

A Bootstrap Test for Time Series Linearity[☆]

Arthur Berg^{*,a}, Efstathios Paparoditis^b, Dimitris N. Politis^{*,c}

^a*Division of Biostatistics, Pennsylvania State University, Hershey, PA 17033, USA*
^b*Department of Mathematics and Statistics, University of Cyprus, P.O.Box 20537, CY
1678 Nicosia, CYPRUS*

^c*Department of Mathematics, University of California, San Diego, La Jolla, CA
92093-0112, USA*

Abstract

A bootstrap algorithm is proposed for testing Gaussianity and linearity in stationary time series, and consistency of the relevant bootstrap approximations is proven rigorously for the first time. Subba Rao and Gabr (1980) and Hinich (1982) have formulated some well-known nonparametric tests for Gaussianity and linearity based on the asymptotic distribution of the normalized bispectrum. The proposed bootstrap procedure gives an alternative way to approximate the finite-sample null distribution of such test statistics. We revisit a modified form of Hinich's test utilizing kernel smoothing, and compare its performance to the bootstrap test on several simulated and two real data sets—the S&P 500 returns and the quarterly US real GNP growth rate. Interestingly, Hinich's test and the proposed bootstrapped version yield substantially different results when testing Gaussianity and linearity of the GNP data.

Key words: Bispectrum, Bootstrap, Gaussianity test, Linearity test
2000 MSC: 62M10, 91B84

1. Introduction

Every discrete covariance-stationary time series, $\mathbf{X} = \{X_t, t \in \mathbb{Z}\}$, with a purely continuous spectral density can be represented as the mean square

[☆]Parzen Special Issue

*Corresponding authors.

Email addresses: `berg@psu.edu` (Arthur Berg), `dpolitis@ucsd.edu` (Dimitris N. Politis)

convergent series

$$X_t = a + \sum_{j=-\infty}^{\infty} a_j \varepsilon_{t-j} \quad (1)$$

where $\{\varepsilon_t\}$ is a white noise process (Priestley, 1983, Chapter 10). If the errors ε_t in representation (1) are independent and identically distributed (iid), then \mathbf{X} is termed a *linear* time series; otherwise, if such a linear representation with respect to iid errors does not exist, \mathbf{X} is termed *nonlinear*. For example, an AR(p) time series is linear only if the innovations driving the autoregressive (AR) model are iid. A general overview of nonlinear time series is given in Tjøstheim (1994) as well as in the books Tong (1993) and Fan and Yao (2003).

The classification of time series as linear versus nonlinear or Gaussian versus non-Gaussian is of great value in the modeling and analysis of time series. In practice, a Gaussian classification would typically indicate an ARMA model with Gaussian innovations could be used, and a linear classification would indicate that an ARMA model with independent but possibly non-Gaussian innovations can be considered. However, the rejection of linearity typically requires the practitioner to carefully select an appropriate nonlinear model characteristic of the time series at hand.

Some parametric and semiparametric tests of linearity designed for a specific nonlinear model as an alternative hypothesis have been constructed, e.g., An et al. (2000); Ashley and Patterson (2009); Chan and Tong (1986, 1990); Chan (1990); Hansen (2000); Harvey and Leybourne (2007); Keenan (1985); Luukkonen et al. (1988); Petrucci (1986, 1990); Terasvirta et al. (1993); Terasvirta (1994); Tsay (1986).

In this article we focus specifically on bootstrapping nonparametric tests of Gaussianity and linearity based on the normalized bispectrum, cf. Birkelund and Hanssen (2009); Brockett et al. (1988); Hinich (1982); Jahan and Harvill (2008); Subba Rao and Gabr (1980, 1984); Yuan (2000). Tests based on the normalized bispectrum are frequently used in practice, especially when analyzing financial time series, e.g., Hinich and Patterson (1985, 1989); Hsieh (1989); Abhyankar et al. (1995, 1997). Note that there exist tests of linearity that are not based on the normalized bispectrum; see e.g., Ashley et al. (1986); Hong-Zhi and Bing (1991); Terdik and Math (1998); Theiler et al. (1992). A nice overview of some of these tests is provided in Corduas (1994).

It should be noted that a different definition of time series linearity has also been considered in the literature (Hannan (1973, 1986)) where a time

series is defined to be linear if the best (in terms of minimizing the mean squared prediction error) one-step ahead predictor is linear; for more details see Kokoszka and Politis (2008) or Cryer and Chan (2008), p. 383. This alternative definition of linearity, however, does not lend itself to handling the problem of testing considered here.

There are many published reports, especially in recent years, that utilize the bootstrap in linearity testing, e.g., Barnett and Wolff (2005), Birkelund and Hanssen (2009), Hinich et al. (2005), Hjellvik and Tjostheim (1995), Kugiumtzis (2008). Our approach is fundamentally different from these previous works although some similarities are present. The approach we take is based on the $AR(\infty)$ –“*sieve*” bootstrap where residuals from an $AR(p)$ fit are bootstrapped while p increases with the sample size (Kreiss, 1988, 1992; Paparoditis and Streitberg, 1992; Bühlmann, 1997; Lahiri, 2003). From the bootstrapped residuals, new ‘pseudo’ time series are constructed, and used to approximate the null distribution of test statistics based on the normalized bispectrum. Under general assumptions on the time series and test statistic (labelled as Assumptions 1, 2, and 3) we prove consistency of our bootstrap test in our Theorem 3.1.

Alternative uses of the bootstrap in the literature of linearity and Gaussianity testing include the surrogate data method (Theiler et al., 1992), a phase scrambling bootstrap for variance estimation (Barnett and Wolff, 2005), bootstrapped residuals to obtain the correct false alarm rate (Hinich et al., 2005; Birkelund and Hanssen, 2009), etc.; several surrogate and bootstrap tests for linearity in time series were compared in Kugiumtzis (2008). The specific use of the $AR(\infty)$ bootstrap—including a proof of consistency for the first time—distinguishes our proposal from the aforementioned approaches.

To offer finite-sample validation to our asymptotic results, we compared a slightly modified version of Hinich’s linearity test (Hinich, 1982) with a bootstrapped version of that test over several different simulated time series processes and two real data sets—quarterly US GNP growth rate and daily S&P 500 returns. Hinich’s test is based on the asymptotic distribution of the test statistic, and therefore one might expect the bootstrap to provide a better approximation to the finite-sample distribution of the test statistic; this was indeed observed in the simulations.

The outline of the rest of the paper is as follows. In Section 2, we describe the normalized bispectrum and indicate how it motivates tests of Gaussianity and linearity. We also describe a ‘kernelized’ form of Hinich’s test that allows more flexibility—and accuracy when considering higher-order kernels—in im-

plementing tests of Gaussianity and linearity. In Section 3, the steps of the bootstrap algorithm are described and a consistency theorem is established whose proof is provided in the Appendix. A simulation study is provided in Section 4 where Hinich's test is compared with the proposed bootstrap test on eight different time series models, and a further comparison is made in Section 5 on real data sets.

2. Gaussianity and linearity testing with the normalized bispectrum

2.1. The normalized bispectrum

Let X_1, \dots, X_n be an observed stretch from a zero mean (or centered), third-order stationary time series $\mathbf{X} = \{X_t, t \in \mathbb{Z}\}$ with autocovariance function $\gamma(\tau) = E(X_t X_{t+\tau})$ and third-order (auto-)cumulant function $\gamma(\tau_1, \tau_2) = E(X_t X_{t+\tau_1} X_{t+\tau_2})$. Let $f(\lambda)$ and $f(\lambda_1, \lambda_2)$ denote the spectral and bispectral density of \mathbf{X} , respectively, given by

$$f(\lambda) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma(\tau) e^{-i\lambda\tau}$$

$$f(\lambda_1, \lambda_2) = \frac{1}{(2\pi)^2} \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \gamma(\tau_1, \tau_2) e^{-i(\lambda_1\tau_1 + \lambda_2\tau_2)}.$$

As was defined in the Introduction, a time series \mathbf{X} is defined to be linear if it admits the representation (1) with respect to an iid sequence $\{\varepsilon_t\}$. Under linearity, it follows that the bispectrum of \mathbf{X} becomes (Zurbenko, 1986; Brillinger, 1965)

$$f(\lambda_1, \lambda_2) = \frac{\mu_3}{(2\pi)^2} H(-\lambda_1 - \lambda_2) H(\lambda_1) H(\lambda_2)$$

where μ_3 is the third moment of \mathbf{X} and $H(\lambda)$ is the transfer function of \mathbf{X} given by

$$H(\lambda) = \sum_{j=-\infty}^{\infty} a_j e^{-ij\lambda}.$$

The transfer function is related to the spectral density by the identity (Brockwell and Davis, 1991) $f(\lambda) = \frac{\sigma^2}{2\pi} |H(\lambda)|^2$. It is then natural to define the normalized bispectrum to be

$$K(\lambda_1, \lambda_2) \triangleq \frac{|f(\lambda_1, \lambda_2)|^2}{f(\lambda_1)f(\lambda_2)f(\lambda_1 + \lambda_2)} \stackrel{\text{linearity}}{=} \frac{(\mu_3)^2}{2\pi\sigma^6}. \quad (2)$$

Therefore if a time series is linear, its normalized bispectrum is constant over all (λ_1, λ_2) . Furthermore, if the process is Gaussian, then, in particular, it is linear (Priestley, 1983, p.867) in which case the normalized bispectrum is identically zero over all (λ_1, λ_2) . Therefore testing for Gaussianity or linearity amounts to testing the normalized bispectrum equalling zero or a non-zero constant. Such linearity tests may lack power with nonlinear time series possessing certain bispectral symmetries, and similarly, such Gaussianity tests lack power with non-Gaussian time series with certain bispectral symmetries and a zero third moment. However, such alternatives are rather specialized and do not impose any practical limitations on the tests.

Symmetries of the third-order cumulant function given by

$$\begin{aligned} \gamma(\tau_1, \tau_2) &= \gamma(\tau_2, \tau_1) = \gamma(-\tau_1, \tau_2 - \tau_1) = \gamma(\tau_1 - \tau_2, -\tau_2) \\ &= \gamma(-\tau_2, \tau_1 - \tau_2) = \gamma(\tau_2 - \tau_1, -\tau_1) \end{aligned} \quad (3)$$

induce symmetries in the bispectrum. From the relationships given in (3) and by noticing the bispectrum is 2π -periodic in each variable, the values of the bispectrum in the entire plane can be determined from the values inside one of the twelve sectors illustrated in Figure 1 (Van Ness, 1966; Subba Rao and Gabr, 1984). Therefore we will limit estimation of the normalized bispectrum to values within just one of the sectors, the first sector, and we define Λ the set of points contained in this sector; i.e.,

$$\Lambda = \{(\lambda_1, \lambda_2) : 0 \leq \lambda_1 \leq \pi, 0 \leq \lambda_2 \leq \min\{\lambda_1, 2(\pi - \lambda_1)\}\}. \quad (4)$$

A formalized test of Gaussianity and linearity was initially proposed by Subba Rao and Gabr (1980), and shortly afterwards Hinich (1982) presented a more robust test providing large-sample approximations of its power. The approach Hinich took was to estimate the normalized bispectrum on a grid of points inside Λ by averaging the bivariate periodogram at Fourier frequencies around the points of interest. Referencing the work of Van Ness (1966), asymptotic distributions of the normalized bispectral estimators were obtained along with the asymptotic distributions of two test statistics for testing Gaussianity and linearity. With the use of our bootstrap procedure that is described in the next section, the asymptotic distributions of Hinich's test statistics can be replaced by data-driven bootstrap distributions yielding potentially better results in the finite sample case. Further details of Hinich's tests are provided in the Section 4 where simulations and comparisons are provided.

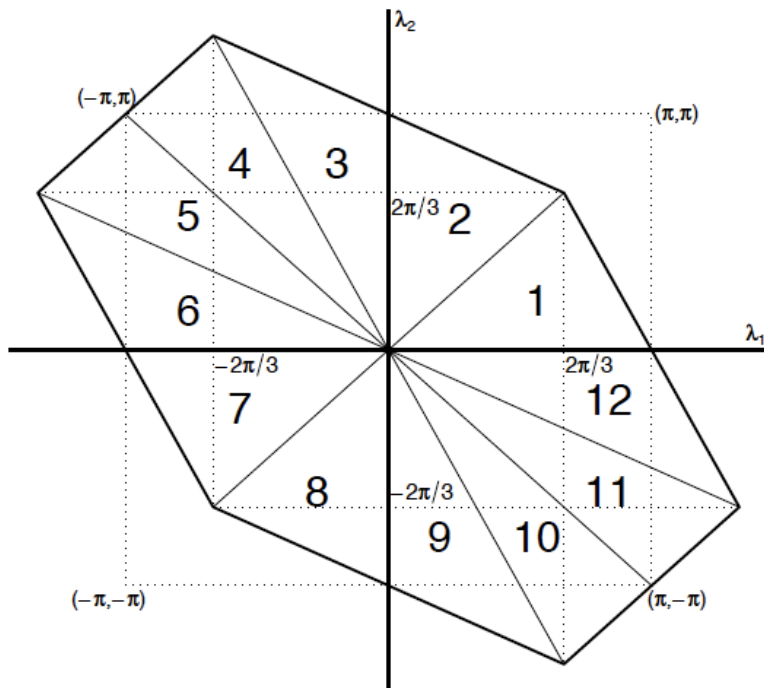


Figure 1: Through its symmetries, values of the bispectrum over the entire plane can be determined by the values in any one of the twelve labelled sectors.

2.2. Estimation of the Normalized Bispectrum

We consider Rosenblatt-Parzen type kernel-smoothed estimators of the spectral and bispectral densities. Estimators of $\gamma(\cdot)$ and $\gamma(\cdot, \cdot)$ are given as

$$\begin{aligned}\hat{\gamma}(\tau) &= \frac{1}{n} \sum_{t=1}^{n-\tau} X_t X_{t+\tau} \\ \hat{\gamma}(\tau_1, \tau_2) &= \frac{1}{n} \sum_{t=1}^{n-\beta} X_{t-\alpha} X_{t-\alpha+\tau_1} X_{t-\alpha+\tau_2}\end{aligned}\tag{5}$$

where $\alpha = \min(0, \tau_1, \tau_2)$ and $\beta = \max(0, \tau_1, \tau_2) - \alpha$. The above estimators are defined to yield zero when there are no terms in the respective sum of (5). In practice, when the data are not assumed to have mean zero, X_t would be centered to the sample mean; such data-based centering does not spoil the asymptotic results—cf. Berg and Politis (2009).

Kernel estimators of the spectrum and bispectrum are given by (Parzen,

1957; Rosenblatt and Van Ness, 1965; Grenander and Rosenblatt, 1984)

$$\begin{aligned}\widehat{f}(\lambda) &= \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} w(\tau/M) \widehat{\gamma}(\tau) e^{-i\lambda\tau} \\ \widehat{f}(\lambda_1, \lambda_2) &= \frac{1}{(2\pi)^2} \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} w(\tau_1/M, \tau_2/M) \widehat{\gamma}(\tau_1, \tau_2) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2}\end{aligned}\tag{6}$$

Assumption 1 provides the necessary conditions on the lag-window functions $w(\cdot)$ and $w(\cdot, \cdot)$ and the truncation parameter M to ensure consistency of the spectral and bispectral densities.

■ **Assumption 1**

- (i) $M = M_n \rightarrow \infty$ and $M^2/n \rightarrow 0$ as $n \rightarrow \infty$.
 - (ii) $w(\cdot)$ and $w(\cdot, \cdot)$ are bounded and continuous functions with compact support and a characteristic exponent (also referred to as kernel order) that is greater than or equal to one.
-

The characteristic exponent, or kernel order, of $w(\cdot)$ is the largest number q , allowing ∞ , such that

$$1 - w(\tau) \leq h_q |\tau|^q \quad \text{for all } \tau$$

where h_q is a constant that may depend on q , and the kernel order of $w(\cdot, \cdot)$ is the largest q such that

$$1 - w(\tau_1, \tau_2) \leq h_q (|\tau_1| + |\tau_2|)^q \quad \text{for all } \tau_1, \tau_2$$

again where h_q is a constant that may depend on q . These bound conditions can be replaced by limit conditions, given the limit exists, which allows for easier verification with specific kernels (Parzen, 1961; Rosenblatt and Van Ness, 1965).

The condition in (ii) requiring compact support can be relaxed with the tradeoff of a more involved theory. The flat-top lag window functions form a class of functions for $w(\cdot)$ and $w(\cdot, \cdot)$ with infinite characteristic exponent that possess certain mean-square error optimality properties (Politis and Romano, 1995; Berg and Politis, 2009; Politis, 2009); these functions do not have

compact support, but do satisfy sufficient criteria such as

$$\begin{aligned} \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{\tau=-n}^n w(\tau/M) &< \infty \\ \lim_{M \rightarrow \infty} \frac{1}{M^2} \sum_{\tau_1=-n}^n \sum_{\tau_2=-n}^n w(\tau_1/M, \tau_2/M) &< \infty \end{aligned}$$

In addition to the above assumptions, very often symmetry conditions are imposed on w and $w(\cdot, \cdot)$ that mimic the symmetry conditions known to hold for $\gamma(\cdot)$ and $\gamma(\cdot, \cdot)$. Specifically, the conditions are

$$\begin{aligned} w(\tau) &= w(-\tau) \\ w(\tau_1, \tau_2) &= w(\tau_2, \tau_1) = w(-\tau_1, \tau_2 - \tau_1) \end{aligned} \tag{7}$$

Imposing the conditions in (7) guarantees that the spectral and bispectral density estimators satisfy the same symmetry conditions as their theoretical counterparts, but the symmetry requirements are not required for consistency or asymptotic normality of the estimators and are therefore not included in Assumption 1. However, just like we do not wish to have variance estimates to be negative, conditions in (7) should be imposed in practice; techniques given in Berg (2008) can be used to symmetrize any function $w(\cdot)$ or $w(\cdot, \cdot)$ to satisfy (7).

2.3. Kernel form of Hinich's Gaussianity and linearity tests

In Hinich's original test, and in a multivariate extension considered by Wong (1997)—, estimators of the normalized bispectrum are computed by explicitly averaging the two-variable periodogram over rectangular regions. Instead of averaging over rectangular regions, Birkelund and Hanssen (2009) consider averaging the two-variable periodogram over hexagonal regions. These estimators of the bispectrum are analogous to estimating the spectral density with the second-order Daniell window (Daniell, 1946; Priestley, 1983). Therefore the averaged two-variable periodogram bispectrum can be viewed as a member in the class of kernel-smoothed estimators of the bispectrum. Here we provide the theory for Hinich's test in a general setting with kernel-smoothed estimators of the normalized bispectrum.

As mentioned above, a test of Gaussianity can be formed by testing if the normalized bispectrum is zero and a test of linearity can be formed by testing if the normalized bispectrum is constant. Due to the symmetries of

the normalized bispectrum as depicted in Figure 1, we can restrict estimation of the normalized bispectrum to the points in Λ defined in (4).

In the first step, k more-or-less equally spaced points are selected inside Λ . One algorithmic approach would be to select the number of rows, r , and build a triangle of points within Λ where the bottom row contains r equally-spaced points, the row above contains $r - 1$ equally-spaced points, so on and so forth, and in the end there would be $k = \binom{r+1}{2}$ somewhat equally-spaced points in the triangle. An example of this approach with 10 rows is presented in Figure 2.

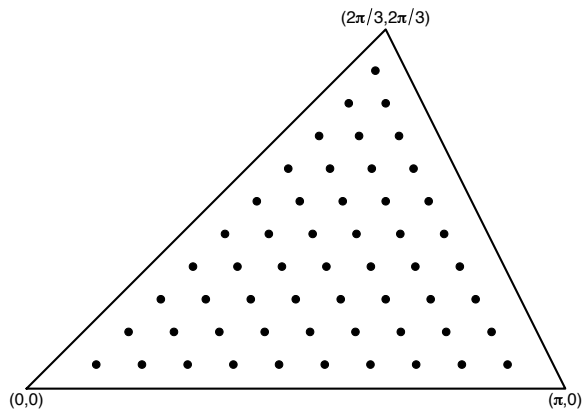


Figure 2: A grid of $\binom{11}{2} = 55$ points used for testing Gaussianity and linearity

Let $\Lambda_0 = \{\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k\} \subset \Lambda$ where $\boldsymbol{\lambda}_j = (\lambda_j^1, \lambda_j^2)$ correspond to the k points selected in Λ , such as the $k = 55$ points in Figure 2. To make the calculations simpler, we do not allow the points in Λ_0 to lie on the boundary of Λ . In general, the bispectral density, $f(\cdot, \cdot)$, is complex-valued, and we express the values of the bispectrum at the points $\boldsymbol{\lambda}_j$ in terms of its real and imaginary components as

$$f(\lambda_j^1, \lambda_j^2) = a_j + ib_j.$$

Van Ness (1966) showed the kernel-estimators, $\hat{f}(\lambda_j^1, \lambda_j^2)$ as in (6) are approximately complex Gaussian; i.e.

$$\hat{f}(\lambda_j^1, \lambda_j^2) \dot{\sim} X + iY$$

where $X \sim \mathcal{N}(a_j, \sigma_j^2)$ and independently $Y \sim \mathcal{N}(b_j, \sigma_j^2)$ with

$$\sigma_j^2 = \frac{M^2 \omega_2}{n} \frac{1}{2\pi} f(\lambda_j^1) f(\lambda_j^2) f(\lambda_j^1 + \lambda_j^2)$$

and

$$\omega_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w^2(\tau_1, \tau_2) d\tau_1 d\tau_2.$$

Now define

$$T_j = \frac{|\hat{f}(\lambda_j^1, \lambda_j^2)|^2}{\hat{\sigma}_j^2} = \frac{n}{M^2} \frac{2\pi}{\omega_2} \frac{|\hat{f}(\lambda_j^1, \lambda_j^2)|^2}{\hat{f}(\lambda_j^1)\hat{f}(\lambda_j^2)\hat{f}(\lambda_j^1 + \lambda_j^2)}$$

which has an approximate non-central chi-squared distribution

$$T_j \sim \left(\frac{X_j}{\sigma_j}\right)^2 + \left(\frac{Y_j}{\sigma_j}\right)^2 \sim \chi_2^2(\lambda_j)$$

where $\lambda_j = (a_j^2 + b_j^2)/\sigma_j^2$. Note that spectral density estimators $\hat{f}(\cdot)$ converge more rapidly than the bispectral estimators, and therefore the error in their estimation of the normalized bispectrum can be ignored.

Summing T_j over all k points in Λ_0 gives the test statistic for Gaussianity, $T_G = \sum_{j=1}^k T_j$, which approximately has the non-central chi-squared distribution $\chi_{2k}^2(\lambda)$ where $\lambda = \sum_{j=1}^k \lambda_j$. Under the assumption of Gaussianity, $a_j = b_j = 0$ for all j , and in this case T_G has a (central) chi-squared distribution with $2k$ degrees of freedom. Therefore rejection regions and p -values for the Hinich-based test of Gaussianity are computed from the χ_{2k}^2 distribution.

Taking the interquartile range of T_j gives the test statistic for linearity, $T_L = \text{IQR}(T_j)$. Under linearity of the time series, the right hand side of (2) implies

$$T_j \sim \chi_2^2(\lambda_0)$$

where

$$\lambda_0 = \frac{n(\mu_3)^2}{M^2\omega_2\sigma^6}.$$

Hence the approximate distribution of T_L , as deduced from the theory of order statistics, is given as (David and Nagaraja, 2004; DasGupta, 2008, page 93)

$$T_L \sim \mathcal{N}\left(\xi_{3/4} - \xi_{1/4}, \frac{1}{16k} \left(\frac{3}{g^2(\xi_{1/4})} + \frac{3}{g^2(\xi_{3/4})} - \frac{2}{g^2(\xi_{1/4})g^2(\xi_{3/4})}\right)\right) \quad (8)$$

where ξ_{\cdot} and $g(\cdot)$ are the quantile and density functions, respectively, of the $\chi_2^2(\lambda_0)$ distribution. Therefore rejection regions and p -values for the

Hinich-based test of linearity are computed from the upper tail of the normal distribution in (8) with λ_0 estimated by

$$\hat{\lambda}_0 = \frac{n\hat{\gamma}^2(0, 0)}{M^2\omega_2\hat{\gamma}^3(0)}.$$

As two asymptotic distributions are utilized—one in the distribution of T_j and the other in the distribution of the interquartile range—there is good chance the data-dependent bootstrapped distributions will significantly improve the finite-sample performance of this test, and this is confirmed in the simulations.

3. Bootstrap Approximation

We now invoke assumptions on the time series to satisfy requirements of AR(∞) bootstrap procedure and allow us to prove consistency of the bootstrap tests. Superiority of the bootstrap approximation in the case of test statistics with a chi-squared distribution often requires some additional adjustments (Babu, 1984).

■ Assumption 2

Let $\mathbf{X} = \{X_t, t \in \mathbb{Z}\}$ be a zero mean, third-order stationary process with autocovariance function satisfying $\gamma(0) > 0$, and $\sum_{\tau=-\infty}^{\infty} \tau^2 |\gamma(\tau)| < \infty$. Also assume that the spectral density f is positive on $[0, \pi]$, and that

$$\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} (1 + \tau_j^2) |\gamma(\tau_1, \tau_2)| < \infty, \quad j \in \{1, 2\}.$$

Although we will be using an AR(∞) bootstrap, it should be stressed that our assumptions do *not* require that the underlying process is linear, i.e., satisfying (1) with *iid* innovations $\{\varepsilon_t\}$. Nevertheless, Wold's decomposition does ensure that every non-deterministic, possibly nonlinear process has a representation of the form

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{9}$$

where $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ and $\{\varepsilon_t\}$ is a *white noise* process with mean zero and variance

$$\sigma^2 = E(X_t - P_{H_{t-1}} X_t)^2;$$

here, H_{t-1} is defined by

$$H_{t-1} = \overline{sp}\{X_j, j < t - 1\}$$

and $P_Z Y$ denotes the projection of Y onto Z . Note that

$$\sigma^2 = \lim_{p \rightarrow \infty} \sigma_p^2 \equiv E(X_t - \sum_{j=1}^p a_{j,p} X_{t-j})^2$$

and that because of stationarity, the sequence $\{\sigma_p^2\}_{p \in \mathbb{N}}$ and its limit σ^2 , are independent of t .

Furthermore, the conditions imposed on the spectral density f in Assumption 1 have several implications which are important in our set-up. To elaborate, the positivity and continuity of f imply that the process \mathbf{X} also exhibits an autoregressive representation as well; i.e., X_t can be expressed as a mean square convergent series (Wiener and Masani, 1967)

$$X_t = \sum_{j=1}^{\infty} a_j X_{t-j} + e_t. \quad (10)$$

Furthermore, the coefficients $\{a_j, j = 1, 2, \dots\}$ in (10) satisfy $\sum_{j=1}^{\infty} |a_j| < \infty$ and

$$a(z) := 1 - \sum_{j=1}^{\infty} a_j z^j \neq 0 \quad \text{for } |z| = 1$$

Also, the convergence of the sequence $\{a_{j,p}, j = 1, 2, \dots, p\}_{p \in \mathbb{N}}$ towards $\{a_j, j = 1, 2, \dots\}$ fulfills the so-called Baxter's inequality (Baxter, 1962; Pourahmadi, 2001); i.e., there exists an integer L and a constant $C > 0$ such that for all $p \geq L$,

$$\sum_{j=1}^p |a_{j,p} - a_j| \leq \sum_{j=p+1}^{\infty} |a_j|.$$

The aforementioned properties and in particular expression (10) justify the use of the above autoregressive bootstrap to approximate the distribution of the spectral estimators under the null hypothesis of linearity (resp. Gaussianity) even if the underlying process \mathbf{X} is nonlinear (resp. non-Gaussian).

Note that in investigating the properties of the bootstrap procedure proposed in the testing set-up considered in this paper, it is not appropriate to focus only on the case that the underlying process \mathbf{X} indeed satisfies the

null hypothesis of interest. In other words, a successful bootstrap procedure should be able to generate pseudo-series $X_1^*, X_2^*, \dots, X_n^*$ that satisfy the null hypothesis whether the true process \mathbf{X} does or not.

To formalize the bootstrap procedure, we formulate a vector of estimators from a time series realization of length n to be

$$V_n = \left(\hat{f}(\lambda_1^1), \dots, \hat{f}(\lambda_k^1), \hat{f}(\lambda_1^2), \dots, \hat{f}(\lambda_k^2), \right. \\ \left. \hat{f}(\lambda_1^1 + \lambda_1^2), \dots, \hat{f}(\lambda_k^1 + \lambda_k^2), \hat{f}(\lambda_1^1, \lambda_1^2), \dots, \hat{f}(\lambda_k^1, \lambda_k^2) \right)$$

Asymptotic properties of V_n and its convergence to a Gaussian distribution has been established under certain regularity conditions by Brillinger and Rosenblatt (1967a,b); Rosenblatt (1985). However, in this paper we are interested in estimating the distribution of V_n under the *null* hypothesis that the underlying process is linear (resp. Gaussian). To approximate this distribution, we propose the following algorithm that is based on the AR(∞)-bootstrap.

Bootstrap algorithm for testing linearity or Gaussianity

Step 1: Fit an AR(p) model to \mathbf{X} with estimated coefficients $\hat{\mathbf{a}}_p = (\hat{a}_{1,p}, \hat{a}_{2,p}, \dots, \hat{a}_{p,p})$; i.e., $\hat{\mathbf{a}}_p$ is an estimator for \mathbf{a}_p where

$$\mathbf{a}_p = (a_{1,p}, a_{2,p}, \dots, a_{p,p}) = \arg \min_{(c_1, \dots, c_p)} \mathbb{E} \left[\left(X_t - \sum_{j=1}^p c_j X_{t-j} \right)^2 \right].$$

Step 2: Let $\mathbf{X}^* = \{X_1^*, X_2^*, \dots, X_n^*\}$ be a series of n pseudo-observations generated by

$$X_t^* = \sum_{j=1}^p \hat{a}_{j,p} X_{t-j}^* + u_t^* \quad (t = 1, \dots, n) \quad (11)$$

where $X_t^* := 0$ for $t \leq 0$ and the u_t^* 's are iid random variables having mean zero and distribution function F_n which is selected based on the purpose of the analysis. One of three distribution functions are selected depending on the null hypothesis under consideration:

Linear null ($H_0^{(1)}$): If the null hypothesis states the time series is linear, then set $F_n = F_n^{(1)}$ to be the empirical distribution function of the centered residuals $\hat{u}_t - \bar{u}_n$, where

$$\hat{u}_t = X_t - \sum_{j=1}^p \hat{a}_{j,p} X_{t-j} \quad (t = p, p+1, \dots, n)$$

and

$$\bar{u}_n = \frac{1}{n-p} \sum_{t=p+1}^n \hat{u}_t.$$

Linear symmetric null ($H_0^{(2)}$): If the null hypothesis states the time series is linear with a symmetric distribution of errors, then set $F_n = F_n^{(2)}$ to be a symmetrized version of $F_n^{(1)}$ obtained by setting $u_t^* = S_t u_t^+$ with $S_t \stackrel{\text{iid}}{\sim} \text{unif}\{-1, 1\}$ (the discrete uniform distribution on -1 and 1) and $u_t^+ \sim F_n^{(1)}$.

Gaussian null ($H_0^{(3)}$): If the null hypothesis states the time series is linear with Gaussian errors, then set $F_n = F_n^{(3)} = N(0, \hat{\sigma}_p^2)$, where

$$\hat{\sigma}_p^2 = \frac{1}{n-p} \sum_{t=p+1}^n (\hat{u}_t - \bar{u}_n)^2.$$

Step 3: $V_n^{*(i)}$ ($i = 1, 2, 3$) is computed analogous to V_n but with \mathbf{X}^* replacing \mathbf{X} ; \mathbf{X}^* is generated from (11) with $u_t^* \stackrel{\text{iid}}{\sim} F_n^{(i)}$.

Depending on the choice of F_n in Step 2, we can then use V_n^* to approximate the distribution of V_n under the three null hypotheses $H_0^{(1)}$, $H_0^{(2)}$, and $H_0^{(3)}$.

The above bootstrap algorithm entails approximating the null distribution of V_n by that of an AR(p) model; the name AR(∞)-bootstrap will be justified since the order p will be allowed to tend to infinity as a function of the sample size n (cf. Assumption 3 below). Since an AR(p) or AR(∞) model with iid errors is a linear time series, the proposed AR(∞)-bootstrap algorithm in effect approximates the null distribution of V_n by the distribution of V_n as computed from a linear series that satisfies the given null hypothesis.

To investigate the asymptotic properties of the bootstrap, the following technical assumptions are imposed that deal with the behavior of the autoregressive order p and the estimators used.

■ **Assumption 3**

- (i) $p = p(n) \in [p_{\min}(n), p_{\max}(n)]$, where the sequences $p_{\min}(n)$ and $p_{\max}(n)$ are nonstochastic and satisfy

$$p_{\max}(n) \geq p_{\min}(n) \xrightarrow{n \rightarrow \infty} \infty$$

and

$$p_{\max}^9(n)(\log(n))^3/n^2 \xrightarrow{n \rightarrow \infty} 0.$$

- (ii) The AR parameter estimators satisfy

$$\max_{1 \leq j \leq p} |\hat{a}_{j,p} - a_{j,p}| = O(\sqrt{\log(n)}/\sqrt{n})$$

uniformly in $p \leq p_n$ where $p_n = o(\sqrt{n/\log(n)})$ such that $\hat{\sigma}_p^2 \rightarrow \sigma^2$ in probability as $n \rightarrow \infty$.

- (iii) The sequence of distribution functions $F_n^{(i)}$, $i = 1, 2, 3$, converge to a distribution function $F^{(i)}$, where $F^{(2)}$ is a symmetric distribution function and $F^{(3)}$ is the distribution function of $\mathcal{N}(0, \sigma^2)$. Furthermore,

$$\int u^r dF_n^{(i)} \xrightarrow{n \rightarrow \infty} \int u^r dF^{(i)} \quad \text{for } r = 1, 2, \dots, 6$$

and $\int u dF^{(1)} = 0$.

Part (ii) of the above assumption imposes a weak condition on the sequence of estimators used in the second step of the bootstrap algorithm. Derivation of such a condition requires additional assumptions on the dependence structure of the underlying process \mathbf{X} than those imposed in Assumption 1. For a linear time series, it is well-known that condition (ii) above is satisfied, for instance, using the least squares or the Yule-Walker estimators of the AR parameters (Hannan and Kavalieris, 1986). However, it is expected that such a stochastic behavior of the parameters of a p^{th} -order autoregressive fit can be established under general dependence conditions (e.g., mixing or weak dependence) on the underlying process. Part (iii) is concerned with the asymptotic behavior of the distribution of pseudo-errors e_t^* .

Now we will begin the setup for proving consistency of the AR(∞) algorithm which is summarized in Theorem 3.1. Define

$$a_p(z) = 1 - \sum_{j=1}^p a_{j,p} z^j$$

and let $\psi_{j,p}$, $j = 1, 2, \dots$ be the coefficients in the power series expansion of $1/a_p(z)$. It follows that

$$1 - \sum_{j=1}^{\infty} \psi_{j,p} z^j \neq 0 \quad \text{for } |z| \leq 1$$

where

$$\psi_{j,p} = \sum_{k=1}^{j \wedge p} \phi_{k,p} \psi_{j-k,p} \quad \text{for } j = 1, 2, \dots$$

and $\psi_{0,p} = 1$ and $\psi_{j,p} = 0$ for $j < 0$.

Let V be a column vector with $4k$ entries analogous to V_n but with true values of spectral and bispectral densities as components instead of estimators as components. Let $\{\psi_j\}$ be the coefficients in the Wold decomposition of \mathbf{X} given in (9), and define new processes $\mathbf{X}^{(i)}$ by

$$X_t^{(i)} = \sum_{j=0}^{\infty} \psi_j e_{t-j} \quad i = 1, 2, 3.$$

where $e_t \stackrel{\text{iid}}{\sim} F^{(i)}$; recall that $F^{(i)}$ is the stochastic limit of the sequence $F_n^{(i)}$ given in Assumption 3(iii). Corresponding to the series $\mathbf{X}^{(i)}$, we define the quantities $V_n^{(i)}$ and $V^{(i)}$ in the same manner as the quantities V_n and V which correspond to \mathbf{X} . Let D_n be the $4k$ diagonal matrix with the first $3k$ elements of the main diagonal equal to $\sqrt{n/M}$ and the remaining k elements equal to \sqrt{n}/M , and define

$$L_n^{(i)} = D_n(V_n^{(i)} - V^{(i)}).$$

In the sequel we show that the bootstrap procedure proposed succeeds in mimicking correctly the distribution of $L_n^{(i)}$ under the null hypothesis $H_0^{(i)}$. To proceed, we consider the bootstrap random variable

$$L_n^{*(i)} = D_n(V_n^* - V^{(i)})$$

where $V^{*(i)}$ is produced in Step 3 of the bootstrap algorithm.

Our main theorem, below (whose proof is provided in the Appendix) establishes the asymptotic validity of the proposed AR(∞)-bootstrap algorithm in that it correctly approximates the distribution of V_n under any of the null hypotheses $H_0^{(1)}$, $H_0^{(2)}$, or $H_0^{(3)}$.

Theorem 3.1. *Suppose Assumptions 1, 2, and 3 are satisfied and $n/M^5 \rightarrow 0$ as $n \rightarrow \infty$. Then,*

$$\lim_{n \rightarrow \infty} \sup_x \left| P_{H_0^{(i)}}(L_n^{(i)} \leq x) - P(L_n^{*(i)} \leq x | X_1, X_2, \dots, X_n) \right| = 0$$

where $P_{H_0^{(i)}}(L_n^{(i)} \leq x)$ denotes the distribution function of $L_n^{(i)}$ under the null hypothesis $H_0^{(i)}$.

The assumption $n/M^5 \rightarrow 0$ in the above theorem ensures that the bias $D_n(\mathbb{E}[V_n] - V)$ in estimating the spectral quantities of interest is asymptotically negligible; i.e., the random variables $D_n(V_n - V)$ and $D_n(V_n - \mathbb{E}[V_n])$ have the same limiting distribution. If in fact infinite-order kernels are being considered and the underlying spectral and bispectral densities are sufficiently smooth, then this assumption can be removed entirely; cf. Berg and Politis (2009).

Theorem 3.1 together with a version of the continuous mapping theorem—see e.g. Shorack (2000)—imply that the bootstrap procedure proposed can be applied to correctly approximate the null distribution of any test statistic that is a continuous function of $L^{(i)}$.

Corollary 3.1. *Let $T(\cdot)$ be an almost everywhere continuous function of a real argument. Then, under the assumptions of Theorem 3.1, for $i \in \{1, 2, 3\}$ we have*

$$\lim_{n \rightarrow \infty} \sup_x \left| P_{H_0^{(i)}}(T(L_n^{(i)}) \leq x) - P(T(L_n^{*(i)}) \leq x | X_1, X_2, \dots, X_n) \right| = 0.$$

Two immediate applications of interest include bootstrapped versions of the test statistics T_G and T_L for Gaussianity and linearity, respectively, as described in Section 2.3. That is, from V_n^* computed in step 3 of the bootstrap algorithm, compute $T_G^* = \sum_{j=1}^k T_j^*$ and $T_L^* = \text{IQR}(T_j^*)$. From the above corollary, the bootstrapped estimates consistently approximate the null distributions of T_G and T_L , respectively.

Repeating the bootstrap algorithm, multiple bootstrapped statistics are produced that yield the bootstrapped distribution of the test statistics of interest, and rejection regions as well as p -values can be computed from the quantiles of bootstrapped distributions. This procedure is performed with simulated and real data sets in the following section.

4. Simulations

Simulations were performed in R to further validate the proposed bootstrap algorithm and compare test statistics based on the bispectrum for Gaussianity and linearity with their bootstrapped equivalents. In addition two real data sets were considered—daily S&P 500 returns and quarterly US real GNP growth rate.

4.1. Simulated series

Eight time series were considered in the simulations and are described below.

iid norm: $X_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. This process is both Gaussian and linear.

iid chisq: $X_t \stackrel{\text{iid}}{\sim} \chi_1^2$. This process is not Gaussian but is linear.

AR(1): $X_t = .9X_{t-1} + \varepsilon_t$ where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. This process is both Gaussian and linear.

ARMA(2,2): $X_t = .8897X_{t-1} - .4858X_{t-2} + \varepsilon_t - .2279\varepsilon_{t-1} + .2488\varepsilon_{t-2}$ where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, .1796^2)$. This ARMA model comes from the first example in the built-in `arma.sim` command in R which is used to simulated an ARMA time series. This process is both Gaussian and linear.

bilinear(1,0,1,1): As is considered in Berg and Politis (2009), data is simulated from the following bilinear model:

$$X_t = .4X_{t-1} + .4X_{t-1}\varepsilon_{t-1} + \varepsilon_t$$

where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. This process is neither Gaussian nor linear.

ARCH(1): As considered in Fan and Yao (2003), data is simulated from the following ARCH model:

$$X_t = \varepsilon_t \sqrt{1.5 + .9X_{t-1}^2}$$

where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. This process is neither Gaussian nor linear.

GARCH(1,3): In modeling daily S&P 500 returns, Fan and Yao (2003) considered the following GARCH(1,3) model with Gaussian error which is also considered in the simulations:

$$X_t = \sigma_t \varepsilon_t \quad \text{and} \quad \sigma_t^2 = .015 + .112X_{t-1}^2 + .492\sigma_{t-1}^2 - .034\sigma_{t-2}^2 + .420\sigma_{t-3}^2$$

where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$. This process is neither Gaussian nor linear.

TAR: In modeling the quarterly US real GNP growth rate, Tiao and Tsay (1994) utilized the following four-regime threshold autoregressive model which is also considered in the simulations:

$$X_t = \begin{cases} -0.015 - 1.076X_{t-1} + \varepsilon_{1,t}, & X_{t-1} \leq X_{t-2} \ \& \ X_{t-2} \leq 0 \\ -0.006 + .630X_{t-1} + \varepsilon_{2,t} - .756X_{t-2}, & X_{t-1} > X_{t-2} \ \& \ X_{t-2} \leq 0 \\ 0.006 + .438X_{t-1} + \varepsilon_{3,t}, & X_{t-1} \leq X_{t-2} \ \& \ X_{t-2} > 0 \\ .004 + .443X_{t-1} + \varepsilon_{4,t}, & X_{t-1} > X_{t-2} \ \& \ X_{t-2} > 0 \end{cases}$$

where $\varepsilon_{1,t} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, .0062^2)$, $\varepsilon_{2,t} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, .0132^2)$, $\varepsilon_{3,t} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, .0094^2)$, and $\varepsilon_{4,t} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, .0082^2)$. This process is neither Gaussian nor linear.

4.2. *Kernels and other user-defined parameters*

Infinite-order, flat-top lag-window functions for $w(\cdot)$ and $w(\cdot, \cdot)$ have been shown to produce higher-order accurate estimators of the spectral and bispectral densities (Politis and Romano, 1995; Berg and Politis, 2009), and therefore we elect to use the suggested kernels in the simulations. In particular, we use a trapezoid-shaped function for $w(\cdot)$ and a right-pyramidal frustrum-shaped function for $w(\cdot, \cdot)$. The specific equations for these functions are given as

$$w(\tau) = 2(1 - |\tau|)^+ - (1 - 2|\tau|)^+ \\ w(\tau_1, \tau_2) = 2w_0(\tau_1, \tau_2) - w_0(2\tau_1, 2\tau_2)$$

where

$$w_0(x, y) = \begin{cases} (1 - \max(|x|, |y|))^+, & -1 \leq x, y \leq 0 \text{ or } 0 \leq x, y \leq 1 \\ (1 - \max(|x + y|, |x - y|))^+, & \text{otherwise} \end{cases}$$

Graphs of $w(\tau)$ and $w(\tau_1, \tau_2)$ are provided in Figure 3.

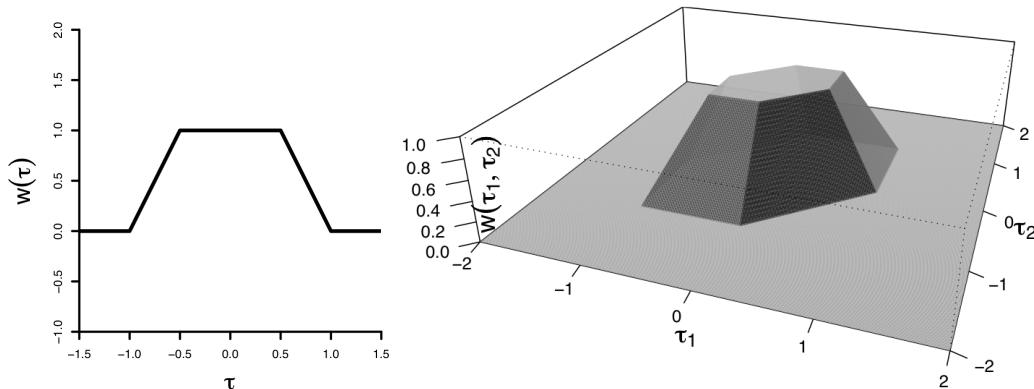


Figure 3: Flat-top functions $w(\cdot)$ and $w(\cdot, \cdot)$ used in (6)

Simulations from three different sample sizes of 250, 500, and 1000 are included. There are a number of user-defined parameters that the Hinich test and the bootstrap algorithm requires. Indeed, results will vary with different choices of the parameters, but it has been noticed in the simulations that the procedure is not very sensitive to the choice of parameters provided the parameters are reasonably selected. Sensitivity of the estimates due to varying the choice of the parameters is considered in testing the two real data sets below.

There were 200 bootstrap replications for each bootstrapped test, and 504 realizations performed across all tests. Other parameters include number of grid points (k) considered, the bispectral bandwidth (M_b) and spectral bandwidth (M_s), and the order of the AR approximation (p). The choices of these parameters over the three different sample sizes are provided in Table 1.

n	k	M_b	M_s	p
250	21	4	8	15
500	36	6	12	20
1000	55	8	15	30

Table 1: User-defined parameters used in the simulations

In general, there is not much sensitivity of the results due to selection of the user-define parameters listed in Table 1. Asymptotically, all of the parameters should tend to infinity. For the grid points, the more points selected, the more robust the test statistic will be, but too many grid points will increase dependencies among the normalized bispectral estimates. For the bandwidths, the bispectral bandwidth should converge to infinity at a slower rate than the spectral bandwidth (Berg and Politis, 2009; Subba Rao and Gabr, 1984); asymptotic rates are also available. For the AR parameter, p should be large enough to robustly approximate the residual time series, but this parameter in general has little effect on the performance of the test.

4.3. Simulation results

The results of the simulations are presented in Figures 4 (testing Gaussianity) and 5 (testing linearity). They depict the number of times, out of 504 realizations, the two test statistics—Hinich and bootstrap—rejected the null hypotheses of Gaussianity and linearity at the $\alpha = .05$ level. If the null is indeed true, then we would ideally see around 5% of the tests to be rejected, whereas if the null were not true, we would ideally expect 100% of the tests to be rejected. The ideal level for each model is depicted in the plots with a bullet, and therefore the closeness of a given test’s performance to the bullet indicates its accuracy. Results among the three different sample sizes of 250, 500, and 1000 are presented together for each of the eight models.

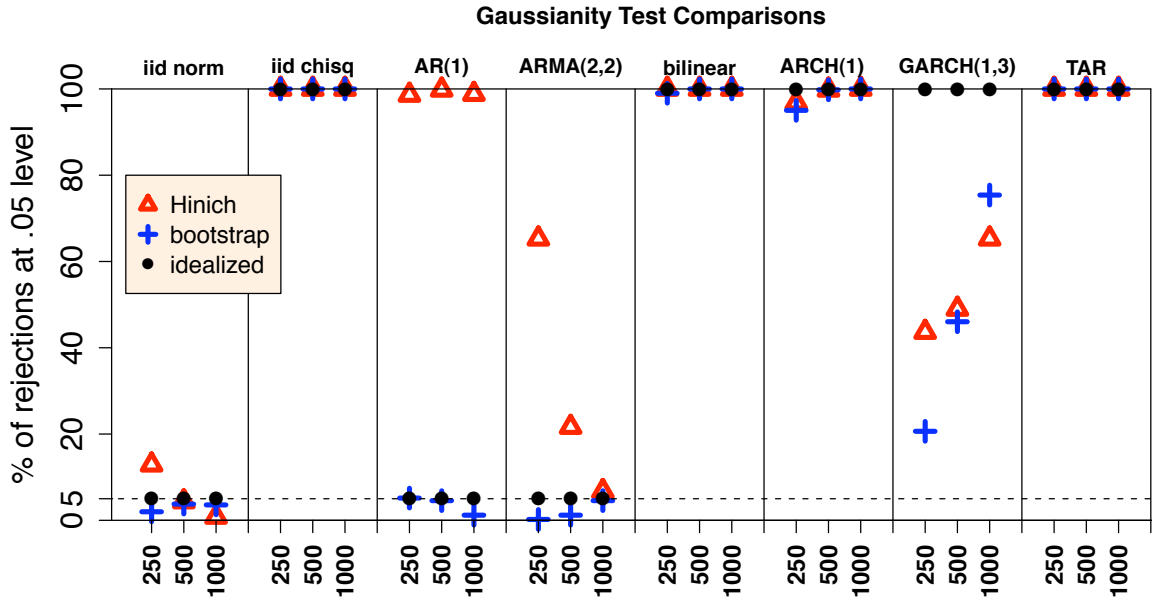


Figure 4: Comparison of Hinich's **Gaussianity** test with a bootstrapped form of the test

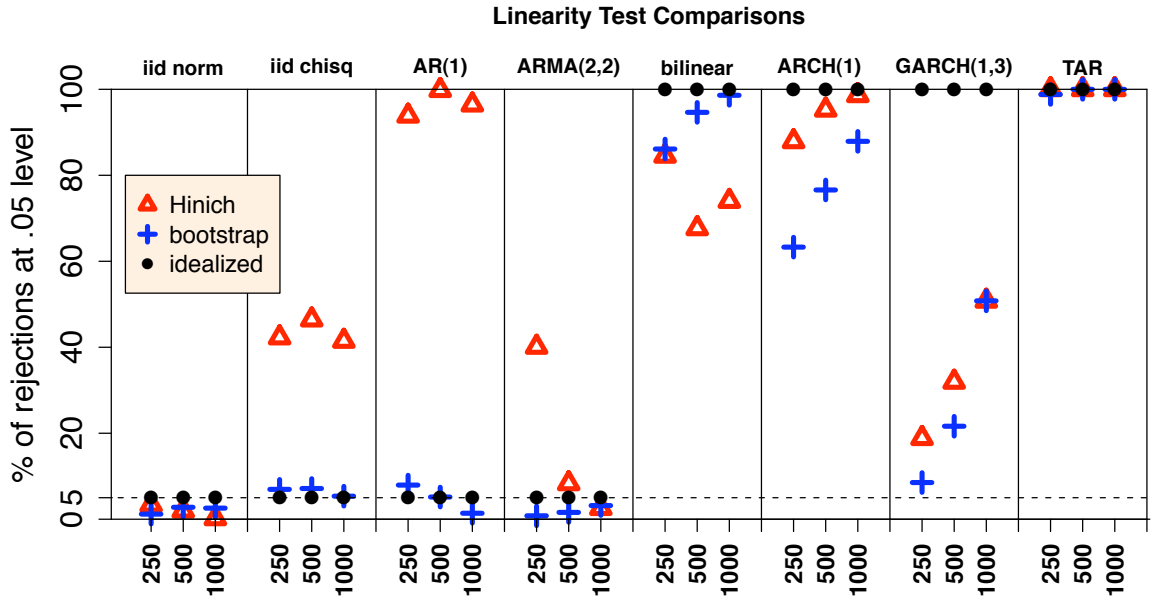


Figure 5: Comparison of Hinich's **linearity** test with a bootstrapped form of the test

The bootstrapped test does a much better job in creating a test procedure with the correct size which is fundamental in testing. The powers of the tests are mostly comparable but it is somewhat unfair to compare the powers of

tests that have different sizes. One approach to re-formulate the powers of tests with different sizes is suggested in Parker et al. (2006).

It is interesting that Hinich's test seems to completely break down on the simple AR(1) model with AR coefficient of .9, even at the very large sample size of 1000. After seeing the results of this model, different parameters were checked to see if performance could be improved. By carefully tweaking the parameters, some amount of improvement could be obtained, but the overall performance of Hinich's test is quite limited in this model.

5. Testing with real data

The simulations above indicate the bootstrapped test can be more conservative than the original test for Gaussianity and linearity, and the Hinich test can produce spurious results for some series like the AR model considered above. We now compare the Hinich tests and its bootstrap form on two real data sets and allow the user-defined parameters to be randomly drawn within reasonable ranges.

5.1. The data – GNP and S&P

Quarterly data of the US real GNP from January 1947 through January 2009 was collected from the St. Louis Reserve FRED database. Growth rate data was computed by differencing the log transform of the seasonally adjusted data. The sample size of this data set (after differencing) is 248. This index, from 1947 through 1991 was modeled in Tiao and Tsay (1994) by the TAR model simulated in the previous section.

Daily S&P 500 returns on the published S&P composite index from January 3, 1972 through December 31, 2008 was obtained from the CRSP database via University of Pennsylvania's WRDS data management system. The sample size of this data set is 9,338. This data is often modeled with the GARCH model to account for the clustered volatility. Fan and Yao (2003) considered the GARCH(1,3) model with Gaussian errors that was used in the simulations of the previous section.

There is ample evidence that S&P return index possesses nonlinear characteristics (Abhyankar et al., 1997; Brock et al., 1991; Hinich and Patterson, 1985, 1989; Hsieh, 1989; Vaidyanathan and Krehbiel, 1992). There is also some research based on Hinich's linearity test indicating the US real GNP growth data is nonlinear (Ashley and Patterson, 1989; Scheinkman

and LeBaron, 1989) and many threshold autoregressive models and generalizations have been used to model the series (Potter, 1995; Scheinkman and LeBaron, 1989; Terasvirta, 1994; Tiao and Tsay, 1994). Plots of these two data sets are provided in Figure 6.

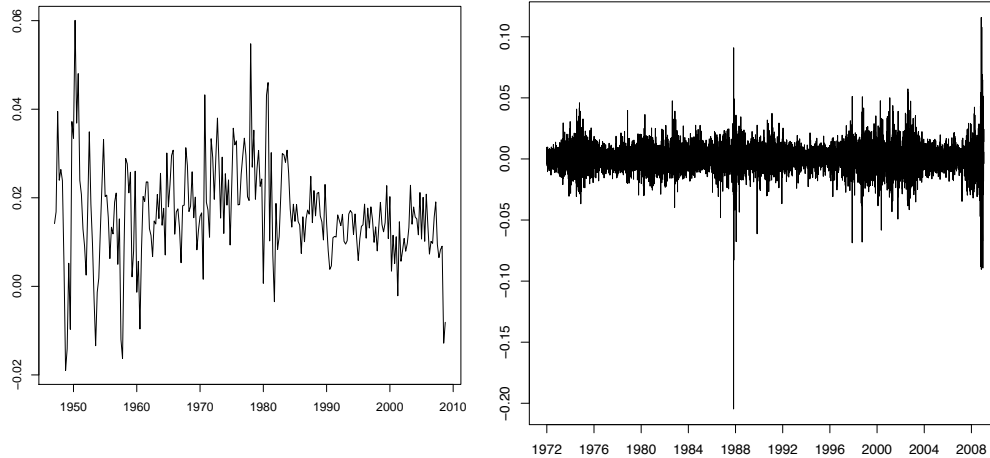


Figure 6: Quarterly US real GNP growth rate (left) and daily S&P 500 returns (right)

5.2. Gaussianity and linearity testing on real data

For the GNP series ($n = 248$), parameters are randomly selected as follows. r is simulated from the discrete uniform distribution on $\{3, 4, 5\}$ which yields k where $k = \binom{r+1}{2}$. M_b is simulated from the discrete uniform distribution on $\{2, 3, \dots, 8\}$. Asymptotically, M_s is larger than M_b , and therefore we take $M_s = cM_b$ where $c \sim \text{unif}(1.5, 3)$. Finally, for the bootstrap AR order parameter p , we let p be drawn from the discrete uniform distribution $\{4, 5, \dots, 15\}$.

For the S&P series ($n = 9,338$), parameters are randomly selected as follows. $r \sim \text{unif}\{5, 6, \dots, 20\}$; $M_b \sim \text{unif}\{8, 9, \dots, 30\}$; $M_s = cM_b$ where $c \sim \text{unif}(2, 6)$; $p \sim \text{unif}\{10, 11, \dots, 40\}$. For both data sets, the bootstrap results are based on 200 bootstrap iterations.

The results for the S&P series were unanimous. Every test, Hinich and bootstrap alike, no matter the choice of parameters, flat-out rejected Gaussianity and Linearity, always with a p -value of 0.

Conclusions for the GNP series, however, were vastly different. At the .05 level, the Hinich test rejected Gaussianity in 84.4% of the 320 randomly selected user-defined parameters, but the bootstrapped test rejected Gaussianity for only 2.8% of the selected parameters. Results also differ with the

linearity test as the Hinich test rejecting linearity for 66.5% of the parameters, and the bootstrap test did not reject linearity for any of the randomly selected parameters.

This inconsistency between the Hinich and bootstrap tests draws striking resemblance to the findings in the AR and ARMA simulations in the previous section. For the simulated AR and ARMA series, the Hinich test frequently but incorrectly rejected both Gaussianity and linearity, whereas the bootstrapped version fully corrected this limitation.

Furthermore the bootstrap test is seen to be less sensitive to the user-specified parameters as it consistently rejected Gaussianity and linearity where as the Hinich test for linearity drew rejections for 66.5% of the randomly selected parameters and therefore failed to reject for the remaining 33.5% of the parameters. Although Hinich's Gaussian test was less variable, it still exhibited greater sensitivity towards the user-defined parameters than the bootstrap test.

Appendix: Proof of Theorem 3.1

For notational simplicity we write L_n and L_n^* for $L_n^{(i)}$ and $L_n^{*(i)}$ respectively. By the conditions imposed on the spectral density f in Assumption 1 and Wiener's Theorem, we get for $\psi(z) = 1/a(z) = \sum_{j=0}^{\infty} \psi_j z^j$ that $\psi(z) \neq 0$ for $|z| = 1$ and $\sum_j |\psi_j| < \infty$. These properties together with $\sum_{j=1}^{\infty} |a_j| < \infty$ and $\sum_{j=0}^{\infty} a_j z^j \neq 0$ for $|z| = 1$, allow us to argue exactly as in the proof of Lemma 8.2 and Lemma 8.3 of Kreiss (1988) and to derive the following useful bounds for $\psi_{j,p}$ and its estimator $\hat{\psi}_{j,p}$.

A constant $C > 0$ exists such that for all large p

$$\sum_{j=1}^{\infty} |\psi_{j,p} - \psi_j| \leq C \sum_{r=p}^{\infty} |a_r|. \quad (12)$$

Furthermore, uniformly in $p \leq p_n$, $p_n^2 = o(n/\log(n))$, and uniformly in $j \in \mathbb{N}$ we have

$$|\hat{\psi}_{j,p} - \psi_{j,p}| \leq p \left(1 + \frac{1}{p}\right)^{-j} O_P\left(\sqrt{\log(n)/n}\right). \quad (13)$$

Using Assumption 1 and Assumption 2, it easily follows that

$$\begin{aligned} & \sqrt{n/M}(\hat{f}(\lambda) - f(\lambda)) \\ &= (2\pi)^{-1} \sum_{\tau} w(\tau/M)(\hat{\gamma}(\tau) - \gamma(\tau))e^{-i\lambda\tau} + o_P(\sqrt{n/M}), \end{aligned} \quad (14)$$

and

$$\begin{aligned} \sqrt{n}/M(\widehat{f}(\lambda_1, \lambda_2) - f(\lambda_1, \lambda_2)) &= (2\pi)^{-2} \sum_{\tau_1} \sum_{\tau_2} w_2(\tau_1/M, \tau_2/M) \\ &\times (\widehat{\gamma}_3(\tau_1, \tau_2) - \gamma(\tau_1, \tau_2)) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} + o_P(\sqrt{n}/M). \end{aligned} \quad (15)$$

To see (15) verify that

$$\begin{aligned} \sqrt{n}/M(\widehat{f}(\lambda_1, \lambda_2) - f(\lambda_1, \lambda_2)) &- (2\pi)^{-2} \sum_{|\tau_1| \leq M} \sum_{|\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) \\ &\times (\widehat{\gamma}_3(\tau_1, \tau_2) - \gamma(\tau_1, \tau_2)) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \\ &= -\frac{\sqrt{n}}{4M\pi^2} \sum_{|\tau_1| \leq M} \sum_{|\tau_2| \leq M} (1 - w_2(\tau_1/M, \tau_2/M)) \gamma(\tau_1, \tau_2) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \\ &\quad - \frac{\sqrt{n}}{4M\pi^2} \sum_{|\tau_1| > M} \sum_{|\tau_2| > M} \gamma(\tau_1, \tau_2) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \\ &= R_{1,n} + R_{2,n}, \end{aligned}$$

with an obvious notation for $R_{1,n}$ and $R_{2,n}$. Now, using Assumption 1 and 2 and expanding $w(u_1, u_2)$ in a Taylor series around $w(0, 0) = 1$, we get

$$\begin{aligned} |R_{1,n}| &\leq C \frac{\sqrt{n}}{4M\pi^2} \sum_{|\tau_1| \leq M} \sum_{|\tau_2| \leq M} [(\tau_1/M)^2 + (\tau_2/M)^2] |\gamma(\tau_1, \tau_2)| \\ &\leq O(n^{1/2}/M^3) \sum_{|\tau_1| \leq M} \sum_{|\tau_2| \leq M} [(\tau_1)^2 + (\tau_2)^2] |\gamma(\tau_1, \tau_2)| \\ &= O(n^{1/2}/M^3) \rightarrow 0. \end{aligned}$$

Furthermore,

$$|R_{2,n}| \leq \sqrt{n/M^5} \sum_{|\tau_1| > M} \sum_{|\tau_2| > M} |\tau_1|^2 |\gamma(\tau_1, \tau_2)| \rightarrow 0,$$

by Assumption 2 and $n/M^5 \rightarrow 0$.

Thus to establish the theorem it suffices to show that

$$\lim_{n \rightarrow \infty} \sup_x \left| P(\widetilde{L}_n \leq x) - P(L_{n|X_1, X_2, \dots, X_n}^* \leq x) \right| = 0,$$

in probability, where \tilde{L}_n is the vector obtained by replacing the elements of L_n by the corresponding first terms on the right hand side of (14) and (15). Denote by $\tilde{L}_{j,n}$ and $L_{j,n}^*$ the j th element, $j = 1, 2, \dots, 2N - 1$, of \tilde{L}_n and L_n^* respectively. By the Cramer-Wold device, the assertion of the theorem is proved by showing that for any $c = (c_1, c_2, \dots, c_{2N-1}) \in \mathbb{C}^{2N-1}$, $c' \tilde{L}_n$ and $c' L_n^*$ converges to the same distribution as $n \rightarrow \infty$. To measure distance between distributions we use Mallow's metric d_2 defined for any two probability measures Q_1 and Q_2 on \mathbb{C}^S , $S \in \mathbb{N}$, as

$$d_2(Q_1, Q_2) = \inf\{E\|Y_1 - Y_2\|^2\}^{1/2},$$

where the infimum is taken over all \mathbb{C}^{2S} -valued random variables (Y_1, Y_2) such that $Y_1 \sim Q_1$ and $Y_2 \sim Q_2$. We then have

$$d_2^2(c' \tilde{L}_n, c' L_n^*) = \inf E \left| c' (\tilde{L}_n - L_n^*) \right|^2 \leq \inf \left\{ C \sum_{j=1}^{2N-1} |c_j|^2 E |\tilde{L}_{j,n} - L_{j,n}^*|^2 \right\}$$

for some appropriate constant $C > 0$, where the infimum is taken over all i.i.d. pairs of random variables (u_t, u_t^*) with $u_t \sim F^{(i)}$ and $u_t^* \sim F_n^{(i)}$. The proof that each one of the $2N - 1$ expectation terms above vanish asymptotically as $n \rightarrow \infty$ uses essentially the same arguments. We demonstrate the arguments used only for the expectation terms corresponding to the bispectral density estimators. Let $\delta_{k,r,s}$ denote Kronecker's delta, i.e., $\delta_{k,r,s} = 1$ if $k = r = s$ and zero otherwise and consider, for instance, the $(N + 1)$ th element of \tilde{L}_n

and L_n^* . We then have

$$\begin{aligned}
& \mathbb{E}|L_{N+1,n}^* - \tilde{L}_{N+1,n}|^2 \\
&= \frac{n}{M^2} \mathbb{E} \left| \frac{1}{(2\pi)^2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} (\hat{\gamma}^*(\tau_1, \tau_2) - \gamma^*(\tau_1, \tau_2)) \right. \\
&\quad \left. - \frac{1}{(2\pi)^2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} (\hat{\gamma}(\tau_1, \tau_2) - \gamma(\tau_1, \tau_2)) \right|^2 \\
&= \frac{n}{M^2} \mathbb{E} \left| \frac{1}{(2\pi)^2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \right. \\
&\quad \times \left(\frac{1}{n} \sum_{t \in n(\tau_1, \tau_2)} \sum_{j_1, j_2, j_3} \hat{\psi}_{j_1, p} \hat{\psi}_{j_2, p} \hat{\psi}_{j_3, p} e_{t-j_1}^* e_{t+\tau_1-j_2}^* e_{t+\tau_2-j_3}^* - \sum_j \hat{\psi}_{j, p} \hat{\psi}_{j+\tau_1, p} \hat{\psi}_{j+\tau_2, p} E^*(e_1^{*3}) \right) \\
&\quad \left. - \frac{1}{(2\pi)^2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \right. \\
&\quad \times \left(\frac{1}{n} \sum_{t \in n(\tau_1, \tau_2)} \sum_{j_1, j_2, j_3} \psi_{j_1} \psi_{j_2} \psi_{j_3} e_{t-j_1} e_{t+\tau_1-j_2} e_{t+\tau_2-j_3} - \sum_j \psi_j \psi_{j+\tau_1} \psi_{j+\tau_2} E(e_1^3) \right) \Big|^2 \\
&\leq 2 \frac{n}{M^2} \mathbb{E} \left| (2\pi)^{-2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \right. \\
&\quad \times \left(\frac{1}{n} \sum_{t \in n(\tau_1, \tau_2)} \sum_{j_1, j_2, j_3} \left(\hat{\psi}_{j_1, p} \hat{\psi}_{j_2, p} \hat{\psi}_{j_3, p} - \psi_{j_1} \psi_{j_2} \psi_{j_3} \right) \right. \\
&\quad \times \left(e_{t-j_1}^* e_{t+\tau_1-j_2}^* e_{t+\tau_2-j_3}^* - \delta_{j_1, j_2 - \tau_1, j_3 - \tau_2} E^*(e_1^{*3}) \right) \Big|^2 \\
&\quad \left. + 2 \frac{n}{M^2} \mathbb{E} \left| \frac{1}{(2\pi)^2} \sum_{|\tau_1|, |\tau_2| \leq M} w_2(\tau_1/M, \tau_2/M) e^{-i\lambda_1\tau_1 - i\lambda_2\tau_2} \left(\frac{1}{n} \sum_{t \in n(\tau_1, \tau_2)} \sum_{j_1, j_2, j_3} \psi_{j_1} \psi_{j_2} \psi_{j_3} \right. \right. \right. \\
&\quad \times \left. \left. \left(e_{t-j_1}^* e_{t+\tau_1-j_2}^* e_{t+\tau_2-j_3}^* - e_{t-j_1} e_{t+\tau_1-j_2} e_{t+\tau_2-j_3} - \delta_{j_1, j_2 - \tau_1, j_3 - \tau_2} (E^*(e_1^{*3}) - E(e_1^3)) \right) \right) \right|^2 \\
&= E_n + H_n
\end{aligned}$$

with an obvious notation for E_n and H_n . Notice that E_n is essentially due to the estimation and approximation error

$$(\hat{\psi}_{j, p} - \psi_j) = (\hat{\psi}_{j, p} - \psi_{j, p}) + (\psi_{j, p} - \psi_j) \quad (16)$$

of the coefficients in the linear representation of X_t resp. X_t^* , while H_n is due to the difference between the error distributions of e_t^* and e_t . Now, using the

convergence properties of the estimator $\widehat{\psi}_{j,p}$ to ψ_j and of the distribution of the errors e_t^* to $F^{(i)}$, we can show that both terms vanish asymptotically.

To see this consider first the term E_n and note that

$$\begin{aligned}
E_n &\leq O(1/(nM^2)) \sum_{|\tau_1|, |\tau_2| \leq M} \sum_{|r_1|, |r_2| \leq M} w_2(\tau_1/M, \tau_2/M) w_2(r_1/M, r_2/M) \\
&\times \sum_{t_1 \in n(\tau_1, \tau_2)} \sum_{t_2 \in n(\tau_1, \tau_2)} \sum_{j_1, j_2, j_3} \sum_{l_1, l_2, l_3} \left| \widehat{\psi}_{j_1, p} \widehat{\psi}_{j_2, p} \widehat{\psi}_{j_3, p} - \psi_{j_1} \psi_{j_2} \psi_{j_3} \right| \left| \widehat{\psi}_{l_1, p} \widehat{\psi}_{l_2, p} \widehat{\psi}_{l_3, p} - \psi_{l_1} \psi_{l_2} \psi_{l_3} \right| \\
&\times \left| E(e_{t_1 - j_1}^* e_{t_1 + \tau_1 - j_2}^* e_{t_1 + \tau_2 - j_3}^* e_{t_2 - l_1}^* e_{t_2 + r_1 - l_2}^* e_{t_2 + r_2 - l_3}^*) \right. \\
&\quad \left. - \delta_{j_1, j_2 - \tau_1, j_3 - \tau_2} \delta_{l_1, l_2 - r_1, l_3 - r_2} (E^*(e_1^{*3}))^2 \right|.
\end{aligned}$$

Now, using equation (16),

$$E(e_{t_1}^* e_{t_2}^* e_{t_3}^* e_{t_4}^* e_{t_5}^* e_{t_6}^*) = \begin{cases} E(e_1^{*6}) & \text{if } t_1 = t_2 = t_3 = t_4 = t_5 = t_6 \\ E(e_1^{*4}) E(e_1^{*2}) & \text{if } t_{i_1} = t_{i_2} \neq t_{i_3} = t_{i_4} = t_{i_5} = t_{i_6} \\ (E(e_1^{*3}))^2 & \text{if } t_{i_1} = t_{i_2} = t_{i_3} \neq t_{i_4} = t_{i_5} = t_{i_6} \\ (E(e_1^{*2}))^3 & \text{if } t_{i_1} = t_{i_2} \neq t_{i_3} = t_{i_4} \neq t_{i_5} = t_{i_6} \neq t_{i_1} \\ 0 & \text{else,} \end{cases}$$

where the t_{i_k} 's are elements of the set $\{t_1, t_2, \dots, t_k\}$ and different for different values of k , and

$$\begin{aligned}
&(a_1 a_2 a_3 - b_1 b_2 b_3) \\
&= (a_1 - b_1)(a_2 - b_2)(a_3 - b_3) + (a_1 - b_1)(a_2 - b_2)b_3 + (a_1 - b_1)b_2(a_3 - b_3) \\
&\quad + b_1(a_2 - b_2)(a_3 - b_3) + (a_1 - b_1)b_2 b_3 + b_1(a_2 - b_2)b_3 + b_1 b_2(a_3 - b_3),
\end{aligned}$$

we can bound E_n by the sum of several terms a dominant of which is

$$\begin{aligned}
E_{1,n} &\leq O(1/(nM^2))(E(e_1^{*2}))^3 \\
&\times \sum_{|\tau_1|, |\tau_2| \leq M} \sum_{|r_1|, |r_2| \leq M} w_2(\tau_1/M, \tau_2/M) w_2(r_1/M, r_2/M) \\
&\times \sum_{t_1 \in n(\tau_1, \tau_2)} \sum_{t_2 \in n(r_1, r_2)} \sum_{j_1, j_3, l_2} |\widehat{\psi}_{j_1, p} - \psi_{j_1}| |\widehat{\psi}_{j_1 + \tau_1, p} - \psi_{j_1 + \tau_1}| |\widehat{\psi}_{j_3, p} - \psi_{j_3}| \\
&\times |\widehat{\psi}_{t_2 - t_1 - \tau_2 + j_3, p} - \psi_{t_2 - t_1 - \tau_2 + j_3}| |\widehat{\psi}_{l_2, p} - \psi_{l_2}| |\widehat{\psi}_{r_2 - r_1 + l_2, p} - \psi_{r_2 - r_1 + l_2}|.
\end{aligned}$$

Now, because of (16) and using (12) and (13) we get after tedious calculations, that this term is bounded by

$$O_P\left(\frac{M^2 p_{\max}^9}{n^3} \log^3(n) + \sum_{j=p_{\min}}^{\infty} |\psi_j|\right),$$

which goes to zero by Assumption 3(i) and the fact that $M^2/n \rightarrow 0$ and $\sum_{j=p_{\min}}^{\infty} |\psi_j| \rightarrow 0$ as $n \rightarrow \infty$.

Consider next the term H_n . After evaluating the expectation term several terms can be obtained a typical of which is

$$\begin{aligned} H_{1,n}^* &\leq \frac{1}{nM^2} \sum_{|\tau_1|, |\tau_2| \leq M} \sum_{|r_1|, |r_2| \leq M} w_2(\tau_1/M, \tau_2/M) w_2(r_1/M, r_2/M) \\ &\quad \times \sum_{t_1 \in n(\tau_1, \tau_2)} \sum_{t_2 \in n(r_1, r_2)} \sum_{j_1, j_2, l_2} |\psi_{j_1}| |\psi_{j_2}| |\psi_{\tau_2 - \tau_1 + j_2}| |\psi_{t_2 - t_1 + j_1}| |\psi_{l_2}| |\psi_{l_2 + r_2 - r_1}| \\ &\quad \times \left\{ (E(e_1^{*2}))^3 - 2E(e_1^2)E(e_1^* e_1)E(e_1^{*2}) + (E(e_1^2))^3 \right\} \\ &\leq O(1) \left\{ (E(e_1^{*2}))^2 |Ee_t^*(e_1^* - e_1)| + (E(e_1^2))^2 |Ee_1(e_1^* - e_1)| \right\} \\ &\leq O(1) d_2(e_1^*, e_1) \rightarrow 0, \quad \text{as } n \rightarrow \infty \end{aligned}$$

by Assumption 3(iii).

References

- Abhyankar, A., Copeland, L., Wong, W., 1995. Nonlinear dynamics in real-time equity market indices: evidence from the united kingdom. *The Economic Journal*, 864–880.
- Abhyankar, A., Copeland, L., Wong, W., 1997. Uncovering nonlinear structure in real-time stock-market indexes: the s&p 500, the dax, the nikkei 225, and the ftse-100. *Journal of Business & Economic Statistics*, 1–14.
- An, H.-Z., Zhu, L.-X., Li, R.-Z., 2000. A mixed-type test for linearity in time series. *Journal of Statistical Planning Inference* 88, 339–353.
- Ashley, R., Patterson, D., 1989. Linear versus nonlinear macroeconomies: a statistical test. *International Economic Review*, 685–704.

- Ashley, R., Patterson, D., 2009. A test of the garch(1,1) specification for daily stock returns, working paper presented at the 17th Society for Nonlinear Dynamics and Econometrics on April 16, 2009).
- Ashley, R., Patterson, D., Hinich, M., 1986. A diagnostic test for nonlinear serial dependence in time series fitting errors. *Journal of Time Series Analysis* 7 (3), 165–178.
- Babu, G., 1984. Bootstrapping statistics with linear combinations of chi-squares as weak limit. *Sankhyā: The Indian Journal of Statistics, Series A* 46 (1), 85–93.
- Barnett, A., Wolff, R., 2005. A time-domain test for some types of nonlinearity. *IEEE Transactions on Signal Processing* 53 (1), 26–33.
- Baxter, G., 1962. An asymptotic result for the finite predictor. *Math. Scand* 10 (2).
- Berg, A., 2008. Multivariate lag-windows and group representations. *Journal of Multivariate Analysis* 99 (10), 2479–2496.
- Berg, A., Politis, D., 2009. Higher-order accurate polyspectral estimation with flat-top lag-windows. *Annals of the Institute of Statistical Mathematics* 61, 1–22.
- Birkelund, Y., Hanssen, A., 2009. Improved bispectrum based tests for gaussianity and linearity. *Signal Processing*.
- Brillinger, D., 1965. An introduction to polyspectra. *The Annals of mathematical statistics*, 1351–1374.
- Brillinger, D., Rosenblatt, M., 1967a. Asymptotic theory of estimates of kth-order spectra. *Proceedings of the National Academy of Sciences* 57 (2), 206–210.
- Brillinger, D., Rosenblatt, M., 1967b. *Spectral Analysis of Time Series*. Wiley, New York, Ch. Asymptotic theory of kth order spectra in Spectral analysis of time series.
- Brock, W., Hsieh, D., LeBaron, B., 1991. *Nonlinear dynamics, chaos, and instability: statistical theory and economic evidence*. MIT press.

- Brockett, P., Hinich, M., Patterson, D., 1988. Bispectral-based tests for the detection of gaussianity and linearity in time series. *Journal of the American Statistical Association*, 657–664.
- Brockwell, P., Davis, R., 1991. *Time series: theory and methods*. Springer, New York.
- Bühlmann, P., 1997. Sieve bootstrap for time series. *Bernoulli* 3 (2), 123–148.
- Chan, K., 1990. Testing for threshold autoregression. *The Annals of Statistics*, 1886–1894.
- Chan, K. S., Tong, H., 1990. On likelihood ratio tests for threshold autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)* 52 (3), 469–476.
- Chan, W., Tong, H., 1986. On tests for non-linearity in time series analysis. *Journal of forecasting* 5 (4), 217–228.
- Corduas, M., 1994. Nonlinearity tests in time series analysis. *Statistical Methods and Applications* 3 (3), 291–313.
- Cryer, J., Chan, K., 2008. *Time series analysis: with applications in R*. Springer Verlag, New York.
- Daniell, P., 1946. Discussion on symposium on autocorrelation in time series. *Journal of the Royal Statistical Society* 8, 88–90.
- DasGupta, A., 2008. *Asymptotic theory of statistics and probability*. Springer, New York.
- David, H., Nagaraja, H., 2004. *Order statistics*. Wiley, New York.
- Fan, J., Yao, Q., 2003. *Nonlinear time series: nonparametric and parametric methods*. Springer, New York.
- Grenander, U., Rosenblatt, M., 1984. *Statistical analysis of stationary time series*. Chelsea Publishing Company.
- Hannan, E., 1973. The asymptotic theory of linear time-series models. *Journal of Applied Probability* 10 (1), 130–145.

- Hannan, E., 1986. Remembrance of things past. The craft of probabilistic modeling. New York: Springer-Verlag.
- Hannan, E., Kavalieris, L., 1986. Regression, autoregression models. *Journal of Time Series Analysis* 7 (1), 27–49.
- Hansen, B., 2000. Testing for linearity. *Surveys in Economic Dynamics*, 47–72.
- Harvey, D., Leybourne, S., 2007. Testing for time series linearity. *Econometrics Journal* 10 (1), 149–165.
- Hinich, M., 1982. Testing for gaussianity and linearity of a stationary time series. *Journal of time series analysis* 3 (3), 169–176.
- Hinich, M., Mendes, E., Stone, L., 2005. Detecting nonlinearity in time series: Surrogate and bootstrap approaches. *Studies in Nonlinear Dynamics & Econometrics* 9 (4), 3.
- Hinich, M., Patterson, D., 1985. Evidence of nonlinearity in daily stock returns. *Journal of Business & Economic Statistics*, 69–77.
- Hinich, M., Patterson, D., 1989. Evidence of nonlinearity in the trade-by-trade stock market return generating process. In: Barnett, W., Geweke, J., Shell, K. (Eds.), *Economic complexity: chaos, sunspots, bubbles and nonlinearity-international symposium in economic theory and econometrics*. pp. 383–409.
- Hjellvik, V., Tjøstheim, D., 1995. Nonparametric tests of linearity for time series. *Biometrika* 82 (2), 351–368.
- Hong-Zhi, A., Bing, C., 1991. A Kolmogorov-Smirnov type statistic with application to test for nonlinearity in time series. *International Statistical Review/Revue Internationale de Statistique*, 287–307.
- Hsieh, D., 1989. Testing for nonlinear dependence in daily foreign exchange rates. *Journal of Business*, 339–368.
- Jahan, N., Harvill, J. 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.

- Keenan, D., 1985. A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika* 72 (1), 39–44.
- Kokoszka, P. S., Politis, D. N., 2008. The variance of sample autocorrelations: does Bartlett's formula work with ARCH data? Tech. rep., Department of Economics, University of California, San Diego, <http://repositories.cdlib.org/ucsdecon/2008-12/>.
- Kreiss, J., 1988. Asymptotic statistical inference for a class of stochastic processes. Habilitationsschrift, Fachbereich Mathematik, Universität Hamburg.
- Kreiss, J., 1992. Bootstrapping and Related Techniques. Vol. 376. Springer-Verlag, Berlin, Ch. Bootstrap procedures for AR (∞)-processes, pp. 107–113.
- Kugiumtzis, D., 2008. Evaluation of surrogate and bootstrap tests for nonlinearity in time series. *Studies in Nonlinear Dynamics & Econometrics* 12 (1), 4.
- Lahiri, S., 2003. Resampling methods for dependent data. Springer, New York.
- Luukkonen, R., Saikkonen, P., Terasvirta, T., 1988. Testing linearity against smooth transition autoregressive models. *Biometrika* 75 (3), 491–499.
- Paparoditis, E., Streitberg, B., 1992. Order identification statistics in stationary autoregressive moving-average models: vector autocorrelations and the bootstrap. *Journal of Time Series Analysis* 13 (5), 415–434.
- Parker, C., Paparoditis, E., Politis, D., 2006. Unit root testing via the stationary bootstrap. *Journal of Econometrics* 133 (2), 601–638.
- Parzen, E., 1957. On consistent estimates of the spectrum of a stationary time series. *The Annals of Mathematical Statistics*, 329–348.
- Parzen, E., 1961. Mathematical considerations in the estimation of spectra. *Technometrics*, 167–190.
- Petrucelli, J., 1990. A comparison of tests for SETAR-type non-linearity in time series. *Journal of Forecasting* 9 (1).

- Petrucelli, J. and Davies, N., 1986. A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika* 73 (3), 687–694.
- Politis, D. N., 2009. Higher-order accurate, positive semi-definite estimation of large-sample covariance and spectral density matrices. Tech. rep., Department of Economics, UCSD. Paper 2005-03R.
URL <http://repositories.cdlib.org/ucsdecon/2005-03R>
- Politis, D. N., Romano, J., 1995. Bias-corrected nonparametric spectral estimation. *Journal of Time Series Analysis* 16 (1), 67–103.
- Potter, S., 1995. A nonlinear approach to us gnp. *Journal of Applied Econometrics*, 109–125.
- Pourahmadi, M., 2001. *Foundations of time series analysis and prediction theory*. Wiley, New York.
- Priestley, M., 1983. *Spectral analysis and time series, Vol 1 and II*. Academic Press.
- Rosenblatt, M., 1985. *Stationary sequences and random fields*. Birkhäuser Boston, Inc.
- Rosenblatt, M., Van Ness, J., 1965. Estimation of the bispectrum. *The Annals of Mathematical Statistics*, 1120–1136.
- Scheinkman, J., LeBaron, B., 1989. Nonlinear dynamics and gnp data. In: *Economic complexity: chaos, sunspots, bubbles, and nonlinearity, Proceedings of the Fourth International Symposium in Economic Theory and Econometrics* (Cambridge University Press, Cambridge). pp. 213–227.
- Shorack, G., 2000. *Probability for statisticians*. Springer Verlag, New York.
- Subba Rao, T., Gabr, M., 1980. A test for linearity of stationary time series. *Journal of Time Series Analysis* 1 (2), 145–158.
- Subba Rao, T., Gabr, M., 1984. *An introduction to bispectral analysis and bilinear time series models. Vol. 24*.
- Terasvirta, T., 1994. Testing linearity and modelling nonlinear time series. *Kybernetika* 30 (3), 319–330.

- Terasvirta, T., Lin, C., Granger, C., 1993. Power of the neural network linearity test. *Journal of Time Series Analysis* 14 (2), 209–220.
- Terdik, G., Math, J., 1998. A new test of linearity of time series based on the bispectrum. *Journal of Time Series Analysis* 19 (6), 737–753.
- Theiler, J., Galdrikian, B., Longtin, A., Eubank, S., Farmer, J., 1992. Testing for nonlinearity in time series: the method of surrogate data. *Physica D* 58, 77–94.
- Tiao, G., Tsay, R., 1994. Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting* 13 (2).
- Tjøstheim, D., 1994. Non-linear time series: a selective review. *Scandinavian Journal of Statistics*, 97–130.
- Tong, H., 1993. *Non-linear time series: a dynamical system approach*. Oxford University Press.
- Tsay, R., 1986. Nonlinearity tests for time series. *Biometrika* 73 (2), 461–466.
- Vaidyanathan, R., Krehbiel, T., 1992. Does the s&p 500 futures mispricing series exhibit nonlinear dependence across time? *Journal of Futures Markets* 12 (6).
- Van Ness, J., 1966. Asymptotic normality of bispectral estimates. *The Annals of Mathematical Statistics*, 1257–1272.
- Wiener, N., Masani, P., 1967. The prediction theory of multivariate stochastic processes, ii. *Acta Mathematica* 99 (1), 93–137.
- Wong, W., 1997. Frequency domain tests of multivariate gaussianity and linearity. *Journal of time series analysis* 18 (2), 181–194.
- Yuan, J., 2000. Testing linearity for stationary time series using the sample interquartile range. *Journal of Time Series Analysis* 21 (6), 713–722.
- Zurbenko, I., 1986. *The spectral analysis of time series*. North-Holland Series in Statistics and Probability.