DISSERTATION

IMPROVED METHOD FOR DISTINGUISHING DETERMINISM FROM RANDOMNESS IN THE PRESENCE OF NOISE

Submitted by
Sean Roberts
Physics Department

In partial fulfillment of the requirements
For the Degree of Doctor of Philosophy
Colorado State University
Fort Collins, Colorado
Spring 2003

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

January 27, 2003

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY SEAN ROBERTS ENTITLED "IMPROVED METHOD FOR DISTINGUISHING DETERMINISM FROM RANDOMNESS IN THE PRESENCE OF NOISE" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Martin P. Gelfand

Raymond A. Shaw

Richard D. Etters

Richard E. Eykholt

Adviser

David A. Krueger

Department Head/Director

Abstract of DISSERTATION

IMPROVED METHOD FOR DISTINGUISHING DETERMINISM FROM RANDOMNESS IN THE PRESENCE OF NOISE

This work presents an improved method for detecting determinism in time series. The method requires several input parameters, and a procedure for determining optimal values for these parameters is given. The method is applied to several types of noise and it is shown to be able to detect determinism for noise levels up to 70% – 80% of the magnitude of the signal of the dynamical system to which the noise is being added. The method is reliable for data sets as small as 256 data points. The method is also applied to data sets for which the source of noise is a chaotic system. In these cases, determinism is detected in the noise source itself.

Sean Roberts
Physics Department
Colorado State University
Fort Collins, CO 80523
Spring 2003

Contents

I. INTRODUCTION

The advent of chaos theory has led to an explosion in the study of nonlinear systems. Chaos theory has been used to describe a wide range of systems, such as the self-generation of the earth's magnetic field [1], nonlinear mode coupling of plasmas [2], solid state physics (e.g., dynamics of particles in an anharmonic potential in the presence of a periodic external field [3]), and the onset of turbulence in fluids [4–6], among many others.

Contrary to popular belief, chaos does not imply a lack of order; in particular, the presence of chaos does not imply a non-deterministic system. In fact, chaotic systems are by definition deterministic. The ability to classify a system as a deterministic system is important for the purposes of modeling and prediction. A random system can be modeled as a stochastic process, but it is possible to model a chaotic system by a set of coupled nonlinear differential equations. If one wishes to model a system that appears to be random, it is important to first determine if the system is a low-dimensional chaotic system, so that it can be modeled by a set of differential equations, rather than a stocastic process.

Although long-term prediction is unreliable for chaotic systems, short-term prediction is possible. The lyaponov exponent is a measure of the rate of separation of nearby trajectories in phase space, and, as such it gives a measure of the time scale for which reliable predictions can be made. The lyaponov exponent is only meaningful for deterministic systems. If one wishes to make predicitions for a given system, one must first determine whether or not the system is deterministic. If the system is found to be deterministic, lyaponov exponents can be calculated to determine a time scale at which predictions can be made.

Because the ability to model a system and make predictions is dependent on whether or not the system is deterministic, it is desireable to have a simple method to distinguish deterministic chaos from randomness. Such a method should be robust in the presence of noise. Some systems can only be measured infrequently,

so it is desirable to be able to detect determinism in small data sets. The method of Salvino and Cawley [7] is a method for distinguising determinism from randomness that is simple to implement. This dissertation presents extensions to this method that make it robust to the presence of noise and useful for small data sets.

Complex systems are often studied by periodically measuring a single observable of the system, which generates a time series. Much research has been done on analyzing such systems. One such avenue of research has focused on methods to detect low-dimensional deterministic chaos by analyzing these time series. The ability to detect determinism in time series has a wide range of applications in a wide range fields. The fields of study for which such methods have been used include physics, Earth science, biology, medical science, metallurgy, and engineering.

Studies in the field of physics include research on tunneling in asymmetric coupled quantum wells [8], lasers [9, 10], and topics in astronomy, including studies on black holes [11] and solar activity [12].

In the Earth sciences, work has been done in field of hydrology, namely studies of river flow and hydrological runnoff [13]-[21]. In the field of atmospheric science, determinism has been explored in studies of weather patterns [22–24], the magnetosphere [25], and ozone concentrations [26]. In the field of geology, studies on soil formation [27], seismic activity [28], and volcanic activity [29] have involved detecting determinism.

Studies in biological fields include a large volume of work in ecology on population dynamics [30]-[44], and several studies in pathology on disease epidemics [45]-[50].

The largest volume of work involving the detection of determinism in time series is in the medical fields. Possibly the most active area of research relates to the human brain via studies in the field of electroneurodiagnostic science, particularly studies of electroencephalograms [51]-[64], as well as work in the field of neurobiology, with studies of neural activity [65]-[71], rhythmic motor control [72] and epilepsy [73]. A good deal of work has also been done in the field of cardiology, in

studies of heart rhythm [74]-[85]. Methods to detect determinism have began to appear in other medical fields as well, such as optokinetics (eye movements) [86, 87], physiology [88, 89], and voice science [90].

Researchers have began to explore the possibility for detecting deterministic chaos in the fields of metallurgy, in studies of blast furnace dynamics [91], in studies on crystal growth [92], and in nautical engineering [93], traffic-flow dynamics [94, 95] and commerce [96].

Since detecting deterministic chaos in time series has such wide-ranging applications, there is a need for a method that is simple to implement, and that is robust to noise and applicable to small data sets. Existing methods for detecting determinism in time series can be classified into three broad categories. The first set of methods relies on measuring geometrical or dynamical quatities meaningful in the context of chaos, such as the fractal dimension, lyapunov exponents, etc. One such method was developed by Packard, Crutchfield, Farmer, and Shaw [97], and it uses the lyapunov exponent to classify a system as chaotic. These methods are unreliable for small data sets, due to the fact that the quantities they are computing (lyapunov exponent, fractal dimension, etc.) are defined as infinite time averages, but are measured over a finite data set [98]. Kennel and Isabelle [99] point out many other problems with these methods, such as biases in dimension calculations, and the effects of false correlations in correlation-dimension calculations. In addition these methods rely on correctly estimating embedding parameters, such as the index lag.

The next set of methods relies on the ability of chaotic systems to make short-term predictions of the future. The most popular of these methods was devised by Sugihara and May [100]. This method uses past patterns in a time series to predict future patterns. For dynamical chaos, the accuracy of the nonlinear forecast falls off with increasing prediction interval; whereas, for uncorrelated random data, the accuracy is independent of the prediction interval. In order to build up the library of past patterns needed to implement this method, a fairly large time series

is needed. In addition, correlated noise can mimic the prediction accuracy of deterministic chaos and cause this method to fail [101]. Finally, this method also requires a good estimate of the index lag.

Several methods related to the method of Sugihara and May have been developed to overcome some of these shortcomings. One of these methods was developed by Kennel and Isabelle [99], and it was the first method to use the concept proposed by Theiler [102] of using surrogate data as a null test. This idea has subsequently been used by many other methods for distinguishing determinism from randomness. Another method that is related to the nonlinear forecasting method of Sugihara and May was devised by Basu and Foufoula–Georgiou [101]. These methods overcome some of the problems associated with noisy systems; however, since they still rely on using past patterns to predict the future they are best used with larger data sets. The method of Kennel and Isabelle is reported to reliably distinguish determinism from randomness for noise levels on the order of 75% for a data set size of 16,384, while the method of Basu and Foufoula–Georgiou reported the ability of distinguish determinism from randomness for noise levels on the order of 50% for data sets of 10,000 data points.

The third, and most recent, set of methods rely on the fact that deterministic trajectories contain a degree of smoothness. The method in this dissertation expands on such a method.

The method presented in this dissertation has its roots in work of Kaplan and Glass [103] and Wayland, Bromley, Pickett, and Passamante [104]. Salvino and Cawley [7] later refined this work. They then used this method to detect deterministic chaos in an experiment by Ditto et al. [105] in which a gravitationally buckled, amorphous magnetoelastic ribbon was driven parameterically by a time-varying magnetic field. Later, their method was used to analyze an experiment on rhythmic finger tapping by Roberts, Eykholt, and Thaut [72], and to analyze electroencephalograms by Jeong, Kim, and Kim [61].

Until now, there has been no detailed investigation of the reliability of the

Salvino and Cawley method in the presence of noise or for small data sets. In fact, the studies on the effects of noise by Kennel and Isabelle and by Basu and Foufoula–Georgiou appear to be only preliminary in nature. This dissertation represents the first detailed quantification of the ability of any method to distinguish determinism from randomness in the presence of noise.

This dissertation also introduces refinements to the method of Salvino and Cawley that increases its reliability in the presence of noise and for small data sets. The ability to distinguish determinism from randomness is quantified, and the improved method is shown to be reliable for noise levels on the order of 100% for data sets as small as 256 data points. In addition, this work offers some practical advice for implementing the method.

The layout of this dissertation is as follows. Chapter II is a review of the method of Salvino and Cawley. Chapter III presents the extensions to the method of Salvino and Cawley. Chapter IV presents results on the ability of the extended method of Salvino and Cawley to detect determinism in the prescence of noise. Chapter V presents a summary and conclusions.

II. DETERMINISTIC CHAOS VERSUS RANDOMNESS

The time series in Fig. 1 appears to fluctuate randomly, but it is actually the result of a deterministic process, so it is chaotic, rather than truly random. This is an important distinction, since a chaotic time series contains a degree of order that is useful for prediction and modeling. This section reviews a test developed by Salvino and Cawley [7] to investigate whether a time series is deterministically chaotic or truly random. In future sections this technique will be expanded on, and the effects of noise on the ability to detect chaos will be explored.

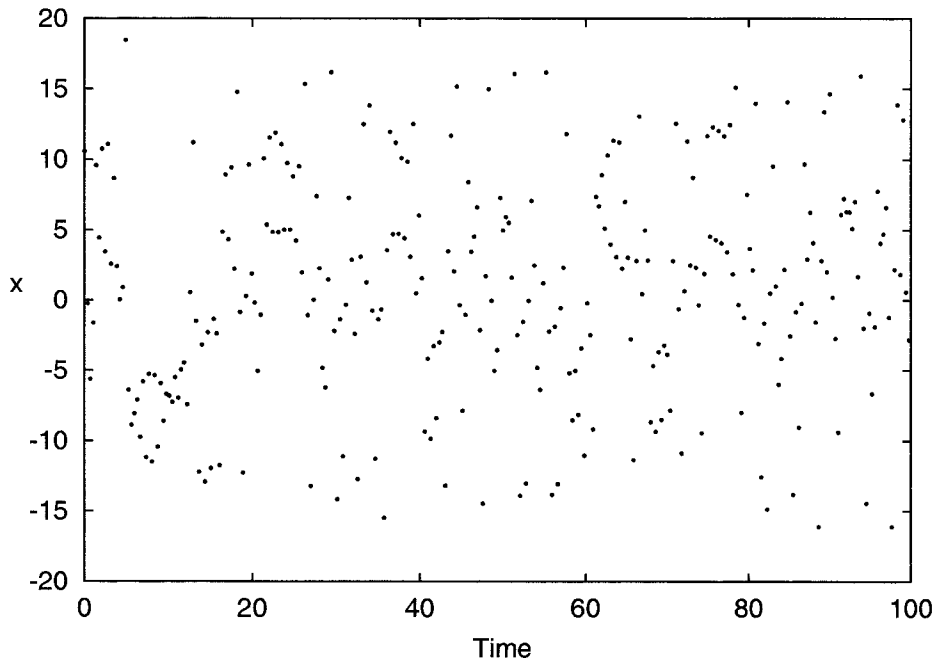

FIG. 1: Time series from a deterministic chaotic process.

A. Time Series and Embedding

Most systems in nature have a very large number of degrees of freedom, but such systems can often be modeled using only a few relevant variables. An example of such a system would be a gas consisting of 10^{23} molecules described using only a few variables, such as pressure, volume, and temperature. Once the relevant variables have been determined, the time evolution of the system may be described by a trajectory in a low-dimensional space with these relevant variables as the axes.

Suppose only a single variable is measured at regular time intervals. This type of sampling is called a time series. Takens [106] developed a method for using a time series to construct a representation of the full trajectory in another low-dimensional space which is topologically equivalent to the true trajectory. Such a representation is called an *embedding* of the trajectory, Takens' method for constructing an embedding from a time series of a single variable will now be reviewed.

To construct an m-dimensional embedding from a time series $\{x_i\}$, select an index lag τ, and form the m-dimensional vectors

$$\vec{v}_i = \left(x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\right) \ . \tag{1}$$

A plot of the vectors $\{\vec{v}_i\}$ in m dimensions gives the desired embedding of the true trajectory for the system. The quality of this embedding depends on the chosen values of m and τ. For example, if the space for the true trajectory is three-dimensional, but we construct a two-dimensional embedding, we will obtain a projection of the trajectory onto only two dimensions. Also, the index lag τ must not be too short or too long compared to the time scale of the dynamics. In the former case, all of the components of the vector $\vec{v}_i$ will be nearly the same, causing the embedded trajectory to be compressed along a single line. In the latter case, the components of the vectors $\vec{v}_i$ become uncorrelated, and the embedded trajectory does not represent the true underlying dynamics of the system. There is an extensive literature regarding the selection of appropriate values for m and

τ. However, the method of Salvino and Cawley does not require a high quality embedding, so these details need not be considered.

As an illustration, consider the trajectory in three dimensions generated by the three coupled Lorenz equations [4]:

$$\begin{aligned} \dot{x} &= \sigma(y - x) \ , \\ \dot{y} &= rx - y - xz \ , \\ \dot{z} &= -bz + xy \ . \end{aligned} \tag{2}$$

The trajectory shown in Fig. 2 is obtained by integrating these equations using a fourth-order Runge-Kutta method with a time step $\delta t = 0.001$, and with the standard parameter values $\sigma = 10$, $r = 28$, and $b = 8/3$. The time series $\{x_i\}$ shown in Fig. 3 is constructed from Fig. 2 by measuring the variable $x(t)$ at the times $0, T, 2T, \ldots$, where $T = 0.05 = 50\delta t$. Figure 1 is in fact a blowup of a portion of Fig. 3; it appears to be random, but, it is actually chaotic, since it was generated by the deterministic Lorenz equations. To reconstruct the trajectory from the single time series $\{x_i\}$ of Fig. 3, form the vectors $\vec{v}_i$ given by Eq. (1), and plot these vectors as points in m dimensions. Using $\tau = 3$ and $m = 3$ yields the embedding shown in Fig. 4, which is topologically equivalent to the true trajectory in Fig. 2.

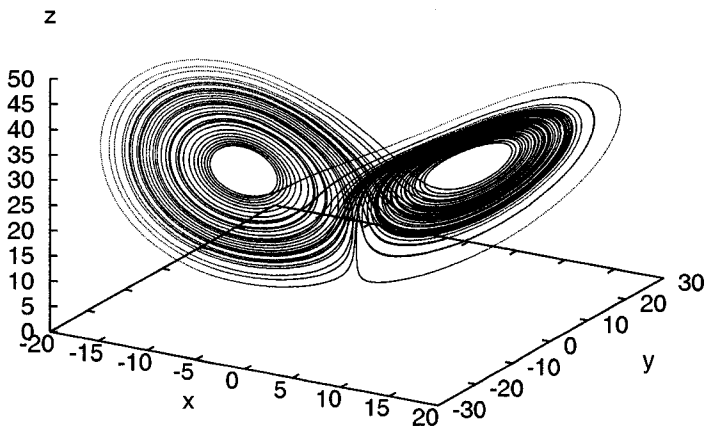

FIG. 2: Trajectory for the Lorenz system of Eqs. (2) for the parameter values $\sigma = 10$, $r = 28$, and $b = 8/3$.

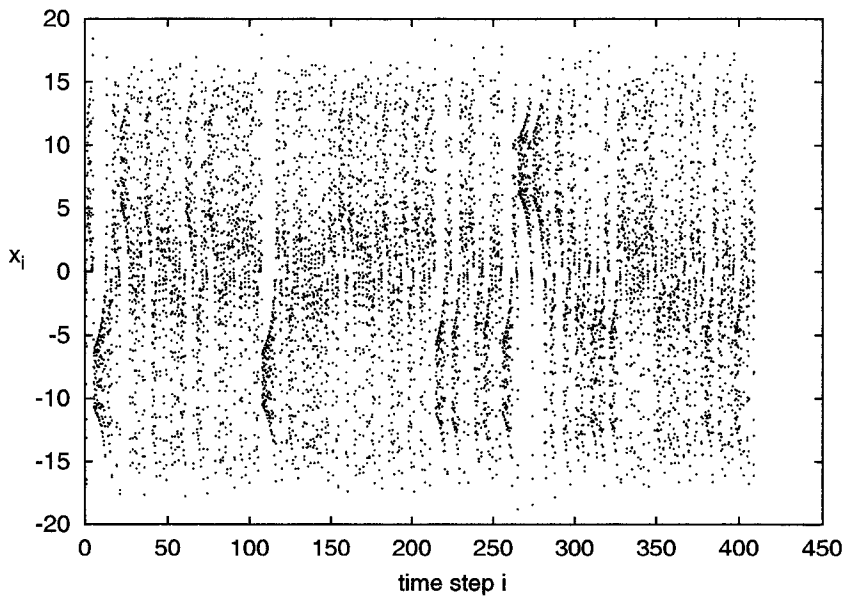

FIG. 3: Time series $\{x_i\}$ for the trajectory in Fig. 2.

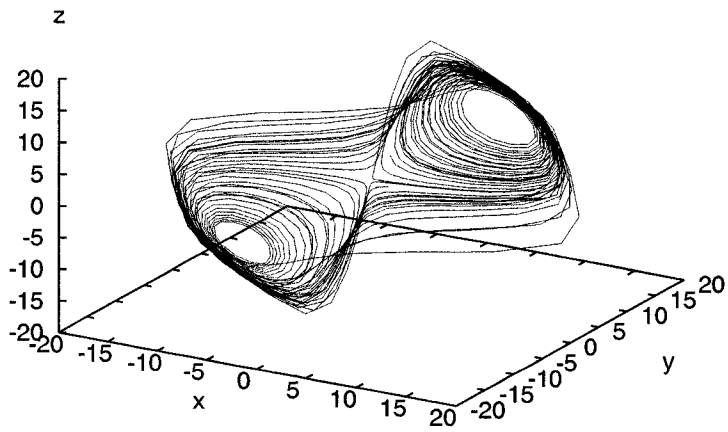

FIG. 4: Reconstruction of the trajectory in Fig. 2 using a three-dimensional embedding with an index lag of 3 time steps, generated from the time series $\{x_i\}$ of Fig. 3. Successive points have been connected by straight lines as a guide to the eye.

B. Method for Distinguishing Between Chaos and Randomness

The basis for the method of Salvino and Cawley is the fact that a deterministic process generates a smooth trajectory like the one in Fig. 2. As a result, the trajectory reconstructed from the embedding will have a degree of smoothness that can be used to distinguish deterministic chaos from randomness.

To illustrate the basic idea, consider the unit tangent vector at each point of an embedded trajectory. For a smooth trajectory, nearby points will have nearly the same unit vectors, and a local average of these vectors will produce a vector with nearly unit length. For a random trajectory, nearby points will produce unit tangent vectors that point in random directions, so a local average of these vectors will tend to cancel, yielding a net vector whose length is much less than one. Thus, the length of the local average of unit tangent vectors in a vector field may be used to distinguish deterministic chaos from randomness.

The choice of tangent vectors at each trajectory point is a simple example of a vector field; other assignments of unit vectors to points on the trajectory produce other vector fields. Salvino and Cawley constructed various vector fields by choosing five numbers c_0, c_1, c_2, c_3, and c_4, and assigning to each embedded trajectory point $\vec{v}_i$ the unit vector

$$\vec{\phi}_i = \sum_{r=0}^{4} c_r \vec{v}_{i+r} \Big/ \left| \sum_{r=0}^{4} c_r \vec{v}_{i+r} \right|. \tag{3}$$

They constructed ten different vector fields using the ten sets of coefficients listed in Table I. Note that the first vector field produces the vectors $\vec{\phi}_i = (\vec{v}_{i+1} - \vec{v}_i)/|\vec{v}_{i+1} - \vec{v}_i|$, which are roughly the same as the unit tangent vectors.

The m-dimensional embedding space is now partitioned into a uniform grid of small m-dimensional cubes. For each cube, the unit vectors $\vec{\phi}_i$ in that cube are averaged to obtain the vector

TABLE I: Coefficients for the ten different vector fields

	c_0	c_1	c_2	c_3	c_4
1	-1.0	1.0	0.0	0.0	0.0
2	-3.0	4.0	-1.0	0.0	0.0
3	2.0	-5.0	3.0	0.0	0.0
4	4.7	-3.0	-1.7	0.0	0.0
5	-2.0	3.0	-4.0	3.0	0.0
6	3.5	-2.7	-1.4	0.6	0.0
7	-3.4	-0.5	-0.1	4.0	0.0
8	1.0	2.0	3.0	-4.0	-2.0
9	0.9	0.8	-3.5	4.0	-2.2
10	3.0	-2.0	0.0	2.0	-3.0

$$\vec{Y}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \vec{\phi}_i \ , \tag{4}$$

where j labels the particular cube, and n_j denotes the number of trajectory points contained in cube j. If the embedding is smooth, the vectors $\vec{\phi}_i$ within a single cube will be approximately equal, so $|\vec{Y}_j|^2$ will be nearly one. However, if the embedding is not smooth, the vectors $\vec{\phi}_i$ within a single cube will point in many different directions, causing them to cancel in the sum, so $|\vec{Y}_j|^2$ will be much less than one. The smoothness of an embedded trajectory is classified by the global weighted average of $|\vec{Y}_j|^2$ over all the boxes of the grid

$$W = \sum_j n_j \, |\vec{Y}_j|^2 \Big/ \sum_j n_j. \tag{5}$$

A value of W near one indicates deterministic chaos, while a value of W much smaller than one indicates randomness.

This method relies on having several vectors in each grid cube, so, with a finite number of data points, these cubes cannot be too small. Since each grid cube must

be of finite size the vectors $\vec{\phi}_i$ will vary over the extent of the cube, even for a smooth deterministic trajectory, and W will be slightly less than one. Alternately, for a random trajectory the vectors within each box will not necessarily cancel completely, since the number of vectors in each box is finite, so W will be small, but nonzero. Because of these constraints, a concrete criterion is needed to decide whether W is close enough to one to indicate determinism or small enough to indicate randomness.

A simple way to test a time series for deterministism or randomness is to see whether or not randomizing the time series reduces the value of W significantly. When randomizing the time series, its basic distributional properties, such as its power spectrum [i.e., the magnitude of its Fourier transform $F(k)$] should not be altered. To randomize a data set, take the Fourier transform, randomize the phases $\theta(k)$, but leave the magnitudes $\mid F(k) \mid$ unchanged. The inverse Fourier transform is then taken to return to return to the original space. The resulting randomized data set is referred to as a surrogate data set. In order for the surrogate data set to remain real, the phase $\theta(k)$ must be odd about $k = 0$. Thus, for each $k > 0$, $\theta(k)$ is chosen to be a random number between $-\pi$ and π, and the remaining phases are then chosen by requiring that $\theta(-k) = -\theta(k)$ (and, thus, $\theta(0) = 0$). A Fast Fourier transform (FFT) routine was used to calculate the necessary Fourier transforms. The FFT used requires the data set to have a size N which is a power of 2. An alternate FFT could be used for data sets of arbitrary size.

For the surrogate data set, the values of W are computed for each of the ten vector fields in Table I. The range of these ten values indicates the range expected for random data sets. The ten values of W are then computed for the original data set. If the original data set is random, then these latter values of W should fall within the same range as the values for the surrogate data set. However, if the original data set is deterministic, then the largest of the ten values of W for the original data set should be significantly above the range of values for the surrogate data set. Since the quality of the embedding depends the index lag τ,

this procedure is repeated for a range of values of τ.

As an example, a data set consisting of $2^{13} = 8192$ successive values from the time series $\{x_i\}$ of Fig. 3 is analyzed. This time series was obtained from the Lorenz system of Eqs. (2) and used to produce the embedding in Fig. 4. For this example, the time series is embedded in $m = 3$ dimensions and coarse grained into a grid of $40 \times 40 \times 40$ boxes. It should be noted that most of these boxes are empty, and, in computing W, only those boxes that contain more than some minimum number of embedded points are considered. Later, the reason for specifying a minimum number of embedding points in a box will be discussed in more detail, but, for now, the minimum number of embedded points in a box will be taken as three. The results are shown in Fig. 5, where the lower two curves are the maximum and minumum values of W for the surrogate data set, and the upper curve is the maximum value of W for the original data set. The maximum value of W for the original data is well removed from the the maximum value of W for the surrogate data set. This demonstrates that the original data set was generated by a deterministic process.

In general, the separation between the upper two curves is not this dramatic. As another example, consider a time series $\{t_i\}$ that consists of the differences in the times at which $x(t)$ for the Lorenz system [Eqs. (2)] reaches a local maximum. This time series is shown in Fig. 6. It samples the trajectory less frequently than the first example, so the evidence of deterministic chaos should be weaker. This is a more realistic analogue to many experiments, since it consists of the times at which a particular event occurs. The results of this latter analysis are shown in Fig. 7. The number of data points is again 8192, the embedding dimension is $m = 3$, and the grid consists of $10 \times 10 \times 10$ boxes. As in the previous example, the lower two curves give the maximum and minimum values of W for the surrogate data set, and the upper curve gives the maximum value of W for the original data set. In this case, the maximum value of W for the original data set is not nearly as close to one as in the previous example. More importantly the separation

between the upper two curves is less than in the previous example, although it is still well-pronounced.

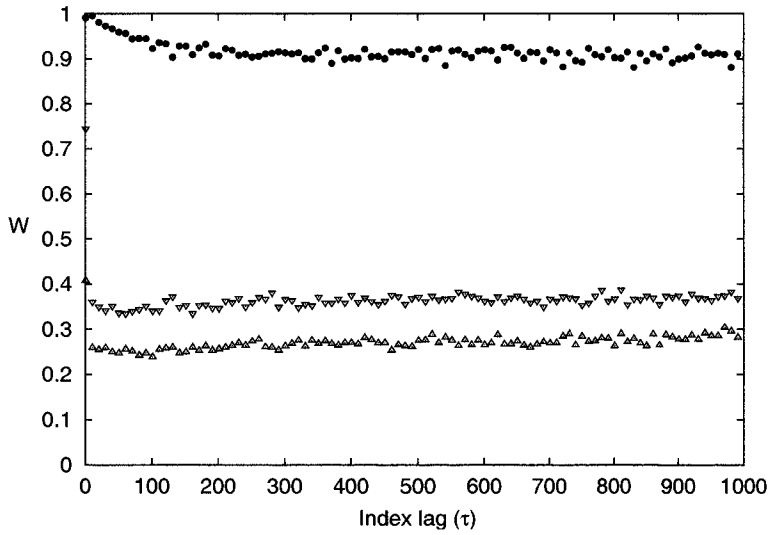

FIG. 5: W versus τ for the Lorenz system of Eqs. (2). The lower two curves are the maximum and minumum values of W for the surrogate data set, and the upper curve is the maximum value of W for the original data set.

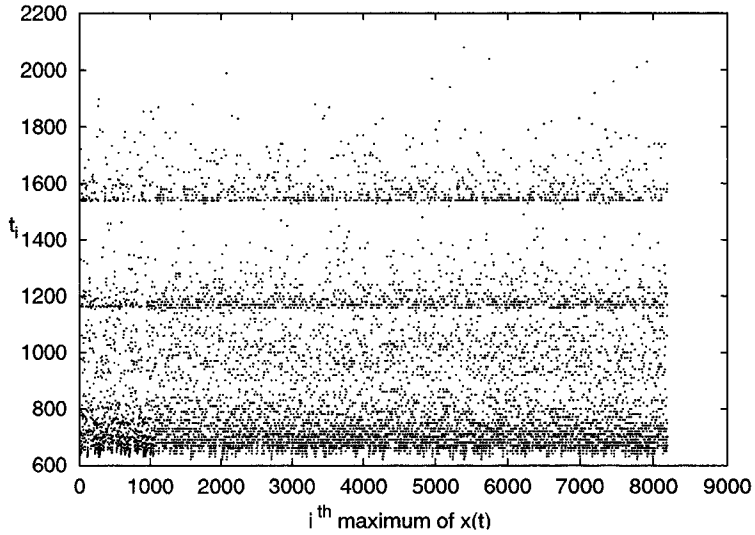

FIG. 6: Time series $\{t_i\}$ generated from the differences in the times when the x component of the trajectory in Fig. 2 reaches a local maximum.

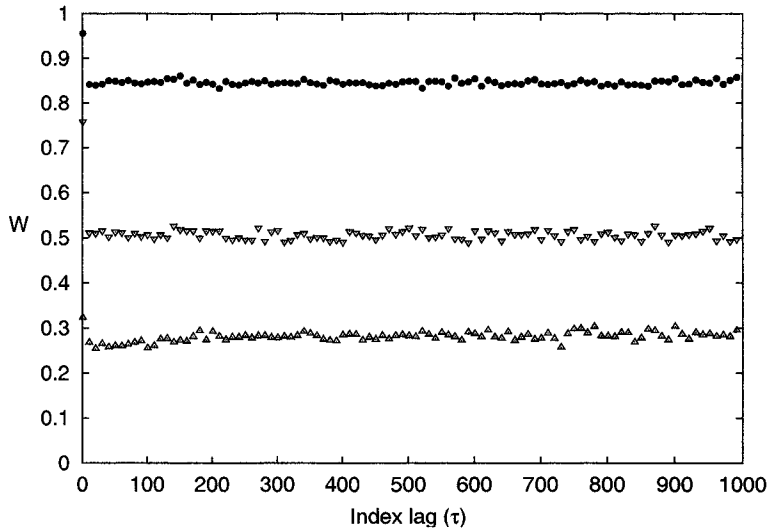

FIG. 7: W versus τ for the time series shown in Fig. 6. The lower two curves are the maximum and minimum values of W for the surrogate data set, and the upper curve is the maximum value of W for the original data set.

III. EXTENDING THE METHOD OF SALVINO AND CAWLEY

This section will discuss our extensions to the method of Salvino and Cawley that make it more useful for small and/or noisy time series. Subsection III A presents a procedure for determining optimal values of the various parameters, and it discusses several details that must be addressed when implementing the method. Subsection III B discusses a problem with the surrogate data for small data sets and a method for overcoming this problem. Subsection III C presents a method for smoothing the W curves, which is beneficial when looking for deterministic chaos in the prescence of noise. Finally, subsection III D gives a method for characterizing the separation between the W curves for real and surrogate data.

A. Determining Optimal Parameter Values for the Salvino and Cawley Method

There are a number of parameters used in the Salvino and Cawley method. In this section, a method for estimating optimal values for these parameters is described. These parameters include the size δ of the grid cubes used to coarse grain the embedded trajectory, a minimum number p of trajectory points that must fall in a given grid cube, and a minimum number b of sufficiently visited grid cubes.

The method of Salvino and Cawley relies on the fact that a local region of the embedding is visited multiple times. Therefore, the most important parameter to determine is the size of the grid to use when coarse graining the embedded trajectory. If the grid is too fine, individual boxes of the grid will not be visited enough times to produce a statistically significant value for the average of the unit vectors $\vec{\phi}$. For example, consider a time series consisting of random numbers. If a very small grid is chosen, each box in the grid will be visited only once. For this case the average of the unit vectors $\vec{\phi}$ will be itself a unit vector, and, thus, $W = 1$,

indicating determinism, when, in fact, the original data set was random. If the grid is too large, a grid cube will cover too large a portion of the attractor, and a local property of the trajectory is no longer being measured. In this case, even for determinism, the unit vectors $\vec{\phi}$ in a single box will vary enough that their average won't be a unit vector.

Even if an optimal size δ for the grid is chosen, there will still be boxes that are not visited at all, and others that contain only one or two embedded points. For random data, averaging the unit vectors $\vec{\phi}$ for these infrequently visited boxes will not produce a cancellation, and the average vector may have a length close to one. These boxes should not be considered in the calculation of W. Therefore, it is required that a box contain a minimum number p of points in order to be considered.

Finally, the method of Salvino and Cawley requires a global average of the lengths of the local averages of the unit vectors $\vec{\phi}$. For this average to be truly global, there must be a minimum number b of sufficiently visited boxes. Index lags for which W cannot be computed are termed invalid index lags, while index lags for which W can be computed are termed valid index lags.

Figure 8 shows the effects of coarse graining with various grid sizes. This figure was generated from 2048 points of a time series of the the x coordinate of the Lorenz system integrated using a fourth-order Runge-Kutta method with a time step of .001 and sampled every 50 time steps. This time series was embedded in 3 dimensions with an index lag of 5. Each plot in Fig. 8 is the most often visited grid cube of the coarse grained embedding for various grid sizes. The vectors in this figure are the unit tangent vectors to the embedded trajectory (field 1 of Table I). Figure 8a shows a large number of points falling in the cube, but the tangent vectors change as we change location within the box. In this case, the grid box is not small enough to give a local property of the attractor. Figs 8b-d show that the number of trajectory points that fall in a given cube decrease as the grid size decreases, while the directions of the unit vectors become more consistent.

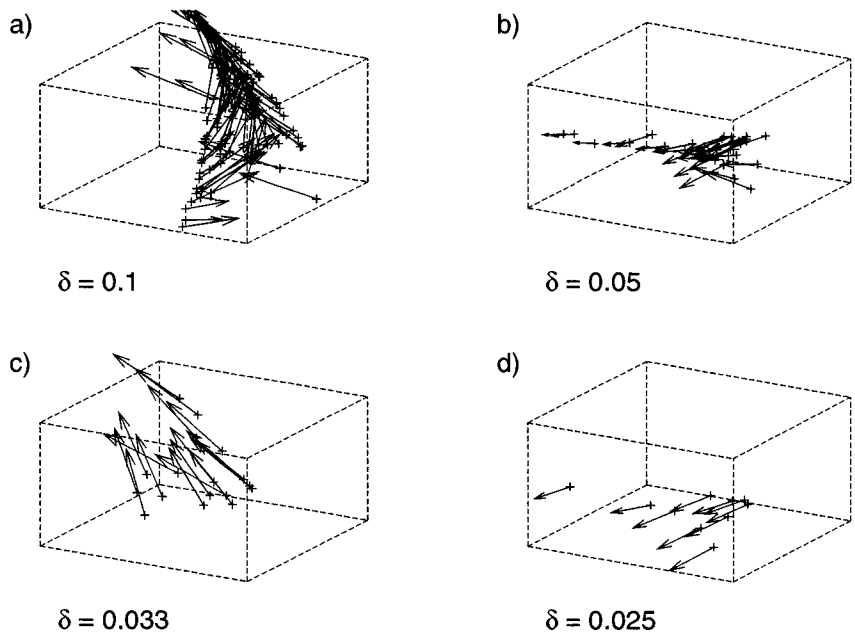

FIG. 8: Effects of coarse graining with various grid sizes. Each plot contains the most visited cube in the grid for that grid size. a) A $10 \times 10 \times 10$ grid covering the attractor ($\delta = 0.1$). b) A $20 \times 20 \times 20$ grid covering the attractor ($\delta = 0.05$). c) A $30 \times 30 \times 30$ grid covering the attractor ($\delta = 0.033$). d) A $40 \times 40 \times 40$ grid covering the attractor ($\delta = 0.025$).

The choice of the grid size δ is affected by two constraints. The first constraint is that a box of the grid is viable only if it is sufficiently visited — at least p times. The second constraint is that there must be at least b of these boxes. An optimal set of parameters is a minimal grid size δ, with maximal visitation constraints p and b, that yield the best coverage of the embedding.

Optimal values of the parameters are calculated by creating an m-dimensional embedding with a given index lag τ. The embedding is covered by a uniform grid of m-dimensional cubes of size δ. For this index lag and grid size, the number of times p_i that box i is visited is counted. The total number of sufficiently visited

boxes is

$$B(\delta, \tau, p) = \sum_i \Theta(p_i - p), \tag{6}$$

where Θ is the unit step function.

$B(\delta, \tau, p)$ is computed for ranges of values of p, τ and δ. To choose these ranges, the user must descide the largest possible grid size δ and smallest possible visitation constraints p and b to be considered. A largest grid size of $\delta = 0.125$ (8 boxes on a size), and smallest visitation constraints of $p = 3$ and $b = 3$ seem to be sufficient.

For given values of δ and p, there may be some values of the index lag τ for which the total number of sufficiently visited boxes $B(\delta, \tau, p)$ will be less than the minimum allowed b. For these index lags a grid of size δ is not a good coverage of the embedding. The number of index lags that are viable is

$$N(\delta, p, b) = \sum_\tau \Theta(B(\delta, \tau, p) - b). \tag{7}$$

The grid that is chosen must be independent of the index lag. Therefore instead of using the number of sufficiently visited boxes for a given index lag consider the average number of sufficiently visited boxes per index lag. This average is only over those index lags for which there are a sufficient number of viable boxes.

$$\bar{B}(\delta, p, b) = \sum_\tau B(\delta, \tau, p)\, \Theta(B(\delta, \tau, p) - b) \Big/ N(\delta, p, b) \tag{8}$$

Both the number of viable index lags $N(\delta, p, b)$ and the average number of viable boxes $\bar{B}(\delta, p, b)$ are measures of the coverage of the embedding. Any linear combination of $N(\delta, p, b)$ and $\bar{B}(\delta, p, b)$ is another measure of the coverage of the embedding. This is the metric used to determine the parameters that yield the best coverage of the embedding.

Since $N(\delta, p, b)$ and $\bar{B}(\delta, p, b)$ have different units they must be normalized before they are used in a linear combination. They will be normalized using maxima over all parameter values.

The largest number of viable index lags $N_{\max}$ found over all the parameters is

$$N_{\max} = \max_{\delta,p,b} \left(N\left(\delta, p, b\right) \right), \tag{9}$$

and the greatest average number of viable boxes per index lag $B_{\max}$ is

$$B_{\max} = \max_{\delta,p,b} \left(\bar{B}\left(\delta, p, b\right) \right). \tag{10}$$

A linear combination of $N\left(\delta, p, b\right)$ and $\bar{B}\left(\delta, p, b\right)$ is

$$S\left(\delta, p, b\right) = \left(\omega_\tau \cdot \frac{N\left(\delta, p, b\right)}{N_{\max}} + \omega_a \cdot \frac{\bar{B}\left(\delta, p, b\right)}{B_{\max}} \right) \Big/ \left(\omega_\tau + \omega_a\right). \tag{11}$$

where ω_τ and ω_a are weighting factors. These are input parameters by the user. However, the grid size produced by this method is not very sensitive to the input values, so $\omega_\tau = 1$ and $\omega_a = 0$ were used in this dissertation.

The grid size that yields the largest value of $S\left(\delta, p, b\right)$ is the best covering of the embedding. The maximum of $S\left(\delta, p, b\right)$ is

$$S_{\max} = \max_{\delta,p,b} \left(S\left(\delta, p, b\right) \right). \tag{12}$$

It is possible that $S\left(\delta, p, b\right)$ will be equal to $S_{\max}$ for several different combinations of δ, p, and b. Any of these combinations will give a good coverage of the embedding. However, the smaller the grid, the better, so, from among these combinations of δ, p, and b, choose those combinations with the smallest grid size $\delta_{\min}$. From these remaining combinations, choose the largest p, and then the largest b. This yields the desired optimal parameters $\delta_{\min}, p_{\max}, b_{\max}$.

B. Surrogate Data for Small Data Sets

When coarse graining an embedding, many boxes will not be visited. A surrogate data set tends to be more evenly distributed over the volume of the attractor than the original data set from which it was created. Therefore, more boxes will be visited by the surrogate data set than by the original data set, so fewer points will

fall in each visited box. For small data sets, this leads to boxes not being visited enough to give good statistics for the averages of unit vectors.

To address this problem, one creates multiple surrogate data sets from the same original data set and coarse grains the embedding of each of these surrogate data sets using the same parameters ($\delta_{\min}$, $p_{\max}$, $b_{\max}$) that were used for the original data set. W is then calculated for each of the ten vector fields in Table I for each of the surrogate data sets.

For the j^{th} surrogate data set and an index lag of τ, let $W_{r,j}(\tau)$ denote the value of W for the r^{th} vector field, and let $B(\delta_{\min}, \tau, p_{\max})_j$ denote the number of boxes in this embedding with greater than $p_{\max}$ points. For each r and τ, calculate the weighted average $W(\tau)$ over all of the surrogate data sets for which this index lag was valid:

$$W_r(\tau) = \sum_j W_{r,j}(\tau)\,\Theta\left(B(\delta_{\min}, \tau, p_{\max})_j - b_{\max}\right) \Big/ \sum_j \Theta\left(B(\delta_{\min}, \tau, p_{\max})_j - b_{\max}\right). \tag{13}$$

For data set sizes smaller than 2^{11}, five surrogate data sets were generally enough to produce a sufficient number of viable index lags. For data set sizes of 2^{11} and 2^{12} three surrogate sets were sufficient, and one surrogate data set was sufficient for data sets larger than 2^{12}.

Figure 9 is an example for a surrogate data set of 128 points. Figure 9a is for a single surrogate data set. Figure 9b is generated from the average of five surrogate data sets. W could be computed for considerably more index lags when the average of five surrogate data sets was used than when only a single surrogate data set was used.

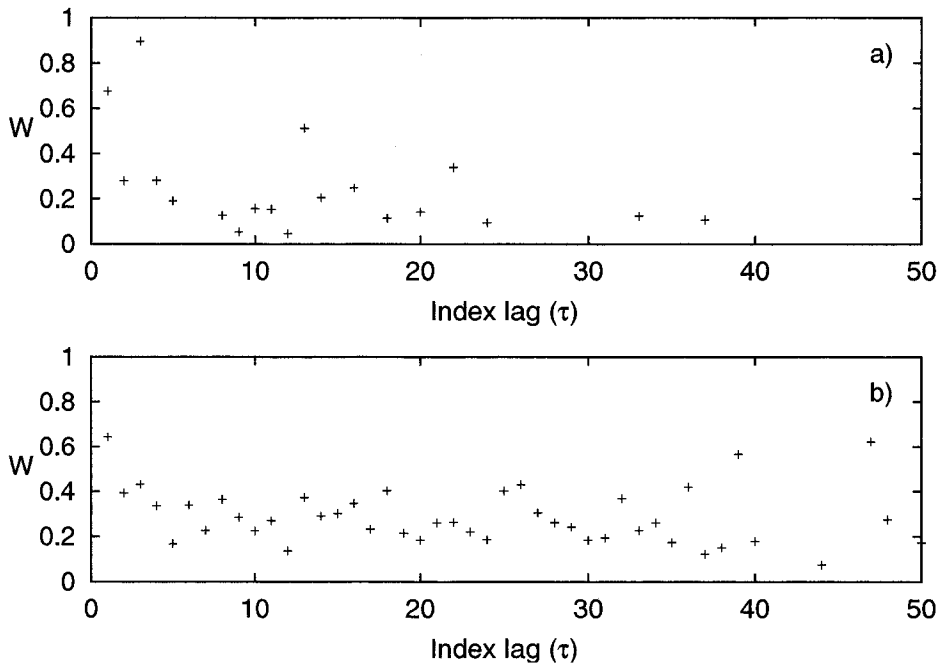

FIG. 9: Calculation of W for small surrogate data sets. a) A single surrogate data set. b) An average over five surrogate data sets.

C. Data Smoothing

In section II B, a method for distinguishing between chaos and randomness was presented. This method requires a comparison of $W(\tau)$ between real and surrogate data sets.

In practice, W varies from index lag to index lag. This is especially pronounced in the presence of noise. These variations arise from two sources. First is an overall functional dependence of W on the index lag τ. Second, W will have small irregular deviations from this functional dependence. The method requires a comparison of the functional dependence of the two curves to determine if a time series is deterministic or random. Irregular deviations in $W(\tau)$ can cause an intermingling of the real and surrogate data that can make drawing conclusions difficult in practice. Figure 10 shows such a case. The closed circles are W for a real data set, and the open squares are for a surrogate data set. The values of W for most index lags lie above those for those for the surrogate data. However, for many index lags the real and surrogate data are intermingled.

To obtain a good comparison between the two curves, the irregular deviations from the functional dependence of W must be minimized. This is achieved by *locally smoothing* the W versus τ curves. The term *locally smoothing* is defined as a process by which the difference between values of the function $W(\tau)$ for nearby values of τ is minimized. This process takes the function $W(\tau)$ defined on a set S of index lags and transforms it to a smoother function $\mathcal{W}(\tau)$ defined on a different set $\mathcal{S}$ of index lags:

$$W(\tau) : S \Rightarrow \mathcal{W}(\tau) : \mathcal{S}. \tag{14}$$

For data from a deterministic process, the displacement of $\mathcal{W}(\tau)$ for the real data from $\mathcal{W}(\tau)$ for the surrogate data will be more distinct than for the unsmoothed functions $W(\tau)$.

The discussion of *locally smoothing* is broken into 4 subsections. In subsection 1, the local neighborhood is defined. Subsection 2 presents the procedure used to

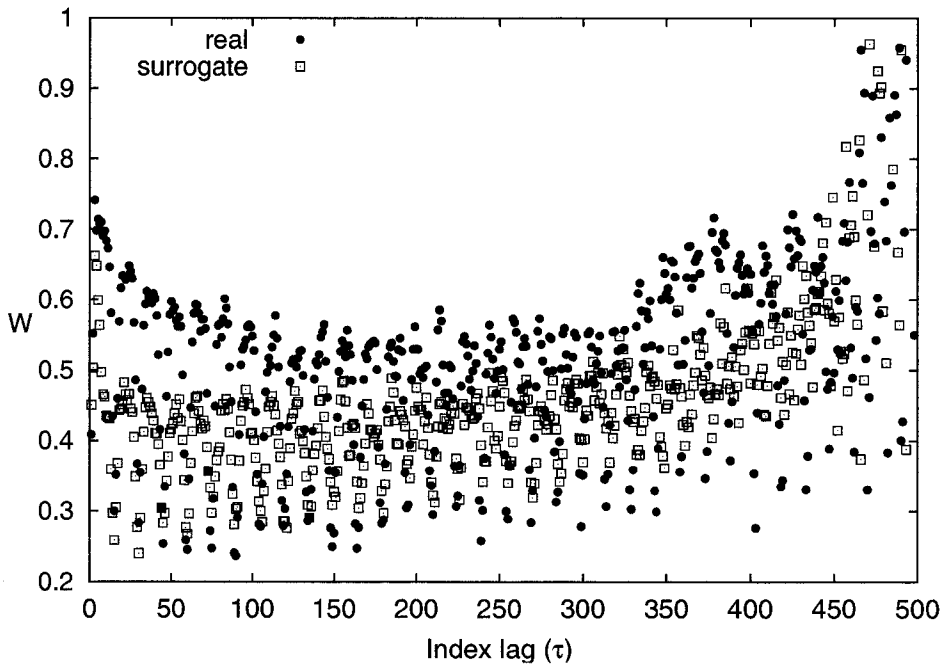

FIG. 10: Intermingling of $W(\tau)$ for real and surrogate data due to irregular deviations in $W(\tau)$. The closed circles are W for a real data set, and the open squares are for a surrogate data set.

perform local smoothing. Subsection 3 defines a metric to measure the effects of *locally smoothing* a data set. Subsection 4 examines the behavior of this metric to estimate an optimal local neighborhood size.

1. *Definition of the Local Neighborhood*

Smoothing the function $W(\tau)$ involves taking a local average of W over several nearby index lags τ. This section defines the local neighborhood over which the average is performed.

The function $W(\tau)$ is defined for a set S of index lags, the smallest of which is τ_0, and the largest of which is τ_L, subject to the constraints of section III A. These constraints were that there are at least $b_{\max}$ usable boxes in the coarse grained

embedding, and that each box contained at least $p_{\max}$ embedding points.

The local neighborhood about a given index lag τ is defined as the set Λ of at most τ_n index lags for which $W(\tau)$ is defined:

$$\Lambda(\tau) = \{\tau \mid \tau \in [\tau_-, \tau_+] \text{ and } B(\delta_{\min}, \tau, p_{\max}) \geq b_{\max}\}, \tag{15}$$

where $B(\delta_{\min}, \tau, p_{\max})$ is given by Eq. (6) in section III A, τ_- is the left endpoint of the local neighborhood, and τ_+ is the right endpoint of the local neighborhood:

$$\tau_- \equiv \begin{cases} \tau - \frac{\tau_n - 1}{2} & (\tau_n \text{ odd}) \\ \tau - \frac{\tau_n - 2}{2} & (\tau_n \text{ even}) \end{cases} \tag{16}$$

$$\tau_+ \equiv \begin{cases} \tau + \frac{\tau_n - 1}{2} & (\tau_n \text{ odd}) \\ \tau + \frac{\tau_n}{2} & (\tau_n \text{ even}) \end{cases} \tag{17}$$

For index lags near the endpoints of $[\tau_0, \tau_L]$ this could lead to $\tau_- < \tau_0$ or $\tau_+ > \tau_L$, which is not acceptable. This restricts the set $\mathcal{S}$ over which $\mathcal{W}(\tau)$ is defined to

$$\mathcal{S} = \{\tau \mid \tau \in [\tau_<, \tau_>] \text{ and } B(\delta_{\min}, \tau, p_{\max}) \geq b_{\max}\}, \tag{18}$$

where $\tau_<$ and $\tau_>$ are defined as

$$\tau_< \equiv \begin{cases} \tau_0 + \frac{\tau_n - 1}{2} & (\tau_n \text{ odd}) \\ \tau_0 + \frac{\tau_n - 2}{2} & (\tau_n \text{ even}) \end{cases} \qquad \tau_> \equiv \begin{cases} \tau_L - \frac{\tau_n - 1}{2} & (\tau_n \text{ odd}) \\ \tau_L - \frac{\tau_n}{2} & (\tau_n \text{ even}) \end{cases} \tag{19}$$

The endpoints τ_- and τ_+ of the local neighborhood about τ in Eqs. (16) and (17) in terms of $\tau_<$ and $\tau_>$ are

$$\tau_- = \tau + \tau_0 - \tau_< \tag{20}$$

$$\tau_+ = \tau + \tau_L - \tau_> \tag{21}$$

2. Smoothing Procedure

For each vector field r in Table I and each index lag in the set $\mathcal{S}$, W is smoothed by averaging over all index lags in the local neighborhood $\Lambda(\tau)$:

$$\bar{W}_r(\tau, \tau_n) = \sum_{\tau' \in \Lambda(\tau)} W_r(\tau') \Big/ \mid \Lambda(\tau) \mid, \tag{22}$$

where $\mid \Lambda(\tau) \mid$ is the number of elements of the set Λ. The error in $\bar{W}_r(\tau, \tau_n)$ is the standard error, which is denoted by $\sigma_r(\tau, \tau_n)$.

An example of this procedure is shown in Fig. 11 for an index lag of $\tau = 9$, for a local neighborhood of size $\tau_n = 5$, and for the first vector field in Table I. The two vertical lines represent the extent of the local neighborhood about $\tau = 9$. The solid dots are the original data points, and the open square at $\tau = 9$ is the average $\bar{W}$ of all the W's in the local region. The error bar is the standard error σ.

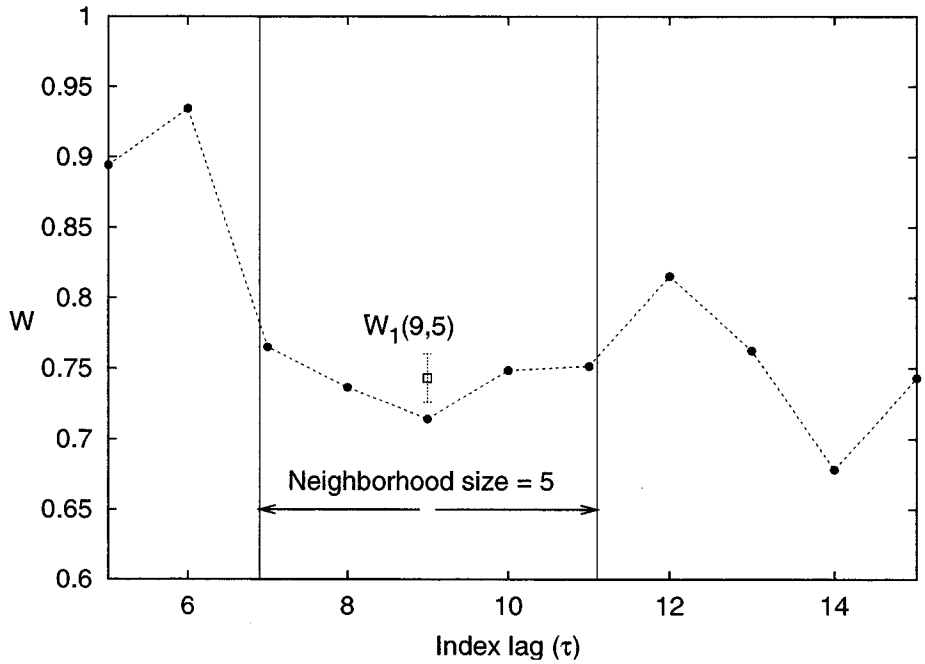

FIG. 11: Local smoothing of W for the first vector field of Table I for $\tau = 9$ and a local neighborhood size of $\tau_n = 5$. The solid dots are the values of W. The 5 points that fall between the vertical lines are averaged to determine $\bar{W}_1(9, 5)$, which is indicated by the open square.

In the next subsection, determining an optimal local neighborhood size $\bar{\tau}_n$ will be discussed. The value of the smoothed function $\mathcal{W}$ for a given index lag τ is given by the value of the smoothed maximum $\bar{W}_{\max}$ at that index lag for the optimal local neighborhood size $\tau_n = \bar{\tau}_n$.

The optimal local neighborhood size is determined by examining the behavior of $\bar{W}$ as a function of local neighborhood size τ_n. Each vector field r of Table I, as well as the maximum $\bar{W}$ from all vector fields, is used to determine the optimal local neighborhood size.

3. A Metric for Smoothness

In order to smooth the function W, a size τ_n for the local neighborhood must be chosen. If too large a local neighborhood is chosen, the functional variation of $W(\tau)$ may get averaged out. Also, the number of elements of $\mathcal{S}$ over which the smoothed function $\mathcal{W}$ is defined decreases as the size of the local neighborhood increases [see Eqs.(18) and (19)]. Thus, for large local neighborhood sizes, the resulting smoothed function $\mathcal{W}$ will be defined at too few points to allow a useful comparison between real data and surrogate data. On the other hand, if the local neighborhood is too small, the averaging will not have much of an effect on the irregular deviations from the functional variation of $W(\tau)$. In this section, a metric is defined that measures the smoothness of $\bar{W}$, and that can be used to choose a local neighborhood size.

Figure 12 demonstrates the effects of the size of the local neighborhood on the smoothness of the function $\bar{W}$. In each plot, the solid line is the original unsmoothed function, and the dashed line is the smoothed function for a given local neighborhood size τ_n. In Fig. 12a, $\tau_n = 3$, and there are still noticeable irregular variations between adjacent index lags for the smoothed function, which means that $\tau_n = 3$ is too small. As the local neighborhood τ_n size increases, these variations decrease. In Fig. 12d, for $\tau_n = 18$, the smoothed function is beginning to diverge from the original function, so $\tau_n = 18$ is too large.

The smoother a function is, the less variation there is in $\bar{W}_r$ between successive index lags. Thus, a measure of smoothness is the average difference in $\bar{W}_r$ between

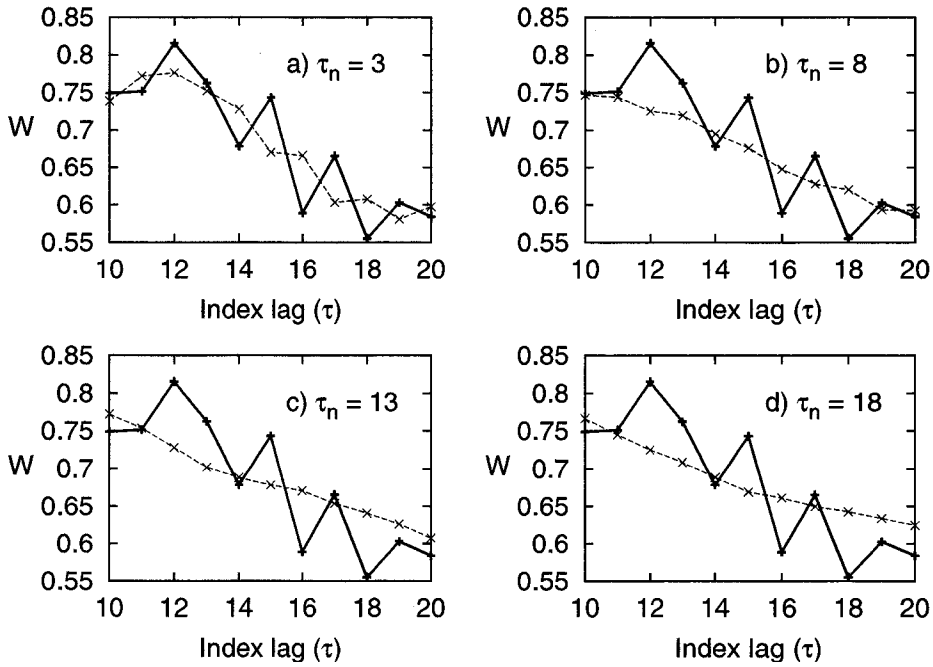

FIG. 12: Comparison of original and smoothed functions for different sizes τ_n of the local neighborhood. The solid line in each plot is the original function W, and the dashed line is the smoothed function $\mathcal{W}$data set.

adjacent index lags. The difference for two adjacent index lags is

$$\Delta\bar{W}_r(\tau,\tau_n) = \left|\bar{W}_r(\tau,\tau_n) - \bar{W}_r(\tau-1,\tau_n)\right|. \tag{23}$$

$\bar{W}_r$ is defined on the set of index lags $\mathcal{S}$, so $\Delta\bar{W}_r$ is defined over a set Υ of index lags for which both τ and $\tau-1$ are contained in $\mathcal{S}$:

$$\Upsilon = \{\tau \in \mathcal{S} \mid \tau - 1 \in \mathcal{S}\} \tag{24}$$

The average difference in $\bar{W}_r$ between adjacent index lags is then

$$\xi_r(\tau_n) = \sum_{\tau' \in \Upsilon} \Delta\bar{W}_r(\tau',\tau_n) \Big/ \mid \Upsilon \mid . \tag{25}$$

ξ_r will be refered to as the smoothness metric. It measures the smoothness of a locallaly smoothed function.

4. Determination of an Optimal Local Neighborhood Size

The function $W(\tau)$ is locally smoothed for a range of local neighborhood sizes $[\tau_{n,i}, \tau_{n,f}]$. The smoothing metric ξ_r is then plotted as a function of local neighborhood size, and this can be used to estimate an optimal local neighborhood size $\bar{\tau}_n$. The optimal local neighborhood size is the one for which larger values of τ_n do not produce a smoother function $\bar{W}$.

As an example, the functional dependence of $\xi_r(\tau_n)$ is shown in Fig. 13 for two vector fields from Table I, and for neighborhood sizes from $\tau_{n,i} = 8$ through $\tau_{n,f} = 25$. Both curves in Fig. 13 show that a local neighborhood size $\tau_n = 15$ produces the smoothest function $\bar{W}$ possible.

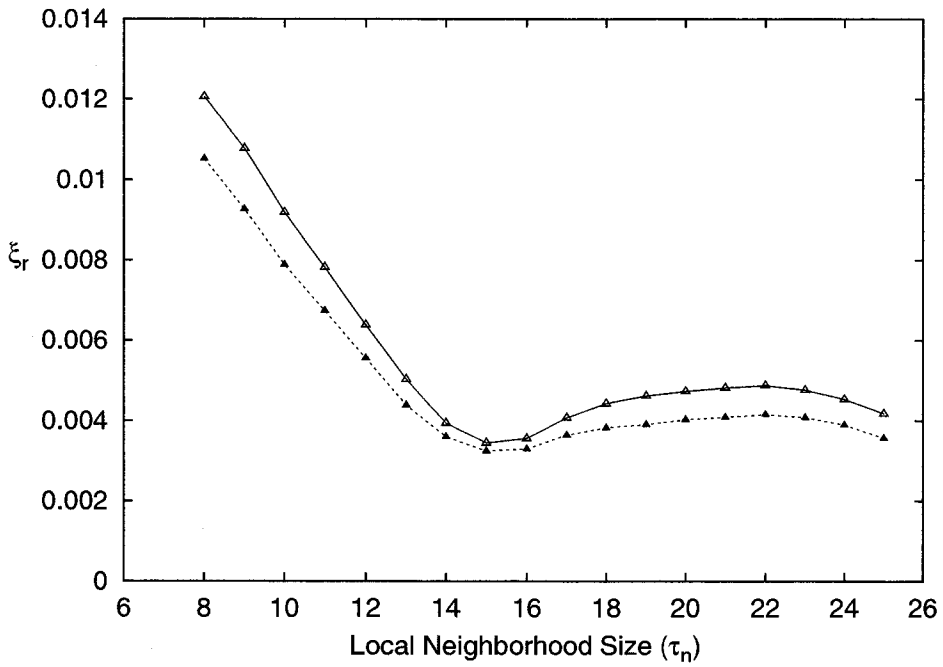

FIG. 13: Smoothness metric ξ_r versus local neighborhood size τ_n. The two curves are for two different vector fields from Table I. In this figure, $\xi_r(\tau_n)$ is calculated for neighborhood sizes from $\tau_{n,i} = 8$ through $\tau_{n,f} = 25$.

Three tests are used to determine the optimal local neighborhood size $\bar{\tau}_n$. Let the result of test m performed on ξ for the r^{th} vector field be denoted $\bar{\tau}_n(m,r)$.

Test 1: $\bar{\tau}_n(1,r)$ is the value of τ_n at the first local minimum of $\xi_r(\tau_n)$.

Test 2: $\bar{\tau}_n(2,r)$ is the value of τ_n at which $\xi_r(\tau)$ drops to $1/e$ of its initial value.

Test 3: $\bar{\tau}_n(3,r)$ is the value of τ_n at which the slope of $\xi_r(\tau)$ drops to $1/e$ of its initial value.

These tests are used to determine when a function has effectively dropped off to zero, and they are used extensively in the study of nonlinear dynamics (e.g., they are used to analyze the autocorrelation function and the mutual information).

The optimal local neighborhood size $\bar{\tau}_n$ is the smallest local neighborhood size obtained from the three tests described above:

$$\bar{\tau}_n = \min_{m,r}\left(\bar{\tau}_n(m,r)\right). \tag{26}$$

The first two tests are straightforward to implement, but the third test requires some explanation. All slopes are calculated using a least-squares fit. Let $\Psi_r(\tau_i,\tau_f)$ denote the slope of a least-squares fit to the points from $(\tau_i,\xi_r(\tau_i))$ through $(\tau_f,\xi_r(\tau_f))$. The region from which the initial slope $\mu_{0,r}$ is calculated is determined from $\tau_{n,i}$ through $\bar{\tau}(2,r)$ obtained from Test 2:

$$\mu_{0,r} = \Psi_r\left(\tau_{n,i},\bar{\tau}(2,r)\right). \tag{27}$$

If ξ_r never drops to $1/e$ of its initial value the entire function $\xi_r(\tau_n)$ from $\tau_{n,i}$ through $\tau_{n,f}$ is used to determine the initial slope.

Figure 14 illustrates the process of finding the initial slope $\mu_{0,r}$ from Fig. 13. The horizontal lines in Fig. 14 are $\xi_r(8)/e$. The first point below $1/e$ of the initial value for both curves occurs at $\bar{\tau}_n(2,r) = 14$. The initial slope for both curves is computed from $\mu_{0,r} = \Psi_r(8,14)$.

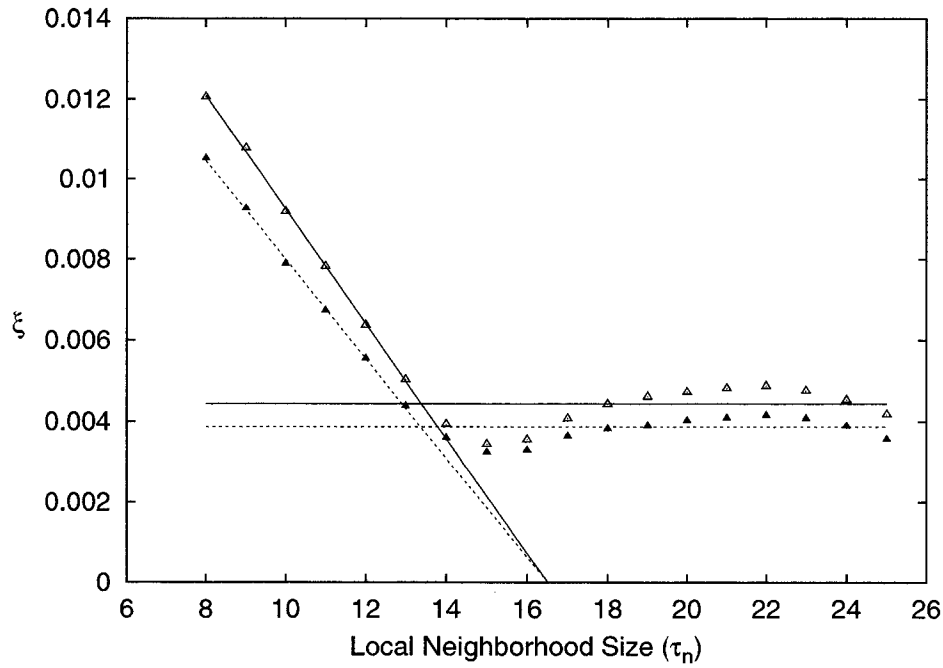

FIG. 14: Determination of the initial value of the slope of $\xi_r(\tau_n)$. A line is fit to the first several points. The number of points to use is determined by Test 2.

Once the initial slope is calculated, the slope at each $\tau_n \in [\tau_{n,i}+1, \tau_{n,f}-1]$ is calculated from

$$\mu_r(\tau_n) = \Psi_r(\tau_n - 1, \tau_n + 1). \tag{28}$$

Figure 15 shows $\mid \mu_r(\tau_n) \mid$. For both curves, the slope drops below $1/e$ of the initial slope $\mu_{0,r}$ at $\bar{\tau}_n(3,r) = 15$.

The conditions that cause a given test to yield a result are not guaranteed to be met over the interval of local neighborhood sizes considered. However, in the hundreds of test cases that were run, there was always at least one test that yielded a result.

The optimal local neighborhood size $\bar{\tau}_n$ is determined using Eq. (26). All values of m and r produce similar results for $\bar{\tau}_n(m,r)$. In addition, the resulting smoothed function is not very sensitive to the value of τ_n used, as long as it is sufficiently large.

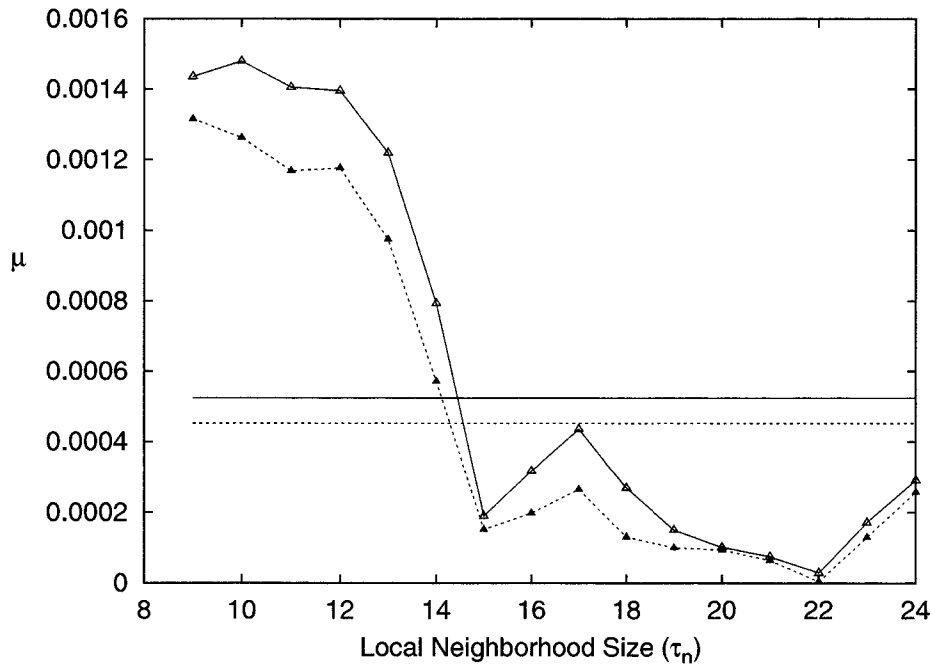

FIG. 15: Slope $\mu(\tau_n)$ of $\xi(\tau_n)$. The horizontal lines are $1/e$ of the initial slope $\mu_{0,r}$.

In summary, a method for transforming the function $W(\tau)$, defined on a set of index lags S, into a smoother function $\mathcal{W}(\tau)$, defined on a smaller set of index lags $\mathcal{S}$, has been described. This is done by smoothing the function W using Eq. (22) over the local neighborhood Λ defined by Eq. (15). The size of the local neighborhood to use is determined by applying three tests to the smoothing metric of Eq. (25) and choosing the smallest of the three results. The smoothed function $\mathcal{W}$ is obtained from the locally smoothed values of the maximum W from the original data set:

$$\mathcal{W}(\tau) = \bar{W}_{\max}(\tau, \bar{\tau}_n), \tag{29}$$

where $\bar{\tau}_n$ is given by Eq. (26), and the error bar for this value is the standard error:

$$\sigma(\tau) = \sigma(\tau, \bar{\tau}_n). \tag{30}$$

In this section, the process of data smoothing was described. This produces a transformation of $W(\tau)$ to $\mathcal{W}(\tau)$, together with an error bar $\sigma(\tau)$, given by

Eqs. (29) and (30), respectively. Figure 16a shows an example of an unsmoothed function W, and Fig. 16b shows the corresponding smoothed function $\mathcal{W}$. The original function $W(\tau)$ varies a lot from index lag to index lag. The smoothed function on the other hand does not vary much from index lag to index lag, but it still follows the basic trend of the original function very well.

Previously both the minimum and the maximum values of W were used when determining if a time series was produced by a deterministic system. The separation between the minimum W and the maximum W generated from surrogate data was used to set a confidence region. The process of smoothing the function W places error bars on the smoothed functions, and these can instead be used to determine a confidence region. Because of this, the minimum W for surrogate data set will no longer be considered.

In the next section the smoothed function $\mathcal{W}$ for original data will be compared to the smoothed function for the surrogate data to create a single measure of the separation between these two curves.

D. Characterizing the Difference Between Real and Surrogate Data

Consider a time series $X_R = \{x_i\}$ and a set T of index lags used to embed the attractor in m-dimensions:

$$T = \{\tau \mid \tau \in [\tau_0, \tau_L]\} \tag{31}$$

The previous sections contained a description of how to use this time series to generate two smoothed functions $\mathcal{W}(\tau)$. The first function is obtained directly from the time series. It will be referred to as the real function and denoted $\mathcal{W}_R$, with an estimated error of σ_R. This function is defined over a set of index lags $\mathcal{S}_R \subset T$. The second function is obtained from multiple surrogate time series. This function will be referred to as the surrogate function and denoted $\mathcal{W}_S$, with an estimated error of σ_S. This function is defined over a set of index lags $\mathcal{S}_S \subset T$.

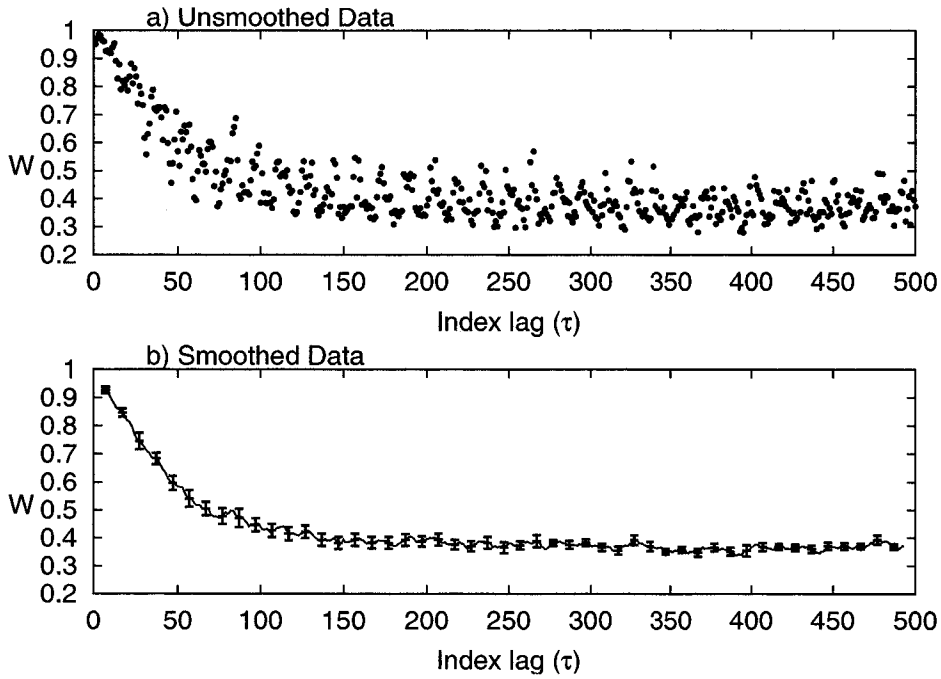

FIG. 16: Comparison of smoothed and unsmoothed functions. Plot (a) is the original unsmoothed function W. Plot (b) is the smoothed function $\mathcal{W}$, together with the error bars σ. The smoothed function follows the original curve well, but with a reduced variation from index lag to index lag.

Figure 17 shows an example of these two functions. Figure 17a shows the original unsmoothed functions W_R (closed circles) and W_S (closed squares). Figure 17b shows the smoothed functions $\mathcal{W}_R$ (upper curve) and $\mathcal{W}_S$ (lower curve). This section describes how to characterize the separation between these two functions.

Define the displacement D of $\mathcal{W}_R$ from $\mathcal{W}_S$:

$$D(\tau) = \mathcal{W}_R(\tau) - \mathcal{W}_S(\tau). \tag{32}$$

Since the estimated errors of these two functions are independent of one another, the associated error in D is

$$\sigma_D(\tau) = \sqrt{\sigma_R^2(\tau) + \sigma_S^2(\tau)}. \tag{33}$$

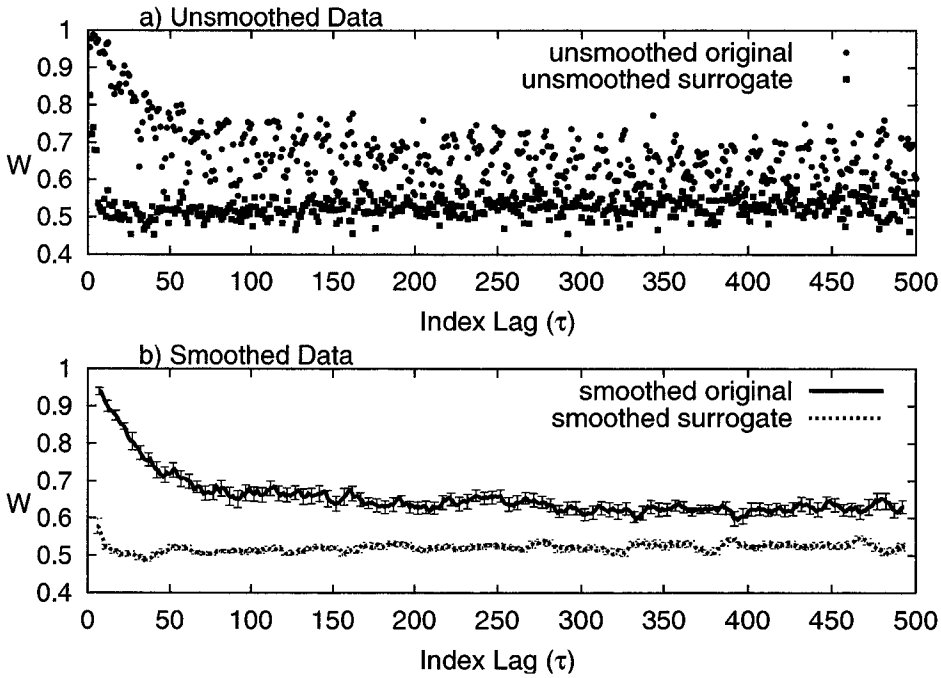

FIG. 17: Comparison of smoothed and unsmoothed functions for both real and surrogate data. Plot (a) shows the unsmoothed functions W_R (closed circles) and W_S (closed squares). Plot (b) shows the corresponding smoothed functions $\mathcal{W}_\mathcal{R}$ (upper curve) and $\mathcal{W}_\mathcal{S}$ (lower curve).

Both $D(\tau)$ and $\sigma(\tau)$ are defined over a set Φ of index lags:

$$\Phi = \mathcal{S}_R \cap \mathcal{S}_S. \tag{34}$$

Figure 18 is an example of $D(\tau)$ obtained from the data of Fig. 17. The error bars are the estimated error from Eq. (33), and the horizontal line is the average displacement over all index lags.

To characterize the separation between $\mathcal{W}_R$ and $\mathcal{W}_S$ define two quantities related to $D(\tau)$ — the percentage ρ_+ of index lags for which the displacement is positive within error, and the average displacement $\mathcal{D}$ over all index lags. The number of index lags for which the displacement is positive within error is

$$N_+ = \sum_{\tau' \in \Phi} \Theta\left(D(\tau') - \sigma_D(\tau')\right). \tag{35}$$

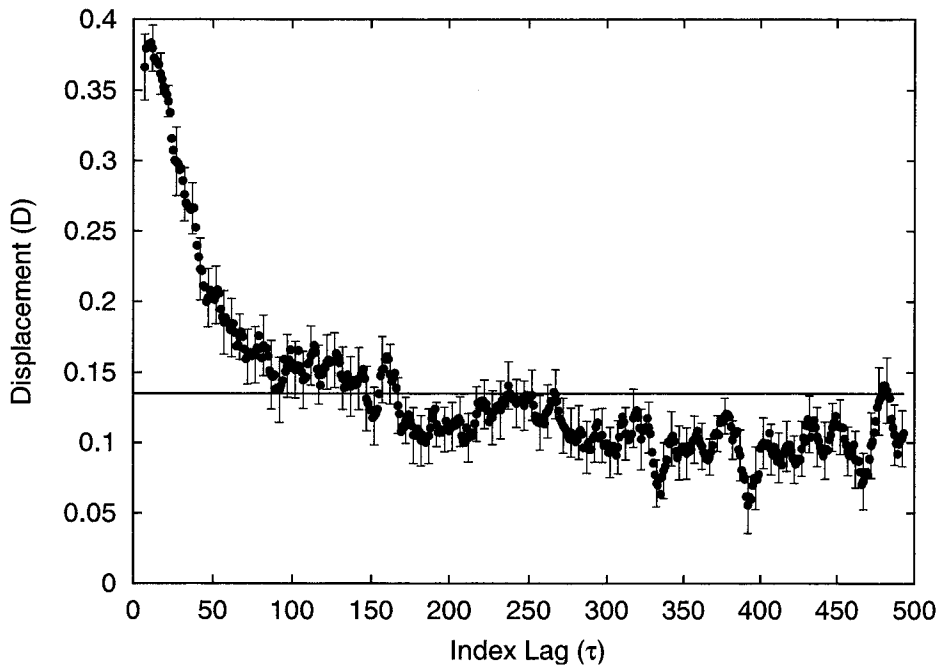

FIG. 18: Difference between real and surrogate data from Fig. 17. The horizontal line just below 0.15 is the average displacement between the two curves in Fig. 17.

The percentage of index lags for which the displacement is positive within error is

$$\rho_+ = 100\left(\frac{N_+}{\mid \Phi \mid}\right). \tag{36}$$

ρ_+ is a measure of the persistence of a positive displacement of $\mathcal{W}_R$ from $\mathcal{W}_S$, and it will be referred to as the percentage of positive displacements.

The average displacement of $\mathcal{W}_R$ from $\mathcal{W}_S$ will be refered to simply as the average displacement:

$$\mathcal{D} = \sum_{\tau' \in \Phi} D\left(\tau'\right) \Big/ \mid \Phi \mid . \tag{37}$$

$\mathcal{D}$ is a measure of the strength of the displacement .

In the next section the method of Salvino and Cawley, and the extensions discussed in this section, will be analyzed to determine their robustness with respect to noise. The quantities ρ_+ and $\mathcal{D}$ will be used to distinguish determinism from randomness.

IV. EFFECTS OF NOISE ON DETECTING DETERMINISM

A successful method for detecting deterministic chaos must be robust to the presence of noise. This section discusses the effects of noise on the ability of the Salvino and Cawley method, with the extensions of section III, to detect deterministic chaos. The following discussion will answer three questions. First, above what noise level can the extended method of Salvino and Cawley no longer detect deterministic chaos? Second, do different types of noise affect the ability of the extended Salvino and Cawley method to detect determinism? Finally, what are the benefits of the extensions to the Salvino and Cawley method discussed in previous sections?

Several methods for adding noise to a system are studied in this chapter, and these methods are described in section IV A. Each method of adding noise was studied at various noise levels and for several data set sizes. Section IV B describes how to interpret the amount of noise added in terms of characteristics of the system to which it is added. Section IV C describes the method used to analyze the results of the Salvino and Cawley method. Sections IV E through IV K presents results of analyzing various noise types.

A. Noise Types and Dynamical Systems

Two ways of adding noise to a system were considered. In one version, noise is combined with the underlying dynamics of the system; this will be refered to as *Dynamical Noise.* In the other version, noise is added to the resulting time series after the dynamics has been observed; this will be refered to as *Additive Noise.* Figure 19 illustrates these two methods.

Physically, Dynamical Noise may indicate an inherently noisy system, or a system with multiple components, some of which are noisy. Additive Noise has no effect on the underlying dynamics of the system. Physically, additive noise is rep-

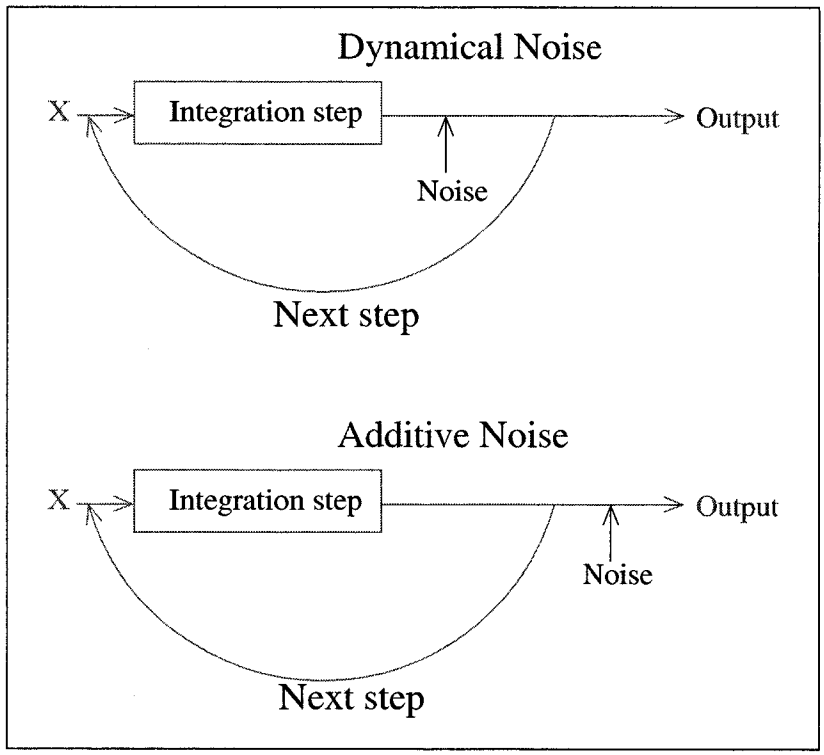

FIG. 19: Two methods for adding noise.

resentative of noise in the measurement process.

For Dynamical Noise, there are two cases which specify whether noise is added at every integration step, or whether it is added intermittently. Noise added at each integration step is referred to as *Continuous Dynamical Noise.* Noise added intermittently is referred to as *Intermittent Dynamical Noise.* Either case of Dynamical Noise effects the time evolution of the trajectory. Because of this, it is possible the noise will eject the trajectory from the basin of attraction, and the trajectory will wander off to infinity.

Two forms of noise were studied. The first form of noise is gaussian white noise generated from a Mersenne Twister algorithm [109]. Gaussian white noise is added to a system using any of the three methods described above. White noise combined with the system at every integration step is referred to as *Continuous Dynamical*

White Noise, white noise combined intermittently is referred to as *Intermittent Dynamical White Noise*, and white noise added to a time series is referred to as *Additive White Noise.*

The other form of noise considered is generated from a chaotic system. This type of noise is studied to determine if two competing chaotic systems will appear to be deterministic to the Salvino and Cawley method, or whether the method will be fooled into thinking the system is noisy.

Two chaotic differential equations were studied; the Lorenz system, given by Eqs. 2 and illustrated in Fig. 2, and the Rössler system [107], given by Eqs. 38 and illustrated in Fig. 20:

$$\begin{aligned} \dot{x} &= -y - z \ , \\ \dot{y} &= x + Ay \ , \\ \dot{z} &= B + xz - Cz \ . \end{aligned} \tag{38}$$

Both differential equations were integrated with a fourth-order Runge-Kutta method with an integration step of $\delta t = 0.001$. The parameters used with the Rössler system are $A = 0.2$, $B = 0.4$, and $C = 5.7$.

Both of these differential equations are used as both dynamical systems and as sources of noise. To generate chaotic noise, the differential equation is integrated, and a time series is produced by sampling the x component of the differential equation every S integration steps. This time series is then added as noise to another system using any of the three methods for adding noise described above.

For example, to use the Rössler system as a source of noise, a time series is generated by sampling the x component of the Rössler system every $S = 200$ integration steps. Combining this time series at every integration step is referred to as *Continuous Dynamical Rössler Noise.* Combining the time series intermittently is referred to as *Intermittent Dynamical Rössler Noise.* Adding the time series to another time series is referred to as *Additive Rössler Noise.* The Lorenz system is

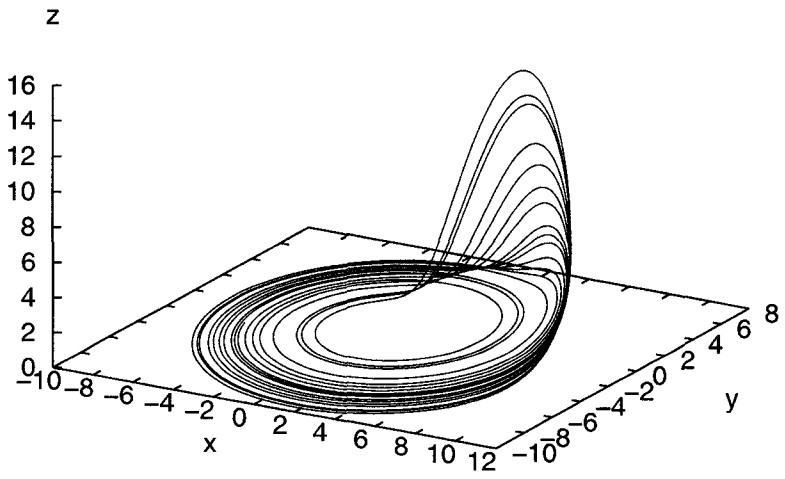

FIG. 20: Attractor for the Rössler system of Eqs. (38) for the parameter values A = 0.2, B = 0.4, and C = 5.7

used as a source of noise in the same way, only the x component of the trajectory is sampled every $S = 50$ integration steps.

Finally, the Hénon attractor, produced from Eqs. 39 and illustrated in Fig. 21, is used as an example of an iterated map. The Hénon attractor will be used as another dynamical system to test the effectiveness of the Salvino and Cawley method for an iterated map.

$$\begin{aligned} x_{n+1} &= y_n + 1 - a{x_n}^2 \,, \\ y_{n+1} &= b x_n \,. \end{aligned} \tag{39}$$

The combination of version, case (where applicable), and form of noise specifies a noise type. Table II lists the noise types studied and which systems were tested for each noise type.

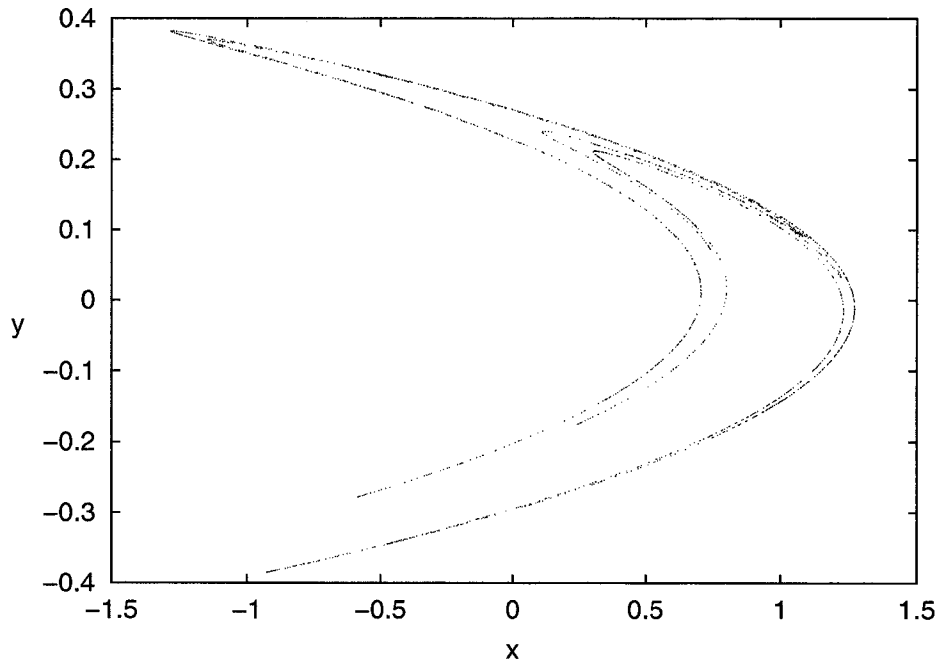

FIG. 21: The Hénon attractor of Eqs. (39) for the parameter values a = 1.4 and b = 0.3

TABLE II: Dynamical Systems and Noise Types

Noise Type	Dynamical Systems	Section
No Noise	Rössler, Lorenz, White Noise	IV E
Continuous Dynamical White Noise	Rössler, Lorenz	IV F
Intermittent Dynamical White Noise	Rössler, Lorenz	IV G
Additive White Noise	Rössler, Lorenz, Hénon	IV I
Continuous Dynamical Rössler Noise	Rössler, White Noise	IV K
Continuous Dynamical Lorenz Noise	Rössler, White Noise	IV K
Intermittent Dynamical Rössler Noise	Rössler, Lorenz	IV K
Intermittent Dynamical Lorenz Noise	Rössler, Lorenz	IV K
Additive Rössler Noise	Rössler, Lorenz	IV K
Additive Lorenz Noise	Rössler, Lorenz	IV K

B. Noise Level

The amount η of white noise added to a system at each time step is obtained from a random number generated from a Marsenne Twister algorithm, with

$$\eta \in [-N, N]\,, \tag{40}$$

where N is the characteristic noise level for a given trial.

Although chaotic systems are not periodic, they often do have a characteristic period, and this is the case for the Rössler and Lorenz systems. The standard way of relating the amount of noise added to a system to the characteristic size of the system is the ratio of the amount of noise added during one characteristic period to the characteristic size of the system.

Let S be the time interval at which the trajectory is sampled (number of integration steps between samples), let M be the number of samples taken in one characteristic period of the system, and let T be the time interval at which noise is added. Then S/T is the number of times noise is added between samples, and MS/T is the number of times noise is added in one characteristic period. Since noise added at one step is independent of noise added at another step, the total amount of noise added between samples is $N_{total} = N\sqrt{MS/T}$. The percentage of noise $\mathcal{P}$ added per sample period is related to the size of the system by $\mathcal{P} = 100\,(N_{total}/\sigma)$, where σ is the characteristic size of the attractor. In terms of the characteristic noise level, this percentage is:

$$\mathcal{P} = 100\left(\frac{N}{\sigma}\cdot\sqrt{\frac{MS}{T}}\right). \tag{41}$$

M, S and σ are characteristics of the system in question. N and T are characteristics of the noise being added.

A standard measure for the size of an attractor is its standard deviation. To obtain the standard deviation of the Rössler and Lorenz attractors, each system was integrated through $n = 10^5$ integration steps to obtain a set of trajectory

points $\left\{\vec{V}_i\right\}$. The center of the attractor $\vec{C}$ is found from the vector average of $\left\{\vec{V}_i\right\}$. The standard deviation is obtained from the distance of each point $\vec{V}_i$ from the center $\vec{C}$:

$$\sigma = \sqrt{\sum_{i=1}^{n} \mid \vec{V}_i - \vec{C} \mid^2 / n}. \tag{42}$$

The characteristic period of a system is determined by finding the average time between successive piercings of a poincaré section. To create a poincaré section, place a plane across the trajectory and mark the position where the trajectory crosses this plane in a given direction.

Table III summarizes the characteristics M, S and σ for the Lorenz and the Rössler systems.

TABLE III: Attractor Characteristics

System	Size (σ)	Sampling Period (S)	Samples/Period (M)
Rössler	7.1	200	30
Lorenz	14.7	50	64

For noise generated from a chaotic system, a time series $\{x_i\}$ is produced from this chaotic system. The amount of chaotic noise added on the i^{th} noise step is a fraction f of x_i. This fraction f will be referred to as the fractional noise level. For a chaotic noise source with a characteristic size σ the magnitude of the noise added on a given noise step is

$$N = f\sigma. \tag{43}$$

This noise level N is suitable to use in Eq. 41 to calculate the percent noise level $\mathcal{P}$ in terms of the size of the system to which it is being added.

C. Analyzing Noisy Data

The following sections will be looking at results for various noise types. This section discusses the simple approximation method used to analyze the results

presented in those sections.

Section III D discusses using the percentage ρ_+ of index lags for which the displacement D of the smoothed function $\mathcal{W}_R$ (for real data) from the smoothed function $\mathcal{W}_S$ (for surrogate data) is positive within error. ρ_+ is referred to as the percentatage of positive displacements. As the noise level increases, there is a drop in the percentage of positive displacements ρ_+. The location where this drop off occurs is an estimate of the maximum noise level for which we can detect determinism, and this will be referred to as the maximum noise cutoff. The present section discusses a method for making a simple approximation of this maximum noise cutoff. After making this simple approximation of the maximum noise cutoff, noise levels above and below the cutoff are examined more closely to confirm that this is in fact the noise level above which determinism can no longer be observed.

To analyze a given noise type and noise level N, a time series of length n and the associated surrogate data sets were generated. The method of Salvino and Cawley, with the extensions discussed in section III, is applied to these time series to generate the smoothed functions $\mathcal{W}_R(\tau)$ and $\mathcal{W}_S(\tau)$. Plots of $\mathcal{W}_R(\tau)$ and $\mathcal{W}_S(\tau)$, as well as the original unsmoothed functions $W_R(\tau)$ and $W_S(\tau)$, are shown in Figs. 22 and 23 for one trial that shows evidence of determinism, and for one trial that does not, repectively.

Next, the displacement $D(\tau)$ of $\mathcal{W}_R(\tau)$ from $\mathcal{W}_S(\tau)$ is computed from Eq. (32) as discussed in section III D. Plots of $D(\tau)$ for the data in Figs. 22 and 23 are shown in Figs. 24 and 25, respectively.

This procedure is repeated for a range of noise levels. For each noise level N, the percentage of positive displacements ρ_+ and the average displacement $\mathcal{D}$ were computed and plotted versus noise level N. Figures. 26 and 27 show typical plots of ρ_+ and $\mathcal{D}$ versus noise level N.

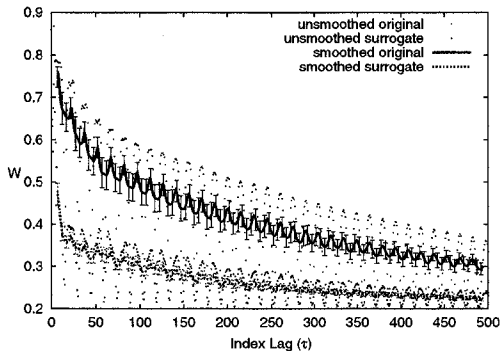

FIG. 22: Comparison of original and surrogate functions for a trial that shows evidence of determinism. Both smoothed and unsmoothed functions are shown. The dots are the original unsmoothed functions. The upper curve is the smoothed function $\mathcal{W}_R$, the lower curve is the smoothed surrogate function $\mathcal{W}_S$.

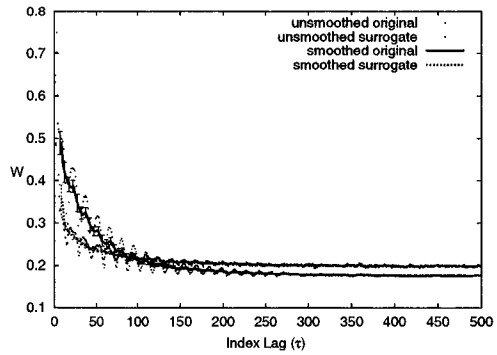

FIG. 23: Comparison of original and surrogate functions for a trial that shows no evidence of determinism and appears random. Both smoothed and unsmoothed functions are shown. The dots are the original unsmoothed functions. The smoothed function $\mathcal{W}_R$ is the upper curve for small index lags, and it switches to the lower curve around an index lag of 100. The other curve is the smoothed function $\mathcal{W}_S$.

1. *Analysis of Percentage of Positive Displacements Versus Noise Level*

In Fig. 26 for noise levels below $N = 0.875$, every displacement of $\mathcal{W}_R$ from $\mathcal{W}_S$ is positive, so, $\rho_+ = 100\%$ for these noise levels. As the noise level increases above $N = 0.875$, ρ_+ drops sharply. This indicates the maximum noise level where deterministic chaos can still be detected. The maximum noise level indicated by the percentage of positive displacements will be denoted by N_{pd}.

The description of the simple approximation of N_{pd} will be divided into two parts. The first part describes the basic ideas for finding N_{pd}. The second part discusses a situation that would lead to false results, and gives a modification to

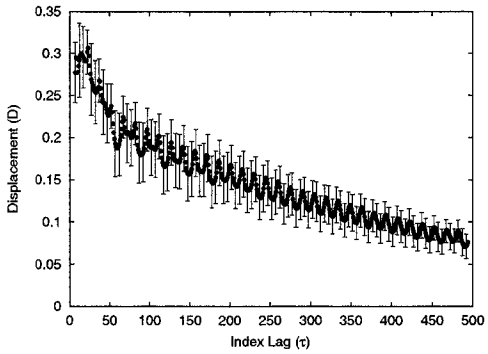

FIG. 24: Displacement $\mathcal{D}$ of $\mathcal{W}_R$ from $\mathcal{W}_S$ versus index lag τ for the data in Fig. 22.

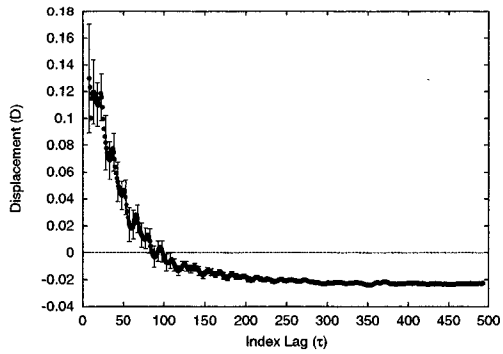

FIG. 25: Displacement $\mathcal{D}$ of $\mathcal{W}_R$ from $\mathcal{W}_S$ versus index lag τ for the data in Fig. 23. The horizontal line indicates a zero displacement.

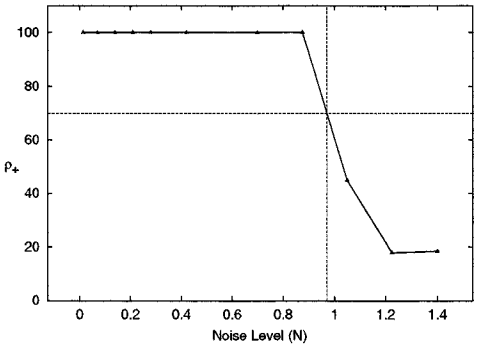

FIG. 26: Percent positive displacement versus noise level.

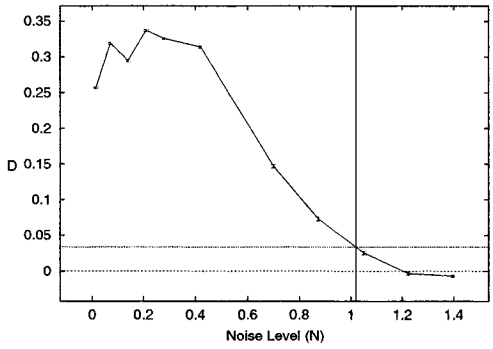

FIG. 27: Average displacement versus noise level.

the method that will help deal with such anomalous cases. It must be stressed that this method only provides a simple approximation for the maximum noise cutoff and it is intended as a way of automating the search for this cutoff. However, noise levels in the region near this approximation must then be examined more closely to find the true noise cutoff and verify that noise is in fact destroying the evidence of determinism. Usually, the noise cutoff produced by this method is a good approximation for the noise level above which determinism can no longer be

distinguished from randomness.

To estimate N_{pd}, locate the largest noise level $N_>$ with a percentage of positive displacements greater than some threshold ρ_{thresh}. In future sections, it will be shown that $\rho_{thresh} = 70\%$ is a good choice for this threshold. The maximum noise cutoff N_{pd} is then determined by interpolating between $N_>$ and the next noise level $N_<$.

In Fig. 26, $N_> = 0.875$ is the largest noise level above the threshold, and $N_< = 1.05$ is the first noise level larger than $N_>$. Linearly interpolating between these two points yields a maximum noise cutoff of ≈ 0.97.

Suppose the plot of the percentage of positive displacements looked like Fig. 28. The method described above would (incorrectly) identify the point at $N = 0.075$ as $N_>$, and the estimate of N_{pd} would be too high. Because of this, the method described above must be modified slightly to correctly deal with these anomalous cases. The necessary modification involves how $N_>$ is found.

Figures 28–30 illustrate the procedure for locating $N_>$. As before, we define a threshold ρ_{thresh}, and we divide the plot of ρ_+ versus N into those points that fall above the threshold and those that fall below the threshold, as shown in Fig. 28. Points above the threshold are divided into sets $S_1, S_2..., S_n$ of consecutive points. Figure 29 shows the example data set divided into 2 sets of consecutive points. $N_>$ is taken to be the largest noise level in the set that contains the most points. Given $N_>$, the determination of N_{pd} proceeds the same as before. In Fig. 30, $N_> = 0.056$ and $N_< = 0.0595$, and interpolation gives a maximum noise cutoff of ≈ 0.059.

2. Analysis of Average Displacement Versus Noise Level

Figure 27 shows the average displacement of $\mathcal{W}_R$ from $\mathcal{W}_S$ as a function of noise level. For low noise levels, the average displacement is large, but it decreases and approaches zero as the noise level increases. This provides another estimate of the maximum noise level where deterministic chaos can still be detected. The

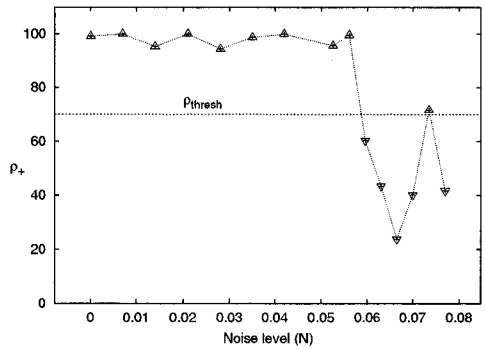

FIG. 28: Separation of a plot of $\rho_+(N)$ into those points above the threshold ρ_{thresh} (upright triangles) and those below (inverted triangles).

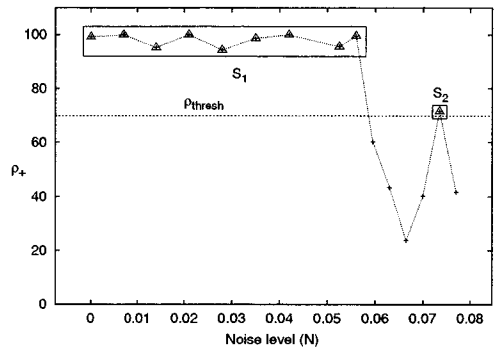

FIG. 29: Separation of data points above the threshold into sets of consecutive noise levels.

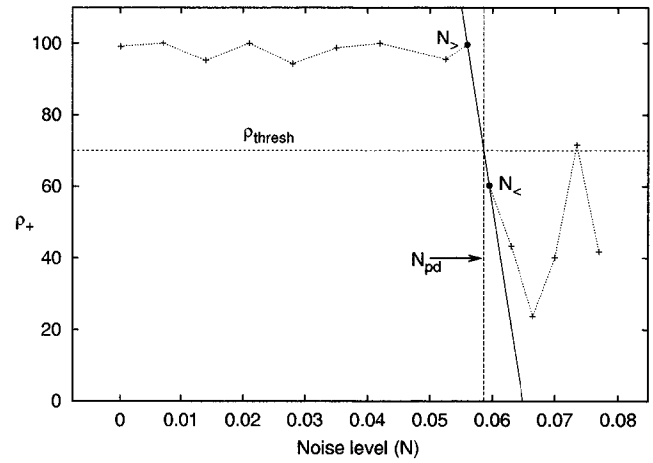

FIG. 30: Determination of N_{pd} from $N_>$ and $N_<$. The linear fit to these points determines the cutoff N_{pd}.

maximum noise level indicated by the average displacement will be denoted by N_{ave}. The estimate of N_{ave} is done in a similar manner as the simple approximation of N_{pd}. The natural value for a threshold would be zero; however, the curves often tend to settle just above zero, or oscillate about it. Thus, the threshold for computing N_{ave} is chosen to be 10% of the largest average displacement. Using

this threshold, one follows the same simple approximation method for determining N_{ave} as for determining N_{pd}.

3. Analysis of Cutoffs Versus Data Set Size

The simple approximation for the maximum noise cutoff gives us an idea of the amount of noise that can be added to a system for a given data set size before the dynamics are obscured enough that determinism can no longer be distinguished from randomness. It is found that, within error bars, the maximum noise cutoff does not depend on the size of the data set. Therefore, the average value of the cutoffs over all data set sizes gives a measure of the amount of noise that can be added to the system before determinism can no longer be distinguished from randomness.

The simple approximations of the cutoffs N_{pd} and N_{ave} are determined for data set sizes 2^n from $2^7 = 128$ through $2^{16} = 65,536$, and they are plotted versus n. Such a plot is shown in Fig. 31. The averages of N_{pd} and N_{ave} over all data set sizes, which are denoted by $\bar{N}_{pd}$ and $\bar{N}_{ave}$, respectively, are computed, along with the standard errors δ_{pd} and δ_{ave}. The ranges $\bar{N}_{pd} \pm \delta_{pd}$ and $\bar{N}_{ave} \pm \delta_{ave}$ give two measures of the maximum noise level above which determinism cannot be detected for a particular noise type. These ranges are indicated by the horizontal lines in Fig. 31. Since the procedure used provides only an approximation, it occasionally gives an estimate that is not in agreement with the estimates obtained for other data set sizes. When computing the average maximum noise cutoffs $\bar{N}_{pd}$ and $\bar{N}_{ave}$, such outliers are discarded.

D. Presentation of Results

Sections IV E through IV K presents results for several noise types. Section IV E gives results for both the Lorenz and Rössler systems in the absence of noise, and

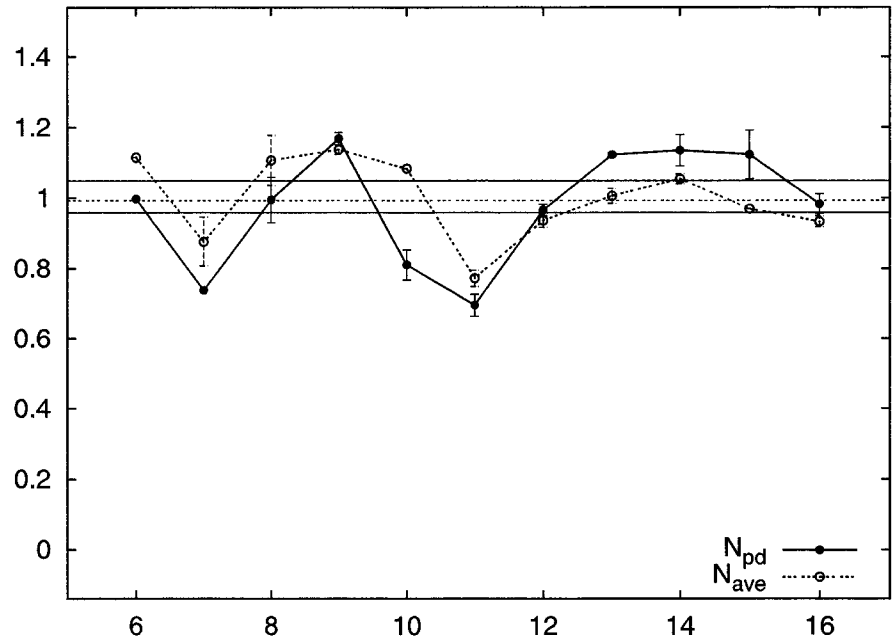

FIG. 31: Maximum noise cutoff versus data set size. The horizontal lines are $\bar{N}_{pd} \pm \delta_{pd}$ and $\bar{N}_{ave} \pm \delta_{ave}$.

for white noise all by itself.

Sections IV F through IV K present results for other noise types. In these sections, the procedure mapped out in section IV C is used to determine the maximum noise level for each noise type. This is done for smoothed data $\mathcal{W}$ as well as for the original unsmoothed data W. This will allow conclusions to be drawn regarding the effectiveness of the exentions to the method of Salvino and Cawley.

In future sections, a convention is adopted of using a script font for values obtained from the smooth functions $\mathcal{W}$ and a regular font for values obtained from the unsmoothed functions W. For example, the maximum noise cutoff in the percentage of positive displacements ρ_+ obtained from the smoothed functions $\mathcal{W}$ is denoted by $\mathcal{N}_{ave}$, and the maximum noise cutoff in ρ_+ obtained from the unsmoothed functions W is denoted by N_{ave}.

E. Results for No Noise

This section presents results for chaotic systems in the absence of noise, as well as results for white noise all by itself. This section is intended as a validation of our extended Salvino and Cawley method to detect determinism in systems without noise, and to validate the ability to detect randomness. This section also illustrates how to interpret plots of ρ_+, $\mathcal{D}$, $\mathcal{W}$, and W. Both the Rössler and the Lorenz systems, in the absence of noise, show evidence for determinism, while pure white noise shows evidence of randomness.

1. The Rössler System in the Absence of Noise

Figure 32a shows the shows percentage of positive displacements ρ_+ for the Rössler system in the absence of noise for data sets of $2^n = 2^7$ through $2^n = 2^{16}$ points. For all data set sizes, the displacement D of the smoothed function $\mathcal{W}_R$ (from real data) from $\mathcal{W}_S$ (from surrogate data) was positive within error, so $\rho_+ \approx 100\%$. Figure 32b is a plot of the average displacement $\mathcal{D}$ of $\mathcal{W}_R$ from $\mathcal{W}_S$ as a function of data set size. For all data set sizes, the average separation is greater than 0.1. A high percentage of positive displacements ρ_+, and average displacements $\mathcal{D}$ of this magnitude, is an indication of determinism.

An understanding of why this is an indication of determinism requires a closer look at several data set sizes. Figure 33a shows the unsmoothed function W (open points) and the smoothed function $\mathcal{W}$ (lines) for real and surrogate data for a data set size of $2^n = 2^7 = 128$. In this figure, the curves for real data are well separated from the curves for surrogate data for all index lags below about $\tau = 40$, which indicates determinism. Above an index lag of $\tau = 40$, the curves for real and surrogate data become intermingled, which would indicate randomness. However, the number of points l in an embedding is

$$l = n - \tau\,(m-1)\,, \tag{44}$$

where n is the length of the time series, and m is the dimension of the embedding. For this case, $n = 2^7 = 128$, and $m = 3$, so embedding with an index lag of $\tau = 40$ produces only 48 embedding points. This illustrates the fact that the separation decreases for larger index lags due to a lack of embedding points. This explains why ρ_+ is not 100% for this data set size.

The error bars on the $\mathcal{W}$ curves are all on the order of 0.05, so an average separation $\mathcal{D}$ greater than 0.1 indicates determinism.

Figure 34a shows the functions W and $\mathcal{W}$ for real and surrogate data for a data set size of $2^9 = 512$. The real unsmoothed function W is intermingled with the surrogate data for several index lags. However, the two smoothed functions $\mathcal{W}$ exhibit a clear separation for all index lags, indicating the process of smoothing allows a better comparison between real and surrogate data. There is a definite separation of these curves, so the underlying system is deterministic.

Figure 35a shows the functions W and $\mathcal{W}$ for real and surrogate data, for a data set size of $2^9 = 8192$. In this case, the separation between real and surrogate data is very well pronounced, indicating determinism.

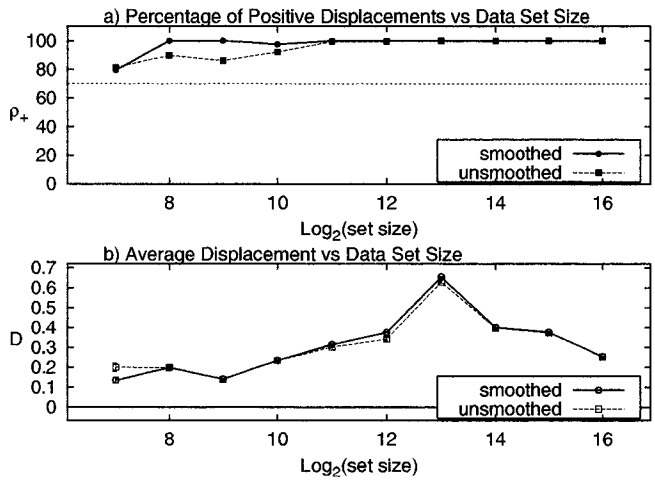

FIG. 32: ρ_+ and $\mathcal{D}$ for the Rössler system in the absence of noise. Plot (a) is the percentage of positive displacements ρ_+, and plot (b) is the average displacement $\mathcal{D}$.

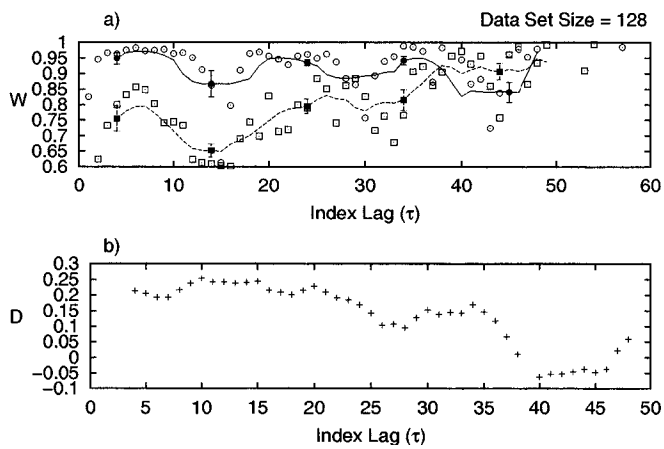

FIG. 33: W and D versus index lag τ for the Rössler system in the absence of noise for a data set of $2^7 = 128$ points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

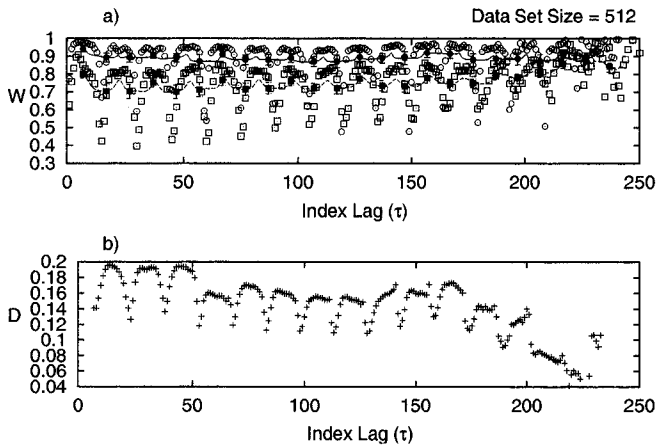

FIG. 34: W and D versus index lag τ for the Rössler system in the absence of noise for a data set of $2^9 = 512$ points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

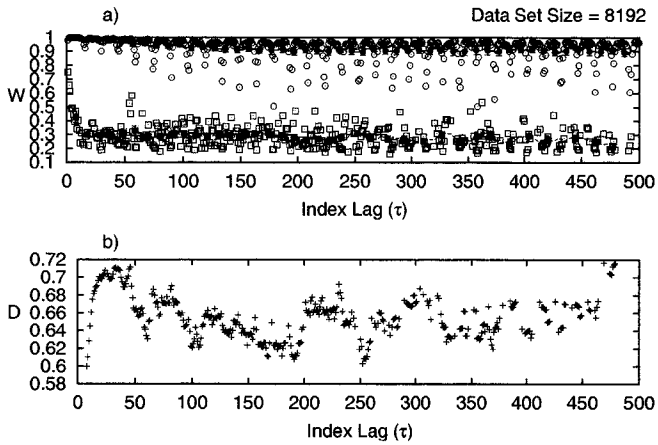

FIG. 35: W and D versus index lag τ for the Rössler system in the absence of noise for a data set of $2^{13} = 8192$ points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

2. The Lorenz System in the Absence of Noise

Figure 36a shows the shows percentage of positive displacements ρ_+ for the Lorenz system in the absence of noise for data sets of $2^n = 2^7$ through $2^n = 2^{16}$ points. As with the Rössler system, the percentage of positive displacements ρ_+ for all data set sizes is above the 70% threshold, indicating determinism. Figure 37a shows the W and $\mathcal{W}$ curves for a data set size of $2^{12} = 4096$. There is a very distinct separation, indicating determinism.

In Fig. 36a, ρ_+ makes a dip at 2^9 to around 80%. Figure 38 shows the W and $\mathcal{W}$ curves for this noise level. In this case, the two curves are separated for most index lags, but are intermingled for others. The smoothed curves $\mathcal{W}$ show a distinct separation for most noise levels. This is a marginal case, but it is still valid to conclude that the system is deterministic.

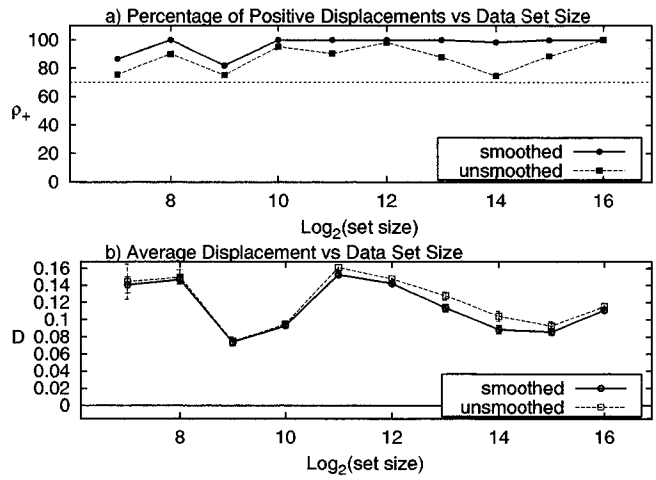

FIG. 36: ρ_+ and $\mathcal{D}$ for the Lorenz system in the absence of noise. Plot (a) is the percentage of positive displacements ρ_+, and plot (b) is the average displacement $\mathcal{D}$.

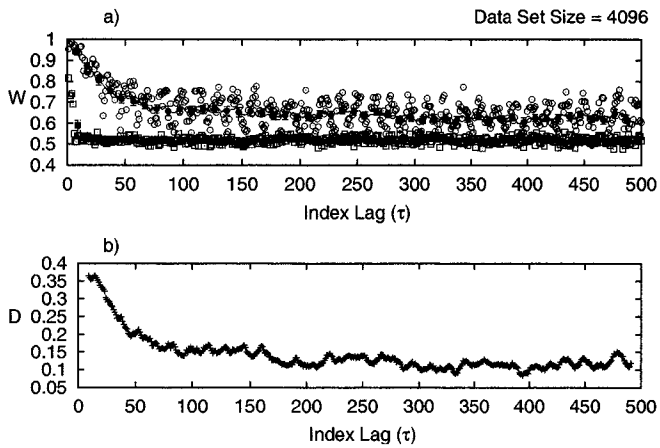

FIG. 37: W and D versus index lag τ for the Lorenz system in the absence of noise for a data set of $2^{12} = 4096$ points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

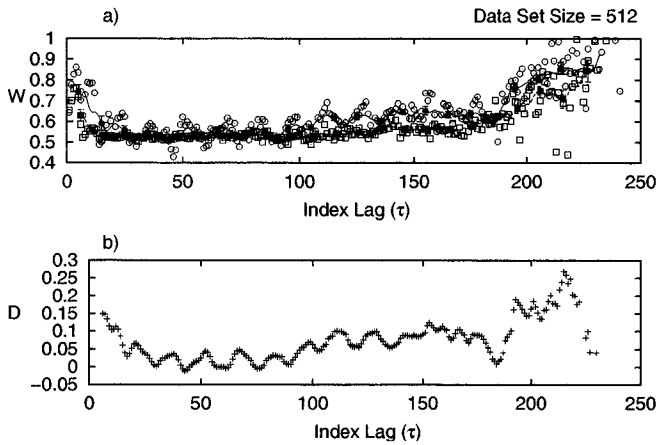

FIG. 38: W and D versus index lag τ for the Lorenz system in the absence of noise for a data set of $2^9 = 512$ points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

3. White Noise

In this subsection, data sets consisting of random numbers are examined. Since there is no a priori notion of the embedding dimension, determinism is checked for by embedding in 1,2, and 3 dimensions. If the noise source is random, there should be no evidence for determinism regardless of the embedding dimension.

Figure 39a shows the percentage of positive displacements ρ_+, for a data set of random numbers embedded in one dimension. The curve is identically zero for all data set sizes. Figure 39b shows that the average separation $\mathcal{D}$ is negative for all data set sizes. These results indicate randomness. The W curves are shown in Fig. 40a for a data set size of 2^8. Embedding in one dimension does not involve an index lag, so, W is the same for all index lags. In this case, W for real data lies below W for surrogate data. The same thing is seen in Fig. 41a for a data set size of 2^9.

Figure 42a shows ρ_+ for random data embedded in two dimensions. For most data set sizes, the percentage of positive displacements ρ_+ is zero. For data set sizes of 2^8 and 2^9, $\rho_+ \sim 30\%$. For all data set sizes, the average separation $\mathcal{D}$ is negative. The low percentage of positive displacements, and the negative average displacements, indicate randomness. Figure 43a shows the W curves for a data set of 2^9 points, which is one of the data set sizes where $\rho_+ \neq 0\%$. The curve for real data and the curve for surrogate data lie on top of one another, which is an indication of randomness. Since the two curves lie on top of one another, one would expect that, for some index lags, the curve for real data is the greater of the two, and, for other index lags the curve for surrogate data is the greater of the two. This explains why the percentage of positive displacements is nonzero.

Figure 44a shows ρ_+ for random data embedded in three dimensions. Figure 44b shows the average separation $\mathcal{D}$ for a three-dimensional embedding of random data. As for one-dimensional and two-dimensional embeddings, there is no evidence for determinism. For completeness, the W curves for random data embedded in three

dimensions are shown in Fig. 45a for a data set of 2^9 points.

In conclusion, for random data, the embedding dimension has no effect on the ability of the extended method of Salvino and Cawley to recognize the system as being random.

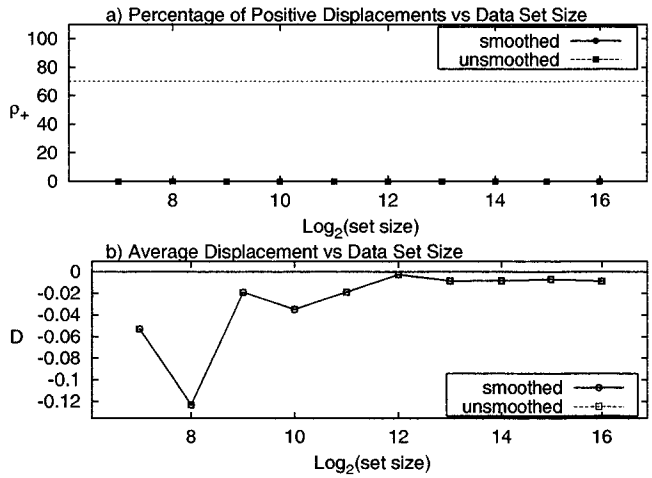

FIG. 39: ρ_+ and $\mathcal{D}$, for a time series of random numbers embedded in 1 dimensions. Plot (a) shows ρ_+, and plot (b) shows $\mathcal{D}$.

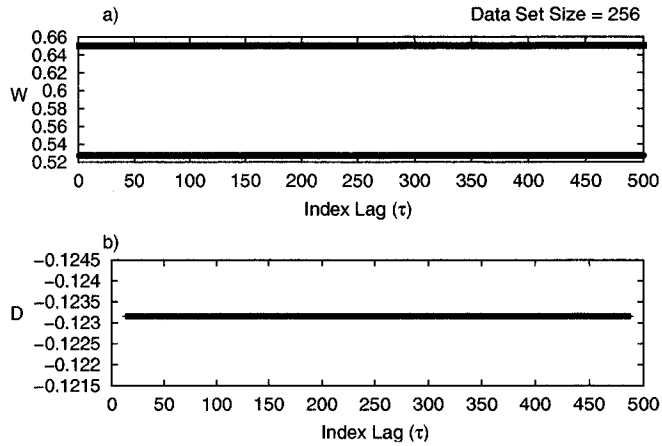

FIG. 40: W and D versus τ for a data set consisting of $2^8 = 256$ random numbers using an embedding dimension of one. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

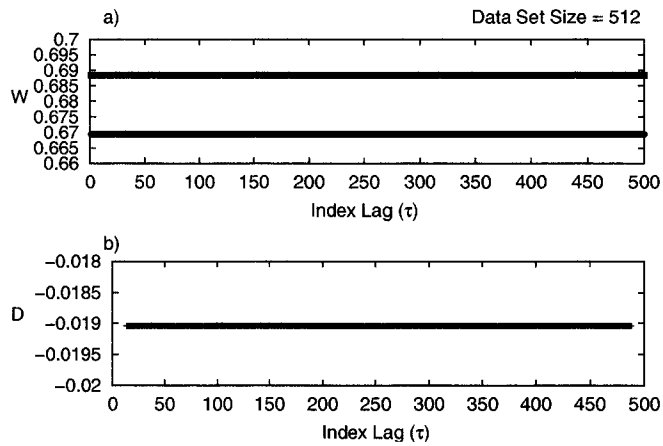

FIG. 41: W and D versus τ for a data set consisting of $2^9 = 512$ random numbers using an embedding dimension of one. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

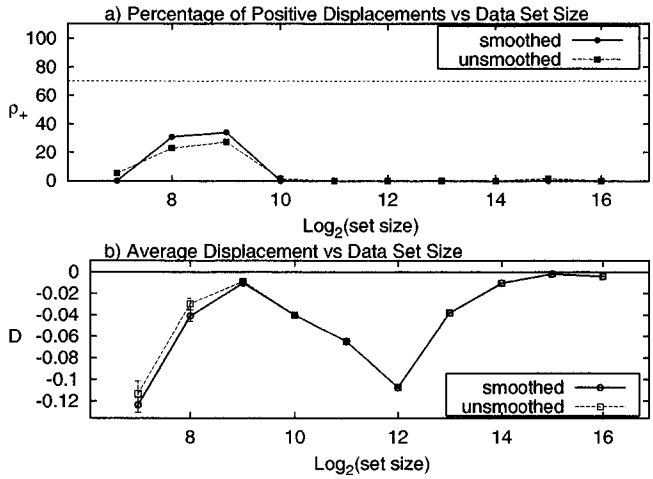

FIG. 42: ρ_+ and $\mathcal{D}$, for a time series of random numbers embedded in 2 dimensions. Plot (a) shows ρ_+, and plot (b) shows $\mathcal{D}$.

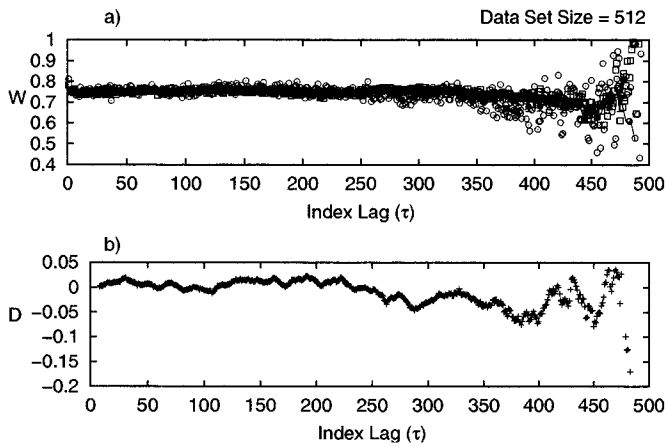

FIG. 43: W and D versus τ for a data set consisting of $2^9 = 512$ random numbers using an embedding dimension of two. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

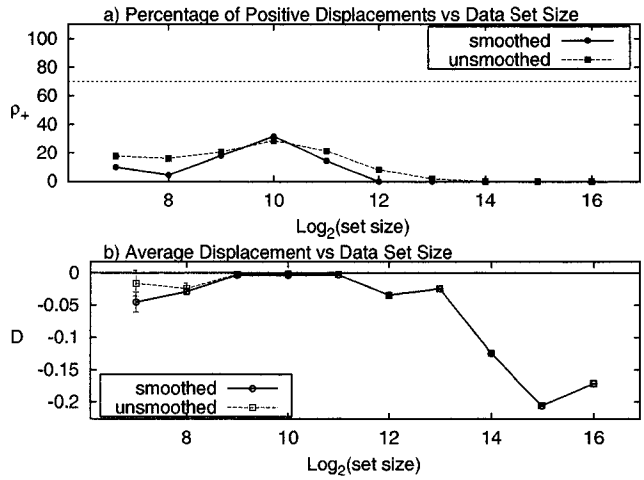

FIG. 44: ρ_+ and $\mathcal{D}$, for a time series of random numbers embedded in 3 dimensions. Plot (a) shows ρ_+, and plot (b) shows $\mathcal{D}$.

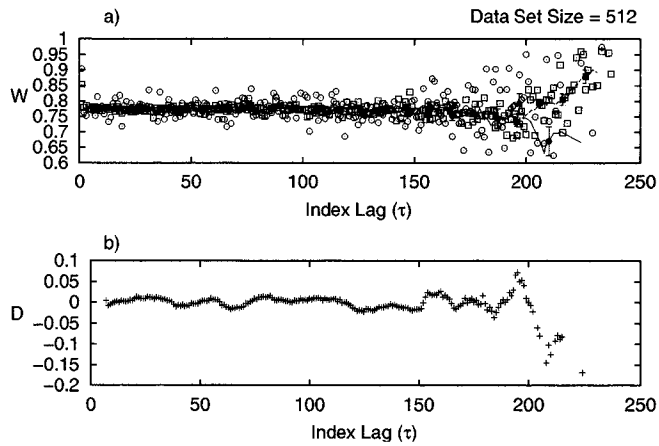

FIG. 45: W and D versus τ for a data set consisting of $2^9 = 512$ random numbers using an embedding dimension of three. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

F. Results for Continuous Dynamical White Noise

Continuous Dynamical White Noise was studied for the Lorenz, Rössler, and Hénon systems. However, applying this type of noise to the Hénon system ejected the trajectory from the basin of attraction, and it ran off to infinity.

Results for Continuous Dynamical White Noise applied to the Rössler and Lorenz systems are presented in subsections IV F 1 and IV F 2, respectively. In the plots of percentage of positive displacements ρ_+ versus noise level N, the solid curve is generated by comparing the smoothed functions $\mathcal{W}$, while the dashed curve is generated by comparing the unsmoothed functions W (for example, see Fig. 46). The same convention is used for plots of average displacement $\mathcal{D}$ versus noise level. Solid vertical lines in either plot are the maximum noise cutoffs for smoothed data, and the dashed vertical lines are maximum noise cutoffs for unsmoothed data.

In this section, the choice of the threshold $\rho_{thresh} = 70\%$ will be justified. The plots of ρ_+ and $\mathcal{D}$ will be analyzed in detail to demonstrate how the maximum noise cutoffs are determined. Plots of $\mathcal{W}(\tau)$ and $W(\tau)$ will often be refered to in order to demonstrate how to distinguish between determinism and randomness.

1. *Continuous Dynamical White Noise Applied to the Rössler System*

Results for a data set size of $2^n = 2^7 = 128$ are shown in Figs. 46–49. These results are a bit atypical. To understand how they are different, the other data set sizes will be discussed first, and then this data set size will be discussed afterwards.

Results for a data set size of $2^8 = 256$ are shown in Figs. 50–55. Figure 50a shows the precentage of positive displacements ρ_+ versus noise level N. The first thing to note is that the curve for the smoothed functions $\mathcal{W}$ (solid line) is slightly above the curve for the unsmoothed functions W (dashed line). This demonstrates that smoothing does in fact create a more persistent displacement of $\mathcal{W}_R$, from $\mathcal{W}_S$. Figure 50b shows the average displacement $\mathcal{D}$ for both smoothed and unsmoothed

functions. Note that these two curves follow each other very closely. This indicates that, on average, the process of smoothing does not increase or decrease the separation between the real and surrogate data. This is the desired result; if the process of smoothing changed the average separation between real and surrogate data, it would be possible to obtain false results.

In Fig. 50a, the curve for ρ_+ makes a sudden drop below the threshold of 70% just before a noise level of $N = 0.06$. To verify that this is in fact the point where noise begins to obscure the dynamics enough to destroy evidence of determinism, noise levels in the region from $N = 0.042$ to $N = 0.077$ will be looked at more closely.

Figure 51a shows both the smoothed functions $\mathcal{W}$ (solid line for real data and dashed line for surrogate data) and the unsmoothed functions W (open circles for real data and open boxes for surrogate data) for a noise level of $N = 0.042$. Figure 51b shows the separation D of the two curves in Fig. 51a as a function of index lag for this noise level. There is a separation between the W curves for real and surrogate data of about $D = 0.2$ for index lags up to about 80. The size of the error bars on the $\mathcal{W}$ curves are on the order of 0.05, so a separation of 0.2 is a significant separation, indicating that the system is deterministic. Above an index lag of $\tau = 80$ the separation between the two curves begins to decrease. This can be attributed to data starvation of the embedding. Using Eq. 44, with $n = 2^8 = 256$, $m = 3$, and an index lag of $\tau = 80$ produces 96 embedding points. The separation D decreases for larger index lags due to lack of embedding points (i.e., the smallness of the effective data set).

In Fig. 50a, the largest noise level where the ρ_+ is above 70% is $N = 0.056$. The functions W and $\mathcal{W}$ are shown in Fig. 52a for a noise level of $N = 0.056$. Comparing this figure to Fig. 51a shows that the curve for unsmoothed real data fluctuates more than for a noise level of $N = 0.042$. This is reflected in the oscillations seen for the smoothed function $\mathcal{W}_R$ in Fig 52a. The curve for $\mathcal{W}_R$ lies above the curve for $\mathcal{W}_S$ for most index lags. The separation between the two

smoothed curves has decreased to about 0.1. This can be seen by referring to the figure for average separation $\mathcal{D}$ [Fig 50b], and by an inspection of the curve of $D(\tau)$ shown in Fig. 52b. A separation of 0.1 is larger than the size of the error bars of $\mathcal{W}_R$, so the system is deterministic. However, the fact that the separation is beginning to approach the size of the error bars, is an indication that the noise is begining to have a noticeable effect on the system.

In Fig. 50a, a noise level of $N = 0.0595$ is the smallest noise level for which ρ_+ is less than 70%, which is our criterion for randomness. Figure 53a shows the W curves for this noise level. Notice that the curve for real data now falls below the curve for surrogate data, in agreement with the conclusion of randomness.

Finally, consider the noise levels $N = 0.07$ and $N = 0.077$ in Fig. 50. The percentage of positive displacements ρ_+ for both of these noise levels fall below the 70% threshold, although it is greater at these two noise levels than it is for nearby noise levels. The average separation $\mathcal{D}$ shown in Fig. 50b for these noise levels is less than the size 0.05 of the error bars for $\mathcal{W}_R$, so the separation is zero within error, in agreement with the conclusion of randomness. An important thing to note is that, even for randomness, a significant percentage of the index lags can yield a positive separation. Figures 54a and 55a show the W curves for $N = 0.07$ and $N = 0.007$, respectively. In both cases, the two curves for real and surrogate data lie on top of each other, indicating randomness.

Results for a data set size of 2^9 are shown in Figs. 56 and 57. As for a data set size of 2^8, the curve of ρ_+ for smoothed data in Fig. 56a has a definite separation from the curve of ρ_+ for unsmoothed data. The maximum noise cutoffs $\mathcal{N}_{pd}$, $\mathcal{N}_{ave}$, N_{pd} and N_{ave} are all near $N = 0.06$, as was the case for the data set size of 2^8. The curve of ρ_+ in Fig. 56a momentarily rises slightly above 70% at a noise level of $N = 0.07$. The plots of W are shown in Fig. 57a for this noise level. In this figure, the unsmoothed W curves are intermingled in many places; however, the smoothed $\mathcal{W}$ curves show a small but persistent separation. At this noise level there is some evidence for determinism, although it is not strong. This is an indication that the

70% threshold is a good value to use to determine the maximum noise cutoff $\mathcal{N}_{pd}$. Similar behavior is seen for data set sizes of 2^{10}, which are shown in Figs. 58–60.

Results for a data set size of 2^{11} are shown in Fig. 61–63. We again see a distinct separation between the curves of ρ_+ for smoothed data and unsmoothed data in Fig 61a. The maximum noise cutoff is around $\mathcal{N}_{pd} \approx 0.06$, in agreement with the results obtained for the data set sizes discussed previously. Two points beyond this maximum noise cutoff rise above the 70% threshold at $N = 0.07$ and at $N = 0.077$. Plots of W for these noise levels are shown in Fig 62 and Fig. 63, respectively. Both of these plots show weak evidence of determinism.

Results for a data set size of 2^{12} are shown in Figs. 64 and 65. For this data set size, the effects of smoothing the data are not as pronounced. This can be seen by noting the separation between the curves of ρ_+ for smoothed and unsmoothed data in Fig. 64a is not as great as was observed for smaller data set sizes. This can also be seen by looking at the W curves in Fig 65a and noting that they do not vary as much from index lag to index lag as for the smaller data set sizes. For smaller data set sizes, there was an increase in the scatter of the W plots as the index lag increased, which was attributed to data starvation of the embedding. For this data set size, this effect is greatly reduced, and the scatter in the W curves is similar for all index lags.

Figures 66–69 shows the results for a data set size of 2^{13}, and they are similar to the results for the data set size of 2^{12}. One thing to notice is that the curves of ρ_+ for smoothed and unsmoothed data begin to separate from one another near the maximum noise cutoff.

Results for a data set size of 2^{14} are shown in Fig. 70–77. In Fig. 70a the percentage of positive displacements for smoothed data is consistently near 100%, but the curve for the unsmoothed data makes small drops below the 70% threshold at $N = 0.0525$ and at $N = 0.07$. The average displacements $\mathcal{D}$ shown in Fig. 70b drop below the threshold of 0.05 at around $N_{ave} = 0.0665$. Referring to Figs. 71–77 shows that this maximum noise cutoff is indeed the noise level for which determin-

ism can no longer be detected. This shows that both average displacement $\mathcal{D}$ and percentage of positive displacements ρ_+ are useful when determining the maximum noise level.

Similar results are seen for data set sizes of 2^{15} and 2^{16}, as seen in Figs. 78 and 79. For these higher data set sizes, smoothing the W functions has less of an effect than for smaller data set sizes. The maximum noise cutoffs are again between $N = 0.06$ and $N = 0.07$.

Returning to Fig. 46, for a data set size of 2^7, note that ρ_+ for this data set size oscillates quite a bit as a function of noise level. This is more pronounced for the unsmoothed data than for the smoothed data. These oscillations are not typical of plots of ρ_+ versus N. If ρ_+ for unsmoothed data is used, these oscillations lead to a significant underestimation of the maximum noise cutoff N_{pd}. However, these oscillations are not as pronounced for smoothed data, which eliminates the problem.

The curves for ρ_+ and the curves for $\mathcal{D}$ in Fig. 46 all make a sudden drop at a noise level of around $N = 0.06$, and then rise again for larger noise levels. Figures 47–49 show that the noise begins to obscure the dynamics enough to prevent distinguishing between determinism and randomness at a noise level of $N = 0.06$, in agreement with the simple approximation for the smoothed functions. In general, distinguishing between determinism and randomness for data sets this small is not reliable. However, in this case, smoothing the data allowed us to determine a maximum noise cutoff that was in agreement with the value found for larger data set sizes.

Trials for a data set size of $2^6 = 64$ data points were also considered. These trials proved to be very unreliable, and, in fact, the calculation of W could not always be carried out. Because of this, results for $2^6 = 64$ data points are not included in this dissertation.

Finally, Fig. 80 shows the results of the simple approximation at maximum cutoffs for all noise levels. For all data set sizes, the maximum noise cutoffs for

both smoothed and unsmoothed data obtained from either ρ_+ or $\mathcal{D}$ are between 0.06 and 0.07. The lone exception is for ρ_+ for unsmoothed data for a data set size of $2^7 = 128$. The maximum noise cutoff in this case was greatly underestimated by the simple approximation, but the actual cutoff was in agreement with the other cutoffs.

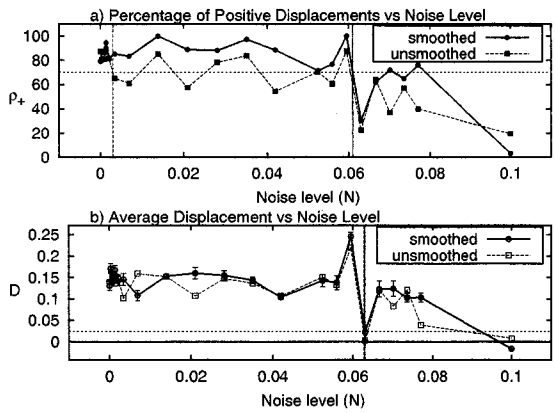

FIG. 46: Continuous Dynamical White Noise applied to the Rössler system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

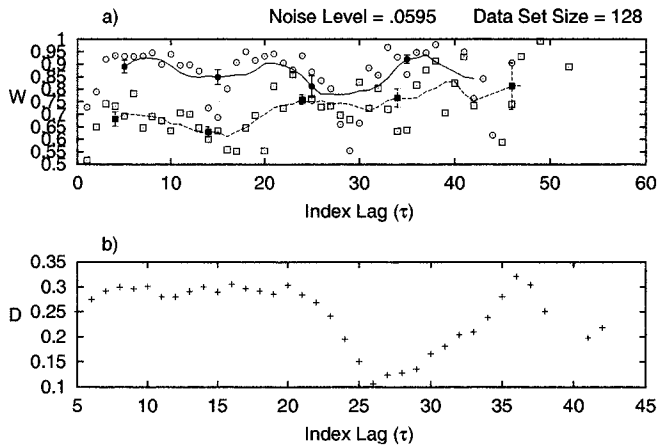

FIG. 47: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^7 data points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

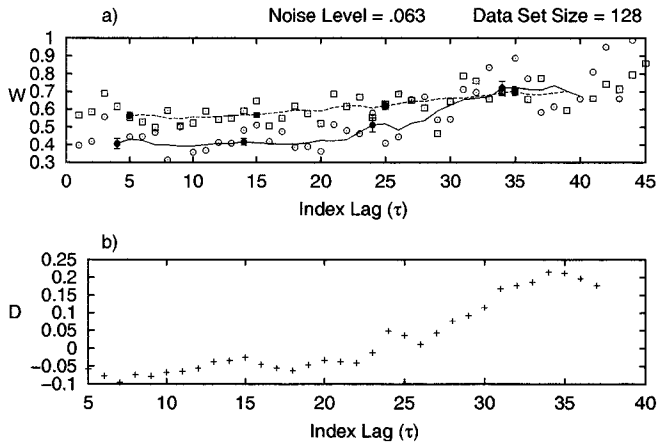

FIG. 48: W and D versus $\tau$for Continuous Dynamical White Noise applied to the Rössler system for 2^7 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

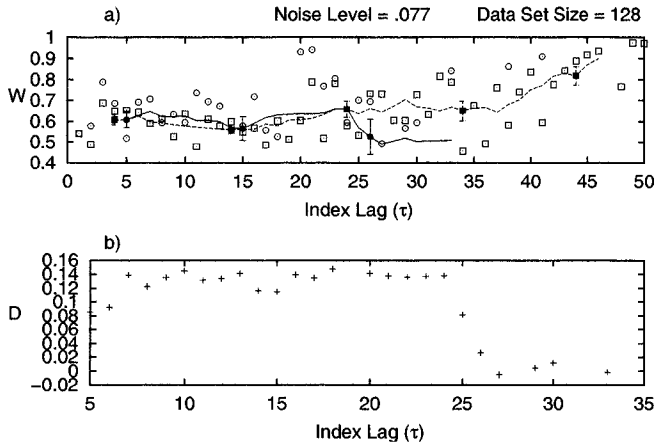

FIG. 49: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^7 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

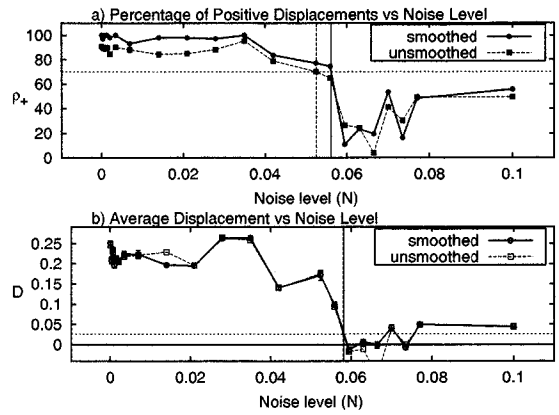

FIG. 50: Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

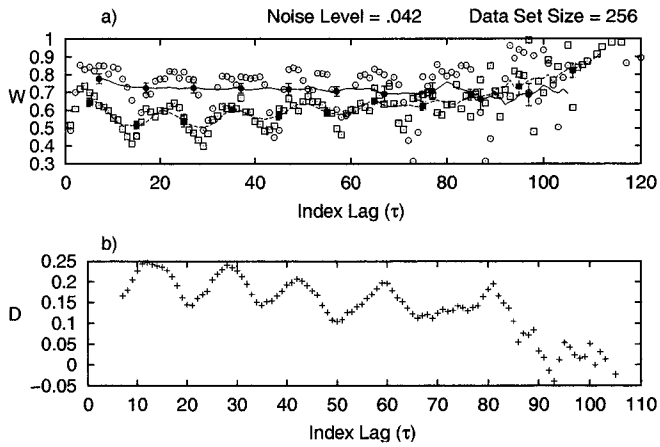

FIG. 51: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

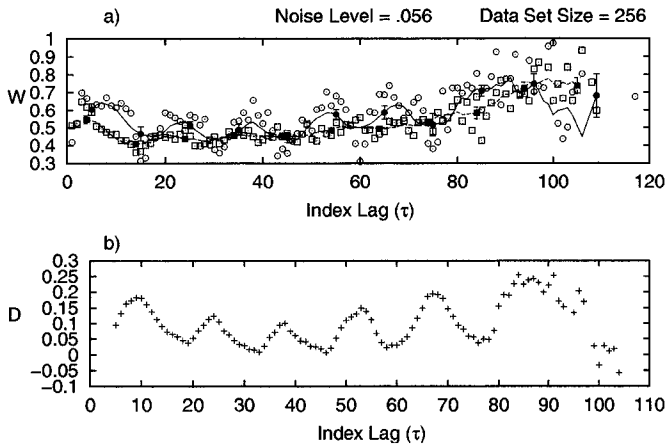

FIG. 52: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

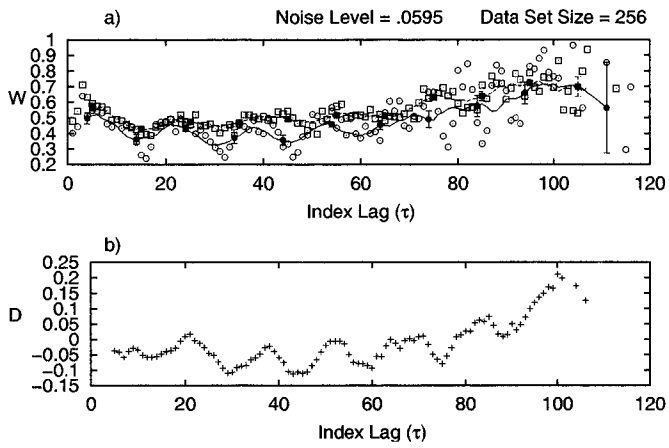

FIG. 53: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

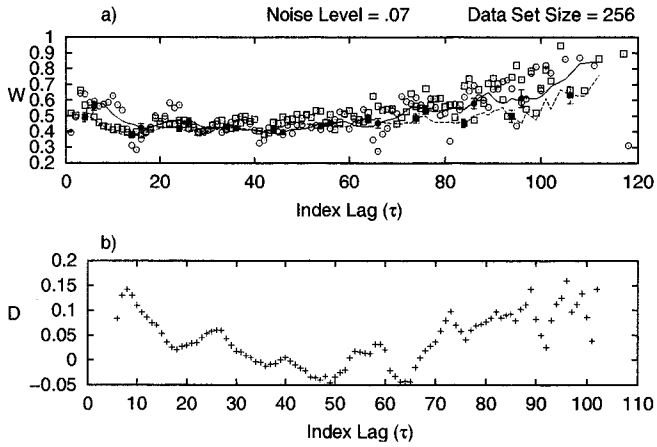

FIG. 54: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

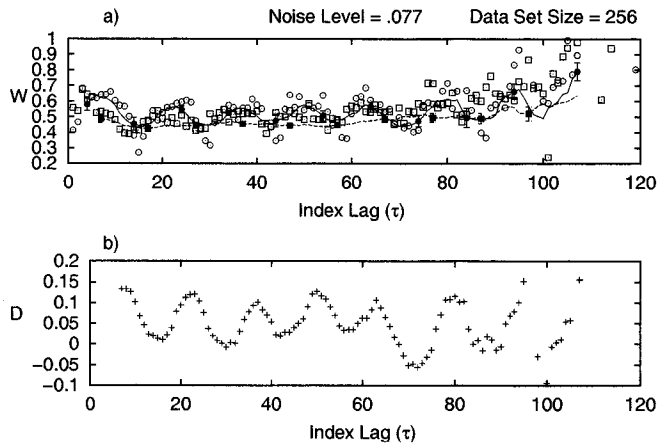

FIG. 55: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^8 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

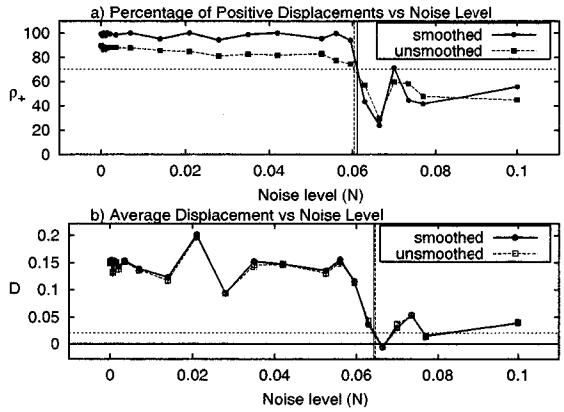

FIG. 56: Continuous Dynamical White Noise applied to the Rössler system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

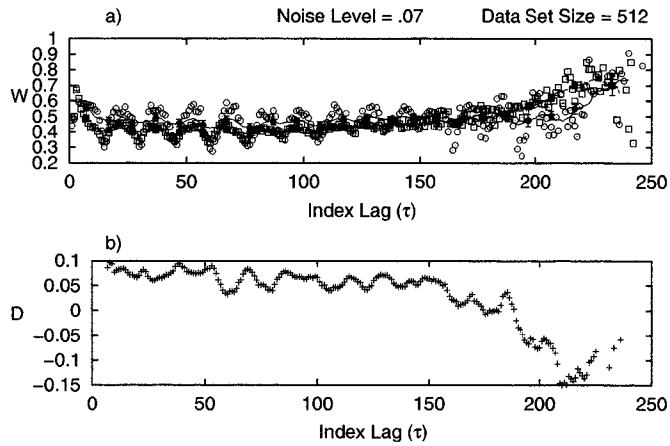

FIG. 57: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^9 data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

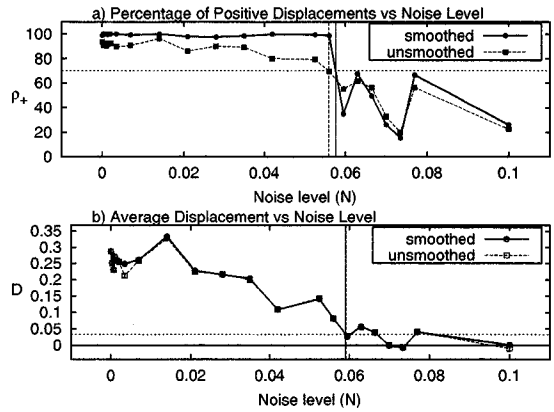

FIG. 58: Continuous Dynamical White Noise applied to the Rössler system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

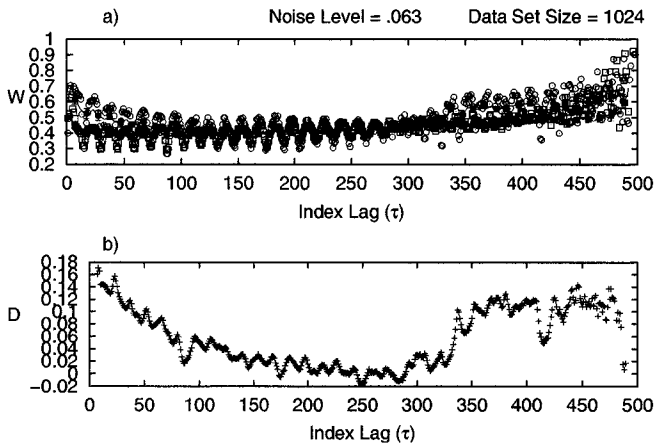

FIG. 59: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{10} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

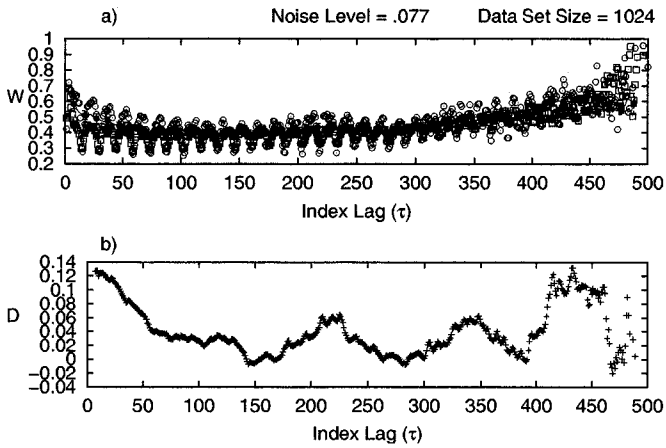

FIG. 60: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{10} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

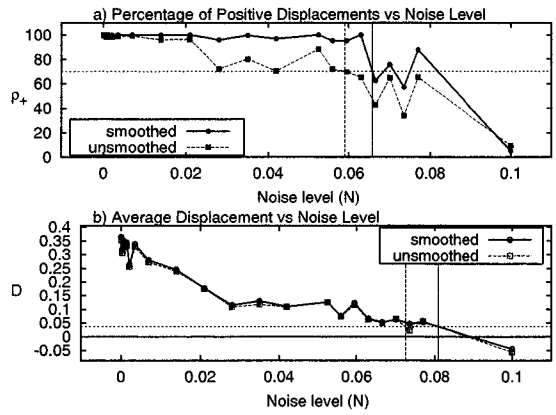

FIG. 61: Continuous Dynamical White Noise applied to the Rössler system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

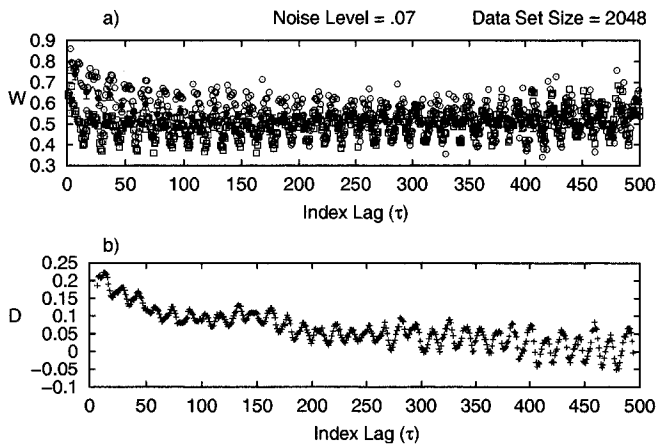

FIG. 62: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{11} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

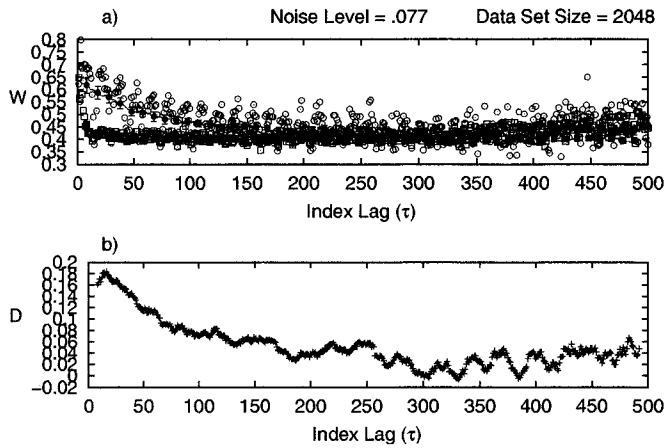

FIG. 63: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{11} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

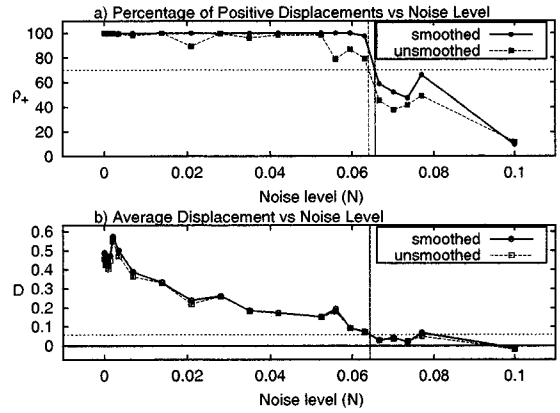

FIG. 64: Continuous Dynamical White Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

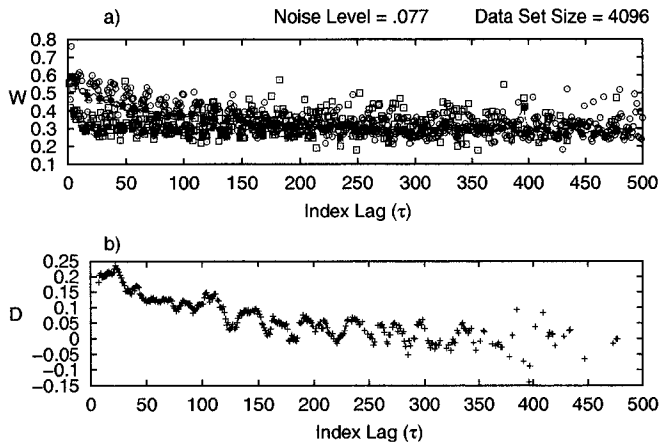

FIG. 65: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{12} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

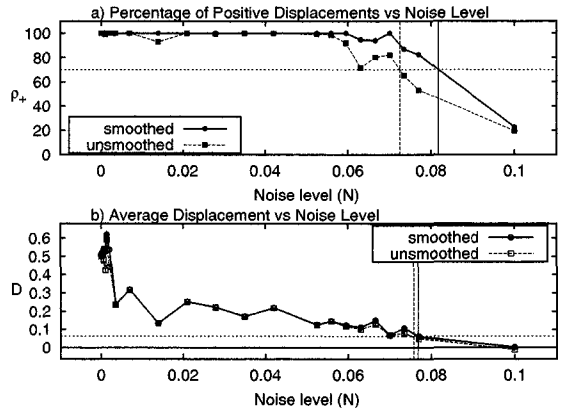

FIG. 66: Continuous Dynamical White Noise applied to the Rössler system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

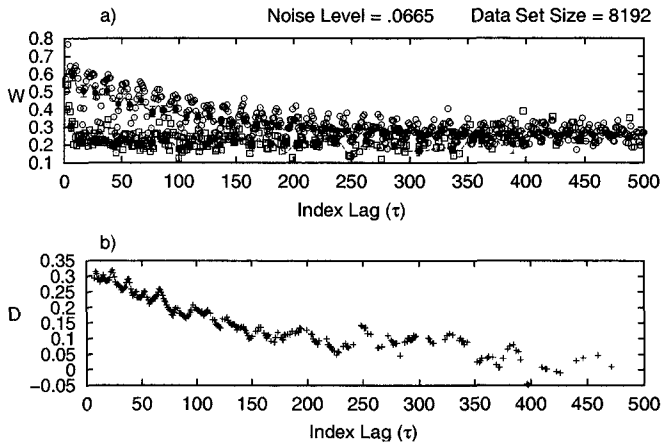

FIG. 67: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{13} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

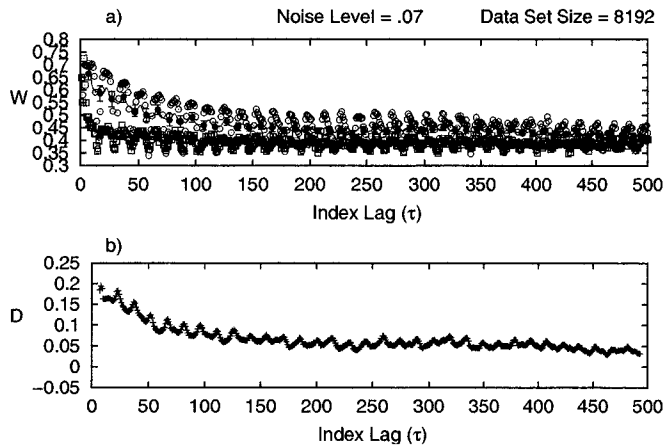

FIG. 68: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{13} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

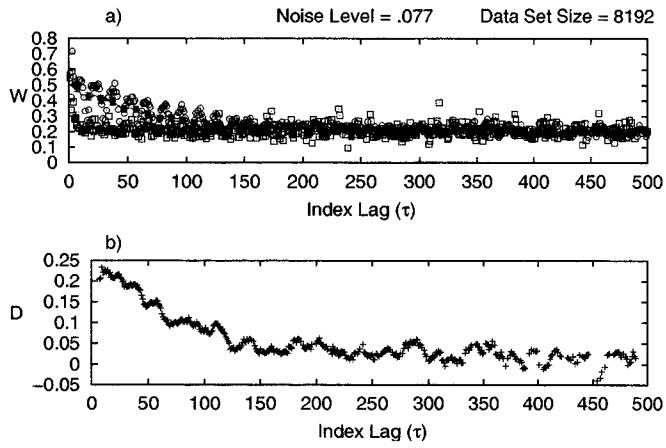

FIG. 69: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{13} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

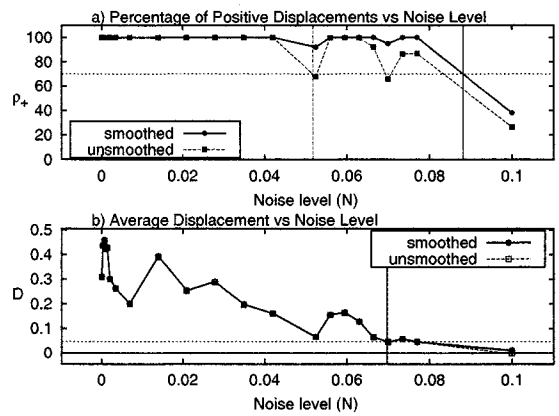

FIG. 70: Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

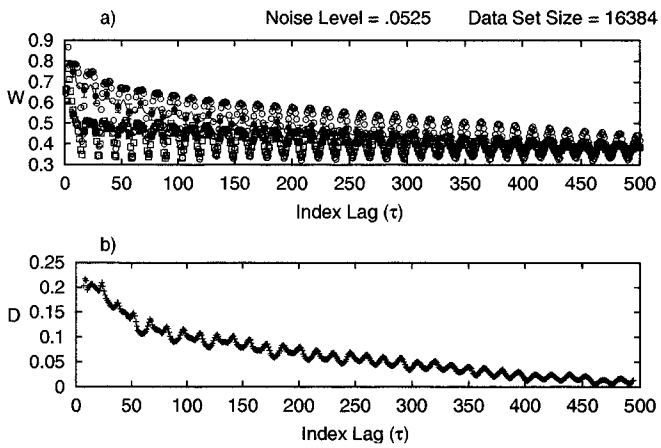

FIG. 71: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

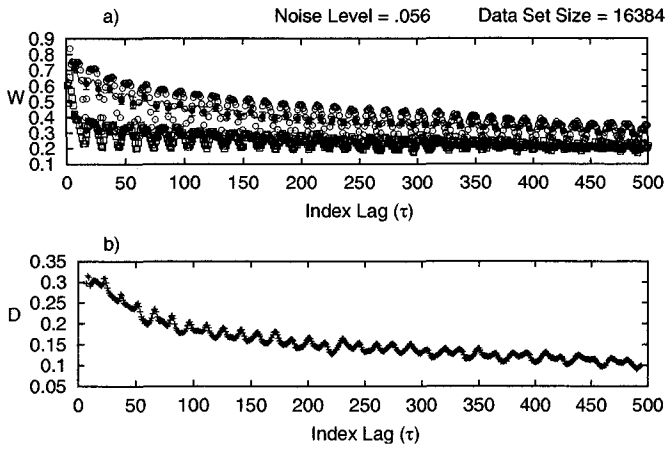

FIG. 72: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

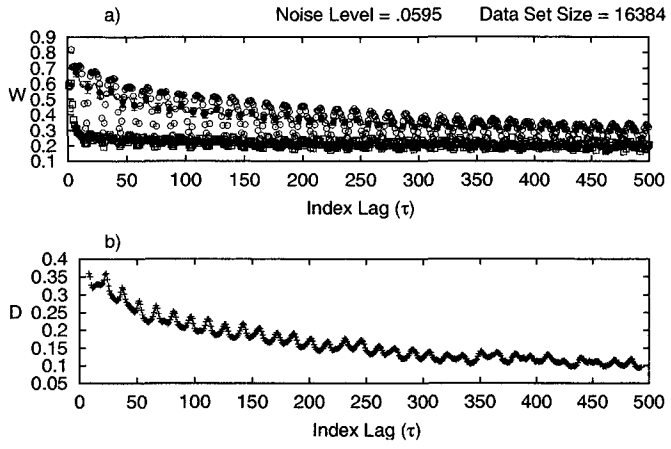

FIG. 73: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

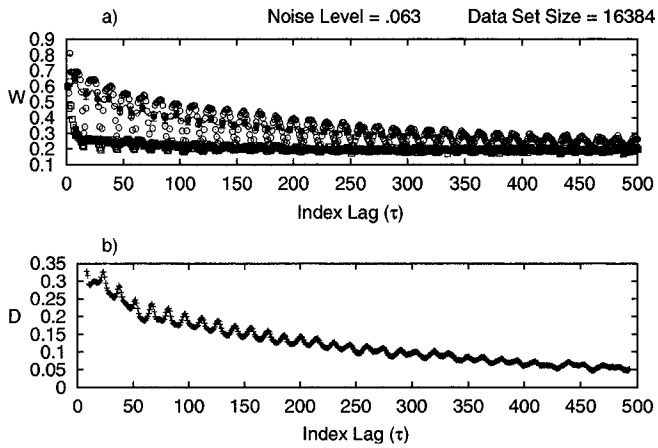

FIG. 74: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

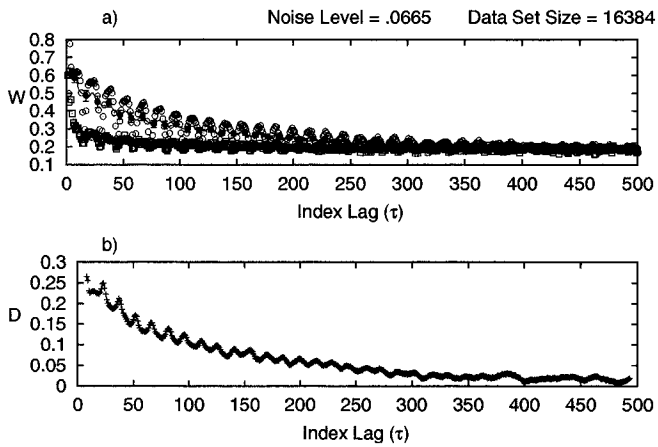

FIG. 75: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

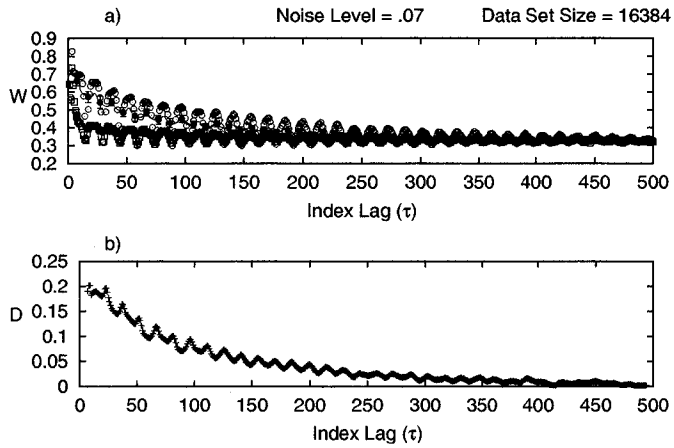

FIG. 76: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

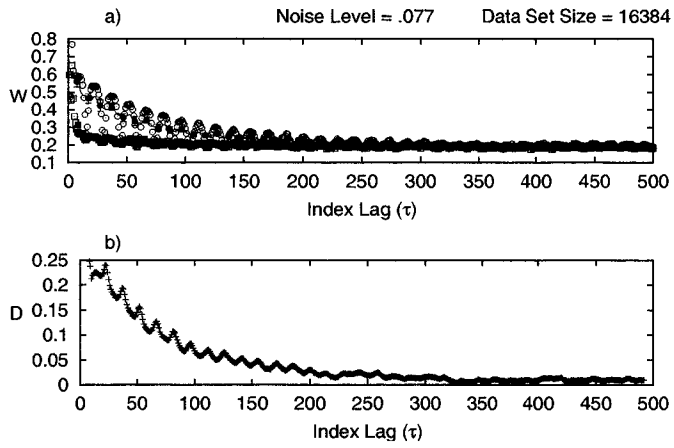

FIG. 77: W and D versus τ for Continuous Dynamical White Noise applied to the Rössler system for 2^{14} data points . Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

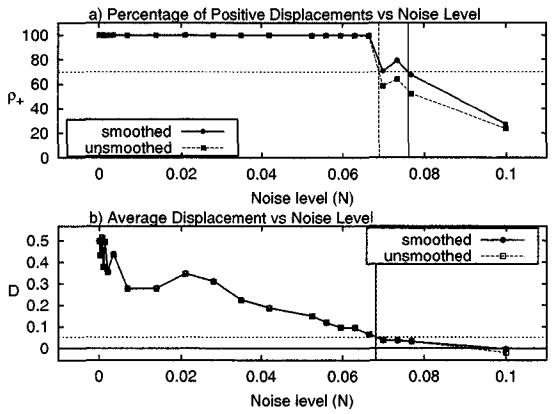

FIG. 78: Continuous Dynamical White Noise applied to the Rössler system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

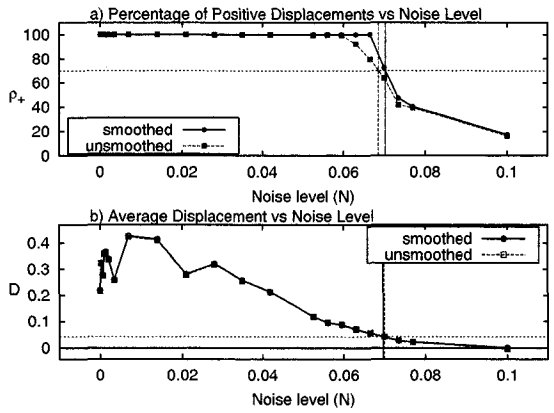

FIG. 79: Continuous Dynamical White Noise applied to the Rössler system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

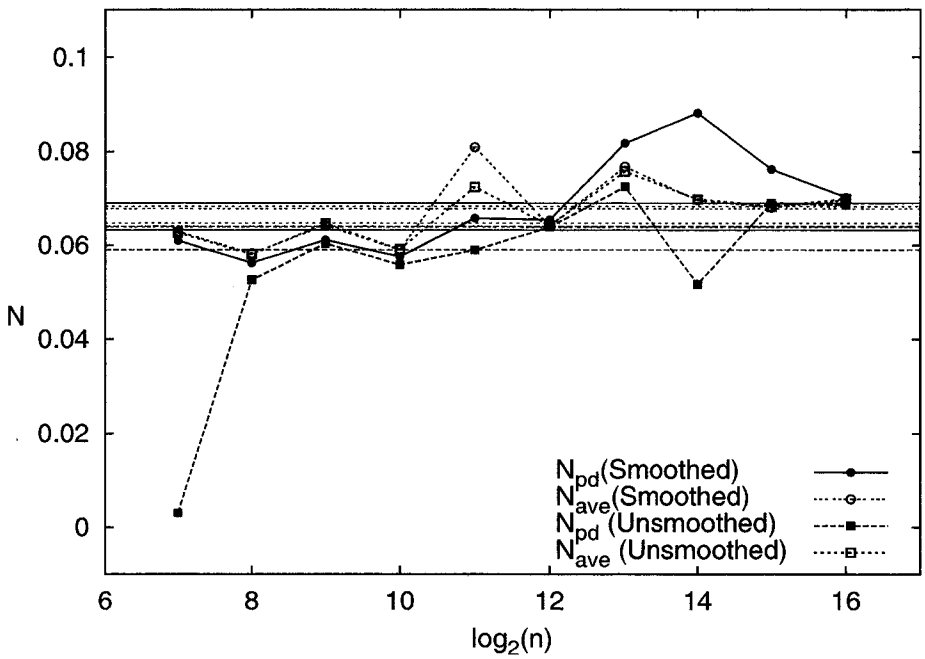

FIG. 80: Maximum noise cutoffs for the Rössler system for Continuous Dynamical White Noise.

2. Continuous Dynamical White Noise Applied to the Lorenz System

This subsection discusses Continuous Dynamical White Noise applied to the Lorenz system. The results shown here are very similar to those seen in the previous subsection for the Rössler system. Because of these similarities, this subsection will focus on confirming that the estimates of the maximum noise cutoffs occur at the correct locations.

As for Continuous Dynamical White Noise applied to the Rössler system, results for data sets of $2^n = 2^7 = 128$ points are unreliable. Figure 81a shows curves of ρ_+ versus N, and Fig. 81b shows curves of $\mathcal{D}$ versus N for Continuous Dynamical White Noise applied to the Lorenz system. In both figures, the curves fluctuate greatly, so there is no definite location where one can conclude the dynamics are obscured by noise to an extent that one can no longer detect determinism.

For a data set size of 2^8, the results are shown in Fig. 82. The cutoff $\mathcal{N}_{pd}$ occurs between $N = 0.2$ and $N = 0.25$. Figs 83a and 84a show the W curves for these noise levels. For $N = 0.2$, there is evidence for determinism, but, for $N = 0.25$, the curves for the real and surrogate data lie on top of one another, indicating randomness.

The situation for a data set size of 2^9 is shown in Figs. 85–87, and it is similar to that for a data set size of 2^8, except that the cutoff $\mathcal{N}_{pd}$ occurs between $N = 0.3$ and $N = 0.35$. This maximum noise cutoff was confirmed by inspecting the W curves.

The results for data set sizes 2^{10}, 2^{11}, and 2^{12} are shown in Figs. 88—90, and they are very similar. The plots of ρ_+ for smoothed and unsmoothed data are distinctly separated. In all cases, the cutoff $\mathcal{N}_{pd}$ for the smoothed data was around $N = 0.35$. This was confirmed by inspecting the plots of W versus index lag. For data set sizes of 2^{11} and 2^{12}, the simple approximation underestimates the true noise cutoff when unsmoothed data is used. However, smoothing the functions W again solves this problem. Also, we again see that a threshold of 70% is a good

choice for distinguishing between determinism and randomness.

The results for data set sizes of 2^{13}, 2^{14}, and 2^{15} are shown in Figs. 91–93. The maximum noise cutoffs are now around $N = 0.15$, which is less than the cutoffs for the smaller data set sizes. This was confirmed by examining the W versus index lag curves for noise levels in this region. For these data set sizes, the simulations were repeated with different initial conditions for the trajectories, and with a different seed for the random number generator. These trials gave the same results as presented here. The reason for the drop in the maximum noise cutoff is unclear.

For a data set size of 2^{16}, the maximum noise cutoff is about 0.3, which is in agreement with the cutoffs seen for data set sizes 2^9 through 2^{12}. As for Continuous Dynamical White Noise applied to the Rössler system, there is very little difference between ρ_+ for smoothed and unsmoothed data.

Finally, Fig. 95 shows the maximum noise cutoffs for all data set sizes. This plot is more scattered than the equivalent one for the Rössler system in Fig. 80. Also note that, for the Lorenz system, the maximum noise cutoffs estimated from ρ_+ for unsmoothed data vary wildly, suggesting that this is not a reliable measure of the maximum noise level for which determinism can be detected. This is more evidence that smoothing the W curves assists in distinguishing determinism from randomness.

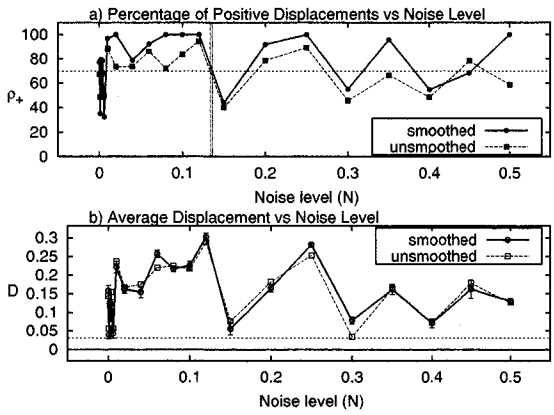

FIG. 81: Continuous Dynamical White Noise applied to the Lorenz system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

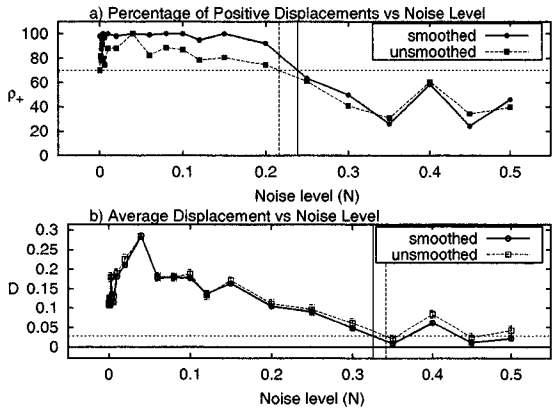

FIG. 82: Continuous Dynamical White Noise applied to the Lorenz system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

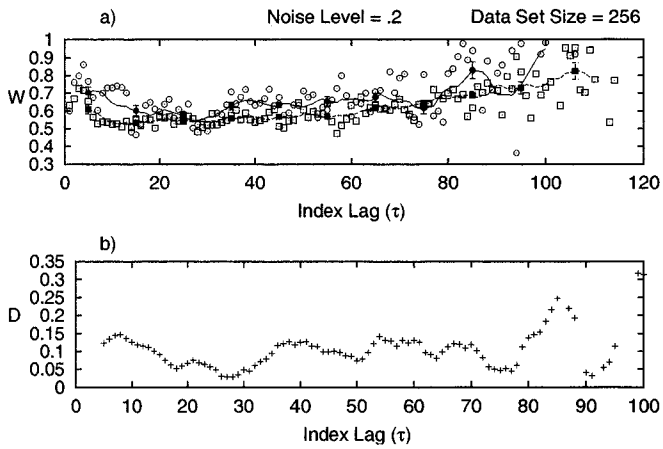

FIG. 83: W and D versus τ for Continuous Dynamical White Noise applied to the Lorenz system for 2^8 data points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

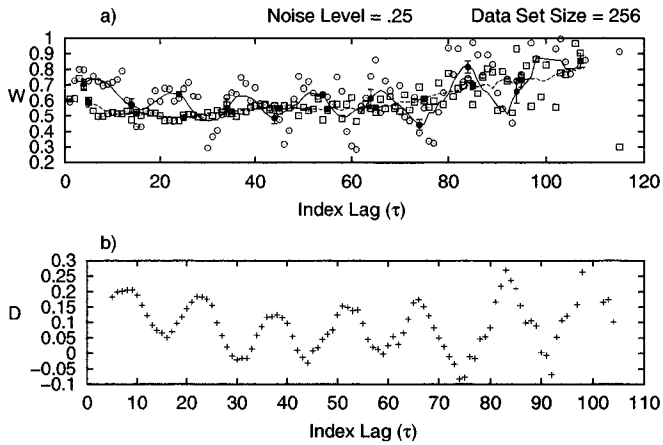

FIG. 84: W and D versus τ for Continuous Dynamical White Noise applied to the Lorenz system for 2^8 data points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

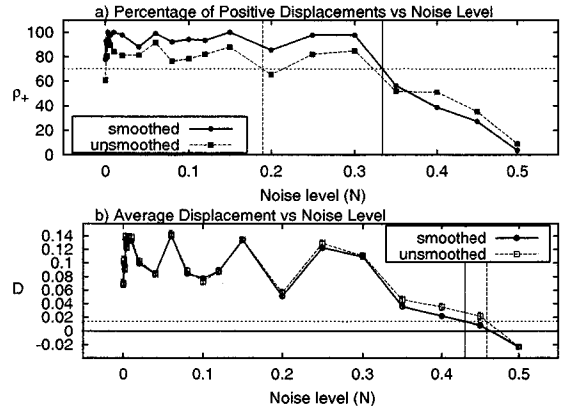

FIG. 85: Continuous Dynamical White Noise applied to the Lorenz system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

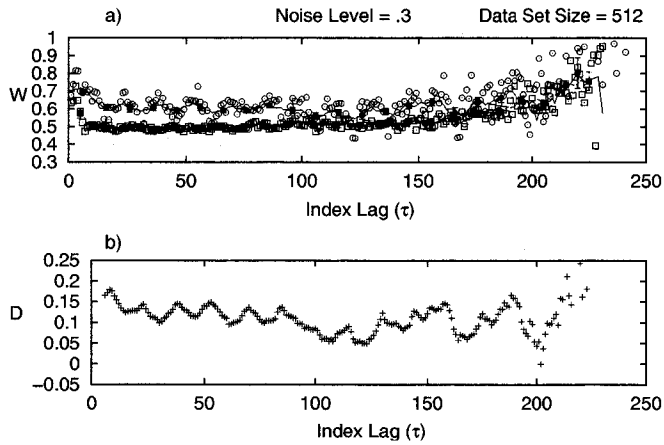

FIG. 86: W and D versus τ for Continuous Dynamical White Noise applied to the Lorenz system for 2^9 data points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

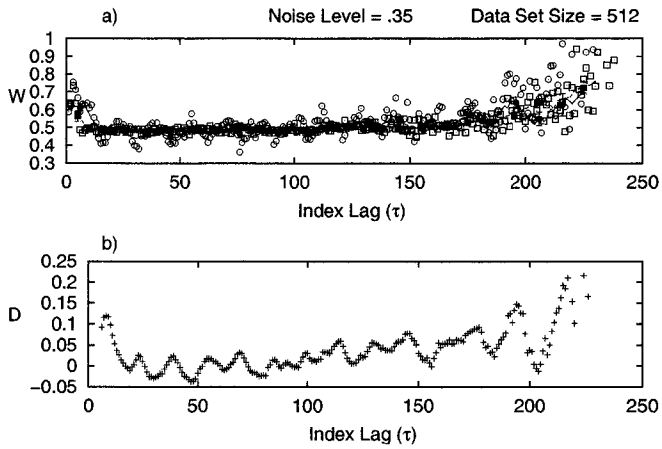

FIG. 87: W and D versus τ for Continuous Dynamical White Noise applied to the Lorenz system for 2^9 data points. Plot (a) shows W and $\mathcal{W}$ for both real and surrogate data. Plot (b) shows the displacement between $\mathcal{W}$ for real and surrogate data versus τ.

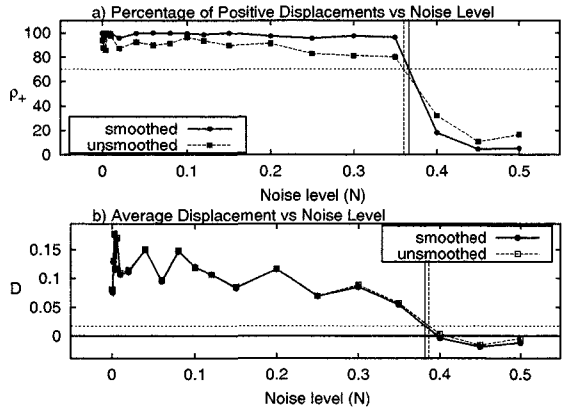

FIG. 88: Continuous Dynamical White Noise applied to the Lorenz system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

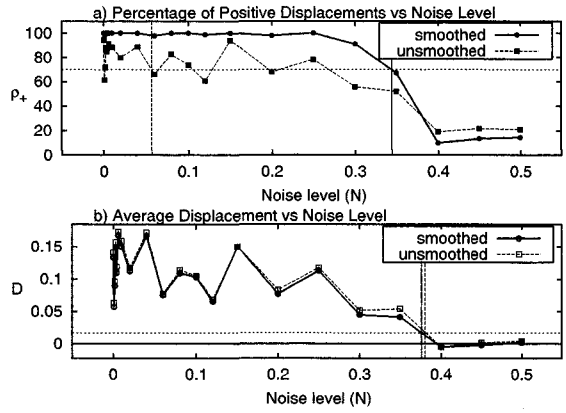

FIG. 89: Continuous Dynamical White Noise applied to the Lorenz system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

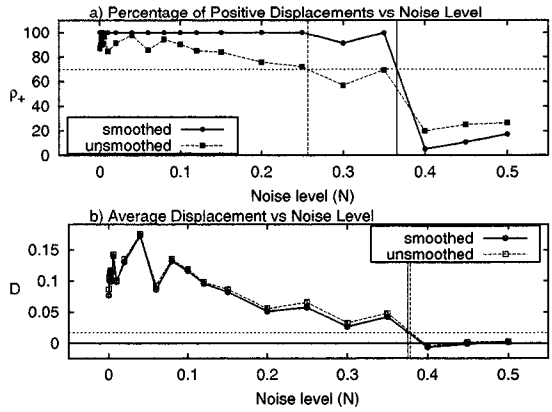

FIG. 90: Continuous Dynamical White Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

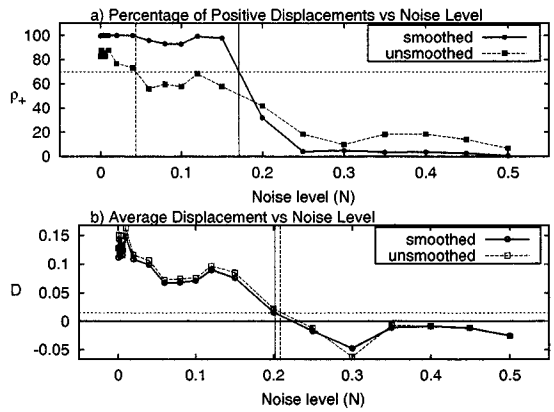

FIG. 91: Continuous Dynamical White Noise applied to the Lorenz system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

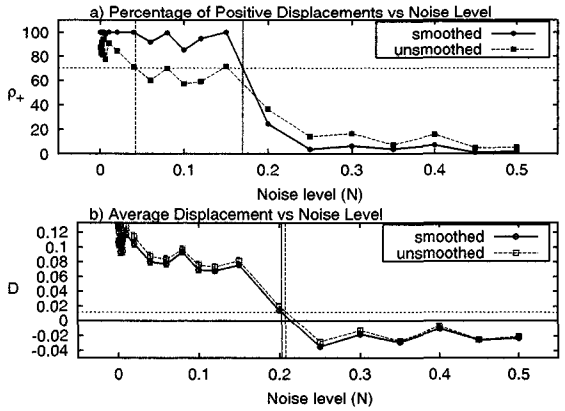

FIG. 92: Continuous Dynamical White Noise applied to the Lorenz system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

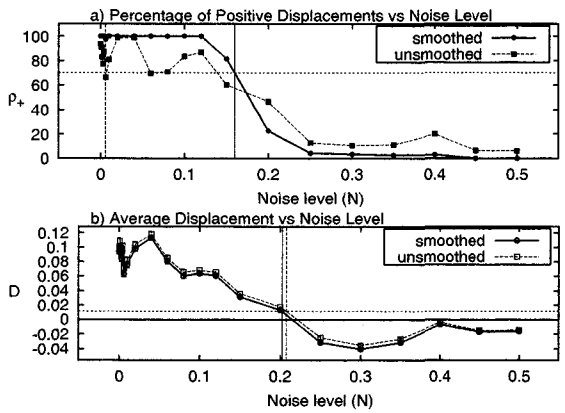

FIG. 93: Continuous Dynamical White Noise applied to the Lorenz system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

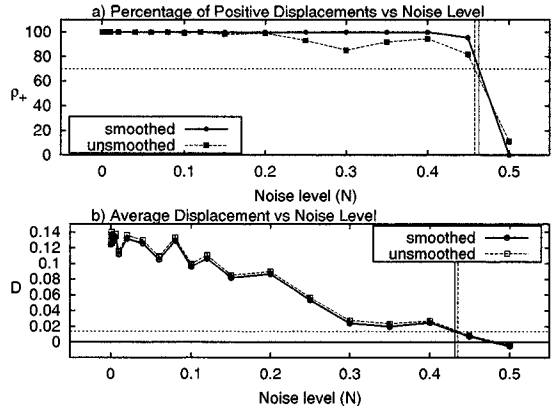

FIG. 94: Continuous Dynamical White Noise applied to the Lorenz system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

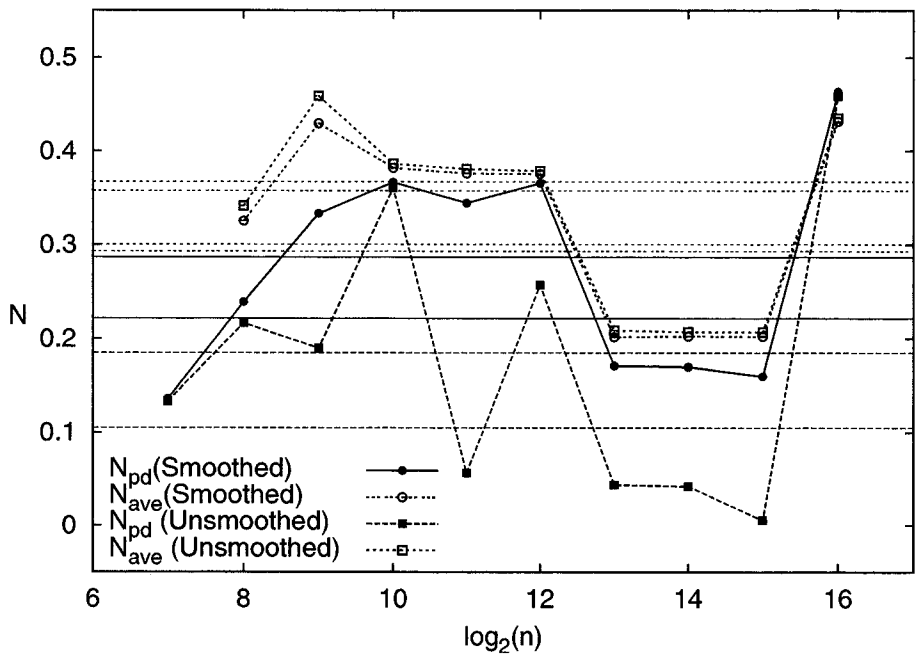

FIG. 95: Maximum noise cutoffs for the Lorenz system for Continuous Dynamical White Noise.

3. Summary of Results for Continuous Dynamical White Noise

This subsection summarizes the results presented in subsections IV F 1 and IV F 2. The maximum noise cutoffs will be related to the characteristic sizes of the systems being considered.

In the previous subsections, it was shown that, for smaller data set sizes, smoothing $W(\tau)$ assists in distinguishing between determinism and randomness. This conclusion is a result of observing a consistent separation between the curves of ρ_+ for smoothed functions and unsmoothed functions. When using the unsmoothed functions W_R and W_S, the maximum noise cutoff N_{pd} is occasionally underestimated, but this did not occur when using the smoothed functions $\mathcal{W}_R$ and $\mathcal{W}_S$. This indicates that the smoothed functions are more reliable in distinguishing between determinism and randomness. In all cases, the average separation is not effected by smoothing, indicating that the process of smoothing preserves the basic structure of $W(\tau)$.

In most cases, the approximation of the maximum noise cutoff N_{pd} obtained from ρ_+ using a threshold of $\rho_{thresh} = 70\%$ was correct. In most cases, there was evidence of determinism when ρ_+ was above 70%, but not when ρ_+ was below 70%. There are some exceptions to this rule of thumb, but inspecting the plot of average separation versus noise level gives the true value for the maximum noise cutoff.

For both the Rössler and Lorenz systems, a data set size of $2^8 = 256$ was a lower limit for obtaining reliable results.

Table IV summarizes the maximum noise cutoffs for both the Rössler and Lorenz systems obtained from Figs. 80 and 95. The average cutoffs $\bar{\mathcal{N}}_{pd}$ and $\bar{\mathcal{N}}_{ave}$ obtained from the smoothed functions $\mathcal{W}$, and the average cutoffs $\bar{N}_{pd}$ and $\bar{N}_{ave}$ obtained from the unsmoothed functions W, are given.

Section IV B described how to relate the noise level $\mathcal{N}$ to the percentage $\mathcal{P}$ of the size of the system to which the noise is being added. Since this type of noise

TABLE IV: Cutoffs for Continuous Dynamical White Noise

Attractor	$\bar{\mathcal{N}}_{pd}$	$\bar{\mathcal{N}}_{ave}$	$\bar{N}_{pd}$	$\bar{N}_{ave}$
Rössler	0.063 — 0.069	0.064 — 0.068	0.059 — 0.064	0.065 — 0.068
Lorenz	0.22 — 0.29	0.29 — 0.36	0.11 — 0.18	0.30 — 0.37

is added at every integration step, then $T = 1$ in Eq. 41, so

$$\mathcal{P} = 100\left(\frac{\mathcal{N}}{\sigma}\cdot\sqrt{MS}\right), \tag{45}$$

where σ is the characteristic size of the dynamical system, S is the sampling interval, and M is the characteristic period of the sytem. Using the values from Table III,

$$\begin{aligned}\mathcal{P}_{rossler} &\approx 1100N\ ,\\ \mathcal{P}_{lorenz} &\approx 400N\ .\end{aligned} \tag{46}$$

Using Eqs. 46 to convert the noise levels in Table IV to percentages of the size of the attractor for the dynamical system being studied yields Table V.

TABLE V: Percent Cutoffs for Continuous Dynamical White Noise

Attractor	$\bar{\mathcal{P}}_{pd}$	$\bar{\mathcal{P}}_{ave}$	$\bar{P}_{pd}$	$\bar{P}_{ave}$	$\bar{\mathcal{P}}_{pd}/\bar{P}_{pd}$	$\bar{\mathcal{P}}_{ave}/\bar{P}_{ave}$
Rössler	70% — 76%	70% — 75%	65% — 70%	71% — 75%	1.1	1.0
Lorenz	89% — 115%	117% — 143%	42% — 74%	120% — 147%	1.8	1.0

The last two columns in this table give the ratios of noise cutoffs obtained from the smoothed functions $\mathcal{W}$ and the unsmoothed functions W. In most cases, smoothing does not allow one to observe determinism with higher noise levels, although it does improve the reliability of the method in determining this noise level.

The results in the following sections are similar to those found here. Thus, most of the figures are deferred to appendices, and the focus is placed on the conclusions.

G. Results for Intermittent Dynamical White Noise

Intermittent Dynamical White Noise was studied for the Lorenz, Rössler, and Hénon systems. However, as for Continuous Dynamical White Noise, applying this type of noise to the Hénon attractor ejected the trajectory from the basin of attraction, and it ran off to infinity.

Results for Intermittent Dynamical White Noise applied to the Rössler and Lorenz systems are presented in subsections IV G 1 and IV G 2, respectively. Figures can be found in appendix VI A.

1. Intermittent Dynamical White Noise Applied to the Rössler System

The results for Intermittent Dynamical White Noise applied to the Rössler system are very similar to those for Continuous Dynamical White Noise applied to the Rössler system. The main difference occurs for a data set size of $2^n = 2^7 = 128$. Unlike before, the curves of ρ_+ and $\mathcal{D}$ versus N (Figs. 115a and 115b, respectively) do not display the large variations seen for Continuous Dynamical White Noise. In this case, it is possible to distinguish between determinism and randomness for a data set of this size. However, as has been seen earlier, data sets this small are generally unreliable.

As before, the curve of ρ_+ for the smoothed functions $\mathcal{W}$ is separated from the curve of ρ_+ for the unsmoothed functions W for data set sizes up to 2^{14}. For larger data set sizes, the two curves follow each other very closely. This indicates that the process of smoothing is most helpful for smaller data sets.

The maximum noise cutoffs $(\mathcal{N}_{pd}, \mathcal{N}_{ave}, N_{pd}, N_{ave})$ are all around a noise level of $N = 1$. This is shown explicitly in Fig. 96, which shows the maximum noise cutoffs as a function of the data set size n.

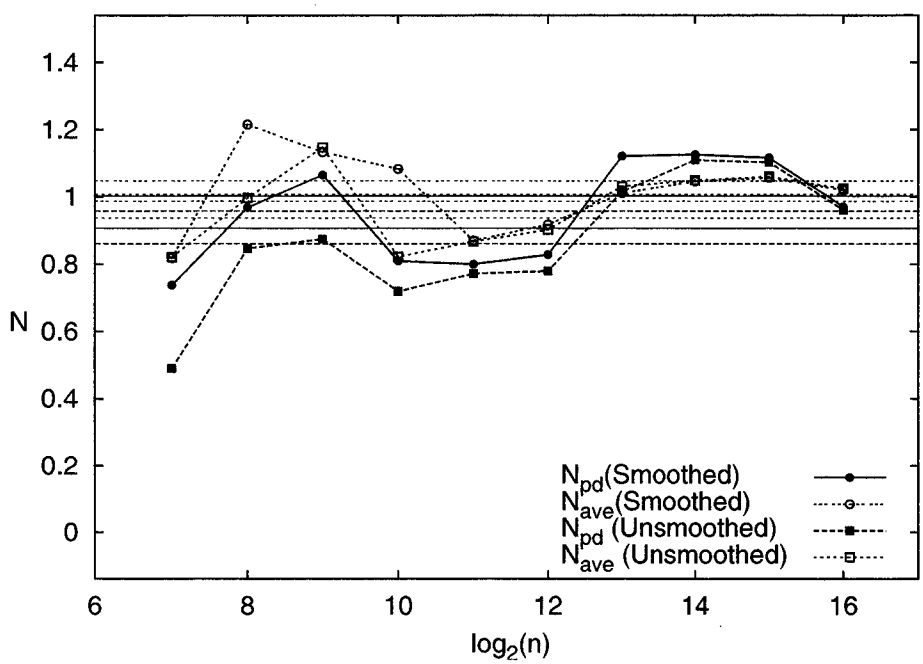

FIG. 96: Maximum noise cutoffs for the Rössler system for Intermittent Dynamical White Noise.

2. Intermittent Dynamical White Noise Applied to the Lorenz System

The results for Intermittent Dynamical White Noise applied to the Lorenz system are similar to those for the other noise types already discussed.

Figure 125a shows ρ_+ for a data set size of $2^n = 2^7$. As for Continuous Dynamical White Noise applied to the Lorenz system in Fig. 81a, fluctuations in ρ_+ make determining a maximum noise cutoff difficult. Again a data set size of 2^7 is insufficient to reliably distinguish between determinism and randomness.

In Figs. 126a, 128a, 129a, 131a, and 132a, the maximum noise cutoff N_{pd} for unsmoothed data obtained from the percentage of positive displacements ρ_+ is an underestimate.

The maximum noise cutoffs $(\mathcal{N}_{pd}, \mathcal{N}_{ave}, N_{ave})$ are all around a noise level of 1.75. The maximum noise cutoff N_{pd} for unsmoothed data obtained from the percentage of positive displacements ρ_+ is just below 1. As noted earlier, this quantity was underestimated for several data set sizes, which explains why it is not in line with the other average values of the maximum noise cutoffs. This is seen explicitly in Fig. 97, which shows the maximum noise cutoffs as a function of the data set size n.

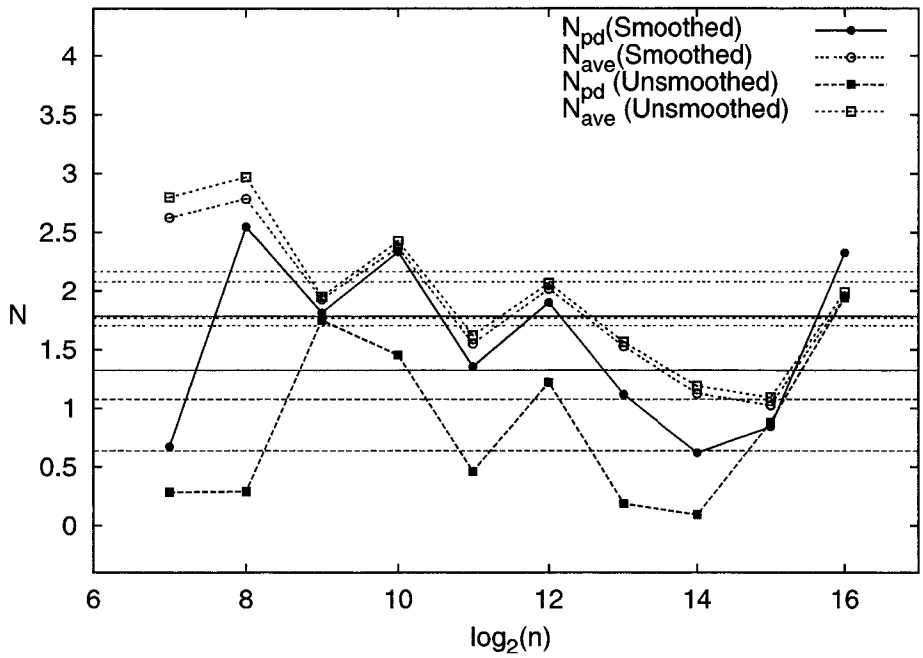

FIG. 97: Maximum noise cutoffs for the Lorenz system for Intermittent Dynamical White Noise.

3. Summary of Results for Intermittent Dynamical White Noise

This subsection summarizes the results observed in subsections IV G 1 and IV G 2. The maximum noise cutoffs will be related to the characteristic size of the systems being considered.

Table VI summarizes the maximum noise cutoffs for both the Rössler and Lorenz systems obtained from Figs. 96 and 97. The average cutoffs $\bar{\mathcal{N}}_{pd}$ and $\bar{\mathcal{N}}_{ave}$ obtained from the smoothed functions $\mathcal{W}$, and the average cutoffs $\bar{N}_{pd}$ and $\bar{N}_{ave}$ obtained from the unsmoothed functions W, are given. Remember that the maximum noise cutoff $\bar{N}_{pd}$ for the Lorenz system was shown to be unreliable.

TABLE VI: Cutoffs for Intermittent Dynamical White Noise

Attractor	$\bar{\mathcal{N}}_{pd}$	$\bar{\mathcal{N}}_{ave}$	$\bar{N}_{pd}$	$\bar{N}_{ave}$
Rössler	0.91 — 1.0	0.99 — 1.0	0.86 — 0.96	0.94 — 1.0
Lorenz	1.3 — 1.8	1.7 — 2.1	0.64 — 1.1	1.8 — 2.2

This type of noise is added on the same integration step that the trajectory is sampled to generate a time series. For this noise type S/T in Eq. 41 equals one, so Eq. 41 becomes

$$\mathcal{P} = 100\left(\frac{\mathcal{N}}{\sigma}\sqrt{M}\right). \tag{47}$$

Using the values from Table III,

$$\begin{aligned} \mathcal{P}_{rossler} &\approx 77N\ , \\ \mathcal{P}_{lorenz} &\approx 56N\ . \end{aligned} \tag{48}$$

Using Eqs. 48 to convert the noise levels in Table VI to percentages of the size of the attractor for the dynamical system being studied yields Table VII.

In general, the values of the maximum noise cutoffs are in agreement with those found for Continuous Dynamical White Noise. The next section will compare Continuous Dynamical White Noise and Intermittent Dynamical White Noise in more detail.

TABLE VII: Cutoffs for Intermittent Dynamical White Noise

Attractor	$\bar{\mathcal{P}}_{pd}$	$\bar{\mathcal{P}}_{ave}$	$\bar{P}_{pd}$	$\bar{P}_{ave}$	$\bar{\mathcal{P}}_{pd}/\bar{P}_{pd}$	$\bar{\mathcal{P}}_{ave}/\bar{P}_{ave}$
Rössler	70% — 77%	76% — 81%	66% — 74%	72% — 78%	1.1	1.0
Lorenz	74% — 100%	95% — 116%	36% — 60%	99% — 121%	1.8	1.0

The last two columns in this table give the ratios of noise cutoffs obtained from the smoothed functions $\mathcal{W}$ and the unsmoothed functions W. In most cases, smoothing W does not allow us to observe determinism with higher noise levels, although it does improve the reliability of the method in determining this noise level.

H. Comparison Between Continuous and Intermittent Dynamical White Noise

This section summarizes and comparesthe results for Continuous Dynamical White Noise and Intermittent Dynamical White Noise presented in sections IV F and IV G. Tables VIII and IX summarize the results for the Rössler and Lorenz systems, respectively.

TABLE VIII: Comparison Between Continuous and Intermittent Dynamical White Noise for the Rössler Attractor

Cutoff Type	Continuous	Intermittent	Cont/Int
$\bar{\mathcal{P}}_{pd}$	70% — 76%	70% — 77%	1.0
$\bar{\mathcal{P}}_{ave}$	70% — 75%	76% — 81%	0.9
$\bar{P}_{pd}$	65% — 70%	66% — 74%	1.0
$\bar{P}_{ave}$	71% — 75%	72% — 78%	1.0

TABLE IX: Comparison between Continuous and Intermittent Dynamical White Noise for the Lorenz Attractor

Cutoff Type	Continuous	Intermittent	Cont/Int
$\bar{\mathcal{P}}_{pd}$	89% — 115%	74% — 100%	1.2
$\bar{\mathcal{P}}_{ave}$	117% — 143%	95% — 116%	1.2
$\bar{P}_{pd}$	42% — 74%	36% — 60%	1.2
$\bar{P}_{ave}$	120% — 147%	99% — 121%	1.2

The last column of each table shows the ratios of the noise cutoffs for Continuous Dynamical White Noise and Intermittent Dynamical White Noise. For the Rössler system, the amount of noise that can be added to a system before the noise obscures the dynamics enough so that determinism can no longer be detected is not dependent on whether noise is added at every integration step or added inter-

mittently. However, for the Lorenz system, determinism can be detected for noise levels 1.2 times higher for continuous noise than for intermittent noise.

For the Rössler system, the two cases of Dynamical Noise yield very similar results. Both cases have maximum noise cutoffs in the range of 70%–80%. For the Lorenz system, there is a difference between the maximum noise cutoffs for continuous noise and intermittent noise. The maximum noise cutoff for Continuous Dynamical White Noise applied to the Lorenz system is in the range of 100%–130%, while it is in the range of 90%–110% for Intermittent Dynamical White Noise. In this case, determinism can be distinguished from randomness even when the noise source is the dominant contributor to the system.

I. Results for Additive White Noise

Additive White Noise was studied for the Lorenz, Rössler, and Hénon systems. These results are presented in subsections IVI1, IVI2, and IVI4, respectively. Figures can be found in appendix VIB.

1. Additive White Noise Applied to the Rössler System

The results for Additive White Noise applied to the Rössler system are very similar to the results for the other noise types already discussed.

Figure 135a shows ρ_+ for a data set size of 2^7. As for other noise types, it is difficult to determine a maximum noise cutoff for data sets this small. Again, a data set size of 2^7 is insufficient to reliably distinguish between determinism and randomness.

The plot of ρ_+ for a data set size of 2^{12} is shown in Fig. 140a. In this case, it is the maximum noise cutoff obtained from the average displacement $\mathcal{D}$, shown in Fig. 140b, that is the actual maximum noise cutoff for this data set size. This behavior was seen before in Fig 70.

The maximum noise cutoffs $(\mathcal{N}_{pd}, \mathcal{N}_{ave}, N_{pd}, N_{ave})$ are all around a noise level of 2. This is shown explicitly in Fig. 98, which shows the maximum noise cutoffs as a function of n.

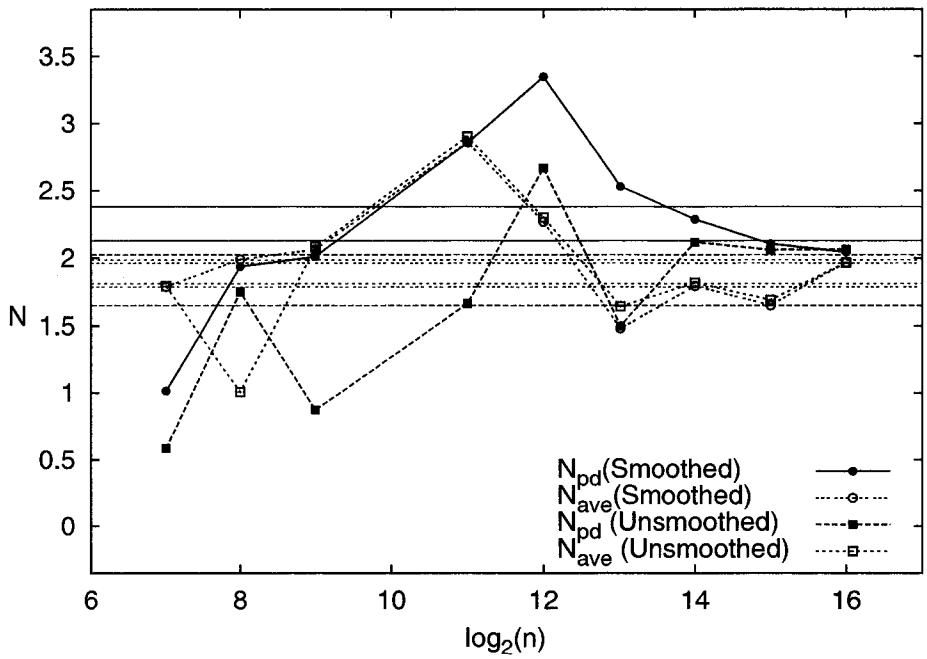

FIG. 98: Maximum noise cutoffs for the Rössler system for Additive White Noise.

2. Additive White Noise Applied to the Lorenz System

The results for Additive White Noise applied to the Lorenz system are very similar to the results for the other noise types already discussed.

Figure 145a shows ρ_+ for a data set size of 2^7. As for other noise types, it is difficult to determine a maximum noise cutoff for data sets this small. Again, a data set size of 2^7 is insufficient to reliably distinguish between determinism and randomness.

The plot of $\mathcal{D}$ for a data set size of 2^{11} is shown in Fig. 150b. In this case, the maximum noise cutoff N_{ave} obtained from the average displacement $\mathcal{D}$ for unsmoothed data is an overestimate, because the curve asymptotes to zero slower than is typical.

The maximum noise cutoffs $(\mathcal{N}_{pd}, \mathcal{N}_{ave}, N_{pd}, N_{ave})$ are all around a noise level of 1.5. This is shown explicitly in Fig. 99, which shows the maximum noise cutoffs as a function of n. The high peak in the curve of N_{ave} at $n = 11$ is because this cutoff was overestimated as noted above. This point was not used to compute the average noise cutoff $\bar{N}_{ave}$.

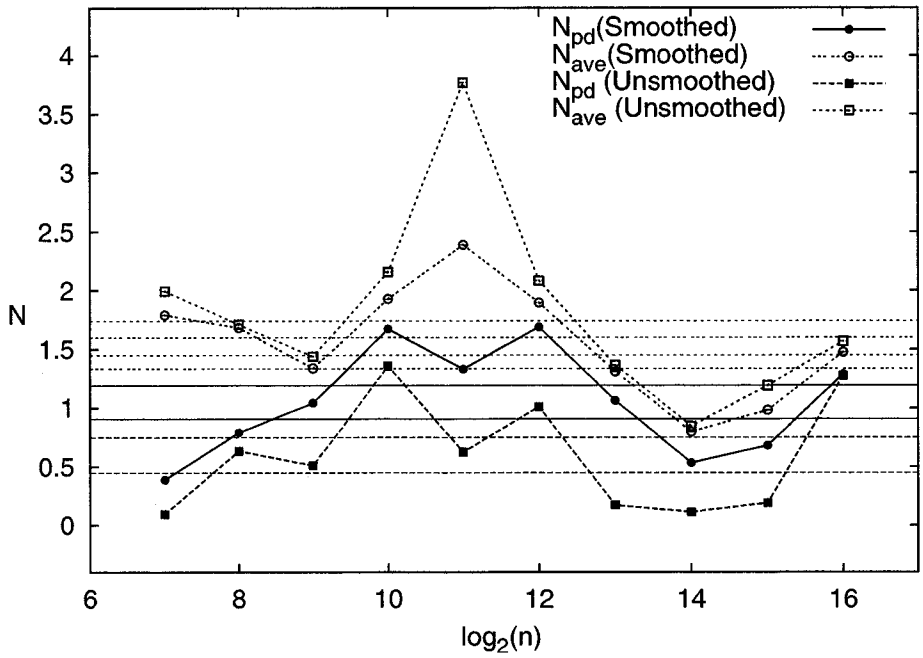

FIG. 99: Maximum noise cutoffs for the Lorenz system for Additive White Noise.

3. Summary of Results for Additive White Noise Applied to the Rössler and Lorenz Systems

This subsection summarizes the results observed in subsections IV I 1 and IV I 2. The maximum noise cutoffs will be related to the characteristic size of the systems being considered.

Table X summarizes the maximum noise cutoffs for both the Rössler and Lorenz systems obtained from Figs. 98 and 99. The average cutoffs $\bar{\mathcal{N}}_{pd}$ and $\bar{\mathcal{N}}_{ave}$ obtained from the smoothed functions $\mathcal{W}$, and the cutoffs $\bar{N}_{pd}$ and $\bar{N}_{ave}$ obtained from the unsmoothed functions W, are given.

TABLE X: Cutoffs for Additive White Noise

Attractor	$\bar{\mathcal{N}}_{pd}$	$\bar{\mathcal{N}}_{ave}$	$\bar{N}_{pd}$	$\bar{N}_{ave}$
Rössler	2.1 — 2.4	1.8 — 2.0	1.6 — 2.0	1.8 — 2.0
Lorenz	0.9 — 1.2	1.3 — 1.6	0.45 — 0.75	1.45 — 1.7

For this noise type, noise is added on the same integration step that the trajectory is sampled to generate a time series, so $S/T = 1$ in Eq. 41 and

$$\mathcal{P} = 100\left(\frac{\mathcal{N}}{\sigma}\sqrt{M}\right). \tag{49}$$

Using the values from Table III,

$$\begin{aligned} \mathcal{P}_{rossler} &\approx 77N\ , \\ \mathcal{P}_{lorenz} &\approx 56N\ . \end{aligned} \tag{50}$$

Using Eqs. 50 to convert the noise levels in Table X to percentages of the size of the attractor for the dynamical system being studied yields Table XI.

Notice that determinism can be distinguished from randomness for Additive White Noise for the Rössler system for higher noise levels than for any other noise type.

TABLE XI: Cutoffs for Additive White Noise

Attractor	$\bar{\mathcal{P}}_{pd}$	$\bar{\mathcal{P}}_{ave}$	$\bar{P}_{pd}$	$\bar{P}_{ave}$	$\bar{\mathcal{P}}_{pd}/\bar{P}_{pd}$	$\bar{\mathcal{P}}_{ave}/\bar{P}_{ave}$
Rössler	164% — 183%	138% — 151%	127% — 156%	139% — 153%	1.2	1.0
Lorenz	51% — 67%	75% — 90%	25% — 42%	81% — 97%	1.8	1.0

The last two columns in this table are the ratios of noise cutoffs obtained from the smoothed functions $\mathcal{W}$ and the unsmoothed functions W. In most cases, smoothing does not allow us to observe determinism with higher noise levels, although it does improve the reliability of the method in determining this noise level.

4. Additive White Noise Applied to the Hénon System

Additive White Noise was the only type of noise that could successfully be applied to the Hénon attractor without causing it to be ejected from the basin of attraction.

Results for data set sizes of 2^7 and 2^8 are shown in Figs. 155 and 156, respectively. These results are very similar to the results seen for data sets of 2^7 points for the Rössler and Lorenz systems. For the Hénon system a larger data set is necessary to obtain reliable results.

Results for data set sizes of 2^9 and 2^{10} are shown in Figs. 157 and 158. The curves of ρ_+ for these data set sizes have some points above the 70% threshold for noise levels higher than the estimated maximum noise cutoff $\mathcal{N}_{pd}$. A closer examination of the W curves revealed that the cutoff $\mathcal{N}_{ave}$ is correct. This was seen before in Fig. 70 for Continuous Dynamical White Noise applied to the Rössler system, and it indicates that there are occasions when both of the quantities ρ_+ and $\mathcal{D}$ are needed to determine the correct value of the maximum noise cutoff.

Results for data sets of 2^{11} points are shown in Fig. 159. The maximum noise cutoffs for smoothed data, $\mathcal{N}_{pd}$ and $\mathcal{N}_{ave}$, are both around $N = 1.75$. The cutoff N_{pd} for unsmoothed data is around $N = 1.25$. Closer examination reveals that the values for the smoothed functions are correct. This behavior has been seen repeatedly before, and it is an indication that smoothing the data yields more reliable results.

The results for data set sizes of 2^{12} and 2^{13} are shown in Figs. 160 and 161, respectively. The maximum noise cutoffs for these noise levels are noticebly lower than those found for the other data set sizes. This behavior was seen before for data set sizes of 2^{13}, 2^{14}, and 2^{15}, shown in Figs 91–93, for Continuous Dynamical White Noise applied to the Lorenz system. As in that case, it is uncertain why this is the case.

For data set sizes of 2^{15} and 2^{16}, the curve for ρ_+ never decreases below the

70% threshold. However, the average separation $\mathcal{D}$ does drop below its threshold. Close inspection of the W curves reveals the cutoffs $\mathcal{D}_{ave}$ are, in fact, correct.

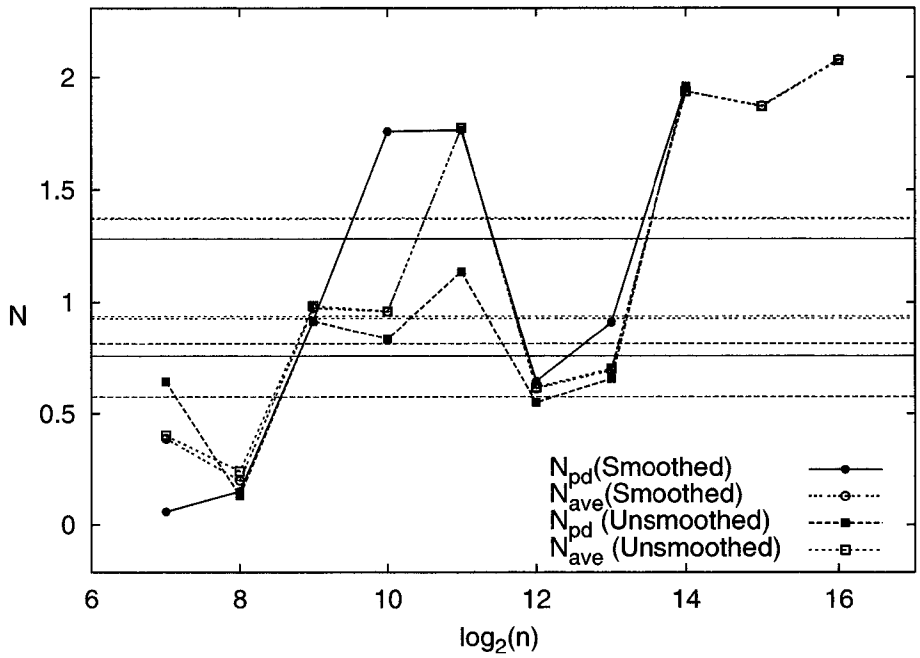

FIG. 100: Maximum noise cutoffs for the Hénon system for Additive White Noise.

J. Comparison Between Intermittent Dynamical White Noise and Additive White Noise

This section summarizes and compares the results for Intermittent Dynamical White Noise and Additive White Noise presented in sections IV G and IV I.

Tables XII and XIII summarize the results for the Rössler and Lorenz systems, respectively.

TABLE XII: Comparison Between Additive White Noise and Intermittent Dynamical White Noise for the Rössler Attractor

Cutoff Type	Additive	Dynamical	Add/Dyn
$\bar{\mathcal{P}}_{pd}$	164% — 183%	70% — 77%	2.4
$\bar{\mathcal{P}}_{ave}$	138% — 151%	76% — 81%	1.8
$\bar{P}_{pd}$	127% — 156%	66% — 74%	2.0
$\bar{P}_{ave}$	139% — 153%	72% — 78%	1.9

TABLE XIII: Comparison Between Additive White Noise and Intermittent Dynamical White Noise for the Lorenz Attractor

Cutoff Type	Additive	Dynamical	Add/Dyn
$\bar{\mathcal{P}}_{pd}$	51% — 67%	74% — 100%	0.7
$\bar{\mathcal{P}}_{ave}$	75% — 90%	95% — 116%	0.8
$\bar{P}_{pd}$	25% — 42%	36% — 60%	0.7
$\bar{P}_{ave}$	81% — 97%	99% — 121%	0.8

The last column of each table shows the ratios of the noise cutoffs for Additive White Noise and Intermittent Dynamical White Noise. For the Rössler system, the amount of noise that can by added to a system is about 2 times higher for Additive White Noise, compared to Intermittent Dynamical White Noise. However, for the Lorenz system, determinism can be distinguished from randomness for slightly

higher noise levels for Intermittent Dynamical White Noise than for Additive White Noise.

Since Additive White Noise does not effect the dynamics of the system, one would expect to be able to detect determinism at higher noise levels for Additive Noise than for Dynamical Noise. This is the case for the Rössler system, but the Lorenz system is just the opposite. The reason for this is uncertain.

K. Results for Chaotic Noise

This section examines the effects of noise generated from a chaotic system. The goal of this section is to determine if two competing chaotic systems will appear deterministic, or if this will fool the method of Salvino and Cawley into thinking the system is random.

In previous sections, it was found that there is not a strong correlation between the size of the data set and the ability to detect determinism, as long as the data set is larger than 2^7. Thus, in this section, only data sets containing $2^{12} = 4096$ points are considered.

The noise levels reported in this section are the fractional noise level f discussed in section IV B. The magnitude N of noise added on a given noise step is related to the fractional noise level f by Eq. 43, where σ in that equation is the characteristic size of the system being used as a source of noise. Table III contains the values of σ for the Rössler and the Lorenz systems.

Initially, the time series was embedded in three dimensions, and in most cases, there was evidence of determinism. However, there were several combinations of noise type, dynamical system, and noise level, that showed evidence of randomness instead. For example, Intermittent Dynamical Lorenz Noise for a fractional noise level of $f = 1.0$ was applied to the Lorenz system to create a time series. The extended method of Salvino and Cawley was applied to this time series using an embedding dimension of $m = 3$. The functions $W(\tau)$ and $\mathcal{W}(\tau)$ are shown in Fig. 101a. Even though both components are deterministic, this figure shows that, for this noise level, embedding in three dimensions does not yield evidence of determinism.

When two chaotic systems are combined, the resulting system has a dimension larger than the dimension of either of the two components. Thus, all analyses were repeated using an embedding dimension of $m = 6$. Figure 102a shows the W curves for a time series generated from the Lorenz system with Intermittent Dynamical

Lorenz Noise as the source of noise, and a fractional noise level of $f = 1.0$. For this trial, the time series was embedded in six dimensions. This figure shows distinct evidence of determinism.

For chaotic Dynamical Noise a large noise level will eject the trajectory of the primary system from the basin of attraction. For all noise levels that could be run without this occurring, there was evidence of determinism when an embedding dimension of $m = 6$ was used.

Table XIV is a summary of the results for chaotic noise. All of the trials showed evidence of determinism. The fractional noise levels reported for these noise types, are the largest that could be used before the trajectories were ejected from the basin of attraction. For Additive Noise, the fractional noise levels given are equivalent to roughly 2700% for the Rössler system, and 4300% for the Lorenz system.

The values of N in Table XIV are obtained from Eq. 43 using the value of σ from Table III. The percent noise P is obtained from Eqs. 46 for Continuous Dynamical Rössler Noise and Continuous Dynamical Lorenz Noise, and from Eqs. 48 for the other noise types.

TABLE XIV: Summary of Results for Non–White Noise

Dynamical System ⇒		Rössler				Lorenz		
Noise Type ⇓	f	N	P	Fig.	f	N	P	Fig.
Continuous Dynamical Rössler Noise	0.075	0.53	605%	103	2.0	14.2	5700%	104
Continuous Dynamical Lorenz Noise	0.001	0.015	16%	105	0.2	2.9	1200%	106
Intermittent Dynamical Rössler Noise	0.02	0.14	11%	107	6.3	45	2560%	108
Intermittent Dynamical Lorenz Noise	0.02	0.29	22%	109	4.0	59	3280%	110
Additive Rössler Noise	5.0	36	2750%	111	11.0	78	4400%	112
Additive Lorenz Noise	2.4	35	2700%	113	5.0	74	4160%	114

There are two interesting things in this table. First, for most combinations of noise type and dynamical system, the largest noise level that could be run is much

larger than 100%. In these cases, the source of noise is the dominant contributor to the time series. This shows that the extended method of Salvino and Cawley can detect determinism for coupled chaotic systems.

The second thing to notice is that, when Continuous Dynamical Lorenz Noise, Intermittent Dynamical Rössler Noise, or Intermittent Dynamical Lorenz Noise, is applied to the Rössler system, the trajectory is ejected from the basin of attraction at low noise levels. At first glance, it would appear that the trajectory is ejected from the basin of attraction more easily for chaotic noise than for white noise. However, this is not the case. For white noise, the largest amount of noise that can be added on a given noise step is N. However, for chaotic noise, N denotes the average amount of noise added per step, and the maximum amount of noise added can be much larger than N

To verify that embedding in higher dimensions does not produce false results, some of the time series of the previous sections were reexamined using an embedding dimension of $m = 6$. Embedding in higher dimensions did not affect the result in any of these cases. Cases that indicated determinism still indicated determinism, and cases that indicated randomness still indicated randomness.

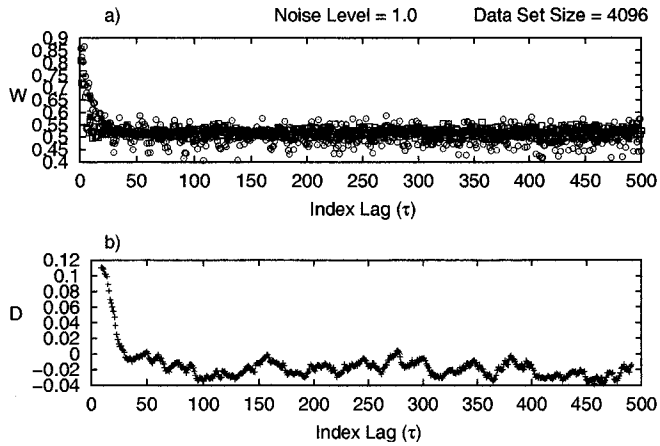

FIG. 101: W and D versus index lag for Intermittent Dynamical Lorenz Noise applied to the Lorenz system embedded in three dimensions for a fractional noise level of $f = 1.0$.

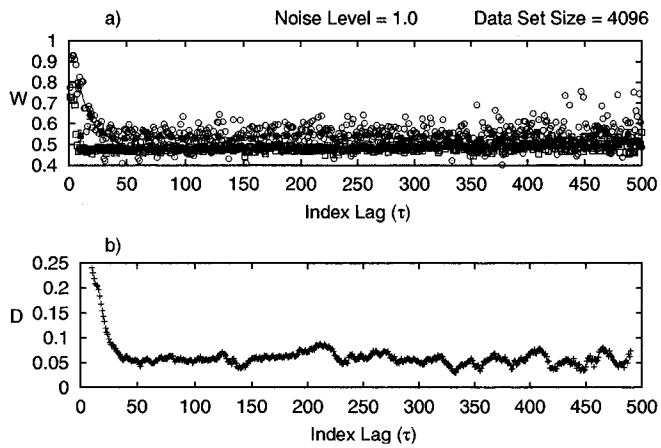

FIG. 102: W and D versus index lag for Intermittent Dynamical Lorenz Noise applied to the Lorenz system embedded in six dimensions for a fractional noise level of $f = 1.0$.

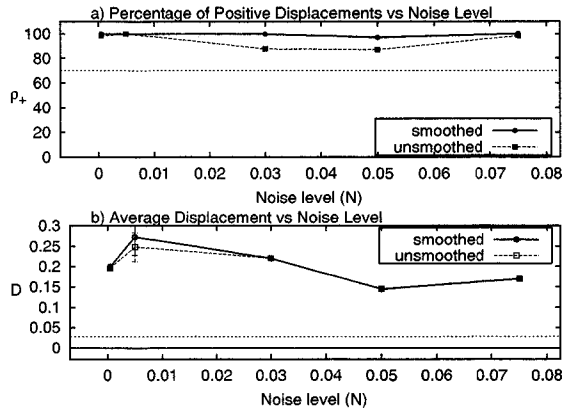

FIG. 103: Continuous Dynamical Rössler Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

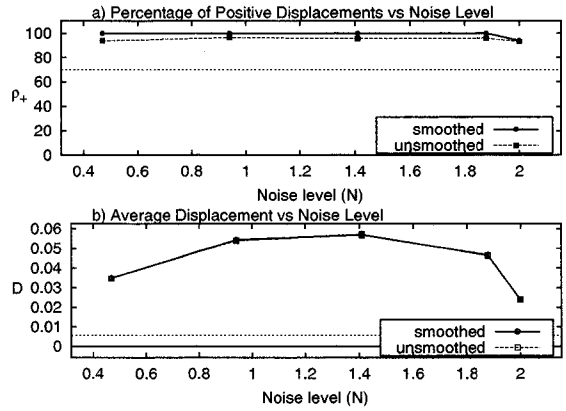

FIG. 104: Continuous Dynamical Rössler Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

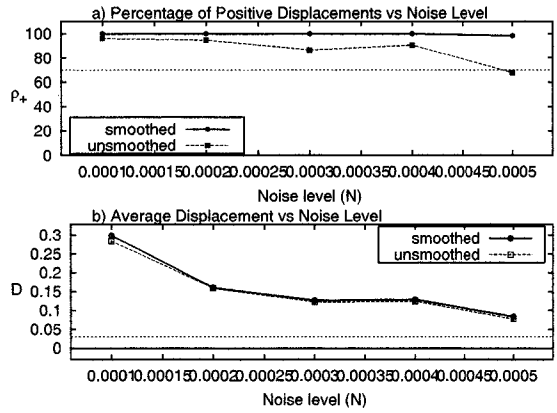

FIG. 105: Continuous Dynamical Lorenz Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

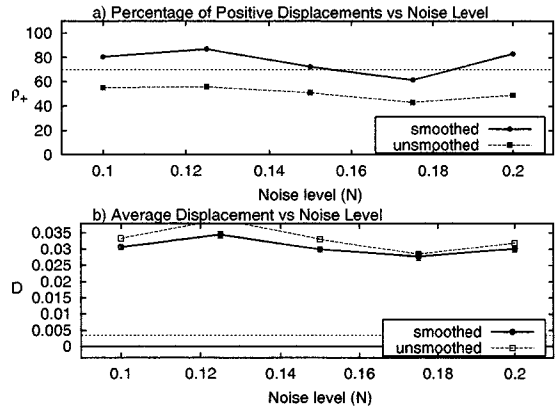

FIG. 106: Continuous Dynamical Lorenz Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

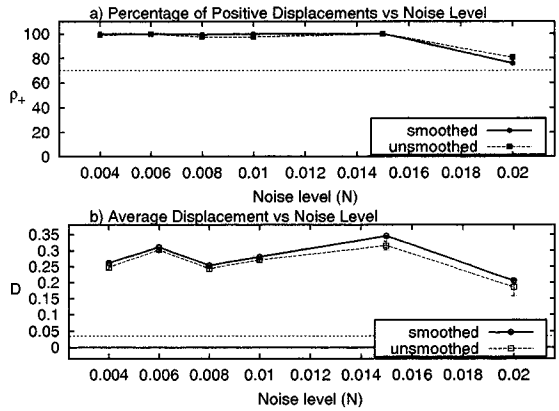

FIG. 107: Intermittent Dynamical Rössler Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

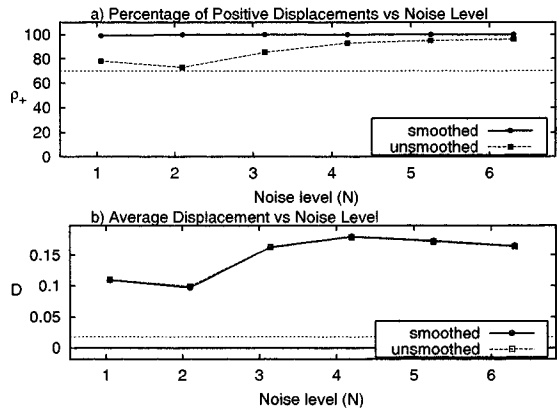

FIG. 108: Intermittent Dynamical Rössler Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

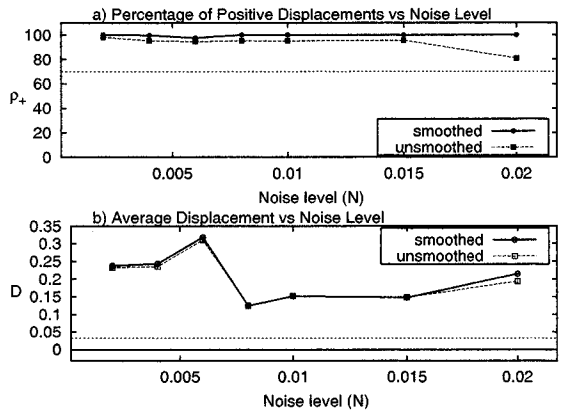

FIG. 109: Intermittent Dynamical Lorenz Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

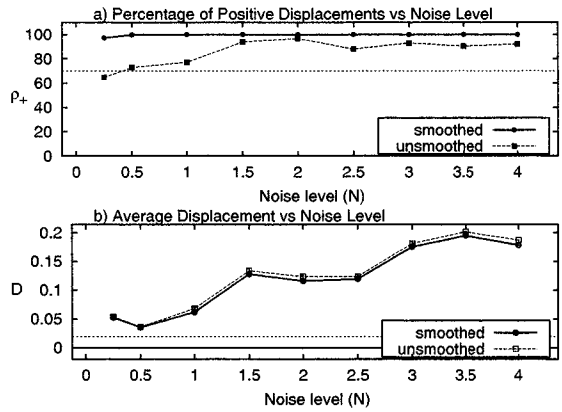

FIG. 110: Intermittent Dynamical Lorenz Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

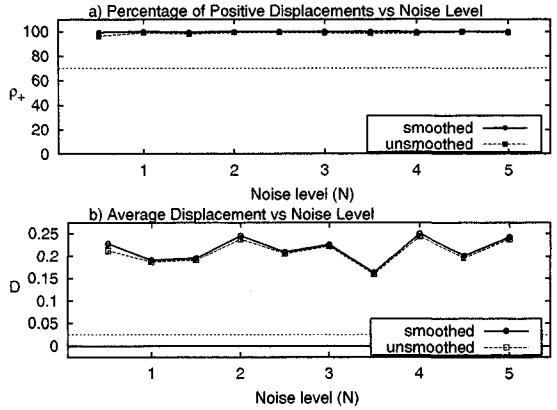

FIG. 111: Additive Rössler Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

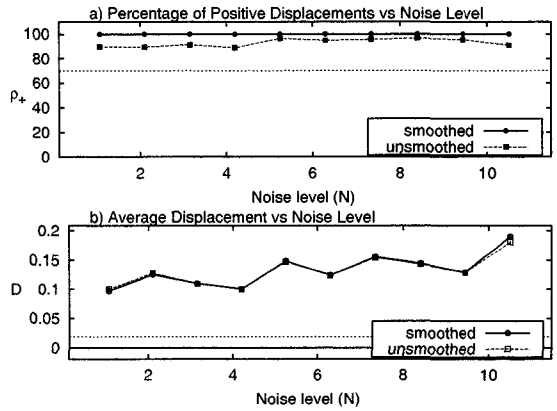

FIG. 112: Additive Rössler Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

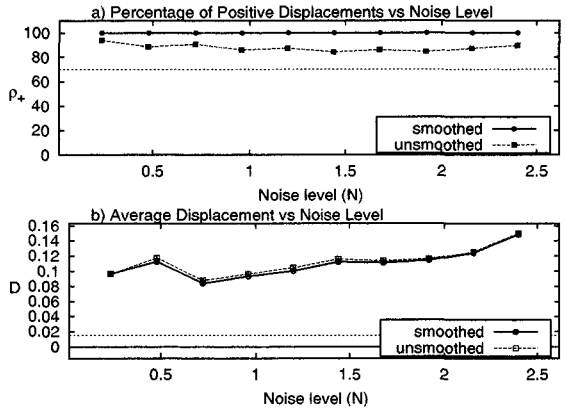

FIG. 113: Additive Lorenz Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

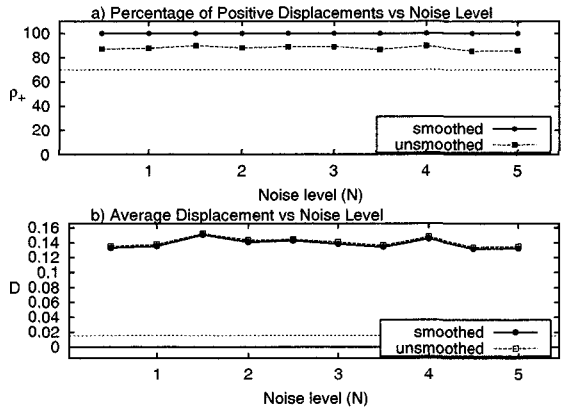

FIG. 114: Additive Lorenz Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

V. SUMMARY AND DISCUSSION

This dissertation has addressed three main subjects: Implementation details for using the method of Salvino and Cawley, extensions to the method of Salvino and Cawley, and effects of noise on detecting determinism. These are considered in turn, followed by suggestions for future research.

A. Implementation Details for Using the Method of Salvino and Cawley

Trajectories produced from deterministic processes possess a measure of smoothness that random trajectories do not. A time series is a one-dimensional sampling of a higher dimensional phase space trajectory. A time delay embedding of a time series allows one to reconstruct a higher dimensional, topologically equivalent trajectory from the one-dimensional time series. One can use this reconstructed trajectory to distinguish determinism from randomness by measuring the smoothness of the embedded trajectory.

The method of Salvino and Cawley uses the statistic W to measure the smoothness of an embedding, and, thus, to distinguish determinism from randomness. W is calculated by coarse-graining an embedded time series into small boxes. A unit vector field is defined over the embedded trajectory, and the vector average of all vectors in a given grid box is a local property of the trajectory. The length of the average vector is nearly one for a deterministic process, and it is nearly zero for a random process. W is the global weighted average of the squares of the lengths of these average vectors, and, as such it is a global measure of the smoothness of a trajectory. For a deterministic process, W is nearly one; whereas for a random process, W is nearly zero.

To use the method of Salvino and Cawley, the embedding must be coarse-grained, which requires a grid size to be chosen. The size of the grid cubes cannot be too large, or else the average of the unit vectors is no longer a local property of

the trajectory. On the other hand, if the grid is too small, there will not be enough embedding points in most of the cubes to yield good statistics. Even for a good choice of grid size, parts of the trajectory are not visited frequently enough, and those grid cubes cannot be used in the calculation of W. Since some grid cubes will not be visited sufficiently, they will not be used in the calculation of W, and so it is possible that there will not be enough sufficiently visited grid cubes to perform the global average needed to compute W. Index lags for which the calculation of W cannot be performed are termed invalid index lags, and those for which W can be calculated are termed valid index lags.

There is usually no *a priori* knowledge of the size of the grid cubes, or what the visitation contraints should be. Because of this, it is desireable to have an algorithm to choose these parameters. Section III A presented an algorithm for determining an optimal grid size that maximizes visitation of the grid cubes without allowing the size of the cubes to become too large. Briefly, the algorithm counts the number of valid index lags, and averages the number of suffiently visited boxes over all index lags. It then computes a statistic S that is a linear combination of these two quantities. The user specifies a range of values for the grid size, the minimum number of points in a box, and the minimum number of sufficiently visited boxes. The algorithm considers all combinations of these parameters to compute S. The set of parameters that maximizes S yields an optimal grid size and optimal visitation constraints.

Because data sets are finite, W will be less than one for deterministic data, and it will be greater than zero for random data. A concrete criterion is needed to distinguish between determinism and randomness. Randomization of a time series that was originally created from a deterministic process will reduce the value of W; whereas, randomization of a time series produced by a random process will not reduce W. To distinguish between determinism and randomness, randomized surrogate data is generated from the original time series, and the value of W for the original time series is compared to the value of W for the surrogate data.

Surrogate data is created by randomizing the phases of the Fourier transform of the time series and then inverse Fourier transforming back to the original space. For deterministic data, there will be a separation between the two values of W for real data and surrogate data, but, for random data, there will be no separation.

Surrogate data tends to be more spread out over the volume of the trajectory. Because of this, more boxes are visited, and so there are fewer points per box. This in turn leads to fewer valid index lags. This is especially pronounced for small data sets. Section III B presented a method for handling small surrogate data sets by considering multiple surrogate data sets and averaging W over all of them. This reduces the number of invalid index lags and allows a better comparison between real and surrogate data.

B. Extensions to the Method of Salvino and Cawley

Adding noise to a deterministic system causes the function $W(\tau)$ to fluctuate from index lag to index lag. These fluctuations make distinguishing between determinism and randomness more difficult and often unreliable.

Section III C presented a method for smoothing $W(\tau)$ that makes distinguishing between determinism and randomness easier, more reliable, and possible to automate. This is important for systems with a noisy component, and it is especially helpful for small data sets.

Smoothing W at a given index lag τ is performed by averaging over a set number of nearby index lags — called the local neighborhood. The size of the local neighborhood must be large enough that smoothing W has an effect, but not so large as to average out the basic structure of the $W(\tau)$ curve.

To determine the optimal local neighborhood size τ_n, a measure of the smoothness of the resulting smoothed curves is needed. The metric ξ was defined as the average deviation between successive index lags, and it is a measure of the smoothness of the $W(\tau)$ curve. A range of local neighborhood sizes is considered,

and the smoothness metric is computed for each neighborhood size, which yields a function $\xi(\tau_n)$. This function is examined to determine the neighborhood size for which it passes through its first local minimum, for which it drops to $1/e$ of its initial value, or for which its slope drops to $1/e$ of its initial value. The smallest of these neighborhood sizes is taken as the optimal local neighborhood size.

Smoothing W enhances the distinction between the two curves for real and surrogate data allowing a reliable characterization of the separation between them. This separation is characterized by two different quantities. The first quantity is the percentage of index lags for which the displacement of the real data from the surrogate data is positive within error this is called the percentage of positive displacements ρ_+, and it is a measure of the persistence of positive separation. The second quantity is the average separation $\mathcal{D}$ between the two curves, and it is a measure of the strength of the separation.

In section IV, it was shown that ρ_+ is larger for smoothed W curves than for unsmoothed curves. This indicates that smoothed W curves provide a more reliable determination of determinism or randomness.

For each noise type, a range of noise levels was studied to determine the noise level above which determinism can no longer be detected. The percentage ρ_+ of positive displacements was used to estimate the maximum noise cutoff N_{pd}. Using the unsmoothed functions W_R and W_S, the approximation of the maximum noise cutoff N_{pd} is occasionally underestimated, but this did not occur when using the smoothed functions $\mathcal{W}_R$ and $\mathcal{W}_S$. This indicates that smoothing W makes the procedure more reliable. It also allows it to be automated reliably.

For random data, there can be a significant percentage of the index lags for which the smoothed curve $\mathcal{W}_R$ is above the smoothed curve $\mathcal{W}_S$. This is not an artifact of smoothing, but is a result of the fact that, for random data, the curves for $\mathcal{W}_R(\tau)$ and $\mathcal{W}_S(\tau)$ are on top of one another. As a result, one expects that, for some index lags, the curve for real data is the greater of the two, while, for other index lags, the curve for surrogate data is the greater of the two.

In contrast, the average separation $\mathcal{D}$ is not effected by smoothing, indicating that the process of smoothing preserves the overall structure of $W(\tau)$. $\mathcal{D}$ is another indicator of determinism or randomness, and it is used to supplement the conclusions based on ρ_+.

While the process of smoothing improves the reliability for distinguishing determinism from randomness, it does not allow us to detect determinism at significantly higher noise levels. This indicates that a positive displacement of W for real data from W for surrogate data is not an artifact of smoothing W. (i.e., the process of smoothing does not produce false positive results).

C. Effects of Noise on Detecting Determinism

Two ways of adding noise were considered, Dynamical Noise and Additive Noise. In addition, there are two cases of Dynamical Noise — intermittent and continuous. For both cases of Dynamical Noise, the extended method of Salvino and Cawley is able to detect determinism for noise levels up to 70% – 80% of the size of the system to which it is being added. This is true for data sets as small as $2^8 = 256$ data points.

For Additive Noise applied to the Rössler system, the extended method of Salvino and Cawley is able to detect determinism for noise levels up to 140% – 150% of the size of the Rössler system, roughly twice as much as for Dynamical Noise. This is to be expected, since Dynamical Noise affects the underlying dynamics of the system, unlike Additive Noise. For Additive Noise applied to the Lorenz system, the amount of noise that can be added is actually less than for noise combined with the dynamics. It is unclear why this is the case.

For chaotic noise, it is generally the case that the trajectory is ejected from the basin of attraction before the evidence of deterministic chaos disappears. When the trajectory is not ejected from the basin of attraction, noise levels much greater than 100% still produce evidence of determinism when the time series is embedded

in sufficiently high dimensions. Therefore, the extended method of Salvino and Cawley can distinguish between determinism and randomness in the noise source itself.

The study of chaotic noise also led to the conclusion that the embedding dimension used in the method of Salvino and Cawley must be sufficiently high in order to detect determinism. Using dimensions higher than the true dimension of the system does not produce false results. In practice, however, the calculation of W for small data sets with high dimensions leads to data starvation of the resulting embedding.

In conclusion, the extended method of Salvino and Cawley can distinguish between determinism and randomness for noise levels on the order of 70% – 80% of the size of the system for data sets as small as 256 data points.

D. Topics for Future Research

Other studies [7, 61] on detecting determinism in time series dealt with noise by employing noise reduction techniques, such as the technique devised by Sauer [110]. Future work could compare the effectiveness of the extended method of Salvino and Cawley to the unmodified method in conjunction with noise-reduction techniques.

The quality of an embedding depends on the embedding dimension, as well as the index lag. Two common methods for determining the embedding dimension are the Grassberger and Procaccia correlation dimension [111] and a method devised by Kennel and Abarbanel using false nearest neighbors [112]. Determination of the correct index lag is almost exclusively done using the autocorrelation function.

In this work, it was seen that embedding in too small of a dimension can cause a deterministic system to appear to be random. When the dimension is high enough a deterministic system will show evidence of such. This indicates that this method could be used to estimate the correct embedding dimension. In fact, it may be closely related to the method of Kennel and Abarbanel using false nearest

neighbors.

There are two properties of the curves of $W(\tau)$ that might be used to estimate the index lag that yields the best embedding. First, the curves of $W(\tau)$ tend to oscillate. Second the curves of $W(\tau)$ begin near one and decay to a constant value. The period of oscillation and/or the decay time of the $W(\tau)$ curve may allow for estimates of the proper index lag and/or the time delay window. The index lag τ is the number of time series points between successive components of the embedding vector. The time delay window is the number of time series points between the first and last components of this vector. It has been argued that the time delay window is the proper delay time to use when creating an embedding; however, methods for determining this quantity are lacking.

The extended method of Salvino and Cawley was studied for two examples of three-dimensional differential equations, and one example of a two-dimensional iterated map. Future work could explore a larger collection of dynamical systems. These should include higher dimensional differential equations, higher dimensional maps, and partial differential equations, in order to ensure that the method is robust with respect to the type of dynamical system.

Finally, the ability of the extended method of Salvino and Cawley to distinguish between determinism and randomness should be studied as a function of the lyapunov exponent. The lyapunov exponent is a measure of the rate of separation of nearby trajectories. It would be interesting to see if the magnitude of the lyapunov exponent has an effect on the ability to detect determinism. It would also be interesting to see if the number of positive lyapunov exponents has an effect on the ability to detect determinism.

In summary, an improved method to distinguish deterministic chaos from randomness in time series was presented. This method was shown to be able to detect determinism in small noisy data sets. Compared to competing methods, it is more sensitive and more straight forward to implement.

[1] K. A. Robbins, Math. Proc. Cambridge Philos. Soc. **82**, 209 (1977).

[2] J. M. Wersinger, J. M. Finn, and E. Ott, Phys. Rev. Lett. **44**, 453 (1980).

[3] B. A. Huberman, and J. P. Crutchfield, Phys. Rev. Lett. **43**, 1743 (1979).

[4] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).

[5] J. B. McLaughlin, and P. C. Martin, Phys. Rev. A **12**, 186 (1975).

[6] D. Ruelle, and F. Takens, Commun. Math. Phys. **20**, 167 (1971).

[7] L. W. Salvino, and R. Cawley, Phys. Rev. Lett. **73**, 1091 (1994).

[8] C. Juang, G. Y. Lee, and J. Juang, Jpn. J. Appl. Phys. Pt. 1 **38**, 6222 (1999).

[9] E. J. Mozdy, and C. R. Pollock, Appl. Phys. Lett. **77**, 1771 (2000).

[10] Y. F. Huang, H. Y. Ueng, C. R. Chung, L. W. Tu, T. S. Lim, W. C. Chen, and J. L. Chern, IEEE J. Quantum Elect. **35**, 757 (1999).

[11] J. Timmer, U. Schwarz, H. U. Voss, I. Wardinski, T. Belloni, G. Hasinger, M. van der Klis, and J. Kurths, Phys. Rev. E **61**, 1342 (2000).

[12] Q. Zhang, Sol. Phys. **178**, 423 (1998).

[13] M. N. Islam, and B. Sivakumar, Adv. Water Resour. **25**, 179 (2002).

[14] K. K. Phoon, M. N. Islam, C. Y. Liaw, and S. Y. Liong, J. Hydrol. Eng. **7**, 116 (2002).

[15] A. W. Jayawardena, W. K. Li, and P. Xu, J. Hydrol. **258**, 40 (2002).

[16] A. Porporato, and L. Ridolfi, J. Hydrol. **248**, 109 (2001).

[17] B. Sivakumar, R. Berndtsson, and M. Persson, Hydrolog. Sci. J. **46**, 377 (2001).

[18] B. Sivakumar, R. Berndtsson, J. Olsson, K. Jinno, and A. Kawamura, Hydrol. Earth Syst. Sci. **4**, 407 (2000).

[19] B. Sivakumar, J. Hydrol. **227**, 1 (2000).

[20] Q. Liu, S. Islam, I. Rodriguez-Iturbe, and Y. Le, Adv. Water Resour. **21**, 463 (1998).

[21] A. Kawamura, A. I. McKerchar, R. H. Spigel, and K. Jinno, J. Hydrol. **204**, 168

(1998).

[22] A. A. Tsonis, Int. J. Bifurcation Chaos **11**, 881 (2001).

[23] G. Sugihara, M. Casdagli, E. Habjan, D. Hess, P. Dixon, and G. Holland, Proc. Nat. Acad. Sci. USA **96**, 14210 (1999).

[24] V. Perez-Munuzuri, J. Appl. Meteorol. **37**, 1430 (1998).

[25] G. P. Pavlos, D. Kugiumtzis, M. A. Athanasiu, N. Hatzigeorgiu, D. Diamantidis, and E. T. Sarris, Nonlinear Process. Geophys. **6**, 79 (1999).

[26] J. L. Chen, S. Islam, and P. Biswas, Atmos. Environ. **34**, 837 (2000).

[27] B. Minasny, and A. B. McBratney, Geoderma **103**, 161 (2001).

[28] W. Marzocchi, G. Vilardo, D. P. Hill, G. P. Ricciardi, and C. Ricco, Bull. Seismol. Soc. Amer. **91**, 191 (2001).

[29] V. Cuomo, V. Lapenna, M. Macchiato, C. Serio, and L. Telesca, Tectonophysics **287**, 279 (1998).

[30] Y. Xiao, D. Z. Cheng, and S. Y. Tang, Chaos Soliton Fract. **14**, 1403 (2002).

[31] J. F. McLaughlin, J. J. Hellmann, C. L. Boggs, and P. R. Ehrlich, Oecologia **132**, 538 (2002).

[32] A. B. Medvinsky, I. A. Tikhonova, S. V. Petrovskii, H. Malchow, and E. Venturino, Zh. Obshch. Biol. **63**, 149 (2002).

[33] M. A. Harrison, L. C. Lai, and R. D. Holt, J Theor. Biol. **213**, 53 (2001).

[34] S. Y. Tang, and L. S. Chen, Chaos Soliton Fract. **13**, 875 (2002).

[35] Z. I. Dimitrova, and N. K. Vitanov, J. Phys-A-Math. Gen. **34**, 7459 (2001).

[36] S. Rinaldi, M. Candaten, and R. Casagrandi, Ecol. Lett. **4**, 610 (2001).

[37] O. C. Lingjaerde, N. C. Stenseth, A. B. Kristoffersen, R. H. Smith, S. J. Moe, J. M. Read, S. Daniels, and K. Simkiss, Ecology **82**, 2645 (2001).

[38] O. N. Bjornstad, and B. T. Grenfell, Science **293**, 638 (2001).

[39] V. Kaitala, J. Ylikarjula, and M. Heino, Ecol. Model. **135**, 127 (2000).

[40] V. Rai, and W. M. Schaffer, Chaos Soliton Fract. **12**, 197 (2001).

[41] R. K. Upadhyay, and V. Rai, Chaos Soliton Fract. **12**, 205 (2001).

[42] M. Pascual, and S. P. Ellner, Ecology **81**, 2767 (2000).

[43] P. A. Dixon, M. J. Milicich, and G. Sugihara, Science **283**, 1528 (1999).

[44] J. L. Gonzalez-Andujar, and J. N. Perry, Ecol. Res. **12**, 301 (1997).

[45] W. W. Murdoch, and C. J. Briggs, Trend. Ecol. Evolut. **17**, 399 (2002).

[46] M. Kamo, and A. Sasaki, Physica D **165**, 228 (2002).

[47] O. N. Bjornstad, B. F. Finkenstadt, and B. T. Grenfell, Ecol. Monogr. **72**, 169 (2002).

[48] P. P. Kanjilal, and J. Bhattacharya, J. Theor. Biol. **197**, 163 (1999).

[49] S. P. Ellner, B. A. Bailey, G. V. Bobashev, A. R. Gallant, B. T. Grenfell, and D. W. Nychka, Amer. Naturalist **151**, 425 (1998).

[50] A. Sumi, N. Ohtomo, and Y. Tanaka, Jpn. J. Appl. Phys. Pt. 1 **36**, 7460 (1997).

[51] J. Jeong, J. C. Gore, and B. S. Peterson, IEEE Trans. Biomed. Eng. **49**, 1374 (2002).

[52] J. Jeong, J. C. Gore, and B. S. Peterson, Biol. Cybern. **86**, 335 (2002).

[53] M. Breakspear, and J. R. Terry, Clin. Neurophysiol. **113**, 735 (2002).

[54] M. J. A. M. van Putten, and C. J. Stam, IEEE Eng. Med. Biol. Mag. **20**, 72 (2001).

[55] P. Faure, and H. Korn, C. R. Acad. Sci. Ser. III-Vie **324**, 773 (2001).

[56] A. Kelly, A. Heathcote, R. Heath, and M. Longstaff, Q. J. Exp. Psychol. A **54**, 805 (2001).

[57] L. Pezard, R. Jech, and E. Ruzicka, Clin. Neurophysiol. **112**, 38 (2001).

[58] M. E. Brandt, A. Ademoglu, and W. S. Pritchard, Int. J. Bifurcation Chaos **10**, 123 (2000).

[59] D. Popivanov, Theor. Biosci. **118**, 231 (1999).

[60] B. Jelles, R. L. M. Strijers, C. Hooijer, C. Jonker, C. J. Stam, and E. J. Jonkman, Acta Neurol. Scand. **100**, 360 (1999).

[61] J. Jeong, M. S. Kim, and S. Y. Kim, Phys. Rev. E **60**, 831 (1999).

[62] M. Le van Quyen, J. Martinerie, C. Adam, and F. J. Varela, Physica D **127**, 250 (1999).

[63] S. Micheloyannis, C. J. Stam, E. Fountoulakis, M. Bourkas, S. Arvanitis, and E. Papanikolaou, Neurosci. Lett. **249**, 167 (1998).

[64] L. I. Aftanas, N. V. Lotova, V. I. Koshkarov, S. A. Popov, and V. P. Makhnev, Brain Topogr. **10**, 155 (1997).

[65] T. Tateno, A. Kawana, and Y. Jimbo, Phys. Rev. **65**, 051924 (2002).

[66] M. Barbi, S. Chillemi, A. Di Garbo, G. Molnar, and A. Szucs, Acta Biol. Hung. **51**, 237 (2000).

[67] C. J. Stam, J. P. M. Pijn, P. Suffczynski, and F. H. L. da Silva, Clin. Neurophysiol. **110**, 1801 (1999).

[68] M. Di Mascio, G. Di Giovanni, V. Di Matteo, and E. Esposito, Neuroscience **92**, 237 (1999).

[69] M. Y. Alvarenga, K. Sameshima, L. A. Baccala, and H. M. Yang, Math. Biosci. **157**, 287 (1999).

[70] Y. F. Gong, J. X. Xu, W. Ren, S. J. Hu, and F. Z. Wang, Biol. Cybern. **78**, 159 (1998).

[71] S. Chillemi, M. Barbi, and A. Di Garbo, Nonlinear Anal.-Theor. Meth. App. **30**, 1625 (1997).

[72] S. Roberts, R. Eykholt, and M. H. Thaut, Phys. Rev. E, **62**, 2597 (2000).

[73] Y. C. Lai, I. Osorio, M. A. F. Harrison, and M. G. Frei, Phys. Rev. E **65**, 031921 (2002).

[74] I. Jekova, J. Dushanova, and D. Popivanov, Physiol. Meas. **23**, 337 (2002).

[75] W. J. Pei, Z. Y. He, L. X. Yang, S. S. Hull, and J. Y. Cheung, Chin. Sci. Bull. **46**, 1568 (2001).

[76] L. Cimponeriu, and A. Bezerianos, Lect. Note. Comput. Sci. **1933**, 88 (2000).

[77] S. Garde, M. G. Regalado. V. L. Schechtman, and M. C. K. Khoo, Am. J. Physiol. Heart C **280**, H2920 (2001).

[78] J. O. Fortrat, D. Sigaudo, R. L. Hughson, A. Maillet, Y. Yamamoto, and C. Gharib, Auton. Neurosci.-Basic Clin. **86**, 192 (2001).

[79] H. S. Kim, S. D. Kim, C. S. Kim, and M. K. Yum, Anaesthesia **55**, 847 (2000).

[80] M. E. D. Gomes, A. V. P. Souza, H. N. Guimaraes, and L. A. Aguirre, Chaos **10**, 398 (2000).

[81] K. M. Stein, J. Waldeen, N. Lippman, and B. B. Lerman, Am. J. Physiol. Heart C **277**, H452 (1999).

[82] R. B. Govindan, K. Narayanan, and M. S. Gopinathan, Chaos **8**, 495 (1998).

[83] M. Barbi, S. Chillemi, A. Di Garbo, R. Balocchi, C. Carpeggiani, M. Emdin, C. Michelassi, and E. Santarcangelo, Chaos Soliton Fract. **9**, 507 (1998).

[84] P. V. Bayly, B. H. KenKnight, J. M. Rogers, E. E. Johnson, R. E. Ideker, and W. M. Smith, Chaos **8**, 103 (1998).

[85] S. Chillemi, M. Barbi, A. Di Garbo, R. Balocchi, C. Michelassi, C. Carpeggiani, and M. Emdin, Methods Inform. Med. **36**, 278 (1997).

[86] P. Trillenberg, C. Gross, and M. Shelhamer, J. Appl. Physiol. **91**, 1750 (2001).

[87] M. Shelhamer, J. Neurosci. Meth. **83**, 45 (1998).

[88] L. Y. Cao, B. G. Kim, J. Kurths, and S. Kim, Int. J. Bifurcation Chaos **8**, 179 (1998).

[89] L. Y. Cao, and A. Mees, Int. J. Bifurcation Chaos **10**, 2767 (2000).

[90] T. Miyano, A. Nagami, I. Tokuda, and K. Aihara, Int. J. Bifurcation Chaos **10**, 1973 (2000).

[91] T. Miyano, S. Kimoto, H. Shibuta, K. Nakashima, Y. Ikenaga, and K. Aihara, Physica D **135**, 305 (2000).

[92] T. Miyano, A. Shintani, and T. Kanda, J. Appl. Phys. **83**, 4180 (1998).

[93] S. Murashige, T. Yamada, and K. Aihara, Phil. Trans. Roy. Soc. London A **358**, 1793 (2000).

[94] B. L. Smith, B. M. Williams, and R. K. Oswald, Transport. Res. C Emer. **10**, 303 (2002).

[95] V. K. Jain, A. K. Srivastava, A. Das, and V. Rai, Chaos Soliton Fract. **12**, 859 (2001).

[96] Y. Agnon, A. Golan, and M. Shearer, Econ. Lett. **65**, 293 (1999).

[97] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. **45**, 712 (1980)

[98] M. Cencini, M. Falcioni, E. Olbrich, H. Kantz and A. Valpiani, Phys. Rev. E **62**, 427 (2000).

[99] M. B. Kennel, and S. Isabelle, Phys. Rev. A **46**, 3111 (1992).

[100] G. Sugihara, and R. M. May, Nature (London) **344**, 734 (1990)

[101] S. Basu, and E. Foufoula-Georgiou, Phys. Lett. A **301**, 413 (2002).

[102] J. Theiler, Phys. Lett. A **155**, 480 (1991).

[103] D. T. Kaplan, and L. Glass, Phys. Rev. Lett. **68**, 427 (1992); D. T. Kaplan, and L. Glass, Physica D (Amsterdam) **64**, 431 (1993).

[104] R. Wayland, D. Bromley, D. Pickett, and A. Passamante, Phys. Rev. Lett. **70**, 580 (1993).

[105] W. L. Ditto, S. N. Rauseo, R. Cawley, C. Grebogi, G. H. Hsu, E. Kostelich, E. Ott, H. T. Savage, R. Segnan, M. L. Spano, and J. A. Yorke, Phys. Rev. Lett. **63**, 923 (1989).

[106] F. Takens, in *Dynamical Systems and Turbulence*, edited by D. A. Rand, and L. S. Young, Lecture Notes in Mathematics, **898** (Springer, Berlin, 1981), p. 336.

[107] O. E. Rössler, Phys. Lett. A **57**, 397 (1976).

[108] M. Hénon, Commun. Math. Phys. **50**, 69 (1976)

[109] M. Matsumoto, and T. Nishimura, Acm. T. Model. Comp. Sim. **8**, 3 (1998)

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

[111] P. Grassberger, and I. Procaccia, Physica D **35**, 189 (1983).

[112] M. B. Kennel, and H. D. I. Abarbanel, Phys. Rev. E **66**, 026209 (2002).

VI. APPENDICES

A. Intermittent Dynamical White Noise

1. *Figures for Intermittent Dynamical White Noise Applied to the Rössler System*

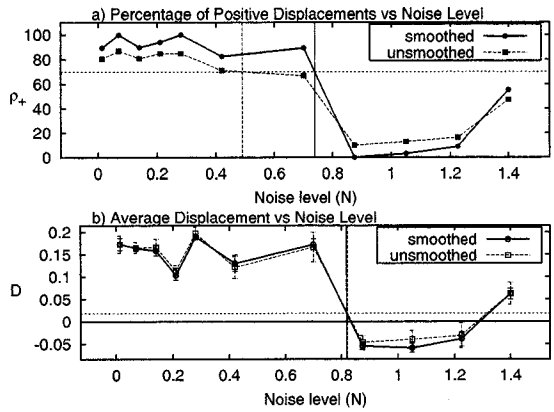

FIG. 115: Intermittent Dynamical White Noise applied to the Rössler system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

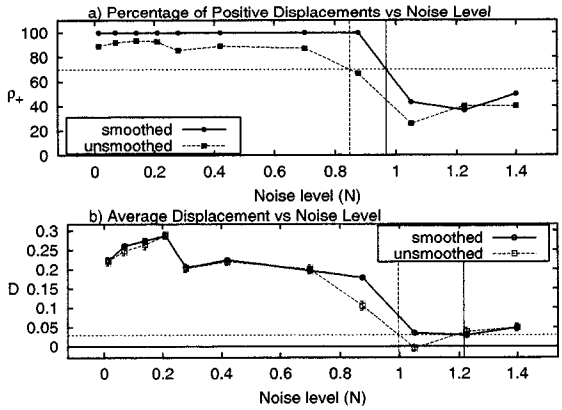

FIG. 116: Intermittent Dynamical White Noise applied to the Rössler system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

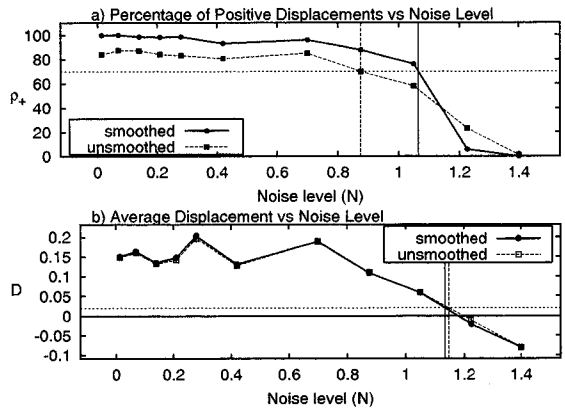

FIG. 117: Intermittent Dynamical White Noise applied to the Rössler system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

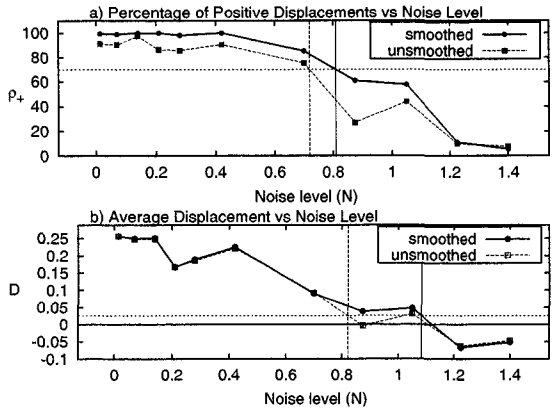

FIG. 118: Intermittent Dynamical White Noise applied to the Rössler system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

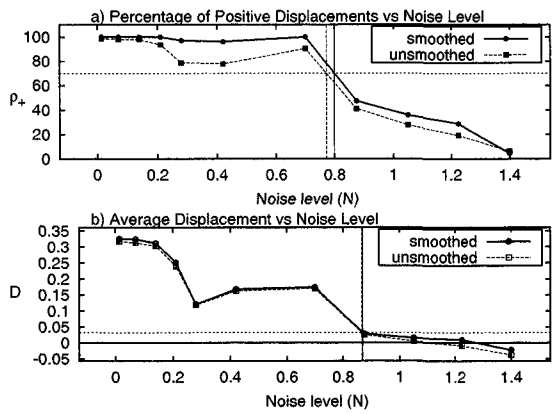

FIG. 119: Intermittent Dynamical White Noise applied to the Rössler system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

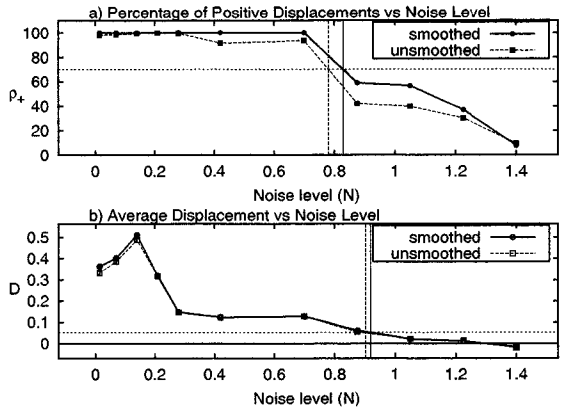

FIG. 120: Intermittent Dynamical White Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

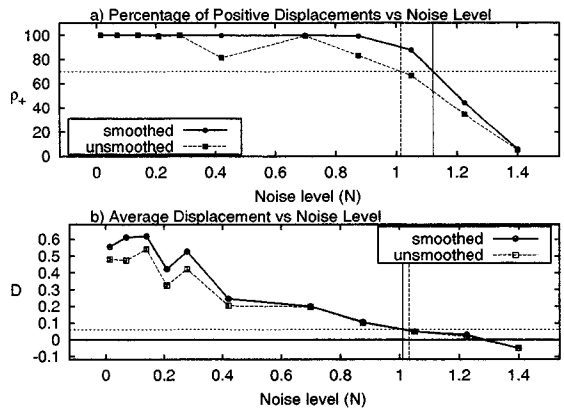

FIG. 121: Intermittent Dynamical White Noise applied to the Rössler system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

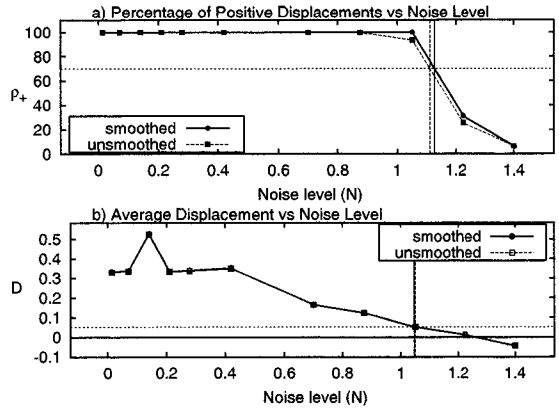

FIG. 122: Intermittent Dynamical White Noise applied to the Rössler system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

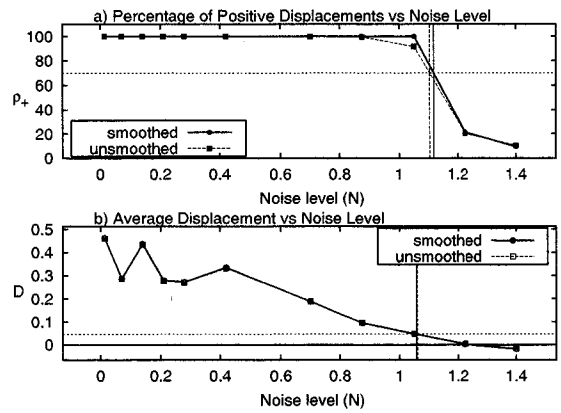

FIG. 123: Intermittent Dynamical White Noise applied to the Rössler system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

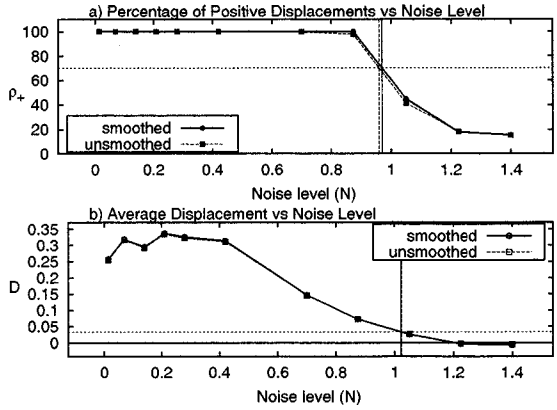

FIG. 124: Intermittent Dynamical White Noise applied to the Rössler system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

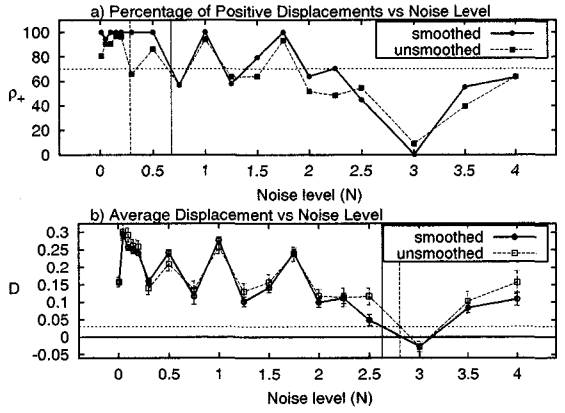

FIG. 125: Intermittent Dynamical White Noise applied to the Lorenz system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

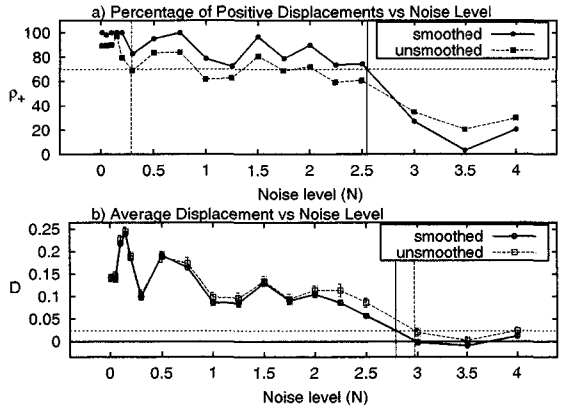

FIG. 126: Intermittent Dynamical White Noise applied to the Lorenz system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

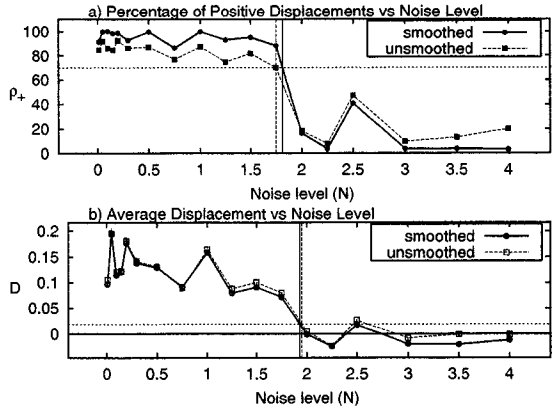

FIG. 127: Intermittent Dynamical White Noise applied to the Lorenz system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

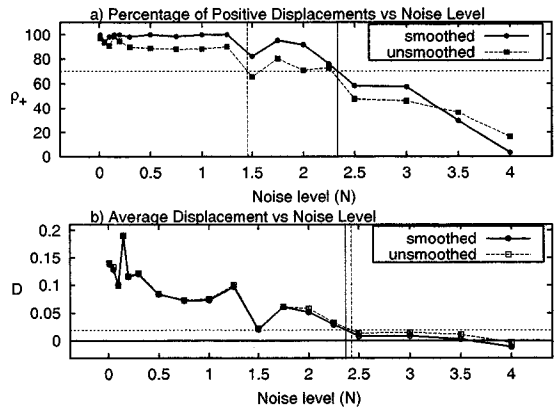

FIG. 128: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

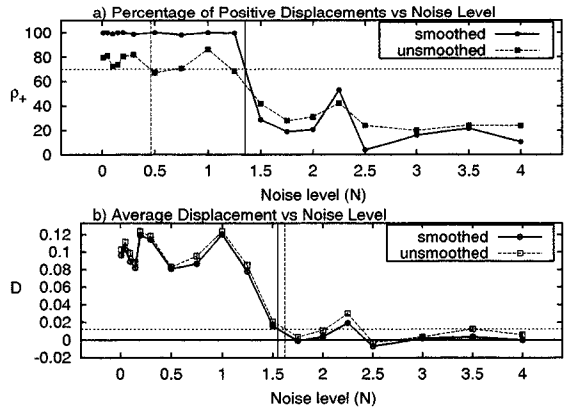

FIG. 129: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

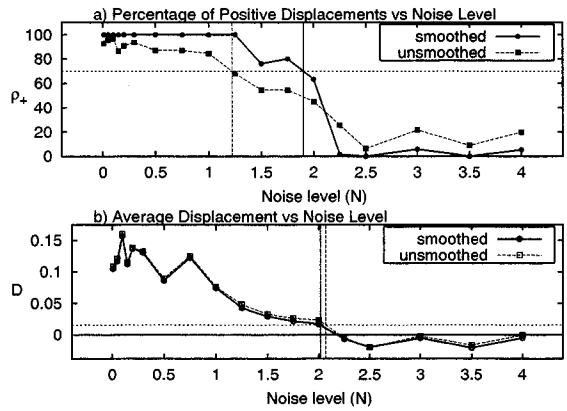

FIG. 130: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

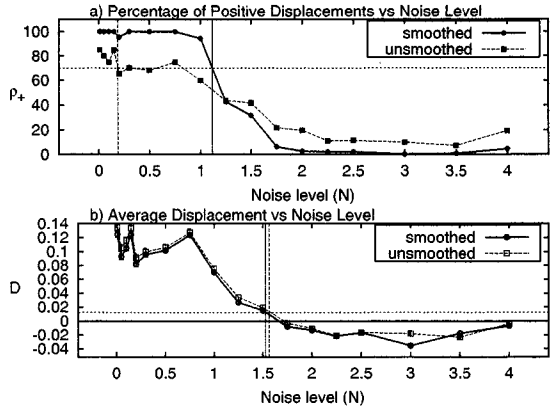

FIG. 131: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

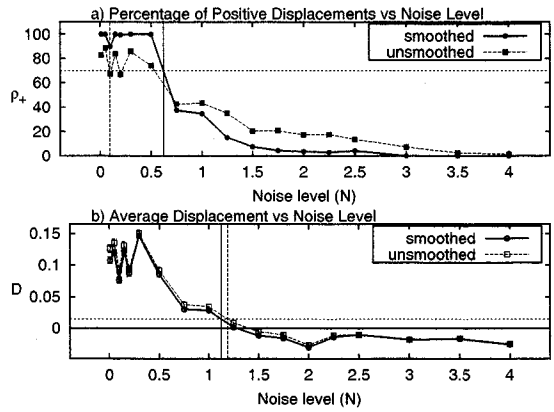

FIG. 132: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

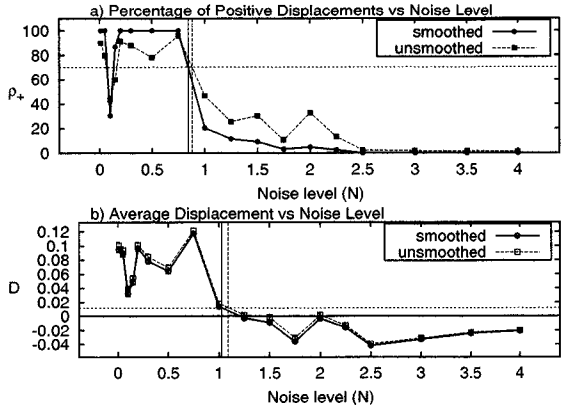

FIG. 133: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

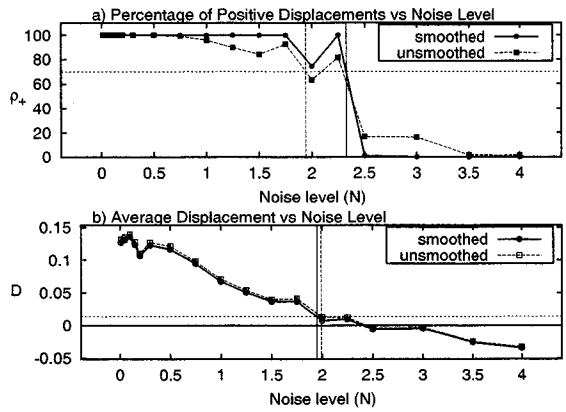

FIG. 134: Intermittent Dynamical White Noise applied to the Lorenz system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

B. Additive White Noise

1. *Figures for Additive White Noise Applied to the Rössler System*

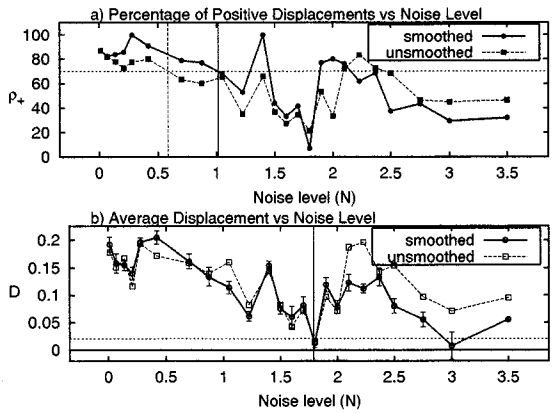

FIG. 135: Additive White Noise applied to the Rössler system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

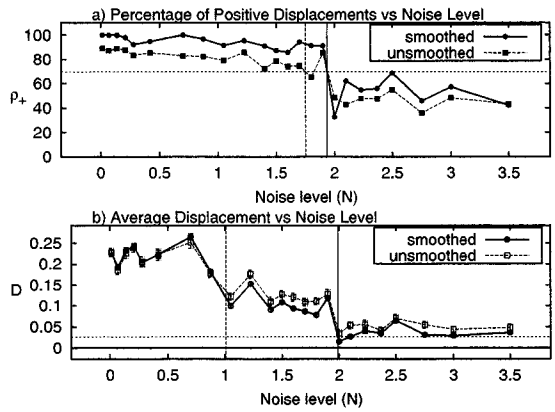

FIG. 136: Additive White Noise applied to the Rössler system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

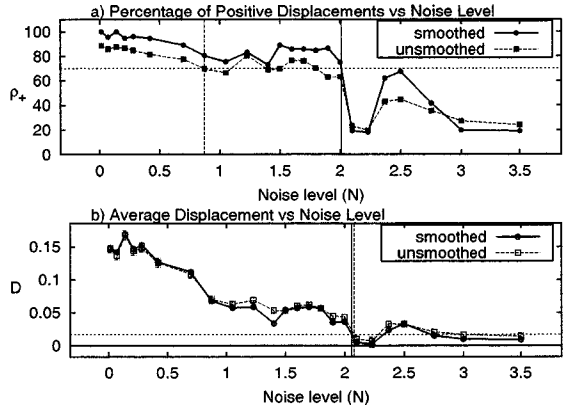

FIG. 137: Additive White Noise applied to the Rössler system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

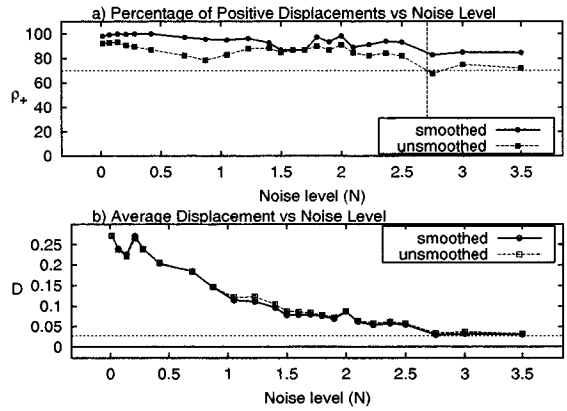

FIG. 138: Additive White Noise applied to the Rössler system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

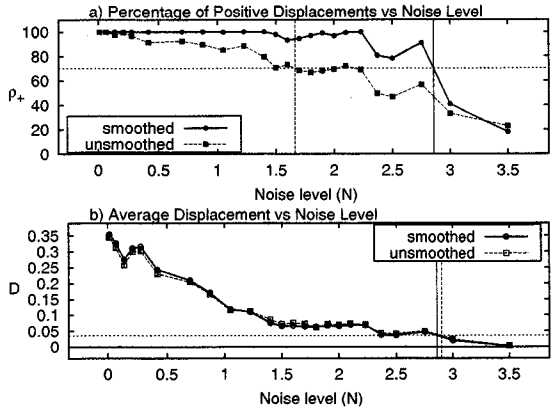

FIG. 139: Additive White Noise applied to the Rössler system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

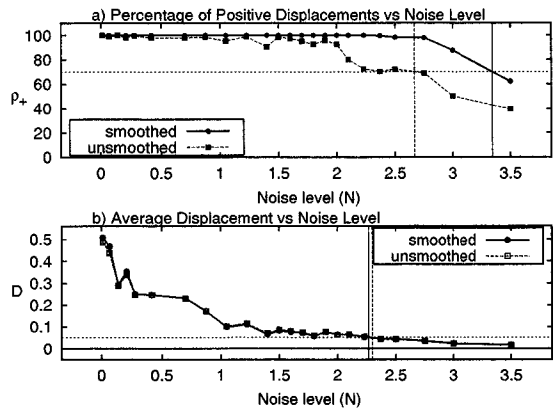

FIG. 140: Additive White Noise applied to the Rössler system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

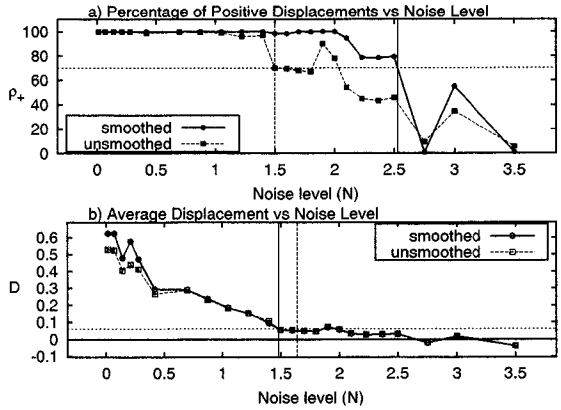

FIG. 141: Additive White Noise applied to the Rössler system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

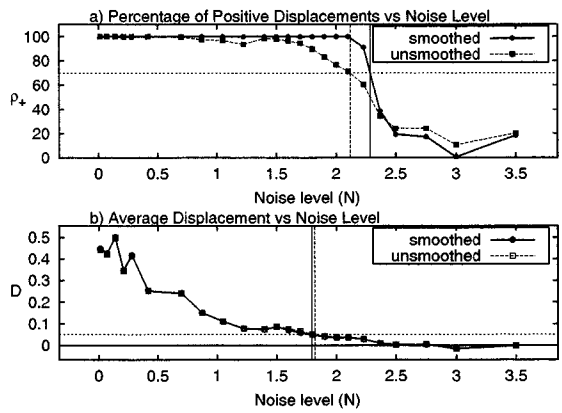

FIG. 142: Additive White Noise applied to the Rössler system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

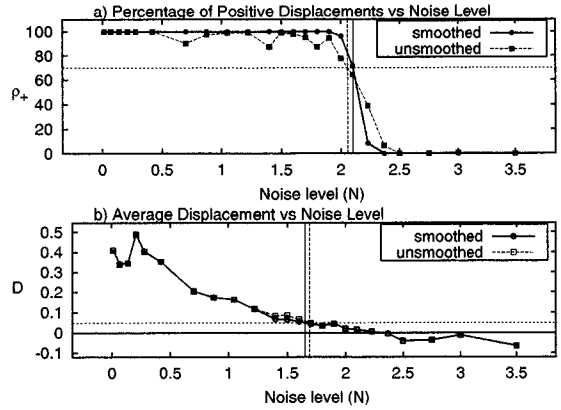

FIG. 143: Additive White Noise applied to the Rössler system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

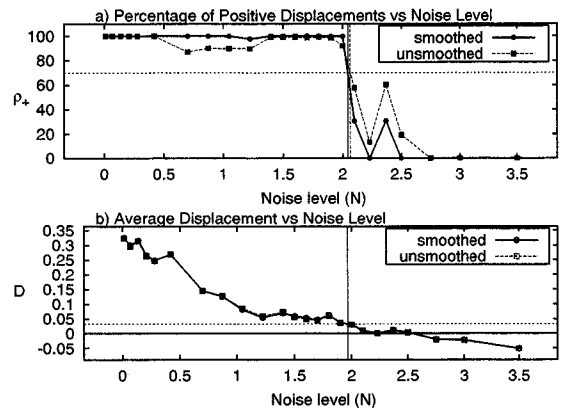

FIG. 144: Additive White Noise applied to the Rössler system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

2. *Figures for Additive White Noise Applied to the Lorenz System*

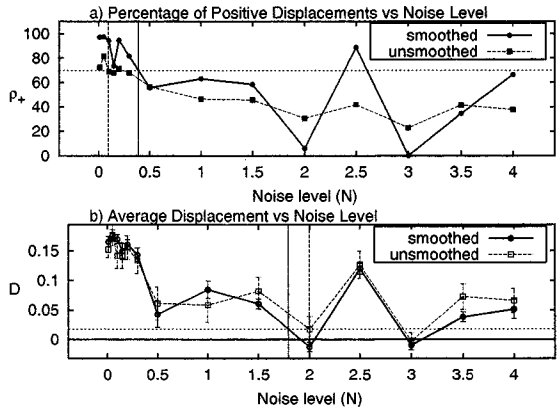

FIG. 145: Additive White Noise applied to the Lorenz system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

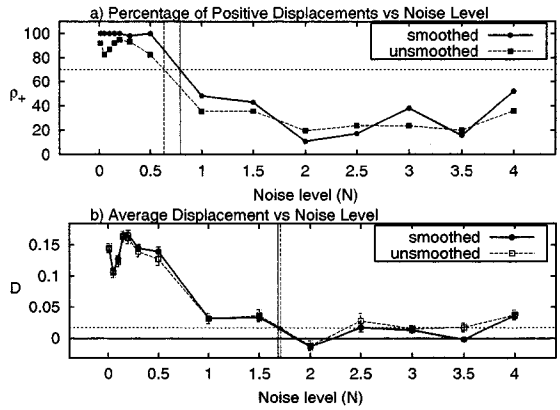

FIG. 146: Additive White Noise applied to the Lorenz system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

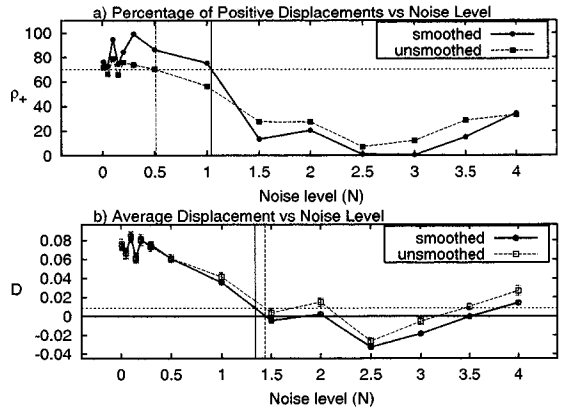

FIG. 147: Additive White Noise applied to the Lorenz system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

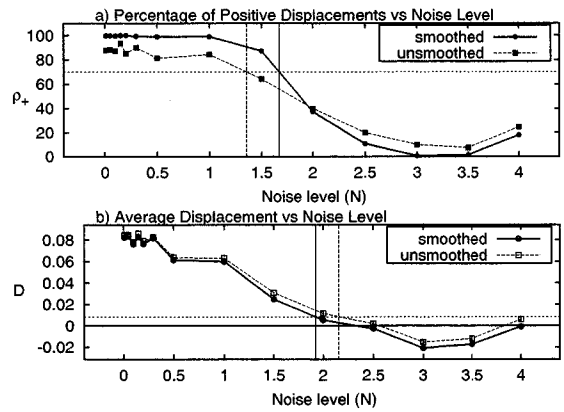

FIG. 148: Additive White Noise applied to the Lorenz system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

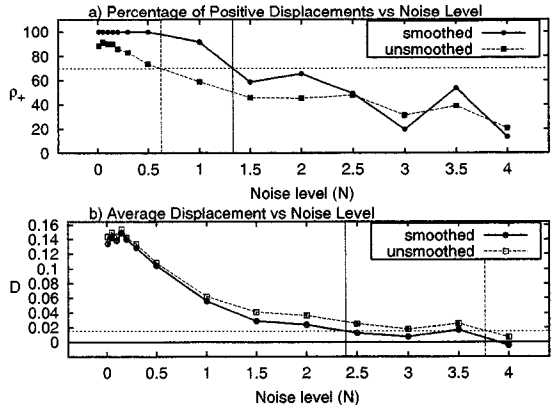

FIG. 149: Additive White Noise applied to the Lorenz system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

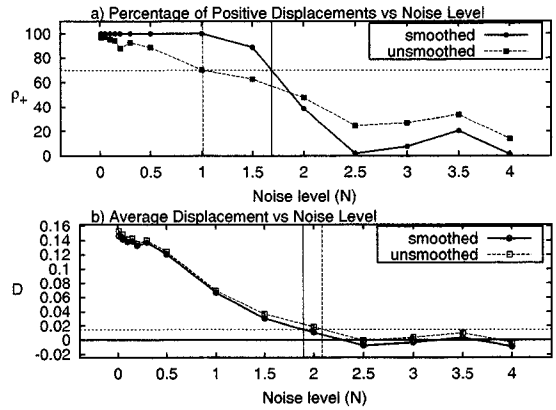

FIG. 150: Additive White Noise applied to the Lorenz system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

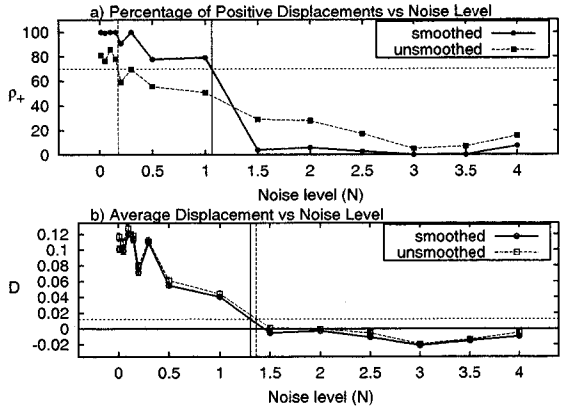

FIG. 151: Additive White Noise applied to the Lorenz system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

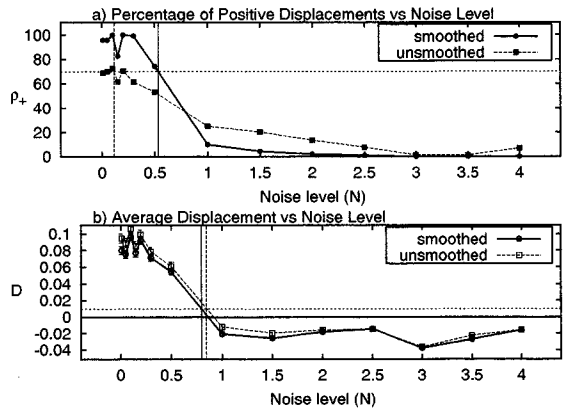

FIG. 152: Additive White Noise applied to the Lorenz system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

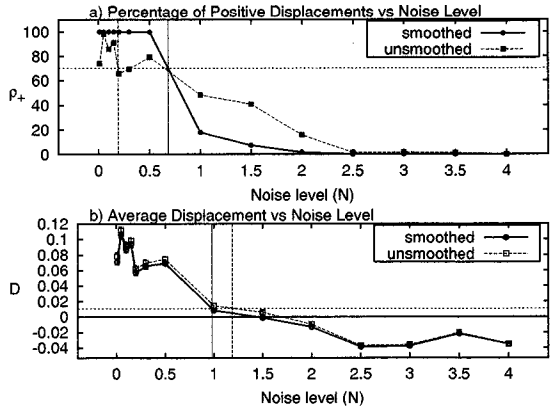

FIG. 153: Additive White Noise applied to the Lorenz system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

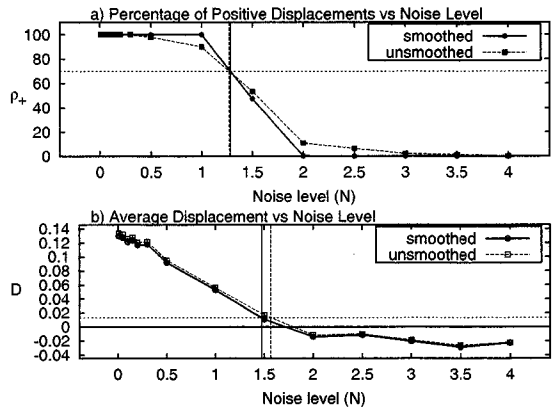

FIG. 154: Additive White Noise applied to the Lorenz system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

3. Figures for Additive White Noise Applied to the Hénon System

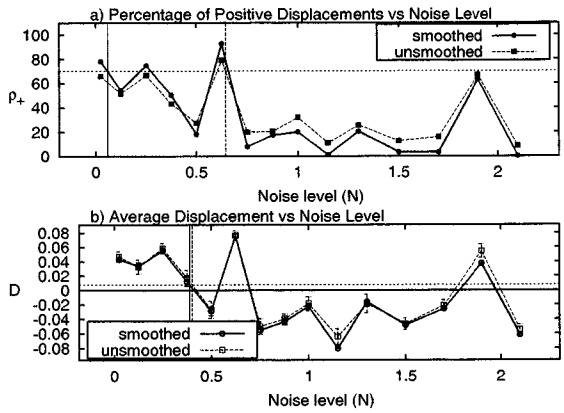

FIG. 155: Additive White Noise applied to the Hénon system for 2^7 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

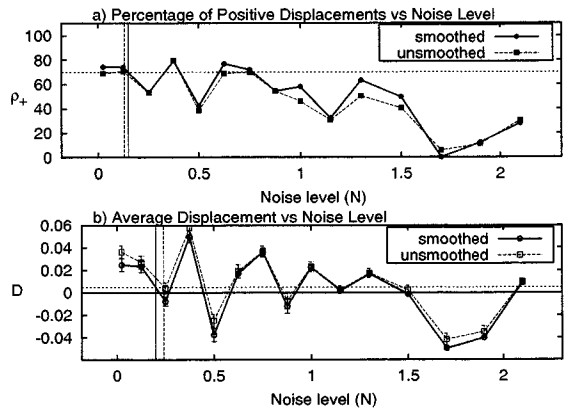

FIG. 156: Additive White Noise applied to the Hénon system for 2^8 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

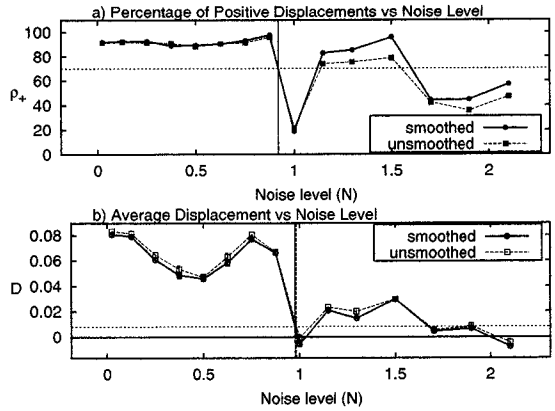

FIG. 157: Additive White Noise applied to the Hénon system for 2^9 data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

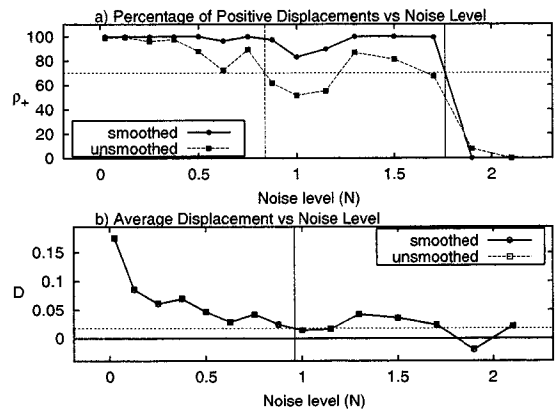

FIG. 158: Additive White Noise applied to the Hénon system for 2^{10} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

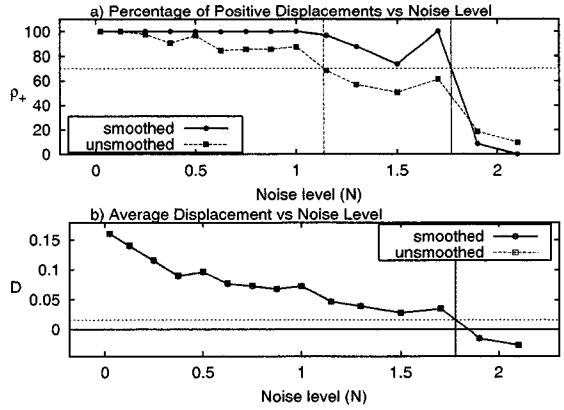

FIG. 159: Additive White Noise applied to the Hénon system for 2^{11} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

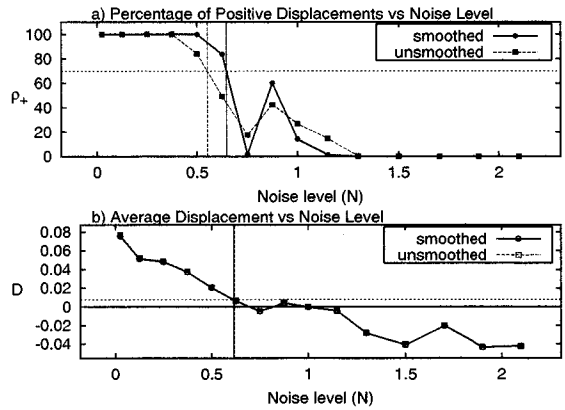

FIG. 160: Additive White Noise applied to the Hénon system for 2^{12} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

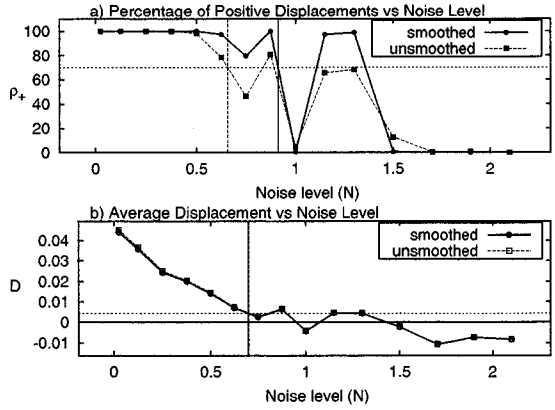

FIG. 161: Additive White Noise applied to the Hénon system for 2^{13} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

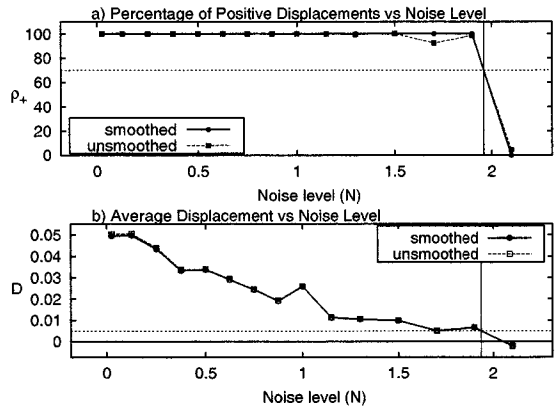

FIG. 162: Additive White Noise applied to the Hénon system for 2^{14} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

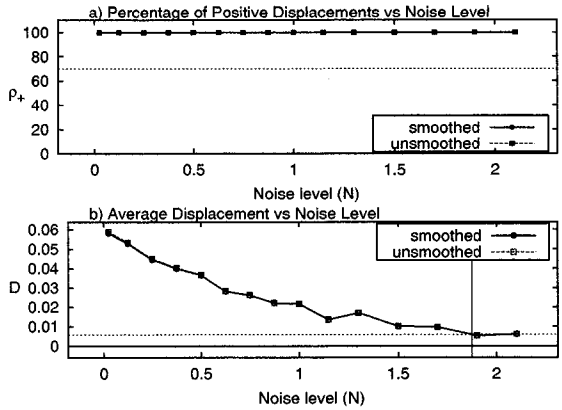

FIG. 163: Additive White Noise applied to the Hénon system for 2^{15} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.

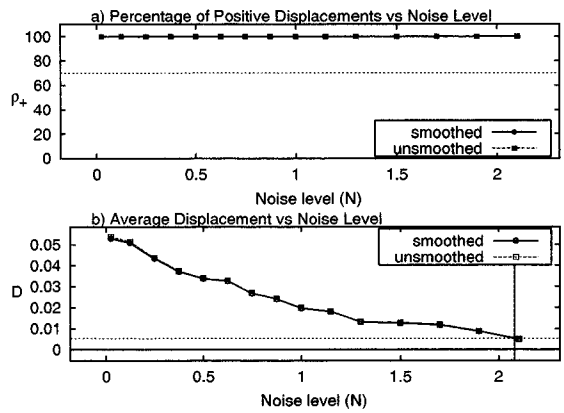

FIG. 164: Additive White Noise applied to the Hénon system for 2^{16} data points. Plot (a) shows ρ_+ versus N. Plot (b) shows $\mathcal{D}$ versus N. The vertical lines in these plots are the values of the maximum noise cutoffs N_{pd} and N_{ave}, respectively.