Innovations Algorithm for Periodically Stationary Time Series

Paul L. Anderson[1]

Department of Mathematics

University of Nevada

Mark M. Meerschaert

Department of Mathematics

University of Nevada

Aldo V. Vecchia

Water Resources Division

U.S. Geological Survey

April 20, 2004

AMS 1991 subject classification: Primary 62M10, 62E20; Secondary 60E07, 60F05.

Key words and phrases: time series, periodically stationary, Yule– Walker estimates, innovations algorithm, heavy tails, regular variation.

[1] On leave from Department of Mathematics, Albion College, Albion MI 49224.

Abstract

Periodic ARMA, or PARMA, time series are used to model periodically stationary time series. In this paper we develop the innovations algorithm for periodically stationary processes. We then show how the algorithm can be used to obtain parameter estimates for the PARMA model. These estimates are proven to be weakly consistent for PARMA processes whose underlying noise sequence has either finite or infinite fourth moment. Since many time series from the fields of economics and hydrology exhibit heavy tails, the results regarding the infinite fourth moment case are of particular interest.

1 Introduction

The innovations algorithm yields parameter estimates for nonstationary time series models. In this paper we show that these estimates are consistent for periodically stationary time series. A stochastic process X_t is called periodically stationary if $\mu_t = EX_t$ and $\gamma_t(h) = EX_t X_{t+h}$ for $h = 0, \pm 1, \pm 2, \ldots$ are all periodic functions of time t with the same period ν. 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 Jones and Brelsford (1967), Pagano (1978), Troutman (1979), Tjostheim and Paulsen (1982), Salas, Tabios, and Bartolini (1985), Vecchia and Ballerini (1991), Anderson and Vecchia (1993), Ula (1993), Adams and Goodwin (1995), and Anderson and Meerschaert (1997).

This paper provides a parameter estimation technique that considers two types of periodic time series models, those with finite fourth moment and the models with finite variance but infinite fourth moment. In the latter case we make the technical assumption that the innovations have regularly varying probability tails. The estimation procedure used adapts the well–known innovations algorithm (see for example Brockwell and Davis (1991) p. 172) to the case of periodically stationary time series. We show that the estimates from the algorithm are weakly consistent. A more formal treatment of the asymptotic behavior for the innovations algorithm will be discussed in a forthcoming paper Anderson, Meerschaert, and Vecchia (1999).

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 $\nu = 1$ and the process has finite fourth moments. For infinite fourth moment time series, our results are new even in the stationary case. Davis and Resnick (1986) establish the consistency of Yule–Walker estimates for a stationary autoregressive process of finite order with finite variance and infinite fourth moments. We extend their result to periodic ARMA processes. However, the Durbin–Levinson algorithm to compute the Yule–Walker estimates does not extend to nonstationary processes, and so these results are primarily of theoretical interest. Mikosch, Gadrich, Klüppenberg and Adler (1995) investigate parameter estimation for ARMA models with infinite variance innovations, but they do not consider the case of finite variance and infinite fourth moment. Time series with infinite fourth moment and finite variance are common in finance and hydrology, see for example Jansen and de Vries (1991), Loretan and Phillips (1994), and Anderson and Meerschaert (1998). The results in this paper provide the first practical method for time series parameter estimation in this important special case.

2 The Innovations Algorithm for Periodically Correlated Processes

Let $\{\tilde{X}_t\}$ be a time series with finite second moments and define its mean function $\mu_t = E(\tilde{X}_t)$ and its autocovariance function $\gamma_t(\ell) = \text{cov}(\tilde{X}_t, \tilde{X}_{t+\ell})$. $\{\tilde{X}_t\}$ is said to be periodically correlated with period ν if, for some positive integer ν and for all integers k and ℓ, (i) $\mu_t = \mu_{t+k\nu}$ and (ii) $\gamma_t(\ell) = \gamma_{t+k\nu}(\ell)$. For a monthly periodic time series it is typical that $\nu = 12$. In this paper, we are especially interested in the periodic ARMA process due to its importance in modeling periodically correlated processes. The periodic ARMA process,

$\{\tilde{X}_t\}$, with period ν (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 model parameters $\phi_t(j)$, $\theta_t(j)$, and σ_t are respectively the periodic autoregressive, periodic moving average, and periodic residual standard deviation parameters. In this paper we will consider models where $E\varepsilon_t^4 < \infty$, and also models in which $E\varepsilon_t^4 = \infty$. We will say that the i.i.d. sequence $\{\varepsilon_t\}$ is RV(α) if $P[|\varepsilon_t| > x]$ varies regularly with index $-\alpha$ and $P[\varepsilon_t > x]/P[|\varepsilon_t| > x] \to p$ for some $p \in [0,1]$. In the case where the noise sequence has infinite fourth moment, we assume that the sequence is RV(α) with $\alpha > 2$. This assumption implies that $E|\varepsilon_t|^\delta < \infty$ if $0 < \delta \leq \alpha$, in particular the variance of ε_t exists. With this technical condition, Anderson and Meerschaert (1997) 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 Samorodnitsky and Taqqu (1994).

There are some restrictions that need to be placed on the parameter space of (1). The first restriction is 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. The absolute summability of the ψ-weights ensures that (2) converges almost surely for all t, and in the mean-square to the same limit. The causality condition places constraints on the autoregressive parameters (see for example Tiao and Grupe (1980)) but these constraints are not the focus of this paper. It should be noted that $\psi_t(j) = \psi_{t+k\nu}(j)$ for all j. Another restriction on the parameter space of (1) is the 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. The invertibility condition places constraints on the moving average parameters in the same way that (2) places constraints on the autoregressive parameters. Again, $\pi_t(j) = \pi_{t+k\nu}(j)$ for all j.

Given N years of data with ν seasons per year, the innovations algorithm allows us to forecast future values of X_t for $t \geq N\nu$ in terms of the observed values $\{X_0, ..., X_{N\nu-1}\}$. Toward this end, we would like to find the best linear combination of $X_0, ..., X_{N\nu-1}$ for predicting $X_{N\nu}$ such that the mean- square distance from $X_{N\nu}$ is minimized. For a periodic time series, the one-step predictors must be calculated for each season i, $i = 0, 1..., \nu - 1$. The remainder of this section develops the innovations algorithm for periodic time series models. We adapt the development of Brockwell and Davis (1991) to this special case, and introduce the notation which will be used throughout the rest of the paper.

2.1 Equations for the One-Step Predictors

Let $\mathcal{H}_{n,i}$ denote the closed linear subspace $\overline{sp}\{X_i, ..., X_{i+n-1}\}, n \geq 1$, and let $\{\hat{X}_{i+n}^{(i)}\}, n \geq 0$, denote the one-step predictors, which are defined by

$$\hat{X}_{i+n}^{(i)} = \begin{cases} 0 & \text{if } n = 0 \\ P_{\mathcal{H}_{n,i}} X_{i+n} & \text{if } n \geq 1. \end{cases} \tag{4}$$

We call $P_{\mathcal{H}_{n,i}} X_{i+n}$ the projection mapping of X_{i+n} onto the space $\mathcal{H}_{n,i}$. Also, define

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

There are two representations of $P_{\mathcal{H}_{n,i}} X_{i+n}$ pertinent to the goals of this paper. The first one relates directly to the innovations algorithm and depends on writing $\mathcal{H}_{n,i}$ as a span of orthogonal components, viz.,

$$\mathcal{H}_{n,i} = \overline{sp}\{X_i - \hat{X}_i^{(i)}, X_{i+1} - \hat{X}_{i+1}^{(i)}, ..., X_{i+n-1} - \hat{X}_{i+n-1}^{(i)}\}, \qquad n \geq 1,$$

so that

$$\hat{X}_{i+n}^{(i)} = \sum_{j=1}^{n} \theta_{n,j}^{(i)} (X_{i+n-j} - \hat{X}_{i+n-j}^{(i)}). \tag{5}$$

The second representation of $P_{\mathcal{H}_{n,i}} X_{i+n}$ is given by

$$\hat{X}_{i+n}^{(i)} = \phi_{n,1}^{(i)} X_{i+n-1} + \cdots + \phi_{n,n}^{(i)} X_i, \quad n \geq 1. \tag{6}$$

The vector of coefficients, $\phi_n^{(i)} = (\phi_{n,1}^{(i)}, \ldots, \phi_{n,n}^{(i)})'$, appears in the prediction equations

$$\Gamma_{n,i} \phi_n^{(i)} = \gamma_n^{(i)} \tag{7}$$

where $\gamma_n^{(i)} = (\gamma_{i+n-1}(1), \gamma_{i+n-2}(2), \ldots, \gamma_i(n))'$ and

$$\Gamma_{n,i} = \left[\gamma_{i+n-1-\ell}(\ell - m)\right]_{\ell, m = 0, \ldots, n-1}, \quad i = 0, ..., \nu - 1. \tag{8}$$

is the covariance matrix of $(X_{i+n-1}, ..., X_i)'$. The condition sufficient for $\Gamma_{n,i}$ to be invertible for all $n \geq 1$ and each $i = 0, 1, \ldots, \nu - 1$ is given in the following proposition. Only the causality condition is required for the proposition to be valid.

Proposition 2.1.1 If $\sigma_i^2 > 0$ for $i = 0, \ldots, \nu - 1$, then for a causal PARMA$_\nu(p, q)$ process the covariance matrix $\Gamma_{n,i}$ in (7) is nonsingular for every $n \geq 1$ and each i.

Proof. See Proposition 4.1 of Lund and Basawa(1999) for a proof.

Remark. Proposition 5.1.1 of Brockwell and Davis (1991) does not extend to general periodically stationary processes. By Proposition 2.1.1, however, if our periodic process is a PARMA process, then we are guaranteed that the covariance matrix $\Gamma_{n,i}$ is nonsingular for every n and each i. To establish this remark consider the periodically stationary process

$\{X_t\}$ of period $\nu = 2$ given by

$$\begin{aligned} X_{2t} &= Z_{2t}, \\ X_{2t+1} &= (X_{2t-1} + X_{2t-2})/\sqrt{2} \end{aligned}$$

where $\{Z_t\}$ is an i.i.d. sequence of standard normal variables. It is easy to show that $\gamma_0(0) = \gamma_1(0) = 1$ and $\gamma_0(1) = 0$. Also, for $n \geq 1$, $\gamma_0(2n) = 0$, $\gamma_0(2n+1) = \frac{1}{2^{n/2}}$, $\gamma_1(2n-1) = 0$, and $\gamma_1(2n) = \frac{1}{2^{n/2}}$. The process $\{X_t\}$ is, by definition, periodically stationary of period $\nu = 2$. Using (8) we let $\Gamma_{n,0} = \left[\gamma_{i+n-1-\ell}(\ell - m)\right]_{\ell,m=0,\ldots,n-1}$ be the covariance matrix of $(X_{n-1}, ..., X_0)'$. Again, it is easy to show that $\Gamma_{2,0}$ and $\Gamma_{3,0}$ are identity matrices, hence nonsingular. However, $\Gamma_{4,0}$ is a singular matrix so that $\Gamma_{n,0}$ is singular for $n \geq 4$. Thus, the process is such that $\Gamma_{2,0}$ is invertible and $\gamma_i(h) \to 0$ as $h \to \infty$ but $\Gamma_{n,0}$ is singular for $n \geq 4$. Note that this process is not a PARMA_2 process.

2.2 The Innovations Algorithm

The proposition that follows is the innovations algorithm for periodically stationary processes. For a proof, see Proposition 5.2.2 in Brockwell and Davis (1991).

Proposition 2.2.1. If $\{X_t\}$ has zero mean and $E(X_\ell X_m) = \gamma_\ell(m - \ell)$, where the matrix $\Gamma_{n,i} = [\gamma_{i+n-1-\ell}(\ell - m)]_{\ell,m=0,\ldots,n-1}$, $i = 0, ..., \nu - 1$, is nonsingular for each $n \geq 1$, then the one-step predictors $\hat{X}_{i+n}$, $n \geq 0$, and their mean-square errors $v_{n,i}$, $n \geq 1$, are given by

$$\hat{X}_{i+n} = \begin{cases} 0 & \text{if } n = 0 \\ \sum_{j=1}^{n} \theta_{n,j}^{(i)} (X_{i+n-j} - \hat{X}_{i+n-j}) & \text{if } n \geq 1 \end{cases} \tag{9}$$

and for $k = 0, 1, \ldots, n-1$

$$
\begin{aligned}
v_{0,i} &= \gamma_i(0) \\
\theta^{(i)}_{n,n-k} &= (v_{k,i})^{-1}\Big[\gamma_{i+k}(n-k) - \textstyle\sum_{j=0}^{k-1} \theta^{(i)}_{k,k-j}\theta^{(i)}_{n,n-j} v_{j,i}\Big] \qquad (10)\\
v_{n,i} &= \gamma_{i+n}(0) - \textstyle\sum_{j=0}^{n-1} (\theta^{(i)}_{n,n-j})^2 v_{j,i}.
\end{aligned}
$$

We solve (10) recursively 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 corollaries which follow in this section require the invertibility condition (3). The first corollary shows that the innovations algorithm provides consistent estimates of the seasonal standard deviations, and the proof also provides the rate of convergence.

Corollary 2.2.1. In the innovations algorithm, for each $i = 0, 1, \ldots, \nu - 1$ we have

$$
v_{m,\langle i-m\rangle} \to \sigma_i^2 \text{ as } m \to \infty,
$$

where

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

and $[\cdot]$ is the greatest integer function. Note that $\langle k\rangle$ denotes the season associated with time k.

Proof. Let $\mathcal{H}_{i+n-1} = \overline{sp}\{X_j, -\infty < j \leq i+n-1\}$. Then

$$
\sigma^2_{i+m} = E(\varepsilon^2_{i+m}) = E(X_{i+m} + \sum_{j=1}^{\infty} \pi_{i+m}(j) X_{i+m-j})^2 = E(X_{i+m} - P_{\mathcal{H}_{i+m-1}} X_{i+m})^2
$$

where

$$
\sum_{j=1}^{\infty} \pi_{i+m}(j) X_{i+m-j} = P_{\mathcal{H}_{i+m-1}}(\varepsilon_{i+m} - X_{i+m}) = -P_{\mathcal{H}_{i+m-1}} X_{i+m}
$$

since $\varepsilon_{i+m} \perp \mathcal{H}_{i+m-1}$. Thus, we have,

$$
\sigma^2_{i+m} = E(X_{i+m} - P_{\mathcal{H}_{i+m-1}} X_{i+m})^2
$$

$$
\begin{aligned}
&\leq E(X_{i+m} - P_{\mathcal{H}_{m,i}} X_{i+m})^2 \\
&= v_{m,i} \\
&\leq E\Big(X_{i+m} + \sum_{j=1}^{m} \pi_{i+m}(j) X_{i+m-j}\Big)^2 \\
&= E\Big(\varepsilon_{i+m} - \sum_{j>m} \pi_{i+m}(j) X_{i+m-j}\Big)^2 \\
&= E(\varepsilon_{i+m})^2 + E\Big(\sum_{j>m} \pi_{i+m}(j) X_{i+m-j}\Big)^2 \\
&= \sigma_{i+m}^2 + E\Big(\sum_{j>m} \pi_{i+m}(j) X_{i+m-j} \sum_{k>m} \pi_{i+m}(k) X_{i+m-k}\Big) \\
&\leq \sigma_{i+m}^2 + \sum_{j,k>m} (|\pi_{i+m}(j)||\pi_{i+m}(k)| E|X_{i+m-j} X_{i+m-k}|) \\
&\leq \sigma_{i+m}^2 + \sum_{j,k>m} \Big(|\pi_{i+m}(j)||\pi_{i+m}(k)| \sqrt{\gamma_{i+m-j}(0)\gamma_{i+m-k}(0)}\Big) \\
&\leq \sigma_{i+m}^2 + \Big(\sum_{j>m} |\pi_{i+m}(j)|\Big)^2 M,
\end{aligned}
$$

where $M = \max\{\gamma_i(0) : i = 0, 1, \ldots, \nu - 1\}$. Since $\langle i - m \rangle + m = i + k\nu$ for all m and some k we write

$$
\sigma_{\langle i-m \rangle + m}^2 \leq v_{m,\langle i-m \rangle} \leq \sigma_{\langle i-m \rangle + m}^2 + M\Big(\sum_{j>m} |\pi_i(j)|\Big)^2
$$

yielding

$$
\sigma_i^2 \leq v_{m,\langle i-m \rangle} \leq \sigma_i^2 + M\Big(\sum_{j>m} |\pi_i(j)|\Big)^2
$$

where $v_{m,\langle i-m \rangle} = E(X_{n\nu+i} - P_{\mathcal{M}} X_{n\nu+i})^2$ and $\mathcal{M}=\overline{sp}\{X_{n\nu+i-1}, \ldots, X_{n\nu+i-m}\}$, n arbitrary. Hence, as $m \to \infty$, $v_{m,\langle i-m \rangle} \to \sigma_i^2$.

Corollary 2.2.2 $\lim_{m\to\infty} \|X_{i+m} - \hat{X}_{i+m}^{(i)} - \varepsilon_{i+m}\| = 0$.

Proof.

$$
E(X_{i+m} - \hat{X}_{i+m}^{(i)} - \varepsilon_{i+m})^2 \;\; = \;\; E(X_{i+m} - \hat{X}_{i+m}^{(i)})^2
$$

$$
\begin{aligned}
&\quad -2E[\varepsilon_{i+m}(X_{i+m} - \hat{X}^{(i)}_{i+m})] + E(\varepsilon^2_{i+m}) \\
&= v_{m,i} - 2\sigma^2_{i+m} + \sigma^2_{i+m} \\
&= v_{m,i} - \sigma^2_{i+m}
\end{aligned}
$$

where the last expression approaches 0 as $m \to \infty$ by Corollary 2.2.1.

Corollary 2.2.3 $\theta^{(\langle i-m \rangle)}_{m,k} \to \psi_i(k)$ as $m \to \infty$ for all $i = 0, 1, \ldots, \nu-1$ and all $k = 1, 2, \ldots,$.

Proof. We know that

$$\theta^{(i)}_{m,k} = v^{-1}_{m-k,i} E\Big(X_{i+m}(X_{i+m-k} - \hat{X}^{(i)}_{i+m-k})\Big)$$

and

$$\psi_{i+m}(k) = \sigma^{-2}_{i+m-k} E(X_{i+m}\varepsilon_{i+m-k}).$$

By the triangle inequality,

$$
\begin{aligned}
|\theta^{(i)}_{m,k} - \psi_{i+m}(k)| &\leq \Big|\theta^{(i)}_{m,k} - \sigma^{-2}_{i+m-k} E\Big(X_{i+m}(X_{i+m-k} - \hat{X}^{(i)}_{i+m-k})\Big)\Big| \\
&\quad + \Big|\sigma^{-2}_{i+m-k} E\Big(X_{i+m}(X_{i+m-k} - \hat{X}^{(i)}_{i+m-k} - \varepsilon_{i+m-k})\Big)\Big| \\
&= \Big|\theta^{(i)}_{m,k} - \sigma^{-2}_{i+m-k}\theta^{(i)}_{m,k} v_{m-k,i}\Big| \\
&\quad + \Big|\sigma^{-2}_{i+m-k} E\Big(X_{i+m}(X_{i+m-k} - \hat{X}^{(i)}_{i+m-k} - \varepsilon_{i+m-k})\Big)\Big| \\
&\leq \Big|\theta^{(i)}_{m,k} - \sigma^{-2}_{i+m-k}\theta^{(i)}_{m,k} v_{m-k,i}\Big| \\
&\quad + |\sigma^{-2}_{i+m-k}|\sqrt{\gamma_{i+m}(0)}\|X_{i+m-k} - \hat{X}^{(i)}_{i+m-k} - \varepsilon_{i+m-k}\|.
\end{aligned}
$$

As $m \to \infty$, the first term on the right-hand side approaches 0 by Corollary 2.2.1 and the fact that $\theta^{(i)}_{m,k}$ is bounded in m. Also, as $m \to \infty$, the second term on the right-hand side approaches 0 by Corollary 2.2.2 and the fact that $\sigma^{-2}_{i+m-k}\sqrt{\gamma_{i+m}(0)}$ is bounded in m. Thus, $|\theta^{(i)}_{m,k} - \psi_{i+m}(k)| \to 0$ as $m \to \infty$, and consequently, $|\theta^{(\langle i-m \rangle)}_{m,k} - \psi_i(k)| \to 0$ as $m \to \infty$, k arbitrary but fixed.

Corollary 2.2.4 $\phi_{m,k}^{(\langle i-m\rangle)} \to -\pi_i(k)$ as $m \to \infty$ for all $i = 0, 1, \ldots, \nu-1$ and $k = 1, 2, \ldots,$.

Proof. Define $\phi_m^{(i)} = (\phi_{m,1}^{(i)}, \ldots, \phi_{m,m}^{(i)})'$ and $\pi_m^{(i)} = (\pi_{i+m}(1), \ldots, \pi_{i+m}(m))'$. We show that $(\phi_m^{(i)} + \pi_m^{(i)}) \to 0$ as $m \to \infty$. From Theorem A.1 in the Appendix we have

$$\begin{aligned}\sum_{j=1}^{m}(\phi_{m,j}^{(i)} + \pi_{i+m}(j))^2 &\leq \frac{1}{2\pi C}(\phi_m^{(i)} + \pi_m^{(i)})'\Gamma_{m,i}(\phi_m^{(i)} + \pi_m^{(i)}) \\ &= \frac{1}{2\pi C}\mathrm{Var}\Big(\sum_{j=1}^{m}(\phi_{m,j}^{(i)} + \pi_{i+m}(j))X_{i+m-j}\Big) \\ &= \frac{1}{2\pi C}\mathrm{Var}\Big(\varepsilon_{i+m} - (X_{i+m} - \hat{X}_{i+m}^{(i)}) \\ &\quad - \sum_{j>m}\pi_{i+m}(j)X_{i+m-j}\Big)\end{aligned}$$

since

$$\varepsilon_{i+m} - (X_{i+m} - \hat{X}_{i+m}^{(i)}) = \sum_{j=1}^{m}(\phi_{m,j}^{(i)} + \pi_{i+m}(j))X_{i+m-j} + \sum_{j=m+1}^{\infty}\pi_{i+m}(j)X_{i+m-j}.$$

Now,

$$\begin{aligned}&\frac{1}{2\pi C}\mathrm{Var}\Big(\varepsilon_{i+m} - (X_{i+m} - \hat{X}_{i+m}^{(i)}) \\ &\quad - \sum_{j>m}\pi_{i+m}(j)X_{i+m-j}\Big) \\ \leq\ & \frac{1}{2\pi C}\cdot 2\Big[\mathrm{Var}\Big(\varepsilon_{i+m} - (X_{i+m} - \hat{X}_{i+m}^{(i)})\Big) + \mathrm{Var}\Big(\sum_{j>m}\pi_{i+m}(j)X_{i+m-j}\Big)\Big] \\ =\ & \frac{1}{\pi C}\Big[v_{m,i} - \sigma_{i+m}^2 + \mathrm{Var}\Big(\sum_{j>m}\pi_{i+m}(j)X_{i+m-j}\Big)\Big] \\ =\ & \frac{1}{\pi C}\Big[\Big(\sum_{j>m}|\pi_{i+m}(j)|\Big)^2 M + \Big(\sum_{j>m}|\pi_{i+m}(j)|\Big)^2 M\Big]\end{aligned}$$

where the first inequality is a result of the fact that $Var(X-Y) \leq 2Var(X) + 2Var(Y)$ and the last inequality follows from the proof of Corollary 2.2.1 recalling that $M = \max\{\gamma_i(0) : i = 0, 1, \ldots, \nu-1\}$. The right hand side of the last inequality approaches 0 as m approaches ∞ since $\sum|\pi_i(j)| < \infty$ for all $i = 0, 1, \ldots, \nu-1$. We have shown, for fixed but arbitrary k, that $|\phi_{m,k}^{(i)} + \pi_{i+m}(k)| \to 0$ as $m \to \infty$. Using the notation of Corollary 2.2.1, our corollary is established.

3 Weak Consistency of Innovation Estimates

Given N years of data $\tilde{X}_0, \tilde{X}_1, \ldots, \tilde{X}_{N\nu-1}$, where ν is the number of seasons per year, the estimated periodic autocovariance at season i and lag ℓ is defined by

$$\gamma_i^*(\ell) = N^{-1} \sum_{j=0}^{N-1} (\tilde{X}_{j\nu+i} - \tilde{\mu}_i)(\tilde{X}_{j\nu+i+\ell} - \tilde{\mu}_{i+\ell})$$

where

$$\tilde{\mu}_i = N^{-1} \sum_{j=0}^{N-1} \tilde{X}_{j\nu+i}$$

and any terms involving $\tilde{X}_t$ are set equal to zero whenever $t > N\nu - 1$. For what follows, it is simplest to work with the function

$$\hat{\gamma}_i(\ell) = N^{-1} \sum_{j=0}^{N-1} X_{j\nu+i} X_{j\nu+i+\ell} \tag{11}$$

where $X_t = \tilde{X}_t - \mu_t$. Since $\gamma_i^*(\ell)$ and $\hat{\gamma}_i(\ell)$ have the same asymptotic properties, we use (11) as our estimate of $\gamma_i(\ell)$. If we replace the autocovariances in the innovations algorithm with their corresponding sample autocovariances we obtain the estimator, $\hat{\theta}_{k,j}^{(\langle i-k \rangle)}$, of $\theta_{k,j}^{(\langle i-k \rangle)}$. We prove in this section that the innovations estimates are weakly consistent in the sense that

$$(\hat{\theta}_{k,1}^{(\langle i-k \rangle)} - \psi_i(1), \ldots, \hat{\theta}_{k,k}^{(\langle i-k \rangle)} - \psi_i(k), 0, 0, \ldots) \xrightarrow{P} 0$$

in R^∞ where $\xrightarrow{P}$ is used to denote convergence in probability. Results are presented for both the finite and infinite fourth moment cases. Theorems 3.1 through 3.4 below relate to the case where we assume the underlying noise sequence has finite fourth moment. Analogously, theorems 3.5 through 3.8 relate to the infinite fourth moment case where we assume the underlying noise sequence is RV(α) with $2 < \alpha \leq 4$ (see first paragraph of section 2). The latter set of theorems require regular variation theory for proof and are therefore treated separately from the first set of theorems. We assume throughout this section that the associated PARMA process is causal and invertible. With this assumption it can be shown that the spectral density matrix of the corresponding vector process, (see Anderson and Meerschaert (1997), pg. 778), is positive definite. We emphasize this fact in the statements of each of the theorems in this section, since it is essential in their proofs. Replacing the autocovariances given in (8) with their corresponding sample autocovariances yields the sample covariance matrix $\hat{\Gamma}_{n,i}$ for season $i = 0, \ldots, \nu - 1$. In theorems 3.1 and 3.2 we make use of the matrix 2-norm given by

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 \tag{12}$$

where $\|x\|_2 = (x'x)^{\frac{1}{2}}$ (see Golub and Van Loan (1989) pg. 56).

Theorem 3.1. Let $\{X_t\}$ be the mean zero PARMA process with period ν given by (1) with $E(\varepsilon_t^4) < \infty$. Assume that the spectral density matrix, $f(\lambda)$, of its equivalent vector ARMA process (see Anderson and Meerschaert (1997), pg. 778) is such that $mzz' \leq z'f(\lambda)z \leq Mzz'$, $-\pi \leq \lambda \leq \pi$, for some m and M such that $0 < m \leq M < \infty$ and for all z in R^ν. 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 $\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2 \xrightarrow{P} 0$.

Proof. The proof of this theorem is patterned after that of Lemma 3 in Berk (1974). Let

$p_{k,i} = \|\Gamma_{k,i}^{-1}\|_2$, $q_{k,i} = \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2$, and $Q_{k,i} = \|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_2$. Then

$$\begin{aligned} q_{k,i} &= \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2 \\ &= \|\hat{\Gamma}_{k,i}^{-1}(\hat{\Gamma}_{k,i} - \Gamma_{k,i})\Gamma_{k,i}^{-1}\|_2 \\ &\leq \|\hat{\Gamma}_{k,i}^{-1}\|_2\|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_2\|\Gamma_{k,i}^{-1}\|_2 \\ &= \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1} + \Gamma_{k,i}^{-1}\|_2\|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_2\|\Gamma_{k,i}^{-1}\|_2 \\ &\leq \Big\{\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2 + \|\Gamma_{k,i}^{-1}\|_2\Big\}\|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_2\|\Gamma_{k,i}^{-1}\|_2 \\ &= (q_{k,i} + p_{k,i})Q_{k,i}p_{k,i}. \end{aligned} \tag{13}$$

Now,

$$Q_{k,i}^2 = \|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_2^2 \leq \sum_{\ell,m=0}^{k-1}\Big[\hat{\gamma}_{i+k-\ell-1}(l-m) - \gamma_{i+k-\ell-1}(\ell-m)\Big]^2.$$

Multiplying the above equation by N and taking expectations yields

$$NE(Q_{k,i}^2) \leq N\sum_{\ell,m=0}^{k-1}\mathrm{Var}\Big(\hat{\gamma}_{i+k-\ell-1}(\ell-m)\Big).$$

Anderson (1989) shows that $N\,\mathrm{Var}(\hat{\gamma}_{i+k-\ell-1}(\ell-m))$ is bounded above by

$$\Big|\eta - 3\Big|\Big(\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}|\psi_{i+k-\ell-1}(m_1)||\psi_{i+k-m-1}(m_2)|\Big)^2 < \infty$$

where $\eta = E(\varepsilon_t^4)$. Define

$$C = \max\Big\{|\eta-3|\Big(\sum_{m_1=0}^{\infty}\sum_{m_2=0}^{\infty}|\psi_i(m_1)||\psi_j(m_2)|\Big)^2, 0 \leq i,j \leq \nu-1\Big\}$$

which is independent of N and k so that we can write

$$NE(Q_{k,i}^2) \leq k^2C$$

which holds for all i. Thus, $E(Q_{k,i}^2) \leq k^2C/N \to 0$ as $k \to \infty$ since $k^2/N \to 0$ as $N \to \infty$. It follows that $Q_{k,i} \xrightarrow{P} 0$ and since $p_{k,i}$ is bounded for all i and k (see Appendix, Theorem A.1), we also have $p_{k,i}Q_{k,i} \xrightarrow{P} 0$. From (13) we can write

$$q_{k,i} \leq \frac{p_{k,i}^2Q_{k,i}}{1 - p_{k,i}Q_{k,i}}$$

if $1 - p_{k,i}Q_{k,i} > 0$, i.e., if $p_{k,i}Q_{k,i} < 1$. Now,

$$\begin{aligned} P(q_{k,i} > \epsilon) &= P(q_{k,i} > \epsilon | p_{k,i}Q_{k,i} < 1)P(p_{k,i}Q_{k,i} < 1) \\ &+ P(q_{k,i} > \epsilon | p_{k,i}Q_{k,i} \geq 1)P(p_{k,i}Q_{k,i} \geq 1) \\ &\leq P\Big(\frac{p_{k,i}^2 Q_{k,i}}{1 - p_{k,i}Q_{k,i}} > \epsilon\Big) \\ &+ P(q_{k,i} > \epsilon | p_{k,i}Q_{k,i} \geq 1)P(p_{k,i}Q_{k,i} \geq 1). \end{aligned}$$

Since $p_{k,i}^2 Q_{k,i} \xrightarrow{P} 0$ and $(1 - p_{k,i}Q_{k,i}) \xrightarrow{P} 1$, then by Theorem 5.1, Corollary 2 of Billingsley (1968) $\frac{p_{k,i}^2 Q_{k,i}}{1 - p_{k,i}Q_{k,i}} \xrightarrow{P} 0$. Also, we know that $\lim_{k\to\infty} P(p_{k,i}Q_{k,i} \geq 1) = 0$ so

$$\begin{aligned} \lim_{k\to\infty} P(q_{k,i} > \epsilon) &\leq \lim_{k\to\infty} P(\frac{p_{k,i}^2 Q_{k,i}}{1 - p_{k,i}Q_{k,i}} > \epsilon) + 1 \cdot \lim_{k\to\infty} P(p_{k,i}Q_{k,i} \geq 1) \\ &= 0 + 1 \cdot 0 \\ &= 0 \end{aligned}$$

and it follows that $q_{k,i} \xrightarrow{P} 0$. This proves the theorem.

Substituting sample autocovariances for autocovariances in (7) yields the Yule-Walker estimators

$$\hat{\phi}_k^{(i)} = \hat{\Gamma}_{k,i}^{-1} \hat{\gamma}_k^{(i)} \tag{14}$$

assuming $\hat{\Gamma}_{k,i}^{-1}$ exists. The next theorem shows that $\hat{\phi}_k^{(i)}$ is consistent for $\phi_k^{(i)}$.

Theorem 3.2. If the hypotheses of Theorem 3.1 hold, then $(\hat{\phi}_k^{(i)} - \phi_k^{(i)}) \xrightarrow{P} 0$.

Proof. Write

$$\begin{aligned} \hat{\phi}_k^{(i)} - \phi_k^{(i)} &= \hat{\Gamma}_{k,i}^{-1}\hat{\gamma}_k^{(i)} - \Gamma_{k,i}^{-1}\gamma_k^{(i)} \\ &= \hat{\Gamma}_{k,i}^{-1}\hat{\gamma}_k^{(i)} - \hat{\Gamma}_{k,i}^{-1}\gamma_k^{(i)} + \hat{\Gamma}_{k,i}^{-1}\gamma_k^{(i)} - \Gamma_{k,i}\gamma_k^{(i)} \\ &= \hat{\Gamma}_{k,i}^{-1}(\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}) + (\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1})\gamma_k^{(i)}. \end{aligned}$$

Then,

$$
\begin{aligned}
\|\hat{\phi}_k^{(i)} - \phi_k^{(i)}\|_2 &\leq \|\hat{\Gamma}_{k,i}^{-1}\|_2 \|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 + \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2 \|\gamma_k^{(i)}\|_2 \\
&= \|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1} + \Gamma_{k,i}^{-1}\|_2 \|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 + q_{k,i}\|\gamma_k^{(i)}\|_2 \\
&\leq \left\{\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_2 + \|\Gamma_{k,i}^{-1}\|_2\right\} \|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 + q_{k,i}\|\gamma_k^{(i)}\|_2 \\
&= (q_{k,i} + p_{k,i})\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 + q_{k,i}\|\gamma_k^{(i)}\|_2.
\end{aligned}
$$

The last term on the right-hand side of the inequality goes to 0 in probability by Theorem 3.1 and the fact that

$$
\|\gamma_k^{(i)}\|_2 = \sum_{j=0}^{k-1} \Big(\gamma_{i+j}(k-j)\Big)^2 \leq \sum_{i=0}^{\nu-1} \sum_{j=0}^{\infty} \gamma_i^2(j) < \infty
$$

by the absolute summability of $\{\gamma_i(k)\}$ for each $i = 0, 1, \ldots, \nu - 1$. The first term on the right-hand side of the inequality goes to 0 in probability if we can show that $\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 \xrightarrow{P} 0$ by Theorem 3.1 and the fact that $p_{k,i}$ is uniformly bounded. Write

$$
\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2^2 = \sum_{j=0}^{k-1} \Big(\hat{\gamma}_{i+j}(k-j) - \gamma_{i+j}(k-j)\Big)^2
$$

which leads to

$$
\begin{aligned}
E\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2^2 &= \sum_{j=0}^{k-1} E\Big(\hat{\gamma}_{i+j}(k-j) - \gamma_{i+j}(k-j)\Big)^2 \\
&\leq \sum_{j=0}^{k-1} C/N \\
&= kC/N
\end{aligned}
$$

where $kC/N \to 0$ by hypothesis and where C is as in the proof of Theorem 3.1. It follows that $\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_2 \xrightarrow{P} 0$ and hence $(\hat{\phi}_k^{(i)} - \phi_k^{(i)}) \xrightarrow{P} 0$.

Theorem 3.3. Under the conditions of Theorem 3.1, we have that

$$
\hat{\phi}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} -\pi_i(j)
$$

for all j.

Proof. From Corollary 2.2.4 we know that $\phi_{k,j}^{(i)} + \pi_{i+k}(j) \to 0$ for all j as $k \to \infty$. From Theorem 3.2 we have $\hat{\phi}_{k,j}^{(i)} - \phi_{k,j}^{(i)} \xrightarrow{P} 0$ for all j so that

$$\begin{aligned} |\hat{\phi}_{k,j}^{(i)} + \pi_{i+k}(j)| &= |\hat{\phi}_{k,j}^{(i)} - \phi_{k,j}^{(i)} + \phi_{k,j}^{(i)} + \pi_{i+k}(j)| \\ &\leq |\hat{\phi}_{k,j}^{(i)} - \phi_{k,j}^{(i)}| + |\phi_{k,j}^{(i)} + \pi_{i+k}(j)| \\ &\xrightarrow{P} 0 \end{aligned}$$

as $k \to \infty$ for all fixed but arbitrary j, by the continuous mapping theorem. Another application of the continuous mapping theorem yields

$$\hat{\phi}_{k,j}^{(\langle i-k \rangle)} + \pi_i(j) - \pi_i(j) \xrightarrow{P} 0 - \pi_i(j) = -\pi_i(j)$$

using the notation of Corollary 2.2.1. This proves the theorem.

Theorem 3.4. Under the conditions in Theorem 3.1, we have that

$$\hat{\theta}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(j)$$

for all j.

Proof. From the representations of $\hat{X}_{i+k}$ given by (5) and (6) and the invertibility of $\Gamma_{k,i}$ for all k and i, one can check that

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

$j = 1, \ldots, k$ if we define $\theta_{k-j,0}^{(i)} = 1$. Also, because of the way the estimates $\hat{\theta}_{k,j}^{(i)}$ and $\hat{\phi}_{k,j}^{(i)}$ are defined we have

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

$j = 1, \ldots, k$ if we define $\hat{\theta}_{k-j,0}^{(i)} = 1$. We propose that, for every n,

$$\hat{\theta}_{k,\ell}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(\ell),$$

$\ell = 1, \ldots, n$ as $k \to \infty$ and $N \to \infty$ according to the hypotheses of the theorem. We use strong induction on n. The proposition is true for $n = 1$ since

$$\hat{\theta}_{k,1}^{(\langle i-k \rangle)} = \hat{\phi}_{k,1}^{(\langle i-k \rangle)} \xrightarrow{P} -\pi_i(1) = \psi_i(1).$$

Now, assume the proposition is true for $n = j-1$,i.e., $\hat{\theta}_{k,\ell}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(\ell)$, $\ell = 1, \ldots, j-1$. Note that $\hat{\theta}_{k-\ell,j-\ell}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(j-\ell)$ as $N \to \infty$ and $k \to \infty$ according to $k^2/N \to 0$ since $(k-\ell)^2/N \to 0$ also. Additionally, $\hat{\phi}_{k,\ell}^{(\langle i-k \rangle)} \xrightarrow{P} -\pi_i(\ell)$, so by the continuous mapping theorem,

$$\hat{\theta}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} \sum_{\ell=1}^{j} -\pi_i(\ell)\psi_i(j-\ell) = \psi_i(j)$$

hence the theorem follows.

Corollary 3.4 $\hat{v}_{k,\langle i-k \rangle} \xrightarrow{P} \sigma_i^2$ where

$$\hat{v}_{k,\langle i-k \rangle} = \hat{\gamma}_i(0) - \sum_{j=0}^{k-1} (\hat{\theta}_{k,k-j}^{(\langle i-k \rangle)})^2 \hat{v}_{j,\langle i-k \rangle}.$$

Proof. Using a strong induction argument similar to that in Theorem 3.4 yields the result.

In Theorems 3.5 and 3.6 the matrix 1-norm is used to obtain the required bounds on the appropriate statistics since these theorems deal with the infinite fourth moment case. The matrix 1-norm is given by

$$\|A\|_1 = \max_{\|x\|_1 = 1} \|Ax\|_1$$

where $\|x\|_1 = |x_1| + \cdots + |x_k|$ (see Golub and Van Loan (1989) pg. 57). We also need to define

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

where

$$a_N^{-1} \sum_{t=0}^{N-1} \varepsilon_{t\nu+i} \Rightarrow S^{(i)},$$

$S^{(i)}$ is an α-stable law, and $\Rightarrow$ denotes convergence in distribution.

Theorem 3.5 Let $\{X_t\}$ be the mean zero PARMA process with period ν given by (1) with $2 < \alpha \leq 4$. Assume that the spectral density matrix, $f(\lambda)$, of its equivalent vector ARMA process is such that $mzz' \leq z' f(\lambda) z \leq Mzz'$, $-\pi \leq \lambda \leq \pi$, for some m and M such that $0 < m \leq M$ and for all z in R^{ν}. 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 $\|\hat{\Gamma}_{k,i}^{-1} - \Gamma_{k,i}^{-1}\|_1 \xrightarrow{P} 0$.

Proof. Define $p_{k,i}$, $q_{k,i}$, and $Q_{k,i}$ as in Theorem 3.1 with the 1-norm replacing the 2-norm. Starting with the equations (13), we want to show that $Q_{k,i} \xrightarrow{P} 0$. Toward this end, it is shown in the Appendix, Theorem A.2, that there exists a constant, C, such that

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

for all $i = 0, 1, \ldots, \nu - 1$, for all $\ell = 0, \pm 1, \pm 2, \ldots$, and for all $N = 1, 2, \ldots$. If we have a random $k \times k$ matrix A with $E|a_{ij}| \leq C$ for all i and j then

$$\begin{aligned} E\|A\|_1 &= E\Big(\max_{1 \leq j \leq k} \sum_{i=1}^{k} |a_{ij}| \Big) \\ &\leq E \sum_{i,j=1}^{k} |a_{ij}| \\ &= k^2 C. \end{aligned}$$

Thus

$$E(Q_{k,i}) = E\|\hat{\Gamma}_{k,i} - \Gamma_{k,i}\|_1 \leq k^2 a_N^2 C / N$$

for all i,k, and N. We therefore have that $Q_{k,i} \xrightarrow{P} 0$ and since

$$p_{k,i} = \|\Gamma_{k,i}^{-1}\|_1 \leq k^{1/2} \|\Gamma_{k,i}^{-1}\|_2$$

then $p_{k,i} Q_{k,i} \xrightarrow{P} 0$ if $k^{5/2} a_N^2 C / N \to 0$ as $k \to \infty$ and $N \to \infty$. To show that $q_{k,i} \xrightarrow{P} 0$ we follow exactly the proof given in Theorem 3.1 and this concludes the proof of our theorem.

Theorem 3.6 Given the hypotheses set forth in Theorem 3.5 we have that $(\hat{\phi}_k^{(i)} - \phi_k^{(i)}) \xrightarrow{P} 0$.

Proof. From the proof of Theorem 3.2 with the 1-norm replacing the 2-norm we start with the inequality

$$\|\hat{\phi}_k^{(i)} - \phi_k^{(i)}\|_1 \le (q_{k,i} + p_{k,i})\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_1 + q_{k,i}\|\gamma_{k,i}\|_1.$$

The last term on the right-hand side of the inequality goes to 0 in probability by Theorem 3.5 and the fact that

$$\|\gamma_k^{(i)}\|_1 = \sum_{j=0}^{k-1} |\gamma_{i+j}(k-j)| \le \sum_{i=0}^{\nu-1}\sum_{j=0}^{\infty} |\gamma_i(j)| < \infty$$

by the absolute summability of $\{\gamma_i(k)\}$ for each $i = 0, 1, \ldots, \nu - 1$. The first term on the right-hand side of the inequality goes to 0 in probability if we can show that $\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_1 \xrightarrow{P} 0$ since we know that $k^{-1/2} p_{k,i}$ is uniformly bounded. By Theorem A.2 in the Appendix

$$\begin{aligned} E\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_1 &= \sum_{j=0}^{k-1} E|\hat{\gamma}_{i+j}(k-j) - \gamma_{i+j}(k-j)| \\ &\le kCa_N^2/N \end{aligned}$$

where the last term approaches 0 by hypothesis. It follows that $\|\hat{\gamma}_k^{(i)} - \gamma_k^{(i)}\|_1 \xrightarrow{P} 0$ and hence $(\hat{\phi}_k^{(i)} - \phi_k^{(i)}) \xrightarrow{P} 0$.

Theorem 3.7 Let $\{X_t\}$ be the mean zero PARMA process with period ν given by (1) with $2 < \alpha \le 4$. Assume that the spectral density matrix, $f(\lambda)$, of its equivalent vector ARMA process is such that $mzz' \le z' f(\lambda) z \le Mzz'$, $-\pi \le \lambda \le \pi$, for some m and M such that $0 < m \le M$ and for all z in R^ν. 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 $\hat{\theta}_{k,j}^{(\langle i-k \rangle)} \xrightarrow{P} \psi_i(j)$ for all j and for every $i = 0, 1, \ldots, \nu - 1$.

Proof. The result follows by mimicking the proofs given in Theorems 3.3 and 3.4.

We state the next corollary without proof since it is completely analogous to Corollary 3.4.

Corollary 3.7 $\hat{v}_{k,\langle i-k\rangle} \xrightarrow{P} \sigma_i^2$ where

$$\hat{v}_{k,\langle i-k\rangle} = \hat{\gamma}_i(0) - \sum_{j=0}^{k-1} (\hat{\theta}_{k,k-j}^{(\langle i-k\rangle)})^2 \hat{v}_{j,\langle i-k\rangle}.$$

Remarks.

1. All of the results in this section hold true for second-order stationary ARMA models since they are a special case of the periodic ARMA models with $\nu = 1$.

2. In Theorem 3.2 and Theorem 3.6, we not only have that $(\hat{\phi}_k^{(i)} - \phi_k^{(i)}) \xrightarrow{P} 0$ in R^∞ but also in ℓ_2.

APPENDIX

Theorem A.1. Let $\{X_t\}$ be a mean zero periodically stationary time series with period $\nu \geq 1$. Also, let $Y_t = (X_{t\nu+\nu-1}, \ldots, X_{t\nu})'$ be the corresponding ν-variate stationary vector process with spectral density matrix, $f(\lambda)$. If there exists constants c and C such that $cz'z \leq z'f(\lambda)z \leq Cz'z$ for all $z \in \mathrm{R}^\nu$ where $0 < c \leq C < \infty$, then $\|\Gamma_{k,i}\|_2 \leq 2\pi C$ and $\|\Gamma_{k,i}^{-1}\|_2 \leq 1/(2\pi c)$ for all k and i. Note that $\|A\|_2$ is the matrix 2-norm defined by (12).

Proof. Let $\Gamma(h) = \mathrm{Cov}(Y_t, Y_{t+h})$, $Y = (Y_{n-1}, \ldots, Y_0)'$, and $\Gamma = \mathrm{Cov}(Y, Y) = [\Gamma(i-j)]_{i,j=0}^{n-1}$ where Y_t is as stated in the theorem. In the notation of (8) we see that $\Gamma = \Gamma_{n\nu,0} = \mathrm{Cov}(X_{n\nu-1}, \ldots, X_0)'$. For fixed i and k let $n = [\frac{k+i}{\nu}] + 1$. Then $\Gamma_{k,i} = \mathrm{Cov}(X_{i+k-1}, \ldots, X_i)'$ is a submatrix of $\Gamma = \Gamma_{n\nu,0}$. It is clear that $\|\Gamma^{-1}\|_2 \geq \|\Gamma_{k,i}^{-1}\|_2$ and $\|\Gamma_{k,i}\|_2 \leq \|\Gamma\|_2$, since $\Gamma_{k,i}$ is the restriction of Γ onto a lower dimensional subspace. The spectral density matrix of Y_t is $f(\lambda) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\lambda h}\Gamma(h)$ so that $\Gamma(h) = \int_{-\pi}^{\pi} e^{i\lambda h} f(\lambda) d\lambda$. Define the fixed but arbitrary vector $y \in \mathrm{R}^{n\nu}$ such that $y = (y_0, y_1, \ldots, y_{n-1})'$ where $y_j = (y_{j\nu}, y_{j\nu+1}, \ldots, y_{j\nu+\nu-1})'$. Then

$$\begin{aligned}
y'\Gamma y &= \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} y_j'\Gamma(j-k)y_k \\
&= \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} y_j'(\int_{-\pi}^{\pi} e^{i\lambda(j-k)}f(\lambda)d\lambda)y_k \\
&= \int_{-\pi}^{\pi}(\sum_{j=0}^{n-1} e^{i\lambda j}y_j)'f(\lambda)(\sum_{k=0}^{n-1} e^{-i\lambda k}y_k)d\lambda \\
&\leq C\int_{-\pi}^{\pi}(\sum_{j=0}^{n-1} e^{i\lambda j}y_j)'(\sum_{k=0}^{n-1} e^{-i\lambda k}y_k)d\lambda \\
&= C\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} y_j'y_k\int_{-\pi}^{\pi} e^{i\lambda(j-k)}d\lambda \\
&= 2\pi C\sum_{j=0}^{n-1} y_j'y_j \\
&= 2\pi C y'y.
\end{aligned}$$

Similarly, $y'\Gamma y \geq 2\pi c y'y$. If $\Gamma y = \lambda y$ then $y'\Gamma y = y'\lambda y = \lambda y'y$ so $2\pi c y'y \leq \lambda y'y \leq 2\pi C y'y$ which shows that every eigenvalue of Γ lies between $2\pi c$ and $2\pi C$ for all n. If we write $\lambda_1 \leq \cdots \leq \lambda_{n\nu}$ for the eigenvalues of Γ then since $\lambda_1 = \frac{1}{\|\Gamma^{-1}\|_2}$ and $\lambda_{n\nu} = \|\Gamma\|_2$ we have

$$\|\Gamma_{k,i}\|_2 \leq \|\Gamma\|_2 = \lambda_{n\nu} \leq 2\pi C$$

and

$$\|\Gamma_{k,i}^{-1}\|_2 \leq \|\Gamma^{-1}\|_2 = \frac{1}{\lambda_1} \leq \frac{1}{2\pi c}.$$

The next result given in the Appendix affirms that $E|Na_N^{-2}(\hat{\gamma}_i(\ell) - \gamma_i(\ell)|$ is uniformly bounded for all $i = 0, 1, \ldots, \nu - 1$, for all $\ell = 0, \pm 1, \pm 2, \ldots$, and for all $N = 1, 2, \ldots$. We assume that (1) and (2) hold and that the i.i.d. sequence $\{\varepsilon_t\}$ is RV(α) with $2 < \alpha < 4$. Then the squared noise $Z_t = \varepsilon_t^2$ belong to the domain of attraction of an $\alpha/2$-stable law. We also have that

$$\begin{aligned} \gamma_i(\ell) &= E(X_{n\nu+i}X_{n\nu+i+\ell}) \\ &= \sum_{j=-\infty}^{\infty} \psi_i(j)\psi_{i+\ell}(j+\ell) \end{aligned}$$

assuming $E(\varepsilon_t) = 0$, and $E(\varepsilon_t^2) = 1$. In preparation for the following two lemmas we define the quantities

$$\begin{aligned} V_\eta(y) &= E|Z_1|^\eta I(|Z_1| \geq y) \\ U_\zeta(y) &= E|Z_1|^\zeta I(|Z_1| \leq y) \end{aligned}$$

and recall that a_N=inf$\{x : P(|\varepsilon_t| > x) < 1/N\}$.

Lemma A.1. Let the i.i.d. sequence $\{Z_t\}$ be in the domain of attraction of an α-stable law where $1 < \alpha < 2$ and $E(Z_t) = 0$. For all $\delta > 0$, there exists some constant K such that

$$P\Big(\Big|\sum_{i=1}^{N} Z_i\Big| > d_N t\Big) \leq K t^{-\alpha+\delta}$$

for all $t > 0$ and $N \geq 1$ where $d_N = a_N^2$ and $NV_0(d_N)\to 1$.

Proof. For fixed but arbitrary $t > 0$ define

$$\begin{aligned} T_N &= \sum_{i=1}^{N} Z_i, \\ T_{NN} &= \sum_{i=1}^{N} Z_i I(|Z_i| \leq d_N t), \\ E_N &= \bigcup_{i=1}^{N} (|Z_i| > d_N t) \\ & \quad \textit{and} \\ G_N &= \{|T_{NN}| > d_N t\}. \end{aligned}$$

Then $P(|T_N| > d_N t) \leq P(E_N) + P(G_N)$. Also,

$$P(E_N) \leq NP(|Z_1| > d_N t) = NV_0(d_N t) \leq C_1 t^{-\alpha+\delta}$$

for all t greater than or equal to some t_0, where the last inequality follows from Potter's Theorem (see Bingham, Goldie, and Teugels (1987), pg.25). Now, by Chebychev's inequality, $P(G_N) \leq E(T_{NN}^2)/(d_N^2 t^2)$ where

$$\begin{aligned} E(T_{NN}^2) &= NEZ_1^2 I(|Z_1| \leq d_N t) \\ &= +N(N-1)E\{Z_1 I(|Z_1| \leq d_N t) Z_2 I(|Z_2| \leq d_N t)\} \\ &= I_N + J_N. \end{aligned}$$

Note that,

$$\frac{I_N}{d_N^2 t^2} = \frac{NU_2(d_N t)}{d_N^2 t^2} = NV_0(d_N t) \frac{U_2(d_n t)}{(d_N t)^2 V_0(d_N t)} \leq C_2 t^{-\alpha+\delta}$$

for all $t \geq t_0$ by Karamata's Theorem (see Feller(1971), pg. 283). Also, for all $t \geq t_0$

$$\begin{aligned} \left|\frac{J_N}{d_N^2 t^2}\right| &\leq \frac{N^2}{d_N^2 t^2}\Big|EZ_1 I(|Z_1| \leq d_N t) EZ_2 I(|Z_2| \leq d_N t)\Big| \\ &= \left\{\frac{N}{d_N t}\Big|EZ_1 I(|Z_1| > d_N t)\Big|\right\}^2 \\ &= \left\{\frac{NV_1(d_N t)}{d_N t}\right\}^2 \\ &= \left\{NV_0(d_N t) \frac{(d_N t)^3 V_1(d_N t)}{U_4(d_N t)} \frac{U_4(d_N t)}{(d_N t)^4 V_0(d_N t)}\right\}^2 \\ &\leq C_3 t^{-\alpha+\delta} \end{aligned}$$

by Karamata's Theorem. Hence $P(|T_N| > d_N t) \leq K t^{-\alpha+\delta}$ for all $t \geq t_0$ with $K = C_1 + C_2 + C_3$. Now, enlarge K if necessary so that $K t_0^{-\alpha+\delta} > 1$. Then

$$P\left(\Big|\sum_{i=1}^N Z_i\Big| > d_N t\right) \leq K t^{-\alpha+\delta}$$

holds for $t > 0$ because $P(|\sum_{i=1}^N Z_i| > d_N t) \leq 1$.

Lemma A.2. Under the conditions of Lemma A.1.,

$$E\Big|d_N^{-1} \sum_{i=1}^N Z_i\Big| \to E|Y|$$

where

$$d_N^{-1} \sum_{i=1}^{N} Z_i \Rightarrow Y.$$

Proof. By Billingsley (1995), pg. 338, it suffices to show that $E|d_N^{-1} T_N|^{1+\epsilon} < \infty$ for all N where $T_N = \sum_{i=1}^{N} Z_i$. By Lemma A.1.,

$$\begin{aligned} E|d_N^{-1} T_N|^{1+\epsilon} &= \int_0^\infty P(|d_N^{-1} T_N|^{1+\epsilon} > t) dt \\ &= \int_0^1 P(|d_N^{-1} T_N|^{1+\epsilon} > t) dt + \int_1^\infty P(|d_N^{-1} T_N|^{1+\epsilon} > t) dt \\ &\leq 1 + \int_1^\infty K(t^{\frac{1}{1+\epsilon}})^{-\alpha+\delta} dt \end{aligned}$$

where the last term is finite.

Theorem A.2. There exists a constant, $C > 0$, such that

$$E|Na_N^{-2}(\hat{\gamma}_i(\ell) - \gamma_i(\ell))| \leq C$$

for all $i = 0, 1, \ldots, \nu - 1$, for all $\ell = 0, \pm 1, \pm 2, \ldots$, and for all $N = 1, 2, \ldots$. Proof. By the proof of Lemma 2.1 of Anderson and Meerschaert (1997) we have

$$Na_N^{-2}\Big(\hat{\gamma}_i(\ell) - N^{-1} \sum_{t=0}^{N-1} \sum_{j=-\infty}^{\infty} \psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon^2_{t\nu+i-j}\Big) = A_1 + A_2 + A_3 + A_4$$

where

$$\begin{aligned} Var(A_1) &\leq Na_N^{-4}\sigma_N^4 K \\ E|A_2| &\leq Na_N^{-2}|\mu_N| K \\ E|A_3| &\leq Na_N^{-2} V_N K \\ |A_4| &\leq Na_N^{-2}\mu_N^2 K \end{aligned}$$

where

$$\mu_N = E\varepsilon_1 I(|\varepsilon_1| \leq a_N)$$

$$
\begin{aligned}
\sigma_N^2 &= E\varepsilon_1^2 I(|\varepsilon_1| \le a_N) \\
V_N &= E|\varepsilon_1| I(|\varepsilon_1| \le a_N) \\
K &= \sum_{j=-\infty}^{\infty} |\psi_i(j)| \sum_{j=-\infty}^{\infty} |\psi_{i+\ell}(j)|
\end{aligned}
$$

and $\psi_i(j) = 0$ for $j < 0$. For $\alpha > 2$ we have $\sigma_N^2 \le 1$, $|\mu_N| \le V_N$, and $V_N \sim \frac{\alpha}{\alpha-1}\frac{a_N}{N}$. Then $Na_N^{-1}V_N \to \frac{\alpha}{\alpha-1}$ implies $Na_N^{-1}V_N \le K$ for all N, so $Na_N^{-1}|\mu_N| \le K$ for all N. Therefore, $E|A_2| \le a_N^{-1}K$, $E|A_3| \le a_N^{-1}K$, and $|A_4| \le Na_N^{-1}|\mu_N|Na_N^{-1}|\mu_N|N^{-1} \le N^{-1}a_N^{-1}K$ for all N, and finally, since

$$
(E|A_1|)^2 \le E(|A_1|^2) = E(A_1^2) = \mathrm{Var}(A_1)
$$

we have

$$
E|A_1| \le \sqrt{\mathrm{Var}(A_1)} \le N^{1/2}a_N^{-2}\sigma^2 K^{1/2}
$$

for all N. Thus, for all N, i, and ℓ, we have

$$
\begin{aligned}
&E\Big|Na_N^{-2}\Big(\hat{\gamma}_i(\ell) - N^{-1}\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon_{t\nu+i-j}^2\Big)\Big| \\
&\le E|A_1| + E|A_2| + E|A_3| + |A_4| \\
&\le N^{1/2}a_N^{-2}\sigma^2K^{1/2} + a_N^{-1}K + a_N^{-1}K + N^{-1}a_N^{-1}K \\
&\le K_0 N^{1/2}a_N^{-2}
\end{aligned}
$$

for all N, i, and ℓ. Next write

$$
\begin{aligned}
& N^{-1}\Big(\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon_{t\nu+i-j}^2\Big) - \gamma_i(\ell) \\
=\; & N^{-1}\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)(\varepsilon_{t\nu+i-j}^2 - 1)
\end{aligned}
$$

and apply Lemma A.2. with $d_N = a_N^2$ and $Z_t = \varepsilon_t^2 - 1$ to see that

$$
E|a_N^{-2}\sum_{t=0}^{N-1}(\varepsilon_{t\nu+i-j}^2 - 1)| \to E|S_{i-j}|
$$

where, as in Anderson and Meerschaert (1997) we have the corresponding weak convergence result

$$a_N^{-2}\sum_{t=0}^{N-1}(\varepsilon^2_{t\nu+r}-1)\Rightarrow S_r$$

for all $r = 0, 1, \ldots, \nu-1$ where $S_0, \ldots, S_{\nu-1}$ are i.i.d. $\alpha/2$-stable laws. Then we have $E|a_N^{-2}\sum_{t=0}^{N-1}(\varepsilon^2_{t\nu-r}-1)| < C^{(r)}$ for $r = 0, \ldots, \nu-1$ since this sequence is convergent, hence bounded. Let $B_0 = \max\{C^{(r)}\}$ and write

$$\begin{aligned}
& E\Big|N^{-1}\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon^2_{t\nu+i-j}-\gamma_i(\ell)\Big| \\
= \;& E\Big|\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\Big[N^{-1}\sum_{t=0}^{N-1}(\varepsilon^2_{t\nu+i-j}-1)\Big]\Big| \\
\le \;& \sum_{j=-\infty}^{\infty}|\psi_i(j)\psi_{i+\ell}(j+\ell)|E\Big|N^{-1}\sum_{t=0}^{N-1}(\varepsilon^2_{t\nu+i-j}-1)\Big| \\
\le \;& \Big(\sum_{j=-\infty}^{\infty}|\psi_i(j)|\Big)\Big(\sum_{j=-\infty}^{\infty}|\psi_{i+\ell}(j)|\Big)B_0a_N^2/N \\
= \;& Ba_N^2/N.
\end{aligned}$$

Finally, we have

$$\begin{aligned}
E|\hat\gamma_i(\ell)-\gamma_i(\ell)| \le \;& \\
& E\Big|\hat\gamma_i(\ell)-N^{-1}\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon^2_{t\nu+i-j}\Big| \\
+ \;& E\Big|N^{-1}\sum_{t=0}^{N-1}\sum_{j=-\infty}^{\infty}\psi_i(j)\psi_{i+\ell}(j+\ell)\varepsilon^2_{t\nu+i-j}-\gamma_i(\ell)\Big| \\
\le \;& K_0N^{1/2}a_N^{-2}N^{-1}a_N^2+Ba_N^2N^{-1} \\
= \;& K_0N^{-1/2}+Ba_N^2N^{-1}
\end{aligned}$$

where a_N^2/N is regularly varying with index $\frac{2}{\alpha}-1$. For $2<\alpha<4$, $N^{-1/2}=o(a_N^2/N)$. Hence, $E|\hat\gamma_i(\ell)-\gamma_i(\ell)| \le Ca_N^2/N$.

Acknowledgements

The authors thank the referees for many helpful suggestions which greatly improved the manuscript.

REFERENCES

Adams, G. and C. Goodwin (1995) Parameter estimation for periodic ARMA models, *J. Time Series Anal.*, 16, 127–145.

Anderson, P. (1989) *Asymptotic Results and Identification for Cyclostationary Times Series*, Doctoral Dissertation, Colorado School of Mines, Golden, Colorado.

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

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

Anderson, P., Meerschaert, M. and A. Vecchia (1998) Asymptotics of the innovations algorithm for periodic time series, in preparation.

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

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

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

Billingsley (1968) *Convergence of Probability Measures*, Wiley, New York.

Billingsley (1995) *Probability and Measure*, 3rd Ed., Wiley, New York.

Bingham, N., C. Goldie, and J. Teugels (1987) *Regular Variation*, Encyclopedia of Mathematics and its Applications, 27, Cambridge University Press.

Brockwell, P. and R. Davis (1988) Simple consistent estimation of the coefficients of a linear filter, *Stoch. Proc. Appl.*, 28, 47–59.

Brockwell, P. and R. Davis (1991) *Time Series: Theory and Methods*, 2nd Ed., Springer–Verlag, New York.

Davis, R. and S. Resnick (1986) Limit theory for the sample covariance and correlation functions of moving averages, *Ann. Statist.*, 14, 533–558.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd Ed., Wiley, New York.

Golub, G. and C. Van Loan (1989) Matrix Computations, 2nd Ed., Johns Hopkins University Press.

Jansen, D. and C. de Vries (1991) On the frequency of large stock market returns: Putting booms and busts into perspective, *Review of Econ. and Statist.*, 23, 18–24.

Jones, R. and W. Brelsford (1967) Time series with periodic structure, *Biometrika*, 54, 403–408.

Loretan, M. and P. Phillips (1994) Testing the covariance stationarity of heavy–tailed time series, *J. Empirical Finance*, 211-248.

Lund, R. and I. Basawa (1999) Recursive Prediction and Likelihood Evaluation for Periodic ARMA Models, *J. Time Series Anal.*, to appear.

Mikosch, T., T. Gadrich, C. Klüppenberg and R. Adler (1995) Parameter estimation for ARMA models with infinite variance innovations, *Ann. Statist.*, 23, 305-326.

Pagano, M. (1978) On periodic and multiple autoregressions, *Ann. Statist.* 6, 1310–1317.

Salas, J., G. Tabios and P. Bartolini (1985) Approaches to multivariate modeling of water resources time series, *Water Res. Bull.* 21, 683–708.

Samorodnitsky, G. and M. Taqqu (1994) *Stable non–Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, London.

Tiao, G. and M. Grupe (1980) Hidden periodic autoregressive–moving average models in time series data, *Biometrika*, 67, 365–373.

Tjostheim, D. and J. Paulsen (1982) Empirical identification of multiple time series, *J. Time Series Anal.*, 3, 265–282.

Troutman, B. (1979) Some results in periodic autoregression, *Biometrika*, 6, 219–228.

Ula, T. (1993) Forecasting of multivariate periodic autoregressive moving average processes, *J. Time Series Anal.*, 14, 645.

Vecchia, A. and R. Ballerini (1991) Testing for periodic autocorrelations in seasonal time series data, *Biometrika*, 78, 53–63.