To appear in *Physica D*

Reconstruction expansion as a geometry-based framework for choosing proper delay times

Michael T. Rosenstein, James J. Collins, and Carlo J. De Luca

NeuroMuscular Research Center and Department of Biomedical Engineering,
Boston University, 44 Cummington Street, Boston, MA 02215, USA

19 November 1993

Running Title:	Reconstruction expansion and delay times
Key Words:	chaos, method of delays, time series analysis
PACS codes:	05.45.+b, 02.50.+s, 02.60.+y

Corresponding Author's Present Address: Michael T. Rosenstein
Dynamical Research
15 Pecunit Street
Canton, MA 02021-1219
U.S.A.

Telephone: (617) 821-6350
Internet: MTR1a@aol.com

Abstract

The quality of attractor reconstruction using the method of delays is known to be sensitive to the delay parameter, τ. Here we develop a new, computationally efficient approach to choosing τ that quantifies reconstruction expansion from the identity line of the embedding space. We show that reconstruction expansion is related to the concept of reconstruction signal strength and that increased expansion corresponds to diminished effects of measurement error. Thus, reconstruction expansion represents a simple, geometrical framework for choosing τ. Furthermore, we describe the role of dynamical error in attractor expansion and argue that algorithms for determining τ should be considered as attempts at estimating an upper bound to the optimal delay.

1. Introduction

Attractor reconstruction is usually the first step in the analysis of dynamical systems. Typically, an experimenter obtains a scalar time series from one observable of a multi-dimensional system; state-space reconstruction is then needed for the indirect measurement of the system's properties, e.g., dimension. Several techniques for attractor reconstruction are currently employed, such as derivative coordinates [1, 2] and principal components (or singular value decomposition) [3]. The method of delays [1, 2] is the most widespread approach because it is the most straightforward and the noise level is constant for each delay component [4]. However, the drawback with the method of delays is that the quality of the reconstruction depends upon the delay parameter, τ, and presently, there is no commonly accepted procedure for choosing τ.

The method of delays reconstructs the attractor dynamics by using *delay coordinates* to form multiple state-space vectors, $\mathbf{X}_i$. That is, the reconstructed trajectory, $\mathbf{X}$, is given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \ldots & \mathbf{X}_M \end{bmatrix}^{\mathrm{T}}. \tag{1}$$

For an *N*-point time series, $\{x_1, x_2, \ldots, x_N\}$, the reconstructed state of the system at each discrete time *i* is

$$\mathbf{X}_i = \begin{bmatrix} x_i & x_{i+J} & \ldots & x_{i+(m-1)J} \end{bmatrix}, \tag{2}$$

where *J* is the *reconstruction delay* or *lag*, and *m* is the *embedding dimension*. If the time series represents a continuous flow with samples taken every Δt seconds, then the delay parameter may be expressed as

$$\tau = J \cdot \Delta t\,. \tag{3}$$

According to Takens' theorem [2], a faithful reconstruction, i.e., an embedding, is guaranteed as long as the embedding dimension is greater than twice the topological dimension, *n*:

$$m > 2n. \tag{4}$$

By "faithful" reconstruction, we refer to one that preserves the system invariants, e.g., dimension, Lyapunov exponents. Unfortunately, Takens' theorem assumes the availability of an infinite amount of noise-free data. (In this ideal, one may choose any value of τ that does not lead to a degenerate reconstruction.)[1] In real experiments, i.e., experiments with noisy, finite data sets, a value of τ that is too small results in little information gain, i.e., *redundance* [4], between successive delay coordinates, and the reconstructed trajectory becomes compressed along the main diagonal, or identity line, of the embedding space. With chaotic systems and large values of τ, successive delay coordinates may become causally unrelated, and the reconstruction is no longer representative of the true dynamics. Casdagli *et al.* [4] called this phenomenon *irrelevance*.

A related difficulty with attractor reconstruction involves the choice of *m*. Normally, one has no *a priori* knowledge regarding the topological dimension, and it is unclear what values of *m* will satisfy Eq. (4).[2] This problem was addressed elsewhere [5-8], and for the purposes of this paper, we assume that a suitable value for *m* has been chosen. This is an important assumption since previous research [3, 8-10] suggests that it may be more appropriate to fix the *reconstruction window*, τ_w, rather than τ alone, where τ_w is the length of the interval spanned by the first and last delay coordinates:

$$\tau_w = \tau \cdot (m-1). \tag{5}$$

For example, Martinerie *et al.* [10] showed that the correlation integral is sensitive to τ_w, but not to τ and *m* individually. Hence, we reformulate the problem addressed in this paper to that of

---

[1] For example, a limit cycle with period *T* is not properly reconstructed whenever τ/T is rational [4].

[2] The usual solution is to increase *m* gradually until a suitable statistic, e.g., dimension, largest Lyapunov exponent, converges for two or three consecutive values of *m*.

choosing the proper τ once the embedding dimension is fixed. (In the remainder of this paper, τ and τ_w are interchanged as dictated by the particular context.)

From the above discussion one can see that choosing proper reconstruction delays is, inherently, an ill-defined problem. Nevertheless, a suitable approach to choosing the reconstruction delay is desirable, and in Section 2 we present a brief review of the existing methods. However, these methods are often inconsistent or too time-consuming for practical applications. Hence, Section 3 presents a computationally-efficient, geometry-based approach that is derived from the consideration of intrinsic reconstruction errors. As shown in Section 4, this method yields values of τ_w that lead to acceptable state-space reconstructions. In Section 5, we discuss some of the theoretical and practical implications of the proposed method. Finally, Section 6 contains a summary of our conclusions.

2. Previous methods

2.1. Autocorrelation

A number of criteria for selecting τ_w depend upon the autocorrelation function, $R_{xx}(\tau)$. The autocorrelation function provides a measure of the similarity between a signal, $x(t)$, and a delayed version of itself; thus, $R_{xx}(\tau)$ is maximized when the delay is zero. The autocorrelation function is expected to provide a reasonable measure of the transition from redundance to irrelevance (as a function of delay). Typically, τ_w is chosen as the delay where $R_{xx}(\tau)$ first drops to a certain fraction of its initial value, e.g., 1/e [9]. Similarly, τ_w may be chosen by locating the first inflection point of $R_{xx}(\tau)$ [11]. A related criterion derived from the Fourier transform of $R_{xx}(\tau)$, i.e., the power spectrum of $x(t)$, is the inverse of the band-limiting frequency [3].

The autocorrelation-based methods have the advantage of short computation times when calculated via the fast Fourier transform (FFT) algorithm. As suggested by the numerous variations, however, these methods tend to be inconsistent [9, 10, 12]. That is, a particular criterion may be superior for one dynamical system and poor for another. This is not surprising

given the ill-defined relationship between the *spatial* distribution of a reconstructed attractor and the *temporal* autocorrelation of a single time series. (In Section 5.3, we provide some insight into this relationship.) For this reason, Fraser and Swinney [12] suggested a spatial measure based on *mutual information.*

2.2. Mutual information

In contrast to the *linear* dependence measured by autocorrelation, mutual information, $I(\tau)$, supplies a measure of *general* dependence [12]. Therefore, $I(\tau)$ is expected to provide a better measure of the shift from redundance to irrelevance with nonlinear systems. Mutual information answers the following question [12]: Given the observation of $x(t)$, how accurately can one predict $x(t+\tau)$? Thus, successive delay coordinates are interpreted as relatively independent when the mutual information is small. Fraser and Swinney [12] associated the greatest independence, i.e., the lowest $I(\tau)$, with the least redundance and, therefore, the best attractor reconstruction. (This assumption is discussed further in Section 3.1.) For this reason, they selected τ as the lag that produces a local minimum of $I(\tau)$. Typically, the first local minimum is the preferred choice [12, 13]. (Note that Fraser and Swinney originally considered this method as establishing the value of τ, not τ_w. However, in light of the work by Martinerie *et al.* [10], it seems more appropriate to interpret the result as τ_w.) Liebert and Schuster [13] showed that the minima of $I(\tau)$ coincide with those of the correlation integral, $C_m(r;\tau)$. The correlation integral, which requires less computation than $I(\tau)$ [13], is defined as

$$C_m(r;\tau) = \frac{2}{M(M-1)} \sum_{i \neq j} \theta\left[r - \left\|\mathbf{X}_i - \mathbf{X}_j\right\|\right], \tag{6}$$

where M is the number of reconstructed points, $\theta[..]$ is the Heavyside function, and $\|..\|$ denotes the Euclidean norm. It follows that an algorithm for calculating correlation dimension [14] is easily adapted to estimate τ. The primary drawback of this approach is the enormous computational costs. For each value of τ, Eq. (6) requires four or five orders of magnitude more computation than the FFT. Furthermore, Martinerie *et al.* [10] showed that mutual information is also inconsistent in identifying the optimal value of τ_w.

2.3. Higher-order correlations

Albano *et al.* [15] developed a method for choosing τ_w by examining higher-order moments of $x(t)$, i.e., correlation functions up to the fourth order. They observed that several such correlation functions (though not always the same ones) exhibit coincident extrema, and the delay at the first coincident extremum, i.e., the "coincident time" (τ_c), may be a characteristic time that serves as a good candidate for τ_w. The main advantages of this approach are twofold: (1) reasonable computation times (roughly one order of magnitude longer than autocorrelation alone), and (2) consistent results. Albano *et al.* [15] found that using τ_c resulted in estimates of the correlation dimension which were consistently within 10% of those calculated with larger and more precise data sets. (However, they also found that τ_c did not always lead to the best estimate of dimension.) Though the empirical results are compelling, the primary drawback of this approach is its weak theoretical basis. Nevertheless, "even in the absence of a rigorous theoretical justification for its choice, [τ_c] can serve as a standard for use in the analysis of experimental data" [15, p. 96].

2.4. Fill factor

Buzug and Pfister [16, 17] proposed the "fill factor" as a spatial measure for determining the reconstruction delay. The fill factor quantifies an attractor's utilization of embedding space as a function of τ. This is accomplished by computing the average volume of numerous *m*-dimensional parallelepipeds, where *m*+1 randomly chosen points are used to define the vertices of each parallelepiped. The optimal τ is then selected as the one that maximizes the fill factor, i.e., the one that leads to the most voluminous reconstruction. (This method also considers τ to be fixed for different embedding dimensions.) One difficulty with this criterion is that it cannot account for the "overfolding" of the attractor once the reconstruction has expanded (with increasing τ) from the main diagonal [16]. A more serious problem is that attractors with more than one unstable focus, e.g., the Lorenz attractor [18], have no significant fill-factor extrema for locating τ [16].

As a solution to the difficulties related to the fill-factor criterion, Buzug and Pfister [16] chose the delay that minimizes a local measure of attractor deformation. However, this method is somewhat more complicated and requires a significant amount of additional computation. In [17], Buzug and Pfister described another, less-complicated local measure that also requires more computation than the fill-factor approach. (The increased computation times associated with these local measures is due to the need to perform a costly nearest-neighbors search.) It is important to note that the computational cost of the fill factor exhibits an (approximately) exponential increase with increasing embedding dimension. Thus, the fill factor, itself, requires a substantial number of calculations for embedding dimensions greater than three or four. (For example, with m=4, the computational cost of the fill-factor algorithm is about two orders of magnitude greater than that of autocorrelation.)

2.5. Wavering product

Like the fill factor, the *wavering product*, $\overline{W}(m,\tau)$, provides a spatial criterion for choosing the optimal delay [19]. The wavering product quantifies the distribution of numerous sets of nearest neighbors for successive values of m. Thus, $\overline{W}(m,\tau)$ is useful for estimating the smallest embedding dimension that preserves the topological invariants. This is accomplished after recognizing that spurious nearest neighbors are present when the embedding dimension is too small. (Previous methods for choosing m, e.g., [5, 7], are based on this approach.) The wavering product is also dependent upon τ, since $\overline{W}(m,\tau)$ utilizes nearest neighbors that are reconstructed using the method of delays. Liebert *et al.* [19] found that the first minimum of $\overline{W}(m,\tau)$ with respect to τ corresponds to a suitable choice for the reconstruction delay. (Like the mutual-information and fill-factor criteria, the wavering product -- as proposed in [19] -- does not account for the embedding dimension when determining the proper delay.) Note that $\overline{W}(m,\tau)$ led to values of τ that are 10-15% less than those previously found by Liebert and Schuster [13] via the mutual-information criterion [19]. Furthermore, the wavering product requires a considerable amount of computation time (20 times more than the fill-factor method [16] or about three orders of magnitude more than autocorrelation).

2.6. Small-window solution

As a follow-up to their work on state-space reconstructions in the presence of noise [4], Gibson *et al.* [20] derived the relationships between three different reconstruction techniques: delays, derivatives, and principal components. Much of their analysis was based on the assumption that the system observable, $x(t)$, is analytic such that for sufficiently small windows, one may replace $x(t + \tau_w)$ with a Taylor expansion about $x(t)$:

$$x(t+\tau_w) = \sum_{i=0}^{\infty} \frac{(\tau_w)^i}{i!} \frac{\mathrm{d}^{(i)}}{\mathrm{d}t^{(i)}} x(t). \tag{7}$$

Eq. (7) allowed Gibson *et al.* [20] to show that under certain conditions, principal components are a rotation of derivative coordinates which, in turn, are a rotation of delay coordinates. They proceeded to show that the "small-window solution" is valid for $\tau_w << \tau_w^*$, where τ_w^* is the *critical window width*:

$$\tau_w^* = 2\sqrt{\frac{3\langle x^2 \rangle}{\langle \dot{x}^2 \rangle}}. \tag{8}$$

($\langle .. \rangle$ denotes an average over the time index.)

Using Eq. (8) and a time series, one may calculate a suitable window length as

$$\tau_w = \mu \tau_w^*, \tag{9}$$

where μ is chosen to fit the particular application and data set. (Based on empirical results [20], $\mu \approx \frac{1}{2}$ appears to be a good estimate.) The advantages of this method are a sound theoretical basis and short computation times (roughly 10 times shorter than autocorrelation). The primary drawback involves the need to calculate $\dot{x}$ in Eq. (8). Since differentiation amplifies noise, one typically requires a suitable noise-reduction technique for experimentally obtained data.

While formulating the small-window solution, Gibson *et al.* [20] also developed a mathematical framework for understanding the relationship between reconstruction errors and τ. In the next section, we derive a similar relationship from a geometry-based perspective and show that the size of the reconstructed attractor is related to the influence of measurement error.

Furthermore, we confirm the result that "good delay reconstructions sit on the upper edge of the small-window solution" [20, p. 18]. Specifically, we show that the optimal delay depends upon the details of the time series (not just the dynamics), and that in general, one can only estimate an upper bound to the optimal delay.

3. Reconstruction expansion and optimal delay

3.1. Reconstruction errors

One may view the method of delays as a bridge between the temporal fluctuations of a single system observable and the spatial characteristics of a dynamical system. However, we reiterate that Takens' theorem provides no assistance when selecting a reconstruction delay for experimentally obtained data. For that reason, one seeks the delay that results in the optimal tradeoff between redundance and irrelevance [4]. Figure 1 shows a schematic of this tradeoff.

**** FIGURE 1 NEAR HERE ****

For small delays, redundance is relatively high since additive measurement error (including quantization error) is comparatively large with respect to the average difference, i.e., information gain, between successive delay coordinates. If one increases the delay, the difference between successive delay coordinates increases, and the relative impact of the measurement error declines, i.e., redundance decreases. Note that the "Redundance Error" indicated in Figure 1 reaches a plateau because attractors are bounded in state space. In other words, there is a limit to the maximum difference between delay coordinates, and, hence, there is a limit to the improvements gained with larger delays. (Note that we associate the term "measurement error" with a time series and the term "redundance error" with a reconstruction. An analogous usage applies to the terms "dynamical error" and "irrelevance error" used below.)

With chaotic systems, i.e., bounded-output systems with sensitivity to initial conditions, small perturbations of a system's state result in a type of error that may grow exponentially. The manifestation of this *dynamical error* is denoted "Irrelevance Error" in Figure 1. (Notice that the irrelevance error also reaches a limit because the system is bounded in state space.) Increases in

reconstruction delay -- which correspond to reduced effects from measurement error -- translate into extended time for exponential growth and, hence, into greater effects from dynamical error. It follows, as illustrated in Figure 1, that the optimal delay balances the redundance and irrelevance errors such that the total error is minimized.

Since one typically cannot measure dynamical error, methods for determining the proper reconstruction delay work with the implicit assumption that irrelevance error is small in comparison to redundance error. Under this assumption, one simply estimates τ as the delay that minimizes the redundance error. However, the validity of this assumption is questionable and may be the source of much of the inconsistency present in the literature. Thus, a more conservative view is to consider previous methods (as well as the approach described in Section 3.3) as techniques for determining an upper bound to the optimal delay.

3.2. Expansion and measurement error

Geometry-based methods for determining τ may be interpreted as various attempts to answer the following question: What value for the delay results in the most space-filling reconstruction? The difficulty with this sort of approach is that one has no guarantee that the reconstruction will become less space-filling as the lag increases beyond the optimal value. Hence, we attempt to answer an alternate question that views the problem from a slightly different perspective: For a given delay, is the reconstruction sufficiently expanded from the line of identity? As shown below, we refer to "sufficient expansion" as an adequate reduction in redundance error.

Before examining redundance error, one requires a concept of *reconstruction signal strength* (RSS). As suggested in the previous section, the two-dimensional RSS is related to the difference between delay coordinates; we quantify this difference as follows:

$$RSS_2(\tau) = RSS_2(J \cdot \Delta t) = \frac{1}{M}\sum_{i=1}^{M}\left|x_{i+J} - x_i\right|, \tag{10}$$

where the subscripted "2" refers to a two-dimensional reconstruction. Figure 2 shows a plot of $RSS_2(\tau)$ for Gaussian noise and the *x*-coordinate time series from the Lorenz attractor. (The

noise time series has a variance equal to one-tenth that of the Lorenz time series.) Notice that for white noise the RSS is independent of lag. Thus, if one considers redundance as a measure of a clean signal's RSS as compared to that of additive Gaussian noise, then minimizing the redundance error is equivalent to maximizing the RSS of the noise-free signal.

**** FIGURE 2 NEAR HERE ****

Geometrically, $RSS(\tau)$ is a measure of an attractor's expansion as a function of delay. Expansion from the identity line is illustrated in Figure 3, which shows two-dimensional reconstructions of the time series used to generate Figure 2. Notice that for increasing lag, the reconstructed attractor expands whereas the noise fills an area of approximately constant size. At a lag of about 0.15 s, the attractor appears to be "sufficiently" expanded. Thus, a larger delay is not expected to yield a substantial reduction in the total reconstruction error.

**** FIGURE 3 NEAR HERE ****

It is conceivable that one may devise numerous methods for quantifying expansion from the main diagonal. (See, for example, the alternative methods described in Sections 5.2 and 5.3.[3]) With regards to established theory, there is no *a priori* reason to assume one approach will be better than another. Hence, we recommend some practical guidelines for selecting techniques to quantify reconstruction expansion. Comparatively, the best methods will: (1) be computationally efficient, (2) work well with noisy data sets, and (3) lead to consistent, accurate estimates of dimension. In the next section, we present the average-displacement method which satisfies each of the above guidelines.

3.3. Average displacement

We have found that expansion from the main diagonal is best quantified by measuring the average displacement, $\langle S_m \rangle$, of the embedding vectors from their original locations on the line of identity. That is, we calculate $\langle S_m \rangle$ as a function of τ such that

---

[3] Also, see the work of Kember and Fowler [21], which was published after the submission of our revised manuscript.

$$\langle S_m(\tau)\rangle = \frac{1}{M}\sum_{i=1}^{M}\left\|\mathbf{X}_i^{\tau} - \mathbf{X}_i^{0}\right\| , \tag{11}$$

where the superscripts denote the delay between successive embedding components. Using the scalar time series, $\langle S_m \rangle$ is computed as follows:

$$\langle S_m(J\cdot\Delta t)\rangle = \frac{1}{M}\sum_{i=1}^{M}\sqrt{\sum_{j=1}^{m-1}\left[x_{i+jJ} - x_i\right]^2} . \tag{12}$$

One may interpret $\langle S_m \rangle$ as a multi-dimensional extension of RSS_2. (See Eq. (10).) Thus, $\langle S_m \rangle$ is useful for quantifying the decrease in redundance error with increasing τ. Figure 4 shows the relation between $\langle S_m \rangle$ and τ for the time series used to generate Figures 2 and 3. (The embedding dimensions are $m = 2$, 5, and 8.) As the lag increases from zero, the reconstructed trajectory expands from the main diagonal and the average displacement increases accordingly. It is important to recognize that with larger values of *m*, reconstruction expansion reaches a plateau at smaller values of τ. This is a convenient geometrical feature that helps maintain an approximately constant embedding window -- and therefore a constant level of dynamical error between the first and last embedding components -- for increasing values of *m*. (The importance of a fixed embedding window is discussed further in Section 5.1.)

**** FIGURE 4 NEAR HERE ****

4. Numerical results

4.1. Correlation dimension

Table I summarizes the dynamical systems examined in this section (which are similar to those in [15]). The differential equations were solved numerically using a fourth-order Runge-Kutta integration with a step size equal to Δt as given in Table I.[4] For each system, a 10,000-point time series was generated for further analysis; with the Lorenz and Rössler attractors, the *x*-

[4] For the Rössler and Mackey-Glass systems, the step size was 10 times smaller than the sampling period. The resulting time series were then downsampled by a factor of 10 to achieve the desired Δt.

coordinate time series was used to reconstruct the dynamics. When implemented on a 25 MHz desktop computer, our method for choosing τ requires less than 5 s of computation time using a 2,500-point time series from the Lorenz attractor and a seven-dimensional embedding space. (This compares to approximately 2 s for the autocorrelation-based methods described in Section 2.1.)

**** TABLES I NEAR HERE ****

Figures 5-8 show curves of $\langle S_m \rangle$ versus τ for the respective dynamical systems, where the embedding dimension was chosen to be the minimum value that satisfies Takens' theorem, i.e., m=7. In each graph, the selected value of τ, which is denoted as τ_S, is distinguished using a dashed vertical line. (For comparative purposes, the points that correspond to the critical window width, τ_w^*, are marked with open circles.) Based on empirical results, we quantified τ_S as the point where the slope first decreased to less than 40% of its initial value.[5] This criterion is somewhat arbitrary since we have no knowledge of the dynamical error needed to estimate the true optimal delay (described in Section 3.1). However, as illustrated for the Lorenz attractor in Figure 9, acceptable reconstructions may be obtained using a range of lags in the vicinity of the 40% slope threshold. By choosing τ near τ_S, one expects the greatest information gain between successive delay coordinates, i.e., the least redundance, while avoiding excessive dynamical error.

**** FIGURES 5-9 NEAR HERE ****

Figure 10 shows plots of correlation dimension (D_2) [14] versus $\ln r / r_{\max}$ for each dynamical system, where $\tau = \tau_S$ as marked in Figures 5-8. For each curve, D_2 was calculated via Eq. (6), i.e., the correlation sum [14], using approximately 10^3 embedding vectors formed

---

[5] As expected, we observed that an anomalous jump in S_m from $\tau = 0$ to $\tau = \Delta t$ (or equivalently from J=0 to J=1) is associated with noisy data sets. Thus, one should exclude $S_m(0)$ when calculating the initial slope.

from the original 10,000-point data set.[6] As expected, one observes a plateau of D_2 near the expected dimension, i.e., $D_2 \approx 2$ for the chaotic systems, and $D_2 = 3$ for the three-torus. Thus, $\langle S_m \rangle$ (and, presumably, other measures of reconstruction expansion) is useful for the implementation of the method of delays with experimentally obtained data.

**** FIGURE 10 NEAR HERE ****

4.2. Redundance error

As detailed in Section 3, reconstruction expansion is related to a notion of reconstruction signal strength. Furthermore, reconstruction expansion, i.e., an increase in RSS, corresponds to a decrease in redundance error. To illustrate this phenomenon, we utilized numerical experiments involving the Lorenz attractor (the *x*-coordinate time series) with superimposed Gaussian noise. (The Lorenz time series had a variance equal to 1000 times that of the noise time series.) By adding noise to the clean data set, we introduced a substantial measurement error. However, as illustrated below, the addition of noise resulted in modest amounts of redundance error with large delays.

Figure 11 shows D_2 versus $\ln r / r_{\max}$ for the noise-free data set and three different values of τ: (1) $\tau = \tau_S$ (heavy, solid line), (2) $\tau = \tau_S/4$ (light, solid line), and (3) $\tau = 4\tau_S$ (dashed line). All three cases yielded acceptable estimates of D_2, although Case 2 clearly underestimated the true dimension (≈2.06). Figure 12 shows the corresponding results from the noisy data set. As expected, the noise appears to have a progressive effect as the lag decreases. In particular, the plateau of D_2 was lost with $\tau = \tau_S/4$, whereas the plateau for $\tau = 4\tau_S$ was virtually unchanged.

**** FIGURES 11 & 12 NEAR HERE ****

4.3. Irrelevance error

---

[6] The vectors had a uniform temporal sampling equal to ten times the sampling period shown in Table I. This strategy allowed us to study the effects of small changes in τ without the burden of the unnecessary computations normally associated with large, over-sampled data sets.

Next we illustrate the impact of irrelevance error using an example similar to the one presented the previous section. Specifically, Figure 11 serves as a "baseline" derived from the noise-free data set, and Figure 13 shows the corresponding results when the dynamical error is purposely exaggerated. Our paradigm for dynamical error was as follows: the Lorenz equations were numerically integrated using a sampling interval of 0.04 s; then, linear interpolation was performed to simulate the sampling interval ($\Delta t = 0.01$ s) used with the noise-free data set. By increasing the step size of the numerical integration, we effectively introduced a perturbation that was allowed to propagate and alter the system's state-space trajectory.

**** FIGURE 13 NEAR HERE ****

It can be seen in Figure 13 that the dynamical error had the greatest effect on the reconstruction from the largest delay. For the intermediate delay, i.e., for $\tau = \tau_S$, the plateau of D_2 was nearly identical to that of the noise-free data set. We also observed an unintended shift in D_2 for the smallest delay. This change may be attributed to redundance error that arises from the inaccuracies related to linear interpolation. Thus, our model incorporated measurement error in addition to dynamical error. As desired, however, the value of τ selected using $\langle S_m \rangle$ balanced both forms of error and yielded the best reconstruction for this example.

5. Discussion

5.1. Embedding window vs. lag

As stated in Section 1, previous research [3, 9, 10] suggests that one should specify the length of the embedding window, τ_w, rather than τ and m separately. In Figure 14 we provide additional evidence that choosing τ_w is of greater relevance. For the investigated systems, we observed a dependence of τ_S on the embedding dimension such that τ_w was roughly independent of m. That is,

$$\tau_w(m) = (m-1) \cdot \tau_S(m) \approx \text{constant}. \qquad (13)$$

Figure 14 illustrates this m-independence of τ_w. From Eq. (13) it follows that one may select τ_w after quantifying reconstruction expansion using just one embedding dimension. However,

note the slight upward trend of the curves in Figure 14. We suspect that this trend is due either to the geometry of the reconstructions or to estimation errors associated with finite data sets. Thus, the more prudent approach is to calculate τ_w explicitly for each value of *m*.

**** FIGURE 14 NEAR HERE ****

In light of Takens' theorem, there is no theoretical basis to expect that τ_w is of greater importance than τ. Thus, we associate this finding with the practical ramifications of analyzing noisy data sets. Figure 14 shows that a fixed window length can lead to near-maximum expansion and, therefore, near-minimum redundance. Similarly, by employing a fixed window length, one ensures that the irrelevance error is held at an acceptable level. That is, the first and last embedding components maintain a certain degree of causality. Increasing the number of components then increases the "precision" with which one may answer the following question: Given the observed measurements, what is the probability that the system is in a particular state?

Conversely, the use of a fixed delay may result in the undesirable blurring of the information from two (or more) states as the number of delay coordinates increases. The notion of "blurred states" becomes important when dealing with approximations to true trajectories of a dynamical system. For example, consider the reconstruction of a system's dynamics with a fixed delay and twice the minimum number of delay coordinates required by Takens' theorem, i.e., $4n+2$ coordinates. (Recall that *n* is the topological dimension.) With true trajectories, the $4n+2$ coordinates could be considered as an overly detailed specification of one state along one trajectory. With approximations to true trajectories, however, the first $2n+1$ coordinates and last $2n+1$ coordinates could be considered as specifications of two different states along two different (albeit nearby) trajectories. (Considered together, the $4n+2$ coordinates for the approximated trajectory represent a blurred state that may not be sufficiently close to any true reconstructed state.)

5.2. Effects of noise

Prior to developing the average-displacement technique, we sought after the most computationally efficient method for quantifying expansion from the main diagonal. As a first attempt, we measured the total reconstructed trajectory length, L_{total}, as a function of delay:

$$L_{total}(\tau) = \sum_{i=1}^{M-1} \left\| \mathbf{X}_{i+1}^{\tau} - \mathbf{X}_{i}^{\tau} \right\|. \tag{14}$$

L_{total} is attractive from a computational standpoint because the number of calculations involved is proportional to M and independent of m.[7] With noise-free data, Eq. (14) yielded curves of L_{total} versus τ that had a similar shape to those of average displacement versus τ. Therefore, in such cases, L_{total} is the preferred method for choosing τ_S. Unfortunately, L_{total} performed poorly with noisy time series. Even a modest amount of additive noise caused an undesirable increase in L_{total} -- undesirable because the effect was greater with smaller values of τ. Thus, as the noise level increased, curves of L_{total} versus τ acquired a flatter appearance such that one could no longer estimate τ_S consistently. Due to this failure of L_{total}, we then turned to measuring displacements of a reconstructed point from the main diagonal.[8]

Figure 15 illustrates the performance of $\langle S_m \rangle$ for the Lorenz attractor and additive Gaussian noise. Before superposition with the original noise-free signal, the Gaussian noise was scaled to achieve a desired signal-to-noise ratio (SNR), where SNR equals the ratio of the powers

---

[7] For i=1 to J-1, computing the Euclidean distance in Eq. (14) carries a computational cost proportional to m. However, for each subsequent value of i, an efficient algorithm uses a fixed amount of additional computation, i.e., three additions, one multiplication, and one square root.

[8] Before settling on $\langle S_m \rangle$, we tested the efficacy of quantifying attractor expansion using the average perpendicular distance from the identity line, $\langle S_m^{\perp} \rangle$. Both $\langle S_m \rangle$ and $\langle S_m^{\perp} \rangle$ gave similar results for the examined systems. However, curves of $\langle S_m^{\perp} \rangle$ typically displayed a less prominent bend in the vicinity of τ_S. Thus, $\langle S_m^{\perp} \rangle$ is expected to be more sensitive to variations in the slope threshold mentioned in Section 4. Furthermore, $\langle S_m^{\perp} \rangle$ carries a larger computational cost than $\langle S_m \rangle$, i.e., three extra additions and two extra multiplications per point.

in the noise-free and pure-noise signals. Even with extremely noisy data sets, i.e., SNR = 1-10, we were able to select τ_S with results almost identical to those of the noise-free data set, i.e., SNR = ∞. The present method works well with noisy data sets because additive noise predominantly affects the accuracy of individual embedding vectors and not the overall volume of the attractor. That is, noise displaces some embedding vectors toward the main diagonal and others away from the diagonal such that the *average* displacement is approximately unaltered. However, notice the upward bias of the $\langle S_m \rangle$ versus τ curve with increasing amounts of noise. We expect this characteristic due to the increased likelihood that a vector will be displaced to the "opposite side" of the main diagonal rather than simply closer to the diagonal.

**** FIGURE 15 NEAR HERE ****

5.3. Average displacement vs. autocorrelation

In this paper we advocate geometry-based methods for choosing the reconstruction delay. To provide further insight into the advantages of quantifying reconstruction expansion, we examined the relationship between average displacement and autocorrelation. Due to the square-root operation in Eq. (12), it is difficult to make a direct examination of this relationship, and therefore, we make an indirect comparison via average *squared* displacement, $\langle S_m^2 \rangle$. More specifically, we briefly describe the qualitative relationship between $\langle S_m \rangle$ and $\langle S_m^2 \rangle$, and then we derive the analytical relationship between $\langle S_m^2 \rangle$ and R_{xx}.

First we define average squared displacement in a similar fashion as average displacement (Eqs. (11) and (12)):

$$\langle S_m^2(\tau) \rangle = \frac{1}{M} \sum_{i=1}^{M} \left\| \mathbf{X}_i^{\tau} - \mathbf{X}_i^{0} \right\|^2, \text{ and} \tag{15}$$

$$\langle S_m^2(J \cdot \Delta t) \rangle = \frac{1}{M} \sum_{i=1}^{M} \sum_{j=1}^{m-1} \left[x_{i+jJ} - x_i \right]^2. \tag{16}$$

From our definitions, it follows that $\langle S_m^2 \rangle$ may be interpreted as a scaled version of $\langle S_m \rangle$, such that

$$\left\langle S_m^2(\tau)\right\rangle = \left\langle f(\tau)\cdot S_m(\tau)\right\rangle, \tag{17}$$

where $f(\tau)$ is a nonlinear function derived from the data set. Geometrically, $f(\tau)$ represents a weighting function that disproportionately contracts and expands individual displacement vectors before the average displacement is computed. For example, the greater the displacement of a given vector from the main diagonal, the greater its influence on $\left\langle S_m^2\right\rangle$.

Before deriving the relationship between $\left\langle S_m^2\right\rangle$ and R_{xx}, we assume that R_{xx} may be approximated from a finite data set, where

$$R_{xx}(J\cdot\Delta t)\approx\frac{1}{N-J}\sum_{i=1}^{N-J}x_i\cdot x_{i+J}\,. \tag{18}$$

Additionally, we assume that the number of scalar data points, *N*, is sufficiently large such that *N* is approximately equal to the number of reconstructed embedding vectors, *M*. Next, we square the quantity in brackets from Eq. (16):

$$\begin{aligned}\left\langle S_m^2(J\cdot\Delta t)\right\rangle &= \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{m-1}\left[x_{i+jJ}^2-2x_{i+jJ}\cdot x_i+x_i^2\right]\\ &= \frac{1}{M}\sum_{j=1}^{m-1}\sum_{i=1}^{M}\left[x_{i+jJ}^2-2x_{i+jJ}\cdot x_i\right]+\text{constant}.\end{aligned} \tag{19}$$

Our third assumption is that *M* is sufficiently large such that

$$\sum_{j=1}^{m-1}\sum_{i=1}^{M}x_{i+jJ}^2\approx\sum_{j=1}^{m-1}\sum_{i=1}^{M}x_i^2=\text{constant}. \tag{20}$$

It follows that

$$\begin{aligned}\left\langle S_m^2(J\cdot\Delta t)\right\rangle &\approx \text{constant}-\frac{1}{M}\sum_{j=1}^{m-1}\sum_{i=1}^{M}\left[2x_{i+jJ}\cdot x_i\right]\\ &\approx \text{constant}-2\sum_{j=1}^{m-1}R_{xx}(jJ\cdot\Delta t)\end{aligned}\,. \tag{21}$$

Since we are primarily interested in curve shapes, we can ignore the constant and multiplication factor in Eq. (21). Hence,

$$\left\langle S_m^2(J\cdot\Delta t)\right\rangle \Leftrightarrow \sum_{j=1}^{m-1} R_{xx}(jJ\cdot\Delta t), \tag{22}$$

where "⇔" means "has the same shape as." From Eq. (22), one readily notices that the autocorrelation function may be interpreted in terms of the geometry. Specifically, R_{xx} has the same shape as $\left\langle S_m^2\right\rangle$ when m=2. This result suggests a possible explanation for the difficulties associated with the use of R_{xx} for defining embedding windows. Since $R_{xx}\Leftrightarrow\left\langle S_2^2\right\rangle$, R_{xx} seems best suited for two-dimensional reconstructions; however, all chaotic flows have a topological dimension of at least three, and two-dimensional reconstructions violate Takens' theorem.

Figure 16 shows an overlay of curves derived from $\left\langle S_m\right\rangle$, $\left\langle S_m^2\right\rangle$, and R_{xx} for the Lorenz attractor and two embedding dimensions (m=2, 7). Note that $\left\langle S_m^2\right\rangle$ and R_{xx} are appropriately scaled and translated to facilitate a shape comparison with $\left\langle S_m\right\rangle$. As expected, $\left\langle S_m^2\right\rangle$ and R_{xx} are nearly identical for m=2. We attribute the slight discrepancies (between $\left\langle S_m^2\right\rangle$ and R_{xx} for m=2) to the effects of working with finite data sets as well as our implementation of Eq. (18). (We used an algorithm based on the fast Fourier transform rather than a slower time-domain approach.) As illustrated in Figure 16, curves of $\left\langle S_m\right\rangle$ typically exhibit a sharper bend, i.e., smaller radius of curvature, with respect to curves of $\left\langle S_m^2\right\rangle$. Therefore, $\left\langle S_m^2\right\rangle$ is expected to be more sensitive to variations in any criterion used to quantify τ_S, e.g., a slope threshold such as that described in Section 4. This difficulty, which may lead to poorer estimates of dimension, seems to outweigh any computational advantages gained by eliminating the square-root operation in Eq. (12).

**** FIGURE 16 NEAR HERE ****

6. Summary

The problem of choosing the optimal window length for the method of delays was addressed in this paper. We discussed the reconstruction errors inherent to the method of delays and showed that reconstruction expansion is related to a reduction in the effects of measurement error. We then developed a simple procedure that quantifies expansion from the identity line of

the embedding space. Such a procedure may be used to estimate an upper bound to the optimal reconstruction delay. From a practical standpoint, this method has several advantages: (1) reliable with small data sets, (2) reliable with noisy data sets, and (3) modest computational costs. In addition, we provided further evidence that suggests one should choose the embedding window length, rather than the reconstruction delay and embedding dimension separately.

The main conclusions of this paper are twofold. First, the optimal delay time is partially understood in terms of the geometry of the reconstruction. Specifically, the greater the size of the reconstruction, the smaller the effects of measurement error. From this statement and the fact that attractors are bounded in state space, it follows that there exists an upper bound to the optimal delay. Second, the optimal delay depends upon the details of the time series as well as the dynamics of the underlying system. For example, a system may possess a characteristic time that serves as a good choice for τ with noise-free conditions; however, if the level of measurement error or dynamical error changes, the optimal delay also changes. Hence, we recommend choosing τ_w using the 40% slope threshold with curves of average displacement; Gibson *et al.* [20], on the other hand, recommend choosing τ_w as one-half the critical window width. The relationship between these criteria remains a problem for future study. Regardless of the method used to estimate τ, our work has shown that choosing the proper τ is not as simple as finding a characteristic time. However, a geometry-based framework, i.e., reconstruction expansion, is useful for making an educated decision about the proper delay.

Acknowledgments

This work was supported by the Rehabilitation Research and Development Service of Veterans Affairs.

References

[1] N. H. Packard, J. P. Crutchfield, J. D. Farmer and R. S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45 (1980) 712.

[2] F. Takens, Detecting strange attractors in turbulence, Lect. Notes in Math. 898 (1981) 366.

[3] D. S. Broomhead and G. P. King, Extracting qualitative dynamics from experimental data, Physica D 20 (1986) 217.

[4] M. Casdagli, S. Eubank, J. D. Farmer and J. Gibson, State space reconstruction in the presence of noise, Physica D 51 (1991) 52.

[5] Z. Aleksic, Estimating the embedding dimension, Physica D 52 (1991) 362.

[6] A. Cenys and K. Pyragas, Estimation of the number of degrees of freedom from chaotic time series, Phys. Lett. A 129 (1988) 227.

[7] M. B. Kennel, R. Brown and H. D. I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A 45 (1992) 3403.

[8] A. I. Mees, P. E. Rapp and L. S. Jennings, Singular-value decomposition and embedding dimension, Phys. Rev. A 36 (1987) 340.

[9] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees and P. E. Rapp, Singular-value decomposition and the Grassberger-Procaccia algorithm, Phys. Rev. A 38 (1988) 3017.

[10] J. M. Martinerie, A. M. Albano, A. I. Mees and P. E. Rapp, Mutual information, strange attractors, and the optimal estimation of dimension, Phys. Rev. A 45 (1992) 7058.

[11] G. P. King, R. Jones and D. S. Broomhead, Phase portraits from a time series: a singular system approach, Nuclear Phys. B 2 (1987) 379.

[12] A. M. Fraser and H. L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33 (1986) 1134.

[13] W. Liebert and H. G. Schuster, Proper choice of the time delay for the analysis of chaotic time series, Phys. Lett. A 142 (1989) 107.

[14] P. Grassberger and I. Procaccia, Characterization of strange attractors, Phys. Rev. Lett. 50 (1983) 346.

[15] A. M. Albano, A. Passamante and M. E. Farrell, Using higher-order correlations to define an embedding window, Physica D 54 (1991) 85.

[16] T. Buzug and G. Pfister, Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behavior of strange attractors, Phys. Rev. A 45 (1992) 7073.

[17] T. Buzug and G. Pfister, Comparison of algorithms calculating optimal parameters for delay time coordinates, Physica D 58 (1992) 127.

[18] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (1963) 130.

[19] W. Liebert, K. Pawelzik and H. G. Schuster, Optimal embeddings of chaotic attractors from topological considerations, Europhys. Lett. 14 (1991) 521.

[20] J. F. Gibson, J. D. Farmer, M. Casdagli and S. Eubank, An analytic approach to practical state space reconstruction, Physica D 57 (1992) 1.

[21] G. Kember and A. C. Fowler, A correlation function for choosing time delays in phase portrait reconstructions, Phys. Lett. A 179 (1993) 72.

[22] O. E. Rössler, An equation for continuous chaos, Phys. Lett. A 57 (1976) 397.

[23] M. C. Mackey and L. Glass, Oscillation and chaos in physiological control systems, Science 197 (1977) 287.

Captions

Table I. Summary of the examined dynamical systems.

Figure 1. Schematic of the reconstruction errors inherent to the method of delays. The optimal delay corresponds to the minimum total error.

Figure 2. Plot of two-dimensional reconstruction signal strength versus τ for the Lorenz attractor (*x*-coordinate time series) and Gaussian noise.

Figure 3. Two-dimensional phase portraits for the time series used to generate Figure 2. Lag equals: (a) 0.05 s; (b) 0.15 s; (c) 0.25 s; (d) 0.40 s.

Figure 4. Typical plot of $\langle S_m \rangle$ versus τ for the same data sets used to generate Figures 2 and 3. (The curves for m=2 are the same as those shown in Figure 2.)

Figure 5. $\langle S_m \rangle$ versus τ for the Lorenz attractor with m=7 and N=2500. The value of τ_S (the delay corresponding to the 40% slope threshold) is distinguished using a dashed vertical line. The point corresponding to τ_w^* (the critical window width) is marked with an open circle.

Figure 6. $\langle S_m \rangle$ versus τ for the Rössler attractor with m=7 and N=2500. The value of τ_S is distinguished using a dashed vertical line. The point corresponding to τ_w^* is marked with an open circle.

Figure 7. $\langle S_m \rangle$ versus τ for the three-torus with m=7 and N=2500. The value of τ_S is distinguished using a dashed vertical line. The point corresponding to τ_w^* is marked with an open circle.

Figure 8. $\langle S_m \rangle$ versus τ for the Mackey-Glass system with m=7 and N=2500. The value of τ_S is distinguished using a dashed vertical line. The point corresponding to τ_w^* is marked with an open circle.

Figure 9. D_2 versus $\ln r / r_{\max}$ for the Lorenz attractor with m=7, N=2500, and τ=0.06, 0.07, 0.08, 0.09, and 0.10 s. (The heavy line corresponds to $\tau = \tau_S = 0.08$ s. Delays of 0.06 s and 0.10 s coincided with slope thresholds of 60% and 20%, respectively.)

Figure 10. D_2 versus $\ln r / r_{\max}$ for the dynamical systems summarized in Table I. The respective values for τ correspond to τ_S as marked in Figures 5-8.

Figure 11. D_2 versus $\ln r / r_{\max}$ for the Lorenz attractor with m=7 and three values of τ: (1) $\tau = \tau_S$=0.08 s (heavy, solid line), (2) $\tau = \tau_S/4$=0.02 s (light, solid line), and (3) $\tau = 4\tau_S$=0.32 s (dashed line).

Figure 12. Effects of redundance error. The change in D_2 (as compared to the noise-free case in Figure 11) is plotted for the Lorenz attractor with m=7 and three values of τ: (1) $\tau = \tau_S$=0.08 s (heavy, solid line), (2) $\tau = \tau_S/4$=0.02 s (light, solid line), and (3) $\tau = 4\tau_S$=0.32 s (dashed line).

Figure 13. Effects of irrelevance error. The change in D_2 (as compared to the noise-free case in Figure 11) is plotted for the Lorenz attractor with m=7 and three values of τ: (1) $\tau = \tau_S$=0.08 s (heavy, solid line), (2) $\tau = \tau_S/4$=0.02 s (light, solid line), and (3) $\tau = 4\tau_S$=0.32 s (dashed line).

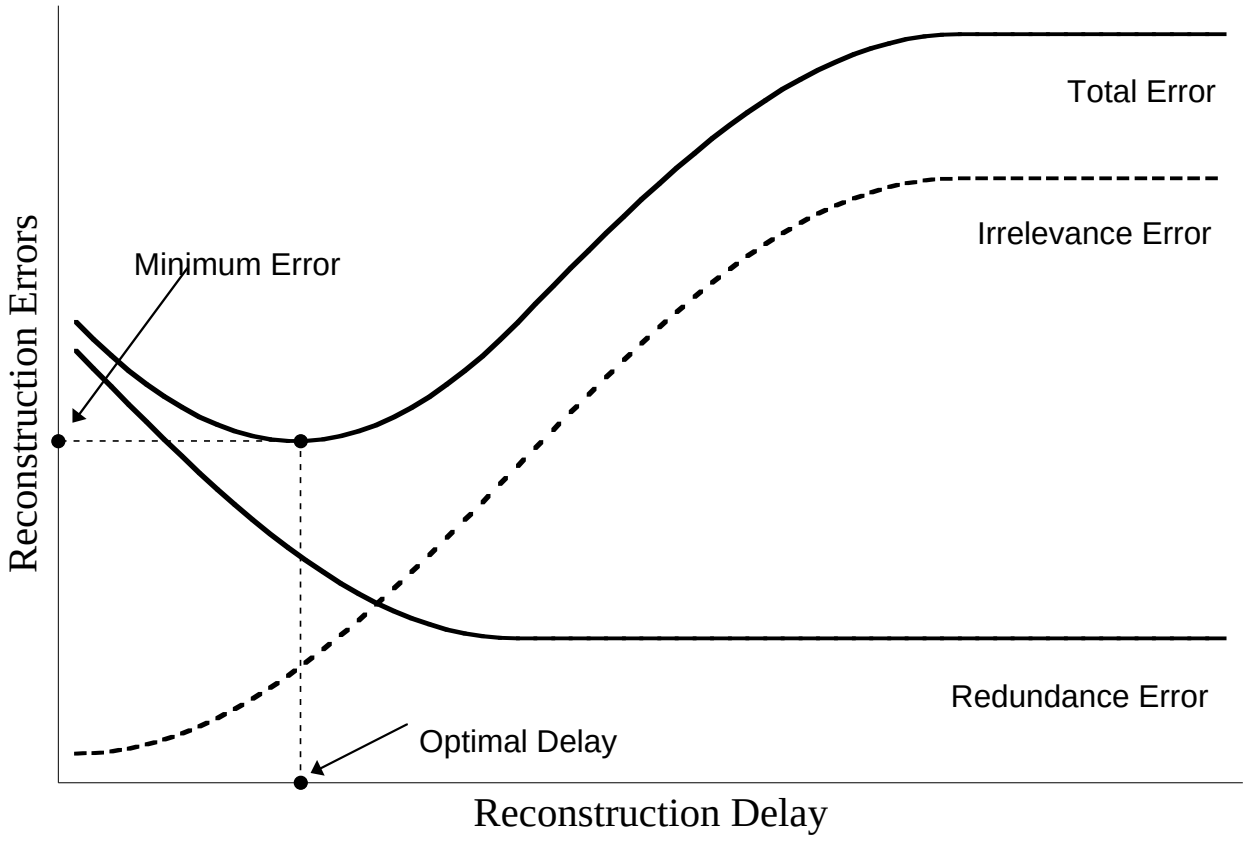
Figure 14. Normalized window length versus embedding dimension for the dynamical systems summarized in Table I. In order to calculate a more precise estimate of τ_w, the algorithm inputs were 10,000-point data sets, each with a sampling interval equal to one-fourth the Δt given in Table I. For each system, the values of τ_w were divided by the mean window length. Group means were as follows: (1) Lorenz attractor, τ_w=0.41 ± 0.02 s, (2) Rössler attractor, τ_w=2.95 ± 0.12 s, (3) three-torus, τ_w=11.1 ± 0.5 s, and (4) Mackey-Glass system, τ_w=21.9 ± 0.6 s.

Figure 15. Effects of noise for the Lorenz attractor with *m*=7. A signal-to-noise ratio (SNR) of ∞ denotes a time series that is noise-free up to the computer precision. The dashed vertical line marks the value of τ_S (0.08 s) chosen using the noise-free data set. With the SNR equal to 1 and 10, τ_S was found to be 0.10 s and 0.09 s, respectively.

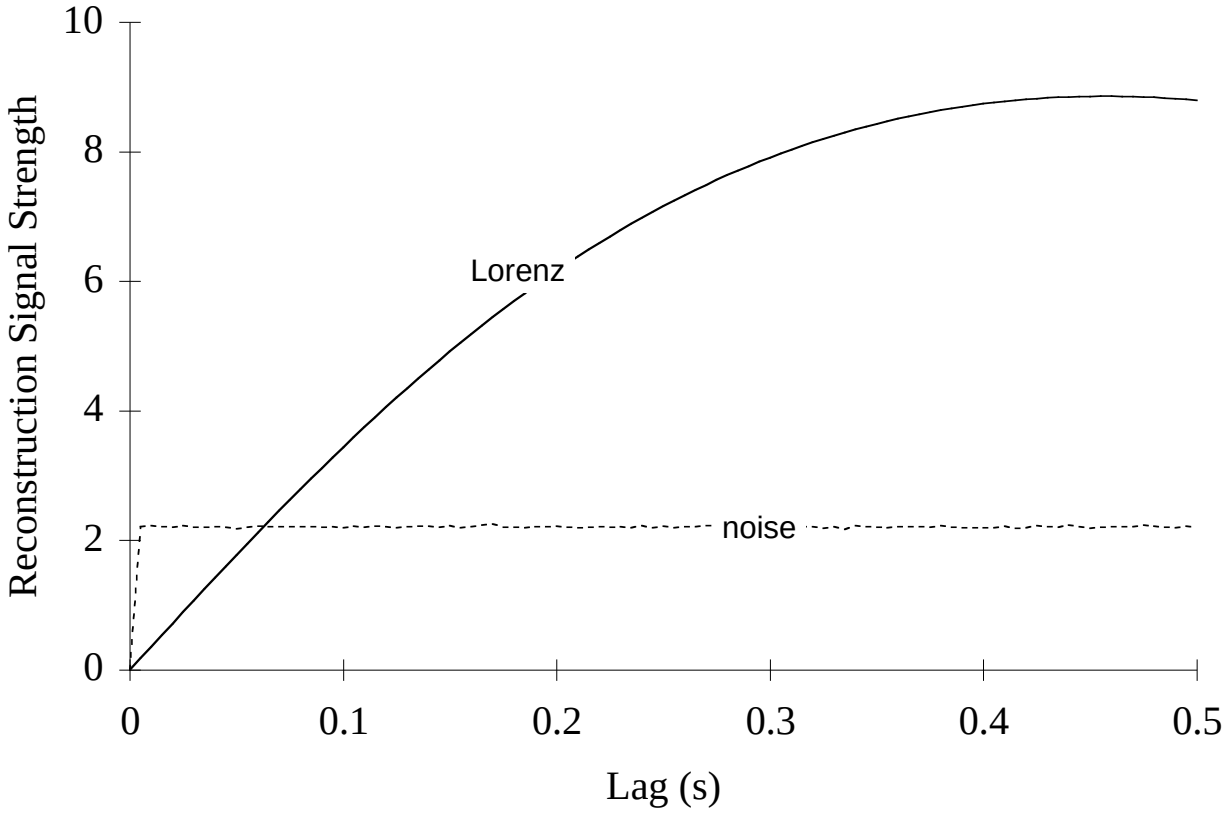
Figure 16. $\langle S_m \rangle$, $\langle S_m^2 \rangle$, and R_{xx} versus τ. The curves are appropriately scaled and translated to facilitate a shape comparison with $\langle S_m \rangle$. For $\langle S_m \rangle$ and $\langle S_m^2 \rangle$, embedding dimension equals: (a) 2; (b) 7.

System [ref.]	Equations	Parameters	Δt (s)
Lorenz [18]	$\dot{x} = \sigma(y - x)$ $\dot{y} = x(R - z) - y$ $\dot{z} = xy - bz$	$\sigma = 10.0$ $R = 28.0$ $b = 2.667$	0.01
Rössler [22]	$\dot{x} = -y - z$ $\dot{y} = x + ay$ $\dot{z} = b + z(x - c)$	$a = 0.20$ $b = 0.40$ $c = 5.7$	0.10
three-torus [15]	$x(t) = \sin(\omega_1 t) + \sin(\omega_2 t) + \sin(\omega_3 t)$	$\omega_1 = 0.060$ $\omega_2 = 0.171$ $\omega_3 = 0.314$	0.25
Mackey-Glass [23]	$\dot{x} = \frac{ax(t-s)}{1+[x(t-s)]^c} - bx(t)$	a=0.2 b=0.1 c=10.0 s=17.0	0.25

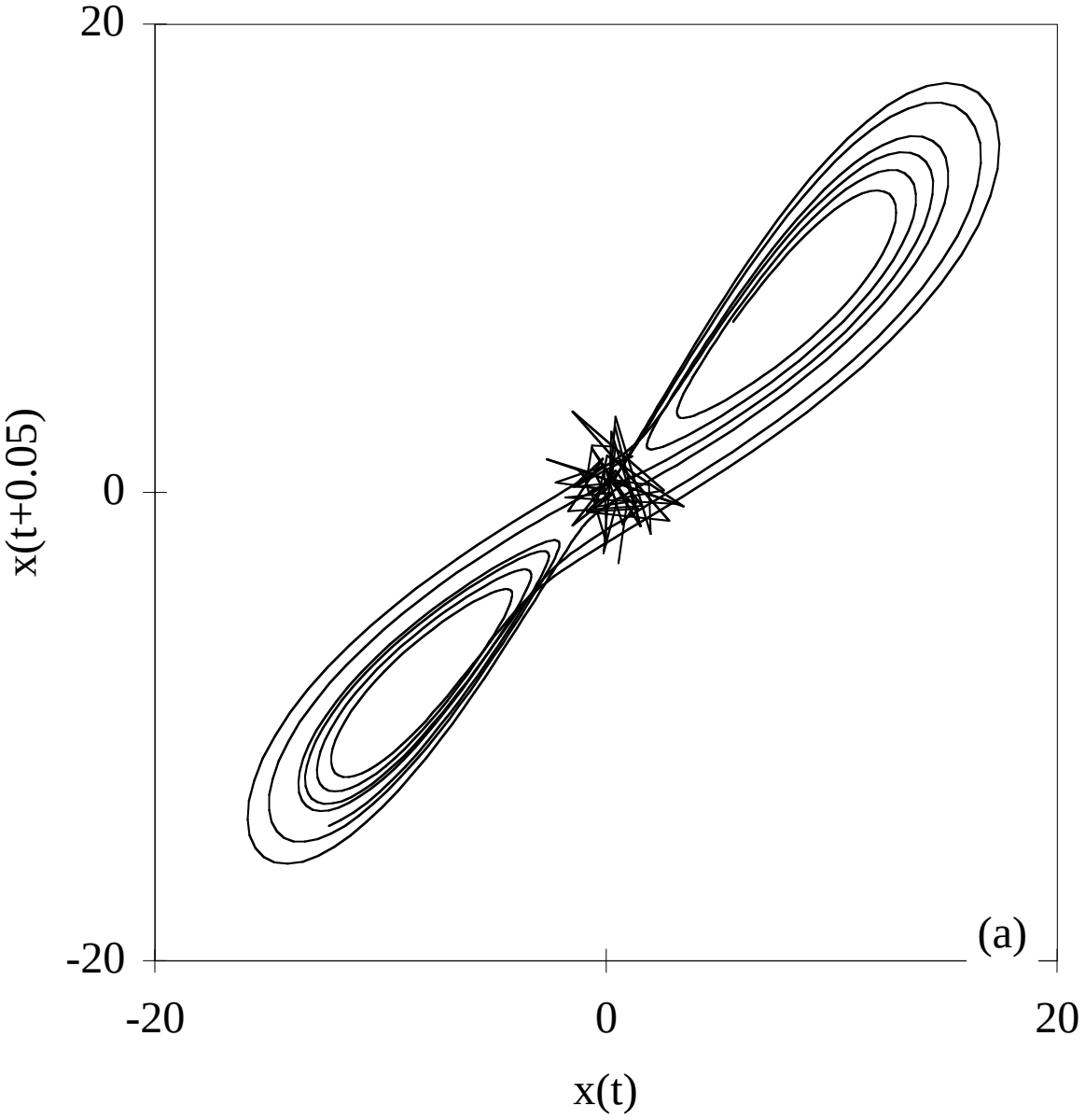
(TABLE I)


Total Error
Irrelevance Error
Minimum Error
Reconstruction Errors
Redundance Error
Optimal Delay
Reconstruction Delay

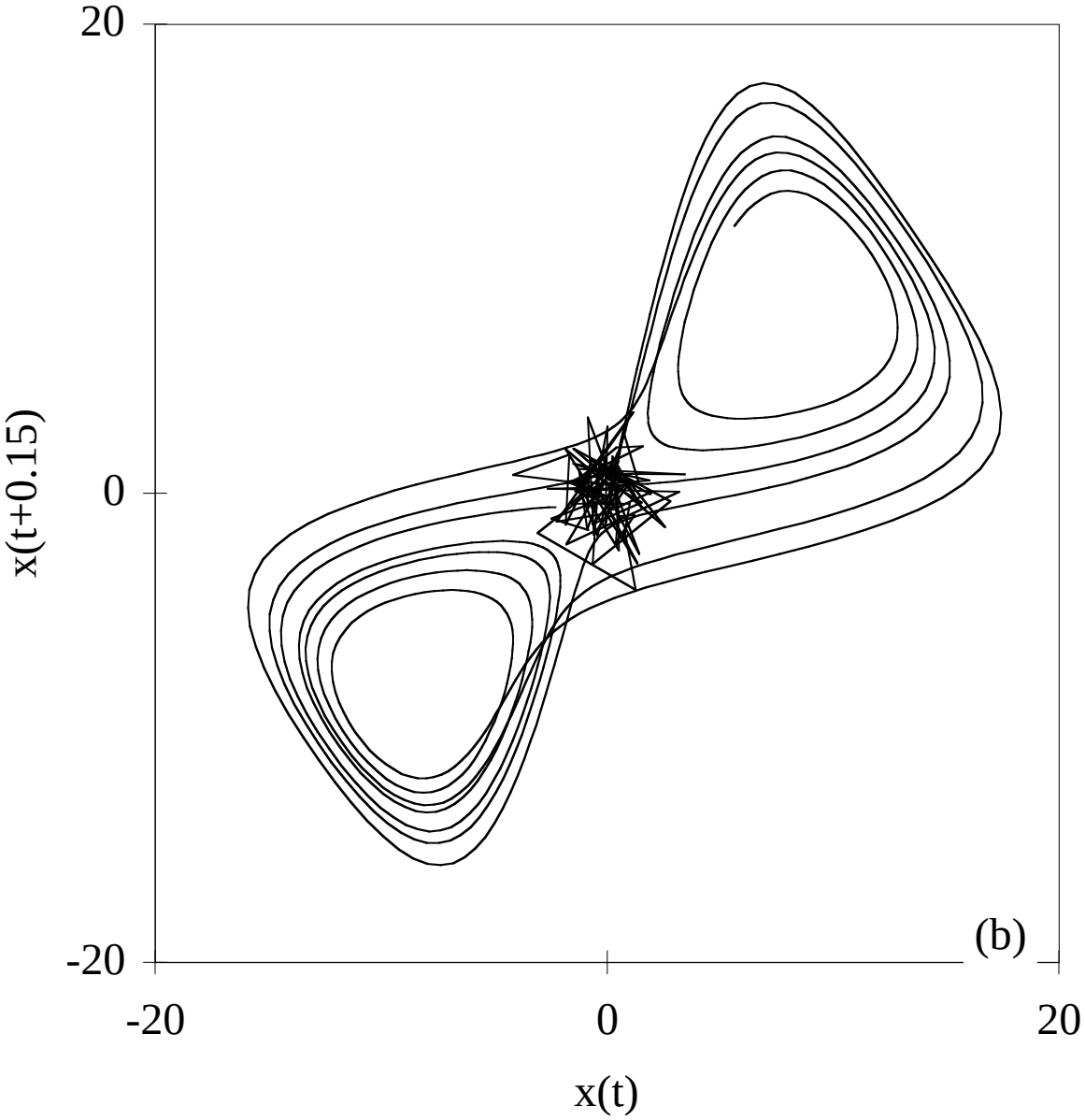
(FIGURE 1)


Reconstruction Signal Strength
10
8
6
4
2
0
Lorenz
noise
0
0.1
0.2
0.3
0.4
0.5
Lag (s)

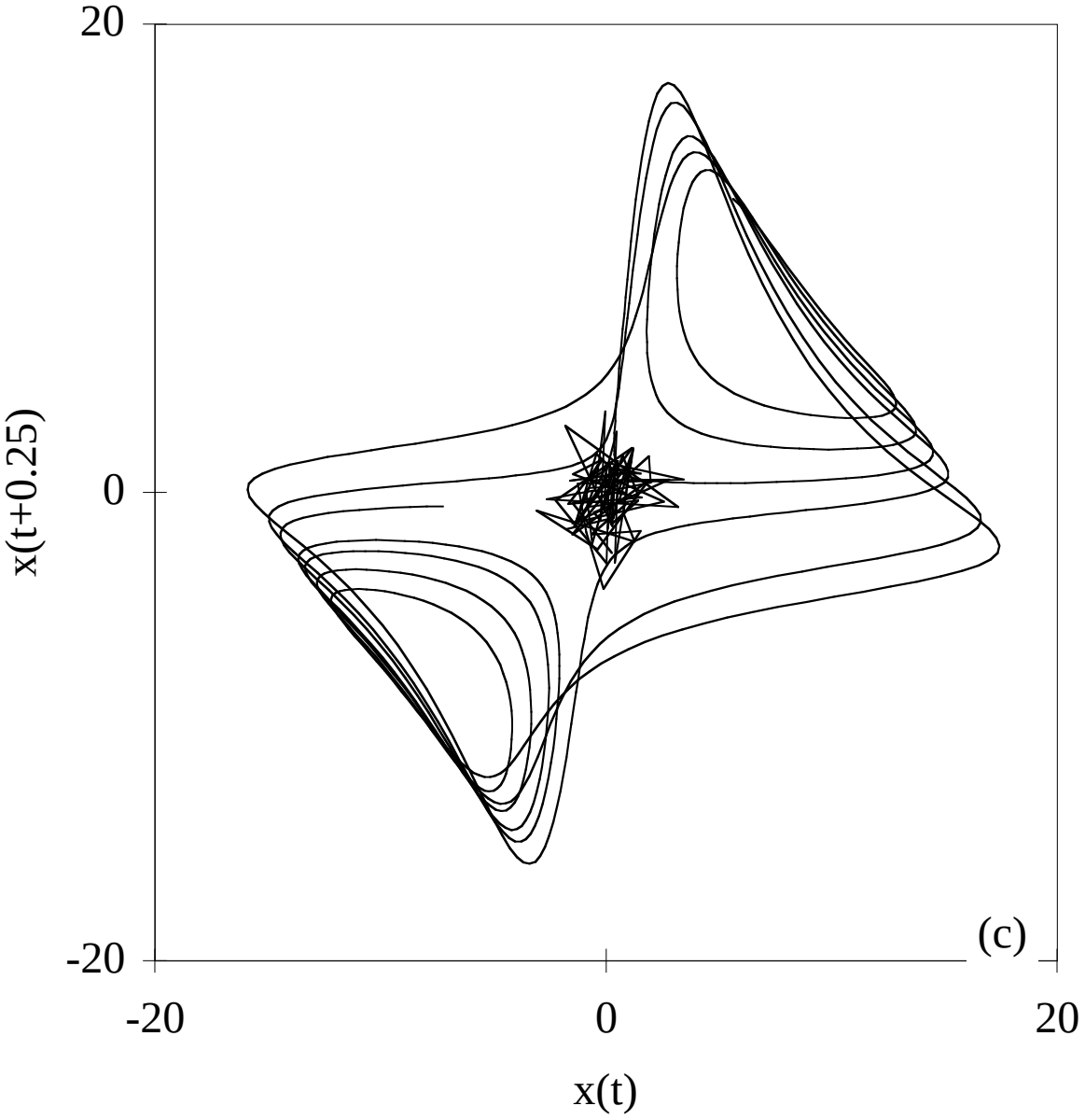
(FIGURE 2)



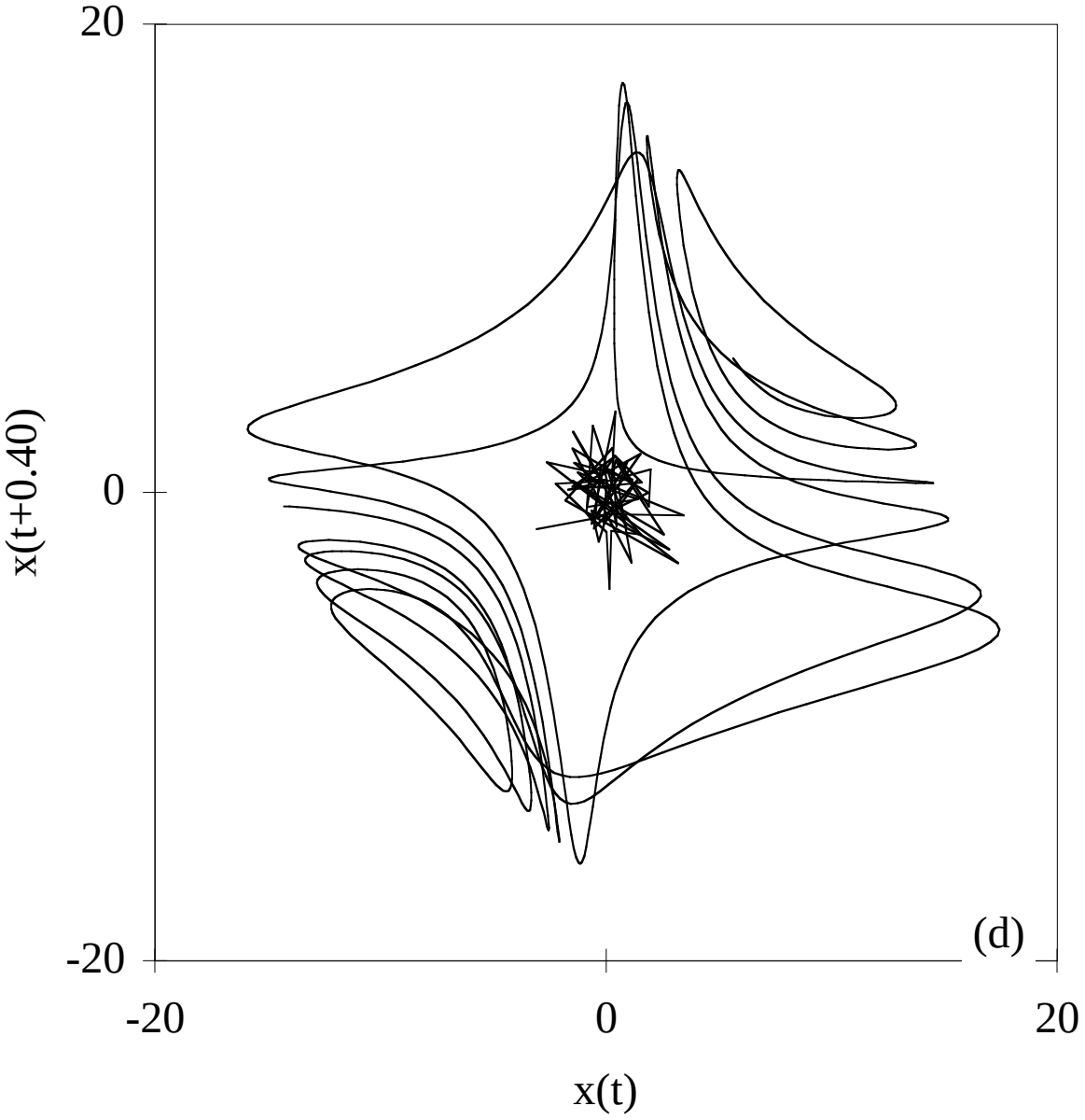
(FIGURE 3a)



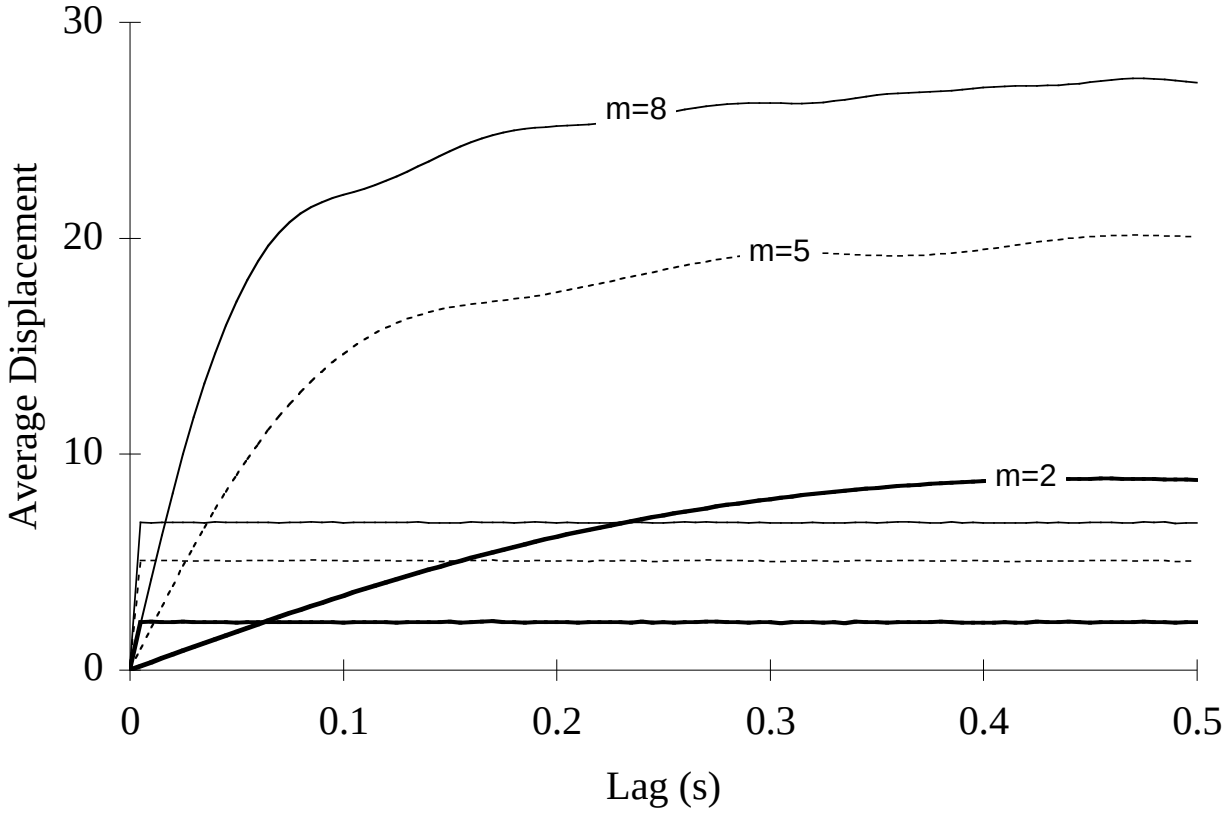
(FIGURE 3b)



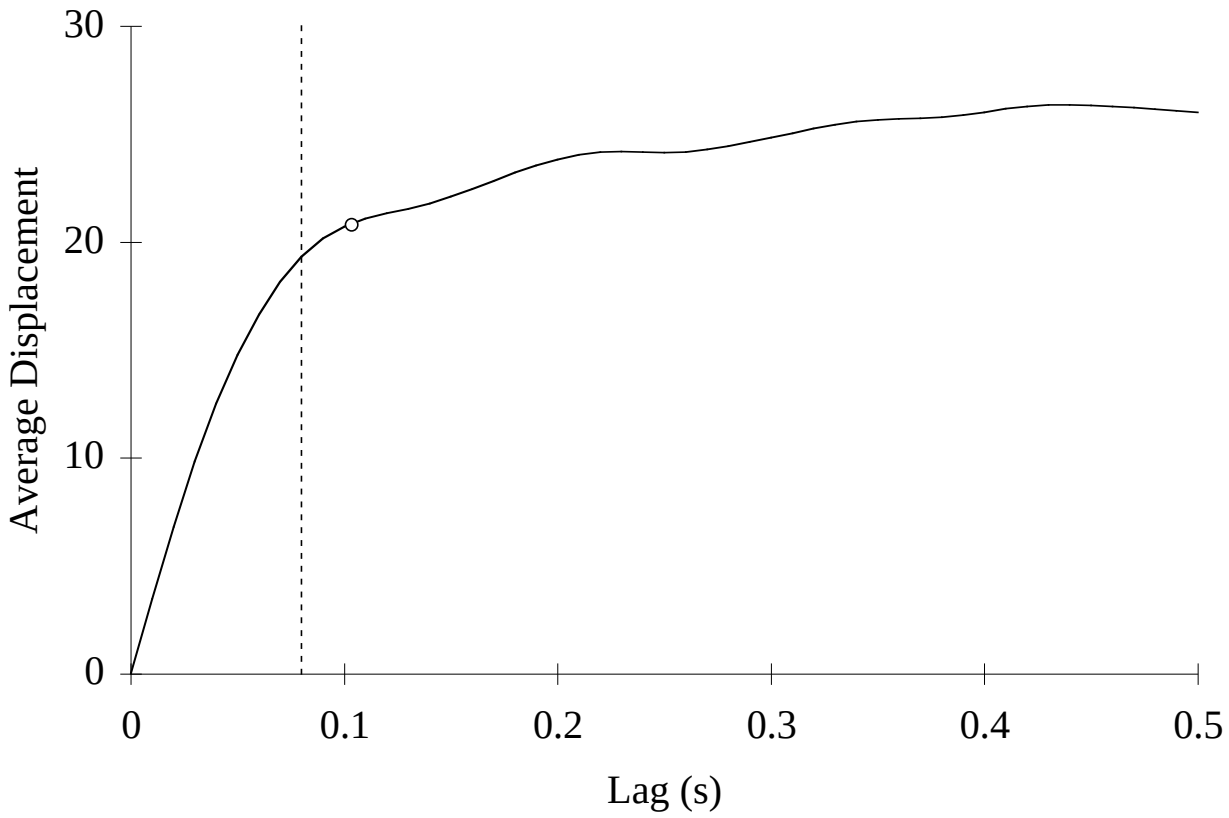
(FIGURE 3c)



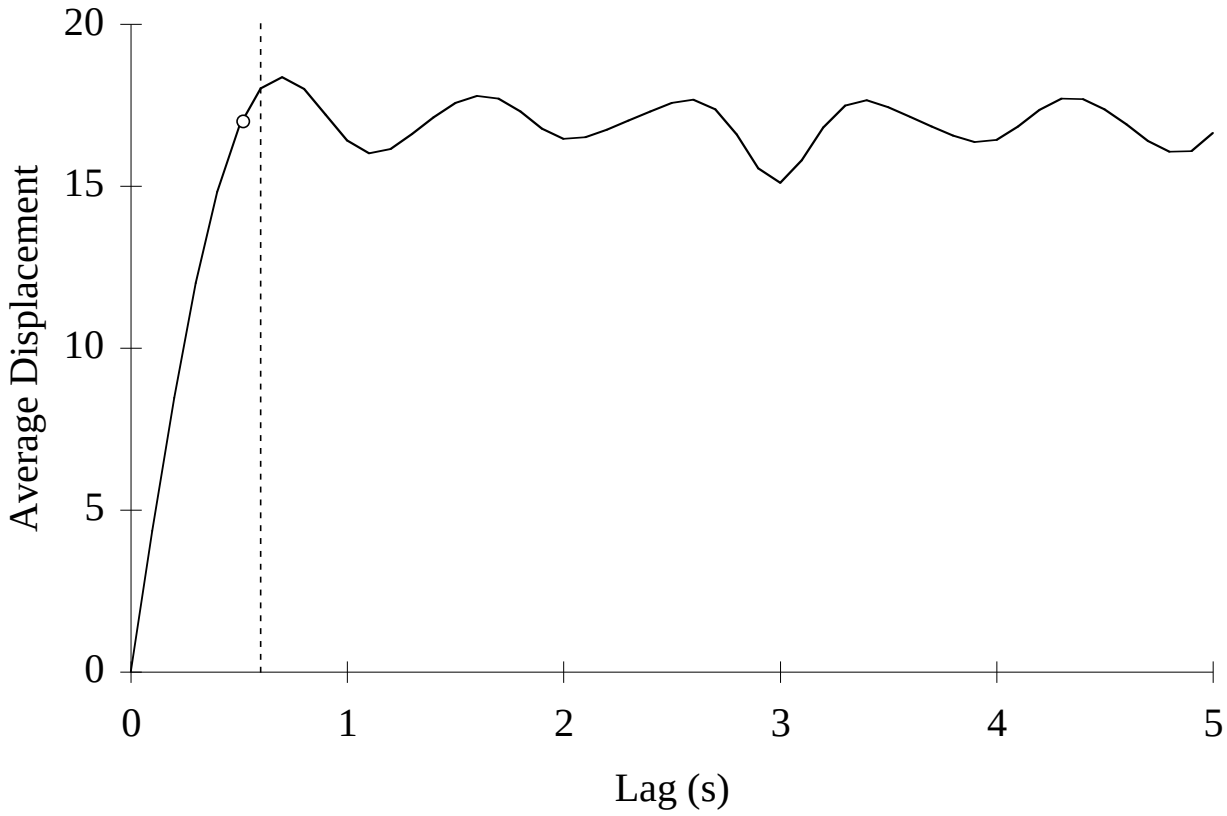
(FIGURE 3d)



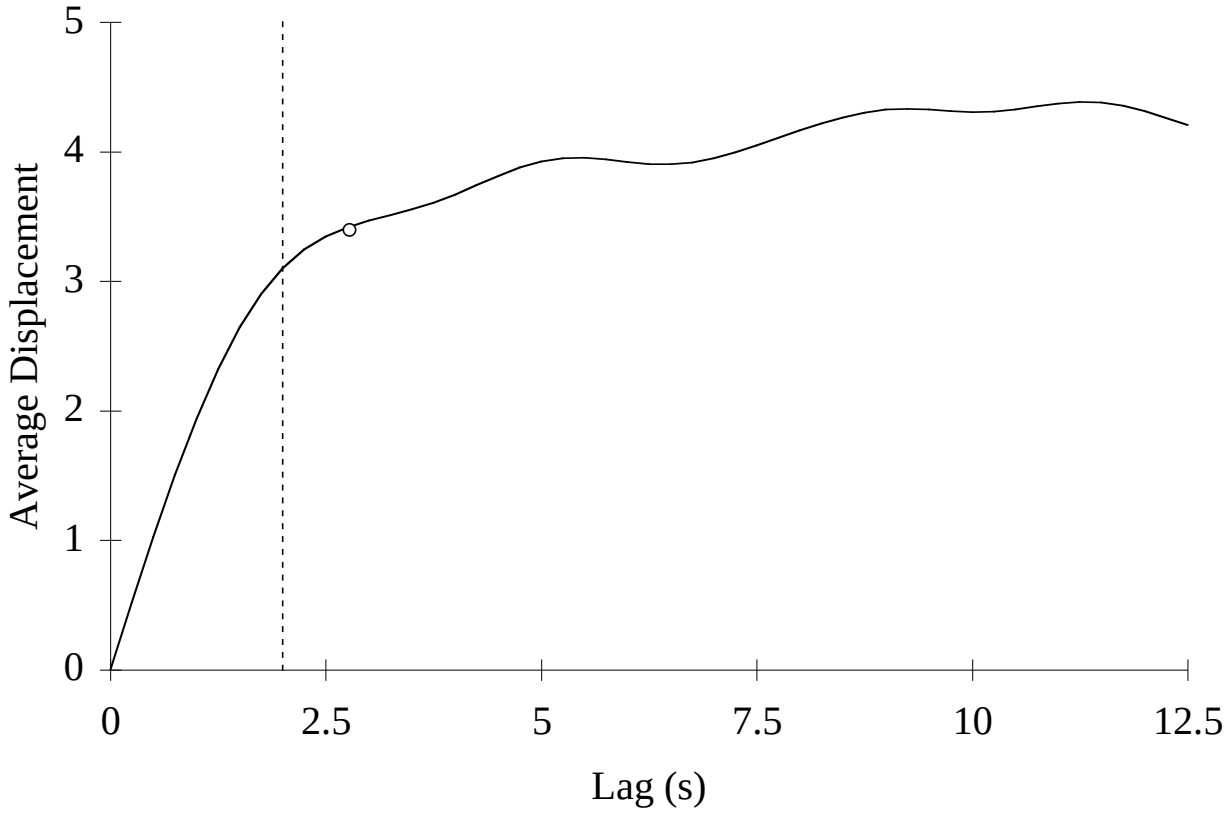
(FIGURE 4)



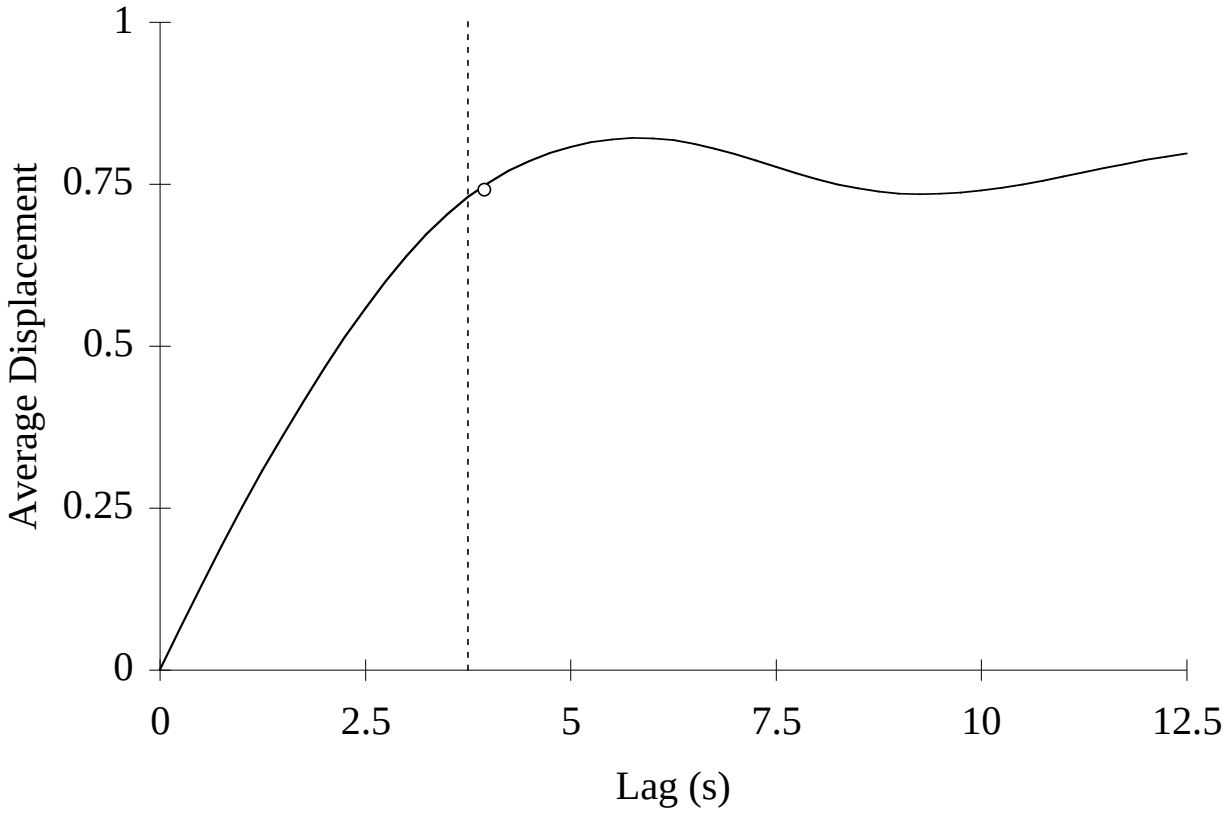
(FIGURE 5)



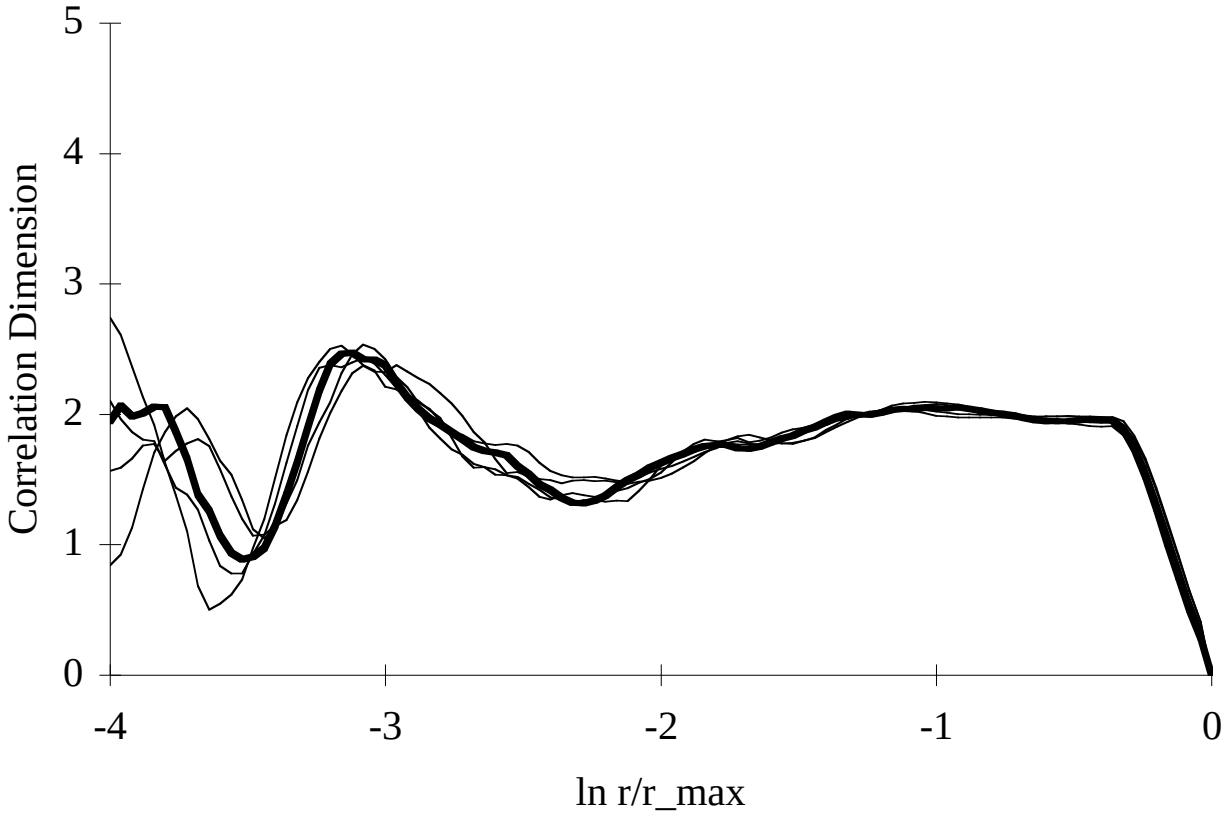
(FIGURE 6)



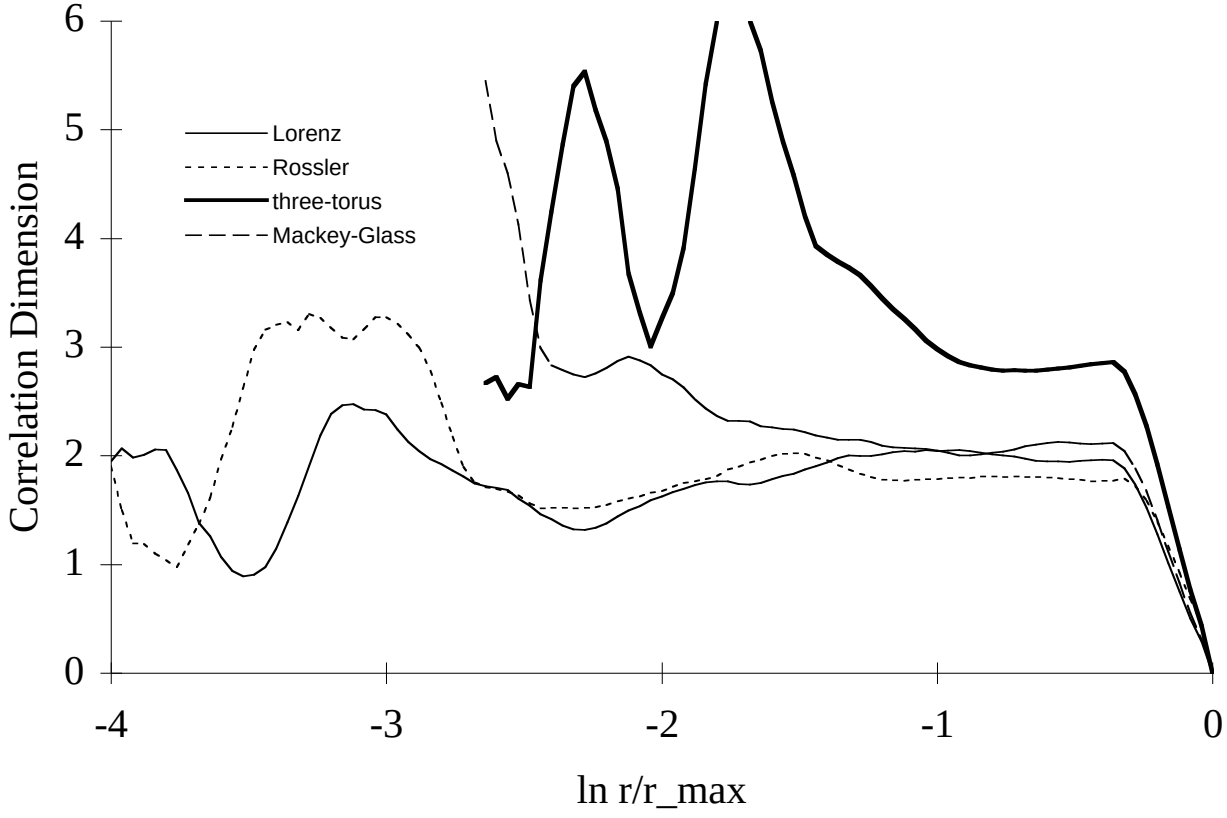
(FIGURE 7)



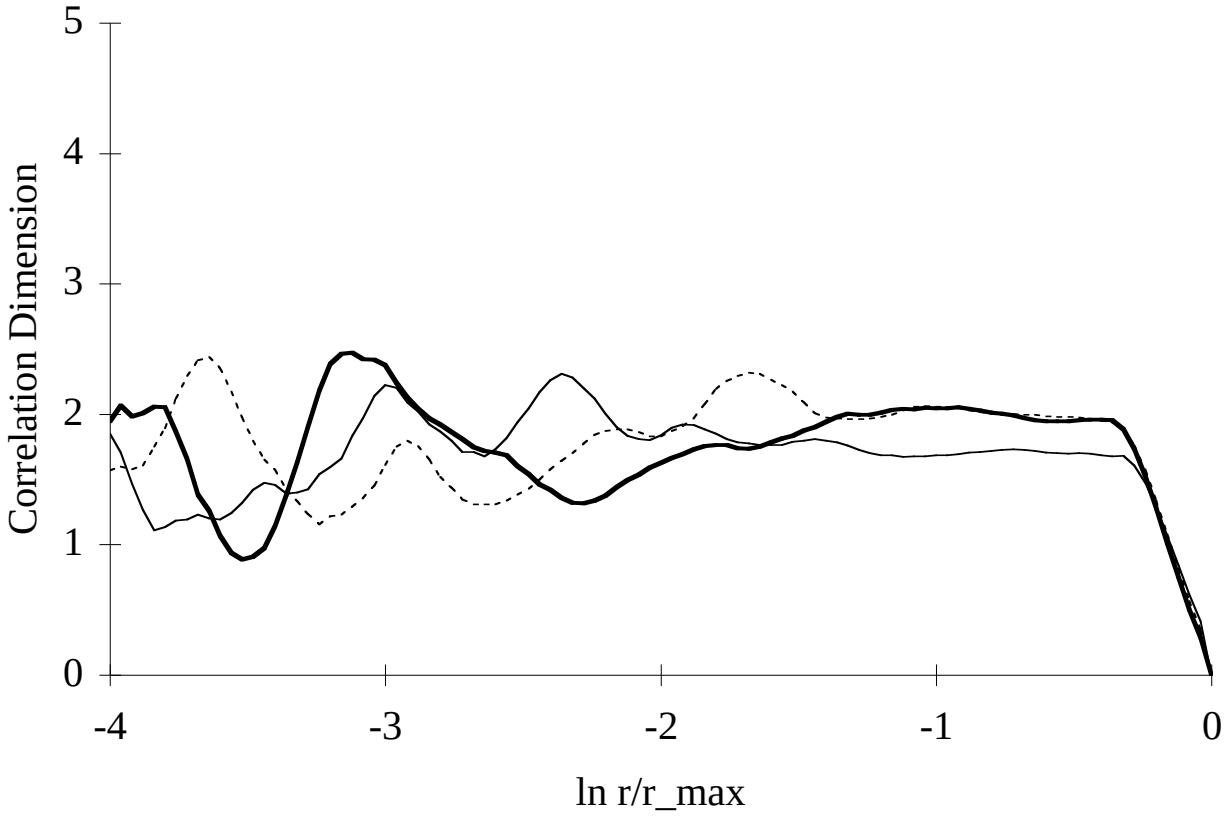
(FIGURE 8)



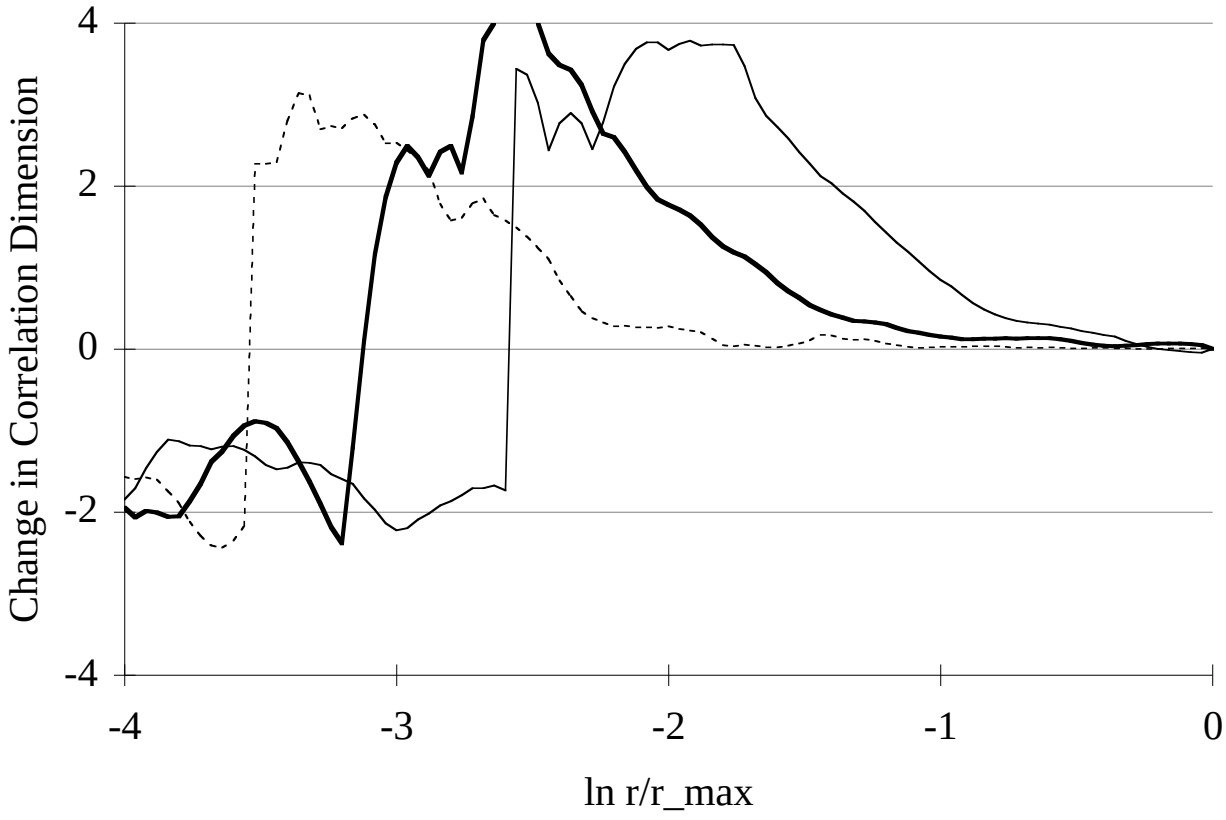
(FIGURE 9)


Correlation Dimension
6
5
4
3
2
1
0
-4
-3
-2
-1
0
ln r/r_max
Lorenz
Rossler
three-torus
Mackey-Glass

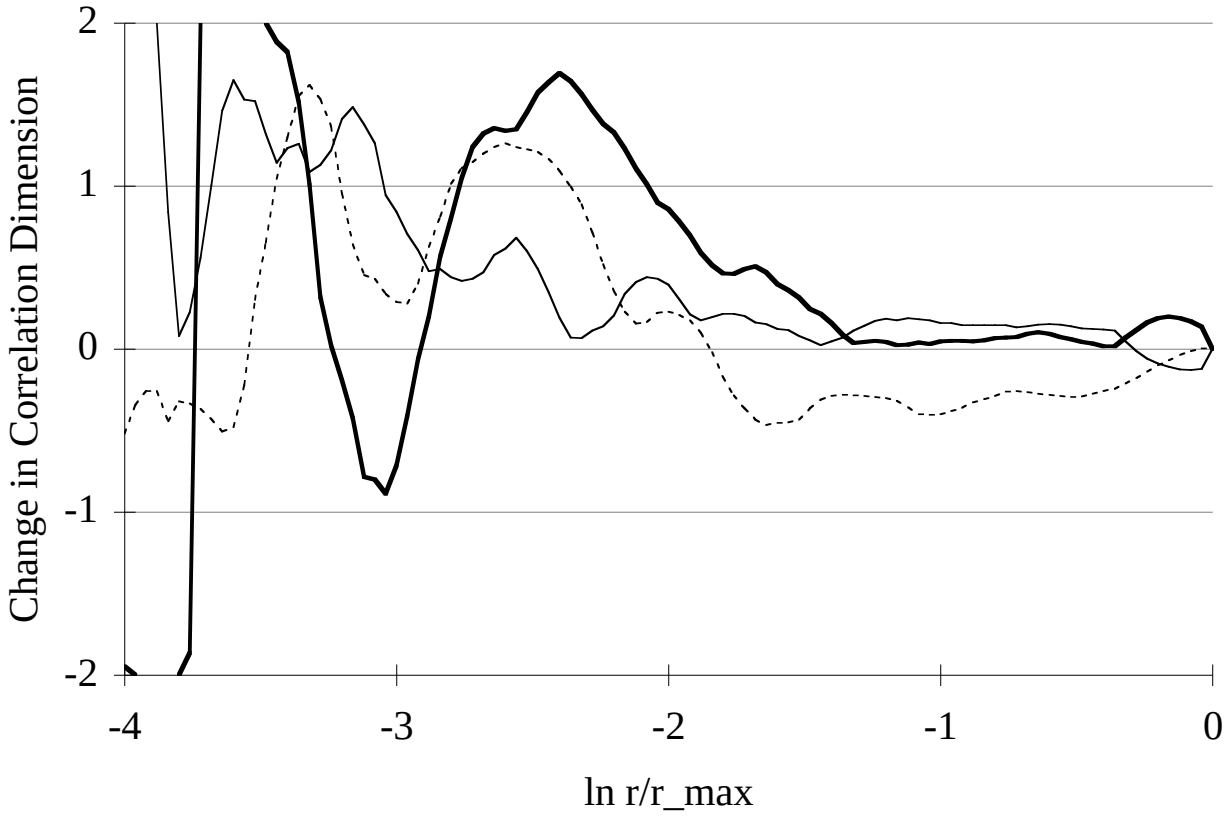
(FIGURE 10)



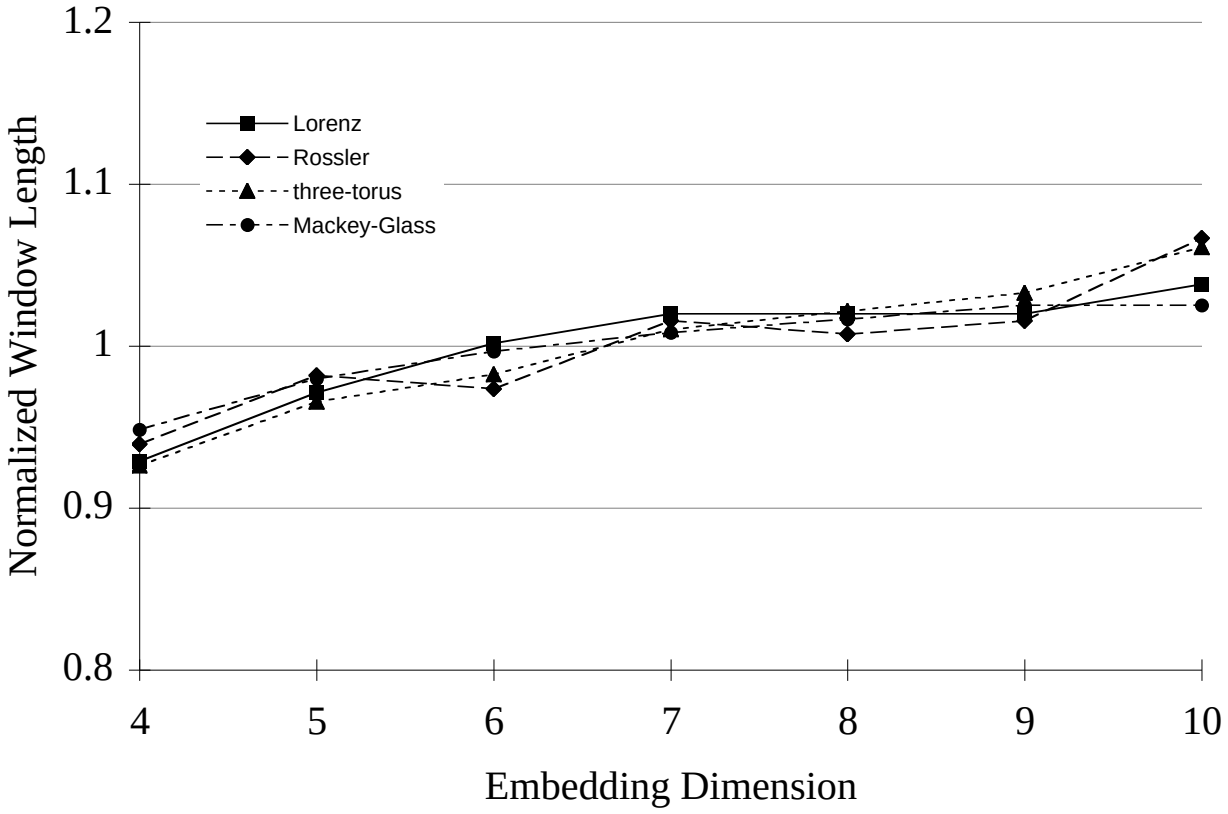
(FIGURE 11)



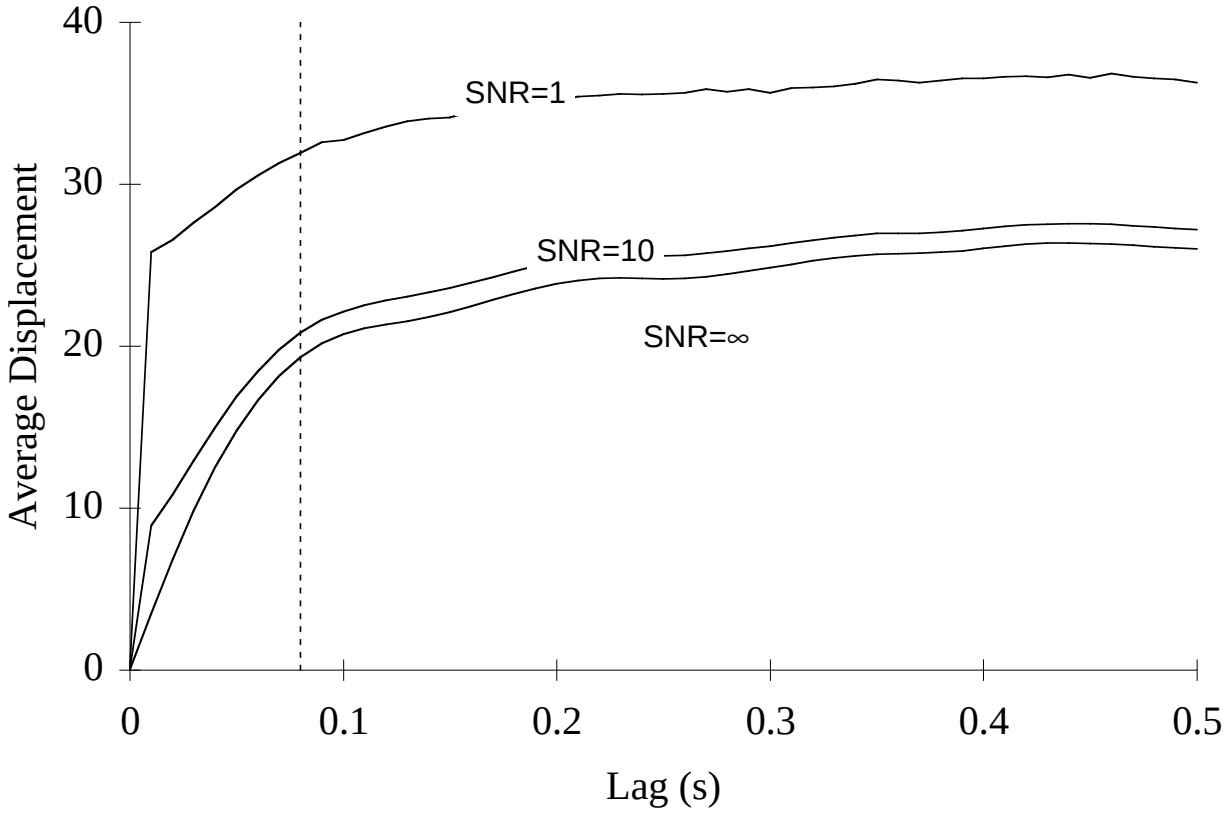
(FIGURE 12)



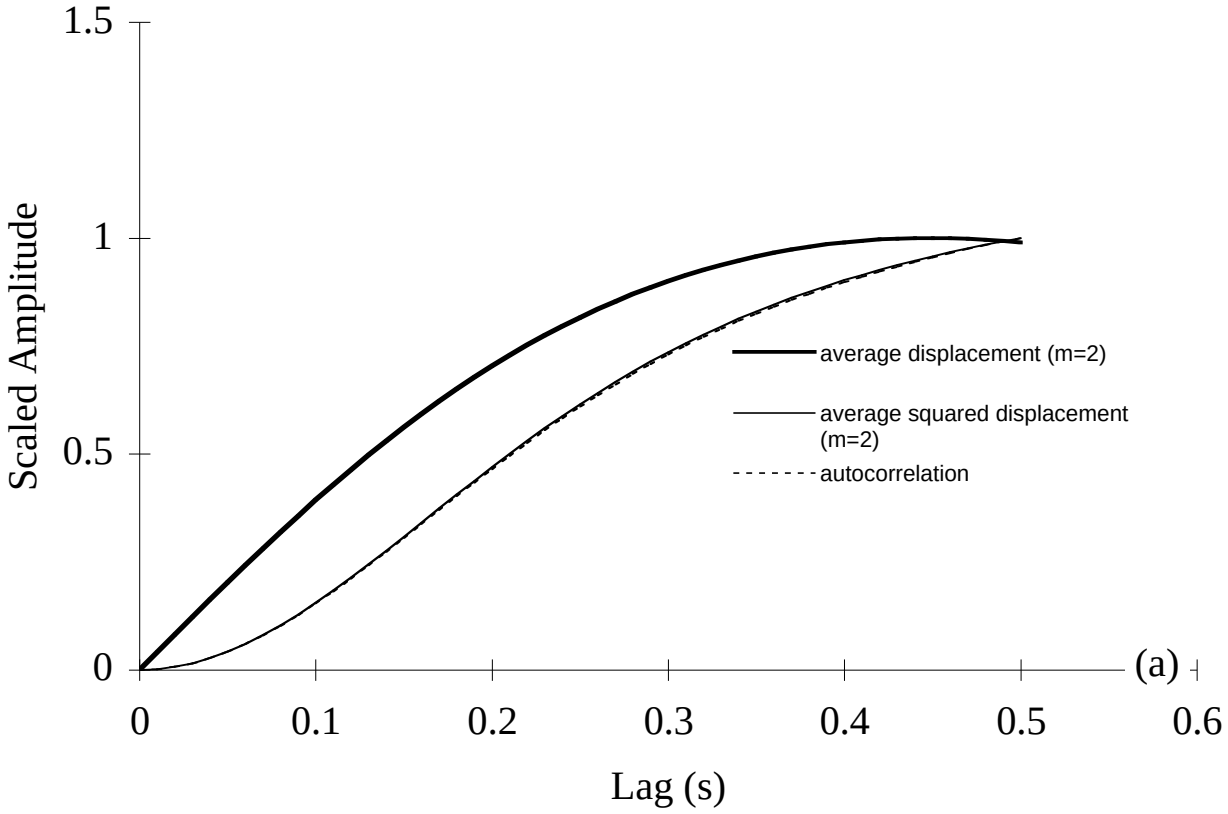
(FIGURE 13)


1.2
1.1
1
0.9
0.8
Normalized Window Length
Lorenz
Rossler
three-torus
Mackey-Glass
4
5
6
7
8
9
10
Embedding Dimension

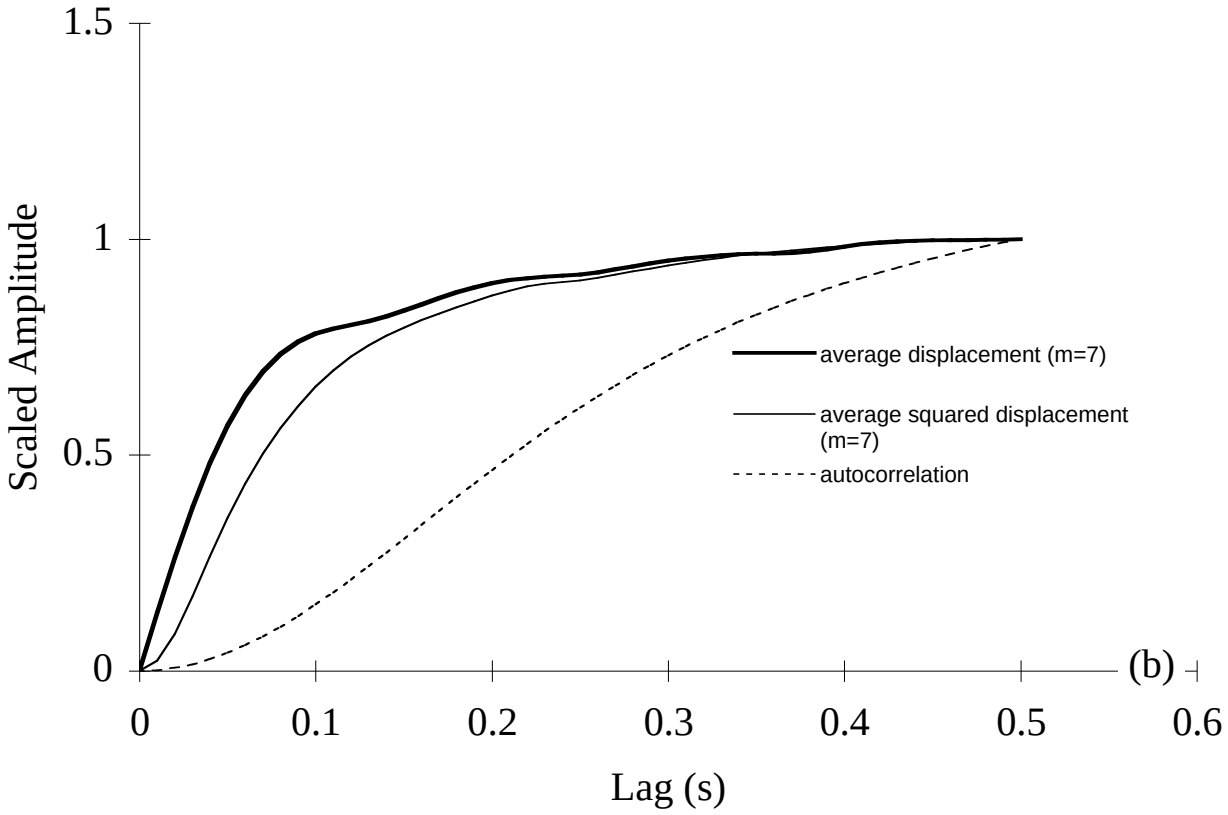
(FIGURE 14)


40
30
20
10
0
Average Displacement
SNR=1
SNR=10
SNR=∞
0
0.1
0.2
0.3
0.4
0.5
Lag (s)

(FIGURE 15)


1.5
1
0.5
0
Scaled Amplitude
average displacement (m=2)
average squared displacement (m=2)
autocorrelation
(a)
0
0.1
0.2
0.3
0.4
0.5
0.6
Lag (s)

(FIGURE 16a)


1.5
1
0.5
0
Scaled Amplitude
average displacement (m=7)
average squared displacement (m=7)
autocorrelation
(b)
0
0.1
0.2
0.3
0.4
0.5
0.6
Lag (s)

(FIGURE 16b)