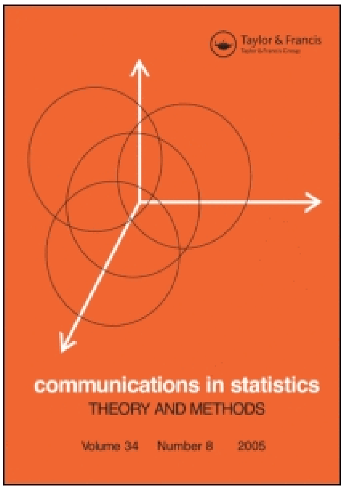

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Bispectral-Based Goodness-of-Fit Tests of Gaussianity and Linearity of Stationary Time Series

Nusrat Jahan [a]; Jane L. Harvill [b]
[a] Department of Mathematics and Statistics, James Madison University, Harrisonburg, Virginia, USA [b] Department of Statistical Science, Baylor University, Waco, Texas, USA

Online Publication Date: 01 December 2008

To cite this Article Jahan, Nusrat and Harvill, Jane L.(2008)'Bispectral-Based Goodness-of-Fit Tests of Gaussianity and Linearity of Stationary Time Series',Communications in Statistics - Theory and Methods,37:20,3216 — 3227

To link to this Article: DOI: 10.1080/03610920802133319

URL: http://dx.doi.org/10.1080/03610920802133319

Communications in Statistics—Theory and Methods, 37: 3216–3227, 2008

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920802133319

Time Series Analysis

Bispectral-Based Goodness-of-Fit Tests of Gaussianity and Linearity of Stationary Time Series

NUSRAT JAHAN[1] AND JANE L. HARVILL[2]

[1]Department of Mathematics and Statistics,
James Madison University, Harrisonburg, Virginia, USA
[2]Department of Statistical Science, Baylor University,
Waco, Texas, USA

Spectral domain tests for time series linearity typically suffer from a lack of power compared to time domain tests. We present two tests for Gaussianity and linearity of a stationary time series. The tests are two-stage procedures applying goodness-of-fit techniques to the estimated normalized bispectrum. We illustrate the performances of the tests are competitive with time domain tests. The new tests typically outperform Hinich's (1982) bispectral based test, especially when the length of the time series is not large.

Keywords Bispectral density function; Frequency domain analysis; Testing time series linearity.

Mathematics Subject Classification 62-07; 62G10; 62G20; 62M10; 62M15; 91B84.

1. Introduction

Contrary to only 30 years ago, the analysis of a time series realization as being from a nonlinear process is no longer uncommon. Many examples appear in the literature with the series being one that is commonly accepted as a realization from a nonlinear dynamic process (see, for example, Tong, 1990). The explosion of parametric and nonparametric methods for nonlinear time series analysis encourages the use of nonlinear models when they are appropriate. However, the

Received February 1, 2008; Accepted April 16, 2008
Address correspondence to Nusrat Jahan, Department of Mathematics and Statistics, James Madison University, Harrisonburg, VA 22807, USA; E-mail: jahannx@jmu.edu

principle of parsimony should always apply. And so, although it is possible to analyze a time series via more complicated nonlinear techniques, if a linear model will suffice, then linear time series analysis methods should be used. All of that being said, it is desirable to have reliable methods for testing the linearity of a time series.

In general, tests for time series linearity can be classified into one of two broad categories: parametric time domain tests or nonparametric frequency domain tests. In general, the time domain tests are parametric in that the alternative hypothesis designates a specific nonlinear model (e.g., Tsay, 1986, 1989), or a class of nonlinear models (e.g., Tsay, 1991). In contrast, frequency domain tests (e.g., Hinich, 1982; Subba Rao and Gabr, 1980) are nonparametric with an alternative specifying only that the underlying time series model is nonlinear. Due in large part to their parametric nature, time domain tests typically outperform the existing frequency domain tests, especially when the underlying model generating the realization is that model specified in the alternative hypothesis. Moreover, the power of spectral domain tests has been shown to be heavily dependent upon the choice of smoothing parameters used in estimating the spectral quantities that make up the test statistics (Chan and Tong, 1986). On the other hand, Harvill (1999) showed frequency domain tests can outperform time domain tests when the model in the time domain alternative is misspecified. Additionally, exploiting the properties of the bispectrum has proven a useful tool in many time series problems; see, for example, Hinich and Rothman (1998), Barnett et al. (1997), or Hinich and Messer (1995).

Motivated by the asymptotic sampling distribution of twice the square modulus of the estimated normalized bispectrum, Z_2^2 say, we develop a two-stage, bispectral-based, goodness-of-fit test for the Gaussianity and linearity of a time series. The two-stage approach was first proposed by Subba Rao and Gabr (1980) and later by Hinich (1982). Between the two, Hinich's test has long been preferred. For both tests, as well as the test we propose, the null hypothesis for the first stage of the test is that the time series is Gaussian. If rejected, the test proceeds to the second stage. In this stage, the null hypothesis is that the time series is linear, but not Gaussian. Hinich's test statistic for linearity computes p-values based on approximate normality of the standardized difference of quantiles (e.g., the interdecile range) of the Z_2^2 across frequencies on a lattice. Harvill and Newton (1995) and Harvill (1999) discussed two primary short-comings with Hinich's approach. First, using the difference of quantiles relies on bispectral estimates at only two frequency pairs (those included in the difference), thus effectively ignoring information contained in the other quantiles. Second, even for a reasonably long time series, the number of frequency pairs in the lattice is small (this is illustrated further in Sec. 5). The small number of frequency pairs necessarily implies that applying normal approximation to the distribution of the difference is questionable practice, and can result in an inflated Type I error probability.

The remainder of the article is organized as follows. In Sec. 2, the normalized bispectral density function is defined, and expressions are developed for the normalized bispectrum when the time series is Gaussian or linear. Section 3 contains a brief discussion of estimating the spectral quantities and the asymptotic properties of the estimators. Based on these properties, the proposed two-stage bispectral-based goodness-of fit test is developed in Sec. 4. In Sec. 5, an empirical study is provided for comparing the performance of the proposed test to existing time and frequency domain tests.

2. Bispectral Density Function

The zero mean process $\{X_t : t \in \mathscr{Z}\}$, where $\mathscr{Z}$ represents the set of integers, is linear if admits the representation

$$X_t = \sum_{j=0}^{\infty} \beta_j \varepsilon_{t-j},$$

where $\sum_{j=0}^{\infty} \beta_j^2 < \infty$ and the $\{\varepsilon_t\}$ is an i.i.d. sequence with finite constant variance σ_ε^2. Under the additional restriction that the process is up to third-order stationary, the autocovariance function and third-moment function can be defined as

$$\gamma_v = \mathrm{E}[X_t X_{t+v}] \quad \text{and} \quad \gamma_{u,v} = \mathrm{E}[X_t X_{t+u} X_{t+v}],$$

respectively. Furthermore, if $\sum_{v=-\infty}^{\infty} |\gamma_v| < \infty$, the spectral density function of $\{X_t\}$ exists and is defined as

$$I(\omega) = \sum_{v=-\infty}^{\infty} \gamma_v e^{-2\pi i v \omega} \quad \text{for } \omega \in [0, 1]. \tag{1}$$

Symmetries allow us to restrict our consideration of $I(\omega)$ to $\omega \in [0, 0.5]$.

If it can be further assumed that the process $\{X_t\}$ is up to sixth-order stationary, and if the third-moment function is absolutely summable, i.e., $\sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} |\gamma_{u,v}| < \infty$, then the bispectral density function of $\{X_t\}$ is the bivariate analogue of (1) given by

$$I(\omega_1, \omega_2) = \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \gamma_{u,v} e^{-2\pi i (u\omega_1 + v\omega_2)}, \quad (\omega_1, \omega_2) \in [0, 1] \times [0, 1]. \tag{2}$$

By again taking note of symmetries, the area of the domain of the bispectrum can be reduced by a factor of $1/12$ to a principal domain of $\mathscr{D} = \{(\omega_1, \omega_2) : 0 \le \omega_2 \le \omega_1 \le 0.5, \omega_1 \le (1 - \omega_2)/2\}$. Finally, the normalized bispectrum is the function $Z(\omega_1, \omega_2)$ defined by

$$Z(\omega_1, \omega_2) = \frac{|I(\omega_1, \omega_2)|^2}{I(\omega_1) I(\omega_2) I(\omega_1 + \omega_2)}, \quad (\omega_1, \omega_2) \in \mathscr{D}. \tag{3}$$

For purposes of clarification, we note that the principal domain of $I(\omega_1, \omega_2)$ is not quite so simple if $\{X(t)\}$ is a discrete-time realization with sampling rate δ^{-1} (where $\delta^{-1} = 1$) of a continuous-time process that is third-order stationary with band limit $\omega_0 \le (2\delta)^{-1}$. In this case, the continuous-time support set of $I(\omega_1, \omega_2)$ is the isosceles right-triangle $\{(\omega_1, \omega_2) : 0 \le \omega_1 \le \omega_0, \omega_2 \le \omega_1, \omega_1 + \omega_2 = \omega_0\}$. However, the principal domain of the discrete-time bispectrum is a larger area consisting of two adjoining triangles: the isoceles triangle $\{(\omega_1, \omega_2) : \omega_2 \le \omega_1, 0 \le \omega_1 + \omega_2 \le (2\delta)^{-1}\}$ and the odd triangle $\{(\omega_1, \omega_2) : \omega_2 \le \omega_2, (2\delta)^{-1} \le \omega_1 + \omega_2 \le \delta^{-1} - \omega_1\}$. For details, see Hinich and Wolinsky (1988) or Hinich and Messer (1995).

It is well known that if $\{X_t\}$ is linear, then the spectral density function in (1) reduces to

$$I(\omega) = \sigma_\varepsilon^2 \,|\, H(\omega)|^2,$$

where $H(\omega) = \sum_{j=0}^{\infty} \beta_j e^{-2\pi i j \omega}$ is the transfer function (Priestley, 1981). Furthermore, it is not difficult to prove that under linearity, the bispectral density function in (2) reduces to

$$I(\omega_1, \omega_2) = \mu_3 H(\omega_1) H(\omega_2) H^*(\omega_1 + \omega_2),$$

where μ_3 is the third moment of the series ε, and $*$ denotes complex conjugate. Combining these two results, if the process $\{X_t\}$ is linear, then the normalized bispectral density function defined in (3) is simply

$$Z(\omega_1, \omega_2) = \frac{\mu_3^2}{\sigma_\varepsilon^6} \quad \text{for all } (\omega_1, \omega_2) \in \mathcal{D}. \tag{4}$$

This fundamental property is the basis of frequency domain tests for the Gaussianity and linearity of a time series.

Inspection of (4) also yields a quick explanation for the two-stage nature of frequency domain tests. Note that if the series X is linear, and the distribution of the innovations ε is symmetric, then $\mu_3 = 0$, and $Z(\omega_1, \omega_2) \equiv 0$ for all frequency pairs. It is often the case that the symmetric distribution of the ε is interpreted as the innovations are Gaussian, although that is unnecessarily restrictive. On the other hand, if the series X is linear and the distribution of the noise is not symmetric, then the normalized bispectrum is a non-zero constant equal to $\mu_3^2/\sigma_\varepsilon^6$ for all frequency pairs.

3. Estimating Spectral Quantities

Let $X_t,\ t = 1, \ldots, n$ represent a realization of a zero-mean sixth-order stationary process $\{X_t\}$. The estimators of the autocovariance and third-moment function functions are

$$\hat{\gamma}(v) = \frac{1}{n} \sum_{t=1}^{n-|v|} X_t X_{t+v} \quad \text{and} \quad \hat{\gamma}(u, v) = \frac{1}{n} \sum_{t=1}^{n-s} X_t X_{t+u} X_{t+v},$$

where $s = \max\{0, u, v\}$.

To estimate $I(\omega)$ and $I(\omega_1, \omega_2)$, the sets of (continuous) frequencies on which the two functions are defined are discretized to the set of natural frequencies $\omega_j = (j-1)/n$. For estimating $I(\omega)$, $j = 1, \ldots, [n/2]+1$, where $[r]$ is the greatest integer value of r. For estimating $I(\omega_1, \omega_2)$, the set of frequency pairs (ω_j, ω_k) is defined in the triangular grid, $D = \{(\omega_j, \omega_k) : 1 \le k \le j \le [n/2]+1, 2j + k - 3 \le n\}$. Crude estimates of the spectral quantities are then the discrete Fourier transforms.

Let $\lambda(\tau)$ be a one-dimensional lag window satisfying $\lambda(0) = 1$ and the symmetry condition $\lambda(\tau) = \lambda(-\tau)$; let $\lambda(\tau_1, \tau_2)$ be a two-dimensional lag window that satisfies

$$\lambda(\tau_1, \tau_2) = \lambda(\tau_2, \tau_1) = \lambda(-\tau_1, \tau_2 - \tau_1) = \lambda(\tau_1 - \tau_2, -\tau_2) = \lambda(\tau_1)\lambda(\tau_2)\lambda(\tau_1 - \tau_2).$$

For some truncation point M, estimators of the spectral quantities are defined by

$$\widehat{I}(\omega_j) = \sum_{v=-M}^{M} \lambda\left(\frac{v}{M}\right) \hat{\gamma}(v) e^{-2\pi i v \omega_j}, \quad \omega_j \in [0, 0.5]$$

and

$$\widehat{I}(\omega_j, \omega_k) = \sum_{u=-M}^{M} \sum_{v=-M}^{M} \lambda\left(\frac{u}{M}, \frac{v}{M}\right)\hat{\gamma}(u, v)e^{-2\pi i(u\omega_j + v\omega_k)}, \quad (\omega_j, \omega_k) \in D. \tag{5}$$

It is well known that, under very general conditions, $\widehat{I}(\omega_j)$ is asymptotically normal (see, for example, Priestley, 1981).

From Van Ness (1966) and Brillinger and Rosenblatt (1967), we have that

$$\mathrm{E}[\widehat{I}(\omega_j, \omega_k)] = I(\omega_j, \omega_k) + O(n^{-1}).$$

Thus

$$Z_2^2(j,k) = 2|\widehat{Z}(\omega_j, \omega_k)|^2 = \frac{|\widehat{I}(\omega_j, \omega_k)|^2}{\widehat{I}(\omega_j)\widehat{I}(\omega_k)\widehat{I}(\omega_j + \omega_k)} \tag{6}$$

converges in distribution to a non central $\chi_2^2(\eta)$, where the non centrality parameter

$$\eta = 2KI(\omega_j, \omega_k),$$

where the constant of proportionality K depends upon M and

$$\lambda_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda^2(u, v)\, du\, dv.$$

See, for example, Hinich (1982), Brockett et al. (1988), or Hinich and Wolinsky (1988).

4. Goodness-of-Fit Test for Gaussianity and Linearity

Generally speaking, goodness-of-fit (GOF) tests are used to determine the likelihood that a random sample was selected from a specified distribution function $F(x, \boldsymbol{\theta})$. The parameter vector $\boldsymbol{\theta}$ may be completely specified or may contain unknown parameters. Traditional GOF tests are based on popular forms of $F(x; \boldsymbol{\theta})$, such as, the uniform, normal, exponential, or gamma distributions.

Let $F_n(x)$ denote the empirical distribution function (EDF). Then the general EDF test statistic of quadratic class, defined as

$$T = n\int_{-\infty}^{\infty} [F_n(x) - F(x; \boldsymbol{\theta})]^2 \psi(x)\, dF(x; \boldsymbol{\theta}) = \|F_n(x) - F(x; \boldsymbol{\theta})\|,$$

is a measure of agreement between the EDF and the underlying distribution function $F(x; \boldsymbol{\theta})$. Well-known forms of the EDF test statistic include the Cramér–von Mises (CVM) statistic ($\psi(u) = 1$) and the Anderson–Darling (AD) statistic ($\psi(u) = F(u; \boldsymbol{\theta})[1 - F(u; \boldsymbol{\theta})]^{-1}$). For the aforementioned popular null distributions, a number of methods have been developed for computing the CVM, AD, and other EDF statistics. (See, for example, D'Agostino and Stephens, 1986.)

To apply EDF GOF statistics to testing time series Gaussianity and linearity, we proceed in two stages. In the first stage, the null hypothesis is that the time series is Gaussian. Under Gaussianity, the non centrality parameter η of the distribution of

$Z_2^2(j,k)$ is identically zero for all j, k, and so the distribution reduces to exponential with mean 2. Therefore, applying the GOF approach to test Gaussianity of the time series is equivalent to testing that the sampling distribution of $Z_2^2(j,k)$ is exponential(2). In contrast, Subba Rao and Gabr (1980) and Hinich (1982) relied on $\sum_{(j,k)\in D} Z_2^2(j,k)$, which is approximately χ^2_{2P} under Gaussianity (P representing the number of frequency pairs in $\mathcal{D}$). Subba Rao and Gabr use the sample estimate of the variance–covariance matrix of $\widehat{I}$, while Hinich uses the asymptotic variance–covariance matrix. For all three procedures, if the null hypothesis is rejected, then the Gaussian null is rejected. If not, then the process may be non Gaussian, but the bispectral estimates are consistent with a zero bispectrum.

If Gaussian null is rejected, the test proceeds to the second stage, which has a null hypothesis stating that the time series is linear (although not Gaussian). Under this hypothesis, the constancy of

$$Z(\omega_1,\omega_2) = \frac{\mu_3^2}{\sigma_\varepsilon^6} \quad \text{for all } (\omega_1,\omega_2) \in \mathcal{D}.$$

can used to develop a test for linearity. The approach taken by Subba Rao and Gabr is that of an F test, which is not robust to outliers. As previously mentioned, Hinich proposed using a standardized difference of quantiles of the $Z_2^2(j,k)$. We propose a GOF EDF test. However, having as the non central $\chi^2_2(\eta)$ as the null distribution results in an problem for which finding the null distribution of the EDF statistic is intractable.

Faced with this situation, we apply a robust transformation from non central $\chi^2_\nu(\eta)$ to standard normal. The primary advantage of this is to have available the known critical values for both Anderson–Darling and Cramér-von Mises test statistic under a standard normality. Abdel-Aty (1954) provided an approximation for non central $\chi^2_\nu(\eta)$ to standard normal by equating the skewness and kurtosis of the transformed normal variable to that of the non central $\chi^2_\nu(\eta)$ distribution. Specifically, if $X \sim \chi^2_\nu(\eta)$, then

$$Y = \left(\frac{X}{\nu+\eta}\right)^h \tag{7}$$

with $h = 1/3$ is approximately standard normal. In 1959, Sankaran suggested taking h in (7) to be

$$h = 1 - \frac{2r(\nu+3\eta)}{3s^2}, \tag{8}$$

where $r = \nu+\eta$ and $s = \nu+2\eta$, proving that Y has an approximate normal distribution with mean and variance given, respectively, by

$$\mu_Y = 1 + h(h-1)\frac{s}{r^2} - h(h-1)(2-h)(1-3h)\frac{s^2}{2r^4}, \tag{9}$$

$$\sigma_Y^2 = h^2\frac{2s}{r^2}\left[1-(1-h)(1-3h)\frac{s}{r^2}\right]. \tag{10}$$

Johnson et al. (1995) provided a numerical comparison of upper and lower 5% points of a good number of normal approximations of the non central χ^2, including the Abdel–Aty with $h = 1/3$ and the approximation by Sankaran (1959). Their numerical comparison indicates Sankaran's approximation is most accurate in the tails. Empirical studies have shown that $(Y - \mu_Y)/\sigma_Y$ converges in distribution to $N(0, 1)$ as long as $\hat{\eta}$ is a consistent estimator of η (Jahan 2006, 2006). Applying this transformation to the $Z_2^2(j, k)$ yields that testing the linearity of the (non Gaussian) time series is equivalent to testing the null that the transformed Z_2^2 have a standard normal distribution. Under the null hypothesis, a consistent estimator of the non centrality parameter η for use in computing (8), and hence (9) and (10) is

$$\hat{\eta} = \begin{cases} \bar{x} - v, & \text{if } \bar{x} > v, \\ 0, & \text{if } \bar{x} \le v. \end{cases} \tag{11}$$

If we let $Q_{(i)}$ denote the quantiles computed from the order statistics of the $Z_2^2(j, k)$, then the Cramér–von Mises statistic may be computed using

$$W^2 = \frac{1}{12P} + \sum_{i=1}^{P} \left[Q_{(i)} - \frac{(2i-1)}{2P} \right]^2, \tag{12}$$

and the Anderson–Darling test statistic can be computed using

$$A^2 = -P - \frac{1}{P} \sum_{i=1}^{P} (2i-1)[\log Q_{(i)} + \log(1 - Q_{(P+1-i)})].$$

Finally, since both mean and variance of the transformed random variable are unknown, the modified Anderson–Darling test statistic

$$A_*^2 = A^2 \left(1 + \frac{3}{4P} + \frac{9}{4P^2} \right), \tag{13}$$

should be compared to the to the appropriate upper-tail critical value.

To summarize, the bispectral-based goodness-of-fit testing procedure is as follows.

Stage 1. Testing Gaussianity.

1. Compute the estimator of twice the square modulus of the normalized bispectrum Z_2^2, as defined in Eq. (6).
2. Apply GOF test of exponential(2) to Z_2^2. This is equivalent to testing a null hypothesis of a symmetric error distribution, which is usually interpretted as time series Gaussianity.
3. If the null is not rejected, then stop. Otherwise, proceed to the second stage of testing.

Stage 2. Testing linearity of a series with non Gaussian errors.

1. Apply the transformation described in Eqs. (7)–(10) to Z_2^2 using $\hat{\eta}$ as defined in Eq. (11) as an estimate of the non centrality parameter
2. Compute GOF statistic for normality; the Cramér–von Mises statistic is given by Eq. (12) and the Anderson–Darling statistic by (13).
3. Compare GOF statistic to appropriate upper-tail critical value.

5. Comparison to Existing Tests

In application, estimation of $Z(\omega_1, \omega_2)$ is accomplished by taking the lag window $\lambda(u, v)$ to be rectangular with width $M = n^c$, $1/2 \leq c \leq 1$. The value of c controls the trade-off between bias and variance. This closely follows the procedure in Hinich (1982). Specific expressions for the asymptotic variance–covariance matrix are found within. Ashley et al. (1986) presented a simulation study that illustrates optimal values of c, with values closer to $c^{1/2}$ performing better for Hinich's test. Accordingly, we take $c = 5/8$. Construct a lattice

$$\mathscr{L} = \left\{ \left(\frac{(2\imath - 1)M}{2n}, \frac{(2\kappa - 1)M}{2n} \right) : \imath = 1, \ldots, \kappa \text{ and } \kappa \leq n/(2M) - [\imath/2] + 0.75 \right\} \in D.$$

Then for each frequency pair $((2\imath - 1)M/(2n), (2\kappa - 1)M/(2n)) \in \mathscr{L}$, the estimated bispectral density function $\widehat{I}(\cdot, \cdot)$ is simply an average of the bivariate Fourier transforms at the natural frequencies in a square of M^2 points centered at $((2\imath - 1)M/(2n), (2\kappa - 1)M/(2n))$, if all the M^2 points are in $\mathscr{D}$. If a square has points outside the set $\mathscr{D}$, then those points are not included in the average. Figure 1 illustrates the lattice constructed for $c = 5/8$ for a time series of length $n = 500$.

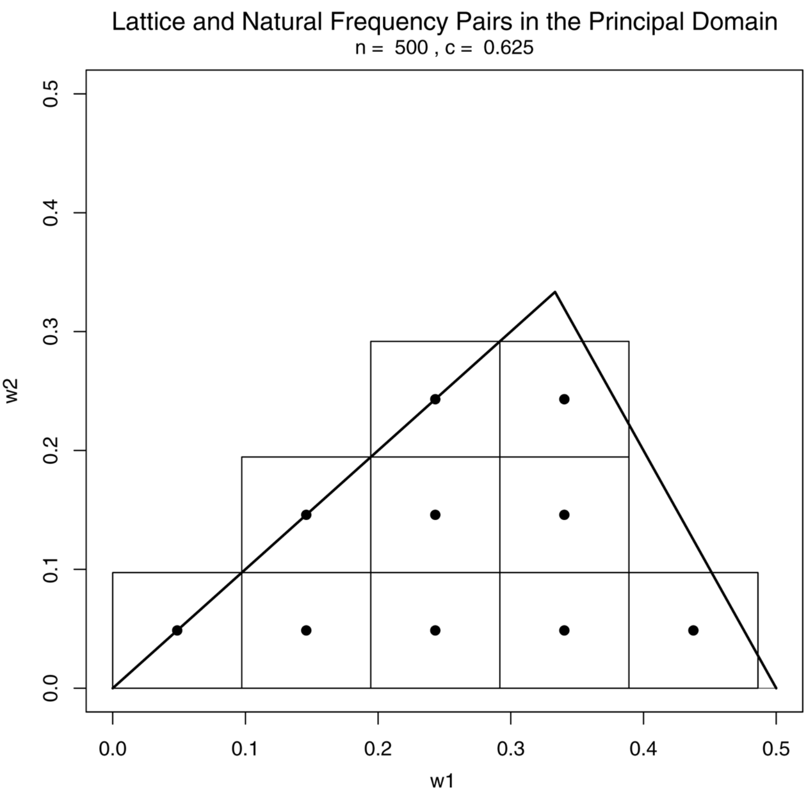

Figure 1. Lattice $\mathscr{L}$ for estimating normalized bispectrum for $c = 5/8$ and $n = 500$.

At each frequency point that is fully contained within a square, the number of bivariate Fourier transforms used to estimate the bispectrum is indeed large. Recall that in the second stage of the test, the test statistic proposed by Hinich (1982) is the standardized difference of quantiles, with p-values computed using the normal approximation for the sampling distribution of the standardized difference. Figure 1 clearly illustrates that the number of Z_2^2 used in computing the difference of quantiles is only 10. Thus, the normal approximation for distribution of the standardized difference of quantiles will be poor.

A numerical investigation was conducted to compare the performance of the spectral domain tests and some popular time domain tests. The time domain tests were the Tukey nonadditivity test based on an added variable approach (Keenan, 1985), original F test (Tsay, 1986), CUSUM test (Petruccelli and Davies, 1986), TAR F test (Tsay, 1989), and the new F test (Tsay, 1991). Most of the time domain tests are constructed under the premise that the alternative is a specific nonlinear model. In particular, the tests by Keenan (1985) and Tsay (1986) specify a bilinear model in the alternative. Tsay's (1986) test is considered an improvement over Keenan's test in that Tsay retains the simplicity of Keenan's approach, but increases the power by examining the residuals of regressions that include the individual nonlinear and quadratic terms up to third order, while Keenan considers residuals of regressions on only the second-order terms. The CUSUM and Tsay's TAR F test specify a threshold model in the alternative. The CUSUM test is based on cumulative sums of standardized one-step-ahead forecast errors from arranged autoregressive fits to the data. Tsay's TAR F test takes the process farther, by performing a second autoregression using those residuals, and constructs an F test based on the sums of squares of those two regressions. Finally, Tsay's New F test is constructed in such a way as to include any of the bilinear, smoothed threshold, and exponential autoregressive models in the alternative. For most of the time domain tests, it is necessary to specify some set of testing parameters. In the simulation studies, the parameters selected were optimal, based on the true model.

The three spectral domain tests included were Hinich's (1982) bispectral based test, the AD GOF test, and the CVM GOF test. The Subba Rao and Gabr (1980) test was not included since Hinich's test performs better, in general. All results are based on 1,000 replications for time series lengths of $n = 100$ or 500, and level of significance $\alpha = 0.05$.

Because the distribution of the test statistics for all spectral-domain tests are based on approximations, an empirical investigation of the size of the tests is warranted. Three of the models considered in that investigation are listed immediately below. For all models, the series ε_t is white noise.

I. $X_t = \varepsilon_t$ (White noise)
II. $X_t - 0.4X_{t-1} + 0.3X_{t-2} = \varepsilon_t$ (Autoregressive)
III. $X_t = -0.4\varepsilon_{t-1} + 0.3\varepsilon_{t-2} + \varepsilon_t$ (Moving average)

Table 1 contains a summary of the results. The empirical sizes of all three tests are slightly higher than the nominal $\alpha = 0.05$. However, this simulation suggests that the size of Hinich's test is inflated more than AD or CVM, with CVM performing better, in general.

The majority of the nonlinear models considered in the study were selected because they were found in the time domain testing literature. Specifically, Model IV is found in Tsay (1986), and Models V–VIII in Tsay (1991). Model IX was added at the suggestion of a referee.

Table 1
Empirical sizes for Hinich, AD, and CVM bispectral-based tests of Gaussianity

Model	n	Hinich	AD	CVM
I	100	0.095	0.076	0.068
	500	0.075	0.064	0.060
II	100	0.083	0.069	0.065
	500	0.071	0.066	0.063
III	100	0.084	0.071	0.064
	500	0.057	0.065	0.062

IV. $X_t = 0.5\varepsilon_{t-1} - 0.6\varepsilon_{t-1}^2 + \varepsilon_t$ (Nonlinear moving average)
V. $X_t = 0.4X_{t-1} - 0.3X_{t-2} + 0.8\varepsilon_{t-1} + 0.5X_{t-1}\varepsilon_{t-1} + \varepsilon_t$ (Bilinear)
VI. (Threshold autoregressive)

$$X_t = \begin{cases} 1 - 0.5X_{t-1} + \varepsilon_t & \text{if } X_{t-1} < 0, \\ -1 - 0.5X_{t-1} + \varepsilon_t & \text{if } X_{t-1} \geq 0. \end{cases}$$

VII. $X_t = 1 - 0.5X_{t-1} - 0.5F(X_{t-1}) + \varepsilon_t$, where $F(u) = [1 + \exp(-u/2)]$ (Smoothed threshold autoregressive)
VIII. $X_t = (0.3 + 100e^{X_{t-1}^2})X_{t-1} + \varepsilon_t$ (Exponential autoregressive)
IX. $X_t = -0.25X_{t-1} + 0.2X_{t-2} + 0.15X_{t-1}^2 - 0.1X_{t-2}^2 + \varepsilon_t$ (Nonlinear autoregressive)

Table 2 summarizes the results. A comparison of the three frequency domain tests clearly illustrates that for smaller values of n, both GOF approaches outperform Hinich's difference of quantile approach. For larger values of n, the three methods are comparable, and in two cases (both with $n = 500$), Hinich's test outperforms the GOF tests.

Table 2
Empirical powers for time and frequency domain tests

Model	n	Keenan	Ori-F	CUSUM	Tar-F	New-F	Hinich	AD	CVM
IV	100	0.319	0.323	0.557	0.680	0.887	0.089	0.350	0.309
	500	0.644	0.645	0.957	0.977	1.000	0.910	0.835	0.783
V	100	0.650	0.740	0.890	0.670	0.810	0.097	0.638	0.573
	500	0.952	0.990	1.000	0.999	1.000	0.751	1.000	0.999
VI	100	0.183	0.186	0.345	0.228	0.700	0.010	0.954	0.735
	500	0.195	0.195	0.945	0.649	1.000	0.074	0.999	0.996
VII	100	0.091	0.098	0.090	0.102	0.114	0.037	0.089	0.085
	500	0.155	0.155	0.308	0.281	0.490	0.078	0.064	0.051
VIII	100	0.047	0.047	1.000	0.096	0.884	0.066	0.232	0.201
	500	0.090	0.090	1.000	0.039	1.000	0.216	0.170	0.156
IX	100	0.454	0.532	0.284	0.400	0.302	0.383	0.953	0.904
	500	0.757	1.000	0.936	0.986	0.967	1.000	1.000	1.000

Other comparisons of interest are the between the frequency and time domain tests. While it is certainly the case that different time domain test outperform all the spectral domain tests for some models, it is not always the case. In fact, for those models considered here, the power of the GOF is comparable to the time domain tests. For example, for the threshold model (VI) and for the nonlinear AR model (IX), the AD and CVM spectral domain tests outperform all time domain tests, including the two time domain tests with a test statistic constructed to detect threshold nonlinearity. On the other hand, no test performs well for the smoothed threshold model. In general, the GOF spectral domain tests outperform Keenan's test and usually outperforms the original F test. The power of these spectral-domain tests rivals that of the CUSUM, TAR F, and New F tests, but without the restriction of a parametric alternative.

6. Conclusion

Spectral domain tests have largely been considered inferior to their time domain counterparts because the spectral domain tests suffer from lack of power. In this article, we presented two bispectral-based test for the Gaussianity and linearity that out perform the popular spectral domain test proposed by Hinich (1982). Moreover, in many cases, the power of the tests we propose rivals that of existing parametric time domain tests. While these new approaches can not be recommended in every circumstance, neither can any of the more popular time domain approaches. The additional information that can be gleaned from plotting the bispectrum may also prove to be useful in explaining the interaction of frequency components due found in nonlinear series.

References

Abdel-Aty, S. H. (1954). Approximate formulae for the percentage points and probability integral of the non central χ^2 distribution. *Biometrika* 41:538–540.

Ashley, R. A., Patterson, D. M., Hinich, M. (1986). A diagnostic test for nonlinear serial dependence in time series fitting errors. *J. Time Ser. Anal.* 7:165–178.

Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., Jensen, M. J. (1997). A single-blind controlled competition among tests for nonlinearity and chaos. *J. Econometrics* 82:157–192.

Brillinger, D. R., Rosenblatt, M. (1967). Asymptotic theory of estimates of kth order spectra. In: Harris, B., ed. *Spectral Analysis of Time Series*. New York: Wiley, pp. 153–188.

Brockett, P. L., Hinich, M. J., Patterson, D. (1988). Bispectral-based tests for the detection of Gaussianity and linearity in time series. *J. Amer. Statist. Assoc.* 83:657–664.

Chan, K., Tong, H. (1986). A note on certain integral equations associated with nonlinear time series analysis. *Probab. Theor. Related Fields* 73:153–159.

D'Agostino, R. B., Stephens, M. A. (1986). *Goodness-of-Fit Techniques*. New York: John Wiley & Sons.

Harvill, J. L. (1999). Testing time series linearity via goodness of fit methods. *J. Statist. Plann. Infer.* 75:331–341.

Harvill, J. L., Newton, H. J. (1995). Saddlepoint approximations for the difference of order statistics. *Biometrika* 82:226–231.

Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* 3:169–176.

Hinich, M. J. (2005). Normalizing bispectra. *J. Statist. Plann. Infer.* 130:405–411.

Hinich, M. J., Wolinsky, M. A. (1988). A test for aliasing using bispectral analysis. *J. Amer. Statist. Assoc.* 83:499–502.

Hinich, M. J., Messer, H. (1995). On the principal domain of the discrete bispectrum of a stationary signal. *IEEE Trans. Signal Process.* 43:2130–2134.

Hinich, M. J., Rothman, P. (1998). Frequency-domain test of time reversibility. *Macroeconomic Dyn.* 2:72–88.

Jahan, N. (2006). Applying Goodness-of-Fit Techniques in Testing Time Series Gaussianity and Linearity. Unpublished doctoral dissertation, Mississippi State University, Mississippi State, MS.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). *Continuous Univariate Distributions*. New York: John Wiley & Sons.

Keenan, D. M. (1985). A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika* 72:39–44.

Petruccelli, J. D., Davies, N. (1986). A Portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika* 73:687–694.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. London: Academic Press.

Rosenblatt, M., Van Ness, J. W. (1965). Estimation of the bispectrum. *Ann. Mathemat. Statist.* 36:1120–1136.

Sankaran, M. (1959). On the non central chi-square distribution. *Biometrika* 46:235–237.

Subba Rao, T., Gabr, M. M. (1980). A test for linearity of stationary time series. *J. Time Ser. Anal.* 1:145–158.

Tong, H. (1990). *Non-linear Time Series Analysis: A Dynamical System Approach*. Oxford: Oxford University Press.

Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika* 73:461–466.

Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes. *J. Amer. Statist. Assoc.* 84:231–240.

Tsay, R. S. (1991). Detecting and modeling nonlinearity in univariate time series analysis. *Statistica Sinica* 1:431–451.

Van Ness, J. W. (1966). Asymptotic normality of bispectral estimates. *Ann. Mathemat. Statist.* 37:1257–1275.