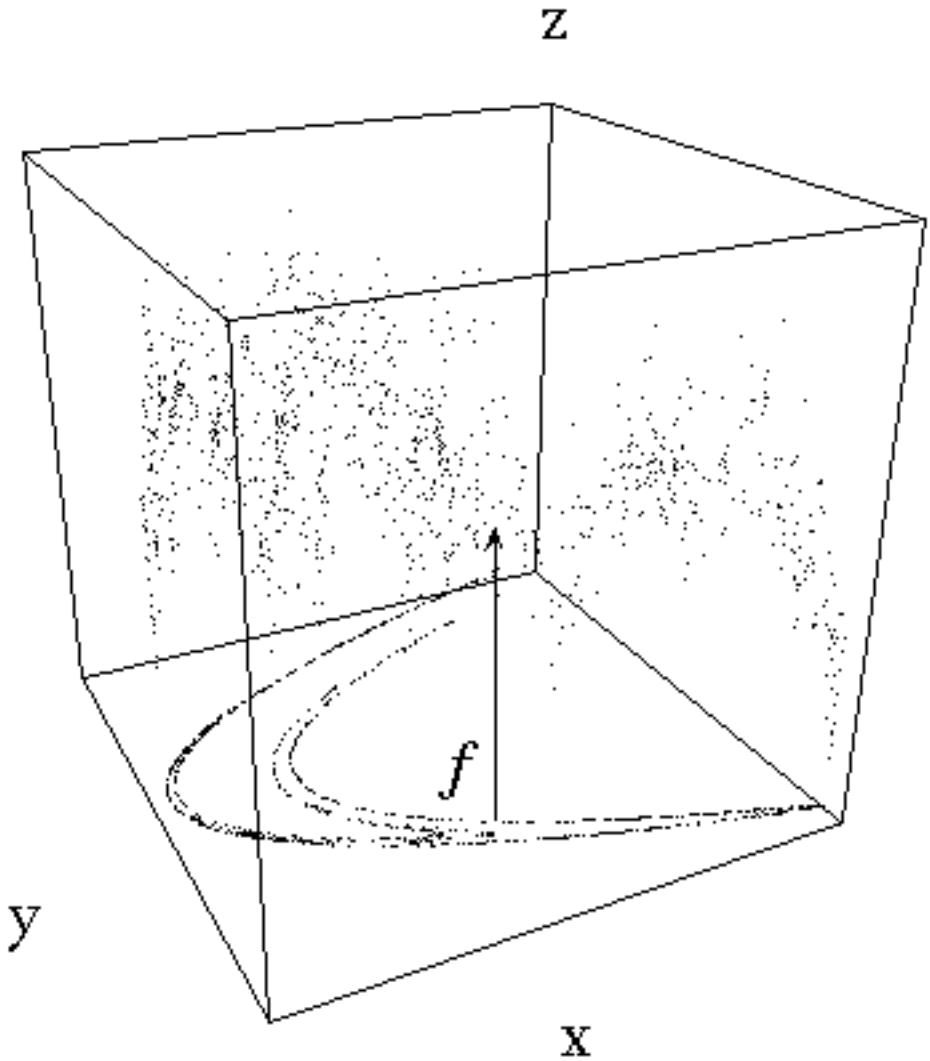

Figure 5. (b) Plot of the Henon attractor in the (*x*,*y*) plane and the filtered values *z* related to each Henon (*x*,*y*) point for acausal filtering above the threshold for dimension increase (η = 0.2).

The power-law scaling of $\Theta_C 0$ strongly suggests that the discontinuities are "dense." Thus, the new 3D object will, in general, have a larger dimension than the Henon attractor. When we reconstruct using filtered time series we are actually sampling the 3D object and not the Henon attractor. Hence, our filtered attractor will have the increased dimension of the 3D object.

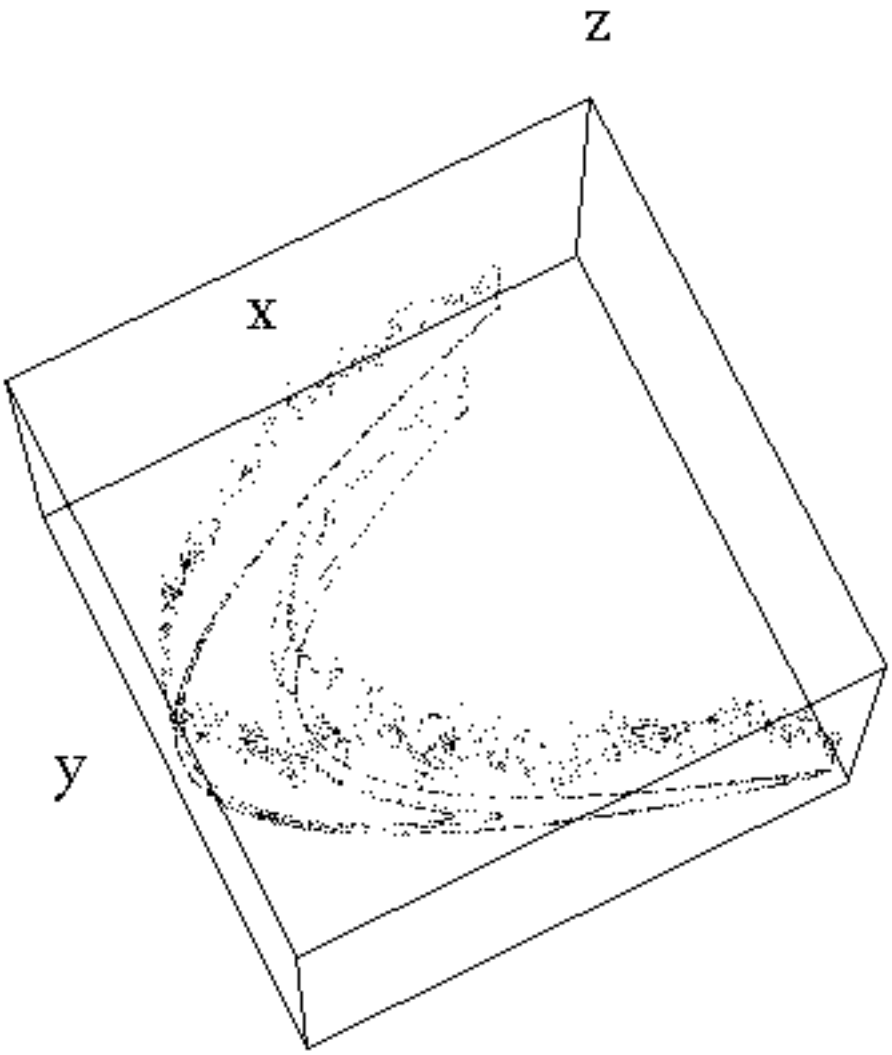

Figure 5. (c) Same as (b) but from a higher *z* viewpoint to show that the augmented phase-space object projects down onto the original Henon map.

Now that we have a way to view the filtering, we can heuristically explain where the discontinuities come from. Consider the application of a filter with a long tail into the past (e.g. LTI filter). Apply the filter at two points, one starting on an unstable periodic orbit (UPO) and the other starting near the UPO on the Henon attractor. Each point will have a filter value z associated with it. We show the scenario in Fig. 6 with a period-2 UPO. The point on the UPO will oscillate back and forth between the two circles. The point nearby at point-0 will drift away from the UPO as we move backward in time. This is a consequence of the negative Lyapunov exponent which acts to make points diverge in reverse time. Hence, the nearby point will sample different areas of the attractor than the point on the period-2 UPO. A long-tailed filter will be a weighted average of these points. If the tail is long enough the filter values will be vastly different, even though both trajectories started very close at point-0. Hence, nearby points (at point-0) will map to different filter (z) values – we have may have a discontinuity.

We note that this does not prove that we truly have discontinuities. It is only a heuristic argument to help explain what the continuity statistic suggests. It can also be argued that the function from the attractor to the z values is actually non-differentiable [18]. To see this we can use Eq. (2) with an LTI filter. In this case the integration over time is from $-\infty$ to 0 and the filter kernel $R(t)$ is just $e^{\eta t}$. Then the difference between filter values Δz coming from nearby points $\mathbf{x}_1$ and $\mathbf{x}_2$ is given rougly by

$$\Delta z = \int_{-\infty}^{0} e^{\eta t}(\Phi_{-t}[\mathbf{x}_1] - \Phi_{-t}[\mathbf{x}_2])dt' \approx \int_{-\infty}^{0} e^{\eta t} D\Phi_{-t} \quad \Delta\mathbf{x}\, dt' \quad , \tag{4}$$

where $\Delta\mathbf{x}=\mathbf{x}_1-\mathbf{x}_2$. In order to get an estimate of the magnitude of Δz we might be tempted to substitue $e^{-\lambda t}$ for $D\Phi_{-t}$, where λ is the smallest (negative) eigenvalue and conclude that when $|\lambda| > |\eta|$ Eq. (4) diverges and no matter how small $\Delta\mathbf{x}$ is we will always have a finite Δz (a discontinuity). But we must remember that the attractor is compact (finite in size) and a better substitution for $D\Phi_{-t}$ would be min { $e^{-\lambda t}$, 1}, where we assume the attractor size is of order 1. Then, when $t \to -\infty$, we have a "bound" on Δz of $1 \cdot |\Delta\mathbf{x}|$.

This argues (but does not prove) that we are not really seeing discontinuities, but rather that the function relating *z* to **x** is non-differentiable.

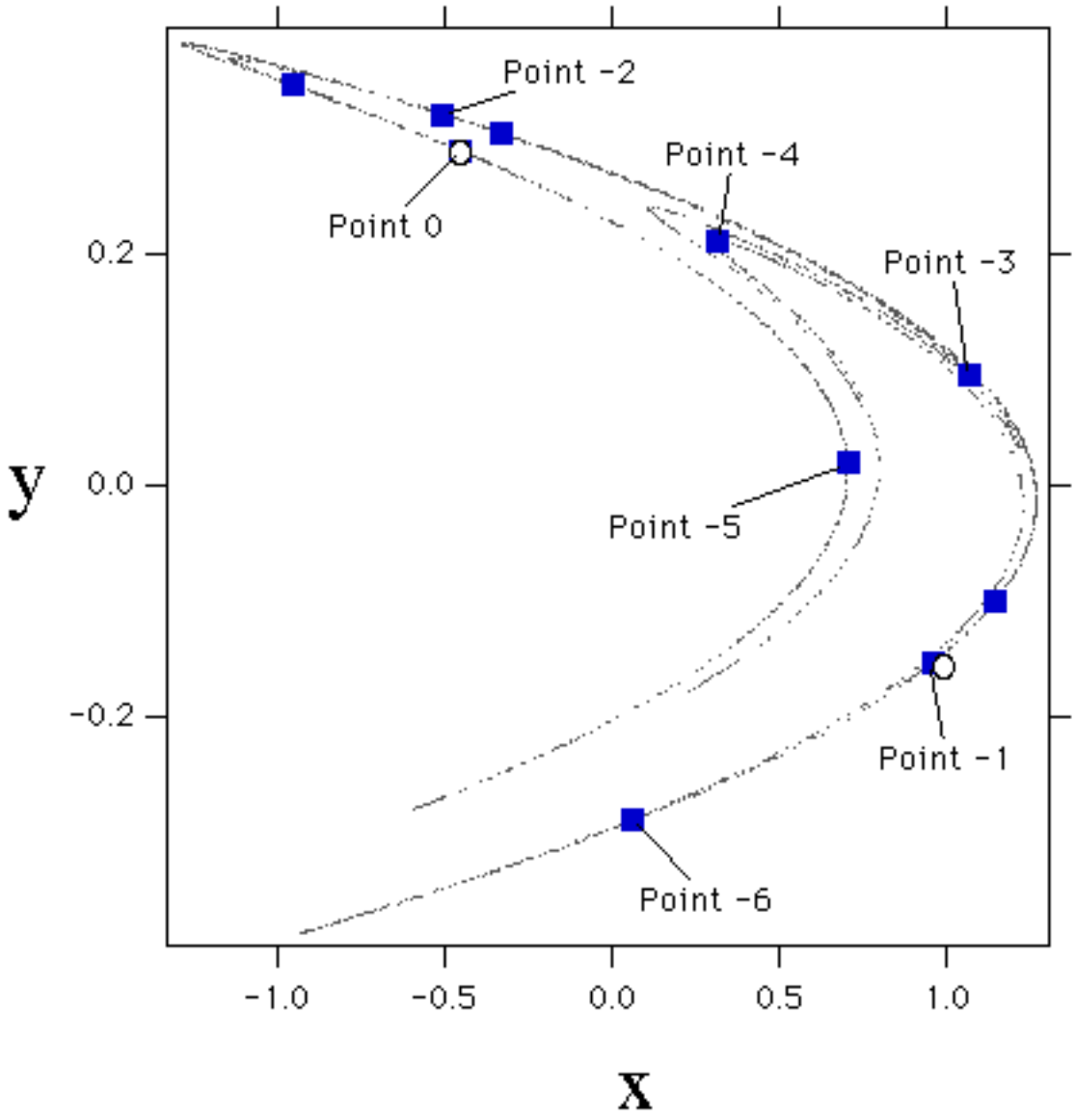

Figure 6. The trajectories in reverse time for two points on the Henon attractor, one on the period-2 UPO and another starting close by. Because the one close by diverges it covers other areas of the attractor not covered by the UPO points and therefore gives a different filtered value *z*.

Why do the continuity statistics remain low? Apparently we are not sampling at enough points to see continuity. If we imagine a non-differentiable function (e.g. the Weirstrass function), we see large and rapid oscillations. If this is not sampled densely enough, we will appear to have a discontinuous function. This is probably what is happening in our examples. This does not mean that the continuity statistic is useless. Either discontinuities or non-differentiabilities can cause dimension increase. The detection of discontinuities between data sets serves to tell us that, for any practical purpose, the relation is truly discontinuous. This is the message one can take away from any statistical result: given the data, for practical purposes, it acts like the statistic says it does.

V. Conclusions and discussion

The filtering-dimension-increase scenario we expose above does not depend on the type of filter, so long as the filter has a long enough tail that drops off more slowly than the smallest Lyapunov exponent into the past and/or the largest Lyapunov exponent into the future. All filters of infinite time extent can potentially cause dimension increase.

Below we offer what we believe are some interesting conjectures regarding the relationships between chaotic attractors, filters, and UPO's.

We note that others (e.g. Rosenstein *et al* [19]) have seen a "thickening" of attractors after the application of LTI filters. For example, the Lorenz attractor appears to bulge out transverse to its usual planes of oscillation around the unstable fixed points [19]. We would conjecture that this is a manifestation of the effect we see in the Henon attractor, Figs. 5.

Since UPO's appear to be involved we might expect isolated discontinuities to begin to appear as we change the filter parameter η . This is because each UPO has its own Lyapunov exponent. The Lyapunov exponents of the attractor are an ergodic average of all the UPO Lyapunov exponents. Hence, we may see discontinuities associated with the more unstable UPO's first, then get a dense set of discontinuities only when we reach

some threshold dependent on the attractor's Lyapunov exponents. The scaling we see in the $\Theta_C 0$ statistic should be predicatable from the density of UPO's of various periods.

We feel that rigorous proofs of many of our conjectures and scenarios may be possible, but we have not been able to accomplish that at this time.

Acknowledgments

We would like to acknowledge helpful conversations with T. Sauer, J. Heagy, J. Yorke, J. Theiler, M. Kennel, D. Broomhead, B. Hunt, and E. Ott regarding the development of statistics to measure mathematical properties, the effects of filters and in understanding diffeomorphisms.

References

[1] Pecora, Carroll, and Heagy, “Statistics for Mathematical Properties of Maps between Time-Series Embeddings,” Physical Review E **52** (4), 3420 (1995).

[2] R. Badii, G. Broggi, B. derighetti *et al*., “Dimension Increase in Filtered Chaotic Signals,” Physical Review Letters **60** (11), 979 (1988).

[3] J.L. Kaplan and J.A. Yorke, “Chaotic behavior of multidimensional difference equations,” in *Functional Differential Equations and Approximations of Fixed Points, Lecture Notes in Mathematics*, edited by H.-O. Peitgen and H.-O. Walther (Springer-Verlag, Berlin, 1979), Vol. 730, pp. 204--227.

[4] P. Paoli, A. Politi, G. Broggi *et al*., “Phase Transitions in Filtered Chaotic Signals,” Physical Review Letters **62** (21), 2429 (1989).

[5] A. Chennaoui, J. Liebler, and H.G. Schuster, “The Mechanism of the Increase of the Generalized Dimension of a Filtered Chaotic Time Series,” Journal of Statistical Physics **59** (5), 1311 (1990).

[6] F. Mitschke, “Acausal filters for chaotic signals,” Physical Review **A 41** (2), 1169 (1990).

[7] S.H. Isabelle, A.V. Oppenheim, and G.W. Wornell, “Effects of Convolution on Chaotic Signals,” *ICASSP-92 Proceedings*, San Francisco, 1992, **5**,IV-133-136 (IEEE,Piscataway, NJ)

[8] D.S. Broomhead, J.P. Huke, and M.R. Muldoon, “Linear Filters and Non-linear Systems,” Journal Royal Statistical Society **B 54** (2), 373-382 (1992).

[9] T. Sauer and J.A. Yorke, “Are the dimensions of a set and its image equal under typical

smooth functions?,” Ergodic Theory and Dynamical Systems **(to be published)** (1995).

[10] D.T. Kaplan, “Exceptional events as evidence for determinism,” Physica D **73**, 38-48 (1994).

[11] D.T. Kaplan and L. Glass, “Direct Test for Determinism in a Time Series,” Physical Review Letters **68**, 427 (1992).

[12] N. Rul'kov, M.M. Sushchik, L.S. Tsimring *et al*., “Generalized Synchronization of Chaos in Directionally Coupled Chaotic Systems,” Physical Review **E 51**, 980 (1995).

[13] M.B. Kennel, R. Brown, and H.D.I. Abarbanel, “Determining Embedding Dimension for Phase Space reconstruction Using the Method of False Nearest Neighbors,” Physical Review **A 45**, 3403 (1992).

[14] M.B. Kennel and H.D.I. Abarbanel, “Local false nearest neighbors and the dynamical dimensions from observed chaotic data,” Physical Review **E 47**, 3057 (1993).

[15] Y. Termonia and Z. Alexandrowicz, “Fractal Dimension of STrange Attractors from Radius versus Size of Arbitrary Clusters,” Physical Review Letters **51**, 1265 (1983).

[16] Jens Feder, *Fractals* (Plenum Press, New York, 1988).

[17] E. Ott, “Strange attractors and chaotic motions of dynamical systems,” Review of Modern Physics **53** (4), 655-671 (1981).

[18] B.Hunt and E. Ott, private communications.

[19] M.T. Rosenstein and J.J. Collins, “Visualizing the effects of filtering chaotic signals,” Computers and Graphics **18**, 587-592 (1994).