# TESTING STATIONARITY USING RESIDUAL

Ohama, Gan
Research Fellow of the Japan Society for the Promotion of Science

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

https://doi.org/10.5109/13459

出版情報：Bulletin of informatics and cybernetics. 29 (1), pp.15-39, 1997-03. Research Association of Statistical Sciences
バージョン：
権利関係：

*Bulletin of Informatics and Cybernetics, Vol. 29, No. 1, 1997*

# TESTING STATIONARITY USING RESIDUAL

By

**Gan OHAMA*† and Takashi YANAGAWA***

**Abstract**

Stationarity is assumed in many conventional analysis of time series data. The aim of this paper is to propose statistical tests for a weakly stationarity.

## 1. Introduction

A time series $\{X_t\}$ is said to be stationary if the mean of $X_t$ is independent of $t$ and the covariance of $X_i$ and $X_j$ depends only on $i-j$. The stationarity is assumed in many important methods in time series analysis. The aim of this paper is to propose statistical tests for the stationarity. Okabe and Nakano (1991) explored recently a statistical method for testing this assumption. They introduced the criterion for deciding whether any given $d$-dimensional data can be regarded as a realization of a stationary time series. The criterion presumes the identity of the sample autocovariance function and population autocovariance function. In this paper we first study this presumption and then develop two tests of the stationarity. The first test modifies the one proposed by Okabe and Nakano (1991) taking into account the multiplicity of the test. The second test is new and based on the periodgrams. In section 2 we summarize the definitions and known results which will be used in subsequent sections. In Section 3, we will give the procedures of the proposed tests. In Section 5, we will show the results of simulations and application to practical examples. This will demonstrate the usefulness of our testing procedure. In Section 4, we will give the approximate distribution of the test statistics in our procedure. Section 6 is devoted to the discussion about this test.

## 2. Preliminary

DEFINITION 2.1. A multivariate time series $\{\mathbf{X}_t\}$ is said to be weakly stationary if each coordinate of $\mathbf{X}_t$ has finite second order moment for all $t$, and satisfies $E[\mathbf{X}_t] = \boldsymbol{\mu}$ for all $t$ and $Cov[\mathbf{X}_i, \mathbf{X}_j] = R(i-j)$ for all $i,j$. $R(\cdot)$ is called autocovariance function.

DEFINITION 2.2. A weakly stationary series $\{\mathbf{Z}_t\}$ is said to be white noise with covariance matrix $\Sigma$, written $WN(0,\Sigma)$, if $\{\mathbf{Z}_t\}$ satisfies $E[\mathbf{Z}_t] = 0$ for all $t$ and $Cov[\mathbf{Z}_i, \mathbf{Z}_j] = \delta_{i,j}\Sigma$ for all $i,j$, where $\delta_{i,j}$ is the Kronecker delta.

* Graduate School of Mathematics, Kyushu University

† Research Fellow of the Japan Society for the Promotion of Science

DEFINITION 2.3. The right-continuous, non-decreasing, and bounded function $F$ on $[-\pi, \pi]$ with $F(-\pi) = 0$ is said to be spectral distribution function of an univariate stationary process $\{X_t\}$ with autocovariance function $\gamma(\cdot)$ if $\gamma(h) = \int_{(-\pi,\pi]} \exp^{ih\nu} dF(\nu)$ for all $h = 0, \pm 1, \cdots$. Furthermore, if $F(\cdot)$ is absolutely continuous, the function $f(\cdot)$ such that $F(\lambda) = \int_{-\pi}^{\lambda} f(\nu) d\nu$, $-\pi \leq \lambda \leq \pi$, is called the spectral density function.

For the exsistence of the spectral distribution function, see Section 4.3 of Brockwell and Davis (1991).

The spectral density of the stationary process with absolutely summable autocovariance function $\gamma(\cdot)$ is represented by (see Corollary 4.3.2 of Brockwell and Davis (1991))

$$f(\lambda) = \frac{1}{2\pi} \sum_{n=-\infty}^{n=\infty} \gamma(n) \exp^{-in\lambda} . \tag{2.1}$$

Therefore, a process $\{\nu_t\}$ is $WN(0, 1)$ if and only if $E[\nu_t] = 0$ and its spectral density function is $(2\pi)^{-1}$.

DEFINITION 2.4. The periodgram of a univariate process $\{Z_t\}_{t=1}^{n}$ for the Fourier frequencies $\omega_j = 2\pi j/n \, (\omega_j \in [-\pi, \pi])$, is defined as follows:

$$I_n(\omega_j) = n^{-1} \left| \sum_{t=1}^{n} Z_t \exp^{-it\omega_j} \right|^2 .$$

The periodgram is extended to any $\omega \in [-\pi, \pi]$ as follows: (see Section 10.3 of Brockwell and Davis (1991))

$$I_n(\omega) = \begin{cases} I_n(\omega_k) & \omega_k - \pi/n < \omega \leq \omega_k + \pi/n, \ \omega \in [0, \pi], \\ I_n(-\omega) & \omega \in [-\pi, 0). \end{cases}$$

It is well known that $I_n(\omega_j)$ may be represented as follows.

$$I_n(\omega_j) = \begin{cases} n \left| \bar{Z}_n \right|^2 & \omega_j = \omega_0 = 0, \\ \sum_{|k|<n} \hat{\gamma}(k) \exp^{-ik\omega_j} & \omega_j \neq 0, \end{cases} \tag{2.2}$$

where $\bar{Z}_n$ and $\hat{\gamma}(k)$ are the sample mean and the sample autocovariance function of $\{Z_t\}_{t=1}^{n}$ respectively. The periodgram is used for estimating the spectral density.

## 3. Procedure of the proposed test

### 3.1. Preliminary

For a $d$-dimensional time series $\{\mathbf{X}_t\}$, we suppose that there exist mean $\boldsymbol{\mu}_t := E[\mathbf{X}_t]$ and covariance matrix $\Gamma_{s,t} := Cov[\mathbf{X}_s, \mathbf{X}_t]$. Let $X_{t,k}$ and $\mu_{t,k}$ be the $k$-th coordinate of $\mathbf{X}_t$ and $\boldsymbol{\mu}_t$, respectively. Since $X_{t,k}$ is considered as an element of a Hilbert space $L^2(\Omega)$, we may introduce the projection operator. Let $P_{\mathcal{M}}$ be the projection operator onto a closed subspace $\mathcal{M}$, and denoted by $\mathcal{M}_s^u$ the closed subspace spaned by $\{X_{t,k} - \mu_{t,k} | 1 \leq$

$k \leq d,\ s \leq t \leq u\}$. We call the "regularity condition" if and only if the covariance matrix of $\mathbf{X}_t - \boldsymbol{\mu}_t - (P_{\mathcal{M}_1^{t-1}}(X_{t,1} - \mu_{t,1}), \cdots, P_{\mathcal{M}_1^{t-1}}(X_{t,d} - \mu_{t,d}))'$ is positive definite for all $t$. If the regularity condition is satisfied, following (I) and (II) are equivalent.

(I) $\{\mathbf{X}_t\}_{t=1}^{\infty}$ is weakly stationaly.

(II) There exist $\boldsymbol{\mu} \in \mathrm{R}^d$ and the function $R(\cdot) : \mathrm{Z} \to M_d(\mathrm{R})$, where $M_d(\mathrm{R})$ denotes the set of $d \times d$ matrix on R, such that:

(i) $S_n := [R(j-i)]\,(i,j = 1, \cdots, n)$ is symmetric and positive definite for all $n$.

(ii) Put $R_n = (R(1), \cdots, R(n))$, $V_1 = R(0)$ and $V_n = R(0) - R_{n-1}S_{n-1}^{-1}R_{n-1}'$, $(n \geq 2)$. Let $W_n$ be the symmetric matrix such that $W_n' W_n = V_n$. Then the series $\{\boldsymbol{\xi}_t\}_{t=1}^{\infty}$ defined as follows is $WN(0, I)$ :

$$\boldsymbol{\xi}_t := \begin{cases} W_1^{-1}(\mathbf{X}_1 - \boldsymbol{\mu}) & t = 1, \\ W_t^{-1}(\mathbf{X}_t - \boldsymbol{\mu} - R_{t-1}S_{t-1}^{-1}(\mathbf{X}_{t-1}' - \boldsymbol{\mu}', \cdots, \mathbf{X}_1' - \boldsymbol{\mu}')') & t \geq 2. \end{cases}$$

REMARK. In (II), the existence of $W_n$ and $W_n^{-1}$ are guranteed by the assumption (II(i)).

This shows that testing the stationarity of $\{\mathbf{X}_t\}$ is equivalent to testing whehter $\{\boldsymbol{\xi}_t\}$ is $WN(0, I)$.

Now suppose that $\{\mathbf{X}_t\}$ is stationary with mean $\boldsymbol{\mu}$ and autocovariance function $R(\cdot)$, and that the regularity condition is satisfied. Put

$$(\mathbf{X}_t - \boldsymbol{\mu})^{\frown} = \begin{cases} 0 & t = 1, \\ R_{t-1}S_{t-1}^{-1}(\mathbf{X}_{t-1}' - \boldsymbol{\mu}', \cdots, \mathbf{X}_1' - \boldsymbol{\mu}')' & t \geq 2. \end{cases} \tag{3.1}$$

Then $(\mathbf{X}_t - \boldsymbol{\mu})^{\frown} = (P_{\mathcal{M}_1^{t-1}}(X_{t,1} - \mu_{t,1}), \cdots, P_{\mathcal{M}_1^{t-1}}(X_{t,d} - \mu_{t,d}))'$ and $V_t$ is the covariance matrix of $\mathbf{X}_t - \boldsymbol{\mu} - (\mathbf{X}_t - \boldsymbol{\mu})^{\frown}$. Since $V_1 \geq V_2 \geq \cdots$ in quadratic form, it follows that the regularity condition is the necessary condition of the following:

$$\tau := \inf_{t \geq 1} \tau_t > 0, \text{ where } \tau_t \text{ is the minimum eigenvalue of } V_t.$$

Recall that $\boldsymbol{\xi}_t$ is the function of true parameter $\boldsymbol{\mu}$ and $R(\cdot)$ which are unknwon. Let sample size be $N$. We estimate $\boldsymbol{\mu}$ and $R(\cdot)$ by the sample mean $\overline{\mathbf{X}}_N = N^{-1}\sum_{t=1}^{N}\mathbf{X}_t$ and the sample autocovariance function $\hat{R}(h) = N^{-1}\sum_{j=1}^{N-h}(\mathbf{X}_{j+h} - \overline{\mathbf{X}}_N)(\mathbf{X}_j - \overline{\mathbf{X}}_N)'$. If $h$ is close to $N$, $\hat{R}(h)$ is not reliable. So we use $\hat{R}(h)$ for $0 \leq h < M$, where $M$ is an appropriate constant integer to be discussed in Subsection 3.4. We estimate $\boldsymbol{\xi}_t$ by $\hat{\boldsymbol{\xi}}_t$ which is obtained by replacing $\boldsymbol{\mu}$ and $R(h)$ by $\overline{\mathbf{X}}_N$ and $\hat{R}(h)$, respectively. We use the notation "$\widehat{\ }$" to show the quantities given by such replacement, for example, $\widehat{S}_{t-1} = \left(\hat{R}(j-i)\right)_{i,j=1,\cdots,t-1}$. An exception is $(\mathbf{X}_t - \boldsymbol{\mu})^{\frown}$; when replaced it is represented by $(\mathbf{X}_t - \boldsymbol{\mu})^{\smile}$.

Since it is not possible to compute $\{\hat{\boldsymbol{\xi}}_t\}_{t>M}$ from $\hat{R}(1),\cdots,\hat{R}(M-1)$, we decompose the original process $\{\mathbf{X}_t\}_{t=1}^N$ into $N-M+1$ blocks, $B(j)$, such that

$$B(j)=\{\mathbf{X}_t^{(j)}\}_{t=1}^M=\{\mathbf{X}_t\}_{t=j}^{M+j-1},$$

as Okabe and Nakano (1991). We can compute $\{\hat{\boldsymbol{\xi}}_t^{(j)}\}_{t=1}^M$ in block $B(j)$. If the original process $\{\mathbf{X}_t\}_{t=1}^N$ is stationary then $\{\boldsymbol{\xi}_t^{(j)}\}_{t=1}^M$ is $WN(0,I)$ for each $j$. It is clear that $\{\boldsymbol{\xi}_t\}$ is $WN(0,I)$ if and only if $\{\xi_t\}$ is $WN(0,1)$, where $\xi_j$ is the $j$-th coordinate of $(\boldsymbol{\xi}_1',\boldsymbol{\xi}_2',\cdots)$. Therefore if the hypothesis $H_0^{(j)}(WN)$ : "$\{\xi_t^{(j)}\}_{t=1}^L$ *is* $WN(0,1)$ " is rejected for some $j$, the hypothesis $H_0(S)$ : "$\{\mathbf{X}_t\}_{t=1}^N$ *is stationary* " is rejected.

Employing all blocks for the test might not be the best method since the blocks are highly correlated. Thus we also consider the employment of selected blocks. The rule of the selection is as follows:

Let $N$ and $M$ be the sample size and the size of each block , and $J$ be the least integer not less than $N/M$. Let $\{j_1,j_2,\cdots,j_J\}$ be the set of indices of the selected blocks and put $R=JM-N$. Let $Q_1$ and $R_1$ be quotient and residue of $N$ divided by $M$ respectively, and $Q_2$ and $R_2$ be the corresponding quantities of $R$ divided by $Q_1$. We define $j_k$ as follows:

$$j_k=\begin{cases}(k-1)(M-1-Q_2)+1 & 1\le k\le R_2+1,\\ (k-1)(M-Q_2)-R_2+1 & R_2+1<k<J,\\ N-M+1 & k=J.\end{cases}$$

For simplify we denote by J the set of indices of the blocks to be tested. Namely,

$$\mathrm{J}=\begin{cases}\{1,2,\cdots,N-M+1\} & \text{if all blocks are employed,}\\ \{j_1,j_2,\cdots,j_J\} & \text{if } j_1,j_2,\cdots,j_J \text{ are selected.}\end{cases}$$

Furthermore we denote by $\mathrm{j}_1,\mathrm{j}_2,\cdots$ the elements of J.

### 3.2. A direct procedure

WN(0,1) is characterized by mean zero, variance one, and the orthogonality of the series. Okabe and Nakano (1991) test these three characteristics separately. They consider $2+K(K+1)/2$ test statistics in each test, where $K$ is the integer part of $2\sqrt{dM}$. Thus $(2+K(K+1)/2)(N-M+1)$ tests are undertaken altogether. If this is the case incidental rejections of the hypothesis occur, in paticular, when $N$ is large. Multiplicity of the tests must be taken into account. We develop a test which tests the three characteristics at once in each block and take into account the multiplicity of the tests throughout the blocks. The procedure is called the modified Okabe and Nakano procedure (MON procedure), including Type A (MON Type A) which uses all blocks and Type B (MON Type B) which uses selected blocks.

Now in the MON procedure, the third characteristics is tested by using the sample autocovariance function of $\{\xi_t^{(j)}\}$ till $K$. Put $L=dM$ and define $\hat{\mu}_k^{(j)}$, $\hat{\gamma}(h)^{(j)}$, $\widehat{\mathbf{Z}}_L^{(j)}$, and

$\widehat{\Sigma}_L^{(j)}$ in block $B(j)$ as follows:

$$\hat{\mu}_k^{(j)} = L^{-1}\sum_{j=1}^{L}(\hat{\xi}_j^{(j)})^k, \qquad \hat{\gamma}(h)^{(j)} = L^{-1}\sum_{j=1}^{L-h}\hat{\xi}_{j+h}^{(j)}\hat{\xi}_j^{(j)},$$

$$\widehat{\mathbf{Z}}_L^{(j)} = \sqrt{L}(\hat{\mu}_1^{(j)}, \hat{\mu}_2^{(j)}, \hat{\gamma}(1)^{(j)}, \cdots, \hat{\gamma}(K)^{(j)})',$$

$$\widehat{\Sigma}_L^{(j)} = \begin{pmatrix} \hat{\mu}_2^{(j)} & \hat{\mu}_3^{(j)} - \hat{\mu}_1^{(j)} & & & \\ \hat{\mu}_3^{(j)} - \hat{\mu}_1^{(j)} & \hat{\mu}_4^{(j)} - 2\hat{\mu}_2^{(j)} + 1 & & 0 & \\ & & 1 & & \\ & 0 & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

The MON Type A procedure for testing $H_0(S)$ with the Type I familywise error $\alpha$ is as follows:

1. Compute $\overline{\mathbf{X}}_N$ and $\hat{R}(h)$ $(0 \leq h < M)$ from $\{\mathbf{X}_t\}_{t=1}^N$.
2. Decompose $\{\mathbf{X}_t\}_{t=1}^N$ into $\{B(k)\}$ $(k \in \mathrm{J})$.
3. Put $i = 1$.
4. Compute $\{\hat{\xi}_t^{(\mathrm{j}_i)}\}_{t=1}^L$ from $B(\mathrm{j}_i)$ and compute $\hat{\mu}_k^{(\mathrm{j}_i)}$, $\hat{\gamma}(h)^{(\mathrm{j}_i)}$, $\widehat{\mathbf{Z}}_L^{(\mathrm{j}_i)}$, and $\widehat{\Sigma}_L^{(\mathrm{j}_i)}$.
5. If $T_{(\mathrm{j}_i)} = (\widehat{\mathbf{Z}}_L^{(\mathrm{j}_i)})'(\widehat{\Sigma}_L^{(\mathrm{j}_i)})^{-1}\widehat{\mathbf{Z}}_L^{(\mathrm{j}_i)} > C_0$, reject $H_0(S)$, else suppose $i = i + 1$ and return to 4. until $i = N - M + 1$.
6. $T_{(\mathrm{j}_i)} \leq C_0$ for all $i$, do not reject $H_0(S)$.

Here the constant $C_0$ is specified to be the $\alpha/(N - M + 1)$ upper quantile of chi-square distribution with $K + 2$ degree of freedom. The MON Type B procedure is for selected blocks, and given by replacing $N - M + 1$ by $J$ in the MON Type A procedure.

### 3.3. Procedure based on the spectral density

As noticed in Section 2, a time series is $WN(0, 1)$ if and only if its mean is 0 and spectral density function is identical to $(2\pi)^{-1}$ We test these characteristics by the periodgram. The procedure is called the SPED procedure, including Type A (SPED Type A) which uses all blocks and Type B (SPED Type B) which uses selected blocks. The SPED Type A procedure is now described. Put $\omega_k = 2\pi k/L$ and let $\hat{I}_L^{(j)}(\cdot)$ be the periodgram of $\{\hat{\xi}_t^{(j)}\}_{t=1}^L$. Let $\mathcal{L}$ be the integer part of $(L - 1)/2$. Type 2-A procedure with the Type I familywise error $\alpha$ is as follows:

1. Same as MON Type A.
2. Same as MON Type A.
3. Same as MON Type A.

4. Compute $\{\hat{\xi}_t^{(j_i)}\}_{t=1}^{L}$ from $B(\mathrm{j_i})$ and compute $\hat{I}_L^{(j_i)}(\omega_k)$ for $k = 1, 2, \cdots, \mathcal{L}$.

5. If $T_{(\mathrm{j_i})} = \max_{1 \le k \le \mathcal{L}} \hat{I}_L^{(j_i)}(\omega_k) \notin (C_1, C_2)$, reject $H_0(\mathrm{S})$, else suppose $i = i + 1$ and return to 4 until $i = N - M + 1$.

6. $T_{(\mathrm{j_i})} \in (C_1, C_2)$ for all $i$, do not reject $H_0(S)$.

Here $C_1$ and $C_2$ are constants such that $G_{\mathcal{L}}(C_1) = \alpha/2(N - M + 1)$, $G_{\mathcal{L}}(C_2) = 1 - \alpha/2(N - M + 1)$ and $G_n(x) = (1 - \exp^{-x})^n$. As similar to the MON Type B, the SPED Type B procedure is given by replacing $N - M + 1$ by $J$ in the SPED Type A procedure.

### 3.4. Constants $M$ and $K$.

Box and Jenkins (1970) postulate that useful sample autocovariances are obtained if sample size $N$ is greater than 50 and time lag is less than $N/4$. Thus one possibility of selecting $M$ and $K$ are the integer part of $N/4$ and $L/4$. Alternatively Okabe and Nakano (1991) suggest to select $M$ and $K$ as the integer part of $3\sqrt{N}/d$ and $2\sqrt{L}$, respectively. Note that the order of $M$ by the Box and Jenkins (1970) criterion is $N$, whereas by the Okabe and Nakano (1991) criterion is $N^{1/2}$. We compare these two criterion by simulation in Section 5.

## 4. Asymptotic properties of test statistics

To establish the asymptotic properties of those test statistics in the preceding section, we set up the following conditions. Remind that $\xi_{t,k}^{(j)}$ is the $k$-th coordinate of $\boldsymbol{\xi}_t^{(j)}$ and that $\{\xi_t^{(j)}\} = \{\xi_{1,1}^{(j)}, \cdots, \xi_{1,d}^{(j)}, \xi_{2,1}^{(j)}, \cdots, \xi_{2,d}^{(j)}, \cdots\cdots\}$.

(C1) $\tau = \inf_{t \ge 0} \tau_t > 0$, where $\tau_t$ is the minimum eigenvalue of $V_t$.

(C2) $\bar{\mathbf{X}}_N = \boldsymbol{\mu} + o_p(1)$ and $\hat{R}(h) = R(h) + o_p(1)$ as $N \to \infty$.

(C3) For each $X_{t,i}$, the $i$ th element of $\mathbf{X}_t$,

$$E[X_{t,i}^4] < \infty, \; n^{-2} \sum_{t=1}^{n} E[X_{t,i}^4] \to 0 \qquad \text{as} \quad n \to \infty.$$

(C4) $\{\boldsymbol{\xi}_t^{(j)}\}$ is independent of $t$ for each $j$ .

(C5) $E[(\xi_t^{(j)})^4] < \infty$ and $\{\xi_t^{(j)}\}$ is independent and identically distributed with respect to $t$ for each $j$.

### 4.1. Asymptotic equivarence

We first prove Lemma 4.1.1 to show Lemma 4.1.2. The lemma extends Proposition 5.1.1 of Brockwell and Davis (1991) to multivariate case.

LEMMA 4.1.1. *Let* $\{X_t\}$ *be a d-dimensional stationary process whose autocovariance function is* $\Gamma(\cdot)$. *Put* $\boldsymbol{\Gamma}_n = [\Gamma(j-i)]_{i,j=1,\cdots,n}$. *If* $\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda} > 0$ *and* $\|\Gamma(h)'\boldsymbol{\lambda}\| \to 0$ *as* $h \to \infty$ *for any* $\boldsymbol{\lambda} \in \mathrm{R}^d$, *then* $\boldsymbol{\Gamma}_n$ *is positive definite for all* $n$, *where* $\|\cdot\|$ *denotes usual norm in Euclid space.*

PROOF. Since $\boldsymbol{\Gamma}$ is the covariance matrix of $(X_n', \cdots, X_1')'$, we have $\det \boldsymbol{\Gamma} \geq 0$ for all $n$. Let $E[X_t] = 0$ without loss of generality. Suppose $\det \Gamma_{r+1} = 0$ for some $r$. If there exists $r' < r$ such that $\det \Gamma_{r'+1} = 0$, replace $r$ by $r'$. Namely, $r$ is the minimum integer such that $\det \Gamma_{r+1} = 0$. By formula (2.27) in Okabe and Nakano (1991), we have

$$\det \Gamma_{r+1} = \det \Gamma_r \det V_{r+1}.$$

Thus $\det V_{r+1} = 0$ since $\det \Gamma_r > 0$. By the similar argument as Subsection 3.1, $V_{r+1}$ is shown to be the covariance matrix of $X_{r+1} - (P_{\mathcal{M}_1^r} X_{r+1,1}, \cdots, P_{\mathcal{M}_1^r} X_{r+1,d})'$, where $X_{t,k}$ is the $k$-th coordinte of $X_t$ and $\mathcal{M}_s^u$ is the corresponding closed subspace. Therefore, there exists some $\boldsymbol{\lambda} \in \mathrm{R}^d$ and $d \times d$ matricses $\boldsymbol{\Phi}_1, \cdots, \boldsymbol{\Phi}_r$ such that

$$\boldsymbol{\lambda}' X_{r+1} = \sum_{j=1}^{r} \boldsymbol{\lambda}' \boldsymbol{\Phi}_j X_{r+1-j} \quad a.e.$$

By stationarity we have

$$\boldsymbol{\lambda}' X_{r+h} = \sum_{j=1}^{r} \boldsymbol{\lambda}' \boldsymbol{\Phi}_j X_{r+1-j} \quad a.e. \qquad \text{for all } h \geq 1.$$

Thus for all $n \geq r+1$ there exist $\boldsymbol{\Phi}_1^{(n)}, \cdots, \boldsymbol{\Phi}_r^{(n)}$ such that

$$\boldsymbol{\lambda}' X_n = \sum_{j=1}^{r} \boldsymbol{\lambda}' \boldsymbol{\Phi}_j^{(n)} X_{r+1-j} = \boldsymbol{\lambda}' \left( \boldsymbol{\Phi}_1^{(n)} \cdots \boldsymbol{\Phi}_r^{(n)} \right) \begin{pmatrix} X_r \\ \vdots \\ X_1 \end{pmatrix}. \tag{4.1}$$

Now from (4.1)

$$\begin{aligned} \boldsymbol{\lambda}' R(0) \boldsymbol{\lambda} = E[\boldsymbol{\lambda}' X_n X_n' \boldsymbol{\lambda}] &= E[\boldsymbol{\lambda}' \left( \boldsymbol{\Phi}_1^{(n)} \cdots \boldsymbol{\Phi}_r^{(n)} \right) \begin{pmatrix} X_r \\ \vdots \\ X_1 \end{pmatrix} (X_r \cdots X_1) \begin{pmatrix} \boldsymbol{\Phi}_1^{(n)} \\ \vdots \\ \boldsymbol{\Phi}_r^{(n)} \end{pmatrix} \boldsymbol{\lambda}] \\ &= \boldsymbol{\lambda}' \left( \boldsymbol{\Phi}_1^{(n)} \cdots \boldsymbol{\Phi}_r^{(n)} \right) \boldsymbol{\Gamma}_r \begin{pmatrix} \boldsymbol{\Phi}_1^{(n)} \\ \vdots \\ \boldsymbol{\Phi}_r^{(n)} \end{pmatrix} \boldsymbol{\lambda}. \end{aligned}$$

Let $\varepsilon$ be the minimum eigenvalue of $\boldsymbol{\Gamma}_r$. Since $\boldsymbol{\Gamma}_r$ is symmetric positive definite matrix, we have $\varepsilon > 0$ and

$$\begin{aligned}\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda} &\geq \varepsilon\boldsymbol{\lambda}'\left(\Phi_1^{(n)}\cdots\Phi_r^{(n)}\right)\begin{pmatrix}\Phi_1^{(n)'}\\ \vdots \\ \Phi_r^{(n)'}\end{pmatrix}\boldsymbol{\lambda}\\ &= \varepsilon\sum_{j=1}^{r}\boldsymbol{\lambda}'\Phi_j^{(n)}\Phi_j^{(n)'}\boldsymbol{\lambda} \;=\; \varepsilon\sum_{j=1}^{r}\|\Phi_j^{(n)'}\boldsymbol{\lambda}\|^2. \end{aligned} \tag{4.2}$$

We also have

$$\begin{aligned}\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda} &= Cov\left[\boldsymbol{\lambda}'\mathrm{X}_n, \boldsymbol{\lambda}'\left(\Phi_1^{(n)}\cdots\Phi_r^{(n)}\right)\begin{pmatrix}\mathrm{X}_r\\ \vdots \\ \mathrm{X}_1\end{pmatrix}\right]\\ &= \boldsymbol{\lambda}'\left(\Gamma(n-r),\cdots,\Gamma(n-1)\right)\begin{pmatrix}\Phi_1^{(n)'}\\ \vdots \\ \Phi_r^{(n)'}\end{pmatrix}\boldsymbol{\lambda}\\ &= \sum_{j=1}^{r}\boldsymbol{\lambda}'\Gamma(n-r-1+j)\Phi_j^{(n)'}\boldsymbol{\lambda} \;\leq\; \sum_{j=1}^{r}|\boldsymbol{\lambda}'\Gamma(n-r-1+j)\Phi_j^{(n)'}\boldsymbol{\lambda}|.\end{aligned}$$

By the Cauchy-Scwartz inequality and (4.2) we have

$$\begin{aligned}\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda} &\leq \sum_{j=1}^{r}|\boldsymbol{\lambda}'\Gamma(n-r-1+j)\Phi_j^{(n)'}\boldsymbol{\lambda}|\\ &\leq \sum_{j=1}^{r}\|\Gamma(n-r-1+j)'\boldsymbol{\lambda}\|\|\Phi_j^{(n)'}\boldsymbol{\lambda}\| \;\leq\; \sqrt{\frac{\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda}}{\varepsilon}}\sum_{j=1}^{r}\|\Gamma(n-r-1+j)'\boldsymbol{\lambda}\|.\end{aligned}$$

The right hand side of the last inequarity converges to 0 as $n \to \infty$. Thus $\boldsymbol{\lambda}'\Gamma(0)\boldsymbol{\lambda} = 0$, and we have contradiction. Therefore $\det \boldsymbol{\Gamma}_n > 0$ for all $n$.

□

LEMMA 4.1.2. *If* (C1) *and* (C2) *hold,* $\hat{V}_t$ *is positive definite for any* $t = 1, \cdots, N$.

PROOF. By an argument similar to section 7.2 of Brockwell and Davis (1991), we may show that $\hat{S}_N$ is non-negative definite. Define $R^*(h)$ and $S_n^*$ as follows:

$$R^*(h) = \begin{cases}\hat{R}(h) & 0 \leq h \leq N-1,\\ 0 & h \geq N.\end{cases} \qquad S_n^* = [R^*(j-i)]_{i,j=1,\cdots,n} \quad (n = 1, 2, \cdots).$$

Then using the same argument again we may show $S_n^*$ is positive definite for all $n$. Therefore, there exists $d$-dimensional stationary process $\{\mathrm{X}_t\}$ whose autocovariance matrix function is $R^*(\cdot)$. We first show $\boldsymbol{\lambda}'R^*(0)\boldsymbol{\lambda} > 0$ for any $\boldsymbol{\lambda}$. Suppose $\boldsymbol{\lambda}'R^*(0)\boldsymbol{\lambda} = 0$

for some $\boldsymbol{\lambda}$ then

$$0 = \boldsymbol{\lambda}' R^*(0)\boldsymbol{\lambda} \;=\; \boldsymbol{\lambda}'\hat{R}(0)\boldsymbol{\lambda} \;=\; \boldsymbol{\lambda}' N^{-1}\sum_{j=1}^{N}\left(\mathbf{X}_j - \overline{\mathbf{X}}_N\right)\left(\mathbf{X}_j - \overline{\mathbf{X}}_N\right)'\boldsymbol{\lambda}$$
$$= N^{-1}\sum_{j=1}^{N} \left| \left(\mathbf{X}_j - \overline{\mathbf{X}}_N\right)'\boldsymbol{\lambda}\right|^2$$

Namely, $\boldsymbol{\lambda}'\mathbf{X}_j = \boldsymbol{\lambda}'\overline{\mathbf{X}}_N$ for any $j = 1, \cdots, N$. From (C2), it follows for any $\varepsilon > 0$ that

$$\begin{aligned}\Pr\left(|\boldsymbol{\lambda}'\mathbf{X}_1 - \boldsymbol{\lambda}'\boldsymbol{\mu}| > \varepsilon\right) &= \Pr\left(|\boldsymbol{\lambda}'\overline{\mathbf{X}}_N - \boldsymbol{\lambda}'\boldsymbol{\mu}| > \varepsilon\right)\\ &= \Pr\left(|\boldsymbol{\lambda}'\left(\overline{\mathbf{X}}_N - \boldsymbol{\mu}\right)| > \varepsilon\right)\\ &\to 0 \qquad \text{as} \quad N \to \infty.\end{aligned}$$

It follows that $\boldsymbol{\lambda}'\left(\mathbf{X}_1 - \boldsymbol{\mu}\right) = 0$ *a.e.* and we have

$$0 = E\left[\left(\boldsymbol{\lambda}'\left(\mathbf{X}_1 - \boldsymbol{\mu}\right)\right)^2\right] \;=\; \boldsymbol{\lambda}' E\left[\left(\mathbf{X}_1 - \boldsymbol{\mu}\right)\left(\mathbf{X}_1 - \boldsymbol{\mu}\right)'\right]\boldsymbol{\lambda}$$
$$= \boldsymbol{\lambda}' R(0)\boldsymbol{\lambda}.$$

However we have $\det V_0 = \det R(0) > 0$ from (C1), thus contradiction. Namely we have $\boldsymbol{\lambda}'R^*(0)\boldsymbol{\lambda} > 0$ for any $\boldsymbol{\lambda}$, so $S_n^*$ must be positive definite for all $n \geq 1$ by Lemma 4.1.1. Since $R^*(h) = \hat{R}(h)$ for $0 \leq h \leq N-1$, $\hat{S}_N$ is positive definite. From formula (2.27) in Okabe and Nakano (1991),

$$\prod_{j=1}^{N} \det \hat{V}_j = \det \hat{S}_N > 0.$$

Thus $\det \hat{V}_t \neq 0$ for any $t = 1, \cdots, N$. Furthermore, since $\hat{V}_t$ is the covariance matrix of $\mathrm{X}_t - \hat{R}_{t-1}\hat{S}_{t-1}^{-1}\left(\mathrm{X}_{t-1}', \cdots, \mathrm{X}_1'\right)'$, $\hat{V}_t$ is positive definite for any $t = 1, \cdots, N$.

□

LEMMA 4.1.3. *If* (C1) *and* (C2) *hold,* $\hat{\boldsymbol{\xi}}_t^{(j)} = \boldsymbol{\xi}_t^{(j)} + o_p(1)$.

PROOF. From Lemma 4.1.2, there exist $\hat{S}_t^{-1}$, and $\hat{W}_t$ is well-defined and positive definite for any $t = 1, \cdots, N$. From (C2), we have $\hat{S}_n = S_n + o_p(1)$ and $\hat{R}_n = R_n + o_p(1)$ as $N \to \infty$, and from the perturbation theory (see Kato (1995)) we have $\hat{S}_n^{-1} = S_n^{-1} + o_p(1)$.

Thus

$$(\mathbf{X}_t^{(j)} - \boldsymbol{\mu})^{\sim} = \hat{R}_{t-1}\hat{S}_{t-1}^{-1}\begin{pmatrix}\mathbf{X}_{t-1}^{(j)} - \overline{\mathbf{X}}_N \\ \vdots \\ \mathbf{X}_1^{(j)} - \overline{\mathbf{X}}_N\end{pmatrix}$$

$$= (R_{t-1} + o_{p(1)})\left(S_{t-1}^{-1} + o_{p(1)}\right)\begin{pmatrix}\mathbf{X}_{t-1}^{(j)} - (\boldsymbol{\mu} + o_{p(1)}) \\ \vdots \\ \mathbf{X}_1^{(j)} - (\boldsymbol{\mu} + o_{p(1)})\end{pmatrix}$$

$$= \left\{R_{t-1}S_{t-1}^{-1} + (R_{t-1} + S_{t-1}^{-1})o_{p(1)} + o_{p(1)}\right\}\left\{\begin{pmatrix}\mathbf{X}_{t-1}^{(j)} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{X}_1^{(j)} - \boldsymbol{\mu}\end{pmatrix} + o_{p(1)}\right\}.$$

Note that $R_{t-1}$ and $S_{t-1}^{-1}$ are not random and each element of $\mathbf{X}_t^{(j)} - \boldsymbol{\mu}$ has a finite variance. So these quantities are $O_p(1)$ and thus

$$(\mathbf{X}_t^{(j)} - \boldsymbol{\mu})^{\sim} = \left(R_{t-1}S_{t-1}^{-1} + o_{p(1)}\right)\left[\begin{pmatrix}\mathbf{X}_{t-1}^{(j)} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{X}_1^{(j)} - \boldsymbol{\mu}\end{pmatrix} + o_{p(1)}\right]$$

$$= R_{t-1}S_{t-1}^{-1}\begin{pmatrix}\mathbf{X}_{t-1}^{(j)} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{X}_1^{(j)} - \boldsymbol{\mu}\end{pmatrix} + o_{p(1)}$$

$$= (\mathbf{X}_t^{(j)} - \boldsymbol{\mu})^{\frown} + o_{p(1)}.$$

Similarly we have $\hat{V}_t = V_t + o_{p(1)}$, and thus $\hat{W}_t^{-1} = W_t^{-1} + o_{p(1)}$. Therefore

$$\begin{aligned}\hat{\boldsymbol{\xi}}_t^{(j)} &= \hat{W}_t(\mathbf{X}_t^{(j)} - \overline{\mathbf{X}}_N - (\mathbf{X}_t^{(j)} - \boldsymbol{\mu})^{\sim}) \\ &= \left(W_t^{-1} + o_{p(1)}\right)\left(\mathbf{X}_t^{(j)} - \boldsymbol{\mu} - (\mathbf{X}_t^{(j)} - \boldsymbol{\mu})^{\frown} + o_{p(1)}\right) \\ &= \boldsymbol{\xi}_t^{(j)} + o_{p(1)}.\end{aligned}$$

□

REMARK. Lemma 4.1.3 may not be true when $t$ depends on $N$.

Define $\bar{\mu}_k^{(j)}$, $\bar{\gamma}(h)^{(j)}$, $\mathbf{Z}_L^{(j)}$ and $I_L^{(j)}(\lambda)$ as follows:

$$\bar{\mu}_k^{(j)} = L^{-1}\sum_{t=1}^{L}(\xi_t^{(j)})^k, \qquad \bar{\gamma}(h)^{(j)} = L^{-1}\sum_{t=1}^{L-h}\xi_{t+h}^{(j)}\xi_t^{(j)},$$

$$\mathbf{Z}_L^{(j)} = \sqrt{L}\left(\bar{\mu}_1^{(j)}, \bar{\mu}_2^{(j)} - 1, \bar{\gamma}(1)^{(j)}, \cdots, \bar{\gamma}(K)^{(j)}\right)'$$

$$I_L^{(j)}(\lambda) = L^{-1}\left|\sum_{t=1}^{L}\xi_t^{(j)}\exp^{-itg(L,\lambda)}\right|^2.$$

Where $g(n,\omega)$ is the multiple of $2\pi/n$ closest to $\omega$ (the smaler one if there are two) if $\omega \in [0,\pi]$, and $g(n,-\omega)$ if $\omega \in [-\pi,0)$.

LEMMA 4.1.4. *If* (C1) *and* (C2) *hold, we have*

$$\widehat{\mathbf{Z}}_L^{(j)} = \mathbf{Z}_L^{(j)} + o_p(1) \qquad \text{and} \qquad \hat{I}_L^{(j)}(\lambda) = I_L^{(j)}(\lambda) + o_p(1),$$

*as* $N \to \infty$.

PROOF. The lemma is straightforward from Lemma 4.1.3 and the fact that the functions

$$L^{-1}\sum_{t=1}^{L} x_t^k, \quad L^{-1}\sum_{t=1}^{L-h} x_{t+h}x_t \quad \text{and} \quad L^{-1}\left|\sum_{t=1}^{L} x_t \exp^{-itg(L,\lambda)}\right|$$

are continuous on $\mathrm{R}^L$ to R.

□

## 4.2. Asymptotic distributions of the statistics

In this subsection we show the asymptotic distributions of the test statistics, which consist of true $\xi_t^{(j)}$ when $L \to \infty$ as $N \to \infty$.

Define $\tilde{\gamma}(h)^{(j)}$ and $\mathbf{Y}_L^{(j)}$ as follows:

$$\tilde{\gamma}(h)^{(j)} = L^{-1}\sum_{t=1}^{L} \xi_{t+h}^{(j)}\xi_t^{(j)}, \quad \mathbf{Y}_L^{(j)} = \sqrt{L}\left(\bar{\mu}_1^{(j)}, \bar{\mu}_2^{(j)} - 1, \tilde{\gamma}(1)^{(j)}, \cdots, \tilde{\gamma}(K)^{(j)}\right)'.$$

Denote $N_p(\boldsymbol{\eta},\Psi)$ by the $p$-dimensinal normal distribution with mean vector $\boldsymbol{\eta}$ and covariance matrix $\Psi$ and by $I_p$ the $p\times p$ unit matrix.

LEMMA 4.2.1. *If* (C5) *holds, then under the null hypoyhesis we have*

$$\mathbf{Y}_L^{(j)} \xrightarrow{\mathcal{D}} N_{K+2}(0,\Sigma) \qquad \text{as} \quad L \to \infty.$$

*The covariance matrix* $\Sigma$ *is as follows:*

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & I_K \end{pmatrix} = \begin{pmatrix} 1 & \mu_3 & 0 \\ \mu_3 & \mu_4 - 1 & \\ 0 & & I_K \end{pmatrix} \tag{4.3}$$

*where* $\mu_k := E[(\xi_t^{(j)})^k]$ *for* $k = 3,4$.

PROOF. For any $\boldsymbol{\lambda} = (\lambda_1,\cdots,\lambda_{K+2})' \in \mathrm{R}^{K+2}$

$$\begin{aligned} \boldsymbol{\lambda}'\mathbf{Y}_L^{(j)} &= \boldsymbol{\lambda}'\left\{L^{-1/2}\sum_{t=1}^{L}\left(\xi_t^{(j)}, (\xi_t^{(j)})^2 - 1, \xi_{t+1}^{(j)}\xi_t^{(j)}, \cdots, \xi_{t+K}^{(j)}\xi_t^{(j)}\right)'\right\} \\ &= L^{-1/2}\sum_t(\lambda_1\xi_t^{(j)} + \lambda_2((\xi_t^{(j)})^2 - 1) + \sum_{h=1}^{K}\lambda_{h+2}\xi_{t+h}^{(j)}\xi_t^{(j)}). \end{aligned}$$

Now, put $\eta_t := \eta_t^{(j)} = \lambda_1\xi_t^{(j)} + \lambda_2((\xi_t^{(j)})^2 - 1) + \sum_{h=1}^K \lambda_{h+2}\xi_{j+h}^{(j)}\xi_t^{(j)}$ then $\{\eta_t\}_{t=1}^L$ is a strictly stationary $K$-dependent sequence since $\{\xi_t^{(j)}\}$ is $IID(0,1)$. Moreover, it follows that

$$\mathbf{Y}_L^{(j)}\mathbf{Y}_L^{(j)\prime} = L^{-1}\sum_{s,t=1}^{L} \Xi_{s,t}^{(j)} ,$$

where $\Xi_{s,t}^{(j)}$ is the matrix such as follows:

$$\begin{pmatrix} \xi_s^{(j)}\xi_t^{(j)} & \xi_s^{(j)}((\xi_t^{(j)})^2-1) & \xi_s^{(j)}\xi_{t+1}^{(j)}\xi_t^{(j)} & \cdots & \xi_s^{(j)}\xi_{t+K}^{(j)}\xi_t^{(j)} \\ & ((\xi_s^{(j)})^2-1)((\xi_t^{(j)})^2-1) & ((\xi_s^{(j)})^2-1)\xi_{t+1}^{(j)}\xi_t^{(j)} & \cdots & ((\xi_s^{(j)})^2-1)\xi_{t+K}^{(j)}\xi_t^{(j)} \\ & & \xi_{s+1}^{(j)}\xi_s^{(j)}\xi_{t+1}^{(j)}\xi_t^{(j)} & \cdots & \xi_{s+1}^{(j)}\xi_s^{(j)}\xi_{t+K}^{(j)}\xi_t^{(j)} \\ & & & \ddots & \vdots \\ & & & & \xi_{s+K}^{(j)}\xi_s^{(j)}\xi_{t+K}^{(j)}\xi_t^{(j)} \end{pmatrix}$$

and that

$$E[\Xi_{s,t}^{(j)}] = \delta_{s,t}\Sigma.$$

Therefore we have

$$E[\mathbf{Y}_L^{(j)}] = 0, \qquad Var[\mathbf{Y}_L^{(j)}] = \Sigma.$$

Thus for $\bar{\eta}_L := L^{-1}\sum_{t=1}^L \eta_t = L^{-1/2}\boldsymbol{\lambda}'\mathbf{Y}_L^{(j)}$, we have

$$E[\bar{\eta}_L] = L^{-1/2}\boldsymbol{\lambda}'E[\mathbf{Y}_L^{(j)}] \;=\; \mathbf{0}$$

and

$$Var[\bar{\eta}_L] = (L^{-1/2}\boldsymbol{\lambda}')Var[\mathbf{Y}_L^{(j)}](L^{-1/2}\boldsymbol{\lambda}) \;=\; L^{-1}\boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}.$$

By the central limit theorem for strictly stationary $K$-dependent sequence, it follows that

$$\sqrt{L}\bar{\eta}_L = \boldsymbol{\lambda}'\mathbf{Y}_L^{(j)} \xrightarrow{\mathcal{D}} N(0, \boldsymbol{\lambda}'\Sigma\boldsymbol{\lambda}).$$

Therefore, using the Cramér-Wold device, we have

$$\mathbf{Y}_L^{(j)} \xrightarrow{\mathcal{D}} N_{K+2}(\mathbf{0}, \Sigma).$$

□

LEMMA 4.2.2. *Under the same conditions of Lemma 4.2.1 and under the null hypothesis, we have*

$$\mathbf{Z}_L^{(j)} = \mathbf{Y}_L^{(j)} + o_p(1) \qquad as \quad L \to \infty$$

PROOF. It suffices to show $L^{-1/2}\sum_{t=L-h+1}^{L}\xi_{t+h}^{(j)}\xi_t^{(j)} = o_p(1)$ for any $h = 1, \cdots, K$ since

$$\mathbf{Y}_L^{(j)} - \mathbf{Z}_L^{(j)} = L^{-1/2}(0, 0, \xi_{L+1}^{(j)}\xi_L^{(j)}, \cdots, \sum_{t=L-h+1}^{L}\xi_{t+h}^{(j)}\xi_t^{(j)}, \cdots, \sum_{t=L-K+1}^{L}\xi_{t+K}^{(j)}\xi_t^{(j)})'.$$

However we have

$$\begin{aligned} E[\{L^{-1/2}\textstyle\sum_{t=L-h+1}^{L}\xi_{t+h}^{(j)}\xi_t^{(j)}\}^2] &= L^{-1}\sum_{s,t=L-h+1}^{L}E[\xi_{s+h}^{(j)}\xi_s^{(j)}\xi_{t+h}^{(j)}\xi_t^{(j)}] \\ &= L^{-1}\sum_{t=L-h+1}^{L}E[(\xi_{t+h}^{(j)})^2(\xi_t^{(j)})^2] \\ &= \frac{h}{L} \to 0 \qquad \text{as} \quad L \to \infty. \end{aligned}$$

Thus $L^{-1/2}\sum_{t=L-h+1}^{L}\xi_{t+h}^{(j)}\xi_t^{(j)}$ converges to 0 in mean square, thus in probability.

□

From Lemma 4.2.1 and 4.2.2, we have the next lemma.

LEMMA 4.2.3. *Under the same conditions of Lemma 4.2.1 and under the null hypothesis, we have*

$$\mathbf{Z}_L^{(j)} \xrightarrow{\mathcal{D}} N_{K+2}(0, \Sigma) \qquad \text{as} \quad L \to \infty$$

LEMMA 4.2.4. *If* (C5) *holds, the matrix* $\Sigma$ *given in* (4.3) *is positive definite.*

PROOF. $\Sigma$ is well-deifined from (C5) and it is clear from the proof of Lemma 4.2.1 that $\Sigma$ is non-negative definite. Thus it suffices to show that $\det\Sigma = \mu_4 - 1 - \mu_3^2 \neq 0$. Suppose that $\mu_4 - 1 - \mu_3^2 \neq 0$ then we have

$$\begin{aligned} Var\left[-\mu_3\xi_t^{(j)} + (\xi_t^{(j)})^2 - 1\right] &= E\left[\mu_3^2\xi_t^{(j)} - 2\mu_3\xi_t^{(j)}((\xi_t^{(j)})^2 - 1) + (\xi_t^{(j)})^4 - 2(\xi_t^{(j)})^2 + 1\right] \\ &= \mu_4 - 1 - \mu_3^2 \ = \ 0 \end{aligned}$$

Thus $-\mu_3\xi_t^{(j)} + (\xi_t^{(j)})^2 - 1 = 0$ *a.e.* It shows $\xi_t^{(j)} = C$ *a.e.* for some constant $C$, but it is a contradiction since $E[\xi_t^{(j)}] = 0$ and $E[(\xi_t^{(j)})^2] = 1$.

□

LEMMA 4.2.5. *Define* $\widetilde{\Sigma}_L^{(j)^*}$ *as follows:*

$$\widetilde{\Sigma}_L^{(j)^*} = \begin{pmatrix} \bar{\mu}_2^{(j)} & \bar{\mu}_3^{(j)} - \bar{\mu}_1^{(j)} \\ \bar{\mu}_3^{(j)} - \bar{\mu}_1^{(j)} & \bar{\mu}_4^{(j)} - 2\bar{\mu}_2^{(j)} + 1 \end{pmatrix}.$$

*If* (C5) *holds,* $\widetilde{\Sigma}_L^{(j)^*}$ *is positive definite and* $\widetilde{\Sigma}_L^{(j)^*} = \Sigma_1 + o_p(1)$ *as* $L \to \infty$*, where* $\Sigma_1$ *is given in* (4.3).

PROOF. It is straightforward to show $\widetilde{\Sigma}_L^{(j)*} = \Sigma_1 + o_p(1)$ by the weak law of large number. We show $\widetilde{\Sigma}_L^{(j)*}$ is positive definite. Put $\boldsymbol{x} = \left(\xi_1^{(j)}, \cdots, \xi_L^{(j)}\right)'$ and $\boldsymbol{y} = \left((\xi_1^{(j)})^2 - 1, \cdots, (\xi_L^{(j)})^2 - 1\right)'$ then

$$\widetilde{\Sigma}_L^{(j)*} = \begin{pmatrix} L^{-1}\|\boldsymbol{x}\|^2 & L^{-1}\langle \boldsymbol{x}, \boldsymbol{y}\rangle \\ L^{-1}\langle \boldsymbol{x}, \boldsymbol{y}\rangle & L^{-1}\|\boldsymbol{y}\|^2 \end{pmatrix}.$$

By the Cauchy-Schwartz inequality we have

$$\det \widetilde{\Sigma}_L^{(j)*} = L^{-2}\left(\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 - |\langle \boldsymbol{x}, \boldsymbol{y}\rangle|^2\right) \geq 0.$$

The equality is satisfied if and only if $\boldsymbol{y} = (\langle \boldsymbol{x}, \boldsymbol{y}\rangle / \|\boldsymbol{x}\|^2)\boldsymbol{x}$. Thus it suffices to show

$$\boldsymbol{y} \neq \bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1)\boldsymbol{x}. \tag{4.4}$$

Suppose that (4.4) is not true, then there exists $t$ such that

$$(\xi_t^{(j)})^2 - 1 = \bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1)\xi_t^{(j)}.$$

Put $t = 1$ without loss of generality. Then we have, for any $\varepsilon_1,\ \varepsilon_2 > 0$,

$$\begin{aligned} \Pr\left(|(\xi_1^{(j)})^2 - 1 - \mu_3\xi_1^{(j)}| > \varepsilon_1\right) &= \Pr\left(|\{\bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1) - \mu_3\}\xi_1^{(j)}| > \varepsilon_1\right) \\ &\leq \Pr\left(|\bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1) - \mu_3| > \varepsilon_1\varepsilon_2\right) + \Pr\left(|\xi_1^{(j)}| > \varepsilon_2^{-1}\right) \\ &\leq \Pr\left(|\bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1) - \mu_3| > \varepsilon_1\varepsilon_2\right) + \varepsilon_2^2 E[(\xi_1^{(j)})^2] \\ &= \Pr\left(|\bar{\mu}_2^{-1}(\bar{\mu}_3 - \bar{\mu}_1) - \mu_3| > \varepsilon_1\varepsilon_2\right) + \varepsilon_2^2 \end{aligned} \tag{4.5}$$

The first term of the right hand side of (4.5) converges to zero as $L \to \infty$ and the second term is arbitrary small. Therefore $(\xi_1^{(j)})^2 - 1 - \mu_3\xi_1^{(j)} = 0$ *a.e*, but we may show that it is contradiction by the same argument as that in the proof of Lemma 4.2.4. So $\widetilde{\Sigma}_L^{(j)*}$ must be positive definite.

□

From Lemma 4.2.3, 4.2.4 and 4.2.5 we have the next theorem.

THEOREM 4.2.6. *Define* $\widetilde{\Sigma}_L^{(j)}$ *as follows:*

$$\widetilde{\Sigma}_L^{(j)} = \begin{pmatrix} \widetilde{\Sigma}_L^{(j)*} & 0 \\ 0 & I_K \end{pmatrix}.$$

*If* (C5) *hold,* $(\mathbf{Z}_L^{(j)})'(\widetilde{\Sigma}_L^{(j)})^{-1}\mathbf{Z}_L^{(j)}$ *converges in distribution to chi-square distribution with* $(K+2)$*-degree of freedom under the null hypothesis.*

Remainder of this subsection is for the SPED procedure. The next lemma gives an inequality between the fourth moment of $\xi_{t,k}^{(j)}$ and $X_{t,k}^{(j)}$.

LEMMA 4.2.7. *If* (C1) *and* (C4) *hold,* $E[(\xi_{t,k}^{(j)})^4] \leq C\sum_{k=1}^{d} E[(X_{t,k}^{(j)})^4]$ *for some constant* $C$.

PROOF. We write $\boldsymbol{\xi}_t^{(j)}$ and $\mathbf{X}_t^{(j)}$ by $\boldsymbol{\xi}_t$ and $\mathbf{X}_t$ for short. We can assume $E[\mathbf{X}_t] = 0$ without loss of generality. From condition (C1), $\boldsymbol{\xi}_t$ is well-defined. Recall that $\hat{\mathbf{X}}_{t+1}$ is the best linear predictor of $\mathbf{X}_{t+1}$ based on $\mathbf{X}_1, \cdots, \mathbf{X}_t$. Thus we may represent $\hat{\mathbf{X}}_{t+1}$ by

$$\hat{\mathbf{X}}_{t+1} = -\Phi_{t,1}\mathbf{X}_t - \Phi_{t,2}\mathbf{X}_{t-1} - \cdots - \Phi_{t,t}\mathbf{X}_1.$$

Therefore

$$\begin{bmatrix} \boldsymbol{\xi}_{t+1} \\ \boldsymbol{\xi}_t \\ \vdots \\ \boldsymbol{\xi}_1 \end{bmatrix} = \begin{bmatrix} W_{t+1}^{-1} & & & \\ & W_t^{-1} & & \\ & & \ddots & \\ & & & W_1^{-1} \end{bmatrix} \begin{bmatrix} I & \Phi_{t,1} & \cdots & \Phi_{t,t} \\ & I & \ddots & \vdots \\ & & \ddots & \Phi_{1,1} \\ & & & I \end{bmatrix} \begin{bmatrix} \mathbf{X}_{t+1} \\ \mathbf{X}_t \\ \vdots \\ \mathbf{X}_1 \end{bmatrix}.$$

Put

$$\begin{bmatrix} I & \Theta_{t,1} & \cdots & \Theta_{t,t} \\ & I & \ddots & \vdots \\ & & \ddots & \Theta_{1,1} \\ & & & I \end{bmatrix} = \begin{bmatrix} I & \Phi_{t,1} & \cdots & \Phi_{t,t} \\ & I & \ddots & \vdots \\ & & \ddots & \Phi_{1,1} \\ & & & I \end{bmatrix}^{-1}$$

then we have $\mathbf{X}_{t+1} = W_{t+1}\boldsymbol{\xi}_{t+1} + \mathbf{Y}_{t+1}$, where $\mathbf{Y}_{t+1} = \sum_{j=1}^{t} \Theta_{t,j} W_{t+1-j}\boldsymbol{\xi}_{t+1-j}$. It follows that

$$\begin{aligned}
(\mathbf{X}_{t+1}'\mathbf{X}_{t+1})^2 &= \left\{(W_{t+1}\boldsymbol{\xi}_{t+1} + \mathbf{Y}_{t+1})'(W_{t+1}\boldsymbol{\xi}_{t+1} + \mathbf{Y}_{t+1})\right\}^2 \\
&= (\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1})^2 + 4(\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1})^2 + (\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1})^2 \\
&\qquad + 2\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1}\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1} + 4\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1} \\
&\qquad + 4\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1}\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1}.
\end{aligned}$$

Because $\{\boldsymbol{\xi}_t\}$ is independent, $\boldsymbol{\xi}_{t+1}$ and $\mathbf{Y}_{t+1}$ are independent. Therefore

$$\begin{aligned}
E[\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1}] &= E[\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1}\boldsymbol{\xi}_{t+1}'W_{t+1}']E[\mathbf{Y}_{t+1}] = 0. \\
E[\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1}\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1}] &= E[\boldsymbol{\xi}_{t+1}']E[W_{t+1}'\mathbf{Y}_{t+1}\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1}] = 0.
\end{aligned}$$

So we have

$$\begin{aligned}
E[(\mathbf{X}_{t+1}'\mathbf{X}_{t+1})^2] &= E[(\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1})^2] + 4E[(\boldsymbol{\xi}_{t+1}'W_{t+1}'\mathbf{Y}_{t+1})^2] \\
&\qquad + E[(\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1})^2] + 2E[\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1}]E[\mathbf{Y}_{t+1}'\mathbf{Y}_{t+1}] \\
&\geq E[(\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1})^2].
\end{aligned}$$

By the definition of $\tau_t$ and $\tau$, we have next inequalities.

$$E[(\boldsymbol{\xi}_{t+1}'V_{t+1}\boldsymbol{\xi}_{t+1})^2] \geq E[(\tau_t\boldsymbol{\xi}_{t+1}'\boldsymbol{\xi}_{t+1})^2] \geq \tau^2 E[(\boldsymbol{\xi}_{t+1}'\boldsymbol{\xi}_{t+1})^2] \geq \tau^2 E[(\xi_{t+1,k})^4].$$

Thus

$$\tau^2 E[(\xi_{t+1,k})^4] \le E[(\mathbf{X}'_{t+1}\mathbf{X}_{t+1})^2] = E[(\sum_{k=1}^{d}(X_{t+1,k})^2)^2]$$
$$\le E[d\sum_{k=1}^{d}(X_{t+1,k})^4].$$

We have $\tau \neq 0$ from the condition (C1) and therefore, for some constant $C$ we have

$$E[(\xi_{t+1,k})^4] \le C\sum_{k=1}^{d} E[(X_{t+1,k})^4] \qquad k = 1, 2, \cdots d.$$

□

Next theorem is similar to Proposition 10.3.2 of Brockwell and Davis (1991), where considered is an independent and identically sequence. Here we consider a sequence which is independent but may not be identical.

THEOREM 4.2.8. *Let $E_1, \cdots, E_k$ be distributed independently and identically as an exponential distribution with mean 1. If* (C1), (C3) *and* (C4) *hold, we have under the null hypothesis*

$$(I_L^{(j)}(\lambda_1), \cdots, I_L^{(j)}(\lambda_k))' \xrightarrow{\mathcal{D}} (E_1, \cdots, E_k)' \qquad as \quad M \to \infty$$

*where $I_L(\cdot)^{(j)}$ is a periodogram of $\{\xi_t^{(j)}\}_{t=1}^L = \{\xi_{1,1}^{(j)}, \cdots, \xi_{1,d}^{(j)}, \cdots, \xi_{M,1}^{(j)}, \cdots, \xi_{M,d}^{(j)}\}$.*

PROOF. For simplicity we suppres index $(j)$. Define $\alpha(\lambda)$ and $\beta(\lambda)$ for any $\lambda \in (0, \pi)$ as follows:

$$\alpha(\lambda) = (2/L)^{1/2}\sum_{t=1}^{L} \xi_t \cos(t\, g_{(L,\lambda)}), \qquad \beta(\lambda) = (2/L)^{1/2}\sum_{t=1}^{L} \xi_t \sin(t\, g_{(L,\lambda)}).$$

where $g(\cdot,\cdot)$ is defined in the proof of Lemma 4.1.4 and $L = dM$. We have $2I_L(\lambda) = \alpha(\lambda)^2 + \beta(\lambda)^2$.
Put $\phi_{t,k} = d(t-1) + k$ then since $\{\xi_t\}_{t=1}^L = \{\xi_{1,1}, \cdots, \xi_{1,d}, \cdots, \xi_{M,1}, \cdots, \xi_{M,d}\}$ we have

$$\sum_{t=1}^{L} \xi_t \cos(t\omega) = \sum_{t=1}^{M}\sum_{k=1}^{d} \xi_{t,k} \cos(\phi_{t,k}\omega).$$

Put $\mathcal{Z}_L = (\alpha(\lambda_1), \beta(\lambda_1), \cdots, \alpha(\lambda_k), \beta(\lambda_k))'$. Let $U_1, U_2, \cdots, U_{2k}$ be random variables distributed independently and identically as a standard normal distribution. It suffices to show that $\mathcal{Z}_L$ converges in distribution to $\mathbf{U} = (U_1, U_2, \cdots, U_{2k})'$ since the function

$$h(u_1, \cdots, u_{2k}) = ((u_1^2 + u_2^2)/2, \cdots, (u_{2k-1}^2 + u_{2k}^2)/2)'$$

is continuous on $\mathrm{R}^{2k}$ to $\mathrm{R}^k$ and $(U_i^2 + U_{i+1}^2)/2$ is distributed exponentially with mean 1. Independence of $E_j$ is followed by the independence of $U_j$. We prove it using the

Cramér-Wold device. For any $\boldsymbol{\zeta} = (\zeta_1, \cdots, \zeta_{2k})' \in \mathrm{R}^{2k}$ we have

$$\begin{aligned}
\boldsymbol{\zeta}'\mathcal{Z}_L &= \zeta_1\alpha(\lambda_1) + \zeta_2\beta(\lambda_1) + \cdots \zeta_{2k-1}\alpha(\lambda_k) + \zeta_{2k}\beta(\lambda_k) \\
&= (2/L)^{1/2}\sum_{t=1}^{M}\sum_{k=1}^{d}\xi_{t,k}\{\zeta_1\cos(\phi_{t,k}g(L,\lambda_1)) + \zeta_2\sin(\phi_{t,k}g(L,\lambda_1)) + \cdots \\
&\qquad \cdots + \zeta_{2k-1}\cos(\phi_{t,k}g(L,\lambda_k)) + \zeta_{2k}\sin(\phi_{t,k}g(L,\lambda_k))\}.
\end{aligned}$$

Let $\psi_{t,k}$ be the coefficient of $\xi_{t,k}$ and put $\eta_t = d^{-1/2}\sum_{k=1}^{d}\psi_{t,k}\xi_{t,k}$. Then

$$\boldsymbol{\zeta}'\mathcal{Z}_L = (2/M)^{1/2}\sum_{t=1}^{M}\eta_t \quad (4.6) \qquad \text{and} \qquad |\psi_{t,k}| \le \sum_{j=1}^{2k}|\zeta_j|. \quad (4.7)$$

Put $B_M^2 = \sum_{t=1}^{M} Var[\eta_t]$. Recall that $\{\boldsymbol{\xi}_t\}$ follows $WN(0, I)$ under the null hypothesis, we have

$$E[\eta_t] = 0, \qquad E[\eta_t^2] = d^{-1}\sum_{k=1}^{d}\psi_{t,k}^2.$$

Furthermore, for sufficiently large $M$, that is, large $L$, we may consider $0 < g(L,\lambda_1) < \cdots < g(L,\lambda_k) < \pi$. Namely there exist the Fourier frequencies $\{\omega_{n_i} | \omega_{n_i} = g(L,\lambda_i),\ i = 1, \cdots, k\}$. Thus $B_M^2$ is represented as follows:

$$\begin{aligned}
B_M^2 &= d^{-1}\sum_{t=1}^{M}\sum_{k=1}^{d}\psi_{t,k}^2 \\
&= d^{-1}\sum_{t=1}^{M}\sum_{k=1}^{d}\{\zeta_1\cos(\phi_{t,k}g(L,\lambda_1)) + \cdots + \zeta_{2k}\sin(\phi_{t,k}g(L,\lambda_k))\}^2 \\
&= d^{-1}\sum_{t=1}^{L}(\zeta_1\cos t\omega_{n_1} + \cdots + \zeta_{2k}\sin t\omega_{n_k})^2.
\end{aligned}$$

From the fundamental property of the Fourier frequencies, we have

$$(2/L)\sum_{t=1}^{L}\cos t\omega_i\,\cos t\omega_j = \delta_{ij}\ , \qquad (2/L)\sum_{t=1}^{L}\sin t\omega_i\,\sin t\omega_j = \delta_{ij}\ ,$$

$$(2/L)\sum_{t=1}^{L}\sin t\omega_i\,\cos t\omega_j = 0.$$

Thus

$$\begin{aligned}
B_M^2 &= d^{-1}(\zeta_1^2\sum_{t=1}^{L}\cos^2 t\omega_{n_1} + \cdots + \zeta_{2k}^2\sum_{t=1}^{L}\sin^2 t\omega_{n_k}) \\
&= (2d)^{-1}L\sum_{j=1}^{2k}\zeta_j^2 = \|\boldsymbol{\zeta}\|^2\, M/2\ .
\end{aligned}$$

Substituing this equality to (4.6) we have

$$\zeta' \mathcal{Z}_L = \|\zeta\| B_M^{-1} \sum_{t=1}^{M} \eta_t .$$

Since $\zeta'\mathbf{U}$ follows the normal distribution with mean zero and variance $\|\zeta\|^2$, if we may show $B_M^{-2}\sum_{t=1}^{M}\eta_t^2$ converges to 1 in probability, $\zeta'\mathcal{Z}_L$ converges to $\zeta'\mathbf{U}$ in distribution by Theorem 2 of Raikov (1938). We apply the Chebyshev theorem to show the convergence of $B_M^{-2}\sum_{t=1}^{M}\eta_t^2$. By the Cauchy-Schwartz inequality and (4.7) it follows that

$$\begin{aligned}
E[(\eta_t)^4] &= E[(d^{-1/2}\sum_{k=1}^{d}\psi_{t,k}\xi_{t,k})^4] \\
&= d^{-2}\sum_{k,l,m,n=1}^{d}\psi_{t,k}\psi_{t,l}\psi_{t,m}\psi_{t,n}E[\xi_{t,k}\xi_{t,l}\xi_{t,m}\xi_{t,n}] \\
&\le d^{-2}\sum|\psi_{t,k}||\psi_{t,l}||\psi_{t,m}||\psi_{t,n}|\sqrt{E[(\xi_{t,k})^2(\xi_{t,l})^2]E[(\xi_{t,m})^2(\xi_{t,n})^2]} \\
&\le d^{-2}\sum C\sqrt[4]{E[(\xi_{t,k})^4]E[(\xi_{t,l})^4]E[(\xi_{t,m})^4]E[(\xi_{t,n})^4]}.
\end{aligned}$$

From Lemma 4.2.7

$$E[(\eta_t)^4] \le C\sum_{k=1}^{d}E[(X_{t,k})^4], \quad \text{where } C \text{ is a some constant.}$$

Thus from the condition (C3), we have

$$\begin{aligned}
M^{-2}\sum_{t=1}^{M}E[(\eta_t)^4] \le C\,M^{-2}\sum_{t=1}^{M}\sum_{k=1}^{d}E[(X_{t,k})^4] &= C\sum_{k=1}^{d}\left(M^{-2}\sum_{t=1}^{M}E[(X_{t+j-1,k})^4]\right) \\
&\le C\sum_{k=1}^{d}\left(M^{-2}\sum_{t=1}^{M+j-1}E[(X_{t,k})^4]\right) \to 0 \qquad \text{as} \quad M\to\infty.
\end{aligned}$$

Since $E[B_M^2\sum_{t=1}^{M}\eta_t^2]=1$, the convergence of $B_M^2\sum_{t=1}^{M}\eta_t^2$ to 1 in probability, and that of $\zeta'\mathcal{Z}_L$ to $\zeta'\mathbf{U}$ in distribution, is proved. Thus we have the theorem by the Cramér-Wold device.

□

COROLLARY 4.2.9. *Under the conditions of Theorem 4.2.8 and under the null hypothesis, we have*

$$\max_{1\le k\le n} I_L^{(j)}(\lambda_k) \xrightarrow{\mathcal{D}} E$$

*where $E$ is the random variables with the distribution function $G_n(x) = (1-\exp^{-x})^n$.*

### 4.3. Approximate distributions of test statistics

The test statistics for MON and SPED procedures are given respectively by

$$T_{MON}(j) = (\widehat{\mathbf{Z}}_L^{(j)})'(\widehat{\Sigma}_L^{(j)})^{-1}\widehat{\mathbf{Z}}_L^{(j)}$$

and

$$T_{SPED}(j) = \max_{1\leq k\leq \mathcal{L}} \hat{I}_L^{(j)}(\omega_k).$$

Assuming condition (C1), (C2) and (C5) it is shown in Lemma 4.1.4 that $\widehat{\mathbf{Z}}_L^{(j)}$ converges in probability to $\mathbf{Z}_L^{(j)}$ as $N \to \infty$. Also we may show by the same lemma that $\widehat{\Sigma}_L^{(j)}$ converges to $\widetilde{\Sigma}_L^{(j)}$ in probability as $N \to \infty$. Furthermore from Theorem 4.2.6, the distribution of $(\mathbf{Z}_L^{(j)})'(\widetilde{\Sigma}_L^{(j)})^{-1}\mathbf{Z}_L^{(j)}$ is approximated by a chi-square distribuiton with $(K+2)$ degree of freedom when $L$ is large. Thus under the null hypothesis $T_{MON}(j)$ follows a chi-square distribution with $(K+2)$ degree of freedom approximately when $N$ and $L$ are large.

For $T_{SPED}(j)$, assuming conditions (C1), (C2), (C3) and (C4) it is shown that $\hat{I}_L^{(j)}(\omega_k)$ converges in probability to $I_L^{(j)}(\omega_k)$ as $N \to \infty$. Furthermore, it is shown in Corollary 4.2.9 that the distribution function of $\max_{1\leq k\leq \mathcal{L}} I_L^{(j)}(\omega_k)$ is approximated by $G_{\mathcal{L}}(x) = (1 - \exp^{-x})^{\mathcal{L}}$. Therefore the distribution of $T_{SPED}(j)$ is approximated by $G_{\mathcal{L}}(x)$ under the null hypothesis when $N$ and $L$ are large.

## 5. Simulation and applications

In this sections, we show the results of simulations, and applications of our proposed procedure to practical data. This will demonstrate the usefulness of our proposed procedure.

### 5.1. Simulation

Simulation is conducted to (a) compare the MON procedure and SPED procedure, (b) to examine the selection of blocks, i.e. all blocks v.s. selected blocks, and (c) to compare the Box and Jenkins (1970) and Okabe and Nakano (1991) criteria for selecting $M$ and $K$. Note that since the test are undertaken in each block, the MON Type A and SPED Type A procedures consist of $(N-M+1)$ tests, and the MON Type B and SPED Type B procedures consist of $J$ tests. The Type I familywise error adjusting for the multiplicity of the test is selected at $\alpha = 0.20$ in the simulation. Sample size considered are $N = 100, 200,$ and $400$, and $M$ and $K$ by the Box and Jenkins (1970) criterion are chosen by $M = [N/4]$ and $K = [L/4]$, and by the Okabe and Nakano (1991) criterion is decided by $M = [3\sqrt{N}/d]$ and $K = [2\sqrt{L}]$, where $[x]$ is the integer part of $x$. As is seen below the models employed are the case of $d = 1$, thus the values of $(N, M, K)$ used are

$$(N, M, K) = (100, 25, 6), (200, 50, 12) \text{ and } (400, 100, 25)$$

in the Box and Jenkins (1970) criterion and

$$(N, M, K) = (100, 30, 10), (200, 42, 12) \text{ and } (400, 60, 15)$$

in the Okabe and Nakano (1991) criterion. The models employed are as follows:

**ARMA(2,2)** $(1 - B/\alpha_1)(1 - B/\alpha_2)X_t = (1 - B/\beta_1)(1 - B/\beta_2)Z_t.$

**ARIMA(2,1,2)** $(1 - B/\alpha_1)(1 - B/\alpha_2)(1 - B)X_t = (1 - B/\beta_1)(1 - B/\beta_2)Z_t.$

Here $Z_t$ is the i.i.d. random variables from the standard normal distribution and $B$ is the the backward shift operator. We generated 300 data sets for each sample size from the models with changing the values of $\alpha_1$, $\alpha_2$, $\beta_1$ and $\beta_2$. The values of $\alpha_j$ $(j = 1, 2)$ and $\beta_j$ $(j = 1, 2)$ are generated from $U(1, 2)$ and $U(0, 2)$, where $U(a, b)$ denotes the uniform distribution over $(a, b)$. It is well known that the ARMA model is stationary, whereas the ARIMA model is not.

TABLE 1 and 2 list the number of trials that rejected the null hypoyhesis among 300 trials for each procedure in the ARMA model, thus show the empirical levels of the procedures. TABLE 3 and 4 are corresponding tables to TABLE 1 and 2 in ARIMA model, showing the empirical power of each procedure. TABLE 1 and 3 are for the Box and Jenkins (1970) criterion, i.e. $M = [N/4]$, and TABLE 2 and 4 are for the Okabe and Nakano (1991) criterion, i.e. $M = [3\sqrt{N}/d]$.

| | 100 | 200 | 400 |
|---|---|---|---|
| MON Type A | 17 (0.057) | 20 (0.067) | 6 (0.020) |
| MON Type B | 35 (0.12) | 25 (0.083) | 11 (0.037) |
| SPED Type A | 19 (0.063) | 36 (0.12) | 36 (0.12) |
| SPED Type B | 47 (0.16) | 72 (0.24) | 103 (0.34) |

TABLE 1: Number of rejected trials among 300 and its frequency for ARMA model, $M = [N/4]$

| | 100 | 200 | 400 |
|---|---|---|---|
| MON Type A | 9 (0.030) | 16 (0.053) | 17 (0.057) |
| MON Type B | 26 (0.087) | 26 (0.087) | 23 (0.077) |
| SPED Type A | 30 (0.10) | 30 (0.10) | 25 (0.083) |
| SPED Type B | 49 (0.16) | 47 (0.16) | 40 (0.13) |

TABLE 2: Number of rejected trials among 300 and its frequency for ARMA model, $M = [3\sqrt{N}/d]$

| | 100 | 200 | 400 |
|---|---|---|---|
| MON Type A | 107 (0.36) | 134 (0.45) | 137 (0.46) |
| MON Type B | 152 (0.51) | 147 (0.49) | 143 (0.48) |
| SPED Type A | 132 (0.44) | 169 (0.56) | 184 (0.61) |

TABLE 3: Number of rejected trials among 300 and its frequency for ARIMA model, $M = [N/4]$

| | 100 | 200 | 400 |
|---|---|---|---|
| MON Type A | 91 (0.30) | 147 (0.49) | 143 (0.48) |
| MON Type B | 137 (0.46) | 166 (0.55) | 154 (0.51) |
| SPED Type A | 150 (0.50) | 181 (0.60) | 178 (0.59) |
| SPED Type B | 160 (0.53) | 206 (0.69) | 186 (0.62) |

TABLE 4: Number of rejected trials among 300 and its frequency for ARIMA model, $M = [3\sqrt{N}/d]$

TABLE 1 shows that the level of the SPED Type B procedure with $M = [N/4]$ exceeds the nominal familywise error, and we omit it from further comparison. TABLE 1 and 2 show that the levels of all the other procedures are less than the nomial familywise error; and that the MON procedures are more conservative than the SPED procedures. TABLE 3 and 4 show that the SPED Type B with $M = [3\sqrt{N}/d]$ has the largest powers at every sample sizes; that it is followed by the SPED Type A with $M = [N/4]$ and $M = [3\sqrt{N}/d]$, and by the MON Type B with $M = [N/4]$; and that the powers of the MON Type B with $M = [N/4]$ and with $M = [3\sqrt{N}/d]$ compete, but they are less than the powers of the other procedures. Regarding the selection of the blocks, TABLE 3 and 4 show that Type B is slightly better than Type A. Regarding the selection of $M$, the same tables show that for the SPED procedures the Okabe and Nakano (1991) criterion is slightly better than the Box and Jenkins (1970) criterion; and that for the MON procedures, the Okabe and Nakano (1991) criterion is slightly better than the Box and Jenkins (1970) criterion when sample sizes are 200 and 400, but this relationship is reversed when sample size is 100. In summary the results of the simulation indicate that the SPED Type B procedure with the Okabe and Nakano (1991) criterion is the best among those procedures we investigated.

## 5.2. Applications

### 5.2.1. Sunspot data

We applied all procedures with $M = [N/4]$ and $M = [3\sqrt{N}/d]$ developed in this paper to the Wölfer sunspot data (TABLE 5, from Brockwell and Davis (1991)). It is well known that the practical data have large variation, so we transformed the data logarithmicly. As the observed number is zero in 1810, we replace it with 1. The sample mean of log-transformed data is 3.391. FIGURE 1 gives the scatter plot of the logarithm of the Wölfer sunspot numbers and its sample autocorrelation function.

Results of the procedures are listed in TABLE 6. The table shows that the sunspot data is stationary. It is well known that the sunspot numbers are well fitted to $AR(2)$ model, so our procedures lead to the reasonable result. Note that if we aplly the procedures to untransformed data, we have TABLE 7 where all the procedures except the

MON Type A shows that the sunspot number is not stationary. We suspect that the non-rejection by the MON Type A would be due to its conservativeness as indicated by the simulation. It is important to transform data appropriately before applying our procedures.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1770-79 | 101 | 82 | 66 | 35 | 31 | 7 | 20 | 92 | 154 | 125 |
| 80-89 | 85 | 68 | 38 | 23 | 10 | 24 | 83 | 132 | 131 | 118 |
| 90-99 | 90 | 67 | 60 | 47 | 41 | 21 | 16 | 6 | 4 | 7 |
| 1800-09 | 14 | 34 | 45 | 43 | 48 | 42 | 28 | 10 | 8 | 2 |
| 10-19 | 0 | 1 | 5 | 12 | 14 | 35 | 46 | 41 | 30 | 24 |
| 20-29 | 16 | 7 | 4 | 2 | 8 | 17 | 36 | 50 | 62 | 67 |
| 30-39 | 71 | 48 | 28 | 8 | 13 | 57 | 122 | 138 | 103 | 86 |
| 40-49 | 63 | 37 | 24 | 11 | 15 | 40 | 62 | 98 | 124 | 96 |
| 50-59 | 66 | 64 | 54 | 39 | 21 | 7 | 4 | 23 | 55 | 94 |
| 60-69 | 96 | 77 | 59 | 44 | 47 | 30 | 16 | 7 | 37 | 74 |

TABLE 5: The Wölfer sunspot numbers, 1770-1869

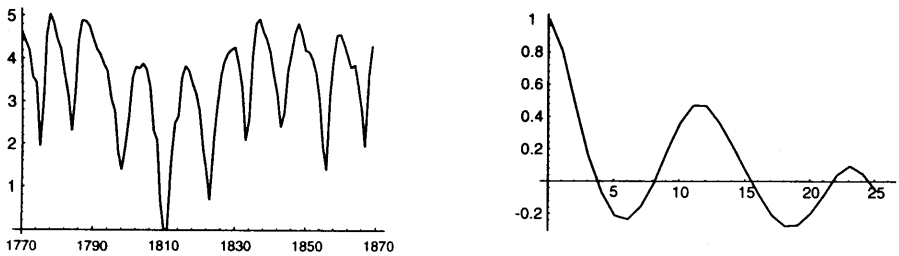


FIGURE 1: The scatter plot of the logarithm of the Wölfer sunspot numbers (left) and the sample autocorrelation function (right)

| | $M = [N/4]$ | $M = [3\sqrt{N}/d]$ |
|---|---|---|
| MON Type A | Not rejected | Not rejected |
| MON Type B | Not rejected | Not rejected |
| SPED Type A | Not rejected | Not rejected |
| SPED Type B | Not rejected | Not rejected |

TABLE 6: Results of log-transformed sunspot data analysis

| | $M = [N/4]$ | $M = [3\sqrt{N}/d]$ |
|---|---|---|
| MON Type A | Not rejected | Not rejected |
| MON Type B | Rejected | Rejected |
| SPED Type A | Rejected | Rejected |
| SPED Type B | Rejected | Rejected |

TABLE 7: Results of untransformed sunspot data analysis

### 5.2.2. Passenger data

The second application of our procedures is to the monthly totals of international airline passengers. (TABLE 8, from Pandit and Wu (1983)) We again transformed the data logarithmicly. FIGURE 2 gives the scatter plot of the log-transformed data and its sample autocorrelation function. The sample mean is 5.542.

TABLE 8 indicates the passenger data is not stationary, since it shows clear trend and seasonality. TABLE 9 summarize the results of our procedures. The table show that all procedures except the MON Type A lead to the same conclusion, i.e. the passengers data are not stationary. The non-rejection by the MON Type A would be due to its conservativeness.

| Year | Jan. | Feb. | Mar. | Apr. | May | Jun. | Jul. | Aug. | Sep. | Oct. | Nov. | Dec. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1949 | 112 | 118 | 132 | 129 | 121 | 135 | 148 | 148 | 136 | 119 | 104 | 118 |
| 1950 | 115 | 126 | 141 | 135 | 125 | 149 | 170 | 170 | 158 | 133 | 114 | 140 |
| 1951 | 145 | 150 | 178 | 163 | 172 | 178 | 199 | 199 | 184 | 162 | 146 | 166 |
| 1952 | 171 | 180 | 193 | 181 | 183 | 218 | 230 | 242 | 209 | 191 | 172 | 194 |
| 1953 | 196 | 196 | 236 | 235 | 229 | 243 | 264 | 272 | 237 | 211 | 180 | 201 |
| 1954 | 204 | 188 | 235 | 227 | 234 | 264 | 302 | 293 | 259 | 229 | 203 | 229 |
| 1955 | 242 | 233 | 267 | 269 | 270 | 315 | 364 | 347 | 312 | 274 | 237 | 278 |
| 1956 | 284 | 277 | 317 | 313 | 318 | 374 | 413 | 405 | 355 | 306 | 271 | 306 |
| 1957 | 315 | 301 | 356 | 348 | 355 | 422 | 465 | 467 | 404 | 347 | 305 | 336 |
| 1958 | 340 | 318 | 362 | 348 | 363 | 435 | 491 | 505 | 404 | 359 | 310 | 337 |
| 1959 | 360 | 342 | 406 | 396 | 420 | 472 | 548 | 559 | 463 | 407 | 362 | 405 |
| 1960 | 417 | 391 | 419 | 461 | 472 | 535 | 622 | 606 | 508 | 461 | 390 | 432 |

TABLE 8: The number of international airline passengers (×1000)

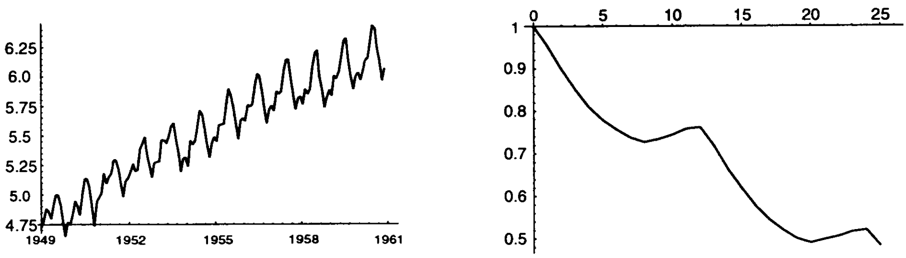


FIGURE 2: The scatter plot of the log-transformed passenger data (left) and the sample autocorrelation function (right)

| | $M = [N/4]$ | $M = [3\sqrt{N/d}]$ |
|---|---|---|
| MON Type A | Not rejected | Not rejected |
| MON Type B | Rejected | Rejected |
| SPED Type A | Rejected | Rejected |
| SPED Type B | Rejected | Rejected |

TABLE 9: Results of the log-transformed passenger data analysis

REMARK. The number of data is 144, so $M = [N/4] = [144/4] = 36$ for Box and Jenkins (1970) criterion and $M = [3\sqrt{N}/d] = [3\sqrt{144}] = 36$ for Okabe and Nakano (1991) criterion. Namley $M$ is same for each case, and thus test statistic of the SPED procedure is also same. Note that the SPED procedure is not dependent on $K$.

## 6. Discussion

Two procedures, called the MON and SPED, for testing the stationarity are developed in this paper. The MON procedure modifies that proposed by Okabe and Nakano (1991) taking into account the multiplicity of the tests. As the Okabe and Nakano (1991) it is assumed in the MON procedure that $\{\xi_t\}$ are distributed independently and identically. However, it is not the case in many examples, as is seen in the univariate AR(1) process $X_t = \phi X_{t-1} + Z_t$, where $\{Z_t\}$ is $WN(0, \sigma)$ such that $E[Z_t^3] \neq 0$. The SPED procedure is developed to free from this assumption. The Bonferoni inequality is used to adjusting for the multiplicity of the tests repeatedly used in the procedures. As is well known, the inequality is conservative, in particular, when the tests are highly correlated. To minimize the consevativeness the idea of the selective blocks (Type B) is introduced to decrease the correlations among the test statistics. Regarding the size of each block, it would not be easy to decide it theoretically, and we considered the Box and Jenkins (1970) and Okabe and Nakano (1991) criteria which have been proposed for its determination. We compared the proposed procedures, selection of the blocks, i.e. Type A and B, and Box and Jenkins (1970) and Okabe and Nakano (1991) criteria by a simulation. Although our simulation is limited because of the explosion of the computational time, the simulation study shows that the SPED Type B procedure with the Okabe and Nakano (1991) criterion is the best among those procedures considered.

## References

Box, G.E.P. and Jenkins, G.M, (1970). *Time series analaysis: Forecasting and control,* Holden-Day, San Fransisco.

Brockwell, P.J. and Davis, R.A, (1991). *Time Series :Theory and Methods*, 2nd edition, Springer-Verlag.

Kato,T, (1995). *Perturbation theory for linear operators*, Reprint of the 1980 edition, Springer-Verlag.

Okabe, Y. and Nakano, Y, (1991). *The theory of $KM_2O$-Langevian equations and its applications to data analysis (I): Stationary analysis*, Hokkaido Math. J. 20, 45-90.

Pandit, S.M. and Wu, S.M, (1983). *Time series and system analysis with applications,* John Wiley and Sons, Inc.

Raikov, D.A, (1938). *On a connection between the central limit theorem in the theory of probability and the law of large numbers,* Izv.Nauk USSR,Sov.Math., 323-338.

*Received January 23, 1997*
*Revised February 28, 1997*