Discussion Paper No. 526

REGRESSION QUANTILES FOR UNSTABLE AUTOREGRESSIVE MODELS

Shiqing Ling
and
Michael McAleer

February 2001

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Regression Quantiles for Unstable Autoregressive Models *

Shiqing Ling
Department of Mathematics
Hong Kong University of Science and Technology

Michael McAleer
Department of Economics
University of Western Australia

Revised: February 2001

Abstract

This paper investigates regression quantiles (RQ) for unstable autoregressive models. The uniform Bahadur representation of the RQ process is obtained. The joint asymptotic distribution of the RQ process is derived in a unified manner for all types of characteristic roots on or outside the unit circle. Unlike the results already available for the regression and stationary autoregression quantiles, the joint asymptotic distribution involves stochastic integrals in terms of a series of independent and identically distributed multivariate Brownian motions with correlated components. The related $L-$estimator is also discussed. As an auxiliary theorem, a weak convergence of a randomly weighted residual empirical process to the stochastic integral of a Kiefer process is established. The results obtained in this paper provide an asymptotic theory for nonstationary time series processes, which can be used to construct robust unit root tests.

1 Introduction

An autoregressive (AR) time series process $\{y_t\}$ of order p is unstable if

$$y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{1.1}$$

*The authors would like to acknowledge the financial support of the Australian Research Council. This paper was revised while the second author was visiting the Institute of Social and Economic Research at Osaka University.

where $\phi_0 = 0$; $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random disturbances with a distribution $F(x)$, zero mean and a finite variance σ^2; y_t is the observation with starting values $(y_0, y_{-1}, \cdots, y_{-p+1})$ independent of $\{\varepsilon_t\}$; and the characteristic polynomial $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ has the decomposition

$$\phi(z) = \psi(z)(1-z)^a(1+z)^b \prod_{k=1}^{l} [(1 - ze^{i\theta_k})(1 - ze^{-i\theta_k})]^{d_k},$$

where a, b, l, d_k, $k = 1, \cdots, l$, are non-negative integers, $0 < \theta_k < \pi$ and $\psi(z)$ is a polynomial of degree $q = p - [a + b + 2(d_1 + \cdots + d_l)]$ with all roots outside the unit circle. Model (1.1) is a general nonstationary autoregressive (AR) time series, which may include real or complex unit roots with various different multiples. Such a model without drift was investigated by Chan and Wei (1988), Jeganathan (1991), Truong-Van and Larramendy (1996), and van der Meer, Pap and van Zuijlen (1999). Recently, Ling and Li (1998, 2001) considered an unstable ARMA model with GARCH errors and an unstable fractionally integrated ARMA model. Such research on unstable time series models is important because it provides a comprehensive understanding of the nature of nonstationary time series processes.

Nonstationary time series have played an important role in both econometric theory and applications over the last fifteen years, and a substantial literature has developed in this field (see Dickey and Fuller (1979), Dickey, Hasza and Fuller (1984), Phillips and Durlauf (1986) and Phillips (1987)). A detailed set of references is given in Phillips and Xiao (1998). Recently, there has been increasing interest in exploring robust estimation methods for nonstationary time series. For example, Cox and Llatas (1991) considered maximum likelihood (ML)-type estimation for a near unit root process; Lucas (1995) investigated M-estimators and related unit root tests for the unit root process with drift; Herce (1996) considered least absolute deviation (LAD) estimation, and showed through simulation that unit root tests based on mixing LAD and lease squares estimators (LSE) are more robust than those based on LSE alone for non-Gaussian unit root processes; and Hasan and Koenker (1997) proposed robust rank tests based on the regression score rank process.

Note that the LAD estimator is a special quantile estimator and the regression score rank process is also related to the regression quantiles (RQ) process (see Koul and Saleh (1995)). According to the same robustness principle, it would be expected that quantile estimators, as well as the L-estimator based on the RQ, will retain the robustness of non-Gaussian nonstationary time series processes. The RQ first developed by Koenker and Bassett (1978) have been popularly accepted as a powerful approach for the robust analysis of linear models, and have led to a number of interesting extensions [cf. Ruppert and Carroll (1980), Bassett and Koenker (1982), Koenker and Bassett (1982), Koenker and D'Orey (1987), and Portnoy and Koenker (1989)]. Recently, Koul and Saleh (1995) extended RQ to stationary AR models, and obtained the uniform Bahadur representation of the autoregression quantile process, and some related asymptotic distributions.

This paper investigates RQ for unstable AR models. The uniform Bahadur representation of the RQ process is obtained. The joint asymptotic distribution of the RQ process is derived in a unified manner for all types of characteristic roots on or outside the unit circle. Unlike the results already available for the regression and stationary AR quantiles, the joint asymptotic distribution involves stochastic integrals in terms of a series of i.i.d. multivariate Brownian motions with correlated components. The related $L-$estimator is also discussed and it is shown that the asymptotic distribution involves the stochastic integrals in terms of a series of i.i.d. bivariate Brownian motions. This is different from the analysis based on the LSE, for which the result depends only on a series of i.i.d. univariate Brownian motions. In Koul and Saleh (1995), an important technique is to apply the uniform closeness of the randomly weighted residual empirical process (RWREP) in Koul and Ossiander (1994) for the RQ process in the stationary AR model. In this paper, we also establish a weak convergence of a RWREP to the stochastic integral of a Kiefer process, so that the uniform closeness can be applied to the RQ process in model (1.1).

The paper proceeds as follows. Section 2 develops two auxiliary theorems. Section 3 presents the main results. Section 4 uses our results to construct unit root tests for some special nonstationary AR models. Section 5 provides the proofs of the main results. Throughout this paper, the following notation is used: A' denotes the transpose of the matrix or vector A; $O_p(1)$ (or $o_p(1)$) denotes a series of random variables that are bounded (or converge to zero) in probability; $\stackrel{p}{\longrightarrow}$ (or $\stackrel{\mathcal{L}}{\longrightarrow}$) denotes convergence in probability (or in distribution); $||\cdot||$ denotes the Euclidean norm; I_k denotes a $k \times k$ identity matrix; $D = D[0,1]$ denotes the space of functions on $[0,1]$ which is defined and equipped with the Skorokhod topology [Billingsley (1968)]; $D^n = D \times D \cdots \times D$ (n factors); and D_2 denotes the space of functions on $[0,1]^2$ which is defined and equipped with the Skorokhod topology in Straf (1970) and Bickel and Wichura (1971).

2 Auxiliary Theorems

This section introduces two auxiliary theorems. The first theorem is the weak convergence of a RWREP, which will be used to establish Theorem 3.1 in Section 3. The second theorem is an invariance principle, which will be used to establish the limiting distribution in Theorem 3.2.

Let $e_t(x)$ be one of the random variables, $I(\varepsilon_t \le x) - F(x)$, $(-1)^t[I(\varepsilon_t \le x) - F(x)]$, $\sin t\theta[I(\varepsilon_t \le x) - F(x)]$ and $\cos t\theta[I(\varepsilon_t \le x) - F(x)]$, where $x \in R$ and $\theta \in (0, \pi)$. Define

$$\mathcal{U}_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} S_n(\frac{t}{n}) e_t(x + \xi_{nt}) \text{ and } \mathcal{U}_n^*(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} S_n(\frac{t}{n}) e_t(x),$$

where $S_n(t/n)$ and ξ_{nt} are two sequences of $\mathcal{F}_{t-1}-$measurable random variables, and $\mathcal{F}_t = \sigma\{\varepsilon_t, \cdots, \varepsilon_0, y_0, \cdots, y_{-p+1}\}$. The following theorem shows the weak convergence of the RWREPs, $\mathcal{U}_n(x)$ and $\mathcal{U}_n^*(x)$.

Theorem 2.1. *Suppose that: (i)* $S_n(\tau) \stackrel{\mathcal{L}}{\longrightarrow} S(\tau)$ *in* D *and* $S(\tau)$ *is continuous in* $\tau \in [0,1]$*; (ii) the finite-dimensional distributions of* $\{\mathcal{U}_n^*(F^{-1}(\alpha)), \alpha \in [0,1]\}$ *converge to those of* $\{\int_0^1 S(\tau) dK(\tau, \alpha), \alpha \in [0,1]\}$ *in distribution; and (iii)* $\max_{1 \le t \le n} |\xi_{nt}| = o_p(1)$*. Assume that* $F(x)$ *has a uniformly continuous and a.e.*

positive density $f(x)$ on $\{x : 0 < F(x) < 1\}$. Then

$$(a) \quad \sup_{x \in R} |\mathcal{U}_n(x) - \mathcal{U}_n^*(x)| = o_p(1),$$
$$(b) \quad \mathcal{U}_n(F^{-1}(\alpha)) \xrightarrow{\mathcal{L}} \int_0^1 S(\tau) dK(\tau, \alpha) \text{ in } D,$$
$$(c) \quad \mathcal{U}_n^*(F^{-1}(\alpha)) \xrightarrow{\mathcal{L}} \int_0^1 S(\tau) dK(\tau, \alpha) \text{ in } D,$$

where $K(\tau, \alpha)$ is a Kiefer process in D_2, a two-parameter Gaussian process with zero mean and covariance cov$(K(\tau_1, \alpha_1) K(\tau_2, \alpha_2)) = (\tau_1 \wedge \tau_2)(\alpha_1 \wedge \alpha_2 - \alpha_1 \alpha_2)$.

Koul and Ossiander (1994) studied the weak convergence of the RWREP, $\mathcal{U}_n(x) = n^{-1/2} \sum_{t=1}^n \gamma_{nt} [I(\varepsilon_t \leq x + \xi_{nt}) - F(x + \xi_{nt})]$ and $\mathcal{U}_n^*(x) = n^{-1/2} \sum_{t=1}^n \gamma_{nt} [I(\varepsilon_t \leq x) - F(x)]$. Under the assumption that $\sum_{t=1}^n \gamma_{nt}^2 / n$ converges to a positive random variable γ^2 in probability, they obtained the asymptotic distribution of $\mathcal{U}_n(x)$ and $U_n^*(x)$, which is the product of γ and a Brownian bridge on D. Here we provide a different condition set, i.e. condition (i) replaces their condition that $\sum_{i=1}^n \gamma_{ni}^2 / n = \gamma^2 + o_p(1)$ with γ^2 being a positive random variable and $n^{-1/2} \max_{1 \leq i \leq n} \gamma_{ni}^2 = o_p(1)$, and obtain a different weak convergence of a RWREP. In Theorem 2.1, if conclusions (a)-(c) are modified as follows: (a) $\sup_{x \in \{x : F(x) \in [\omega_1, \omega_2]\}} |\mathcal{U}_n(x) - \mathcal{U}_n^*(x)| = o_p(1)$, (b) $\mathcal{U}_n(F^{-1}(\alpha)) \xrightarrow{\mathcal{L}} \int_0^1 S_t(\tau) dK(\tau, \alpha)$ in $D[\omega_1, \omega_2]$, and (c) $\mathcal{U}_n^*(F^{-1}\ (\alpha)) \xrightarrow{\mathcal{L}} \int_0^1 S_t(\tau) dK(\tau, \alpha)$ in $D[\omega_1, \omega_2]$, where $[\omega_1, \omega_2] \subset (0, 1)$, then the uniform continuity of $F(x)$ can be weakened as the assumptions in Theorem 3.1 in the next section. The proof of this theorem is similar to those of Koul and Ossiander (1994), and hence is omitted.

Before giving the second theorem, we need the following notation: $\mathbf{A}_t = [\varepsilon_t, \mathbf{B}_t']'$ and $\mathbf{B}_t = [I(\varepsilon_t \leq F^{-1}(\alpha_1)) - \alpha_1, \cdots, I(\varepsilon_t \leq F^{-1}(\alpha_m)) - \alpha_m]'$, where $0 < \alpha_1 < \cdots < \alpha_m < 1$. Furthermore, define $W_i(\tau, \alpha) = [B_i(\tau), \mathbf{K}_i'(\tau, \alpha)]$ as a sequence of i.i.d. $(m+1)$-dimensional Brownian motions with mean zero and covariance

$$\tau \tilde{\Omega} = \tau \begin{pmatrix} \sigma^2 & \Omega_{11} \\ \Omega_{11}' & \Omega \end{pmatrix}, \tag{2.1}$$

$\Omega = (\sigma_{ij})_{m \times m}$, $\sigma_{ij} = \alpha_i - \alpha_i \alpha_j$, $1 \leq i \leq j \leq m$, $\Omega_{11} = (\sigma_1, \cdots, \sigma_m)$ and $\sigma_i = \int_0^{\alpha_i} F^{-1}(s) ds$, where $\mathbf{K}_i(\tau, \alpha) = [K_i(\tau, \alpha_1), \cdots, K_i(\tau, \alpha_m)]'$, and $i = 1, \cdots, 2l + 2$.

Here, each $\mathbf{K}_i(\tau,\alpha)$ is a Kiefer processes with varying τ in [0, 1] and fixed set $\{\alpha_1,\cdots,\alpha_m\}$.

Theorem 2.2. *Let $\{z_t : t=1,\cdots,n\}$ be generated by the AR(q) model, $z_t = \sum_{i=1}^q \psi_i z_{t-i}+\varepsilon_t$, with all roots of the polynomial $1-\sum_{i=1}^q \psi_i B^i$ outside the unit circle. Denote $S_t = [\mathbf{A}_t', (-1)^t\mathbf{A}_t', \sin t\theta_1\mathbf{A}_t', \cos t\theta_1\mathbf{A}_t', \cdots, \sin t\theta_l\mathbf{A}_t', \cos t\theta_l\mathbf{A}_t', \mathbf{B}_t'\otimes Z_{t-1}']'$, where $\theta_i\in(0,\pi)$, $\theta_i\neq\theta_j$ if $i\neq j$, $i,j=1,\cdots,l$, and $Z_{t-1}=(z_{t-1},\cdots,z_{t-q})'$. Then*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]} S_t \xrightarrow{\mathcal{L}} W(\tau,\alpha) \;\; in \;\; D^{2(m+1)(l+1)+qm}, \tag{2.2}$$

where $W(\tau,\alpha) = [W_1'(\tau,\alpha),\ W_2'(\tau,\alpha),\ \cdots,\ W_{2l+1}'(\tau,\alpha),\ W_{2l+2}'(\tau,\alpha),\ N'(\tau)]'$, and $N(\tau)$ is an $mq-$dimensional Brownian motion independent of $W_i(\tau,\alpha)$, and has mean zero and covariance $\tau\Omega\otimes\Sigma$, with $\Sigma=E(Z_{t-1}Z_{t-1}')$.

Proof. Let $\lambda=(\lambda_1,\cdots,\lambda_{2(m+1)(l+1)},\lambda_{mq}')'$ be a $[2(m+1)(l+1)+mq]-$ dimensional constant vector with $\lambda'\lambda\neq 0$, where λ_{mq} is an $mq-$dimensional constant. Denote $S_t^*=\lambda' S_t=\sum_{t=1}^{[n\tau]} a_t$. It is straightforward to show that $\tilde{\Omega}$ is positive definite and

$$\frac{1}{n}\sum_{t=1}^{[n\tau]} E(a_t^2|\mathcal{F}_{t-1}) \longrightarrow \tau\lambda'\Omega^*\lambda > 0 \;\; a.s., \tag{2.3}$$

where $\Omega^*=\mathrm{diag}(I_{2(l+1)}\otimes\tilde{\Omega},\Omega\otimes\Sigma)$. Denote $\tilde{a}_t=c_0+c_1|\varepsilon_t|+c_2\sum_{i=1}^q|z_{t-i}|$, where c_0, c_1 and c_2 are constant such that $a_t^2\leq\tilde{a}_t^2$. Since $E\tilde{a}_t^2<\infty$, for any small η,

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{[n\tau]} E[a_t^2 I(|a_t|>\sqrt{n}\eta)] &\leq \frac{1}{n}\sum_{t=1}^{n} E[\tilde{a}_t^2 I(|\tilde{a}_t|>\sqrt{n}\eta)]\\ &= E[\tilde{a}_t^2 I(|\tilde{a}_t|>\sqrt{n}\eta)] = \int_{x>\sqrt{n}\eta}^{\infty} x^2 dP \to 0,\end{aligned} \tag{2.4}$$

where P is the distribution of $\tilde{a}_t$. Applying the invariance principle in Helland (1982, Theorem 3.2) and the Gramer advice completes the proof. □

3 Main Results

Denote $X_{t-1}=(1,y_{t-1},\cdots,y_{t-p})'$ and $h_\alpha(s)=\alpha sI(s>0)-(1-\alpha)sI(s\leq 0)$, where $s\in R$, $\alpha\in(0,1)$ and $I(\cdot)$ is the indicator function. Following Koenker and Bassett

(1978) and Koul and Saleh (1995), define the $\alpha-$th regression quantile (RQ) as any member $\hat{\phi}_n(\alpha)$ of the set

$$\hat{\mathcal{R}}_n(\alpha) = \Big\{\lambda \in R^{p+1} : \sum_{t=1}^{n} h_\alpha(y_t - X'_{t-1}\lambda) = minimum\Big\},$$

and refer to $\{\hat{\phi}_n(\alpha) : 0 < \alpha < 1\}$ as an RQ process. In practice, $\hat{\phi}_n(\alpha)$ can be obtained using a linear programming version of the minimization problem above, as given in Koenker and D'Orey (1987, 1993). $\hat{\phi}_n(1/2)$ is the important LAD estimator of ϕ, where $\phi = (\phi_0, \phi_1, \cdots, \phi_p)'$. Denote $\phi(\alpha) = \phi + (F^{-1}(\alpha), 0, \cdots, 0)'$ and $q(\alpha) = f(F^{-1}(\alpha))$. Define

$$\mathbf{T}_n(s, \alpha) = \sum_{t=1}^{n} X_{t-1}[I(\varepsilon_t \le F^{-1}(\alpha) + s'\delta_n' X_{t-1}) - \alpha], \tag{3.1}$$

where $\alpha \in [0, 1]$, $\delta_n = G' J_n^{-1}$, $J_n = \mathrm{diag}(\sqrt{n}, N_1, \cdots, N_{l+2}, \sqrt{n} I_q)$, $N_1 = \mathrm{diag}(n^a, n^{a-1}, \cdots, n)$, $N_2 = \mathrm{diag}(n^b, n^{b-1}, \cdots, n)$, $N_{k+2} = \mathrm{diag}(nI_2, \cdots, n^{d_k} I_2)$, $k = 1, \cdots, l$,1 and G is a constant matrix defined as in (5.2) in Section 5.

The following theorem gives the Bahadur representation of the RQ $\hat{\phi}_n(\alpha)$.

Theorem 3.1. *Under model (1,1), if it is assumed that $F(x)$ has a continuous and positive density function $f(x)$ on $\{x : 0 < F(x) < 1\}$, then*

$$\hat{\phi}_n(\alpha) - \phi(\alpha) = -[q(\alpha) \sum_{t=1}^{n} X_{t-1} X'_{t-1}]^{-1} \mathbf{T}_n(0, \alpha) + o_p(\delta_n),$$

where $o_p(\cdot)$ holds uniformly for $\alpha \in \omega(\epsilon) = [\epsilon, 1-\epsilon]$ with any $\epsilon \in (0, 1/2]$.

The following notation is needed to state the limiting distribution of $\hat{\phi}_n(\alpha) - \phi(\alpha)$:

$$\xi(\alpha) = \Big(\int_0^1 \Gamma_{a-1}(s) d\mathbf{K}_1'(s, \alpha), \cdots, \int_0^1 \Gamma_0(s) d\mathbf{K}_1'(s, \alpha)\Big)',$$
$$\Gamma_0(\tau) = B_1(\tau),\ \ \Gamma_j(\tau) = \int_0^\tau \Gamma_{j-1}(s) ds,\ \ \Gamma = (\vartheta_{ij})_{a \times a},\ \ \vartheta_{ij} = \int_0^1 \Gamma_i(s)\Gamma_j(s) ds;$$
$$\eta(\alpha) = -\Big(\int_0^1 \tilde{\Gamma}_{b-1}(\tau) d\mathbf{K}_2'(\tau, \alpha), \cdots, \int_0^1 \tilde{F}_0(\tau) d\mathbf{K}_2'(\tau, \alpha)\Big)',$$
$$\tilde{\Gamma}_0(\tau) = B_2(\tau),\ \ \tilde{\Gamma}_j(\tau) = \int_0^\tau \tilde{\Gamma}_{j-1}(s) ds,\ \ \tilde{\Gamma} = (\tilde{\vartheta}_{ij})_{a \times a},\ \ \tilde{\vartheta}_{ij} = \int_0^1 \tilde{\Gamma}_i(s)\tilde{\Gamma}_j(s) ds;$$
$$\zeta_k = (\xi_1, \cdots, \xi_{2d_k})',\ \ H_k = (\sigma_{ij})_{2d_k \times 2d_k},\ f_0(\tau) = B_{2k+1}(\tau),\ \ g_0(\tau) = B_{2k+2}(\tau),$$
$$f_j(\tau) = \frac{1}{2\sin\theta}\{\sin\theta \int_0^\tau f_{j-1}(s) ds - \cos\theta \int_0^\tau g_{j-1}(s) ds\},$$

$$
\begin{aligned}
g_j\{\tau) &= \frac{1}{2\sin\theta}\{\cos\theta\int_0^\tau f_{j-1}(s)ds + \sin\theta\int_0^\tau g_{j-1}(s)ds\},\\
\xi_{2j-1} &= \frac{1}{2\sin\theta}\{\int_0^1 f_{j-1}(s)d\mathbf{K}'_{2k+2}(s,\alpha) - \int_0^1 g_{j-1}(s)d\mathbf{K}'_{2k+1}(s,\alpha)\},\\
\xi_{2j} &= \frac{1}{2\sin\theta}\{\cos\theta[\int_0^1 f_{j-1}(s)d\mathbf{K}'_{2k+2}(s,\alpha) - \int_0^1 g_{j-1}(s)d\mathbf{K}'_{2k+1}(s,\alpha)]\\
&\qquad - \sin\theta[\int_0^1 f_{j-1}(s)d\mathbf{K}'_{2k+1}(s,\alpha) + \int_0^1 g_{j-1}(s)d\mathbf{K}'_{2k+2}(s,\alpha)]\},\\
\sigma_{2i-1,2j-1} &= \sigma_{2i,2j} = \frac{1}{4\sin^2\theta}\{\int_0^1 f_{i-1}(s)f_{j-1}(s)ds + \int_0^1 g_{i-1}(s)g_{j-1}(s)ds\},\\
\sigma_{2i-1,2j} &= \sigma_{2j,2i-1} = \frac{1}{4\sin^2\theta}\{\cos\theta[\int_0^1 f_{i-1}(s)f_{j-1}(s)ds + \int_0^1 g_{i-1}(s)g_{j-1}(s)ds]\\
&\quad - \sin\theta[\int_0^1 f_{j-1}(s)g_{i-1}(s)ds - \int_0^1 g_{j-1}(s)f_{i-1}(s)ds]\},
\end{aligned}
$$

where $i,j = 1,\cdots,d_k$, $k = 1,\cdots,l$, and $[B_i, \mathbf{K}_i(s,\alpha)]$ is defined in Theorem 2.2.

Theorem 3.2. *Under the assumption of Theorem 3.1,*

$$
\begin{aligned}
\delta_n^{-1}[\hat\phi_n(\alpha_1)-\phi(\alpha_1),\cdots,\hat\phi_n(\alpha_m)-\phi(\alpha_m)] &\xrightarrow{\mathcal{L}} [\ (\mathbf{K}_1(1,\alpha),\xi'(\alpha))\begin{pmatrix} 1 & \xi^* \\ \xi^{*'} & \Gamma \end{pmatrix}^{-1},\\
(\tilde\Gamma^{-1}\eta(\alpha))', (H_1^{-1}\zeta_1(\alpha))',\ \cdots,\ (H_l^{-1}\zeta_l(\alpha))', N_\alpha'\ ]'\, diag[\frac{1}{q(\alpha_1)},\cdots,\frac{1}{q(\alpha_m)}],
\end{aligned}
$$

for any $0<\alpha_1<\alpha_2<\cdots<\alpha_m<1$, *where* $\xi^* = (\int_0^1\Gamma_{a-1}(s)ds,\cdots,\int_0^1\Gamma_0(s)ds)'$ *and* N_α *is a* $q\times m-$*variate normal matrix independent of* $[B_i(\tau),\ \mathbf{K}_i'(\alpha,\tau)]$, *and has a null mean matrix and covariance matrix* $\Omega\otimes\Sigma^{-1}$, *with* Ω *and* Σ *defined in Theorem 2.2.*

Let ν be a finite signed measure with compact support on (0,1). The $L-$estimator of ϕ is defined by

$$
\hat\phi_n^\nu = \int_0^1 \hat\phi_n(\alpha)d\nu(\alpha).
$$

Denote $\phi(\nu,F) = \phi\int_0^1 d\nu(\alpha) + (\int_0^1 F^{-1}(\alpha)d\nu(\alpha),0,\cdots,0)'$. The following theorem follows directly from Theorems 3.1-3.2.

Theorem 3.3. *Under the assumption of Theorem 3.1,*

(a) $\hat\phi_n^\nu - \phi(\nu,F) = -[\sum_{t=1}^n X_{t-1}X_{t-1}']^{-1}\int_0^1[\mathbf{T}_n(0,\alpha)/q(\alpha)]d\nu(\alpha) + o_p(\delta_n)$;

(b) $\delta_n^{-1}[\hat\phi_n^\nu - \phi(\nu,F)] \xrightarrow{\mathcal{L}}$

$$
[\ (K_1^\nu(1),\xi'(\nu))\begin{pmatrix} 1 & \xi^* \\ \xi^{*'} & \Gamma \end{pmatrix}^{-1},\ (\tilde\Gamma^{-1}\eta(\nu))', (H_1^{-1}\zeta_1(\nu))',\ \cdots,\ (H_l^{-1}\zeta_l(\nu))', N_\nu'\ ]',
$$

where $\xi(\nu)$, $\eta(\nu)$ and $\zeta_k(\nu)$ are defined as $\xi(\alpha)$, $\eta(\alpha)$ and $\zeta_k(\alpha)$ in Theorem 3.2, with $[B_i(\tau,\alpha), \mathbf{K}_i'(\tau,\alpha)]$ replaced by $[B_i^\nu(\tau), K_i^\nu(\tau)]$ which are a series of i.i.d. bivariate Brownian motions with mean zero and covariances given by $\tau\Omega_\nu = \tau\begin{pmatrix} \sigma^2 & \sigma_{\varepsilon\nu} \\ \sigma_{\varepsilon\nu} & \sigma_\nu^2 \end{pmatrix}$,

$$\sigma_{\varepsilon\nu} = -\int_0^1 [\int_0^\alpha F^{-1}(s)ds/q(\alpha)]d\nu(\alpha), \tag{3.2}$$

$$\sigma_\nu^2 = \int_0^1 \int_0^1 [(s \wedge \alpha - s\alpha)/q(\alpha)q(s)]d\nu(\alpha)d\nu(s), \tag{3.3}$$

and N_ν is a q-dimensional normal random vector with mean zero and covariance $\sigma_\nu^2\Sigma^{-1}$.

The assumptions and the joint asymptotic distribution of the RQ process corresponding to the stationary componentwise argument in Theorem 3.2 are the same as those given in Koul and Saleh (1995). The joint asymptotic distributions corresponding to the nonstationary componentwise arguments are new results and involve a series of i.i.d. $m-$dimensional Brownian motions. The asymptotic distribution of the $L-$estimator involves a series of i.i.d. bivariate Brownian motions. It is different from the asymptotic distributions of the LSE given by Chan and Wei (1988), Jeganathan (1991), Truong-Van and Larramendy (1996), and van der Meer, Pap and van Zuijlen (1999), which depend only on a series of i.i.d. univariate Brownian motions. The result here is similar to that given by Ling and Li (1998) for ML estimators, which also involves a series of i.i.d. bivariate Brownian motions, but with a different covariance.

4 Two Special Cases

In this section, we apply the results in Section 3 to two special nonstationary AR models and construct corresponding unit root tests.

4.1 AR(1) model

Consider the AR(1) model,

$$y_t = \phi_0 + \phi y_{t-1} + \varepsilon_t, \tag{4.1}$$

where $\phi_0 = 0$ and $\phi = 1$. This model is a special case of model (1.1) with $a = 1$, $b = l = 0$ and $\psi(z) = 1$. Suppose that $\hat{\phi}_n(\alpha)$ and $\hat{\phi}_n^\nu$ are the $\alpha-$th RQ and the $L-$estimator of ϕ, respectively, and $\int_0^1 d\nu(\alpha) = 1$. Then we can obtain directly from Theorem 3.1-3.3,

$$n[\hat{\phi}_n(\alpha) - 1] \xrightarrow{\mathcal{L}} \rho(\alpha) \equiv \frac{\int_0^1 B(\tau)dK(\tau,\alpha) - K(1,\alpha)\int_0^1 B(\tau)d\tau}{\int_0^1 B^2(\tau)d\tau - (\int_0^1 B(\tau)d\tau)^2} \tag{4.2}$$

and

$$n(\hat{\phi}_n^\nu - 1) \xrightarrow{\mathcal{L}} \rho(\nu) \equiv \frac{\int_0^1 B(\tau)dK(\tau,\nu) - K(1,\nu)\int_0^1 B(\tau)d\tau}{\int_0^1 B^2(\tau)d\tau - (\int_0^1 B(\tau)d\tau)^2}, \tag{4.3}$$

where $(B(\tau), K(\tau,\alpha))$ and $(B(\tau), K(\tau,\nu))$ are two bivariate Brownian motions with covariances $\tau\tilde{\Omega}$ and $\tau\Omega_\nu$, $\tilde{\Omega} = \begin{pmatrix} \sigma^2 & \sigma_1 \\ \sigma_1 & \sigma_{11}^2 \end{pmatrix}$ with $\sigma_1 = \int_0^\alpha F^{-1}(s)ds/\ q(\alpha)$ and $\sigma_{11}^2 = (\alpha - \alpha^2)/q^2(\alpha)$, and Ω_ν is defined as in Theorem 3.3.

Let

$$w_1(\tau) = \frac{1}{\sigma}B(\tau) \quad \text{and} \quad w_2(\tau) = -\frac{\sigma_1}{\sigma^2}\sqrt{\frac{\sigma^2}{\sigma^2\sigma_{11}^2 - \sigma_1^2}}B(\tau) + \sqrt{\frac{\sigma^2}{\sigma^2\sigma_{11}^2 - \sigma_1^2}}K(\tau,\alpha).$$

Then $w_1(\tau)$ and $w_2(\tau)$ are two independent standard Brownian motions. As shown in Herce (1996), we have

$$\begin{aligned} n[\hat{\phi}_n(\alpha) - 1] \quad \xrightarrow{\mathcal{L}} \quad & \frac{\sigma_1[\int_0^1 w_1(\tau)dw_1(\tau) - w_1(1)\int_0^1 w_1(\tau)d\tau]}{\sigma^2[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]} \\ & + \frac{\sqrt{\sigma^2\sigma_{11}^2 - \sigma_1^2}}{\sigma^2}\frac{\int_0^1 w_1(\tau)dw_2(\tau) - w_2(1)\int_0^1 w_1(\tau)d\tau}{[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]}. \end{aligned} \tag{4.4}$$

The second term in (4.4) can be simplified to $[\sqrt{\sigma^2\sigma_{11}^2 - \sigma_1^2}/\sigma^2]\ [\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi$, where Φ is a standard normal random variable independent of $\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2$ (see Phillips, 1989). Thus, $n[\hat{\phi}_n(\alpha) - 1]$ can be approximated by

$$\begin{aligned} n[\hat{\phi}_n(\alpha) - 1] \quad \xrightarrow{\mathcal{L}} \quad & \frac{\sigma_1[\int_0^1 w_1(\tau)dw_1(\tau) - w_1(1)\int_0^1 w_1(\tau)d\tau]}{\sigma^2[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]} \\ & + \frac{\sqrt{\sigma^2\sigma_{11}^2 - \sigma_1^2}}{\sigma^2}[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi. \end{aligned} \tag{4.5}$$

If it is further assumed that ε_t has median zero, then $\sigma_{11}^2 = 1/[4f^2(0)]$ and $\sigma_1 = E(|\varepsilon_t|)/[2f(0)]$. In this case, $n[\hat{\phi}_n(1/2) - 1]$, as well as its asymptotic distribution above, are the same as those given by Herce (1996).

Similarly, let

$$w_1(\tau) = \frac{1}{\sigma}B(\tau) \quad \text{and} \quad \tilde{w}_2(\tau) = -\frac{\sigma_{\varepsilon\nu}}{\sigma^2}\sqrt{\frac{\sigma^2}{\sigma^2\sigma_\nu^2 - \sigma_{\varepsilon\nu}^2}}B(\tau) + \sqrt{\frac{\sigma^2}{\sigma^2\sigma_\nu^2 - \sigma_{\varepsilon\nu}^2}}K(\tau,\nu).$$

Then $w_1(\tau)$ and $\tilde{w}_2(\tau)$ are two independent standard Brownian motions. It can be shown that

$$\begin{aligned} n(\hat{\phi}_n^\nu - 1) \quad \xrightarrow{\mathcal{L}} \quad & \frac{\sigma_{\varepsilon\nu}[\int_0^1 w_1(\tau)dw_1(\tau) - w_1(1)\int_0^1 w_1(\tau)d\tau]}{\sigma^2[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]} \\ & + \frac{\sqrt{\sigma^2\sigma_\nu^2 - \sigma_{\varepsilon\nu}^2}}{\sigma^2}[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi. \end{aligned} \tag{4.6}$$

As the limiting distributions in (4.5) and (4.6) include nuisance parameters. there are two methods of constructing unit root tests.

The first method is to combine the LSE so that the nuisance can be cancelled. Denote $\hat{\phi}_{LS}$ as the LSE of ϕ. It is well known that $n(\hat{\phi}_{LS} - 1) \xrightarrow{\mathcal{L}} [\int_0^1 w_1(\tau)dw_1(\tau)$ $-w_1(1)\int_0^1 w_1(\tau)d\tau]/\ [\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]$. Define

$$\begin{aligned} M_\alpha &= \frac{\sigma^2}{\sqrt{\sigma^2\sigma_{11}^2 - \sigma_1^2}}\{n[\hat{\phi}_n(\alpha) - 1] - (\sigma_1/\sigma^2)[n(\hat{\phi}_{LS} - 1)]\}, \\ M_\nu &= \frac{\sigma^2}{\sqrt{\sigma^2\sigma_\nu^2 - \sigma_{\varepsilon\nu}^2}}\{n(\hat{\phi}_n^\nu - 1) - (\sigma_{\varepsilon\nu}/\sigma^2)[n(\hat{\phi}_{LS} - 1)]\}, \\ M_{\alpha,t} &= (\frac{1}{n^2}\sum_{i=1}^n (y_{i-1} - \bar{y})^2)^{1/2}M_\alpha \text{ and } M_{\nu,t} = (\frac{1}{n^2}\sum_{i=1}^n (y_{i-1} - \bar{y})^2)^{1/2}M_\nu, \end{aligned}$$

where $\bar{y} = \sum_{i=1}^n y_{i-1}/n$. It is straightforward to show that

$$\begin{aligned} & M_\alpha \xrightarrow{\mathcal{L}} [\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi, \\ & M_\nu \xrightarrow{\mathcal{L}} [\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi, \\ & M_{\alpha,t} \xrightarrow{\mathcal{L}} \Phi \text{ and } M_{\nu,t} \xrightarrow{\mathcal{L}} \Phi. \end{aligned}$$

Herce (1996) derived the limiting distributions of $M_{1/2}$ and $M_{1/2,t}$. The results above provide a more general asymptotic theory. The statistics M_α, $M_{\alpha,t}$, M_ν and $M_{\nu,t}$ can be used to test for a unit root in model (4.1). From the simulation results given in Lucas (1995) and Herce (1996), these tests should be more robust, especially for a non-Gaussian unit root process. Note that these asymptotic distributions are invariant to α and ν, so that the critical values given by Herce (1996) can still be used.

As the LSE is used in the above method, it may not be quite robust. Another method of accommodating the nuisance parameters is given in Hansen (1995). Let $\tilde{M}_\alpha = n[\hat{\phi}_n(\alpha) - 1]\sigma/\sigma_{11}$ and $\tilde{M}_\nu = n[\hat{\phi}_n^\nu - 1]\sigma/\sigma_\nu$. Then

$$
\begin{aligned}
\tilde{M}_\alpha \quad & \xrightarrow{\mathcal{L}} \quad \frac{r_\alpha[\int_0^1 w_1(\tau)dw_1(\tau) - w_1(1)\int_0^1 w_1(\tau)d\tau]}{[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]} \\
& \quad + \sqrt{1-r_\alpha^2}[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi, \\
\tilde{M}_\nu \quad & \xrightarrow{\mathcal{L}} \quad \frac{r_\nu[\int_0^1 w_1(\tau)dw_1(\tau) - w_1(1)\int_0^1 w_1(\tau)d\tau]}{[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]} \\
& \quad + \sqrt{1-r_\nu^2}[\int_0^1 w_1^2(\tau)d\tau - (\int_0^1 w_1(\tau)d\tau)^2]^{-1/2}\Phi.
\end{aligned}
$$

where $r_\alpha = \sigma_\alpha/\sigma\sigma_{11}$ and $r_\nu = \sigma_{\epsilon\nu}/\sigma\sigma_\nu$. It is easy to see that r_α and $r_\nu \in (0,1)$. Similarly, let $\tilde{M}_t = n[\hat{\phi}_n(\alpha) - 1]\sum_{i=1}^n (y_{i-1} - \bar{y})^2)^{1/2}/\sigma_{11}$ and $\tilde{M}_{\nu,t} = n[\hat{\phi}_n^\nu - 1]\sum_{i=1}^n (y_{i-1} - \bar{y})^2)^{1/2}/\sigma_\nu$, so that we can write down their limiting distributions. These distributions include a nuisance parameter so that the critical values can be determined by the simulation method for different r_α and r_ν (see Hansen (1995)).

4.2 AR(p) model with one unit root

Consider the model

$$
\phi(B)y_t = \phi_0 + \varepsilon_t, \tag{4.7}
$$

where $\phi_0 = 0$ and $\phi(B) = (1-B)\phi^*(B)$, with all the roots of $\phi^*(B)$ outside the unit circle. Reparameterize (4.7) as

$$
y_t = \phi_0 + \gamma_1 y_{t-1} + \sum_{i=2}^p \gamma_i (y_{t-i+1} - y_{t-i}) + \varepsilon_t,
$$

where $\gamma_1 = \sum_{i=1}^p \phi_i$ and $\gamma_j = -\sum_{i=j}^p \phi_i$, $j = 2, \cdots, p$. Suppose that $\hat{\phi}_n(\alpha)$ and $\hat{\phi}_n^\nu$ are the $\alpha-$th RQ and the $L-$estimator of the parameter $\phi = (\phi_1, \cdots, \phi_p)'$, respectively, and $\int_0^1 d\nu(\alpha) = 1$. Denote $\gamma = (\gamma_1, \cdots, \gamma_p)'$ and $\hat{\gamma}_n(\alpha) = (\hat{\gamma}_1, \cdots, \hat{\gamma}_p)'$, with $\hat{\gamma}_1 = \sum_{i=1}^p \hat{\phi}_i$ and $\hat{\gamma}_j = -\sum_{i=j}^p \hat{\phi}_i$, $j = 2, \cdots, p$, where $\hat{\phi}_i$ is the $i-$th element of $\hat{\phi}_n(\alpha)$, and similarly define $\hat{\gamma}_n^{\nu}$. Then, by Theorems 3.1-3.2, as in Ling and Li (1998), we can show that

$$G_n^{-1}[\hat{\gamma}_n(\alpha)-\gamma] \xrightarrow{\mathcal{L}} [c\rho(\alpha), N_\alpha^{'}]' \text{ and } G_n^{-1}(\hat{\gamma}_n^{\nu}-\gamma) \xrightarrow{\mathcal{L}} [c\rho(\nu), N_\nu^{'}]',$$

where $G_n = \text{diag}(1/n, I_{(p-1)\times(p-1)}/\sqrt{n})$, $c = 1/(1-\sum_{i=2}^{p}\gamma_i)$, $\rho(\alpha)$ and $\rho(\nu)$ are defined as in (4.2)-(4.3), and N_α and N_ν are normal random vectors with zero means and covariances $\sigma_\alpha^2 E(Z_{t-1}Z_{t-1}')$ and $\sigma_\nu^2 E(Z_{t-1}Z_{t-1}')$, respectively, $Z_{t-1} = (z_{t-1},\cdots,z_{t-p+1})'$ and $z_t = y_t - y_{t-1}$. As in Section 4.1, the asymptotic distributions, $\rho(\alpha)$ and $\rho(\nu)$, can be used to construct robust unit root tests of $\gamma_1 = 1$.

5 Proofs

Let $u_t = \phi(B)(1-B)^{-a}y_t$, $v_t = \phi(B)(1+B)^{-b}y_t$, $z_t = \phi(B)\psi^{-1}(B)y_t$, and $x_t = \phi(B)(1-2B\cos\theta_k+B^2)^{-d_k}y_t$, where B is the backward shift operator and $k = 1,\cdots,l$. Then $(1-B)^a u_t = \varepsilon_t$, $(1+B)^b v_t = \varepsilon_t$, and $(1-2B\cos\theta_k+B^2)^{d_k}x_t = \varepsilon_t$. Denote $\mathbf{u}_t = (u_t,\cdots,u_{t-a+1})'$, $\mathbf{v}_t = (v_t,\cdots,v_{t-b+1})'$, $Z_t = (z_t,\cdots,z_{t-q+1})'$, and $\mathbf{x}_t(k) = (x_t,\cdots,x_{t-d_k+1})', k = 1,\cdots,l$. As shown in (3.2) of Chan and Wei (1988), abbreviated hereafter as CW, there exists a non-singular matrix Q such that

$$Q\tilde{X}_t = (\mathbf{u}_t', \mathbf{v}_t', \mathbf{x}_t'(1),\cdots,\mathbf{x}_t'(l), Z_t')', \tag{5.1}$$

where $\tilde{X}_t = (y_t,\cdots,y_{t-p+1})'$. Furthermore, let $U_t(j) = (1-B)^{a-j}u_t$ for $j = 0,1,\cdots,a$, $U_t = (U_t(a),\cdots,U_t(1))'$, $V_t(j) = (1+B)^{b-j}v_t$ for $j = 0,1,\cdots,b$, $V_t = (V_t(b),\cdots, V_t(1))'$, $Y_t(k,j) = (1-2B\cos\theta_k+B^2)^{d_k-j}x_t$ for $j = 0,1,\cdots,d_k$, and $Y_t(k) = (Y_t(k,1), Y_{t-1}(k,1), \cdots, Y_t(k,d_k), Y_{t-1}(k,d_k))'$, where $k = 1,\cdots,l$. Then there exist non-singular matrices M, $\tilde{M}$ and C_k, such that

$$M\mathbf{u}_t = U_t,\ \tilde{M}\mathbf{v}_t = V_t,\ C_k\mathbf{x}_t(k) = Y_t(k),\ k = 1,\cdots,l.$$

Denoting $G = \text{diag}(1, M, \tilde{M}, C_1,\cdots,C_l, I_q)\text{diag}(1,Q)$, then

$$GX_t = (1, U_t', V_t', Y_t'(1),\cdots,Y_t'(l), Z_t')'. \tag{5.2}$$

Thus, X_t has been decomposed into various nonstationary componentwise argument vectors corresponding to the locations of unit roots and the stationary componentwise argument vector. Before giving the proofs of our results, we will need the following seven lemmas.

Lemma 5.1. *Suppose that $\{y_t\}$ is generated by model (1.1). Then*

$$\begin{array}{ll} (a) & \sup_{1\le t\le n} ||\delta_n' X_{t-1}|| = o_p(1), \\ (b) & \frac{1}{\sqrt{n}}\sum_{t=1}^{n} ||\delta_n' X_{t-1}|| = O_p(1), \\ (c) & \sum_{t=1}^{n} ||\delta_n' X_{t-1}||^2 = O_p(1). \end{array}$$

Proof. A direct application of Lemma 2.1 in Ling (1998) completes the proof.□

Let γ_{nt} be an $\mathcal{F}_{t-1}$-measurable random variable and assume that the following condition is satisfied:

$$\sum_{t=1}^{n}[|\gamma_{nt}| \cdot ||\delta_n' X_{t-1}||] = O_p(1). \tag{5.3}$$

Denote $R_\epsilon = R \bigcap \{x : \epsilon \le F(x) \le 1-\epsilon\}$, $\gamma_{nt}^+ = \max\{0, \gamma_{nt}\}$, $\gamma_{nt}^- = \gamma_{nt}^+ - \gamma_{nt}$,

$$g_t(s,\lambda) = s'\delta_n' X_{t-1} + \lambda ||\delta_n' X_{t-1}|| \tag{5.4}$$

and

$$\begin{aligned} \tilde{\mathcal{Z}}_n^{\pm}(x,s,\lambda) &= \sum_{t=1}^{n} \gamma_{nt}^{\pm}[I(\varepsilon_t \le x + g_t(s,\lambda)) \\ &- F(x+g_t(s,\lambda)) - I(\varepsilon_t \le x) + F(x)], \end{aligned} \tag{5.5}$$

where $\epsilon \in (0, 1/2]$, $s \in R^{p+1}$ and $\lambda \in R$.

Lemma 5.2. *Let $\mathcal{Z}_n^{\pm}(x,s) = \tilde{\mathcal{Z}}_n^{\pm}(x,s,0)$ and $\mathcal{Z}_n(x,s) = \mathcal{Z}_n^{+}(x,s) - \mathcal{Z}_n^{-}(x,s)$. Under the assumption of Theorem 3.1 and (5.3), if $\sup_{x\in R_\epsilon} |\tilde{\mathcal{Z}}_n^{\pm}(x,s,\lambda)| = o_p(1)$ for any $s \in R^{p+1}$ and $\lambda \in R$, then*

$$\sup_{s\in D_\Delta} \sup_{x\in R_\epsilon} |\mathcal{Z}_n(x,s)| = o_p(1),$$

where $D_\Delta = [-\Delta, \Delta]^{p+1} \subset R^{p+1}$.

The proof of Lemma 5.2 is similar to that of Lemma 3.2 in Koul (1996) (also see Koul (1991)). The main difference is to use Lemma 5.1 to replace Koul's Lemma 3.1, and hence the details are omitted. In the following, we will state three lemmas. Lemmas 5.1-5.2 and these three lemmas are used to prove Lemma 5.6. In addition, these three lemmas will be used to derive the limiting distribution in Theorem 3.2.

Denote $U_t^+(j) = \max\{0, U_t(j)\}$, $U_t^-(j) = U_t^+(j) - U_t(j)$, $\Gamma_j^+(\tau) = \max\{0, \Gamma_j(\tau)\}$ and $\Gamma_j^-(\tau) = \Gamma_j(\tau) - \Gamma_j^+(\tau)$. For the process $\{U_t\}$ defined in (5.2), we have the following lemma.

Lemma 5.3. *Under the assumption of Theorem 3.1,*

$$(a) \quad \frac{1}{\sqrt{n}} N_1^{-1} \sum_{t=1}^{n} U_{t-1} \xrightarrow{\mathcal{L}} \xi^*,$$

$$(b) \quad N_1^{-1} \sum_{t=1}^{n} U_{t-1} \mathbf{B}_t' \xrightarrow{\mathcal{L}} \xi(\alpha),$$

$$(c) \quad \sum_{t=1}^{n} N_1^{-1} U_{t-1} U_{t-1}' N_1^{-1} \xrightarrow{\mathcal{L}} \Gamma,$$

$$(d) \quad n^{-j} \sum_{t=1}^{n} U_{t-1}^{\pm}(j) \mathbf{B}_t \xrightarrow{\mathcal{L}} \int_0^1 \Gamma_{j-1}^{\pm}(\tau) d\mathbf{K}_1(\tau, \alpha), j = 1, \cdots, a.$$

Proof. For (a), note that

$$U_t(1) = \sum_{i=1}^{t} U_i(0) = \sum_{i=1}^{t} \varepsilon_i,\ U_t(j+1) = \sum_{k=1}^{t} U_k(j),$$

where $j = 0, \cdots, a-1$. By Theorem 2.3 of CW and Theorem 2.2,

$$n^{\frac{1}{2}-j} U_{[n\tau]}(j) \xrightarrow{\mathcal{L}} \Gamma_{j-1}(\tau) \text{ in } D \text{ for } j = 1, \cdots, a. \tag{5.6}$$

Again, by Theorem 2.3 of CW, we obtain

$$\sqrt{n} N_1^{-1} U_{[n\tau]} \xrightarrow{\mathcal{L}} (\Gamma_{a-1}(\tau), \cdots, \Gamma_0(\tau))' \text{ in } D^a. \tag{5.7}$$

By (5.7) and the continuous mapping theorem (Billingsley, 1968, Theorem 5.1),

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} N_1^{-1} U_{t-1} = \frac{1}{n} \sum_{t=1}^{n} (\sqrt{n} N_1^{-1} U_{t-1}) \xrightarrow{\mathcal{L}} \xi^* \text{ in } D^a, \tag{5.8}$$

so that (a) holds. By (5.7) and Theorem 2.2, applying Theorem 2.4 of CW, (b) holds. By (5.7) and the continuous mapping theorem, it is easy to show that (c) holds. Again by (5.7) and the continuous mapping theorem, we have

$$n^{\frac{1}{2}-j}U^{\pm}_{[n\tau]}(j) \xrightarrow{\mathcal{L}} \Gamma^{\pm}_{j-1}(\tau) \text{ in } D \text{ for } j = 1, \cdots, a. \tag{5.9}$$

Furthermore, by Theorem 2.4 of CW and Theorem 2.2, we know that (d) holds. This completes the proof. □

Denote $V_t^+(j) = \max\{0, (-1)^t V_t(j)\}$, $V_t^-(j) = V_t^+(j) - (-1)^t V_t(j)$, $\tilde{\Gamma}_j^+(\tau) = \max\{0, \tilde{\Gamma}_j(\tau)\}$ and $\tilde{\Gamma}_j^-(\tau) = \tilde{\Gamma}_j^+(\tau) - \tilde{\Gamma}_j(\tau)$. For the process $\{V_t\}$ defined in (5.2), we have the following lemma.

Lemma 5.4. *Under the assumptions of Theorem 3.1,*

$$(a) \quad \frac{1}{\sqrt{n}} N_2^{-1} \sum_{t=1}^{n} V_{t-1} \xrightarrow{p} 0,$$

$$(b) \quad N_2^{-1} \sum_{t=1}^{n} V_{t-1}\mathbf{B}_t' \xrightarrow{\mathcal{L}} \eta(\alpha),$$

$$(c) \quad \sum_{t=1}^{n} N_2^{-1} V_{t-1} V_{t-1}' N_2^{-1} \xrightarrow{\mathcal{L}} \tilde{\Gamma},$$

$$(d) \quad n^{-j} \sum_{t=1}^{n} V_{t-1}^{\pm}(j)(-1)^t \mathbf{B}_t \xrightarrow{\mathcal{L}} \int_0^1 \tilde{\Gamma}^{\pm}_{j-1}(\tau) d\mathbf{K}_2(\tau, \alpha), j = 1, \cdots, b.$$

Proof. It is similar to the proof of Lemma 5.3, and hence is omitted. □

In the following, we will show the asymptotic properties of the process $\{Y_t(k)\}$ defined in (5.2), where $k = 1, \cdots, l$. Let

$$S_t(k, j) = \sum_{i=1}^{t} Y_i(k, j) \sin\theta_k \text{ and } T_t(k, j) = \sum_{i=1}^{t} Y_i(k, j) \cos\theta_k.$$

Denote $S_t^+(k, j) = \max\{0, S_t(k, j)\}$ and $S_t^-(k, j) = S_t^+(k, j) - S_t(k, j)$, and similarly define $T_t^{\pm}(k, j)$, where $k = 1, \cdots, l$, $j = 0, \cdots, d_k$.

Lemma 5.5. *Under the assumption of Theorem 3.1,*

$$(a) \quad \frac{1}{\sqrt{n}} N_{k+2}^{-1} \sum_{t=1}^{n} Y_{t-1}(k) \xrightarrow{p} 0,$$

$$(b) \quad N_{k+2}^{-1} \sum_{t=1}^{n} Y_{t-1}(k)\mathbf{B}_t' \xrightarrow{\mathcal{L}} \zeta_k(\alpha),$$

$$(c) \quad N_{k+2}^{-1}\sum_{t=1}^{n} Y_{t-1}(k)Y'_{t-1}(k)N_{k+2}^{-1} \xrightarrow{\mathcal{L}} H_k,$$

$$(d) \quad n^{-j}\sum_{t=1}^{n}\begin{pmatrix} S_{t-1}^{\pm}(k,j) \\ T_{t-1}^{\pm}(k,j)\end{pmatrix}(\cos t\theta_k \mathbf{B}'_t, \sin t\theta_k \mathbf{B}'_t) \xrightarrow{\mathcal{L}} \int_0^1 \begin{pmatrix} f_{kj}^{\pm}(\tau) \\ g_{kj}^{\pm}(\tau)\end{pmatrix} d(\mathbf{K}'_{2k+1}(\tau,\alpha), \mathbf{K}'_{2k+2}(\tau,\alpha)),$$

where $f_{kj}^{+}(\tau) = \max\{0, f_{kj}\}$ *and* $f_{kj}^{-}(\tau) = f_{kj}^{+}(\tau) - f_{kj}(\tau)$, *and similarly define* $g_{kj}^{\pm}(\tau)$.

Proof. By direct verification, we have

$$Y_t(k,j)\sin\theta_k = S_t(k,j-1)\sin(t+1)\theta_k - T_t(k,j-1)\cos(t+1)\theta_k, \tag{5.10}$$

where $j = 1, \cdots, d_k$. By Lemma 3.3.7 of CW,

$$\sqrt{2}n^{-j-1/2}(S_{[n\tau]}(k,j), T_{[ns]}(k,j)) \xrightarrow{\mathcal{L}} (f_{kj}(\tau), g_{kj}(s)) \text{ in } D^2, \tag{5.11}$$

where $k = 1, \cdots, l$, $j = 0, \cdots, d_k - 1$. By Proposition 8 of Jeganathan (1991), we obtain

$$\max_{1\le i\le n} |\frac{1}{n}\sum_{t=1}^{i} n^{-(j-1)-1/2} S_{t-1}(k,j-1)\sin t\theta_k| = o_p(1), \tag{5.12}$$

$$\max_{1\le i\le n} |\frac{1}{n}\sum_{t=1}^{i} n^{-(j-1)-1/2} T_{t-1}(k,j-1)\cos t\theta_k| = o_p(1), \tag{5.13}$$

where $j = 1, \cdots, d_k$. By (5.10) and (5.12)-(5.13), we have

$$||\frac{1}{\sqrt{n}}\sum_{t=1}^{n} N_{k+2}^{-1} Y_{t-1}(k)|| = o_p(1),\ k = 1, \cdots, l,$$

so that (a) holds. By Theorem 2.2 and (5.12)-(5.13), the proofs of (b)-(c) are similar to those given in CW, and hence are omitted. By Theorem 2.4 of CW, (5.11), Theorem 2.2, and the continuous mapping theorem, we can show that (d) holds. This completes the proof. □

Lemma 5.6. *Under the assumption of Theorem 3.1, for any constant* $M \ge 0$,

$$(a) \sup_{\alpha\in\omega(\epsilon), ||s||\le M} ||\delta'_n[\mathbf{T}_n(s,\alpha) - \mathbf{T}_n(0,\alpha) - \sum_{t=1}^{n} X_{t-1}X'_{t-1}\delta_n s q(\alpha)]|| = o_p(1),$$

$$(b) \sup_{\alpha\in\omega(\epsilon)} ||\delta'_n \mathbf{T}_n(0,\alpha)|| = O_p(1).$$

Proof. By (5.2),

$$\delta_n' X_t = \left[n^{-1/2}, (N_1^{-1}U_t)', (N_2^{-1}V_t)', (N_3^{-1}Y_t(1))', \cdots, (N_{l+2}^{-1}Y_t(l))', n^{-1/2}Z_t'\right]'. \quad (5.14)$$

Let $S_n(\tau) = n^{1/2-j}U^{\pm}_{[n\tau]}(j)$ and $\xi_{nt} = g_t(s,0)$. By (5.9), Lemma 5.3 (d), and Lemma 5.1 (a), $\{S_n(\tau), \tau \in [0,1]\}$ and ξ_{nt} satisfy the conditions of Theorem 2.1. Thus, by Theorem 2.1, for any $s \in R^{p+1}$,

$$\sup_{\alpha\in\omega(\epsilon)} \Bigg| n^{-j}\sum_{t=1}^{n} U^{\pm}_{t-1}(j)[I(\varepsilon_t \le F^{-1}(\alpha) + g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0)) - I(\varepsilon_t \le F^{-1}(\alpha)) + \alpha]\Bigg| = o_p(1), \quad (5.15)$$

$$n^{-j}\sum_{t=1}^{n} U^{\pm}_{t-1}(j)[I(\varepsilon_t \le F^{-1}(\alpha)) - \alpha] \xrightarrow{\mathcal{L}} \int_0^1 \Gamma^{\pm}_{j-1}(\tau)dK_1(\tau,\alpha) \text{ in } D[\omega(\epsilon)], \quad (5.16)$$

where $K_1(\tau,\alpha)$ is a Kiefer process in D_2 with the finite-dimensional distribution $\mathbf{K}_1(\tau,\alpha)$. Let $\gamma_{nt} = n^{-j}U_{t-1}(j)$. By Lemma 5.1 (c), we know that (5.3) is satisfied. By Lemma 5.2, we have

$$\sup_{\alpha\in\omega(\epsilon),||s||\le M} |n^{-j}\sum_{t=1}^{n} U_{t-1}(j)[I(\varepsilon_t \le F^{-1}(\alpha) + g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0)) - I(\varepsilon_t \le F^{-1}(\alpha)) + \alpha]| = o_p(1). \quad (5.17)$$

By (5.16) and the continuous mapping theorem,

$$\sup_{\alpha\in\omega(\epsilon)} |n^{-j}\sum_{t=1}^{n} U_{t-1}(j)[I(\varepsilon_t \le F^{-1}(\alpha)) - \alpha]| \xrightarrow{\mathcal{L}} \sup_{\alpha\in\omega(\epsilon)} |\int_0^1 \Gamma_{j-1}(\tau)dK_1(\tau,\alpha)|,$$

that is,

$$\sup_{\alpha\in\omega(\epsilon)} |n^{-j}\sum_{t=1}^{n} U_{t-1}(j)[I(\varepsilon_t \le F^{-1}(\alpha)) - \alpha]| = O_p(1), \quad (5.18)$$

where $j = 1, \cdots, a$. By the triangle inequality and (5.17)-(5.18), we obtain

$$\sup_{\alpha\in\omega(\epsilon),||s||\le M} ||N_1^{-1}\sum_{t=1}^{n} U_{t-1}[I(\varepsilon_t \le F^{-1}(\alpha) + g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0)) - I(\varepsilon_t \le F^{-1}(\alpha)) + \alpha]|| = o_p(1), \quad (5.19)$$

$$\sup_{\alpha\in\omega(\epsilon)} ||N_1^{-1}\sum_{t=1}^{n} U_{t-1}[I(\varepsilon_t \le F^{-1}(\alpha)) - \alpha]|| = O_p(1). \quad (5.20)$$

Similarly, by Lemma 5.4 (d), Theorem 2.1 and Lemmas 5.1-5.2, we can show that

$$\sup_{\alpha\in\omega(\epsilon),||s||\leq M}||N_2^{-1}\sum_{t=1}^{n}V_{t-1}[I(\varepsilon_t\leq F^{-1}(\alpha)+g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0))-I(\varepsilon_t\leq F^{-1}(\alpha))+\alpha]||=o_p(1), \tag{5.21}$$

$$\sup_{\alpha\in\omega(\epsilon)}||N_2^{-1}\sum_{t=1}^{n}V_{t-1}(k)[I(\varepsilon_t\leq F^{-1}(\alpha))-\alpha]||=O_p(1). \tag{5.22}$$

In a similar manner to (5.17)-(5.18), by (5.11), Lemma 5.5 (d), Lemmas 5.1-5.2 and Theorem 2.1, we can show that

$$\sup_{\alpha\in\omega(\epsilon),||s||\leq M}||n^{-j}\sum_{t=1}^{n}\begin{pmatrix} S_{t-1}(k,j)\\ T_{t-1}(k,j)\end{pmatrix}(\cos t\theta_k,\sin t\theta_k)[I(\varepsilon_t\leq F^{-1}(\alpha)+g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0))-I(\varepsilon_t\leq F^{-1}(\alpha))+\alpha]||=o_p(1), \tag{5.23}$$

$$\sup_{\alpha\in\omega(\epsilon)}||n^{-j}\sum_{t=1}^{n}\begin{pmatrix} S_{t-1}(k,j)\\ T_{t-1}(k,j)\end{pmatrix}(\cos t\theta_k,\sin t\theta_k)[I(\varepsilon_t\leq F^{-1}(\alpha))-\alpha]||=O_p(1), \tag{5.24}$$

where $k=1,\cdots,l$ and $j=1,\cdots,d_k$. Using the equation:

$$Y_{t-1}(k,j)\sin\theta_k=\cos\theta_k[S_{t-1}(k,j)\cos(t+1)\theta_k \\ -T_{t-1}(k,j)\sin(t+1)\theta_k]-\sin\theta_k[S_{t-1}(k,j)\sin(t+1)\theta_k+T_{t-1}(k,j)\cos(t+1)\theta_k]$$

for $k=1,\cdots,l$ and $j=1,\cdots,d_k$, and by (5.10), (5.23)-(5.24), and the triangle inequality, we can show that

$$\sup_{\alpha\in\omega(\epsilon),||s||\leq M}||N_{k+2}^{-1}\sum_{t=1}^{n}Y_{t-1}(k)[I(\varepsilon_t\leq F^{-1}(\alpha)+g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0))-I(\varepsilon_t\leq F^{-1}(\alpha))+\alpha]||=o_p(1), \tag{5.25}$$

$$\sup_{\alpha\in\omega(\epsilon)}||N_{k+2}^{-1}\sum_{t=1}^{n}Y_{t-1}(k)[I(\varepsilon_t\leq F^{-1}(\alpha))-\alpha]||=O_p(1). \tag{5.26}$$

Let $\tilde{\gamma}_t$ be 1 or any element of Z_{t-1}. Since $\{Z_{t-1}\}$ is stationary and ergodic, we have $(n^{-1}\sum_{t=1}^{n}\tilde{\gamma}_t^2)^{1/2}=\tilde{\gamma}+o_p(1)$, where $\tilde{\gamma}$ is a positive constant. By Lemma 5.1 (a), $n^{-1/2}\max_{1\leq t\leq n}|\tilde{\gamma}_t|=o_p(1)$ and $\max_{1\leq t\leq n}|g_t(s,0)|=o_p(1)$. Now applying Theorem 1.1 of Koul and Ossiander (1994) and Lemma 5.2, we can show that

$$\sup_{\alpha\in\omega(\epsilon),||s||\leq M}\frac{1}{\sqrt{n}}|\sum_{t=1}^{n}\tilde{\gamma}_{t-1}[I(\varepsilon_t\leq F^{-1}(\alpha)+g_t(s,0)) \\ -F(F^{-1}(\alpha)+g_t(s,0))-I(\varepsilon_t\leq F^{-1}(\alpha))+\alpha]|=o_p(1), \tag{5.27}$$

$$\sup_{\alpha\in\omega(\epsilon)}\frac{1}{\sqrt{n}}|\sum_{t=1}^{n}\tilde{\gamma}_{t-1}[I(\varepsilon_t\leq F^{-1}(\alpha))-\alpha]|=O_p(1). \tag{5.28}$$

By (5.14), (5.19), (5.21), (5.25), (5.27) and the triangle inequality, we can show that (a) holds. Similarly by (5.14), (5.20), (5.22), (5.26) and (5.28), we can show that (b) holds. This completes the proof. $\Box$

Lemma 5.7. *Under the assumption of Theorem 3.1,*

$$\sup_{\alpha\in\omega(\epsilon)} ||\delta_n' \mathbf{T}_n(\delta_n^{-1}[\hat{\phi}_n(\alpha)-\phi(\alpha)],\alpha)|| = o_p(1).$$

Proof. Denote $\mathcal{W}_n = [X_1,\cdots,X_n]'$ and $\mathcal{Y}_n = [1, y_1,\cdots,y_n]'$. Under model (1.1), the rows of $\mathcal{W}_n$ are linearly independent a.s. and the columns of $\mathcal{W}_n$ are also linearly independent a.s. (otherwise, ε_t will be $\mathcal{F}_{t-1}-$measurable). Let $\mathbf{h}$ be a subset of $\{1,\cdots,n\}$ of size $p+1$ and $\mathcal{W}_h$ (or $\mathcal{Y}_h$) be the subdesign matrix (or subresponse vector) with row $X_{i-1}', i\in \mathbf{h}$ (or coordinates $y_i, i\in\mathbf{h}$). Then $\mathcal{W}_h$ is invertible a.s.. By a linear programming algorithm given by Koenker and Bassett (1978) and Koul and Saleh (1995), $\hat{\phi}_n(\alpha)$ is a solution of the form $b=\mathcal{W}_h^{-1}\mathcal{Y}_h$. Furthermore, note that $\mathbf{T}_n(\delta_n^{-1}[\hat{\phi}_n(\alpha)-\phi(\alpha)],\alpha) = \sum_{t=1}^n X_{t-1}\{I(\varepsilon_t\le[\hat{\phi}_n(\alpha)-\phi(\alpha)]'X_{t-1}+F^{-1}(\alpha))-\alpha\} = \sum_{t=1}^n X_{t-1}\{I(y_t-\hat{\phi}_n'(\alpha)X_{t-1}\le 0)-\alpha\}$. In a similar manner to Koul and Saleh (1995), by the inequality in (3.1) of Theorem 3.3 of Koenker and Bassett (1978), we can show that

$$\sup_{\alpha\in\omega(\epsilon)} ||\delta_n' \mathbf{T}_n(\delta_n^{-1}[\hat{\phi}_n(\alpha)-\phi(\alpha)],\alpha)|| \le 2(p+1)\max_{1\le t\le n}||\delta_n' X_{t-1}||.$$

By Lemma 5.1 (a), this completes the proof. $\Box$.

Proof of Theorem 3.1. Denote $\Upsilon_n(\alpha)=\delta_n^{-1}[\hat{\phi}_n(\alpha)-\phi(\alpha)]$. For any $\varepsilon,\eta>0$, by Lemma 5.7, there exists an integer $n_1>0$ such that, when $n>n_1$,

$$P\Big\{\sup_{\alpha\in\omega(\epsilon)} ||\delta_n'\mathbf{T}_n(\Upsilon_n(\alpha),\alpha)|| > \eta\Big\} < \varepsilon.$$

Thus, for a positive constant M, when $n>n_1$,

$$\begin{aligned} &P\left\{||\Upsilon_n(\alpha)||\ge M, \forall\alpha\in\omega(\epsilon)\right\} \\ &\quad\le P\left\{||\Upsilon_n(\alpha)||\ge M, ||\delta_n'\mathbf{T}_n(\Upsilon_n(\alpha),\alpha)||\le\eta, \forall\alpha\in\omega(\epsilon)\right\} \\ &\qquad + P\left\{||\delta_n'\mathbf{T}_n(\Upsilon_n(\alpha),\alpha)||\ge\eta,\forall\alpha\in\omega(\epsilon)\right\} \\ &\quad\le P\Big\{\inf_{||s_1||\ge M}||\delta_n'\mathbf{T}_n(s_1,\alpha)||\le\eta,\forall\alpha\in\omega(\epsilon)\Big\}+\varepsilon. \end{aligned} \tag{5.29}$$

Note that $s_1'\delta_n'\mathbf{T}_n(\lambda s_1, \alpha)$ is a non-decreasing function of λ for any $\alpha \in (0,1)$ and $s_1 \in R^{p+1}$. Writing s_1 as $s_1 = \lambda s$ with $\lambda \geq 1$ and $||s|| = M$ for any $||s_1|| \geq M$, by the Cauchy-Schwarz inequality, we have

$$\inf_{||s||=M} |s'\delta_n'\mathbf{T}_n(s,\alpha)| \leq \inf_{||s||=M, \lambda\geq 1} |s'\delta_n'\mathbf{T}_n(\lambda s,\alpha)| \leq M \inf_{||s_1||\geq M} ||\delta_n'\mathbf{T}_n(s_1,\alpha)||.$$

Thus, by (5.29),

$$\begin{aligned} &P\left\{||\Upsilon_n(\alpha)|| \geq M, \forall \alpha \in \omega(\epsilon)\right\} \\ &\leq P\Big\{\inf_{||s||=M} ||s'\delta_n'\mathbf{T}_n(s,\alpha)|| \leq \eta M, \forall \alpha \in \omega(\epsilon)\Big\} + \varepsilon. \end{aligned} \tag{5.30}$$

Denote

$$\Omega_n = \delta_n' \sum_{t=1}^{n} X_{t-1}X_{t-1}'\delta_n, \quad q_\epsilon = \inf_{\alpha\in\omega(\epsilon)} q(\alpha),$$
$$R_n(\alpha) = \sup_{||s||=M} |s'\delta_n'[\mathbf{T}_n(s,\alpha) - \mathbf{T}_n(0,\alpha)] - s'\Omega_n s q(\alpha)|.$$

Since

$$|s'\delta_n'\mathbf{T}_n(s,\alpha)| \geq \inf_{||s||=M}[s'\Omega_n s q(\alpha)] - R_n(\alpha) - \sup_{||s||=M} |s'\delta_n'\mathbf{T}_n(0,\alpha)|,$$

by (5.30),

$$\begin{aligned} &P\Big\{||\Upsilon_n(\alpha)|| \geq M, \forall \alpha \in \omega(\epsilon)\Big\} \leq P\Big\{R_n(\alpha) \geq \\ &\inf_{||s||=M}[s'\Omega_n s q(\alpha)] - \sup_{||s||=M} |s'\delta_n'\mathbf{T}_n(0,\alpha)| - \eta M, \forall \alpha \in \omega(\epsilon)\Big\} + \varepsilon. \end{aligned} \tag{5.31}$$

By Theorem 3.5.1 of CW and (5.37) below, Ω_n converges to a matrix Ω_x in distribution and Ω_x is positive definite a.s.. Denote λ_n and λ_0 as the minimum eigenvalues of Ω_n and Ω_x, respectively. Then λ_n converges to λ_0 in distribution with $\lambda_0 > 0$ a.s.. For the above ε, there exists a constant $c_0 > 0$ such that $P(\lambda_0 < c_0) < \varepsilon/2$. Furthermore, there exists an integer n_2 such that, when $n > n_2$,

$$P(\inf_{||s||=M} s'\Omega_n s < c_0 M^2) \leq P(\lambda_n < c_0) < P(\lambda_0 < c_0) + \varepsilon/2 < \varepsilon. \tag{5.32}$$

By Lemma 5.6 (b), there exists a large constant M_1 and an integer n_3 such that, when $n > n_3$,

$$\begin{aligned} &P(\sup_{||s||=M} |s'\delta_n'\mathbf{T}_n(0,\alpha)| > MM_1, \forall \alpha \in \omega(\epsilon)) \\ &\qquad \leq P(||\delta_n'\mathbf{T}_n(0,\alpha)|| > M_1, \forall \alpha \in \omega(\epsilon)) < \varepsilon. \end{aligned} \tag{5.33}$$

Thus, by (5.32)-(5.33), when $n > \max\{n_2, n_3\}$,

$$\begin{aligned}
&P\Big\{R_n(\alpha) \geq \inf_{||s||=M}[s'\Omega_n s q(\alpha)] - \sup_{||s||=M}|s'\delta_n'\mathbf{T}_n(0,\alpha)| - \eta M, \forall \alpha \in \omega(\epsilon)\Big\} \\
&\quad \leq P\Big\{R_n(\alpha) \geq \inf_{||s||=M}[s'\Omega_n s q(\alpha)] - \sup_{||s||=M}|s'\delta_n'\mathbf{T}_n(0,\alpha)| - \eta M, \\
&\qquad \sup_{||s||=M}|s'\delta_n'\mathbf{T}_n(0,\alpha)| \leq MM_1, \inf_{||s||=M}[s'\Omega_n s] \geq c_0 M^2, \forall \alpha \in \omega(\epsilon)\Big\} \\
&\quad + P\Big(\sup_{||s||=M}|s'\delta_n'\mathbf{T}_n(0,\alpha)| > MM_1, \forall \alpha \in \omega(\epsilon)\Big) + P\Big(\inf_{||s||=M}[s'\Omega_n s] < c_0 M^2\Big) \\
&\quad \leq P\Big\{R_n(\alpha) \geq c_0 M^2 q_\epsilon - MM_1 - \eta M, \forall \alpha \in \omega(\epsilon)\Big\} + 2\varepsilon. \qquad (5.34)
\end{aligned}$$

We may choose M large enough such that $c = c_0 M q_\epsilon - M_1 - \eta > 0$. For the constant c, by Lemma 5.6 (a), there exists an integer n_4 such that, when $n > n_4$,

$$\begin{aligned}
&P\Big\{R_n(\alpha) \geq Mc, \forall \alpha \in \omega(\epsilon)\Big\} \\
&\leq P\Big\{\sup_{||s||=M}||\delta_n'[\mathbf{T}_n(s,\alpha) - \mathbf{T}_n(0,\alpha)] - s'\Omega_n s q(\alpha)]|| \geq c, \forall \alpha \in \omega(\epsilon)\Big\} < \varepsilon. \qquad (5.35)
\end{aligned}$$

Thus, by (5.31) and (5.34)-(5.35), when $n > \max\{n_1, n_2, n_3, n_4\}$, $P\{||\ \Upsilon_n(\alpha)|| \geq M, \forall \alpha \in \omega(\epsilon)\} < 4\varepsilon$. Finally, by Lemmas 5.6 (a) and 5.7 , we have

$$\hat{\phi}_n(\alpha) - \phi(\alpha) = -[q(\alpha)\sum_{t=1}^{n} X_{t-1}X_{t-1}']^{-1}\mathbf{T}_n(0,\alpha) + o_p(\delta_n),$$

where $o_p(\cdot)$ holds uniformly for $\alpha \in \omega(\epsilon)$. This completes the proof.□

Proof of Theorem 3.2. Since Z_t is a stationary and ergodic time series, by the ergodic theorem, $n^{-1}\sum_{t=1}^{n} Z_{t-1} = o_p(1)$ and $n^{-1}\sum_{t=1}^{n} Z_{t-1}Z_{t-1}' = \Sigma + o_p(1)$. By Theorems 3.4.1 and 3.4.2 in CW, the quantities $\sum_{t=1}^{n}(N_1^{-1}U_{t-1}V_{t-1}'N_2^{-1})$, $\sum_{t=1}^{n}(N_1^{-1}U_{t-1}Y_{t-1}'(k)N_{k+2}^{-1})$, $\sum_{t=1}^{n}(N_2^{-1}V_{t-1}Y_{t-1}'(k)N_{k+2}^{-1})$, $n^{-1/2}\sum_{t=1}^{n}(N_1^{-1}U_{t-1}Z_{t-1}')$, $n^{-1/2}\sum_{t=1}^{n}(N_2^{-1}V_{t-1}Z_{t-1}')$ and $n^{-1/2}\sum_{t=1}^{n}(N_{k+2}^{-1}Y_{t-1}(k)Z_{t-1}')$ converge to zero in probability, where $k = 1, \cdots, l$. Furthermore, by Lemmas 5.3-5.5 (a)-(c), we have

$$\delta_n^{-1}\sum_{t=1}^{n} X_{t-1}\mathbf{B}_t' \xrightarrow{\mathcal{L}} \{\mathbf{K}_1'(1,\alpha), \xi(\alpha), \eta(\alpha), \zeta_1(\alpha), \cdots, \zeta_l(\alpha), N_\alpha\}, \qquad (5.36)$$

$$\delta_n^{-1}\sum_{t=1}^{n} X_{t-1}X_{t-1}'\delta_n^{-1} \xrightarrow{\mathcal{L}} \text{diag}\Big\{\begin{pmatrix} 1 & \xi^* \\ \xi^{*\prime} & F \end{pmatrix}, \tilde{F}, H_1, \cdots, H_l, \Sigma\Big\}. \qquad (5.37)$$

By Theorem 3.5.1 of CW, the limiting matrix of (5.37) is positive definite a.s.. By Theorem 3.1,

$$\delta_n^{-1}[\hat{\phi}_n(\alpha_1) - \phi(\alpha_1), \cdots, \hat{\phi}_n(\alpha_m) - \phi(\alpha_m)] = \qquad (5.38)$$

$$\Big[\delta_n^{-1}\sum_{t=1}^n X_{t-1}X_{t-1}'\delta_n^{-1}\Big]^{-1}\Big[\delta_n^{-1}\sum_{t=1}^n X_{t-1}\mathbf{B}_t'\Big]\text{diag}\Big[\frac{1}{q(\alpha_1)},\cdots,\frac{1}{q(\alpha_m)}\Big]+o_p(1).$$

Note that the random matrices and vectors involved in (5.36)-(5.37) are functionals of the corresponding process of (2.2). By (5.36)-(5.37) and the continuous mapping theorem, this completes the proof. □

REFERENCES

Bassett, G. W. and Koenker, R. W. (1982). An empirical quantiles function for linear models with i.i.d. errors. *J. Amer. Statist. Assoc.* **77**, 407-415.

Bickel, P. J. and Wichura, M. J. (1971). Convergence for multiparameter stochastic processes and some applications. *Ann. Statist.* **42**, 1656-1670.

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

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

Cox, D. D. and Llatas, I. (1991). Maximum likelihood type estimation for nearly nonstationary autoregressive time series. *Ann. Statist.* **19**, 1109-1128.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* **74**, 427-431.

Dickey, D. A., Hasza, D. P. and Fuller, W. A. (1984). Testing for units roots in seasonal time series. *J. Amer. Statist. Assoc.* **79**, 355-367.

Hansen, B. E. (1995). Rethinking the univariate approach to unit root testing. Using covariates to increase power. Trending multiple time series. *Econometric Theory.* **11**, 1148-1171.

Hasan M. N. and Koenker, R.(1997). Robust rank tests of the unit root hypothesis. *Econometrica* **65**, 133-61.

Herce, M. A. (1996). Asymptotic theory of LAD estimation in a unit root process with finite variance errors. *Econometric Theory* **12**, 129-153.

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

Jeganathan, P. (1991). On the asymptotic behavior of least squares estimators in AR time series with roots near the unit circle. *Econometric Theory* **7**, 269-306.

Koenker, R. W. and Bassett, G. W. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

Koenker, R. W. and Bassett, G. W. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica* **50**, 43-61.

Koenker, R. W. and D'Orey, V. (1987). Algorithm AS 229: Computing regression quantiles. *J. Roy. Statist. Soc. Ser. C* **36**, 383-393.

Koenker, R. W. and D'Orey, V. (1993). Remark AS229: Computing regression quantiles. *J. Roy. Statist. Soc. Ser. C* **43**, 410-414.

Koul, H. L. (1991). A weak convergence result useful in robust autoregression. *J. Statist. Plann. Inference* **29**, 291-308.

Koul, H. L.(1996). Asymptotics of some estimators and sequential residual empiricals in nonlinear time series. *Ann. Statist.*, **24**, 380-404.

Koul, H. L. and Ossiander, M. (1994). Weak convergence of randomly weighted dependent residual empiricals with applications to autoregression. *Ann. Statist.* **22**, 540-562.

Koul, H. L. and Saleh, A. K. Md. E. (1995). Autoregression quantiles and related rank scores processes. *Ann. Statist.* **23**, 670-689.

Ling , S. (1998). Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models. *Ann. Statist.* **26**, 741-754.

Ling , S. and Li, W. K. (1998). Limiting distributions of maximum likelihood estimators for unstable ARMA time series with GARCH errors. *Ann. Statist.* **26**, 84-125.

Ling , S. and Li, W. K. (2001). Asymptotic inference of nonstationary fractional ARIMA models. To appear in *Econometric Theory.*

Lucas, A. (1995). Unit root tests based on $M-$estimators. *Econometric Theory* **11**, 331-346.

van der Meer, T., Pap, G. and van Zuijlen, M. C. A. (1999). Asymptotic inference for nearly unstable AR(p) processes. *Econometric Theory*, **15**, 184-217.

Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica* **55**, 277-301.

Phillips, P. C. B. (1989). Partially identified econometric models. *Econometric Theory* **5**, 181-240.

Phillips, P. C. B. and Durlauf, S. N. (1986). Multiple time series regression with integrated processes. *Review of Economic Studies* **53**, 473-495.

Phillips, P. C. B. and Xiao, Z. (1998). A primer on unit root testing. *Journal of Economic Surveys* **12**, 423-470. Reprinted in M. McAleer and L. Oxley (eds.), *Practical Issues in Cointegration Analysis.* Blackwell, Oxford, 1999.

Portnoy, S. and Koenker, R. W. (1989). Adaptive $L-$estimation for linear models. *Ann. Statist.* **17**, 362-381.

Ruppert, D. and Carroll, R. J. (1980). Trimmed least squares estimation in the linear model. *J. Amer. Statist. Assoc.* **75**, 828-838.

Straf, M. J. (1970). Weak convergence of stochastic process with several parameters. *Proc. Fourth Berkeley Symp. Math. Statist. and Probab.* **4**, 187-221. Univ. California Press, Berkeley.

Truong-Van, B. and Larramendy, I. (1996). Asymptotic distribution of least squares estimators for purely unstable ARMA (m,∞). *Statistics* **28**, 307-346.