Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 3, 640–655
DOI: 10.1214/14-BJPS237

Asymptotic distribution of the estimated parameters of an ARMA(p, q) process with mixing innovations

Sankha Bhattacharya[a] **and Sugata Sen Roy**[b]

[a]*R.K.M.R. College*
[b]*University of Calcutta*

Abstract. In this paper, we consider an ARMA(p, q) model with stationary, ϕ-mixing error variables having uniformly bounded fourth-order moments. Both the autoregressive and moving average components of the model involve stable and explosive roots. Estimating the autoregressive parameters using the instrumental variable technique and the moving average parameters using a derived autoregressive process, we derive the asymptotic distribution of the estimators.

1 Introduction

The limiting distribution of the least squares estimators of an autoregressive process with identically and independently distributed (i.i.d.) errors have been studied by several authors like Mann and Wald (1943), White (1958), Anderson (1959), Jeganathan (1988) and Chan and Wei (1988) for both the stable and explosive roots. Basu and Sen Roy (1993) considered all forms of the roots and derived the asymptotic distribution of the estimator assuming ϕ-mixing error variables.

However, very few such studies have been extended to an ARMA(p, q) model. In a recent paper (Sen Roy and Bhattacharya, 2012), we had derived the asymptotic distribution of the estimators of the parameters of a model with i.i.d. innovations and having both stable and explosive roots. In the present paper, we seek to extend those results to a model with dependent innovations. Since the ordinary least squares estimator of the AR parameters is inconsistent even for the i.i.d. case, we use the instrumental variable technique to estimate the autoregressive parameters and a derived autoregressive process to estimate the moving average parameters.

Consider the ARMA(p, q) model,

$$\mathrm{X}_t - \alpha_1 \mathrm{X}_{t-1} - \alpha_2 \mathrm{X}_{t-2} - \cdots - \alpha_p \mathrm{X}_{t-p} = e_t - \beta_1 e_{t-1} - \cdots - \beta_q e_{t-q}, \quad (1.1)$$

where X_t is the observation at time t, $t = 1, \ldots, N$ and e_t is a stationary ϕ-mixing sequence with mean zero, $E(e_t^2) = \sigma^2$ and

$$E(e_t^4) < \infty. \quad (1.2)$$

Key words and phrases. ARMA process, explosive roots, ϕ-mixing errors, asymptotic distribution.

Received May 2013; accepted January 2014.

The ϕ-mixing function $\phi(n)$ is a decreasing function of n, with $\sum_{n=1}^{\infty}\phi(n)^{1/2} < \infty$. This means that the dependence between the errors decreases as the distance between the corresponding time points increases. Also the initial conditions are assumed to be zero, that is, $e_t = 0$ for $t \leq 0$.

The autoregressive (AR) component is stable or explosive according as the roots of the characteristic polynomial $\Phi(z) = 1 - \alpha_1 z - \alpha_2 z^2 - \cdots - \alpha_p z^p$ are greater than or less than unity in absolute value. Similarly, the moving average (MA) component is stable or explosive according as the roots of the characteristic polynomial $\Theta(z) = 1 - \beta_1 z - \beta_2 z^2 - \cdots - \beta_q z^q$ are greater than or less than unity in absolute value.

Using a backward shift operator B, model (1.1) can be rewritten as

$$\Phi(\mathrm{B})\mathrm{X}_t = \Theta(\mathrm{B})e_t. \tag{1.3}$$

Here, we study the asymptotic distribution of the ARMA(p,q) process as defined in (1.1) under the above conditions. A problem here is that under these conditions even the instrumental variable estimator is inconsistent. To circumvent this difficulty, the condition

$$\mathrm{E}(\mathbf{x}_t e_j) = \mathbf{0}_p \qquad \text{for all } t \text{ and integers } j > t, \tag{1.4}$$

where $\mathbf{x}_t = (\mathrm{X}_t, \mathrm{X}_{t-1}, \ldots, \mathrm{X}_{t-p+1})'$ and $\mathbf{0}_n$ is a n-dimensional vector of zero elements, needs to be imposed.

In practice, this means that if $\xi(h)$ is the hth-order autocovariance function of e_t,

$$\Theta(\mathrm{B})\xi(h) = 0, \tag{1.5}$$

that is, (1.4) translates into a restriction on the autocovariance function of e_t. A particular and plausible choice of $\xi(h)$ is

$$\xi(h) = \xi^h, \qquad 0 < \xi < 1,$$

that is, $\xi(h)$ is exponentially decreasing in h.

In studying the limiting distribution, a component-wise break-up according to stable and explosive roots is made using techniques similar to that of Chan and Wei (1988). Then using suitably chosen norming matrices, the limiting distribution of each component is found separately. The results are then put together in the final theorem. However, since the norming matrices involve the parameters of the model, it is further shown that the asymptotic results hold even if these parameters are substituted by their estimators.

Since some of the results are similar to those for the i.i.d. case, we simply state such results for the sake of completeness and omit their proofs. In Section 2, a componentwise break up of the process is made. Section 3 considers the asymptotic distributions of the estimators componentwise, while Section 4 contains the main theorem. Some concluding remarks are made in Section 5.

In the sequel I_n denotes an identity matrix of order n. $\mathrm{diag}((\cdot))$ denotes a block diagonal matrix. $\sim$ implies "asymptotically equivalent to." The norm of a vector refers to Euclidean norm, while for a matrix A, $\|\mathrm{A}\| = \sup_{\|x\|=1} \|\mathrm{A}x\|$. c_i's, $i = 0, 1, \ldots$ denote constants.

2 A componentwise break-up of the process

For $r + s = p$ and $|\rho_i| > 1$, $i = 1, 2, \ldots, r$ and $|\gamma_j| < 1$, $j = 1, 2, \ldots, s$, $\Phi(z)$ can be rewritten as

$$\Phi(z) = \prod_{i=1}^{r} (1 - \rho_i^{-1} z) \prod_{j=1}^{s} (1 - \gamma_j^{-1} z), \tag{2.1}$$

where ρ_i are the r stable roots and γ_j are the s explosive roots of $\Phi(z) = 0$.

Similarly, $\Theta(z)$ can be written as

$$\Theta(z) = \prod_{i=1}^{c} (1 - \pi_i^{-1} z) \prod_{j=1}^{d} (1 - \eta_j^{-1} z), \tag{2.2}$$

where π_i are the stable roots and η_j are the explosive roots of $\Theta(z)$, with $|\pi_i| > 1$, $i = 1, 2, \ldots, c$, $|\eta_j| < 1$, $j = 1, 2, \ldots, d$ and $c + d = q$. All roots are assumed to be distinct.

Model (1.3) can be rewritten as

$$\Phi(\mathrm{B})\mathrm{X}_t = u_t, \tag{2.3}$$

where

$$u_t = \Theta(\mathrm{B}) e_t \tag{2.4}$$

is a MA(q) process.

Defining $\mathbf{u}_t = (u_t, \mathbf{0}'_{p-1})'$, and $\mathbf{A} = \begin{pmatrix} \alpha_1 \ \ldots \ \alpha_{p-1} & \alpha_p \\ \mathrm{I}_{p-1} & \mathbf{0}_{p-1} \end{pmatrix}$, (2.3) can be rewritten as

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{u}_t, \qquad t = 1, 2, \ldots. \tag{2.5}$$

Since $\mathbf{x}_{t-1}$ is correlated with u_t through $e_{t-1}, \ldots, e_{t-q}$ the least squares estimator of the AR parameter $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_p)'$ will be inconsistent. Taking $n = N - q - 1$ and following Basu et al. (2005), the instrumental variable estimator of $\boldsymbol{\alpha}$ is

$$\hat{\boldsymbol{\alpha}}_n = \left(\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}'_{t+q} \right)^{-1} \left(\sum_{t=1}^{n} \mathbf{x}_t \mathrm{X}_{t+q+1} \right). \tag{2.6}$$

To estimate the parameters of the MA component, let $\mathrm{Y}_{t-i} = \frac{de_t}{d\beta_i}$, be the partial derivative of e_t with respect to β_i. Then following Tsay (1993), we obtain the derived AR(q) process

$$\Theta(\mathrm{B})\mathrm{Y}_t = e_t, \qquad t = 1, 2, 3, \ldots. \tag{2.7}$$

Defining $\mathbf{y}_t = (\mathrm{Y}_t, \ldots, \mathrm{Y}_{t-q+1})'$, $\mathbf{v}_t = (e_t, \mathbf{0}'_{q-1})'$, and $\mathbf{C} = \begin{pmatrix} \beta_1 \ \ldots \ \beta_{q-1} & \beta_q \\ \mathrm{I}_{q-1} & \mathbf{0} \end{pmatrix}$, (2.7) can be rewritten as

$$\mathbf{y}_t = \mathbf{C}\mathbf{y}_{t-1} + \mathbf{v}_t, \qquad t = 1, 2, 3, \ldots. \tag{2.8}$$

Then the least squares estimator of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_q)'$, based on n observations, is

$$\hat{\boldsymbol{\beta}}_n = \left(\sum_{t=1}^{n} \mathbf{y}_{t+q}\mathbf{y}'_{t+q}\right)^{-1} \left(\sum_{t=1}^{n} \mathbf{y}_{t+q}\mathrm{Y}_{t+q+1}\right). \tag{2.9}$$

Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}')'$, $\hat{\boldsymbol{\theta}}_n = (\hat{\boldsymbol{\alpha}}'_n, \hat{\boldsymbol{\beta}}'_n)'$, $\mathbf{z_t} = (\mathbf{x}'_t u_{t+q+1}, \mathbf{y}'_{t+q} e_{t+q+1})'$, and

$$\mathbf{D}_n = \begin{pmatrix} \sum_{t=1}^{n} \mathbf{x}_t\mathbf{x}'_{t+q} & \mathbf{0} \\ \mathbf{0} & \sum_{t=1}^{n} \mathbf{y}_{t+q}\mathbf{y}'_{t+q} \end{pmatrix}.$$

Then

$$(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \mathbf{D}_n^{-1}\left(\sum_{t=1}^{n} \mathbf{z}_t\right). \tag{2.10}$$

Denoting by B the backshift operator, the different components are segregated as

$$\mathrm{R}_t = \Phi(\mathrm{B})\prod_{i=1}^{r}(1 - \rho_i^{-1}\mathrm{B})^{-1}\mathrm{X}_t, \tag{2.11}$$

$$\mathrm{S}_t = \Phi(\mathrm{B})\prod_{i=1}^{s}(1 - \gamma_i^{-1}\mathrm{B})^{-1}\mathrm{X}_t, \tag{2.12}$$

$$\mathrm{Q}_t = \Theta(\mathrm{B})\prod_{i=1}^{c}(1 - \pi_i^{-1}\mathrm{B})^{-1}\mathrm{Y}_t, \tag{2.13}$$

$$\mathrm{P}_t = \Theta(\mathrm{B})\prod_{i=1}^{d}(1 - \eta_i^{-1}\mathrm{B})^{-1}\mathrm{Y}_t. \tag{2.14}$$

Let $\mathbf{r}_t = (\mathrm{R}_t, \ldots, \mathrm{R}_{t-r+1})$, $\mathbf{s}_t = (\mathrm{S}_t, \ldots, \mathrm{S}_{t-s+1})$, $\mathbf{q}_t = (\mathrm{Q}_t, \ldots, \mathrm{Q}_{t-c+1})$ and $\mathbf{p}_t = (\mathrm{P}_t, \ldots, \mathrm{P}_{t-d+1})$. Following (2.1) and (2.11), R_t can be written as

$$\mathrm{R}_t = \prod_{i=1}^{s}(1 - \gamma_i^{-1}\mathrm{B})\mathrm{X}_t = \mathrm{X}_t - \gamma_1^*\mathrm{X}_{t-1} - \cdots - \gamma_s^*\mathrm{X}_{t-s} \tag{2.15}$$

so that for the $r \times p$ matrix

$$\mathrm{T}_1 = \begin{pmatrix} 1 & -\gamma_1^* & \ldots & -\gamma_s^* & 0 & 0 & 0 & 0 \\ 0 & 1 & -\gamma_1^* & \ldots & -\gamma_s^* & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & 1 & -\gamma_1^* & \ldots & -\gamma_s^* \end{pmatrix},$$

$\mathrm{T}_1\mathbf{x}_t = \mathbf{r}_t$. Similarly, following (2.1) and (2.12) we may find a $s \times p$ matrix T_2 so that $\mathrm{T}_2\mathbf{x}_t = \mathbf{s}_t$. Hence, there exists a $p \times p$ matrix $\mathrm{T}^{(1)} = (\mathrm{T}_1', \mathrm{T}_2')'$ such that $\mathrm{T}^{(1)}\mathbf{x}_t = (\mathbf{r}_t', \mathbf{s}_t')'$. Similarly, following (2.2) and (2.13) we may define a $c \times q$ matrix T_3 for which $\mathrm{T}_3\mathbf{y}_t = \mathbf{q}_t$ and following (2.2) and (2.14) a $d \times q$ matrix T_4 such that $\mathrm{T}_4\mathbf{y}_t = \mathbf{p}_t$. Combining these, we define the $q \times q$ matrix $\mathrm{T}^{(2)} = (\mathrm{T}_3', \mathrm{T}_4')'$ with $\mathrm{T}^{(2)}\mathbf{y}_t = (\mathbf{q}_t', \mathbf{p}_t')'$. Finally, let $\mathbf{T} = \mathrm{diag}(\mathrm{T}^{(1)}, \mathrm{T}^{(2)})$. We next derive the componentwise limiting distributions.

3 Componentwise asymptotic distributions

3.1 The AR stable component

We first consider the stable component of the autoregressive part, $\prod_{i=1}^{r}(1 - \rho_i^{-1}\mathrm{B})\mathrm{R}_t = u_t$. Following (2.11), this can be reconstructed as

$$\mathrm{R}_t = \rho_1^*\mathrm{R}_{t-1} + \rho_2^*\mathrm{R}_{t-2} + \cdots + \rho_r^*\mathrm{R}_{t-r} + u_t, \tag{3.1}$$

where $\boldsymbol{\rho}^* = (\rho_1^*, \rho_2^*, \ldots, \rho_r^*)$ are the parameters of the process with roots ρ_j, $j = 1, \ldots, r$. Define $\mathbf{L}_1 = \begin{pmatrix} \rho_1^* \cdots \rho_{r-1}^* & \rho_r^* \\ \mathrm{I}_{r-1} & \mathbf{0}_{r-1} \end{pmatrix}$ and $\mathbf{u}_{1t} = (u_t, \mathbf{0}_{r-1}')$.

Then (3.1) can be rewritten as

$$\mathbf{r}_t = \mathbf{L}_1\mathbf{r}_{t-1} + \mathbf{u}_{1t}, \qquad t = 1, 2, 3, \ldots. \tag{3.2}$$

Let $\check{\rho}_1 = \max_{1\le j\le r} |\rho_j^{-1}| < 1$. Then

$$\|\mathbf{L}_1^n\| \sim c_0\check{\rho}_1^n \qquad \text{as } n \to \infty. \tag{3.3}$$

Let $\mathbf{J}_n = n^{-1/2}\mathrm{I}_n$ and $\Sigma_1 = \mathrm{E}(\mathbf{r}_n\mathbf{r}_{n+q}')$, Σ_1 positive definite. Define $\mathbf{w}_t = \mathbf{r}_t' u_{t+q+1}$ and $\mathbf{R}_n = n^{-1}\sum_{t=1}^{n} \mathbf{r}_t\mathbf{r}_{t+q}'$.

Lemma 3.1. *Under (3.3) and bounded fourth-order moments of the innovations,*

$$n^{-1/2}\sum_{t=1}^{n} \mathbf{w}_t \xrightarrow{d} \mathrm{N}(0, \Sigma_1^*), \tag{3.4}$$

where

$$\Sigma_1^* = \mathrm{E}(\mathbf{w}_1\mathbf{w}_1') + \sum_{k=1}^{\infty} \mathrm{E}(\mathbf{w}_1\mathbf{w}_{k+1}') + \sum_{k=1}^{\infty} \mathrm{E}(\mathbf{w}_{k+1}\mathbf{w}_1') \tag{3.5}$$

and the elements of Σ_1^* are convergent.

Proof. Similar to that of Sen Roy and Bhattacharya (2012). □

Lemma 3.2. *Under fourth-order bounded moment condition of the innovations, for any constant c_1 and for all $\varepsilon > 0$,*

$$P\big[\|\mathbf{R}_n - \Sigma_1\| > \varepsilon\big] < c_1 n^{-1}\varepsilon^{-1}. \tag{3.6}$$

Proof. Similar to that of Sen Roy and Bhattacharya (2012). □

Theorem 3.1. *Under conditions* (1.2) *and* (3.3),

$$\text{(i)} \quad \mathbf{J}_n \sum_{i=1}^{n} \mathbf{r}_t \mathbf{r}'_{t+q} \mathbf{J}'_n \xrightarrow{p} \Sigma_1 \tag{3.7}$$

and

$$\text{(ii)} \quad (\mathbf{J}'_n)^{-1} \left(\sum_{t=1}^{n} \mathbf{r}_t \mathbf{r}'_{t+q} \right)^{-1} \left(\sum_{t=1}^{n} \mathbf{r}_t u_{t+q+1} \right) \xrightarrow{d} \mathrm{N}_r(0, \Sigma_1^{-1} \Sigma_1^* \Sigma_1^{-1}). \tag{3.8}$$

Proof. The proof follows from Lemmas 3.1 and 3.2. □

3.2 The AR explosive component

Next, consider the explosive component of the autoregressive part, $\prod_{i=1}^{s}(1 - \gamma_i^{-1}\mathrm{B})\mathrm{S}_t = u_t$ which from (2.12) can be rewritten as

$$\mathrm{S}_t = \gamma_1^* \mathrm{S}_{t-1} + \gamma_2^* \mathrm{S}_{t-2} + \cdots + \gamma_s^* \mathrm{S}_{t-s} + u_t \qquad \text{for } t = 1, 2, \ldots, \tag{3.9}$$

where $\boldsymbol{\gamma}^* = (\gamma_1^*, \ldots, \gamma_s^*)$ are the parameters of the process with roots γ_j for $j = 1, 2, \ldots, s$. Defining $\mathbf{F} = \begin{pmatrix} \gamma_1^* \cdots \gamma_{s-1}^* & \gamma_s^* \\ \mathrm{I}_{s-1} & \mathbf{0} \end{pmatrix}$ and $\mathbf{u}_{2t} = (u_t, \mathbf{0}'_{s-1})'$, the model (3.22) can be rewritten as

$$\mathbf{s}_t = \mathbf{F}\mathbf{s}_{t-1} + \mathbf{u}_{2t}, \qquad t = 1, 2, \ldots. \tag{3.10}$$

Let $\check{\gamma}_1 = \min_{1 \le j \le s} |\gamma_j^{-1}| > 1$ and $\check{\gamma}_2 = \max_{1 \le j \le s} |\gamma_j^{-1}| > 1$. Then $\|\mathbf{F}^n\| \sim c_2 \check{\gamma}_2^n$ and

$$\|\mathbf{F}^{-n}\| \sim c_3 \check{\gamma}_1^{-n} \qquad \text{as } n \to \infty. \tag{3.11}$$

Let

$$\mathbf{s}_n^* = \mathbf{F}^{-(n-1)}\mathbf{s}_n = \sum_{t=1}^{n} \mathbf{F}^{-(t-1)}\mathbf{u}_{2t} = \sum_{t=1}^{n} \mathrm{f}_t u_{2t}, \tag{3.12}$$

where f_t denotes the first column of $\mathbf{F}^{-(t-1)}$.

Following Longnecker and Serfling (1978), and because of (1.2) and

$$\sum_{t=1}^{\infty} \|\mathbf{F}^{-t}\| < \infty,$$

it follows that $\mathbf{s}_n^*$ converges a.s. Let

$$\lim_{n \longrightarrow \infty} \mathbf{s}_n^* = \mathbf{s}^* = \sum_{t=1}^{\infty} \mathbf{F}^{-(t-1)}\mathbf{u}_{2t}. \tag{3.13}$$

The next two lemmas are similar to those of Sen Roy and Bhattacharya (2012).

Lemma 3.3. $\mathbf{s}_n^* \xrightarrow{L_2} \mathbf{s}^*$, *and hence* $\mathbf{s}_n^* \xrightarrow{p} \mathbf{s}^*$.

Lemma 3.4. *For* $\mathbf{d}_n = \mathbf{F}^{-(n-1)} \sum_{t=1}^n \mathbf{s}_t u_{t+q+1}$ *and* $\mathbf{h}_n = \sum_{t=1}^n \mathbf{F}^{-(t-1)} \mathbf{s}_n^* u_{n+q+2-t}$,

$$\mathbf{d}_n - \mathbf{h}_n \xrightarrow{p} 0.$$

Let $\mathbf{K}$ *be a nonsingular matrix such that* $\mathbf{KFK}^{-1} = \mathrm{diag}(\gamma_1^{-1}, \ldots, \gamma_s^{-1})$. *Writing* $\mathbf{G} = \mathrm{diag}(\gamma_1, \ldots, \gamma_s)$, *we have* $\mathbf{F}^{-n} = \mathbf{K}^{-1}\mathbf{G}^n\mathbf{K}$ *where*

$$\|\mathbf{G}^n\| \sim c_4 \check{\gamma}_1^{-n} \qquad \text{as } n \to \infty. \tag{3.14}$$

Also let $\mathbf{S}_n$ *and* $\mathbf{S}$ *be* $s \times s$ *diagonal matrices with* i*th diagonal element equal to the* i*th element of* $\mathbf{Ks}_n^*$ *and* $\mathbf{Ks}^*$, *respectively, and let* $\boldsymbol{\vartheta}_n = (v_1, \ldots, v_s)'$, *where* $v_j = \sum_{i=1}^n \gamma_j^{(i-1)} u_{n+q+2-i}$ *for* $j = 1, 2, \ldots, s$. *Then* $\mathbf{h}_n$ *can be written in the form*

$$\mathbf{h}_n = \mathbf{K}^{-1} \sum_{t=1}^n \mathbf{G}^{t-1} \mathbf{Ks}_n^* u_{n+q+2-t} = \mathbf{K}^{-1}\mathbf{S}_n\boldsymbol{\vartheta}_n. \tag{3.15}$$

Define the $s \times s$ *diagonal matrix* $\mathbf{S}_n^*$ *with* i*th diagonal element equal to the* i*th element of* $\mathbf{S}_n^{**} = \mathbf{K}\sum_{t=1}^{[n/3]} \mathrm{f}_t u_t$ *and* $\boldsymbol{\vartheta}_n^* = (v_1^*, v_2^*, \ldots, v_s^*)$, *where for* $j = 1, 2, \ldots, s$ $v_j^* = \sum_{i=1}^{[n/3]} \gamma_j^{(i-1)} u_{n+q+2-i}$. *Here,* $\mathbf{S}_n^*$ *and* $\boldsymbol{\vartheta}_n^*$ *are partial sums consisting of only* $[n/3]$ *of the* u_i*'s. However,* $\bar{\mathbf{S}}_n^*$ *depends on the first* $[n/3]$ *observations of* u_t, *while* $\boldsymbol{\vartheta}_n^*$ *depends on the last* $[n/3]$ *observations.* $\mathbf{S}_n^*$ *and* $\boldsymbol{\vartheta}_n^*$ *are separated by* $[n/3] + q + 1$ *intervening* u_i*'s.*

Lemma 3.5. $\mathbf{S}_n$ *and* $\boldsymbol{\vartheta}_n$ *are asymptotically independent.*

Proof. Under bounded second-order moment of u_t's and since u_t's of $\mathbf{S}_n^{**}$ are separated from those of $\boldsymbol{\vartheta}_n^*$ by at least length $[n/3]+q+1$, using the lemma (page 170) of Billingsley (1968),

$$\begin{aligned}
&\|\mathrm{E}(\mathbf{S}_n^{**}\boldsymbol{\vartheta}_n^{*\prime})\| \\
&\quad \le \|\mathbf{K}\| \sum_{i=1}^{[n/3]} \sum_{j=[2n/3]+1}^{n} \|d_i\| \|\mathbf{G}^{n-j}\| |\mathrm{E}(u_i u_{q+1+j})| \\
&\quad \le \|\mathbf{K}\| \sum_{i=1}^{[n/3]} \sum_{j=[2n/3]+1}^{n} \|d_i\| 2\phi^{1/2}([n/3]+q+1)\mathrm{E}(u_i^2)\|\mathbf{G}^{n-j}\| \\
&\quad \le c_5 \left(\sum_{i=1}^{[n/3]} \|d_i\|\right)\left(\sum_{j=[2n/3]+1}^{n} \|\mathbf{G}^{n-j}\|\right) \to 0 \qquad \text{as } n \to \infty.
\end{aligned} \tag{3.16}$$

Since $\mathrm{E}(\mathbf{S}_n^{**}) = 0$ and $\mathrm{E}(\boldsymbol{\vartheta}_n^*) = 0$, $\mathbf{S}_n^*$ and $\boldsymbol{\vartheta}_n^*$ are asymptotically uncorrelated.

Now, following (1.2) and (3.11), we have

$$\begin{aligned}\mathrm{E}\|(\mathbf{S}_n-\mathbf{S}_n^*)(\mathbf{S}_n-\mathbf{S}_n^*)'\| &\le \mathrm{E}\left(\sum_{i=[n/3]+1}^{n}\|\mathbf{K}\|\|\mathbf{F}^{-(i-1)}\|\|\mathbf{u}_{2i}\|\right)^2\\ &\le c_6\|\mathbf{F}^{-2[n/3]}\|\to 0 \qquad \text{as } n\to\infty.\end{aligned}$$

Hence, $\mathbf{S}_n-\mathbf{S}_n^*\xrightarrow{L_2}0$ which implies $\mathbf{S}_n-\mathbf{S}_n^*\xrightarrow{p}0$.

Following (1.2) and (3.14),

$$\begin{aligned}\mathrm{E}\|(\boldsymbol{\vartheta}_n-\boldsymbol{\vartheta}_n^*)(\boldsymbol{\vartheta}_n-\boldsymbol{\vartheta}_n^*)'\| &\le \mathrm{E}\left(\sum_{i=1}^{n-[n/3]}\|\mathbf{G}^{n-i}\|\|u_{q+1+i}\|\right)^2\\ &\le c_7\|\mathbf{G}^{2[n/3]}\|\to 0 \qquad \text{as } n\to\infty.\end{aligned}$$

Hence, $\boldsymbol{\vartheta}_n-\boldsymbol{\vartheta}_n^*\xrightarrow{L_2}0$ which implies $\boldsymbol{\vartheta}_n-\boldsymbol{\vartheta}_n^*\xrightarrow{p}0$.

Since $\mathbf{S}_n^*$ and $\boldsymbol{\vartheta}_n^*$ are Gaussian, they are asymptotically independent. Hence, $\mathbf{S}_n$ and $\boldsymbol{\vartheta}_n$ are asymptotically independent. □

Lemma 3.6. $\mathbf{S}_n^*\xrightarrow{L_2}\mathbf{S}$ *and* $\boldsymbol{\vartheta}_n^*\xrightarrow{L_2}\boldsymbol{\vartheta}$, *where* $\boldsymbol{\vartheta}=(\bar{v}_1,\bar{v}_2,\dots,\bar{v}_s)$, *with* $\bar{v}_j=\sum_{i=1}^{\infty}\gamma_j^{(i-1)}u_{n+q+2-i}$ *for* $j=1,2,\dots,s$.

Proof. The proof is similar to that of Lemma 3.5. □

Next, define

$$\Gamma=\begin{pmatrix}(1-\gamma_1^2)^{-1} & (1-\gamma_1\gamma_2)^{-1} & \dots & (1-\gamma_1\gamma_s)^{-1}\\ \dots & \dots & \dots & \dots\\ (1-\gamma_1\gamma_s)^{-1} & (1-\gamma_2\gamma_s)^{-1} & \dots & (1-\gamma_s^2)^{-1}\end{pmatrix}$$

and $\mathbf{F}^*=\sum_{i=1}^{\infty}\mathbf{F}^{-(i-1)}\mathbf{s}^*\mathbf{s}^{*\prime}\mathbf{F}^{-(i-1)\prime}$. Then with $\gamma^{(i-1)}=(\gamma_1^{i-1},\dots,\gamma_s^{i-1})'$ we observe that

$$\begin{aligned}\mathbf{F}^* &=\sum_{i=1}^{\infty}\mathbf{K}^{-1}\mathbf{G}^{(i-1)}\mathbf{K}\mathbf{s}^*\mathbf{s}^{*\prime}\mathbf{K}'\mathbf{G}^{(i-1)\prime}\mathbf{K}^{-1\prime}\\ &=\mathbf{K}^{-1}\sum_{i=1}^{\infty}\mathbf{S}\gamma^{(i-1)}\gamma^{(i-1)\prime}\mathbf{S}'\mathbf{K}^{-1\prime}=\mathbf{K}^{-1}\mathbf{S}\Gamma\mathbf{S}'\mathbf{K}^{-1\prime}.\end{aligned}\tag{3.17}$$

Taking $\mathbf{K}_n=\mathbf{F}^{-(n+q-1)}$, we have the following theorem.

Theorem 3.2. *Under* (1.2) *and* (3.11),

$$\text{(i)}\quad \mathbf{K}_{n+q-1}\sum_{t=1}^{n}\mathbf{s}_t\mathbf{s}_{t+q}'\mathbf{K}_n'\xrightarrow{p}\mathbf{F}^*.\tag{3.18}$$

If in addition e_t's are Gaussian, $\mathbf{F}^*$ *is positive definite a.s. and*

$$\text{(ii)} \quad (\mathbf{K}_n')^{-1}\left(\sum_{t=1}^{n}\mathbf{s}_t\mathbf{s}_{t+q}'\right)^{-1}\left(\sum_{t=1}^{n}\mathbf{s}_t u_{t+q+1}\right) \xrightarrow{d} \mathrm{N}_1^*, \tag{3.19}$$

where $\mathrm{N}_1^* = \mathbf{K}'\mathbf{S}^{-1}\Gamma^{-1}\boldsymbol{\vartheta}$, $\boldsymbol{\vartheta}$ *being a s-variate Gaussian variable with mean zero and dispersion matrix* $\mathbf{V} = ((v_{ij}))$ *with*

$$v_{ij} = \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\gamma_i^{(k-1)}\gamma_j^{(l-1)}\mathrm{E}(u_{n+q+2-k}u_{n+q+2-l}).$$

Also $\boldsymbol{\vartheta}$ *is independent of* $\mathbf{K}'\mathbf{S}^{-1}\Gamma^{-1}$.

Proof. Under Lemmas 3.3–3.6, the proof follows similarly as in Sen Roy and Bhattacharya (2012). □

3.3 The MA stable component

Following (2.13), the stable part of the moving average component, $\prod_{i=1}^{c}(1 - \pi_i^{-1}\mathrm{B})\mathrm{Q}_t = e_t$, can be rewritten as

$$\mathrm{Q}_t = \pi_1^*\mathrm{Q}_{t-1} + \cdots + \pi_c^*\mathrm{Q}_{t-c} + e_t, \tag{3.20}$$

where $\boldsymbol{\pi}^* = (\pi_1^*, \ldots, \pi_c^*)$ are the parameters of the process with roots π_j, $j = 1, \ldots, c$. Defining $\mathbf{L}_2 = \begin{pmatrix} \pi_1^* \ \ldots \ \pi_{c-1}^* & \pi_c^* \\ \mathrm{I}_{c-1} & \mathbf{0}_{c-1} \end{pmatrix}$ and $\mathbf{v}_{1t} = (e_t, \mathbf{0}_{c-1}')'$, model (3.20) reduces to

$$\mathbf{q}_t = \mathbf{L}_2\mathbf{q}_{t-1} + \mathbf{v}_{1t}, \qquad t = 1, 2, \ldots. \tag{3.21}$$

Let $\check{\pi}_1 = \max_{1\le j\le c}|\pi_j^{-1}| < 1$. Then

$$\|\mathbf{L}_2^n\| \sim c_8\check{\pi}_1^n \qquad \text{as } n \to \infty. \tag{3.22}$$

Let $\mathbf{M}_n = n^{-1/2}\mathrm{I}_n$ and $\Sigma_2 = \mathrm{E}(\mathbf{q}_n\mathbf{q}_n')$.

Theorem 3.3. *Under the conditions* (1.2) *and* (3.22),

$$\text{(i)} \quad \mathbf{M}_n\sum_{i=1}^{n}\mathbf{q}_{t+q}\mathbf{q}_{t+q}'\mathbf{M}_n' \xrightarrow{p} \Sigma_2 \tag{3.23}$$

and

$$\begin{aligned}\text{(ii)} \quad (\mathbf{M}_n')^{-1}\left(\sum_{i=1}^{n}\mathbf{q}_{t+q}\mathbf{q}_{t+q}'\right)^{-1}\left(\sum_{i=1}^{n}\mathbf{q}_{t+q}e_{t+q+1}\right) \\ \xrightarrow{d} \mathrm{N}_c(0, \Sigma_2^{-1}\Sigma_2^*\Sigma_2^{-1}),\end{aligned} \tag{3.24}$$

where $\Sigma_2^* = \mathrm{E}(\mathbf{q}_{q+1}\mathbf{q}_{q+1}'e_{q+2}^2)$.

Proof. The proof is similar to that of Theorem 3.1. □

3.4 The MA explosive component

From (2.14), the explosive component of the moving average part $\prod_{i=1}^{d}(1 - \eta_i^{-1}\mathrm{B})\mathrm{P}_t = e_t$ can be rewritten as

$$\mathrm{P}_t = \eta_1^*\mathrm{P}_{t-1} + \cdots + \eta_d^*\mathrm{P}_{t-d} + e_t \qquad \text{for } t = 1, 2, \ldots, \tag{3.25}$$

where $\boldsymbol{\eta}^* = (\eta_1^*, \ldots, \eta_d^*)$ are the parameters of the process with roots η_j for $j = 1, 2, \ldots, d$. Define, $\widetilde{\mathbf{F}} = \begin{pmatrix} \eta_1^* \cdots \eta_{d-1}^* & \eta_d^* \\ \mathrm{I}_{d-1} & \mathbf{0}_{d-1} \end{pmatrix}$ and $\mathbf{v}_{2t} = (e_t, \mathbf{0}'_{d-1})'$, and rewrite (3.25) as

$$\mathbf{p}_t = \widetilde{\mathbf{F}}\mathbf{p}_{t-1} + \mathbf{v}_{2t}, \qquad t = 1, 2, \ldots. \tag{3.26}$$

Let $\check{\eta}_1 = \min_{1 \le j \le d} |\eta_j^{-1}| > 1$ and $\check{\eta}_2 = \max_{1 \le j \le d} |\eta_j^{-1}| > 1$. Then $\|\widetilde{\mathbf{F}}^n\| \sim c_9\check{\eta}_2^n$ and

$$\|\widetilde{\mathbf{F}}^{-n}\| \sim c_{10}\check{\eta}_1^{-n} \qquad \text{as } n \to \infty. \tag{3.27}$$

Let $\widetilde{\mathbf{K}}$ be a nonsingular matrix such that $\widetilde{\mathbf{K}}\widetilde{\mathbf{F}}\widetilde{\mathbf{K}}^{-1} = \mathrm{diag}(\eta_1^{-1}, \ldots, \eta_d^{-1})$ and $\widetilde{\mathbf{s}} = \sum_{t=1}^{\infty} \widetilde{\mathbf{F}}^{-(t-1)}\mathbf{v}_{2t}$. Define the $d \times d$ diagonal matrix $\widetilde{\mathbf{S}}$ whose ith diagonal element is the ith element of $\widetilde{\mathbf{K}}\widetilde{\mathbf{s}}$. Let $\mathbf{N}_n = \widetilde{\mathbf{F}}^{-(n+q-1)}$,

$$\Lambda = \begin{pmatrix} (1-\eta_1^2)^{-1} & \ldots & (1-\eta_1\eta_d)^{-1} \\ \ldots & \ldots & \ldots \\ (1-\eta_1\eta_d)^{-1} & \ldots & (1-\eta_d^2)^{-1} \end{pmatrix}$$

and $\widetilde{\mathbf{F}}^* = \sum_{i=1}^{\infty} \widetilde{\mathbf{F}}^{-(i-1)}\widetilde{\mathbf{s}}\widetilde{\mathbf{s}}'\widetilde{\mathbf{F}}^{-(i-1)'} = \widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{S}}\Lambda\widetilde{\mathbf{S}}'\widetilde{\mathbf{K}}^{-1'}$.

Theorem 3.4. *Under* (1.2) *and* (3.27),

$$\text{(i)} \quad \mathbf{N}_n \sum_{t=1}^{n} \mathbf{p}_{t+q}\mathbf{p}'_{t+q}\mathbf{N}'_n \xrightarrow{p} \widetilde{\mathbf{F}}^*. \tag{3.28}$$

In addition if e_t's are Gaussian, $\widetilde{\mathbf{F}}^*$ is positive definite a.s. Also

$$\text{(ii)} \quad (\mathbf{N}'_n)^{-1}\left(\sum_{t=1}^{n} \mathbf{p}_{t+q}\mathbf{p}'_{t+q}\right)^{-1} \sum_{t=1}^{n} \mathbf{p}_{t+q}e_{t+q+1} \xrightarrow{d} \mathrm{N}_2^*, \tag{3.29}$$

where $\mathrm{N}_2^* = \widetilde{\mathbf{K}}'\widetilde{\mathbf{S}}^{-1}\Lambda^{-1}\widetilde{\boldsymbol{\vartheta}}$, $\widetilde{\boldsymbol{\vartheta}}$ being a d-variate Gaussian variable with mean zero and dispersion matrix $\tilde{\mathbf{V}} = ((\tilde{v}_{ij}))$ with $\tilde{v}_{ij} = \sigma^2 \sum_{k=1}^{\infty}\sum_{l=1}^{\infty} \eta_i^{(k-1)}\eta_j^{(l-1)}$. Also $\widetilde{\boldsymbol{\vartheta}}$ is independent of $\widetilde{\mathbf{K}}'\widetilde{\mathbf{S}}^{-1}\Lambda^{-1}$.

Proof. The proof is similar to that of Theorem 3.2. □

4 The main theorem

We first show the consistency of $\hat{\boldsymbol{\theta}}_n$. Although, like for the i.i.d. errors or the martingale difference errors, the consistency can be shown directly, in this case we take advantage of the discussions in Section 3 to do so.

Let $\mathbf{G}_n = \text{diag}((\mathbf{J}_n, \mathbf{K}_n, \mathbf{M}_n, \mathbf{N}_n))$. Then we have the following theorem.

Theorem 4.1. *Under conditions* (1.2), (3.3), (3.11), (3.22) *and* (3.27),

$$(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = o_p(1). \tag{4.1}$$

Proof. Consider the different components of $(\mathbf{T}'\mathbf{G}_n')^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$.

For the stable component of the AR part, it follows from Theorem 3.1(ii) that

$$(\mathbf{J}_n')^{-1}\left(\sum_{t=1}^{n} \mathbf{r}_t \mathbf{r}_{t+q}'\right)^{-1}\left(\sum_{t=1}^{n} \mathbf{r}_t u_{t+q+1}\right) = O_p(1).$$

Next, for the explosive component of the AR part, defining $\mathbf{d}_n$ as in Lemma 3.4,

$$\begin{aligned}
&\left(\mathbf{F}^{-(n+q-1)'}\right)^{-1}\left(\sum_{t=1}^{n} \mathbf{s}_t \mathbf{s}_{t+q}'\right)^{-1} \sum_{t=1}^{n} \mathbf{s}_t u_{t+q+1} \\
&\quad = \left(\mathbf{F}^{-(n-1)} \sum_{t=1}^{n} \mathbf{s}_t \mathbf{s}_{t+q}' \mathbf{F}^{-(n+q-1)'}\right)^{-1} \mathbf{d}_n.
\end{aligned}$$

Now under the stationarity of the sequence u_t, for some $c_{11} > 0$,

$$\begin{aligned}
\mathrm{E}\|\mathbf{d}_n\| &\le n\left\|\mathbf{F}^{-(n-1)}\right\| n^{-1} \sum_{t=1}^{n} \mathrm{E}\|\mathbf{s}_t u_{t+q+1}\| \\
&= c_{11} n \left\|\mathbf{F}^{-(n-1)}\right\| \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty.
\end{aligned}$$

Hence, $\mathbf{d}_n = o_p(1)$. This along with Theorem 3.2(ii) gives

$$\left(\mathbf{F}^{-(n-1)} \sum_{t=1}^{n} \mathbf{s}_t \mathbf{s}_{t+q}' \mathbf{F}^{-(n+q-1)'}\right)^{-1} \mathbf{d}_n = o_p(1).$$

Similar results hold for the stable and explosive components of the MA part. Hence, using Proposition 6.1.2 of Brockwell and Davis (1991), we have

$$\left\|(\mathbf{T}'\mathbf{G}_n')^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})\right\| = O_p(1)$$

so that

$$\left\|(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})\right\| = o_p\big(\|\mathbf{T}\mathbf{G}_n\|\big).$$

Hence, we have the following theorem. □

Theorem 4.2. *Under conditions* (1.2), (3.3), (3.11), (3.22) *and* (3.27), *as* $n \longrightarrow \infty$,

$$\text{(i)} \quad \mathbf{G}_n \mathbf{T} \mathbf{D}_n \mathbf{T}' \mathbf{G}_n' \overset{p}{\sim} \operatorname{diag}\big((\Sigma_1, \mathbf{F}^*, \Sigma_2, \widetilde{\mathbf{F}}^*)\big) \tag{4.2}$$

and

$$\text{(ii)} \quad \big(\mathbf{T}'\mathbf{G}_n'\big)^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \overset{d}{\sim} (\mathrm{N}_r, \mathrm{N}_1^*, \mathrm{N}_c, \mathrm{N}_2^*)', \tag{4.3}$$

where the stable and explosive components are asymptotically independent of each other, but the two stable components and the two explosive components of the AR and MA parts are not.

To prove Theorem 4.2, we require the following lemmas.

Lemma 4.1. *Under conditions* (1.2), (3.3), (3.11), (3.22) *and* (3.27),

$$\text{(i)} \quad \mathbf{J}_n \sum_{t=1}^{n} \mathbf{r}_t \mathbf{s}_{t+q}' \mathbf{K}_n' \xrightarrow{p} 0$$

and

$$\text{(ii)} \quad \mathbf{M}_n \sum_{t=1}^{n} \mathbf{q}_{t+q} \mathbf{p}_{t+q}' \mathbf{N}_n' \xrightarrow{p} 0.$$

Proof. The proof is similar to that of Sen Roy and Bhattacharya (2012). □

We next state (without proof) a lemma by Helland (1982).

Lemma 4.2. *Let* $(\mathrm{X}_{n,k}^{(1)}, \ldots, \mathrm{X}_{n,k}^{(k)})$, $k = 1, 2, \ldots, m = 1, 2, \ldots$ *be a sequence of* m*-dimensional, stationary* ϕ*-mixing array with* $\sum_{n=1}^{\infty} \{\phi^{(k)}(n)\}^{1/2} < \infty$ *for each* k. *For some stopping rule* $s_n(t)$, *let* $\mathrm{X}_n(t) = (\mathrm{X}_n^{(1)}(t), \ldots, \mathrm{X}_n^{(k)}(t))$, *where* $\mathrm{X}_n^{(j)}(t) = \sum_{k=1}^{s_n(t)} \mathrm{X}_{n,k}^{(j)}$. *Also suppose that* $\mathrm{W}_1, \mathrm{W}_2, \ldots, \mathrm{W}_m$ *are* m *independent Gaussian processes and* $\mathrm{f}_1, \mathrm{f}_2, \ldots, \mathrm{f}_m$ *are independent nonnegative measurable functions such that for all* $t > 0$ *and* $k = 1, 2, \ldots, m$, $\int_0^t f_k^2(s)\, ds < \infty$. *Then under the conditions*

$$\sum_{k=1}^{s_n(t)} (\mathrm{X}_{n,k}^{(i)})^2 \xrightarrow{p} \int_0^t f_i^2(s)\, ds \qquad \text{for all } i = 1, 2, \ldots, m \tag{4.4}$$

and

$$\sum_{k=1}^{s_n(t)} \mathrm{X}_{n,k}^{(i)} \mathrm{X}_{n,k}^{(j)} \xrightarrow{p} 0 \qquad \text{for all } i \neq j = 1, \ldots, m, \tag{4.5}$$

$$\mathrm{X}_n \xrightarrow{p} \Big(\int \mathrm{f}_1\, dW_1, \ldots, \int \mathrm{f}_m\, dW_m\Big). \tag{4.6}$$

Lemma 4.3. *$\mathbf{J}_n \sum_{t=1}^{n} \mathbf{r}_t u_{t+q+1}$ and $\mathbf{K}_n \sum_{t=1}^{n} \mathbf{s}_t u_{t+q+1}$ are asymptotically independent.*

Proof. In fact, it can be shown that

$$\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \mathbf{r}_k u_{k+q+1}, \mathbf{F}^{-(n-1)} \sum_{k=1}^{[nt]} \mathbf{s}_k u_{k+q+1}\right) \xrightarrow{d} (\mathrm{N}_s, \mathrm{N}_1^*), \tag{4.7}$$

where N_s is a normal vector with mean zero and variance Σ_1^* and $\mathrm{N}_1^* = \mathbf{K}^{-1}\bar{\mathbf{S}}\hat{\boldsymbol{\vartheta}}$ is the product of two independent normal variates, N_s being independent of N_1^*.

In Theorems 3.1 and 3.2, we have shown that each of the two components in (4.7) converge to the corresponding marginal distributions, that is, condition (4.4) holds. Hence, we need only to show that the cross product term converges in probability to zero.

Under the assumption of bounded fourth-order moment,

$$\begin{aligned}
&\mathrm{E}\left\| \frac{1}{\sqrt{n}} \mathbf{F}^{-(n-1)} \sum_{k=1}^{[nt]} \mathbf{r}_k u_{k+q+1} \mathbf{s}_k' u_{k+q+1} \right\| \\
&\quad = \mathrm{E}\left\| \frac{1}{\sqrt{n}} \mathbf{F}^{-(n-1)} \sum_{k=1}^{[nt]} u_{k+q+1}^2 \left(\sum_{j=1}^{k} \mathrm{L}_1^{k-j} \mathbf{u}_j\right)\left(\sum_{l=1}^{k} \mathbf{F}^{k-l} \mathbf{u}_l\right)' \right\| \\
&\quad \le c_{12} \frac{1}{\sqrt{n}} \|\mathbf{F}^{-(n-1)}\| \sum_{k=1}^{[nt]} (1 - \|\mathbf{L}_1\|^k)(\|\mathbf{F}\|^k - 1) \qquad (4.8) \\
&\quad \le c_{12} \frac{1}{\sqrt{n}} \|\mathbf{F}^{-(n-1)}\| \left(\sum_{k=1}^{n} \|\mathbf{F}^k\| - [nt]\right) \\
&\quad = c_{12} \frac{1}{\sqrt{n}} \|\mathbf{F}^{-(n-2)}\| \left(\frac{(\|\mathbf{F}^n\| - 1)}{\|\mathbf{F}\| - 1} - [nt]\right) \to 0 \qquad \text{as } n \to \infty.
\end{aligned}$$

Hence, condition (4.5) holds. Thus, Lemma 4.2 leads to Lemma 4.3 and (4.7). □

Corollary 4.1. *The following are asymptotically independent:*

(i) $\mathbf{M}_n \sum_{t=1}^{n} \mathbf{q}_{t+q} e_{t+q+1}$ *and* $\mathbf{K}_n \sum_{t=1}^{n} \mathbf{s}_t u_{t+q+1}$,
(ii) $\mathbf{J}_n \sum_{t=1}^{n} \mathbf{r}_t u_{t+q+1}$ *and* $\mathbf{N}_n \sum_{t=1}^{n} \mathbf{p}_{t+q} e_{t+q+1}$,
(iii) $\mathbf{M}_n \sum_{t=1}^{n} \mathbf{q}_{t+q} e_{t+q+1}$ *and* $\mathbf{N}_n \sum_{t=1}^{n} \mathbf{p}_{t+q} e_{t+q+1}$.

Proof. The proofs are similar to that of Lemma 4.3. □

Lemma 4.4. *The following terms are dependent even for large n:*

(i) $\mathbf{J}_n \sum_{t=1}^{n} \mathbf{r}_t u_{t+q+1}$ *and* $\mathbf{M}_n \sum_{t=1}^{n} \mathbf{q}_{t+q} e_{t+q+1}$,
(ii) $\mathbf{K}_n \sum_{t=1}^{n} \mathbf{s}_t u_{t+q+1}$ *and* $\mathbf{N}_n \sum_{t=1}^{n} \mathbf{p}_{t+q} e_{t+q+1}$.

Proof. For large n and any $k_n \ni k_n/n \to 0$, both the terms in (i) depend on e_i, $k_n + 1 \le i \le n - k_n$. Hence, the two terms are dependent.

Similarly, the terms in (ii) are dependent since for large n, $\mathbf{S}_n$ and $\widetilde{\mathbf{S}_n}$ depend on e_i for $1 \le i \le k_n$ and $\boldsymbol{\vartheta}_n$ and $\widetilde{\boldsymbol{\vartheta}_n}$ depend on e_i for $n - k_n + 1 \le i \le n$. Hence, the lemma. □

Proof of Theorem 4.2. Theorem 4.2(i) and (ii) follows from Theorems 3.1, 3.2, 3.3 and 3.4, Lemmas 4.1, 4.3 and 4.4 and Corollary 4.1. □

Remark. $\mathbf{T}$ and $\mathbf{G}_n$ in Theorem 4.2 involves parameters, which in practical situations need to be estimated. This can be done as follows. An estimator $\hat{\mathbf{A}}$ of $\mathbf{A}$ as defined in (2.5) can be obtained using the instrumental variable estimator $\hat{\boldsymbol{\alpha}}_n$ of α. This would lead to the estimators $\hat{\rho}_i^{-1}, i = 1, \ldots, r$ and $\hat{\gamma}_j^{-1}, j = 1, \ldots, s$ of the roots of $\mathbf{A}$, and hence to

$$\hat{\rho}_i^* = (-)^j \sum_{i_1<} \sum_{i_2<}^{r} \cdots \sum_{<i_j=1} \hat{\rho}_{i_1}^{-1} \hat{\rho}_{i_2}^{-1} \cdots \hat{\rho}_{i_k}^{-1}, \qquad i = 1, \ldots, r \tag{4.9}$$

and the corresponding expression of $\hat{\gamma}_j^*$, $j = 1, \ldots, s$. $\hat{\pi}_i^*$'s and $\hat{\eta}_j^*$ can be similarly estimated by using $\hat{\boldsymbol{\beta}}_n$ in $\mathbf{C}$ as defined in (2.8).

Since $\|(\hat{\mathbf{A}} - \mathbf{A})\| = o_p(1)$ and $\|(\hat{\mathbf{B}} - \mathbf{B})\| = o_p(1)$, it follows that $\hat{\rho}_i^*$'s, $\hat{\gamma}_j^*$'s, $\hat{\pi}_k^*$'s and $\hat{\eta}_l^*$'s are consistent estimators of ρ_i^*'s, γ_j^*'s, π_k^*'s and η_l^*'s, respectively. Hence, $\|\hat{\mathbf{L}}_1 - \mathbf{L}_1\| = o_p(1)$, $\|\hat{\mathbf{L}}_2 - \mathbf{L}_2\| = o_p(1)$, $\|\hat{\mathbf{F}} - \mathbf{F})\| = o_p(1)$ and $\|\hat{\widetilde{\mathbf{F}}} - \widetilde{\mathbf{F}}\| = o_p(1)$.

Using these in $\mathbf{T}$ and $\mathbf{G}_n$, we get the estimators $\hat{\mathbf{T}}$ and $\hat{\mathbf{G}}_n$.

Theorem 4.3. *Under conditions* (1.2), (3.3), (3.11), (3.22) *and* (3.27), *as* $n \longrightarrow \infty$,

$$(\hat{\mathbf{T}}'\hat{\mathbf{G}}_n')^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \overset{d}{\sim} (\mathrm{N}_r, \mathrm{N}_1^*, \mathrm{N}_c, \mathrm{N}_2^*)'.$$

Proof. To prove the theorem, we write

$$\begin{aligned}
&(\hat{\mathbf{T}}'\hat{\mathbf{G}}_n')^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \\
&\quad = \{(\hat{\mathbf{T}}'\hat{\mathbf{G}}_n')^{-1} - (\mathbf{T}'\mathbf{G}_n')^{-1}\}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) + (\mathbf{T}'\mathbf{G}_n')^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}).
\end{aligned} \tag{4.10}$$

Now,

$$\begin{aligned}
&\|(\hat{\mathbf{G}}_n\hat{\mathbf{T}}) - (\mathbf{G}_n\mathbf{T})\| \\
&\quad = \|(\hat{\mathbf{G}}_n - \mathbf{G}_n)(\hat{\mathbf{T}} - \mathbf{T}) + \mathbf{G}_n(\hat{\mathbf{T}} - \mathbf{T}) + \mathbf{T}(\hat{\mathbf{G}}_n - \mathbf{G}_n)\| \\
&\quad \le \|(\hat{\mathbf{G}}_n - \mathbf{G}_n)\|\|\hat{\mathbf{T}} - \mathbf{T}\| + \|\mathbf{G}_n\|\|(\hat{\mathbf{T}} - \mathbf{T})\| + \|\mathbf{T}\|\|(\hat{\mathbf{G}}_n - \mathbf{G}_n)\|,
\end{aligned} \tag{4.11}$$

where for some $c_{13} > 0$,

$$\|\hat{\mathbf{G}}_n - \mathbf{G}_n\| \le c_{13} \max\big(\|\hat{\mathbf{F}}^{-n} - \mathbf{F}^{-n}\|, \|\widetilde{\mathbf{F}}^{-n} - \mathbf{F}^{-n}\|\big). \tag{4.12}$$

Since $\|\hat{\mathbf{A}} - \mathbf{A}\| = o_p(1)$, $\|\hat{\mathbf{F}} - \mathbf{F}\| = o_p(1)$, since maximum eigenvalue of $\hat{\mathbf{F}}$ and $\mathbf{F}$ are the same as those of $\hat{\mathbf{A}}$ and $\mathbf{A}$, respectively. Hence,

$$\begin{aligned} \|\hat{\mathbf{F}}^{-1} - \mathbf{F}^{-1}\| &= \|\hat{\mathbf{F}}^{-1}\| \|\mathrm{I}_s - \hat{\mathbf{F}}\mathbf{F}^{-1}\| \\ &= \|\mathbf{F}^{-1}\| \|\mathrm{I}_s - (\mathbf{F}\mathbf{F}^{-1} + o_p(1))\| = o_p(1). \end{aligned} \tag{4.13}$$

Also since for any positive integer $m > 0$,

$$\hat{\mathbf{F}}^{-m} - \mathbf{F}^{-m} = \sum_{j=0}^{m-1} \binom{m}{j} (\hat{\mathbf{F}}^{-1} - \mathbf{F}^{-1})^{m-j} \mathbf{F}^{-j},$$

it follows from (4.13) that

$$\|\hat{\mathbf{F}}^{-n} - \mathbf{F}^{-n}\| = o_p(1).$$

Similarly, $\|\widetilde{\mathbf{F}}^{-n} - \mathbf{F}^{-n}\| = o_p(1)$, so that from (4.12) we have

$$\|\hat{\mathbf{G}}_n - \mathbf{G}_n\| = o_p(1). \tag{4.14}$$

Again the nonzero terms of the matrix $\hat{\mathbf{T}} - \mathbf{T}$ converges in probability to zero, so that $\|\hat{\mathbf{T}} - \mathbf{T}\| = o_p(1)$. Using this along with (4.14) in (4.11) gives

$$\|(\hat{\mathbf{G}}_n \mathbf{T}) - (\mathbf{G}_n \mathbf{T})\| = o_p(1), \tag{4.15}$$

from which arguments similar to (4.13) leads to

$$\{(\hat{\mathbf{T}}' \hat{\mathbf{G}}_n')^{-1} - (\mathbf{T}' \mathbf{G}_n')^{-1}\} = o_p(1). \tag{4.16}$$

Theorem 4.1 along with (4.16) show that the first term in (4.10) converges to zero in probability, and hence the theorem follows. □

5 Concluding remarks

In this paper, we derive the asymptotic distribution of the estimated ARMA parameters taking the instrumental variable estimator for the AR component and the derived AR process estimator for the MA component. The latter is unobservable, and hence cannot be directly used to estimate β. As suggested by Chan and Tsay (1996), the derived process $Y_t(\beta^0)$ can be constructed from an initial value β^0 of β and then iterated to obtain the final solution $\hat{\beta}$.

The proofs show that unlike for i.i.d. or martingale difference errors, for ϕ-mixing errors, conditions like (1.4) are necessary to derive the asymptotic distributions. The results, however, come out in similar form. The implication is that for more stringent dependence structure more conditions are necessary to bring about similar results.

Acknowledgments

The authors are grateful to the referees for their useful comments which led to an improvement of the paper.

References

Anderson, T. W. (1959). On asymptotic distribution of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.* **30**, 676–687. MR0107347

Basu, A. K. and Sen Roy, S. (1993). Asymptotic theory of estimation of parameters in autoregressive models under general set-up of the roots. *Comm. Statist. Theory Methods* **22**, 2223–2237. MR1238164

Basu, A. K., Sen Roy, S. and Bhattacharya, S. (2005). Prediction in a causal invertible ARMA(p, q) process with martingale difference innovations. *J. Stat. Theory Appl.* **4**, 256–271. MR2268097

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd ed. New York: Springer. MR1093459

Chan, N. H. and Tsay, R. S. (1996). Asymptotic inference for non-invertible moving average time series. *J. Time Series Anal.* **17**, 1–17. MR1380893

Chan, N. H. and Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16**, 367–401. MR0924877

Helland, I. S. (1982). Central limit theorems for martingales with discrete or continuous time. *Scand. J. Stat.* **9**, 79–94. MR0668684

Jeganathan, P. (1988). On the strong approximation of the distributions of estimators in linear stochastic models, I and II: Stationary and explosive AR models. *Ann. Statist.* **16**, 1283–1314. MR0959203

Longnecker, M. and Serfling, R. J. (1978). Moment inequalities for S_n under general dependence restrictions with application. *Z. Wahrsch. Verw. Gebiete* **43**, 1–21. MR0493478

Mann, H. B. and Wald, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrika* **11**, 173–220. MR0009291

Sen Roy, S. and Bhattacharya, S. (2012). Asymptotic distribution of the estimated parameters of an ARMA(p, q) process in the presence of explosive roots. *Appl. Math.* **39**(3), 257–272. MR2981996

Tsay, R. S. (1993). Testing for non-invertible models with applications. *J. Bus. Econom. Statist.* **11**, 225–233.

White, J. S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29**, 1188–1197. MR0100952

R.K.M.R. College
Narendrapur 700103
India

Department of Statistics
University of Calcutta
35, Ballygunge Circular Road
Calcutta 700019
India