PARAMETER ESTIMATION FOR PERIODICALLY STATIONARY TIME SERIES

By Paul L. Anderson and Mark M. Meerschaert

Albion College and University of Otago

First Version received December 2002

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 innovations have a finite fourth moment. These asymptotic results are useful to determine which model parameters are significant. In the process, we also develop asymptotics for the Yule–Walker estimates.

Keywords. Time series; periodically stationary; innovations algorithm.
AMS 1991 subject classifications. Primary 62M10, 62E20; secondary 60E07, 60F05.

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 $h = 0, \pm 1, \pm 2, \ldots$ are all periodic functions of time t with the same period $v \geq 1$. If $v = 1$, 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 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 in Adams and Goodwin (1995), Anderson and Vecchia (1993), Anderson and Meerschaert (1997, 1998), Anderson *et al.* (1999), Jones and Brelsford (1967), Lund and Basawa (2000), Pagano (1978), Salas *et al.* (1985), Tjostheim and Paulsen (1982), Troutman (1979), Vecchia and Ballerini (1991) and Ula (1993).

Anderson *et al.* (1999) develop the innovations algorithm for periodic ARMA model parameters. In this paper, we provide the asymptotic estimates necessary to determine which of these estimates are statistically different from zero, under the classical assumption that the innovations have finite fourth moment. Brockwell and Davis (1988), discuss asymptotics of the innovations algorithm for stationary time series, using results of Berk (1974) and Bhansali (1978). Our results reduce to theirs, when the period $v = 1$. Since our technical approach extends that of Brockwell and Davis (1988), we also need to develop periodically stationary

0143-9782/05/04 489–518 JOURNAL OF TIME SERIES ANALYSIS Vol. 26, No. 4

analogues of results in Berk (1974) and Bhansali (1978). In particular, we obtain asymptotics for the Yule–Walker estimates of a periodically stationary process. Although the innovations estimates are more useful in practice, the asymptotics of the Yule–Walker estimates are also of some independent interest.

2. THE INNOVATIONS ALGORITHM FOR PARMA PROCESSES

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 SD σ_t such that $\{\sigma_t^{-1}\varepsilon_t\}$ is i.i.d. The autoregressive (AR) parameters $\phi_t(j)$, the moving average parameters $\theta_t(j)$, and the residual SDs σ_t are all periodic functions of t with the same period $\nu \geq 1$. In this paper, we will make the classical assumption $E\varepsilon_t^4 < \infty$, which leads to normal asymptotics for the parameter estimates. 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.

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

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

Then

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

where the vector of coefficients $\phi_k^{(i)} = (\phi_{k,1}^{(i)}, \ldots, \phi_{k,k}^{(i)})'$ solves the prediction equation

$$\Gamma_{k,i} \phi_k^{(i)} = \gamma_k^{(i)} \tag{5}$$

with $\gamma_k^{(i)} = (\gamma_{i+k-1}(1), \gamma_{i+k-2}(2), \ldots, \gamma_i(k))'$ 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 $\hat{\phi}_{k,j}^{(i)}$ of $\phi_{k,j}^{(i)}$.

Because the process is nonstationary, the Durbin–Watson algorithm for computing $\hat{\phi}_{k,j}^{(i)}$ does not apply. However, the innovations algorithm still applies to a nonstationary process. Writing

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

yields the one-step predictors in terms of the innovations $X_{i+k-j} - \hat{X}_{i+k-j}^{(i)}$. Proposition 4.1 of Lund and Basawa (2000), shows that if $\sigma_i^2 > 0$ for $i = 0, \ldots, v-1$, then for a causal PARMA$_v(p, q)$ process the covariance matrix $\Gamma_{k,i}$ is nonsingular for every $k \geq 1$ and each i. Anderson *et al.* (1999) shows that if $EX_t = 0$ and $\Gamma_{k,i}$ is nonsingular for each $k \geq 1$, then the one-step predictors $\hat{X}_{i+k}$, $k \geq 0$, and their mean-square errors (MSEs)$v_{k,i}$, $k \geq 1$, are given by

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

where (9) is solved in the order $v_{0,i}$, $\theta_{1,1}^{(i)}$, $v_{1,i}$, $\theta_{2,2}^{(i)}$, $\theta_{2,1}^{(i)}$, $v_{2,i}$, $\theta_{3,3}^{(i)}$, $\theta_{3,2}^{(i)}$, $\theta_{3,1}^{(i)}$, $v_{3,i}, \ldots$ The results in Anderson *et al.* (1999) show that

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

as $k \to \infty$ for all i, j, where

$$\langle j \rangle = \begin{cases} j - v[j/v] & \text{if } j = 0, 1, \ldots, \\ v + j - v[j/v + 1] & \text{if } j = -1, -2, \ldots \end{cases}$$

and $[\cdot]$ is the greatest integer function.

If we replace the autocovariances in (9) with the corresponding sample autocovariances (7), we obtain the innovations estimates $\hat{\theta}_{k,\ell}^{(i)}$ and $\hat{v}_{k,i}$. Similarly, replacing the autocovariances in (5) with the corresponding sample autocovariances yields the Yule–Walker estimators $\hat{\phi}_{k,\ell}^{(i)}$. The consistency of these estimators was also established in Anderson *et al.* (1999). Suppose that $\{X_t\}$ is the mean zero PARMA process with period v given by (1) and that $E(\varepsilon_t^4) < \infty$. Assume that the spectral density matrix $f(\lambda)$ of the equivalent vector ARMA process is such that $mz'z \leq z'f(\lambda)z \leq Mz'z$, $-\pi \leq \lambda \leq \pi$, for some m and M such that $0 < m \leq M < \infty$ and for all z in R^v. If k is chosen as a function of the sample size N, so that $k^2/N \to 0$ as $N \to \infty$ and $k \to \infty$, then the results in Anderson *et al.* (1999) also show that

$$\begin{aligned} &\hat{\theta}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(j), \\ &\hat{v}_{k,\langle i-k \rangle} \xrightarrow{P} \sigma_i^2, \\ &\hat{\phi}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} -\pi_i(j) \end{aligned} \tag{11}$$

for all i, j. This yields a practical method for estimating the model parameters, in the classical case of finite 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

We compute the asymptotic distribution for the innovations estimates of the parameters in a periodically stationary time series (2) with period $v \geq 1$. In the process, we also obtain the asymptotic distribution of the Yule–Walker 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

$$Y_t = \sum_{j=-\infty}^{\infty} \Psi_j Z_{t-j}, \tag{12}$$

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 Yule–Walker estimates. A similar result was obtained by Lewis and Reinsel (1985) for vector AR 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 1. *Suppose that the periodically stationary moving average* (2) *is causal, invertible,* $E\varepsilon_t^4 < \infty$, *and that for some* $0 < g \leq G < \infty$ *we have* $gz'z \leq z'f(\lambda)z \leq Gz'z$ *for all* $-\pi \leq \lambda \leq \pi$, *and all* z *in* $\mathbb{R}^\nu$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (12). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3/N \to 0$ *and*

$$N^{1/2} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \to 0 \quad \text{for } \ell = 0, 1, \ldots, \nu - 1 \tag{13}$$

then for any fixed positive integer D

$$N^{1/2}\left(\pi_i(u) + \hat{\phi}_{k,u}^{\langle i-k \rangle} : 1 \leq u \leq D, i = 0, \ldots, \nu - 1\right) \Rightarrow N(0, \Lambda), \tag{14}$$

where

$$\Lambda = \mathrm{diag}(\sigma_0^2 \Lambda^{(0)}, \sigma_1^2 \Lambda^{(1)}, \ldots, \sigma_{\nu-1}^2 \Lambda^{(\nu-1)}) \tag{15}$$

with

$$(\Lambda^{(i)})_{u,v} = \sum_{s=0}^{m-1} \pi_{i-m+s}(s) \pi_{i-m+s}(s + |v - u|) \sigma_{i-m+s}^{-2} \tag{16}$$

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

In Theorem 1, note that $\Lambda^{(i)}$ is a $D \times D$ matrix and the $D\nu$-dimensional vector given in (14) is ordered

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

Next we present our main result, giving asymptotics for innovations estimates of a periodically stationary time series.

THEOREM 2. *Suppose that the periodically stationary moving average* (2) *is causal, invertible,* $E\varepsilon_t^4 < \infty$, *and that for some* $0 < g \leq G < \infty$ *we have* $gz'z \leq z'f(\lambda)z \leq Gz'z$ *for all* $-\pi \leq \lambda \leq \pi$, *and all* z *in* $\mathbb{R}^\nu$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (12). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3/N \to 0$ *and*

$$N^{1/2} \sum_{j=1}^{\infty} |\pi_\ell(k+j)| \to 0 \quad \text{for } \ell = 0, 1, \ldots, \nu - 1 \tag{17}$$

then

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-k \rangle)} - \psi_i(u) : u = 1, \ldots, D, i = 0, \ldots, \nu - 1) \Rightarrow N(0, V), \tag{18}$$

where

$$V = A \ \mathrm{diag}(\sigma_0^2 D^{(0)}, \ldots, \sigma_{\nu-1}^2 D^{(\nu-1)}) A', \tag{19}$$

$$A = \sum_{n=0}^{D-1} E_n \Pi^{[D\nu - n(D+1)]}, \tag{20}$$

$$\begin{aligned} E_n = \operatorname{diag}(&\overbrace{0,\ldots,0}^{n}, \overbrace{\psi_0(n),\ldots,\psi_0(n)}^{D-n}, \overbrace{0,\ldots,0}^{n}, \\ &\underbrace{\psi_1(n),\ldots,\psi_1(n)}_{D-n},\ldots,\underbrace{0,\ldots,0}_{n},\underbrace{\psi_{\nu-1}(n),\ldots,\psi_{\nu-1}(n)}_{D-n}), \end{aligned} \tag{21}$$

$$D^{(i)} = \operatorname{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2}) \tag{22}$$

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

$$\Pi = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{23}$$

Note that Π^0 is the $D\nu \times D\nu$ identity matrix and $\Pi^{-\ell} \equiv (\Pi')^{\ell}$. Matrix multiplication yields Corollary 1.

COROLLARY 1. *Regarding Theorem* 2, *in particular*, *we have that*

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-k \rangle)} - \psi_i(u)) \Rightarrow N(0, \sigma_{i-u}^{-2} \sum_{n=0}^{u-1} \sigma_{i-n}^2 \psi_i^2(n)). \tag{24}$$

REMARK. Corollary 1 also holds the asymptotic result for the second-order stationary process, where the period is just $\nu = 1$. In this case $\sigma_i^2 = \sigma^2$ so (24) becomes

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

which agrees with Theorem 2.1 in Brockwell and Davis (1988).

4. PROOFS

The proof of Theorem 1 requires some preliminaries. First we show that $(\Lambda^{(i)})_{u,v}$ is the limit of the u, v entry in the inverse covariance matrix (6).

LEMMA 1. *Suppose that the periodically stationary moving average* (2) *is causal*, *invertible*, *and that* $E\varepsilon_t^4 < \infty$. *If* $\Gamma_{k,i}$ *is given by* (6) *and* $(A)_{u,v}$ *denotes the* u, v *entry of the matrix* A, *then for any* $i = 0, 1, \ldots, \nu - 1$ *we have*

$$(\Gamma_{k,\langle i-k\rangle}^{-1})_{u,v} \to (\Lambda^{(i)})_{u,v} = \sum_{s=0}^{m-1} \pi_{i-m+s}(s)\pi_{i-m+s}(s+|v-u|)\sigma_{i-m+s}^{-2} \tag{25}$$

as $k \to \infty$, *where* $m = \min(u, v)$.

PROOF. The prediction equation (5) yield

$$\gamma_{i+\ell}(k-\ell) - \phi_{k,1}^{(i)}\gamma_{i+\ell}(k-1-\ell) - \cdots - \phi_{k,k}^{(i)}\gamma_{i+\ell}(-\ell) = 0 \tag{26}$$

for $0 \le \ell \le k-1$ and since $v_{k,i} = \langle X_{i+k}, X_{i+k} - \hat{X}_{i+k}^{(i)} \rangle$ we also have

$$v_{k,i} = \gamma_{i+k}(0) - \phi_{k,1}^{(i)}\gamma_{i+k}(-1) - \cdots - \phi_{k,k}^{(i)}\gamma_{i+k}(-k). \tag{27}$$

Define

$$F = \begin{pmatrix} 1 & -\phi_{k-1,1}^{(i)} & -\phi_{k-1,2}^{(i)} & \cdots & -\phi_{k-1,k-1}^{(i)} \\ 0 & 1 & -\phi_{k-2,1}^{(i)} & \cdots & -\phi_{k-2,k-2}^{(i)} \\ \vdots & & \ddots & & \vdots \\ & & & & -\phi_{1,1}^{(i)} \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

and use the fact that $\gamma_r(s) = \gamma_{r+s}(-s)$, so that $(\Gamma_{k,i})_{j,\ell} = \gamma_{i+k-\ell}(\ell - j)$, to compute that for $1 \le j < \ell$ and $2 \le \ell \le k$ the (j, ℓ) element of the matrix $F\Gamma_{k,i}$ is

$$\gamma_{i+k-\ell}(\ell-j) - \phi_{k-j,1}^{(i)}\gamma_{i+k-\ell}(\ell-j-1) - \cdots - \phi_{k-j,k-j}^{(i)}\gamma_{i+k-\ell}(\ell-k).$$

Substitute $n' = k - j$ and $\ell' = k - \ell$ to obtain

$$(F\Gamma_{k,i})_{j,\ell} = \gamma_{i+\ell'}(n'-\ell') - \phi_{n',1}^{(i)}\gamma_{i+\ell'}(n'-\ell'-1) - \cdots - \phi_{n',n'}^{(i)}\gamma_{i+\ell'}(-\ell') = 0$$

in view of (26), so that $F\Gamma_{k,i}$ is lower triangular. Also (27) yields $(F\Gamma_{k,i})_{j,j} = v_{k-j,i}$ for $1 \le j \le k$. Since the transpose F' is also lower triangular, the matrix $F\Gamma_{k,i}F'$ is still lower triangular, with (j, j) element $v_{k-j,i}$. But $F\Gamma_{k,i}F'$ is a symmetric matrix, so it is diagonal. In fact, $F\Gamma_{k,i}F' = H^{-1}$, where $H = \operatorname{diag}(v_{k-1,i}^{-1}, v_{k-2,i}^{-1}, \ldots, v_{0,i}^{-1})$. Then $\Gamma_{k,i}^{-1} = F'HF$, and a computation shows that

$$(\Gamma_{k,i}^{-1})_{u,v} = \sum_{\ell=1}^{\min(u,v)} \phi_{k-\ell,u-\ell}^{(i)}\phi_{k-\ell,v-\ell}^{(i)}v_{k-\ell,i}^{-1}, \tag{28}$$

where the number of summands is fixed and finite for all k, and we set $\phi_{k,0}^{(i)} \equiv -1$ to simplify the notation. Substitute $\langle i - k\rangle$ for i in (28) and apply (10) using $\langle i - k\rangle = \langle (i - \ell) - (k - \ell)\rangle$ to obtain

$$\begin{aligned} v_{k-\ell,\langle i-k\rangle} &\to \sigma_{i-\ell}^2, \\ \phi_{k-\ell,u-\ell}^{(\langle i-k\rangle)} &\to -\pi_{i-\ell}(u-\ell), \\ \phi_{k-\ell,v-\ell}^{(\langle i-k\rangle)} &\to -\pi_{i-\ell}(v-\ell) \end{aligned} \tag{29}$$

as $k \to \infty$, so that

$$(\Gamma^{-1}_{k,\langle i-k\rangle})_{u,v} \to \sum_{\ell=1}^{\min(u,v)} \pi_{i-\ell}(u-\ell)\pi_{i-\ell}(v-\ell)\sigma^{-2}_{i-\ell}$$

as $k \to \infty$. Substitute $m = \min(u, v)$ and commute the first two terms if $m = v$ to finish the proof of Lemma 1. □

In order to prove Theorem 1, it will be convenient to use a slightly different estimator $f^{(i)}_{k,j}$ for the coefficients $\phi^{(i)}_{k,j}$ of the one-step predictor (4), obtained by minimizing

$$w_{k,i} = (N-k)^{-1} \sum_{j=0}^{N-k-1} (X_{jv+i+k} - f^{(i)}_{k,1} X_{jv+i+k-1} - \cdots - f^{(i)}_{k,k} X_{jv+i})^2 \tag{30}$$

for $i = 0, \ldots, v-1$. For i fixed, take partials in (30) with respect to $f^{(i)}_{k,\ell}$ for each $\ell = 1, \ldots, k$ to obtain

$$\sum_{j=0}^{N-k-1} (X_{jv+i+k} - f^{(i)}_{k,1} X_{jv+i+k-1} - \cdots - f^{(i)}_{k,k} X_{jv+i}) X_{jv+i+k-\ell} = 0$$

and rearrange to get

$$f^{(i)}_{k,1}\hat{s}^{(i)}_{k-1,k-\ell} + \cdots + f^{(i)}_{k,k}\hat{s}^{(i)}_{0,k-\ell} = \hat{s}^{(i)}_{k,k-\ell} \tag{31}$$

for each $\ell = 1, \ldots, k$, where

$$\hat{s}^{(i)}_{m,n} = (N-k)^{-1} \sum_{j=0}^{N-k-1} X_{jv+i+m} X_{jv+i+n}.$$

Now define $\hat{r}^{(i)}_k = (\hat{s}^{(i)}_{k,k-1}, \ldots, \hat{s}^{(i)}_{k,0})'$ and $f^{(i)}_k = (f^{(i)}_{k,1}, \ldots, f^{(i)}_{k,k})'$ and let

$$\hat{R}^{(i)}_k = \begin{pmatrix} \hat{s}^{(i)}_{k-1,k-1} & \cdots & \hat{s}^{(i)}_{0,k-1} \\ \vdots & & \vdots \\ \hat{s}^{(i)}_{k-1,0} & \cdots & \hat{s}^{(i)}_{0,0} \end{pmatrix}$$

so that (31) becomes

$$\hat{R}^{(i)}_k f^{(i)}_k = \hat{r}^{(i)}_k \tag{32}$$

analogous to the prediction equations (5).

Theorem 1 and Lemma 1 depend on modulo v arithmetic, which requires our $\langle i-k\rangle$-notation. Since Lemmas 2–8 do not have this dependence, we proceed with the less cumbersome i-notation.

LEMMA 2. *Let* $\pi^{(\mathrm{i})}_k = (\pi_{\mathrm{i+k}}(1), \ldots, \pi_{\mathrm{i+k}}(\mathrm{k}))'$ *and* $X^{(\mathrm{i})}_j(\mathrm{k}) = (\mathrm{X}_{(\mathrm{j-k})v+\mathrm{i+k}-1}, \ldots, X_{(\mathrm{j-k})v+\mathrm{i}})'$. *Then for all* $i = 0, \ldots, v-1$ *and* $k \geq 1$ *we have*

$$\pi_k^{(i)} + f_k^{(i)} = (\hat{R}_k^{(i)})^{-1} \frac{1}{N-k} \sum_{j=k}^{N-1} X_j^{(i)}(k) \varepsilon_{(j-k)v+i+k,k}, \tag{33}$$

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

PROOF. Note that

$$\hat{R}_k^{(i)} = \frac{1}{N-k} \sum_{j=k}^{N-1} X_j^{(i)}(k) X_j^{(i)}(k)' \quad \text{and} \quad \hat{r}_k^{(i)} = \frac{1}{N-k} \sum_{j=k}^{N-1} X_j^{(i)}(k) X_{(j-k)v+i+k}$$

and apply (32) to obtain

$$\begin{aligned} \pi_k^{(i)} + f_k^{(i)} &= \pi_k^{(i)} + (\hat{R}_k^{(i)})^{-1} \hat{r}_k^{(i)} \\ &= (\hat{R}_k^{(i)})^{-1} \left[\hat{R}_k^{(i)} \pi_k^{(i)} + \hat{r}_k^{(i)} \right] \\ &= (\hat{R}_k^{(i)})^{-1} \frac{1}{N-k} \sum_{j=k}^{N-1} X_j^{(i)}(k) X_j^{(i)}(k)' \pi_k^{(i)} + X_j^{(i)}(k) X_{(j-k)v+i+k} \\ &= (\hat{R}_k^{(i)})^{-1} \frac{1}{N-k} \sum_{j=k}^{N-1} X_j^{(i)}(k) \left[X_j^{(i)}(k)' \pi_k^{(i)} + X_{(j-k)v+i+k} \right], \end{aligned}$$

which is equivalent to (33). □

LEMMA 3. *For all* $i = 0, \ldots, v-1$ *and* $k \geq 1$ *we have*

$$w_{k,i} = \hat{r}_{i+k}(0) - f_k^{(i)\prime} \hat{r}_k^{(i)}, \tag{34}$$

where $\hat{r}_i(0) = (N-k)^{-1} \sum_{j=k}^{N-1} X_{(j-k)v+i}^2$.

PROOF. The right-hand side of (34) equals

$$\begin{aligned} &\hat{r}_{i+k}(0) - \frac{1}{N-k} \sum_{j=k}^{N-1} f_k^{(i)\prime} X_j^{(i)}(k) X_{(j-k)v+i+k} \\ &= \frac{1}{N-k} \sum_{j=k}^{N-1} \left[X_{(j-k)v+i+k} - f_{k,1}^{(i)} X_{(j-k)v+i+k-1} - \cdots - f_{k,k}^{(i)} X_{(j-k)v+i} \right] X_{(j-k)v+i+k}, \end{aligned}$$

while $w_{k,i} = (N-k)^{-1} \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2$, where we let $\mathbf{Y} = (X_{i+k}, X_{v+i+k}, \ldots, X_{(N-k-1)v+i+k})'$, $\mathbf{X}_t = (X_{i+k-t}, X_{v+i+k-t}, \ldots, X_{(N-k-1)v+i+k-t})'$ for $t = 1, \ldots, k$ and $\hat{\mathbf{Y}} = f_{k,1}^{(i)} \mathbf{X}_1 + \cdots + f_{k,k}^{(i)} \mathbf{X}_k$. Since $\langle \hat{\mathbf{Y}}, \mathbf{Y} - \hat{\mathbf{Y}} \rangle = 0$, we also have

$$\begin{aligned} (N-k) w_{k,i} &= \langle \mathbf{Y} - \hat{\mathbf{Y}}, \mathbf{Y} - \hat{\mathbf{Y}} \rangle = \langle \mathbf{Y} - \hat{\mathbf{Y}} + \hat{\mathbf{Y}}, \mathbf{Y} - \hat{\mathbf{Y}} \rangle = \langle \mathbf{Y}, \mathbf{Y} - \hat{\mathbf{Y}} \rangle \\ &= \sum_{j=k}^{N-1} \left[X_{(j-k)v+i+k} - f_{k,1}^{(i)} X_{(j-k)v+i+k-1} - \cdots - f_{k,k}^{(i)} X_{(j-k)v+i} \right] X_{(j-k)v+i+k} \end{aligned}$$

and (34) follows easily. □

LEMMA 4. *For all* $i = 0, \ldots, v-1$ *and* $k \geq 1$ *we have*

$$\begin{aligned} w_{k,i} - \sigma_{i+k}^2 &= (\hat{r}_{i+k}(0) - \gamma_{i+k}(0)) - (f_k^{(i)} + \pi_k^{(i)})'\gamma_k^{(i)} \\ &\quad + \pi_k^{(i)'}(\hat{r}_k^{(i)} - \gamma_k^{(i)}) - (f_k^{(i)} + \pi_k^{(i)})'(\hat{r}_k^{(i)} - \gamma_k^{(i)}) \\ &\quad - \sum_{j=k+1}^{\infty} \pi_{i+k}(j)\gamma_{i+k-j}(j), \end{aligned} \tag{35}$$

where $\gamma_k^{(i)} = (\gamma_{i+k-1}(1), \gamma_{i+k-2}(2), \ldots, \gamma_i(k))'$.

PROOF. From (34) we have

$$\begin{aligned} w_{k,i} - \sigma_{i+k}^2 &= \hat{r}_{i+k}(0) - f_k^{(i)'}\hat{r}_k^{(i)} - \sigma_{i+k}^2 \\ &= (\hat{r}_{i+k}(0) - \gamma_{i+k}(0)) - (f_k^{(i)} + \pi_k^{(i)})'\gamma_k^{(i)} \\ &\quad + \pi_k^{(i)'}(\hat{r}_k^{(i)} - \gamma_k^{(i)}) - (f_k^{(i)} + \pi_k^{(i)})'(\hat{r}_k^{(i)} - \gamma_k^{(i)}) \\ &\quad + \gamma_{i+k}(0) + \pi_k^{(i)'}\gamma_k^{(i)} - \sigma_{i+k}^2 \end{aligned}$$

since the remaining terms cancel. Define $\hat{X}_t = X_t - \varepsilon_t$ so that $\hat{X}_{i+k} = X_{i+k} - \varepsilon_{i+k} = -\pi_{i+k}(1)X_{i+k-1} - \pi_{i+k}(2)X_{i+k-2} - \cdots$ and $\sigma_{i+k}^2 = E(\varepsilon_{i+k}^2) = E[(X_{i+k} - \hat{X}_{i+k})^2]$. Since $E[\hat{X}_{i+k}(X_{i+k} - \hat{X}_{i+k})] = 0$ it follows that $\sigma_{i+k}^2 = E[X_{i+k}(X_{i+k} - \hat{X}_{i+k})] = \gamma_{i+k}(0) + \pi_{i+k}(1)\gamma_{i+k-1}(1) + \pi_{i+k}(2)\gamma_{i+k-2}(2) + \cdots$ and (35) follows easily. □

LEMMA 5. *Let* $c_i(\ell)$, $d_i(\ell)$ *for* $\ell = 0, 1, 2, \ldots$ *be arbitrary sequences of real numbers, 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

$$C_i^2 = \sum_{k=0}^{\infty} c_i(k)^2 \quad \text{and} \quad D_j^2 = \sum_{m=0}^{\infty} d_j(m)^2$$

for $0 \leq i, j < v$. *Then*

$$E\left[\left(\sum_{t=1}^{M} u_{tv+i}v_{tv+j}\right)^2\right] \leq 4M^2C^2D^2\eta, \tag{36}$$

where $C^2 = \max(C_i^2)$, $D^2 = \max(D_j^2)$, $\eta = \max(\eta_t)$, *and* $\eta_t = E(\varepsilon_t^4)$.

PROOF. Since $\sigma_t^{-1}\varepsilon_t$ are i.i.d., $\eta = \max(\eta_t\colon -\infty < t < \infty) = \max(\eta_t\colon 0 \leq i < v) < \infty$. The left-hand side of (36) consists of M^2 terms of the form

$$E(u_iv_ju_{\ell v+i}v_{\ell v+j}) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty} c_i(k)d_j(m)c_i(r)d_j(s)E(\varepsilon_{i-k}\varepsilon_{j-m}\varepsilon_{\ell v+i-r}\varepsilon_{\ell v+j-s}). \tag{37}$$

The terms in (37) with $i - k = j - m = \ell v + i - r = \ell v + j - s$ sum to

$$
\begin{aligned}
&\sum_{k=0}^{\infty} c_i(k) d_j(k+j-i) c_i(k+\ell v) d_j(k+\ell v+j-i) E(\varepsilon_{i-k}^4) \\
&\quad \leq \eta \sum_{k=0}^{\infty} c_i(k) d_j(k+j-i) c_i(k+\ell v) d_j(k+\ell v+j-i) \\
&\quad \leq \eta \sum_{k=0}^{\infty} |c_i(k) d_j(k)|^2 \\
&\quad \leq \eta \sum_{k=0}^{\infty} c_i(k)^2 \sum_{m=0}^{\infty} d_j(m)^2 \\
&\quad \leq \eta C^2 D^2
\end{aligned}
$$

using the fact that for $a_n, b_n > 0$ we always have $(\sum a_n b_n)^2 \leq \max(a_n)^2 \sum b_n^2 \leq \sum a_n^2 \sum b_n^2$. The terms in (37) with $i-k = j-m \neq \ell v + i - r = \ell v + j - s$ sum to

$$
\begin{aligned}
&\sum_{k=0}^{\infty} \sum_{r=0}^{\infty} c_i(k) d_j(k+j-i) c_i(r) d_j(r+j-i) E(\varepsilon_{i-k}^2 \varepsilon_{\ell v+i-r}^2) \\
&\quad \leq \eta \sum_{k=0}^{\infty} \sum_{r=0}^{\infty} |c_i(k) d_j(k+j-i) c_i(r) d_j(r+j-i)| \\
&\quad \leq \eta \sum_{k=0}^{\infty} |c_i(k) d_j(k)| \sum_{r=0}^{\infty} |c_i(r) d_j(r)| \\
&\quad \leq \eta \sum_{k=0}^{\infty} c_i(k)^2 \sum_{m=0}^{\infty} d_j(m)^2 \\
&\quad \leq \eta C^2 D^2
\end{aligned}
$$

since $E(\varepsilon_{i-k}^2 \varepsilon_{\ell v+i-r}^2) = E(\varepsilon_{i-k}^2) E(\varepsilon_{i-r}^2) \leq \sqrt{E(\varepsilon_{i-k}^4) E(\varepsilon_{i-r}^4)} = \sqrt{\eta_{i-k} \eta_{i-r}} \leq \eta$ by the Schwarz inequality. Similarly, the terms in (37) with $i-k = \ell v + i - r \neq j - m = \ell v + j - s$ sum to

$$
\sum_{k=0}^{\infty} \sum_{m=0}^{\infty} c_i(k) d_j(m) c_i(\ell v+k) d_j(\ell v+m) E(\varepsilon_{i-k}^2 \varepsilon_{j-m}^2) \leq \eta C^2 D^2
$$

and the terms in (37) with $i-k = \ell v + j - s \neq j - m = \ell v + i - r$ sum to

$$
\sum_{k=0}^{\infty} \sum_{m=0}^{\infty} c_i(k) d_j(m) c_i(\ell v+m+i-j) d_j(\ell v+k+j-i) E(\varepsilon_{i-k}^2 \varepsilon_{j-m}^2) \leq \eta C^2 D^2,
$$

while the remaining terms are zero. Then $E(u_i v_j u_{\ell v+i} v_{\ell v+j}) < 4\eta C^2 D^2$ and (36) follows immediately. □

LEMMA 6. *For $\varepsilon_{t,k}$ as in Lemma 2 and u_{tv+i} as in Lemma 5 we have*

$$E\left[\left(\sum_{t=k}^{N-1} u_{(t-k)v+i}\left(\varepsilon_{(t-k)v+\ell,k} - \varepsilon_{(t-k)v+\ell}\right)\right)^2\right] \leq 4\eta(N-k)^2 C^2 B \max_{\ell} \sum_{j=1}^{\infty} \pi_\ell(k+j)^2, \tag{38}$$

where C^2, η are as in Lemma 5 and $B = \left(\sum_{i=0}^{v-1} \sum_{\ell=0}^{\infty} |\psi_i(\ell)|\right)^2$.

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,

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

is finite, and hence we also have $\sum_{j=1}^{\infty} d_{\ell,k}(k+j)^2 < \infty$. Now apply Lemma 5 with $v_{tv+\ell} = \varepsilon_{tv+\ell,k} - \varepsilon_{tv+\ell}$ to see that

$$\begin{aligned}
E\left[\left(\sum_{t=0}^{N-k-1} u_{tv+i}\left(\varepsilon_{tv+\ell,k} - \varepsilon_{tv+\ell}\right)\right)^2\right] &= E\left[\left(\sum_{t=k}^{N-1} u_{(t-k)v+i}\left(\varepsilon_{(t-k)v+\ell,k} - \varepsilon_{(t-k)v+\ell}\right)\right)^2\right] \\
&\leq 4(N-k)^2 C^2 D_{\ell,k}^2 \eta \\
&\leq 4(N-k)^2 C^2 D_k^2 \eta,
\end{aligned}$$

where $D_k^2 = \max(D_{\ell,k}^2 : 0 \leq \ell < v)$ and $D_{\ell,k}^2 = \sum_{j=1}^{\infty} d_{\ell,k}(k+j)^2$. Next compute

$$
\begin{aligned}
D_{\ell,k}^2 &= \sum_{j=1}^{\infty}\left(\sum_{r=1}^{j}\pi_\ell(k+r)\psi_{\ell-(k+r)}(j-r)\right)\left(\sum_{s=1}^{j}\pi_\ell(k+s)\psi_{\ell-(k+s)}(j-s)\right)\\
&= \sum_{j=1}^{\infty}\left(\sum_{p=0}^{j-1}\pi_\ell(k+j-p)\psi_{\ell-(k+j-p)}(p)\right)\left(\sum_{q=0}^{j-1}\pi_\ell(k+j-q)\psi_{\ell-(k+j-q)}(q)\right)\\
&= \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{j=\max(p,q)+1}^{\infty}\psi_{\ell-(k+j-p)}(p)\psi_{\ell-(k+j-q)}(q)\pi_\ell(k+j-p)\pi_\ell(k+j-q)\\
&\le \sum_{i=0}^{v-1}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}|\psi_i(p)\psi_{i+q-p}(q)|\sum_{j=\max(p,q)+1}^{\infty}|\pi_\ell(k+j-p)\pi_\ell(k+j-q)|,
\end{aligned}
$$

where (without loss of generality we suppose $p \ge q$)

$$
\begin{aligned}
\sum_{j=\max(p,q)+1}^{\infty}|\pi_\ell(k+j-p)\pi_\ell(k+j-q)| &= \sum_{j=1}^{\infty}|\pi_\ell(k+j)\pi_\ell(k+j+p-q)|\\
&\le \sqrt{\sum_{j=1}^{\infty}\pi_\ell(k+j)^2\sum_{j=1}^{\infty}\pi_\ell(k+j+p-q)^2}\\
&\le \sum_{j=1}^{\infty}\pi_\ell(k+j)^2
\end{aligned}
$$

and

$$
\sum_{i=0}^{v-1}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}|\psi_i(p)\psi_{i+q-p}(q)| \le \left(\sum_{i=0}^{v-1}\sum_{\ell=0}^{\infty}|\psi_i(\ell)|\right)^2 = B.
$$

Then $D_{\ell,k}^2 \le B\sum_{j=1}^{\infty}\pi_\ell(k+j)^2$ and (38) follows easily. □

LEMMA 7. *For $\varepsilon_{t,k}$ and $X_j^{(i)}(k)$ as in Lemma 2 we have for some real constant $A > 0$ that*

$$
E\left[\left\|(N-k)^{-1}\sum_{j=k}^{N-1}X_j^{(i)}(k)(\varepsilon_{(j-k)v+i+k,k}-\varepsilon_{(j-k)v+i+k})\right\|^2\right] \le Ak\max_\ell\sum_{j=1}^{\infty}\pi_\ell(k+j)^2. \tag{39}
$$

PROOF. Rewrite the left-hand side of (39) in the form

$$
(N-k)^{-2}\sum_{s=0}^{k-1}E\left[\left(\sum_{t=k}^{N-1}X_{(t-k)v+i+s}(\varepsilon_{(t-k)v+i+k,k}-\varepsilon_{(t-k)v+i+k})\right)^2\right]
$$

and apply Lemma 6, k times with $u_{(t-k)v+i} = X_{(t-k)v+i+s}$ for each $s = 0, \ldots, k-1$ to obtain the upper bound of (39) with $A = 4\eta C^2 B$. □

LEMMA 8. *Suppose that the periodically stationary moving average* (2) *is causal, invertible,* $E\varepsilon_t^4 < \infty$, *and that for some* $0 < g \le G < \infty$ *we have* $gz'z \le z'f(\lambda)z \le Gz'z$ *for all* $-\pi \le \lambda \le \pi$, *and all* z *in* $\mathbb{R}^v$, *where* $f(\lambda)$ *is the spectral density matrix of the equivalent vector moving average process* (12). *If* $k = k(N) \to \infty$ *as* $N \to \infty$ *with* $k^3/N \to 0$ *and* (13) *holds then*

$$(N-k)^{1/2} b(k)'(\pi_k^{(i)} + f_k^{(i)}) - (N-k)^{-1/2} b(k)' \Gamma_{k,i}^{-1} \sum_{j=k}^{N-1} X_j^{(i)}(k)\varepsilon_{(j-k)v+i+k} \xrightarrow{P} 0 \quad (40)$$

for any $b(k) = (b_{k1}, \ldots, b_{kk})'$ *such that* $\|b(k)\|^2$ *remains bounded, where* $f_k^{(i)}$ *is from* (32) *and* $X_j^{(i)}(k)$ *is from Lemma* 4.2.

PROOF. Using (33) the left-hand side of (40) can be written as

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

where

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

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

so that

$$|I_1| \le (N-k)^{-1/2} \|b(k)\| \cdot \|(\hat{R}_k^{(i)})^{-1} - \Gamma_{k,i}^{-1}\| \cdot \left\| \sum_{j=k}^{N-1} X_j^{(i)}(k)\varepsilon_{(j-k)v+i+k,k} \right\| = J_1 \cdot J_2 \cdot J_3,$$

where $J_1 = \|b(k)\|$ is bounded by assumption,

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

and $J_2 = k^{1/2}\|(\hat{R}_k^{(i)})^{-1} - \Gamma_{k,i}^{-1}\| \to 0$ in probability by an argument similar to Theorem 1 in Anderson *et al.* (1999). In fact, if we let

$$M = \max\left\{ \left|\eta - 3\right| \left(\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} |\psi_i(m_1)||\psi_j(m_2)| \right)^2 < \infty,\ 0 \le i, j < v \right\}, \quad (41)$$

where $\eta = E(\varepsilon_t^4)$ and if we let $Q_{k,i} = \|\hat{R}_k^{(i)} - \Gamma_{k,i}\|$ then $(N-k)\mathrm{Var}(\hat{s}_{m,n}^{(i)}) \le M$ in general and hence $E(k^{1/2}Q_{k,i}^2) \le k^3 M/(N-k) \to 0$ as $k \to \infty$ since $k^3/N \to 0$ as $N \to \infty$, and the remainder of the argument is exactly the same as Theorem 3.1 in Anderson *et al.* (1999). Using the inequality $\mathrm{Var}(X+Y) \le 2(\mathrm{Var}(X) + \mathrm{Var}(Y))$ we obtain

$$E(J_3^2) = \left(\frac{N-k}{k}\right) E\left[\left\|\frac{1}{N-k}\sum_{j=k}^{N-1} X_j^{(i)}(k)\varepsilon_{(j-k)\nu+i+k,k}\right\|^2\right] \le 2(T_1 + T_2),$$

where

$$T_1 = \left(\frac{N-k}{k}\right) E\left[\left\|\frac{1}{N-k}\sum_{j=k}^{N-1} X_j^{(i)}(k)\varepsilon_{(j-k)\nu+i+k}\right\|^2\right]$$

$$T_2 = \left(\frac{N-k}{k}\right) E\left[\left\|\frac{1}{N-k}\sum_{j=k}^{N-1} X_j^{(i)}(k)(\varepsilon_{(j-k)\nu+i+k,k} - \varepsilon_{(j-k)\nu+i+k})\right\|^2\right].$$

Lemma 7 implies that

$$T_2 \le A(N-k)\max_{\ell}\sum_{j=1}^{\infty}\pi_\ell(k+j)^2$$

and

$$\begin{aligned}
T_1 &= k^{-1}(N-k)^{-1}\sum_{t=0}^{k-1} E\left[\left(\sum_{j=k}^{N-1} X_{(j-k)\nu+i+t}\varepsilon_{(j-k)\nu+i+k}\right)^2\right] \\
&= k^{-1}(N-k)^{-1}\sum_{t=0}^{k-1} E\left[\left(\sum_{j=k}^{N-1} X_{(j-k)\nu+i+t}\varepsilon_{(j-k)\nu+i+k}\right)\left(\sum_{r=k}^{N-1} X_{(r-k)\nu+i+t}\varepsilon_{(r-k)\nu+i+k}\right)\right] \\
&= k^{-1}(N-k)^{-1}\sum_{t=0}^{k-1}\sum_{j=k}^{N-1}\sum_{r=k}^{N-1} E(X_{(j-k)\nu+i+t}X_{(r-k)\nu+i+t}\varepsilon_{(j-k)\nu+i+k}\varepsilon_{(r-k)\nu+i+k}) \\
&= k^{-1}(N-k)^{-1}\sum_{t=0}^{k-1}\sum_{r=k}^{N-1} E(X^2_{(j-k)\nu+i+t})E(\varepsilon^2_{(j-k)\nu+i+k}) \\
&= k^{-1}\sum_{t=0}^{k-1}\gamma_{i+t}(0)\cdot(N-k)^{-1}\sum_{r=k}^{N-1}\sigma^2_{i+k}
\end{aligned}$$

so that $T_1 \le D = \max_i(\gamma_i(0)) \cdot \max_i(\sigma_i^2)$. Hence

$$E(J_3^2) \le 2D + 2A(N-k)\max_{\ell}\sum_{j=1}^{\infty}\pi_\ell(k+j)^2,$$

where $(N-k)\max_\ell \sum_{j=1}^{\infty}\pi_\ell(k+j)^2 \to 0$ in view of (13), so that J_3 is bounded in probability. Since $J_2 \to 0$ in probability, it follows that $I_1 \to 0$ in probability.

Apply Lemma 6 with $\ell = i + k$ and $u_{(t-k)\nu+i} = b(k)'\Gamma_{k,i}^{-1}X_t^{(i)}(k)$ to see that

$$E(I_2^2) \le 4\eta(N-k)C^2B\max_{\ell}\sum_{j=1}^{\infty}\pi_\ell(k+j)^2,$$

where $C^2 = \max_i \sum_{k=0}^{\infty} c_i(k)^2$ comes from the representation $u_{(t-k)v+i} = \sum_k c_i(k)\varepsilon_{(t-k)v+i-k}$ so that C^2 is finite if and only if $\mathrm{Var}(u_{(t-k)v+i}) < \infty$. Since $\Gamma_{k,i} = E[X_t^{(i)}(k)X_t^{(i)}(k)']$ we have $\mathrm{Var}(u_{(t-k)v+i}) = b(k)'\Gamma_{k,i}^{-1}E[X_t^{(i)}(k)X_t^{(i)}(k)']\Gamma_{k,i}^{-1}b(k) = b(k)'\Gamma_{k,i}^{-1}b(k)$ and Theorem A.1 in Anderson *et al.* (1999) shows that $\|\Gamma_{k,i}^{-1}\| \le (2\pi g)^{-1}$, when the assumed spectral bounds hold, so $C^2 < \infty$. Then it follows from (13) that $I_2 \to 0$ in probability, completing the proof of Lemma 8.□

PROOF OF THEOREM 1. Let $X_{tm} = \sum_{j=0}^{m} \psi_t(j)\varepsilon_{t-j}$ and define $e_u(k)$ to be the k dimensional vector with 1 in the uth place and zero entries elsewhere. Let

$$t_{Nm,k}^{(\langle i-k\rangle)}(u) = (N-k)^{-1/2}e_u(k)'(\Gamma_{k,\langle i-k\rangle}^{(m)})^{-1}\sum_{j=k}^{N-1} X_{jm}^{(\langle i-k\rangle)}(k)\varepsilon_{(j-k)v+\langle i-k\rangle+k}$$

so that

$$t_{Nm,k}^{(\langle i-k\rangle)}(u) = (N-k)^{-1/2}\sum_{j=k}^{N-1} w_{uj,k}^{(\langle i-k\rangle)}, \tag{42}$$

where

$$w_{uj,k}^{(\langle i-k\rangle)} = e_u(k)'(\Gamma_{k,\langle i-k\rangle}^{(m)})^{-1}X_{jm}^{(\langle i-k\rangle)}(k)\varepsilon_{(j-k)v+\langle i-k\rangle+k}, \tag{43}$$

$X_{jm}^{(\langle i-k\rangle)}(k) = (X_{(j-k)v+\langle i-k\rangle+k-1,m}, \ldots, X_{(j-k)v+\langle i-k\rangle,m})'$ and

$$\Gamma_{k,\langle i-k\rangle}^{(m)} = E[X_{jm}^{(\langle i-k\rangle)}(k)X_{jm}^{(\langle i-k\rangle)}(k)']. \tag{44}$$

For each $0 \le r < k$, $X_{(j-k)v+\langle i-k\rangle+r,m}$ is a linear combination of $(\varepsilon_{(j-k)v+\langle i-k\rangle+r-s} : 0 \le s \le m)$. Hence we write

$$w_{uj,k}^{(\langle i-k\rangle)} = L(\varepsilon_{(j-k)v+\langle i-k\rangle-m}, \ldots, \varepsilon_{(j-k)v+\langle i-k\rangle+k-1})\varepsilon_{(j-k)v+\langle i-k\rangle+k}, \tag{45}$$

where $L(\varepsilon_1, \ldots, \varepsilon_q)$ denotes a linear combination of $\varepsilon_1, \ldots, \varepsilon_q$. Note that for i, u, k fixed, $w_{uj,k}^{(\langle i-k\rangle)}$ are identically distributed for all j since the first two terms in (43) are nonrandom and do not depend on j, while the last two terms are identically distributed for all j by definition, using the fact that X_t is periodically strictly stationary. Also, $w_{uj,k}^{(\langle i-k\rangle)}$ are uncorrelated since $E[w_{uj,k}^{(\langle i-k\rangle)}w_{vj',k}^{(\langle i'-k\rangle)}] = 0$ unless $(j-k)v + \langle i-k\rangle + k = (j'-k)v + \langle i'-k\rangle + k$, which requires $i = i' \bmod v$. Since $0 \le i < v$, this implies that $i = i'$. Then $E[w_{uj,k}^{(\langle i-k\rangle)}w_{vj',k}^{(\langle i-k\rangle)}] = 0$ if $j \ne j'$ and otherwise

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

from (44). It follows immediately from (42) that the covariance matrix of the vector

$$t_{Nm,k} = (t_{Nm,k}^{(\langle i-k\rangle)}(u) : 1 \le u \le D, 0 \le \langle i-k\rangle \le v-1)'$$

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

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

and $1 \le u,\ v \le D$. Note that

$$t_{Nm,k} = (t_{Nm,k}^{(\langle 0-k\rangle)}(1),\ldots,t_{Nm,k}^{(\langle 0-k\rangle)}(D),\ldots,t_{Nm,k}^{(\langle v-1-k\rangle)}(1),\ldots,t_{Nm,k}^{(\langle v-1-k\rangle)}(D))'.$$

We also have for any $\lambda \in \mathbb{R}^{Dv}$ that $\text{Var}(\lambda' t_{Nm,k}) = \lambda'\Gamma_m\lambda$. Apply Lemma 1 to the periodically stationary process X_{tm} to see that

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

as $N \to \infty$, where

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

and $\varepsilon_t = \sum_{j=0}^{\infty}\pi_{t,m}(j)X_{t-j,m}$ (with $\pi_{t,m}(0)=1$). Since $\Gamma_m \to \Lambda_m$ as $k \to \infty$, then

$$\lim_{N\to\infty}\text{Var}(\lambda' t_{Nm,k}) = \lambda'\Lambda_m\lambda,$$

where

$$\Lambda_m = \text{diag}(\sigma_0^2\Lambda_m^{(0)},\sigma_1^2\Lambda_m^{(1)},\ldots,\sigma_{v-1}^2\Lambda_m^{(v-1)}).$$

Next, we want to use the Lindeberg–Lyapounov central limit theorem to show that

$$\lambda' t_{Nm,k} \Rightarrow N(0,\lambda'\Lambda_m\lambda).$$

Towards this end, define the vector

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

for any $\lambda \in \mathbb{R}^{Dv}$ so that

$$\lambda' t_{Nm,k} = (N-k)^{-1/2}\sum_{j=k}^{N-1} w_{j,k}.$$

Then, $\text{Var}(w_{j,k}) = \text{Var}(\lambda' t_{Nm,k}) = \lambda'\Gamma_m\lambda$. Also, $\{w_{j,k} : j = k,\ldots,N-1\}$ are mean zero, identically distributed and uncorrelated. Now, let $K = [(k+m)/v] + 2$, where again $[\cdot]$ is the greatest integer function and let N_1 be an integer such that $N_1/(N-K) \to 0$ and $K/N_1 \to 0$ (the largest integer less than $((N-K)K)^{1/2}$ suffices). Let M_1 be the greatest integer less than or equal to $(N-K)/N_1$, so that $M_1N_1/N \to 1$ and $M_1K/N \to 0$. Define the random variables

$$
\begin{aligned}
z_{1N,k} &= (w_{k,k} + \cdots + w_{k+N_1-K-1,k})/N_1^{1/2}, \\
z_{2N,k} &= (w_{k+N_1,k} + \cdots + w_{k+2N_1-K-1,k})/N_1^{1/2} \\
&\vdots \\
z_{M_1N,k} &= (w_{k+(M_1-1)N_1,k} + \cdots + w_{k+M_1N_1-K-1,k})/N_1^{1/2}.
\end{aligned}
$$

The choice of K ensures that $z_{1N,k}, \ldots, z_{M_1N,\,k}$ are i.i.d. with mean zero and variance $s_{M_1}^2 = ((N_1 - K)/N_1)\lambda'\Gamma_m\lambda$. Hence, $M_1^{-1/2}\sum_{j=1}^{M_1} z_{jN,k}$ is also mean zero with variance $s_{M_1}^2$, where $s_{M_1}^2 \to s^2$ as $M_1 \to \infty$ and $s^2 = \lambda'\Lambda_m\lambda$. The Lindeberg–Lyapounov central limit theorem (see, e.g. Billingsley, 1968, Thm 7.3) implies that

$$
M_1^{-1/2}\sum_{j=1}^{M_1} z_{jN,k} \Rightarrow N(0, s^2) \tag{46}
$$

if we can show that

$$
s_{M_1}^{-4}\sum_{j=1}^{M_1} E\Big((M_1^{-1/2}z_{jN,k})^4\Big) \to 0
$$

as $N \to \infty$. Letting $\tilde{N}_1 = N_1 - K$, we have

$$
\begin{aligned}
E(z_{jN,k}^4) &= N_1^{-2}\Big(\tilde{N}_1 E(w_{j,k}^4) + (\tilde{N}_1{}^2 - \tilde{N}_1)E(w_{j,k}^2 w_{\ell,k}^2)\Big) \\
&\le N_1^{-2}\Big(\tilde{N}_1 E(w_{j,k}^4) + (\tilde{N}_1{}^2 - \tilde{N}_1)E(w_{j,k}^4)\Big) \\
&\le E(w_{j,k}^4),
\end{aligned}
$$

where $E(w_{j,k}^4) < \infty$. Hence,

$$
\begin{aligned}
s_{M_1}^{-4}\sum_{j=1}^{M_1} E((M_1^{-1/2}z_{jN,k})^4) &\le s_{M_1}^{-4}M_1^{-1}E(w_{j,k}^4) \\
&\sim s^{-4}M_1^{-1}E(w_{j,k}^4) \\
&\to 0
\end{aligned}
$$

as $N \to \infty$ (hence $N_1 \to \infty$ and $M_1 \to \infty$). Thus, (46) holds. Also,

$$
M_1^{-1/2}\sum_{j=1}^{M_1} z_{jN,k} - (N-k)^{-1/2}\sum_{j=k}^{N-1} w_{j,k}
$$

has limiting variance 0, since its variance is no more than

$$\frac{(M_1+1)K(\mathrm{Var}(w_{t,k}))}{N-k} \sim \frac{(M_1+1)K(\mathrm{Var}(w_{t,k}))}{N},$$

which approaches 0 as $N \to \infty$. We can therefore conclude that

$$\lambda' t_{Nm,k} \Rightarrow N(0, s^2)$$

recalling that $s^2 = \lambda'\Lambda_m\lambda$. Now, an application of the Cramer–Wold device yields

$$t_{Nm,k} \Rightarrow N(0, \Lambda_m). \tag{47}$$

Next we want to show that $\Lambda_m \to \Lambda$ as $m \to \infty$. This is equivalent to showing that $\Lambda_m^{(i)} \to \Lambda^{(i)}$ as $m \to \infty$. Recall that $\psi_t(0) = 1$ and write

$$\begin{aligned} X_t &= \sum_{\ell=0}^{\infty} \psi_t(\ell)\varepsilon_{t-\ell} \\ &= \sum_{\ell=0}^{\infty} \psi_t(\ell) \sum_{j=0}^{\infty} \pi_{t-\ell}(j) X_{t-\ell-j} \\ &= \varepsilon_t + \sum_{r=1}^{\infty} \left(\sum_{\ell=1}^{r} \psi_t(\ell)\pi_{t-\ell}(r-\ell) \right) X_{t-r} \end{aligned}$$

so that $\pi_t(r) = -\sum_{\ell=1}^{r} \psi_t(\ell)\pi_{t-\ell}(r-\ell)$ for $r \geq 1$. Similarly

$$\begin{aligned} X_{tm} &= \sum_{\ell=0}^{m} \psi_t(\ell) \sum_{j=0}^{\infty} \pi_{t-\ell,m}(j) X_{t-\ell-j,m} \\ &= \varepsilon_t + \sum_{r=1}^{\infty} \left(\sum_{\ell=1}^{\min(r,m)} \psi_t(\ell)\pi_{t-\ell,m}(r-\ell) \right) X_{t-r,m} \end{aligned}$$

so that $\pi_{t,m}(r) = -\sum_{\ell=1}^{\min(r,m)} \psi_t(\ell)\pi_{t-\ell,m}(r-\ell)$ for $r \geq 1$. Apply these two formulae recursively to see that $\pi_{t,m}(r) = \pi_t(r)$ for $1 \leq r \leq m$. Then we actually have $\Lambda_m^{(i)} = \Lambda^{(i)}$ for all m sufficiently large.

Now we want to show that

$$\lim_{m\to\infty} \limsup_{N\to\infty} P(|t_{N,k}^{(\langle i-k\rangle)}(u) - t_{Nm,k}^{(\langle i-k\rangle)}(u)| > \delta) = 0 \tag{48}$$

for any $\delta > 0$, where

$$t_{N,k}^{(\langle i-k\rangle)}(u) = (N-k)^{-1/2} e_u(k)' \Gamma_{k,\langle i-k\rangle}^{-1} \sum_{j=k}^{N-1} X_j^{(\langle i-k\rangle)}(k)\varepsilon_{(j-k)v+\langle i-k\rangle+k}.$$

Apply the Chebyshev inequality to see that the probability in (48) is bounded above by

$$\delta^{-2}\mathrm{Var}(t_{N,k}^{(\langle i-k\rangle)}(u) - t_{Nm,k}^{(\langle i-k\rangle)}(u)) = \delta^{-2}\mathrm{Var}\left[(N-k)^{-1/2} \sum_{j=k}^{N-1} A_j \varepsilon_{(j-k)v+\langle i-k\rangle+k}\right],$$

where $A_j = e_u(k)'\Gamma^{-1}_{k,\langle i-k\rangle}X_j^{(\langle i-k\rangle)}(k) - e_u(k)'(\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}X_{jm}^{(\langle i-k\rangle)}(k)$. Since the summands are uncorrelated, we have

$$\begin{aligned}\mathrm{Var}(t_{N,k}^{(\langle i-k\rangle)}(u) - t_{Nm,k}^{(\langle i-k\rangle)}(u)) &= (N-k)^{-1}\sum_{j=k}^{N-1} E\left[A_j^2\varepsilon^2_{(j-k)v+\langle i-k\rangle+k}\right],\\ &= \sigma_i^2(N-k)^{-1}\sum_{j=k}^{N-1}E(A_j^2)\\ &= \sigma_i^2 E(A_j^2)\end{aligned}$$

because each A_j has the same distribution. Write $E(A_j^2) = \mathrm{Var}(I_1 + I_2)$, where

$$\begin{aligned}I_1 &= e_u(k)'\Gamma^{-1}_{k,\langle i-k\rangle}\left[X_j^{(\langle i-k\rangle)}(k) - X_{jm}^{(\langle i-k\rangle)}(k)\right],\\ I_2 &= e_u(k)'\left[\Gamma^{-1}_{k,\langle i-k\rangle} - (\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}\right]X_{jm}^{(\langle i-k\rangle)}(k),\end{aligned}$$

so that $\mathrm{Var}(I_1) = e_u(k)'\Gamma^{-1}_{k,\langle i-k\rangle}R_{k,\langle i-k\rangle}\Gamma_{k,\langle i-k\rangle}e_u(k)$, where $R_{k,\langle i-k\rangle}$ is the covariance matrix of the random vector $(X_{tv+\langle i-k\rangle+k-1} - X_{tv+\langle i-k\rangle+k-1,\,m}, \ldots, X_{tv+\langle i-k\rangle} - X_{tv+\langle i-k\rangle,\,m})'$. Since $\|e_u(k)\| = 1$, Theorem A.1 in Anderson *et al.* (1999) implies that $\mathrm{Var}(I_1) \le (G/g)\|R_{k,\langle i-k\rangle}\|$. In order to show that $\|R_{k,\langle i-k\rangle}\| \to 0$ as $m \to \infty$, we consider the spectral density matrix $f_d(\lambda)$ of the vector moving average $W_t = (X_{tv+v-1} - X_{tv+v-1,m}, \ldots, X_{tv} - X_{tv,m})'$. If we let $Y_t = \sum_{\ell=0}^{\infty}\Psi_\ell Z_{t-\ell}$, where $Y_t = (X_{tv+v-1}, \ldots, X_{tv})'$, $Z_t = (\varepsilon_{tv+v-1}, \ldots, \varepsilon_{tv})'$, and $(\Psi_\ell)_{ij} = \psi_{v-1-i}(\ell v - i + j)$ and $Y_{tm} = (X_{tv+v-1,m}, \ldots, X_{tv,m})'$, then $W_t = Y_t - Y_{tm}$. Since $X_t - X_{tm} = \sum_{j>m}\psi_t(j)\varepsilon_{t-j}$, the moving average

$$W_t = \sum_{\ell=\lceil\frac{m+2}{v}\rceil-1}^{\infty}\tilde{\Psi}_\ell Z_{t-\ell},$$

where $\lceil x\rceil$ is the smallest integer greater than or equal to x and $(\tilde{\Psi}_\ell)_{ij} = \psi_{v-1-i}(\ell v - i + j)$ except that we zero out the entries with $\ell v - i + j \le m$. Then the spectral density matrix

$$f_d(\lambda) = \frac{1}{2\pi}\left(\sum_{\ell=\lceil\frac{m+2}{v}\rceil-1}^{\infty}\tilde{\Psi}_\ell e^{i\lambda\ell}\right)\Sigma\left(\sum_{\ell=\lceil\frac{m+2}{v}\rceil-1}^{\infty}\tilde{\Psi}_\ell e^{i\lambda\ell}\right)',$$

where Σ is the covariance matrix of Z_t. As in the proof of Theorem A.1 in Anderson *et al.* (1999), we now define $\Gamma_w(h) = \mathrm{Cov}(W_t, W_{t+h})$, $W = (W_{n-1}, \ldots, W_0)'$ and

$$\Gamma = \mathrm{Cov}(W, W) = [\Gamma_w(i-j)]_{i,j=0}^{n} = \mathrm{Cov}(X_{nv-1} - X_{nv-1,m}, \ldots, X_0 - X_{0,m})'.$$

For fixed i and k, let $n = [(k + \langle i-k\rangle)/v] + 1$. Then $R_{k,\langle i-k\rangle}$ is a submatrix of $\Gamma = R_{nv,0}$. Fix an arbitrary vector $y = (y_0, \ldots, y_{n-1})'$ in $\mathbb{R}^{nv}$, whose jth entry $y_j \in \mathbb{R}^v$. Then

$$
\begin{aligned}
y'\Gamma y &= \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} y_j'\Gamma_w(j-k)y_k \\
&= \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} y_j'\left(\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda(j-k)} f_d(\lambda)\mathrm{d}\lambda\right)y_k \\
&= \int_{-\pi}^{\pi}\left(\sum_{j=0}^{n-1}\mathrm{e}^{\mathrm{i}\lambda j}y_j'\right)' f_d(\lambda)\left(\sum_{k=0}^{n-1}\mathrm{e}^{-\mathrm{i}\lambda k}y_k\right)\mathrm{d}\lambda \\
&\leq \int_{-\pi}^{\pi}\delta_m\left(\sum_{j=0}^{v-1}y_j'y_j\right)\mathrm{d}\lambda = 2\pi\delta_m y'y
\end{aligned}
$$

so that $\|\Gamma\| \leq 2\pi\delta_m$, where δ_m is the largest modulus of the elements of $f_d(\lambda)$. But $\delta_m \to 0$ in view of the causality condition (2), hence $\|R_{k,\langle i-k\rangle}\| \leq \|\Gamma\| \leq 2\pi\delta_m \to 0$, which implies that $\mathrm{Var}(I_1) \to 0$ as well.

Next write $\mathrm{Var}(I_2) = e_u(k)'[\Gamma^{-1}_{k,\langle i-k\rangle} - (\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}]\Gamma^{(m)}_{k,\langle i-k\rangle}[\Gamma^{-1}_{k,\langle i-k\rangle} - (\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}]$ $e_u(k)$ and recall that $\Gamma^{(m)}_{k,\langle i-k\rangle}$ is the covariance matrix of $(X_{tv+\langle i-k\rangle\ +\ k-1,m}, \ldots, X_{tv+\langle i-k\rangle,m})'$ for each t. Then as in the preceeding paragraph write

$$
Y_{tm} = (X_{tv+v-1,m}, \ldots, X_{tv,m}) = \sum_{\ell=0}^{[(m-1)/v]+1} \bar{\Psi}_\ell Z_{t-\ell},
$$

where $(\bar{\Psi}_\ell)_{ij} = \psi_{v-1-i}(\ell v - i + j)$ except that we zero out the entries with $\ell v - i + j > m$, and let $f_m(\lambda)$ denote the spectral density matrix of Y_{tm}. Let $B(Z_t) = Z_{t-1}$ denote the backward shift operator, and write $Y_t = \Psi(B)Z_t$ and similarly $Y_{tm} = \bar{\Psi}(B)Z_t$. Then $f(\lambda) = (2\pi)^{-1}\Psi(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\Psi'(\mathrm{e}^{-\mathrm{i}\lambda})$ and $f_m(\lambda) = (2\pi)^{-1}\bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\bar{\Psi}'(\mathrm{e}^{-\mathrm{i}\lambda})$, where as before Σ is the covariance matrix of Z_t. Using the Frobenius norm $\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2}$, we have

$$
\begin{aligned}
\|f(\lambda) - f_m(\lambda)\|_F &= (4\pi^2)^{-1}\|\Psi(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\Psi'(\mathrm{e}^{-\mathrm{i}\lambda}) - \bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\bar{\Psi}'(\mathrm{e}^{-\mathrm{i}\lambda})\|_F \\
&\leq (4\pi^2)^{-1}\|\Psi(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\Psi'(\mathrm{e}^{-\mathrm{i}\lambda}) - \bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\Psi'(\mathrm{e}^{-\mathrm{i}\lambda})\|_F \\
&\quad + (4\pi^2)^{-1}\|\bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\Psi'(\mathrm{e}^{-\mathrm{i}\lambda}) - \bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\Sigma\bar{\Psi}'(\mathrm{e}^{-\mathrm{i}\lambda})\|_F \\
&\leq (4\pi^2)^{-1}\|\Psi(\mathrm{e}^{\mathrm{i}\lambda}) - \bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\|_F \cdot \|\Sigma\|_F \cdot \|\Psi'(\mathrm{e}^{-\mathrm{i}\lambda})\|_F \\
&\quad + (4\pi^2)^{-1}\|\bar{\Psi}(\mathrm{e}^{\mathrm{i}\lambda})\|_F \cdot \|\Sigma\|_F \cdot \|\Psi'(\mathrm{e}^{-\mathrm{i}\lambda}) - \bar{\Psi}'(\mathrm{e}^{-\mathrm{i}\lambda})\|_F
\end{aligned}
$$

where $\|\Psi'(\mathrm{e}^{\mathrm{i}\lambda}) - \bar{\Psi}'(\mathrm{e}^{\mathrm{i}\lambda})\|_F \to 0$ as $m \to \infty$ in view of the causality condition (2) and the remaining norms are bounded independent of m, so that $|z'[f(\lambda) - f_m(\lambda)]z| \leq \delta(m)z'z$ for any $z \in \mathbb{R}^v$, where $\delta(m) \to 0$ as $m \to \infty$. Now Theorem A.1 in Anderson *et al.* (1999) yields

$$
z'(g - \delta(m))z \leq z'f(\lambda)z - z'\delta(m)z \leq z'f_m(\lambda)z \leq z'f(\lambda)z + z'\delta(m)z \leq z'(G + \delta(m))z.
$$

Let $\Gamma(h) = \mathrm{Cov}(Y_t, Y_{t+h})$, $Y = (Y_{n-1}, \ldots, Y_0)'$ and $\Gamma_{nv,0} = \mathrm{Cov}(Y, Y) = [\Gamma(i-j)]_{i,j=0}^{n-1}$, the covariance matrix of $(X_{nv-1}, \ldots, X_0)'$. For fixed i and k, let

$n = [(k + \langle i-k \rangle)/v] + 1$ as before. Similarly, let $\Gamma_m(h) = \mathrm{Cov}(Y_{tm}, Y_{t+h,m})$, $Y_m = (Y_{n-1,m}, \ldots, Y_{0m})'$ and $\Gamma^{(m)}_{nv,0} = \mathrm{Cov}(Y_m, Y_m)$, the covariance matrix of $(X_{nv-1,m}, \ldots, X_{0m})'$. Since $\Gamma_{k,\langle i-k \rangle}$ is a submatrix of $\Gamma_{nv,0}$ and $\Gamma^{(m)}_{k,\langle i-k \rangle}$ is a submatrix of $\Gamma^{(m)}_{nv,0}$, it follows that

$$\begin{aligned}\|\Gamma^{(m)}_{k,\langle i-k\rangle} - \Gamma_{k,\langle i-k\rangle}\| &\le \|\Gamma^{(m)}_{nv,0} - \Gamma_{nv,0}\| \\ &= \sup_{\|y\|=1} |y'(\Gamma^{(m)}_{nv,0} - \Gamma_{nv,0})y| \\ &= \sup_{\|y\|=1} \left| \sum_{j=0}^{n-1} \sum_{\ell=0}^{n-1} y_j'(\Gamma_m(j-\ell) - \Gamma(j-\ell))y_\ell \right| \\ &= \sup_{\|y\|=1} \left| \sum_{j=0}^{n-1} \sum_{\ell=0}^{n-1} y_j' \left(\int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\lambda(j-\ell)} (f_m(\lambda) - f(\lambda)) \mathrm{d}\lambda \right) y_\ell \right| \\ &= \sup_{\|y\|=1} \left| \int_{-\pi}^{\pi} \left(\sum_{j=0}^{n-1} \mathrm{e}^{\mathrm{i}\lambda j} y_j \right)' (f_m(\lambda) - f(\lambda)) \left(\sum_{\ell=0}^{n-1} \mathrm{e}^{-\mathrm{i}\lambda\ell} y_\ell \right) \mathrm{d}\lambda \right|\end{aligned}$$

is bounded above by $2\pi\delta(m)$, hence

$$\begin{aligned}\|\Gamma^{-1}_{k,\langle i-k\rangle} - (\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}\| &= \|\Gamma^{-1}_{k,\langle i-k\rangle}(\Gamma^{(m)}_{k,\langle i-k\rangle} - \Gamma_{k,\langle i-k\rangle})(\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}\| \\ &\le \|\Gamma^{-1}_{k,\langle i-k\rangle}\| \cdot \|\Gamma^{(m)}_{k,\langle i-k\rangle} - \Gamma_{k,\langle i-k\rangle}\| \cdot \|(\Gamma^{(m)}_{k,\langle i-k\rangle})^{-1}\| \\ &\le \frac{1}{2\pi g} \cdot 2\pi\delta(m) \cdot \frac{1}{2\pi(g - \delta(m))}.\end{aligned}$$

Then

$$\mathrm{Var}(I_2) \le \left(\frac{1}{2\pi g} \cdot 2\pi\delta(m) \cdot \frac{1}{2\pi(g-\delta(m))} \right)^2 2\pi(G + \delta(m)) \to 0$$

as $m \to \infty$, and so $\mathrm{Var}(I_1 + I_2) \to 0$ as well, which proves (48). Together with (47), the fact that $\Lambda^{(i)}_m \to \Lambda^{(i)}$ as $m \to \infty$, and Theorem 4.2 in Billingsley (1968), this proves that

$$t_{N,k} \Rightarrow N(0, \Lambda) \tag{49}$$

as $N \to \infty$, where

$$t_{N,k} = (t^{(\langle i-k\rangle)}_{N,k}(u) : 1 \le u \le D, 0 \le \langle i-k \rangle \le v-1)'.$$

Applying Lemma 8 yields $(N-k)^{1/2} e_u(k)'(\pi^{(\langle i-k\rangle)}_k + f^{(\langle i-k\rangle)}_k) - t^{(\langle i-k\rangle)}_{N,k}(u) \to 0$ in probability. Note that $\pi^{(\langle i-k\rangle)}_k = (\pi_i(1), \ldots, \pi_i(k))'$ and $f^{(\langle i-k\rangle)}_k = (f^{(\langle i-k\rangle)}_{k,1}, \ldots, f^{(\langle i-k\rangle)}_{k,k})$. Combining statement (49) with $e_u(k)'(\pi^{(\langle i-k\rangle)}_k + f^{(\langle i-k\rangle)}_k) = \pi_i(u) + f^{(\langle i-k\rangle)}_{k,u}$ and the fact that $(N-k)/N \to 1$, implies that

$$N^{1/2}(\pi_i(u) + f_{k,u}^{(\langle i-k\rangle)}) - t_{N,k}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0 \tag{50}$$

as $N \to \infty$, where $f_{k,u}^{(\langle i-k\rangle)}$ is defined as in (30). Then (14) holds with $\hat{\phi}_{k,u}^{\langle i-k\rangle}$ replaced by $f_{k,u}^{\langle i-k\rangle}$.

Finally, we must show that

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

Write

$$\begin{aligned} N^{1/2}\|f_{k,u}^{(\langle i-k\rangle)} - \hat{\phi}_{k,u}^{(\langle i-k\rangle)}\| &= N^{1/2}\|(\hat{R}_{k,\langle i-k\rangle})^{-1}\hat{r}_k^{(\langle i-k\rangle)} - \hat{\Gamma}_{k,\langle i-k\rangle}^{-1}\hat{\gamma}_k^{(\langle i-k\rangle)}\| \\ &\leq N^{1/2}\|(\hat{R}_{k,\langle i-k\rangle})^{-1} - \hat{\Gamma}_{k,\langle i-k\rangle}^{-1}\| \cdot \|\hat{r}_k^{(\langle i-k\rangle)}\| \\ &\quad + N^{1/2}\|\hat{\Gamma}_{k,\langle i-k\rangle}^{-1}\| \cdot \|\hat{r}_k^{(\langle i-k\rangle)} - \hat{\gamma}_k^{(\langle i-k\rangle)}\| \end{aligned}$$

and recall that $(N-k)\mathrm{Var}(\hat{s}_{k,j}^{(\langle i-k\rangle)}) \leq M$, where M is given by (41), so that $N^{1/2}\|\hat{r}_k^{(\langle i-k\rangle)}\|$ is bounded in probability. Also, recall from the proof of Lemma 8 that $k^{1/2}\|\hat{R}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\| \to 0$ in probability, so the same is true with i replaced with $\langle i-k\rangle$. A very similar argument yields $k^{1/2}\|\hat{\Gamma}_{k,\langle i-k\rangle}^{-1} - \Gamma_{k,\langle i-k\rangle}^{-1}\| \to 0$ in probability, so that $k^{1/2}\|\hat{R}_{k,\langle i-k\rangle}^{-1} - \hat{\Gamma}_{k,\langle i-k\rangle}^{-1}\| \to 0$ in probability as well. Next observe that $\|\hat{\Gamma}_{k,\langle i-k\rangle}^{-1}\| \leq \|\hat{\Gamma}_{k,\langle i-k\rangle}^{-1} - \Gamma_{k,\langle i-k\rangle}^{-1}\| + \|\Gamma_{k,\langle i-k\rangle}^{-1}\|$, where $\|\Gamma_{k,\langle i-k\rangle}^{-1}\|$ is uniformly bounded by Theorem A.1 of Anderson *et al.* (1999). Since $\hat{s}_{k,j}^{(\langle i-k\rangle)}$ and $\hat{\gamma}_i(j-k)$ differ by only k terms and a factor $(N-k)/N$, it is easy to check that $N^{1/2}\|\hat{r}_k^{(\langle i-k\rangle)} - \hat{\gamma}_k^{(\langle i-k\rangle)}\| \to 0$ in probability. It follows that $N^{1/2}(f_{k,u}^{(\langle i-k\rangle)} - \hat{\phi}_{k,u}^{(\langle i-k\rangle)}) \to 0$ in probability, which completes the proof of Theorem 1. □

Proof of Theorem 2. From the two representations of $\hat{X}_{i+k}^{(i)}$ given by (4) and (8), it follows that

$$\theta_{k,j}^{(\langle i-k\rangle)} = \sum_{\ell=1}^{j} \phi_{k,\ell}^{(\langle i-k\rangle)}\theta_{k-\ell,j-\ell}^{(\langle i-k\rangle)} \tag{52}$$

for $j = 1, \ldots, k$ if we define $\theta_{k-j,0}^{(\langle i-k\rangle)} = 1$ and replace i with $\langle i-k\rangle$. Equation (52) can be modified and written as

$$\begin{pmatrix} \theta_{k,1}^{(\langle i-k\rangle)} \\ \theta_{k,2}^{(\langle i-k\rangle)} \\ \theta_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \theta_{k,D}^{(\langle i-k\rangle)} \end{pmatrix} = R_k^{(\langle i-k\rangle)} \begin{pmatrix} \phi_{k,1}^{(\langle i-k\rangle)} \\ \phi_{k,2}^{(\langle i-k\rangle)} \\ \phi_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \phi_{k,D}^{(\langle i-k\rangle)} \end{pmatrix}, \tag{53}$$

where

$$R_k^{(\langle i-k\rangle)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \theta_{k-1,1}^{(\langle i-k\rangle)} & 1 & 0 & \cdots & 0 & 0 \\ \theta_{k-1,2}^{(\langle i-k\rangle)} & \theta_{k-2,1}^{(\langle i-k\rangle)} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ \theta_{k-1,D-1}^{(\langle i-k\rangle)} & \theta_{k-2,D-2}^{(\langle i-k\rangle)} & \cdots & & \theta_{k-D+1,1}^{(\langle i-k\rangle)} & 1 \end{pmatrix} \tag{54}$$

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

$$\begin{pmatrix} \hat{\theta}_{k,1}^{(\langle i-k\rangle)} \\ \hat{\theta}_{k,2}^{(\langle i-k\rangle)} \\ \hat{\theta}_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \hat{\theta}_{k,D}^{(\langle i-k\rangle)} \end{pmatrix} = \hat{R}_k^{(\langle i-k\rangle)} \begin{pmatrix} \hat{\phi}_{k,1}^{(\langle i-k\rangle)} \\ \hat{\phi}_{k,2}^{(\langle i-k\rangle)} \\ \hat{\phi}_{k,3}^{(\langle i-k\rangle)} \\ \vdots \\ \hat{\phi}_{k,D}^{(\langle i-k\rangle)} \end{pmatrix}, \tag{55}$$

where $\hat{R}_k^{(\langle i-k\rangle)}$ is defined as in (54) with $\hat{\theta}_{k,u}^{(\langle i-k\rangle)}$ replacing $\theta_{k,u}^{(\langle i-k\rangle)}$. From (11) 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,

$$\hat{R}_k^{(\langle i-k\rangle)} \xrightarrow{P} R^{(i)}, \tag{56}$$

where

$$R^{(i)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \psi_{i-1}(1) & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \psi_{i-1}(D-1) & \psi_{i-2}(D-2) & \cdots & \psi_{i-D+1}(1) & 1 \end{pmatrix}. \tag{57}$$

We have

$$\hat{\theta}^{(\langle i-k\rangle)} - \theta^{(\langle i-k\rangle)} = \hat{R}_k^{(\langle i-k\rangle)}(\hat{\phi}^{(\langle i-k\rangle)} - \phi^{(\langle i-k\rangle)}) + (\hat{R}_k^{(\langle i-k\rangle)} - R_k^{(\langle i-k\rangle)})\phi^{(\langle i-k\rangle)}, \tag{58}$$

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

$$\begin{aligned} (\hat{R}_k^{(\langle i-k\rangle)} - R_k^{(\langle i-k\rangle)})\phi^{(\langle i-k\rangle)} = {} & (\hat{R}_k^{(\langle i-k\rangle)} - \hat{R}_k^{(\langle i-k\rangle)*})\phi^{(\langle i-k\rangle)} \\ & + (\hat{R}_k^{(\langle i-k\rangle)*} - R_k^{(\langle i-k\rangle)*})\phi^{(\langle i-k\rangle)} \\ & + (R_k^{(\langle i-k\rangle)*} - R_k^{(\langle i-k\rangle)})\phi^{(\langle i-k\rangle)}, \end{aligned} \tag{59}$$

where

$$R_k^{(\langle i-k\rangle)*} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ \theta_{k,1}^{(\langle i-1-k\rangle)} & 1 & 0 & \cdots & 0 & 0 \\ \theta_{k,2}^{(\langle i-1-k\rangle)} & \theta_{k,1}^{(\langle i-2-k\rangle)} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ \theta_{k,D-1}^{(\langle i-1-k\rangle)} & \theta_{k,D-2}^{(\langle i-2-k\rangle)} & \cdots & & \theta_{k,1}^{(\langle i-D+1-k\rangle)} & 1 \end{pmatrix} \tag{60}$$

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

$$N^{1/2}(\theta_{k,u}^{(\langle i-\ell-k\rangle)} - \theta_{k-\ell,u}^{(\langle i-k\rangle)}) \to 0 \tag{61}$$

and

$$N^{1/2}(\hat{\theta}_{k,u}^{(\langle i-\ell-k\rangle)} - \hat{\theta}_{k-\ell,u}^{(\langle i-k\rangle)}) \xrightarrow{P} 0 \tag{62}$$

for $\ell = 1, \ldots, D-1$ and $u = 1, \ldots, D$. Using estimates from the proof of Anderson *et al*. (1999, Cor. 2.2.4) and condition (13) of Theorem 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

$$N^{1/2}(\phi_{k,u}^{(\langle i-\ell-k\rangle)} + \pi_{i-\ell}(u)) \to 0 \tag{63}$$

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,

$$N^{1/2}(\phi_{k-\ell,u}^{(\langle i-k\rangle)} + \pi_{i-\ell}(u)) \to 0 \tag{64}$$

as $k \to \infty$. Subtracting (64) from (63) yields

$$N^{1/2}(\phi_{k,u}^{(\langle i-\ell-k\rangle)} - \phi_{k-\ell,u}^{(\langle i-k\rangle)}) \to 0, \tag{65}$$

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 (61) with $u = 1$. The cases $u = 2, \ldots, D$ follow iteratively using (10), (53), and (65). Thus, (61) is established. To prove (62), we need Lemma 9. □

LEMMA 9. *For all* $\ell = 1, \ldots, u-1$ *and* $u = 1, \ldots, k$, *we have*

$$N^{1/2}(\hat{\phi}_{k,u}^{(\langle i-\ell-k\rangle)} - \hat{\phi}_{k-\ell,u}^{(\langle i-k\rangle)}) \xrightarrow{P} 0. \tag{66}$$

PROOF. Starting from (42), we need to show that $t_{Nm,k}^{(\langle i-\ell-k\rangle)}(u) - t_{Nm,k-\ell}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0$, where

$$
\begin{aligned}
t_{Nm,k}^{(\langle i-\ell-k\rangle)}(u) &= (N-k)^{-1/2}\sum_{j=k}^{N-1} w_{uj,k}^{(\langle i-\ell-k\rangle)} \\
t_{Nm,k-\ell}^{(\langle i-k\rangle)}(u) &= (N-k+\ell)^{-1/2}\sum_{j=k-\ell}^{N-1} w_{uj,k-\ell}^{(\langle i-k\rangle)} \\
w_{uj,k}^{(\langle i-\ell-k\rangle)} &= e_u(k)'(\Gamma_{k,\langle i-\ell-k\rangle}^{(m)})^{-1}X_{jm}^{(\langle i-\ell-k\rangle)}(k)\varepsilon_{(j-k)v+\langle i-\ell-k\rangle+k} \\
w_{uj,k-\ell}^{(\langle i-k\rangle)} &= e_u(k-\ell)'(\Gamma_{k-\ell,\langle i-k\rangle}^{(m)})^{-1}X_{jm}^{(\langle i-k\rangle)}(k-\ell)\varepsilon_{(j-k+\ell)v+\langle i-k\rangle+k-\ell}
\end{aligned}
$$

and

$$
\begin{aligned}
X_{jm}^{(\langle i-\ell-k\rangle)}(k) &= (X_{(j-k)v+\langle i-\ell-k\rangle+r,m} : r = 0,\ldots,k-1)', \\
X_{jm}^{(\langle i-k\rangle)}(k-\ell) &= (X_{(j-k+\ell)v+\langle i-k\rangle+r,m} : r = 0,\ldots,k-\ell-1)',
\end{aligned}
$$

$\Gamma_{k,\langle i-\ell-k\rangle}^{(m)}$ is the covariance matrix of $X_{jm}^{(\langle i-\ell-k\rangle)}(k)$, and $\Gamma_{k-\ell,\langle i-k\rangle}^{(m)}$ is the covariance matrix of $X_{jm}^{(\langle i-k\rangle)}(k-\ell)$. Note that $E(w_{uj,k}^{(\langle i-\ell-k\rangle)} w_{uj',k-\ell}^{(\langle i-k\rangle)}) = 0$ unless

$$
j' = j - \ell + (\langle i-\ell-k\rangle + \ell - \langle i-k\rangle)/v, \tag{67}
$$

which is always an integer. For each j, there is at most one j' that satisfies (67) for $j' \in \{k-\ell,\ldots,N-1\}$. If such a j' exists then

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

where $C = E(X_{jm}^{(\langle i-\ell-k\rangle)}(k)X_{j'm}^{(\langle i-k\rangle)}(k-\ell)')$. Note that the $(k-\ell)$-dimensional vector $X_{j'm}^{(\langle i-k\rangle)}(k-\ell)$ is just the first $(k-\ell)$ of the k entries of the vector $X_{jm}^{(\langle i-\ell-k\rangle)}(k)$. Hence, the matrix C is just $\Gamma_{k,\langle i-\ell-k\rangle}^{(m)}$ with the last ℓ columns deleted. Then $(\Gamma_{k,\langle i-\ell-k\rangle}^{(m)})^{-1}C = I(k,\ell)$, 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 $e_u(k)'(\Gamma_{k,\langle i-\ell-k\rangle}^{(m)})^{-1}C = e_u(k-\ell)'$. Then

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

Consequently,

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

by Lemma 1 applied to the finite moving average. Thus,

$$\mathrm{Var}(t_{Nm,k}^{(\langle i-\ell-k\rangle)}(u) - t_{Nm,k-\ell}^{(\langle i-k\rangle)}(u)) \le (N-k)^{-1}\mathrm{Var}(w_{uj,k}^{(\langle i-\ell-k\rangle)} - w_{uj',k-\ell}^{(\langle i-k\rangle)})(N-k) \to 0$$

since for each $j = k, \ldots, N-1$ there is at most one j' satisfying (67) along with $j' = k-\ell, \ldots, N-1$. Then Chebyshev's inequality shows that $t_{Nm,k}^{(\langle i-\ell-k\rangle)}(u) - t_{Nm,k-\ell}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0$. Since $t_{Nm,k}^{(\langle i-k\rangle)}(u) - t_{N,k}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0$ we also get $t_{N,k}^{(\langle i-k\rangle)}(u) - t_{N,k-\ell}^{(\langle i-k\rangle)}(u) \xrightarrow{P} 0$. Now the Lemma follows easily using (50) along with (51).

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 (62) with $u = 1$. The cases $u = 2, \ldots, D$ follow iteratively using (55), Lemma 9 and (52) with θ, ϕ replaced by $\hat{\theta}, \hat{\phi}$. Thus, (62) is established. From (58) to (62), it follows that

$$\begin{aligned}\hat{\theta}^{(\langle i-k\rangle)} - \theta^{(\langle i-k\rangle)} = {} & \hat{R}_k^{(\langle i-k\rangle)}(\hat{\phi}^{(\langle i-k\rangle)} - \phi^{(\langle i-k\rangle)}) \\ & + (\hat{R}_k^{(\langle i-k\rangle)*} - R_k^{(\langle i-k\rangle)*})\phi^{(\langle i-k\rangle)} + o_P(N^{1/2}).\end{aligned} \tag{68}$$

To accommodate the derivation of the asymptotic distribution of $\hat{\theta}^{(i)} - \theta^{(i)}$, we need to rewrite (68). Define

$$\begin{aligned}\hat{\theta} - \theta = {} & (\hat{\theta}_{k,1}^{(\langle 0-k\rangle)} - \theta_{k,1}^{(\langle 0-k\rangle)}, \ldots, \hat{\theta}_{k,D}^{(\langle 0-k\rangle)} - \theta_{k,D}^{(\langle 0-k\rangle)}, \ldots, \\ & \hat{\theta}_{k,1}^{(\langle v-1-k\rangle)} - \theta_{k,1}^{(\langle v-1-k\rangle)}, \ldots, \hat{\theta}_{k,D}^{(\langle v-1-k\rangle)} - \theta_{k,D}^{(\langle v-1-k\rangle)})' \\ & \text{and} \\ \phi = {} & (\phi_{k,1}^{(\langle 0-k\rangle)}, \ldots, \phi_{k,D}^{(\langle 0-k\rangle)}, \ldots, \phi_{k,1}^{(\langle v-1-k\rangle)}, \ldots, \phi_{k,D}^{(\langle v-1-k\rangle)})'.\end{aligned} \tag{69}$$

Using (69) we can rewrite (68) as

$$\hat{\theta} - \theta = \hat{R}_k(\hat{\phi} - \phi) + (\hat{R}_k^* - R_k^*)\phi + o_P(N^{1/2}), \tag{70}$$

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

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

and

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

noting that both R_k and R_k^* are $Dv \times Dv$ 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

$$C_k = \sum_{n=1}^{D-1} B_{n,k}\Pi^{[Dv-n(D+1)]} \tag{71}$$

and

$$B_{n,k} = \text{diag}(\overbrace{0,\ldots,0}^{n}, \overbrace{\phi_{k,n}^{(\langle 0-k\rangle)},\ldots,\phi_{k,n}^{(\langle 0-k\rangle)}}^{D-n}, \overbrace{0,\ldots,0}^{n}, \underbrace{\phi_{k,n}^{(\langle 1-k\rangle)},\ldots,\phi_{k,n}^{(\langle 1-k\rangle)}}_{D-n},\ldots,\underbrace{0,\ldots,0}_{n}, \underbrace{\phi_{k,n}^{(\langle \nu-1-k\rangle)},\ldots,\phi_{k,n}^{(\langle \nu-1-k\rangle)}}_{D-n})$$

with Π the orthogonal $D\nu \times D\nu$ cyclic permutation matrix (23). Thus, we write equation (70) as

$$\hat{\theta} - \theta = \hat{R}_k(\hat{\phi} - \phi) + C_k(\hat{\theta} - \theta) + o_P(N^{1/2}). \tag{72}$$

Then,

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

so that

$$\hat{\theta} - \theta = (I - C_k)^{-1}\hat{R}_k(\hat{\phi} - \phi) + o_P(N^{1/2}). \tag{73}$$

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 (57). Equation (56) shows that $\hat{R}_k \xrightarrow{P} R$ and then Theorem 1 along with equation (73) yield

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

where

$$V = (I - C)^{-1} R \Lambda R' [(I - C)^{-1}]' \tag{74}$$

and Λ is as in (15). Let

$$S^{(i)} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ \pi_{i-1}(1) & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ \pi_{i-1}(D-1) & \pi_{i-2}(D-2) & \cdots & \pi_{i-D+1}(1) & 1 \end{pmatrix}. \tag{75}$$

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,

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

and it immediately follows that

$$R\Lambda R' = \text{diag}(\sigma_0^2 D^{(0)}, \ldots, \sigma_{v-1}^2 D^{(v-1)}),$$

where $D^{(i)} = \text{diag}(\sigma_{i-1}^{-2}, \sigma_{i-2}^{-2}, \ldots, \sigma_{i-D}^{-2})$. Thus, eqn (74) becomes

$$V = (I - C)^{-1}\text{diag}(\sigma_0^2 D^{(0)}, \ldots, \sigma_{v-1}^2 D^{(v-1)})[(I - C)^{-1}]'. \quad (76)$$

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

$$(I - C)^{-1} = \sum_{n=0}^{D-1} E_n \Pi^{[Dv-n(D+1)]},$$

where E_n is defined in (21). Using estimates from the proof of Corollary 2.2.3 in Anderson *et al.* (1999) along with condition (17) 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 $\psi = (\psi_0(1), \ldots, \psi_0(D), \ldots, \psi_{v-1}(1), \ldots, \psi_{v-1}(D))'$. We have proved the theorem. □

ACKNOWLEDGEMENTS

Mark M. Meerschaert is partially supported by NSF Grants DES-9980484, DMS-0139927, DMS-0417869 and by the Marsden Foundation in New Zealand.

REFERENCES

Adams, G. and Goodwin, C. (1995) Parameter estimation for periodic ARMA models. *Journal of Time Series Analysis* 16, 127–45.

Anderson, P. and Vecchia, A. (1993) Asymptotic results for periodic autoregressive moving–average processes. *Journal of Time Series Analysis* 14, 1–18.

Anderson, P. and Meerschaert, M. (1997) Periodic moving averages of random variables with regularly varying tails. *The Annals of Statistics* 25, 771–85.

Anderson, P. and Meerschaert M. (1998) Modeling river flows with heavy tails. *Water Resources Research* 34, 2271–80.

Anderson, P., Meerschaert, M. and Vecchia, A. (1999) Innovations algorithm for periodically stationary time series. *Stochastic Processes and Applications* 83, 149–69.

Berk, K. (1974) Consistent autoregressive spectral estimates. *The Annals of Statistics* 2, 489–502.

Bhansali, R. (1978) Linear prediction by autoregressive model fitting in the time domain. *The Annals of Statistics* 6, 224–31.

BILLINGSLEY, P. (1968) *Convergence of Probability Measures*. New York: Wiley.
BROCKWELL, P. and DAVIS, R. (1988) Simple consistent estimation of the coefficients of a linear filter. *Stochastic Processes and Applications* 28, 47–59.
JONES, R. and BRELSFORD, W. (1967) Time series with periodic structure, *Biometrika* 54, 403–8.
LEWIS, R. and REINSEL G. (1985) Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis* 16, 393–411.
LUND, R. and BASAWA, I. V. (2000) Recursive prediction and likelihood evaluation for periodic ARMA models. *Journal of Time Series Analysis* 21, 75–93.
PAGANO, M. (1978) On periodic and multiple autoregressions. *The Annals of Statistics* 6, 1310–7.
SALAS, J., TABIOS, G. and BARTOLINI, P. (1985) Approaches to multivariate modeling of water resources time series. *Water Research Bulletin* 21, 683–708.
TJOSTHEIM, D. and PAULSEN, J. (1982) Empirical identification of multiple time series. *Journal Time Series Analysis* 3, 265–82.
TROUTMAN, B. (1979) Some results in periodic autoregression. *Biometrika*, 6, 219–28.
ULA, T. (1993) Forecasting of multivariate periodic autoregressive moving average processes. *Journal of Time Series Analysis* 14, 645.
VECCHIA, A. and BALLERINI, R. (1991) Testing for periodic autocorrelations in seasonal time series data. *Biometrika* 78, 53–63.