Some Theory of Statistical Inference for Nonlinear Science: Expanded Version

William A. Brock
Ehung G. Baek

SFI WORKING PAPER: 1991-02-009

SFI Working Papers contain accounts of scientific work of the authors and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author’s copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
SOME THEORY OF STATISTICAL INference FOR NONLINEAR SCIENCE:
EXPANDED VERSION

William A. Brock
University of Wisconsin-Madison

Ehung G. Baek
Iowa State University

January 1991
Abstract

This article shows how standard errors can be estimated for a measure of the number of excited degrees of freedom (the correlation dimension), and a measure of the rate of information creation (a proxy for the Kolmogorov entropy), and a measure of instability. These measures are motivated by nonlinear science and chaos theory. The main analytical method is central limit theory of U-statistics for mixing processes. This paper takes a step toward formal hypothesis testing in nonlinear science and chaos theory.

Acknowledgments

The first author thanks the Guggenheim Foundation, the Wisconsin Graduate School, the Wisconsin Alumni Research Foundation, and the National Science Foundation (Grant # SEC-8420872) for essential financial support for this work. We thank two referees and the editor, C. Bean, for very helpful comments and Taeho Lee for the numerical integrations. Support for this paper was provided in part by the Santa Fe Institute Economic Research Program which is funded by grants from Citicorp/Citibank and the Russell Sage Foundation and by grants to SFI from the John D. and Catherine T. MacArthur Foundation, the National Science Foundation (PHY-8714918) and the U.S. Department of Energy (ER-FG05-88ER25054).
1. INTRODUCTION


Much of the excitement has to do with the potentiality of quantifying such vague notions as "level of complexity", "degree of instability", and "number of active nonlinear degrees of freedom". At a general level nonlinear science has a rich storehouse of ideas to inspire the field of nonlinear time series analysis, and, vice versa.

Most of the work to date has relied on diagnostics such as correlation dimension, Kolmogorov entropy, and Lyapunov exponents. Expository papers in this area are Brock (1986), Frank and Stengos (1988b) for economics, and Eckmann and Ruelle (1985), Theiler (1990b) for natural science. Eckman and Ruelle (1985) is an especially detailed and comprehensive review of nonlinear science. Brock (1986) contains some applications to economics and a discussion of some pitfalls to avoid. Frank and Stengos (1988b) surveys some of the useful literature and techniques and studies empirical chaos in economics by using daily rates of return on gold.

Unfortunately no formal theory of statistical inference exists for the dimension measures and the instability measures of nonlinear science. Brock, Dechert, and Scheinkman, hereafter, BDS (1987) developed some statistical theory (discussed below) for the
correlation integral of Grassberger/Procaccia/Takens (a measure of spatial nonlinear correlation) and used this theory to formulate a test of the null hypothesis of independently and identically distributed (IID) for a univariate series against an unspecified alternative. This work was extended to the vector case by Baek and Brock (1988). Brock and Dechert (1988a) provided some ergodic theorems for the correlation integral and some convergence theorems for a Lyapunov exponent estimation algorithm.

The new contribution of this paper is to provide some statistical inference theory for dimension measures and Kolmogorov entropy. Central limit theorems for weakly dependent stochastic processes and for U-statistics provide the tools needed for this paper. They are presented in section two. Asymptotic standard errors of dimension and Kolmogorov entropy estimates are derived as applications of the theory. Nuisance parameter problems occurring in these measures are discussed. In section three we calculate the correlation dimension estimates, the Kolmogorov entropy estimates, and their standard errors by using returns on weekly stock market index studied by Scheinkman and LeBaron (1989a). Final remarks and conclusions are in section four.

2. THEORY OF STATISTICAL INFERENCES

Let \( \{a_t\}, \ t=1,2,\ldots,T \) be a sample from a strictly stationary and ergodic stochastic process which we abuse notation by also denoting
by \{a_t\}, or a deterministic chaos with unique, ergodic invariant measure as in Brock (1986). This assumption allows us to replace all limiting time averages by corresponding phase averages. Also the limiting value of all time averages will be independent of initial conditions. The data, \{a_t\} can be "embedded" in m-space by constructing "m-futures"

\[ a^m_t = (a_t, \ldots, a_{t+m-1}), \quad t=1,2,\ldots,T-m+1. \]

The correlation integral for embedding dimension \(m\) is defined by

\[ C(\epsilon,m,T) = \frac{1}{T_m(T_m-1)} \sum_{1 \leq t \neq s \leq T_m} I(a^m_t, a^m_s; \epsilon) \frac{1}{T_m(T_m-1)}. \]

where \(T_m = T-m+1\), \(I(x,y;\epsilon) = 1\) if \(\|x-y\| \leq \epsilon\) and 0 otherwise, \(\|x\|\) denotes the maximum norm, i.e. \(\|x\| = \max_{0 \leq i \leq m-1} |x_i|\) on \(\mathbb{R}^m\). The correlation integral measures the fraction of total number of pairs \((a^m_t, a^m_s)\) such that the distance between \(a^m_t\) and \(a^m_s\) is no more than \(\epsilon\). In other words, it is a measure of spatial correlation. Note that \(C(\epsilon,m,T)\) is a double average of an indicator function. Hence one expects it to converge as \(T \to \infty\). Denker and Keller (1986, Theorem 1 and (3.9)) and Brock and Dechert (1988a) show that

\[ C(\epsilon,m,T) \overset{d}{\to} C(\epsilon,m). \]

It is worthwhile to give some intuition into the measure \(C(\epsilon,m)\).
Let \( x \equiv (x_0, x_1, \ldots, x_{m-1}) \), \( y \equiv (y_0, y_1, \ldots, y_{m-1}) \), and

\[
F_m(x_0, x_1, \ldots, x_{m-1}) = \text{Prob}\{a_t \leq x_0, a_{t+1} \leq x_1, \ldots, a_{t+m-1} \leq x_{m-1}\} = \text{Prob}\{a_t \leq x\}.
\]

Then \( C(\epsilon, m) \) is given by

\[
(2.3) \quad C(\epsilon, m) = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} I(x, y; \epsilon) dF_m(x) dF_m(y).
\]

For example, look at \( C(\epsilon, 1) \)

\[
(2.4) \quad C(\epsilon, 1) = \int_{\mathbb{R}} \int_{\mathbb{R}} I(x_0, y_0; \epsilon) dF_1(x_0) dF_1(y_0)
\]

\[
= \int_{\mathbb{R}} [F_1(x_0 + \epsilon) - F_1(x_0 - \epsilon)] dF_1(x_0).
\]

In case \( \{a_t\} \) is IID, \( F_m(x) = \prod_{i=0}^{m-1} F_1(x_i) \), hence, by (2.3) we have

\[
(2.5) \quad C(\epsilon, m) = [C(\epsilon, 1)]^m.
\]

In general \( C(\epsilon, m) \) measures the concentration of the joint distribution of \( m \) consecutive observations, \( a_t^m \). It describes the mean volume of a ball of radius \( \epsilon \). The elasticity of \( C(\epsilon, m) \) describes the mean percentage of new neighbors of the center of a ball of radius \( \epsilon \) that are captured as the radius of the ball increases by one percent. The measure \( C(\cdot) \) is an example of a gauge function (Mayer-Kress (1987)). Its elasticity is a measure of dimension which is discussed below.
The information dimension (Eckmann and Ruelle (1985)) is estimated by measuring, for each embedding dimension \( m \), the slope of \( \log(C(\epsilon, m, T)) \) plotted against \( \log(\epsilon) \) in a zone where the slope appears constant (Ramsey and Yuan (1989, 1990), Scheinkman and LeBaron (1989a)). One then looks to see if these estimated dimensions become independent of \( m \) as \( m \) increases. An alternative measure of dimension is the point elasticity \( \frac{d[\log(C(\epsilon, m, T))]}{d[\log(\epsilon)]} = \frac{C'(\epsilon, m, T)\epsilon}{C(\epsilon, m, T)} \) where \( C'(\epsilon, m, T) \) is the derivative of \( C(\epsilon, m, T) \) with respect to \( \epsilon \). We will focus on the point elasticity here because it (cf. (2.18) below) can be written as a function of \( U \)-statistics. If \( \{a_t\} \) is IID the correlation integral takes a simple form. A useful nonparametric test of the null hypothesis of IID, which uses the correlation integral and which illustrates the methods to be used in this paper is in Brock, Dechert, and Scheinkman (1987).

BDS (1987) proved

**Theorem 2.1:** Let \( \{a_t\} \) be IID and assume \( V > 0 \) in (2.7) below, then

\[
T^{1/2}[C(\epsilon, m, T) - [C(\epsilon, 1, T)]^m] \xrightarrow{d} N(0, V), \text{ as } T \to \infty.
\]

Here "\( \xrightarrow{d} N(0, V) \)" means "convergence in distribution to \( N(0, V) \)" and \( N(0, V) \) denotes the normal distribution with mean 0 and variance \( V \) where

\[
V = 4[K^m + 2 \sum_{i=1}^{m-1} K^{m-i} 2i + (m-1)^2 C^{2m} - m^2 KC^{2m-2}],
\]
(2.8) \( C = EI(a_i, a_j; \epsilon) \) and \( I(a_i, a_j; \epsilon) = \begin{cases} 1 & \text{if } |a_i - a_j| \leq \epsilon \\ 0 & \text{otherwise} \end{cases} \)

where \(|x|\) is the absolute value of the real number \(x\),

(2.9) \( K = EI(a_i, a_j; \epsilon)I(a_j, a_k; \epsilon) \).

This theorem was used by BDS to build a nonparametric test for IID that had good size and power characteristics (especially against deterministic chaos) in comparison to some other popular tests for independence (cf. Hsieh and LeBaron (1988a,b)). The proof of Theorem 2.1 uses the theory of U-statistics. U-statistics are a type of generalized time average. \( C(\epsilon, m, T) \) is an example. They behave enough like simple time averages that a central limit theory exists for U-statistics that parallels the central limit theory for simple time averages like the sample mean.

**U-STATISTICS**

We will follow Sen (1972), Serfling (1980), and Denker and Keller (1983). A measureable function \( h: \mathbb{R}^n \to \mathbb{R} \) is called a kernel for \( \theta = Eh \) if it is symmetric in its \( n \) arguments. Typically \( \mathbb{R} = \mathbb{R}^k \) for some positive integer \( k \). A U-statistic for estimating \( \theta \) is given by

(2.10) \( U_T = \frac{\sum h(a_{t_1}, \ldots, a_{t_n})}{T_n^*} \)
where \( \mathcal{E} \) is taken over \( 1 \leq t_1 \leq \ldots \leq t_n \leq T \), \( T^*_n \) is the number of \( n \)-subsets of \( \{1, \ldots, T\} \), and \( \{a_t\} \) is a strictly stationary stochastic process with values in \( \mathcal{A} \). \( U \)-statistics are interesting because (i) they have many of the desirable properties of the simple time average \( u_T = \frac{1}{T} \sum_{t=1}^{T} a_t \), including central limit theorems and laws of large numbers, (ii) in a certain context they are minimum variance estimators of \( \theta \) in the class of all unbiased estimators of \( \theta \) (Serfling (1980, p. 176)), (iii) they converge rapidly to normality (Serfling (1980, p. 193, Theorem B)), and (iv) many useful statistics can be written in \( U \)-statistic form (Serfling (1980, Chapter 5)). We will only use the case \( n=2 \). So from now on \( n \) is fixed at 2. Before going on we stress that \( \{a_t\} \) can be an \( \mathbb{R}^k \) valued stochastic process in the general theory below.

The projection method of Hoeffding is applied by Denker and Keller (1983) to obtain the decomposition

\[(2.11) \ U_T = \theta + \frac{2}{T} \mathbb{E}\{h_1(a_t) - \theta\} + R(T), \text{ if } n=2\]

where \( h_1(a) = \mathbb{E}\{h(a_1, a_2) | a_1 = a\} \), and \( \theta = \mathbb{E}h_1(a) \).

\( \mathbb{E} \) runs from 1 to \( T \), and \( R(T) \) is a remainder that goes to 0 in distribution when multiplied by \( \sqrt{T} \) as \( T \rightarrow \infty \). Let us denote by

\[(2.12) \ \sigma_T^2 = \mathbb{E}\left\{ \sum_{t=1}^{T} g_1(a_t) \right\}^2 \text{ where } g_1(a) = h_1(a) - \theta,\]


the exact variance of the leading term in the above decomposition. Let us denote by

\begin{equation}
\sigma^2 = E\{g_1(a_1)^2 + 2 \sum_{1<i} g_1(a_1)g_1(a_i)\}
\end{equation}

its asymptotic variance, provided the sum converges absolutely. In this case

\begin{equation}
\sigma^2 = \lim_{T \to \infty} \sigma_T^2 / T.
\end{equation}

We state part of one of Denker and Keller's theorems below.

**Theorem 2.2:** (Denker and Keller (1983, p. 507)) Let \( \sigma^2 > 0 \), then,

\begin{equation}
T/(2\sigma_T)[U_T - \theta] \xrightarrow{d} N(0,1), \text{ as } T \to \infty,
\end{equation}

provided that the following condition is satisfied: The strictly stationary stochastic process \( \{a_t\} \) is absolutely regular with coefficients \( \beta_t \) satisfying

\begin{equation}
\sum_{t} \beta_t \delta/(2+\delta) < \infty, \text{ for some } \delta > 0, \sigma^2 > 0, \text{ and } \sup \{E|h(a_{t_1}, \ldots, a_{t_n})|^{2+\delta}\} < \infty.
\end{equation}
Here the "sup" is taken over \( 1 \leq t_1 < \ldots < t_n < T \). "Absolutely regular" asks that

\[
\beta_t = \sup E[\sup\{|P(A|G(1, s)) - P(A|A \in G(s + t, \omega))\}]
\]

tends to zero as \( t \to \omega \). Here the outside "sup" is taken over \( s \) in \( \{1, 2, \ldots\} \) and the inside "sup" is taken over \( A \) in \( G(s + k, \omega) \). The symbol "\( G(s, v) \)" denotes the sigma algebra generated by \( \{a_t | s \leq t < v\} \) \((1 \leq s, v \leq \omega)\). Other mixing conditions besides (2.17) including two by Denker and Keller yield similar results. The point is that we need some type of condition on the rate of decay of dependence over time, i.e. a mixing condition, in order to get the central limit theorem for dependent processes. Condition (2.17) seems as useful as any.

In the applications to follow we use Theorem 2.2 and the delta method (Serfling (1980, p. 124)) to obtain central limit theorems for differentiable functions \( H(z_1(T), \ldots, z_k(T)) \) of the \( k \)-vector of U-statistics \( z(T) = (z_1(T), \ldots, z_k(T)) \) where each \( z_i(T) \) has symmetric kernel function \( h_i(a_t, a_s) = h_i(a_s, a_t) \). Here is the basic method. Let \( z = Ez(t) \). Provided \( H(x) \) has non-zero derivative at \( z \) and the component U statistics have nondegenerate asymptotic distributions then we know from the delta method (Serfling (1980, p. 124)) that

\[
T^{1/2}(H(z(T)) - H(z)) \text{ has the same limit law as } T^{1/2}(DH(z) \cdot (z(T) - z))
\]

where \( DH(z) \) is the derivative of \( H \) evaluated at \( z \) and "\( \cdot \)" denotes scalar product. Put \( g(a_1, a_2) = DH(z) \cdot (h(a_1, a_2) - z) \), \( g_1(a) = \)
\text{E}\{g(a_t, a_s) | a_t = a\}. Then the formula (2.13) can be used to calculate the asymptotic variance of the limit distribution and Theorem 2.2 applies. With this background turn now to applications.

APPLICATIONS

In the applications below we will assume \{a_t\} is IID to simplify calculations of asymptotic variances from (2.13). But the methods apply to any general process to which Theorem 2.2 applies.

(1) STANDARD ERRORS OF DIMENSION ESTIMATES

The statistical properties of dimension calculations are investigated by Ramsey and Yuan (1989, 1990) and Theiler (1990a). As Ramsey and Yuan point out the point estimate of correlation dimension is typically derived from ordinary least squares (OLS) regression over an apparent constant slope zone on a log-log plot (Ramsey and Yuan (1990, p. 157, p. 160-161), Scheinkman and LeBaron (1989a)). Problems of subjectivity in the choice of the apparent constant slope zone together with the mathematical form of the OLS estimator lead us to focus upon the elasticity measure of dimension. We also wanted to see how well our methodology would perform on the most volatile measure of dimension. Derivatives are well known to be noisy and difficult to estimate. Another reason for concentration on this form is that we can write various estimators of the elasticity as a function of U-statistics. We calculate the slope of \([\log(C)/\log(\epsilon)]\)
of two nearby points for a point estimate of correlation dimension. Sample properties of our estimate are discussed in section 3. Using this dimension concept enables us to apply the theory of U-statistics. While Denker and Keller (1986) use U-statistics theory to derive asymptotic standard errors for a Grassberger-Procaccia type of correlation dimension estimate, our work was done independently.²

Let \{a_t\} be an IID stochastic process with finite moments as in (2.16).³ Then \{a^m_t\} satisfies the mixing condition of Theorem 2.2.

The dimension estimate, which is intended to approximate the elasticity, \(d[\log(C)]/d[\log(\varepsilon)]\), that we will examine is defined as follows:

\[
\log C(\varepsilon+\Delta\varepsilon,m,T) - \log C(\varepsilon,m,T)
\]
\[
= \frac{\log C(\varepsilon+\Delta\varepsilon,m,T) - \log C(\varepsilon,m,T)}{\log(\varepsilon+\Delta\varepsilon) - \log(\varepsilon)}
\]

Since \(C(\varepsilon,m,T) \xrightarrow{d} C(\varepsilon,m)\) therefore,

\[
(2.19) \quad d_m(\varepsilon,\Delta\varepsilon,T) \xrightarrow{d} d_m(\varepsilon,\Delta\varepsilon) \quad \text{as} \quad T \xrightarrow{\infty}.
\]

Note that \(d_m(\varepsilon,\Delta\varepsilon,T)\) is a function of two quantities, \(C(\varepsilon+\Delta\varepsilon,m,T)\) and \(C(\varepsilon,m,T)\), i.e. \(d_m(\varepsilon,\Delta\varepsilon,T)\equiv D(C(\varepsilon+\Delta\varepsilon,m,T),C(\varepsilon,m,T))\). By (2.11), we have

\[
(2.20) \quad C(\varepsilon+\Delta\varepsilon,m,T) - \theta(\varepsilon+\Delta\varepsilon,m) = \frac{2}{T}\Sigma(h_1(a^m_t,\varepsilon+\Delta\varepsilon) - \theta(\varepsilon+\Delta\varepsilon,m)) + R_1,
\]
\[(2.21) \quad C(\epsilon, m, T) - \theta(\epsilon, m) = \frac{2}{T} \Sigma(h_1(a_t^m, \epsilon) - \theta(\epsilon, m)) + R_2\]

where \(\theta(\epsilon, m) = C(\epsilon, m) = EC(\epsilon, m, T)\). Hence, we may apply the delta method (Serfling (1980, p. 124)) to prove the following theorem. The proof is in the Appendix.

**Theorem 2.3:** Assume \(\{a_t\}\) is IID and satisfies the moment condition in \(2.16\). Suppose the differential of \(D(\cdot, \cdot)\) is nonzero at \((C(\epsilon + \Delta \epsilon, m), C(\epsilon, m))\), and the covariance matrix of \((C(\epsilon + \Delta \epsilon, m, T), C(\epsilon, m, T))\) is nonsingular, and \(V_{D_m}\) defined below is positive. Then

\[(2.22) \quad T^{1/2}[D\{C(\epsilon + \Delta \epsilon, m, T), C(\epsilon, m, T)\} - D\{\theta(\epsilon + \Delta \epsilon, m), \theta(\epsilon, m)\}] \sim N(0, V_{D_m})\]

where

\[(2.23) \quad V_{D_m} = 4\gamma^2[A^m + B^m - 2C^m + \sum_{j=1}^{m-1}(A^{m-j} + B^{m-j} - 2C^{m-j})] ,\]

\[\gamma = [\log(\epsilon + \Delta \epsilon) - \log(\epsilon)]^{-1},\]

\[A = K(\epsilon + \Delta \epsilon)/C(\epsilon + \Delta \epsilon)^2,\]

\[B = K(\epsilon)/C(\epsilon)^2,\]

\[C = W(\epsilon + \Delta \epsilon, \epsilon)/(C(\epsilon + \Delta \epsilon)C(\epsilon)),\]

\(C(\cdot), K(\cdot)\) are defined in \(2.8\), \(2.9\) and \(W(\epsilon + \Delta \epsilon, \epsilon) = EI(a_i; a_j; \epsilon + \Delta \epsilon)I(a_i, a_j; \epsilon)\).

Theorem 2.3 is a basis for setting up hypothesis testing
concerning dimension. For example Scheinkman and LeBaron (1989a) produced a point estimate of about 6 for the correlation dimension of stock returns. This number has been widely cited. With our methods one can now estimate a standard error for such point estimates of dimension. This was not possible before. We investigate this problem in section three.

(2) STANDARD ERRORS OF KOLMOGOROV ENTROPY

The standard error of the approximate Kolmogorov entropy $K_m(\epsilon) \equiv \log[C(\epsilon,m)/C(\epsilon,m+1)]$ can be derived following the procedure of Theorem 2.3 since the sample estimator of $K_m(\epsilon), K_m(\epsilon,T)$, is a differentiable function of two $U$ statistics, $C(\epsilon,m,T), C(c,m+1,T)$.

The Kolmogorov entropy of a deterministic dynamical system, $y_{t+1}=f(y_t), y_t \in \mathbb{R}^k, f: \mathbb{R}^k \rightarrow \mathbb{R}^k$, is a measure of how fast a pair of states become distinguishable to a measuring apparatus with fixed precision under forward iteration (Eckmann and Ruelle (1985, p. 637)). For example if $\{a_t\}$ is IID the limit of the approximate Kolmogorov entropy, $K_m(\epsilon)$, is infinity as $\epsilon$ goes to zero. For finite $\epsilon, \{a_t\}$ IID implies $K_m(\epsilon) = -\log(C(\epsilon,1)) \rightarrow \infty, \epsilon \rightarrow 0$. The proof of the following theorem is found in the Appendix.

Theorem 2.4: Make the same assumptions as in Theorem 2.3. If $\{a_t\}$ is an IID process,
where $K_m(\epsilon,T)$ is the sample estimate of the Kolmogorov entropy,

\[(2.24) \quad T^{1/2}K_m(\epsilon,T) \xrightarrow{d} N(-\log C(\epsilon), VK_m),\]

\[(2.25) \quad T^{1/2}[K_m(\epsilon,T)+\log(C(\epsilon,1,T))] \xrightarrow{d} N(0, VK'_m)\]

where $K_m(\epsilon,T)$ is the sample estimate of the Kolmogorov entropy,

\[(2.26) \quad VK_m = 4[[K(\epsilon)/C(\epsilon)^2]^{m+1}-\{K(\epsilon)/C(\epsilon)^2\}^m] \quad \text{and}\]

\[(2.27) \quad VK'_m = 4[[K(\epsilon)/C(\epsilon)^2]^{m+1}-\{K(\epsilon)/C(\epsilon)^2\}^m+\{K(\epsilon)/C(\epsilon)^2 - 1\}].\]

For some applications, the invariance property, i.e. the first order asymptotics of the correlation dimension and the Kolmogorov entropy evaluated at estimated residuals are the same for true residuals, can be shown. We sketch this idea here.

First, in many applications we replace the series of observations $\{a_t\}$ by the standardized series in an attempt to scale the series so that its mean is zero and its variance is unity. But this introduces two nuisance parameters that are estimated by the sample mean and the sample standard deviation, which may change. These nuisance parameters may change the asymptotic distributions above.

A second fundamental concern is that many times we are really interested in testing the estimated residuals of some parametric
model such as an Autoregressive Moving Average (ARMA) model or an Autoregressive Conditional Heteroscedastic (ARCH) model for temporal dependence or instability. But then the distribution of the estimated residuals is contaminated by the estimation procedure. Some limited results are discussed below.

Consider null models of the form

$$(2.28) \ y_t = G(y_{t-1}, y_{t-2}, \ldots, y_{t-q}; b) + \epsilon_t$$

where $G$ is $C^2$ (twice continuously differentiable), the parameter vector $b$ is estimated $\sqrt{T}$ consistently and $\{\epsilon_t\}$ is IID with mean zero and unit variance. Then under modest regularity conditions the argument in Brock and Dechert (1988b) can be extended to show that, under (2.28), if $E h_1(u)=0$, the limit law of $\sqrt{T}(C(\epsilon, m, T) - C(\epsilon, m))$ is the same whether $C(\epsilon, m, T)$ is evaluated at the true $\{\epsilon_t\}$ or the estimated $\{\epsilon_t\}$. The full details are in Brock and Dechert (1988b) for the case $m=1$ and $G$ is a linear autoregression. We call this property "the invariance property". A similar argument can be developed to show the invariance property, under (2.28), for the limit law of

$$(2.29) \ \sqrt{T}[F\{C(\epsilon, 1, T), \ldots, C(\epsilon, k, T)\} - F\{C(\epsilon, 1), \ldots, C(\epsilon, k)\}]$$

where $F$ is $C^2$. This includes the correlation dimension estimate and the Kolmogorov entropy estimate discussed in this paper.
Remark: For the indicator kernel \( h_1(u) = E[I(u,v;\epsilon)|u] = F_e(u+\epsilon) - F_e(u-\epsilon) \). So \( E h_1'(u) = 0 \) in this case. Here \( F_e(u) = \text{Prob}(e_t \leq u) \).

These results are of limited usefulness in applications. First, they do not cover all \( F \) in (2.29) when the variance of \( e_t \) must be estimated. An invariance result for the BDS (1987) statistic is in Brock (1989). Second, they apply only to the estimated residuals of null models where the true residuals are assumed IID. We would like to get away with assuming weaker maintained assumptions on these residuals. Unfortunately we have not obtained any useful results under more general assumptions.

3. EMPIRICAL APPLICATION

In this section we apply the theory. Scheinkman and LeBaron (1989a) estimated the correlation dimension for 1226 weekly observations on the CRSP value weighted U.S. stock returns index starting in the early 1960's. They arrived at roughly a dimension of 6. They then calculated another estimate of dimension due to Takens which was also close to 6. Here we provide asymptotic estimates of standard errors for the elasticity estimate of dimension for Scheinkman and LeBaron's data set.

The embedding dimension is increased from 1 to 14, and the resolution parameters \( \epsilon + \Delta \epsilon \), \( \epsilon \) are adjusted from 0.9, 0.9^2 to 0.9^4, 0.9^5. For each embedding space and parameter value, a point estimate
of the correlation dimension is reported in Table 1. Dimension estimates are between 7 and 9 in high embedding dimensions, and their standard errors are low enough to make the test statistic values significant at the 5% level under the assumption of asymptotic normality. Note that the null hypothesis of IID is rejected in favor of a lower dimensional alternative. This is consistent with the results of Scheinkman and LeBaron.

Table 2 reports the dimension estimates and their associated standard errors computed for 1226 IID observations to compare with the 1226 actual weekly stock returns. An IMSL standard normal subroutine DRNNOA was used to generate the pseudo random numbers. Since the correlation integral loses too many comparable pairs in high embedding dimensions, we report the results for embeddings only up to 8 dimensions. We can see the correlation dimension estimate and the embedding dimension go almost together as they should. When the resolution parameter $\epsilon$ is too small, we lose comparable pairs very fast. The interesting fact from Table 2 is that most of the test statistics are insignificant at conventional significance levels, i.e. we fail to reject the null hypothesis that returns are IID. This is encouraging since we know the artificial returns are IID.

To estimate the speed of information creation, $K_m$, we also estimate the approximate Kolmogorov entropy. Tables 3 and 4 are entropy estimates and their standard errors computed from actual stock returns and standard normal random numbers. Theoretically if the stock returns process is IID then the entropy estimate should be
close to the value, \(-\log(C(\epsilon,1,T))\). The Kolmogorov entropy estimate becomes smaller than \(-\log(C(\epsilon,1,T))\) when actual values are used in Table 3, but this is not true when random numbers are used in Table 4. Actual data generate statistically significant test statistics, in other words, the test rejects the null hypothesis that stock returns are generated by an IID stochastic process.

Note from Table 4 that the \(K_m\) estimates are all positive even though the process is IID. Eckmann and Ruelle (1985, Sections 4 and 5) point out that \(K_m\) is a lower bound to the true Kolmogorov entropy and positive Kolmogorov entropy is associated with chaos. Our results caution the investigator that stochastic processes such as IID processes are also consistent with positive Kolmogorov entropy. This indeterminacy brings up a natural question: What do we learn when we reject the null hypothesis of IID with the \(K_m\) based test statistic as in Table 3? Let us explain.

Note that \(\log(C(\epsilon,m)/C(\epsilon,m+1)) = \log(C(\epsilon,1))^{-1}\) if and only if \(C(\epsilon,m+1)/C(\epsilon,m)=C(\epsilon,1)\), i.e. \(\text{Prob}(X_{t+1}|X_t,X_{t-1},\ldots,X_{t-(m-1)}) = \text{Prob}(X_{t+1})\) when "\(X_t\)" is shorthand for the event \(|X_t-X_s| \leq \epsilon\). Hence failure to reject the null of IID under a \(K_m\) based test is consistent with the \(m\)-past \((X_t,X_{t-1},\ldots,X_{t-(m-1)})\) having no predictive power for the future, \(X_{t+1}\). We say more about this in Baek and Brock (1990) where we show that this kind of testing methodology based upon \(K_m\) leads naturally to tests of whether one series \(\{Y_t\}\) helps predict
Monte Carlo experiments were done to examine the quality of normal approximation of the test statistics and small sample bias pointed out by Ramsey and Yuan (1989). 2500 samples were replicated to generate a sampling distribution. The same experiments were performed with different values of the sample size and the parameter $\epsilon$. But we only report the results where the sample size is 1000, and $\epsilon+\Delta \epsilon$, $\epsilon$ are 0.9, 0.92 for the correlation dimension and $\epsilon=0.9$ for the Kolmogorov entropy in Table 5 and 6.

In Table 5, the second column shows that the correlation dimension estimate is biased downward which makes the test statistic take negative values. Also histograms of the standardized estimates of the correlation dimension in Figure 1-9 (top plot) show this downward biasness. The average empirical asymptotic standard errors (ASE) which are computed by the 2500 empirical ASE’s based on (2.23) are in the third column. The fourth column contains the true ASE’s of the test statistic, $[d_m(\epsilon, \Delta \epsilon, T)-m]$, computed by using numerically calculated $C(\epsilon+\Delta \epsilon)$, $C(\epsilon)$, $K(\epsilon+\Delta \epsilon)$, $K(\epsilon)$ and $W(\epsilon+\Delta \epsilon, \epsilon)$. Even though $C$, $K$, and $W$ are consistently estimated, there is a big deviation between the mean ASE and true ASE (See the proof of Theorem 2.3 in Appendix for notations.). The main reason for this is that the $\gamma$ parameter exaggerates the ASE in our dimension calculation method. For instance $\gamma^2 \approx 90$ when $\epsilon+\Delta \epsilon=0.9$ and $\epsilon=0.9^2$. If there is a 1% discrepancy between the true value and the estimated value except for the factor, $4\gamma^2$, in the variance formula then we expect there will be
a 360% difference in the variance. As long as the normal approximation is good, there should not be a big problem to use the empirical ASE for hypothesis testing. If the sampling distribution is well approximated by a normal distribution, $\sqrt{T \cdot SE}$ is close to the mean of the empirical ASE. By comparing the third column and the last column, we may see how good the normal approximation is. In the high embedding dimensions 9 and 10, the approximation is quite good. However, when the data is embedded in low dimensions, the test statistic has smaller dispersion than the standard normal distribution. Since previous studies such as Scheinkman and LeBaron (1989a) indicate the meaningful embedding dimension range is high dimensional space, we think that our previous application to stock returns is suggestive.

It is important to realize that even though the test of IID based upon the derivative measure of dimension is capable of rejecting the null hypothesis of IID for stock returns in favor of some "lower dimensional" alternative. This does not necessarily mean chaos is present. There are many stochastic processes where close $m$-histories tend to be followed by close descendants that must be ruled out before one can claim chaos. Also, consistent with Ramsey and Yuan (1989,1990), biases appear in the dimension estimates and the asymptotic standard errors grow dramatically with the embedding dimension. Although theory implies that these biases disappear in the limit bootstrapping and bias reduction techniques along the lines of Efron (1982) have potential to improve performance. We suspect that bootstrapping, rather than using asymptotics will improve
performance of all the statistics discussed in this paper because bootstrapping apparently helps approximate some of the higher order terms in the Edgeworth expansion (Efron (1982), Efron and Tibshirani (1986) and references) whereas our methods capture only the first order terms.

Another technique that may improve performance of dimension based tests is to fix a zone of epsilons of the log(C(ε,m)) vs. log(ε) plot and follow Denker and Keller (1986) to estimate the slope of the log(C(ε,m)) vs. log(ε) plot over this zone. Since our derivative estimate of dimension performed better than we thought (even though it performed poorly) we believe the Denker and Keller procedure may perform well.\(^9\) Turn now to the \(K_m\)-based test which performed much better.

We analyzed the reliability of the \(K_m\)-based test in a similar way in Table 6. From the second column, there is no clear evidence that the entropy estimate is biased. Also the consistent estimates of \(C\) and \(K\) bring the mean ASE based on (2.27) very close to its true value. By comparing the third and last column we can say that the normal approximation is good for high embedding dimensions. The histogram constructed by the standardized sampling distribution in Figure 1 - 9 (bottom plot) also shows evidence of good normal approximation.

Finally we turn to the two examples which show how our dimension test can be applied without such a serious bias problem.
Example 1: Suppose you have two time series \( \{a_{1t}\} \) and \( \{a_{2t}\} \) and you want to test whether their dimensions are the same. I.e. you want to capture, in some reasonable way, the notion of a statistically significant difference in the number of "possibly nonlinear factors" or active modes in the two series.

To do this one could, under the maintained hypothesis of stationarity and mixing as in Theorem 2.2, set up the null hypothesis that the dimension for \( m, \varepsilon \) for the two series is the same against the alternative that it is not. The asymptotic variance under the null for the difference of the two dimension estimates could be derived as in the proof of Theorem 2.3, but the variance formula will need to be modified to include a string of covariance terms.

As a special case we will construct a test of the null hypothesis that \( \{a_{1t}\} \) and \( \{a_{2t}\} \) are both IID and mutually independent. This test is based upon comparison of the dimension estimates. The IID null leads to asymptotic null standard normality being achieved for a test statistic with a simple asymptotic variance formula but it opens a gap between the null hypothesis of the same dimension and alternatives of different dimension. That is to say, the test developed under the IID null may reject the null because IID does not hold even though the dimensions are the same. This can happen through the change in the variance formula. At the risk of repeating, however, under a suitable weak dependence condition, it would be possible to construct a test of the more desirable null hypothesis that the dimensions are the same which has a limiting null
standard distribution but a more complicated variance formula. We have

**Theorem 3.1:** Let $d_m^1(\epsilon, \Delta \epsilon)$ and $d_m^2(\epsilon, \Delta \epsilon)$ be the correlation dimension of the first series and the second series respectively. The null hypothesis that both series are IID and mutually independent with common distribution function is tested by $d_m^1(\epsilon, \Delta \epsilon) = d_m^2(\epsilon, \Delta \epsilon)$, and the alternative hypothesis is $d_m^1(\epsilon, \Delta \epsilon) \neq d_m^2(\epsilon, \Delta \epsilon)$. Then under the null hypothesis,

$$T^{1/2} \left[ d_m^1(\epsilon, \Delta \epsilon, T) - d_m^2(\epsilon, \Delta \epsilon, T) \right] / (\hat{V}D_1)^{1/2} \overset{d}{\to} N(0,1)$$

where

- $d_m^1(\epsilon, \Delta \epsilon, T) = \text{the correlation dimension estimator of the first sample}$,
- $d_m^2(\epsilon, \Delta \epsilon, T) = \text{the correlation dimension estimator of the second sample}$,
- $\hat{V}D_1 = \text{the consistent estimator of the variance, } 8\gamma^2 [ \hat{A}^m + \hat{B}^m - 2\hat{C}^m + \sum_{j=1}^{m-1} (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j}) ]$, $A$, $B$ and $C$ are defined in (2.23).

**Proof:** The proof is similar to the proof of Theorem 2.3.

For practical purposes we computed $\hat{V}D_1$ from the first sample because the two series have the same distribution under the null hypothesis. A histogram of the statistic (3.1) with 2500 iterations
which is drawn in Figure 10 it did not show downward bias because the bias factors in the numerator cancel and the $\gamma$ factors are cancelled from the numerator and the denominator of (3.1). It also showed that the sampling distribution has thin tails relative to the standard normal distribution.

Example 2: The second application is designed to test whether the given series $\{a_t\}$ is IID or not. The IID test based on the dimension estimate is constructed in a similar way as the first example. First of all a bootstrap sample is generated from the given sample with replacement. If the series $\{a_t\}$ has a chaotic attractor which shows a low dimension estimate, the difference between the original and the bootstrap sample correlation estimates should be statistically significant since the chaotic structure is destroyed by shuffling. However if $\{a_t\}$ is IID, the difference between them should not be significantly large. The formal test statistic will be the following.

Theorem 3.2: Let $d^*(E, \Delta \epsilon)$ and $d_0^*(E, \Delta \epsilon)$ be the correlation dimension of the original and the bootstrap series respectively. The null hypothesis of IID is tested by $d^*(E, \Delta \epsilon) = d_0^*(E, \Delta \epsilon)$ against the alternative hypothesis, $d^*(E, \Delta \epsilon) < d_0^*(E, \Delta \epsilon)$ to set up a one-tail test. Then under the null hypothesis,
\[ (3.2) \quad T^{1/2} [d_m(\varepsilon, \Delta \varepsilon, T) - d_m^*(\varepsilon, \Delta \varepsilon, T)] / (\hat{V}D_2)^{1/2} \xrightarrow{d} N(0,1) \]

where

\[ d_m(\varepsilon, \Delta \varepsilon, T) = \text{the correlation dimension estimator of the original sample}, \quad d_m^*(\varepsilon, \Delta \varepsilon, T) = \text{the correlation dimension estimator of the bootstrap sample}, \quad \hat{V}D_2 = \text{the consistent estimator of the variance}, \]

\[ 8\gamma^2 [\hat{A}_m^m + \hat{B}_m^m - 2\hat{C}_m^m + 2 \sum_{j=1}^{m-1} \{\hat{A}^{m-j}_m + \hat{B}^{m-j}_m - 2\hat{C}^{m-j}_m\}], \quad A, B \text{ and } C \text{ are defined in (2.23)}. \]

Proof: The proof is similar to that of Theorem 2.3.

We computed \( \hat{V}D_1 \) from the original sample because the bootstrap sample gives a close value under IID assumption. A histogram based on 2500 iterations of this experiment in Figure 11 shows no clear evidence of downward bias. The same cancelling of the \( \gamma \) factor that occurred in (3.1) also occurred in (3.2). \(^{10}\)

4. CONCLUSION AND FUTURE RESEARCH

This paper has shown that central limit theory for U-statistics under assumptions of weak dependence may be fruitfully applied to provide inference theory using objects of nonlinear science such as the correlation dimension and the approximate Kolmogorov entropy.
For example we derived asymptotic standard errors for correlation dimension estimates and estimates of approximate Kolmogorov entropy. We then estimated these quantities for stock returns. Dimension estimates appear rather unstable. Kolmogorov entropy estimates were better behaved.

The performance of the dimension estimate was poor, due to a bias in the dimension estimate itself and bias in the standard error estimate. But inference was improved by use of bias reduction techniques. U-statistic theory can also be applied to provide inference theory for measures of instability (See Appendix 3.).

Our methods are general. For example the correlation integral can be used to build tests for nonlinear "Granger/Wiener" causality which is explain in Appendix 4, as well as for "instability." This work is in progress and is touched upon in Baek and Brock (1990).
FOOTNOTES

1. "Log" denotes the natural logarithm in this paper.

2. After our work was completed Dee Dechert told us about Denker and Keller (1986). They choose 5 values $\epsilon_i = 0.08 \times 2^{-i}$, $i=0,1,2,3,4$ and write the OLS estimators of $a$, $\beta$ in the OLS regression, $\log C(\epsilon_i,m) = a + \beta \log(\epsilon_i) + \eta_i$, in U-statistic form. In this way they obtain an estimate of $\beta$ from the vector of estimates $\{C(\epsilon_i,m)\}$ and a standard error for $\beta$. They show the results are very good for a certain dynamical system on the plane.

3. The assumption of IID is not needed for any of the applications. It is used to simplify the variance formulae for the statistics to be treated below. If one imposes the mixing assumptions of Denker and Keller (1983, 1986) one can develop a variance formula like (2.13) which is an infinite sum of relevant covariances. One can then use a consistent estimator of this infinite sum to develop the general theory along the lines of the special case of IID developed here. For example the general theory can be used to estimate confidence intervals for estimates of objects like Kolmogorov entropy and dimension.

4. A continuous function of random variables which converge in distribution also converges in distribution. (Serfling, 1980, p. 24,
Theorem).

5. We thank Pedro DeLima for help with this formula. A similar independence test based upon the Kolmogorov entropy was independently developed by Hiemstra (1990). We highly recommend this excellent study to the reader. It contains not only a study of independence tests but also applications to testing the efficient markets hypothesis.

6. For the case of standardized $t$ and standard normal distributions, $\text{DIM}$ calculations are approximately equal to $m$ for each embedding dimension $m$. Technically $\text{DIM} = m[(dC(\varepsilon)/d\varepsilon)\varepsilon]/C(\varepsilon)$ for embedding dimension $m$ under the null of IID distribution. The following table reports the numerical calculations of $[(dC(\varepsilon)/d\varepsilon)\varepsilon]/C(\varepsilon)$ for the $t$ and standard normal distributions. A fat tailed $t$ distribution of degree of freedom 3 was chosen from Hsieh and LeBaron (1988a) since we assume the underlying structure of the Scheinkman and LeBaron data is approximated by the $t$ distribution.

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>$x/\sqrt{3}$ where $x\sim t(3)$</th>
<th>$x$ where $x\sim N(0,1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.9^2$</td>
<td>0.8849</td>
<td>0.8721</td>
</tr>
<tr>
<td>$0.9^3$</td>
<td>0.9052</td>
<td>0.8954</td>
</tr>
<tr>
<td>$0.9^4$</td>
<td>0.9221</td>
<td>0.9145</td>
</tr>
<tr>
<td>$0.9^4$</td>
<td>0.9362</td>
<td>0.9303</td>
</tr>
</tbody>
</table>

Therefore we use $m$ for the approximate value of $\text{DIM}$ under the given null hypothesis in Table 1 and 2. It can be shown, putting $C(\varepsilon,1)=C(\varepsilon)$, and using (2.4) that $[(dC(\varepsilon)/d\varepsilon)\varepsilon]/C(\varepsilon) \rightarrow 1$, $\varepsilon \rightarrow 0$. One can evaluate the quality of the approximation near $\varepsilon=0$ by
computing the Taylor expansion of \( \left( \frac{dC(\epsilon)}{d\epsilon} \right) / \left( \frac{C(\epsilon)}{\epsilon} \right) \) from (2.4) around \( \epsilon = 0 \). It is quite good as the table shows.

7. The true ASE of the test statistic uses numerically integrated values of \( K, C, \) and \( W \) for the calculation of

\[
ASE = \left[ 4 \gamma^2 \left\{ A^m + B^m - 2C^m + \sum_{j=1}^{m-1} (A^{m-j} + B^{m-j} - 2C^{m-j}) \right\} \right]^{1/2},
\]

where

\[
\gamma = \left[ \log(\epsilon+\Delta \epsilon) - \log(\epsilon) \right]^{-1}, \quad A = K(\epsilon+\Delta \epsilon) / C(\epsilon+\Delta \epsilon)^2, \quad B = K(\epsilon) / C(\epsilon)^2, \quad C = W(\epsilon+\Delta \epsilon, \epsilon ) / (C(\epsilon+\Delta \epsilon)C(\epsilon)).
\]

### True \( K, C, \) and \( W \) (by numerical integration)

<table>
<thead>
<tr>
<th>( \epsilon )</th>
<th>( K(\epsilon) )</th>
<th>( C(\epsilon) )</th>
<th>( W(\epsilon', \epsilon) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>0.2511</td>
<td>0.4755</td>
<td>0.2295 ((\epsilon'=0.9, \epsilon=0.9^2))</td>
</tr>
<tr>
<td>0.9^2</td>
<td>0.2098</td>
<td>0.4332</td>
<td>0.1912 ((\epsilon'=0.9^2, \epsilon=0.9^3))</td>
</tr>
<tr>
<td>0.9^3</td>
<td>0.1743</td>
<td>0.3938</td>
<td></td>
</tr>
</tbody>
</table>

Notes:

- Error function \( E(z) = (2/\sqrt{2\pi}) \int_0^z \exp(-t^2/2) \, dt \) was used to calculate \( K, C, \) and \( W \). Let \( f(z) \) be the probability density function of a standard normal random variable, i.e., \( f(z) = (1/\sqrt{2\pi}) \exp(-z^2/2) \).

Then \( C(\epsilon) = (1/2) \int_{-\infty}^{\infty} [E((x+\epsilon)/\sqrt{2}) - E((x-\epsilon)/\sqrt{2})] f(x) \, dx \), and \( K(\epsilon) = (1/4) \int_{-\infty}^{\infty} [E((x+\epsilon)/\sqrt{2}) - E((x-\epsilon)/\sqrt{2})] \left[ E((x+\epsilon)/\sqrt{2}) - E((x-\epsilon)/\sqrt{2}) \right] f(x) \, dx \).

8. The table below shows that the sample estimates of $K$, $C$, and $W$ converge to their true values consistently.

<table>
<thead>
<tr>
<th>$T$</th>
<th>$\epsilon$</th>
<th>$K$</th>
<th>$C$</th>
<th>$W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>0.9</td>
<td>0.251 (0.0052)</td>
<td>0.475 (0.0037)</td>
<td>0.228 (0.0049)</td>
</tr>
<tr>
<td>1000</td>
<td>$0.9^2$</td>
<td>0.210 (0.0047)</td>
<td>0.433 (0.0036)</td>
<td>0.190 (0.0044)</td>
</tr>
<tr>
<td>1000</td>
<td>$0.9^3$</td>
<td>0.175 (0.0042)</td>
<td>0.394 (0.0035)</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.9</td>
<td>0.252 (0.0074)</td>
<td>0.476 (0.0053)</td>
<td>0.229 (0.0070)</td>
</tr>
<tr>
<td>500</td>
<td>$0.9^2$</td>
<td>0.210 (0.0067)</td>
<td>0.434 (0.0052)</td>
<td>0.191 (0.0063)</td>
</tr>
<tr>
<td>500</td>
<td>$0.9^3$</td>
<td>0.175 (0.0060)</td>
<td>0.394 (0.0050)</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>0.9</td>
<td>0.252 (0.0104)</td>
<td>0.476 (0.0076)</td>
<td>0.229 (0.0099)</td>
</tr>
<tr>
<td>250</td>
<td>$0.9^2$</td>
<td>0.211 (0.0095)</td>
<td>0.434 (0.0073)</td>
<td>0.191 (0.0088)</td>
</tr>
<tr>
<td>250</td>
<td>$0.9^3$</td>
<td>0.176 (0.0085)</td>
<td>0.395 (0.0070)</td>
<td></td>
</tr>
</tbody>
</table>

Notes:

- $T$ is the sample size.
- $K$, $C$, and $W$ are calculated by 2500 replications. The standard errors for $K$, $C$, and $W$ are reported in parentheses.

9. Note that the bad performance of our dimension based statistic is due to the "magnification" quantity $4\gamma^2$. In a comparison test of whether true dimension estimates were significantly different this quantity could be cancelled which may lead to better performance.

10. Since there are well known alternative tests for IID besides
dimension based tests and Kolmogorov entropy based tests, the issue of comparison arises. A serious discussion of this issue is beyond the scope of this paper. A rather extensive discussion of the power and size properties of the closely related BDS test for IID is in Brock, Hsieh, and LeBaron (1991). A comparison (with moment generating function tests, Kendall’s tau, and Blum, Kiefer, Rosenblatt’s test) of the size and power properties of a vector version of the BDS test is in Baek (1988).

The Kolmogorov entropy test is treated in Hiemstra (1990). Hiemstra’s general conclusion is that the Kolmogorov entropy test performs quite poorly in comparison with the optimal test especially against weak linearly dependent alternatives in conditional mean and conditional variance. In general one must expect nonparametric tests like those treated in this paper to do poorly in power properties against specific parametric alternatives when compared with tests that are designed to be optimal against specific parametric alternatives. Based upon work with the closely related BDS test we expect the Kolmogorov entropy based test to do well against highly nonlinear alternatives that are predictable in the short term using nonlinear prediction schemes such as nearest neighbors. See Brock, W., Hsieh, D., and LeBaron, B., (1990) for the general argument and Monte Carlo evidence. A serious comparison study of the tests discussed in this paper must be left to future work.
APPENDIX

1. Proof of Theorem 2.3

Put $\epsilon' = \epsilon + \Delta \epsilon$. Applying the delta method (Serfling (1980, p. 124)) to $[D\{C(\epsilon', m, T), C(\epsilon, m, T)\} - D\{\theta(\epsilon', m), \theta(\epsilon, m)\}]$, we have

\[
(A.1) \quad dD = D_1 \frac{2}{T} E\{h_1(a_t^m, \epsilon') - \theta(\epsilon', m)\} + D_2 \frac{2}{T} E\{h_1(a_t^m, \epsilon) - \theta(\epsilon, m)\} + o_p(T^{-1/2})
\]

where

\[
D_1 = \frac{1}{\ln(\epsilon') - \ln(\epsilon)} \frac{1}{\theta(\epsilon', m)} \quad D_2 = \frac{1}{\ln(\epsilon') - \ln(\epsilon)} \frac{1}{\theta(\epsilon, m)}.
\]

The formula for the variance will be derived as in (2.13) after using the delta method. Put

\[
(A.2) \quad g_1(a_t^m) = 2[D_1\{h_1(a_t^m, \epsilon') - \theta(\epsilon', m)\} + D_2\{h_1(a_t^m, \epsilon) - \theta(\epsilon, m)\}].
\]

By Theorem 2.2, $T^{1/2} dD \xrightarrow{d} N(0, VD_m)$ where $VD_m = E[g_1(a_t^m)^2 + m^{-1} \sum \sum g_1(a_t^m)^2 g_1(a_t^m)]$. Under the IID assumption on the $\{a_t\}$ process, $\theta(\epsilon', m)$ and $\theta(\epsilon, m)$ are $C(\epsilon')^m$ and $C(\epsilon)^m$. Denote $[\log(\epsilon') - \log(\epsilon)]^{-1}$ by $\gamma$. Then $D_1, D_2$ equal $\gamma[C(\epsilon')]^{-m}, -\gamma[C(\epsilon)]^{-m}$. First, recalling

\[
h_1(a_t^m; \epsilon) = E[I(a_t^m, a_S^m; \epsilon)|a_t^m]
\]

\[
= \prod_{i=0}^{m-1} [F(a_{t+i} + \epsilon) - F(a_{t+i} - \epsilon)]
\]
where \( F(x) \equiv \text{Prob}\{a \leq x\} \), we compute

\[
E[g_1(a_t^m)^2] = 4\gamma^2 \left[ (K(\epsilon')/C(\epsilon'))^2 \right] m + \left( K(\epsilon)/C(\epsilon) \right)^2 m
2 \{ Eh_1(a_t^m, \epsilon') h_1(a_t^m, \epsilon)/C(\epsilon') C(\epsilon)^m \}. \]

Next, compute

\[
E[g_1(a_t^m)g_1(a_{t+j}^m)] = 4\gamma^2 \left[ (K(\epsilon')/C(\epsilon'))^2 \right] m-j + \left( K(\epsilon)/C(\epsilon) \right)^2 m-j
Eh_1(a_t^m, \epsilon') h_1(a_{t+j}^m, \epsilon')/(C(\epsilon') C(\epsilon))^m - Eh_1(a_t^m, \epsilon) h_1(a_{t+j}^m, \epsilon')/(C(\epsilon') C(\epsilon))^m.
\]

Then

\[
E[g_1(a_t^m)^2] + 2 \sum_{j=1}^{m-1} E[g_1(a_t^m)g_1(a_{t+j}^m)] = 4\gamma^2 \left[ (K(\epsilon')/C(\epsilon'))^2 \right] m + \left( K(\epsilon)/C(\epsilon) \right)^2 m - j
(E(\epsilon')/C(\epsilon))^m - 2 \sum \{ (K(\epsilon')/C(\epsilon'))^m - j + (K(\epsilon)/C(\epsilon))^m - j \} + R,
\]

where

\[
R = 4\gamma^2 \left[ -2/(C(\epsilon') C(\epsilon)) \right] \left[ Eh_1(a_t^m, \epsilon') h_1(a_t^m, \epsilon) + \sum_{j=1}^{m-1} Eh_1(a_t^m, \epsilon') h_1(a_{t+j}^m, \epsilon)
+ \sum_{j=1}^{m-1} Eh_1(a_t^m, \epsilon) h_1(a_{t+j}^m, \epsilon') \right].
\]

Moreover \( Eh_1(a_t^m, \epsilon') h_1(a_{t+j}^m, \epsilon) = Eh_1(a_{t+j}^m, \epsilon') h_1(a_t^m, \epsilon) \) can be shown easily. Based on this, \( R \) can be further simplified to

\[
-8\gamma^2 \left[ C(\epsilon') C(\epsilon) \right] m \left[ Eh_1(a_t^m, \epsilon') h_1(a_t^m, \epsilon) + \sum_{j=1}^{m-1} Eh_1(a_t^m, \epsilon') h_1(a_{t+j}^m, \epsilon) \right].
\]

Now let \( W(\epsilon', \epsilon) \) be \( Eh_1(a_t^m, \epsilon') h_1(a_t^m, \epsilon) \). Then \( Eh_1(a_t^m, \epsilon') h_1(a_{t+j}^m, \epsilon) = [C(\epsilon') C(\epsilon)]^j W(\epsilon', \epsilon) m-j \) by the IID assumption. Hence \( VD_m = \)

\[
4\gamma^2 \left[ (K(\epsilon')/C(\epsilon'))^2 \right] m + \left( K(\epsilon)/C(\epsilon) \right)^2 m - 2 \left[ W(\epsilon', \epsilon)/C(\epsilon') C(\epsilon) \right]^m + \sum_{j=1}^{m-1} \left[ (K(\epsilon')/C(\epsilon'))^m - j + (K(\epsilon)/C(\epsilon))^m - j \right] - 2 \left\{ W(\epsilon', \epsilon)/C(\epsilon') C(\epsilon) \right\}^{m-j}
= 4\gamma^2 \left[ A^m + B^m - 2C^m + 2 \sum_{j=1}^{m-1} \{ A^{m-j} + B^{m-j} - 2C^{m-j} \} \right]
\]

where

\( A = K(\epsilon')/C(\epsilon')^2, B = K(\epsilon)/C(\epsilon)^2, \) and \( C = W(\epsilon', \epsilon)/(C(\epsilon') C(\epsilon)) \). q.e.d.
2. Proof of Theorem 2.4

We show the asymptotic property for (2.24). Equation (2.21) is similarly derived. Under the null hypothesis of IID, $K_m(\epsilon)$ converges to its mean $-\log C(\epsilon)$. Let

$$K_m[C(\epsilon, m, T), C(\epsilon, m+1, T), C(\epsilon, 1, T)]$$

$$= \log \left[ \frac{C(\epsilon, m, T)}{C(\epsilon, m+1, T)} \right] + \log C(\epsilon, 1, T),$$

and define

$$dK = K_m[C(\epsilon, m, T), C(\epsilon, m+1, T), C(\epsilon, 1, T)] - K_m[C(\epsilon, m), C(\epsilon, m+1), C(\epsilon, 1)].$$

By Denker-Keller's (1983, p. 507) decomposition,

$$dK = K_1(2/T) \sum [h_1(a_t, \epsilon) - \theta(\epsilon, m)] + K_2(2/T) \sum [h_1(a_{t+1}, \epsilon) - \theta(\epsilon, m+1)] + K_3(2/T) \sum [h_1(a_t, \epsilon) - \theta(\epsilon, 1)] + o_p(T^{-1/2}),$$

where $K_1 = [C(\epsilon)]^{-m}$, $K_2 = [C(\epsilon)]^{-m-1}$, and $K_3 = [C(\epsilon)]^{-1}$. Let

$$g_1(a_t^m) = 2[K_1[h_1(a_t^m, \epsilon) - \theta(\epsilon, m)] + K_2[h_1(a_t^{m+1}, \epsilon) - \theta(\epsilon, m+1)] + K_3[h_1(a_t, \epsilon) - \theta(\epsilon, 1)]]$$

By the delta method $(T)^{1/2} dK \xrightarrow{d} N(0, VK'_m)$

where $VK'_m = E[g_1(a_t^m)^2 + 2 \sum_{j=1}^{m-1} g_1(a_t^m)g_1(a_{t+j}^m)]$. We can easily show that

$$E[g_1(a_t^m)^2] = 4[(K(\epsilon)/C(\epsilon)^2)^{m+1} - \{K(\epsilon)/C(\epsilon)^2\}^m + \{K(\epsilon)/C(\epsilon)^2 - 1\}]$$

and $E[g_1(a_t^m)g_1(a_{t+j}^m)] = 0$ for $j=1, \ldots, m-1$. q.e.d.
3. Instability Measure

In the study of algorithms to compute Lyapunov exponents (Eckmann and Ruelle (1985), and Wolf et al. (1985)), quantities that measure the rate of spreading of two nearby trajectories emerge. For example

\[(A.5) \quad R(t,s,q,m) = \frac{\|a_t^m - a_s^m\|}{\|a_t^m - a_s^m\|}.\]

The idea is that if \(\{a_t\}\) is generated by a deterministic chaos the ratio \(R\) should be greater than one and increase with \(q\) for small \(\|a_t^m - a_s^m\|\). Here \(\|\cdot\|\) denotes a norm on \(m\)-dimensional space. This suggests testing that the quantities

\[(A.6) \quad R(q,m,\epsilon) = \mathbb{E}\{R(t,s,q,m) \mid \|a_t^m - a_s^m\| \leq \epsilon\},\]

\[(A.7) \quad R(t,q,m,\epsilon) = \mathbb{E}\{R(t,s,q,m) \mid \|a_t^m - a_s^m\| \leq \epsilon, a_t^m\}\]

are bigger than one and increase in \(q\) for \(\epsilon\) small and for \(m\) large enough to "reconstruct" the dynamics (if they are chaotic).

Unfortunately, as we will see below, there is a tendency for \(R\) to be bigger than one even if the data is IID. Therefore we need some way of testing whether \(R\) is "significantly" greater than one in order to adduce evidence of instability or chaos.

The quantity \(R(q,m,\epsilon)\) measures the average rate of local spreading over all pairs \(a_t^m, a_s^m\) whereas the quantity \(R(t,q,m,\epsilon)\)
measures the average rate of local spreading over trajectories passing through the $\epsilon$-neighborhood of $a^m_t$.

A natural benchmark for measuring instability at resolution level $\epsilon$ is to compare $R$ for your data to $R$ for an IID sequence with the same stationary distribution as your data and see if they are "significantly" different. One could also test for a change in the level of instability by computing $R$ in (A.7) at two different histories $a^m_t$, $a^m_r$ and see if it is significantly different. This requires statistical theory for $R$ which we provide below. First we need to estimate $R$. Ideally we would like to provide a theory of statistical inference for Lyapunov exponent estimation algorithms such as Wolf et al. (1985) and Eckmann et al. (1986) but the distribution theory of such estimates is difficult to work out. While using Efron's (1982) bootstrap under the null of IID with the same stationary distribution is a plausible strategy this is beyond the scope of the current paper. Therefore we will start with the simpler quantities $R$ which, nevertheless, capture some of the spirit of Lyapunov exponent estimation. Let us take care of (A.7) first.

Consider the statistic, $S(t,q,m,\epsilon,T)$ defined by

$$\mathbb{E}\{N(t,q,m,\epsilon,T)/\mathbb{E}\{D(t,1,m,\epsilon,T)\}\}$$

where

(A.8) $N(t,q,m,\epsilon,T)=\{R(t,s,q,m)I(a^m_t,a^m_s;\epsilon)\}$,

(A.9) $D(t,1,m,\epsilon,T)=\{I(a^m_t,a^m_s;\epsilon)\}$
where \( I(a,b;\epsilon) = 1 \) whenever \(||a-b|| \leq \epsilon\), and 0 otherwise. Here, each \( \Sigma \) runs from \( s=1,...,T_m \). Note that since \( a^m_t \) is fixed both \( N \) and \( D \) can be written, under the null that \( \{a_t\} \) is IID, as a function of the \( m-1 \) dependent process \( \{a^m_s\} \). Note also that the statistic \( S \) is a ratio of averages. We are now set up to use the delta method and Denker and Keller's Theorem 2.2. But there is a practical problem that must be taken care of before we proceed.

The ratio \( R \) causes practical problems of computational instability for pairs of points \( a^m_t, a^m_s \) that are extremely close to each other. For such a point pair the ratio will be huge and will overwhelm all of the other ratios in the computation of the statistic. Furthermore even in the case of a deterministic chaos with small noise we would expect to draw a few point pairs that are extremely close to each other solely by chance. We fix this by computing \( R \) only for pairs \( a^m_t, a^m_s \) such that \( \eta<||a^m_t-a^m_s||<\epsilon \) where \( \eta \) is small relative to \( \epsilon \) but \( \epsilon \) is still small enough to pick up systematic local spreading of nearby trajectories. Obviously this requires experimentation or commitment to a specific \( H_0, H_A \) pair of null and alternative hypotheses so that a theory of choice of \( \eta, \epsilon \) to maximize power of the test of \( H_0 \) against \( H_A \) can be developed. Here we content ourselves with a study of the value of \( R \) under the null that \( \{a_t\} \) is IID.

Consider the following string of inequalities: For \( q \geq m \), for \( \Gamma \) any event dependent upon \( a^m_t, a^m_s \), under the null that \( \{a_t\} \) is
IID, we have

\[(A.10) \quad E\{\|a_t^m - a_s^m\| / \|a_t^m - a_s^m\| \mid \Gamma\} \]

\[= E\{\|a_t^m - a_s^m\| \} E\{1 / \|a_t^m - a_s^m\| \mid \Gamma\} \].

But, by Jensen's inequality,

\[(A.11) \quad E\{1 / \|a_t^m - a_s^m\| \mid \Gamma\} \geq 1 / E\{\|a_t^m - a_s^m\| \mid \Gamma\} \].

Apply (A.11) to get a lower bound for the R.H.S. of (A.10) under the null of IID for \(q \geq m\). For events \(\Gamma \equiv \{0 < c < \|a_t^m - a_s^m\| < d < \infty\}\) but with \(d\) small this lower bound can be arbitrarily large. Show this by taking \(c, d\) close to zero but holding the ratio constant. Finish it off with a l'Hospital's rule argument. This makes the point that Lyapunov-like estimates can be very large and positive even for random numbers.

Some preliminary discussion of this problem was in Brock (1986).

In the case that \(q < m\) there is overlap between the numerator and the denominator of \(R\) that does not disappear asymptotically. This causes asymptotic dependence which destroys the simple argument made above. Nevertheless the point is made that, in many cases, \(R\) is bigger than one, on average. Let us set up a test of the null hypothesis of IID using observed values of \(R\). We shall stick to the case \(q \geq m\). This test of IID should have power against dependent alternatives like Blanchard and Watson (1982) quasi rational bubbles as well as "unstable" alternatives.
**Theorem A.1:** Assume that \( \{a_t\} \) is IID, \( q \geq m \). Consider the statistic

\[
(A.12) \quad S_1(q) = (T_m)^{1/2}[\Sigma N_2/\Sigma D_1 - (\Sigma N_1/T_m)(\Sigma N_3/\Sigma D_1)],
\]

where,

\[
D_1 = I(a^m_t, a^m_s; \epsilon), \quad D_2 = \|a^m_t - a^m_s\|, \quad N_1 = \|a^m_{t+q} - a^m_s + q\|, \quad N_2 = (D_1 N_1)/\|a^m_t - a^m_s\|
\]

\[
= D_1 N_1/D_2, \quad N_3 = D_1/\|a^m_t - a^m_s\| = D_1/D_2.
\]

Here the m-history \( a^m_t \) is fixed so that the \( \Sigma \)'s run from \( s = 1, 2, \ldots, T_m \) in \( S_1(q) \). Denote by \( S_2(q) \) the same statistic in (A.12) except that all \( \Sigma \)'s are double sums so that \( S_2(q) \) is a function of \( U \)-statistics of order two. Then under the null hypothesis \( S_1(q) \) and \( S_2(q) \) both converges to random variables with finite variance.

**Remark:** The proof follows from Denker and Keller's Theorem 2.2 and the delta method. Since the expression for the variance is very messy we are bootstrapping the variance in actual applications that are underway. The procedure follows Freedman and Peters (1984) and represents research in progress that is beyond the scope of the current paper.

4. Nonlinear Granger/Wiener Causality
As we saw in the text $K_m = \log(C(\epsilon, m)/C(\epsilon, m+1)) = -\log C(\epsilon, 1)$ if and only if $C(\epsilon, m+1)/C(\epsilon, m) = C(\epsilon, 1)$, i.e., $\text{Prob}(X_{t+1}|X_t, \ldots, X_{t-(m-1)}) = \text{Prob}(X_{t+1})$ where "$X_t$" is shorthand for the event $|X_t - X_{t-1}| \leq \epsilon$. Hence $K_m = -\log(C(\epsilon, 1))$ precisely when two $m$-histories being close i.e.,

$$
\|(X_t-X_s, \ldots, X_{t-(m-1)}-X_{s-(m-1)})\| \leq \epsilon
$$

predict nothing about $|X_{t+1}-X_{s+1}| \leq \epsilon$. Thus when the null hypothesis of IID is rejected by the test based on $K_m$ a statement is made that close $m$-pasts tend to be followed by close descendents. This idea can be generalized to test whether one series $\{Y_t\}$ helps predict another series $\{X_t\}$. Turn now to a brief discussion how the correlation integral can be 

generalized to do this.

Many conventional time series tests for "Granger/Wiener" causality (Geweke (1984, p. 1102)) are restricted to the linear class of models. If two variables have a linear relationship these tests can detect it. However if two variables are related in a nonlinear way to each other, the conventional tests may have weak power since the class of models against which they have strong power is linear. Now when we are equipped with the correlation integral in the vector case and the asymptotic developed in this paper, a general class of Granger/Wiener type causality, can be formulated. Even though our work is at the early stage, the main idea can be described in the following way.

The null hypothesis is that "$Y$ does not Granger cause $X$" which we formulate as, letting $Y_t$ be shorthand for the event $|Y_t - Y_s| \leq \epsilon$, 
(A.13) \[ \text{Prob}(X_{t+1}|X_t, X_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots) = \text{Prob}(X_{t+1}|X_t, X_{t-1}, \ldots). \]

By definition of conditional probability, (A.13) can be written as

(A.14) \[ \frac{\text{Prob}(X_{t+1}, X_t, \ldots, Y_t, Y_{t-1}, \ldots)}{\text{Prob}(X_t, X_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots)} = \frac{\text{Prob}(X_{t+1}, X_t, \ldots)}{\text{Prob}(X_t, X_{t-1}, \ldots)}. \]

From a sample of length T, estimate each "Prob" by the corresponding correlation integral, take the difference, multiply by \( \sqrt{T} \), and finally take the limit to get the distribution under the null hypothesis. One has to assume mixing conditions as in Denker and Keller (1983). When "Prob" is replaced by the corresponding correlation integral, the null hypothesis (A.14) can be written as

(A.15) \[ C(X_{t+1}, X_t, X_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots)C(X_t, X_{t-1}, \ldots) - C(X_t, X_{t-1}, \ldots, Y_t, \ldots, Y_{t-1}, \ldots)C(X_{t+1}, X_t, \ldots) = 0. \]

For the practical implementation of the test, we truncate X at \( t_i \) and Y at \( t_j \). In order to simplify the variance formula we replace the null by: \( \{X_t\}, \{Y_t\} \) IID and independent. We abbreviate this null by \IIDI. Then

(A.16) \[ T^{1/2} \left[ \hat{C}(X_{t+1}, X_t, \ldots, X_{t-i}, Y_t, \ldots, Y_{t-j})\hat{C}(X_t, X_{t-1}, \ldots, X_{t-i}) - \hat{C}(X_t, X_{t-1}, \ldots, Y_t, \ldots, Y_{t-j})\hat{C}(X_{t+1}, X_t, \ldots, X_{t-i}) \right] \xrightarrow{d} \text{N}(0, V) \]
where $V = 4(C_x^{i+2} - (K_x^{i+1})^2[K_x^j - (C_x^j)^2])$, 

where $C_x = EI(X_t, X_s; \varepsilon)$, $C_y = EI(Y_t, Y_s; \varepsilon)$, $K_x = EI(X_t, X_s; \varepsilon)I(X_s, X_r; \varepsilon)$, $K_y = EI(Y_t, Y_s; \varepsilon)I(Y_s, Y_r; \varepsilon)$, and $\hat{\cdot}$ denotes the sample estimate.

Note that in order to get the simple variance formula (A.16) we replaced the null we were interested in by IID. However we constructed the test of IID so that it has zero power against dependence of $X_{t+1}$ on past X’s, but has power against dependence of $X_{t+1}$ on past Y’s. Monte Carlo work is probably the most effective way of evaluating whether we made a wise move in trading off a complicated variance formula and an accurate null for a simple variance formula and a less accurate null. This is all "research in progress". We mention it here only because it illustrates the general theme of this paper: The measure of spatial correlation from chaos theory, called the correlation integral, can be combined with U-statistics theory to produce useful statistical tests for the presence of nonlinear predictability.
REFERENCES


BAEK, E. and BROCK, W.A. (1990), "A General Test for Nonlinear Granger Causality" (Department of Economics, Iowa State University and Department of Economics, University of Wisconsin-Madison).


BROCK, W.A. and DECHERT, W.D. (1988a), "Theorems on Distinguishing


EFRON, B. (1982), *The Jacknife, the Bootstrap, and Other Resampling
**Planes** (Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania).


GENNOTTE, G. and MARSH, T. (1987), "Variations in Economic Uncertainty and Risk Premiums on Capital Assets" (School of Business Administration, University of California, Berkeley).


HSIEH, D. and LEBARON, B. (1988a), "Finite Sample Properties of the


Note: All Monte Carlo experiments are based on 2500 iterations in the following tables and figures.

CONTENTS
Table 1: Correlation dimension estimates of weekly returns
Table 2: Correlation dimension estimates of standard normal random numbers
Table 3: Kolmogorov entropy estimates of weekly stock returns
Table 4: Kolmogorov entropy estimates of standard normal random numbers
Table 5: Correlation dimension estimates and standard errors
   sample size=1000; ε+Δε=0.9; ε=0.9
Table 6: Kolmogorov entropy estimates and standard errors
   sample size=1000; ε=0.9
Table 7: Correlation dimension estimates and standard errors
   sample size=500; ε+Δε=0.9; ε=0.9
Table 8: Kolmogorov entropy estimates and standard errors
   sample size=500; ε=0.9
Table 9: Correlation dimension estimates and standard errors
   sample size=250; ε+Δε=0.9; ε=0.9
Table 10: Kolmogorov entropy estimates and standard errors
   sample size=250; ε=0.9
Table 11: Correlation dimension estimates and standard errors
   sample size=1000; ε+Δε=0.9; ε=0.9
Table 12: Kolmogorov entropy estimates and standard errors
   sample size=1000; ε=0.9
Table 13: Correlation dimension estimates and standard errors
   sample size=500; ε+Δε=0.9; ε=0.9
Table 14: Kolmogorov entropy estimates and standard errors
   sample size=500; ε=0.9
Table 15: Correlation dimension estimates and standard errors
   sample size=250; ε+Δε=0.9; ε=0.9
Table 16: Kolmogorov entropy estimates and standard errors
   sample size=250; ε=0.9
Figure 1: Histogram of the Correlation dimension estimates
   sample size=1000; ε+Δε=0.9; ε=0.9; m=2
Figure 2: Histogram of the Kolmogorov entropy estimates
   sample size=1000, ε=0.9; m=2
Figure 3: Histogram of the Correlation dimension estimates
   sample size=1000; ε+Δε=0.9; ε=0.9; m=4
Figure 4: Histogram of the Kolmogorov entropy estimates
   sample size=1000, ε=0.9; m=4
sample size=1000; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=8
Histogram of the Kolmogorov entropy estimates
sample size=1000, \epsilon=0.9; m=8

Figure 4: Histogram of the Correlation dimension estimates
sample size=500; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=2
Histogram of the Kolmogorov entropy estimates
sample size=500, \epsilon=0.9; m=2

Figure 5: Histogram of the Correlation dimension estimates
sample size=500; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=4
Histogram of the Kolmogorov entropy estimates
sample size=500, \epsilon=0.9; m=4

Figure 6: Histogram of the Correlation dimension estimates
sample size=500; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=8
Histogram of the Kolmogorov entropy estimates
sample size=500, \epsilon=0.9; m=8

Figure 7: Histogram of the Correlation dimension estimates
sample size=250; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=2
Histogram of the Kolmogorov entropy estimates
sample size=250, \epsilon=0.9; m=2

Figure 8: Histogram of the Correlation dimension estimates
sample size=250; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=4
Histogram of the Kolmogorov entropy estimates
sample size=250, \epsilon=0.9; m=4

Figure 9: Histogram of the Correlation dimension estimates
sample size=250; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=8
Histogram of the Kolmogorov entropy estimates
sample size=250, \epsilon=0.9; m=8

Figure 10: Histogram of the difference of two dimension estimates
sample size=500; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=6

Figure 11: Histogram of the difference of original and bootstrap samples
sample size=500; \epsilon+\Delta \epsilon=0.9; \epsilon=0.9^2; m=6
### Table 1

**CORRELATION DIMENSION ESTIMATES OF WEEKLY STOCK RETURNS**

<table>
<thead>
<tr>
<th>m</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.82</td>
<td>0.060</td>
<td>-3.00</td>
<td>0.86</td>
<td>0.069</td>
<td>-2.03</td>
</tr>
<tr>
<td>6</td>
<td>4.28</td>
<td>0.451</td>
<td>-3.81</td>
<td>4.42</td>
<td>0.526</td>
<td>-3.00</td>
</tr>
<tr>
<td>8</td>
<td>5.24</td>
<td>0.663</td>
<td>-4.16</td>
<td>5.36</td>
<td>0.778</td>
<td>-3.39</td>
</tr>
<tr>
<td>10</td>
<td>5.89</td>
<td>0.915</td>
<td>-4.49</td>
<td>5.96</td>
<td>1.081</td>
<td>-3.74</td>
</tr>
<tr>
<td>12</td>
<td>6.55</td>
<td>1.214</td>
<td>-4.49</td>
<td>6.53</td>
<td>1.445</td>
<td>-3.79</td>
</tr>
<tr>
<td>14</td>
<td>7.19</td>
<td>1.570</td>
<td>-4.34</td>
<td>7.09</td>
<td>1.882</td>
<td>-3.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.88</td>
<td>0.077</td>
<td>-1.56</td>
<td>0.89</td>
<td>0.083</td>
<td>-1.33</td>
</tr>
<tr>
<td>6</td>
<td>4.52</td>
<td>0.588</td>
<td>-2.52</td>
<td>5.01</td>
<td>0.647</td>
<td>-1.53</td>
</tr>
<tr>
<td>8</td>
<td>5.53</td>
<td>0.876</td>
<td>-2.82</td>
<td>6.34</td>
<td>0.968</td>
<td>-1.71</td>
</tr>
<tr>
<td>10</td>
<td>6.46</td>
<td>1.225</td>
<td>-2.89</td>
<td>7.50</td>
<td>1.362</td>
<td>-1.84</td>
</tr>
<tr>
<td>12</td>
<td>7.52</td>
<td>1.649</td>
<td>-2.72</td>
<td>7.31</td>
<td>1.845</td>
<td>-2.54</td>
</tr>
<tr>
<td>14</td>
<td>8.47</td>
<td>2.162</td>
<td>-2.56</td>
<td>7.73</td>
<td>2.434</td>
<td>-2.58</td>
</tr>
</tbody>
</table>

**Notes:**

- 1226 standardized weekly stock returns are used, i.e. \( z_i = (r_i - \bar{r}) / \hat{\sigma} \) for \( i = 1, 2, \ldots, 1226 \) where \( \bar{r} = \frac{\sum r_i}{T}, \hat{\sigma} = \sqrt{\frac{\Sigma (r_i - \bar{r})^2}{T - 1}} \) and \( T = 1226 \).
- \( m \) is the embedding dimension, \( \text{DIM} \) is the dimension estimate, and \( \text{ASE} \) is the empirical asymptotic standard error from the data

\[
\text{ASE} = \left[ 4 \gamma^2 \left\{ A^m + B^m - 2C^m + 2 \sum_{j=1}^{m-1} (A^{m-j} + B^{m-j} - 2C^{m-j}) \right\} \right]^{1/2}
\]

where

\[
\gamma = \left[ \log(\epsilon + \Delta \epsilon) - \log(\epsilon) \right]^{-1}, \quad A = \frac{K(\epsilon + \Delta \epsilon)}{C(\epsilon + \Delta \epsilon)}, \quad B = \frac{K(\epsilon)}{C(\epsilon)}, \quad C = W(\epsilon + \Delta \epsilon, \epsilon) / (C(\epsilon + \Delta \epsilon)C(\epsilon)), \quad W(\epsilon + \Delta \epsilon, \epsilon) = EI(a_i, a_j; \epsilon + \Delta \epsilon)l(a_i, a_j; \epsilon).
\]

- \( \text{TEST} \) is the test statistic, i.e. \( \text{TEST} = \sqrt{T} (\text{DIM} - m) / \text{ASE} \).
Table 2

CORRELATION DIMENSION ESTIMATES OF STANDARD NORMAL RANDOM NUMBERS

<table>
<thead>
<tr>
<th>m</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.89</td>
<td>0.063</td>
<td>-1.75</td>
<td>0.91</td>
<td>0.067</td>
<td>-1.34</td>
</tr>
<tr>
<td>2</td>
<td>1.78</td>
<td>0.128</td>
<td>-1.72</td>
<td>1.82</td>
<td>0.138</td>
<td>-1.30</td>
</tr>
<tr>
<td>6</td>
<td>5.29</td>
<td>0.438</td>
<td>-1.62</td>
<td>5.66</td>
<td>0.474</td>
<td>-0.72</td>
</tr>
<tr>
<td>8</td>
<td>7.00</td>
<td>0.625</td>
<td>-1.60</td>
<td>8.30</td>
<td>0.678</td>
<td>0.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
<th>DIM</th>
<th>ASE/√T</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.94</td>
<td>0.072</td>
<td>-0.83</td>
<td>0.93</td>
<td>0.074</td>
<td>-0.95</td>
</tr>
<tr>
<td>2</td>
<td>1.86</td>
<td>0.147</td>
<td>-0.95</td>
<td>1.88</td>
<td>0.153</td>
<td>-0.78</td>
</tr>
<tr>
<td>6</td>
<td>5.59</td>
<td>0.509</td>
<td>-0.81</td>
<td>4.97</td>
<td>0.528</td>
<td>-1.95</td>
</tr>
<tr>
<td>8</td>
<td>8.29</td>
<td>0.730</td>
<td>0.40</td>
<td>5.13</td>
<td>0.760</td>
<td>-3.78</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNOA was called to generate 1226 standard normal random numbers, and RNSET was called to set an initial seed.
- m is the embedding dimension, DIM is the dimension estimate, and ASE is the empirical asymptotic standard error from the data
  \[ ASE = \left[ 4\gamma^2 \{ A_m + B_m - 2C_m + 2 \sum_{j=1}^{m-1} (A_{m-j} + B_{m-j} - 2C_{m-j}) \} \right]^{1/2}, \]
  \[ \gamma = \frac{[\log(\epsilon+\Delta \epsilon) - \log(\epsilon)]^{-1}}{\log(\epsilon+\Delta \epsilon)}, \]
  \[ A = K(\epsilon+\Delta \epsilon)/C(\epsilon+\Delta \epsilon)^2, \]
  \[ B = K(\epsilon)/C(\epsilon)^2, \]
  \[ C = W(\epsilon+\Delta \epsilon, \epsilon)/(C(\epsilon+\Delta \epsilon)C(\epsilon)), \]
  \[ W(\epsilon+\Delta \epsilon, \epsilon) = EI(a_i, a_j; \epsilon+\Delta \epsilon)I(a_i, a_j; \epsilon). \]
- TEST is the test statistic, i.e. \[ TEST = \sqrt{DIM-m}/ASE. \]
- Less than 0.01% of the total number of pairs are available to calculate the dimension estimates beyond \( m = 8 \).
### Table 3

**KOLMOGOROV ENTROPY ESTIMATES OF WEEKLY STOCK RETURNS**

<table>
<thead>
<tr>
<th>$\epsilon=0.9$</th>
<th>$\epsilon=0.9^2$</th>
<th>$\epsilon=0.9^3$</th>
<th>$\epsilon=0.9^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>ENTROPY</td>
<td>ASE/$\sqrt{T}$</td>
<td>TEST</td>
</tr>
<tr>
<td>1</td>
<td>0.592</td>
<td>0.031</td>
<td>-1.74</td>
</tr>
<tr>
<td>6</td>
<td>0.405</td>
<td>0.038</td>
<td>-6.34</td>
</tr>
<tr>
<td>8</td>
<td>0.360</td>
<td>0.042</td>
<td>-6.84</td>
</tr>
<tr>
<td>10</td>
<td>0.306</td>
<td>0.046</td>
<td>-7.37</td>
</tr>
<tr>
<td>12</td>
<td>0.269</td>
<td>0.051</td>
<td>-7.35</td>
</tr>
<tr>
<td>13</td>
<td>0.254</td>
<td>0.054</td>
<td>-7.23</td>
</tr>
</tbody>
</table>

**Notes:**

- 1226 standardized weekly stock returns are used, i.e. $z_i=(r_i-\bar{r})/\hat{\sigma}$ for $i=1,2,\ldots,1226$ where $\bar{r}=E(r_i)/T$, $\hat{\sigma}=E((r_i-\bar{r})^2)/(T-1)$ and $T=1226$.
- $m$ is the embedding dimension, ENTROPY is the entropy estimate, and ASE is the empirical asymptotic standard error from the data
  $$ASE=\left[4\{(K(\epsilon)/C(\epsilon)^2)^{m+1}-(K(\epsilon)/C(\epsilon)^2)^m+K(\epsilon)/C(\epsilon) -1\}\right]^{1/2}.$$  
- TEST is the test statistic, i.e.
  $$TEST=\sqrt{T}[\text{ENTROPY}+\log(C(\epsilon,1,T))]/\text{ASE}.$$
Table 4

<table>
<thead>
<tr>
<th>m</th>
<th>ENTROPY</th>
<th>ASE/$\sqrt{T}$</th>
<th>TEST</th>
<th>ENTROPY</th>
<th>ASE/$\sqrt{T}$</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.758</td>
<td>0.026</td>
<td>-0.05</td>
<td>0.852</td>
<td>0.027</td>
<td>-0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.756</td>
<td>0.027</td>
<td>-0.12</td>
<td>0.850</td>
<td>0.028</td>
<td>-0.10</td>
</tr>
<tr>
<td>6</td>
<td>0.756</td>
<td>0.030</td>
<td>-0.11</td>
<td>0.840</td>
<td>0.032</td>
<td>-0.40</td>
</tr>
<tr>
<td>8</td>
<td>0.751</td>
<td>0.032</td>
<td>-0.25</td>
<td>0.892</td>
<td>0.034</td>
<td>1.17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>m</th>
<th>ENTROPY</th>
<th>ASE/$\sqrt{T}$</th>
<th>TEST</th>
<th>ENTROPY</th>
<th>ASE/$\sqrt{T}$</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.948</td>
<td>0.028</td>
<td>-0.02</td>
<td>1.046</td>
<td>0.028</td>
<td>-0.06</td>
</tr>
<tr>
<td>2</td>
<td>0.945</td>
<td>0.029</td>
<td>-0.14</td>
<td>1.040</td>
<td>0.029</td>
<td>-0.24</td>
</tr>
<tr>
<td>6</td>
<td>0.956</td>
<td>0.033</td>
<td>0.22</td>
<td>1.070</td>
<td>0.033</td>
<td>0.68</td>
</tr>
<tr>
<td>8</td>
<td>1.154</td>
<td>0.035</td>
<td>5.88</td>
<td>1.150</td>
<td>0.036</td>
<td>2.86</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 1000 standard normal random numbers, and RNSET was called to set an initial seed.
- m is the embedding dimension, ENTROPY is the entropy estimate, and ASE is the empirical asymptotic standard error from the data
  \[ ASE = \left[ 4 \left( \frac{K(\epsilon)}{C(\epsilon)} \right)^{m+1} - \left( \frac{K(\epsilon)}{C(\epsilon)} \right)^m + \frac{K(\epsilon)}{C(\epsilon)} - 1 \right]^{1/2}. \]
- TEST is the test statistic, i.e.
  \[ TEST = \sqrt{T} \left[ \frac{\text{ENTROPY} + \log(C(\epsilon, m, T))}{\text{ASE}} \right]. \]
- Less than 0.01% of the total number of pairs are available to calculate the Kolmogorov Entropy estimates beyond m=8.
Table 5: Monte Carlo Simulation

CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS
Number of Replications = 2500; Sample Size = 1000; $\epsilon+\Delta \epsilon = 0.9$, $\epsilon = 0.9^2$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Dim Est</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.885 (0.008)</td>
<td>2.221 (0.030)</td>
<td>0.526</td>
<td>0.246</td>
</tr>
<tr>
<td>2</td>
<td>1.770 (0.019)</td>
<td>4.566 (0.066)</td>
<td>1.087</td>
<td>0.600</td>
</tr>
<tr>
<td>3</td>
<td>2.655 (0.036)</td>
<td>7.080 (0.110)</td>
<td>1.696</td>
<td>1.131</td>
</tr>
<tr>
<td>4</td>
<td>3.540 (0.063)</td>
<td>9.776 (0.166)</td>
<td>2.359</td>
<td>1.980</td>
</tr>
<tr>
<td>5</td>
<td>4.425 (0.102)</td>
<td>12.667 (0.236)</td>
<td>3.077</td>
<td>3.235</td>
</tr>
<tr>
<td>6</td>
<td>5.310 (0.166)</td>
<td>15.765 (0.321)</td>
<td>3.857</td>
<td>5.241</td>
</tr>
<tr>
<td>7</td>
<td>6.198 (0.271)</td>
<td>19.084 (0.425)</td>
<td>4.702</td>
<td>8.561</td>
</tr>
<tr>
<td>8</td>
<td>7.083 (0.437)</td>
<td>22.640 (0.549)</td>
<td>5.617</td>
<td>13.822</td>
</tr>
<tr>
<td>9</td>
<td>7.978 (0.701)</td>
<td>26.448 (0.897)</td>
<td>6.609</td>
<td>22.161</td>
</tr>
<tr>
<td>10</td>
<td>8.906 (1.115)</td>
<td>30.525 (0.869)</td>
<td>7.682</td>
<td>35.271</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 1000 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE's,
  \[
  ASE = [4 \gamma^2 \{\hat{A}^m + \hat{B}^m - 2C^m + 2 \sum_{j=1}^{m-1} \hat{A}^m \hat{B}^m \hat{C}^m \hat{D}^m \}]^{1/2},
  \]
  where
  \[
  \gamma = \left[ \log(\epsilon + \Delta \epsilon) - \log(\epsilon) \right]^{-1}, \quad \hat{A} = K(\epsilon + \Delta \epsilon)/C(\epsilon + \Delta \epsilon), \quad \hat{B} = K(\epsilon)/C(\epsilon),
  \]
  \[
  C = W(\epsilon + \Delta \epsilon, \epsilon)/(C(\epsilon + \Delta \epsilon)C(\epsilon)).
  \]
  Standard normal random numbers of size of 1000 were used to calculate $K(\epsilon + \Delta \epsilon)$, $K(\epsilon)$, $C(\epsilon + \Delta \epsilon)$, $C(\epsilon)$ and $W(\epsilon + \Delta \epsilon, \epsilon)$ at which $\hat{A}$, $\hat{B}$ and $\hat{C}$ are evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But $\hat{A}$, $\hat{B}$ and $\hat{C}$ were evaluated at numerically calculated values of $K(\epsilon + \Delta \epsilon) = 0.2511$, $K(\epsilon) = 0.2098$, $C(\epsilon + \Delta \epsilon) = 0.4755$, $C(\epsilon) = 0.4332$, and $W(\epsilon + \Delta \epsilon, \epsilon) = 0.2295$ for $\epsilon + \Delta \epsilon = 0.9$ and $\epsilon = 0.9^2$.
- $\sqrt{T}$ SD = $\sqrt{T} \times$ standard error of $[D(\cdot) - m]$ of the 2500 replications.
Table 6: Monte Carlo Simulation

KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS

Number of Replications=2500; Sample Size=1000; \( \varepsilon = 0.9 \)

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy Mean ASE</th>
<th>True ASE</th>
<th>( \sqrt{T} ) SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0006 (0.0073)</td>
<td>0.967 (0.029)</td>
<td>0.966 (0.029)</td>
</tr>
<tr>
<td>2</td>
<td>0.0011 (0.0105)</td>
<td>0.994 (0.031)</td>
<td>0.994 (0.031)</td>
</tr>
<tr>
<td>3</td>
<td>0.0013 (0.0135)</td>
<td>1.024 (0.034)</td>
<td>1.024 (0.034)</td>
</tr>
<tr>
<td>4</td>
<td>0.0012 (0.0164)</td>
<td>1.057 (0.037)</td>
<td>1.057 (0.037)</td>
</tr>
<tr>
<td>5</td>
<td>0.0019 (0.0196)</td>
<td>1.092 (0.041)</td>
<td>1.092 (0.041)</td>
</tr>
<tr>
<td>6</td>
<td>0.0019 (0.0238)</td>
<td>1.129 (0.044)</td>
<td>1.129 (0.044)</td>
</tr>
<tr>
<td>7</td>
<td>0.0025 (0.0295)</td>
<td>1.169 (0.049)</td>
<td>1.169 (0.049)</td>
</tr>
<tr>
<td>8</td>
<td>0.0031 (0.0379)</td>
<td>1.212 (0.053)</td>
<td>1.212 (0.053)</td>
</tr>
<tr>
<td>9</td>
<td>0.0053 (0.0512)</td>
<td>1.259 (0.059)</td>
<td>1.259 (0.059)</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNDNOA was called to generate 1000 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate \([K_m(\cdot) + \log C(\varepsilon,1,T)]\) of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's,
  \[\text{ASE} = \sqrt{\frac{4\{K(\varepsilon)/C(\varepsilon)^2\}^{m+1} - (K(\varepsilon)/C(\varepsilon)^2)^m + K(\varepsilon)/C(\varepsilon)^2 - 1\}}}{1/2}.\]
  Standard normal random numbers of size of 1000 were used to calculate \(K(\varepsilon), C(\varepsilon)\) at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of \(K(\varepsilon)=0.2511, C(\varepsilon)=0.4755\) for \(\varepsilon=0.9\).
- \(\sqrt{T} \text{ SD} = \sqrt{T} \times \text{standard error of } [K_m(\cdot) + \log C(\varepsilon,1,T)] \) of the 2500 replications.
### Table 7: Monte Carlo Simulation

**CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS**

Number of Replications = 2500; Sample Size = 500; \( \epsilon+\Delta \epsilon = 0.9 \), \( \epsilon = 0.9^2 \)

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Dim Est</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>( \sqrt{T} ) SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.885 (0.014)</td>
<td>2.216 (0.052)</td>
<td>1.526</td>
<td>0.320</td>
</tr>
<tr>
<td>2</td>
<td>1.771 (0.035)</td>
<td>4.566 (0.111)</td>
<td>1.087</td>
<td>0.785</td>
</tr>
<tr>
<td>3</td>
<td>2.658 (0.068)</td>
<td>7.063 (0.182)</td>
<td>1.696</td>
<td>1.516</td>
</tr>
<tr>
<td>4</td>
<td>3.546 (0.121)</td>
<td>9.776 (0.268)</td>
<td>2.359</td>
<td>2.705</td>
</tr>
<tr>
<td>5</td>
<td>4.434 (0.201)</td>
<td>12.641 (0.236)</td>
<td>3.077</td>
<td>4.484</td>
</tr>
<tr>
<td>6</td>
<td>5.331 (0.331)</td>
<td>15.735 (0.372)</td>
<td>3.857</td>
<td>7.410</td>
</tr>
<tr>
<td>7</td>
<td>6.229 (0.544)</td>
<td>19.051 (0.498)</td>
<td>4.702</td>
<td>12.168</td>
</tr>
<tr>
<td>8</td>
<td>7.135 (0.881)</td>
<td>22.605 (0.648)</td>
<td>5.617</td>
<td>19.702</td>
</tr>
<tr>
<td>9</td>
<td>8.066 (1.410)</td>
<td>26.412 (0.827)</td>
<td>6.609</td>
<td>31.516</td>
</tr>
<tr>
<td>10</td>
<td>9.088 (2.291)</td>
<td>30.525 (1.283)</td>
<td>7.682</td>
<td>51.237</td>
</tr>
</tbody>
</table>

**Notes:**

- IMSL subroutine DRNNOA was called to generate 500 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE’s,
  
  \[
  \text{ASE} = [4\gamma^2 \left\{ \hat{A}^m + \hat{B}^m - 2\hat{C}^m + 2 \sum_{j=1}^{m-1} (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j}) \right\}]^{1/2},
  \]

  \( \gamma = [\log(\epsilon+\Delta \epsilon) - \log(\epsilon)]^{-1} \), \( \hat{A} = K(\epsilon+\Delta \epsilon)/C(\epsilon+\Delta \epsilon) \), \( \hat{B} = K(\epsilon)/C(\epsilon) \), \( C = W(\epsilon+\Delta \epsilon, \epsilon)/C(\epsilon+\Delta \epsilon)C(\epsilon) \).

  Standard normal random numbers of size of 500 were used to calculate \( K(\epsilon+\Delta \epsilon) \), \( K(\epsilon) \), \( C(\epsilon+\Delta \epsilon) \), \( C(\epsilon) \) and \( W(\epsilon+\Delta \epsilon, \epsilon) \) at which \( \hat{A} \), \( \hat{B} \) and \( \hat{C} \) are evaluated. The standard error of the mean ASE is reported in parenthesis.

- True ASE is obtained from the same ASE formula which is given above. But \( \hat{A}, \hat{B} \) and \( \hat{C} \) were evaluated at numerically calculated values of \( K(\epsilon+\Delta \epsilon) = 0.2511 \), \( K(\epsilon) = 0.2098 \), \( C(\epsilon+\Delta \epsilon) = 0.4755 \), \( C(\epsilon) = 0.4332 \), and \( W(\epsilon+\Delta \epsilon, \epsilon) = 0.2295 \) for \( \epsilon+\Delta \epsilon = 0.9 \) and \( \epsilon = 0.9^2 \).

- \( \sqrt{T} \) SD = \( \sqrt{T} \times \text{standard error of } [D(\cdot) - m] \) of the 2500 replications.
Table 8: Monte Carlo Simulation

KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS

Number of Replications=2500; Sample Size=500; $\epsilon=0.9$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0006 (0.0106)</td>
<td>0.968 (0.041)</td>
<td>0.966</td>
<td>0.237</td>
</tr>
<tr>
<td>2</td>
<td>0.0013 (0.0157)</td>
<td>0.995 (0.044)</td>
<td>0.994</td>
<td>0.352</td>
</tr>
<tr>
<td>3</td>
<td>0.0019 (0.0206)</td>
<td>1.026 (0.034)</td>
<td>1.024</td>
<td>0.462</td>
</tr>
<tr>
<td>4</td>
<td>0.0024 (0.0255)</td>
<td>1.058 (0.052)</td>
<td>1.056</td>
<td>0.571</td>
</tr>
<tr>
<td>5</td>
<td>0.0031 (0.0312)</td>
<td>1.093 (0.057)</td>
<td>1.091</td>
<td>0.698</td>
</tr>
<tr>
<td>6</td>
<td>0.0030 (0.0395)</td>
<td>1.131 (0.062)</td>
<td>1.128</td>
<td>0.884</td>
</tr>
<tr>
<td>7</td>
<td>0.0043 (0.0518)</td>
<td>1.172 (0.068)</td>
<td>1.168</td>
<td>1.158</td>
</tr>
<tr>
<td>8</td>
<td>0.0081 (0.0698)</td>
<td>1.215 (0.075)</td>
<td>1.211</td>
<td>1.502</td>
</tr>
<tr>
<td>9</td>
<td>0.0139 (0.0980)</td>
<td>1.262 (0.082)</td>
<td>1.257</td>
<td>2.191</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 500 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate $[K_m(\cdot)+\log C(\epsilon,1,T)]$ of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's,
  $$\text{ASE} = [4\{(K(\epsilon)/C(\epsilon))^2 \cdot m+1 - (K(\epsilon)/C(\epsilon))^m + K(\epsilon)/C(\epsilon) \cdot 2 \cdot 1\}]^{1/2}.$$  
  Standard normal random numbers of size of 500 were used to calculate $K(\epsilon), C(\epsilon)$ at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of $K(\epsilon)=0.2511, C(\epsilon)=0.4755$ for $\epsilon=0.9$.
- $\sqrt{T}$ SD = $\sqrt{T} \times \text{standard error of } [K_m(\cdot)+\log C(\epsilon,1,T)]$ of the 2500 replications.
Table 9: Monte Carlo Simulation

CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS
Number of Replications=2500; Sample Size=250; \( \epsilon+\Delta \epsilon=0.9, \epsilon=0.9^2 \)

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Dim Est</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>( \sqrt{T} ) SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.886 (0.027)</td>
<td>2.202 (0.052)</td>
<td>0.526</td>
<td>0.428</td>
</tr>
<tr>
<td>2</td>
<td>1.774 (0.068)</td>
<td>4.527 (0.198)</td>
<td>1.087</td>
<td>1.070</td>
</tr>
<tr>
<td>3</td>
<td>2.663 (0.133)</td>
<td>7.022 (0.316)</td>
<td>1.696</td>
<td>2.106</td>
</tr>
<tr>
<td>4</td>
<td>3.558 (0.242)</td>
<td>9.698 (0.453)</td>
<td>2.359</td>
<td>3.827</td>
</tr>
<tr>
<td>5</td>
<td>4.454 (0.420)</td>
<td>12.568 (0.613)</td>
<td>3.077</td>
<td>6.642</td>
</tr>
<tr>
<td>6</td>
<td>5.360 (0.702)</td>
<td>15.646 (0.780)</td>
<td>3.857</td>
<td>11.096</td>
</tr>
<tr>
<td>7</td>
<td>6.230 (1.153)</td>
<td>18.946 (1.020)</td>
<td>4.702</td>
<td>18.230</td>
</tr>
<tr>
<td>8</td>
<td>7.278 (1.888)</td>
<td>22.483 (1.277)</td>
<td>5.617</td>
<td>29.536</td>
</tr>
<tr>
<td>9</td>
<td>8.470 (3.184)</td>
<td>26.273 (1.576)</td>
<td>6.609</td>
<td>50.345</td>
</tr>
<tr>
<td>10</td>
<td>9.511 (5.034)</td>
<td>30.334 (1.923)</td>
<td>7.682</td>
<td>79.599</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 250 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE’s,
  \[
  \text{ASE} = \left[4 \gamma^2 \left\{ \hat{A}^m + \hat{B}^m - 2\hat{C}^m + 2 \sum_{j=1}^{m-1} (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j}) \right\} \right]^{1/2},
  \]
  where \( \gamma = [\log(\epsilon+\Delta \epsilon) - \log(\epsilon)]^{-1}, \hat{A} = K(\epsilon+\Delta \epsilon)/C(\epsilon+\Delta \epsilon)^2, \hat{B} = K(\epsilon)/C(\epsilon)^2, \)
  \( C = W(\epsilon+\Delta \epsilon, \epsilon)/C(\epsilon+\Delta \epsilon)C(\epsilon) \). Standard normal random numbers of size of 250 were used to calculate \( K(\epsilon+\Delta \epsilon), K(\epsilon), C(\epsilon+\Delta \epsilon), C(\epsilon) \) and \( W(\epsilon+\Delta \epsilon, \epsilon) \) at which \( \hat{A}, \hat{B} \) and \( \hat{C} \) are evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But \( \hat{A}, \hat{B} \) and \( \hat{C} \) were evaluated at numerically calculated values of \( K(\epsilon+\Delta \epsilon)=0.2511, K(\epsilon)=0.2098, C(\epsilon+\Delta \epsilon)=0.4755, C(\epsilon)=0.4332, \)
  and \( W(\epsilon+\Delta \epsilon, \epsilon)=0.2295 \) for \( \epsilon+\Delta \epsilon=0.9 \) and \( \epsilon=0.9^2 \).
- \( \sqrt{T} \) SD = \( \sqrt{T} \times \) standard error of [D(.) - m] of the 2500 replications.
Table 10: Monte Carlo Simulation

KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS

Number of Replications=2500; Sample Size=250; \( \varepsilon = 0.9 \)

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>( \sqrt{T} ) SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0013 (0.0160)</td>
<td>0.967 (0.057)</td>
<td>0.966</td>
<td>0.253</td>
</tr>
<tr>
<td>2</td>
<td>0.0026 (0.0245)</td>
<td>0.995 (0.062)</td>
<td>0.994</td>
<td>0.387</td>
</tr>
<tr>
<td>3</td>
<td>0.0040 (0.0317)</td>
<td>1.025 (0.067)</td>
<td>1.024</td>
<td>0.502</td>
</tr>
<tr>
<td>4</td>
<td>0.0053 (0.0405)</td>
<td>1.058 (0.073)</td>
<td>1.056</td>
<td>0.640</td>
</tr>
<tr>
<td>5</td>
<td>0.0072 (0.0524)</td>
<td>1.093 (0.080)</td>
<td>1.091</td>
<td>0.829</td>
</tr>
<tr>
<td>6</td>
<td>0.0114 (0.0711)</td>
<td>1.131 (0.088)</td>
<td>1.128</td>
<td>1.125</td>
</tr>
<tr>
<td>7</td>
<td>0.0182 (0.0982)</td>
<td>1.172 (0.096)</td>
<td>1.168</td>
<td>1.552</td>
</tr>
<tr>
<td>8</td>
<td>0.0326 (0.1480)</td>
<td>1.216 (0.105)</td>
<td>1.211</td>
<td>2.341</td>
</tr>
<tr>
<td>9</td>
<td>0.0687 (0.2403)</td>
<td>1.262 (0.116)</td>
<td>1.257</td>
<td>3.780</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 250 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate \([K_m(\cdot)+\log C(\varepsilon,1,T)]\) of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's,
  \[\text{ASE} = \left\{\frac{K(\varepsilon)/C(\varepsilon)^2}{m+1} + \frac{K(\varepsilon)/C(\varepsilon)^2}{m+1} \cdot \frac{K(\varepsilon)/C(\varepsilon)^2}{m+1} \right\}^{1/2} \]
  Standard normal random numbers of size of 250 were used to calculate \(K(\varepsilon), C(\varepsilon)\) at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of \(K(\varepsilon)=0.2511, C(\varepsilon)=0.4755\) for \(\varepsilon = 0.9\).
- \(\sqrt{T} \) SD = \(\sqrt{T} \times \) standard error of \([K_m(\cdot) + \log C(\varepsilon,1,T)]\) of the 2500 replications.
Table 11: Monte Carlo Simulation

CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS
Number of Replications=2500; Sample Size=1000; \(\epsilon+\Delta\epsilon=0.9^2\), \(\epsilon=0.9^3\)

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Dim Est</th>
<th>Mean ASE</th>
<th>True ASE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.043)</td>
<td>(0.348)</td>
</tr>
<tr>
<td>1</td>
<td>0.960</td>
<td>2.423</td>
<td>0.348</td>
</tr>
<tr>
<td>2</td>
<td>1.812 (0.019)</td>
<td>4.987 (0.095)</td>
<td>0.721 (0.607)</td>
</tr>
<tr>
<td>3</td>
<td>2.718 (0.038)</td>
<td>7.746 (0.162)</td>
<td>1.130 (1.212)</td>
</tr>
<tr>
<td>4</td>
<td>3.623 (0.071)</td>
<td>10.715 (0.246)</td>
<td>1.567 (2.252)</td>
</tr>
<tr>
<td>5</td>
<td>4.529 (0.124)</td>
<td>13.909 (0.350)</td>
<td>2.064 (3.922)</td>
</tr>
<tr>
<td>6</td>
<td>5.437 (0.214)</td>
<td>17.343 (0.475)</td>
<td>2.597 (6.711)</td>
</tr>
<tr>
<td>7</td>
<td>6.351 (0.369)</td>
<td>21.035 (0.626)</td>
<td>3.179 (11.671)</td>
</tr>
<tr>
<td>8</td>
<td>7.280 (0.621)</td>
<td>25.003 (0.804)</td>
<td>3.812 (19.637)</td>
</tr>
<tr>
<td>9</td>
<td>8.220 (1.036)</td>
<td>29.266 (1.014)</td>
<td>4.503 (32.766)</td>
</tr>
<tr>
<td>10</td>
<td>9.221 (1.740)</td>
<td>33.846 (1.259)</td>
<td>5.255 (55.035)</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 1000 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE's,

\[
\text{ASE} = \left[4\gamma^2 \left\{ \hat{A}^m + \hat{B}^m - 2\hat{C}^m + 2E (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j}) \right\} \right]^{1/2},
\]

where

\[
\gamma = \frac{\log(\epsilon+\Delta\epsilon) - \log(\epsilon)}{\epsilon+\Delta\epsilon}, \quad \hat{A} = \frac{K(\epsilon+\Delta\epsilon)}{C(\epsilon+\Delta\epsilon)} K(\epsilon), \quad \hat{B} = \frac{K(\epsilon)}{C(\epsilon)} K(\epsilon), \quad \hat{C} = \frac{W(\epsilon+\Delta\epsilon, \epsilon)}{C(\epsilon+\Delta\epsilon) C(\epsilon)},
\]

and \(W(\epsilon+\Delta\epsilon, \epsilon)\) at which \(\hat{A}, \hat{B}\) and \(\hat{C}\) are evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But \(\hat{A}, \hat{B}\) and \(\hat{C}\) were evaluated at numerically calculated values of \(K(\epsilon+\Delta\epsilon)=0.2098, K(\epsilon)=0.1743, C(\epsilon+\Delta\epsilon)=0.4332, C(\epsilon)=0.3938,\) and \(W(\epsilon+\Delta\epsilon, \epsilon)=0.1902\) for \(\epsilon+\Delta\epsilon=0.9^2\) and \(\epsilon=0.9^3\).
- \(\sqrt{T} \text{ SD} = \sqrt{T} \times \text{ standard error of } [D(\cdot) - m] \text{ of the 2500 replications.}\)
Table 12: Monte Carlo Simulation

**KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS**

Number of Replications=2500; Sample Size=1000; $\varepsilon=0.9^2$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0006 (0.0078)</td>
<td>0.999 (0.033)</td>
<td>1.000</td>
<td>0.245</td>
</tr>
<tr>
<td>2</td>
<td>0.0011 (0.0113)</td>
<td>1.030 (0.035)</td>
<td>1.030</td>
<td>0.358</td>
</tr>
<tr>
<td>3</td>
<td>0.0014 (0.0147)</td>
<td>1.063 (0.039)</td>
<td>1.064</td>
<td>0.466</td>
</tr>
<tr>
<td>4</td>
<td>0.0012 (0.0180)</td>
<td>1.100 (0.043)</td>
<td>1.100</td>
<td>0.569</td>
</tr>
<tr>
<td>5</td>
<td>0.0020 (0.0225)</td>
<td>1.139 (0.047)</td>
<td>1.138</td>
<td>0.711</td>
</tr>
<tr>
<td>6</td>
<td>0.0021 (0.0288)</td>
<td>1.181 (0.052)</td>
<td>1.180</td>
<td>0.909</td>
</tr>
<tr>
<td>7</td>
<td>0.0025 (0.0388)</td>
<td>1.226 (0.057)</td>
<td>1.225</td>
<td>1.226</td>
</tr>
<tr>
<td>8</td>
<td>0.0042 (0.0537)</td>
<td>1.257 (0.063)</td>
<td>1.274</td>
<td>1.699</td>
</tr>
<tr>
<td>9</td>
<td>0.0100 (0.0792)</td>
<td>1.327 (0.069)</td>
<td>1.326</td>
<td>2.504</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 1000 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate $[K_m(\cdot)+\log C(\varepsilon,1,T)]$ of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's,
  
  $$ASE = \left[4\left\{\left(\frac{K(\varepsilon)}{C(\varepsilon)}\right)^2 - 1\right\}\right]^{1/2}.$$ 

  Standard normal random numbers of size 1000 were used to calculate $K(\varepsilon)$, $C(\varepsilon)$ at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of $K(\varepsilon)=0.2098$, $C(\varepsilon)=0.4332$ for $\varepsilon=0.9^2$.
- $\sqrt{T}$ SD = $\sqrt{T} \times$ standard error of $[K_m(\cdot) + \log C(\varepsilon,1,T)]$ of the 2500 replications.
Table 13: Monte Carlo Simulation

CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS

Number of Replications=2500; Sample Size=500; $\epsilon + \Delta \epsilon = 0.9^2$, $\epsilon = 0.9^3$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Dim Est</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T} SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.906 (0.015)</td>
<td>2.417 (0.067)</td>
<td>0.348</td>
<td>0.330</td>
</tr>
<tr>
<td>2</td>
<td>1.813 (0.037)</td>
<td>4.976 (0.147)</td>
<td>0.721</td>
<td>0.830</td>
</tr>
<tr>
<td>3</td>
<td>2.718 (0.076)</td>
<td>7.730 (0.246)</td>
<td>1.130</td>
<td>1.691</td>
</tr>
<tr>
<td>4</td>
<td>3.622 (0.140)</td>
<td>10.695 (0.368)</td>
<td>1.567</td>
<td>3.122</td>
</tr>
<tr>
<td>5</td>
<td>4.525 (0.245)</td>
<td>13.885 (0.516)</td>
<td>2.064</td>
<td>5.482</td>
</tr>
<tr>
<td>6</td>
<td>5.428 (0.420)</td>
<td>17.317 (0.694)</td>
<td>2.597</td>
<td>9.393</td>
</tr>
<tr>
<td>7</td>
<td>6.341 (0.715)</td>
<td>21.007 (0.907)</td>
<td>3.179</td>
<td>15.998</td>
</tr>
<tr>
<td>8</td>
<td>7.284 (1.229)</td>
<td>24.976 (1.158)</td>
<td>3.812</td>
<td>27.488</td>
</tr>
<tr>
<td>9</td>
<td>8.332 (2.157)</td>
<td>29.241 (1.453)</td>
<td>4.503</td>
<td>48.229</td>
</tr>
<tr>
<td>10</td>
<td>9.975 (3.997)</td>
<td>33.825 (1.796)</td>
<td>5.255</td>
<td>89.365</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 500 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE's,
  \[
  ASE = \left[4\gamma^2\left(\hat{A}^m + \hat{B}^m - 2\hat{C}^m + 2 \sum_{j=1}^{m-1} (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j})\right)\right]^{1/2},
  \]
  where \( \gamma = \frac{\log(\epsilon + \Delta \epsilon) - \log(\epsilon)}{1} \), \( \hat{A} = K(\epsilon + \Delta \epsilon)/C(\epsilon + \Delta \epsilon)^2 \), \( \hat{B} = K(\epsilon)/C(\epsilon)^2 \), \( C = W(\epsilon + \Delta \epsilon, \epsilon)/(C(\epsilon + \Delta \epsilon)C(\epsilon)) \). Standard normal random numbers of size of 500 were used to calculate \( K(\epsilon + \Delta \epsilon) \), \( K(\epsilon) \), \( C(\epsilon + \Delta \epsilon) \), \( C(\epsilon) \) and \( W(\epsilon + \Delta \epsilon, \epsilon) \) at which \( \hat{A}, \hat{B} \) and \( \hat{C} \) are evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But \( \hat{A}, \hat{B} \) and \( \hat{C} \) were evaluated at numerically calculated values of \( K(\epsilon + \Delta \epsilon) = 0.2098 \), \( K(\epsilon) = 0.1743 \), \( C(\epsilon + \Delta \epsilon) = 0.4332 \), \( C(\epsilon) = 0.3938 \), and \( W(\epsilon + \Delta \epsilon, \epsilon) = 0.1902 \) for \( \epsilon + \Delta \epsilon = 0.9^2 \) and \( \epsilon = 0.9^3 \).
- \( \sqrt{T} SD = \sqrt{T} \times \text{standard error of } [D(\cdot) - m] \text{ of the 2500 replications.} \)
Table 14: Monte Carlo Simulation

**KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS**

Number of Replications=2500; Sample Size=500; $\epsilon=0.9^2$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0006 (0.0114)</td>
<td>1.001 (0.046)</td>
<td>1.000</td>
<td>0.256</td>
</tr>
<tr>
<td>2</td>
<td>0.0015 (0.0171)</td>
<td>1.032 (0.050)</td>
<td>1.030</td>
<td>0.383</td>
</tr>
<tr>
<td>3</td>
<td>0.0022 (0.0229)</td>
<td>1.065 (0.055)</td>
<td>1.064</td>
<td>0.511</td>
</tr>
<tr>
<td>4</td>
<td>0.0027 (0.0291)</td>
<td>1.101 (0.060)</td>
<td>1.100</td>
<td>0.651</td>
</tr>
<tr>
<td>5</td>
<td>0.0043 (0.0377)</td>
<td>1.141 (0.066)</td>
<td>1.138</td>
<td>0.843</td>
</tr>
<tr>
<td>6</td>
<td>0.0043 (0.0503)</td>
<td>1.183 (0.072)</td>
<td>1.180</td>
<td>1.125</td>
</tr>
<tr>
<td>7</td>
<td>0.0065 (0.0718)</td>
<td>1.229 (0.080)</td>
<td>1.225</td>
<td>1.606</td>
</tr>
<tr>
<td>8</td>
<td>0.0129 (0.1042)</td>
<td>1.278 (0.088)</td>
<td>1.274</td>
<td>2.329</td>
</tr>
<tr>
<td>9</td>
<td>0.0284 (0.1637)</td>
<td>1.331 (0.097)</td>
<td>1.326</td>
<td>3.660</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 500 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate $[K_m(\cdot)+\log C(\epsilon,1,T)]$ of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's, $\text{ASE}=[4\{(K(\epsilon)/C(\epsilon)^2)^m+1-(K(\epsilon)/C(\epsilon)^2)^m+K(\epsilon)/C(\epsilon)^2-1\}]^{1/2}$. Standard normal random numbers of size of 500 were used to calculate $K(\epsilon)$, $C(\epsilon)$ at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of $K(\epsilon)=0.2098$, $C(\epsilon)=0.4332$ for $\epsilon=0.9^2$.
- $\sqrt{T}$ SD = $\sqrt{T}$ x standard error of $[K_m(\cdot)+\log C(\epsilon,1,T)]$ of the 2500 replications.
### Table 15: Monte Carlo Simulation

**CORRELATION DIMENSION ESTIMATES AND STANDARD ERRORS**

Number of Replications=2500; Sample Size=250; $\epsilon+\Delta \epsilon=0.9^2$, $\epsilon=0.9^3$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Mean True</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dim Est</td>
<td>ASE</td>
</tr>
<tr>
<td>1</td>
<td>0.908 (0.028)</td>
<td>2.405 (0.115)</td>
</tr>
<tr>
<td>2</td>
<td>1.816 (0.073)</td>
<td>4.950 (0.246)</td>
</tr>
<tr>
<td>3</td>
<td>2.724 (0.148)</td>
<td>7.692 (0.403)</td>
</tr>
<tr>
<td>4</td>
<td>3.625 (0.278)</td>
<td>10.643 (0.590)</td>
</tr>
<tr>
<td>5</td>
<td>4.524 (0.496)</td>
<td>13.820 (0.814)</td>
</tr>
<tr>
<td>6</td>
<td>5.452 (0.867)</td>
<td>17.239 (1.078)</td>
</tr>
<tr>
<td>7</td>
<td>6.453 (1.561)</td>
<td>20.917 (1.390)</td>
</tr>
</tbody>
</table>

**Notes:**
- IMSL subroutine DRNNOA was called to generate 250 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Dimension Estimates = mean of the correlation dimension estimates of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis for given dimension.
- Mean ASE = mean of the 2500 empirical ASE's,
  \[
  ASE = \left(4 \gamma^2 \left\{ \hat{A}^m + \hat{B}^m - 2\hat{C}^m + 2 \sum_{j=1}^{m-1} (\hat{A}^{m-j} + \hat{B}^{m-j} - 2\hat{C}^{m-j}) \right\} \right)^{1/2},
  \]
  where
  \[
  \gamma = \left[ \log(\epsilon+\Delta \epsilon) - \log(\epsilon) \right]^{-1}, \quad \hat{A} = K(\epsilon+\Delta \epsilon)/C(\epsilon+\Delta \epsilon)^2, \quad \hat{B} = K(\epsilon)/C(\epsilon)^2, \quad C = W(\epsilon+\Delta \epsilon, \epsilon)/(C(\epsilon+\Delta \epsilon)C(\epsilon)).
  \]
  Standard normal random numbers of size of 250 were used to calculate $K(\epsilon+\Delta \epsilon)$, $K(\epsilon)$, $C(\epsilon+\Delta \epsilon)$, $C(\epsilon)$ and $W(\epsilon+\Delta \epsilon, \epsilon)$ at which $\hat{A}$, $\hat{B}$ and $\hat{C}$ are evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But $\hat{A}$, $\hat{B}$ and $\hat{C}$ were evaluated at numerically calculated values of $K(\epsilon+\Delta \epsilon)=0.2098$, $K(\epsilon)=0.1743$, $C(\epsilon+\Delta \epsilon)=0.4332$, $C(\epsilon)=0.3938$, and $W(\epsilon+\Delta \epsilon, \epsilon)=0.1902$ for $\epsilon+\Delta \epsilon=0.9^2$ and $\epsilon=0.9^3$.
- $\sqrt{T}$ SD = $\sqrt{T} \times$ standard error of $[D(.) - m]$ of the 2500 replications.
Table 16: Monte Carlo Simulation

KOLMOGOROV ENTROPY ESTIMATES AND STANDARD ERRORS

Number of Replications=2500; Sample Size=250; $\varepsilon=0.9^2$

<table>
<thead>
<tr>
<th>Embedding Dimension</th>
<th>Average of Entropy</th>
<th>Mean ASE</th>
<th>True ASE</th>
<th>$\sqrt{T}$ SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0014 (0.0173)</td>
<td>1.000 (0.064)</td>
<td>1.000</td>
<td>0.274</td>
</tr>
<tr>
<td>2</td>
<td>0.0029 (0.0269)</td>
<td>1.031 (0.070)</td>
<td>1.030</td>
<td>0.426</td>
</tr>
<tr>
<td>3</td>
<td>0.0047 (0.0365)</td>
<td>1.065 (0.077)</td>
<td>1.064</td>
<td>0.577</td>
</tr>
<tr>
<td>4</td>
<td>0.0066 (0.0493)</td>
<td>1.101 (0.084)</td>
<td>1.100</td>
<td>0.780</td>
</tr>
<tr>
<td>5</td>
<td>0.0093 (0.0673)</td>
<td>1.141 (0.092)</td>
<td>1.138</td>
<td>1.064</td>
</tr>
<tr>
<td>6</td>
<td>0.0160 (0.0963)</td>
<td>1.184 (0.102)</td>
<td>1.180</td>
<td>1.523</td>
</tr>
<tr>
<td>7</td>
<td>0.0289 (0.1460)</td>
<td>1.230 (0.112)</td>
<td>1.225</td>
<td>2.308</td>
</tr>
</tbody>
</table>

Notes:
- IMSL subroutine DRNNOA was called to generate 250 standard normal random numbers, and RNSET was called to set an initial seed.
- Average of Entropy = mean of the Kolmogorov Entropy estimate $[K_m(\cdot)+\log C(\varepsilon,1,T)]$ of the 2500 replications. The standard error of sample mean out of the 2500 replications is reported in parenthesis.
- Mean ASE = mean of the 2500 empirical ASE's,
  $$\text{ASE} = \sqrt{\frac{4}{n-1}} \left[ \left( \frac{K(\varepsilon)}{C(\varepsilon)} \right)^2 \frac{m+1}{m+2} \right]$$
  Standard normal random numbers of size of 250 were used to calculate $K(\varepsilon)$, $C(\varepsilon)$ at which ASE is evaluated. The standard error of the mean ASE is reported in parenthesis.
- True ASE is obtained from the same ASE formula which is given above. But it was evaluated at numerically calculated values of $K(\varepsilon)=0.2098$, $C(\varepsilon)=0.4332$ for $\varepsilon=0.9^2$.
- $\sqrt{T}$ SD = $\sqrt{T} \times$ standard error of $[K_m(\cdot)+\log C(\varepsilon,1,T)]$ of the 2500 replications.
Figure 1

HISTOGRAM OF CORR. DIMENSION

T=1000, ε=0.9, ε+4ε=0.9^2, m=2

HISTOGRAM OF KOL. ENTROPY

T=1000, ε=0.9, m=2
Figure 2

HISTOGRAM OF CORR. DIMENSION

$\#$ Frequency
0 40 80 120 160 200 240 280 320 360
-2.2 -1.8 -1.4 -1.0

$T=1000, \epsilon=0.9, \epsilon+\Delta \epsilon=0.9^2, m=4$

HISTOGRAM OF KOL. ENTROPY

$\#$ Frequency
0 40 80 120 160 200 240 280 320 360 400
-2.2 -1.8 -1.4 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0

$T=1000, \epsilon=0.9, m=4$
Figure 3

HISTOGRAM OF CORR. DIMENSION

Midpoint

T=1000, \( \epsilon = 0.9 \), \( \epsilon + \Delta \epsilon = 0.9^2 \), m=8

HISTOGRAM OF KOL. ENTROPY

Midpoint

T=1000, \( \epsilon = 0.9 \), m=8
Figure 4

HISTOGRAM OF CORR. DIMENSION

Midpoint

T=500, $\varepsilon=0.9$, $\varepsilon+4\varepsilon=0.9^2$, $m=2$

HISTOGRAM OF KOL. ENTROPY

Midpoint

T=500, $\varepsilon=0.9$, $m=2$
Figure 5

HISTOGRAM OF CORR. DIMENSION

T=500, $\epsilon=0.9$, $\epsilon+4\epsilon=0.9^2$, $m=4$

HISTOGRAM OF KOL. ENTROPY

T=500, $\epsilon=0.9$, $m=4$
Figure 6

HISTOGRAM OF CORR. DIMENSION

T=500, ε=0.9, ε+Δε=0.9^2, m=8

HISTOGRAM OF KOL. ENTROPY

T=500, ε=0.9, m=8
Figure 7

HISTOGRAM OF CORR. DIMENSION

$T=250, \epsilon=0.9, \epsilon+\Delta\epsilon=0.9^2, m=2$

HISTOGRAM OF KOL. ENTROPY

$T=250, \epsilon=0.9, m=2$
Figure 8

HISTOGRAM OF CORR. DIMENSION

T=250, $\epsilon=0.9$, $\epsilon+4\epsilon=0.9^2$, $m=4$

HISTOGRAM OF KOL. ENTROPY

T=250, $\epsilon=0.9$, $m=4$
Figure 9

HISTOGRAM OF CORR. DIMENSION

Midpoint

T=250, \( \epsilon = 0.9, \epsilon + \Delta \epsilon = 0.9^2, m = 8 \)

HISTOGRAM OF KOL. ENTROPY

Midpoint

T=250, \( \epsilon = 0.9, m = 8 \)
Figure 10
HISTOGRAM OF THE DIFFERENCE OF TWO DIMENSION ESTIMATES

Notes:
• The histogram was generated under the null hypothesis that \( \{a_{1t}\} \)
  and \( \{a_{2t}\} \) are IID with common distribution function.
• Number of Replications=2500; Size of each sample=500; \( m=6; \)
  \( \epsilon + \Delta \epsilon = 0.9; \ \epsilon = 0.9^2 \)
• IMSL subroutine DRNNOA was called to generate the first and the
  second half of each sample separately.
• The test statistic was standardized.
Figure 11

HISTOGRAM OF THE DIFFERENCE OF ORIGINAL AND BOOTSTRAP SAMPLES

Notes:
- The histogram was generated under the null hypothesis that \( \{a_t\} \) is IID
- Number of Replications=2500; Sample Size=500; \( m=6; \epsilon+\Delta\epsilon=0.9; \epsilon=0.9^2 \)
- IMSL subroutine DRNNOA was called to generate all the original samples.
- The bootstrap samples were constructed with replacement