Research Article

Estimation for Non-Gaussian Locally Stationary Processes with Empirical Likelihood Method

Hiroaki Ogata

Waseda University, Tokyo 169-8050, Japan

Correspondence should be addressed to Hiroaki Ogata, hiroakiogata@aoni.waseda.jp

An application of the empirical likelihood method to non-Gaussian locally stationary processes is presented. Based on the central limit theorem for locally stationary processes, we give the asymptotic distributions of the maximum empirical likelihood estimator and the empirical likelihood ratio statistics, respectively. It is shown that the empirical likelihood method enables us to make inferences on various important indices in a time series analysis. Furthermore, we give a numerical study and investigate a finite sample property.

1. Introduction

The empirical likelihood is one of the nonparametric methods for a statistical inference proposed by Owen [1, 2]. It is used for constructing confidence regions for a mean, for a class of M-estimates that includes quantile, and for differentiable statistical functionals. The empirical likelihood method has been applied to various problems because of its good properties: generality of the nonparametric method and effectiveness of the likelihood method. For example, we can name applications to the general estimating equations, [3] the regression models [4–6], the biased sample models [7], and so forth. Applications are also extended to dependent observations. Kitamura [8] developed the blockwise empirical likelihood for estimating equations and for smooth functions of means. Monti [9] applied the empirical likelihood method to linear processes, essentially under the circular Gaussian assumption, using a spectral method. For short- and long-range dependence, Nordman and Lahiri [10] gave the asymptotic properties of the frequency domain empirical likelihood. As we named above, some applications to time series analysis can be found but it seems that they were mainly for stationary processes. Although stationarity is the most fundamental assumption when we are engaged in a time series analysis, it is also known that real time series data are generally nonstationary (e.g., economics analysis). Therefore we need to

use nonstationary models in order to describe the real world. Recently Dahlhaus [11–13] proposed an important class of nonstationary processes, called locally stationary processes. They have so-called time-varying spectral densities whose spectral structures smoothly change in time.

In this paper we extend the empirical likelihood method to non-Gaussian locally stationary processes with time-varying spectra. First, We derive the asymptotic normality of the maximum empirical likelihood estimator based on the central limit theorem for locally stationary processes, which is stated in Dahlhaus [13, Theorem A.2]. Next, we show that the empirical likelihood ratio converges to a sum of Gamma distribution. Especially, when we consider a stationary case, that is, the time-varying spectral density is independent of a time parameter, the asymptotic distribution becomes the chi-square.

As an application of this method, we can estimate an extended autocorrelation for locally stationary processes. Besides we can consider the Whittle estimation which is stated in Dahlhaus [13].

This paper is organized as follows. Section 2 briefly reviews the stationary processes and explains about the locally stationary processes. In Section 3, we propose the empirical likelihood method for non-Gaussian locally stationary processes and give the asymptotic properties. In Section 4 we give numerical studies on confidence intervals of the autocorrelation for locally stationary processes. Proofs of theorems are given in Section 5.

2. Locally Stationary Processes

The stationary process is a fundamental setting in a time series analysis. If the process $\{X_t\}_{t \in \mathbb{Z}}$ is stationary with mean zero, it is known to have the spectral representation:

$$X_t = \int_{-\pi}^{\pi} \exp(i\lambda t) A(\lambda) d\xi(\lambda),$$ (2.1)

where $A(\lambda)$ is a 2π-periodic complex-valued function with $A(-\lambda) = \overline{A(\lambda)}$, called transfer function, and $\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with $\xi(-\lambda) = \overline{\xi(\lambda)}$ and

$$E[d\xi(\lambda)] = 0, \qquad \mathrm{Cov}(d\xi(\lambda_1), d\xi(\lambda_2)) = \eta(\lambda_1 - \lambda_2),$$ (2.2)

where $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the 2π-periodic extension of the Dirac delta function. If the process is stationary, the covariance between X_t and X_{t+k} is independent of time t and a function of only the time lag k. We denote it by $\gamma(k) = \mathrm{Cov}(X_t, X_{t+k})$. The Fourier transform of the autocovariance function

$$g(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) \exp(-ik\lambda)$$ (2.3)

is called spectral density function. In the expression of (2.1), the spectral density function is written by $g(\lambda) = |A(\lambda)|^2$. It is estimated by the periodogram, defined by $I_T(\lambda) = (2\pi)^{-1} |\sum_{t=1}^{T} X_t \exp(-i\lambda t)|^2$. If one wants to change the weight of each data, we can insert

the function $h(x)$ defined on $[0,1]$ into the periodogram: $I_T(\lambda) = (2\pi \sum_{t=1}^{T} h(t/T)^2)^{-1} |\sum_{t=1}^{T} h(t/T) X_t \exp(-i\lambda t)|^2$. The function $h(x)$ is called data taper. Now, we give a simple example of the stationary process below.

Example 2.1. Consider the following AR(p) process:

$$\sum_{j=0}^{p} a_j X_{t-j} = \varepsilon_t, \tag{2.4}$$

where ε_t are independent random variables with mean zero and variance 1. In the form of (2.1), this is obtained by letting

$$A(\lambda) = \frac{1}{\sqrt{2\pi}} \left(\sum_{j=0}^{p} a_j \exp(-i\lambda j) \right)^{-1}. \tag{2.5}$$

As an extension of the stationary process, Dahlhaus [13] introduced the concept of locally stationary. An example of the locally stationary processes is the following time-varying AR(p) process:

$$\sum_{j=0}^{p} a_j \left(\frac{t}{T} \right) X_{t-j,T} = \varepsilon_t, \tag{2.6}$$

where $a_j(u)$ is a function defined on $[0,1]$ and ε_t are independent random variables with mean zero and variance 1. If all $a_j(u)$ are constant, the process (2.6) reduces to stationary. To define a general class of the locally stationary processes, we can naturally consider the time-varying spectral representation

$$X_{t,T} = \int_{-\pi}^{\pi} \exp(i\lambda t) A\left(\frac{t}{T}, \lambda \right) d\xi(\lambda). \tag{2.7}$$

However, it turns out that (2.6) has not exactly but only approximately a solution of the form of (2.7). Therefore, we only require that (2.7) holds approximately. The following is the definition of the locally stationary processes given by Dahlhaus [13].

Definition 2.2. A sequence of stochastic processes $X_{t,T}$ ($t = 1, \ldots, T$) is called locally stationary with mean zero and transfer function $A_{t,T}^{\circ}$, if there exists a representation

$$X_{t,T} = \int_{-\pi}^{\pi} \exp(i\lambda t) A_{t,T}^{\circ}(\lambda) d\xi(\lambda), \tag{2.8}$$

where the following holds.

(i) $\xi(\lambda)$ is a stochastic process on $[-\pi, \pi]$ with $\xi(-\lambda) = \overline{\xi(\lambda)}$ and

$$\mathrm{cum}\{d\xi(\lambda_1), \ldots, d\xi(\lambda_k)\} = \eta\left(\sum_{j=1}^{k} \lambda_j\right) q_k(\lambda_1, \ldots, \lambda_{k-1}) d\lambda_1 \cdots d\lambda_k, \qquad (2.9)$$

where $\mathrm{cum}\{\cdots\}$ denotes the cumulant of kth order, $q_1 = 0$, $q_2(\lambda) = 1$, $|q_k(\lambda_1, \ldots, \lambda_{k-1})| \leq \mathrm{const}_k$ for all k and $\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ is the 2π-periodic extension of the Dirac delta function.

(ii) There exists a constant K and 2π-periodic function $A : [0,1] \times \mathbf{R} \rightarrow \mathbf{C}$ with $A(u, -\lambda) = \overline{A(u, \lambda)}$ which satisfies

$$\sup_{t,\lambda} \left| A_{t,T}^{\circ}(\lambda) - A\left(\frac{t}{T}, \lambda\right) \right| \leq KT^{-1} \qquad (2.10)$$

for all T; $A(u, \lambda)$ is assumed to be continuous in u.

The time-varying spectral density is defined by $g(u, \lambda) := |A(u, \lambda)|^2$. As an estimator of $g(u, \lambda)$, we define the local periodogram $I_N(u, \lambda)$ (for even N) as follows:

$$d_N(u, \lambda) = \sum_{s=1}^{N} h\left(\frac{s}{N}\right) X_{[uT]-N/2+s,T} \exp(-i\lambda s),$$

$$H_{k,N} = \sum_{s=1}^{N} h\left(\frac{s}{N}\right)^k, \qquad (2.11)$$

$$I_N(u, \lambda) = \frac{1}{2\pi H_{2,N}} |d_N(u, \lambda)|^2.$$

Here, $h : \mathbf{R} \rightarrow \mathbf{R}$ is a data taper with $h(x) = 0$ for $x \notin [0,1]$. Thus, $I_N(u, \lambda)$ is nothing but the periodogram over a segment of length N with midpoint $[uT]$. The shift from segment to segment is denoted by S, which means we calculate I_N with midpoints $t_j = S(j-1) + N/2$ ($j = 1, \ldots, M$), where $T = S(M-1) + N$, or, written in rescaled time, at time points $u_j := t_j/T$. Hereafter we set $S = 1$ rather than $S = N$. That means the segments overlap each other.

3. Empirical Likelihood Approach for Non-Gaussian Locally Stationary Processes

Consider an inference on a parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbf{R}^q$ based on an observed stretch $X_{1,T}, \ldots, X_{T,T}$. We suppose that information about $\boldsymbol{\theta}$ exists through a system of general estimating equations. For short- or long-memory processes, Nordman and Lahiri [10] supposed that $\boldsymbol{\theta}_0$, the true value of $\boldsymbol{\theta}$, is specified from the following spectral moment condition:

$$\int_{-\pi}^{\pi} \boldsymbol{\phi}(\lambda, \boldsymbol{\theta}_0) g(\lambda) d\lambda = \mathbf{0}, \qquad (3.1)$$

where $\boldsymbol{\phi}(\lambda, \boldsymbol{\theta})$ is an appropriate function depending on $\boldsymbol{\theta}$. Following this manner, we naturally suppose that $\boldsymbol{\theta}_0$ satisfies the following time-varying spectral moment condition:

$$\int_0^1 \int_{-\pi}^{\pi} \boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta}_0) g(u, \lambda) d\lambda \, du = \mathbf{0} \tag{3.2}$$

in a locally stationary setting. Here $\boldsymbol{\phi} : [0, 1] \times [-\pi, \pi] \times \mathbf{R}^q \to \mathbf{C}^q$ is a function depending on $\boldsymbol{\theta}$ and satisfies Assumption 3.4(i). We give brief examples of $\boldsymbol{\phi}$ and corresponding $\boldsymbol{\theta}_0$.

Example 3.1 (autocorrelation). Let us set

$$\phi(u, \lambda, \theta) = \theta - \exp(i\lambda k). \tag{3.3}$$

Then (3.2) leads to

$$\theta_0 = \frac{\int_0^1 \int_{-\pi}^{\pi} \exp(i\lambda k) g(u, \lambda) d\lambda \, du}{\int_0^1 \int_{-\pi}^{\pi} g(u, \lambda) d\lambda \, du}. \tag{3.4}$$

When we consider the stationary case, that is, $g(u, \lambda)$ is independent of the time parameter u, (3.4) becomes

$$\theta_0 = \frac{\int_{-\pi}^{\pi} \exp(i\lambda k) g(\lambda) d\lambda}{\int_{-\pi}^{\pi} g(\lambda) d\lambda} = \frac{\gamma(k)}{\gamma(0)} = \rho(k), \tag{3.5}$$

which corresponds to the autocorrelation with lag k. So, (3.4) can be interpreted as a kind of autocorrelation with lag k for the locally stationary processes.

Example 3.2 (Whittle estimation). Consider the problem of fitting a parametric spectral model to the true spectral density by minimizing the disparity between them. For the stationary process, this problem is considered in Hosoya and Taniguchi [14] and Kakizawa [15]. For the locally stationary process, the disparity between the parametric model $g_\theta(u, \lambda)$ and the true spectral density $g(u, \lambda)$ is measured by

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{4\pi} \int_0^1 \int_{-\pi}^{\pi} \left\{ \log g_\theta(u, \lambda) + \frac{g(u, \lambda)}{g_\theta(u, \lambda)} \right\} d\lambda \, du \tag{3.6}$$

and we seek the minimizer

$$\boldsymbol{\theta}_0 = \arg \min_{\boldsymbol{\theta} \in \Theta} \mathcal{L}(\boldsymbol{\theta}). \tag{3.7}$$

Under appropriate conditions, $\boldsymbol{\theta}_0$ in (3.7) is obtained by solving the equation $\partial \mathcal{L}(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = \mathbf{0}$. Suppose that the fitting model is described as $g_\theta(u, \lambda) = \sigma^2(u) f_\theta(u, \lambda)$, which means $\boldsymbol{\theta}$ is free from innovation part $\sigma^2(u)$. Then, by Kolmogorov's formula (Dahlhaus [11, Theorem 3.2])

we can see that $\int_{-\pi}^{\pi} \log g_{\theta}(u, \lambda) d\lambda$ is independent of $\boldsymbol{\theta}$. So the differential condition on $\boldsymbol{\theta}_0$ becomes

$$\int_0^1 \int_{-\pi}^{\pi} \frac{\partial}{\partial \boldsymbol{\theta}} g_{\theta}(u, \lambda)^{-1} \bigg|_{\theta = \theta_0} g(u, \lambda) d\lambda \, du = \mathbf{0}. \tag{3.8}$$

This is the case when we set

$$\boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} g_{\theta}(u, \lambda)^{-1}. \tag{3.9}$$

Now, we set

$$\mathbf{m}_j(\boldsymbol{\theta}) = \int_{-\pi}^{\pi} \boldsymbol{\phi}(u_j, \lambda, \boldsymbol{\theta}) I_N(u_j, \lambda) d\lambda \quad (j = 1, \dots, M) \tag{3.10}$$

as an estimating function and use the following empirical likelihood ratio function $\mathcal{R}(\boldsymbol{\theta})$ defined by

$$\mathcal{R}(\boldsymbol{\theta}) = \max_w \left\{ \prod_{j=1}^M M w_j \mid \sum_{j=1}^M w_j \mathbf{m}_j(\boldsymbol{\theta}) = \mathbf{0}, \ w_j \geq 0, \ \sum_{j=1}^M w_j = 1 \right\}. \tag{3.11}$$

Denote the maximum empirical likelihood estimator by $\widetilde{\boldsymbol{\theta}}$, which maximizes the empirical likelihood ratio function $\mathcal{R}(\boldsymbol{\theta})$.

Remark 3.3. We can also use the following alternative estimating function:

$$\mathbf{m}_j^{(T)}(\boldsymbol{\theta}) = \frac{2\pi}{T} \sum_{t=1}^T \boldsymbol{\phi}(u_j, \Delta_t, \boldsymbol{\theta}) I_N(u_j, \Delta_t) \quad \left(\Delta_t = \frac{2\pi t}{T} \right) \tag{3.12}$$

instead of $\mathbf{m}_j(\boldsymbol{\theta})$ in (3.10). The asymptotic equivalence of $\mathbf{m}_j(\boldsymbol{\theta})$ and $\mathbf{m}_j^{(T)}(\boldsymbol{\theta})$ can be proven if

$$E \left| \mathbf{m}_j^{(T)}(\boldsymbol{\theta}) - \mathbf{m}_j(\boldsymbol{\theta}) \right| = o(1) \tag{3.13}$$

is satisfied for any j, and this is shown by straightforward calculation.

To show the asymptotic properties of $\widetilde{\boldsymbol{\theta}}$ and $\mathcal{R}(\boldsymbol{\theta}_0)$, we impose the following assumption.

Assumption 3.4. (i) The functions $A(u, \lambda)$ and $\boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta})$ are 2π-periodic in λ, and the periodic extensions are differentiable in u and λ with uniformly bounded derivative $(\partial/\partial u)(\partial/\partial \lambda) A$ ($\boldsymbol{\phi}$, resp.).

(ii) The parameters N and T fulfill the relations $T^{1/4} \ll N \ll T^{1/2}/\log T$.

(iii) The data taper $h : \mathbf{R} \to \mathbf{R}$ with $h(x) = 0$ for all $x \notin (0,1)$ is continuous on $\mathbf{R}$ and twice differentiable at all $x \notin p$ where p is a finite set and $\sup_{x \notin p}|h''(x)| < \infty$.

(iv) For $k = 1, \ldots, 8$,

$$q_k(\lambda_1, \ldots, \lambda_{k-1}) = c_k \quad \text{(constant)}. \tag{3.14}$$

Remark 3.5. Assumption 3.4(ii) seems to be restrictive. However, this is required to use the central limit theorem for locally stationary processes (cf. Assumption A.1 and Theorem A.2 of Dahlhaus [13]) (Most of the restrictions on N result from the $\sqrt{T}$-unbiasedness in the central limit theorem). See also A.3. Remarks of Dahlhaus [13] for the detail.

Now we give the following theorem.

Theorem 3.6. *Suppose that Assumption 3.4 holds and $X_{1,T}, \ldots, X_{T,T}$ is realization of the locally stationary process which has the representation* (2.8). *Then,*

$$\sqrt{M}\left(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}) \tag{3.15}$$

as $T \to \infty$, where

$$\boldsymbol{\Sigma} = 4\pi \left(\boldsymbol{\Sigma}_3' \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_3\right)^{-1} \boldsymbol{\Sigma}_3' \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_3 \left(\boldsymbol{\Sigma}_3' \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\Sigma}_3\right)^{-1}. \tag{3.16}$$

Here $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are the q by q matrices whose (i, j) elements are

$$(\boldsymbol{\Sigma}_1)_{ij} = \frac{1}{2\pi} \int_0^1 \left[\iint_{-\pi}^{\pi} \phi_i(u, \lambda, \boldsymbol{\theta}_0)\{\phi_j(u, \lambda, \boldsymbol{\theta}_0) + \phi_j(u, -\lambda, \boldsymbol{\theta}_0)\}g(u, \lambda)^2 d\lambda \right.$$
$$\left. + c_4 \int_{-\pi}^{\pi} \phi_i(u, \lambda, \boldsymbol{\theta}_0)g(u, \lambda)d\lambda \int_{-\pi}^{\pi} \phi_j(u, \mu, \boldsymbol{\theta}_0)g(u, \mu)d\mu \right] du, \tag{3.17}$$

$$(\boldsymbol{\Sigma}_2)_{ij} = \frac{1}{2\pi} \int_0^1 \left[\iint_{-\pi}^{\pi} \phi_i(u, \lambda, \boldsymbol{\theta}_0)\{\phi_j(u, \lambda, \boldsymbol{\theta}_0) + \phi_j(u, -\lambda, \boldsymbol{\theta}_0)\}g(u, \lambda)^2 d\lambda \right.$$
$$\left. + \int_{-\pi}^{\pi} \phi_i(u, \lambda, \boldsymbol{\theta}_0)g(u, \lambda)d\lambda \int_{-\pi}^{\pi} \phi_j(u, \mu, \boldsymbol{\theta}_0)g(u, \mu)d\mu \right] du, \tag{3.18}$$

respectively, and $\boldsymbol{\Sigma}_3$ is the q by q matrix which is defined as

$$\boldsymbol{\Sigma}_3 = \int_0^1 \int_{-\pi}^{\pi} \frac{\partial \boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} g(u, \lambda)d\lambda \, du. \tag{3.19}$$

In addition, we give the following theorem on the asymptotic property of the empirical likelihood ratio $\mathcal{R}(\boldsymbol{\theta}_0)$.

Theorem 3.7. *Suppose that Assumption 3.4 holds and $X_{1,T}, \ldots, X_{T,T}$ is realization of a locally stationary process which has the representation* (2.8). *Then,*

$$-\frac{1}{\pi} \log \mathcal{R}(\boldsymbol{\theta}_0) \xrightarrow{d} (\mathbf{FN})'(\mathbf{FN}) \tag{3.20}$$

as $T \to \infty$, where $\mathbf{N}$ is a q-dimensional normal random vector with zero mean vector and covariance matrix $\mathbf{I}_q$ (identity matrix) and $\mathbf{F} = \boldsymbol{\Sigma}_2^{-1/2}\boldsymbol{\Sigma}_1^{1/2}$. Here $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are same matrices in Theorem 3.6.

Remark 3.8. Denote the eigenvalues of $\mathbf{F}'\mathbf{F}$ by $a_1, \ldots, a_q$, then we can write

$$(\mathbf{FN})'(\mathbf{FN}) = \sum_{i=1}^{q} Z_i, \tag{3.21}$$

where Z_i is distributed as $\mathrm{Gamma}(1/2, 1/(2a_i))$, independently.

Remark 3.9. If the process is stationary, that is, the time-varying spectral density is independent of the time parameter u, we can easily see that $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ and the asymptotic distribution becomes the chi-square with degree of freedom q.

Remark 3.10. In our setting, the number of the estimating equations and that of the parameters are equal. In that case, the empirical likelihood ratio at the maximum empirical likelihood estimator, $\mathcal{R}(\widetilde{\boldsymbol{\theta}})$, becomes one (cf. [3, page 305]). That means the test statistic in Theorem 3.7 becomes zero when we evaluate it at the maximum empirical likelihood estimator.

4. Numerical Example

In this section, we present simulation results of the estimation of the autocorrelation in locally stationary processes which is stated in Example 3.1. Consider the following time-varying AR(1) process:

$$X_{t,T} - a\left(\frac{t}{T}\right)X_{t-1,T} = \varepsilon_t \quad \text{for } t \in \mathbb{Z}, \tag{4.1}$$

where $\varepsilon_t \overset{i.i.d.}{\sim} \mathrm{Gamma}(3/\pi, (3/\pi)^{1/2}) - (3/\pi)^{1/2}$ and $a(u) = (u - b)^2$, $b = 0.1, 0.5, 0.9$. The observations $X_{1,T}, \ldots, X_{T,T}$ are generated from the process (4.1), and we make the confidence intervals of the autocorrelation with lag $k = 1$, which is expressed as

$$\theta_0 = \frac{\int_0^1 \int_{-\pi}^{\pi} e^{i\lambda} g(u, \lambda) d\lambda \, du}{\int_0^1 \int_{-\pi}^{\pi} g(u, \lambda) d\lambda \, du}, \tag{4.2}$$

based on the result of Theorem 3.7. The several combinations of the sample size T and the window length N are chosen: $(T, N) = (100, 10), (500, 10), (500, 50), (1000, 10), (1000, 100)$, and the data taper is set as $h(x) = (1/2)\{1 - \cos(2\pi x)\}$. Then we calculate the values of the test statistic $-\pi^{-1} \log \mathcal{R}(\theta)$ at numerous points θ and obtain confidence intervals by

Table 1: 90% confidence intervals of the autocorrelation with lag $k = 1$.

(T, N)	Lower bound	Upper bound	Interval length	Successful rate
		$b = 0.1$, $\theta_0 = 0.308$		
$(100, 10)$	0.057	0.439	0.382	0.854
$(500, 10)$	0.172	0.382	0.210	0.866
$(500, 50)$	0.203	0.332	0.129	0.578
$(1000, 10)$	0.203	0.356	0.154	0.826
$(1000, 100)$	0.225	0.308	0.084	0.444
		$b = 0.5$, $\theta_0 = 0.085$		
$(100, 10)$	−0.087	0.225	0.312	0.890
$(500, 10)$	0.001	0.169	0.168	0.910
$(500, 50)$	0.028	0.104	0.076	0.515
$(1000, 10)$	0.023	0.139	0.116	0.922
$(1000, 100)$	0.047	0.087	0.040	0.384
		$b = 0.9$, $\theta_0 = 0.308$		
$(100, 10)$	0.060	0.449	0.388	0.841
$(500, 10)$	0.176	0.393	0.216	0.871
$(500, 50)$	0.201	0.332	0.131	0.586
$(1000, 10)$	0.203	0.359	0.156	0.827
$(1000, 100)$	0.226	0.310	0.083	0.467

collecting the points θ which satisfy $-\pi^{-1} \log \mathcal{R}(\theta) < z_\alpha$ where z_α, is α-percentile of the asymptotic distribution in Theorem 3.7. We admit that Assumption 3.4. (ii) is hard to hold in a finite sample experiment, but this Monte Carlo simulation is purely illustrative and just for investigating how the sample size and the window length affect the results of confidence intervals.

We set a confidence level as $\alpha = 0.90$ and carry out the above procedure 1000 times for each case. Table 1 shows the averages of lower and upper bounds, lengths of the intervals, and the successful rates. Looking at the results, we find out that the larger sample size gives the shorter length of the interval, as expected. Furthermore, the results indicate that the larger window length leads to the worse successful rate. We can predict that the best rate N/T lies around 0.02 because the combination $(T, N) = (500, 10)$ seems to give the best result among all.

5. Proofs

5.1. Some Lemmas

In this subsection we give the three lemmas to prove Theorems 3.6 and 3.7. First of all, we introduce the following function $L_N : \mathbf{R} \to \mathbf{R}$, which is defined by the 2π-periodic extension of

$$L_N(\alpha) := \begin{cases} N, & |\alpha| \leq \dfrac{1}{N} \\ \dfrac{1}{|\alpha|}, & \dfrac{1}{N} \leq |\alpha| \leq \pi. \end{cases} \tag{5.1}$$

The properties of the function L_N are described in Lemma A.4 of Dahlhaus [13].

Lemma 5.1. *Suppose* (3.2) *and Assumption 3.4 hold. Then for* $1 \leq k \leq 8$,

$$\operatorname{cum}\{d_N(u_1, \lambda_1), \ldots, d_N(u_k, \lambda_k)\}$$

$$= (2\pi)^{k-1} c_k \left\{ \prod_{j=1}^{k} A(u_j, \lambda_j) \right\} \exp\left\{ -i \sum_{j=1}^{k} \lambda_j ([u_k T] - [u_j T]) \right\}$$

$$\times \sum_{s=1}^{N} \left\{ \prod_{j=1}^{k} h\left(\frac{s + [u_k T] - [u_j T]}{N} \right) \right\} \exp\left\{ -i \left(\sum_{j=1}^{k} \lambda_j \right) s \right\}$$

$$+ O\left(\frac{N^2}{T} \right) + O\left((\log N)^{k-1} \right) = O\left(L_N \left(\sum_{j=1}^{k} \lambda_j \right) \right) + O\left(\frac{N^2}{T} \right) + O\left((\log N)^{k-1} \right). \tag{5.2}$$

Proof. Let $\Pi = (-\pi, \pi]$ and let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_k)$. Since

$$\operatorname{cum}(X_{t_1,T}, \ldots, X_{t_k,T}) = c_k \int_{\Pi^k} \exp\left(i \sum_{j=1}^{k} \omega_j t_j \right) \left(\prod_{j=1}^{k} A^{\circ}_{t_j,T}(\omega_j) \right) \eta\left(\sum_{j=1}^{k} \omega_j \right) d\boldsymbol{\omega}, \tag{5.3}$$

the kth cumulant of d_N is equal to

$$c_k \int_{\Pi^k} \exp\left\{ i \sum_{j=1}^{k} \omega_j \left([u_j T] - \frac{N}{2} \right) \right\} \eta\left(\sum_{j=1}^{k} \omega_j \right)$$

$$\times \prod_{j=1}^{k} \sum_{s=1}^{N} h\left(\frac{s}{N} \right) A^{\circ}_{[u_j T] - N/2 + s, T}(\omega_j) \exp\{ -i(\lambda_j - \omega_j) s \} d\boldsymbol{\omega}. \tag{5.4}$$

As in the proof of Theorem 2.2 of Dahlhaus [12] we replace $A^{\circ}_{[u_1 T] - N/2 + s_1, T}(\omega_1)$ by $A(u_1 + (-N/2 + s_1)/T, \lambda_1)$ and we obtain

$$\left| \sum_{s=1}^{N} h\left(\frac{s}{N} \right) \left\{ A^{\circ}_{[u_1 T] - N/2 + s, T}(\omega_1) - A\left(u_1 + \frac{-N/2 + s}{T}, \lambda_1 \right) \right\} \exp\{ -i(\lambda_1 - \omega_1) s \} \right| \leq K \tag{5.5}$$

with some constant K while

$$\left| \sum_{s=1}^{N} h\left(\frac{s}{N} \right) A^{\circ}_{[u_j T] - N/2 + s, T}(\omega_j) \exp\{ -i(\lambda_j - \omega_j) s \} \right| \leq K L_N(\lambda_j - \omega_j) \tag{5.6}$$

for $j = 2, \ldots, k$. The replacement error is smaller than

$$K \int_{\Pi^k} \prod_{j=2}^{k} L_N(\lambda_j - \omega_j) d\boldsymbol{\omega} \le K (\log N)^{k-1}. \tag{5.7}$$

In the same way we replace $A^{\circ}_{[u_jT]-N/2+s_j,T}(\omega_j)$ by $A(u_j + (-N/2 + s_j)/T, \lambda_j)$ for $j = 2, \ldots, k$, and then we obtain

$$c_k \sum_{s_1,\ldots,s_k=1}^{N} \left\{ \prod_{j=1}^{k} h\left(\frac{s_j}{N}\right) A\left(u_j + \frac{-N/2+s_j}{T}, \lambda_j\right) \right\} \exp\left(-i\sum_{j=1}^{k}\lambda_j s_j\right)$$

$$\times \int_{\Pi^k} \eta\left(\sum_{j=1}^{k}\omega_j\right) \exp\left\{i\sum_{j=1}^{k}\omega_j\left([u_jT] - \frac{N}{2} + s_j\right)\right\} d\boldsymbol{\omega} + O\left((\log N)^{k-1}\right). \tag{5.8}$$

The integral part is equal to

$$\prod_{j=1}^{k-1} \int_{\Pi} \exp\{i\omega_j([u_jT] - [u_kT] + s_j - s_k)\} d\omega_j. \tag{5.9}$$

So we get

$$(2\pi)^{k-1} c_k \sum_{s=1}^{N} \left\{ \prod_{j=1}^{k} h\left(\frac{s + [u_kT] - [u_jT]}{N}\right) A\left(u_j + \frac{-N/2 + s + [u_kT] - [u_jT]}{T}, \lambda_j\right) \right\}$$

$$\times \exp\left\{-i\sum_{j=1}^{k}\lambda_j(s + [u_kT] - [u_jT])\right\} + O\left((\log N)^{k-1}\right). \tag{5.10}$$

Since $h(x) = 0$ for $x \notin (0,1)$, we only have to consider the range of s which satisfies $1 \le s + [u_kT] - [u_jT] \le N - 1$. Therefore we can regard $(-N/2 + s + [u_kT] - [u_jT])/T$ as $O(N/T)$, and Taylor expansion of A around u_j gives the first equation of the desired result. Moreover, as in the same manner of the proof of Lemma A.5 of Dahlhaus [13] we can see that

$$\sum_{s=1}^{N} \left\{ \prod_{j=1}^{k} h\left(\frac{s + [u_kT] - [u_jT]}{N}\right) \right\} \exp\left\{-i\left(\sum_{j=1}^{k}\lambda_j\right)s\right\} = O\left(L_N\left(\sum_{j=1}^{k}\lambda_j\right)\right), \tag{5.11}$$

which leads to the second equation. $\qquad\square$

Lemma 5.2. *Suppose* (3.2) *and Assumption 3.4 hold. Then,*

$$\mathbf{P}_M := \frac{1}{2\pi\sqrt{M}}\sum_{j=1}^{M}\mathbf{m}_j(\boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_1). \tag{5.12}$$

Proof. We set

$$\mathbf{J}_T(\boldsymbol{\phi}) := \frac{1}{M}\sum_{j=1}^{M}\int_{-\pi}^{\pi}\boldsymbol{\phi}(u_j,\lambda,\boldsymbol{\theta}_0)I_N(u_j,\lambda)d\lambda,$$

$$\mathbf{J}(\boldsymbol{\phi}) := \int_0^1\int_{-\pi}^{\pi}\boldsymbol{\phi}(u,\lambda,\boldsymbol{\theta}_0)g(u,\lambda)d\lambda\,du. \tag{5.13}$$

Henceforth we denote $\boldsymbol{\phi}(u,\lambda,\boldsymbol{\theta}_0)$ by $\boldsymbol{\phi}(u,\lambda)$ for simplicity. This lemma is proved by proving the convergence of the cumulants of all orders. Due to Lemma A.8 of Dahlhaus [13] the expectation of $\mathbf{P}_M$ is equal to

$$\frac{\sqrt{M}}{2\pi}\left\{\mathbf{J}(\boldsymbol{\phi}) + o\left(T^{-1/2}\right)\right\}. \tag{5.14}$$

By (3.2) and $O(M) = O(T)$, this converges to zero.

Next, we calculate the covariance of $\mathbf{P}_M$. From the relation $T = M + N - 1$ we can rewrite

$$\mathbf{P}_M = \sqrt{\frac{M}{T}}\frac{\sqrt{T}}{2\pi}\mathbf{J}_T(\boldsymbol{\phi}) = \sqrt{1 - \frac{N+1}{T}}\frac{\sqrt{T}}{2\pi}\mathbf{J}_T(\boldsymbol{\phi}). \tag{5.15}$$

Then the (α,β)-element of the covariance matrix of $\mathbf{P}_M$ is equal to

$$\frac{1}{(2\pi)^2}\left(1 - \frac{N+1}{T}\right)T\,\mathrm{cov}\{J_T(\phi_\alpha), J_T(\phi_\beta)\}. \tag{5.16}$$

Due to Lemma A.9 of Dahlhaus [13], this converges to

$$\frac{1}{2\pi}\int_0^1\left[\int_{-\pi}^{\pi}\phi_i(u,\lambda)\{\phi_j(u,\lambda) + \phi_j(u,-\lambda)\}g(u,\lambda)^2d\lambda\right.$$

$$\left. + \int\int_{-\pi}^{\pi}\phi_i(u,\lambda)\phi_j(u,\mu)g(u,\lambda)g(u,\mu)q_4(\lambda,-\lambda,\mu)d\lambda\,d\mu\right]du. \tag{5.17}$$

By Assumption 3.4(iv) the covariance tends to $\boldsymbol{\Sigma}_1$.

The kth cumulant for $k \geq 3$ tends to zero due to Lemma A.10 of Dahlhaus [13]. Then we obtain the desired result. $\square$

Lemma 5.3. *Suppose* (3.2) *and Assumption 3.4 hold. Then,*

$$\mathbf{S}_M := \frac{1}{2\pi M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)' \xrightarrow{p} \boldsymbol{\Sigma}_2. \tag{5.18}$$

Proof. First we calculate the mean of (α, β)-element of $\mathbf{S}_M$:

$$E\left[\frac{1}{2\pi M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)'\right]_{\alpha\beta}$$

$$= \frac{1}{2\pi M} \sum_{j=1}^{M} \int \int_{-\pi}^{\pi} \phi_\alpha(u_j, \lambda) \phi_\beta(u_j, \mu) E[I_N(u_j, \lambda) I_N(u_j, \mu)] d\lambda \, d\mu \tag{5.19}$$

$$= \frac{1}{2\pi M} \sum_{j=1}^{M} \int \int_{-\pi}^{\pi} \phi_\alpha(u_j, \lambda) \phi_\beta(u_j, \mu)$$

$$\times \left[\operatorname{cov}\{I_N(u_j, \lambda), I_N(u_j, \mu)\} + EI_N(u_j, \lambda) EI_N(u_j, \mu)\right] d\lambda \, d\mu.$$

Due to Dahlhaus [12, Theorem 2.2 (i)] the second term of (5.19) becomes

$$\frac{1}{2\pi M} \sum_{j=1}^{M} \int_{-\pi}^{\pi} \phi_\alpha(u_j, \lambda) \left\{ g(u_j, \lambda) + O\left(\frac{N^2}{T^2}\right) + O\left(\frac{\log N}{N}\right) \right\} d\lambda$$

$$\times \int_{-\pi}^{\pi} \phi_\beta(u_j, \mu) \left\{ g(u_j, \mu) + O\left(\frac{N^2}{T^2}\right) + O\left(\frac{\log N}{N}\right) \right\} d\mu \tag{5.20}$$

$$= \frac{1}{2\pi} \int_{0}^{1} \left\{ \int_{-\pi}^{\pi} \phi_\alpha(u, \lambda) g(u, \lambda) \, d\lambda \int_{-\pi}^{\pi} \phi_\beta(u, \mu) g(u, \mu) d\mu \right\}$$

$$+ O\left(\frac{1}{M}\right) + O\left(\frac{N^2}{T^2}\right) + O\left(\frac{\log N}{N}\right).$$

Next we consider

$$\operatorname{cov}\{I_N(u_j, \lambda), I_N(u_j, \mu)\}$$

$$= \frac{1}{(2\pi H_{2,N})^2} \left[\operatorname{cum}\{d_N(u_j, \lambda), d_N(u_j, \mu)\}\operatorname{cum}\{d_N(u_j, -\lambda), d_N(u_j, -\mu)\} \right. \tag{5.21}$$

$$+ \operatorname{cum}\{d_N(u_j, \lambda), d_N(u_j, -\mu)\}\operatorname{cum}\{d_N(u_j, -\lambda), d_N(u_j, \mu)\}$$

$$\left. + \operatorname{cum}\{d_N(u_j, \lambda), d_N(u_j, -\lambda), d_N(u_j, \mu), d_N(u_j, -\mu)\}\right].$$

We calculate the three terms separately. From **Lemma 5.1** the first term of (5.21) is equal to

$$
\frac{1}{(2\pi H_{2,N})^2}\left\{2\pi A(u_j,\lambda)A(u_j,\mu)\sum_{s=1}^{N}h\left(\frac{s}{N}\right)^2\exp\{-i(\lambda+\mu)s\}+O\left(\frac{N^2}{T}\right)+O(\log N)\right\}
$$
$$
\times\left\{2\pi A(u_j,-\lambda)A(u_j,-\mu)\sum_{s=1}^{N}h\left(\frac{s}{N}\right)^2\exp\{-i(-\lambda-\mu)s\}+O\left(\frac{N^2}{T}\right)+O(\log N)\right\}.
$$

$$(5.22)$$

It converges to zero when $\lambda\neq-\mu$ and is equal to

$$
g(u_j,\lambda)^2+O\left(\frac{N}{T}\right)+O\left(\frac{\log N}{N}\right)
$$

$$(5.23)$$

when $\lambda=-\mu$. Similarly the second term of (5.21) converges to zero when $\lambda\neq\mu$ and is equal to (5.23) when $\lambda=\mu$. We can also apply **Lemma 5.1** to the third term of (5.21), and analogous calculation shows that it converges to zero. After all we can see that (5.19) converges to $(\Sigma_2)_{\alpha\beta}$, the (α,β)-element of Σ_2.

Next we calculate the second-order cumulant:

$$
\text{cum}\left\{\left[\frac{1}{2\pi M}\sum_{j=1}^{M}\mathbf{m}_j(\boldsymbol{\theta}_0)\mathbf{m}_j(\boldsymbol{\theta}_0)'\right]_{\alpha_1\beta_1},\left[\frac{1}{2\pi M}\sum_{j=1}^{M}\mathbf{m}_j(\boldsymbol{\theta}_0)\mathbf{m}_j(\boldsymbol{\theta}_0)'\right]_{\alpha_2\beta_2}\right\}.
$$

$$(5.24)$$

This is equal to

$$
(2\pi M)^{-2}(2\pi H_{2,N})^{-4}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\int\int\int\int_{-\pi}^{\pi}\phi_{\alpha_1}(u_{j_1},\lambda_1)\phi_{\beta_1}(u_{j_1},\mu_1)\phi_{\alpha_2}(u_{j_2},\lambda_2)\phi_{\beta_2}(u_{j_2},\mu_2)
$$

$$(5.25)$$

$$
\times\text{cum}\{d_N(u_{j_1},\lambda_1)d_N(u_{j_1},-\lambda_1)d_N(u_{j_1},\mu_1)d_N(u_{j_1},-\mu_1),
$$
$$
d_N(u_{j_2},\lambda_2)d_N(u_{j_2},-\lambda_2)d_N(u_{j_2},\mu_2)d_N(u_{j_2},-\mu_2)\}d\lambda_1 d\mu_1 d\lambda_2 d\mu_2.
$$

Using the product theorem for cumulants (cf. [16, Theorem 2.3.2]) we have to sum over all indecomposable partitions $\{P_1,\dots,P_m\}$ with $|P_i|\geq 2$ of the scheme

$$
\begin{array}{cccc}
d_N(u_{j_1},\lambda_1) & d_N(u_{j_1},-\lambda_1) & d_N(u_{j_1},\mu_1) & d_N(u_{j_1},-\mu_1)\\
d_N(u_{j_2},\lambda_2) & d_N(u_{j_2},-\lambda_2) & d_N(u_{j_2},\mu_2) & d_N(u_{j_2},-\mu_2).
\end{array}
$$

$$(5.26)$$

We can apply **Lemma 5.1** to all cumulants which is seen in (5.25), and the dominant term of the cumulants is $o(N^4)$ so (5.25) tends to zero. Then we obtain the desired result. $\square$

5.2. Proof of Theorem 3.6

Using the lemmas in Section 5.1, we prove Theorem 3.6. To find the maximizing weights $w'_j s$ of (3.11), we proceed by the Lagrange multiplier method. Write

$$G = \sum_{j=1}^{M} \log(Mw_j) - M\boldsymbol{\alpha}' \sum_{j=1}^{M} w_j \mathbf{m}_j(\boldsymbol{\theta}) + \gamma \left(\sum_{j=1}^{M} w_j - 1 \right), \tag{5.27}$$

where $\boldsymbol{\alpha} \in \mathbf{R}^q$ and $\gamma \in \mathbf{R}$ are Lagrange multipliers. Setting $\partial G / \partial w_j = 0$ gives

$$\frac{\partial G}{\partial w_j} = \frac{1}{w_j} - M\boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta}) + \gamma = 0. \tag{5.28}$$

So the equation $\sum_{j=1}^{M} w_j(\partial G / \partial w_j) = 0$ gives $\gamma = -M$. Then, we may write

$$w_j = \frac{1}{M} \frac{1}{1 + \boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta})}, \tag{5.29}$$

where the vector $\boldsymbol{\alpha} = \boldsymbol{\alpha}(\boldsymbol{\theta}_0)$ satisfies q equations given by

$$\frac{1}{M} \sum_{j=1}^{M} \frac{\mathbf{m}_j(\boldsymbol{\theta})}{1 + \boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta})} = \mathbf{0}. \tag{5.30}$$

Therefore, $\tilde{\boldsymbol{\theta}}$ is a minimizer of the following (minus) empirical log likelihood ratio function

$$l(\boldsymbol{\theta}) := \sum_{j=1}^{M} \log\{1 + \boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta})\} \tag{5.31}$$

and satisfies

$$\begin{aligned}
\mathbf{0} = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \bigg|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} &= \sum_{j=1}^{M} \frac{(\partial \boldsymbol{\alpha}'(\boldsymbol{\theta})/\partial \boldsymbol{\theta})\mathbf{m}_j(\boldsymbol{\theta}) + \left(\partial \mathbf{m}'_j(\boldsymbol{\theta})/\partial \boldsymbol{\theta}\right)\boldsymbol{\alpha}(\boldsymbol{\theta})}{1 + \boldsymbol{\alpha}'(\boldsymbol{\theta})\mathbf{m}_j(\boldsymbol{\theta})} \bigg|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}} \\
&= \sum_{j=1}^{M} \frac{\left(\partial \mathbf{m}'_j(\boldsymbol{\theta})/\partial \boldsymbol{\theta}\right)\boldsymbol{\alpha}(\boldsymbol{\theta})}{1 + \boldsymbol{\alpha}'(\boldsymbol{\theta})\mathbf{m}_j(\boldsymbol{\theta})} \bigg|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}.
\end{aligned} \tag{5.32}$$

Denote

$$
Q_{1M}(\boldsymbol{\theta}, \boldsymbol{\alpha}) := \frac{1}{M} \sum_{j=1}^{M} \frac{\mathbf{m}_j(\boldsymbol{\theta})}{1 + \boldsymbol{\alpha}'(\boldsymbol{\theta})\mathbf{m}_j(\boldsymbol{\theta})},
$$

$$
Q_{2M}(\boldsymbol{\theta}, \boldsymbol{\alpha}) := \frac{1}{M} \sum_{j=1}^{M} \frac{1}{1 + \boldsymbol{\alpha}'(\boldsymbol{\theta})\mathbf{m}_j(\boldsymbol{\theta})} \frac{\partial \mathbf{m}'_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \boldsymbol{\alpha}(\boldsymbol{\theta}).
$$

$$(5.33)$$

Then, from (5.30) and (5.32), we have

$$
\begin{aligned}
\mathbf{0} &= Q_{1M}\left(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\alpha}}\right) \\
&= Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0}) + \frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) + \frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\alpha}'}(\widetilde{\boldsymbol{\alpha}} - \mathbf{0}) + o_p(\delta_M),
\end{aligned}
$$

$$(5.34)$$

$$
\begin{aligned}
\mathbf{0} &= Q_{2M}\left(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\alpha}}\right) \\
&= Q_{2M}(\boldsymbol{\theta}_0, \mathbf{0}) + \frac{\partial Q_{2M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) + \frac{\partial Q_{2M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\alpha}'}(\widetilde{\boldsymbol{\alpha}} - \mathbf{0}) + o_p(\delta_M),
\end{aligned}
$$

$$(5.35)$$

where $\widetilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha}(\widetilde{\boldsymbol{\theta}})$ and $\delta_M = \|\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\| + \|\widetilde{\boldsymbol{\alpha}}\|$. Let us see the asymptotic properties of the above four derivatives. First,

$$
\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'} = \frac{1}{M} \sum_{j=1}^{M} \frac{\partial \mathbf{m}_j(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'} = \frac{1}{M} \sum_{j=1}^{M} \int_{-\pi}^{\pi} \frac{\partial \boldsymbol{\phi}(u_j, \lambda, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} I_N(u_j, \lambda) d\lambda.
$$

$$(5.36)$$

From Lemmas A.8 and A.9 of Dahlhaus [13], we have

$$
E\left[\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'}\right] = \int_0^1 \int_{-\pi}^{\pi} \frac{\partial \boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} g(u, \lambda) d\lambda du + o\left(M^{-1/2}\right),
$$

$$
\mathrm{cov}\left[\left[\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'}\right]_{\alpha_1 \beta_1}, \left[\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'}\right]_{\alpha_2 \beta_2}\right] = O\left(M^{-1}\right),
$$

$$(5.37)$$

which leads to

$$
\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'} \xrightarrow{p} \int_0^1 \int_{-\pi}^{\pi} \frac{\partial \boldsymbol{\phi}(u, \lambda, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} g(u, \lambda) d\lambda \, du = \boldsymbol{\Sigma}_3.
$$

$$(5.38)$$

Similarly, we have

$$
\frac{\partial Q_{2M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\alpha}'} = \frac{1}{M} \sum_{j=1}^{M} \frac{\partial \mathbf{m}_j(\boldsymbol{\theta}_0)'}{\partial \boldsymbol{\theta}} \xrightarrow{p} \boldsymbol{\Sigma}'_3.
$$

$$(5.39)$$

Next, from Lemma 5.3, we obtain

$$\frac{\partial Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\alpha}'} = -\frac{1}{M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)' \xrightarrow{p} -2\pi \boldsymbol{\Sigma}_2. \tag{5.40}$$

Finally, we have

$$\frac{\partial Q_{2M}(\boldsymbol{\theta}_0, \mathbf{0})}{\partial \boldsymbol{\theta}'} = \mathbf{0}. \tag{5.41}$$

Now, (5.34), (5.35) and (5.38)–(5.41) give

$$\begin{pmatrix} \tilde{\boldsymbol{\alpha}} \\ \tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \end{pmatrix} = \begin{pmatrix} \dfrac{\partial Q_{1M}}{\partial \boldsymbol{\alpha}'} & \dfrac{\partial Q_{1M}}{\partial \boldsymbol{\theta}'} \\ \dfrac{\partial Q_{2M}}{\partial \boldsymbol{\alpha}'} & \dfrac{\partial Q_{2M}}{\partial \boldsymbol{\theta}'} \end{pmatrix}_{(\boldsymbol{\theta}_0, \mathbf{0})}^{-1} \begin{pmatrix} -Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0}) + o_p(\delta_M) \\ o_p(\delta_M) \end{pmatrix}, \tag{5.42}$$

where

$$\begin{pmatrix} \dfrac{\partial Q_{1M}}{\partial \mathbf{s}\boldsymbol{\alpha}'} & \dfrac{\partial Q_{1M}}{\partial \boldsymbol{\theta}'} \\ \dfrac{\partial Q_{2M}}{\partial \boldsymbol{\alpha}'} & \dfrac{\partial Q_{2M}}{\partial \boldsymbol{\theta}'} \end{pmatrix}_{(\boldsymbol{\theta}_0, \mathbf{0})} \xrightarrow{p} \begin{pmatrix} -2\pi \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_3 \\ \boldsymbol{\Sigma}_3' & \mathbf{0} \end{pmatrix}. \tag{5.43}$$

Because of Lemma 5.2, we have

$$Q_{1M}(\boldsymbol{\theta}_0, \mathbf{0}) = \frac{1}{M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) = O_p\left(M^{-1/2}\right). \tag{5.44}$$

From this and the relation (5.42), (5.43), we can see that $\delta_M = O_p(M^{-1/2})$. Again, from (5.42), (5.43), and Lemma 5.2, direct calculation gives that

$$\sqrt{M}\left(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}). \tag{5.45}$$

5.3. Proof of Theorem 3.7

Using the lemmas in Section 5.1, we prove Theorem 3.7. The proof is the same as that of Theorem 3.6 up to (5.30). Let $\boldsymbol{\alpha} = \|\alpha\|\mathbf{e}$ where $\|\mathbf{e}\| = 1$, and we introduce

$$Y_j := \boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta}_0), \qquad Z_M^* := \max_{1 \leq j \leq M} \|\mathbf{m}_j(\boldsymbol{\theta}_0)\|. \tag{5.46}$$

Note $1/(1 + Y_j) = 1 - Y_j/(1 + Y_j)$ and from (5.30) we find that

$$\mathbf{e}' \left\{ \frac{1}{M} \sum_{j=1}^{M} \left(1 - \frac{Y_j}{1 + Y_j} \right) \mathbf{m}_j(\boldsymbol{\theta}_0) \right\} = 0,$$

$$\mathbf{e}' \left(\frac{1}{M} \sum_{j=1}^{M} \frac{\boldsymbol{\alpha}' \mathbf{m}_j(\boldsymbol{\theta}_0)}{1 + Y_j} \mathbf{m}_j(\boldsymbol{\theta}_0) \right) = \mathbf{e}' \left(\frac{1}{M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) \right), \tag{5.47}$$

$$\|\boldsymbol{\alpha}\| \mathbf{e}' \left(\frac{1}{M} \sum_{j=1}^{M} \frac{\mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)'}{1 + Y_j} \right) \mathbf{e} = \mathbf{e}' \left(\frac{1}{M} \sum_{j=1}^{M} \mathbf{m}_j(\boldsymbol{\theta}_0) \right).$$

Every $w_j > 0$, so $1 + Y_j > 0$, and therefore by (5.47), we get

$$\|\boldsymbol{\alpha}\| \mathbf{e}' \mathbf{S}_M \mathbf{e} \leq \|\boldsymbol{\alpha}\| \mathbf{e}' \left(\frac{1}{2\pi M} \sum_{j=1}^{M} \frac{\mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)'}{1 + Y_j} \right) \mathbf{e} \cdot \left(1 + \max_j Y_j \right)$$

$$\leq \|\boldsymbol{\alpha}\| \mathbf{e}' \left(\frac{1}{2\pi M} \sum_{j=1}^{M} \frac{\mathbf{m}_j(\boldsymbol{\theta}_0) \mathbf{m}_j(\boldsymbol{\theta}_0)'}{1 + Y_j} \right) \mathbf{e} \cdot (1 + \|\boldsymbol{\alpha}\| Z_M^*) \tag{5.48}$$

$$= \mathbf{e}' M^{-1/2} \mathbf{P}_M (1 + \|\boldsymbol{\alpha}\| Z_M^*),$$

where $\mathbf{S}_M$ and $\mathbf{P}_M$ are defined in Lemmas 5.2 and 5.3. Then by (5.48), we get

$$\|\boldsymbol{\alpha}\| \left\{ \mathbf{e}' \mathbf{S}_M \mathbf{e} - Z_M^* \mathbf{e}' \left(M^{-1/2} \mathbf{P}_M \right) \right\} \leq \mathbf{e}' \left(M^{-1/2} \mathbf{P}_M \right). \tag{5.49}$$

From Lemmas 5.2 and 5.3 we can see that

$$M^{-1/2} \mathbf{P}_M = O_p \left(M^{-1/2} \right), \qquad \mathbf{S}_M = O_p(1). \tag{5.50}$$

We evaluate the order of Z_M^*. We can write

$$Z_M^* \leq \max_{1 \leq j \leq M} \int_{-\pi}^{\pi} \left\| \boldsymbol{\phi}_{\boldsymbol{\theta}_0}(u_j, \lambda) \right\| I_N(u_j, \lambda) d\lambda =: \max_{1 \leq j \leq M} m_j^*(\boldsymbol{\theta}_0) \quad \text{(say)}. \tag{5.51}$$

Then, for any $\varepsilon > 0$,

$$P\left(\max_{1 \leq j \leq M} m_j^*(\boldsymbol{\theta}_0) > \varepsilon\sqrt{M}\right) \leq \sum_{j=1}^{M} P\left(m_j^*(\boldsymbol{\theta}_0) > \varepsilon\sqrt{M}\right)$$

$$= \sum_{j=1}^{M} P\left(m_j^*(\boldsymbol{\theta}_0)^3 > \left(\varepsilon\sqrt{M}\right)^3\right) \leq \sum_{j=1}^{M} \frac{1}{\varepsilon^3 M^{3/2}} E\left|m_j^*(\boldsymbol{\theta}_0)\right|^3$$

$$= \frac{1}{\varepsilon^3 M^{3/2}} \sum_{j=1}^{M} \int\int\int_{-\pi}^{\pi} \|\boldsymbol{\phi}_\theta(u_j, \lambda_1)\boldsymbol{\phi}_\theta(u_j, \lambda_2)\boldsymbol{\phi}_\theta(u_j, \lambda_3)\| \tag{5.52}$$

$$\times E\left[I_N(u_j, \lambda_1)I_N(u_j, \lambda_2)I_N(u_j, \lambda_3)\right]d\lambda_1 d\lambda_2 d\lambda_3.$$

The above expectation is written as

$$E\left[I_N(u_j, \lambda_1)I_N(u_j, \lambda_2)I_N(u_j, \lambda_3)\right]$$

$$= \frac{1}{(2\pi H_{2,N})^3} \text{ cum } \left[d_N(u_j, \lambda_1)d_N(u_j, -\lambda_1)d_N(u_j, \lambda_2)\right. \tag{5.53}$$

$$\left.\times d_N(u_j, -\lambda_2)d_N(u_j, \lambda_3)d_N(u_j, -\lambda_3)\right].$$

From Lemma 5.1 this is of order $O_p(1)$, so we can see that (5.52) tends to zero, which leads

$$Z_M^* = o_p\left(M^{1/2}\right). \tag{5.54}$$

From (5.49), (5.50), and (5.54), it is seen that

$$\|\boldsymbol{\alpha}\|\left[O_p(1) - o_p\left(M^{-1/2}\right)O_p\left(M^{-1/2}\right)\right] \leq O_p\left(M^{-1/2}\right). \tag{5.55}$$

Therefore,

$$\|\boldsymbol{\alpha}\| = O_p\left(M^{-1/2}\right). \tag{5.56}$$

Now we have from (5.54) that

$$\max_{1 \leq t \leq T}|Y_j| = O_p\left(M^{-1/2}\right)o_p\left(M^{1/2}\right) = o_p(1) \tag{5.57}$$

and from (5.30) that

$$
\begin{aligned}
\mathbf{0} &= \frac{1}{M}\sum_{j=1}^{M}\mathbf{m}_j(\boldsymbol{\theta}_0)\frac{1}{1+Y_j} \\
&= \frac{1}{M}\sum_{j=1}^{M}\mathbf{m}_j(\boldsymbol{\theta}_0)\left(1 - Y_j + \frac{Y_j^2}{1+Y_j}\right) \\
&= 2\pi M^{-1/2}\mathbf{P}_M - 2\pi\mathbf{S}_M\boldsymbol{\alpha} + \frac{1}{M}\sum_{j=1}^{M}\frac{\mathbf{m}_j(\boldsymbol{\theta}_0)Y_j^2}{1+Y_j}.
\end{aligned}
\tag{5.58}
$$

Noting that

$$
\frac{1}{M}\sum_{j=1}^{M}\left\|\mathbf{m}_j(\boldsymbol{\theta}_0)\right\|^3 \le \frac{1}{M}\sum_{j=1}^{M}Z_M^*\left\|\mathbf{m}_j(\boldsymbol{\theta}_0)\right\|^2 = o_p\left(M^{1/2}\right),
\tag{5.59}
$$

we can see that the final term in (5.58) has a norm bounded by

$$
\frac{1}{M}\sum_{j=1}^{M}\left\|\mathbf{m}_j(\boldsymbol{\theta}_0)\right\|^3\|\boldsymbol{\alpha}\|^2|1+Y_j|^{-1} = o_p\left(M^{1/2}\right)O_p\left(M^{-1}\right)O_p(1) = o_p\left(M^{-1/2}\right).
\tag{5.60}
$$

Hence, we can write

$$
\boldsymbol{\alpha} = M^{-1/2}\mathbf{S}_M^{-1}\mathbf{P}_M + \boldsymbol{\epsilon},
\tag{5.61}
$$

where $\boldsymbol{\epsilon} = o_p(M^{-1/2})$. By (5.57), we may write

$$
\log(1 + Y_j) = Y_j - \frac{1}{2}Y_j^2 + \eta_j,
\tag{5.62}
$$

where for some finite K

$$
\Pr\left(|\eta_j| \le K|Y_j|^3, 1 \le j \le M\right) \longrightarrow 1 \quad (T \longrightarrow \infty).
\tag{5.63}
$$

We may write

$$
\begin{aligned}
-\frac{1}{\pi} \log \mathcal{R}(\boldsymbol{\theta}_0) &= -\frac{1}{\pi} \sum_{j=1}^{M} \log(T w_j) = \frac{1}{\pi} \sum_{j=1}^{M} \log(1 + Y_j) \\
&= \frac{1}{\pi} \sum_{j=1}^{M} Y_j - \frac{1}{2\pi} \sum_{j=1}^{M} Y_j^2 + \frac{1}{\pi} \sum_{j=1}^{M} \eta_j \\
&= \mathbf{P}_M' \mathbf{S}_M^{-1} \mathbf{P}_M - M \boldsymbol{\epsilon}' \mathbf{S}_M \boldsymbol{\epsilon} + \frac{1}{\pi} \sum_{j=1}^{M} \eta_t \\
&= (A) - (B) + (C) \quad \text{(say)}.
\end{aligned}
\tag{5.64}
$$

Here it is seen that

$$
\begin{aligned}
(B) &= M o_p\left(M^{-1/2}\right) O_p(1) o_p\left(M^{-1/2}\right) = o_p(1), \\
(C) &\le K \|\boldsymbol{\alpha}\|^3 \sum_{j=1}^{M} \|\mathbf{m}_j(\boldsymbol{\theta}_0)\|^3 = O_p\left(M^{-3/2}\right) o_p\left(M^{3/2}\right) = o_p(1).
\end{aligned}
\tag{5.65}
$$

And finally from Lemmas 5.2 and 5.3, we can show that

$$
\begin{aligned}
(A) &\xrightarrow{d} \left(\boldsymbol{\Sigma}_2^{-1/2} \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_1^{-1/2} \mathbf{P}_M\right)' \left(\boldsymbol{\Sigma}_2^{-1/2} \boldsymbol{\Sigma}_1^{1/2} \boldsymbol{\Sigma}_1^{-1/2} \mathbf{P}_M\right) \\
&\xrightarrow{d} (\mathbf{FN})'(\mathbf{FN}).
\end{aligned}
\tag{5.66}
$$

Then we can obtain the desired result.

Acknowledgments

References

[1] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.

[2] A. B. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.

[3] J. Qin and J. Lawless, "Empirical likelihood and general estimating equations," *The Annals of Statistics*, vol. 22, no. 1, pp. 300–325, 1994.

[4] A. B. Owen, "Empirical likelihood for linear models," *The Annals of Statistics*, vol. 19, no. 4, pp. 1725–1747, 1991.

[5] S. X. Chen, "On the accuracy of empirical likelihood confidence regions for linear regression model," *Annals of the Institute of Statistical Mathematics*, vol. 45, no. 4, pp. 621–637, 1993.

[6] S. X. Chen, "Empirical likelihood confidence intervals for linear regression coefficients," *Journal of Multivariate Analysis*, vol. 49, no. 1, pp. 24–40, 1994.

[7] J. Qin, "Empirical likelihood in biased sample problems," *The Annals of Statistics*, vol. 21, no. 3, pp. 1182–1196, 1993.

[8] Y. Kitamura, "Empirical likelihood methods with weakly dependent processes," *The Annals of Statistics*, vol. 25, no. 5, pp. 2084–2102, 1997.

[9] A. C. Monti, "Empirical likelihood confidence regions in time series models," *Biometrika*, vol. 84, no. 2, pp. 395–405, 1997.

[10] D. J. Nordman and S. N. Lahiri, "A frequency domain empirical likelihood for short- and long-range dependence," *The Annals of Statistics*, vol. 34, no. 6, pp. 3019–3050, 2006.

[11] R. Dahlhaus, "On the Kullback-Leibler information divergence of locally stationary processes," *Stochastic Processes and Their Applications*, vol. 62, no. 1, pp. 139–168, 1996.

[12] R. Dahlhaus, "Asymptotic statistical inference for nonstationary processes with evolutionary spectra," in *Proceedings of the Athens Conference on Applied Probability and Time Series Analysis*, vol. 115 of *Lecture Notes in Statistics*, pp. 145–159, Springer, 1996.

[13] R. Dahlhaus, "Fitting time series models to nonstationary processes," *The Annals of Statistics*, vol. 25, no. 1, pp. 1–37, 1997.

[14] Y. Hosoya and M. Taniguchi, "A central limit theorem for stationary processes and the parameter estimation of linear processes," *The Annals of Statistics*, vol. 10, no. 1, pp. 132–153, 1982, Correction: vol 21, pp. 1115–1117, 1993.

[15] Y. Kakizawa, "Parameter estimation and hypothesis testing in stationary vector time series," *Statistics & Probability Letters*, vol. 33, no. 3, pp. 225–234, 1997.

[16] D. R. Brillinger, *Time Series: Data Analysis and Theory*, Holden-Day, San Francisco, Calif, USA, 2001.

Advances in
Operations Research

Advances in
Decision Sciences

Mathematical Problems
in Engineering

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Journal of
Complex Analysis

Journal of
Mathematics

International Journal of
Stochastic Analysis

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

Journal of
Applied Mathematics

Journal of
Optimization