Characterization of noisy symbolic time series

Christopher W. Kulp and Suzanne Smith
Department of Astronomy and Physics, Lycoming College, Williamsport, Pennsylvania 17701, USA
(Received 7 September 2010; revised manuscript received 23 November 2010; published 10 February 2011)

The 0-1 test for chaos is a recently developed time series characterization algorithm that can determine whether a system is chaotic or nonchaotic. While the 0-1 test was designed for deterministic series, in real-world measurement situations, noise levels may not be known and the 0-1 test may have difficulty distinguishing between chaos and randomness. In this paper, we couple the 0-1 test for chaos with a test for determinism and apply these tests to noisy symbolic series generated from various model systems. We find that the pairing of the 0-1 test with a test for determinism improves the ability to correctly distinguish between chaos and randomness from a noisy series. Furthermore, we explore the modes of failure for the 0-1 test and the test for determinism so that we can better understand the effectiveness of the two tests to handle various levels of noise. We find that while the tests can handle low noise and high noise situations, moderate levels of noise can lead to inconclusive results from the two tests.

DOI: 10.1103/PhysRevE.83.026201 PACS number(s): 05.45.Tp

I. INTRODUCTION

The goal of time series analysis is to understand the behavior of a system from which a time series has been measured. There are many time series analysis algorithms [1,2] that provide useful pieces of information about a system by analyzing the data measured from that system. In this paper, we are interested in the problem of characterization, where one attempts to determine whether the behavior of a system is regular (periodic or quasiperiodic), chaotic, or random by analyzing symbolic time series measured from a noisy system.

One of the standard tests for chaos is the estimation of the maximum Lyapunov exponent [3,4] for a system. The Lyapunov exponent measures the rate of divergence between two initially close trajectories in the system's phase space. If the maximum Lyapunov exponent is greater than zero, then initially close trajectories diverge exponentially in time and the system is chaotic (this is a reflection of the chaotic system's sensitive dependence on initial conditions). If the maximum Lyapunov exponent is less than zero, the system is not chaotic.

Recently, another test for chaos has been proposed. This test is called the 0-1 *Test for Chaos* [5,6] and is intended for use on deterministic systems. The 0-1 test will be described in detail in Sec. II. Unlike methods that determine the Lyapunov exponent, the 0-1 test does not depend on a reconstruction of the system's phase space. Furthermore, the test is easy to interpret since the test results in a 1 if the system is chaotic or a 0 if the system is not chaotic. Questions about the reliability of the 0-1 test were raised [7] and were addressed [8]. One of the comments made in Ref. [7] was that the 0-1 test might not be useful for data with little *a priori* knowledge and that such a problem could lead to misclassification of a time series. A misclassification could arise if the 0-1 test is applied to a time series that is stochastic (or suffers from very high levels of noise) because the 0-1 test will return a 1 if the system is not deterministic. Hence, a noisy system could be misclassified as chaotic. This kind of situation could be easily encountered when analyzing real-world data where little is known about the system, *a priori*. Furthermore, in our experience with noisy systems, the 0-1 test can also produce values between 0 and 1; we will discuss those results in Sec. III. To address the issues of misclassification, we propose to couple the 0-1 test with a test for determinism. The 0-1 test has been applied to experimental data [9]; however, in Ref. [9], the experimental system was well understood and the results were, therefore, easier to interpret. For example, the authors knew that the data were measured from a deterministic system; hence, a result of 1 from the 0-1 test implies the system was chaotic as opposed to random. In Ref. [9], no test for determinism was performed on the data. Such a test was not needed because the source of the data was well understood. Thus, in this paper we combine a test for determinism with the 0-1 test.

Our characterization algorithm will consist of two tests. A test for determinism must first be done on the time series to see if the 0-1 test is applicable. If the series fails the test for determinism, then the series is stochastic (or very noisy) and the 0-1 test cannot be applied to that series. There are many different tests for determinism in the literature such as Refs. [10–13]. Each test has its own way of detecting determinism. In Refs. [12] and [13], an embedding is used. The work done in Ref. [11] analyzes the singular-value spectrum of a trajectory matrix constructed from the time series. In Ref. [10], a symbolic time series is partitioned into nonoverlapping subsets and the test compares the symbol spectrum of the subsets. In this paper, we refer to the work done in Ref. [10] as the *symbol spectrum test*. In Sec. II, we will discuss the work in Ref. [10] as it is the test for determinism used in this paper. We chose the symbol spectrum test because it is computationally efficient and requires little *a priori* knowledge of the system such as the dimension of the system's phase space. Furthermore, because we are working with symbolic series, a test for determinism that uses symbolic analysis seems to be a natural fit for our problem. However, we have found that interpreting the results generated from the symbol spectrum test can be difficult, as we will show in Sec. III.

The goal of this paper is to explore the modes of failure for a characterization algorithm that combines the symbol spectrum test for determinism with the 0-1 test. We are combining a test for determinism with a test for chaos because of the 0-1 test's tendency to produce a 1 when a series is stochastic. Hence, when dealing with experimental time series measured from systems whose dynamics may be largely unknown, it will be important to determine whether the data are deterministic or stochastic (possibly due to high noise) before applying the 0-1 test to classify the dynamics. In order for these tests to be reliably applied to experimental data, we must understand under what conditions the tests misclassify time series as chaotic. Of course, if we want to understand the conditions in which the symbol spectrum test and the 0-1 test misclassify a time series, we must know something about the system generating the time series. In this paper, we used contrived time series whose dynamics are well known in order to test the modes of failure for the combined symbol spectrum/0-1 test algorithm. We generate our time series from model systems (the Logistic map, Duffing equation, and a complex continuous regular system) with known dynamics and noise levels. Next, we apply the symbol spectrum and 0-1 tests, and then use the results of the tests to classify the time series. Our approach in this paper is different from previous work in that we perform something similar to a blind test. As we will see in Sec. III, we look at the results of the symbol spectrum test and the 0-1 test and then we ask, according to the results of these tests, what conclusion should be made about the nature of the system. We then compare that result with the actual known dynamics to determine whether or not the symbol spectrum test and the 0-1 test misclassified the series. By using this kind of methodology on known systems, we hope to obtain a better understanding of how these tests perform on experimental time series in which little may be known about the system's dynamics *a priori*.

The structure of the paper is as follows. In Sec. II, we give a brief description of the symbol spectrum test and the 0-1 test and present references that provide more details for the interested reader. In Sec. III, we discuss the methodology and the results of our study. Finally, Sec. IV wraps up with some concluding remarks as well as comments on future work.

II. THE CHARACTERIZATION TESTS

In this section, we briefly outline the symbol spectrum test and the 0-1 test for chaos. For more details on each test, the interested reader is referred to the references in this section.

A. Symbol spectrum test

We chose to use the symbol spectrum test [10] as our test for determinism for two reasons. First, the test uses symbolic series. Since we are working with symbolized series, the test appeared to be a natural fit. Second, as we will show in this section, the test is generally easy to interpret, especially for noise-free periodic series. The symbol spectrum test is the first test in our characterization method because it should determine whether or not the 0-1 test for chaos is appropriate for the series we wish to characterize because the 0-1 test was designed for deterministic series and not random ones.

Consider a symbolized series of length N. For simplicity, we will use a binary alphabet for our series. However, such a restriction is not necessary for the symbol spectrum test. The symbol spectrum test begins by partitioning the series into disjoint subsets of length l. The choice of l will be discussed later and is dependent upon the next step. Next, we choose which level, L, of the symbol tree we wish to use. The choice of L groups the elements of each partition into "words" of length L. For example, consider the case $L = 2$. Then, for a binary series, there will be four possible words in the partition: 00, 01, 10, and 11. For a binary series, there will be 2^L words at the Lth level of the symbol tree. If the series has an alphabet of length, A (i.e., A different symbols in the series), then the number of words will be A^L for a given level, L, of the symbol tree. Our next step is to convert each word to base 10. While this step is not necessary, it is done to speed up the computation of the symbol spectrum test [14]. For each partition of length l, there will be $l - (L - 1)$ words (the second element of one word is the first element of the next). Figure 1 illustrates these steps for one partition of a binary series. The top line of Fig. 1 is a partition of some binary series. Note that $l = 12$ in Fig. 1. The middle line of Fig. 1 uses the second level of the symbol tree to break up the series into words of length $L = 2$. The length of the middle line is $l - (L - 1) = 11$. Finally, the third line shows the conversion to base 10.

{0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0}

↓

{01, 10, 01, 11, 10, 00, 00, 01, 11, 11, 10}

↓

{1, 2, 1, 3, 2, 0, 0, 1, 3, 3, 2}

FIG. 1. Demonstration of the transformation done to a binary series during the symbol spectrum test.

Next, we plot the number of times each base-10 "word" appears in the partition. This is the symbol spectrum for the partition. Finally, we plot the symbol spectrum for each partition on the same graph. If a series is deterministic, then the symbol spectrum from each partition will be similar. However, if the series is stochastic, there should be little to no similarities between the symbol spectra for each partition. Figure 2 illustrates some possible results of the symbol spectrum test for three different series, each with $N = 20\,000$, $l = 1000$, and $L = 6$. Figure 2(a) is the symbol spectrum test result of a periodic series generated from the Logistic map, $x_{n+1} = r x_n (x_n - 1)$ with $r = 3.55$. The thick line of the plot is actually the overlapping of 20 symbol spectra. Figure 2(b) is the symbol spectrum test result of a chaotic series generated from the Logistic map with $r = 3.91$. The regularity of the chaotic symbol spectra is not as clean as the periodic case, which, in our experience, can lead to some difficulties in the regime of high noise. We will discuss those issues in Sec. III. Figure 2(c) is the symbol spectrum test result for a random series with a uniform distribution. Notice that there is no common regularity shared between the symbol spectra.

At this point, we can address the question of how to choose the partition size l and the symbol tree level L. While there is no formula for choosing these parameters, our experience with the test has helped us develop a few guidelines. The authors of the test [10] use very long time series ($N = 50\,000$) with $l = 1000$ and $L = 6$. It is often not possible to work

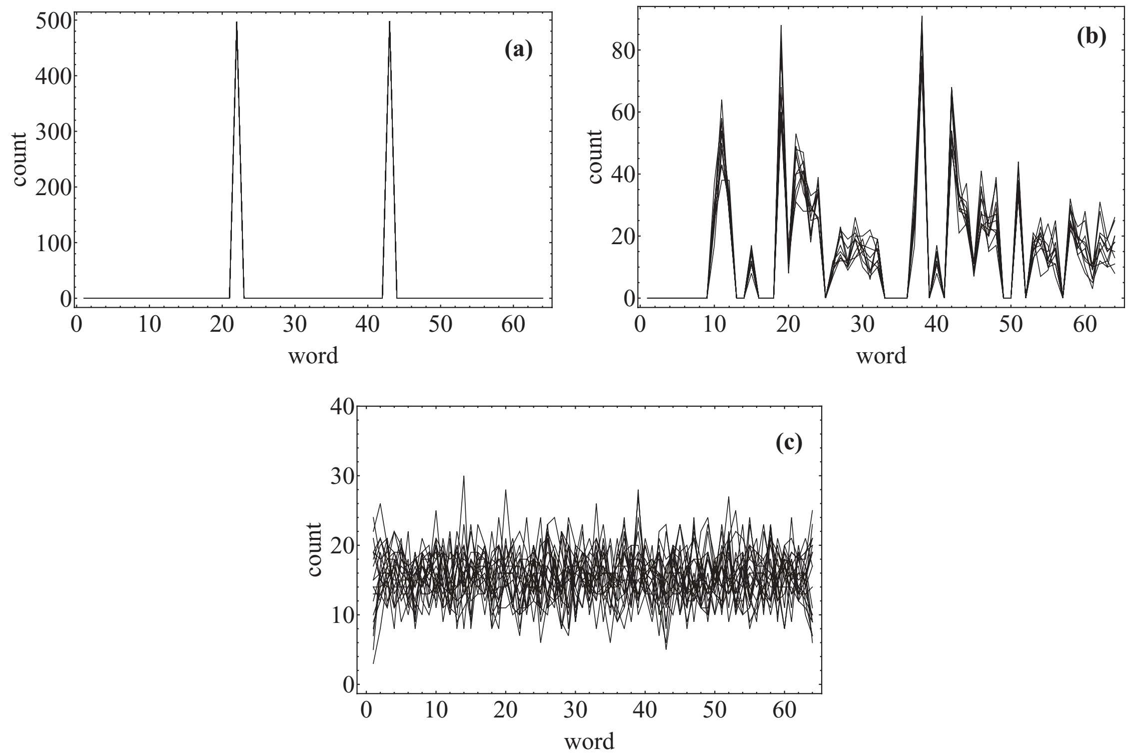

FIG. 2. Some results of the symbol spectrum test for the periodic Logistic map (a), the chaotic Logistic map (b), and a uniform random series (c). Notice that the spectra are more similar in (a) and (b) (more overlap) as compared to the spectra in (c) (less overlap).

with such long time series. Hence, we focus on generating a specific number of symbol spectra and let that choice dictate l and L. We have found that 20 spectra tend to be enough to determine whether the spectra are similar or not. If one uses too few spectra, then it is difficult to distinguish between a deterministic or a stochastic series, because it is difficult to establish a pattern (or lack there of) between the spectra. If one chooses too many partitions, then the partition length might be too short to produce a reliable spectrum. Once the partition length is determined, then one can choose L, which essentially determines the number of different words in the spectrum. We want to choose an L such that 2^L (or A^L for the general case) is significantly less than l. In this way we can obtain a good spectrum from each partition. If 2^L is too close to l, we find it difficult to establish patterns in the symbol spectra, especially for chaotic series. In this paper, we use $N = 10\,000$ for each series with $l = 500$ and $L = 5$. If a series is found to be deterministic when using the symbol spectrum test, then we can get meaningful results from the 0-1 test that will tell us if the series is chaotic or not.

B. 0-1 test for chaos

The 0-1 test for chaos first appeared in Ref. [5]. As mentioned previously, the 0-1 test is designed to distinguish chaotic behavior from regular (periodic or quasiperiodic) behavior in deterministic systems. The test results in a 1 if the system is chaotic and a 0 if the system is regular. Theoretical justification was given to the 0-1 test in Ref. [15]. An implementation guide for the 0-1 test was developed [6] that has been very useful in the work presented in this paper. For reference, we provide a brief outline of the 0-1 test below. The material for this outline comes from the work in Ref. [6]. For more details, the interested reader is directed to Ref. [6], and references therein.

Consider a time series of length N. The 0-1 test begins with a computation of two variables, $p_c(n)$ and $q_c(n)$:

$$p_c(n) = \sum_{j=1}^{n} \phi(j)\cos(jc),$$
$$q_c(n) = \sum_{j=1}^{n} \phi(j)\sin(jc), \tag{1}$$

where $\phi(j)$ is an observable constructed from the time series and $c \in (0,\pi)$ is randomly chosen. In this paper, we will choose $\phi(j)$ to be the jth value of the time series. The quantities $p_c(n)$ and $q_c(n)$ are referred to as "translation variables" in Ref. [6] and it is the behavior of these translation variables (either diffusive or nondiffusive) that determines whether the dynamics of the system are regular or chaotic.

To determine the behavior the translation variables, the test uses their mean square displacement

$$M_c(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} [p_c(j+n) - p_c(j)]^2 + [q_c(j+n) - q_c(j)]^2. \tag{2}$$

The theory of the test developed in Ref. [15] shows that if the dynamics of the system is regular, then (2) is bounded in time. However, if the dynamics of the system is chaotic, then (2) scales linearly in time. In order to numerically handle the limit in (2), we must use $n \ll N$. Hence we calculate $M_c(n)$ for $n < n_{\mathrm{cut}} = N/10$, as recommended in Ref. [6].

As mentioned earlier, (2) has two possible behaviors; either it is bounded in time or it scales linearly in time. As explained in Ref. [6], as long as the autocorrelations of the series are absolutely summable, then for a given value of c, $M_c(n)$ takes the form

$$M_c(n) = V(c)n + V_{\mathrm{osc}}(c,n) + e(c,n), \tag{3}$$

where $e(c,n)$ is an error term [$e(c,n)/n \to 0$ as $n \to \infty$] and

$$V_{\mathrm{osc}}(c,n) = \langle\phi\rangle^2 \frac{1-\cos(nc)}{1-\cos(c)}. \tag{4}$$

The term $\langle\phi\rangle$ is the expectation value of the time series. Hence, ignoring the error term, $M_c(n)$ has the form of a cosine with a slope given by $V(c)$, which is constant for a given value of c.

To determine the scaling behavior of $M_c(n)$, we calculate its asymptotic growth rate K_c. We can find the asymptotic growth rate directly from $M_c(n)$ by using a linear regression of a log-log plot of $M_c(n)$. However, in Ref. [6] it is pointed out that we can get a better estimate of K_c if we subtract the oscillatory term from $M_c(n)$, creating a modified mean square displacement

$$D_c(n) = M_c(n) - V_{\mathrm{osc}}(c,n). \tag{5}$$

In Ref. [6], there is one more recommended correction to $D_c(n)$ for noisy series. There it was found that the 0-1 test gives better results with noisy data if a modified version of $D_c(n)$, denoted as $D_c^*(n)$, is used,

$$D_c^*(n) = D_c(n) - \alpha V_{\mathrm{damp}}(n), \tag{6}$$

where $V_{\mathrm{damp}}(n) = \langle\phi\rangle^2 \sin(\sqrt{2}n)$. In Ref. [6], it is stated that the $\sqrt{2}$ was chosen arbitrarily. In this paper, we use the same frequency in V_{damp}, $\sqrt{2}$, as in Ref. [6] since we are interested in testing their algorithm. In Ref. [6] it is stated that α controls the sensitivity of the test to weak noise and weak chaos. Hence, by increasing α, the 0-1 test loses the sensitivity to weak noise but also increases its robustness to noise. In Ref. [6], $\alpha = 2.5$ was used to characterize data from the Logistic map that had uniformly distributed additive noise with an amplitude of 0.1. In this paper, we use $\alpha = 10$. We chose to use a higher α because some of the noise levels with which we will be working are significantly higher and we are not working with weakly chaotic systems. We are effectively losing sensitivity to weak chaos but gaining sensitivity to noise. Our choice of α is the only change we made to the algorithm presented in Ref. [6].

Finally, we find the asymptotic growth rate K_c of the modified mean square displacement $D_c^*(n)$. While we could use a similar linear regression method as previously mentioned, the authors of the 0-1 test state that, in practice, using a correlation method to estimate K_c works better than the linear regression method. Hence, K_c is the correlation coefficient of the vectors,

$$\xi = (1,2,\ldots,n_{\mathrm{cut}}), \tag{7}$$

$$\Delta = D_c^*(1), D_c^*(2), \ldots, D_c^*(n_{\mathrm{cut}}). \tag{8}$$

Recall that $M_c(n)$ is bounded for regular dynamics and scales linearly in time for chaotic dynamics; hence, $K_c = 1$ for chaotic dynamics and $K_c = 0$ for nonchaotic dynamics.

Note the subscripts of c in the preceding equations. Each of the foregoing quantities depends on the chosen value of c. As mentioned in Ref. [6], while most values of c give $K_c = 0$ for periodic systems, there are some isolated values of c for which a resonance occurs in the test that leads to a periodic system producing $K_c = 1$ when using the correlation method. Hence, we need to compute K_c for many values of c in order to get an accurate characterization. One way to avoid some, but not all, of the resonant values of c is to restrict the choice of c to $c \in (\pi/5, 4\pi/5)$ [6]. In Ref. [16], it was found that 100 different values of c chosen at random is sufficient to characterize a system. The final result for the characterization is K, which is the median of the K_c's. The choice of the median is used to further suppress any resonances that may not have been avoided by the restricted choice of c. Using the median becomes important if one does not restrict the values of c as recommended and/or if one uses the regression method in which resonances can lead to values of $K_c = 2$. There are other ways K can be chosen from the spectrum of K_c values; however, for the work described in this paper, we chose to follow what is done in Ref. [6] since that is the algorithm we are testing in conjunction with the symbol spectrum test. While the median value, K, of K_c is used for characterization, it is often helpful to inspect the "spectrum" of K_c. A sample spectrum of a random binary series is given in Fig. 3(a). Note that, in this paper, we will use the notation that K_c is the asymptotic growth rate for a particular value of c while K is the median of the K_c's.

It was commented in Ref. [8] that the 0-1 test was not intended for stochastic series. We include Fig. 3 for two reasons. First, it serves as an example spectrum to illustrate the results of the 0-1 test. Second, it also demonstrates the necessity of pairing a test for determinism with the 0-1 test. If one were to calculate a spectrum similar to the one shown in Fig. 3(a) from a system whose dynamics were not known *a priori*, one might assume the series was measured from a chaotic system. However, this series fails the test for determinism, as shown in Fig. 3(b). Because our series fails the test for determinism, we can conclude either that the series must have a high level of noise or that it is random.

To further illustrate the need for understanding the modes of failure for the symbol spectrum/0-1 test algorithm, in Fig. 4 we present the results of the algorithm applied to apparent magnitude data taken from the variable star S Persei. S Persei is a variable star in the constellation Perseus. The magnitude data shown in the top left graph of Fig. 4 are available at the American Association of Variable Star Observers (AAVSO) website [17]. We took 20 000 measurements from the AAVSO database for S Persei and symbolized the magnitude data into a binary series by using the series mean as the threshold for symbolization. We then ran the symbol spectrum test by using $L = 6$ and a partition size of 1000. The results of the symbol spectrum test appear in the top right of Fig. 4. While the symbol spectrum appears irregular, there are some regular tall peaks. Note that the whole range of the vertical axis is not shown in Fig. 4 because there are very tall peaks that exist at either end of the graph that wash out the detail in the middle of the plot. The 0-1 test results are shown in the bottom graph of Fig. 4. Note the K_c values are at or very close to 1 for all values of c shown. The

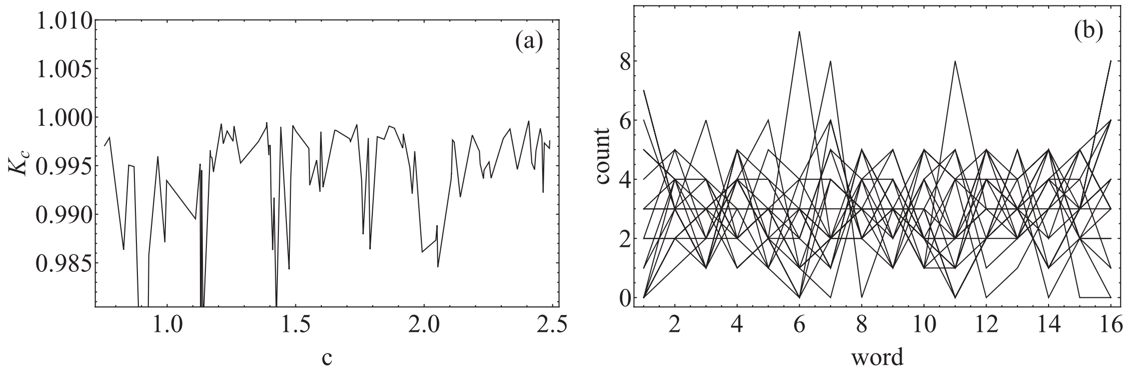

FIG. 3. Test results for a random binary series with a uniform distribution and a length of 1000. The left graph is a K_c spectrum for 100 random values of $c \in (\pi/5, 4\pi/5)$. The right graph is the result of a symbol spectrum test for the same series. Here we use a partition size of 50 and a word length of 4 (hence values run from 0 to $2^4 = 16$).

median of the K_c spectrum is 1. The 0-1 test implies that the star's variability is chaotic. However, the interpretation of the symbol spectrum test is difficult. While irregularity certainly exists among the symbol spectra, the regular large peaks hint toward determinism. While S Persei's variability is known to be semiregular and consists of several frequencies [18], can we say that these results suggest that the star's variability is chaotic? We know that there is measurement error (noise) and irregularity in sampling. How do these affect the symbol spectrum/0-1 test algorithm? The S Persei example motivates the goal of this paper. We explore the modes of failure of the symbol spectrum/0-1 test algorithm for systems whose dynamics and noise characteristics are known so that we can hopefully better understand results such as those obtained for S Persei on real-world data. We should note that the effect of irregular sampling on these algorithms is still an open question.

III. RESULTS

In this section, we look at our results of applying the symbol spectrum test and the 0-1 test to three model systems: the Logistic map, the Duffing equation, and a complex continuous

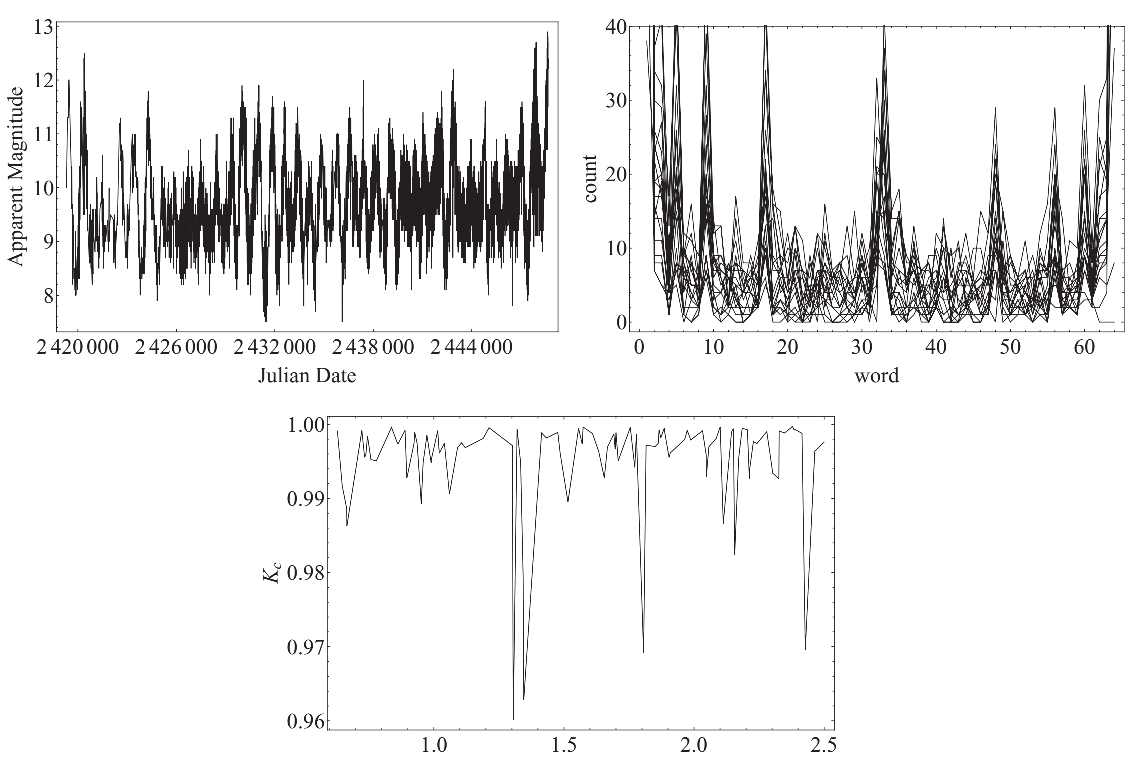

FIG. 4. Results of the symbol spectrum/0-1 test algorithm for apparent magnitude data measured from the variable star S Persei. The top left graph shows the time series before symbolization. The upper right graph shows the results of the symbol spectrum test (with $L = 6$ and a partition size of 1000). The bottom graph is the K_c spectrum for 100 random values of $c \in (\pi/5, 4\pi/5)$.

regular oscillation. We will discuss each system in detail in the following subsections. We chose the Logistic map and the Duffing equation because these systems can display both chaotic and periodic behavior. The complex continuous regular oscillation is the superposition of two sine waves with incommensurate frequencies and is a different type of behavior from the periodic oscillations of the Duffing equation, which contains only one frequency. However, the 0-1 test should produce a 0 for a complex continuous regular system. Our choice of systems includes an example of a discrete system (Logistic map) and two continuous systems (Duffing equation and complex continuous regular). In Ref. [6], it was shown that with continuous systems, the sampling rate is important. If the sampling rate is too high, a chaotic system can produce a false 0. The details of this are explained in Ref. [6].

We begin by generating a noisy series from a model system. While each model system is different, we generated symbolic time series from each system in a similar way. Here, we will outline how a noisy symbolized series is generated from the model systems. Let x_i be the ith element of a time series generated from one of our model systems. Next, we add noise to the series,

$$X_i = x_i + \eta\xi_i, \tag{9}$$

where ξ_i are independent and identically distributed random variables from either a uniform distribution on $[-1,1]$ or a normal distribution with a mean of zero and a standard deviation of 1.0. In this paper, we study both a uniform noise distribution with varying η's and a normal noise distribution with varying η's. After a noisy series is generated, we symbolize the series. Symbolization assigns a symbol to each element of the series. In this paper, we work with binary series; hence, the symbols are either 0 or 1. Symbolization can be done for a variety of reasons. One of the most common reasons to symbolize is to increase the speed of calculations done on the series. Because a symbolic series has fewer unique elements, calculations such as probability estimates are much more computationally efficient. Another reason symbolization is done is for noise suppression. We will see the effects of noise suppression later in this section. We symbolize each series by looking at the range of values contained in the unsymbolized noisy time series. We then choose our threshold for symbolization to be the mean of the range. This method was chosen because it is easily generalized to symbolic series with larger alphabets, in which case we break the range of the series up into equal parts and assign a symbol to each part. The characterization of noisy symbolic series with longer alphabets is part of our planned future work.

After the series is generated and symbolized, we apply the classification algorithms. First, we apply the symbol spectrum test to determine whether or not the series is deterministic. Regardless of the results of the test for determinism, we next apply the 0-1 test to the series. If the series is found not to be deterministic (i.e., the series "fails" the test for determinism) we still apply the 0-1 test even though it is not designed for nondeterministic series. We apply the 0-1 test to series that fail the test for determinism to demonstrate the possible results the 0-1 test can provide for random series.

TABLE I. Results of the symbol spectrum and 0-1 tests for the Logistic map with uniform noise. An asterisk denotes an apparent pass that is different from the $\eta = 0$ case.

	Periodic ($r = 3.55$)		Chaotic ($r = 3.91$)	
η	Symbol spectrum	K	Symbol spectrum	K
0	pass	−0.012	pass	1
0.01	pass	−0.012	pass	1
0.05	pass	−0.012	pass	1
0.10	pass*	0.79	pass	1
0.50	fail	0.99	fail	1
1.0	fail	1	fail	1

A. Logistic map

The Logistic map

$$x_{n+1} = rx_n(x_n - 1) \tag{10}$$

was chosen for two reasons. First, it is the system most frequently used in much of the 0-1 test literature. Second, it is a well-understood example of a discrete system that displays both chaotic and periodic behaviors. In this paper, we use the same parameter values as those used in Ref. [6], $r = 3.55$ for periodic behavior and $r = 3.91$ for chaotic behavior. In each case, the initial condition for the series is $x_1 = 0.01$ and we begin the time series after 20 000 iterations to allow transients to decay.

The results of the symbol spectrum test and the 0-1 test for the Logistic map with noise having a uniform distribution and a normal distribution are displayed in Tables I and II, respectively.

Recall that our approach is to assume that we know nothing about the dynamics *a priori*. We ask the question, how accurately do the symbol spectrum test and the 0-1 test characterize the series if the dynamics of the system is unknown? Let us begin by discussing the data summarized in Table I, which contains data for the case in which a uniform noise distribution was used. The detailed results from Table I appear in Figs. 5 and 6 for the cases of $r = 3.55$ and $r = 3.91$, respectively. For each figure, the first column contains the unsymbolized time series, the second column contains the results of the symbol spectrum test, and the third column contains the results of the 0-1 test in the form of the K_c spectrum for 100 randomly chosen values of $c \in (\pi/5, 4\pi/5)$.

TABLE II. Results of the symbol spectrum and 0-1 tests for the Logistic map with normally distributed noise. An asterisk denotes an apparent pass that is different from the $\eta = 0$ case.

	Periodic ($r = 3.55$)		Chaotic ($r = 3.91$)	
η	Symbol spectrum	K	Symbol spectrum	K
0	pass	−0.012	pass	1.0
0.01	pass	−0.012	pass	1.0
0.05	pass*	0.73	pass	1.0
0.10	pass*	0.98	pass	1.0
0.50	fail	1.0	fail	1.0
1.0	fail	1.0	fail	1.0

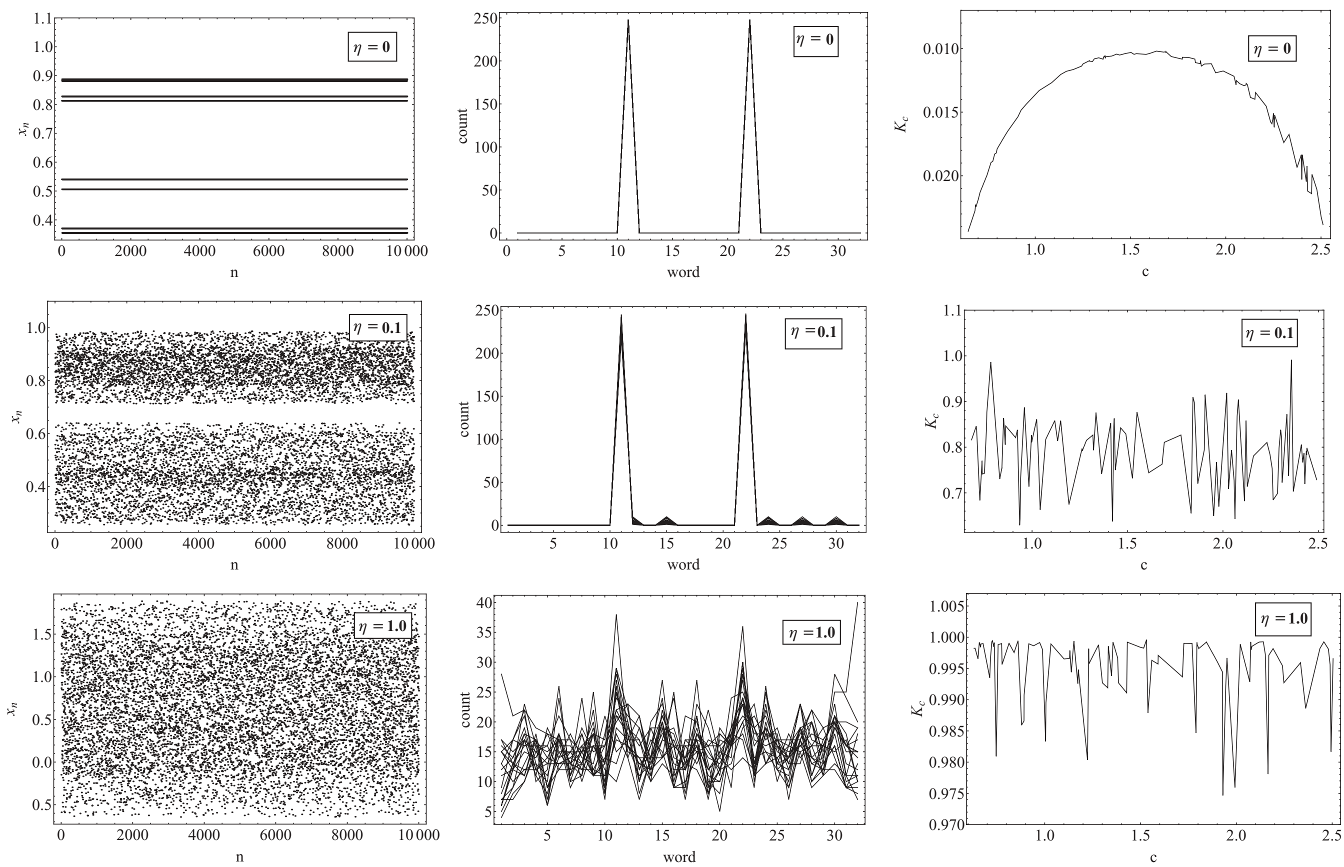

FIG. 5. Results of the symbol spectrum test and the 0-1 test for the Logistic map with $r = 3.55$. The first column contains the unsymbolized time series and the second and third columns contain the results of the symbol spectrum test and the 0-1 test, respectively.

For the case $r = 3.55$, the Logistic map is periodic and we expect the 0-1 test to give a result of 0. The top line of Fig. 5 shows the results for the noise-free periodic Logistic map. Note that in the time series for $\eta = 0$ there are eight different values in the series; the top two overlap due to the scale of the graph in Fig. 5. The symbol spectra (top middle graph of Fig. 5) all lie on top of one another; hence, the series is deterministic. Furthermore, the K_c spectrum (upper right graph in Fig. 5) is consistently around 0 for all values of c.

In Table I, we see that for $r = 3.55$ we have good performance for our test for low and high noise. For low noise, $\eta = 0.01$ and 0.05, the series manages to pass the symbol spectrum test (symbol spectra resemble the $\eta = 0$ case) and the 0-1 test produces a 0. The low noise results are not surprising because it has been reported that the 0-1 test works well with low-noise series [6]. In the case of high noise, $\eta = 0.5$ and 1.0, the symbol spectrum test shows that the time series is stochastic and even though we get a 1 for the 0-1 test, the symbol spectrum test informs us that this series either is random or has a high amount of noise. The bottom row of Fig. 5 contains some of the details for the case of $\eta = 1.0$. Notice that if we had not done the symbol spectrum test, we could have falsely concluded that the series for the cases of $r = 3.55$ and $\eta = 0.5, 1.0$ are chaotic. The case $\eta = 0.1$ (middle of Fig. 5) shows some interesting results. The symbol spectrum result for $\eta = 0.1$ is similar to when $\eta = 0$ with the addition of small "bumps" in the spectrum. We denote this behavior as "pass*" in Table I. In this paper, we say that the result of the symbol spectrum test is "pass*" if the symbol spectrum is different from the $\eta = 0$ case but could still pass as a deterministic symbol spectrum. We found "pass*" results to be difficult to characterize. For example, in the case of $r = 3.55$ and $\eta = 0.1$, one may see this symbol spectrum and, with no *a priori* knowledge of the system, conclude that the series is deterministic. However, inspection of the K_c spectrum suggests an inconclusive result. The median of the spectrum is 0.79 and the spectrum itself varies wildly between $K_c = 0.7$ and 0.9. This suggests that while the tests can handle low noise and clearly fail at high noise, there exists some "transition" region of noise where results can be inconclusive or possibly give an incorrect characterization.

Next, we will analyze the results for the chaotic Logistic map, $r = 3.91$, in Table I. The analysis for the periodic case was much easier to do than in the chaotic case, $r = 3.91$. For the periodic case, there are two values that can change, the result of the symbol spectrum and the 0-1 test result. However, for the chaotic Logistic map the 0-1 test will provide a 1 for chaotic series *and* for random series. Hence, interpretation of the symbol spectrum test takes on even more importance. As shown in Fig. 6, the time series and K_c spectra are very similar for $\eta = 0, 0.1$, and 1.0. In addition, the symbol spectrum for the $\eta = 0$ case is not as clean as the noise-free periodic Logistic map. We have found that this can lead to difficulties in interpreting the results of the symbol spectrum test. As

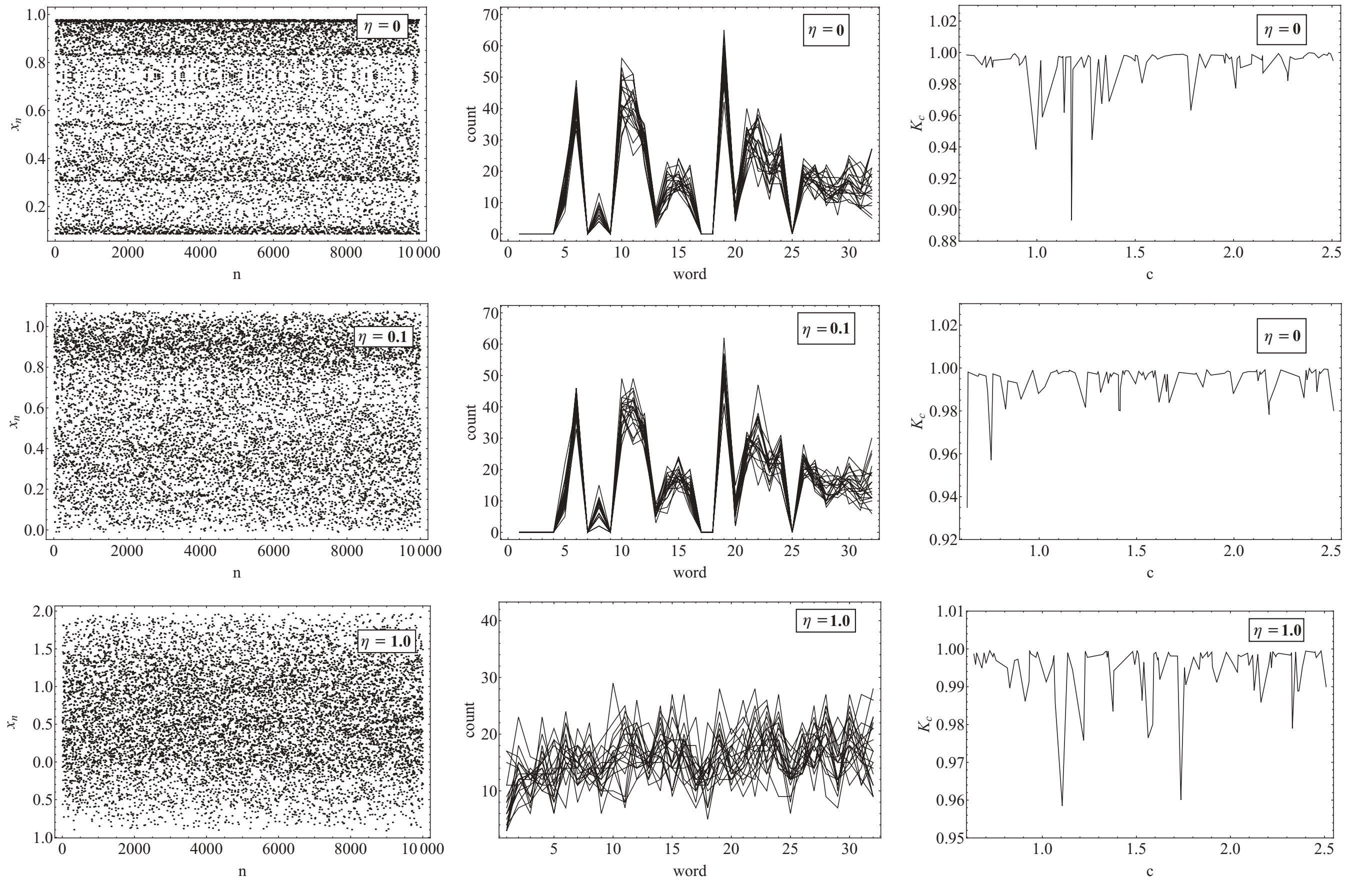

FIG. 6. Results of the symbol spectrum test and the 0-1 test for the Logistic map with $r = 3.91$.The first column contains the unsymbolized time series and the second and third columns contain the results of the symbol spectrum test and the 0-1 test, respectively.

in the periodic case, the data in Table I suggest that the test works well for low noise ($\eta = 0.01, 0.05$) and clearly fails in the case of high noise ($\eta = 0.5, 1.0$). In fact, the results even suggest that the tests can correctly characterize the case $\eta = 0.1$, which could not be done for the periodic series. This does raise one question: If $\eta = 0.1$ is a sufficiently high noise level to change the symbol spectrum result and the 0-1 test result for the periodic case, can we trust these results for the chaotic case? Notice that there are some minor differences between the symbol spectrum results for $\eta = 0$ and 0.1 in Fig. 6, just as there are minor differences between the two symbol spectra in Fig. 5. The minor difference in the periodic case leads to major differences in the K_c spectrum. In the chaotic case, the K_c spectra for $\eta = 0$ and 0.1 are not very different. If one were to use the knowledge gained from the periodic series, a conservative analysis would suggest that the results of the tests for $r = 3.91$, $\eta = 0.1$ are inconclusive. However, it is easy to see how no *a priori* knowledge of the system could lead one to characterize the series as chaotic because of the nature of the symbol spectrum.

Table II summarizes the results of the Logistic map with noise from a normal distribution. Again, we will begin with the periodic results. We see that for the normal noise distribution, we begin to get inconclusive results for lower values of η. Again, with the periodic data, we see a transition region between low noise and high noise where the 0-1 test results are difficult to interpret. If one was not aware of the previous results, it would be possible to mischaracterize the $\eta = 0.1$ results as chaotic data for the periodic case. Of course, in a situation in which one has no *a priori* knowledge of the system, such misclassifications would be easy to make; we avoid such pitfalls only because we know the noise levels and because of our experience with the other Logistic maps series analyzed in this paper. Likewise, in the case of the chaotic Logistic map with normally distributed noise, we do not see a clear transition region. Again, this is due to the difficulty of interpreting the symbol spectrum test for chaotic data and the fact that the 0-1 test will produce a 1 for random (or high noise) data. Such transitions probably also exist for the chaotic regime $r = 3.91$ but are more difficult to detect in the symbol spectrum for the same reasons as the flat noise case.

B. Duffing equation

Our next model system is the Duffing equation

$$m\ddot{x} + \delta\dot{x} + \omega x + \beta x^3 = A\cos(\Omega t + \phi), \qquad (11)$$

where $x = x(t)$. The value of the parameters m, δ, ω, β, A, Ω, and ϕ determine the dynamics of the Duffing equation. The Duffing equation displays periodic behavior (with a period of 2π) when $m = 1.0$, $\delta = 0.08$, $\omega = 0.0$, $\beta = 1.0$, $A = 0.2$, $\Omega = 1.0$, and $\phi = 0$ [19]. The Duffing equation displays chaotic behavior when $m = 1.0$, $\delta = 0.4$, $\omega = -1.0$, $\beta = 1.0$, $A = 0.4$, $\Omega = 1.0$, and $\phi = -\pi/4$ [19]. There are other parameter values [19] that produce periodic and chaotic behavior in the Duffing equation; however, the previously mentioned values are the ones used in this paper.

The Duffing equation was chosen as a model system because it is a well-understood continuous system. Continuous systems are different from discrete systems in that, for continuous systems, there is no obvious natural sampling time. For discrete systems, sampling typically occurs at each iteration. If a continuous signal is undersampled, one may not be able to reconstruct the original signal from the sampled series. Furthermore, it was shown in Ref. [6] that oversampling can lead to $K = 0$ for chaotic signals. To generate our time series, we use the technical computing software MATHEMATICA to numerically integrate the Duffing equation (with the necessary choice of parameters), using the **NDSolve** command with the initial conditions of $x(0) = -0.21$ and $\dot{x}(0) = 0.02$ for the periodic series and $x(0) = 1.0$ and $\dot{x}(0) = 0$ for the chaotic series [19]. After waiting 500 time units for transients to decay, a time series is produced at a fixed sampling rate until a time series of length 10 000 is achieved. For the periodic series, we use the Nyquist frequency, which leads to a sampling interval of approximately π. For the chaotic series, the choice of sampling time is not as straightforward and it is important to avoid oversampling. One method is simply "visual reconstruction." In other words, how large can the sampling time be and still have the sampled series look similar to the original solution of the Duffing equation? We choose a sampling time interval of 2 because it was small enough that the sampled series still had the appearance of the original series, but it was not so oversampled that the 0-1 test returned $K_c = 0$ instead of $K_c = 1$.

The results of the symbol spectrum test and the 0-1 test for the Duffing equation with noise having a uniform distribution and a normal distribution are displayed in Tables III and IV, respectively.

We will begin by discussing the results in Table III. Figures 7 and 8 illustrate some of the results found in Table III for the periodic and chaotic Duffing equation, respectively. Figures 7 and 8 are laid out in a similar way as Figs. 5 and 6 in that the first column is the unsymbolized time series, the second column is the symbol spectrum test result, and the third column is the K_c spectrum from the 0-1 test. Notice that the graphs in the first column of Figs. 7 and 8 have lines that join one element of time series to the next. These lines are added to guide the eye along the graph of the time series.

The data for the periodic series in Table III seem to suggest that our characterization method is robust to noise. In fact, the characterization method never seems to fail (i.e., produce a false positive for chaos). These results are less of a statement about the effectiveness of the symbol spectrum test and the 0-1 test and more of a demonstration of the ability of symbolization to suppress noise. The original time series for the periodic case has the form of a sine wave with an amplitude slightly greater than 1.2; see Fig. 9. The threshold for symbolization is approximately located at $x = 0$. In the periodic noise-free case, Fig. 9(a), the series is sampled consistently approximately at $x = \pm 0.6$; when we add noise,

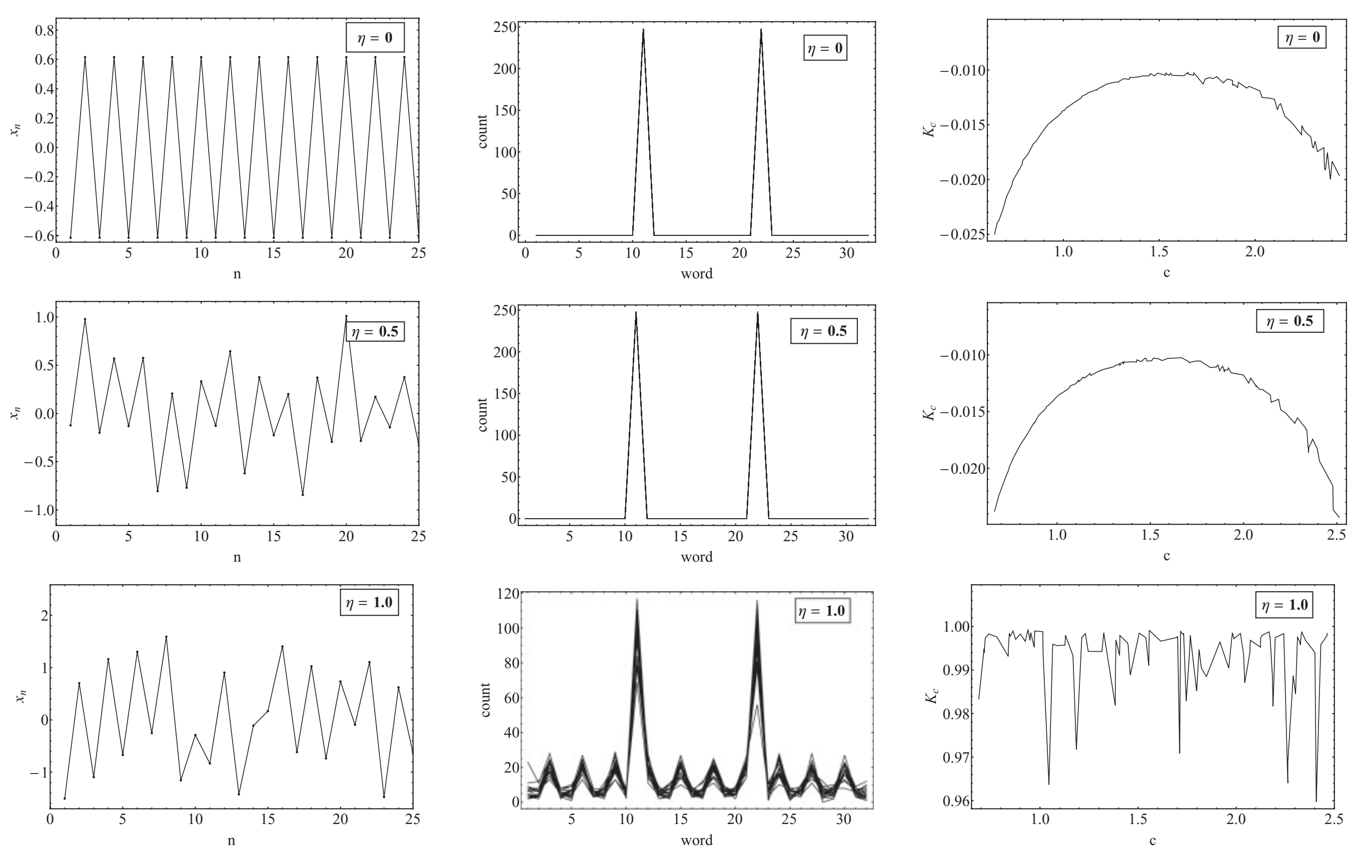

FIG. 7. Results of the symbol spectrum test and the 0-1 test for the periodic Duffing equation. The first column contains the unsymbolized time series and the second and third columns contain the results of the symbol spectrum test and the 0-1 test, respectively.

TABLE III. Results of the symbol spectrum and 0-1 tests for the Duffing equation with uniform noise.

η	Periodic		Chaotic	
	Symbol spectrum	K	Symbol spectrum	K
0	pass	−0.012	pass	0.98
0.01	pass	−0.012	pass	0.97
0.05	pass	−0.012	pass	0.98
0.10	pass	−0.013	pass	0.98
0.50	pass	−0.012	pass*	0.99
1.0	fail	1.0	fail	0.99

TABLE IV. Results of the symbol spectrum and 0-1 tests for the Duffing equation with normally distributed noise.

η	Periodic		Chaotic	
	Symbol spectrum	K	Symbol spectrum	K
0	pass	−0.012	pass	0.98
0.01	pass	−0.012	pass	0.98
0.05	pass	−0.012	pass	0.99
0.10	pass	−0.013	pass	0.98
0.50	fail	0.99	pass*	1.0
1.0	fail	1.0	fail	1.0

we see that a noise amplitude of $\eta = 0.5$ would not be enough to change a noise-free value from being greater than zero to being less than zero (or vice versa). Hence, noise is not causing the noisy symbolized series to be different from the noise-free symbolized series. Noise effects should not alter the symbolized series until $\eta \approx 0.6$. As we see in Table III, the symbol spectrum finds that the symbolic series are deterministic until $\eta = 1.0$. The symbol spectrum result for the $\eta = 1.0$ case is shown in Fig. 7. While the peaks in the symbol spectra occur at the same locations for the $\eta = 1.0$ case, we see that there is little overlap in the individual spectra; hence, this result is interpreted as a failure for the test for determinism. Note that we see a similar pattern for the periodic data in Table IV. The series does not fail the test for determinism until $\eta = 0.5$. A large standard deviation for the normal distribution can cause symbol changes between the noise-free and the noisy symbolic series. It is important to note here that if we had chosen a longer alphabet length for our symbolic series (i.e., more than two symbols in our series), the noise tolerance of symbolization would decrease.

As we found with the Logistic map, the chaotic Duffing equation is a bit more difficult to interpret than the periodic Duffing equation. Similar to the periodic case, the symbol spectrum test finds that the series are deterministic until $\eta = 0.5$ for the flat distribution. Furthermore, even though the spectra for the periodic and chaotic cases are different, their symbol spectrum test results are similar for the normal distribution. Like the Logistic map, the noise-free chaotic symbol spectra have less overlap than the noise-free periodic case. However, there is still a distinct difference in the symbol

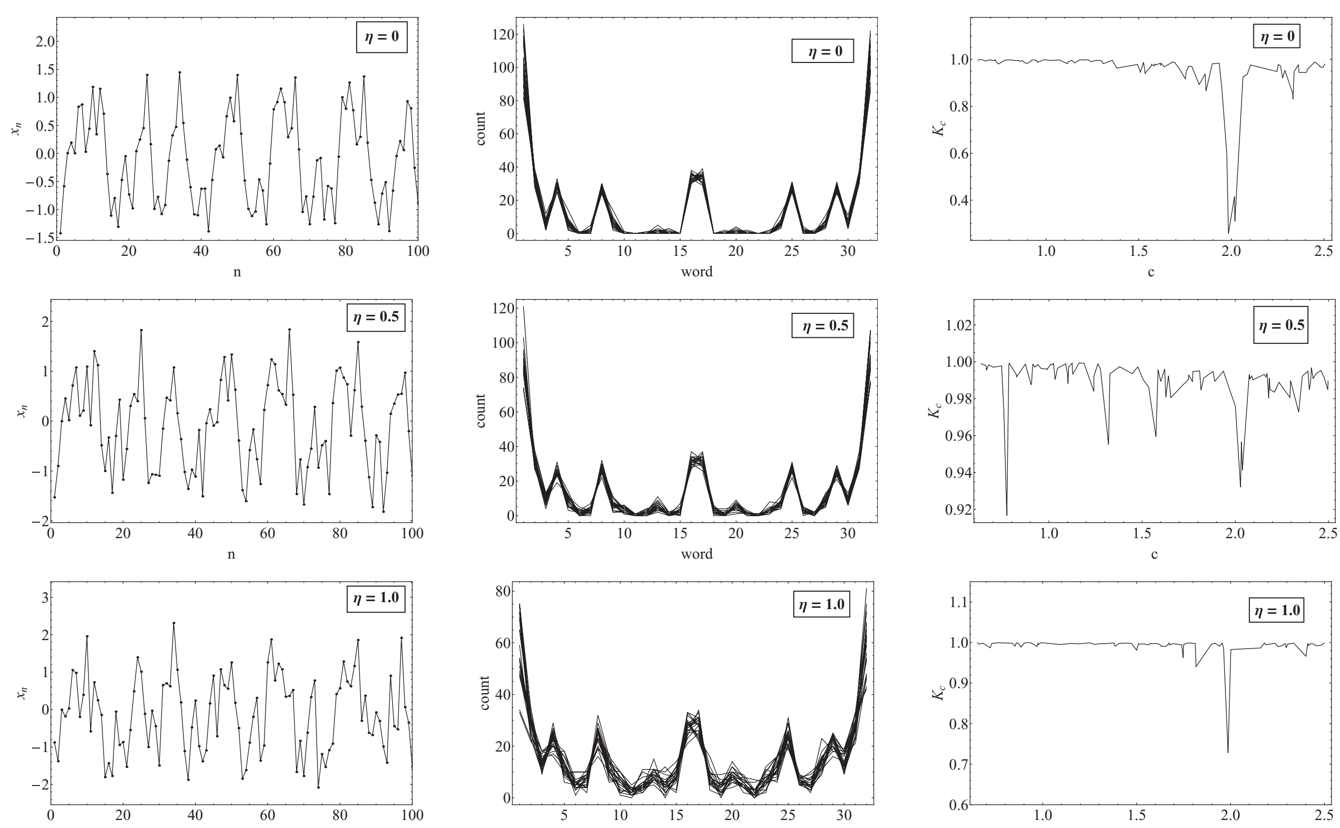

FIG. 8. Results of the symbol spectrum test and the 0-1 test for the chaotic Duffing equation.The first column contains the unsymbolized time series and the second and third columns contain the results of the symbol spectrum test and the 0-1 test, respectively.

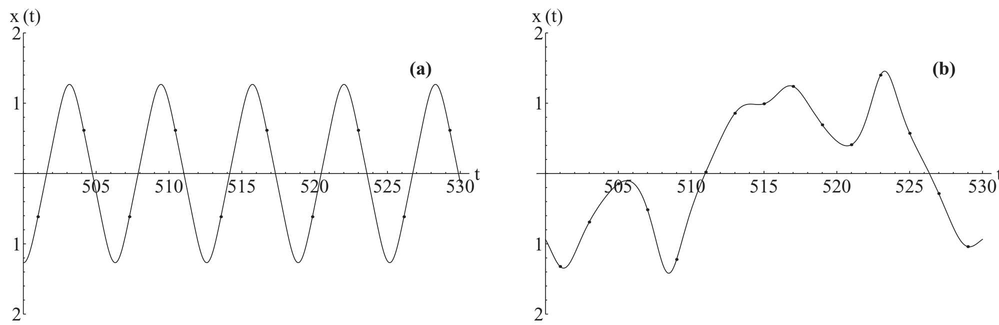

FIG. 9. Sampled solution of the periodic (a) and chaotic (b) solutions of the Duffing equation. For each case, $\eta = 0$. The sampling rate and parameter values are the same as discussed in the text.

spectra between the $\eta = 0$ and $\eta = 1.0$ case as shown in Fig. 8. There is little overlap between much of the spectra for the $\eta = 1.0$ case; hence, the series is stochastic (since we know the noise is high in this series, that result is expected). Similarly to the periodic case, the high noise tolerance for the symbol spectrum test is likely due to symbolization. Although Fig. 9(b) shows some points of the noise-free time series close to the threshold at $x \approx 0$, the vast majority of the points in the series are closer to $x = \pm 1.0$. Because of the regularity of the symbol spectra even in high noise series, we feel that it is safe to claim that the symbol spectrum test and the 0-1 test accurately determine that the series are chaotic up to $\eta = 0.5$ and that the $\eta = 1.0$ case is stochastic. Hence, we see that for the Duffing equation, symbolization is a powerful tool to suppress noise and still accurately captures the dynamics of the system.

C. Complex continuous regular system

Our final model system consists of a superposition of two sine waves with incommensurate frequencies, $\sqrt{2}$ and $\sqrt{3}$:

$$x(t) = \sin(2\pi\sqrt{2}t) + \sin(2\pi\sqrt{3}t). \tag{12}$$

In this paper, we refer to (12) as an example of a "complex continuous regular system," where "complex" means that there is more than one frequency contributing to the underlying dynamics and the term "regular" implies that the system is not chaotic. We use a sampling time of $\frac{1}{2\sqrt{3}}$, based on the Nyquist frequency. We chose this system because it is a continuous system that contains more than one frequency. Whereas our periodic Duffing equation was continuous, it consisted of only one frequency. Many continuous systems consist of multiple frequencies, so it makes sense to test our algorithm on a continuous system that has more than one frequency contributing to the underlying dynamics. While (12) is a simple example of a complex continuous regular series, it serves as a first step in testing the 0-1 algorithm on other systems that have larger numbers of frequencies contributing to the dynamics of the system. Figure 10 shows the graph of (12) and the first few elements of the noise-free time series (the dots).

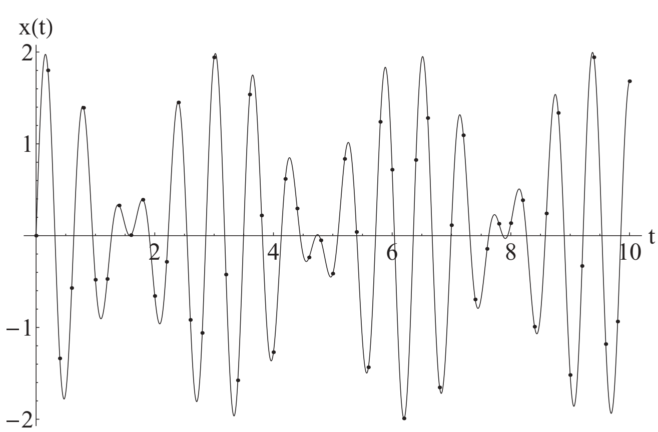

FIG. 10. Sampled noise-free complex continuous regular system (12).

The results of the symbol spectrum test and the 0-1 test for the complex continuous regular system with noise having a uniform distribution and a normal distribution are displayed in Tables V and VI, respectively.

The results for the uniform distribution, as summarized in Table V and illustrated in Fig. 11, are different from the uniform noise results for the Logistic map and Duffing equation. For the complex continuous regular system, the median of the K_c spectrum, K, slowly increases to 0.99 as η increases. This is unlike the Logistic map and Duffing equation cases, in which there is a sharp increase from a low value of K to a high value as η increases. Recall that in this paper, our interest is in being able to accurately characterize a time series with no *a priori* knowledge of the system. The complex continuous regular series produces many intermediate results for the 0-1 test that would be difficult to interpret for a series about which we know little. If we look at the original series,

TABLE V. Results of the symbol spectrum and 0-1 tests for the complex continuous regular system (12) with uniform noise. An asterisk denotes an apparent pass that is different from the $\eta = 0$ case.

η	Symbol spectrum	K
0	pass	0.077
0.01	pass	0.11
0.05	pass	0.49
0.1	pass	0.76
0.5	pass*	0.98
1.0	fail	0.99

TABLE VI. Results of the symbol spectrum and 0-1 tests for the complex continuous regular system (12) with normally distributed noise.

η	Symbol spectrum	K
0	pass	0.077
0.01	pass	0.16
0.05	pass	0.74
0.1	pass	0.90
0.5	fail	0.99
1.0	fail	1.0

the upper left graph of Fig. 11, we see that the symbolization threshold would be at approximately $x = 0$ and that there are significantly many elements of the series between $x = 0$ and $x = \pm 0.5$. Hence, noise amplitudes on the order of 0.5 should cause threshold crossings that do not exist in the original series. This is supported by Table V, where K jumps from 0.76 when $\eta = 0.1$ to 0.98 when $\eta = 0.5$. In fact, further inspection of the original series shows that there are some elements of the series between $x = 0$ and $x = \pm 0.1$, which explains the relatively high value of K for $\eta = 0.1$. Although not shown in Fig. 11, the symbol spectrum for the $\eta = 0.1$ case is very similar to the $\eta = 0$ case (shown in Fig. 11), which is why we interpret the result of the symbol spectrum test as a "pass" suggesting the series is deterministic. However, because $K = 0.76$ for $\eta = 0.1$, we should conclude that the results of the combined symbol spectrum test and the 0-1 test are inconclusive. The $\eta = 0.5$ results (middle row of Fig. 11) can also be difficult to interpret. Because of our knowledge of the noise-free data and the nature of the system, we know that this series is very noisy and, hence, random. However, without such prior knowledge, one could interpret the symbol spectrum test result (middle graph in Fig. 11) as a deterministic system because it appears to be similar to some of the results for chaotic series with which we have worked before. Hence, we could incorrectly characterize the series as chaotic because of the consistently high K_c spectrum for $\eta = 0.5$. The symbol spectrum test identifies the $\eta = 1.0$ case to be stochastic due to the relatively low degree of overlap of the symbol spectra. This, of course, is consistent with the knowledge of the original series. As we have worked with this test, we have learned that symbol spectrum results similar to those of the $\eta = 0.5$ case are common "transitional" results as the noise amplitude increases from low to high noise and usually suggests a very noisy (hence stochastic) series. The previous statement is a result of our experience with the symbol spectrum test and is not a rigorous formal result.

The results from the normal noise distribution in Table VI parallel the uniform noise distribution results of Table V. We see that the normal distribution does produce some inconclusive results where the series appears to be deterministic (passes the symbol spectrum test), but produces a middle-range value for the 0-1 test.

The symbol spectrum test and the 0-1 test combined seem to work well for the complex continuous regular series in the

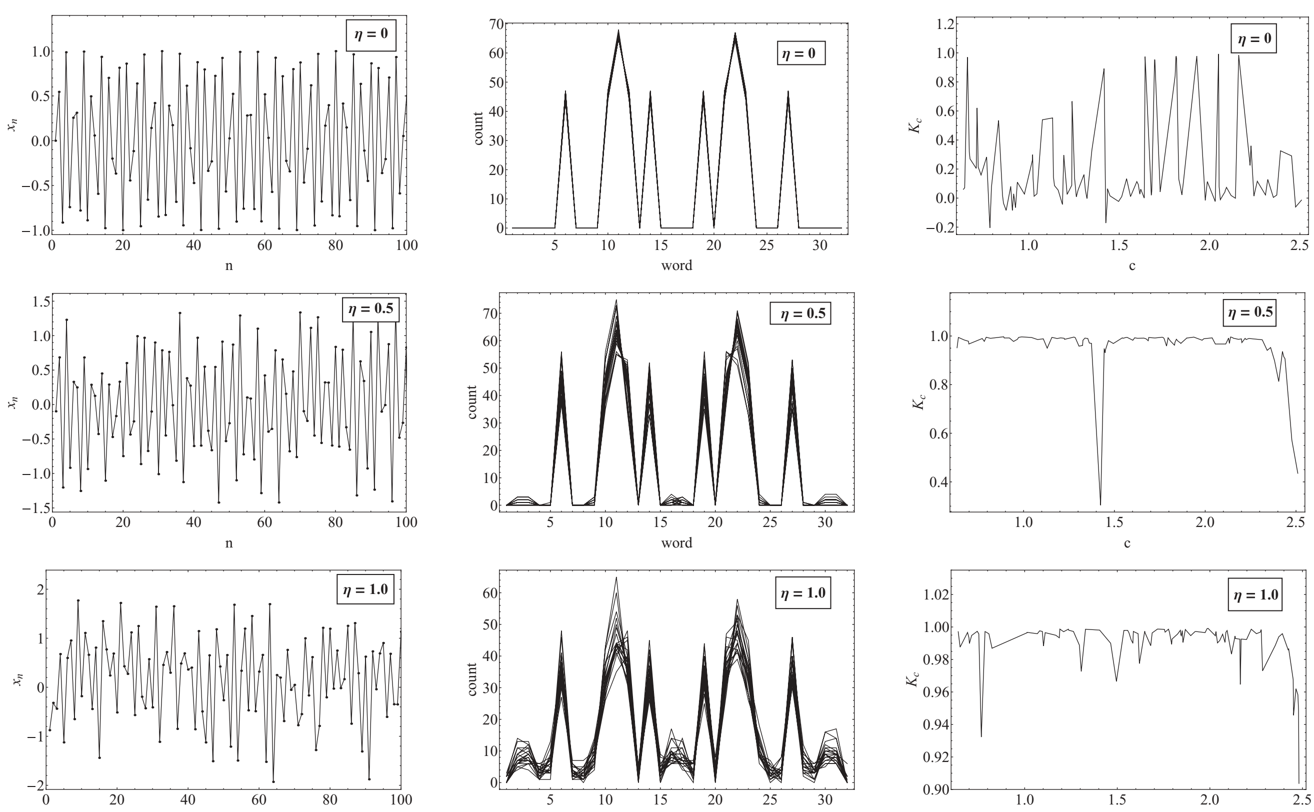

FIG. 11. Results of the symbol spectrum test and the 0-1 test for the complex continuous regular system (12). The first column contains the unsymbolized time series and the second and third columns contain the results of the symbol spectrum test and the 0-1 test, respectively.

cases of low noise and high noise. However, just as with the Logistic map, the test can be inconclusive for "moderate" noise levels.

D. S Persei

Based on our work in the preceding sections, what can we say about the results obtained from the S Persei data illustrated in Fig. 4? The symbol spectrum test has some hints of regularity between the spectra, especially when it comes to the location of the peaks in the spectra. However, there is also much variability in the spectra between the peaks. The 0-1 test returns $K = 1$ and the lowest K_c is around 0.95. This, in fact, could be noisy chaotic data; however, the symbol spectrum is simply not clean enough to characterize the data as chaotic with certainty. The regularities of the peaks suggest that there may be some hint of deterministic structure in the series. However, the series is irregularly sampled, as many real-world data sets are, and the effects of irregular sampling on these tests are unknown. It is possible that irregular sampling may affect the tests in a way similar to noise. Understanding the effects of irregular sampling on these tests is part of our planned future work. Based on our results, we believe a conservative characterization is warranted here. The results on their own are inconclusive. A more definitive characterization could be made if we had more information about how irregular sampling affects the tests.

IV. CONCLUSION

In this paper, we have studied the modes of failure for the symbol spectrum test and the 0-1 test when attempting to classify noisy binary time series generated from three model systems. The symbol spectrum test and the 0-1 test combined can serve as an effective tool for the characterization of noisy binary time series, especially when noise levels are low or very high. Furthermore, the tests seem more sensitive to normally distributed noise than to uniformly distributed noise. We have also found that these tests could incorrectly characterize a series as chaotic. Understanding the modes of failure is an important prerequisite for using any test on time series of which little is known.

We approached the problem of characterization supposing that we had no *a priori* knowledge of the system before running the characterization tests. Of course, this is typically not the case. Often, one has some estimate or model of the amount of noise present in a measurement. Furthermore, it is not uncommon to have working models of the systems from which the data are being measured. We chose to approach the problem assuming no knowledge of the system in order to test the characterization algorithm in the worst-case scenario. We believe the characterization algorithm worked well even with such strict assumptions. Of course, the more one knows about the system from which the time series is measured, the easier it is to interpret results from tests such as the symbol spectrum test and the 0-1 test.

Our results suggest that we should consider tolerances for the 0-1 test result. Our experience with these tests suggests that a $K < 0.5$ can be used to characterize a system as regular (i.e., neither chaotic nor random). This threshold has been identified as a guideline based on our experience of using the 0-1 test in conjunction with the symbol spectrum test and is not formally derived. Values of $K > 0$ for regular dynamics are sometimes due to working with short time series. We have also found cases in which poor choice of sampling times for regular continuous systems can lead to $0 < K < 0.5$. The complex continuous regular and Logistic map results suggest that this guideline is at least a good starting point for such a threshold. Refining the $K < 0.5$ guideline will be the subject for future work. In our experience, a similar lower bound for chaotic series is not as necessary. The chaotic series we have studied with the 0-1 test so far have produced $K > 0.95$.

While we believe that our model systems display a good range of different behaviors from which to assess the characterization tests, it is of course impossible to test any algorithm on all possible systems. However, we believe that the results of this paper can serve as a guide to implementing these tests on data taken from "real-world" systems. In the future, we plan to apply the combined symbol spectrum test and 0-1 test to "real-world" data taken from systems whose dynamics are known. The 0-1 test has been successfully applied to some laboratory data [9] from a system known to be deterministic. However, if the dynamics of the system are not known *a priori*, a test for determinism such as the symbol spectrum test should be done to properly characterize the dynamics. Understanding how well the symbol spectrum test works with such laboratory data will be a critical next step in being able to confidently apply the symbol spectrum test and the 0-1 test to systems whose dynamics are largely unknown.

ACKNOWLEDGMENTS

The authors would like to thank David G. Fisher and Eugene Sprechini of Lycoming College and Eugene Tracy of The College of William and Mary, Gail R. Kulp, and the reviewers for their helpful comments on this manuscript. The authors would also like to thank Charles Finney and C. Stuart Daw of Oak Ridge National Laboratory, whose comments were of great help with the symbol tree test algorithm. We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research. C.W.K. was partially supported by Lycoming College's Professional Development Grant.

[1] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data* (Springer, New York, 1996).
[2] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge, London, 2004).
[3] H. Kantz, Phys. Lett. A **185**, 77 (1994).
[4] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Physica D **16**, 285 (1985).
[5] G. A. Gottwald and I. Melbourne, Proc. R. Soc. London Ser. A **460**, 603 (2004).

[6] G. A. Gottwald and I. Melbourne, SIAM J. Appl. Dynam. Syst. **8**, 129 (2009).
[7] J. Hu, W. Tung, J. Gao, and Y. Cao, Phys. Rev. E **72**, 056207 (2005).
[8] G. A. Gottwald and I. Melbourne, Phys. Rev. E **77**, 028201 (2008).
[9] I. Falconer, G. A. Gottwald, I. Melbourne, and K. Wormnes, SIAM J. Appl. Dynam. Syst. **6**, 395 (2007).
[10] Z. Yang and G. Zhou, *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 20, (IEEE, New York, 1998), pp.2670–2673.
[11] J.-S. Lih, J.-Y. Ko, J.-L. Chern, and I.-M. Jiang, Europhys. Lett. **40**, 7 (1997).
[12] D. T. Kaplan and L. Glass, Phys. Rev. Lett. **68**, 427 (1992).
[13] S. Lu, K. H. Ju, J. K. Kanters, and K. H. Chon, in *Proceedings of the 28th IEEE EMBS Annual International Conference* (IEEE, New York, 2006), pp. 763–766.
[14] C. Finney (private communication).
[15] G. A. Gottwald and I. Melbourne, Nonlinearity **22**, 1367 (2009).
[16] G. A. Gottwald and I. Melbourne, Physica D **212**, 100 (2005).
[17] American Association of Variable Star Observers website: [http://www.aavso.org].
[18] I. D. Howarth, J. Br. Astron. Assoc. **86**, 210 (1976).
[19] J. M. T. Thompson and H. B. Stewart, *Nonlinear Dynamics and Chaos*, 2nd ed. (Wiley, West Sussex, England, 2002).