

Statistical tests for deterministic effects in broad band time series

Kesheng Wu^{a,1}, Robert Savit^a and William Brock^b

^aPhysics Department, University of Michigan, Ann Arbor, MI 48109, USA

^bDepartment of Economics, 1180 Observatory Dr., University of Wisconsin, Madison, WI 53706, USA

Received 20 July 1992

Revised manuscript received 12 April 1993

Accepted 19 April 1993

Communicated by G. Ahlers

We derive a normalized version of the indicators of Savit and Green, and prove that these normalized statistics have, asymptotically, a normal distribution with a mean of zero and standard deviation of one if the time series is random in the sense of being IID (independent and identically distributed). We verify this result numerically, and study the magnitude of the finite size effects. We also show that these statistics are very sensitive to the existence of deterministic effects in the series, even if the underlying deterministic structure is complex, such as those generated by a chaotic system. We show that with moderate amounts of data, the statistics can easily indicate the presence of an underlying attractor even in the presence of IID noise which is as large as, or greater than the signal. Finally, we discuss the generalization of our approach to include (1) other null hypotheses besides IID which express assumptions of specific dependencies and (2) the study of deterministic effects between more than one time series.

1. Introduction

The analysis of systems with broad band output is notoriously difficult. Common techniques based on the calculation of Fourier spectra or correlation functions often provide little insight into the underlying dynamics of such a system. Indeed, using these traditional methods it is often difficult to even distinguish between random and nonrandom systems. With the growth in interest in nonlinear systems, the need to understand the underlying dynamics of systems generating complicated time series becomes even more pressing.

Recently, an approach to this problem has

been developed [1,2]. This approach, which was inspired by the work of Brock, Dechert and Scheinkman [3,4], is based on the construction of conditional probabilities for the persistence of closeness of sets of variables in a time series. In its simplest incarnation, one can define a set of quantities which indicate the degree of deterministic dependence of an element of a time series on the j th lag of the series, given that one has used the information in the intervening lags.

Consider a time series $x(i)$. Define the d -dimensional vector, whose coordinates are d sequential elements of the time series as

$$\begin{aligned} v(i) &= (x(i), x(i-1), \dots, x(i-d+1)) \\ &\equiv (v_1(i), v_2(i), \dots, v_d(i)). \end{aligned} \quad (1.1)$$

Now choose a positive number ε , and define the Grassberger–Procaccia correlation integral as

¹ Current address: Computer Science Department, University of Minnesota, Minneapolis, MN 55455, USA.

$$C_d(\varepsilon) = \frac{1}{\nu} \sum_{\text{pairs}} I_{ij}(\varepsilon), \tag{1.2}$$

with

$$I_{ij}(\varepsilon) = \prod_{k=1}^d \Theta(\varepsilon - |v_k(i) - v_k(j)|). \tag{1.3}$$

ν is the number of distinct pairs of d -dimensional vectors constructed from the time series according to (1.1) and $\Theta(x)$ is the Heaviside function. In words, $C_d(\varepsilon)$ is the probability that each pair of the d Cartesian components of the two randomly chosen d -vectors are within ε of each other, respectively.

We can now construct a set of indicators, δ_j , defined as

$$\delta_j = 1 - \frac{C_j^2}{C_{j-1}C_{j+1}}. \tag{1.4}$$

In terms of conditional probabilities, it is not difficult to show [1] that

$$\delta_j = 1 - \frac{P(t_{j+1}|t_j, \dots, t_2)}{P(t_{j+1}|t_j, \dots, t_1)}, \tag{1.5}$$

where t_k stands for the statement $|v_{d-k+1}(m) - v_{d-k+1}(n)| \leq \varepsilon$, for a randomly chosen m and n . The extent to which δ_j differs from zero is a measure of the extent to which new deterministic information is contained in the j th lag of the series, which is not contained in the 1st, 2nd, . . . , $(j - 1)$ st lags. More precisely, suppose we consider two sets of $j - 1$ successive elements of the time series, and we require that each element of one set be within ε of the corresponding element of the other set. We then calculate the probability that the next elements in the two sets will also be within ε of each other. We repeat the calculation with j , instead of $j - 1$ elements and compare the results. If these conditional probabilities are equal, then $\delta_j = 0$, and we say that there is no new information contained in the j th lag. The extent to which they are not equal is a measure of the extent to

which there is statistical dependence in the time series on the j th lag. The larger δ_j is, the more information there is in the j th lag. This is a quick summary of the meaning of the δ_j 's. These same ideas can be generalized to study statistical dependencies among different time series. For more information consult refs. [1,2].

This method has been shown to work extremely well in a wide range of broad band time series [5-7]. It can often discriminate very well between series that are random in the sense of being IID (independent and identically distributed) and those that have an underlying deterministic structure, even if they are highly chaotic. The methods also work well in the presence of noise. Even with noise which is the same order of magnitude as or larger than an underlying deterministic process, these methods are quite sensitive to the underlying signal.

The existence of nonrandomness in the sense discussed above is indicated by nonzero values of some of the δ_j 's. But in any realistic finite time series the δ_j 's will not be identically zero even if the series has an IID nature. This raises the statistical question of what constitutes a significant nonzero value of δ_j . One approach is to compare computed values of δ_j for the time series of interest, with a distribution of values of δ_j computed for times series of the same length and same overall probability distribution, but which are known to be IID. This can readily be accomplished by scrambling the series of interest with a good random number generator a number of times, calculating δ_j for each scrambled series. One thus produces a realistic null distribution for δ_j . This is an example of a method known in the statistical literature as bootstrapping. In section 4 we shall briefly discuss this further. In addition, however, the δ_j 's have some very nice asymptotic statistical properties. It turns out that the δ_j 's are related to quantities which are known as U -statistics and for which powerful asymptotic theorems exist [8]. In particular, it is possible to derive a kind of central limit theorem for properly normalized versions of the δ_j 's. One can then

prove that if a very long time series is IID, these normalized δ_j 's are well behaved, normally distributed statistics with mean of zero and standard deviation of one. In addition to their aesthetic and theoretical appeal, these asymptotic results are very useful reference points in the empirical analysis of finite size data sets.

Before presenting the technical details of our work, it may be useful to place our results and their significance in context. Those readers interested in a wide ranging discussion of methods for testing nonlinearities in time series are encouraged to consult refs. [9–11]. Here we remark on the relationship of our work to previous work on three aspects of the problem of analyzing nonlinear effects in time series. First, we shall briefly discuss the use of Lyapunov exponent estimates as a method for indicating the presence of chaos in a time series. This approach is quite distinct from ours, but is used to analyze many of the same systems and so is worth briefly discussing. Second, we shall comment on more general methods of testing for “neglected structure” in time series. The work in the present paper falls into this category. Third, we shall discuss the relative merits of bootstrapping and analytic asymptotic results in determining the significance of statistical indicators.

First, there has been a good deal of work on the use of estimates of Lyapunov exponents as indicators of chaotic structure in time series. The statistical inference issues of this approach, as well as comparisons with other methods is discussed in ref. [12] (hereinafter referred to as BGHJ). They argue that a good way to test the null hypothesis of chaos consists of three steps: (i) Estimate the deterministic part, $F(\cdot)$, of the dynamical system

$$x(t+1) = F[x(t), \dots, x(t-L)] + e(t+1), \quad (1.6)$$

where $\{e(t)\}$ is independently and identically distributed with mean 0, using a consistent estimation method (i.e., one that converges to the true deterministic dynamic, $F(\cdot)$, as sample

length tends to infinity). Call this estimate \hat{F} (\hat{F} denotes an estimate of F). (ii) Estimate the largest-Lyapunov exponent of \hat{F} by a consistent method. Call it $\lambda(\hat{F})$ and show this converges to $\lambda(F)$ as sample length goes to infinity. (iii) Test the null hypothesis of chaos by testing $H_0: \lambda(F) > 0$. The BGHJ strategy delivers a test of the null hypothesis of chaos, such that probability of rejection of chaos when chaos is true goes to α for a level α test as sample length goes to infinity. A level α test rejects the null when the test statistic lies in the $1-\alpha$ tail of the null distribution^{*1}. For any alternative hypothesis H_x with deterministic part G such that $\lambda(G) < 0$, the BGHJ strategy is to deliver a test that will reject H_x with probability one as sample length tends to infinity.

Methods for estimating Lyapunov exponents include the works in refs. [14–16] and their references to the works of Eckmann, Ruelle, Wolf and others. Methods based on neural net estimation of the underlying deterministic map as discussed in refs. [14,15] appear to converge faster. It is beyond the scope of this paper to discuss the relative merits of different Lyapunov estimation procedures. The advantage of a Lyapunov exponent test for chaos is that it is a direct test for chaos. One major disadvantage is this: The sampling distribution of Lyapunov exponent estimates is unknown therefore “error bars” for estimates have to be attached on an *ad hoc* basis.

A second approach to the analysis of complex time series involves methods of testing for “neglected structure” of any form, not necessarily chaotic. This approach uses an estimated best fitting model in a given null hypothesis class and passes the estimated residuals through a testing procedure designed to detect the neglected structure. For example, one can test the null hypothesis class of functions that depend on L -lags by

^{*1} See ref. [13] for a discussion of this kind of classical hypothesis testing in the context of stochastic processes relevant to the discussion here.

estimating F in (1.6) and testing the residuals for IID using various tests, such as the BDS test. Brock and Dechert in ref. [9] discuss when this approach is valid and extend it to cases where the $\{e(t)\}$ are not independent. When F is the class of linear functions this gives a procedure to test the null class of linear models driven by IID errors (noise), which we call IID-linear. The test of Brock, Dechert, Scheinkman, and LeBaron [3], is especially adaptable to testing the hypothesis of IID-linearity. This is because the first order asymptotic distributions of the test on the estimated residuals of best fitting linear models are the same as on the true residuals for a large class of IID-linear processes. Call this the “invariance result”^{#2}. Brock and Dechert in ref. [9] show the invariance result holds for a large class of F 's for systems (1.6). This is a very useful property because most testing methods based on estimated residuals require a different correction term for each different F in (1.6). This invariance property makes the BDS methods particularly useful in testing the adequacy of fit of estimated models. (See ref. [17] and the article by Sayers in ref. [18].)

However, the BDS test, while it can reject the null hypothesis of IID on the estimated residuals, $\hat{e}(t+1) = x(t+1) - \hat{F}(x(t), \dots, x(t-L))$, gives little information about the nature of the non-IID structure of the series being tested. In particular, in the context of the form (1.6), the BDS test cannot isolate the variables in the estimated F, \hat{F} , which are responsible for the rejection of IID for the residuals.

Savit and Green [1,2] developed a set of indicators which help detect at which lag the dependence is concentrated. But no distribution theory was developed in those papers. The current paper develops distribution theory for an

appropriately scaled version of the Savit–Green indicators under the IID-null. Furthermore, since the statistics treated in this paper are functions of correlation integral statistics, the same “Taylor series” expansion arguments used in the derivation of the BDS statistic (see ref. [4], Appendix D) may be used to show that the normalized Savit–Green statistic derived here also preserves the invariance property^{#3}. This invariance property, as well as the ability to detect at which lag the dependence occurs, makes our contribution a useful addition to the literature on detecting structure in time series observations.

A third aspect of the analysis of time series which is relevant to our work is the method of bootstrapping or surrogate data [11,20]. Here one posits a specific null hypothesis for a series, and “bootstraps” from the data itself the distribution of statistics of interest under this null class. For example, if the null hypothesis for a time series $x(n)$ is that it is IID, then one can simply randomly scramble the elements $x(n)$ to create an IID series with the same overall distribution statistics. The statistics can then be computed on a set of such null series to determine a null distribution for the statistics. An up-to-date guide to bootstrapping can be found in ref. [21]^{#4}. Depending upon the application, it may be useful to bootstrap the Savit–Green statistics without the normalization factor derived in this paper. Nevertheless, the fact that the normalized Savit–Green statistics derived in this paper satisfies the invariance property, and the fact that the asymptotic results are well controlled, speaks to the utility of the normalized statistics derived here. Moreover, since bootstrapping relies on the computer rather than

^{#2} See Brock and Potter in ref. [10] for a discussion of testing different notions of linearity and avoiding common misunderstandings concerning the Wold representation theorem.

^{#3} The two papers of ref. [19] do the most rigorous job of developing this kind of argument for a broad class of statistics which are functions of the correlation integral type statistics used in this paper.

^{#4} See Lahiri's chapter in [21] for bootstrapping under conditions of weak dependence rather than IID.

analytics to work out approximations to null distributions, a sole reliance on bootstrapping obscures much of the structure that shines through in the analytics.

In summary, the work described in this paper is useful for the following reasons: (1) Our statistics can be used to test goodness-of-fit of estimated models like (1.6) which may be used as input into other methods such as Lyapunov exponent estimation. (2) Our statistics satisfy the invariance property like the BDS statistic but, unlike BDS and many other “residuals based” tests, our statistics give information on which lag the left out dependence occurs. (3) Like BDS our statistics are simple to compute and are analytically transparent for the IID case.

The rest of this paper is organized as follows: In the next section we describe the δ_j 's in a little more detail, discuss their relation to the BDS statistic, and introduce the idea of these quantities as vehicles for tests of various null hypotheses. In section 3 we present our main result, which is the derivation of the asymptotic behavior of the δ_j 's. In Section 3.1 we introduce the philosophy of the calculations, set the notation, and define a number of possibly unfamiliar concepts. With this background, we show in section 3.2 how to normalize the δ_j 's so that they become asymptotically normally distributed statistics with a mean of zero and a standard deviation of one, under the null hypothesis of IID. During the course of this derivation, it will become clear how to derive normalization factors so that statistics proportional to the deltas can be derived with normal asymptotic behavior under nulls that are less restrictive than IID. We also briefly compare our asymptotic results with those of the BDS statistic. In section 4 we present results of computer experiments showing that our normalized δ_j 's are, indeed, normally distributed for long IID time series. We also describe the practical application of these statistical methods to real finite data sets, and point out some interesting exceptional cases. A summary and conclusion constitute section 5.

2. The δ_j 's as statistics

2.1. Review of the BDS statistic

The BDS statistic [3,4] is a statistical quantity, the evaluation of which may be considered to be a test against a null hypothesis that a sequence of numbers is IID. The construction of the statistic proceeds from the observation that for an asymptotically long sequence of numbers that are IID

$$T_d(\varepsilon) = C_d(\varepsilon) - [C_1(\varepsilon)]^d \cong 0 \quad (2.1)$$

in a statistical sense, for any d no matter how large. The contribution of Brock, Dechert and Scheinkman was to derive the correct normalization factor, $K_d(\varepsilon)$ by which to divide (2.1) to be able to make a precise statistical statement. They showed that in the limit that the time series gets infinitely long,

$$\xi_d(\varepsilon) = \frac{T_d(\varepsilon)}{K_d(\varepsilon)} \rightarrow \mathcal{N}(0, 1), \quad N \rightarrow \infty, \quad (2.2)$$

i.e., $\xi_d(\varepsilon)$ is a statistic asymptotically normally distributed with a mean of zero and a standard deviation of one for any embedding dimension, d , if the sequence of numbers is IID. Here N is the length of the series. The normalization factor, $K_d(\varepsilon)$, is somewhat complicated. It is derived in refs. [3,4]. The values of $\xi_d(\varepsilon)$ are used in the following way. Suppose one calculates $\xi_d(\varepsilon)$ for some values of d and ε for an asymptotically long time series, and suppose, for example that the value is 3. Using the probabilities associated with a normal distribution, one can then say that there is only about a 1% chance that that sequence of numbers could have been produced by an IID process. Thus, the computation of (2.2) constitutes a test against the null hypothesis that the sequence of numbers is IID.

The BDS statistic has been extensively studied and has been shown to have good power against a wide class of alternative hypotheses, with

particularly good power against deterministic chaos [3.4]. Thus, many sequences generated by deterministic maps, but which pass standard tests of randomness based on the calculation of autocorrelation functions, can be shown to be nonrandom using the BDS test. $\xi_d(\varepsilon)$ may be thought of as a quantity that measures the “clumpiness” (other than that due to a non-flat probability distribution) in d embedding dimensions. Typically, when applied to a low dimensional deterministic chaotic sequence, $\xi_d(\varepsilon)$ becomes significantly larger than one for values of d in which the attractor begins to show some non-trivial structure, and increases with increasing d . For example, when applied to the tent map with ε chosen to be one-half the standard deviation of the values of the sequence (about 0.144) $\xi_2(\varepsilon)$ is already about 300, $\xi_3(\varepsilon)$ is about 700, and $\xi_d(\varepsilon)$ continues to grow for larger values of d . For a noisy deterministic system with a higher dimensional attractor, $\xi_d(\varepsilon)$ will typically be fairly small for low embedding dimensions and will begin to increase in dimensions in which the attractor begins to show some non-IID structure.

2.2. The δ_j 's and tests of null hypotheses

The BDS test is very powerful. However, once $\xi_d(\varepsilon)$ takes on non-null values, it will generally continue to show non-null behavior for increasing d , regardless of whether or not new information is contained in higher dimensional embeddings of the series. A case in point is the tent map described above, for which no new information is contained in embeddings above two dimensions, particularly as $\varepsilon \rightarrow 0$. The $\delta_j(\varepsilon)$, on the other hand, since they are related to conditional probabilities, can indicate the existence of new information in higher dimensions.

A significant nonzero value of δ_j can be used as an indication of conditional probabilistic dependence in the sense described in section 1. Many important qualitative aspects of a series can be reasonably deduced from this positive use

of the δ_j 's, as is evident in refs. [1,2]. From a statistical point of view, however, the δ_j 's can be used more precisely as tests against various null hypotheses, in a manner similar to the way the BDS statistic can be used. For example, each δ_j can be thought of as a separate test against the null hypothesis that a sequence of numbers is IID. On the other hand, it is clear that the δ_j 's do contain more information, than just a measure of the probability that a sequence of numbers is not IID. Suppose, for example, that δ_3 is the only δ_j that is significantly nonzero. This tells us not only that the sequence is unlikely to be IID, but that it fails to be IID because of conditional probabilistic dependence on the third lag. This suggests that the δ_j 's can be used as tests against alternate null hypotheses. For example, one might form the null hypothesis that the only conditional probabilistic dependence in the series is on the third lag. This means that for all $j \neq 3$,

$$P(t_{j+1}|t_j, \dots, t_2) = P(t_{j+1}|t_j, \dots, t_1). \quad (2.3)$$

As will become clear in the next section, one can formulate such null hypotheses which are more specific than the IID null, and one can derive normalized versions of the δ_j 's such that those statistics are asymptotically normally distributed with a mean of zero and a standard deviation of one. The price one pays for such more specific null hypotheses is that the normalization factors are more complex. The IID null is important in simplifying the results derived in the next section. In lieu of deriving normalization factors for more specific nulls, one may alternatively use the IID null, and rely on the fact that such normalized δ_j 's have power against various alternative hypotheses. In the example above, if the δ_j 's are normalized appropriate to the IID null, and if only (the so normalized) δ_3 is nonzero, that is strong evidence in favor of the hypothesis that (2.3) fails only for $j=3$. Thus, the nonzero values of the IID normalized δ_j 's contain significant detailed information about the underlying structure of the time series, which can often be

useful, even in the absence of forming statistically precise alternate null tests.

We turn now to a derivation of the properly normalized version of the δ_j 's with asymptotically simple behavior under the IID null.

3. Asymptotic behavior of the δ_j 's

3.1. Background

The results of BDS [3,4] were derived by observing that the correlation integral, $C_d(\varepsilon, N)$, computed from a sample of length N , is a U -statistic. The theory of U -statistics for weakly dependent stochastic processes has been developed by Denker and Keller [8]. This theory is the basis for our development here.

U-statistics:

Let $\{Y_t\}$ be an \mathbb{R}^k -valued stochastic process. U -statistics of order 2 are defined thus:

$$U(T) = \frac{2}{T(T-1)} \sum_{s=1}^T \sum_{t=s}^T h(Y_t, Y_s), \tag{3.1}$$

where $h: \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$. Moreover, $h(y, z) = h(z, y)$, is a symmetric function and is called a "kernel". Serfling [9] proves (i) $U(T)$ is a minimum variance unbiased estimator of $E[h(y, z)]$, the expectation value of h , over the class of all unbiased estimators of $E[h(y, z)]$. (ii) An asymptotic theory, parallel to the more familiar case of single sum statistics, exists for U -statistics. (iii) The expectation of $U(T)$ is the same as the expectation of h .

Since the integral for embeddings in d -dimensional space contains " d -overlapping" dependence in the kernel, even under the null hypothesis of IID, we need the more general theory of Denker and Keller [8]. We state Denker and Keller's Theorem 1 below for general statistics of order m , although we shall only need it for U -statistics of order 2.

Theorem. Let $h: X^m \rightarrow \mathbb{R}$ be a non-degenerate kernel and let $\{X(n)\}$ be a strictly stationary stochastic process. Then the asymptotic distribution of $(N/m_{\sigma_N})(U_N(h) - \theta)$ is $\mathcal{N}(0, 1)$ provided one of the following conditions is satisfied:

(a) $(X_n)_{n \geq 1}$ is uniformly mixing in both directions of time, $\sigma_N^2 \rightarrow \infty$, and for some $\alpha > 0$

$$\sup_{1 \leq t_1 < t_2 < \dots < t_m} E|h(X_{t_1}, \dots, X_{t_m})|^{2+\alpha} < \infty.$$

(b) $(X_n)_{n \geq 1}$ is uniformly mixing in both directions of time with mixing coefficients $\varphi(n)$ satisfying $\sum \varphi(n) < \infty$, $\sigma^2 \neq 0$ and

$$\sup_{1 \leq t_1 < t_2 < \dots < t_m} E|h(X_{t_1}, \dots, X_{t_m})|^2 < \infty.$$

(c) $(X_n)_{n \geq 1}$ is absolutely regular with coefficients $\beta(n)$ satisfying $\sum \beta(n)^{\alpha/(2+\alpha)} < \infty$ for some $\alpha > 0$, $\sigma^2 \neq 0$ and

$$\sup_{1 \leq t_1 < t_2 < \dots < t_m} E|h(X_{t_1}, \dots, X_{t_m})|^{2+\alpha} < \infty.$$

3.2. Derivation of asymptotic results

In this section we shall use the theorem above to prove that, under the IID assumption, properly normalized δ_j are asymptotically normally distributed statistics with a mean of zero and standard deviation of one. We shall also derive the required normalization factor.

For our purposes, we will not need the more general weak dependence captured in (a)–(c) above. Since the indicator kernel in the correlation integral is bounded between zero and one, the moment conditions needed by Denker and Keller are trivially satisfied. In addition, under the IID assumption the weak dependence conditions are also satisfied. We define the following (finite N) U -statistics and functions of U -statistics. We shall suppress ε in the notation, since it is fixed,

$$C(j, N) \equiv C_j(\varepsilon) \quad \text{in (1.2)} \tag{3.2}$$

for a series of length N ,

$$\delta(j, N) = \delta_i \text{ in (1.4) with } C_i \text{ replaced by } C(i, N) \text{ for } i = j - 1, j, j + 1, \tag{3.3}$$

$$g(j, N) \equiv C(j + 1, N) C(j - 1, N) - C^2(j, N) \equiv G[C(j - 1, N), C(j, N), C(j + 1, N)]. \tag{3.4}$$

Note that $\delta(j, N) \equiv D[C(j - 1, N), C(j, N), C(j + 1, N)]$ and $g(j, N)$ are functions of U -statistics. Since $C(1, N)$ converges in probability to $C(1) \equiv C$ under the IID as $N \rightarrow \infty$, it follows that $\delta(j, N)$ and $g(j, N)$ converge in probability to zero as $N \rightarrow \infty$. To put it another way, under IID, D and G are zero at population values.

Let D_k, G_k denote the partial derivatives with respect to the k th argument of D and G respectively ($k = 1, 2, 3$), evaluated at population values, $C(i) = C^i, i = j - 1, j, j + 1$.

We now apply the “ δ -method” of Serfling [22]. I.e., we expand the functions D and G around the population values in a Taylor series and take the limits as $N \rightarrow \infty$. This yields, for any smooth function H such that $H[C(j - 1), C(j), C(j + 1)] = 0$

$$\begin{aligned} \text{Lim}\{N^{1/2}H[C(j - 1, N), C(j, N), C(j + 1, N)]\} \\ = \text{Lim}\{N^{1/2}[H_1 \times (C(j - 1, N) - C^{j-1}) \\ + H_2 \times (C(j, N) - C^j) \\ + H_3 \times (C(j + 1, N) - C^{j+1})]\} \\ = \mathcal{N}(0, V(H)), \end{aligned} \tag{3.5}$$

where all partial derivative $H_i, i = j - 1, j, j + 1$ are evaluated at (C^{j-1}, C^j, C^{j+1}) , “Lim” denotes limit in distribution as $N \rightarrow \infty$, and $\mathcal{N}(0, V)$ denotes normal distribution with mean zero and variance, V . Note that the right hand side of (3.5) is a linear combination of U -statistics, and hence a U -statistic for $H = D, G$. Let $\hat{g}(j, N)$ and $\hat{\delta}(j, N)$ denote the linear terms in (3.5) for $H = G, D$ respectively. In particular

$$\hat{g}(j, N) = C(j + 1)[C(j - 1, N) - C(j + 1)] + C(j - 1)[C(j + 1, N) - C(j + 1)]$$

$$- 2C(j)[C(j, N) - C(j)], \tag{3.6}$$

$$\hat{\delta}(j, N) = \frac{g(j, N)}{[C(j + 1)C(j - 1)]}. \tag{3.7}$$

Equation (3.5) implies $\text{Lim}\{N^{1/2}\hat{g}(j, N)\} = \mathcal{N}(0, V(G)), \text{Lim}\{N^{1/2}\hat{\delta}(j, N)\} = \mathcal{N}(0, V(D))$, where

$$\begin{aligned} V(G) &= \lim E[\{N^{1/2}\hat{g}(j, N)\}^2], \\ V(D) &= \lim E[\{N^{1/2}\hat{\delta}(j, N)\}^2]. \end{aligned} \tag{3.8}$$

The limits in (3.8) are easy to calculate. Just square the terms, take the mathematical expectation, and use the null hypothesis of IID to simplify the expressions. Let us explain. Define

$$\begin{aligned} K^\varepsilon(j + 1) &\equiv K^\varepsilon[u(t, j + 1), u(s, j + 1)] \\ &= C(j + 1)[K^C(j - 1) - C(j - 1)] \\ &\quad + C(j - 1)[K^C(j + 1) - C(j + 1)] \\ &\quad - 2C(j)[K^C(j) - C(j)], \end{aligned} \tag{3.9}$$

where,

$$\begin{aligned} K^C(j) &\equiv K^C[u(s, j), u(t, j)] \\ &= \theta(\varepsilon - |u(s, j) - u(t, j)|), \end{aligned} \tag{3.10}$$

and $u(s, j) = (x(s), x(s - 1), \dots, x(s - j + 1))$. Here θ is the j -dimensional Heaviside function. I.e., the right hand side of (3.10) is one if $|x(s - k) - x(t - k)| < \varepsilon$ for all $0 \leq k \leq j - 1$, and is zero otherwise. All of the component U -statistics are of degree 2, hence the weighted sum is also a U -statistic of degree two. Therefore, “ m ” is equal to 2 in Denker and Keller’s Theorem 1 [8]. Thus,

$$\begin{aligned} 4\sigma_N^2(g) &\equiv V(G) \\ &= 4E\left(\{K_1^\varepsilon[u(t, j + 1)]\}^2 \right. \\ &\quad \left. + 2 \sum_{i \geq 1} K_1^\varepsilon[u(t, j + 1)] K_1^\varepsilon[u(t + i, j + 1)]\right); \end{aligned} \tag{3.11}$$

here

$$\begin{aligned} K_1^\varepsilon[u(t + i, j + 1)] \\ \equiv E\{K^\varepsilon[u(t + i, j + 1), u(s + i, j + 1)]|u(t + i, j + 1)\} \end{aligned}$$

$$\begin{aligned} &= \int K^g[u(t+1, j+1), u(s+i, j+1)] \\ &\quad \times dF[u(s+i, j+1)], \end{aligned} \tag{3.12}$$

and the $(j+1)$ -dimensional invariant measure is $dF[u(s+i, j+1)] \equiv dF[x(s+i)] \cdots dF[x(s+i-j)]$.

Given the embedding in d dimensions, the overlap can be at most d , hence at most d of the ‘‘cross terms’’ in (3.11) are nonzero under the IID hypothesis. The regularity conditions imposed by Denker and Keller [8] which are stated above ensure that the series (3.11) is absolutely convergent.

We can now compute the right hand side of (3.11) to obtain

$$\begin{aligned} \sigma^2(g[j]) &= C(2j-2) \text{Var}\{K_1^C(j+1)\} \\ &\quad + C(2j+2) \text{Var}\{K_1^C(j-1)\} - 2C(2j) \\ &\quad \times \text{Var}\{K_1^C(j)\}, \end{aligned} \tag{3.13}$$

where $K_1^C(j)$ is computed from $K^C(j)$ as K_1^g is computed from K^g in (3.12).

We now define $K^2 \equiv E\{K^C[x(r), x(s)]K^C[x(s), x(t)]\}$ with K^C defined as in (3.10) for $j=1$. Here the expectation value is taken over all three elements $x(r)$, $x(s)$ and $x(t)$. Then, it is not difficult to show that $\text{Var}\{K_1^C(j)\} = K^{2j} - C^{2j}$. Using this and $C^j = C(j)$ in (3.13) we have

$$\sigma(g) = C^{j-1}K^{j-1}(K^2 - C^2). \tag{3.14}$$

We can now compute the same quantities for $\delta(j, N)$. We have

$$\aleph_j \equiv \left[\frac{N^{1/2}}{2\sigma[\delta(j)]} \right] \delta(j, N) \rightarrow \mathcal{N}(0, 1), \quad N \rightarrow \infty, \tag{3.15}$$

where the variance formula can be computed following the same procedure as above. Doing this one obtains

$$\begin{aligned} \sigma^2[\delta(j)] &= \frac{\text{Var}[K_1^C(j+1)]}{C(2j+2)} + \frac{\text{Var}[K_1^C(j-1)]}{C(2j-2)} \\ &\quad - 2 \frac{\text{Var}[K_1^C(j)]}{C(2j)} \end{aligned} \tag{3.16}$$

and so

$$\sigma[\delta(j)] = \left[\frac{K}{C} \right]^{j-1} \left(\left[\frac{K}{C} \right]^2 - 1 \right). \tag{3.17}$$

We turn now to numerical results and applications.

4. Numerical results, practical applications, and some exceptions

In this section we will present the results of some calculations of the \aleph_j which are the δ_j normalized to have an asymptotic distribution of $\mathcal{N}(0, 1)$. We first examine some IID series, and in section 4.2, we turn our attention to some non-IID series generated by dynamical maps.

4.1. IID series

Since the result derived in the last section is asymptotic, we are particularly interested in seeing the effects of finite size data sets. As we shall see, the finite size effects have some dependence on the overall probability distribution. Moreover, there are some exceptional cases of IID distributions for which the result (3.15) is not valid for a technical reason.

To begin, turn to figs. 1–3 in which we plot the results of calculating the \aleph_j for $1 < j < 3$ for an IID time series of numbers drawn from a Gaussian distribution. Define

$$\varepsilon = \mu \times (\text{standard deviation of the unconditional probability distribution of the elements in the time series}). \tag{4.1}$$

In these plots, μ is set equal to $\frac{1}{2}$. One hundred iterations of three different length time series (1000, 5000, and 20000 points) were per-

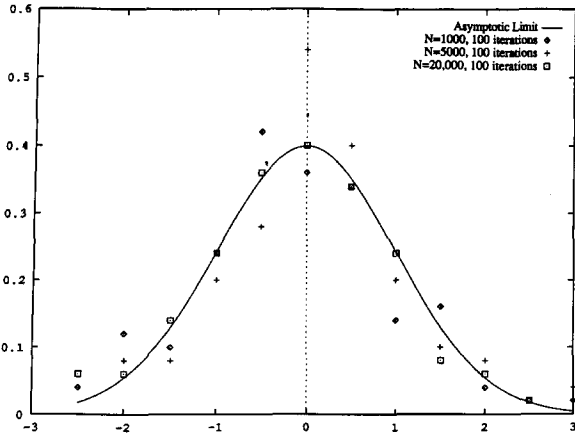


Fig. 1. \aleph_1 for an IID sequence with an unconditional Gaussian probability distribution. Results from three calculations consisting of 100 iterations each of time series of length 1000, 5000 and 20 000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

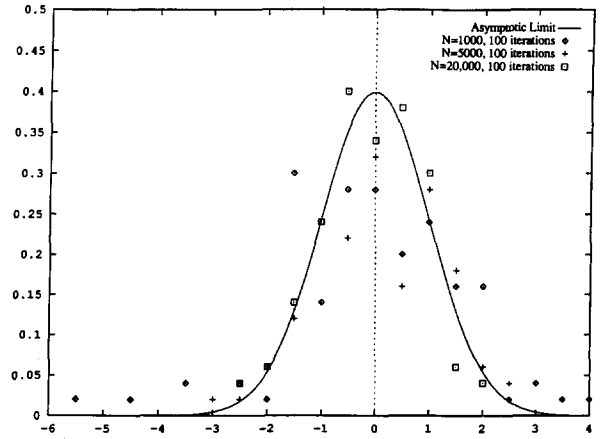


Fig. 3. \aleph_3 for an IID sequence with an unconditional Gaussian probability distribution. Results from three calculations consisting of 100 iterations each of time series of length 1000, 5000 and 20 000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

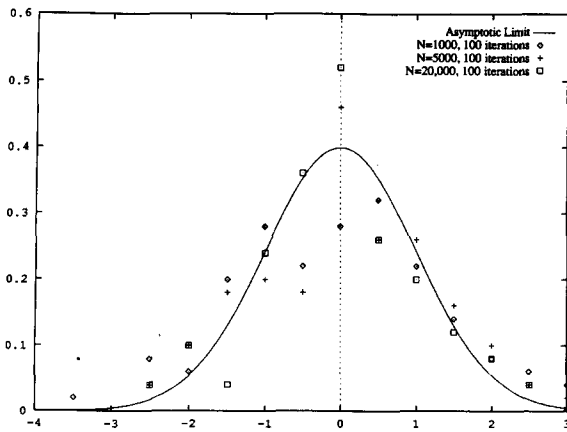


Fig. 2. \aleph_2 for an IID sequence with an unconditional Gaussian probability distribution. Results from three calculations consisting of 100 iterations each of time series of length 1000, 5000 and 20 000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

formed, and the \aleph_j for each iteration were computed. Each of the graphs presents three histograms of the 100 iterations for each length of time series for a given value of j . The theoretical asymptotic normal distribution is also shown. From these graphs we see (1) fairly good agreement with the asymptotic result even for time series of 1000 points iterated 100 times, (2) an approach to the asymptotic limit as the length

of the time series increases, and (3) closer agreement with the asymptotic result for smaller j , for the same length time series. These results are consistent with our expectations. In particular, the somewhat slower approach to asymptopia for larger j just reflects the fact that increasingly more data is required to clearly discern structure in higher dimensional embeddings of a time series.

That said, however, it is still quite remarkable how close the agreement is to the asymptotic result with only moderate amounts of data, even for larger j values^{#5}. To see this yet more clearly, refer to table 1 where the mean and standard deviations for this system are tabulated up to $j = 5$ for $\mu = \frac{1}{2}$.

The finite size behavior of these statistics is strongly dependent on the value of μ . As μ decreases, more data is needed to render the statistics reliable, since for smaller values of μ there are fewer points within any ϵ sized box. To

^{#5} Because U -statistics have a normal asymptotic distribution, it is easy to construct measures of their efficiency for finite data sets. See ref. [23]. See also ref. [22], p. 176, for a discussion of the relative efficiency of U -statistics. Those comments may be relevant to an explanation of the close agreement between our numerical results and the asymptotic limit.

Table 1
Mean values and standard deviations of the \aleph_j for a random Gaussian distribution, $\mu = \frac{1}{2}$.

Size	Iterations	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>						
250	100	-0.23	-0.15	-0.60	-1.15	-4.41
500		-0.14	0.05	-0.22	-0.36	-0.82
1000		-0.17	-0.05	-0.24	-0.15	-0.19
5000		0.029	-0.05	0.19	0.06	-0.32
250	500	-0.17	-0.11	-0.17	-0.79	-3.48
500		-0.12	-0.015	-0.095	-0.11	-1.05
1000		-0.04	-0.05	-0.14	-0.253	-0.08
5000		-0.005	-0.005	0.03	-0.01	-0.05
<i>Standard deviations</i>						
250	100	1.28	2.02	2.87	5.40	14.87
500		1.24	1.61	2.20	3.61	6.58
1000		1.08	1.33	1.66	2.73	4.67
5000		0.99	0.93	1.19	1.47	2.04
250	500	1.29	1.81	2.95	5.34	11.47
500		1.17	1.47	2.17	5.54	6.72
1000		1.11	1.24	1.60	2.46	4.57
5000		0.97	1.00	1.16	1.43	2.09

put it another way, as μ decreases, we are probing increasingly detailed behavior of the system. A determination of the nature of this more detailed structure naturally requires more data. For example, for the IID distributions in table 1, 100 iterations of time series of length 5000 yields an average value of δ_3 of 0.36 and a standard deviation of 3.05 for $\mu = 0.25$, while for $\mu = 0.1$ these same quantities are -0.06 and 16.05. Notice, in particular in this example the increasing standard deviations of the δ 's with decreasing μ .

As a second example, refer to figs. 4–6. These graphs are similar to figs. 1–3 and show the \aleph_j calculated for time series with $N = 1000$ and 5000, but now for an IID time series with an overall probability for the elements defined by

$$p(x) \propto \begin{cases} x, & 0 \leq x \leq 1, \\ x^{-4} & 1 \leq x \leq \infty. \end{cases} \quad (4.2)$$

In these figures $\mu = \frac{1}{2}$. Because of its power law tail, this probability distribution differs qualitatively from the Gaussian, and so is a good test

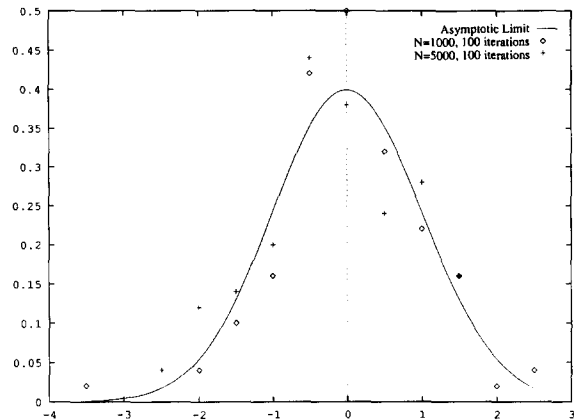


Fig. 4. \aleph_1 for an IID sequence with the unconditional probability distribution (4.2). Results from three calculations consisting of 100 iterations each of time series of length 1000 and 5000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

of the sensitivity of finite size effects on the asymptotic behavior of the probability distribution. As we see in the figures, the \aleph_j have the same qualitative behavior as a function of N as for the Gaussian case. Although they are certain to differ in the details, the general behavior of

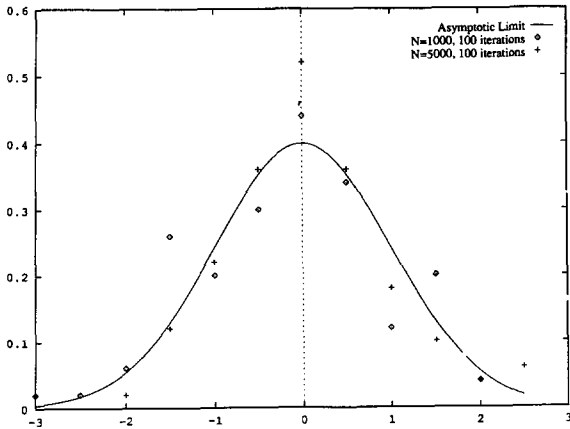


Fig. 5. \aleph_2 for an IID sequence with the unconditional probability distribution (4.2). Results from three calculations consisting of 100 iterations each of time series of length 1000 and 5000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

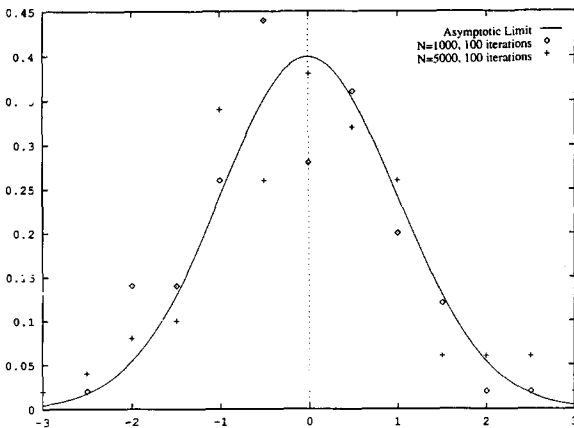


Fig. 6. \aleph_3 for an IID sequence with the unconditional probability distribution (4.2). Results from three calculations consisting of 100 iterations each of time series of length 1000 and 5000 elements are plotted. The solid line is the theoretical asymptotic limit $\mathcal{N}(0, 1)$.

the \aleph_j for finite length series appears to be relatively stable with respect to different overall probability distributions. In fact, it is quite generic. We have studied a number of IID series with various distributions including binomial, Poisson, linear with $p(x) \propto x$ for $0 < x < 1$, as well as an IID sequence with the overall probability distribution of the logistic map. In all these cases, the behavior of the \aleph_j are consistent with

the results presented in table 1 for the Gaussian case.

The third numerical case we wish to present is something of an exception. The derivation in section 3 is a kind of Taylor expansion about the asymptotic limit. The result, eq. (3.15) depends on the existence of a nonzero second term in that expansion. The coefficient of this second term is proportional to the variance, in the denominator of (3.15). In a recent paper, in a discussion of the BDS statistic, Theiler [24] pointed out that for an IID sequence with a uniform probability distribution on the unit ring, (i.e. $P(x) = 1$ for $0 \leq x \leq 1$ with zero and one identified) the coefficient of the second term in the Taylor expansion leading to the BDS statistic is zero so that the variance associated with the BDS statistic goes to zero as $N \rightarrow \infty$. Thus the asymptotic behavior of the BDS statistic is not necessarily $\mathcal{N}(0, 1)$, and one must compute higher order terms to determine the asymptotic form. Similarly, in such a case the \aleph_j will also not necessarily be normally distributed for large N , and the result (3.15) may not be valid. On the other hand, it is interesting to remark that for an IID sequence with a uniform distribution on the open interval $(0, 1)$, the variance in (3.15) is nonzero, and the \aleph_j do appear to approach an asymptotic $\mathcal{N}(0, 1)$ distribution, although somewhat more slowly than for a nonuniform distribution. To see this numerically, refer to tables 2 and 3, in which we repeat the calculations of table 1 for the uniform distribution on the open unit interval, and the uniform distribution on the circle. Notice in table 2 that the size of the series seems to limit the asymptotic results. The standard deviations do not change substantially as the number of iterations is increased from 100 to 500 if the size of the series is fixed. This observation is also consistent with the results in table 1. The relatively slow approach to $\mathcal{N}(0, 1)$ in table 2 can be understood by noting that although the second term in the Taylor series is nonzero, it is relatively (and nongenerically) small compared to higher order terms, so that one must go to larger

Table 2
Mean values and standard deviations of the \aleph_j for the uniform distribution on the interval $(0, 1)$, $\mu = \frac{1}{2}$.

Size	Iterations	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>						
250	100	-0.48	0.25	-0.45	-2.20	-21.40
500		-0.45	0.075	-1.63	-4.77	-10.14
1000		0.25	-0.49	0.05	-2.23	-1.87
5000		-0.09	0.04	-0.27	0.93	-0.19
250	500	0.038	-0.90	-1.66	-5.94	-22.47
500		-0.18	-0.20	-0.86	-4.13	-11.25
1000		-0.05	0.01	0.30	-0.33	-3.06
5000		0.07	-0.14	0.18	0.10	-0.61
<i>Standard deviations</i>						
250	100	4.28	8.64	16.40	37.14	78.86
500		3.64	5.97	11.63	29.79	51.23
1000		2.76	4.40	8.38	19.70	33.60
5000		1.65	2.64	4.13	7.93	16.07
250	500	4.62	8.50	15.80	33.48	83.73
500		5.28	6.68	12.55	26.52	49.80
1000		2.74	5.19	9.32	17.53	33.73
5000		1.66	2.73	4.28	8.86	15.24

Table 3
Mean values and standard deviations of the \aleph_j for the uniform distribution on the unit ring with $\mu = \frac{1}{2}$.

Size	Iterations	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>						
250	100	-1.24	-5.83	0.95	-16.40	-0.88
500		2.36	-11.48	16.21	-38.22	-20.86
1000		-0.47	-1.78	8.53	-12.91	22.65
5000		-0.04	10.04	19.22	17.94	17.57
250	500	-1.56	-2.01	-0.71	-14.74	-42.90
500		0.30	-2.47	0.87	-11.06	-20.72
1000		-0.75	-4.44	2.81	-15.17	-26.16
5000		-1.53	4.88	-0.07	-1.78	0.74
<i>Standard deviations</i>						
250	100	14.98	33.34	51.18	112.2	224.9
500		49.23	69.05	107.8	213.5	388.4
1000		21.96	44.52	72.21	140.0	264.9
5000		45.66	86.23	152.3	279.3	565.2
250	500	19.31	38.05	86.03	156.0	414.2
500		36.23	55.81	106.31	194.7	334.1
1000		22.82	41.53	74.67	153.7	257.7
5000		47.73	86.29	161.9	309.6	603.2

N to see the emergence of the asymptotic distribution. In table 3, corresponding to the uniform distribution on a ring, there is no tendency for the results to approach $\mathcal{N}(0, 1)$ asymptotically. In fact, the standard deviations grow as the size of the time series increases because the denominator of (3.15) goes to zero as $N \rightarrow \infty$. Again, note that in table 3 the results are relatively stable as the number of iterations increases from 100 to 500 with fixed size for the time series.

4.2. Non-IID series

We now discuss the behavior of the \aleph_j for time series that are not IID. In particular, we shall

consider several cases of simple iterative maps. First, in table 4 we display the results of calculations of the \aleph_j for the simple tent map. With x restricted to $[0, 1]$, the tent map is defined by

$$x(n + 1) = \begin{cases} 2x(n), & x(n) \leq \frac{1}{2} \\ -2x(n) + 2, & x(n) \geq \frac{1}{2} \end{cases} \equiv T(x(n)). \tag{4.3}$$

Notice first the relative independence of the results on the number of iterations. Second, notice that, even though the tent map depends on only one lag, the \aleph_j are clearly nonzero even for $j > 1$. As stated earlier, the \aleph_j are only guaranteed to be asymptotically $\mathcal{N}(0, 1)$ for an IID series. Although there is but one lag depen-

Table 4
Mean values and standard deviations of the \aleph_j for the tent map, $\mu = \frac{1}{2}$.

Size	Iterations	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>						
1000	100	342.5	3.26	-3.10	2.76	0.54
5000		828.5	7.73	-7.27	7.20	0.64
1000	500	344.2	3.26	-3.27	2.98	0.49
5000		827.1	7.43	-7.67	7.02	0.91
<i>Standard deviations</i>						
1000	100	4.77	2.90	2.49	3.45	3.82
5000	100	5.11	2.79	1.90	2.65	2.37
1000	500	5.32	2.82	2.54	3.43	3.70
5000	500	5.45	2.58	1.87	2.83	2.52

Table 5
Mean values and standard deviations of the \aleph_j for the Hénon map, $\mu = \frac{1}{2}$.

Size	Iterations	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>						
1000	100	64.12	27.60	-4.280	6.378	2.692
5000		143.4	61.35	-9.557	14.09	5.855
1000	500	63.79	27.35	-4.14	6.15	2.68
5000		144.43	62.07	-9.39	14.10	6.12
<i>Standard deviations</i>						
1000	100	0.76	1.63	1.27	1.06	0.99
5000		0.80	1.69	1.21	1.02	0.88
1000	500	0.85	1.70	1.23	1.11	1.06
5000		0.83	1.74	1.25	1.07	1.06

dence in the tent map, the series violates the IID assumption so that \aleph_j for higher j values may be nonzero. In this case, for example, it appears that \aleph_3 will be significantly negative asymptotically, while \aleph_2 and \aleph_4 will be significantly positive. This is in contradistinction to the unnormalized δ_j which, in some sense, approaches zero asymptotically if there are no conditional probabilistic dependencies, as discussed in refs. [1,2]. The nonzero values of \aleph_j for $j > 1$ does contain information about the series. While the information is more difficult to untangle in the \aleph_j 's than in the δ_j 's because of the relatively complex nature of the normalization factor in (3.15), the \aleph_j 's do satisfy the invariance property discussed in the introduction.

As another example of a non-IID series, we turn to the Hénon map. This map has two lag dependence. As shown in table 5, \aleph_j is very large for $j = 1$ and 2, but is still significantly nonzero for higher values of j . As with the tent map, this is due to the fact that the \aleph_j are defined to be null under the IID hypothesis, so that even though there is no explicit dependence for $j > 2$, these \aleph_j are not guaranteed to be $\mathcal{N}(0, 1)$ asymptotically^{#6}. Finally, we note the interesting result that the standard deviations of the \aleph_j for $j > 2$ are all near one in table 5. We have no simple explanation for this observation, but suspect that the answer involves detailed properties of the Hénon map.

4.3. Noise

We now present some results that demonstrate the sensitivity of the \aleph_j to the presence of noise

^{#6} We should point out that even for maps in which there is only explicit dependence up to a given lag, the δ_j 's may be statistically nonzero for larger j 's, if ϵ is not too small. This is explained in detail in ref. [1]. Thus one, might suppose that the nonzero values of \aleph_j for large j in tables 4 and 5 could be due to a large value of the numerator in (3.15). However, as discussed in ref. [1], this effect is not particularly important in the tent map, in which there is no curvature. So, we conclude that it is the failure of the IID hypothesis which is primarily responsible for the nonzero values of the \aleph_j for larger j in tables 4 and 5.

Table 6
Mean values and standard deviations of the \aleph_j for the logistic map with and without observer noise, $\mu = \frac{1}{2}$.

σ	\aleph_1	\aleph_2	\aleph_3	\aleph_4	\aleph_5
<i>Mean values</i>					
0.0	590.4	44.90	-27.40	-8.71	-4.39
0.1	632.1	57.9	-24.7	-7.84	-2.68
0.5	182.7	22.62	2.85	1.29	0.71
1.0	11.0	2.90	0.37	-0.033	-0.026
10.0	-0.14	-0.036	0.10	-0.91	-1.22
<i>Standard deviations</i>					
0.0	1.14	2.95	2.96	3.25	3.97
0.1	2.87	2.98	2.67	3.36	3.29
0.5	5.23	3.97	3.58	3.25	3.88
1.0	1.70	1.25	1.30	1.99	3.51
10.0	1.36	2.15	3.70	7.71	12.6

added to an otherwise deterministic, chaotic system. For our example we have chosen to study the logistic map. In table 6 we present results for the values of the \aleph_j for a time series of 5000 points generated by the logistic map with and without observer noise. That is, a time series is generated completely deterministically from the chaotic logistic map:

$$y_{j+1} = 4y_j(1 - y_j). \tag{4.4}$$

IID noise drawn from a random, flat distribution with $0 \leq \eta_j \leq 1$ is then added to the y_j with strength σ to produce the resultant time series,

$$x_j = y_j + \sigma \eta_j. \tag{4.5}$$

In table 6 standard deviations and averages of the various \aleph_j are computed using 100 iterations of the relevant time series. We see from these results that the \aleph_j are quite robust to the addition of noise to a deterministic, even chaotic system. Even for $\sigma = 1.0$, which corresponds to the addition of noise which is 100% of the signal, the \aleph_j can still clearly distinguish the underlying chaotic structure. Only when the noise is many times the strength of the signal do the \aleph_j fail to detect the embedded deterministic structure. These results are consistent with those of ref. [1] in which it was shown that the δ_j are also

sensitive indicators of underlying determinism, even in the presence of substantial IID noise.

5. Summary

In this paper we have shown that the δ_j 's derived in refs. [1,2] have elegant asymptotic statistical properties. In particular, one can define a normalized version of the δ_j 's (called \aleph_j) and can prove that the \aleph_j 's are asymptotically normally distributed with a mean of zero and standard deviation of one for an IID series. The δ_j 's and the \aleph_j 's are very sensitive indicators of deviations from randomness (in the sense of IID) and are particularly useful in analyzing broad band time series without a simple Fourier spectrum.

We presented numerical results which verify that the asymptotic behavior is indeed $\mathcal{N}(0, 1)$ and which also gives some indication of the finite size effects and the approach to asymptopia. Generically the \aleph_j 's are quite well behaved with moderately small finite size effects. One exception is an IID sequence with a flat unconditional probability distribution. It can be shown that for such a series with a flat distribution on a ring, the asymptotic behavior is not $\mathcal{N}(0, 1)$ because the coefficient of the second term in the Taylor expansion about the asymptotic result is zero. Moreover, for a uniform distribution on an open interval, the asymptotic behavior is $\mathcal{N}(0, 1)$, but the approach to asymptopia is slower than in the generic case. This is due to a relatively small second term in the Taylor expansion, compared to higher order terms.

We also presented numerical results which show how the \aleph_j 's behave for non-IID series. As expected, the \aleph_j 's generically indicate the existence of underlying determinism in a time series, even if the series is chaotic. Unlike their unnormalized counterparts, the δ_j 's, the \aleph_j 's are not direct measures of conditional probabilities. Therefore, when applied to systems with depen-

dencies on a certain number of lags, the \aleph_j 's may not be $\mathcal{N}(0, 1)$ asymptotically even for those j 's corresponding to lags which do not provide new deterministic information about the underlying dynamics of the series. In order to ensure that a set of statistics will have such behavior, one must derive results analogous to those derived in section 3, with alternate null hypotheses which express more specifically which dependencies are included in the null and which are not.

We also studied the sensitivity of the \aleph_j 's to the addition of noise in an otherwise deterministic series. We found that even if the noise is the same order of magnitude as the underlying (chaotic) deterministic signal, the \aleph_j 's can still easily detect the presence of the deterministic structure. Only when the noise is much larger than the signal (in our example, an order of magnitude larger), do the \aleph_j 's fail to indicate the presence of the signal in time series of length 5000 with $\mu = \frac{1}{2}$.

The analytic and numerical results presented in this paper should be very useful in the analysis of complex time series. The statistics derived here are sensitive indicators of deviations from randomness in the sense of IID in time series. In addition, the methods used here can be generalized in two important ways. First, it is possible to derive normalized versions of the generalizations of the δ_j 's discussed in refs. [1,2]. Thus it is possible to construct statistics analogous to the \aleph_j 's, with simple asymptotic behavior, which, for example, can be used to test for independence of the elements of one time series on a second, different time series. Second, it is possible to generalize the derivation of section 3, to address other, more restrictive null hypotheses. For instance, one can seek statistics which have an asymptotic normal distribution under a null hypothesis that the n th term in the time series depends only on the $(n-1)$ st and $(n-2)$ nd terms.

The methods and results discussed here are significant new tools in the analysis of complex systems. They have already been fruitfully ap-

plied to a number of problems from diverse disciplines, and have the potential for providing the analyst with many more important new insights.

Acknowledgements

We are grateful to Matthew Green for help with the numerical computations. The work of R.S. and K.W. was supported in part by the US Department of Energy under grant no. DE-FG02-85ER45189; and the work of W.B. by National Science Foundation grant no. SES-9122344 and the Vilas Trust.

References

- [1] R. Savit and M. Green, *Physica D* 50 (1991) 95.
- [2] M. Green and R. Savit, *Physica D* 50 (1991) 521.
- [3] W. Brock, W. Dechert, J. Scheinkman and B. LeBaron, A Test for Independence Based Upon the Correlation Dimension, preprint, Department of Economics, The University of Wisconsin, Madison, The University of Houston, and The University of Chicago (1990).
- [4] W.A. Brock, D.A. Hsieh and B. LeBaron, *Nonlinear Dynamics, Chaos and Instability: Statistical Theory and Economic Evidence* (MIT, Cambridge, 1991).
- [5] L. Iasimides, C. Sackleres and R. Savit, University of Michigan preprint (1992).
- [6] C. Pentzell, C. Akerlof, M. Green and R. Savit, *New Statistical Tests of the Randomness of Radioactive Decay*, University of Michigan preprint no. UM HE 92-8 (1992).
- [7] C. Hiemstra and J. Jones, *Detection and Description of Nonlinear Dynamics in Daily Stock Returns and Volume*, Department of Economics, Loyola College and US Securities Commission preprint (1993); H. Pi and C. Peterson, *Finding the Embedding Dimension and Variable Dependences in Time Series*, University of Lund preprint no. LU TP 93-4 (1993).
- [8] M. Denker and D. Keller, *Z. Wahrscheinlichkeitstheor. Verw. Gebiet.* 64 (1983) 505.
- [9] N. Abraham, A. Albano, A. Passamante and P. Rapp, eds., *Measures of Dynamical Complexity and Chaos* (Plenum, New York, 1989).
- [10] M. Casdagli and S. Eubank, eds., *Nonlinear Modeling and Forecasting*, Santa Fe Institute Studies in the Sciences of Complexity (Addison-Wesley, Redwood City, CA, 1992).
- [11] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and D. Farmer, *Physica D* 58 (1992) 77.
- [12] W. Barnett, A. Gallant, M. Hinich and M. Jensen, *Robustness of Nonlinearity and Chaos Test to Measurement Effort, Inference Method, and Sample Size*, preprint prepared for Conf. on Nonlinear Dynamics in Economics (Florence, Italy, June 1992).
- [13] I. Basawa and B. Prakasa Rao, *Statistical Inference for Stochastic Processes* (Academic Press, New York, 1980).
- [14] R. Gencay and W. Dechert, *Physica D* 59 (1992) 142.
- [15] D. Nychka, S. Ellner, R. Gallant and D. McCaffrey, *J. R. Stat. Soc. B* 54 (1992) 399.
- [16] H. Abarbanel, R. Brown and M. Kennel, *Int. J. Mod. Phys. B* 5 (1991) 1347.
- [17] H. Pesaran and S. Potter, eds., *Special Issue on Nonlinear Dynamics and Econometrics: J. Appl. Econometr.* (Wiley, New York, 1992).
- [18] S. Krasner, *The Ubiquity of Chaos*, American Association for the Advancement of Science, Washington DC, Publication No. 89-15S (1990).
- [19] P. de Lima, *Testing Nonlinearities Under Moment Condition Failure*, preprint, Department of Economics, John Hopkins University (1992); *On the Robustness of Nonlinearity Tests to Moment Condition Failure*, preprint, Department of Economics, John Hopkins University (1992).
- [20] W. Brock, J. Lakonishok and B. LeBaron, *J. Finance*, XLVII (1992) 1731.
- [21] R. LePage and L. Billard, eds., *Exploring the Limits of Bootstrap* (Wiley, New York, 1992).
- [22] R. Serfling, *Approximation Theorems of Mathematical Statistics* (Academic Press, New York, 1980).
- [23] M. Denker and D. Keller, *J. Stat. Phys.* 44 (1986) 67.
- [24] J. Theiler, *Phys. Rev. A* 41 (1990) 3038.