Fall 2012

On wavelet-based testing for serial correlation of unknown form using Fan's adaptive Neyman method

Shan Yao
University of New Hampshire, Durham

ON WAVELET-BASED TESTING FOR SERIAL CORRELATION OF UNKNOWN FORM USING FAN'S ADAPTIVE NEYMAN METHOD

BY

SHAN YAO

B.S., Beijing Normal University, CHINA, 2006
M.S., University of New Hampshire, USA, 2008

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Mathematics: Statistics Concentration

September, 2012

This dissertation has been examined and approved.

Dissertation Director, Dr. Linyuan Li

Associate Professor of Mathematics and Statistics

Dr. Ernst Linder

Professor of Mathematics and Statistics

Dr. Philip Ramsey

Lecturer of Mathematics and Statistics

Dr. Edward Hinson

Associate Professor of Mathematics and Statistics

Dr. Donald Hadwin

Professor of Mathematics and Statistics

7/17/2012

Date

DEDICATION

To my parents, grandparents, and Kenny

ACKNOWLEDGEMENTS

This dissertation would not have been accomplished without the help and support that I gained at the University of New Hampshire. I would like to express my appreciation and recognition to the many individuals who have helped me made thus far.

First I would like to give special thanks to my advisor Dr. Linyuan Li, for all of his inspiring guidance and great wisdom, who helped me through my rigorous doctoral research. I would also like to thank my committee members: Dr. Ernst Linder, Dr. Philip Ramsey, Dr. Edward Hinson, and Dr. Donald Hadwin, who provided me helpful suggestions and valuable advise throughout my study at University of New Hampshire. I am also grateful for Dr. Rita Hibschweiler, Dr. Tom Zhang, Jan Jankowski, April Flore, and Ellen O'Keefe for their time and various forms of support. Finally, I would like to express my gratitude to my fellow graduate students: Eric Laflamme, Neil Bornstein, Chengwei Yuan, Zhaoyu Yin, Kewei Lu, Yibin Pan, Tianjiao Dai, Yanni Chen, Hyung Kim, Beth Roberts, and Vincent Mateescu for their great friendship.

I have received generous financial and academic support from the University of New Hampshire during my graduate study. I appreciate the teaching assistantship offered by the Department of Mathematics and Statistics at University of New Hampshire throughout these six years. I am also thankful to the summer teaching fellowship granted by the Graduate School in 2011 which made my summer research possible and achievable.

Overall I would like to say that this dissertation would not have been accomplished without the support and encouragement from my family and friends. Many thanks to all that gave their support, I would not be where I am today without all of you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

ON WAVELET-BASED TESTING FOR SERIAL CORRELATION OF UNKNOWN FORM USING FAN'S ADAPTIVE NEYMAN METHOD

by

Shan Yao

University of New Hampshire, September 2012

Advisor: Dr. Linyuan Li

Test procedures for serial correlation of unknown form with wavelet methods are investigated in this dissertation. The new wavelet-based consistent test is motivated using Fan's (1996) canonical multivariate normal hypothesis testing model. In our framework, the test statistic relies on empirical wavelet coefficients of a wavelet-based spectral density estimator. We advocate the choice of the simple Haar wavelet function, since evidence demonstrates that the choice of the wavelet function is not critical. Under the null hypothesis of no serial correlation, the asymptotic distribution of a vector of empirical wavelet coefficients is derived, which is the multivariate normal distribution in the limit. It is also shown that the wavelet coefficients are asymptotically uncorrelated. The proposed test statistic presents the serious advantage to be completely data-driven or adaptive, which avoids the need to select any smoothing parameters. Furthermore, under a suitable class of local alternatives, the wavelet-based method is consistent against serial correlation of

unknown form. The test statistic is expected to exhibit better power than the current test statistics when the true spectral density displays significant spatial inhomogeneity, such as seasonal or cycle periodicities. However, the convergence of the test statistic toward its respective asymptotic distribution is expected to be relatively slow. Thus, Monte Carlo methods are investigated to determine the corresponding critical value. In a small simulation study, the new method is compared with several current test statistics, with respect to their empirical levels and powers.

INTRODUCTION

Testing for serial correlation has been a long-standing problem in statistics and econometrics. Many test statistics for serial correlation have been proposed, including the popular Box-Pierce-Ljung portmanteau test statistics developed in the seminal works of Box and Pierce (1970) and Ljung and Box (1978). These portmanteau test statistics have been generalized using a spectral density approach by Hong (1996), where the testing procedures relied on a normalized distance between a kernel-based spectral density estimator and the spectral density under the null hypothesis of no serial correlation. Wavelet methods represent an alternative approach to kernel-based spectral density estimators. Using a wavelet expansion of the spectral density, Lee and Hong (2001) proposed a wavelet-based spectral density estimator and they obtained a consistent test statistic for serial correlation using quadratic integrated measure. In Duchesne, Li and Vandermeerschen (2010), a similar test statistic has been investigated, using wavelet thresholding of the wavelet coefficients.

In Hong's (1996) spectral density approach, a kernel function $k(\cdot)$ needs to be specified and the user has to also specify a smoothing parameter or a truncation parameter p_n, depending on the nature of the kernel function. Interestingly, it provided an interpretation for Box-Pierce-Ljung test statistics, which can be considered as a particular case of Hong's statistic using the truncated uniform kernel and a truncation parameter. For the kernel-based test of Hong (1996), the selection of the kernel functions has very little impact on the performance of the test statistic, except for the truncated uniform kernel where p_n is in fact a lag order. However, theoretical and empirical evidence suggest that the selection of p_n can have a significant impact on the power of the spectral test statistic. From a theoretical point of view, the test statistic of Hong (1996) is consistent under the assumptions $p_n/n \to 0$ and $p_n \to \infty$, n being the sample size. However, in practice, p_n is fixed and the test statistic with small values of p_n, when p_n is denoted as a lag order, may miss high order dependence, due for example to seasonality. On the other hand, when p_n corresponds to a smoothing parameter (not a lag order), it may be difficult to specify in practical applications. Alternatively, a wavelet basis can be used to describe the spectral density. The test statistic of Lee and Hong (2001) was constructed using a quadratic distance measure between a wavelet-based spectral density estimator and the null spectral density. In that framework, a finest scale J_n needs to be selected. The finest scale J_n used in the wavelet-based test statistic also has significant impact on the performance of the test statistic. As a spatially adaptive estimation method, wavelet method has its major strength in detecting local characteristics and global alternations such as peaks and spikes. As a result, the wavelet-based test statistics of Lee and Hong (2001) are expected to reach better power than the kernel-based test statistics of Hong (1996) if the spectral density displays significant spatial inhomogeneity. Both Hong's (1996) test statistic and Lee and Hong's (2001) method involve the selection of smoothing parameters p_n and J_n, which are chosen either by subjective approaches or data-driven methods such as the method given in Walter (1994). Cross-validation or data-driven methods may be appealing, but they are computationally intensive. Furthermore, the additional variability due to the data-driven selection may affect the finite sample performance of the test statistics. These issues may be viewed as serious disadvantages, see Li (2004, pp. 104 and 168), among others. The Duchesne, Li and Vandermeerschen (2010) wavelet thresholding test statistic was also motivated using a quadratic distance measure between a wavelet-based spectral density estimator and the null spectral density. Using an appropriate thresholding parameter, shrinkage rules were applied to the empirical wavelet coefficients by vanishing those which are smaller than the threshold parameter. They found that the thresholding rule was particularly appealing when most of the energy was concentrated on few dimensions with unknown locations.

In this dissertation, we also consider using wavelet coefficients and a wavelet-based spectral density approach. The new test statistic for testing for serial correlation of unknown form is motivated using Fan's (1996) adaptive Neyman method. Neyman's fundamental testing problem is for a location parameter in a multivariate normal framework. If the large coefficients of the location parameters are concentrated on the first few dimensions, a test statistic based on the first few components of the random vector is expected to be powerful. Fan (1996) proposed a simple and powerful procedure to select the number of dimensions based on power consideration. That approach is comparable to thresholding methods, since in both approaches the test statistics are based on the significant few dimensions. In our framework, the random components are the wavelet coefficients. Based on the theoretical and empirical results of Lee and Hong (2001) and Duchesne, Li and Vandermeerschen (2010), the choice of the wavelet function is not critical. Thus, we use the simple Haar wavelet function to compute the wavelet coefficients and the test statistic. The proposed test statistic is expected to display high power when the true spectral density has significant spatial inhomogeneity, such as seasonal or cycle periodicities often encountered in economic and financial time series. A clear advantage of the proposed test statistic is that it is completely automatic, or adaptive, which avoids the need to select smoothing parameters or finest scales. We study the asymptotic distributions of the wavelet coefficients and the asymptotic distribution of the test statistic is also investigated. That problem was also considered by Duchesne, Li and Vandermeerschen (2010), but the results were stated without proof. Here, detailed proofs are provided, which are useful in their own right. As for the test statistics based on thresholding rules, the convergence of the test statistic based on Fan's approach toward its asymptotic distribution is expected to be slow. Thus, a Monte Carlo method is applied in order to find the critical values. Empirical evidence confirms that the proposed test statistic has reasonable properties under the null hypothesis and it displays high power under a large number of alternatives.

The organization of the dissertation is as follows. In Chapter 1, we introduce the basic framework including the introduction of the serial correlation, the wavelet analysis, and Fan's adaptive Neyman approach. And then we discuss how they can be used to develop the new testing procedure for serial correlation in a time series framework. The asymptotic distributions of the wavelet coefficients under the null hypothesis of no serial correlation is studied. We also provide the consistency of the proposed test statistic under fixed alternatives. Chapter 2 presents a small simulation study under the null hypothesis and for several alternative hypotheses. We demonstrate empirically that the proposed wavelet-based adaptive test statistic is powerful compared to current spectral-based test statistics. All computations were done using the R statistical software version 2.15.0 (http://cran.r-project.org/). Related scripts can be found in the Appendix. Chapter 3 offers some concluding remarks and Chapter 4 provides the proofs of the main results.

CHAPTER I

PRELIMINARIES AND THE TESTING PROBLEM

1.1 Serial Correlation and the Testing Problem

Serial correlation is also known as autocorrelation. It refers to the correlation of a time series with its own past and future values. Serial correlation has many applications in various fields. In signal processing, serial correlation can give information about repeating events like musical beats (for example, to determine tempo) or pulsar frequencies. It can also be used to estimate the pitch of a musical tone. In statistics, spatial autocorrelation between sample locations also helps one estimate mean value uncertainties when sampling a heterogeneous population. In Astrophysics, autocorrelation is used to study and characterize the spatial distribution of galaxies in the universe and in multi-wavelength observations of Low Mass X-ray Binaries.

Let $X = \{X_t,\ t \in \mathbb{Z}\}$ be a covariance stationary real-valued time series with normalized spectral density $f_X(w)$, $w \in [-\pi, \pi]$. Assuming $\sum_{h=-\infty}^{\infty} |R_X(h)| < \infty$, where the lag-$h$ autocovariance is defined by $R_X(h) = \text{Cov}(X_t, X_{t-|h|})$, $h \in \mathbb{Z}$, the spectral density can be written as

$$f_X(w) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \rho_X(h) e^{-ihw}, \qquad w \in [-\pi, \pi],$$

where $\rho_X(h) = R_X(h)/R_X(0)$ denotes the lag-h autocorrelation.

The hypothesis of interest states that the stochastic process X corresponds to a white noise process, against the alternative hypothesis of serial correlation of arbitrary form. More precisely, the null and alternative hypotheses in the time domain can be written as:

$$\begin{aligned} H_0 &: \rho_X(h) = 0, \qquad \text{for all } h,\ h \neq 0, \\ H_1 &: \rho_X(h) \neq 0, \qquad \text{for some } h,\ h \neq 0. \end{aligned}$$

The hypotheses H_0 and H_1 can be formulated using the spectral density $f_X(w)$ of X. Under the null hypothesis H_0, all $\rho_X(h) = 0$ for $h \neq 0$ and $\rho_X(h) = 1$ for $h = 0$. As a result,

$$f_X(\omega) = \frac{1}{2\pi}\left(\sum_{h\neq 0}\rho_X(h)e^{-ih\omega} + \rho_X(0)e^{-i0\omega}\right) = \frac{1}{2\pi}.$$

Hence, the null hypotheis $\rho_X(h) = 0$, for all $h \neq 0$ is equivalent to $f_X(\omega) = \frac{1}{2\pi}$, $\omega \in [-\pi, \pi]$. However, under the alternative hypothesis of serial correlation of arbitrary form, the spectral density $f_X(w)$ is not identically equal to the constant $(2\pi)^{-1}$. That alternative formulation in the frequency domain provides the main motivation to develop a test statistic for serial correlation using a spectral approach. Therefore, the original hypotheses of interest can be stated in terms of the normalized spectral density function $f_X(\omega)$ as follows:

$$\begin{aligned} H_0 &: f_X(\omega) = \frac{1}{2\pi}, \qquad \text{for any } \omega \in [-\pi, \pi], \\ H_1 &: f_X(\omega) \neq \frac{1}{2\pi}, \qquad \text{for some } \omega \in [-\pi, \pi]. \end{aligned}$$

It is possible to express the normalized spectral density function $f_X(w)$ using a wavelet basis (Lee and Hong, 2001). We now consider a wavelet representation of the normalized spectral density function $f_X(w)$.

1.2 Wavelet Analysis

Wavelet theory is applied in many disciplines: statistics, mathematics, geophysics, astronomy, signal processing, medical imaging, and numerical analysis. From a historical point of view, wavelet analysis is a relatively new method, given the fact that its mathematical foundation dates back to Fourier analysis in the nineteenth century. Fourier analysis is a methodology for the frequency domain while wavelet analysis is for both the frequency domain and time domain. The first mentioning of wavelets was in a thesis by Alfred Haar in 1910. Haar showed that any continuous function $f(x)$ on $[0, 1]$ can be approximated by a set of wavelet base using the Haar wavelet, which has the property of being compactly supported. In the 1930s, prototypes of wavelets first appeared in Lusin's work. In the 1980s, Grossman and Morlet, a physicist and an engineer, broadly defined wavelets in the context of quantum physics. In the mid-1980s, Mallat gave wavelets an additional jump-start through his work in digital signal processing. Inspired by Mallat's results, Meyer (1985) constructed the first non-trivial wavelets. Unlike the Haar wavelets, the Meyer wavelets are continuously differentiable. However they do not have compact support. Several years later, Daubechies (1988) constructed a set of wavelet orthonormal basis functions which have become the cornerstone of wavelet applications today.

Wavelet analysis can be viewed as a generalization of Fourier analysis. The two mathematical techniques are often compared with each other and the main difference is that wavelet analysis is localized in both time and frequency whereas Fourier analysis is only localized in frequency. Wavelets have a gender: father wavelets ϕ and mother wavelets ψ which satisfy:

$$\int \phi(x)\, dx = 1, \qquad \int \psi(x)\, dx = 0.$$

Father wavelets are good at representing the smooth and low-frequency parts of a signal and mother wavelets are good at representing the detail and high-frequency parts of a signal. A complete orthonormal wavelet basis $\{\phi_{jk}(\cdot), \{\psi_{jk}(\cdot)\}$ of the $L^2(\mathcal{R})$ space can be generated from the father and mother wavelets as follows:

$$\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k),$$
$$\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),$$

where the integer j denotes a resolution level and k denotes a translation parameter.

Now we consider a wavelet expansion of the normalized spectral density function $f_X(\omega)$, $\omega \in [-\pi, \pi]$. Since $f_X(w)$ is a 2π-periodic function over $\mathbb{R}$, a wavelet basis $\{\Phi_{jk}(\cdot),\ \Psi_{jk}(\cdot)\}$ for the $L_2(\Pi)$-space of 2π-periodic functions needs to be constructed, where $\Pi = [-\pi,\ \pi]$. Given an orthonormal wavelet basis $\{\phi_{jk}(\cdot),\ \psi_{jk}(\cdot)\}$ of $L_2(\mathbb{R})$, we can construct the 2π-periodic orthonormal wavelet basis $\{\Phi_{jk}(\cdot),\ \Psi_{jk}(\cdot)\}$ from $\{\phi_{jk}(\cdot),\ \psi_{jk}(\cdot)\}$ via the expressions:

$$\Phi_{jk}(\omega) = (2\pi)^{-1/2} \sum_{m=-\infty}^{\infty} \phi_{jk}\Big(\frac{\omega}{2\pi} + m\Big),$$
$$\Psi_{jk}(\omega) = (2\pi)^{-1/2} \sum_{m=-\infty}^{\infty} \psi_{jk}\Big(\frac{\omega}{2\pi} + m\Big),$$

where $-\infty < w < \infty$. Both $\Phi_{jk}(\cdot)$ and $\Psi_{jk}(\cdot)$ are real valued and periodic functions with period 2π. An example is the Haar wavelets ϕ and ψ, which are defined as:

$$\phi(x) = \begin{cases} 1, & x \in [0,1), \\ 0, & \text{otherwise.} \end{cases}$$

$$\psi(x) = \begin{cases} 1, & x \in [0,1/2), \\ -1, & x \in [1/2,1). \\ 0, & \text{otherwise.} \end{cases}$$

Other compactly supported wavelets and their properties are given in Vidakovic (1999) and

Daubechies (1992), among others. Haar wavelets are going to be used to construct our proposed test statistic W_{AN}.

For later use, the Fourier transformations and inverse Fourier transformations for several functions are defined here:

$$
\begin{aligned}
\hat{\phi}(z) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \phi(x) e^{-izx}\, dx, \\
\hat{\psi}(z) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \psi(x) e^{-izx}\, dx, \\
\hat{\phi}_{jk}(h) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \phi_{jk}(\omega) e^{-i\omega h}\, d\omega, \\
\hat{\psi}_{jk}(h) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \psi_{jk}(\omega) e^{-i\omega h}\, d\omega, \\
\hat{\Phi}_{jk}(h) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \Phi_{jk}(\omega) e^{-i\omega h}\, d\omega, \\
\hat{\Psi}_{jk}(h) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} \Psi_{jk}(\omega) e^{-i\omega h}\, d\omega, \\
\Phi_{jk}(\omega) &= (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} \hat{\Phi}_{jk}(h) e^{i\omega h}, \\
\Psi_{jk}(\omega) &= (2\pi)^{-1/2} \sum_{h=-\infty}^{\infty} \hat{\Psi}_{jk}(h) e^{i\omega h},
\end{aligned}
$$

where

$$
\hat{\Phi}_{jk}(h) = (2\pi)^{1/2} \hat{\phi}_{jk}(2\pi h) = \left(\frac{2\pi}{2^j}\right)^{1/2} e^{-i2\pi hk/2^j} \hat{\phi}\left(\frac{2\pi h}{2^j}\right),
$$

$$
\hat{\Psi}_{jk}(h) = (2\pi)^{1/2} \hat{\psi}_{jk}(2\pi h) = \left(\frac{2\pi}{2^j}\right)^{1/2} e^{-i2\pi hk/2^j} \hat{\psi}\left(\frac{2\pi h}{2^j}\right).
$$

Because a periodic wavelet basis $\{\Phi_{jk}(\cdot), \Psi_{jk}(\cdot)\}$ is used, the normalized spectral density function has the following wavelet expansion:

$$
f_X(w) = \beta_{00} \Phi_{00}(w) + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(w), \qquad w \in [-\pi, \pi],
$$

where $\beta_{00} = \int_{-\pi}^{\pi} f_X(w)\Phi_{00}(w)dw$ and $\alpha_{jk} = \int_{-\pi}^{\pi} f_X(w)\Psi_{jk}(w)dw$ for all $j \geq 0$ and $-\infty < k < \infty$. Note that the wavelet coefficients $\alpha_{j,k}$ are periodic with period 2^j, that is, $\alpha_{j,k} = \alpha_{j,2^j l+k}$ for all j, k and integers l. This explains why the summation over k is from 0 to $2^j - 1$.

Since $\sum_{m=-\infty}^{\infty} \phi_{00}(w+m) = 1$ for all w, we have $\Phi_{00}(w) = (2\pi)^{-1/2}$ for all $w \in [-\pi, \pi]$. Thus we have $\beta_{00} = (2\pi)^{-1/2}$. Therefore the normalized spectral density function can be written as:

$$f_X(w) = (2\pi)^{-1} + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}\Psi_{jk}(w), \qquad w \in [-\pi, \pi].$$

Under the null hypothesis H_0, $f_X(w) =: f_{X0}(w) = (2\pi)^{-1}$, $w \in [-\pi, \pi]$. Thus, we have

$$\alpha_{jk} = \int_{-\pi}^{\pi} f_X(w)\Psi_{jk}(w)dw = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_{jk}(w)dw = 0,$$

for all $j \geq 0$, $k = 0, 1, \ldots, 2^j - 1$. Hence, the original hypotheses in our testing problem can be expressed using the wavelet coefficients α_{jk}, j, $k \in \mathbb{Z}$:

$$H_0 : \alpha_{jk} = 0, \text{ for all } j \text{ and } k,$$

$$H_1 : \alpha_{jk} \neq 0, \text{ for at least one couple } (j, k).$$

Since

$$\begin{aligned}
\alpha_{jk} &= \int_{-\pi}^{\pi} f_X(w)\Psi_{jk}(w)dw \\
&= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \rho_X(h)\, e^{-ihw}\, \Psi_{jk}(w)dw \\
&= \frac{1}{\sqrt{2\pi}} \sum_{h=-\infty}^{\infty} \rho_X(h) \cdot \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \Psi_{jk}(w)\, e^{-ihw} dw \\
&= \frac{1}{\sqrt{2\pi}} \sum_{h=-\infty}^{\infty} \rho_X(h)\hat{\Psi}_{jk}(h) \\
&= \sum_{h=-\infty}^{\infty} \rho_X(h)\hat{\psi}_{jk}(2\pi h),
\end{aligned}$$

where $\hat{\Psi}_{jk}(\cdot)$ and $\hat{\psi}_{jk}(\cdot)$ are the Fourier transformations of $\Psi_{jk}(\cdot)$ and $\psi_{jk}(\cdot)$. A natural consistent estimator for α_{jk} is given by:

$$\hat{\alpha}_{jk} = \sum_{h=-(n-1)}^{n-1} \hat{\rho}_X(h)\hat{\psi}_{jk}(2\pi h) = \sum_{h=1}^{n-1} \hat{\rho}_X(h)\Big[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\Big],$$

in which the second equality can be derived by using the property $\hat{\psi}_{jk}(0) = 0$, for all $j = 0, 1, \ldots, J$ and $k = 0, 1, \ldots, 2^j - 1$, where J satisfies $2^{J+1} = n$ and $\hat{R}_X(h) = n^{-1}\sum_{t=|h|+1}^{n}(X_t - \bar{X})(X_{t-|h|} - \bar{X})$, $\hat{\rho}_X(h) = \hat{R}_X(h)/\hat{R}_X(0)$, $\bar{X} = n^{-1}\sum_{t=1}^{n} X_t$. And a wavelet-based estimator for the spectral density f_X can be expressed as:

$$\hat{f}_X^{(J)}(w) = (2\pi)^{-1} + \sum_{j=0}^{J}\sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}\Psi_{jk}(w), \qquad w \in [-\pi, \pi].$$

For the above wavelet coefficients α_{jk} and empirical wavelet coefficients $\hat{\alpha}_{jk}$, we have the following properties.

Theorem 1. *If the time series $X = \{X_t,\ t \in \mathbb{Z}\}$ is second-order stationary, the wavelet coefficients α_{jk} corresponding to the Haar wavelet ψ satisfy, for all $j = 1, 2, \ldots,$ and $k_1, k_2 = 0, 1, \cdots, 2^j - 1$,*

$$\alpha_{00} = 0 \quad \textit{and} \quad \alpha_{jk_1} = -\,\alpha_{jk_2}, \quad \textit{if} \quad k_1 + k_2 = 2^j - 1.$$

Similarly, the empirical wavelet coefficients $\hat{\alpha}_{jk}$ corresponding to the Haar wavelet ψ satisfy, for all $j = 1, 2, \ldots, J$ and $k_1, k_2 = 0, 1, \cdots, 2^j - 1$,

$$\hat{\alpha}_{00} = 0 \quad \textit{and} \quad \hat{\alpha}_{jk_1} = -\,\hat{\alpha}_{jk_2}, \quad \textit{if} \quad k_1 + k_2 = 2^j - 1.$$

The result was stated in Duchesne, Li and Vandermeerschen (2010) without proof. This dissertation provides the proof in Chapter IV. From Theorem 1, at most half of the empirical wavelet coefficients $\hat{\alpha}_{jk},\ k = 0, 1, \dots,$ $2^{j-1} - 1$ are needed to construct the test statistic, at each resolution level j, $j = 1, 2, \dots, J$. When one uses wavelets such as Haar, Franklin and second-order spline wavelets, the first coefficient $\hat{\alpha}_{00}$ could also be dropped since $\hat{\alpha}_{00} = 0$. See Lee and Hong (2001) and Duchesne, Li and Vandermeerschen (2010) for additional details.

In order to derive the null limit distribution of the empirical wavelet coefficients, we suppose the following assumption.

Assumption 1. *The stochastic process $X = \{X_t,\ t \in \mathbb{Z}\}$ is independent and identically distributed with $E(X_t) = \mu$, $E(X_t - \mu)^2 = \sigma^2$ and $E(X_t - \mu)^4 = \mu_4 < \infty$. A random sample $\{X_t\}_{t=1}^{n}$ of size $n \in \mathbb{Z}^+$ is observed.*

Assumption 1 was also assumed in Lee and Hong (2001). It allows for non-Gaussian processes which are common for economic and financial time series. For the empirical wavelet coefficients, we have the following asymptotic distributions. The proof is given in Chapter IV.

Theorem 2. *Under Assumption 1, half of the empirical wavelet coefficients $\hat{\alpha}_{jk}$ converge toward normal distribution asymptotically. Furthermore, they are asymptotically uncorrelated. More precisely, under Assumption 1, we have, as $n \to \infty$,*

$$(2\pi n)^{1/2}\,\hat{\alpha}_{jk} \longrightarrow_d \mathcal{N}(0,\ 1), \quad \textit{for all } j = 1, 2, \dots, J,\ k = 0, 1, \dots, 2^{j-1} - 1,$$

$$Cov\,(\hat{\alpha}_{j_1 k_1},\ \hat{\alpha}_{j_2 k_2}) = o(n^{-1}), \quad \textit{for all } j_1 \neq j_2 \textit{ or } k_1 \neq k_2,$$

where $j = 1, 2, \dots, J,\ k_1 = 0, 1, \dots, 2^{j_1 - 1} - 1,\ k_2 = 0, 1, \dots, 2^{j_2 - 1} - 1$.

The next theorem states that any finite-dimensional subset of the empirical wavelet coefficients $\hat{\alpha}_{jk}$ converge jointly toward a multivariate normal distribution asymptotically. More precisely, let

$$\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_{10}, \hat{\alpha}_{20}, \hat{\alpha}_{21}, \hat{\alpha}_{30}, \cdots, \hat{\alpha}_{33}, \hat{\alpha}_{40}, \hat{\alpha}_{41}, \cdots, \hat{\alpha}_{47}, \cdots, \hat{\alpha}_{\tilde{J}0}, \hat{\alpha}_{\tilde{J}1}, \cdots, \hat{\alpha}_{\tilde{J}2^{\tilde{J}-1}-1})^{\mathrm{T}}.$$

Then we have the following result:

Theorem 3. *Under Assumption 1, for any fixed* $\tilde{J}$ *such that* $1 \leq \tilde{J} < J$*, it follows, as* $n \to \infty$*:*

$$(2\pi n)^{1/2}\, \hat{\boldsymbol{\alpha}} \longrightarrow_d \mathcal{N}\left(0,\ \mathbf{I}_{(2^{\tilde{J}}-1)\times(2^{\tilde{J}}-1)}\right),$$

where $\mathbf{I}_{n\times n}$ *corresponds to the* $n \times n$ *identity matrix.*

Theorems 2 and 3 are related to a result stated in Duchesne, Li and Vandermeerschen (2010). A detailed prooof is provided in Chapter IV.

1.3 Fan's Adaptive Neyman Method

Fan (1996) considered the following canonical high dimensional testing problem: Let $\mathbf{X} \sim \mathcal{N}(\theta, \mathbf{I}_{n\times n})$ be an n-dimensional normal random vector. Consider the classical location testing problem:

$$H_0 : \theta = \mathbf{0} \qquad \text{versus} \qquad H_1 : \theta \neq \mathbf{0}.$$

Given a general alternative $H_1 : \theta = \theta_0 \neq \mathbf{0}$, we can use the Neyman-Pearson fundamental theorem to find the test statistic $\theta_0^\top \mathbf{X}$. We reject H_0 when $\theta_0^\top \mathbf{X}$ is too large. However we do not know the value of θ_0. So Fan used $\mathbf{X}$ to estimate θ_0, and constructed the test statistic $\|\mathbf{X}\|^2 = \sum_{i=1}^n X_i^2$.

Given the significance level α, we first compute the critical value c based on the test statistic $\sum_{i=1}^n X_i^2$.

Under H_0,

$$\sum_{i=1}^n X_i^2 \overset{H_0}{\sim} \chi^2(n) \xrightarrow{d} \mathcal{N}(n, 2n),$$

i.e., the test statistic $\sum_{i=1}^n X_i^2$, under H_0, follows a chi-squre distribution with degrees of freedom n, which could be approximated by a normal distribution with mean n and variance $2n$ for large n. Hence we compute the critical value c as below:

$$\begin{aligned}
\alpha &= P(\text{reject } H_0 | H_0) \\
&= P\Big(\sum_{i=1}^n X_i^2 > c \mid \sum_{i=1}^n X_i^2 \approx \mathcal{N}(n, 2n)\Big) \\
&= P\left(\frac{\sum_{i=1}^n X_i^2 - n}{\sqrt{2n}} > \frac{c-n}{\sqrt{2n}}\right) \\
&\approx 1 - \Phi\left(\frac{c-n}{\sqrt{2n}}\right).
\end{aligned}$$

Thus the critical value of the test statistic is $c = n + \sqrt{2n}Z_{1-\alpha}$. And then we compute the power of the test under the alternative hypothesis.

Under H_1,

$$\sum_{i=1}^{n} X_i^2 \longrightarrow_d N\left(E\left(\sum_{i=1}^{n} X_i^2\right), Var\left(\sum_{i=1}^{n} X_i^2\right)\right),$$

i.e., the test statistic $\sum_{i=1}^{n} X_i^2$, under H_1, asymptotically follows a normal distribution with mean $E\left(\sum_{i=1}^{n} X_i^2\right)$ and variance $Var\left(\sum_{i=1}^{n} X_i^2\right)$, which can be shown by the Lindeberg-Feller Theorem. Also one can easily derive that $E\left(\sum_{i=1}^{n} X_i^2\right) = n + ||\theta_0||^2$ and $Var\left(\sum_{i=1}^{n} X_i^2\right) = 2n + 4||\theta_0||^2$. Thus the power of the test is computed as follows:

$$\begin{aligned}
\text{power} &= P(\text{reject } H_0|H_1) \\
&= P\left(\sum_{i=1}^{n} X_i^2 > c \mid \sum_{i=1}^{n} X_i^2 \longrightarrow_d \mathcal{N}(n + ||\theta_0||^2,\ 2n + 4||\theta_0||^2)\right) \\
&= P\left(\frac{\sum_{i=1}^{n} X_i^2 - n - ||\theta_0||^2}{\sqrt{2n + 4||\theta_0||^2}} > \frac{n + \sqrt{2n}Z_{1-\alpha} - n - ||\theta_0||^2}{\sqrt{2n + 4||\theta_0||^2}}\right) \\
&\approx 1 - \Phi\left(\frac{Z_{1-\alpha} - \frac{||\theta_0||^2}{\sqrt{2n}}}{\sqrt{1 + \frac{2||\theta_0||^2}{n}}}\right) \\
&\approx 1 - \Phi\left(Z_{1-\alpha} - \frac{||\theta_0||^2}{\sqrt{2n}}\right) \\
&\approx 1 - \Phi\left(Z_{1-\alpha}\right) \\
&= 1 - (1 - \alpha) \\
&= \alpha,
\end{aligned}$$

provided that $||\theta_0||^2 = o(\sqrt{n})$. As one can see, the power of the test tends to α. So Fan argued that testing on all the n dimensions is not a good idea. Neyman (1937) proposed testing on the first m-dimensional sub-space, leading to the test statistic $\sum_{i=1}^{m} X_i^2$. Based on the power consideration, Fan proposed an adaptive Neyman test statistic T_{AN}^* which is to maximize the power of the test. That means:

$$\max_{1\le m\le n}\left\{1-\Phi\left(Z_{1-\alpha}-\frac{\sum_{i=1}^{m}\theta_{0i}^2}{\sqrt{2m}}\right)\right\} \;\Rightarrow\; \min_{1\le m\le n}\Phi\left(Z_{1-\alpha}-\frac{\sum_{i=1}^{m}\theta_{0i}^2}{\sqrt{2m}}\right)$$
$$\Rightarrow\; \max_{1\le m\le n}\left\{\frac{\sum_{i=1}^{m}\theta_{0i}^2}{\sqrt{2m}}\right\}$$
$$\Rightarrow\; \max_{1\le m\le n}\left\{\frac{\sum_{i=1}^{m}X_i^2-m}{\sqrt{2m}}\right\},$$

noting that $(2m)^{-1/2}\left(\sum_{i=1}^{m}X_i^2-m\right)$ is an unbiased estimator of $(2m)^{-1/2}\sum_{i=1}^{m}\theta_{0i}^2$.

Therefore Fan's adaptive Neyman test statistic T_{AN}^* was constructed to be:

$$T_{AN}^* = \max_{1\le m\le n}\frac{1}{\sqrt{2m}}\sum_{i=1}^{m}(X_i^2-1).$$

Large values of the above test T_{AN}^* result in rejection of the null hypothesis $H_0:\theta=\mathbf{0}$.

Using results of Darlin and Erdös (1956), T_{AN}^* can be normalized as:

$$T=\sqrt{2\log\log(n)}\,T_{AN}^*-\left[2\log\log(n)+0.5\log\log\log(n)-0.5\log(4\pi)\right],$$

which converges asymptotically to the following distribution under H_0:

$$P_{H_0}(T<x)\to\exp\{-\exp(-x)\},\quad \text{as } n\to\infty.$$

1.4 The Construction of the Test Statistic W_{AN}

In section 1.1, we concluded that for a stochastic process X, the original hypotheses in our testing problem

$$
\begin{aligned}
H_0 &: \rho_X(h) = 0, \qquad \text{for all } h,\ h \neq 0,\\
H_1 &: \rho_X(h) \neq 0, \qquad \text{for some } h,\ h \neq 0,
\end{aligned}
$$

can be expressed using the wavelet coefficients $\alpha_{jk},\ j,\ k \in \mathbb{Z}$:

$$
\begin{aligned}
&H_0 : \alpha_{jk} = 0, \text{ for all } j \text{ and } k,\\
&H_1 : \alpha_{jk} \neq 0, \text{ for at least one couple } (j, k).
\end{aligned}
$$

To construct the new testing procedure, noting that $\alpha_{00} = 0$ and $\hat{\alpha}_{00} = 0$, we consider the quadratic distance measure between the wavelet-based spectral density representation $f_X(\omega)$ and the null spectral density $f_{X0}(\omega) = (2\pi)^{-1}$:

$$
\begin{aligned}
Q(f_X, f_{X0}) &= \int_{-\pi}^{\pi} \{f_X(\omega) - f_{X0}(\omega)\}^2 \, dx\\
&= \int_{-\pi}^{\pi} \Big\{\frac{1}{2\pi} + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(\omega) - \frac{1}{2\pi}\Big\}^2 dx\\
&= \int_{-\pi}^{\pi} \Big\{\sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(\omega)\Big\}^2 dx\\
&= \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2 > 0.
\end{aligned}
$$

The last equality comes from the orthonormality property of the 2π-periodic wavelet basis $\big\{\Phi_{jk}(\cdot), \Psi_{jk}(\cdot)\big\}$.

Based on a suitable J_n, a natural estimator of that quadratic distance relies on the expression:

$$Q(\hat{f}_X^{(J_n)}, f_{X0}) = \int_{-\pi}^{\pi} \{\hat{f}_X^{(J_n)}(w) - f_{X0}(w)\}^2 \, dw = \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}^2.$$

From Theorems 2 and 3, it appears reasonable to propose the following test statistic V_n for our hypothesis testing problem:

$$V_n = 2\pi n \sum_{j=1}^{J} \sum_{k=0}^{2^{j-1}-1} \hat{\alpha}_{jk}^2.$$

Under the null hypothesis H_0, all the theoretical wavelet coefficients vanish, that is $\alpha_{jk} = 0$. Since the empirical wavelet coefficients $\hat{\alpha}_{jk}$'s are consistent estimators of the α_{jk}'s, the test statistic V_n is expected to reject the null hypothesis H_0 when it is too large. From Theorem 3, for any fixed $\tilde{J}$ ($1 \leq \tilde{J} < J$), we obtain the following corollary.

Corollary. *Under Assumption 1, for any fixed* $\tilde{J}$ *such that* $1 \leq \tilde{J} < J$, *we have, as* $n \to \infty$,

$$2\pi n \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \hat{\alpha}_{jk}^2 \longrightarrow_d \chi^2(2^{\tilde{J}} - 1).$$

Intuitively, the test statistic V_n can be interpreted as a Cramér-Von Mises test statistic, which measures the integrated mean squared error between the wavelet estimator $\hat{f}_X^{(J)}$ and the null spectral density f_{X0}. However, based on discussions presented in Fan (1996), the test statistic V_n is not expected to be powerful, the reason being that it involves too many individual terms (a total of $n/2-1$ terms or $n/2-1$ hypotheses). Thus, stochastic errors are accumulated, and therefore variations in the test statistic are too large. More precisely, Fan (1996) considered a canonical high dimensional testing problem. Let $\mathbf{X} \sim \mathcal{N}(\theta, \mathbf{I}_{n \times n})$ be an n-dimensional normal random vector. Consider the classical location testing problem:

$$H_0 : \theta = \mathbf{0} \quad \text{versus} \quad H_1 : \theta \neq \mathbf{0}.$$

Fan (1996) showed that the test statistic based on the norm of $\mathbf{X}$, that is $||\mathbf{X}||^2$, is expected to reach very low power for the general alternative $\theta = \theta_0 \neq \mathbf{0}$. The seminal work of Neyman (1937) proposed testing the first m-dimensional sub-space, leading to the test statistic $\sum_{i=1}^{m} X_i^2$, which relies however on the choice of m. Based on theoretical power consideration, Fan (1996) proposed an adaptive Neyman test statistic:

$$T_{AN}^* = \max_{1 \leq m \leq n} \frac{1}{\sqrt{2m}} \sum_{i=1}^{m} (X_i^2 - 1).$$

When large values of the test statistic T_{AN}^* are observed, the null hypothesis H_0 is rejected. With theoretical power calculation and empirical simulation studies, Fan (1996) showed that the adaptive Neyman test statistic reaches higher power than the Kolmogorov-Smirnov and Cramér-Von Mises test statistics. Using results from Darlin and Erdös (1956), it is possible to establish that the test statistic T_{AN}^* converges asymptotically toward the following limit distribution:

$$P_{H_0}(T_{AN}^* < x) \to \exp\{-\exp(-x)\}, \text{ as } n \to \infty,$$

under the null hypothesis H_0 in the location testing problem.

Although Fan (1996) considered hypothesis testing on an idealized statistical framework, that is an n-dimensional multinormal distribution, the general idea behind that methodology can be used in other testing problems as well. From Theorems 2 and 3, it appears that our problem is asymptotically equivalent to his testing problem, in the sense that from

Theorem 3, the random vector

$$\sqrt{2\pi n}\,(\hat{\alpha}_{10}, \hat{\alpha}_{20}, \hat{\alpha}_{21}, \hat{\alpha}_{30}, \ldots, \hat{\alpha}_{33}, \hat{\alpha}_{40}, \hat{\alpha}_{41}, \ldots, \hat{\alpha}_{47}, \ldots, \hat{\alpha}_{\tilde{J}0}, \hat{\alpha}_{\tilde{J}1}, \ldots, \hat{\alpha}_{\tilde{J}2^{\tilde{J}-1}-1}, \ldots)^{\top}$$

plays the role of $\mathbf{X}$ in a certain asymptotic sense. This kind of asymptotic approximation or equivalence has been used in nonparametric regression (see Härdle et al., 1998, p.202 and Donoho and Johnstone, 1998). Hence, it is reasonable to apply Fan's (1996) idea in our framework to motivate a new test statistic. For the sake of simpler exposition, let i such that $i = 2^{j-1} + k$, where $1 \leq j \leq J$, $0 \leq k < 2^{j-1}$. Thus with that numbering system $1 \leq i \leq N$, where $N = 2^{J-1} + 2^{J-1} - 1 = 2^J - 1 = n/2 - 1$, using the relation $2^{J+1} = n$. Denote

$$\theta = (\theta_1, \theta_2, \ldots, \theta_N)^{\top} = \sqrt{2\pi n}\,(\hat{\alpha}_{10}, \hat{\alpha}_{20}, \hat{\alpha}_{21}, \ldots, \hat{\alpha}_{J0}, \hat{\alpha}_{J1}, \ldots, \hat{\alpha}_{J2^{J-1}-1})^{\top}.$$

Using that notation, we propose a new wavelet-based adaptive Neyman test for serial correlation:

$$W_{AN}^{*} = \max_{1 \leq m \leq N} \frac{1}{\sqrt{2m}} \sum_{i=1}^{m} (\theta_i^2 - 1).$$

Following Fan (1996), the test statistic can be normalized as follows:

$$W_{AN} = \sqrt{2\log\log(N)}\, W_{AN}^{*} - [2\log\log(N) + .5\log\log\log(N) - .5\log(4\pi)].$$

Therefore, for the hypotheses:

$$H_0 : \alpha_{jk} = 0, \text{ for all } j \text{ and } k,$$

$$H_1 : \alpha_{jk} \neq 0, \text{ for at least one couple } (j, k),$$

the null hypothesis is rejected if large values of W_{AN} are observed. The approximate limit distribution

$$P_{H_0}(W_{AN} < x) \to \exp\{-\exp(-x)\}, \text{ as } n \to \infty,$$

can be used to determine the rejection region at a given significance level. For instance, at significance level α, the critical region is $W_{AN} > c_\alpha$, where $c_\alpha = -\log(-\log(1-\alpha))$. However, the above approximation is not expected to be satisfying in finite samples and the rate of convergence of the test statistic W_{AN} toward its asymptotic distribution is expected to be low.

As an illustration, we generated independent and identically distributed normal random variables with sample size $n = 256$ and performed $N = 10,000$ simulations. The distribution of the $10,000$ simulated test statistics under the null hypothesis is presented using a kernel density estimate. We used the common kernel density estimator defined as

$$\hat{f}_h(x) = N^{-1} \sum_{i=1}^{N} h^{-1} K\{(Z_i - x)/h\},$$

where K is the Gaussian kernel function $K(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, and h is the bandwidth. We used the value $h = 1.06 * s_N * N^{-1/5}$ as the bandwidth, where s_N is the standard deviation of $Z_1, Z_2, \ldots, Z_N$ (For more details on the choices of K and bandwidth h, see Chapter 5 of Fan and Yao, 2003). Taking the testing procedure W_{AN}, we calculated $Z_1 = W_{AN,1}$, $Z_2 = W_{AN,2}, \ldots, Z_{10,000} = W_{AN,10,000}$ from the $10,000$ simulations. The density estimator is presented in Figure 1 as dashed line. We also provided a simulated distribution for test statistic W_{AN} with sample size $n = 512$ and $N = 10,000$ presented as dotted line. The solid line represents the theoretical limit distribution. From the estimated distributions of our test statistic W_{AN}, one can see that the finite-sample distributions are not close to the theoretical limit distribution. When the sample size is $n = 256$, the 95-th quantile of the $10,000$ W_{AN}'s is $W_{0.05}(256) = 3.70$. When the sample size is $n = 512$, the 95-th quantile of the $10,000$ W_{AN}'s is $W_{0.05}(512) = 3.58$. Note that these critical values can be used to calculate the empirial powers of our test statistic under the alternative hypotheses.

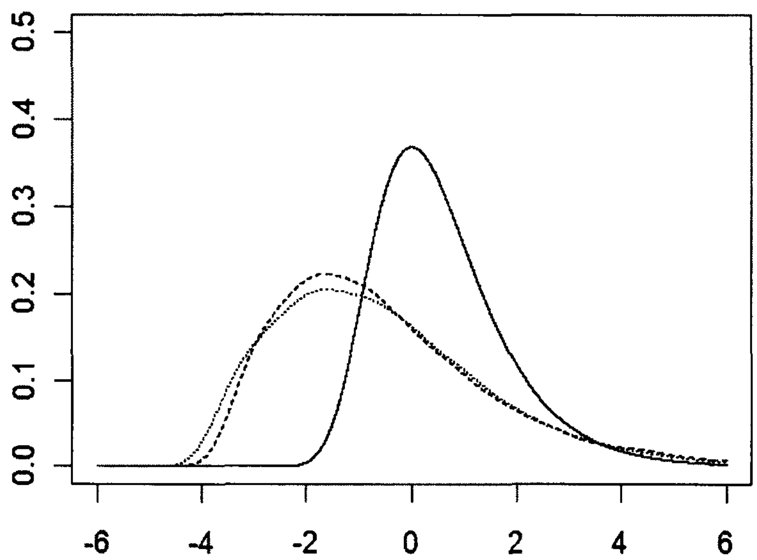

Figure 1: *The estimated density for test statistic W_{AN} under the null hypothesis for n=256 (dashed line) and n=512 (dotted line) based on 10,000 simulations. The solid line represents the theoretical limit distribution.*

Since the random vector $\theta = (\theta_1, \theta_2, \ldots, \theta_N)^\top$ asymptotically converges toward a multinormal distribution, but that the convergence of the test statistic W_{AN} seems to be slow, we propose to use Monte Carlo methods to determine the rejection region given the finite sample size (n) under H_0. We elaborate more on the Monte Carlo methods in the next chapter.

1.5 Consistency of the Test Statistic W_{AN}

In this section, we study the consistency of the proposed test statistic W_{AN}. Let $\kappa(j,k,l)$ be the fourth order cumulant of the joint distribution of $\{X_t, X_{t+j}, X_{t+k}, X_{t+l}\}$, where $j,k,l \in \mathbb{Z}$. For fixed indices j, k and l, it is defined as follows:

$$k(j,k,l) = E(X_t X_{t+j} X_{t+k} X_{t+l}) - E(\tilde{X}_t \tilde{X}_{t+j} \tilde{X}_{t+k} \tilde{X}_{t+l}),$$

where $\{\tilde{X}_t, t \in \mathbb{Z}\}$ represents a Gaussian stochastic process with the same mean and covariance function as $\{X_t, t \in \mathbb{Z}\}$. For more information on the properties of the cumulants, see, e.g., Hannan (1970) and Brillinger (1981). To study the behavior of the proposed test statistic W_{AN} under the alternative hypothesis H_1, we impose in Assumption 2 the temporal dependence of $\{X_t, t \in \mathbb{Z}\}$. The temporal dependence of $\{X_t, t \in \mathbb{Z}\}$ is supposed to satisfy Assumption 2.

Assumption 2. *It is assumed that* $\{X_t,\ t \in \mathbb{Z}\}$ *is a fourth order stationary process with autocovariance function satisfying* $\sum_{h=-\infty}^{\infty} R^2(h) < \infty$ *and such that the cumulants satisfy the following summability assumption:* $\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |\kappa(j,k,l)| < \infty$, *for all* $j,k,l \in \mathbb{Z}$.

The following result shows that our test statistic W_{AN} has an asymptotic power which tends to one at any fixed alternative which belongs to Assumption 2. More precisely, let $f_{X0} = (2\pi)^{-1} \in H_0$ and $f_X \in H_1$ be a spectral density function for time series $\{X_t, t \in \mathbb{Z}\}$ satisfying Assumption 2. A similar assumption has been supposed in Lee and Hong (2001). The following Theorem states the asymptotic power of the test statistic based on the critical region $W_{AN} > c_\alpha$, where $c_\alpha = -\log(-\log(1-\alpha))$.

Theorem 4. *The proposed test statistic W_{AN} has an asymptotic power at alternative $f_X \in H_1$ at least given by the following formula:*

$$P_{H_1}\left(2\pi\, Q(\hat{f}_X^{(J_n)}, f_{X0}) \geq \left[2^{J_n+1}n^{-1} + 2^{3/2}2^{J_n/2}n^{-1}\sqrt{2\log\log(n)}\right]\,(1+o(1))\right),$$

for any $1 \leq J_n \leq J$. In addition, let J_n be such that $J_n \to \infty$ with $2^{2J_n}/n \to 0$. Then $Q(\hat{f}_X^{(J_n)}, f_{X0}) \to Q(f, f_{X0}) > 0$ in probability.

From Theorem 4, the proposed test statistic has a power function which tends toward one, that is $P_{H_1}(W_{AN} > c_\alpha) \to 1$, when the finest scale is supposed to satisfy $J_n \to \infty$ with $2^{2J_n}/n \to 0$. The proof of the result is provided in Chapter IV.

CHAPTER II

SIMULATION RESULTS

In the previous chapter we introduced the new test statistic W_{AN} for serial correlation using Fan's (1996) adaptive Neyman approach. Here, we compare the test statistic W_{AN} with several current test statistics, which are introduced in Section 2.1. More precisely, the finite sample performance of several test statistics in terms of their empirical levels and powers are investigated. In Section 2.2 which is about the level study, we examined the empirical frequencies of rejection of the null hypothesis when it is in fact true. In Section 2.3 which is about the power study, we compute the empirical frequencies of rejection of the null hypothesis under several alternatives. The common $\alpha = 5\%$ significance level has been adopted and two sample sizes $n = 256$ and 512 are considered. All computations were done using scripts written in R 2.15.0. which can be found in the Appendix.

2.1 Several Current Test Statistics

2.1.1 Test Statistic Q_m

The classical Ljung-Box test statistic, denoted as Q_m, is included in the experiments. Consider a time series $\{X_t\}_{t=1}^n$ generated by a ARMA(p, q) model, written as

$$\phi(B)X_t = \theta(B)\epsilon_t,$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p, \quad \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, $B^k X_t = X_{t-k}, \quad \epsilon_t \overset{iid}{\sim} N(0,1)$.

After a model of this form has been fitted to the data, the residuals of the model, written as $\{\hat{\epsilon}_t\}_{t=1}^n$, are examined. If the fit is appropriate, the residuals should be white noise. So the hypotheses of interest are:

$$H_0 \;:\; \{\hat{\epsilon}_t\}_{t=1}^n \text{ are white noise,}$$
$$H_1 \;:\; \{\hat{\epsilon}_t\}_{t=1}^n \text{ are not white noise.}$$

Now consider their autocorrelations

$$\hat{\rho}(h) = \sum_{t=h+1}^{n} \hat{\epsilon}_t \hat{\epsilon}_{t-h} \Big/ \sum_{t=1}^{n} \hat{\epsilon}_t^2, \quad h = 1, 2, \cdots, n-1.$$

Let $\rho = (\rho(1), \rho(2), \cdots, \rho(n-1))$ be the vector form of the theoretical autocorrelations, where

$$\rho(h) = \sum_{t=h+1}^{n} \epsilon_t \epsilon_{t-h} \Big/ \sum_{t=1}^{n} \epsilon_t^2, \quad h = 1, 2, \cdots, n-1.$$

According to the result of Anderson 1942; Anderson & Walker 1964, the limiting distribution of ρ is a multivariate normal distribution with mean vector **0** and $Var(\rho(h)) = \frac{n-h}{n(n+2)}$ and $Cov(\rho(h), \rho(k)) = 0$, for $h \neq k$.

Then the test statistic Q_m was constructed to be:

$$Q_m = n(n+2) \sum_{h=1}^{m} \frac{\hat{\rho}(h)^2}{n-h},$$

where m, fixed with respect to n and satisfying $1 \leq m \leq n-1$, is called the lag order. Under the null hypothesis, the test statistic Q_m converges in distribution to a chi-square distribution with degrees of freedom $m-p-q$. However in our experiments, we skipped the model fitting process and directly generate $\{\hat{\rho}(h)\}_{h=1}^{n}$ as white noise. Thus Q_m converges in distribution to a chi-square with degrees of freedom m under the null hypothesis of no serial correlation. We considered three choices for m: $m = 1, 2$ and 3 in the simulation studies. Low values of m are often recommended to detect low order dependence.

2.1.2 Test Statistic K_n

We also include the kernel-based test statistic of Hong (1996), denoted as K_n. This test statistic is based on the quadratic distance $Q(\hat{f}_X, f_{X0})$ between the spectral density estimator

$$\hat{f}_X(\omega) = (2\pi)^{-1} \sum_{h=-n+1}^{n-1} \kappa(h/p_n)\hat{\rho}_X(h)cos(h\omega), \quad \omega \in [-\pi, \pi],$$

and the null spectral density $f_{X0}(w) = (2\pi)^{-1}$. Then the kernel-based test statistic K_n is constructed to be the standardized version of $Q(\hat{f}_X, f_{X0})$ as:

$$K_n = \frac{n \sum_{h=1}^{n-1} \kappa^2(h/p_n)\hat{\rho}_X^2(h) - M_n(\kappa)}{\sqrt{2V_n(\kappa)}},$$

where $M_n(\kappa) = \sum_{h=1}^{n-1}(1-h/n)\kappa^2(h/p_n)$, $V_n(\kappa) = \sum_{h=1}^{n-2}(1-h/n)(1-(h+1)/n)\kappa^4(h/p_n)$,

$\kappa(\cdot)$ represents a kernel function and p_n denotes the smoothing parameter. According to Hong (1996), the choice of the kernel function has little impact on the size and power properties. For our Monte Carlo experiments, we choose the Daniell kernel defined as $\kappa(z) = \sin(\pi z)/(\pi z)$, $z \in (-\infty, \infty)$. However, the choice of p_n may have significant effects on the size and power. As in Hong (1996), we retain the same rates for p_n: (i) $p_n = [\log(n)]$, (ii) $p_n = [3n^{0.2}]$ and (iii) $p_n = [3n^{0.3}]$, where $[x]$ denotes the integer closest to the real number x. The rates deliver $p_n = 6, 9$ and 16 for $n = 256$; and 6, 10 and 19 for $n = 512$. Under the null hypothesis of no serial correlation, the test statistic K_n converges in distribution to a standard normal distribution.

2.1.3 Test Statistic W_n

We also include the wavelet-based test statistic of Lee and Hong (2001), denoted as W_n. Its construction is based on the distance estimator $\sum_{j=1}^{J} \sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}^2$, which is similar to the construction of our test statistic W_{AN}. The test statistic W_n is constructed by properly standardizing the distance estimator as:

$$W_n = \frac{2\pi n \sum_{j=1}^{J} \sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}^2 - (2^{J+1} - 1)}{4(2^{J+1} - 1)},$$

where $2^{J+1} - 1$ and $4(2^{J+1} - 1)$ are approximately the mean and variance of $\sum_{j=1}^{J} \sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}^2$ according to Lee and Hong (2001). J is the finest scale level which has significant impact on the performance of the test. We select $J = 2, 3$ and 4 for $n = 256$ and $n = 512$ in the simulation study. The test statistic W_n converges in distribution to a standard normal distribution under the null hypothesis for suitable choices of J.

2.1.4 Test Statistic T_n

Finally the test statistic using wavelet thresholding of Duchesne, Li and Vandermeerschen (2010), denoted as T_n, is included in the simulation study. Its construction is also based on the distance estimator $\sum_{j=1}^{J}\sum_{k=0}^{2^{j-1}-1}\hat{\alpha}_{jk}^2$. The idea is to shrink the empirical wavelet coefficients to 0 which are large enough such that $|\sqrt{2\pi n}\hat{\alpha}_{jk}| > \delta_n$, where δ_n is a thresholding parameter. This leads to the test statistic:

$$T_n = \frac{2\pi n \sum_{j=1}^{J}\sum_{k=0}^{2^{j-1}-1}\hat{\alpha}_{jk}^2 I\{|\sqrt{2\pi n}\hat{\alpha}_{jk}| > \delta_n\} - \mu_n}{\sigma_n},$$

where $\mu_n = (2\pi)^{-1/2}a_n^{-1}\delta_n(1+\delta_n^{-2})$, $\sigma_n^2 = (2\pi)^{-1/2}a_n^{-1}\delta_n^3(1+3\delta_n^{-2})$, $\delta_n = \{2\log((n/2)a_n)\}^{1/2}$, and $a_n = c\{\log(n/2)\}^{-d}$ for some positive constants c and d. Two combinations for (c, d) are included in the simulation study: $(c, d) = (1, 2)$ and $(c, d) = (1, 5/2)$ following the choices of Duchesne, Li and Vandermeerschen (2010). The test statistic T_n converges in distribution to a standard normal distribution under the null hypothesis.

2.2 Level Study

As discussed in Chapter 1, the theoretical limit distribution for the test statistic W_{AN} is not a satisfactory approximation for finite sample sizes. Given the modern computing resources, we propose to use Monte Carlo methods to find the critical values and the rejection regions for a given finite sample size n under H_0. Monte Carlo methods can also be used to calculate the empirical powers of the test statistics under a given alternative H_1.

We compute the empirical levels using the asymptotic critical values (denoted as ACV in the Tables) and the empirical critical values (denoted as ECV in the Tables). We illustrate the steps for Monte Carlo computation of the empirical levels using the asymptotic critical values as follows:

1. For a specific test statistic, find the ACV which is the 95-th quantile of the limiting distribution of the test statistic under the null hypothesis.
2. Generate a random sample $\{X_t\}_{t=1}^n$ under the null hypothesis, where n denotes the sample size and $X_t \sim \mathcal{N}(0, 1)$.
3. Compute the test statistic under the null hypothesis based on the random sample $\{X_t\}_{t=1}^n$ generated in step 2.
4. Repeat steps 2 and 3 for $N = 10,000$ times to derive $10,000$ test statistics under the null hypothesis.
5. Compute the level which is the percentage of the $10,000$ test statistics that are larger than the ACV.

To compute the number of rejections using the empirical critical values, the steps are largely similar. For example, for the test statistic W_{AN}, the only difference is that we used the empirical critical values $W_{0.05}(256) = 3.70$ and $W_{0.05}(512) = 3.58$ instead of the asymptotic critical values to compute the levels in step 5.

Table 1: *Level study.*

$\alpha = 5\%$		Q_m			K_n			W_n			T_n		W_{AN}
		$m=1$	$m=2$	$m=3$	$[\log(n)]$	$[3n^{0.2}]$	$[3n^{0.3}]$	$J=2$	$J=3$	$J=4$	$(1,2)$	$(1,\frac{5}{2})$	
$n=$	ACV	0.051	0.048	0.047	0.071	0.069	0.068	0.046	0.046	0.046	0.167	0.141	0.073
256	ECV	0.048	0.052	0.054	0.051	0.053	0.052	0.051	0.052	0.054	0.046	0.047	0.043
$n=$	ACV	0.048	0.048	0.050	0.073	0.072	0.070	0.047	0.050	0.047	0.217	0.175	0.071
512	ECV	0.049	0.048	0.050	0.048	0.050	0.049	0.050	0.053	0.053	0.047	0.048	0.049

Table 1 reports the results of the level study at significance level $\alpha = 5\%$. Based on the results presented in Table 1, the new test statistic W_{AN} displays reasonable levels using the empirical critical values. Using the theoretical limit distribution, W_{AN} exhibited some over-rejection. This finding is not surprising given the discussion in Chapter 1, and is in agreement with the results reported in Fan (1996), stating that the theoretical limit distribution for the test statistic W_{AN} is not a good approximation for finite samples, and that the convergence rate of W_{AN} toward its theoretical limit distribution is relatively slow. The test statistic T_n displayed large over-rejection for both $T_n(1,2)$ and $T_n(1,5/2)$ using the asymptotic critical values, but when using the empirical critical values, the levels are reasonable for both sample sizes. These conclusions are similar to those reported in Duchesne, Li and Vandermeerschen (2010). Note that only two methods are fully automatic: the wavelet-based test T_n using thresholding and the new test statistic W_{AN} using Fan's approach. If one decides to use asymptotic critical values and a fully automatic procedure, it appears preferable to use the new test statistic W_{AN}. The test statistic W_n exhibits a little under-rejection at levels when using the asymptotic critical values, and a little over-rejection at levels when using the empirical critical values for choices $J = 3$ and $J = 4$ for both sample sizes. The kernel-based test exhibits relatively small over-rejection at levels when using the asymptotic critical values, but satisfactory levels when using the empirical critical values. These findings are in line with previous results about the kernel-based test. The test statistic Q_m has reasonable levels when using both the asymptotic critical values and the empirical critical values for both sample sizes. It appears that the choice of m does not have observable impact on the levels. This is explained by the fact that a white noise process represents spatially homogeneous features. Consequently, including one, or two, or three autocorrelation terms should not have observable impact on the performance of the test statistic.

2.3 Power Study

Tables 2 and 3 report the results of the power study. As in the level study, empirical powers have been calculated using both the asymptotic critical values and the empirical critical values. The use of empirical critical values allows us to be able to compare the powers of all the test statistics on an equal basis. Seven models are included in the power analysis with several choices of the model constants (which are specified in the tables):

Model 1: $\text{AR}(1) : (1 - \phi B)X_t = a_t,$

Model 2: $\text{AR}(4) : (1 - \phi B^4)X_t = a_t,$

Model 3: $\text{ARMA}(1, 0) \times (1, 0)_{12} : (1 - \phi_1 B^{12})(1 - \phi_2 B)X_t = a_t,$

Model 4: $\text{ARMA}(0, 1) \times (1, 0)_{12} : (1 - \phi B^{12})X_t = (1 + \theta B)a_t,$

Model 5: $\text{ARMA}(0, 0) \times (1, 0)_{12} : (1 - \phi B^{12})X_t = a_t,$

Model 6: $\text{ARMA}(0, 0) \times (2, 0)_{12} : (1 - \phi_1 B^{12} - \phi_2 B^{24})X_t = a_t,$

Model 7: $\text{ARMA}(0, 0) \times (1, 1)_{12} : (1 - \phi B^{12})X_t = (1 + \theta B^{12})a_t,$

where $B^s X_t = X_{t-s}$, $s \geq 1$, and $\{a_t, t \in \mathbb{Z}\}$ corresponds to a Gaussian white noise. All the alternatives have been chosen based on the general shape of the theoretical spectral density function. Under Model 1, an AR(1) alternative is considered, and there is no peaks or spikes in the spectral density; that alternative shows spatially homogeneous features. For all the other alternatives, the spectral densities exhibit spatially inhomogeneous features. These models are motivated by seasonal time series models, which are quite common in real applications. Model 2 is a pure autoregressive seasonal time series model, which could be used for modelling quarterly data. Similarly, Models 5 and 6 are pure autoregressive seasonal time series model, which could be used for modelling montly data. Models 3 and 4 are seasonal ARIMA time series models, which include a pure seasonal factor and an additional factor to describe local characteristics. Finally, Model 7 include seasonal autoregressive and moving-average factors. The features of the theoretical spectral density functions of the seven models can be seen clearly from their spectral density plots below:

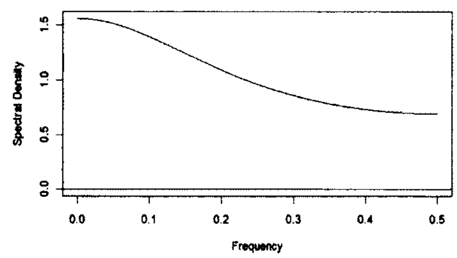

Figure 2: *The spectral density plot of model* $AR(1) : (1 - 0.2B)X_t = a_t$

Figure 2 shows that the spectral density function of model $\mathrm{AR}(1) : (1-0.2B)X_t = a_t$ offers no peaks or spikes which represents spatially regular features.

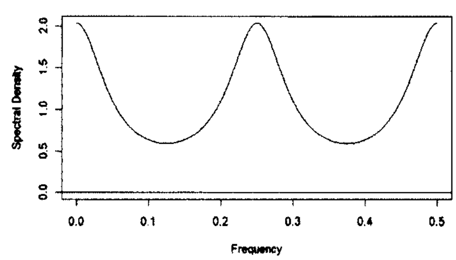

Figure 3: *The spectral density plot of model* $AR(4) : (1 - 0.3B^4)X_t = a_t$

Figure 3 shows that the spectral density function of model $\mathrm{AR}(4) : (1-0.3B^4)X_t = a_t$ offers moderate alternations.

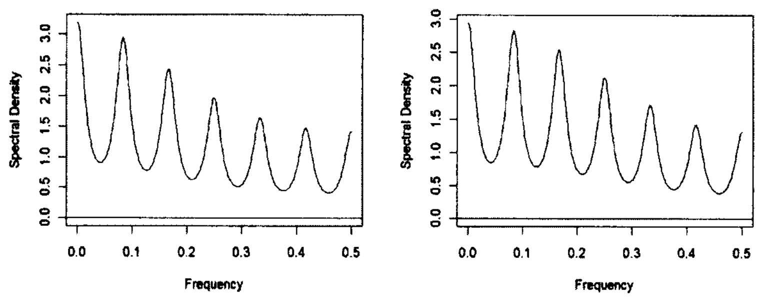

Figure 4: *The spectral density plots of seasonal model group I: plot on the left is the spectral density plot of ARMA*$(1,0) \times (1,0)_{12} : (1-0.3B^{12})(1-0.2B)X_t = a_t$*; plot on the right is the spectral density plot of ARMA*$(0,1) \times (1,0)_{12} : (1-0.3B^{12})X_t = (1+0.2B)a_t$.

Figure 4 shows that the spectral density functions of seasonal model group I offer large alternations at low frequencies.

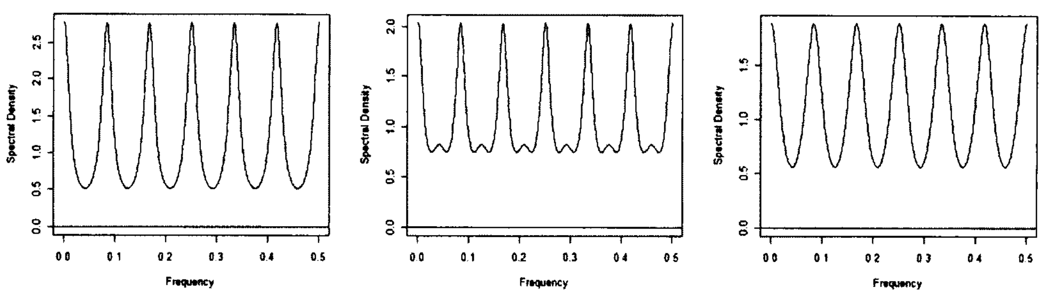

Figure 5: *The spectral density plots of seasonal model group II: plot on the left is the spectral density plot of ARMA*$(0,0) \times (1,0)_{12} : (1-0.4B^{12})X_t = a_t$*; plot in the middle is the spectral density plot of ARMA*$(0,0) \times (2,0)_{12} : (1-0.2B^{12}-0.1B^{24})X_t = a_t$*; plot on the right is the spectral density plot of ARMA*$(0,0) \times (1,1)_{12} : (1-0.2B^{12})X_t = (1+0.1B^{12})a_t$.

Figure 5 shows that the spectral density functions of seasonal model group II offer strong alternations overall.

To compute the empirical powers based on the asymptotic (empirical) critical values, the following steps have been implemented:

1. Generate a random sample $\{X_t\}_{t=1}^{n}$ under an alternative hypothesis, n being the sample size.

2. Compute the test statistic based on the random sample $\{X_t\}_{t=1}^{n}$ generated in step 1.

3. Repeat steps 1 and 2 for $N = 4000$ times to derive 4000 test statistics.

4. Compute the empirical power which is the percentage of the 4000 test statistics that are larger than the asymptotic (empirical) critical values.

Except for the test statistic T_n, the empirical powers calculated using the asymptotic critical values and the empirical critical values are reasonably close. That conclusion was excepted, since the empirical levels of T_n were not satisfying using the asymptotic critical values. Now we concentrate the discussion on the empirical powers using the empirical critical values.

Under Model 1, the spectral densities of the AR(1) processes offer spatially regular features and its spectral density offers no peaks or spikes. As expected, the test statistics Q_m and K_n are powerful in this particular situation; these two test statistics reach high power when the spectral density is relatively smooth. However, the choice of the smoothing parameter needs to be selected carefully, since the power decreases with p_n for K_n, and the power decreases with J in the case of the wavelet-based test W_n. Specifically, for W_n, the highest power is reached at $J = 2$ and the lowest power is reached at $J = 4$. Using wavelet thresholding was inefficient under that alternative and the test statistics T_n were inferior under both choices $(c, d) = (1, 2)$ and $(c, d) = (1, \frac{5}{2})$. This empirical finding is in agreement with the fact that T_n should exhibit high power in detecting sharp peaks and high frequency alternations; under the AR(1) alternative, the spectral density was very smooth. Interestingly, the adaptive test W_{AN} displayed high power. Without any subjective choice of the smoothing parameter or the finest scale, the empirical powers of W_{AN} were very similar to those of K_n with best p_n, or W_n with best J.

Under Model 2, seasonal AR(4) processes are simulated, and the spectral densities under these alternatives offer moderate alternations. The test statistic Q_m offers the lowest power among all the tests, which shows the inability to capture the important characteristic of the spectral density of AR(4). This is due to a too low value of the lag order m. Larger values of m are necessary to obtain larger power for this test statistic but the choice of m remains subjective. The test statistics W_n and K_n achieve the highest empirical powers under this alternative. For K_n, the choice $p_n = [3n^{0.3}]$ is optimal. For W_n, the finest scale $J = 3$ represents the optimal choice. The new test W_{AN} achieves very comparable power to the other spectral-based test statistics. Compared to the test T_n based on thresholding, the test statistic W_{AN} is much more powerful. Comparing the results presented in Tables 2 and 3, the empirical powers of W_{AN} improves substantially when the sample size increases from $n = 256$ to $n = 512$.

Under Models 3 and 4, stochastic processes ARMA$(1,0) \times (1,0)_{12}$ and ARMA$(0,1) \times (1,0)_{12}$ were simulated. For these alternatives, the spectral densities offer large alternations at low frequencies. Under these situations, the new adaptive test statistic W_{AN} delivers interesting power properties. When the sample size $n = 256$, the test W_{AN} offers better power than the test statistics T_n and K_n, and it offers comparable power to the highest powers of Q_m and W_n. For Q_m, the choice $m = 1$ is optimal. For W_n, the choice $J = 4$ is optimal. For the test statistic T_n, the choice $(c, d) = (1, 5/2)$ achieves better power than the choice $(c, d) = (1, 2)$. This is in agreement with theoretical results of Fan (1996): a smaller choice of a_n would improve the normal approximation of the test statistic, but more noise would pass in the thresholding process. When sample size increases to $n = 512$, W_{AN} achieves the best power among all the tests except W_n at choice $J = 4$. However the two highest powers are very similar.

Under Models 5, 6 and 7, stochastic processes ARMA$(0,0) \times (1,0)_{12}$, ARMA$(0,0) \times (2,0)_{12}$, and ARMA$(0,0) \times (1,1)_{12}$ were simulated. Under these alternatives the spectral densities offer strong alternations. When the sample size $n = 256$, the test statistics Q_m and K_n offer the lowest empirical powers. The adaptive test statistic W_{AN} achieves very comparable power to T_n at both choices $(c, d) = (1, 2)$ and $(c, d) = (1, 5/2)$. W_{AN} also achieves comparable power to W_n with best finest scale $J = 4$, and higher power than W_n at choices $J = 2$ and $J = 3$. When the sample size $n = 512$, the test statistic

W_{AN} exhibit high power, very comparable to the one of W_n with finest scale $J = 4$.

Overall, without any choice of the smoothing parameters or finest scales, the proposed test statistic W_{AN} offers very interesting power. Compared to the test statistic T_n of Duchesne, Li and Vandermeerschen (2010), the proposed test statistic W_{AN} seems to display better power properties than wavelet thresholding T_n: from our simulation experiments, Fan's adaptive approach delivers high power for a larger class of alternatives. From their experiments and those presented in this empirical study, wavelet thresholding T_n was not powerful if the spectral density did not offer bumps or alternations. From the simulation experiments presented in this dissertation, the adaptive test statistic W_{AN} was usually among the most powerful test statistics, without any need to select a smoothing parameter or a finest scale.

Table 2: *Power study for sample size* $n = 256$.

$n = 256$		Q_m			K_n			W_n			T_n		W_{AN}
		$m=1$	$m=2$	$m=3$	$[\log(n)]$	$[3n^{0.2}]$	$[3n^{0.3}]$	$J=2$	$J=3$	$J=4$	$\delta_n(1,2)$	$\delta_n(1,\frac{5}{2})$	
model 1	ACV	0.897	0.829	0.770	0.841	0.807	0.722	0.720	0.602	0.442	0.536	0.468	0.779
(0.2)	ECV	0.900	0.831	0.785	0.821	0.766	0.678	0.739	0.627	0.501	0.295	0.284	0.733
model 1	ACV	0.370	0.278	0.232	0.331	0.290	0.235	0.202	0.145	0.108	0.255	0.211	0.264
(0.1)	ECV	0.363	0.282	0.253	0.278	0.241	0.188	0.213	0.160	0.124	0.090	0.092	0.222
model 2	ACV	0.076	0.135	0.141	0.677	0.952	0.972	0.073	0.891	0.845	0.846	0.820	0.847
(0.3)	ECV	0.069	0.139	0.140	0.591	0.923	0.952	0.080	0.901	0.865	0.640	0.638	0.786
model 2	ACV	0.059	0.094	0.088	0.309	0.602	0.671	0.059	0.490	0.409	0.504	0.459	0.425
(0.2)	ECV	0.054	0.096	0.099	0.236	0.518	0.598	0.063	0.513	0.449	0.287	0.277	0.343
model 3	ACV	0.873	0.813	0.779	0.835	0.849	0.839	0.695	0.703	0.871	0.902	0.895	0.908
(0.3, 0.2)	ECV	0.874	0.817	0.774	0.802	0.810	0.795	0.711	0.720	0.889	0.765	0.779	0.872
model 3	ACV	0.365	0.296	0.265	0.327	0.336	0.328	0.204	0.232	0.362	0.529	0.499	0.426
(0.2, 0.1)	ECV	0.364	0.297	0.261	0.277	0.283	0.277	0.218	0.249	0.404	0.297	0.327	0.351
model 4	ACV	0.864	0.794	0.757	0.825	0.838	0.820	0.677	0.677	0.867	0.890	0.888	0.897
(0.3, 0.2)	ECV	0.860	0.790	0.747	0.793	0.805	0.790	0.690	0.698	0.886	0.735	0.767	0.857
model 4	ACV	0.367	0.288	0.260	0.339	0.343	0.334	0.204	0.236	0.367	0.533	0.500	0.396
(0.2, 0.1)	ECV	0.354	0.287	0.251	0.271	0.273	0.269	0.214	0.252	0.406	0.291	0.311	0.346
model 5	ACV	0.091	0.113	0.115	0.162	0.333	0.524	0.116	0.349	0.926	0.962	0.966	0.899
(0.4)	ECV	0.097	0.100	0.117	0.125	0.265	0.450	0.126	0.368	0.943	0.875	0.900	0.859
model 5	ACV	0.078	0.074	0.084	0.115	0.200	0.300	0.065	0.185	0.601	0.773	0.762	0.574
(0.3)	ECV	0.080	0.080	0.085	0.086	0.145	0.230	0.070	0.201	0.641	0.571	0.592	0.487
model 6	ACV	0.101	0.135	0.158	0.201	0.390	0.601	0.139	0.352	0.882	0.979	0.976	0.923
(0.3, 0.2)	ECV	0.118	0.134	0.162	0.167	0.340	0.545	0.148	0.375	0.898	0.936	0.943	0.890
model 6	ACV	0.067	0.075	0.078	0.108	0.154	0.231	0.067	0.128	0.364	0.613	0.585	0.370
(0.2, 0.1)	ECV	0.063	0.070	0.072	0.074	0.106	0.165	0.072	0.137	0.402	0.402	0.417	0.299
model 7	ACV	0.100	0.109	0.130	0.190	0.446	0.654	0.129	0.442	0.986	0.989	0.993	0.977
(0.3, 0.2)	ECV	0.104	0.124	0.133	0.158	0.347	0.576	0.138	0.459	0.990	0.956	0.970	0.962
model 7	ACV	0.066	0.072	0.076	0.109	0.172	0.254	0.073	0.191	0.589	0.732	0.727	0.530
(0.2, 0.1)	ECV	0.067	0.079	0.084	0.085	0.144	0.228	0.086	0.206	0.626	0.505	0.533	0.453

Table 3: *Power study for sample size $n = 512$.*

$n = 512$		Q_m			K_n			W_n			T_n		W_{AN}
		$m=1$	$m=2$	$m=3$	$[\log(n)]$	$[3n^{0.2}]$	$[3n^{0.3}]$	$J=2$	$J=3$	$J=4$	$\delta_n(1,2)$	$\delta_n(1,\frac{5}{2})$	
model 1 (0.2)	ACV	0.994	0.987	0.978	0.992	0.985	0.960	0.967	0.927	0.829	0.781	0.711	0.974
	ECV	0.997	0.990	0.978	0.987	0.973	0.944	0.970	0.932	0.842	0.461	0.440	0.969
model 1 (0.1)	ACV	0.610	0.511	0.453	0.582	0.502	0.403	0.407	0.300	0.204	0.360	0.305	0.476
	ECV	0.627	0.504	0.437	0.518	0.437	0.345	0.417	0.316	0.220	0.113	0.116	0.433
model 2 (0.3)	ACV	0.077	0.135	0.139	0.986	1.000	1.000	0.082	1.000	0.998	0.984	0.974	0.998
	ECV	0.071	0.141	0.144	0.959	1.000	1.000	0.086	1.000	0.998	0.890	0.903	0.996
model 2 (0.2)	ACV	0.062	0.102	0.104	0.587	0.932	0.941	0.058	0.862	0.810	0.727	0.687	0.790
	ECV	0.060	0.094	0.098	0.494	0.901	0.915	0.064	0.870	0.822	0.419	0.438	0.731
model 3 (0.3, 0.2)	ACV	0.990	0.981	0.973	0.988	0.989	0.999	0.959	0.962	0.999	0.997	0.995	1.000
	ECV	0.988	0.980	0.974	0.984	0.982	0.998	0.961	0.966	0.999	0.957	0.969	1.000
model 3 (0.2, 0.1)	ACV	0.613	0.513	0.469	0.575	0.566	0.719	0.417	0.450	0.753	0.788	0.761	0.752
	ECV	0.617	0.514	0.471	0.531	0.508	0.651	0.429	0.460	0.770	0.470	0.519	0.694
model 4 (0.3, 0.2)	ACV	0.988	0.981	0.965	0.984	0.984	0.999	0.950	0.954	0.999	0.992	0.991	1.000
	ECV	0.989	0.980	0.967	0.981	0.982	0.998	0.952	0.959	0.999	0.948	0.968	0.998
model 4 (0.2, 0.1)	ACV	0.602	0.520	0.455	0.559	0.549	0.711	0.412	0.439	0.742	0.768	0.752	0.765
	ECV	0.610	0.510	0.459	0.502	0.485	0.637	0.421	0.456	0.755	0.476	0.503	0.707
model 5 (0.4)	ACV	0.096	0.112	0.128	0.170	0.412	0.999	0.118	0.553	1.000	0.999	1.000	1.000
	ECV	0.092	0.112	0.131	0.130	0.334	0.996	0.123	0.566	1.000	0.993	0.996	1.000
model 5 (0.3)	ACV	0.068	0.084	0.087	0.115	0.223	0.848	0.085	0.337	0.954	0.959	0.958	0.947
	ECV	0.074	0.084	0.088	0.090	0.169	0.781	0.088	0.352	0.960	0.816	0.861	0.922
model 6 (0.3, 0.2)	ACV	0.115	0.140	0.154	0.220	0.536	0.993	0.154	0.516	0.998	1.000	1.000	1.000
	ECV	0.117	0.139	0.161	0.168	0.439	0.982	0.162	0.527	0.999	0.998	0.999	0.999
model 6 (0.2, 0.1)	ACV	0.065	0.083	0.074	0.109	0.174	0.531	0.073	0.189	0.721	0.869	0.858	0.739
	ECV	0.066	0.072	0.076	0.077	0.132	0.449	0.077	0.200	0.737	0.642	0.674	0.678
model 7 (0.3, 0.2)	ACV	0.101	0.126	0.138	0.193	0.508	1.000	0.122	0.682	1.000	1.000	1.000	1.000
	ECV	0.110	0.121	0.141	0.152	0.406	1.000	0.129	0.694	1.000	1.000	1.000	1.000
model 7 (0.2, 0.1)	ACV	0.069	0.078	0.080	0.114	0.204	0.825	0.073	0.305	0.936	0.942	0.944	0.930
	ECV	0.067	0.079	0.084	0.078	0.153	0.749	0.079	0.319	0.943	0.768	0.822	0.898

CHAPTER III

CONCLUSION

In this dissertation, we developed a wavelet-based adaptive test statistic W_{AN} for serial correlation of unknown form. The construction of the test was based on the properties of the empirical wavelet coefficients and asymptotic equivalence between our testing problem and Fan's (1996) canonical high dimensional testing problem. We first derived the asymptotic multivariate normal distribution of any finite-dimensional subset of the empirical wavelet coefficients under the null hypothesis of no serial correlation, then we showed that they are also asymptotically uncorrelated.

A serious advantage of our proposed test is that it avoids the need to select any smoothing parameters. Thus the test is completely data-diven or adaptive. Our simulation studies reveal that the proposed methodology offers very competitive empirical power compared to other test statistics when the true spectral densities have significant spatial inhomogeneity, such as peaks, bumps and alternations (due, for example, to seasonality). Therefore it is hoped that the proposed test statistic W_{AN} will represent a useful complement to the current test statistics for serial correlation.

CHAPTER IV

PROOF OF THEOREMS

3.1 Proof of Theorem 1

Here we only provide a proof for the empirical wavelet coefficients. The proof for the theoretical wavelet coefficients is largely similar. Since we use the Haar wavelet ψ, it is easy to show that the Fourier transformation $\hat{\psi}$ of ψ satisfies $\hat{\psi}(0) = 0$:

$$\hat{\psi}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x)\, e^{-iw0} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x)\, dx = 0.$$

The last equality comes from the orthonomality of ψ. For $w \neq 0$, we have

$$\begin{aligned}
\hat{\psi}(w) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x)\, e^{-iwx} dx \\
&= \frac{1}{\sqrt{2\pi}} \Big(\int_0^{\frac{1}{2}} e^{-iwx} dx - \int_{\frac{1}{2}}^{1} e^{-iwx} dx \Big) \\
&= \frac{i}{\sqrt{2\pi}\, w} \big(e^{-iw/2} - 1 - e^{-iw} + e^{-iw/2} \big) \\
&= \frac{i}{\sqrt{2\pi}\, w} \Big[\big(e^{-iw/2} - 1 \big) - e^{-iw/2} \big(e^{-iw/2} - 1 \big) \Big] \\
&= -\frac{i}{\sqrt{2\pi}\, w} \big(1 - e^{-iw/2} \big)^2 \\
&= -\frac{i}{\sqrt{2\pi}\, w} \Big[e^{-iw/4} \big(e^{iw/4} - e^{-iw/4} \big) \Big]^2 \\
&= -\frac{i}{\sqrt{2\pi}\, w}\, e^{-iw/2} \Big[2i \sin\Big(\frac{w}{4}\Big) \Big]^2 \\
&= \frac{i}{\sqrt{2\pi}}\, e^{-iw/2}\, \frac{\sin^2(w/4)}{w/4}.
\end{aligned}$$

From the definition of $\hat{\alpha}_{jk}$, using $\hat{\rho}_X(h) = \hat{\rho}_X(-h)$ and $\hat{\psi}_{jk}(2\pi h) = e^{-i2\pi hk/2^j} 2^{-j/2} \hat{\psi}(2\pi h/2^j)$, and through straightforward but tedious algebra, we have

$$\begin{aligned}
\hat{\alpha}_{jk} &= \sum_{h=-n+1}^{n-1} \hat{\rho}_X(h) \hat{\psi}_{jk}(2\pi h) \\
&= \sum_{h=-n+1}^{n-1} \hat{\rho}_X(h) \cdot e^{-i2\pi kh/2^j} 2^{-j/2} \cdot \frac{i}{\sqrt{2\pi}} e^{-i\frac{2\pi h/2^j}{2}} \cdot \frac{\sin^2\left(\frac{2\pi h/2^j}{4}\right)}{\frac{2\pi h/2^j}{4}} \\
&= 2^{-j/2} \cdot \frac{i}{\sqrt{2\pi}} \cdot 2^{j+2} \sum_{h=-n+1}^{n-1} \hat{\rho}_X(h)\, e^{-i2\pi kh/2^j} e^{-i2\pi h/2^{j+1}} \cdot \frac{\sin^2\left(\frac{2\pi h}{2^{j+2}}\right)}{2\pi h} \\
&= 2^{j/2+2} \cdot \frac{i}{\sqrt{2\pi}} \sum_{h=-n+1}^{n-1} \hat{\rho}_X(h)\, e^{-i2\pi h(k+1/2)/2^j} \cdot \frac{\sin^2\left(\frac{2\pi h}{2^{j+2}}\right)}{2\pi h} \\
&= 2^{j/2+2} \cdot \frac{i}{\sqrt{2\pi}} \cdot 2i \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \mathrm{Im}\left(e^{-i2\pi h(k+1/2)/2^j}\right) \cdot \frac{\sin^2\left(\frac{2\pi h}{2^{j+2}}\right)}{2\pi h} \\
&= -\frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \left[-\sin\left(\frac{2\pi h(k+1/2)}{2^j}\right) \right] \cdot \frac{\sin^2\left(\frac{2\pi h}{2^{j+2}}\right)}{2\pi h} \\
&= \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \sin\left(\frac{2\pi h(k+1/2)}{2^j}\right) \cdot \frac{\sin^2\left(\frac{2\pi h}{2^{j+2}}\right)}{2\pi h}.
\end{aligned}$$

From the above equation, it is easy to see that $\hat{\alpha}_{00} = 0$.

We also have $\hat{\alpha}_{jk_1} = -\hat{\alpha}_{jk_2}$ as long as $k_1 + k_2 = 2^j - 1$, which can be proved as below:

$$\begin{aligned}
\hat{\alpha}_{jk_2} &= \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \sin\left(\frac{2\pi h(k_2+1/2)}{2^j}\right) \cdot \frac{\sin^2(2\pi h/2^{j+2})}{2\pi h} \\
&= \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \sin\left(\frac{2\pi h(2^j - 1 - k_1 + 1/2)}{2^j}\right) \cdot \frac{\sin^2(2\pi h/2^{j+2})}{2\pi h} \\
&= \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \sin\left(\frac{2\pi h\big(2^j - (k_1 + 1/2)\big)}{2^j}\right) \cdot \frac{\sin^2(2\pi h/2^{j+2})}{2\pi h} \\
&= \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \sin\left(2\pi h - 2\pi h(k_1+1/2)/2^j\right) \cdot \frac{\sin^2(2\pi h/2^{j+2})}{2\pi h}.
\end{aligned}$$

Hence

$$\hat{\alpha}_{jk_2} = \frac{2^{j/2+3}}{\sqrt{2\pi}} \sum_{h=1}^{n-1} \hat{\rho}_X(h) \cdot \big(-\sin(2\pi h(k_1 + 1/2)/2^j)\big) \cdot \frac{\sin^2\big(2\pi h/2^{j+2}\big)}{2\pi h}$$

$$= -\hat{\alpha}_{jk_1}.$$

Thus we proved Theorem 1.

3.2 Proof of Theorem 2

To illustrate the proof of Theorem 2, we introduce a lemma first.

Lemma. *Let $\hat{\psi}$ be the Fourier transformation of Haar wavelet ψ, then for all $1 \leq j_1 \leq J$, $1 \leq j_2 \leq J$ and $0 \leq k_1 < 2^{j_1-1}$, $0 \leq k_2 < 2^{j_2-1}$, we have*

$$\sum_{h=-\infty}^{\infty} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) = (2\pi)^{-1}\delta_{j_1,j_2}\,\delta_{k_1,k_2},$$

$$\sum_{h=-\infty}^{\infty} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h) = 0,$$

where $\delta_{j,k} = 1$ if $j = k$ and $\delta_{j,k} = 0$ if $j \neq k$.

Proof of Lemma:

First we have

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_{j_1k_1}(w)\Psi_{j_2k_2}(w)dw &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Big(\frac{1}{\sqrt{2\pi}}\sum_{h=-\infty}^{\infty} \hat{\Psi}_{j_1k_1}(h)e^{iwh}\Big)\Big(\frac{1}{\sqrt{2\pi}}\sum_{l=-\infty}^{\infty} \hat{\Psi}_{j_2k_2}(l)e^{iwl}\Big)\,dw \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi} \sum_{h}\sum_{l} \hat{\Psi}_{j_1k_1}(h)\hat{\Psi}_{j_2k_2}(l)\, e^{iw(h+l)}\,dw \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi} \sum_{h=-\infty}^{\infty} \hat{\Psi}_{j_1k_1}(h)\hat{\Psi}_{j_2k_2}(-h)\,dw \\
&= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \hat{\Psi}_{j_1k_1}(h)\hat{\Psi}_{j_2k_2}(-h) \\
&= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \sqrt{2\pi}\,\hat{\psi}_{j_1k_1}(2\pi h)\cdot\sqrt{2\pi}\,\hat{\psi}_{j_2k_2}(-2\pi h) \\
&= \sum_{h=-\infty}^{\infty} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h).
\end{aligned}$$

So

$$\sum_{h=-\infty}^{\infty} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) = (2\pi)^{-1} \int_{-\pi}^{\pi} \Psi_{j_1k_1}(w)\, \Psi_{j_2k_2}(w)\, dw.$$

From the orthogonality of wavelet basis Ψ_{jk}, we obtain the first equality. For the second equality, we first consider the particular case $j_1 = j_2 = j$ and $k_1 = k_2 = k$. Similarly to the proof of the first equality, we have

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} \Psi_{jk}(w)\Psi_{jk}(-w)dw &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Big(\frac{1}{\sqrt{2\pi}}\sum_{h=-\infty}^{\infty} \hat{\Psi}_{jk}(h)e^{iwh}\Big)\Big(\frac{1}{\sqrt{2\pi}}\sum_{l=-\infty}^{\infty} \hat{\Psi}_{jk}(l)e^{-iwl}\Big)\, dw \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi} \sum_h \sum_l \hat{\Psi}_{jk}(h)\hat{\Psi}_{jk}(l)\, e^{iw(h-l)}\, dw \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi} \sum_{h=-\infty}^{\infty} \hat{\Psi}_{jk}(h)\hat{\Psi}_{jk}(h)\, dw \\
&= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \hat{\Psi}_{jk}(h)\hat{\Psi}_{jk}(h) \\
&= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \sqrt{2\pi}\,\hat{\psi}_{jk}(2\pi h)\cdot\sqrt{2\pi}\,\hat{\psi}_{jk}(2\pi h) \\
&= \sum_{h=-\infty}^{\infty} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h).
\end{aligned}$$

So we have the relations:

$$\sum_{h=-\infty}^{\infty} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h) = (2\pi)^{-1} \int_{-\pi}^{\pi} \Psi_{jk}(w)\Psi_{jk}(-w)dw.$$

Recall the identity $\Psi_{jk}(w) = (2\pi)^{-1/2}\sum_{m=-\infty}^{\infty} \psi_{jk}(w/(2\pi)+m)$, which can be derived from the periodization technique. Since we advocate using the Haar wavelet $\psi(\cdot)$ in this dissertation, which is compactly supported over $[0,\ 1]$, it is not hard to see that, when $0 \le k < 2^{j-1}$, we have $\psi_{jk}(w/(2\pi)+m) = 0$ for all $j \ge 1$ and any $m \ne 0$. We also note, when $w \in (-\pi,\ \pi)$, $\Psi_{jk}(w) = (2\pi)^{-1/2}\psi_{jk}(w/(2\pi))$. Because of this property, the right hand side of the above equation equals to

$$(2\pi)^{-2}\int_{-\pi}^{\pi}\psi_{jk}(w/(2\pi))\sum_{n=-\infty}^{\infty}\psi_{jk}(-w/(2\pi)+n)dw = (2\pi)^{-1}\int_{-1/2}^{1/2}\psi_{jk}(u)\sum_{n=-\infty}^{\infty}\psi_{jk}(-u+n)du,$$

$$= (2\pi)^{-1}\int_{0}^{1}\psi_{jk}(u)\sum_{n=-\infty}^{\infty}\psi_{jk}(-u+n)du,$$

which can be derived by simply replacing $w/(2\pi)$ by u. Using the compact support property for $\psi(\cdot)$ on $[0,\ 1]$ again and $0 \le k < 2^{j-1}$, one could show that, when $u \in [0,1]$, $\psi_{jk}(u)\sum_{n=-\infty}^{\infty}\psi_{jk}(-u+n) \equiv 0$. Thus, we proved the special case. For the general case, note that

$$\sum_{h=-\infty}^{\infty}\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h) = (2\pi)^{-1}\int_{-\pi}^{\pi}\Psi_{j_1k_1}(w)\Psi_{j_2k_2}(-w)dw.$$

Again, when $0 \le k_1 < 2^{j_1-1}$, we have

$$\Psi_{j_1k_1}(w) = (2\pi)^{-1/2}\sum_{m=-\infty}^{\infty}\psi_{j_1k_1}(w/(2\pi)+m) = (2\pi)^{-1/2}\psi_{j_1k_1}(w/(2\pi)),$$

when $w \in [-\pi,\ \pi]$. Using the compact support property of ψ on $[0,\ 1]$ and $0 \le k_2 < 2^{j_2-1}$, we can show that the above integrand in the right hand side is zero for all $j_1,\ j_2 \ge 1$, when $w \in [-\pi,\ \pi]$. Thus we proved the Lemma.

Proof of Theorem 2:

In what follows we use C to denote any generic positive finite constant. To simplify the presentation of the proof, without loss of generality, we assume that $E(X_t) = \mu_X = 0$. Since $\hat{R}_X(0) - \sigma_X^2 = O_p(n^{-1/2})$, we may assume that the variance σ_X^2 of the random variable X_t is known (note that the limit distribution of $\hat{\alpha}_{jk}$ is the same as that with μ_X and σ_X^2 replaced with their estimators; In practice, one simply replaces σ_X^2 with its estimator

$\hat{R}_X(0)$ and μ_X with $\bar{X}_n$.) Thus, in the following proof, we only need to consider

$$\hat{\rho}_X(h) = \sigma_X^{-2}\,\hat{R}_X(h) = n^{-1}\,\sigma_X^{-2}\sum_{t=|h|+1}^{n} X_t X_{t-|h|}.$$

First notice that

$$\begin{aligned}
\hat{\psi}_{jk}(0) &= e^0\,2^{-\frac{j}{2}}\,\hat{\psi}(0)\\
&= 2^{-\frac{j}{2}}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x)\,e^{-iw0}dx\\
&= 2^{-\frac{j}{2}}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi(x)\,dx\\
&= 0.
\end{aligned}$$

Replacing $\hat{\rho}_X(h)$ in $\hat{\alpha}_{jk}$, and exchanging the order of summations, we have the relations:

$$\begin{aligned}
\hat{\alpha}_{jk} &= \sum_{h=-(n-1)}^{n-1}\hat{\rho}_X(h)\hat{\psi}_{jk}(2\pi h)\\
&= \sum_{h=-(n-1)}^{n-1}\Big(n^{-1}\,\sigma_X^{-2}\sum_{t=|h|+1}^{n}X_tX_{t-|h|}\Big)\hat{\psi}_{jk}(2\pi h)\\
&= n^{-1}\,\sigma_X^{-2}\sum_{h=-(n-1)}^{-1}\sum_{t=-h+1}^{n}X_tX_{t+h}\,\hat{\psi}_{jk}(2\pi h) + n^{-1}\,\sigma_X^{-2}\sum_{h=1}^{n-1}\sum_{t=h+1}^{n}X_tX_{t-h}\,\hat{\psi}_{jk}(2\pi h)\\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\sum_{h=-t+1}^{-1}X_tX_{t+h}\,\hat{\psi}_{jk}(2\pi h) + n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}X_tX_{t-h}\,\hat{\psi}_{jk}(2\pi h)\\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}X_tX_{t-h}\,\hat{\psi}_{jk}(-2\pi h) + n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}X_tX_{t-h}\,\hat{\psi}_{jk}(2\pi h)\\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}X_tX_{t-h}\Big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\Big].
\end{aligned}$$

We write $n^{1/2}\,\hat{\alpha}_{jk}$ as the following sum:

$$n^{1/2}\,\hat{\alpha}_{jk} = n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n}X_tU_t,$$

where $U_t = \sum_{h=1}^{t-1}X_{t-h}\Big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\Big]$.

Based on Assumption 1 that $\{X_t\}_{t=-\infty}^{\infty}$ is independent and identically distributed with $E(X_t) = 0$, we have $E(n^{1/2}\,\hat{\alpha}_{jk}) = 0$, since

$$
\begin{aligned}
E(n^{1/2}\,\hat{\alpha}_{jk}) &= n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n} E\big(X_t U_t\big) \\
&= n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n} E\big(X_t\big)E\big(U_t\big) \\
&= 0.
\end{aligned}
$$

We then evaluate the second moment of $n^{1/2}\,\hat{\alpha}_{jk}$ as follows:

$$
\begin{aligned}
E(n^{1/2}\,\hat{\alpha}_{jk})^2 &= E\big(n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n} X_t U_t\big)^2 \\
&= E\big(n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n}\sum_{s=2}^{n} X_t U_t X_s U_s\big) \\
&= n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n}\sum_{s=2}^{n} E\big(X_t U_t X_s U_s\big) \\
&= n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n} E\big(X_t^2 U_t^2\big) \\
&= n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n} E\big(X_t^2\big)E\big(U_t^2\big) \\
&= n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n} \sigma_X^2 E\big(U_t^2\big) \\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n} E\big(U_t^2\big).
\end{aligned}
$$

From the preceding derivation, we need an expression for the second moment of the random variable U_t as:

$$
\begin{aligned}
E(U_t^2) &= E\Big\{\sum_{h=1}^{t-1} X_{t-h}\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]\Big\}^2 \\
&= E\Big\{\sum_{h=1}^{t-1}\sum_{l=1}^{t-1} X_{t-h}X_{t-l}\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]\left[\hat{\psi}_{jk}(2\pi l)+\hat{\psi}_{jk}(-2\pi l)\right]\Big\} \\
&= \sum_{h=1}^{t-1} E\big(X_{t-h}^2\big)\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]^2 \\
&= \sigma_X^2 \sum_{h=1}^{t-1}\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]^2 \\
&= \sigma_X^2 \sum_{h=1}^{t-1}\Big[\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)+2\,\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&\qquad + \hat{\psi}_{jk}(-2\pi h)\hat{\psi}_{jk}(-2\pi h)\Big] \\
&= \sigma_X^2 \sum_{h=1}^{t-1}\left[\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)\right] \\
&\qquad + \sigma_X^2 \sum_{h=1}^{t-1}\left[\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)+\hat{\psi}_{jk}(-2\pi h)\hat{\psi}_{jk}(-2\pi h)\right] \\
&= \sigma_X^2 \sum_{h=1}^{t-1}\left[\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)\right] \\
&\qquad + \sigma_X^2 \sum_{h=-t+1}^{-1}\left[\hat{\psi}_{jk}(-2\pi h)\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)\right] \\
&= \sigma_X^2 \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) + \sigma_X^2 \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h).
\end{aligned}
$$

Replace the above expression of $E\big(U_t^2\big)$ in the expression of $E(n^{1/2}\,\hat{\alpha}_{jk})^2$, we have

$$
\begin{aligned}
E(n^{1/2}\,\hat{\alpha}_{jk})^2 &= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n} E\big(U_t^2\big) \\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n}\Big[\sigma_X^2 \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&\qquad + \sigma_X^2 \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)\Big] \\
&= n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) + n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{t-1} \hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h).
\end{aligned}
$$

For the first term $n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{t-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)$ in the above summation, we seperate the second summation and then exchange the order of summations, we have

$$
\begin{aligned}
& n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{t-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= n^{-1}\sum_{t=2}^{n}\Big[\sum_{h=-(t-1)}^{-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)+\sum_{h=1}^{t-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)\Big] \\
&= n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)+n^{-1}\sum_{t=2}^{n}\sum_{h=1}^{t-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= n^{-1}\sum_{h=-n+1}^{-1}\sum_{t=-h+1}^{n}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)+n^{-1}\sum_{h=1}^{n-1}\sum_{t=h+1}^{n}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= n^{-1}\sum_{h=-n+1}^{-1}\sum_{t=|h|+1}^{n}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h)+n^{-1}\sum_{h=1}^{n-1}\sum_{t=|h|+1}^{n}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= n^{-1}\sum_{h=-n+1}^{n-1}\sum_{t=|h|+1}^{n}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= n^{-1}\sum_{h=-n+1}^{n-1}(n-|h|-1+1)\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&= \sum_{h=-n+1}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h).
\end{aligned}
$$

Similarly the second term in the summation of $E(n^{1/2}\,\hat{\alpha}_{jk})^2$ can be written as:

$$
n^{-1}\sum_{t=2}^{n}\sum_{h=-(t-1)}^{t-1}\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h)=\sum_{h=-n+1}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h).
$$

Hence $E(n^{1/2}\,\hat{\alpha}_{jk})^2$ can be expressed as a sum of two terms:

$$
\begin{aligned}
E(n^{1/2}\,\hat{\alpha}_{jk})^2 &= \sum_{h=-(n-1)}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(-2\pi h) \\
&\quad+\sum_{h=-(n-1)}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{jk}(2\pi h)\hat{\psi}_{jk}(2\pi h) \\
&=: I_{1n}+I_{2n}.
\end{aligned}
$$

From the Lemma and applying dominated convergence theorem to I_{1n}, we have $I_{1n} \to (2\pi)^{-1}$ as $n \to \infty$. As to the second term I_{2n}, using Lemma and applying dominated convergence theorem again, we have $I_{2n} \to 0$ as $n \to \infty$. Thus we prove that $E(n^{1/2}\,\hat{\alpha}_{jk})^2 \to (2\pi)^{-1}$.

In order to show the asymptotic normality result, we apply Brown's (1971) martingale limit theorem. We want to show

$$(2\pi n)^{1/2}\,\hat{\alpha}_{jk} \longrightarrow_d \mathcal{N}(0,\,1),$$

where $(2\pi n)^{1/2}\,\hat{\alpha}_{jk} = \sqrt{\frac{2\pi}{n}}\,\sigma_X^{-2}\sum_{t=2}^{n} X_t U_t$. In the present context, the following two conditions must be verified:

$$\frac{2\pi}{n\sigma_X^4}\sum_{t=2}^{n} E\Big[X_t^2 U_t^2 I\Big(|X_t U_t| \geq \frac{\varepsilon n^{1/2}\sigma_X^2}{(2\pi)^{1/2}}\Big)\Big] \to 0, \quad \text{for all} \quad \varepsilon > 0,$$

and

$$\frac{2\pi}{n\sigma_X^4}\sum_{t=2}^{n} E\Big[X_t^2 U_t^2 \big| \mathcal{F}_{t-1}\Big] \longrightarrow_p 1,$$

where $\to_p$ denotes convergence in probability, $\mathcal{F}_t$ represents the σ-field consisting of $\{X_s,\, s \leq t\}$ and $\{U_t,\, \mathcal{F}_{t-1}\}$ is an adapted martingale difference sequence.

For the first condition, let I_{3n} be the left hand side of it. Then we have

$$\begin{aligned}
I_{3n} &\leq \frac{2\pi}{n\sigma_X^4}\sum_{t=2}^{n} E\left(\frac{X_t^2 U_t^2 X_t^2 U_t^2\, 2\pi}{\varepsilon^2 n\sigma_X^4}\right)\\
&= \frac{4\pi^2}{n^2\varepsilon^2\sigma_X^8}\sum_{t=2}^{n} E(X_t^4 U_t^4)\\
&= Cn^{-2}\sum_{t=2}^{n} E(U_t^4)\\
&= Cn^{-2}\sum_{t=2}^{n} E\Big(\sum_{h=1}^{t-1} X_{t-h}\left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right]\Big)^4\\
&= Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1} E(X_{t-h}^4)\left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right]^4 + Cn^{-2}\\
&\quad \cdot \sum_{t=2}^{n}\sum\sum_{h\neq l} E(X_{t-h}^2)E(X_{t-l}^2)\left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right]^2\left[\hat{\psi}_{jk}(2\pi l) + \hat{\psi}_{jk}(-2\pi l)\right]^2\\
&=: I_{31n} + I_{32n}.
\end{aligned}$$

In the following, we show that $I_{31n} \to 0$ and $I_{32n} \to 0$ as $n \to \infty$.

From $|\hat{\psi}(w)| \le C(1+|w|)^{-1}$ for Haar wavelet ψ, and $\hat{\psi}_{jk}(2\pi h) = e^{-i2\pi hk/2^j} 2^{-j/2}\hat{\psi}(2\pi h/2^j)$, we have

$$\begin{aligned}
|\hat{\psi}_{jk}(2\pi h)| &= \left|e^{-i2\pi hk/2^j} 2^{-j/2}\hat{\psi}(2\pi h/2^j)\right| \\
&\le C\,2^{-j/2}|\hat{\psi}(2\pi h/2^j)| \\
&\le C\,2^{-j/2}\left(1+|2\pi h/2^j|\right)^{-1} \\
&= C\,2^{-j/2}\frac{2^j}{2^j+2\pi h} \\
&= C\,2^{j/2}(2^j+2\pi h)^{-1}.
\end{aligned}$$

Also, we note that

$$\begin{aligned}
\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]^4 &\le 2^4|\hat{\psi}_{jk}(2\pi h)|^4 \\
&\le C\,2^{2j}(2^j+2\pi h)^{-4}.
\end{aligned}$$

Therefore, it is possible to bound I_{31n} as follows:

$$\begin{aligned}
I_{31n} &= Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1} E(X_{t-h}^4)\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]^4 \\
&\le Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}\frac{2^{2j}}{(2^j+2\pi h)^4} \\
&= Cn^{-2}\sum_{t=2}^{n}2^{2j}\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^4}.
\end{aligned}$$

Notice

$$\begin{aligned}
\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^4} &= \sum_{h=1}^{2^j}\frac{1}{(2^j+2\pi h)^4}+\sum_{h=2^j+1}^{t-1}\frac{1}{(2^j+2\pi h)^4} \\
&\le \sum_{h=1}^{2^j}\frac{1}{2^{4j}}+\int_{2^j}^{\infty}\frac{1}{x^4}\,dx \\
&= 2^{-3j}+\frac{1}{3}2^{-3j} \\
&= C\,2^{-3j}.
\end{aligned}$$

So

$$\begin{aligned}
I_{31n} &\le Cn^{-2}\sum_{t=2}^{n} 2^{2j}\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^4}\\
&\le Cn^{-2}\sum_{t=2}^{n} 2^{2j}\,2^{-3j}\\
&\le Cn^{-2}\sum_{t=2}^{n} 2^{-j}\\
&\le Cn^{-2}n\,2^{-j}\\
&= Cn^{-1}2^{-j}.
\end{aligned}$$

Thus, we proved $I_{31n} \to 0$ as $n \to \infty$.

As to the term I_{32n}, the arguments are largely similar to those for I_{31n}. To find a bound, we use the inequalities:

$$\begin{aligned}
I_{32n} &= Cn^{-2}\sum_{t=2}^{n}\sum\sum_{h\neq l} E(X_{t-h}^2)E(X_{t-l}^2)\left[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\right]^2\\
&\quad\cdot\left[\hat{\psi}_{jk}(2\pi l)+\hat{\psi}_{jk}(-2\pi l)\right]^2\\
&\le Cn^{-2}\sum_{t=2}^{n}\left[\sum_{h=1}^{t-1}\frac{2^j}{(2^j+2\pi h)^2}\right]^2\\
&= Cn^{-2}\sum_{t=2}^{n}\left[2^j\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^2}\right]^2.
\end{aligned}$$

Notice

$$\begin{aligned}
\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^2} &= \sum_{h=1}^{2^j}\frac{1}{(2^j+2\pi h)^2}+\sum_{h=2^j+1}^{t-1}\frac{1}{(2^j+2\pi h)^2}\\
&\le \sum_{h=1}^{2^j}\frac{1}{2^{2j}}+\int_{2^j}^{\infty}\frac{1}{x^2}\,dx\\
&= 2^{-j}+2^{-j}\\
&= C\,2^{-j}.
\end{aligned}$$

So

$$
\begin{aligned}
I_{32n} &\le Cn^{-2}\sum_{t=2}^{n}\Big[2^j\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^2}\Big]^2\\
&\le Cn^{-2}\sum_{t=2}^{n}[2^j\,C\,2^{-j}]^2\\
&\le Cn^{-2}n\\
&= Cn^{-1}.
\end{aligned}
$$

Thus, we have $I_{32n} \to 0$. Therefore the proof for the first condition in Brown's (1971) theorem is completed.

Next, we show the second condition in Brown's (1971) theorem. From $E[X_t^2U_t^2|\mathcal{F}_{t-1}] = \sigma_X^2 U_t^2$, we need to show the following condition:

$$
\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}U_t^2 \longrightarrow_p 1,
$$

which is equivalent to the second condition. By using Markov's inequality, it is sufficient for us to prove the mean squared convergence condition:

$$
E\Big[\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}U_t^2-1\Big]^2 \longrightarrow 0, \quad \text{as} \quad n\to\infty.
$$

The left hand side, denoted with I_{4n}, can be written as

$$
\begin{aligned}
I_{4n} &:= E\Big[\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}U_t^2-1\Big]^2\\
&= E\Big[\Big(\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}U_t^2\Big)^2 - 2\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}U_t^2+1\Big]\\
&= \frac{4\pi^2}{n^2\sigma_X^4}\sum_{t=2}^{n}E(U_t^4) + \frac{4\pi^2}{n^2\sigma_X^4}\sum_{t=2}^{n}\sum_{s\ne t}^{n}E(U_t^2)E(U_s^2) - 2\frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n}E(U_t^2)+1\\
&=: I_{41n}+I_{42n}-2I_{43n}+1.
\end{aligned}
$$

Similar to the previous proof, we have $I_{43n}\to 1$. We also have$I_{42n}\to 1$:

$$
\begin{aligned}
I_{42n} &:= \frac{4\pi^2}{n^2\sigma_X^4}\sum_{t=2}^{n}\sum_{s\neq t}^{n} E(U_t^2)E(U_s^2) \\
&= \frac{4\pi^2}{n^2\sigma_X^4}\sum_{t=2}^{n} EU_t^2\Big(\sum_{s=2}^{n} EU_s^2 - EU_t^2\Big) \\
&= \frac{2\pi}{n\sigma_X^2}\sum_{t=2}^{n} EU_t^2\Big(\frac{2\pi}{n\sigma_X^2}\sum_{s=2}^{n} EU_s^2 - \frac{2\pi}{n\sigma_X^2}EU_t^2\Big) \\
&\to 1\cdot(1-0) \\
&= 1.
\end{aligned}
$$

From arguments used to establish the term I_{3n}, we have $I_{41n} \to 0$, as $n \to \infty$. Therefore,

$$
I_{4n} \to 0 + 1 - 2\cdot 1 + 1 = 0, \quad \text{as } n \to \infty.
$$

i.e., the second condition is verified. These arguments establish the asymptotic normal limit distribution.

In order to complete the proof of Theorem 2, we need to show that the random variables $n^{1/2}\,\hat{\alpha}_{jk}$ are asymptotically uncorrelated. From the definition of covariance, we have

$$
\begin{aligned}
\mathrm{Cov}\big(n^{1/2}\,\hat{\alpha}_{j_1k_1}, n^{1/2}\,\hat{\alpha}_{j_2k_2}\big) &= E\big(n^{1/2}\,\hat{\alpha}_{j_1k_1}\cdot n^{1/2}\,\hat{\alpha}_{j_2k_2}\big) - E\big(n^{1/2}\,\hat{\alpha}_{j_1k_1}\big)\cdot E\big(n^{1/2}\,\hat{\alpha}_{j_2k_2}\big) \\
&= E\Big[\Big(n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n} X_tU_{t,j_1k_1}\Big)\cdot\Big(n^{-1/2}\,\sigma_X^{-2}\sum_{s=2}^{n} X_sU_{s,j_2k_2}\Big)\Big] \\
&\quad - 0\cdot 0 \\
&= n^{-1}\sigma_X^{-4}\sum_{t=2}^{n}\sum_{s=2}^{n} E\big[X_tX_sU_{t,j_1k_1}U_{s,j_2k_2}\big] \\
&= n^{-1}\sigma_X^{-2}\sum_{t=2}^{n} E\big[U_{t,j_1k_1}\,U_{t,j_2k_2}\big],
\end{aligned}
$$

where

$$
\begin{aligned}
U_{t,j_1k_1} &= \sum_{h=1}^{t-1} X_{t-h}\Big[\hat{\psi}_{j_1k_1}(2\pi h) + \hat{\psi}_{j_1k_1}(-2\pi h)\Big], \\
U_{t,j_2k_2} &= \sum_{l=1}^{t-1} X_{t-l}\Big[\hat{\psi}_{j_2k_2}(2\pi l) + \hat{\psi}_{j_2k_2}(-2\pi l)\Big].
\end{aligned}
$$

Since

$$
\begin{aligned}
E\big[U_{t,j_1k_1}U_{t,j_2k_2}\big] &= E\Big[\sum_{h=1}^{t-1} X_{t-h}\left(\hat{\psi}_{j_1k_1}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\right)\Big] \\
&\quad \cdot \Big[\sum_{l=1}^{t-1} X_{t-l}\left(\hat{\psi}_{j_2k_2}(2\pi l)+\hat{\psi}_{j_2k_2}(-2\pi l)\right)\Big] \\
&= \sum_{h=1}^{t-1} E[X_{t-h}^2]\Big(\hat{\psi}_{j_1k_1}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\Big)\cdot\Big(\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_2k_2}(-2\pi h)\Big) \\
&= \sigma_X^2 \sum_{h=1}^{t-1}\Big(\hat{\psi}_{j_1k_1}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\Big)\Big(\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_2k_2}(-2\pi h)\Big).
\end{aligned}
$$

Hence we have the relations:

$$
\begin{aligned}
\operatorname{Cov}\big(n^{1/2}\,\hat{\alpha}_{j_1k_1},\, n^{1/2}\,\hat{\alpha}_{j_2k_2}\big) &= n^{-1}\sigma_X^{-2}\sum_{t=2}^{n} E\big[U_{t,j_1k_1}\,U_{t,j_2k_2}\big] \\
&= \frac{1}{n}\sum_{t=2}^{n}\sum_{h=1}^{t-1}\Big[\hat{\psi}_{j_1k_1}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\Big]\Big[\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_2k_2}(-2\pi h)\Big] \\
&= \frac{1}{n}\sum_{t=2}^{n}\sum_{h=1}^{t-1}\Big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) \\
&\qquad + \hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big] \\
&= \frac{1}{n}\sum_{t=2}^{n}\sum_{h=1}^{t-1}\Big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big] \\
&\qquad + \frac{1}{n}\sum_{t=2}^{n}\sum_{h=-t+1}^{-1}\Big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)\Big] \\
&= \frac{1}{n}\sum_{t=2}^{n}\sum_{h=-t+1}^{t-1}\Big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big] \\
&= \sum_{h=-(n-1)}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) \\
&\qquad + \sum_{h=-(n-1)}^{n-1}\Big(1-\frac{|h|}{n}\Big)\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h) \\
&=: I_{5n}+I_{6n}.
\end{aligned}
$$

The second to the last equality can be derived by exchanging the order of summations.

From the Lemma and applying the dominated convergence theorem to I_{5n} and I_{6n}, we conclude that $I_{5n} \to 0$ and $I_{6n} \to 0$ as $n \to \infty$. Therefore, $n^{1/2}\,\hat{\alpha}_{jk},\ j = 1, 2, \cdots, J,\ k = 0, 1, \cdots, 2^{j-1} - 1$ are asymptotically uncorrelated. This concludes the proof of Theorem 2.

3.3 Proof of Theorem 3

We can take advantage of the proof of Theorem 2 and apply the Cramer-Wold device to transform the problem from a multi-dimensional problem to a one-dimensional problem. That is, we need to show that for any arbitrary vector

$$\boldsymbol{\lambda} = (\lambda_{10}, \lambda_{20}, \lambda_{21}, \lambda_{30}, \cdots, \lambda_{33}, \cdots, \lambda_{\tilde{J}0}, \lambda_{\tilde{J}1}, \cdots, \lambda_{\tilde{J}2^{\tilde{J}-1}-1})^{\top} \in \mathbb{R}^{2^{\tilde{J}}-1},$$

we have $n^{1/2}\boldsymbol{\lambda}^{\top}\hat{\boldsymbol{\alpha}} \to_d \mathcal{N}(0, (2\pi)^{-1}||\boldsymbol{\lambda}||^2)$, where $||\boldsymbol{\lambda}||^2 = \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \lambda_{jk}^2$. Then, by the Cramer-Wold device, we prove the Theorem.

In order to do that, we first write

$$\begin{aligned} n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}} &= \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \lambda_{jk} \cdot n^{1/2}\,\hat{\alpha}_{jk} \\ &= \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \lambda_{jk} \cdot n^{-1/2}\,\sigma_X^{-2} \sum_{t=2}^{n} \sum_{h=1}^{t-1} X_t X_{t-h} \left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right] \\ &= n^{-1/2}\,\sigma_X^{-2} \sum_{t=2}^{n} X_t \sum_{h=1}^{t-1} X_{t-h} \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \lambda_{jk} \left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right] \\ &= n^{-1/2}\,\sigma_X^{-2} \sum_{t=2}^{n} X_t W_t, \end{aligned}$$

where $W_t = \sum_{h=1}^{t-1} X_{t-h} \sum_{j=1}^{\tilde{J}} \sum_{k=0}^{2^{j-1}-1} \lambda_{jk} \left[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\right]$.

From assumption of independence on the process $\{X_t\}$, we have $E(n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}}) = 0$, which can be seen as follows:

$$\begin{aligned} E(n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}}) &= E\big(n^{-1/2}\,\sigma_X^{-2} \sum_{t=2}^{n} X_t W_t\big) \\ &= n^{-1/2}\,\sigma_X^{-2} \sum_{t=2}^{n} E(X_t) E(W_t) \\ &= 0. \end{aligned}$$

The second moment of $n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}}$ is computed as:

$$\begin{aligned}
E(n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}})^2 &= E\big(n^{-1/2}\,\sigma_X^{-2}\sum_{t=2}^{n} X_t W_t\big)^2 \\
&= E\big(n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n}\sum_{s=2}^{n} X_t X_s W_t W_s\big) \\
&= n^{-1}\,\sigma_X^{-4}\sum_{t=2}^{n} E(X_t^2)E(W_t^2) \\
&= n^{-1}\,\sigma_X^{-2}\sum_{t=2}^{n} E(W_t^2).
\end{aligned}$$

Similar to the proof of Theorem 2, we have the following expressions:

$$\begin{aligned}
E(W_t^2) &= E\Big(\sum_{h=1}^{t-1} X_{t-h}\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\Big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\Big]\Big)^2 \\
&= \sum_{h=1}^{t-1} EX_{t-h}^2\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^2 \\
&= \sigma_X^2\sum_{h=1}^{t-1}\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^2 \\
&= \sigma_X^2\sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2}\lambda_{j_1k_1}\lambda_{j_2k_2}\sum_{h=1}^{t-1}\big[\hat{\psi}_{j_1k_1}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\big] \\
&\quad\cdot\big[\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_2k_2}(-2\pi h)\big] \\
&= \sigma_X^2\sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2}\lambda_{j_1k_1}\lambda_{j_2k_2}\sum_{h=1}^{t-1}\big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) \\
&\quad+\hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\big] \\
&= \sigma_X^2\sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2}\lambda_{j_1k_1}\lambda_{j_2k_2}\sum_{h=1}^{t-1}\big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)+\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\big] \\
&\quad+\sigma_X^2\sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2}\lambda_{j_1k_1}\lambda_{j_2k_2}\sum_{h=1}^{t-1}\big[\hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(2\pi h) \\
&\quad+\hat{\psi}_{j_1k_1}(-2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\big].
\end{aligned}$$

So

$$E(W_t^2) = \sigma_X^2 \sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2} \lambda_{j_1k_1}\lambda_{j_2k_2} \sum_{h=1}^{t-1} \big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h) + \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\big]$$

$$+ \sigma_X^2 \sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2} \lambda_{j_1k_1}\lambda_{j_2k_2} \sum_{h=-t+1}^{-1} \big[\hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h) + \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)\big]$$

$$= \sigma_X^2 \sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2} \lambda_{j_1k_1}\lambda_{j_2k_2} \Big[\sum_{h=-(t-1)}^{t-1} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)$$

$$+ \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big].$$

Thus, by exchanging the order of summations, we have

$$E(n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}})^2 = n^{-1}\,\sigma_X^{-2} \sum_{t=2}^{n} E(W_t^2)$$

$$= n^{-1} \sum_{t=2}^{n}\Big(\sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2} \lambda_{j_1k_1}\lambda_{j_2k_2} \Big[\sum_{h=-(t-1)}^{t-1} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)$$

$$+ \sum_{h=-(t-1)}^{t-1} \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big]\Big)$$

$$= \sum_{j_1}\sum_{k_1}\sum_{j_2}\sum_{k_2} \lambda_{j_1k_1}\lambda_{j_2k_2} \Big[\sum_{h=-(n-1)}^{n-1} \Big(1 - \frac{|h|}{n}\Big) \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(2\pi h)$$

$$+ \sum_{h=-(n-1)}^{n-1} \Big(1 - \frac{|h|}{n}\Big) \hat{\psi}_{j_1k_1}(2\pi h)\hat{\psi}_{j_2k_2}(-2\pi h)\Big]$$

$$\to (2\pi)^{-1}\,||\boldsymbol{\lambda}||^2,$$

where the last limit follows from the Lemma and the dominated convergence theorem.

In order to show the asymptotic normal limit distribution, we apply the martingale limit theorem of Brown (1971) again. We want to show that

$$n^{1/2}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}} \longrightarrow_d \mathcal{N}\Big(0,\ \frac{||\boldsymbol{\lambda}||^2}{2\pi}\Big),$$

i.e.,

$$\frac{\sqrt{2\pi n}}{||\boldsymbol{\lambda}||}\boldsymbol{\lambda}'\hat{\boldsymbol{\alpha}} \longrightarrow_d \mathcal{N}\Big(0,\ 1\Big),$$

which is also equivalent to

$$\sqrt{\frac{2\pi}{n}}\,\sigma_X^{-2}\,||\boldsymbol{\lambda}||^{-1}\sum_{t=2}^{n} X_t W_t \longrightarrow_d \mathcal{N}\Big(0,\ 1\Big).$$

So in the present context, the following two conditions must be verified:

$$\frac{2\pi}{n\sigma_X^4\,||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n} E\Big[X_t^2 W_t^2 I\Big(|X_t W_t| \geq \frac{\varepsilon n^{1/2}\sigma_X^2\,||\boldsymbol{\lambda}||}{(2\pi)^{1/2}}\Big)\Big] \to 0, \quad \text{for all} \quad \varepsilon > 0,$$

and

$$\frac{2\pi}{n\sigma_X^4\,||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n} E\Big[X_t^2 W_t^2 | \mathcal{F}_{t-1}\Big] \longrightarrow_p 1.$$

For the first condition, let T_{3n} be the left hand side of it . Similar to the proof of Theorem 2, we have

$$\begin{aligned}
T_{3n} &\leq \frac{2\pi}{n\sigma_X^4||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n} E\Big(\frac{X_t^2 W_t^2 X_t^2 W_t^2\, 2\pi}{\varepsilon^2 n\sigma^4||\boldsymbol{\lambda}||^2}\Big) \\
&= Cn^{-2}\sum_{t=2}^{n} E(W_t^4) \\
&= Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1} E(X_{t-h}^4)\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\big]\Big]^4 \\
&\quad + Cn^{-2}\sum_{t=2}^{n}\sum\sum_{h\neq l} E(X_{t-h}^2)E(X_{t-l}^2)\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h) + \hat{\psi}_{jk}(-2\pi h)\big]\Big]^2 \\
&\quad \times \Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi l) + \hat{\psi}_{jk}(-2\pi l)\big]\Big]^2 \\
&=: T_{31n} + T_{32n}.
\end{aligned}$$

In the following, we try to show that $T_{31n} \to 0$ and $T_{32n} \to 0$ as $n \to \infty$. From the

Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^4\\
&\le\Big[\Big(\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}|\lambda_{jk}|^2\Big)\Big(\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\big|\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big|^2\Big)\Big]^2\\
&=||\boldsymbol{\lambda}||^4\Big(\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\big|\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big|^2\Big)^2.
\end{aligned}$$

Also note that $|\hat{\psi}_{jk}(2\pi h)|\le C\,2^{j/2}(2^j+2\pi h)^{-1}$. Thus we have the following inequalities:

$$\begin{aligned}
T_{31n}&=Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}E(X_{t-h}^4)\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^4\\
&\le Cn^{-2}\sum_{t=2}^{n}\sum_{h=1}^{t-1}||\boldsymbol{\lambda}||^4\Big(\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\frac{2^j}{(2^j+2\pi h)^2}\Big)^2\\
&=Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\sum_{h=1}^{t-1}\Big(\sum_{j=1}^{\tilde{J}}\frac{2^j2^{j-1}}{(2^j+2\pi h)^2}\Big)^2\\
&\le Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\sum_{h=1}^{t-1}\Big(\sum_{j=1}^{\tilde{J}}\frac{2^{2j}}{(2^j+2\pi h)^2}\Big)^2.
\end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\Big(\sum_{j=1}^{\tilde{J}}\frac{2^{2j}}{(2^j+2\pi h)^2}\Big)^2&\le\sum_{j=1}^{\tilde{J}}1^2\sum_{j=1}^{\tilde{J}}\Big(\frac{2^{2j}}{(2^j+2\pi h)^2}\Big)^2\\
&=\tilde{J}\sum_{j=1}^{\tilde{J}}\frac{2^{4j}}{(2^j+2\pi h)^4}.
\end{aligned}$$

Hence

$$\begin{aligned}
T_{31n}&\le Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\sum_{h=1}^{t-1}\tilde{J}\sum_{j=1}^{\tilde{J}}\frac{2^{4j}}{(2^j+2\pi h)^4}\\
&\le Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{h=1}^{n-1}\Big(1-\frac{h}{n}\Big)\sum_{j=1}^{\tilde{J}}\frac{2^{4j}}{(2^j+2\pi h)^4},
\end{aligned}$$

where the last inequality can be derived by exchanging the order of summations.

However,

$$
\begin{aligned}
Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{h=1}^{n-1}\sum_{j=1}^{\tilde{J}}\frac{2^{4j}}{(2^j+2\pi h)^4} &= Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{j=1}^{\tilde{J}}2^{4j}\sum_{h=1}^{n-1}\frac{1}{(2^j+2\pi h)^4} \\
&\leq Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{j=1}^{\tilde{J}}2^{4j}\Big[\sum_{h=1}^{2^j}\frac{1}{(2^j+2\pi h)^4}+\sum_{h=2^j+1}^{n-1}\frac{1}{(2^j+2\pi h)^4}\Big] \\
&\leq Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{j=1}^{\tilde{J}}2^{4j}\Big[\sum_{h=1}^{2^j}\frac{1}{2^{4j}}+\int_{2^j}^{\infty}\frac{dx}{2^{4x}}\Big] \\
&\leq Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}\sum_{j=1}^{\tilde{J}}2^{4j}2^{-3j} \\
&\leq Cn^{-1}||\boldsymbol{\lambda}||^4\tilde{J}2^{\tilde{J}}.
\end{aligned}
$$

Thus, since $\tilde{J}$ is fixed, $n \to \infty$, and using the dominated convergence theorem, we have $T_{31n} \to 0$ as $n \to \infty$.

As to the term T_{32n}, the arguments are very similar to those for I_{31n}. We have

$$
\begin{aligned}
T_{32n} &= Cn^{-2}\sum_{t=2}^{n}\sum\sum_{h\neq l} E(X_{t-h}^2)E(X_{t-l}^2)\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^2 \\
&\quad\times\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi l)+\hat{\psi}_{jk}(-2\pi l)\big]\Big]^2 \\
&\leq Cn^{-2}\sum_{t=2}^{n}\bigg(\sum_{h=1}^{t-1}E(X_{t-h}^2)\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\lambda_{jk}\big[\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big]\Big]^2\bigg)^2 \\
&\leq Cn^{-2}\sum_{t=2}^{n}\bigg(\sum_{h=1}^{t-1}\Big[\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}|\lambda_{jk}|^2\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\big|\hat{\psi}_{jk}(2\pi h)+\hat{\psi}_{jk}(-2\pi h)\big|^2\bigg)^2 \\
&\leq Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{h=1}^{t-1}\sum_{j=1}^{\tilde{J}}\sum_{k=0}^{2^{j-1}-1}\frac{2^j}{(2^j+2\pi h)^2}\Big]^2 \\
&\leq Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{h=1}^{t-1}\sum_{j=1}^{\tilde{J}}\frac{2^{2j}}{(2^j+2\pi h)^2}\Big]^2 \\
&= Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{j=1}^{\tilde{J}}2^{2j}\sum_{h=1}^{t-1}\frac{1}{(2^j+2\pi h)^2}\Big]^2 \\
&= Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{j=1}^{\tilde{J}}2^{2j}\Big(\sum_{h=1}^{2^j}\frac{1}{(2^j+2\pi h)^2}+\sum_{h=2^j+1}^{t-1}\frac{1}{(2^j+2\pi h)^2}\Big)\Big]^2 \\
&\leq Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{j=1}^{\tilde{J}}2^{2j}\Big(\sum_{h=1}^{2^j}\frac{1}{2^{2j}}+\int_{2^j}^{\infty}\frac{dx}{2^{2x}}\Big)\Big]^2 \\
&\leq Cn^{-2}||\boldsymbol{\lambda}||^4\sum_{t=2}^{n}\Big[\sum_{j=1}^{\tilde{J}}2^{2j}2^{-j}\Big]^2 \\
&\leq C||\boldsymbol{\lambda}||^4 n^{-1}2^{2\tilde{J}}.
\end{aligned}
$$

Thus, since $\tilde{J}$ is fixed, $n\to\infty$, we conclude that $I_{32n}\to 0$. Therefore we complete the proof for the first condition.

Next, we show the second condition in Brown's (1971) theorem. Similar to the proof of Theorem 2, it is sufficient for us to show

$$E\Big[\frac{2\pi}{n\sigma_X^2||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n}W_t^2-1\Big]^2 \longrightarrow 0, \quad \text{as} \quad n\to\infty.$$

The left hand side of it, denoted with T_{4n}, can be written as

$$\begin{aligned}
T_{4n} &:= E\Big[\frac{2\pi}{n\sigma_X^2||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n}W_t^2-1\Big]^2 \\
&= E\Big[\Big(\frac{2\pi}{n\sigma_X^2||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n}W_t^2\Big)^2-2\frac{2\pi}{n\sigma_X^2||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n}W_t^2+1\Big] \\
&= \frac{4\pi^2}{n^2\sigma_X^4||\boldsymbol{\lambda}||^4}\sum_{t=2}^{n}E(W_t^4)+\frac{4\pi^2}{n^2\sigma_X^4||\boldsymbol{\lambda}||^4}\sum_{t=2}^{n}\sum_{s\neq t}^{n}E(W_t^2)E(W_s^2) \\
&\quad -2\frac{2\pi}{n\sigma_X^2||\boldsymbol{\lambda}||^2}\sum_{t=2}^{n}E(W_t^2)+1 \\
&=: T_{41n}+T_{42n}-2T_{43n}+1.
\end{aligned}$$

Similar to the previous proof, we have $T_{42n}\to 1$, $T_{43n}\to 1$. From arguments used to establish the term T_{3n}, we have $T_{41n}\to 0$, as $n\to\infty$. Therefore, we prove $T_{4n}\to 0$, i.e., the second condition is established. Thus we complete the proof for the normal limit distribution, as well as the proof of Theorem 3.

3.4 Proof of Theorem 4

To simplify the presentation of the proof, like the proof of Theorem 2, we assume that $E(X_t) = \mu_X = 0$ and the variance σ_X^2 of the random variable X_t is known. Thus, in the following proof, we only need to consider $\hat{\rho}_X(h) = \sigma_X^{-2}\hat{R}_X(h) = n^{-1}\sigma_X^{-2}\sum_{t=|h|+1}^{n} X_t X_{t-|h|}$.

First note that

$$\begin{aligned}
\log\log(N) &= \log\log\left(\frac{n}{2} - 1\right) \\
&= \log\log\frac{n}{2}\left(1 + o(1)\right) \\
&= \log\log(n) - \log\log(2) + o(1) \\
&= \log\log(n)\left(1 + o(1)\right).
\end{aligned}$$

Then observe that

$$\begin{aligned}
P_{H_1}(W_{AN} > c_\alpha) &= P_{H_1}(\sqrt{2\log\log(N)}\, W_{AN}^* - \{2\log\log(N) + .5\log\log\log(N) \\
&\qquad - .5\log(4\pi)\} > c_\alpha) \\
&= P_{H_1}(\sqrt{2\log\log(N)}\, W_{AN}^* > \{2\log\log(N) + .5\log\log\log(N) \\
&\qquad - .5\log(4\pi)\} + c_\alpha) \\
&= P_{H_1}(W_{AN}^* > \sqrt{2\log\log(N)}\,(1 + o(1))) \\
&= P_{H_1}(W_{AN}^* \geq \sqrt{2\log\log(n)}\,(1 + o(1))),
\end{aligned}$$

where $W_{AN}^* = (2m_n)^{-1/2}\sum_{i=1}^{m_n}(\theta_i^2 - 1)$ and $m_n = \underset{1\leq m\leq N}{\text{argmax}}\,(2m)^{-1/2}\sum_{i=1}^{m}(\theta_i^2 - 1)$.

Thus, for any $1 \leq J_n \leq J$, we have

$$W_{AN}^* \geq (2^{J_n+1} - 2)^{-1/2}\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j-1}-1}(2\pi n\hat{\alpha}_{jk}^2 - 1).$$

From $\hat{\alpha}^2_{jk_1} = \hat{\alpha}^2_{jk_2}$ as in Theorem1, we have

$$
\begin{aligned}
(2^{J_n+1}-2)^{-1/2}\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j-1}-1}(2\pi n\hat{\alpha}^2_{jk}-1) &= \frac{2}{2\sqrt{2^{J_n+1}-2}}\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j-1}-1}(2\pi n\hat{\alpha}^2_{jk}-1)\\
&= \frac{1}{2\sqrt{2^{J_n+1}-2}}\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j}-1}(2\pi n\hat{\alpha}^2_{jk}-1),
\end{aligned}
$$

which is derived by multiplying 2 by both the numerator and the denominator in the first step. Thus

$$
W^*_{AN} \geq \frac{1}{2\sqrt{2^{J_n+1}-2}}\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j}-1}(2\pi n\hat{\alpha}^2_{jk}-1).
$$

Therefore the power of our test has, noticing $n = 2N+2$,

$$
\begin{aligned}
P_{H_1}\big(W_{AN} > c_\alpha\big) &= P_{H_1}(W^*_{AN} \geq \sqrt{2\log\log(n)}\,(1+o(1)))\\
&\geq P_{H_1}\Big(\frac{\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}(2\pi n\hat{\alpha}^2_{jk}-1)}{2\sqrt{2^{J_n+1}-2}} \geq \sqrt{2\log\log(n)}\,(1+o(1))\Big)\\
&= P_{H_1}(\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j}-1}(2\pi n\hat{\alpha}^2_{jk}-1) \geq 2\sqrt{2^{J_n+1}-2}\,\sqrt{2\log\log(n)}\,(1+o(1)))\\
&= P_{H_1}(\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j}-1}2\pi n\hat{\alpha}^2_{jk} \geq 2^{J_n+1}-2+2\sqrt{2^{J_n+1}-2}\,\sqrt{2\log\log(n)}\\
&\quad\cdot(1+o(1)))\\
&= P_{H_1}\Big(2\pi\sum_{j=1}^{J_n}\sum_{k=0}^{2^{j}-1}\hat{\alpha}^2_{jk} \geq (2^{J_n+1}-2)n^{-1}\\
&\quad + 2n^{-1}(2^{J_n+1}-2)^{1/2}\sqrt{2\log\log(n)}\,(1+o(1))\Big).
\end{aligned}
$$

Hence

$$
\begin{aligned}
P_{H_1}\big(W_{AN} > c_\alpha\big) \geq P_{H_1}\Big(2\pi\, Q(\hat{f}^{J_n}_X, f_0) \geq \big[2^{J_n+1}n^{-1} + 2^{3/2}2^{J_n/2}n^{-1}\sqrt{2\log\log(n)}\big]\\
\cdot\big(1+o(1)\big)\Big).
\end{aligned}
$$

If we consider J_n such that $J_n \to \infty$, $2^{2J_n}/n \to 0$, the Theorem is proved if one can show that $Q(\hat{f}_X^{J_n}, f_0) \to Q(f, f_0) > 0$ in probability. Its proof is very similar to that of Theorem 2 in Lee and Hong (2001). We write

$$\begin{aligned}
Q(\hat{f}_X^{J_n}, f_0) - Q(f, f_0) &= \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} \hat{\alpha}_{jk}^2 - \sum_{j=1}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2 \\
&= \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} (\hat{\alpha}_{jk}^2 - \alpha_{jk}^2) - \sum_{j=J_n+1}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2 \\
&= \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} [(\hat{\alpha}_{jk} - \alpha_{jk})^2 + 2\,(\hat{\alpha}_{jk} - \alpha_{jk})\,\alpha_{jk}] - \sum_{j=J_n+1}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2 \\
&= \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} (\hat{\alpha}_{jk} - \alpha_{jk})^2 + 2 \sum_{j=1}^{J_n} \sum_{k=0}^{2^j-1} (\hat{\alpha}_{jk} - \alpha_{jk})\alpha_{jk} - \sum_{j=J_n+1}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2 \\
&=: Q_{1n} + Q_{2n} + Q_{3n}.
\end{aligned}$$

Notice we have

$$\begin{aligned}
\int_{-\pi}^{\pi} f^2(w)\,dw &= \int_{-\pi}^{\pi} \Big((2\pi)^{-1} + \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(w)\Big)^2\,dw \\
&= \int_{-\pi}^{\pi} \Big\{ \frac{1}{4\pi^2} + \frac{2}{2\pi} \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(w) + \Big(\sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \Psi_{jk}(w)\Big)^2 \Big\}\,dw \\
&= \frac{1}{2\pi} + \frac{2}{2\pi} \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk} \int_{-\pi}^{\pi} \Psi_{jk}(w)\,dw \\
&\quad + \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \sum_{k_1=0}^{2^{j_1}-1} \sum_{k_2=0}^{2^{j_2}-1} \alpha_{j_1k_1} \alpha_{j_2k_2} \int_{-\pi}^{\pi} \Psi_{j_1k_1}(w)\,\Psi_{j_2k_2}(w)\,dw \\
&= (2\pi)^{-1} + \sum_{j=1}^{\infty} \sum_{k=0}^{2^j-1} \alpha_{jk}^2.
\end{aligned}$$

We also have

$$\begin{aligned}\int_{-\pi}^{\pi} f^2(w)\,dw &= \int_{-\pi}^{\pi}\Big(\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\rho_X(h)e^{-ihw}\Big)^2 dw\\ &= \frac{1}{4\pi^2}\sum_{h=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}\rho_X(h)\rho_X(l)\int_{-\pi}^{\pi} e^{-i(h+l)w}\,dw\\ &= \frac{1}{4\pi^2}\sum_{h=-\infty}^{\infty}\rho_X(h)\rho_X(-h)\cdot 2\pi\\ &= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\rho_X^2(h) < \infty.\end{aligned}$$

So we have

$$\infty > \int_{-\pi}^{\pi} f^2(w)\,dw = (2\pi)^{-1} + \sum_{j=1}^{\infty}\sum_{k=0}^{2^j-1}\alpha_{jk}^2.$$

Thus $Q_{3n} \to 0$, from $J_n \to \infty$.

From the Cauchy-Schwarz inequality, we have

$$Q_{2n}^2 \le 4Q_{1n}\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\alpha_{jk}^2.$$

Thus, in order to prove the Theorem, it suffices to show $Q_{1n} \to 0$ in probability.

Observe that

$$\begin{aligned}\hat{\alpha}_{jk} - \alpha_{jk} &= \sum_{h=-n+1}^{n-1}\hat{\rho}_X(h)\hat{\psi}_{jk}(2\pi h) - \sum_{h=-\infty}^{\infty}\rho_X(h)\hat{\psi}_{jk}(2\pi h)\\ &= \sum_{h=-(n-1)}^{h=n-1}\big[\hat{\rho}_X(h) - \rho_X(h)\big]\hat{\psi}_{jk}(2\pi h) - \sum_{|h|\ge n}\rho_X(h)\hat{\psi}_{jk}(2\pi h).\end{aligned}$$

we have

$$
\begin{aligned}
Q_{1n} &= \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}(\hat{\alpha}_{jk}-\alpha_{jk})^2 \\
&\leq 2\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\Bigg[\sum_{h=-(n-1)}^{h=n-1}\big[\hat{\rho}_X(h)-\rho_X(h)\big]\hat{\psi}_{jk}(2\pi h)\Bigg]^2 + 2\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\Bigg[\sum_{|h|\geq n}\rho_X(h)\hat{\psi}_{jk}(2\pi h)\Bigg]^2 \\
&=: 2Q_{1n1} + 2Q_{1n2}.
\end{aligned}
$$

Applying the Cauchy-Schwarz inequality to Q_{1n2}, we have

$$
\begin{aligned}
Q_{1n2} &= \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\Bigg[\sum_{|h|\geq n}\rho_X(h)\hat{\psi}_{jk}(2\pi h)\Bigg]^2 \\
&\leq \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\Bigg[\sum_{|h|\geq n}\rho_X^2(h)\sum_{|h|\geq n}|\hat{\psi}_{jk}(2\pi h)|^2\Bigg].
\end{aligned}
$$

Since

$$
|\hat{\psi}_{jk}(2\pi h)|^2 \leq C2^j(2^j+2\pi h)^{-2},
$$

we have

$$
\begin{aligned}
\sum_{|h|\geq n}|\hat{\psi}_{jk}(2\pi h)|^2 &\leq 2\int_n^\infty \frac{C2^j}{(2^j+2\pi x)^2}\,dx \\
&\leq 2\int_n^\infty \frac{C2^j}{x^2}\,dx \\
&= \frac{C2^j}{n}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
Q_{1n2} &\leq C\sum_{|h|\geq n}\rho_X^2(h)\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}2^j/n \\
&= C\sum_{|h|\geq n}\rho_X^2(h)\sum_{j=1}^{J_n}2^{2j}/n \\
&= C\sum_{|h|\geq n}\rho_X^2(h)2^{2J_n}/n.
\end{aligned}
$$

Using the facts that $\sum_{|h|\geq n}\rho_X^2(h)\to 0$, and $2^{2J_n}/n\to 0$, We have

$$Q_{1n2}\to 0.$$

As to Q_{1n1}, applying the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
Q_{1n1} &= \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\left[\sum_{h=-(n-1)}^{h=n-1}[\hat{\rho}_X(h)-\rho_X(h)]\hat{\psi}_{jk}(2\pi h)\right]^2 \\
&\leq \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\left[\sum_{|h|<n}[\hat{\rho}_X(h)-\rho_X(h)]^2\sum_{|h|<n}|\hat{\psi}_{jk}(2\pi h)|^2\right].
\end{aligned}$$

So

$$EQ_{1n1}\leq \sup_{0<h<n}Var\{\hat{\rho}_X(h)\}\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\sum_{|h|<n}|\hat{\psi}_{jk}(2\pi h)|^2.$$

Like Q_{1n2}, we can show $\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\sum_{|h|<n}|\hat{\psi}_{jk}(2\pi h)|^2\leq C\,2^{2J_n}$ as below:

Firse we have

$$\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\sum_{|h|<n}|\hat{\psi}_{jk}(2\pi h)|^2\leq\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\sum_{|h|<n}\frac{C2^j}{(2^j+2\pi h)^2}.$$

Since

$$\begin{aligned}
\sum_{|h|<n}\frac{1}{(2^j+2\pi h)^2} &\leq \sum_{|h|<n}\frac{1}{h^2} \\
&\leq \sum_{h=-\infty}^{\infty}\frac{1}{h^2} \\
&< \infty.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1}\sum_{|h|<n}|\hat{\psi}_{jk}(2\pi h)|^2 &\le \sum_{j=1}^{J_n}\sum_{k=0}^{2^j-1} C2^j \\
&= \sum_{j=1}^{J_n} C2^{2j} \\
&= C\sum_{j=1}^{J_n} 4^j \\
&\le C4^{J_n} \\
&= C2^{2J_n}.
\end{aligned}$$

From Lee and Hong (2001, p.417), we have

$$\sup_{0<h<n} Var\{\hat{\rho}_X(h)\} = O(n^{-1}).$$

Thus we have

$$EQ_{1n1} = O(2^{2J_n}/n) \to 0,$$

which can be derived from our assumption on J_n.

Thus, from Markov's inequality, we conclude that $Q_{1n1} \to 0$ in probability. Therefore the proof of Theorem 4 is completed.

LIST OF REFERENCES

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *J. Amer. Statist. Assoc.* **65**, 1509–1526.

Brillinger, R. (1981). *Time series. Data analysis and theory.* Second edition. Holden-Day, Inc., Oakland, California.

Brown, B. M. (1971). Martingale central limit theorems. *Ann. Statist.* **42**, 59–66.

Darling, D. A. and Erdős, P. (1956). A limit theorem for the maximum of normalized sums of independent random variables. *Duke Math. J.* **23**, 143–155.

Daubechies, I. (1992). *Ten Lectures on Wavelets.* SIAM, Philadelphia.

Donoho, D. L. and Johnstone, I. M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26**, 879–921.

Donoho, D. L. and Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinking. *J. Amer. Statist. Assoc.* **90**, 1200–1224.

Duchesne, P. (2007), On consistent testing for serial correlation in seasonal time series models. *The Canadian Journal of Statistics* **35**, 193–213.

Duchesne, P., Li, L. and Vandermeerschen, J. (2010). On testing for serial correlation of unknown form using wavelet thresholding. *Comput. Statist. Data Anal.* **54**, 2512–2531.

Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Amer. Statist. Assoc.* 91, 674–688.

Fan, J. and Yao, Q. (2003). *Nonlinear time series. Nonparametric and parametric methods.* Springer Series in Statistics. Springer-Verlag, New York.

Hannan, E. J. (1970). *Multiple time series.* John Wiley and Sons, Inc., New York-London-Sydney.

Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). Wavelets, Approximation and Statistical Applications. *Lecture Notes in Statistics* **129**, Springer, New York.

Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica* **64**, 837–864.

Lee, J. and Hong, Y. (2001). Testing for serial correlation of unknown form using wavelet methods. *Econometric Theory* 17, 386–423.

Li, W. K. (2004). *Diagnostic Checks in Time Series*, Chapman & Hall/CRC, New York.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65**, 297–303.

Neyman, J. (1937). Smooth Test for Goodness of Fit. *Skandinavisk Aktuarietiskrift* **20**, 149-199.

Vidakovic, B. (1999). *Statistical modeling by wavelets.* John Wiley & Sons, Inc., New York.

Walter, G. (1994). Wavelets and Other Orthogonal Systems with Applications. CRC Press, Boca Raton.

APPENDIX

All calculations in simulation studies are realized using scripts written in R 2.15.0.

```
##############################################################
#####    compute the empirical critical value of Qm     #####
##############################################################


N <- 10000          # number of simulation
n <- 256            # sample size
Q <- rep(0,N)       # initiate N=10000 Qm's under H0
m <- 1              # window length parameter for Qm
                    # ranges from 1 to 3

for (i in 1:N)
{       x <- rnorm(n)    # generate data under H0
        r <- rep(0,m)    # initiate sample autocorrelations

        for (k in 1:m)   # compute sample autocorrelations
        {       for (t in (k+1):n)
                {       r[k] <- r[k] + x[t]*x[t-k]
                }
                r[k] <- r[k]/sum(x^2)
        }
        rm(k)
        rm(t)

        for (k in 1:m)   # compute Qm
        {       Q[i] <- Q[i]+ (r[k]^2)/(n-k)
        }
        rm(k)
        Q[i] <- n*(n+2)*Q[i]
}

t <- quantile(Q, probs=0.95)     # compute ECV of Qm
```

```
################################################################
####   computing the empirical critical value of Kn    ####
################################################################


N <- 10000        # number of simulation
n <- 256          # sample size
pn1 <- 6          # parameter pn1 for n=256
pn2 <- 9          # parameter pn2 for n=256
pn3 <- 16         # parameter pn3 for n=256
#pn1 <- 6         # parameter pn1 for n=512
#pn2 <- 10        # parameter pn2 for n=512
#pn3 <- 19        # parameter pn3 for n=512
M1 <- rep(0,N)    # initiate N=10000 Kn's using pn1 under H0
M2 <- rep(0,N)    # initiate N=10000 Kn's using pn2 under H0
M3 <- rep(0,N)    # initiate N=10000 Kn's using pn3 under H0

for (i in 1:N)
{       x <- rnorm(n)               # generate data under H0
        average <- mean(x)          # compute sample mean
        gamma <- rep(0,n-1)
        # initiate sample autocovariances
        for (j in 1:(n-1))
        # compute sample autocovariances
        {       for (t in (abs(j)+1):n)
                {       gamma[j] <- gamma[j] + (x[t]-average)
                                *(x[t-abs(j)]-average)
                }
        }
        gamma <- gamma/n
        gamma0 <- mean((x-average)^2)
        r <- gamma/gamma0
        # compute sample autocorrelations
        rm(j)
        rm(t)

        kappa <- function(z)
        # define Daniell kernel function
        {       sin(pi*z)/(pi*z)
        }

        C <- 0
        # compute the second term in the numerator of Kn
        for (j in 1:(n-1))
        {       C <- C + (1-j/n)*(kappa(j/pn1))^2
```

```
}
rm(j)

D <- 0
# compute the denominator of Kn
for (j in 1:(n-2))
{        D <- D + (1-j/n)*(1-(j+1)/n)
                  *(kappa(j/pn1))^4
}
rm(j)

temp <- 0
for(j in 1:(n-1))
{        temp <- temp + (kappa(j/pn1)*r[j])^2
}
temp <- temp*n
rm(j)
M1[i] <- (temp - C)/sqrt(2*D)
# compute Kn using pn1

C <- 0
for (j in 1:(n-1))
{        C <- C + (1-j/n)*(kappa(j/pn2))^2
}
rm(j)

D <- 0
for (j in 1:(n-2))
{        D <- D + (1-j/n)*(1-(j+1)/n)
                  *(kappa(j/pn2))^4
}
rm(j)

temp <- 0
for(j in 1:(n-1))
{        temp <- temp + (kappa(j/pn2)*r[j])^2
}
temp <- temp*n
rm(j)
M2[i] <- (temp - C)/sqrt(2*D)

C <- 0
for (j in 1:(n-1))
{        C <- C + (1-j/n)*(kappa(j/pn3))^2
}
```

```
        rm(j)

        D <- 0
        for (j in 1:(n-2))
        {       D <- D + (1-j/n)*(1-(j+1)/n)
                        *(kappa(j/pn3))^4
        }
        rm(j)

        temp <- 0
        for(j in 1:(n-1))
        {       temp <- temp + (kappa(j/pn3)*r[j])^2
        }
        temp <- temp*n
        rm(j)
        M3[i] <- (temp - C)/sqrt(2*D)
}

t1 <- quantile(M1, probs=0.95)
# compute ECV of Kn using pn1

t2 <- quantile(M2, probs=0.95)
# compute ECV of Kn using pn2

t3 <- quantile(M3, probs=0.95)
# compute ECV of Kn using pn3
```

```
###############################################################
####   computing the empirical critical value of Wn   ####
###############################################################


library(gdata)
N <- 10000        # number of simulation
n <- 256          # sample size

J <- log2(n)-1    # number of resolution levels
                  # for wavelet coefficients

Jn2 <- 2          # parameter of Wn
Jn3 <- 3          # parameter of Wn
Jn4 <- 4          # parameter of Wn

Wn2 <- rep(0,N) # initiate N=10000 Wn's using Jn2 under H0
Wn3 <- rep(0,N) # initiate N=10000 Wn's using Jn3 under H0
Wn4 <- rep(0,N) # initiate N=10000 Wn's using Jn4 under H0

for (i in 1:N)
{       x <- rnorm(n)             # generate data under H0
        average <- mean(x)        # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)

        alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
                # initiate wavelet coefficients
                # and record them in matrix "alpha"
        for (j in 1:J)
        {       for (k in 1:(2^j))
                {       for (h in 1:(n-1))
                        {       alpha[j,k] <- alpha[j,k] +
```

```
                        rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                        *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                        }
                        alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                        /sqrt(2*pi)
                }
        }
        rm(j)
        rm(k)
        rm(h)

        temp <- 0
        # initiate temp and use it to compute the summation
        # of alpha[j,k]^2 from level 1 to level Jn
        for (j in 1:Jn2)
        {       for (k in 1:(2^j))
                {       temp <- temp + alpha[j,k]*alpha[j,k]
                }
        }

        Wn2[i] <- (2*pi*n*temp-2^(Jn2+1)+1)/sqrt(2^(Jn2+3)-4)
        # compute test statistic Wn using Jn2
        # and write it into vector Wn2

        temp <- 0
        for (j in 1:Jn3)
        {       for (k in 1:(2^j))
                {       temp <- temp + alpha[j,k]*alpha[j,k]
                }
        }

        Wn3[i] <- (2*pi*n*temp-2^(Jn3+1)+1)/sqrt(2^(Jn3+3)-4)

        temp <- 0
        for (j in 1:Jn4)
        {       for (k in 1:(2^j))
                {       temp <- temp + alpha[j,k]*alpha[j,k]
                }
        }

        Wn4[i] <- (2*pi*n*temp-2^(Jn4+1)+1)/sqrt(2^(Jn4+3)-4)
}

t2 <- quantile(Wn2, probs=0.95)
# compute ECV of Wn2
```

```
t3 <- quantile(Wn3, probs=0.95)
# compute ECV of Wn3

t4 <- quantile(Wn4, probs=0.95)
# compute ECV of Wn4
```

```
##############################################################
####   computing the empirical critical value of Tn    ####
##############################################################


library(gdata)
N <- 10000        # number of simulation
n <- 256          # sample size

J <- log2(n)-1    # number of resolution levels
                  # for wavelet coefficients

c <- 1            # parameter for Tn
d1 <- 2           # parameter for Tn
d2 <- 2.5         # parameter for Tn

T1 <- rep(0,N)
# initiate N=10000 Tn's using c=1, d1=2 under H0
T2 <- rep(0,N)
# initiate N=10000 Tn's using c=1, d2=2.5 under H0

for (i in 1:N)
{       x <- rnorm(n)   # generate data under H0
        average <- mean(x)      # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)

        alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
                # initiate wavelet coefficients
                # and record them in matrix "alpha"
        for (j in 1:J)
        {       for (k in (2^(j-1)):(2^j-1))
                {       for (h in 1:(n-1))
```

```
                    {       alpha[j,k] <- alpha[j,k] +
                    rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                    *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                    }
                    alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                 /sqrt(2*pi)
            }
        }
        rm(j)
        rm(k)
        rm(h)

        alpha_half <- unmatrix(alpha, byrow=TRUE)
        # convert matrix "temp" into
        # a vector named "alpha_half"
        alpha_half <- alpha_half[alpha_half != 0]
        # remove zeros from "alpha_half"

        an1 <- c*(log(n/2))^((-1)*d1)
        delta1 <- sqrt(2*log(an1*n/2))
        mu1 <- (2*pi)^(-1/2)*an1^(-1)*delta1
                    *(1+delta1^(-2))
        var1 <- (2*pi)^(-1/2)*an1^(-1)*delta1^3
                    *(1+3*delta1^(-2))

        temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
                > delta1)
        T1[i] <- (2*pi*n*sum(temp^2) - mu1)/sqrt(var1)
        # compute test statistic T1 using c=1, d1=2
        rm(temp)

        an2 <- c*(log(n/2))^((-1)*d2)
        delta2 <- sqrt(2*log(an2*n/2))
        mu2 <- (2*pi)^(-1/2)*an2^(-1)*delta2
                    *(1+delta2^(-2))
        var2 <- (2*pi)^(-1/2)*an2^(-1)*delta2^3
                    *(1+3*delta2^(-2))

        temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
                > delta2)
        T2[i] <- (2*pi*n*sum(temp^2) - mu2)/sqrt(var2)
        # compute test statistic T2 using c=1, d2=2.5
        rm(temp)
}
```

```
t1 <- quantile(T1, probs=0.95)
# compute ECV of Tn using c=1, d1=2

t2 <- quantile(T2, probs=0.95)
# compute ECV of Tn using c=1, d2=2.5
```

```
###########################################################
###   computing the empirical critical value of Wan   ###
###########################################################


library(gdata)
N <- 10000        # number of simulation
n <- 256          # sample size

J <- log2(n)-1    # number of resolution levels
                  # for wavelet coefficients

T <- rep(0,N)     # initiate N=10000 Wan's under H0

for (i in 1:N)
{       x <- rnorm(n)    # generate data under H0
        average <- mean(x)       # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)

        alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
                # initiate wavelet coefficients
                # and record them in matrix "alpha"
        for (j in 1:J)
        {       for (k in (2^(j-1)):(2^j-1))
                {       for (h in 1:(n-1))
                        {       alpha[j,k] <- alpha[j,k] +
                        rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                        *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                        }
                        alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                        /sqrt(2*pi)
                }
```

```
        }
        rm(j)
        rm(k)
        rm(h)

        alpha_half <- unmatrix(alpha, byrow=TRUE)
        alpha_half <- as.vector(alpha_half)
        # convert matrix "alpha" into
        # a vector named "alpha_half"
        alpha_half <- alpha_half[alpha_half != 0]
        # remove zeros from "alpha_half"

        V <- rep(0,n/2-1)
        # scan through all the values
        # by recording them into V
        # to find the maximum
        # and let it be the test statistic
        for (j in 1:(n/2-2))
        {       temp <- alpha_half
                for (k in (j+1):(n/2-1))
                {       temp[k] <- 0
                }
                V[j] <- ((2*pi*n)*sum(temp^2)-j)/sqrt(2*j)
        }
        V[n/2-1] <- ((2*pi*n)*sum(alpha_half^2)
                        -(n/2-1))/sqrt(2*(n/2-1))
        rm(j)
        rm(k)

        m <- (1:(n/2-1))[V==max(V)]
        # find the location of where max is derived
        T[i] <- sqrt(2*log(log(n/2-1)))*V[m]
                -(2*log(log(n/2-1))+0.5*log(log(log(n/2-1)))
                -0.5*log(4*pi))
        # compute test statistic Wan
        # and write it into vector T
}

t <- quantile(T, probs=0.95)
# compute ECV of Wan
```

```
##############################################################
###############    computing level of Qm    ###############
##############################################################


N <- 10000                 ### number of simulation
n <- 256                   ### sample size
Q <- rep(0,N)              ### initiate N=10000 Qm's under H0
m <- 1                     ### window length of Qm
c <- 3.84                  ### ACV of Qm when m=1
#c <- 3.83                 ### ECV of Qm when m=1 for n=256
#c <- 3.85                 ### ECV of Qm when m=1 for n=512
#c <- 5.99                 ### ACV of Qm when m=2
#c <- 5.94                 ### ECV of Qm when m=2 for n=256
#c <- 6.01                 ### ECV of Qm when m=2 for n=512
#c <- 7.81                 ### ACV of Qm when m=3
#c <- 7.83                 ### ECV of Qm when m=3 for n=256
#c <- 7.76                 ### ECV of Qm when m=3 for n=512

for (i in 1:N)
{       x <- rnorm(n)    # generate data under H0
        r <- rep(0,m)    # initiate sample autocorrelations
        for (k in 1:m)   # compute sample autocorrelations
        {       for (t in (k+1):n)
                {       r[k] <- r[k] + x[t]*x[t-k]
                }
                r[k] <- r[k]/sum(x^2)
        }
        rm(k)
        rm(t)

        for (k in 1:m)
        {       Q[i] <- Q[i]+ (r[k]^2)/(n-k)
        }
        rm(k)
        Q[i] <- n*(n+2)*Q[i]     # compute Qm
}

level <- mean((Q > c))  # compute level of Qm
```

```
###########################################################
###############   computing level of Kn   ###############
###########################################################


N <- 10000        # number of simulation
n <- 256          # sample size
pn1 <- 6          # parameter pn1 for n=256
pn2 <- 9          # parameter pn2 for n=256
pn3 <- 16         # parameter pn3 for n=256
#pn1 <- 6         # parameter pn1 for n=512
#pn2 <- 10        # parameter pn2 for n=512
#pn3 <- 19        # parameter pn3 for n=512
M1 <- rep(0,N)    # initiate N=10000 Kn's using pn1 under H0
M2 <- rep(0,N)    # initiate N=10000 Kn's using pn2 under H0
M3 <- rep(0,N)    # initiate N=10000 Kn's using pn3 under H0
c <- 1.645        # ACV of Kn
c1 <- 1.94        # ECV of Kn using pn1 and n=256
c2 <- 1.92        # ECV of Kn using pn2 and n=256
c3 <- 1.90        # ECV of Kn using pn3 and n=256
#c1 <- 2.00       # ECV of Kn using pn1 and n=512
#c2 <- 1.98       # ECV of Kn using pn2 and n=512
#c3 <- 1.94       # ECV of Kn using pn3 and n=512

for (i in 1:N)
{       x <- rnorm(n)              # generate data under H0
        average <- mean(x)         # compute sample mean
        gamma <- rep(0,n-1)
        # initiate sample autocovariances
        for (j in 1:(n-1))
        # compute sample autocovariances
        {       for (t in (abs(j)+1):n)
                {       gamma[j] <- gamma[j] + (x[t]-average)
                                *(x[t-abs(j)]-average)
                }
        }
        gamma <- gamma/n
        gamma0 <- mean((x-average)^2)
        r <- gamma/gamma0
        # compute sample autocorrelations
        rm(j)
        rm(t)

        kappa <- function(z)
        # define Daniell kernel function
```

```
{       sin(pi*z)/(pi*z)
}

C <- 0
# compute the second term in the numerator of Kn
for (j in 1:(n-1))
{       C <- C + (1-j/n)*(kappa(j/pn1))^2
}
rm(j)

D <- 0
# compute the denominator of Kn
for (j in 1:(n-2))
{       D <- D + (1-j/n)*(1-(j+1)/n)
                *(kappa(j/pn1))^4
}
rm(j)

temp <- 0
for(j in 1:(n-1))
{       temp <- temp + (kappa(j/pn1)*r[j])^2
}
temp <- temp*n
rm(j)
M1[i] <- (temp - C)/sqrt(2*D)
# compute Kn using pn1

C <- 0
for (j in 1:(n-1))
{       C <- C + (1-j/n)*(kappa(j/pn2))^2
}
rm(j)

D <- 0
for (j in 1:(n-2))
{       D <- D + (1-j/n)*(1-(j+1)/n)
                *(kappa(j/pn2))^4
}
rm(j)

temp <- 0
for(j in 1:(n-1))
{       temp <- temp + (kappa(j/pn2)*r[j])^2
}
temp <- temp*n
```

```
        rm(j)
        M2[i] <- (temp - C)/sqrt(2*D)

        C <- 0
        for (j in 1:(n-1))
        {       C <- C + (1-j/n)*(kappa(j/pn3))^2
        }
        rm(j)

        D <- 0
        for (j in 1:(n-2))
        {       D <- D + (1-j/n)*(1-(j+1)/n)
                        *(kappa(j/pn3))^4
        }
        rm(j)

        temp <- 0
        for(j in 1:(n-1))
        {       temp <- temp + (kappa(j/pn3)*r[j])^2
        }
        temp <- temp*n
        rm(j)
        M3[i] <- (temp - C)/sqrt(2*D)
}

level1_ACV <- mean((M1 > c))
# level of Kn using ACV and pn1

level2_ACV <- mean((M2 > c))
# level of Kn using ACV and pn2

level3_ACV <- mean((M3 > c))
# level of Kn using ACV and pn3

level1_ECV <- mean((M1 > c1))
# level of Kn using ECV and pn1

level2_ECV <- mean((M2 > c2))
# level of Kn using ECV and pn2

level3_ECV <- mean((M3 > c3))
# level of Kn using ECV and pn3
```

```
###########################################################
###############   computing level of Wn   ###############
###########################################################


library(gdata)
N <- 10000        # number of simulation
n <- 256          # sample size

J <- log2(n)-1    # number of resolution levels
                  # for wavelet coefficients

Jn2 <- 2          # parameter of Wn
Jn3 <- 3          # parameter of Wn
Jn4 <- 4          # parameter of Wn

Wn2 <- rep(0,N) # initiate N=10000 Wn's using Jn2 under H0
Wn3 <- rep(0,N) # initiate N=10000 Wn's using Jn3 under H0
Wn4 <- rep(0,N) # initiate N=10000 Wn's using Jn4 under H0

t_acv <- 1.645  # ACV of Wn
t2_ecv <- 1.56  # ECV of Wn using Jn2 for n=256
t3_ecv <- 1.55  # ECV of Wn using Jn3 for n=256
t4_ecv <- 1.48  # ECV of Wn using Jn4 for n=256
#t2_ecv <- 1.59 # ECV of Wn using Jn2 for n=512
#t3_ecv <- 1.58 # ECV of Wn using Jn3 for n=512
#t4_ecv <- 1.57 # ECV of Wn using Jn4 for n=512

for (i in 1:N)
{       x <- rnorm(n)          # generate data under H0
        average <- mean(x)     # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)
```

```
alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
        # initiate wavelet coefficients
        # and record them in matrix "alpha"
for (j in 1:J)
{       for (k in 1:(2^j))
        {       for (h in 1:(n-1))
                {       alpha[j,k] <- alpha[j,k] +
                rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                }
                alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                /sqrt(2*pi)
        }
}
rm(j)
rm(k)
rm(h)

temp <- 0
# initiate temp and use it to compute the summation
# of alpha[j,k]^2 from level 1 to level Jn
for (j in 1:Jn2)
{       for (k in 1:(2^j))
        {       temp <- temp + alpha[j,k]*alpha[j,k]
        }
}
Wn2[i] <- (2*pi*n*temp-2^(Jn2+1)+1)/sqrt(2^(Jn2+3)-4)
# compute test statistic Wn using Jn2
# and write it into vector Wn2

temp <- 0
for (j in 1:Jn3)
{       for (k in 1:(2^j))
        {       temp <- temp + alpha[j,k]*alpha[j,k]
        }
}
Wn3[i] <- (2*pi*n*temp-2^(Jn3+1)+1)/sqrt(2^(Jn3+3)-4)

temp <- 0
for (j in 1:Jn4)
{       for (k in 1:(2^j))
        {       temp <- temp + alpha[j,k]*alpha[j,k]
        }
}
```

```
        Wn4[i] <- (2*pi*n*temp-2^(Jn4+1)+1)/sqrt(2^(Jn4+3)-4)
}

level2_acv <- mean((Wn2 > t_acv))
# level of Wn using ACV and Jn2

level3_acv <- mean((Wn3 > t_acv))
# level of Wn using ACV and Jn3

level4_acv <- mean((Wn4 > t_acv))
# level of Wn using ACV and Jn4

level2_ecv <- mean((Wn2 > t2_ecv))
# level of Wn using ECV and Jn2

level3_ecv <- mean((Wn3 > t3_ecv))
# level of Wn using ECV and Jn3

level4_ecv <- mean((Wn4 > t4_ecv))
# level of Wn using ECV and Jn4
```

```
###########################################################
###############    computing level of Tn    ###############
###########################################################


library(gdata)
N <- 10000          # number of simulation
n <- 256            # sample size

J <- log2(n)-1      # number of resolution levels
                    # for wavelet coefficients

c <- 1              # parameter for Tn
d1 <- 2             # parameter for Tn
d2 <- 2.5           # parameter for Tn

T1 <- rep(0,N)
# initiate N=10000 Tn's using c=1, d1=2 under H0
T2 <- rep(0,N)
# initiate N=10000 Tn's using c=1, d2=2.5 under H0

t <- 1.645          # ACV of Tn
t1 <- 3.07          # ECV of Tn using c=1, d1=2 and n=256
t2 <- 2.67          # ECV of Tn using c=1, d1=2.5 and n=256
#t1 <- 3.55         # ECV of Tn using c=1, d1=2 and n=512
#t2 <- 2.97         # ECV of Tn using c=1, d1=2.5 and n=512

for (i in 1:N)
{       x <- rnorm(n)   # generate data under H0
        average <- mean(x)      # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)
```

```
alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
        # initiate wavelet coefficients
        # and record them in matrix "alpha"
for (j in 1:J)
{       for (k in (2^(j-1)):(2^j-1))
        {       for (h in 1:(n-1))
                {       alpha[j,k] <- alpha[j,k] +
                rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                }
                alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                /sqrt(2*pi)
        }
}
rm(j)
rm(k)
rm(h)

alpha_half <- unmatrix(alpha, byrow=TRUE)
# convert matrix "temp" into
# a vector named "alpha_half"
alpha_half <- alpha_half[alpha_half != 0]
# remove zeros from "alpha_half"

an1 <- c*(log(n/2))^((-1)*d1)
delta1 <- sqrt(2*log(an1*n/2))
mu1 <- (2*pi)^(-1/2)*an1^(-1)*delta1
                *(1+delta1^(-2))
var1 <- (2*pi)^(-1/2)*an1^(-1)*delta1^3
                *(1+3*delta1^(-2))

temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
        > delta1)
T1[i] <- (2*pi*n*sum(temp^2) - mu1)/sqrt(var1)
# compute test statistic T1 using c=1, d1=2
rm(temp)

an2 <- c*(log(n/2))^((-1)*d2)
delta2 <- sqrt(2*log(an2*n/2))
mu2 <- (2*pi)^(-1/2)*an2^(-1)*delta2
                *(1+delta2^(-2))
var2 <- (2*pi)^(-1/2)*an2^(-1)*delta2^3
                *(1+3*delta2^(-2))

temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
```

```
                > delta2)
        T2[i] <- (2*pi*n*sum(temp^2) - mu2)/sqrt(var2)
        # compute test statistic T2 using c=1, d2=2.5
        rm(temp)
}

level1_ACV <- mean((T1 > t))
# level of Tn using ACV and c=1, d1=2

level2_ACV <- mean((T2 > t))
# level of Tn using ACV and c=1, d1=2.5

level1_ECV <- mean((T1 > t1))
# level of Tn using ECV and c=1, d1=2

level2_ECV <- mean((T2 > t2))
# level of Tn using ECV and c=1, d1=2.5
```

```
#########################################################
###############   computing level of Wan   ##############
#########################################################


library(gdata)

N <- 10000       # number of simulation
n <- 256         # sample size

J <- log2(n)-1   # number of resolution levels
                 # for wavelet coefficients

T <- rep(0,N)    # initiate N=10000 Wan's under H0

t_ACV <- 2.97    # ACV of Wan
t_ECV <- 3.70    # ECV of Wan for n=256
#t_ECV <- 3.58   # ECV of Wan for n=512

for (i in 1:N)
{       x <- rnorm(n)   # generate data under H0

        average <- mean(x)      # compute sample mean
        var <- mean((x-average)^2)
        # compute sample autocovariance at h=0

        rho <- rep(0,n-1)
        # compute sample autocorrelations
        for (h in 1:(n-1))
        {       for (t in (h+1):n)
                {       rho[h] <- rho[h] + (x[t]-average)
                                *(x[t-h]-average)
                }
                rho[h] <- rho[h]/n
                rho[h] <- rho[h]/var
        }
        rm(h)

        alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
                # initiate wavelet coefficients
                # and record them in matrix "alpha"
        for (j in 1:J)
        {       for (k in (2^(j-1)):(2^j-1))
                {       for (h in 1:(n-1))
                        {       alpha[j,k] <- alpha[j,k] +
```

```
                        rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                        *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                        }
                        alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                        /sqrt(2*pi)
                }
        }
        rm(j)
        rm(k)
        rm(h)

        alpha_half <- unmatrix(alpha, byrow=TRUE)
        alpha_half <- as.vector(alpha_half)
        # convert matrix "alpha" into
        # a vector named "alpha_half"
        alpha_half <- alpha_half[alpha_half != 0]
        # remove zeros from "alpha_half"

        V <- rep(0,n/2-1)
        # scan through all the values
        # by recording them into V
        # to find the maximum
        # and let it be the test statistic
        for (j in 1:(n/2-2))
        {       temp <- alpha_half
                for (k in (j+1):(n/2-1))
                {       temp[k] <- 0
                }
                V[j] <- ((2*pi*n)*sum(temp^2)-j)/sqrt(2*j)
        }
        V[n/2-1] <- ((2*pi*n)*sum(alpha_half^2)
                        -(n/2-1))/sqrt(2*(n/2-1))
        rm(j)
        rm(k)

        m <- (1:(n/2-1))[V==max(V)]
        # find the location of where max is derived

        T[i] <- sqrt(2*log(log(n/2-1)))*V[m]
                -(2*log(log(n/2-1))+0.5*log(log(log(n/2-1)))
                -0.5*log(4*pi))
        # compute test statistic Wan
        # and write it into vector T
}
```

```
level_ACV <- mean((T > t_ACV))
# level of Wan using ACV

level_ECV <- mean((T > t_ECV))
# level of Wan using ECV
```

```
#################################################################
###############    computing power of Qm    ###############
#################################################################


N <- 4000          ### number of simulation
n <- 256           ### sample size
Q <- rep(0,N)      ### initiate N=2000 Qm's under H1
m <- 1             ### window length of Qm
c <- 3.84          ### ACV of Qm when m=1
#c <- 3.83         ### ECV of Qm when m=1 for n=256
#c <- 3.85         ### ECV of Qm when m=1 for n=512
#c <- 5.99         ### ACV of Qm when m=2
#c <- 5.94         ### ECV of Qm when m=2 for n=256
#c <- 6.01         ### ECV of Qm when m=2 for n=512
#c <- 7.81         ### ACV of Qm when m=3
#c <- 7.83         ### ECV of Qm when m=3 for n=256
#c <- 7.76         ### ECV of Qm when m=3 for n=512

for (i in 1:N)
{       x <- arima.sim(list(order=c(1,0,0), ar=0.2), n)
        # AR(1)

        #x <- arima.sim(list(order=c(1,0,0), ar=0.1), n)
        # AR(1)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.3)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.2)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.2,rep(0,10),0.3,-0.06)), n)
        # ARMA(1,0)*(1,0)12

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.1,rep(0,10),0.2,-0.02)), n)
        # ARMA(1,0)*(1,0)12

        #x <- arima.sim(list(order=c(12,0,1),
        ar=c(rep(0,11),0.3), ma=0.2), n)
        # ARMA(0,1)*(1,0)12
```

```
#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.2), ma=0.1), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.4)), n)
# AR(12)

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.3)), n)
# AR(12)

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.3,rep(0,11),0.2)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.2,rep(0,11),0.1)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.3), ma=c(rep(0,11),0.2)), n)
# ARMA(0,0)*(1,1)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.2), ma=c(rep(0,11),0.1)), n)
# ARMA(0,0)*(1,1)12

r <- rep(0,m)
# compute sample autocorrelations
for (k in 1:m)
{       for (t in (k+1):n)
        {       r[k] <- r[k] + x[t]*x[t-k]
        }
        r[k] <- r[k]/sum(x^2)
}
rm(k)
rm(t)

for (k in 1:m)
{       Q[i] <- Q[i]+ (r[k]^2)/(n-k)
}
rm(k)
Q[i] <- n*(n+2)*Q[i]
```

```
            # compute Qm
}

power <- mean((Q > c))
# compute power of Qm
```

```
#############################################################
###############   computing power of Kn    ###############
#############################################################


N <- 4000          # number of simulation
n <- 256           # sample size
pn1 <- 6           # parameter pn1 for n=256
pn2 <- 9           # parameter pn2 for n=256
pn3 <- 16          # parameter pn3 for n=256
#pn1 <- 6          # parameter pn1 for n=512
#pn2 <- 10         # parameter pn2 for n=512
#pn3 <- 19         # parameter pn3 for n=512
M1 <- rep(0,N)     # initiate N=2000 Kn's using pn1 under H1
M2 <- rep(0,N)     # initiate N=2000 Kn's using pn2 under H1
M3 <- rep(0,N)     # initiate N=2000 Kn's using pn3 under H1
c <- 1.645         # ACV of Kn
c1 <- 1.94         # ECV of Kn using pn1 and n=256
c2 <- 1.92         # ECV of Kn using pn2 and n=256
c3 <- 1.90         # ECV of Kn using pn3 and n=256
#c1 <- 2.00        # ECV of Kn using pn1 and n=512
#c2 <- 1.98        # ECV of Kn using pn2 and n=512
#c3 <- 1.94        # ECV of Kn using pn3 and n=512

for (i in 1:N)
{       x <- arima.sim(list(order=c(1,0,0), ar=0.2), n)
        # AR(1)

        #x <- arima.sim(list(order=c(1,0,0), ar=0.1), n)
        # AR(1)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.3)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.2)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.2,rep(0,10),0.3,-0.06)), n)
        # ARMA(1,0)*(1,0)12

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.1,rep(0,10),0.2,-0.02)), n)
```

```
# ARMA(1,0)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.3), ma=0.2), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.2), ma=0.1), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.4)), n)
# AR(12)

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.3)), n)
# AR(12)

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.3,rep(0,11),0.2)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.2,rep(0,11),0.1)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.3), ma=c(rep(0,11),0.2)), n)
# ARMA(0,0)*(1,1)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.2), ma=c(rep(0,11),0.1)), n)
# ARMA(0,0)*(1,1)12

average <- mean(x)       # compute sample mean
gamma <- rep(0,n-1)
# initiate sample autocovariances
for (j in 1:(n-1))
# compute sample autocovariances
{       for (t in (abs(j)+1):n)
        {       gamma[j] <- gamma[j] +
                (x[t]-average)*(x[t-abs(j)]-average)
        }
}
gamma <- gamma/n
```

```
gamma0 <- mean((x-average)^2)
r <- gamma/gamma0
# compute sample autocorrelations
rm(j)
rm(t)

kappa <- function(z)
# define Daniell kernel function
{       sin(pi*z)/(pi*z)
}

C <- 0
# compute the second term in the numerator of Kn
for (j in 1:(n-1))
{       C <- C + (1-j/n)*(kappa(j/pn1))^2
}
rm(j)

D <- 0
# compute the denominator of Kn
for (j in 1:(n-2))
{       D <- D + (1-j/n)*(1-(j+1)/n)
                *(kappa(j/pn1))^4
}
rm(j)

temp <- 0
for(j in 1:(n-1))
{       temp <- temp + (kappa(j/pn1)*r[j])^2
}
temp <- temp*n
rm(j)
M1[i] <- (temp - C)/sqrt(2*D)
# compute Kn using pn1

C <- 0
for (j in 1:(n-1))
{       C <- C + (1-j/n)*(kappa(j/pn2))^2
}
rm(j)

D <- 0
for (j in 1:(n-2))
{       D <- D + (1-j/n)*(1-(j+1)/n)
                *(kappa(j/pn2))^4
```

```
        }
        rm(j)

        temp <- 0
        for(j in 1:(n-1))
        {       temp <- temp + (kappa(j/pn2)*r[j])^2
        }
        temp <- temp*n
        rm(j)
        M2[i] <- (temp - C)/sqrt(2*D)

        C <- 0
        for (j in 1:(n-1))
        {       C <- C + (1-j/n)*(kappa(j/pn3))^2
        }
        rm(j)

        D <- 0
        for (j in 1:(n-2))
        {       D <- D + (1-j/n)*(1-(j+1)/n)
                        *(kappa(j/pn3))^4
        }
        rm(j)

        temp <- 0
        for(j in 1:(n-1))
        {       temp <- temp + (kappa(j/pn3)*r[j])^2
        }
        temp <- temp*n
        rm(j)
        M3[i] <- (temp - C)/sqrt(2*D)
}

power1_ACV <- mean((M1 > c))
# power of Kn using ACV and pn1

power2_ACV <- mean((M2 > c))
# power of Kn using ACV and pn2

power3_ACV <- mean((M3 > c))
# power of Kn using ACV and pn3

power1_ECV <- mean((M1 > c1))
# power of Kn using ECV and pn1
```

```
power2_ECV <- mean((M2 > c2))
# power of Kn using ECV and pn2

power3_ECV <- mean((M3 > c3))
# power of Kn using ECV and pn3
```

```
##############################################################
###############    computing power of Wn    ###############
##############################################################


library(gdata)
N <- 4000          # number of simulation
n <- 256           # sample size

J <- log2(n)-1     # number of resolution levels
                   # for wavelet coefficients

Jn2 <- 2           # parameter of Wn
Jn3 <- 3           # parameter of Wn
Jn4 <- 4           # parameter of Wn

Wn2 <- rep(0,N) # initiate N=2000 Wn's using Jn2 under H1
Wn3 <- rep(0,N) # initiate N=2000 Wn's using Jn3 under H1
Wn4 <- rep(0,N) # initiate N=2000 Wn's using Jn4 under H1

t_acv <- 1.645  # ACV of Wn
t2_ecv <- 1.56  # ECV of Wn using Jn2 for n=256
t3_ecv <- 1.55  # ECV of Wn using Jn3 for n=256
t4_ecv <- 1.48  # ECV of Wn using Jn4 for n=256
#t2_ecv <- 1.59 # ECV of Wn using Jn2 for n=512
#t3_ecv <- 1.58 # ECV of Wn using Jn3 for n=512
#t4_ecv <- 1.57 # ECV of Wn using Jn4 for n=512

for (i in 1:N)
{       x <- arima.sim(list(order=c(1,0,0), ar=0.2), n)
        # AR(1)

        #x <- arima.sim(list(order=c(1,0,0), ar=0.1), n)
        # AR(1)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.3)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.2)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.2,rep(0,10),0.3,-0.06)), n)
```

```
# ARMA(1,0)*(1,0)12

#x <- arima.sim(list(order=c(13,0,0),
ar=c(0.1,rep(0,10),0.2,-0.02)), n)
# ARMA(1,0)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.3), ma=0.2), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.2), ma=0.1), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.4)), n)
# AR(12)

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.3)), n)
# AR(12)

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.3,rep(0,11),0.2)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.2,rep(0,11),0.1)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.3), ma=c(rep(0,11),0.2)), n)
# ARMA(0,0)*(1,1)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.2), ma=c(rep(0,11),0.1)), n)
# ARMA(0,0)*(1,1)12

average <- mean(x)      # compute sample mean
var <- mean((x-average)^2)
# compute sample autocovariance at h=0

rho <- rep(0,n-1)
# compute sample autocorrelations
for (h in 1:(n-1))
```

```
{       for (t in (h+1):n)
        {       rho[h] <- rho[h] + (x[t]-average)
                        *(x[t-h]-average)
        }
        rho[h] <- rho[h]/n
        rho[h] <- rho[h]/var
}
rm(h)

alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
        # initiate wavelet coefficients
        # and record them in matrix "alpha"
for (j in 1:J)
{       for (k in 1:(2^j))
        {       for (h in 1:(n-1))
                {       alpha[j,k] <- alpha[j,k] +
                rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                }
                alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                /sqrt(2*pi)
        }
}
rm(j)
rm(k)
rm(h)

temp <- 0
# initiate temp and use it to compute the summation
# of alpha[j,k]^2 from level 1 to level Jn
for (j in 1:Jn2)
{       for (k in 1:(2^j))
        {       temp <- temp + alpha[j,k]*alpha[j,k]
        }
}
Wn2[i] <- (2*pi*n*temp-2^(Jn2+1)+1)/sqrt(2^(Jn2+3)-4)
# compute test statistic Wn using Jn2
# and write it into vector Wn2

temp <- 0
for (j in 1:Jn3)
{       for (k in 1:(2^j))
        {       temp <- temp + alpha[j,k]*alpha[j,k]
        }
}
```

```
        Wn3[i] <- (2*pi*n*temp-2^(Jn3+1)+1)/sqrt(2^(Jn3+3)-4)

        temp <- 0
        for (j in 1:Jn4)
        {       for (k in 1:(2^j))
                {       temp <- temp + alpha[j,k]*alpha[j,k]
                }
        }
        Wn4[i] <- (2*pi*n*temp-2^(Jn4+1)+1)/sqrt(2^(Jn4+3)-4)
}

power2_acv <- mean((Wn2 > t_acv))
# power of Wn using ACV and Jn2

power3_acv <- mean((Wn3 > t_acv))
# power of Wn using ACV and Jn3

power4_acv <- mean((Wn4 > t_acv))
# power of Wn using ACV and Jn4

power2_ecv <- mean((Wn2 > t2_ecv))
# power of Wn using ECV and Jn2

power3_ecv <- mean((Wn3 > t3_ecv))
# power of Wn using ECV and Jn3

power4_ecv <- mean((Wn4 > t4_ecv))
# power of Wn using ECV and Jn4
```

```
#########################################################
###############   computing power of Tn   ###############
#########################################################


library(gdata)
N <- 4000          # number of simulation
n <- 256           # sample size

J <- log2(n)-1     # number of resolution levels
                   # for wavelet coefficients

c <- 1             # parameter for Tn
d1 <- 2            # parameter for Tn
d2 <- 2.5          # parameter for Tn

T1 <- rep(0,N)
# initiate N=2000 Tn's using c=1, d1=2 under H1
T2 <- rep(0,N)
# initiate N=2000 Tn's using c=1, d2=2.5 under H1

t <- 1.645         # ACV of Tn
t1 <- 3.07         # ECV of Tn using c=1, d1=2 and n=256
t2 <- 2.67         # ECV of Tn using c=1, d1=2.5 and n=256
#t1 <- 3.55        # ECV of Tn using c=1, d1=2 and n=512
#t2 <- 2.97        # ECV of Tn using c=1, d1=2.5 and n=512

for (i in 1:N)
{       x <- arima.sim(list(order=c(1,0,0), ar=0.2), n)
        # AR(1)

        #x <- arima.sim(list(order=c(1,0,0), ar=0.1), n)
        # AR(1)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.3)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.2)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.2,rep(0,10),0.3,-0.06)), n)
        # ARMA(1,0)*(1,0)12
```

```
#x <- arima.sim(list(order=c(13,0,0),
ar=c(0.1,rep(0,10),0.2,-0.02)), n)
# ARMA(1,0)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.3), ma=0.2), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.2), ma=0.1), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.4)), n)
# AR(12)

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.3)), n)
# AR(12)

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.3,rep(0,11),0.2)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.2,rep(0,11),0.1)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.3), ma=c(rep(0,11),0.2)), n)
# ARMA(0,0)*(1,1)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.2), ma=c(rep(0,11),0.1)), n)
# ARMA(0,0)*(1,1)12

average <- mean(x)       # compute sample mean
var <- mean((x-average)^2)
# compute sample autocovariance at h=0

rho <- rep(0,n-1)
# compute sample autocorrelations
for (h in 1:(n-1))
{       for (t in (h+1):n)
```

```
{       rho[h] <- rho[h] + (x[t]-average)
                *(x[t-h]-average)
        }
        rho[h] <- rho[h]/n
        rho[h] <- rho[h]/var
}
rm(h)

alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
        # initiate wavelet coefficients
        # and record them in matrix "alpha"
for (j in 1:J)
{       for (k in (2^(j-1)):(2^j-1))
        {       for (h in 1:(n-1))
                {       alpha[j,k] <- alpha[j,k] +
                rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                }
                alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                /sqrt(2*pi)
        }
}
rm(j)
rm(k)
rm(h)

alpha_half <- unmatrix(alpha, byrow=TRUE)
# convert matrix "temp" into
# a vector named "alpha_half"
alpha_half <- alpha_half[alpha_half != 0]
# remove zeros from "alpha_half"

an1 <- c*(log(n/2))^((-1)*d1)
delta1 <- sqrt(2*log(an1*n/2))
mu1 <- (2*pi)^(-1/2)*an1^(-1)*delta1
                *(1+delta1^(-2))
var1 <- (2*pi)^(-1/2)*an1^(-1)*delta1^3
                *(1+3*delta1^(-2))

temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
        > delta1)
T1[i] <- (2*pi*n*sum(temp^2) - mu1)/sqrt(var1)
# compute test statistic T1 using c=1, d1=2
rm(temp)
```

```
        an2 <- c*(log(n/2))^((-1)*d2)
        delta2 <- sqrt(2*log(an2*n/2))
        mu2 <- (2*pi)^(-1/2)*an2^(-1)*delta2
                        *(1+delta2^(-2))
        var2 <- (2*pi)^(-1/2)*an2^(-1)*delta2^3
                        *(1+3*delta2^(-2))

        temp <- alpha_half*(abs(sqrt(2*pi*n)*alpha_half)
                > delta2)
        T2[i] <- (2*pi*n*sum(temp^2) - mu2)/sqrt(var2)
        # compute test statistic T2 using c=1, d2=2.5
        rm(temp)
}

power1_ACV <- mean((T1 > t))
# power of Tn using ACV and c=1, d1=2

power2_ACV <- mean((T2 > t))
# power of Tn using ACV and c=1, d1=2.5

power1_ECV <- mean((T1 > t1))
# power of Tn using ECV and c=1, d1=2

power2_ECV <- mean((T2 > t2))
# power of Tn using ECV and c=1, d1=2.5
```

```
############################################################
###############   computing power of Wan   ##############
############################################################


library(gdata)
N <- 4000          # number of simulation
n <- 256           # sample size

J <- log2(n)-1     # number of resolution levels
                   # for wavelet coefficients

T <- rep(0,N)      # initiate N=2000 Wan's under H1

t_ACV <- 2.97      # ACV of Wan
t_ECV <- 3.70      # ECV of Wan for n=256
#t_ECV <- 3.58     # ECV of Wan for n=512

for (i in 1:N)
{       x <- arima.sim(list(order=c(1,0,0), ar=0.2), n)
        # AR(1)

        #x <- arima.sim(list(order=c(1,0,0), ar=0.1), n)
        # AR(1)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.3)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(4,0,0),
        ar=c(rep(0,3),0.2)), n)
        # AR(4)

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.2,rep(0,10),0.3,-0.06)), n)
        # ARMA(1,0)*(1,0)12

        #x <- arima.sim(list(order=c(13,0,0),
        ar=c(0.1,rep(0,10),0.2,-0.02)), n)
        # ARMA(1,0)*(1,0)12

        #x <- arima.sim(list(order=c(12,0,1),
        ar=c(rep(0,11),0.3), ma=0.2), n)
        # ARMA(0,1)*(1,0)12
```

```
#x <- arima.sim(list(order=c(12,0,1),
ar=c(rep(0,11),0.2), ma=0.1), n)
# ARMA(0,1)*(1,0)12

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.4)), n)
# AR(12)

#x <- arima.sim(list(order=c(12,0,0),
ar=c(rep(0,11),0.3)), n)
# AR(12)

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.3,rep(0,11),0.2)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(24,0,0),
ar=c(rep(0,11),0.2,rep(0,11),0.1)), n)
# ARMA(0,0)*(2,0)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.3), ma=c(rep(0,11),0.2)), n)
# ARMA(0,0)*(1,1)12

#x <- arima.sim(list(order=c(12,0,12),
ar=c(rep(0,11),0.2), ma=c(rep(0,11),0.1)), n)
# ARMA(0,0)*(1,1)12

average <- mean(x)      # compute sample mean
var <- mean((x-average)^2)
# compute sample autocovariance at h=0

rho <- rep(0,n-1)
# compute sample autocorrelations
for (h in 1:(n-1))
{       for (t in (h+1):n)
        {       rho[h] <- rho[h] + (x[t]-average)
                        *(x[t-h]-average)
        }
        rho[h] <- rho[h]/n
        rho[h] <- rho[h]/var
}
rm(h)

alpha <- matrix(rep(0,J*2^J),nrow=J,ncol=2^J)
```

```
        # initiate wavelet coefficients
        # and record them in matrix "alpha"
for (j in 1:J)
{       for (k in (2^(j-1)):(2^j-1))
        {       for (h in 1:(n-1))
                {       alpha[j,k] <- alpha[j,k] +
                rho[h]*sin(2*pi*h/(2^j)*(1/2+k))
                *(sin(2*pi*h/(2^(j+2))))^2/(2*pi*h)
                }
                alpha[j,k] <- alpha[j,k]*2^(j/2+3)
                                /sqrt(2*pi)
        }
}
rm(j)
rm(k)
rm(h)

alpha_half <- unmatrix(alpha, byrow=TRUE)
alpha_half <- as.vector(alpha_half)
# convert matrix "alpha" into
# a vector named "alpha_half"
alpha_half <- alpha_half[alpha_half != 0]
# remove zeros from "alpha_half"

V <- rep(0,n/2-1)
# scan through all the values
# by recording them into V
# to find the maximum
# and let it be the test statistic
for (j in 1:(n/2-2))
{       temp <- alpha_half
        for (k in (j+1):(n/2-1))
        {       temp[k] <- 0
        }
        V[j] <- ((2*pi*n)*sum(temp^2)-j)/sqrt(2*j)
}
V[n/2-1] <- ((2*pi*n)*sum(alpha_half^2)
                -(n/2-1))/sqrt(2*(n/2-1))
rm(j)
rm(k)

m <- (1:(n/2-1))[V==max(V)]
# find the location of where max is derived
T[i] <- sqrt(2*log(log(n/2-1)))*V[m]
        -(2*log(log(n/2-1))+0.5*log(log(log(n/2-1)))
```

```
            -0.5*log(4*pi))
        # compute test statistic Wan
        # and write it into vector T
}

power_ACV <- mean((T > t_ACV))
# power of Wan using ACV

power_ECV <- mean((T > t_ECV))
# power of Wan using ECV
```