Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 94 – 116
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Innovations algorithm asymptotics for periodically stationary time series with heavy tails

Paul L. Anderson[a], Laimonis Kavalieris[b], Mark M. Meerschaert[c,*,1]

[a]Department of Mathematics, Albion College, MI, USA
[b]Department of Mathematics & Statistics, University of Otago, Dunedin, New Zealand
[c]Department of Statistics and Probability, Michigan State University, East Lansing, MI 48823, USA

Received 9 February 2006
Available online 7 March 2007

Abstract

The innovations algorithm can be used to obtain parameter estimates for periodically stationary time series models. In this paper we compute the asymptotic distribution for these estimates in the case where the underlying noise sequence has infinite fourth moment but finite second moment. In this case, the sample covariances on which the innovations algorithm are based are known to be asymptotically stable. The asymptotic results developed here are useful to determine which model parameters are significant. In the process, we also compute the asymptotic distributions of least squares estimates of parameters in an autoregressive model.

AMS 1991 subject classification: primary 62M10; 62E20; secondary 60E07; 60F05

Keywords: Time series; Periodically stationary; Innovations algorithm

1. Introduction

A stochastic process X_t is called periodically stationary (in the wide sense) if $\mu_t = EX_t$ and $\gamma_t(h) = EX_tX_{t+h}$ for $h = 0, \pm 1, \pm 2, \dots$ are all periodic functions of time t with the same period $\nu \geqslant 1$. If $\nu = 1$ then the process is stationary. Periodically stationary processes manifest themselves in such fields as economics, hydrology, and geophysics, where the observed time series are characterized by seasonal variations in both the mean and covariance structure. An important

* Corresponding author. Fax: +1 517 432 1405.
E-mail address: mcubed@stt.msu.edu (M.M. Meerschaert).

[1] Partially supported by NSF grants DES-9980484, DMS-0139927, DMS-0417869.

doi:10.1016/j.jmva.2007.02.005

class of stochastic models for describing periodically stationary time series are the periodic ARMA models, in which the model parameters are allowed to vary with the season. Periodic ARMA models are developed by many authors including [1,2,4–7,20,22–24,26,28,30,31,33–41].

Anderson et al. [5] develop the innovations algorithm for periodic ARMA model parameters. Anderson and Meerschaert [4] develop the asymptotics necessary to determine which of these estimates are statistically different from zero, under the classical assumption that the noise sequence has finite fourth moment. In this paper, we extend those results to the case where the noise sequence has finite second moment but infinite fourth moment. This case is important in applications to river flows, see for example Anderson and Meerschaert [3]. In that case, Anderson and Meerschaert [2] proved that the sample autocovariances, the basis for the innovations algorithm estimates of the model parameters, are asymptotically stable. Surprisingly, the innovations estimates themselves turn out to be asymptotically normal, although the rate of convergence (in terms of the number of iterations of the innovations algorithm) is slower than in the finite fourth moment case. Brockwell and Davis [13] discuss asymptotics of the innovations algorithm for stationary time series, using results of Berk [8] and Bhansali [10]. However, all of these results assume a finite fourth moment for the noise sequence. Hence our results seem to be new even in the stationary case when the period $\nu = 1$. Since our technical approach extends that of [13], we also need to develop periodically stationary analogues of results in [8,10] for the infinite fourth moment case. In particular, we obtain asymptotics for the least squares estimates of a periodically stationary process. Although the innovations estimates are more useful in practice, the asymptotics of the least squares estimates are also of some independent interest.

2. The innovations algorithm

The periodic ARMA process $\{\tilde{X}_t\}$ with period ν (denoted by PARMA$_\nu(p, q)$) has representation

$$X_t - \sum_{j=1}^{p} \phi_t(j) X_{t-j} = \varepsilon_t - \sum_{j=1}^{q} \theta_t(j) \varepsilon_{t-j}, \tag{1}$$

where $X_t = \tilde{X}_t - \mu_t$ and $\{\varepsilon_t\}$ is a sequence of random variables with mean zero and standard deviation σ_t such that $\{\sigma_t^{-1}\varepsilon_t\}$ is i.i.d. The autoregressive parameters $\phi_t(j)$, the moving average parameters $\theta_t(j)$, and the residual standard deviations σ_t are all periodic functions of t with the same period $\nu \geqslant 1$. We also assume that the model admits a causal representation

$$X_t = \sum_{j=0}^{\infty} \psi_t(j) \varepsilon_{t-j}, \tag{2}$$

where $\psi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\psi_t(j)| < \infty$ for all t, and satisfies an invertibility condition

$$\varepsilon_t = \sum_{j=0}^{\infty} \pi_t(j) X_{t-j}, \tag{3}$$

where $\pi_t(0) = 1$ and $\sum_{j=0}^{\infty} |\pi_t(j)| < \infty$ for all t. We will say that the i.i.d. noise sequence $\delta_t = \sigma_t^{-1}\varepsilon_t$ is RV(α) if $P[|\delta_t| > x]$ varies regularly with index $-\alpha$ and $P[\delta_t > x]/P[|\delta_t| > x] \to p$ for some $p \in [0, 1]$. The case where $E|\delta_t|^4 < \infty$ was treated in Anderson and Meerschaert [4]. In this paper, we assume that the noise sequence $\{\delta_t\}$ is RV(α) for some $2 < \alpha < 4$. This assumption implies that $E|\delta_t|^p < \infty$ if $0 < p < \alpha$, in particular the variance of ε_t exists. With this technical condition, Anderson and Meerschaert [2] show that the sample autocovariance is a consistent estimator of the autocovariance, and asymptotically stable with tail index $\alpha/2$. Stable laws and processes are comprehensively treated in, e.g., Feller [16], Samorodnitsky and Taqqu [29], and Meerschaert and Scheffler [25].

Let $\hat{X}^{(i)}_{i+k} = P_{\mathcal{H}_{k,i}} X_{i+k}$ denote the one-step predictors, where $\mathcal{H}_{k,i} = \overline{\mathrm{sp}}\{X_i, \ldots, X_{i+k-1}\}$, $k \geqslant 1$, and $P_{\mathcal{H}_{k,i}}$ is the orthogonal projection onto this space, which minimizes the mean squared error

$$v_{k,i} = \|X_{i+k} - \hat{X}^{(i)}_{i+k}\|^2 = E(X_{i+k} - \hat{X}^{(i)}_{i+k})^2.$$

Recall that $\gamma_t(h) = \mathrm{Cov}(X_t, X_{t+h})$, $h = 0, \pm 1, \pm 2, \ldots$. Then

$$\hat{X}^{(i)}_{i+k} = \phi^{(i)}_{k,1} X_{i+k-1} + \cdots + \phi^{(i)}_{k,k} X_i, \quad k \geqslant 1, \tag{4}$$

where the vector of coefficients $\phi^{(i)}_k = (\phi^{(i)}_{k,1}, \ldots, \phi^{(i)}_{k,k})'$ solves the prediction equations

$$\Gamma_{k,i} \phi^{(i)}_k = \gamma^{(i)}_k \tag{5}$$

with $\gamma^{(i)}_k = (\gamma_{i+k-1}(1), \gamma_{i+k-2}(2), \ldots, \gamma_i(k))'$ and

$$\Gamma_{k,i} = \left[\gamma_{i+k-\ell}(\ell - m)\right]_{\ell,m=1,\ldots,k} \tag{6}$$

is the covariance matrix of $(X_{i+k-1}, \ldots, X_i)'$ for each $i = 0, \ldots, v-1$. Let

$$\hat{\gamma}_i(\ell) = N^{-1} \sum_{j=0}^{N-1} X_{jv+i} X_{jv+i+\ell} \tag{7}$$

denote the (uncentered) sample autocovariance, where $X_t = \tilde{X}_t - \mu_t$. If we replace the autocovariances in the prediction equation (5) with their corresponding sample autocovariances, we obtain the estimator $\hat{\phi}^{(i)}_{k,j}$ of $\phi^{(i)}_{k,j}$.

Because the scalar-valued process X_t is non-stationary, the Durbin–Levinson algorithm (see, e.g., [14, Proposition 5.2.1]) for computing $\hat{\phi}^{(i)}_{k,j}$ does not apply. However, the innovations algorithm (see, e.g., [14, Proposition 5.2.2]) still applies to a non-stationary process. Writing

$$\hat{X}^{(i)}_{i+k} = \sum_{j=1}^{k} \theta^{(i)}_{k,j} (X_{i+k-j} - \hat{X}^{(i)}_{i+k-j}) \tag{8}$$

yields the one-step predictors in terms of the innovations $X_{i+k-j} - \hat{X}^{(i)}_{i+k-j}$. Proposition 4.1 of Lund and Basawa [22] shows that if $\sigma_i^2 > 0$ for $i = 0, \ldots, v-1$, then for a causal $\mathrm{PARMA}_v(p, q)$ process the covariance matrix $\Gamma_{k,i}$ is non-singular for every $k \geqslant 1$ and each i. Anderson et al. [5] show that if $EX_t = 0$ and $\Gamma_{k,i}$ is nonsingular for each $k \geqslant 1$, then the one-step predictors $\hat{X}_{i+k}$,

$k \geqslant 0$, and their mean-square errors $v_{k,i}$, $k \geqslant 1$, are given by

$$\begin{aligned}
v_{0,i} &= \gamma_i(0),\\
\theta^{(i)}_{k,k-\ell} &= (v_{\ell,i})^{-1}\left[\gamma_{i+\ell}(k-\ell) - \sum_{j=0}^{\ell-1} \theta^{(i)}_{\ell,\ell-j}\theta^{(i)}_{k,k-j} v_{j,i}\right],\\
v_{k,i} &= \gamma_{i+k}(0) - \sum_{j=0}^{k-1} (\theta^{(i)}_{k,k-j})^2 v_{j,i},
\end{aligned} \tag{9}$$

where (9) is solved in the order $v_{0,i}$, $\theta^{(i)}_{1,1}$, $v_{1,i}$, $\theta^{(i)}_{2,2}$, $\theta^{(i)}_{2,1}$, $v_{2,i}$, $\theta^{(i)}_{3,3}$,$\theta^{(i)}_{3,2}$, $\theta^{(i)}_{3,1}$, $v_{3,i}, \ldots$. The results in [5] show that

$$\begin{aligned}
\theta^{(\langle i-k\rangle)}_{k,j} &\to \psi_i(j),\\
v_{k,\langle i-k\rangle} &\to \sigma_i^2,\\
\phi^{(\langle i-k\rangle)}_{k,j} &\to -\pi_i(j)
\end{aligned} \tag{10}$$

as $k \to \infty$ for all i, j, where $\langle j\rangle = j \bmod \nu$.

If we replace the autocovariances in (9) with the corresponding sample autocovariances (7), we obtain the innovations estimates $\hat{\theta}^{(i)}_{k,\ell}$ and $\hat{v}_{k,i}$. Similarly, replacing the autocovariances in (5) with the corresponding sample autocovariances yields the least squares or Yule–Walker estimators $\hat{\phi}^{(i)}_{k,\ell}$. The consistency of these estimators was also established in [5]. Suppose that $\{X_t\}$ is the mean zero PARMA process with period ν given by (1). Assume that the spectral density matrix $f(\lambda)$ of the equivalent vector ARMA process is such that $mz'z \leqslant z'f(\lambda)z \leqslant Mz'z$, $-\pi \leqslant \lambda \leqslant \pi$, for some m and M such that $0 < m \leqslant M < \infty$ and for all z in R^ν. Recall that the i.i.d. noise sequence $\delta_t = \sigma_t^{-1}\varepsilon_t$ is RV(α) for some $2 < \alpha < 4$, viz., the noise sequence has infinite fourth moment but finite variance, and define

$$a_N = \inf\{x : P(|\delta_t| > x) < 1/N\} \tag{11}$$

a regularly varying sequence with index $1/\alpha$, see for example Proposition 6.1.37 in [25]. If k is chosen as a function of the sample size N so that $k^{5/2}a_N^2/N \to 0$ as $N \to \infty$ and $k \to \infty$, then the results in Theorems 3.5–3.7 and Corollary 3.7 of [5], specific to the infinite fourth moment case, also show that

$$\begin{aligned}
\hat{\theta}^{(\langle i-k\rangle)}_{k,j} &\xrightarrow{P} \psi_i(j),\\
\hat{v}_{k,\langle i-k\rangle} &\xrightarrow{P} \sigma_i^2,\\
\hat{\phi}^{(\langle i-k\rangle)}_{k,j} &\xrightarrow{P} -\pi_i(j)
\end{aligned} \tag{12}$$

for all i, j. This yields a practical method for estimating the model parameters, in the case of infinite fourth moments. The results of Section 3 can then be used to determine which of these model parameters are statistically significantly different from zero.

3. Asymptotic results

In this section, we compute the asymptotic distribution for the innovations estimates of the parameters in a periodically stationary time series (2) with period $v \geqslant 1$. In the process, we also obtain the asymptotic distribution of the least squares estimates. For any periodically stationary time series, we can construct an equivalent (stationary) vector moving average process in the following way: Let $Z_t = (\varepsilon_{tv}, \ldots, \varepsilon_{(t+1)v-1})'$ and $Y_t = (X_{tv}, \ldots, X_{(t+1)v-1})'$, so that

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}, \tag{13}$$

where Ψ_j is the $v \times v$ matrix with $i\ell$ entry $\psi_i(tv + i - \ell)$, and we number the rows and columns $0, 1, \ldots, v-1$ for ease of notation. Also, let $N(m, C)$ denote a Gaussian random vector with mean m and covariance matrix C, and let $\Rightarrow$ indicate convergence in distribution. Our first result gives the asymptotics of the least squares estimates in the case where the noise sequence has heavy tails with an infinite fourth moment but finite second moment. The corresponding result in the case where the noise sequence has finite fourth moments was obtained by Anderson and Meerschaert [4]. A similar result was obtained in the finite fourth moment case by Lewis and Reinsel [21] for vector autoregressive models, however, the prediction problem here is different. For example, suppose that (2) represents monthly data with $v = 12$. For a periodically stationary model, the prediction equations (4) use observations for earlier months in the same year. For the equivalent vector moving average model, the prediction equations use only observations from past years.

Theorem 3.1. *Suppose that the periodically stationary moving average* (2) *is causal, invertible, the i.i.d. noise sequence* $\delta_t = \sigma_t^{-1}\varepsilon_t$ *is* $RV(\alpha)$, *and that for some* $0<m\leqslant M<\infty$ *we have* $mz'z \leqslant z'f(\lambda)z \leqslant Mz'z$ *for all* $-\pi \leqslant \lambda \leqslant \pi$, *and all* z *in* $\mathbb{R}^v$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (13). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3 a_N^2/N \to 0$ *where* a_N *is defined by* (11), *and if*

$$N^{1/2} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \to 0 \quad \text{for } \ell = 0, 1, \ldots, v-1 \tag{14}$$

then for any fixed positive integer D

$$N^{1/2}\left(\pi_i(u) + \hat{\phi}_{k,u}^{\langle i-k\rangle} : 1 \leqslant u \leqslant D, i = 0, \ldots, v-1\right) \Rightarrow N(0, \Lambda) \tag{15}$$

where

$$\Lambda = \operatorname{diag}(\sigma_0^2 \Lambda^{(0)}, \sigma_1^2 \Lambda^{(1)}, \ldots, \sigma_{v-1}^2 \Lambda^{(v-1)}) \tag{16}$$

with

$$(\Lambda^{(i)})_{u,v} = \sum_{s=0}^{m-1} \pi_{i-m+s}(s) \pi_{i-m+s}(s + |v-u|) \sigma_{i-m+s}^{-2} \tag{17}$$

and $m = \min(u, v)$, $1 \leqslant u, v \leqslant D$.

In Theorem 3.1, note that $\Lambda^{(i)}$ is a $D \times D$ matrix and the $D\nu$-dimensional vector given in (15) is ordered

$$N^{1/2}(\pi_0(1) + \hat{\phi}_{k,1}^{\langle 0-k\rangle}, \ldots, \pi_0(D) + \hat{\phi}_{k,D}^{\langle 0-k\rangle}, \ldots, \pi_{\nu-1}(1) + \hat{\phi}_{k,1}^{\langle \nu-1-k\rangle}, \ldots, \pi_{\nu-1}(D) + \hat{\phi}_{k,D}^{\langle \nu-1-k\rangle})'.$$

Note also that a_N is roughly on the order of $N^{1/\alpha}$ for some $2 < \alpha < 4$ so that the condition on k is essentially that k^3 grows slower than $N^{1-2/\alpha}$. In practice, on the boundary $\alpha = 2 + \eta$, $\eta > 0$, we look for the value k in the innovations algorithm where the estimates have stabilized. Next we present our main result, giving asymptotics for innovations estimates of a periodically stationary time series.

Theorem 3.2. *Suppose that the periodically stationary moving average* (2) *is causal*, *invertible*, *the i.i.d. noise sequence* $\delta_t = \sigma_t^{-1}\varepsilon_t$ *is* $RV(\alpha)$, *and that for some* $0<m\leqslant M<\infty$ *we have* $mz'z \leqslant z' f(\lambda) z \leqslant M z'z$ *for all* $-\pi \leqslant \lambda \leqslant \pi$, *and all* z *in* $\mathbb{R}^\nu$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (13). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3 a_N^2/N \to 0$ *where* a_N *is defined by* (11), *and if*

$$N^{1/2} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \to 0 \quad \text{for } \ell = 0, 1, \ldots, \nu - 1 \tag{18}$$

then

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-k\rangle)} - \psi_i(u) : u = 1, \ldots, D, i = 0, \ldots, \nu - 1) \Rightarrow N(0, V) \tag{19}$$

where

$$V = A \,\mathrm{diag}(\sigma_0^2 D^{(0)}, \ldots, \sigma_{\nu-1}^2 D^{(\nu-1)}) A', \tag{20}$$

$$A = \sum_{n=0}^{D-1} E_n \Pi^{[D\nu - n(D+1)]}, \tag{21}$$

$$E_n = \mathrm{diag}\Bigg(\overbrace{0, \ldots, 0}^{n}, \overbrace{\psi_0(n), \ldots, \psi_0(n)}^{D-n}, \overbrace{0, \ldots, 0}^{n}, \underbrace{\psi_1(n), \ldots, \psi_1(n)}_{D-n}, \ldots, \underbrace{0, \ldots, 0}_{n}, \underbrace{\psi_{\nu-1}(n), \ldots, \psi_{\nu-1}(n)}_{D-n}\Bigg), \tag{22}$$

$$D^{(i)} = \mathrm{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2}), \tag{23}$$

and Π *an orthogonal* $D\nu \times D\nu$ *cyclic permutation matrix*,

$$\Pi = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{24}$$

Note that Π^0 is the $Dv \times Dv$ identity matrix and $\Pi^{-\ell} \equiv (\Pi')^{\ell}$. Matrix multiplication yields the following corollary.

Corollary 3.3. *Regarding Theorem* 3.2, *in particular*, *we have that*

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-k\rangle)} - \psi_i(u)) \Rightarrow N\left(0, \sigma_{i-u}^{-2}\sum_{n=0}^{u-1}\sigma_{i-n}^2\psi_i^2(n)\right) \tag{25}$$

Remark. Corollary 3.3 also holds the asymptotic result for the second-order stationary process where the period is just $v = 1$. In this case $\sigma_i^2 = \sigma^2$ so (25) becomes

$$N^{1/2}(\hat{\theta}_{k,u} - \psi(u)) \Rightarrow N\left(0, \sum_{n=0}^{u-1}\psi^2(n)\right)$$

which extends Theorem 2.1 in [13] to the case where the noise sequence has only moments of order $2+\eta$, $\eta > 0$.

4. Proofs

Theorem 3.1 depends on modulo v arithmetic which requires our $\langle i-k\rangle$-notation. Since the lemmas in this section do not have this dependence, we proceed with the less cumbersome i-notation.

Lemma 4.1. *Let* $\pi_k^{(i)} = (\pi_{i+k}(1), \ldots, \pi_{i+k}(k))'$ *and* $X_j^{(i)}(k) = (X_{jv+i+k-1}, \ldots, X_{jv+i})'$. *Then for all* $i = 0, \ldots, v-1$ *and* $k \geqslant 1$ *we have*

$$\pi_k^{(i)} + \hat{\phi}_k^{(i)} = (\hat{\Gamma}_{k,i})^{-1}\frac{1}{N}\sum_{j=0}^{N-1}X_j^{(i)}(k)\varepsilon_{jv+i+k,k} \tag{26}$$

where $\varepsilon_{t,k} = X_t + \pi_t(1)X_{t-1} + \cdots + \pi_t(k)X_{t-k}$.

Proof. The least squares equations are

$$\hat{\phi}_k^{(i)} = \left(\hat{\Gamma}_{k,i}\right)^{-1}\hat{\gamma}_k^{(i)} \tag{27}$$

where

$$\hat{\Gamma}_{k,i} = \frac{1}{N}\sum_{j=0}^{N-1}\begin{pmatrix} X_{jv+i+k-1} \\ \vdots \\ X_{jv+i}\end{pmatrix}(X_{jv+i+k-1}, \ldots, X_{jv+i})$$

and

$$\hat{\gamma}_k^{(i)} = \frac{1}{N}\sum_{j=0}^{N-1}\begin{pmatrix} X_{jv+i+k-1} \\ \vdots \\ X_{jv+i}\end{pmatrix}X_{jv+i+k}$$

from Eq. (7). Thus,

$$\hat{\Gamma}_{k,i}\pi_k^{(i)} + \hat{\gamma}_k^{(i)} = \frac{1}{N}\sum_{j=0}^{N-1}\begin{pmatrix} X_{jv+i+k-1} \\ \vdots \\ X_{jv+i} \end{pmatrix} \times \left\{ (X_{jv+i+k-1}, \ldots, X_{jv+i}) \begin{pmatrix} \pi_{i+k}(1) \\ \vdots \\ \pi_{i+k}(k) \end{pmatrix} + X_{jv+i+k} \right\}$$

$$= \frac{1}{N}\sum_{j=0}^{N-1}\begin{pmatrix} X_{jv+i+k-1} \\ \vdots \\ X_{jv+i} \end{pmatrix}\varepsilon_{t,k}$$

since $X_{jv+i+k} + \pi_{i+k}(1)X_{jv+i+k-1} + \cdots + \pi_{i+k}(k)X_{jv+i} = \varepsilon_{t,k}$ by definition. Then, using the least squares equations (27), we have

$$\pi_k^{(i)} + \hat{\phi}_k^{(i)} = \left(\hat{\Gamma}_{k,i}\right)^{-1}\left(\hat{\Gamma}_{k,i}\pi_k^{(i)} + \hat{\gamma}_k^{(i)}\right)$$

$$= \left(\hat{\Gamma}_{k,i}\right)^{-1}\frac{1}{N}\sum_{j=0}^{N-1}\begin{pmatrix} X_{jv+i+k-1} \\ \vdots \\ X_{jv+i} \end{pmatrix}\varepsilon_{t,k}$$

$$= \left(\hat{\Gamma}_{k,i}\right)^{-1}\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{t,k}$$

which finishes the proof of Lemma 4.1. □

Lemma 4.2. *Let $c_i(\ell), d_i(\ell)$ for $\ell = 0, 1, 2, \ldots$ and $i = 0, \ldots, v-1$ be arbitrary sequences of real numbers such that $\sum_{\ell=0}^{\infty}|c_i(\ell)| < \infty$ and $\sum_{\ell=0}^{\infty}|d_i(\ell)| < \infty$, and let*

$$u_{tv+i} = \sum_{k=0}^{\infty} c_i(k)\varepsilon_{tv+i-k} \quad \text{and} \quad v_{tv+j} = \sum_{m=0}^{\infty} d_j(m)\varepsilon_{tv+j-m}$$

and set

$$C_i = \sum_{k=0}^{\infty}|c_i(k)| \quad \text{and} \quad D_j = \sum_{m=0}^{\infty}|d_j(m)|$$

for $0 \leqslant i, j < v$. Then

$$E\left|\sum_{t=1}^{M} u_{tv+i}v_{tv+j}\right| \leqslant MCD\eta, \tag{28}$$

where $C = \max(C_i)$, $D = \max(D_j)$, and $\eta = \max_{t,t'}(\sigma_t\sigma_{t'}\mu_1^2, \sigma_t^2)$ where $\mu_1 = E(|\delta_t|)$.

Proof. Write

$$E\left|\sum_{t=1}^{M} u_{tv+i}v_{tv+j}\right| \leqslant E\sum_{t=1}^{M}\left|u_{tv+i}v_{tv+j}\right|,$$

where

$$\begin{aligned}\left|u_{tv+i}v_{tv+j}\right| &= \left|\sum_{k=0}^{\infty}\sum_{m=0}^{\infty} c_i(k)d_j(m)\varepsilon_{tv+i-k}\varepsilon_{tv+j-m}\right| \\ &\leqslant \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\left|c_i(k)d_j(m)\right|\left|\varepsilon_{tv+i-k}\varepsilon_{tv+j-m}\right|.\end{aligned}$$

Then

$$E\left|u_{tv+i}v_{tv+j}\right| \leqslant \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\left|c_i(k)d_j(m)\right| \ E\left|\varepsilon_{tv+i-k}\varepsilon_{tv+j-m}\right|,$$

where

$$E\left|\varepsilon_{tv+i-k}\varepsilon_{tv+j-m}\right| = \begin{cases} E|\varepsilon_{tv+i-k}|E|\varepsilon_{tv+j-m}| & \text{if } i-k \neq j-m, \\ E(\varepsilon^2_{tv+i-k}) & \text{if } i-k = j-m. \end{cases}$$

Since $E|\varepsilon_{tv+i-k}| = E|\sigma_{i-k}\ \sigma^{-1}_{i-k}\varepsilon_{tv+i-k}| = E|\sigma_{i-k}\delta_{tv+i-k}| = \sigma_{i-k}\mu_1$, we have that

$$E\left|\varepsilon_{tv+i-k}\varepsilon_{tv+j-m}\right| = \begin{cases} \sigma_{i-k}\sigma_{j-m}\mu_1^2 & \text{if } i-k \neq j-m, \\ \sigma^2_{i-k} & \text{if } i-k = j-m. \end{cases}$$

Hence,

$$E\left|u_{tv+i}v_{tv+j}\right| \leqslant \eta \sum_{k=0}^{\infty}|c_i(k)| \sum_{m=0}^{\infty}\left|d_j(m)\right| \leqslant CD\eta$$

for $0 \leqslant i,\ j < v$, and then (28) follows easily, which finishes the proof of Lemma 4.2. □

Lemma 4.3. *For $\varepsilon_{t,k}$ as in Lemma 4.1 and u_{tv+i} as in Lemma 4.2 we have*

$$E\left|\sum_{t=0}^{N-1} u_{tv+i}(\varepsilon_{tv+\ell,k} - \varepsilon_{tv+\ell})\right| \leqslant \eta NCB \max_{\ell} \sum_{j=1}^{\infty} \pi_\ell(k+j), \tag{29}$$

where C, η are as in Lemma 4.2 and $B = \sum_{i=0}^{v-1}\sum_{\ell=0}^{\infty}|\psi_i(\ell)|$.

Proof. Write

$$\begin{aligned}\varepsilon_{tv+\ell} - \varepsilon_{tv+\ell,k} &= \sum_{m=k+1}^{\infty} \pi_\ell(m)X_{tv+\ell-m} \\ &= \sum_{m=k+1}^{\infty} \pi_\ell(m)\sum_{r=0}^{\infty}\psi_{\ell-m}(r)\varepsilon_{tv+\ell-m-r} \\ &= \sum_{j=1}^{\infty} d_{\ell,k}(k+j)\varepsilon_{tv+\ell-k-j},\end{aligned}$$

where

$$d_{\ell,k}(k+j) = \sum_{s=1}^{j}\pi_\ell(k+s)\psi_{\ell-(k+s)}(j-s).$$

Since $\{X_t\}$ is causal and invertible,

$$\begin{aligned}\sum_{j=1}^{\infty}|d_{\ell,k}(k+j)| &= \sum_{j=1}^{\infty}\left|\sum_{s=1}^{j}\pi_\ell(k+s)\psi_{\ell-(k+s)}(j-s)\right| \\ &\leqslant \sum_{s=1}^{\infty}|\pi_\ell(k+s)|\sum_{j=s}^{\infty}|\psi_{\ell-(k+s)}(j-s)| \\ &= \sum_{s=1}^{\infty}|\pi_\ell(k+s)|\sum_{r=0}^{\infty}|\psi_{\ell-(k+s)}(r)|\end{aligned}$$

is finite, and hence we have $\sum_{j=1}^{\infty}|d_{\ell,k}(k+j)| < \infty$. Now apply Lemma 4.2 with $v_{tv+\ell} = \varepsilon_{tv+\ell,k} - \varepsilon_{tv+\ell}$ to see that

$$E\left|\sum_{t=0}^{N-1}u_{tv+i}(\varepsilon_{tv+\ell,k}-\varepsilon_{tv+\ell})\right| \leqslant NCD_{\ell,k}\eta \leqslant NCD_k\eta,$$

where $D_{\ell,k} = \sum_{j=1}^{\infty}|d_{\ell,k}(k+j)|$ and $D_k = \max(D_{\ell,k} : 0 \leqslant \ell < v)$. Next compute

$$\begin{aligned}D_{\ell,k} = \sum_{j=1}^{\infty}|d_{\ell,k}(k+j)| &\leqslant \sum_{s=1}^{\infty}|\pi_\ell(k+s)|\sum_{r=0}^{\infty}|\psi_{\ell-(k+s)}(r)| \\ &\leqslant \sum_{s=1}^{\infty}|\pi_\ell(k+s)|\sum_{j=0}^{v-1}\sum_{r=0}^{\infty}|\psi_i(r)| \\ &= B\sum_{s=1}^{\infty}|\pi_\ell(k+s)|\end{aligned}$$

and (29) follows easily.

The next lemma employs the matrix 1-norm given by

$$\|M\|_1 = \max_{\|x\|_1=1}\|Mx\|_1,$$

where $\|x\|_1 = |x_1| + \cdots + |x_k|$ is the vector 1-norm (see, e.g., [17]). □

Lemma 4.4. *For $\varepsilon_{t,k}$ and $X_j^{(i)}(k)$ as in Lemma 4.1 we have that*

$$E\left\|N^{-1}\sum_{j=0}^{N-1}X_j^{(i)}(k)(\varepsilon_{jv+i+k,k}-\varepsilon_{jv+i+k})\right\|_1 \leqslant Ak\max_\ell\sum_{j=1}^{\infty}|\pi_\ell(k+j)|, \tag{30}$$

where $A = \eta B\max_i\sum_{\ell=0}^{\infty}|\psi_i(\ell)|$ and B is from Lemma 4.3.

Proof. Rewrite the left-hand side of (30) in the form

$$N^{-1}\sum_{s=0}^{k-1}E\left\|\sum_{t=0}^{N-1}X_{tv+i+s}(\varepsilon_{tv+i+k,k}-\varepsilon_{tv+i+k})\right\|_1$$

and apply Lemma 4.3 k times with $u_{tv+i} = X_{tv+i+s}$ for each $s = 0, \ldots, k-1$ and $C = \max_i\sum_{\ell=0}^{\infty}|\psi_i(\ell)|$ to obtain the upper bound of (30). □

Lemma 4.5. *Suppose that the periodically stationary moving average* (2) *is causal, invertible, and that the noise sequence* $\delta_t = \sigma_t^{-1}\varepsilon_t$ *is RV*(α) *with* $2 < \alpha < 4$. *Assume that for some* $0 < m \leqslant M < \infty$ *we have* $mz'z \leqslant z'f(\lambda)z \leqslant Mz'z$ *for all* $-\pi \leqslant \lambda \leqslant \pi$, *and all* z *in* $\mathbb{R}^\nu$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (13). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3 a_N^2/N \to 0$ *where* a_N *is defined by* (11) *and* (14) *holds then*

$$N^{1/2} b(k)'(\pi_k^{(i)} + \hat{\phi}_k^{(i)}) - N^{-1/2} b(k)' \Gamma_{k,i}^{-1} \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k} \xrightarrow{P} 0 \tag{31}$$

for any $b(k) = (b_{k1}, \ldots, b_{kk})'$ *such that* $\|b(k)\|_1$ *remains, of* k *and* $X_j^{(i)}(k)$ *is from Lemma* 4.1.

Proof. Using (26) the left-hand side of (31) can be written as

$$N^{-1/2} b(k)' \left[\hat{\Gamma}_{k,i}^{-1} \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k,k} - \Gamma_{k,i}^{-1} \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k} \right] = I_1 + I_2,$$

where

$$I_1 = N^{-1/2} b(k)' \left[\left(\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1} \right) \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k,k} \right],$$

$$I_2 = N^{-1/2} b(k)' \left[\Gamma_{k,i}^{-1} \sum_{j=0}^{N-1} X_j^{(i)}(k)(\varepsilon_{j\nu+i+k,k} - \varepsilon_{j\nu+i+k}) \right]$$

so that

$$|I_1| \leqslant N^{-1/2} \|b(k)\|_1 \cdot \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_1 \cdot \left\| \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k,k} \right\|_1 = J_1 \cdot J_2 \cdot J_3,$$

where $J_1 = \|b(k)\|_1$ is bounded by assumption, $J_2 = k\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_1$, and $J_3 = \frac{N^{1/2}}{k} \left\| \frac{1}{N} \sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{j\nu+i+k,k} \right\|_1$.

Next we will show that $J_2 = k\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_1 \to 0$ in probability. The proof is similar to Theorem 3.1 in Anderson et al. [5]. Define $p_{ki} = \|\Gamma_{k,i}^{-1}\|_1$, $q_{ki} = \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_1$, and $Q_{k,i} = \|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_1$. Then we have

$$q_{ki} \leqslant (q_{ki} + p_{ki}) Q_{ki} p_{ki} \tag{32}$$

exactly as in the proof of Theorem 3.1 in [5], and we want to show that $kq_{ki} \to 0$ in probability. From Theorem A.2 in [5] we have, for some $C > 0$, that

$$E\left| N a_N^{-2}(\hat{\gamma}_i(\ell) - \gamma_i(\ell)) \right| \leqslant C$$

uniformly over all $i = 0, \ldots, \nu - 1$, all integers ℓ and all positive integers N. Then, using the bound

$$E\|A\|_1 = E \max_{1 \leqslant j \leqslant k} \sum_{i=1}^{k} |a_{ij}| \leqslant k \max_{i,j} E|a_{ij}| \leqslant kC$$

in the proof of Theorem 3.5 in [5] it follows that $EQ_{ki} \leqslant ka_N^2 C/N$ for all i, k, and N. Since $p_{ki} \leqslant k^{1/2}\|\Gamma_{k,i}^{-1}\|_2$ we also have using Theorem A.1 from [5] that for some $m > 0$, $p_{ki} \leqslant k^{1/2}/(2\pi m)$ for all i and k. Then

$$E(p_{ki}Q_{ki}) = p_{ki}EQ_{ki} \leqslant \frac{k^{3/2}a_N^2 C}{2\pi N m} \to 0$$

as $k, N \to \infty$ since we are assuming $k^3 a_N^2/N \to 0$. Then it follows from the Markov inequality that $p_{ki}Q_{ki} \to 0$ in probability. If $p_{ki}Q_{ki} < 1$ then $1 - p_{ki}Q_{ki} > 0$, and then it follows easily from (32) that

$$kq_{ki} \leqslant \frac{kp_{ki}^2 Q_{ki}}{1 - p_{ki}Q_{ki}}.$$

Since $p_{ki} \leqslant k^{1/2}/2\pi m$, we have

$$E\left(kp_{ki}^2 Q_{ki}\right) \leqslant k \cdot \frac{k^2 a_N^2 C}{(2\pi m)^2 N} \to 0$$

so that $kp_{ki}^2 Q_{ki} \to 0$ in probability. As $p_{ki}Q_{ki} \to 0$ in probability, it follows that

$$\frac{kp_{ki}^2 Q_{ki}}{1 - p_{ki}Q_{ki}} \xrightarrow{P} 0.$$

Now the remainder of the proof that $J_2 \to 0$ in probability is exactly the same conditioning argument as in Theorem 3.1 of [5]. As for the remaining term in I_1, write

$$J_3 = \frac{N^{1/2}}{k}\left\|\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{jv+i+k,k}\right\|_1$$

so that $E(J_3) \leqslant E(J_{31}) + E(J_{32})$ where

$$J_{31} = \frac{N^{1/2}}{k}\left\|\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{jv+i+k}\right\|_1 \quad \text{and}$$

$$J_{32} = \frac{N^{1/2}}{k}\left\|\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)(\varepsilon_{jv+i+k,k} - \varepsilon_{jv+i+k})\right\|_1.$$

Lemma 4.4 implies that

$$E(J_{32}) \leqslant \frac{N^{1/2}}{k} Ak \max_{\ell} \sum_{s=1}^{\infty} |\pi_\ell(k+s)| = N^{1/2} A \max_{\ell} \sum_{s=1}^{\infty} |\pi_\ell(k+s)| \to 0,$$

where the maximum is taken over $\ell = 0, \ldots, v-1$. Also

$$\begin{aligned} E(J_{31}) &\leqslant \frac{N^{1/2}}{k} k^{1/2} E\left\|\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{jv+i+k}\right\|_2 \\ &\leqslant \left(\frac{N}{k}\right)^{1/2} \sqrt{E\left\|\frac{1}{N}\sum_{j=0}^{N-1} X_j^{(i)}(k)\varepsilon_{jv+i+k}\right\|_2^2} \end{aligned}$$

$$= \left(\frac{N}{k}\right)^{1/2} \sqrt{\frac{1}{N^2} \sum_{t=0}^{k-1} E\left[\left(\sum_{j=0}^{N-1} X_{jv+i+t}\varepsilon_{jv+i+k}\right)^2\right]}$$

$$= \sqrt{\frac{1}{kN} \sum_{t=0}^{k-1} \sum_{j=0}^{N-1} E(X^2_{jv+i+t}) E(\varepsilon^2_{jv+i+k})}$$

$$= \sqrt{k^{-1} \sum_{t=0}^{k-1} \gamma_{i+t}(0) \cdot N^{-1} \sum_{j=0}^{N-1} \sigma^2_{i+k}}$$

so that $E(J_{31}) \leqslant \sqrt{D}$ where $D = \max_i \gamma_i(0) \cdot \max_i \sigma_i^2$. Thus $E(J_3) < \infty$, J_1 is bounded, and $J_2 \to 0$ in probability. Then it is easy to show that $I_1 \to 0$ in probability.

Next write

$$I_2 = N^{-1/2} \sum_{j=0}^{N-1} u_{jv+i+k-1}(\varepsilon_{jv+i+k,k} - \varepsilon_{jv+i+k}),$$

where $u_{jv+i+k-1} = b(k)'\Gamma_{k,i}^{-1} X_j^{(i)}(k)$. Then

$$\begin{aligned}\operatorname{Var}(u_{jv+i+k-1}) = E(u^2_{jv+i+k-1}) &= b(k)'\Gamma_{k,i}^{-1} E[X_j^{(i)}(k) X_j^{(i)}(k)']\Gamma_{k,i}^{-1} b(k) \\ &= b(k)'\Gamma_{k,i}^{-1} b(k) \leqslant \|b(k)\|_2^2/(2\pi m)\end{aligned}$$

by Theorem A.1 in [5], since $\Gamma_{k,i}$ is the covariance matrix of $X_j^{(i)}(k)$. Hence by Lemma 4.3,

$$E|I_2| = \text{Constant} \cdot N^{1/2} \max_{\ell} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \to 0$$

for $\ell = 0, \ldots, v-1$ by (14). Then $I_2 \to 0$ in probability, which finishes the proof of Lemma 4.5. □

Proof of Theorem 3.1. Define $e_u(k)$ to be the k dimensional vector with 1 in the uth place and zeros elsewhere. Let

$$w_{uj,k}^{(\langle i-k\rangle)} = e_u(k)'\Gamma_{k,\langle i-k\rangle}^{-1} X_j^{(\langle i-k\rangle)}(k)\varepsilon_{jv+\langle i-k\rangle+k} \tag{33}$$

so that

$$t_{N,k}^{(\langle i-k\rangle)}(u) = N^{-1/2} \sum_{j=0}^{N-1} w_{uj,k}^{(\langle i-k\rangle)}. \tag{34}$$

Here $X_j^{(\langle i-k\rangle)}(k) = [X_{jv+\langle i-k\rangle+k-1}, \ldots, X_{jv+\langle i-k\rangle}]'$ and

$$\Gamma_{k,\langle i-k\rangle} = E\{X_j^{(\langle i-k\rangle)}(k) X_j^{(\langle i-k\rangle)}(k)'\} \tag{35}$$

thus for i, u, k fixed, $w_{uj,k}^{(\langle i-k\rangle)}$ are stationary martingale differences since the first two terms in (33) are non-random and do not depend on j while the third term is in the linear span of ε_s,

$s < jv + \langle i-k \rangle + k$ due to the causality assumption. Then $E\{w_{uj,k}^{(\langle i-k \rangle)} w_{vj',k}^{(\langle i'-k \rangle)}\} = 0$ unless

$$jv + \langle i-k \rangle = j' + \langle i'-k \rangle$$

because, if $jv + \langle i-k \rangle > j' + \langle i'-k \rangle$, $X_j^{(\langle i-k \rangle)}(k)$, $X_{j'}^{(\langle i'-k \rangle)}(k)$ and $\varepsilon_{j'+\langle i'-k \rangle}$ are in the linear span of ε_s for $s < jv + \langle i-k \rangle$. Otherwise $X_j^{(\langle i-k \rangle)}(k) = X_{j'}^{(\langle i'-k \rangle)}(k)$ and

$$E\{w_{uj,k}^{(\langle i-k \rangle)} w_{vj,k}^{(\langle i-k \rangle)}\} = \sigma_i^2 e_u(k)' \Gamma_{k,\langle i-k \rangle}^{-1} e_v(k).$$

Take

$$w_{jk} = [w_{1j,k}^{(\langle 0-k \rangle)}, \ldots, w_{Dj,k}^{(\langle 0-k \rangle)}, \ldots, w_{1j,k}^{(\langle v-1-k \rangle)}, \ldots, w_{Dj,k}^{(\langle v-1-k \rangle)}]'.$$

It follows immediately that the covariance matrix of the vector

$$t_{N,k} = N^{-1/2} \sum_{j=0}^{N-1} w_{jk}$$

is $\Lambda_k = \operatorname{diag}\{\sigma_0^2 \Lambda_k^{(0)}, \sigma_1^2 \Lambda_k^{(1)}, \ldots, \sigma_{v-1}^2 \Lambda_k^{(v-1)}\}$ where

$$(\Lambda_k^{(i)})_{u,v} = e_u(k)' \Gamma_{k,\langle i-k \rangle}^{-1} e_v(k), \tag{36}$$

and $1 \leqslant u, v \leqslant D$. Apply Lemma 1 of [4] to see that

$$e_u(k)' \Gamma_{k,\langle i-k \rangle}^{-1} e_v(k) \longrightarrow (\Lambda^{(i)})_{u,v}$$

as $k \to \infty$ where, taking $m = \min\{u, v\}$,

$$(\Lambda^{(i)})_{u,v} = \sum_{s=0}^{m-1} \pi_{i-m+s}(s) \pi_{i-m+s}(s + |v-u|) \sigma_{i-m+s}^{-2}.$$

Then, provided that $k = k(N)$ increases to ∞ with N,

$$\lim_{N \to \infty} \operatorname{Var}\{t_{N,k(N)}\} = \Lambda, \tag{37}$$

where Λ is given in (16).

Next we want to use the martingale central limit theorem (Theorem 3.2, p.58 in Hall and Heyde [18]) to show that

$$\lambda' t_{N,k} \Rightarrow \mathcal{N}(0, \lambda' \Lambda_k \lambda)$$

for a fixed k and any $\lambda \in \mathbb{R}^{Dv}$. Consider the triangular array of summands $X_N(j) = N^{-1/2} \lambda' w_{jk}$, $j = 0, \ldots, N-1$, $N = 1, 2, \ldots$. For each fixed k, it is sufficient to show that

(i) $\max_{0 \leqslant j < N} X_N(j)^2 \xrightarrow{P} 0$,
(ii) $\sum_{j=0}^{N-1} X_N(j)^2 \xrightarrow{P} V$ for $V > 0$, and
(iii) $\sup_N E\{\max_{0 \leqslant j < N} X_N(j)^2\} \leqslant c < \infty$.

As w_{jk} is stationary with a finite second moment, $w_{jk} = o(j^{1/2})$ and therefore $X_j(j)^2 = o(1)$, where both expressions hold almost surely [27, Lemma 7.5.1]. Now

$$\max_{0 \leqslant j < n} X_N(j)^2 \leqslant \max_{0 \leqslant j < n} X_j(j)^2$$

so that for any $\varepsilon > 0$ and $n \leqslant N$, the Markov inequality obtains

$$P\left\{\max_{0 \leqslant j < N} X_N(j)^2 > \epsilon\right\} \leqslant \sum_{j=0}^{n} P\{X_N(j)^2 > \epsilon\} + P\left\{\max_{n < j < N} X_j(j)^2 > \epsilon\right\}$$
$$\leqslant \frac{n}{N} \frac{E\{(\lambda' w_{jk})^2\}}{\epsilon} + \sum_{j=n}^{\infty} P\{X_j(j)^2 > \epsilon\}.$$

Upon taking n an increasing function of N so that $n/N \to 0$, and noting that $\sum_{j=1}^{\infty} P\{X_j(j)^2 > c\} < \infty$, the RHS converges to zero, which establishes (i). Moreover

$$E\left\{\max_{0 \leqslant j < N} X_N(j)^2\right\} \leqslant \sum_{j=0}^{N-1} E\{X_N(j)^2\} = \lambda' E\{w_{jk} w'_{jk}\}\lambda < \infty$$

so that (iii) also holds. To establish (ii) note that w_{jk} is ergodic (see the discussion in [15, p. 458]) and consequently

$$\sum_{j=0}^{N-1} X_N(j)^2 \longrightarrow \lambda' E\{w_{jk} w'_{jk}\}\lambda = \lambda' \Lambda_k \lambda,$$

where the RHS is the positive quantity V in (ii). Thus the conditions of Hall and Heyde [18], Theorem 3.2 are satisfied and therefore, for each λ and fixed k, $\lambda' t_{N,k}$ converges to a normal distribution with zero mean and variance $\lambda' \Lambda_k \lambda$. Then an application of the Cramér–Wold device [12, p. 48] yields

$$t_{N,k} \Rightarrow \mathcal{N}(0, \Lambda_k).$$

To extend the central limit theorem to the case where $k = k(N) \to \infty$ as $N \to \infty$ we use a result due to Bernstein [9] that we refer to as Bernstein's Lemma, which is proved in Hannan [19, p. 242]. *Let x_N be a sequence of vector valued random variables with zero mean such that for every $\epsilon > 0$, $\zeta > 0$, $\eta > 0$ there exist sequences of random vectors $y_N(\epsilon)$, $z_N(\epsilon)$ so that $x_N = y_N(\epsilon) + z_N(\epsilon)$ where $y_N(\epsilon)$ has a distribution converging to the multivariate normal distribution with zero mean and covariance matrix $V(\epsilon)$, and*

$$\lim_{\epsilon \to 0} V(\epsilon) = V, \quad P\{z_N(\epsilon)' z_N(\epsilon) > \zeta\} < \eta.$$

Then the distribution of x_N converges to the multivariate normal distribution with covariance matrix V.

Using the notation of Bernstein's Lemma, take $x_N = t_{N,k(N)}$ where now $k = k(N)$ is explicitly written as a function of N. For any $\epsilon > 0$ take $k = [\epsilon^{-1}]$ and $y_N(\epsilon) = t_{N,k}$. We have shown that $y_N(\epsilon)$ converges to a multivariate normal distribution with zero mean and variance $\mathrm{Var}\{t_{Nk}\} = \Lambda_k = V(\epsilon)$ and, in (37), that $\Lambda_k \to \Lambda$ as $k \to \infty$ thus verifying the first condition of Bernstein's

Lemma. Consider

$$z_N(\epsilon) = x_N - y_N(\epsilon) = t_{N,k(N)} - t_{N,k}.$$

The second condition of Bernstein's Lemma follows if $E\{z_N(\epsilon)' z_N(\epsilon)\} \to 0$ as $k, N \to \infty$ which holds if the variance of each component of $z_N(\epsilon)$ converges to zero.

For a given k, take N sufficiently large so that $k(N) > k$. The components of $z_N(\epsilon)$ are $t_{N,k(N)}^{(\langle i-k(N)\rangle)}(u) - t_{N,k}^{(\langle i-k\rangle)}(u)$, $i = 0, \ldots, v$, $u = 1, \ldots, D$ which have variance

$$\mathrm{Var}\{t_{N,k(N)}^{(\langle i-k(N)\rangle)}(u)\} + \mathrm{Var}\{t_{N,k}^{(\langle i-k\rangle)}(u)\} - 2\mathrm{Cov}\{t_{N,k(N)}^{(\langle i-k(N)\rangle)}(u), t_{N,k}^{(\langle i-k\rangle)}(u)\}.$$

From (36) the first two terms are $\sigma_i^2(\Lambda_{k(N)}^{(i)})_{uu}$ and $\sigma_i^2(\Lambda_k^{(i)})_{uu}$, respectively. The covariance term is the expectation of

$$N^{-1} \sum_{j=0}^{N-1} \sum_{j'=0}^{N-1} e_u(k(N))' \Gamma_{k(N),\langle i-k(N)\rangle}^{-1} X_j^{(\langle i-k(N)\rangle)}(k(N)) \varepsilon_{jv+\langle i-k(N)\rangle+k(N)}$$
$$\times \varepsilon_{j'v+\langle i-k\rangle+k} (X_{j'}^{(\langle i-k\rangle)}(k))' \Gamma_{k,\langle i-k\rangle}^{-1} e_u(k). \tag{38}$$

The only summands with non-zero expectation occur when $jv + \langle i - k(N)\rangle + k(N) = j'v + \langle i-k\rangle + k = s$, say, and then

$$X_j^{(\langle i-k(N)\rangle)}(k(N)) = [X_{s-1}, \ldots, X_{s-k(N)}]',$$
$$X_{j'}^{(\langle i-k\rangle)}(k) = [X_{s-1}, \ldots, X_{s-k}]'$$

so that $E\{X_j^{(\langle i-k(N)\rangle)}(k(N))(X_{j'}^{(\langle i-k\rangle)}(k))'\}$ is the matrix consisting of the first k columns of $\Gamma_{k(N),\langle i-k(N)\rangle}$. Thus the mean of those summands with non-zero expectation in (38) may be evaluated (compare the argument for Lemma 4.6 later in this paper) as $\sigma_i^2(\Lambda_k^{(i)})_{uu}$ and hence

$$\mathrm{Var}\{t_{N,k(N)}^{(\langle i-k(N)\rangle)}(u) - t_{N,k}^{(\langle i-k\rangle)}(u)\} = \sigma_i^2 \left[(\Lambda_{k(N)}^{(i)})_{uu} - N^{-1} C_N (\Lambda_k^{(i)})_{uu}\right],$$

where C_N is the number of those summands. Since $jv + \langle i-k(N)\rangle + k(N) = j'v + \langle i-k\rangle + k$ can occur for at most one value of j' for each $j = 0, \ldots, N-1$ it is not hard to check that $N^{-1}C_N \to 1$ as $N \to \infty$. Lemma 1 from Anderson et al. [4] shows that the (u, u)th elements of $\Lambda_{i,\langle i-k(N)\rangle}^{-1}$ and $\Lambda_{i,\langle i-k\rangle}^{-1}$ converge to

$$\sigma_{i+s}^2 \sum_{s=0}^{u} \pi_{i-j+u}(u)^2$$

as $k, k(N) \to \infty$ and therefore $E\{z_N(\epsilon)' z_N(\epsilon)\} \to 0$. Thus for any $\varepsilon > 0$, $\zeta > 0$ and $\eta > 0$ we may take k, N sufficiently large, $k > [\epsilon^{-1}]$ so that $E\{z_N(\epsilon)' z_N(\epsilon)\} < \eta\zeta$ and an application of Markov's inequality establishes the second condition of Bernstein's Lemma. Thus we conclude

$$t_{N,k(N)} \Rightarrow \mathcal{N}(0, \Lambda), \tag{39}$$

where $\lim_{k\to\infty} \Lambda_k = \Lambda$.

Applying Lemma 4.5 yields $N^{1/2} e_u(k)'(\pi_k^{(\langle i-k\rangle)} + \hat{\phi}_k^{(\langle i-k\rangle)}) - t_{N,k}^{(\langle i-k\rangle)}(u) \to 0$ in probability. Note that $\pi_k^{(\langle i-k\rangle)} = (\pi_i(1), \ldots, \pi_i(k))'$ and $\hat{\phi}_k^{(\langle i-k\rangle)} = (\hat{\phi}_{k,1}^{(\langle i-k\rangle)}, \ldots, \hat{\phi}_{k,k}^{(\langle i-k\rangle)})$. Combining (39) with $e_u(k)'(\pi_k^{(\langle i-k\rangle)} + \hat{\phi}_k^{(\langle i-k\rangle)}) = \pi_i(u) + \hat{\phi}_{k,u}^{(\langle i-k\rangle)}$ implies that

$$N^{1/2}(\pi_i(u) + \hat{\phi}_{k,u}^{(\langle i-k\rangle)}) - t_{N,k}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0 \tag{40}$$

as $N \to \infty$. Then Theorem 3.1 follows from (39) and (40) and the continuous mapping theorem. □

Proof of Theorem 3.2. From the two representations of $\hat{X}_{i+k}^{(i)}$ given by (4) and (8) it follows that

$$\theta_{k,j}^{(\langle i-k\rangle)} = \sum_{\ell=1}^{j} \phi_{k,\ell}^{(\langle i-k\rangle)} \theta_{k-\ell,j-\ell}^{(\langle i-k\rangle)} \tag{41}$$

for $j = 1, \ldots, k$ if we define $\theta_{k-j,0}^{(\langle i-k\rangle)} = 1$ and replace i with $\langle i-k\rangle$. Eq. (41) can be modified and written as

$$\begin{pmatrix} \theta_{k,1}^{(\langle i-k\rangle)} \\ \theta_{k,2}^{(\langle i-k\rangle)} \\ \theta_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \theta_{k,D}^{(\langle i-k\rangle)} \end{pmatrix} = R_k^{(\langle i-k\rangle)} \begin{pmatrix} \phi_{k,1}^{(\langle i-k\rangle)} \\ \phi_{k,2}^{(\langle i-k\rangle)} \\ \phi_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \phi_{k,D}^{(\langle i-k\rangle)} \end{pmatrix}, \tag{42}$$

where

$$R_k^{(\langle i-k\rangle)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \theta_{k-1,1}^{(\langle i-k\rangle)} & 1 & 0 & \cdots & 0 & 0 \\ \theta_{k-1,2}^{(\langle i-k\rangle)} & \theta_{k-2,1}^{(\langle i-k\rangle)} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ \theta_{k-1,D-1}^{(\langle i-k\rangle)} & \theta_{k-2,D-2}^{(\langle i-k\rangle)} & \cdots & & \theta_{k-D+1,1}^{(\langle i-k\rangle)} & 1 \end{pmatrix} \tag{43}$$

for fixed lag D. From the definitions of $\hat{\theta}_{k,u}^{(i)}$ and $\hat{\phi}_{k,u}^{(i)}$ we also have

$$\begin{pmatrix} \hat{\theta}_{k,1}^{(\langle i-k\rangle)} \\ \hat{\theta}_{k,2}^{(\langle i-k\rangle)} \\ \hat{\theta}_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \hat{\theta}_{k,D}^{(\langle i-k\rangle)} \end{pmatrix} = \hat{R}_k^{(\langle i-k\rangle)} \begin{pmatrix} \hat{\phi}_{k,1}^{(\langle i-k\rangle)} \\ \hat{\phi}_{k,2}^{(\langle i-k\rangle)} \\ \hat{\phi}_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \hat{\phi}_{k,D}^{(\langle i-k\rangle)} \end{pmatrix}, \tag{44}$$

where $\hat{R}_k^{(\langle i-k\rangle)}$ is defined as in (43) with $\hat{\theta}_{k,u}^{(\langle i-k\rangle)}$ replacing $\theta_{k,u}^{(\langle i-k\rangle)}$. From (12) we know that $\hat{\theta}_{k,u}^{(\langle i-k\rangle)} \xrightarrow{P} \psi_i(u)$, hence for fixed ℓ with $k' = k - \ell$, we have $\hat{\theta}_{k-\ell,u}^{(\langle i-k\rangle)} = \hat{\theta}_{k',u}^{(\langle i-\ell-k'\rangle)} \xrightarrow{P} \psi_{i-\ell}(u)$. Thus,

$$\hat{R}_k^{(\langle i-k\rangle)} \xrightarrow{P} R^{(i)}, \tag{45}$$

where

$$R^{(i)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \psi_{i-1}(1) & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \psi_{i-1}(D-1) & \psi_{i-2}(D-2) & \cdots & \psi_{i-D+1}(1) & 1 \end{pmatrix}. \tag{46}$$

We have

$$\hat{\theta}^{(\langle i-k \rangle)} - \theta^{(\langle i-k \rangle)} = \hat{R}_k^{(\langle i-k \rangle)}(\hat{\phi}^{(\langle i-k \rangle)} - \phi^{(\langle i-k \rangle)}) + (\hat{R}_k^{(\langle i-k \rangle)} - R_k^{(\langle i-k \rangle)})\phi^{(\langle i-k \rangle)}, \tag{47}$$

where $\theta^{(\langle i-k \rangle)} = (\theta_{k,1}^{(\langle i-k \rangle)}, \ldots, \theta_{k,D}^{(\langle i-k \rangle)})'$, $\phi^{(\langle i-k \rangle)} = (\phi_{k,1}^{(\langle i-k \rangle)}, \ldots, \phi_{k,D}^{(\langle i-k \rangle)})'$, and $\hat{\theta}^{(\langle i-k \rangle)}$ and $\hat{\phi}^{(\langle i-k \rangle)}$ are the respective estimators of $\theta^{(\langle i-k \rangle)}$ and $\phi^{(\langle i-k \rangle)}$. Note that

$$\begin{aligned} (\hat{R}_k^{(\langle i-k \rangle)} - R_k^{(\langle i-k \rangle)})\phi^{(\langle i-k \rangle)} &= (\hat{R}_k^{(\langle i-k \rangle)} - \hat{R}_k^{(\langle i-k \rangle)*})\phi^{(\langle i-k \rangle)} \\ &\quad + (\hat{R}_k^{(\langle i-k \rangle)*} - R_k^{(\langle i-k \rangle)*})\phi^{(\langle i-k \rangle)} \\ &\quad + (R_k^{(\langle i-k \rangle)*} - R_k^{(\langle i-k \rangle)})\phi^{(\langle i-k \rangle)}, \end{aligned} \tag{48}$$

where

$$R_k^{(\langle i-k \rangle)*} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \theta_{k,1}^{(\langle i-1-k \rangle)} & 1 & 0 & \cdots & 0 & 0 \\ \theta_{k,2}^{(\langle i-1-k \rangle)} & \theta_{k,1}^{(\langle i-2-k \rangle)} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ \theta_{k,D-1}^{(\langle i-1-k \rangle)} & \theta_{k,D-2}^{(\langle i-2-k \rangle)} & \cdots & & \theta_{k,1}^{(\langle i-D+1-k \rangle)} & 1 \end{pmatrix} \tag{49}$$

and $\hat{R}_k^{(\langle i-k \rangle)*}$ is the corresponding matrix obtained by replacing $\theta_{k,u}^{(\langle i-k \rangle)}$ with $\hat{\theta}_{k,u}^{(\langle i-k \rangle)}$ for every season i and lag u. We next need to show that $R_k^{(\langle i-k \rangle)*} - R_k^{(\langle i-k \rangle)} = o(N^{1/2})$ and $\hat{R}_k^{(\langle i-k \rangle)} - \hat{R}_k^{(\langle i-k \rangle)*} = o_P(N^{1/2})$. This is equivalent to showing that

$$N^{1/2}(\theta_{k,u}^{(\langle i-\ell-k \rangle)} - \theta_{k-\ell,u}^{(\langle i-k \rangle)}) \to 0 \tag{50}$$

and

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-\ell-k \rangle)} - \hat{\theta}_{k-\ell,u}^{(\langle i-k \rangle)}) \xrightarrow{P} 0 \tag{51}$$

for $\ell = 1, \ldots, D-1$ and $u = 1, \ldots, D$. Using estimates from the proof of [5] Corollary 2.2.4 and condition (14) of Theorem 3.1, it is not hard to show that $N^{1/2}(\phi_{k,u}^{(\langle i-k \rangle)} + \pi_i(u)) \to 0$ as $N \to \infty$ for any $u = 1, \ldots, k$. This leads to

$$N^{1/2}(\phi_{k,u}^{(\langle i-\ell-k \rangle)} + \pi_{i-\ell}(u)) \to 0 \tag{52}$$

by replacing i with $i - \ell$ for fixed ℓ. Letting $a_k = N^{1/2}(\phi_{k,u}^{(\langle i-\ell-k \rangle)} + \pi_{i-\ell}(u))$ and $b_k = N^{1/2}(\phi_{k-\ell,u}^{(\langle i-k \rangle)} + \pi_{i-\ell}(u))$ we see that $b_k = a_{k-\ell}$. Since $a_k \to 0$ then $b_k \to 0$ as $k \to \infty$. Hence,

$$N^{1/2}(\phi_{k-\ell,u}^{(\langle i-k \rangle)} + \pi_{i-\ell}(u)) \to 0 \tag{53}$$

as $k \to \infty$. Subtracting (53) from (52) yields

$$N^{1/2}(\phi_{k,u}^{(\langle i-\ell-k\rangle)} - \phi_{k-\ell,u}^{(\langle i-k\rangle)}) \to 0 \tag{54}$$

which holds for $\ell = 1, \ldots, u-1$ and $u = 1, \ldots, k$. Since

$$\theta_{k,1}^{(\langle i-\ell-k\rangle)} - \theta_{k-\ell,1}^{(\langle i-k\rangle)} = \phi_{k,1}^{(\langle i-\ell-k\rangle)} - \phi_{k-\ell,1}^{(\langle i-k\rangle)}$$

we have (50) with $u = 1$. The cases $u = 2, \ldots, D$ follow iteratively using (10), (42), and (54). Thus, (50) is established. To prove (51), we need the following lemma. □

Lemma 4.6. *For all $\ell = 1, \ldots, u-1$ and $u = 1, \ldots, k$ we have*

$$N^{1/2}(\hat{\phi}_{k,u}^{(\langle i-\ell-k\rangle)} - \hat{\phi}_{k-\ell,u}^{(\langle i-k\rangle)}) \xrightarrow{P} 0. \tag{55}$$

Proof. Starting from (34), we need to show that $t_{N,k}^{(\langle i-\ell-k\rangle)}(u) - t_{N,k-\ell}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0$. Note that $E(w_{uj,k}^{(\langle i-\ell-k\rangle)} w_{uj',k-\ell}^{(\langle i-k\rangle)}) = 0$ unless

$$j' = j - \ell + (\langle i-\ell-k\rangle + \ell - \langle i-k\rangle)/v \tag{56}$$

which is always an integer. For each j there is at most one j' that satisfies (56) for $j' \in \{k-\ell, \ldots, N-1\}$. If such a j' exists then

$$E(w_{uj,k}^{(\langle i-\ell-k\rangle)} w_{uj',k-\ell}^{(\langle i-k\rangle)}) = \sigma_{i-\ell}^2 e_u(k)'(\Gamma_{k,\langle i-\ell-k\rangle})^{-1} C (\Gamma_{k-\ell,\langle i-k\rangle})^{-1} e_u(k-\ell),$$

where $C = E(X_j^{(\langle i-\ell-k\rangle)}(k) X_{j'}^{(\langle i-k\rangle)}(k-\ell)')$. Note that the $(k-\ell)$-dimensional vector $X_{j'}^{(\langle i-k\rangle)}(k-\ell)$ is just the first $(k-\ell)$ of the k entries of the vector $X_j^{(\langle i-\ell-k\rangle)}(k)$. Hence, the matrix C is just $\Gamma_{k,\langle i-\ell-k\rangle}$ with the last ℓ columns deleted. Then $(\Gamma_{k,\langle i-\ell-k\rangle})^{-1} C = I(k,\ell)$ which is the $k \times k$ identity matrix with the last ℓ columns deleted. But then for any fixed u, for all k large we have $e_u(k)'(\Gamma_{k,\langle i-\ell-k\rangle})^{-1} C = e_u(k-\ell)'$. Then

$$\begin{aligned} E(w_{uj,k}^{(\langle i-\ell-k\rangle)} w_{uj',k-\ell}^{(\langle i-k\rangle)}) &= \sigma_{i-\ell}^2 e_u(k-\ell)'(\Gamma_{k-\ell,\langle i-k\rangle})^{-1} e_u(k-\ell) \\ &= \sigma_{i-\ell}^2 (\Gamma_{k-\ell,\langle i-k\rangle})_{uu}^{-1}. \end{aligned}$$

Consequently,

$$\begin{aligned} \operatorname{Var}(w_{uj,k}^{(\langle i-\ell-k\rangle)} - w_{uj',k-\ell}^{(\langle i-k\rangle)}) &= E[(w_{uj,k}^{(\langle i-\ell-k\rangle)} - w_{uj',k-\ell}^{(\langle i-k\rangle)})^2] \\ &= \sigma_{i-\ell}^2 (\Gamma_{k,\langle i-\ell-k\rangle})_{uu}^{-1} + \sigma_{i-\ell}^2 (\Gamma_{k-\ell,\langle i-k\rangle})_{uu}^{-1} \\ &\quad -2\sigma_{i-\ell}^2 (\Gamma_{k-\ell,\langle i-k\rangle})_{uu}^{-1} \\ &= \sigma_{i-\ell}^2 [(\Gamma_{k,\langle i-\ell-k\rangle})_{uu}^{-1} - (\Gamma_{k-\ell,\langle i-k\rangle})_{uu}^{-1}] \\ &\to 0 \end{aligned}$$

by Lemma 1 of [4]. Thus,

$$\begin{aligned} \operatorname{Var}(t_{N,k}^{(\langle i-\ell-k\rangle)}(u) - t_{N,k-\ell}^{(\langle i-k\rangle)}(u)) &\leqslant N^{-1} \operatorname{Var}(w_{uj,k}^{(\langle i-\ell-k\rangle)} - w_{uj',k-\ell}^{(\langle i-k\rangle)}) N \\ &\to 0 \end{aligned}$$

since for each $j = k, \ldots, N-1$ there is at most one j' satisfying (56) along with $j' = k-\ell, \ldots, N-1$. Then Chebyshev's inequality shows that $t_{N,k}^{(\langle i-\ell-k \rangle)}(u) - t_{N,k-\ell}^{(\langle i-k \rangle)}(u) \xrightarrow{P} 0$. Now (40) yields

$$N^{1/2}(\pi_{i-\ell}(u) + \hat{\phi}_{k,u}^{(\langle i-\ell-k \rangle)}) - t_{N,k}^{(\langle i-\ell-k \rangle)}(u) \xrightarrow{P} 0$$

and another application of (40) using $i - \ell - (k-\ell) = i - k$ gives

$$N^{1/2}(\pi_{i-\ell}(u) + \hat{\phi}_{k-\ell,u}^{(\langle i-k \rangle)}) - t_{N,k-\ell}^{(\langle i-k \rangle)}(u) \xrightarrow{P} 0$$

and then Lemma 4.6 follows easily.

Now, since

$$\hat{\theta}_{k,1}^{(\langle i-\ell-k \rangle)} - \hat{\theta}_{k-\ell,1}^{(\langle i-k \rangle)} = \hat{\phi}_{k,1}^{(\langle i-\ell-k \rangle)} - \hat{\phi}_{k-\ell,1}^{(\langle i-k \rangle)}$$

we have (51) with $u = 1$. The cases $u = 2, \ldots, D$ follow iteratively using (44), Lemma 4.6, and (41) with θ, ϕ replaced by $\hat{\theta}, \hat{\phi}$. Thus, (51) is established. From (47), (48), (50), and (51) it follows that

$$\begin{aligned}\hat{\theta}^{(\langle i-k \rangle)} - \theta^{(\langle i-k \rangle)} = \hat{R}_k^{(\langle i-k \rangle)}(\hat{\phi}^{(\langle i-k \rangle)} - \phi^{(\langle i-k \rangle)}) + (\hat{R}_k^{(\langle i-k \rangle)*} - R_k^{(\langle i-k \rangle)*})\phi^{(\langle i-k \rangle)} \\ + o_P(N^{1/2}).\end{aligned} \tag{57}$$

To accommodate the derivation of the asymptotic distribution of $\hat{\theta}^{(i)} - \theta^{(i)}$, we need to rewrite (57). Define

$$\begin{aligned}\hat{\theta} - \theta = (&\hat{\theta}_{k,1}^{(\langle 0-k \rangle)} - \theta_{k,1}^{(\langle 0-k \rangle)}, \ldots, \hat{\theta}_{k,D}^{(\langle 0-k \rangle)} - \theta_{k,D}^{(\langle 0-k \rangle)}, \ldots, \\ &\hat{\theta}_{k,1}^{(\langle \nu-1-k \rangle)} - \theta_{k,1}^{(\langle \nu-1-k \rangle)}, \ldots, \hat{\theta}_{k,D}^{(\langle \nu-1-k \rangle)} - \theta_{k,D}^{(\langle \nu-1-k \rangle)})'\end{aligned}$$

and

$$\phi = (\phi_{k,1}^{(\langle 0-k \rangle)}, \ldots, \phi_{k,D}^{(\langle 0-k \rangle)}, \ldots, \phi_{k,1}^{(\langle \nu-1-k \rangle)}, \ldots, \phi_{k,D}^{(\langle \nu-1-k \rangle)})'. \tag{58}$$

Using (58) we can rewrite (57) as

$$\hat{\theta} - \theta = \hat{R}_k(\hat{\phi} - \phi) + (\hat{R}_k^* - R_k^*)\phi + o_P(N^{1/2}), \tag{59}$$

where $\hat{\phi}$ is the estimator of ϕ and

$$R_k = \operatorname{diag}(R_k^{(\langle 0-k \rangle)}, R_k^{(\langle 1-k \rangle)}, \ldots, R_k^{(\langle \nu-1-k \rangle)})$$

and

$$R_k^* = \operatorname{diag}(R_k^{(\langle 0-k \rangle)*}, R_k^{(\langle 1-k \rangle)*}, \ldots, R_k^{(\langle \nu-1-k \rangle)*})$$

noting that both R_k and R_k^* are $D\nu \times D\nu$ matrices. The estimators of R_k and R_k^* are, respectively, $\hat{R}_k$ and $\hat{R}_k^*$. Now write $(\hat{R}_k^* - R_k^*)\phi = C_k(\hat{\theta} - \theta)$ where

$$C_k = \sum_{n=1}^{D-1} B_{n,k} \Pi^{[D\nu - n(D+1)]} \tag{60}$$

and

$$B_{n,k} = \text{diag}\Bigg(\overbrace{0,\ldots,0}^{n}, \overbrace{\phi_{k,n}^{(\langle 0-k\rangle)},\ldots,\phi_{k,n}^{(\langle 0-k\rangle)}}^{D-n}, \overbrace{0,\ldots,0}^{n},$$

$$\underbrace{\phi_{k,n}^{(\langle 1-k\rangle)},\ldots,\phi_{k,n}^{(\langle 1-k\rangle)}}_{D-n},\ldots,\underbrace{0,\ldots,0}_{n}, \underbrace{\phi_{k,n}^{(\langle \nu-1-k\rangle)},\ldots,\phi_{k,n}^{(\langle \nu-1-k\rangle)}}_{D-n}\Bigg)$$

with Π the orthogonal $D\nu \times D\nu$ cyclic permutation matrix (24). Thus, we write Eq. (59) as

$$\hat{\theta} - \theta = \hat{R}_k(\hat{\phi} - \phi) + C_k(\hat{\theta} - \theta) + o_P(N^{1/2}). \tag{61}$$

Then,

$$(I - C_k)(\hat{\theta} - \theta) = \hat{R}_k(\hat{\phi} - \phi) + o_P(N^{1/2})$$

so that

$$\hat{\theta} - \theta = (I - C_k)^{-1}\hat{R}_k(\hat{\phi} - \phi) + o_P(N^{1/2}). \tag{62}$$

Let $C = \lim_{k\to\infty} C_k$ so that C is C_k with $\phi_{k,u}^{(\langle i-k\rangle)}$ replaced with $-\pi_i(u)$. Also, let $R = \lim_{k\to\infty} R_k$ where

$$R = \text{diag}(R^{(0)}, \ldots, R^{(\nu-1)})$$

and $R^{(i)}$ as defined in (46). Eq. (45) shows that $\hat{R}_k \xrightarrow{P} R$, and then Theorem 3.1 along with Eq. (62) yield

$$N^{1/2}(\hat{\theta} - \theta) \Rightarrow N(0, V)$$

where

$$V = (I - C)^{-1} R \Lambda R' [(I - C)^{-1}]' \tag{63}$$

and Λ is as in (16). Let

$$S^{(i)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \pi_{i-1}(1) & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \pi_{i-1}(D-1) & \pi_{i-2}(D-2) & \cdots & \pi_{i-D+1}(1) & 1 \end{pmatrix}. \tag{64}$$

It can be shown that

$$\Lambda^{(i)} = S^{(i)} \text{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2}) S^{(i)'}.$$

From the equation, $\psi_i(u) = \sum_{\ell=1}^{u} -\pi_i(\ell)\psi_{i-\ell}(u-\ell)$, it follows that $R^{(i)}S^{(i)} = I_{D\times D}$, the $D \times D$ identity matrix. Therefore,

$$\begin{aligned} R^{(i)}\Lambda^{(i)}R^{(i)'} &= R^{(i)}S^{(i)}\text{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2}) S^{(i)'} R^{(i)'} \\ &= \text{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2}) \end{aligned}$$

and it immediately follows that

$$R\Lambda R' = \mathrm{diag}(D^{(0)}, \ldots, D^{(\nu-1)})$$

where $D^{(i)} = \mathrm{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2})$. Thus, Eq. (63) becomes

$$V = (I - C)^{-1}\,\mathrm{diag}(D^{(0)}, \ldots, D^{(\nu-1)})[(I - C)^{-1}]'. \tag{65}$$

Also, from the relation $\psi_i(u) = \sum_{\ell=1}^{u} -\pi_i(\ell)\psi_{i-\ell}(u-\ell)$, it can be shown that

$$(I - C)^{-1} = \sum_{n=1}^{D-1} E_n \Pi^{[D\nu - n(D+1)]}$$

where E_n is defined in (22). Using estimates from the proof of Corollary 2.2.3 in [5] along with condition (18) it is not hard to show that $N^{1/2}(\psi - \theta) \to 0$. Then it follows that

$$N^{1/2}(\hat{\theta} - \psi) \Rightarrow N(0, V),$$

where $\psi = (\psi_0(1), \ldots, \psi_0(D), \ldots, \psi_{\nu-1}(1), \ldots, \psi_{\nu-1}(D))'$. We have proved the theorem. $\square$

References

[1] G. Adams, C. Goodwin, Parameter estimation for periodic ARMA models, J. Time Ser. Anal. 16 (1995) 127–145.

[2] P. Anderson, M. Meerschaert, Periodic moving averages of random variables with regularly varying tails, Ann. Statist. 25 (1997) 771–785.

[3] P. Anderson, M. Meerschaert, Modeling river flows with heavy tails, Water Resources Res. 34 (1998) 2271–2280.

[4] P. Anderson, M. Meerschaert, Parameter estimation for periodically stationary time series, J. Time Ser. Anal. 26 (4) (2005) 489–518.

[5] P. Anderson, M. Meerschaert, A. Vecchia, Innovations algorithm for periodically stationary time series, Stochastic Process. Appl. 83 (1999) 149–169.

[6] P. Anderson, A. Vecchia, Asymptotic results for periodic autoregressive moving-average processes, J. Time Ser. Anal. 14 (1993) 1–18.

[7] I.V. Basawa, R.B. Lund, Large sample properties of parameter estimates from periodic ARMA models, J. Time Ser. Anal. 22 (2001) 651–663.

[8] K. Berk, Consistent autoregressive spectral estimates, Ann. Statist. 2 (1974) 489–502.

[9] S. Bernstein, Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes, Math. Ann. 97 (1926) 1–59.

[10] R. Bhansali, Linear prediction by autoregressive model fitting in the time domain, Ann. Statist. 6 (1978) 224–231.

[12] P. Billingsley, Probability and Measure, third ed., Wiley, New York, 1995.

[13] P. Brockwell, R. Davis, Simple consistent estimation of the coefficients of a linear filter, Stochast. Process. Appl. 28 (1988) 47–59.

[14] P. Brockwell, R. Davis, Time Series: Theory and Methods, second ed., Springer-Verlag, New York, 1991.

[15] J.L. Doob, Stochastic Processes, Wiley, New York, 1953.

[16] W. Feller, An introduction to Probability Theory and its Applications, second ed., Wiley, New York, 1971.

[17] G. Golub, C. Van Loan, Matrix Computations, second ed., Johns Hopkins University Press, Baltimore, MD, 1989.

[18] P. Hall, C.C. Heyde, Martingale Limit Theory and its Application, Academic Press, New York, 1980.

[19] E.J. Hannan, Multiple Time Series, Wiley, New York, 1970.

[20] R. Jones, W. Brelsford, Time series with periodic structure, Biometrika 54 (1967) 403–408.

[21] R. Lewis, G. Reinsel, Prediction of multivariate time series by autoregressive model fitting, J. Multivariate Anal. 16 (1985) 393–411.

[22] R. Lund, I. Basawa, Recursive prediction and likelihood evaluation for periodic ARMA models, J. Time Ser. Anal. 21 (2000) 75–93.

[23] R.B. Lund, I.V. Basawa, Modeling for periodically correlated time series, In: Asymptotics, Nonparametrics, and Time Series, 1999, pp. 37–62.

[24] R.B. Lund, H. Hurd, P. Bloomfield, R.L. Smith, Climatological time series with periodic correlation, J. Climate 11 (1995) 2787–2809.
[25] M.M. Meerschaert, H.P. Scheffler, Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice, Wiley Interscience, New York, 2001.
[26] M. Pagano, On periodic and multiple autoregressions, Ann. Statist. 6 (1978) 1310–1317.
[27] S. Resnick, A Probability Path, Birkhüser, Boston, 2001.
[28] J. Salas, G. Tabios, P. Bartolini, Approaches to multivariate modeling of water resources time series, Water Res. Bull. 21 (1985) 683–708.
[29] G. Samorodnitsky, M. Taqqu, Stable non-Gaussian Random Processes, Chapman & Hall, New York, 1994.
[30] Q. Shao, R.B. Lund, Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models, J. Time Ser. Anal. 25 (3) (2004) 359–372.
[31] Y.G. Tesfaye, M.M. Meerschaert, P.L. Anderson, Identification of PARMA Models and Their Application to the Modeling of Riverflows, Water Resources Research, 2006, in press. Preprint available at ⟨www.maths.otago.ac.nz/%7Emcubed/Fraser.pdf⟩.
[33] D. Tjostheim, J. Paulsen, Empirical identification of multiple time series, J. Time Ser. Anal. 3 (1982) 265–282.
[34] B. Troutman, Some results in periodic autoregression, Biometrika 6 (1979) 219–228.
[35] T.A. Ula, Forecasting of multivariate periodic autoregressive moving average processes, J. Time Series Anal. 14 (1993) 645.
[36] T.A. Ula, A.A. Smadi, Periodic stationarity conditions for periodic autoregressive moving average processess as eigenvalue problems, Water Resources Research 33 (8) (1997) 1929–1934.
[37] T.A. Ula, A.A. Smadi, Identification of periodic moving average models, Communications in Statistics: Theory and Methods 32 (12) (2003) 2465–2475.
[38] A.V. Vecchia, Periodic autoregressive-moving average (PARMA) modelling with applications to water resources, Water Resour. Bull. 21 (1985) 721–730.
[39] A.V. Vecchia, Maximum likelihood estimation for periodic moving average models, Technometrics 27 (4) (1985) 375–384.
[40] A.V. Vecchia, R. Ballerini, Testing for periodic autocorrelations in seasonal time series data, Biometrika 78 (1) (1991) 53–63.
[41] A.V. Vecchia, J.T.B. Obeysekera, J.D. Salas, D.C. Boes, Aggregation and estimation for low-order periodic ARMA models, Water Resources Research 19 (5) (1983) 1297–1306.