

Research Article

An Independence Test Based on Symbolic Time Series

Wiston Adrián Risso

Instituto de Economía (IECON), Universidad de la República, Joaquín Requena 1375, 11200 Montevideo, Uruguay

Correspondence should be addressed to Wiston Adrián Risso; arisso@iecon.ccee.edu.uy

Received 11 October 2013; Revised 30 December 2013; Accepted 6 January 2014; Published 24 February 2014

Academic Editor: Flavia Pennini

Copyright © 2014 Wiston Adrián Risso. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An independence test based on symbolic time series analysis (STSA) is developed. Considering an independent symbolic time series there is a statistic asymptotically distributed as a CHI-2 with $n - 1$ degrees of freedom. Size and power experiments for small samples were conducted applying Monte Carlo simulations and comparing the results with BDS and runs test. The introduced test shows a good performance detecting independence in nonlinear and chaotic systems.

1. Introduction

Economics presents a variety of dynamical processes. Linearity, nonlinearity, deterministic chaos, and stochastic models have been applied when modeling a complex reality. Even more as highlighted by Gabr and Fathey [1] the increasing use of time series data has initiated a great deal of research and developing attempts in the field of data mining. Classification of time series data has a wide range of applications and has attracted researches from a wide range of disciplines. Detecting dependence in time series is an essential task for econometricians and applied economists. In particular, as highlighted by Brooks [2], testing for nonlinearity dependence has become a relevant area of research due to its implications for model adequacy and predictability. Moreover, the importance of testing randomness also at a microeconomic level is already asserted by Wald and Wolfowitz [3]. They designed a runs test and remarked that the problem of testing randomness arises frequently in quality control of manufactured products.

As mentioned by Shahwan and Said [4] obtaining accurate stock prices forecast is one of the main goals of finance and academic research institutions. Therefore, finding evidence of nonlinearity in time series means that forecasting can be improved by switching from a linear to a nonlinear model. Furthermore, detecting dependence in the residual of a linear model is considered as evidence of nonaccuracy representation of the data. As Granger et al. [5] assert, even though economics is rich in dynamical processes, most commonly used test statistics are functions of

correlation motivated by linear relations involving continuous variables and/or Gaussian process. They remark that numerous diagnostics are used to examine model residuals for departure from “independence,” i.i.d., reversibility, martingale difference, and other properties.

There is extensive literature about testing independence and nonlinearity. Correlation tests are widely applied (see King [6] for a survey), but as highlighted by [7] correlation tests are not consistent against alternatives with zero autocorrelation. ARCH model, bilinear, nonlinear moving average (NLMA) process, and iterative logistic maps are examples of serially dependent processes with zero correlation. The Durbin-Watson (DW) test proposed by Durbin and Watson [8, 9] is by far the most common test for first-order autocorrelated AR(1) errors and extended by [10, 11]. However, DW is a test specialized in linear dependence, having low sensitivity to nonlinear dependent processes (see Azzalini and Bowman [12]).

Reference [13] makes a bibliography of a series of non-parametric tests based on instruments such as runs, signs, ranks, permutations, frequency counts, records, and quotas. The well-known runs test was proposed by Wald and Wolfowitz [3] and is based on the repeated occurrence of the same value or category of a variable such as the sign. Runs test of randomness assumes that the mean and variance are constant and the probability is independent.

Reference [14] conducts a competition among the best of the available tests for nonlinearity and chaos. The one proposed by Hinich [15] has zero power against some forms of nonlinearity. The Lyapunov exponent test suggested by

[16] is a test of chaos and it does not detect other types of nonlinearity. White test (see [17]) is a test of nonlinearity under the hypothesis of linearity in the mean. For instance, it correctly accepts linearity in the mean of the ARCH and GARCH processes, even if they are nonlinear processes. The null hypothesis of Kaplan's test [18] is linearity of the process.

Reference [19] proposes the BDS test to detect chaotic processes. However, it also serves as an independence test or nonlinear test showing a high power against a vast class of nonlinear alternatives.

Entropy has also been applied to test independence. References [20–22] suggest measures of serial dependence based on entropy. A normalized smoothed nonparametric entropy measure of serial dependence is proposed by Granger and Liu [23]. Granger et al. [5] introduce transformed metric entropy of dependence. A Kernell-based nonparametric entropy estimator of serial dependence is suggested by Hong and White [24]. Reference [7] constructs a test of independence by using symbolic dynamics and permutation entropy commented and criticized by Elsinger [25].

Reference [26] presents a method based on observable ordinal patterns to discriminate white noise from deterministic time series corrupted with high levels of white noise. Reference [27] introduces a multiscale symbolic information-theory approach for discriminating nonlinear deterministic and stochastic dynamics from time series associated with complex systems. References [28–30] suggest a test for distinguishing regular from chaotic dynamics in deterministic dynamical systems. In [31] the authors present a theoretical justification of the 0-1 test for chaos. In [32] the authors applied the 0-1 test to detect numerically the chaos in the proposed fractional order financial system.

Recently, Cánovas et al. [33] have made a comparative study of tests to detect whether a time series comes from an IID random variable or not. They have found that BDS and permutation type tests fail in detecting dependence. According to the authors testing independence is very complicated and, in their opinion, the combination of several tests is necessary.

In the present paper, a new simple and powerful test is proposed based on symbolic time series analysis (STSA). The STSA approach considered in the present work is extensively explained in [34–37]. The introduced test shows a good performance detecting dependence in nonlinear and chaotic systems.

The paper is organized as follows. The next section presents the symbolic time series analysis approach. Section 3 proposes and derives the symbolic random test. Size and power tests are conducted in Section 4. Independence and nonlinearity in financial time series are tested in Section 5. Finally, Section 6 draws some conclusions and further line of research.

2. Symbolic Time Series Analysis

As mentioned by Finney et al. [38] the concept of symbolization has its roots in dynamical-systems theory, particularly in the study of nonlinear systems which can exhibit bifurcation

and chaos. Besides the computational efficiency, symbolic methods are also robust when noise is present. Williams [39] highlights that symbolic dynamics is a method for studying nonlinear discrete-time systems by taking a previously codified trajectory using sequences of symbols from a finite set also called alphabet. However, as Piccardi [35] remarks symbolic dynamics should be differentiated from symbolic analysis. The former denotes theoretical investigation on dynamical systems. The latter is suggested when data are characterized by low degree of precision. The idea in symbolic analysis is that by discretizing the data with the right partition we obtain a symbolic sequence. This sequence is able to detect the very dynamic of the process when data are highly affected by noise.

Data symbolization implies transforming an original series of measurements into a limited number of discrete symbols. The resulting symbolic series can be analyzed for nonrandom temporal patterns. It means that, given a time series $\{x_t\}_{t=1,2,\dots,T^*}$, we study the dependence by translating the problem into a symbolic time series $\{s_t\}_{t=1,2,\dots,T^*}$.

Let us consider a time series $\{x_t\}_{t=1,2,\dots,T^*}$ where T^* is the sample size. Symbolic time series analysis (STSA) approach suggests as first step to take a partition such that the individual occurrence of each symbol is equiprobable with all others. The result is $\{s_t\}_{t=1,2,\dots,T^*}$, a symbolized time series. For instance, imagine that $\{x_t\}_{t=1,2,\dots,T^*}$ is a time series generated by a Gaussian white noise; we can define a discretization of two regions by establishing $s_t = 0$ when x_t takes a value in the first 50% of the density function and $s_t = 1$ in the other case. The new discrete time series of events $\{s_t\}_{t=1,2,\dots,T^*}$ would be similar to a series generated by tossing a coin. Of course, different discretization could be applied; for example, six equally likely symbols could be interpreted as tossing a die.

Theory calls the set of a symbols alphabet (A) and the sequences of length w words. As a second step, symbol sequences of different lengths should be computed generating new symbolic time series for each w . The first symbolization $w = 1$ is the trivial one, the symbolization of the original time series. A second symbolization is applied for $w = 2$ and so on. Consider the last example about the symbolized Gaussian variable in two symbols ($a = 2$); the sequence of two consecutive symbols ($w = 2$) produces four possible events ($a^w = 4$), the sequence of three consecutive symbols ($w = 3$) produces eight ($a^w = 8$), and so on. Because of an equiprobable partition, the relative frequency of each possible sequence for truly random data will be equal. Following the example, for $a = 2$ and $w = 2$ we know that each event from $\{(0,0), (0,1), (1,0), (1,1)\}$ has probability $1/4$ and for $a = 2$ and $w = 3$ we have $\{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}$ with probability $1/8$ for each event. In general, for a symbolized random process the probability of each sequence of length w is a^{-w} ; let us call these total possible events n .

Once the symbolized time series is obtained, the third step implies computing all the relative frequencies for all the computed symbol sequences in the data. Note that in the present case, since the partition is defined by dividing the frequency in equally probable regions, analyzing the frequency for $w = 1$ is trivial. The interest of the present work

will be for $w > 1$; it is considering subsequences of two or more symbols.

Since in practice we do not have infinite time series, we have to work with finite sample and the differences between relative frequencies and theoretical probability can be due to finite sample. Therefore it is worthy to obtain statistics in order to test the independence null hypothesis and to analyze the finite sample problem.

3. Symbolic Independence Test

The purpose of the present section is to derive a simple statistic and its asymptotic distribution when the process under study is independent. In particular, the statistic performance is analyzed for finite sample.

Let us consider a finite time series generated by an independent or random process sized $T^* \{x_t\}_{t=1,2,\dots,T^*}$. Define a partition in the series in “ a ” equiprobable regions obtaining the symbolized time series $\{s_t\}_{t=1,2,\dots,T^*}$ where each symbol s_t takes a symbolic value from the alphabet $A = \{A_1, A_2, \dots, A_a\}$.

Since we want to derive a general statistic for different alphabet sizes a and different subsequences lengths w , we have to make two considerations. (1) n is the quantity of possible events. That is, $n = a^w$, where for the simplest case ($w = 1$) implies $n = a$ and then the quantity of events is equal to the symbol set size. (2) In practice, we have a finite sample size T^* ; there is no problem for $w = 1$, but when we compute subsequences or time windows w of consecutive symbols we lose observations. For example, when we compute the frequency for 2 consecutive symbols, we have a total sample size $T^* - 1$. In general, we can define the sample size $T = T^* + w - 1$; again for the trivial case $w = 1$, $T^* = T$.

Note that, defining S_i for $i = 1, 2, \dots, n$ as the sum of the total i events in the time series, we can derive the multidimensional variable $S = \{S_i/T\}$ being distributed as a multinomial with $(S_i/T) = (1/n)$, $\text{Var}(S_i/T) = (1/n)(n-1)/nT$, and $\text{Cov}(S_i, S_j) = -(1/n)(1/nT)$ for all $i \neq j$. In the example of tossing a coin we have two symbols for the two events S_1 and S_2 ($n = 2$) and S is a distributed binomial with mean $1/2$ and variance $(1/4T)$. However, for consecutive events of length $w = 2$ we have a variable S distributed multinomial with mean $1/4$ and variance $(3/16T)$.

As we will see, frequencies of the events should be important in the statistic and the vector of the n frequencies S_i/T could be approximated by a multivariate normal distribution $N(1/n, \sigma^2 \Sigma)$ where σ^2 is $(1/nT)$ and Σ is an idempotent matrix as

$$\Sigma_{n \times n} \equiv \begin{bmatrix} \frac{(n-1)}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{(n-1)}{n} & \dots & -\frac{1}{n} \\ \dots & \dots & \dots & \dots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & \frac{(n-1)}{n} \end{bmatrix}. \quad (1)$$

For convenience we can define the vector variable $\{\varepsilon_i\} = \{(S_i/T) - (1/n)\}_{i=1,2,\dots,n}$ having a multivariate normal distribution $N(\phi, \sigma^2 \Sigma)$, ϕ being the null vector. Then the statistic can be defined as

$$\left\{ \frac{\sum_{i=1}^{i=n} \varepsilon_i^2}{\sigma^2} \right\}. \quad (2)$$

The term in brackets in (2) is a quadratic form in random normal variables. As Mathai and Provost [40] assert, the distribution of quadratic forms in normal variables has been extensively studied by many authors. Various representations of the distribution function have been derived and several different procedures have been given for computing the distribution and preparing appropriate tables. Approximated distributions have been proposed by [41–45]. In the present paper the following theorem in [40, page 197] is applied.

Consider the vector X distributed as a multivariate normal distribution with mean vector ϕ and possibly singular covariance matrix Σ . A quadratic form $X'AX$ is distributed as a CHI-2 with r degrees of freedom if the following necessary and sufficient conditions are satisfied:

- (i) $(A\Sigma)^2 = (A\Sigma)^3$ and $\text{tr}(A\Sigma) = r$
- (ii) $\text{tr}(A\Sigma)^2 = \text{tr}(A\Sigma) = r$ and $\rho(\Sigma A \Sigma) = r$.

Note that the theorem can be applied in the present case where vector $X = (\varepsilon_1/\sigma, \varepsilon_2/\sigma, \dots, \varepsilon_n/\sigma)$ is distributed multivariate normal $N(\phi, \Sigma)$. In this case A is the identity matrix I and Σ is symmetric, singular, and idempotent. Since $\text{tr}(A\Sigma) = n - 1$, $X'AX$ distributes CHI-2 with $(n - 1)$ degrees of freedom.

Considering that $\sigma^2 = (1/nT)$, we obtain that the distribution of the symbolic randomness statistic (SRS) as

$$\text{SRS} \equiv Tn \left\{ \sum_{i=1}^{i=n} \left(\frac{S_i}{T} - \frac{1}{n} \right)^2 \right\} \quad (3)$$

asymptotically distributes χ_{n-1}^2 .

Note that computing the statistic is very simple. We just have to know the sample size (T), the symbols (a), and subsequences or length (w) and compute the frequencies for each event ($n = a^w$) in the time series (S_i).

In summary, the test works as follows.

Step 1. Considering time series $\{x_t\}_{t=1,2,\dots,T^*}$, compute the empirical distribution and define equiprobable regions according to the quantity of symbols or the alphabet size.

Step 2. According to the partition, translate $\{x_t\}_{t=1,2,\dots,T^*}$ into $\{s_t\}_{t=1,2,\dots,T^*}$, the symbolic time series when $w = 1$.

Step 3. Compute different symbolic time series for different lengths w ; remember that the obtained series in Step 2 corresponds to $w = 1$.

Step 4. For each w , compute the frequency of the n different events S_i/T for $i = 1, 2, \dots, n$.

Step 5. For each w , compute the $SRS(a, w) = Tn\{\Sigma(S_i/T - 1/n)^2\}$ as shown in (3).

Step 6. Compare the $SRS(a, w)$ with the CHI-2 with $n - 1$ degrees of freedom at 0.05 of significance, under the independence null hypothesis. When $SRS(a, w)$ is larger than the critical value, we reject the null hypothesis.

As an example consider a time series sized $T = 50$ of tossing a coin ($n = 2$ events), the SRS is computed as $50 \cdot 2 \cdot \{[(S_1/50) - (1/2)]^2 + [(S_2/50) - (1/2)]^2\}$ where S_1 and S_2 are the times event 1 (cross) and event 2 (pile) appearing in the time series, respectively. In this case SRS should be compared with the critical value 3.84 from a Chi2 with 1 degree of freedom ($\alpha = 0.05$). Since the process is random, SRS should be less than the critical value. For consecutive symbols of length $w = 2$, we have $(50 - 1) \cdot 4 \cdot \{[(S_1/50) - (1/4)]^2 + [(S_2/50) - (1/4)]^2 + [(S_3/50) - (1/4)]^2 + [(S_4/50) - (1/4)]^2\}$ compared with the critical value 7.81 from a Chi2 with 3 degrees of freedom ($\alpha = 0.05$).

It could be shown that the SRS coincides with the Pearson independence statistic applied to the symbolic time series. Even more, this statistic is related to the Shannon entropy. In order to see the latter, let us consider H as the normalized Shannon entropy (H) where $H = 0$ for a certain process and $H = 1$ for a random process. Substituting the probabilities in H for the frequencies corresponding to the symbolic series (S_i/T), we have

$$H \equiv \frac{-1}{\log_2(n)} \sum_{i=1}^{i=n} \left(\frac{S_i}{T} \right) \log_2 \left(\frac{S_i}{T} \right). \quad (4)$$

The distribution of (S_i/T) is known, but the logarithmic function introduces some difficulty to obtain the exact distribution of H . However, we can take a linear approximation when the variable is less than 1 in absolute value. In the present case, we have defined the multinomial distributed vector $\{\varepsilon_i\} = \{(S_i/T) - (1/n)\}_{i=1,2,\dots,n}$ substituting (S_i/T) for $\varepsilon_i + (1/n)$ for all $i = 1, 2, \dots, n$ in H , we have

$$H \equiv \frac{-1}{\log_2(n)} \sum_{i=1}^{i=n} \left(\varepsilon_i + \frac{1}{n} \right) \log_2 \left(\varepsilon_i + \frac{1}{n} \right). \quad (5)$$

Note that $|\varepsilon_i| \leq 1$, its mean is 0, and the variance is $(n-1)/Tn^2$ which is decreasing with the sample size and the number of events. Given that $\log_2(\varepsilon_i + 1/n) = \log_2(1 + n\varepsilon_i) - \log_2(n)$, a linear approximation for $\varepsilon_i = 0$ of $\log_2(1 + n\varepsilon_i) \approx n\varepsilon_i$ and then $\log_2(\varepsilon_i + 1/n) = n\varepsilon_i - \log_2(n)$ where $c = 1/\ln(2)$. Equation (6) develops H when the expressions mentioned before are introduced:

$$H \approx \frac{-1}{\log_2(n)} \left\{ \sum_{i=1}^{i=n} \left(\frac{n\varepsilon_i + 1}{n} \right) (n\varepsilon_i - \log_2(n)) \right\},$$

$$H \approx \frac{-1}{n \cdot \log_2(n)} \left\{ n^2 c \sum_{i=1}^{i=n} \varepsilon_i^2 - n \cdot \log_2(n) \sum_{i=1}^{i=n} \varepsilon_i + nc \sum_{i=1}^{i=n} \varepsilon_i - n \cdot \log_2(n) \right\}. \quad (6)$$

Note that $\Sigma \varepsilon_i = 0$ since $\Sigma \varepsilon_i = \Sigma \{\Sigma(S_i/T) - \Sigma(1/n)\} = \Sigma\{(T/T) - (n/n)\} = 0$ and multiplying and dividing by σ^2 the H is expressed in

$$H \approx 1 - \frac{nc}{\log_2(n)} \sigma^2 \left\{ \frac{\sum_{i=1}^{i=n} \varepsilon_i^2}{\sigma^2} \right\}. \quad (7)$$

4. Size and Power of the Symbolic Randomness Test

Once the statistic and its asymptotic distribution were obtained, we can proceed to study the performance in finite small samples and the power detecting different forms of dependency (stochastic and deterministic).

As shown, $SRS(a, w)$ is asymptotically distributed as a CHI-2 with $n - 1 = a^w - 1$ degrees of freedom, where a is the number of symbols and w represents the subsequence length of consecutive symbols in the time series.

In practice, the SRS can be applied to residual series of a fitted model to detect independence or nonlinearity, similar to the BDS test. For instance, to test nonlinearity it can be applied to the residuals of a fitted linear model such as an ARMA (p, q) model.

Experiments were conducted in order to show the size and power of the introduced SRS test. The BDS test was applied in order to compare the results. As known, BDS test depends on two parameters: embedding dimension (m) and epsilon (ε). We applied combinations of m and ε suggested by Kanzler [46] who after conducting a series experiments obtains that $m = 2$ and $\varepsilon = 1.5$ appear to give the best approximation. Liu et al. [47] conduct experiments of size and power obtaining a combination of $m = 3$ and ε around 0.26 as the best size. In addition, critical values simulated for small sample by Kanzler [46] were applied.

The following experiment was conducted to study the SRS size. 10,000 Monte Carlo simulations were conducted for time series of a pseudorandom Gaussian i.i.d. (0,1) for different sample sizes ($T = 50, T = 200, T = 500$, and $T = 2,000$). Note that, at first, our interest is the performance in small sample as is common in economic time series. The BDS test, runs test and the present SRS test were applied considering critical values with significance levels $\alpha = 0.01, \alpha = 0.05$, and $\alpha = 0.10$. Therefore we computed the percent of rejection of the independence null hypothesis over the 10,000 Monte Carlo simulations for each test, sample size, and significance level. If the critical values are unbiased, the percentage of rejection of null hypothesis of independence when the process is truly independent should be near the significance levels. It means that for $\alpha = 0.05$ the percentage of null hypothesis rejection should be near 5% for a truly random process.

Table 1 shows the percentage of the null hypothesis rejection for a random process. Note that the columns indicate the different applied symbols (2, 3, 4, and 5) and the significance levels ($\alpha = 1\%, 5\%$, and 10%). The rows show the different sample sizes ($T = 50, 200, 500, 2000$) and lengths (2, 3, 4, and 5). Two general comments can be mentioned: (1) for a time series size less than 2,000 it is not advisable to apply more

than 4 symbols; (2) the smaller the sample size, the smaller the number of symbols and length than should be applied.

In general the test seems to be conservative rejecting the null hypothesis less times than expected. Being a conservative test, it should be contrasted with its power detecting independence in nonlinear and chaotic systems as will be shown later.

Considering the four sample sizes, selecting 2 symbols and length of 4 presents decent results in most of the cases.

Selecting 3 symbols seems to be a relative good option for size of 200 or larger and 4 symbols for a sample size of 500 or larger.

There is not a best measure of the deviation from alpha in practice. However, note that, from the set of results presented in Table 1, the best results are given for a sample of 2,000 applying 3 symbols and length of 4 (SRS(3,4)) and for 2 symbols and length of 5 (SRS(2,5)). In the case of SRS(3,4) the results are conservative since the percentages of rejections (1.14%, 4.19%, and 6.94%) are the best deviation from below 1%, 5%, and 10%. On the other hand, SRS(2,5) shows the best deviation (2.41, 5.79, and 8.64) from above. We consider that combination showing the best results from below (SRS(3,4)) is better because it reduces the risk of accepting nonrandom process when the process is truly random. This will be important when power tests detecting different processes were conducted.

To obtain comparable results, the same Monte Carlo simulations applied to the SRS test were considered when computing the BDS test and runs test. Table 2 shows the percentage of null hypothesis rejection for the runs and the BDS tests. Note that the columns indicate different significant levels ($\alpha = 1\%$, 5% , 10%) for different sample sizes ($T = 50$, 200 , 500 and 2000). The rows indicate the different tests, three versions of BDS test and the runs test. The results suggest that both tests are also conservative.

In the case of the BDS test three suggested combinations of m and ε were applied as mentioned before. In this case for a sample size of 50 the simulated critical values for small samples conducted by Kanzler [46] present the best results. For a sample size of 200, simulated critical values suggested by Kanzler [46] are the best. Parameters suggested by Liu et al. [47] also indicate acceptable results. For samples of 500 and 2,000 parameters suggested by Kanzler [46] and his simulated critical values for small samples present results near the significance levels.

Although runs test is also conservative, the rejection percentages are nearer to the expected. Note also that as the sample size increases, rejection percentages are better.

The next experiment of power will try to show how powerful these tests are detecting independence in nonlinear and chaotic systems.

At this point some generator processes suggested by the literature were considered (see, e.g., [7, 12, 14, 25, 47, 48]). The next 20 generator processes were applied.

- (1) *Normal*. Random process generated by a Normal(0,1).
- (2) *CHI-2(4)*. Random process generated by a Chi2 with 4 degrees of freedom.
- (3) *t-Student(4)*. Random process generated by a *t*-student(4).

- (4) *Truncated Normal Distribution*. Random process generated by a truncated normal distribution (0,1) at the range $[-1.75, +1.75]$.
- (5) *Beta(1/2, 1/2)*. Random process generated by a Beta distribution(1/2, 1/2).
- (6) *Uniform(0,1)*. Random process generated by a Uniform(0,1).
- (7) *AR(1)*. Consider $X_t = 0.45X_{t-1} + \varepsilon_t$.
- (8) *MA(2)*. Consider $X_t = \varepsilon_t - 0.1\varepsilon_{t-1} + 0.2\varepsilon_{t-2}$.
- (9) *Logistic*. Consider $X_t = 4X_{t-1}(1 - X_{t-1})$.
- (10) *Henon*. Consider $X_t = 1 + Y_{t-1} - 1.4X_{t-1}^2$ and $Y_t = 0.3X_{t-1}$, with initial conditions Y_1 generated randomly by $N(0, 0.01)$ and $X_1 = 1$. Time series X_t is considered in the study.
- (11) *Anosov*. Consider $X_t = \text{mod}(X_{t-1} + Y_{t-1})_1$ and $Y_t = \text{mod}(x_{t-1} + 2Y_{t-1})_1$, with initial conditions X_1 and Y_1 generated randomly. Time series X_t is considered in the study.
- (12) *Lorenz*. Consider $X_t = (1 + (1.2 * 0.8))X_{t-1} - 0.8X_{t-1}Y_{t-1}$ and $Y_t = (1 - 0.8)Y_{t-1} + 0.8X_{t-1}^2$, with initial conditions X_1 and Y_1 generated randomly. This is a discrete version of the Lorenz process and time series X_t is considered in the study.
- (13) *TAR (Threshold Autoregressive)*. Consider $X_t = 0.9X_{t-1} + \varepsilon_t$ for $|X_{t-1}| \leq 1$ and $X_t = -0.3X_{t-1} + \varepsilon_t$ for $|X_{t-1}| > 1$.
- (14) *NLSIGN (Nonlinear Sign)*. Consider $X_t = 0.45 \text{sign}(X_{t-1}) + \varepsilon_t$.
- (15) *Bilinear*. Consider $X_t = 0.7X_{t-1}\varepsilon_{t-2} + \varepsilon_t$.
- (16) *NLAR (Autoregressive Nonlinear)*. Consider $X_t = (0.7|X_{t-1}|)/(2 + |X_{t-1}|) + \varepsilon_t$.
- (17) *NLMA (Nonlinear Moving Average)*. Consider $X_t = \varepsilon_t - 0.4\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + 0.5\varepsilon_t\varepsilon_{t-2}$.
- (18) *BLMA (Bilinear Moving Average)*. Consider $X_t = 0.4X_{t-1} - 0.3X_{t-2} + 0.5X_{t-1}\varepsilon_{t-1} + 0.8\varepsilon_{t-1} + \varepsilon_t$.
- (19) *Modular*. Consider $X_t = |(\text{mod}(t)_3 - 1)(-1)^t|(-1)^t + 1 + \varepsilon_t$.
- (20) *EGARCH*. Consider $h_t = 0.25 + 0.5X_{t-1}^2 + 0.6h_{t-1}$ when $\varepsilon_t < 0$ or $h_t = 0.25 + 0.5X_{t-1}^2 + 0.6h_{t-1}$ for $\varepsilon_t > 0$ and $X_t = \varepsilon_t\sqrt{h_t}$ is an exponential GARCH.

Processes (1)–(6) indicate different distribution functions; (7)–(8) are two linear stationary processes; (9)–(12) are deterministic chaotic processes; (13)–(20) are nonlinear stochastic processes. As suggested, the tests were applied to the raw series in the cases of (1)–(6) and the chaotic processes (9)–(12). However, the tests are conducted on the residual of autoregressive processes in the cases of (7)–(8) and (13)–(20) to eliminate the linear dependence. These processes were produced using MatLab R2010a.

Tables 3 to 6 show the percentage of null hypothesis rejection for the different sample sizes. The first five rows are the results for the random processes. Since the results

TABLE 1: Size of the symbolic random test. Percentage of the null hypothesis rejection.

Length	2 symbols			3 symbols			4 symbols			5 symbols		
	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
	$T = 50$			$T = 50$			$T = 50$			$T = 50$		
2	0.01	0.41	1.03	0.28	0.47	0.97	24.68	24.68	24.75	24.68	24.68	24.68
3	0.89	1.95	3.27	98.00	98.00	98.00	100.00	100.00	100.00	100.00	100.00	100.00
4	40.10	40.39	40.68	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
5	99.80	99.80	99.80	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	$T = 200$			$T = 200$			$T = 200$			$T = 200$		
2	0.05	0.59	1.14	0.06	0.42	1.06	0.05	0.32	0.82	0.00	0.00	0.00
3	0.54	1.97	3.64	1.18	2.44	3.76	89.67	89.67	89.69	89.67	89.67	89.67
4	1.31	3.74	6.17	99.84	99.84	99.84	100.00	100.00	100.00	100.00	100.00	100.00
5	7.84	10.57	13.06	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	$T = 500$			$T = 500$			$T = 500$			$T = 500$		
2	0.06	0.54	1.27	0.09	0.47	1.03	0.05	0.31	0.80	0.01	0.01	0.03
3	0.52	1.91	3.49	0.52	1.86	3.69	1.89	3.27	4.73	1.52	1.52	1.52
4	1.28	3.55	6.16	12.23	14.02	16.02	100.00	100.00	100.00	100.00	100.00	100.00
5	2.15	5.65	8.93	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	$T = 2000$			$T = 2000$			$T = 2000$			$T = 2000$		
2	0.08	0.44	1.25	0.06	0.33	0.90	0.02	0.42	1.05	0.00	0.00	0.00
3	0.50	2.03	3.45	0.44	1.88	3.79	0.34	1.72	3.51	0.00	0.00	0.00
4	1.32	3.77	6.32	1.14	4.19	6.94	8.78	11.22	13.88	8.01	8.01	8.01
5	2.41	5.79	8.64	8.13	11.63	14.97	100.00	100.00	100.00	100.00	100.00	100.00

Note: 10,000 Monte Carlo simulations were conducted applying the normally distributed pseudorandom numbers from MatLab R2010a.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

TABLE 2: Size of the BDS and runs tests.

Test	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
	$T = 50$			$T = 200$			$T = 500$			$T = 2000$		
BDS [36]	9.30	19.83	27.83	2.57	8.82	15.55	1.27	6.16	12.07	1.22	5.56	10.55
BDS [37]	4.07	10.80	16.79	1.12	4.86	8.95	0.58	3.53	7.36	0.61	3.05	6.30
BDS [36] small sample	0.76	4.69	9.45	1.03	4.74	9.99	0.83	4.82	10.37	1.25	5.33	10.48
Runs test	0.59	3.25	6.65	0.93	4.23	8.99	0.95	4.41	9.31	0.88	4.50	9.38

Note: 10,000 Monte Carlo simulations were conducted applying the normally distributed pseudorandom numbers from MatLab R2010a.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

of these five models work as a size test for an $\alpha = 5\%$ and a conservative approach will be considered, a percentage of rejection equal to or less than 5% will be considered as a good result. For the nonlinear model we will follow the approach by Liu et al. [47]; a rejection percentage larger than 90% when the process is not independent suggests a very good power, and a percentage less than 50% means that the test does not easily detect nonlinearity.

Table 3 shows the results of the power experiment when a very small sample size $T = 50$ is considered and computing 5,000 Monte Carlo simulations. SRS test presents the best results considering 3 symbols and a length of 2, SRS(3,2). The five processes generated by the random distribution functions show a percentage of rejection less than 5% showing the conservative character of the test. However, consider that it is able to detect three of the four deterministic chaotic processes with a percentage of rejection larger than 95% for Lorenz, Logistic, and Henon. The BDS test performance is

worse than SRS. Note that simulated critical values for small samples have good performance for three of the five random distributions, but it is not able to detect truncated normal distribution, Uniform, and Beta(1/2,1/2) for this sample size. BDS test is able to detect two of the chaotic processes. On the other hand, runs test has the best performance detecting the five random distributions, but it just detects one of the four chaotic processes.

In summary, no test with a good size is able to detect the stochastic nonlinear processes when the sample size is 50 with a percentage larger than 50%. SRS(3,2) detects 11/20 processes, BDS 5/20, and runs test 8/20.

Table 4 shows the results for the experiment when $T = 200$. As expected with a sample larger than 50 the results improve. For the SRS the best results are obtained with 3 symbols and a length of 3. The six random processes and the residuals of AR(1) and MA(1) are detected by the test. Four of five chaotic processes are correctly rejected at

TABLE 3: Power test for a time series size of 50.

Generator process	SRS(2,2)	SRS(2,3)	SRS(3,2)	SRS(4,2)	SRS(5,2)	BDS (a)	BDS (b)	BDS (s)	Runs test
Normal	0.32	1.66	0.46	24.72	91.02	20.48	11.14	5.28	3.40
CHI-2(4)	0.46	2.14	0.44	25.36	91.36	15.26	7.56	3.04	4.12
<i>t</i> -Student(4)	0.48	1.86	0.42	24.58	91.44	14.72	7.80	3.32	3.72
Truncated norm.	0.28	1.62	0.50	24.92	91.62	32.94	18.30	13.30	3.06
Beta(1/2,1/2)	0.30	1.80	0.46	24.12	91.06	35.12	18.16	14.38	3.40
Uniform(0,1)	0.34	1.68	0.52	24.42	91.26	32.70	18.10	12.88	3.26
AR(1)	0.10	0.60	0.26	22.34	90.90	21.22	11.16	5.16	0.90
MA(2)	0.02	0.44	0.32	20.54	90.18	12.40	7.30	2.40	0.86
Logistic	10.54	40.20	100.00	100.00	100.00	59.36	68.42	37.94	22.88
Henon	53.94	77.84	100.00	100.00	100.00	99.18	82.34	96.44	22.38
Anosov	0.98	2.14	0.96	24.50	89.82	31.12	17.78	12.60	4.80
Lorenz	90.72	96.84	95.84	99.78	100.00	96.82	95.82	88.08	98.94
TAR	0.24	1.44	0.50	33.80	95.42	24.32	14.28	6.76	5.04
NLSIGN	0.06	0.94	0.40	23.10	90.28	21.50	12.44	5.32	2.40
Bilinear	0.16	5.38	0.48	28.50	93.30	63.08	60.18	44.64	1.54
NLAR	0.02	0.28	0.40	21.28	90.30	21.50	11.44	5.32	0.58
NLMA	0.14	1.18	0.26	20.80	89.56	15.88	7.56	2.82	1.58
BLMA	0.76	2.34	1.44	39.74	96.46	57.24	51.62	35.58	10.56
Modular	0.00	20.18	0.06	55.68	96.02	70.28	66.10	47.84	19.60
GARCH	0.10	0.88	3.12	37.40	93.52	59.14	58.42	38.22	3.24

Note: 5,000 Monte Carlo simulations were conducted applying the pseudorandom numbers from MatLab R2010a. SRS(a,w) indicates the SRS test for a symbolization of a symbols and a length of w . BDS (a), (b), and (s) applied the Kanzler [46] optimal parameters, the Liu et al. [47] best parameters, and the Kanzler [46] simulated critical values for small sample.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

a is the alphabet size or number of symbols, w is known as the word or a subsequence of symbols.

TABLE 4: Power test for a time series size of 200.

Generator process	SRS(2,3)	SRS(2,4)	SRS(3,3)	SRS(4,2)	SRS(5,2)	BDS (a)	BDS (b)	BDS (s)	Runs test
Normal	1.92	3.70	2.40	0.28	0.32	8.42	4.92	4.60	3.72
CHI-2(4)	1.62	3.34	2.22	0.20	0.32	7.26	3.54	3.56	3.74
<i>t</i> -Student(4)	1.72	3.50	2.18	0.18	0.30	6.56	3.16	3.30	4.04
Truncated norm.	2.02	3.66	2.42	0.26	0.30	12.86	7.48	7.96	3.96
Beta(1/2,1/2)	1.96	3.26	2.36	0.24	0.28	13.52	6.66	7.92	4.02
Uniform(0,1)	1.88	3.64	2.74	0.38	0.30	12.22	6.64	7.56	4.20
AR(1)	0.76	2.00	1.74	0.14	0.26	8.88	4.62	4.72	1.36
MA(2)	0.36	1.28	1.30	0.14	0.14	6.54	3.72	3.22	0.76
Logistic	20.62	33.74	100.00	100.00	100.00	69.00	62.80	64.26	15.22
Henon	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	83.36
Anosov	2.66	5.32	7.32	1.10	1.02	14.36	7.92	9.20	5.38
Lorenz	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
TAR	17.16	21.50	6.26	15.04	40.02	17.42	10.16	11.46	38.44
NLSIGN	2.74	4.98	4.12	0.82	1.28	9.70	5.24	5.26	8.78
Bilinear	67.50	84.08	81.76	11.52	15.66	99.02	99.02	98.50	7.20
NLAR	0.36	1.42	3.40	0.50	0.20	9.14	4.96	5.30	1.08
NLMA	4.68	12.94	32.72	0.58	0.70	7.94	5.46	4.38	5.60
BLMA	17.38	30.90	60.06	39.26	42.48	98.74	98.68	97.78	39.66
Modular	84.48	100.00	96.64	79.66	61.44	77.48	80.82	71.18	91.84
GARCH	5.04	7.60	64.18	59.30	63.22	99.54	99.62	99.18	15.30

Note: 5,000 Monte Carlo simulations were conducted applying the pseudorandom numbers from MatLab R2010a. SRS(a,w) indicates the SRS test for a symbolization of a symbols and a length of w . BDS (a), (b), and (s) applied the Kanzler [46] optimal parameters, the Liu et al. [47] best parameters, and the Kanzler [36] simulated critical values for small sample.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

TABLE 5: Power test for a time series size of 500.

Generator process	SRS(2,3)	SRS(2,4)	SRS(3,3)	SRS(4,3)	SRS(5,2)	BDS (a)	BDS (b)	BDS (s)	Runs test
Normal	2.32	3.94	1.98	3.22	0.22	6.32	3.30	5.12	4.72
CHI-2(4)	2.18	4.28	2.06	3.00	0.30	5.72	3.24	4.70	4.90
<i>t</i> -Student(4)	1.90	3.62	1.84	2.78	0.40	5.84	3.26	4.78	4.12
Truncated norm.	1.64	3.44	1.32	2.84	0.16	8.28	4.46	6.84	4.50
Beta(1/2,1/2)	1.82	3.64	1.94	3.38	0.32	8.16	4.02	6.84	4.18
Uniform(0,1)	1.42	3.30	1.64	2.94	0.32	7.58	4.08	6.50	3.80
AR(1)	0.56	1.78	1.48	2.30	0.14	6.02	3.14	4.88	1.44
MA(2)	0.34	1.46	0.92	2.38	0.04	5.84	3.50	4.60	1.08
Logistic	17.84	31.28	100.00	100.00	100	84.90	65.48	84.02	13.84
Henon	100.00	100.00	100.00	100.00	100	100.00	100.00	100.00	99.92
Anosov	3.18	6.22	6.30	93.46	0.82	11.04	6.02	9.16	6.04
Lorenz	100.00	100.00	100.00	100.00	100	100.00	100.00	100.00	100.00
TAR	70.60	72.10	20.32	86.56	97.7	23.08	13.94	21.54	88.74
NLSIGN	11.04	14.56	8.38	17.48	3.84	7.18	4.00	6.18	25.24
Bilinear	99.72	99.94	100.00	100.00	66.48	100.00	100.00	100.00	13.22
NLAR	0.46	1.82	6.24	6.30	1.34	7.16	4.34	5.86	1.34
NLMA	13.68	42.06	95.60	97.22	1.32	8.32	7.56	7.02	14.90
BLMA	61.44	85.64	99.40	99.90	98.98	100.00	100.00	100.00	77.34
Modular	99.98	100.00	100.00	100.00	99.8	94.42	96.52	93.88	100.00
GARCH	14.44	17.06	97.42	99.60	99.2	100.00	100.00	100.00	32.58

Note: 5,000 Monte Carlo simulations were conducted applying the pseudorandom numbers from MatLab R2010a. SRS(a,w) indicates the SRS test for a symbolization of a symbols and a length of w . BDS (a), (b), and (s) applied the Kanzler [46] optimal parameters, the Liu et al. [47] best parameters, and the Kanzler [46] simulated critical values for small sample.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

TABLE 6: Power test for a time series size of 2,000.

Generator process	SRS(2,4)	SRS(3,3)	SRS(3,4)	SRS(4,3)	SRS(5,3)	BDS	BDS Liu	BDS small	Runs test
Normal	3.72	1.90	3.78	2.06	1.84	4.86	2.76	4.64	4.78
CHI-2(4)	3.32	2.08	4.38	1.80	2.02	5.10	2.88	4.92	4.30
<i>t</i> -Student(4)	3.76	1.76	3.86	1.98	1.76	4.78	2.90	4.62	4.60
Truncated norm.	3.76	1.70	3.68	1.88	1.80	5.62	3.12	2.72	5.32
Beta(1/2,1/2)	3.70	1.80	4.12	1.72	1.88	5.44	2.96	5.18	5.32
Uniform(0,1)	3.66	2.14	4.08	1.74	2.10	6.32	3.64	6.18	5.06
AR(1)	1.78	1.20	2.82	1.36	1.22	5.32	2.92	5.08	1.72
MA(2)	2.20	0.86	3.16	1.22	1.34	5.20	2.94	4.94	1.50
Logistic	29.84	100.00	100.00	100.00	100.00	99.74	79.36	99.72	11.38
Henon	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Anosov	6.50	8.08	100.00	100.00	100.00	7.92	3.94	7.62	6.12
Lorenz	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
TAR	100.00	98.52	97.42	100.00	100.00	57.44	41.62	56.56	100.00
NLSIGN	69.70	73.22	83.50	92.74	92.72	8.04	4.56	7.54	84.84
Bilinear	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	21.20
NLAR	2.94	37.30	36.76	34.64	27.62	6.54	3.72	6.30	3.36
NLMA	99.64	100.00	100.00	100.00	100.00	14.04	14.36	13.78	48.36
BLMA	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.98
Modular	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
GARCH	40.58	100.00	100.00	100.00	100.00	100.00	100.00	100.00	61.86

Note: 5,000 Monte Carlo simulations were conducted applying the pseudorandom numbers from MatLab R2010a. SRS(a,w) indicates the SRS test for a symbolization of a symbols and a length of w . BDS (a), (b), and (s) applied the Kanzler [46] optimal parameters, the Liu et al. [47] best parameters, and the Kanzler [46] simulated critical values for small sample.

Bold values refer to reasonable percentage of rejections considering the process and the critical values.

a is the alphabet size or number of symbols, w is known as the word or a subsequence of symbols.

100% as random. The Anosov process seems to be the more difficult process to be detected. In the case the nonlinear stochastic process detection of dependence improves, SRS is able to detect nonlinearity in Bilinear, BLMA, Modular, and exponential GARCH at a rate larger than 60%. The BDS still has problem to detect some random processes; truncated normal, Uniform, and Beta(1/2,1/2) are rejected at a percentage larger than 7%. As in the case of SRS, it is difficult to detect the Anosov process. The Logistic process is detected with a rate of 64% in the case of BDS, but SRS detects at 100%. The BDS detects the same nonlinear stochastic process as SRS but with a rate larger than 80%. Runs test presents the worst performance; even if it is able to detect the six random processes, it is not able to detect chaotic processes such as Logistic and Anosov. Moreover, runs test is able to detect just one nonlinear stochastic process.

In summary, for a sample size of 200 with a good size, SRS detects 15/20 processes, BDS 12/20, and runs test 11/20.

Table 5 shows the power results for a size of 500. Even if SRS for 3 symbols and a length of 3 still presents good results, applying 4 symbols and a length of 3 improves the results. Remember that a larger symbolization has acceptable results if the sample is large. In this case more processes are detected. Note that the chaotic Anosov process is rejected at 93.46% as a random process, and six over eight nonlinear stochastic processes are detected. BDS test applying the parameters suggested by Liu et al. [47] is finally able to detect all the random process but still has problem detecting nonlinearity in the chaotic Anosov process. In addition, it is able to detect dependence in four over eight nonlinear stochastic processes. Runs test still has problem detecting chaotic Logistic model and is able to detect three over eight nonlinear stochastic processes. NLSIGN and NLAR are the more difficult processes to be detected by the three tests at the sample size of 500. Note also that Anosov is only detected by BDS applying a symbolization of four symbols at a rate of 93.46%. The other tests are able to detect the process less than 12% of the time.

In summary, for a sample size of 500 with a good size, SRS is able to detect 18/20 processes, BDS 15/20, and runs test 13/20.

Finally, Table 6 shows the results for a sample size of 2,000. BDS test presents the best results when considering 4 symbols and a length of 3. However, other symbolizations present good results, such as 3 symbols and $w = 3$ or 4. In the present case the only process that is not detected by SRS is the nonlinearity generated by the NLAR; maybe this process requires a larger sample size. The BDS test cannot detect nonlinearity in 4 processes: chaotic Anosov and the stochastic processes NLSIGN, NLAR, and NLMA. The runs test cannot detect five nonlinear processes: Logistic, Anosov, Bilinear, NLAR and NLMA. However, note that it is able to detect NLSIGN which is not detected by BDS. The proposed SRS is the only one able to detect chaotic Anosov and nonlinear process NLMA when $T = 2000$. In summary, for $T = 2000$ with a good size the SRS test is able to detect 19/20, the BDS test 15/20, and the runs test 12/20.

5. Detecting Independence and Nonlinearity in Financial Time Series

As asserted by Brooks [2] testing for nonlinear dependence is important in financial econometrics due to its profound implications for model adequacy, market efficiency, and predictability. For instance, Shahwan and Said [4] find that modelization with artificial neural networks (ANN) is more relevant to fit a high-dimensional chaotic process than the Bayesian and ARIMA methods. They mention that most recent empirical work implies that the presence of low-dimensional deterministic chaos increases the complexity of the financial time series behavior. Therefore it is important in financial time series to determine the existence of dependence.

The test is applied to financial time series studying the performance in practice. Four asset prices from the New York Stock Exchange, six stock indices, and five exchange rates were considered at different frequencies (daily, weekly, and monthly). Table 7 summarizes some statistics related to the returns of the series.

Test of normality proposed by Jarque and Bera [49] is applied to the returns and null hypothesis of normality is widely rejected for almost all the series. The exchange rate between US dollar and euro at a monthly level is not detected. The latter could be related to the small sample size $T = 171$.

The next step is to apply the different tests to the raw returns of financial time series and to the residuals of a GARCH model. Considering the results obtained in the size and power experiment and the sample size of the financial series, the SRS(2,4), SRS(3,3), SRS(3,4), and SRS(4,3) were selected. In addition, BDS test and runs test are applied in order to compare the results.

Table 8 shows the results of the tests for the different financial returns. Note that the BDS test rejects randomness for almost all the cases but the NIKKEI index at a monthly frequency and the exchange rate between US dollar and euro for a monthly frequency. The SRS rejects randomness in less cases than the BDS test, in particular for the monthly frequency, maybe due to considering few data. This is clear for Coca Cola, IBM, Caterpillar, S&P 500, FTSE, DAX, and NIKKEI. Note that Coca Cola for a daily frequency and the exchange rate between dollar and yuan highlight a strong dependency. In the case of the exchange rate between US dollar and Chinese yuan the economic monetary policy should be considered. China has controlled the exchange rate and just in the last years it has introduced some flexibility. In the same way, the SHANGHAI index presents the largest degree of dependency and it is due to controls and constraints that have been presented in the Chinese stock market. The runs test rejects only 16 over 40 time series as generated by an independent process.

Table 9 shows the results for the different tests after applying a GARCH(1,1) model to the financial returns. The runs test does not reject independence in 28 over 40 cases. BDS test does not reject independence in residuals in 26 over 40 cases. However, note that the conservative SRS test in this case only rejects independence in 13 over 40 cases. This indicates that after applying a GARCH model to the financial

TABLE 7: Summary of the returns of financial time series.

Series	Frequency	T	Mean	Variance	Skewness	Kurtosis	Normality test statistic
Citigroup	Daily	9170	0.0536%	0.0007	1.2025	53.4066	973018.60*
	Weekly	1895	0.2794%	0.0040	3.7869	89.8884	600633.18*
	Monthly	436	1.0121%	0.0119	0.1130	10.5866	1046.54*
Coca Cola	Daily	12926	0.0570%	0.0003	-2.9897	90.2182	4116265.26*
	Weekly	2678	0.2620%	0.0013	-1.7079	26.7327	64150.50*
	Monthly	616	1.0907%	0.0047	-1.3229	13.1023	2799.09*
IBM	Daily	12926	0.0470%	0.0003	0.0294	13.0155	54027.14*
	Weekly	2678	0.2226%	0.0012	0.0999	6.1514	1112.62*
	Monthly	616	0.9686%	0.0049	0.2140	4.8790	95.32*
Caterpillar	Daily	12927	0.0517%	0.0004	-1.2125	37.7400	653214.89*
	Weekly	2678	0.2534%	0.0019	-0.2678	10.9607	7103.35*
	Monthly	616	1.0656%	0.0076	-0.1546	6.0064	234.43*
S&P 500	Daily	15940	0.0335%	0.0001	-0.6504	24.2393	300735.36*
	Weekly	3304	0.1601%	0.0004	-0.3453	7.9369	3421.03*
	Monthly	760	0.6904%	0.0018	-0.4240	4.7084	115.20*
FTSE	Daily	7350	0.0305%	0.0001	-0.2089	10.9092	19211.26*
	Weekly	1518	0.1467%	0.0006	-0.7059	11.0775	4252.88*
	Monthly	349	0.6105%	0.0021	-0.8005	6.1044	177.41*
DAX	Daily	5684	0.0413%	0.0002	0.0486	7.8679	5614.43*
	Weekly	1171	0.1987%	0.0010	-0.3384	7.1055	844.73*
	Monthly	270	0.8400%	0.0038	-0.5284	4.9958	57.38*
BOVESPA	Daily	4955	0.1856%	0.0006	0.9411	16.2854	37171.74*
	Weekly	1045	0.8837%	0.0029	0.5525	6.6351	628.51*
	Monthly	241	4.1075%	0.0198	2.0642	12.5057	1078.50*
NIKKEI	Daily	7213	0.0155%	0.0002	-0.0492	11.0987	19715.15*
	Weekly	1520	0.0650%	0.0008	-0.4998	7.7195	1473.98*
	Monthly	352	0.2828%	0.0037	-0.2933	3.7348	12.97*
SHANGHAI	Daily	5756	0.0844%	0.0007	12.2275	456.4849	49464803.30*
	Weekly	1160	0.4507%	0.0052	10.5452	189.3924	1700702.47*
	Monthly	269	2.1082%	0.0301	5.6259	52.8034	29219.85*
Ex. Canada/USA	Daily	10644	0.0008%	0.0000	-0.0052	15.2475	66525.07*
	Monthly	507	0.0111%	0.0002	0.8962	13.9378	2595.15*
Ex. China/USA	Daily	8133	0.0190%	0.0000	55.8492	3703.1319	4643757374.76*
	Monthly	387	0.3930%	0.0009	13.0613	211.0026	708653.19*
Ex. US dollar/euro	Daily	1680	-0.0540%	0.0006	-37.1096	1471.7500	151391452.12*
	Monthly	171	0.0994%	0.0006	0.0775	2.9535	0.19
Ex. USA/UK	Daily	10644	-0.0022%	0.0000	-0.1398	7.7514	10046.90*
	Monthly	507	-0.0605%	0.0006	-0.1249	4.5997	55.38*
Ex. Japan/USA	Daily	10644	-0.0099%	0.0000	-0.5942	12.1587	37827.76*
	Monthly	507	-0.2199%	0.0007	-0.3986	3.8298	27.97*

Note: own computations. * Denotes rejection of the Jarque-Bera null hypothesis of normality at 5% of significance.

series SRS still detects some dependence or nonlinearity. As an example, after applying GARCH(1,1) to the S&P 500 at different frequencies (daily, weekly, and monthly) the BDS test does not reject the fact that the residuals are independent and so GARCH(1,1) is a good model. However, the SRS test still detects some dependence in residuals and GARCH(1,1) is not the best model.

6. Conclusions

In the present study a new independence test is derived; it can be also applied to detect nonlinearity in time series. The test has the advantage of being easy to compute and it is powerful in detecting dependence generated by different nonlinearity and chaotic systems. The introduced test is based

TABLE 8: SRS, BDS, and runs tests on the raw returns of assets, stock indices, and exchange rates.

Generator process	T (a)	SRS(2,4)	SRS(3,3)	SRS(3,4)	SRS(4,3) (b)	BDS (c)	Runs test (d)
Citigroup (daily)	9170	1.98	539.28*	1039.33*	1223.32*	0.00*	No rejection
Citigroup (weekly)	1895	19.83	64.39*	148.24*	405.51*	0.00*	No rejection
Citigroup (monthly)	436	5.87	40.55*	125.77*	67.09	0.00*	No rejection
Coca Cola (daily)	12926	13253.99*	694795.51*	20514840.97*	24089.11*	0.00*	Rejection
Coca Cola (weekly)	12678	797.78*	86.03*	182.59*	782.92*	0.00*	Rejection
Coca Cola (monthly)	616	8.86	32.50	84.82	60.40	0.00*	No rejection
IBM (daily)	12926	207.11*	205.26*	430.68*	547.90*	0.00*	No rejection
IBM (weekly)	2678	24.50	26.60	98.89	104.22*	0.00*	No rejection
IBM (monthly)	616	4.16	23.71	85.61	70.61	0.00*	No rejection
Caterpillar (daily)	12927	829.84*	338.36*	617.75*	1029.65*	0.00*	Rejection
Caterpillar (weekly)	2678	86.68*	34.33	95.75	84.37*	0.00*	Rejection
Caterpillar (monthly)	616	14.34	14.21	49.93	50.18	0.01*	No rejection
S&P 500 (daily)	15940	245.05*	513.09*	979.56*	883.17*	0.00*	Rejection
S&P 500 (weekly)	3304	10.13	100.03*	228.23*	186.48*	0.00*	No rejection
S&P 500 (monthly)	760	10.30	25.50	78.15	49.01	0.00*	No rejection
FTSE (daily)	7350	25.64*	116.49*	272.55*	209.85*	0.00*	No rejection
FTSE (weekly)	1518	11.82	53.53*	129.33*	93.12*	0.00*	No rejection
FTSE (monthly)	349	14.23	20.50	54748.54 (N)	63.01	0.01*	No rejection
DAX (daily)	5684	13.04	105.82*	248.25*	286.08*	0.00*	No rejection
DAX (weekly)	1171	11.67	81.25*	194.02*	152.88*	0.00*	No rejection
DAX (monthly)	270	6.92	21.95	84432.94 (N)	16661.79 (N)	0.03*	No rejection
BOVESPA (daily)	4955	19.27	139.95*	328.87*	166.32*	0.00*	No rejection
BOVESPA (weekly)	1045	20.94	45.83*	141.13*	81.08	0.00*	No rejection
BOVESPA (monthly)	241	21.23	41.05*	56522.63 (N)	29754.49 (N)	0.00*	Rejection
NIKKEI (daily)	7213	17.00	184.61*	468.53*	363.29*	0.00*	No rejection
NIKKEI (weekly)	1520	14.26	77.85*	196.49*	119.78*	0.00*	No rejection
NIKKEI (monthly)	352	9.46	14.58	27634.89 (N)	70.57	0.17	No rejection
SHANGHAI (daily)	5756	194.96*	396.92*	1064.22*	854.35*	0.00*	Rejection
SHANGHAI (weekly)	1160	176.95*	58.17*	153.21*	178.04*	0.00*	Rejection
SHANGHAI (monthly)	269	25.13*	40.72*	63142.02 (N)	33166.27 (N)	0.00*	Rejection
Ex. Canada/USA (daily)	10644	40.12*	700.20*	1645.48*	1296.32*	0.00*	Rejection
Ex. Canada/USA (monthly)	507	25.97*	64.03*	39963.21*	109.02*	0.00*	Rejection
Ex. China/USA (daily)	8133	14387.35*	3913576.79*	41867540.66*	8623222.42*	0.00*	Rejection
Ex. China/USA (monthly)	387	809.75*	837.29*	366410.86 (N)	384311.27 (N)	0.00*	Rejection
Ex. US dollar/euro (daily)	1680	9.22	14.31	60.32	53.36	0.00*	No rejection
Ex. US dollar/euro (monthly)	171	22.10	8508.98 (N)	278836.18 (N)	94399.98 (N)	0.08	Rejection
Ex. USA/UK (daily)	10644	4.29	1311.28*	3392.54*	782.87*	0.00*	No rejection
Ex. USA/UK (monthly)	507	66.98*	98.78*	159444.32 (N)	31462.36*	0.00*	Rejection
Ex. Japan/USA (daily)	10644	74.14*	1802.25*	4770.12*	1066.80*	0.00*	No rejection
Ex. Japan/USA (monthly)	507	87.75*	92.69*	79848.32 (N)	137.91*	0.00*	Rejection

Note: own calculations. *Indicate rejection of the independence null hypothesis at 5% of significance. (a) Number of observations. (b) SRS test applying selected number of symbols and length according to the power of size experiments. (c) P value of the BDS test. (d) Rejection or no rejection of randomness null hypothesis. (N) Not enough sample size for the SRS test.

on the symbolic time series analysis usually considered when studying series highly contaminated by noise.

There are many tests of independence, randomness, and nonlinearity in the literature and the topic is a growing area of investigation due to the necessity generated by applied economics and statistics.

The present paper suggests that the test can be applied to detect nonlinear dependence in time series through transforming the series to a symbolic one and computing a statistic that is asymptotically distributed as a Chi-square with $n - 1$ degrees of freedom, where n is the quantity of possible events. Moreover, the test is related to the Pearson independence test

TABLE 9: SRS, BDS, and runs tests on the residuals of a GARCH(1, 1) model applied to assets, stock indices, and exchange rates.

Generator process	T (a)	SRS(2,4)	SRS(3,3)	SRS(3,4)	SRS(4,3) (b)	BDS (c)	runs test (d)
Citigroup (daily)	9170	1.98	99.98*	210.67*	426.33*	0.00*	No rejection
Citigroup (weekly)	1895	19.83	31.22	93.44	279.40*	0.43	No rejection
Citigroup (monthly)	436	7.06	20.15	76.76	51.16	0.18	No rejection
Coca Cola (daily)	12926	11512.60*	694785.94*	23577565.97*	23719.11*	0.00*	Rejection
Coca Cola (weekly)	2678	797.78*	82.32*	168.36*	715.33*	0.39	No rejection
Coca Cola (monthly)	616	8.96	29.07	88.52	48.51	0.26	No rejection
IBM (daily)	12926	207.11*	27.22	88.28	230.90*	0.02*	No rejection
IBM (weekly)	2678	25.24*	28.80	107.19*	62.37*	0.96	No rejection
IBM (monthly)	616	4.37	23.80	91.42	66.03	0.93	No rejection
Caterpillar (daily)	12927	829.84*	205.44*	384.42*	1350.71*	0.00*	Rejection
Caterpillar (weekly)	2678	86.68*	33.22	85.21	78.68	0.81	No rejection
Caterpillar (monthly)	616	13.82	19.49	65.52	56.64	0.74	No rejection
S&P 500 (daily)	15940	224.88*	402.18*	695.57*	595.81*	0.27	Rejection
S&P 500 (weekly)	3304	9.34	85.07*	187.74*	178.62*	0.80	No rejection
S&P 500 (monthly)	760	11.36	39.53*	101.26*	72.82	0.47	No rejection
FTSE (daily)	7350	6.85	60.06*	152.68*	163.52*	0.00*	No rejection
FTSE (weekly)	1518	10.09	52.96*	123.77*	85.27*	0.18	No rejection
FTSE (monthly)	349	16.73	10.85	54732.62 (N)	62.64	0.47	No rejection
DAX (daily)	5684	11.08	101.52*	197.55*	186.67*	0.00*	No rejection
DAX (weekly)	1171	10.63	59.21*	131.33*	69.67	0.18	No rejection
DAX (monthly)	270	4.16	24.57	42260.90 (N)	50.57	0.91	No rejection
BOVESPA (daily)	4955	15.99	134.38*	279.73*	212.50*	0.10	No rejection
BOVESPA (weekly)	1045	17.62	36.40*	111.90*	85.99*	0.06	No rejection
BOVESPA (monthly)	241	19.88	20.27	75278.34 (N)	58.51	0.42	No rejection
NIKKEI (daily)	7213	11.96	106.49*	245.12*	160.36*	0.00*	No rejection
NIKKEI (weekly)	1520	8.84	41.13*	127.62*	71.46	0.71	No rejection
NIKKEI (monthly)	352	9.46	20.13	27638.60 (N)	58.50	0.59	No rejection
SHANGHAI (daily)	5756	194.96*	206.46*	582.42*	506.30*	0.18	Rejection
SHANGHAI (weekly)	1160	176.95*	62.92*	174.07*	149.97*	0.01*	Rejection
SHANGHAI (monthly)	269	18.03	33.44	84149.32 (N)	16620.42 (N)	0.37	No rejection
Ex. Canada/USA (daily)	10644	40.12*	62.44*	140.98*	109.43*	0.42	Rejection
Ex. Canada/USA (monthly)	507	24.76*	45.10*	39926.89*	74.55	0.03*	Rejection
Ex. China/USA (daily)	8133	14387.35*	3913576.79*	41867540.66*	8623222.42*	0.00*	Rejection
Ex. China/USA (monthly)	387	809.75*	826.35*	366368.25 (N)	503663.60 (N)	0.00*	Rejection
Ex. US dollar/euro (daily)	1680	9.22	24.74	77.71	82.80*	0.00*	No rejection
Ex. US dollar/euro (monthly)	171	22.10	4291.65 (N)	252306.64(N)	104887.07*	0.80	Rejection
Ex. USA/UK (daily)	10644	3.56	344.80*	955.78*	216.87*	0.07	No rejection
Ex. USA/UK (monthly)	507	66.98*	104.24*	119642.14 (N)	31433.72 (N)	0.03*	Rejection
Ex. Japan/USA (daily)	10644	74.14*	689.43*	1832.69*	509.97*	0.29	No rejection
Ex. Japan/USA (monthly)	507	87.75*	86.49*	79840.61 (N)	137.41*	0.05*	Rejection

Note: own calculations. *Indicate rejection of the independence null hypothesis at 5% of significance. (a) Number of observations. (b) SRS test applying selected number of symbols and length according to the power of size experiments. (c) P value of the BDS test. (d) Rejection or no rejection of randomness null hypothesis. (N) Not enough sample size for the SRS test.

and it is connected with the Shannon entropy widely studied in information theory when measuring uncertainty in data.

Size and power experiments were conducted for small samples and the results are compared with the well-known BDS test and runs test. Experiments indicate that for a sample size less than 2,000 it is not advisable to apply more than four symbols. The three tests are conservative even if SRS test

seems to be the most conservative. The test power detecting randomness and independence in nonlinear and chaotic systems was studied. Results suggested that even if the test is conservative it presented the best performance detecting the 20 different processes. In very small samples (about 50 observations) SRS test was able to detect dependence in most of the chaotic processes. In particular it was the only one

identifying nonlinearity generated by the very simple logistic process. The SRS test is able to detect the chaotic process of Anosov and the stochastic model NLMA for a sample size of 2,000, whereas no other test can do it. It is also observed that BDS and runs test present different performances, even if the latter is better. On the one hand, NLSING was detected by the runs test but not by BDS; on the other hand Bilinear was detected by BDS test and runs test cannot identify this process; SRS test was able to detect both of the models. It should be mentioned that for a sample size of 2,000 no test detected the NLAR process. It is important to remark that the objective was to study the performance in small sample size as is common to find in economics. However, note that all tests performance improves as the sample size increases.

The SRS test was applied to financial series such as stock asset, stock indices, and exchange rates, considering different frequencies (daily, weekly, and monthly). The test was applied to the raw returns and to the residuals of a GARCH(1,1). Comparing with other tests it was noted that the SRS test rejected independence less times than the BDS in the raw returns case, in particular when there are monthly data or few data. However, when the tests are applied to the residuals of a GARCH(1,1), the results change. In this case the BDS rejected the independence few times whereas the SRS test still detects nonlinearity in the residuals. This suggests that BDS considers that the GARCH(1,1) model is a good model most of the time but the SRS test would suggest that GARCH(1,1) is not a good model considering all the nonlinear components.

In applied economics it is generally worked with finite samples. For this reason it is necessary to design a more powerful test detecting nonlinearity or dependence in small samples. In this sense, the symbolic time series analysis can be an approach to go further. A future research line is to develop the test for multidimensional time series. Symbolization permits transforming a time series with many dimensions to a one dimensional series permitting the simplification of the analysis.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] M. Gabr and L. Fathey, "Time series classification," *Journal of Statistics Applications & Probability*, vol. 2, no. 2, pp. 123–133, 2013.
- [2] C. Brooks, "Testing for non-linearity in daily sterling exchange rates," *Applied Financial Economics*, vol. 6, no. 4, pp. 307–317, 1996.
- [3] A. Wald and J. Wolfowitz, "An exact test for randomness in the non-parametric case based on serial correlation," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 378–388, 1943.
- [4] T. Shahwan and R. Said, "A comparison of Bayesian methods and artificial neural networks for forecasting chaotic financial time series," *Journal of Statistics Applications & Probability*, vol. 1, no. 2, pp. 89–100, 2012.
- [5] C. W. Granger, E. Maasoumi, and J. Racine, "A dependence metric for possibly nonlinear processes," *Journal of Time Series Analysis*, vol. 25, no. 5, pp. 649–669, 2004.
- [6] M. King, "Testing for autocorrelation in linear regression model: a survey," in *Specification Analysis in the Linear Regression Model*, M. King and D. E. A. Giles, Eds., Rutledge and Kegan Paul, London, UK.
- [7] M. Matilla-García and M. Ruiz Marín, "A non-parametric independence test using permutation entropy," *Journal of Econometrics*, vol. 144, no. 1, pp. 139–155, 2008.
- [8] J. Durbin and G. S. Watson, "Testing for serial correlation in least squares regression—II," *Biometrika*, vol. 38, no. 1-2, pp. 159–178, 1951.
- [9] J. Durbin and G. S. Watson, "Testing for serial correlation in least squares regression—I," *Biometrika*, vol. 37, no. 3-4, pp. 409–428, 1950.
- [10] K. Wallis, "Testing for fourth order autocorrelation in quarterly regression equations," *Econometrica*, vol. 40, pp. 617–636, 1972.
- [11] H. Vinod, "Generalization of the Durbin-Watson statistic for higher order autoregressive processes," *Communications in Statistics*, vol. 2, pp. 115–144, 1973.
- [12] A. Azzalini and A. Bowman, "On the use of nonparametric regression for checking linear relationships," *Journal of the Royal Statistical Society B*, vol. 55, no. 2, pp. 549–557, 1993.
- [13] J. Dufour, Y. Lepage, and H. Zeidan, "Nonparametric testing for time series: a bibliography," *The Canadian Journal of Statistics*, vol. 10, pp. 1–38, 1982.
- [14] W. A. Barnett, A. R. Gallant, M. J. Hinich, J. A. Jungeilges, D. T. Kaplan, and M. J. Jensen, "A single-blind controlled competition among tests for nonlinearity and chaos," *Journal of Econometrics*, vol. 82, no. 1, pp. 157–192, 1998.
- [15] M. Hinich, "Testing for Gaussianity and linearity of a stationary time series," *Journal of Time Series Analysis*, vol. 3, no. 3, pp. 169–176, 1982.
- [16] D. Nychka, S. Ellner, R. Gallant, and D. McCaffrey, "Finding chaos in noisy systems," *Journal of the Royal Statistical Society B*, vol. 54, pp. 399–426, 1992.
- [17] H. White, "Additional hidden unit test for neglected nonlinearity in multilayer feedforward networks," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '89)*, vol. 2, pp. 451–455, IEEE Press, New York, NY, USA, June 1989.
- [18] D. T. Kaplan, "Exceptional events as evidence for determinism," *Physica D*, vol. 73, no. 1-2, pp. 38–48, 1994.
- [19] W. Brock, W. Dechert, B. LeBaron, and J. Scheinkman, "A test for independence based on the correlation dimension," *Econometric Reviews*, vol. 15, pp. 197–235, 1996.
- [20] H. Joe, "Relative entropy measures of multivariate dependence," *Journal of the American Statistical Association*, vol. 84, pp. 157–164, 1989.
- [21] H. Joe, "Estimation of entropy and other functionals of a multivariate density," *Annals of the Institute of Statistical Mathematics*, vol. 41, no. 4, pp. 683–697, 1989.
- [22] P. Robinson, "Consistent nonparametric entropy-based testing," *Review of Economic Studies*, vol. 58, pp. 437–453, 1991.
- [23] C. Granger and J. Liu, "Using the mutual information coefficient to identify lags in nonlinear models," *Journal of Time Series Analysis*, vol. 15, pp. 371–384, 1994.
- [24] Y. Hong and H. White, "Asymptotic distribution theory for nonparametric entropy measures of serial dependence," *Econometrica*, vol. 73, no. 3, pp. 837–901, 2005.

- [25] H. Elsinger, "Comment on a new test for chaos and determinism based on symbolic dynamics," *Journal of Economic Behavior and Organization*, 2013.
- [26] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, "Detecting determinism in time series with ordinal patterns: a comparative study," *International Journal of Bifurcation and Chaos*, vol. 20, no. 9, pp. 2915–2924, 2010.
- [27] L. Zunino, M. Soriano, and O. Rosso, "Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach," *Physical Review E*, vol. 86, Article ID 046210, 2012.
- [28] G. A. Gottwald and I. Melbourne, "A new test for chaos in deterministic systems," *Proceedings of the Royal Society A*, vol. 460, no. 2042, pp. 603–611, 2004.
- [29] G. A. Gottwald and I. Melbourne, "Testing for chaos in deterministic systems with noise," *Physica D*, vol. 212, no. 1-2, pp. 100–110, 2005.
- [30] G. A. Gottwald and I. Melbourne, "On the implementation of the 0-1 test for chaos," *SIAM Journal on Applied Dynamical Systems*, vol. 8, no. 1, pp. 129–145, 2009.
- [31] G. A. Gottwald and I. Melbourne, "On the validity of the 0-1 test for chaos," *Nonlinearity*, vol. 22, no. 6, pp. 1367–1382, 2009.
- [32] B. Xin and Y. Li, "0-1 test for chaos in a fractional order financial system with investment incentive," *Abstract and Applied Analysis*, vol. 2013, Article ID 876298, 10 pages, 2013.
- [33] J. Cánovas, A. Guillamón, and S. Vera, "Testing for independence: permutation based tests versus BDS test," *The European Physical Journal Special Topics*, vol. 222, pp. 275–288, 2013.
- [34] C. S. Daw, C. E. A. Finney, and E. R. Tracy, "A review of symbolic analysis of experimental data," *Review of Scientific Instruments*, vol. 74, no. 2, pp. 915–930, 2003.
- [35] C. Piccardi, "On the control of chaotic systems via symbolic time series analysis," *Chaos*, vol. 14, no. 4, pp. 1026–1034, 2004.
- [36] C. E. A. Finney, K. Nguyen, C. S. Daw, and J. S. Halow, "Symbol-sequence statistics for monitoring fluidization," in *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, pp. 405–412, Anaheim, Calif, USA, November 1998.
- [37] C. S. Daw, C. E. A. Finney, K. Nguyen, and J. S. Halow, "Symbol statistics: a new tool for understanding multiphase flow phenomena," in *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, pp. 221–229, November 1998.
- [38] C. Finney, J. Green, and C. Daw, "Symbolic time-series analysis of engine combustion measurements," SAE Paper 980624, 1998.
- [39] S. Williams, "Symbolic dynamics and its applications," in *Proceeding of Symposia in Applied Mathematics*, vol. 60, p. 150, 2004.
- [40] A. Mathai and S. Provost, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, 1992.
- [41] P. B. Patnaik, "The non-central chi²- and F-distributions and their applications," *Biometrika*, vol. 36, no. 1-2, pp. 202–232, 1949.
- [42] E. Pearson, "Note on an approximation to the distribution of noncentral χ^2 ," *Biometrika*, vol. 46, p. 364, 1959.
- [43] M. Siddiqui, "Approximations to the distribution of quadratic forms," *The Annals of Mathematical Statistics*, vol. 26, pp. 677–682, 1965.
- [44] H. Solomon and M. Stephens, "Approximation to density functions using Pearson curves," *Journal of the American Statistical Association*, vol. 73, pp. 153–160, 1978.
- [45] S. Oman and S. Zacks, "A mixture approximation to the distribution of weighted sum of chi-squared variables," *Journal of Statistical Computation and Simulation*, vol. 13, pp. 215–224, 1981.
- [46] L. Kanzler, "Very fast and correctly sized estimation of the BDS statistic," *Social Science Research Network*, 1999.
- [47] T. Liu, C. Granger, and W. Heller, "Using the correlation exponent to decide whether an economic series is chaotic," *Journal of Applied Econometrics*, vol. 7, pp. S25–S39, 1992.
- [48] D. Hsieh, "Chaos and nonlinear dynamics: application to financial markets," *The Journal of Finance*, vol. 46, no. 5, pp. 1839–1877.
- [49] C. M. Jarque and A. K. Bera, "Efficient tests for normality, homoscedasticity and serial independence of regression residuals," *Economics Letters*, vol. 6, no. 3, pp. 255–259, 1980.

