The Annals of Statistics
1998, Vol. 26, No. 1, 84–125

LIMITING DISTRIBUTIONS OF MAXIMUM LIKELIHOOD ESTIMATORS FOR UNSTABLE AUTOREGRESSIVE MOVING-AVERAGE TIME SERIES WITH GENERAL AUTOREGRESSIVE HETEROSCEDASTIC ERRORS

BY SHIQING LING AND W. K. LI[1]

University of Hong Kong

This paper investigates the maximum likelihood estimator (MLE) for unstable autoregressive moving-average (ARMA) time series with the noise sequence satisfying a general autoregressive heteroscedastic (GARCH) process. Under some mild conditions, it is shown that the MLE satisfying the likelihood equation exists and is consistent. The limiting distribution of the MLE is derived in a unified manner for all types of characteristic roots on or outside the unit circle and is expressed as a functional of stochastic integrals in terms of Brownian motions. For various types of unit roots, the limiting distribution of the MLE does not depend on the parameters in the moving-average component and hence, when the GARCH innovations reduce to usual white noises with a constant conditional variance, they are the same as those for the least squares estimators (LSE) for unstable autoregressive models given by Chan and Wei (1988). In the presence of the GARCH innovations, the limiting distribution will involve a sequence of independent bivariate Brownian motions with correlated components. These results are different from those already known in the literature and, in this case, the MLE of unit roots will be much more efficient than the ordinary least squares estimation.

1. Introduction. Consider the autoregressive moving-average (ARMA) time series y_t, $t = 1, 2, 3, \ldots$, with the general autoregressive heteroscedastic (GARCH) error process given by

$$\phi(B)y_t = \psi(B)\varepsilon_t, \tag{1.1}$$

$$\varepsilon_t = \eta_t\sqrt{h_t}, \qquad h_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i}, \tag{1.2}$$

where η_t is a sequence of independently and identically distributed (i.i.d.) random variables with zero mean and variance 1; $\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\psi(B) = 1 + \sum_{i=1}^{q} \psi_i B^i$ are polynomials in the backshift operator B with $\phi_p \neq 0$ and $\psi_q \neq 0$ and have no common root; $\alpha_0 > 0$, $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s \geq 0$ and the polynomials $\alpha(B) = \sum_{i=1}^{r} \alpha_i B^i$ and $\beta(B) = 1 - \sum_{i=1}^{s} \beta_i B^i$ have no common root; $\mathcal{F}_t$ denotes the σ-field generated by the information set

Received September 1996; revised June 1997.
[1]Supported by UGC Grant 338 017 0003.
AMS 1991 *subject classifications*. Primary 62M10, 62E20; secondary 60F17.
Key words and phrases. Bivariate Brownian motion, GARCH models, limiting distribution, maximum likelihood estimation, stochastic integral, unstable ARMA models.

$\{\eta_t, \eta_{t-1}, \ldots\}$. We shall assume that the starting value $X_0^* = (y_0, \ldots, y_{-p+1}, \varepsilon_0, \ldots, \varepsilon_{-q^*+1}, h_0, \ldots, h_{-s+1})$ is known and is $\mathcal{F}_0$-measurable, where $q^* = \max\{r, q\}$. As will be seen later, the starting value has no effect on the asymptotic properties considered. We shall say that the ARMA model (1.1) is unstable if $\phi(z)$ has at least a root on the unit circle.

The nonstationary time series have been extensively investigated for the last decade. Some important results for the nonstationary autoregressive AR(p) models can be found in Fuller (1976), Dickey and Fuller (1979), Hasza and Fuller (1979), Dickey, Hasza and Fuller (1984), Tsay and Tiao (1984), Phillips (1987), Chan and Wei (1987, 1988), Tsay and Tiao (1990) and Jeganathan (1991). Among these authors, Chan and Wei (1988) first obtained general results in a unified manner for all types of unit roots in unstable autoregressive AR(p) models. Jeganathan (1991) derived general results for near-nonstationary AR(p) models. However, two important cases are not yet investigated completely and satisfactorily.

On the one hand, because of practical motivations from applications, recently the nonstationary ARMA models have been studied by many statisticians and econometricians. Tsay and Tiao (1990) discussed the asymptotic properties of the least squares estimation (LSE) for general (multivariate) nonstationary ARMA time series and proved that if the AR part contains stationary components and the moving-average (MA) part is nontrivial, the LSE of the parameters in the AR part will be inconsistent. Pantula and Hall (1991) used an instrumental variable approach to estimate and test the regular unit root in an ARMA model, that is, the presence of the factor $(1 - B)$ in the AR polynomial. Yap and Reinsel (1995a, b) considered a Gaussian likelihood estimation of the ARMA models with regular unit root. They showed that Gaussian likelihood estimation for the unit root is more efficient than Pantula and Hall's instrumental variable approach and their simulation results also demonstrate that the performance of the unit root test based on the Gaussian likelihood estimation is better than that based on the instrumental variable approach. However, the asymptotic properties of the maximum likelihood estimation (MLE) for the general nonstationary ARMA model, that is, $\phi(z)$ with roots 1, -1, $e^{i\theta}$ and $e^{-i\theta}$, have not been obtained. In this paper, our results cover this case.

On the other hand, research on nonstationary time series is almost always limited to innovations with constant conditional variances. In the framework of Phillips and Durlauf (1986) and Phillips (1987), the long-run variance and the innovation variances are equal in the presence of heteroscedasticity, but it does not include many conditional heteroscedastic processes as defined in (1.2). The autoregressive conditional heteroscedastic (ARCH) model, that is, model (1.2) with $s = 0$, was proposed by Engle (1982) and generalized by Bollerslev (1986) as the popular GARCH model (1.2). This is a very important class of time series and they have been widely investigated and applied in the financial and econometric literature. These models are able to model the real situation better and hence result in more efficient estimation and statistical inference. Some excellent surveys on the subject can be found in Bollerslev,

Chou and Kroner (1992) and Bollerslev, Engle and Nelson (1994). There have already been several papers which attempt to link nonstationary time series with ARCH processes. Pantula (1989) derived the asymptotic distribution of the least squares estimator for the AR(p) model with one unit root under a first-order ARCH process for the innovation sequence. He demonstrated that the Dickey–Fuller test can still be employed in that case. In fact, the work of Chan and Wei (1987) can also be applied to the LSE of the nonstationary AR(1) model with GARCH innovations. However, a well-known advantage of stationary time series with ARCH/GARCH innovations is that the MLE is more efficient than the LSE. It seems natural to expect that this advantage still carries over to nonstationary time series. If so, in the presence of the ARCH/GARCH innovations, the MLE will be important for nonstationary time series since one can obtain more satisfactory estimation and inference procedures, especially better unit root tests. Although there are not many results in comparing unit root tests based on the MLE, Peters and Veloce (1988) and Kim and Schmidt (1993) provided simulation results showing that the Dickey–Fuller tests based on the LSE are often too sensitive. Unfortunately, as far as we know, there have not been any asymptotic results for the MLE in the presence of ARCH type errors.

In this paper, our aim is to investigate the MLE for unstable ARMA time series with GARCH processes which links the popular GARCH models and the nonstationary ARMA models. Under some mild conditions, it is shown that the MLE satisfying the likelihood equation exists and is consistent. The limiting distribution of the MLE is derived in a unified manner for all types of characteristic roots on or outside the unit circle and is expressed as a functional of stochastic integrals in terms of Brownian motions. For various types of unit roots, the limiting distributions of the MLE do not depend on the parameters in the moving-average components. Hence, when the GARCH innovations reduce to the usual white noise with a constant conditional variance, they are the same as those of the LSE for unstable autoregressive models given by Chan and Wei (1988). When the GARCH innovations are present, the limiting distribution will involve a sequence of independent bivariate Brownian motions with correlated components. These results are different from those already known and, in this case, the MLE of unit roots will be much more efficient than the ordinary LSE. These asymptotic results not only provide the basis for constructing new unit root tests or other applications, but also help us to understand more comprehensively the nature of nonstationary time series. The method for obtaining these asymptotic results can be applied to other ARCH type innovations and near-nonstationary cases as well as multivariate cases.

The paper proceeds as follows. Section 2 introduces the MLE and main result. Section 3 gives some auxiliary theorems. Section 4 derives asymptotic properties of nonstationary componentwise arguments corresponding to the locations of various unit roots. Section 5 gives the proof of the main result.

Throughout the paper, we use the following notation: U' denotes the transpose of the vector U; $o(1)$ ($o_p(1)$) denotes a series of numbers (random

numbers) converging to zero (in probability); $O(1)$ ($O_p(1)$) denotes a series of numbers (random numbers) that are bounded (in probability); $\to_p$ and $\to_{\mathscr{L}}$ denote convergence in probability and in distribution, respectively. $D = D[0,1]$ denotes the space of functions $f(s)$ on $[0,1]$, which is defined and equipped with the Skorokhod topology [Billingsley (1968)]; $D^n = D \times D \cdots \times D$ (n factors). $\|\cdot\|$ denotes the Euclidean norm.

2. The MLE and main result. Suppose that the observations $y_1, \ldots, y_n$ are generated by the model (1.1)–(1.2). The log-likelihood conditional on the starting value X_0^* is

$$Li_n(\lambda) = \sum_{t=1}^{n} l_t(\lambda) \quad \text{and} \quad l_t = -\frac{1}{2}\ln h_t - \frac{1}{2}\frac{\varepsilon_t^2}{h_t}, \tag{2.1}$$

where $\lambda = (\phi', \psi', \delta')$, $\phi = (\phi_1, \ldots, \phi_p)'$, $\psi = (\psi_1, \ldots, \psi_q)'$ and $\delta = (\alpha_0, \alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s)'$; ε_t and h_t are treated as functions of λ, although ε_t is only a function of (ϕ', ψ'); $\lambda \in \Theta$, which is a compact set in the real space $R^{p+q+r+s+1}$ and λ_0 is the true value of λ. The true errors will be denoted by ε_{0t}, and h_t evaluated at $\lambda = \lambda_0$ is denoted by h_{0t}. The MLE, $\hat{\lambda}_n$, of λ_0 is defined as $\lambda \in \Theta$, which maximizes $Li_n(\lambda)$.

Define the random variables

$$D_t(\lambda) = \frac{\partial l_t(\lambda)}{\partial \lambda} \quad \text{and} \quad I_t(\lambda) = \frac{\partial^2 l_t(\lambda)}{\partial \lambda\, \partial \lambda'}, \tag{2.2}$$

where the formulas of $D_t(\lambda)$ and $I_t(\lambda)$ can be found in Appendix A. We shall use D_t and I_t to denote $D_t(\lambda_0)$ and $I_t(\lambda_0)$, respectively. To obtain the MLE of λ_0, we employ Taylor's expansion to write

$$\frac{\partial Li_n(\lambda)}{\partial \lambda} = \sum_{t=1}^{n} D_t + \sum_{t=1}^{n} I_t \cdot (\lambda - \lambda_0) + \sum_{t=1}^{n} \left[I_t(\lambda_n^*) - I_t \right](\lambda - \lambda_0), \tag{2.3}$$

where $\lambda_n^* = \lambda_0 + v(\lambda - \lambda_0)$ with $v = v(n, \lambda)$ satisfying $|v| \le 1$. Throughout this paper, we suppose that, when $\lambda = \lambda_0$, the following assumptions hold.

ASSUMPTION 1. The characteristic polynomial $\phi(z)$ has the decomposition

$$\phi(z) = (1 - z)^a (1 + z)^b \prod_{k=1}^{l} \left(1 - 2\cos\theta_k z + z^2\right)^{d_k} \phi^*(z),$$

where a, b, l and d_k are nonnegative integers, $\theta_k \in (0, \pi)$, $\phi^*(z) = 1 - \sum_{i=1}^{p^*} \phi_i^* z^i$ with all roots outside the unit circle and $p^* = p - (a + b + 2d_1 + \cdots + 2d_l)$.

ASSUMPTION 2. All roots of $\psi(z)$ are outside the unit circle.

ASSUMPTION 3. $\sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{s} \beta_j < 1$.

ASSUMPTION 4. η_t has a symmetrical distribution.

ASSUMPTION 5. $\rho(E(A_t \otimes A_t)) < 1$, where $\rho(A) = \max\{|\text{all eigenvalues of } A|\}$,

$$A_t = \left(\begin{array}{ccc|ccc} \alpha_1\eta_t^2 & \cdots & \alpha_r\eta_t^2 & \beta_1\eta_t^2 & \cdots & \beta_s\eta_t^2 \\ & I_{(r-1)\times(r-1)} \quad O_{(r-1)\times 1} & & & O_{(r-1)\times s} & \\ \hline \alpha_1 & \cdots & \alpha_r & \beta_1 & \cdots & \beta_s \\ & O_{(s-1)\times r} & & & I_{(s-1)\times(s-1)} \quad O_{(s-1)\times 1} & \end{array}\right),$$

$I_{k\times k}$ is the $k \times k$ identity matrix and $\otimes$ denotes the Kronecker product.

Assumptions 3 and 5 are the second-order stationarity condition given by Bollerslev (1986) and the fourth moment condition given by Ling (1995), respectively, for the GARCH process (1.2). Denote $l_{zt}(\lambda^*) = -(1/2)\ln h_t - (1/2)\varepsilon_t^2/h_t$ with $\varepsilon_t = \psi(B)^{-1}\phi^*(B)z_t$, $\lambda^* = (m^{*\prime}, \delta')'$, $m^* = (\phi^{*\prime}, \psi')'$ and $\phi^* = (\phi_1^*, \ldots, \phi_{p^*}^*)'$. Define $\partial l_{zt}(\lambda^*)/\partial m^*$ and $\partial l_{zt}(\lambda^*)/\partial\delta$ as in (A.1) and (A.2) with $l_t(\lambda)$ and m replaced by $l_{zt}(\lambda^*)$ and m^*, respectively. Corresponding to λ_0, the true value of λ^* is denoted by λ_0^*.

To transform $\{y_t\}$ into various componentwise arguments corresponding to the locations of their roots, now let $u_t = (1-B)^{-a}\phi_0(B)y_t$, $v_t = (1+B)^{-b}\phi_0(B)y_t$, $z_t = \phi_0^{*-1}(B)\phi_0(B)y_t$ and $x_{t,k} = (1 - 2\cos\theta_k B + B^2)^{-d_k}\phi_0(B)y_t$, $k = 1, \ldots, l$, where $\phi_0(B) = \phi(B)|_{\lambda=\lambda_0}$ and similarly for $\psi_0(B)$ and $\phi_0^*(B)$. Then $(1-B)^a u_t = \psi_0(B)\varepsilon_t$, $(1+B)^b v_t = \psi_0(B)\varepsilon_t$, $\phi_0^*(B)z_t = \psi_0(B)\varepsilon_t$, $(1 - 2\cos\theta_k B + B^2)^{d_k}x_{t,k} = \psi_0(B)\varepsilon_t$, $k = 1, \ldots, l$, where a, b, d_k and $\phi^*(B)$ are defined as in Assumption 1. Define $\mathbf{u}_t = (u_t, \ldots, u_{t-a+1})'$, $\mathbf{v}_t = (v_t, \ldots, v_{t-b+1})'$, $\mathbf{z}_t = (z_t, \ldots, z_{t-p^*+1})'$, $\mathbf{x}_{t,k} = (x_{t,k}, \ldots, x_{t-d_k+1,k})'$, $k = 1, \ldots, l$. As shown in Chan and Wei (1988), there exists a nonsingular matrix Q^* such that

$$Q^*\mathbf{y}_t = (\mathbf{u}_t', \mathbf{v}_t', \mathbf{x}_{t,1}', \ldots, \mathbf{x}_{t,l}', \mathbf{z}_t')', \tag{2.4}$$

where $\mathbf{y}_t = (y_t, \ldots, y_{t-p+1})'$. Define

$$\begin{aligned} Q &= \operatorname{diag}\left(Q^*, I_{m_1\times m_1}\right), \\ G_n &= \operatorname{diag}\left(J_n, \tilde{J}_n, L_{1n}, \ldots, L_{ln}, n^{-1/2}I_{m_2\times m_2}\right), \end{aligned} \tag{2.5}$$

where J_n, $\tilde{J}_n$, L_{kn}, $k = 1, \ldots, l$, are defined as in Section 4, $m_1 = q + r + s + 1$ and $m_2 = p^* + q + r + s + 1$. Our main results can be stated as the following theorem.

THEOREM 2.1. *Under Assumptions* 1–5 *of the model* (1.1)–(1.2):

(a) *there exists a sequence* $\{\hat{\lambda}_n\}$ *of solutions satisfying the likelihood equation* $\partial Li_n(\lambda)/\partial\lambda = 0$ *such that*

$$\frac{1}{\sqrt{n}}(Q'G_n')^{-1}\left(\hat{\lambda}_n - \lambda_0\right) = o_p(1);$$

(b) *for such a sequence*,

$$(Q'G_n')^{-1}(\hat{\lambda}_n - \lambda_0) \to_{\mathscr{L}} \left(\frac{1}{K}(F^{-1}\xi)', \frac{1}{K}(\tilde{F}^{-1}\tilde{\xi})', \frac{1}{K}(H_1^{-1}\zeta_1)', \ldots, \frac{1}{K}(H_l^{-1}\zeta_l)', N' \right)',$$

where $K, (F, \xi), (\tilde{F}, \tilde{\xi}), (H_k, \zeta_k)$, $k = 1, \ldots, l$, *are defined as in Section* 4; N *is a* $(p^* + q + r + s + 1)$*-dimension normal random vector with mean zero and covariance* $\Sigma^{*-1}\Sigma\Sigma^{*-1}$, $\Sigma^* = \operatorname{diag}(E[\partial^2 l_{zt}(\lambda_0^*)/\partial m^* \partial m^{*\prime}], E[\partial^2 l_{zt}(\lambda_0^*)/\partial\delta\,\partial\delta'])$ *and* $\Sigma = \operatorname{diag}(E[(\partial l_{zt}(\lambda_0^*)/\partial m^*)(\partial l_{zt}(\lambda_0^*)/\partial m^{*\prime})], E[(\partial l_{zt}(\lambda_0^*)/\partial\delta)(\partial l_{zt}(\lambda_0^*)/\partial\delta')])$.

REMARK. If η_t is not normal, the MLE's obtained by Theorem 2.1 are only quasi-maximum likelihood estimators. From Theorem 2.1, we see that the asymptotic distributions of the MLE of various types of unit roots do not depend on the parameters in the moving-average part. As the GARCH innovations reduce to usual white noises with a finite conditional variance (see Theorems 4.1–4.3), the limiting distributions are the same as those given by Chan and Wei (1988) and, in fact, are also the same as Tsay and Tiao's (1990) results in the univariate case. As the GARCH process is present, the limiting distribution will depend on the parameters in the GARCH process and involve a series of independent bivariate Brownian motions with correlated components. These limiting distributions are different from results of Chan and Wei (1988) and Tsay and Tiao (1990). In Example 2.1 below, we will illustrate that the MLE of unit roots is more efficient than the LSE in a special case. In addition, from here to Section 4, all true parameter values λ_0 and λ_0^* are denoted as λ and λ^*, respectively, for simplicity of notation.

EXAMPLE 2.1. Consider the model,

$$y_t = \phi y_{t-1} + \varepsilon_t, \tag{2.6}$$

$$\varepsilon_t | \mathscr{F}_{t-1} \sim N(0, h_t), \quad h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \tag{2.7}$$

where $\phi = 1$, $\alpha_1 \neq 0$ and $E\varepsilon_t^4 < \infty$. Suppose that $\hat{\phi}$ is the MLE of ϕ. Then we can obtain directly from Theorem 2.1,

$$n(\hat{\phi}_{ML} - \phi) \to_{\mathscr{L}} \xi_{ML} = \int_0^1 w_1(t)\, dw_2(t) \Big/ \left[K \int_0^1 w_1^2(t)\, dt \right], \tag{2.8}$$

where (w_1, w_2) is a bivariate Brownian motion with mean zero and covariance $t\begin{pmatrix} \sigma^2 & 1 \\ 1 & K \end{pmatrix}$, $\sigma^2 = Eh_t$ and $K = E(1/h_t) + 2\alpha_1^2 \sum_{i=1}^{\infty} \beta_1^{2(i-1)} E(\varepsilon_{t-i}^2/h_t^2)$.

Denote the LSE of ϕ as $\hat{\phi}_{LS}$, then (see the remark after Theorem 3.3)

$$n(\hat{\phi}_{LS} - 1) \to_{\mathscr{L}} \xi_{LS} = \frac{\int_0^1 B_1(t)\, dB_1(t)}{\int_0^1 B_1^2(t)\, dt}, \tag{2.9}$$

where $B_1(t)$ is defined as in (2.10) below. To compare the efficiency of $\hat{\phi}_{ML}$ with $\hat{\phi}_{LS}$, we normalize the bivariate Brownian motion (w_1, w_2) in (2.8) by letting

$$B_1(t) = \frac{1}{\sqrt{\sigma^2}}\omega_1(t)$$

and

$$B_2(t) = -\frac{1}{\sigma^2}\sqrt{\frac{\sigma^2}{\sigma^2K-1}}\,\omega_1(t) + \sqrt{\frac{\sigma^2}{\sigma^2K-1}}\,\omega_2(t).$$

Then, by Itô's formula [Chung and Williams (1990), page 109],

$$\xi_{ML} = \frac{\int_0^1 B_1(t)\,dB_1(t)}{K\sigma^2\int_0^1 B_1^2(t)\,dt} + \frac{\sqrt{\sigma^2K-1}\int_0^1 B_1(t)\,dB_2(t)}{K\sigma^2\int_0^1 B_1^2(t)\,dt}. \tag{2.10}$$

Since $B_1(t)$ is independent of $B_2(t)$, by the definition of the stochastic integral [cf. Chung and Williams (1990), Chapter 2], it is not difficult to obtain

$$\begin{aligned} &E\left[\frac{\int_0^1 B_1(t)\,dB_1(t)\int_0^1 B_1(t)\,dB_2(t)}{\left(\int_0^1 B_1^2(t)\,dt\right)^2}\right] \\ &\quad = E\left\{\frac{\int_0^1 B_1(t)\,dB_1(t)}{\left(\int_0^1 B_1^2(t)\,dt\right)^2}E\left[\int_0^1 B_1(t)\,dB_2(t)\,|\{B_1(t), 0\le t\le 1\}\right]\right\} = 0. \end{aligned} \tag{2.11}$$

Furthermore,

$$\begin{aligned} &E\left(\frac{\int_0^1 B_1(t)\,dB_2(t)}{\int_0^1 B_1^2(t)\,dt}\right)^2 \\ &\quad = E\left\{\left[\int_0^1 B_1^2(t)\,dt\right]^{-2}E\left[\left(\int_0^1 B_1(t)\,dB_2(t)\right)^2\middle|\{B_1(t), 0\le t\le 1\}\right]\right\} \\ &\quad = E\left[\int_0^1 B_1^2(t)\,dt\right]^{-1}. \end{aligned} \tag{2.12}$$

By (2.10)–(2.12),

$$E\xi_{LM}^2 = \frac{1}{K^2\sigma^4}E\left(\frac{\int_0^1 B_1(t)\,dB_1(t)}{\int_0^1 B_1^2(t)\,dt}\right)^2 + \frac{\sigma^2K-1}{K^2\sigma^4}E\left[\int_0^1 B_1^2(t)\,dt\right]^{-1}. \tag{2.13}$$

Further, by (2.9),

$$\frac{E\xi_{ML}^2}{E\xi_{LS}^2} = \frac{1}{K^2\sigma^4} + \frac{\sigma^2K-1}{K^2\sigma^4}c, \tag{2.14}$$

where $c = E[\int_0^1 B_1^2(t)\,dt]^{-1}/E[\int_0^1 B_1(t)\,dB_1(t)/\int_0^1 B_1^2(t)\,dt]^2$. By simulating the nonstationary AR(1) process with 10,000 replications, we estimate $c =$

0.4489, 0.4307, 0.4247, respectively, for $n = 100, 200, 500$. By the Cauchy–Schwarz inequality, it is easy to show that $K\sigma^2 \geq 1$, and hence,

$$\frac{E\xi_{ML}^2}{E\xi_{LS}^2} = \frac{c\sigma^2 K + 1 - c}{K^2\sigma^4} \leq \frac{K\sigma^2}{K^2\sigma^4} = \frac{1}{K\sigma^2} \leq 1,$$

with equality if and only if $K\sigma^2 = 1$, that is, h_t = a constant. In particular, when $\beta_1 = 0$ and $\alpha_1 \to 1$, $K\sigma^2 \to \infty$, as shown by Engle (1982). This demonstrates that the MLE for the unit root is much more efficient than the LSE and the gain in efficiency could be very large.

REMARK. Based on the asymptotic distribution in (2.8), Ling and Li (1997b) proposed some new unit root tests and presented simulation results showing that these new tests have better performance than Dickey–Fuller tests based solely on LSE. We also believe that the asymptotic theory in Theorem 2.1 can be applied in more general cases. For the simplest GARCH process (2.7), assumption 5 is equal to $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$, that is, Bollerslev's (1986) condition. Under the weaker condition (Nelson, 1990) $E[\ln(\alpha_1\eta_t^2 + \beta_1)] < 0$, which allows $\alpha_1 + \beta_1 = 1$ (in this case, the variance is infinite), Lee and Hansen (1994) derived the asymptotic distribution of the MLE for the pure GARCH process (2.7) with $\beta_1 \neq 0$. For the model (2.6)–(2.7), whether or not corresponding results exist will be an interesting research problem.

EXAMPLE 2.2. Consider the model,

$$\phi(B)y_t = \psi(B)\varepsilon_t, \tag{2.15}$$

where $\varepsilon_t|\mathcal{F}_{t-1} \sim N(0, h_t)$, h_t and $\psi(B)$ are defined as in model (1.1)–(1.2) and $\phi(B) = (1 - B)\phi^*(B)$ with all roots of $\phi^*(B)$ outside the unit circle. Reparametrize (2.15) as

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

where $\gamma_1 = \sum_{i=1}^{p} \phi_i$ and $\gamma_j = -\sum_{i=j}^{p} \phi_i$, $j = 2, \ldots, p$. Suppose that $\hat{\phi}$ is the MLE of the parameter $\phi = (\phi_1, \ldots, \phi_p)'$. Define $\hat{\gamma}_1 = \sum_{i=1}^{p} \hat{\phi}_i$ and $\hat{\gamma}_j = -\sum_{i=j}^{p} \hat{\phi}_i$, $j = 2, \ldots, p$. Then

$$G_n^{-1}(\hat{\gamma} - \gamma) \to_{\mathcal{L}} (c\xi_{ML}, N') \tag{2.17(2.16)}$$

where $G_n = \text{diag}(1/n, I_{(p-1)\times(p-1)}/\sqrt{n})$, $c = 1/(1 - \sum_{i=2}^{p} \gamma_i)$, $\gamma = (\gamma_1, \ldots, \gamma_p)'$, ξ_{ML} is defined as in (2.8) and N is a normal random vector with mean zero and covariance

$$\Sigma = E\big[(1/h_t)(\partial\varepsilon_t/\partial\gamma)(\partial\varepsilon_t/\partial\gamma')\big] + E\big[(1/2)\big(1/h_t^2\big)(\partial h_t/\partial\gamma)(\partial h_t/\partial\gamma')\big].$$

Let $u_t = (1-B)^{-1}\phi(B)y_t$ and $z_t = \phi^{*-1}(B)\phi(B)y_t$. Then $u_t = \phi^*(B)y_t$ and $z_t = (1-B)y_t$. Denote

$$Q = \begin{pmatrix} 1 & -\gamma_2 & -\gamma_3 & \cdots & -\gamma_{p-1} & -\gamma_p \\ 1 & -1 & 0 & \cdots & 0 & 0 \\ & & & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}.$$

Then $(u_t, z_t, \ldots, z_{t-p+2}) = (y_t, \ldots, y_{t-p+1})Q'$. By Theorem 2.1,

$$(Q'G_n')^{-1}\big(\hat{\phi} - \phi\big) \to_{\mathscr{L}} \operatorname{diag}(\xi_{ML}, N). \tag{2.17}$$

On the other hand, we have the relationship, $\gamma = P\phi$ and $\hat{\gamma} = P\hat{\phi}$, where

$$P = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Thus,

$$G_n^{-1}(\hat{\gamma} - \gamma) = G_n^{-1}P\big(\hat{\phi} - \phi\big) = G_n^{-1}PQ'G_n\Big[(Q'G_n')^{-1}\big(\hat{\phi} - \phi\big)\Big]. \tag{2.18}$$

By direct calculation,

$$G_n^{-1}PQ'G_n = G_n^{-1}\begin{pmatrix} c & 0 \\ * & I_{(p-1)\times(p-1)} \end{pmatrix}G_n \to \begin{pmatrix} c & 0 \\ 0 & I_{(p-1)\times(p-1)} \end{pmatrix}, \tag{2.19}$$

where $*$ is composed of elements not depending on n. Further, by (2.17)–(2.19), (2.16) holds.

REMARK. As ε_t reduces to white noise with a constant conditional variance, the asymptotic distribution given by (2.16) is the same as that given by Yap and Reinsel (1995b) and further, as $q = 0$, it is also the result in Fuller (1976) and Dickey and Fuller (1979). This shows that the estimators obtained by the reparametrization method used by Fuller (1976) and Yap and Reinsel (1995b) and the estimator obtained by the componentwise argument method used by Chan and Wei (1988) are asymptotically equivalent.

3. Auxiliary theorems.

THEOREM 3.1. *Let* $\{S_n(t),\ 0 \le t \le 1\}$ *and* $\{\xi_k,\ k = 1, 2, \ldots\}$ *be two sequences of random processes such that*:

(a) $$S_n(t) \to_{\mathscr{L}} S(t) \text{ in } D;$$

(b) $$\frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k \to_{\mathscr{L}} \xi(t) \text{ in } D;$$

(c) $$\max_{1 \le k \le n} \left| \xi_k / \sqrt{n} \right| \to_p 0;$$

(d) $$\frac{1}{n} \sum_{t=1}^{n} |\xi_t| = O_p(1);$$

and almost all trajectories of $S(t)$ and $\xi(t)$ are continuous. Then

$$\sup_{0\le t\le 1}\left|\frac{1}{n}\sum_{k=1}^{[nt]} S_n\left(\frac{k}{n}\right)\xi_k\right| \to_p 0 \quad \text{as } n\to\infty.$$

REMARK. Theorem 3.1 together with Theorem 3.4 below will be an important tool for Lemmas 4.1–4.4.

PROOF. First, by conditions (a)–(b) and Theorem 15.2 in Billingsley [(1968), page 125], there exists a constant M such that, in probability,

$$\sup_{0\le t\le 1}|S_n(t)| < M, \tag{3.1}$$

$$\sup_{1\le j\le n}\left|\frac{1}{\sqrt{n}}\sum_{k=1}^{j}\xi_k\right| < M, \tag{3.2}$$

$$\sup_{|u-v|<1/\sqrt{n}}|S_n(u)-S_n(v)| \to 0, \tag{3.3}$$

$$\sup_{|t_1-t_2|<1/\sqrt{n}}\left|\frac{1}{\sqrt{n}}\sum_{k=[nt_1]}^{[nt_2]}\xi_k\right| \to 0. \tag{3.4}$$

Now, for each n, let $N(n)=[\sqrt{n}]+1$, $m=[n/N(n)]$ and $n_k = kN(n)$, $1\le k\le m$. Then $1=n_0<n_1<n_2<\cdots<n_m\le n$. Denote $s=\sup\{k, n_k<j\}$. Then we have

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{j} S_n\left(\frac{k}{n}\right)\xi_k &= \frac{1}{n}\sum_{l=1}^{s}\sum_{k=n_{l-1}}^{n_l-1} S_n\left(\frac{k}{n}\right)\xi_k + \frac{1}{n}\sum_{k=n_s}^{j} S_n\left(\frac{k}{n}\right)\xi_k \\ &= \frac{1}{n}\sum_{l=1}^{s}\sum_{k=n_{l-1}}^{n_l-1}\left[S_n\left(\frac{k}{n}\right)-S_n\left(\frac{n_{l-1}}{n}\right)\right]\xi_k \\ &\quad + \frac{1}{n}\sum_{l=1}^{s} S_n\left(\frac{n_{l-1}}{n}\right)\sum_{k=n_{l-1}}^{n_l-1}\xi_k + \frac{1}{n}\sum_{k=n_s}^{j} S_n\left(\frac{k}{n}\right)\xi_k \\ &= I_1+I_2+I_3, \quad \text{say.}\end{aligned} \tag{3.5}$$

By condition (d) and (3.3),

$$|I_1| \le \sup_{|u-v|<1/\sqrt{n}}|S_n(u)-S_n(v)|\frac{1}{n}\sum_{k=1}^{n}|\xi_k| \to_p 0. \tag{3.6}$$

By (3.1) and (3.4),

$$\begin{aligned}|I_2| &\le \sup_{0\le t\le 1}|S_n(t)|\max_{1\le l\le s}\left|\frac{1}{\sqrt{n}}\sum_{k=n_{l-1}}^{n_l-1}\xi_k\right|\frac{s}{\sqrt{n}} \\ &\le \sup_{0\le t\le 1}|S_n(t)|\max_{|t_1-t_2|\le 1/\sqrt{n}}\left|\frac{1}{\sqrt{n}}\sum_{k=[nt_1]}^{[nt_2]}\xi_k\right|\frac{s}{\sqrt{n}} \to_p 0.\end{aligned} \tag{3.7}$$

By condition (c) and (3.1),

$$
\begin{aligned}
|I_3| &\le \sup_{0\le t\le 1} |S_n(t)| \frac{1}{n} \sum_{t=n_s}^{j} |\xi_k| \\
&\le \sup_{0\le t\le 1} |S_n(t)| \max_{1\le k\le n} |\xi_k| \frac{j-n_s}{n} \qquad (3.8) \\
&\le \sup_{0\le t\le 1} |S_n(t)| \max_{1\le k\le n} |\xi_k| \frac{2}{\sqrt{n}} \to_p 0.
\end{aligned}
$$

By (3.5)–(3.8), we complete the proof. □

THEOREM 3.2. *Let* $\{\xi_t\colon t = 1, 2, \dots\}$ *be a series of* $\mathscr{F}_t$*-measurable random variables and* $\mathscr{F}_t$ *is the* σ*-field generated by the i.i.d. random variables* $\{\eta_i, i = t, t-1, \dots\}$. *Suppose that the following conditions are satisfied*:

C1. $\sup_t E\xi_t^2 < \infty$ *and* $E\xi_t = 0$, $t = 1, 2, \dots$;

C2. $E|\xi_t - E(\xi_t|G_{t-m}^{t+m})|^2 = O(m^{-2v})$ *for some* $v > 1/2$, *where* $G_{t-m}^{t+m} = \sigma\{\eta_{t+m}, \dots, \eta_{t-m}\}$;

C3. $\{\xi_t\colon t = 1, 2, \dots\}$ *is uniformly integrable*;

C4. $\sigma_n^2 = E(\sum_{k=1}^n \xi_k/\sqrt{n})^2 \to \sigma^2$, *as* $n \to \infty$.

Then

$$
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n\tau]} \xi_t \to_{\mathscr{L}} \sigma W(\tau) \quad \text{in } D \quad \text{as } n \to \infty, \qquad (3.9)
$$

where W *is a standard Brownian motion*.

REMARK. Theorem 3.2 is an extension of Theorem 21.1 in Billingsley (1968), where ξ_t is a fixed measurable function of $\{\eta_t\}$, that is, $\xi_t = f(\eta_t, \eta_{t-1}, \dots)$. Here, the measurable function f can depend on time t. In particular, when $\sigma = 0$, (3.9) still holds.

PROOF. We only need to verify that Assumptions A.1, B.2, A.3, A.4 and A.5 of Theorem 2.11 in Wooldridge and White (1988) are satisfied.

Let $X_{nt} = \xi_t/\sqrt{n}$. By C1, Assumption B.2(i) holds. Taking $d_{nt} = 1/\sqrt{n}$, by C2, Assumption B.2(ii) holds. Assumptions A.1 and B.2(iii)–(iv) are obviously satisfied. Let $c_{nt} = (1/\sqrt{n})\max\{1, (E\xi_t^2)^{1/2}\}$. Then $X_{nt}^2/c_{nt}^2 \le \xi_t^2$ and hence, Assumption A.3 holds by C3. Now, for $\Delta < \infty$ but sufficiently large, by C1,

$$
\sum_{[n\tau]}^{[n(\tau+\epsilon)]} c_{nt}^2 \le \left[\max\left\{1, \left(\sup_t E\xi_t^2\right)^{1/2}\right\}\right]^2 n^{-1}([n(\tau+\epsilon)] - [n\tau]) \le 3\Delta\epsilon.
$$

Hence, for all $0 < \epsilon < 1 - \tau$ and $0 \le \tau < 1$,

$$
\lim_{n\to\infty} \sup\left(\epsilon^{-1} \sum_{[n\tau]}^{[n(\tau+\epsilon)]} c_{nt}^2\right) \le 3\Delta < \infty.
$$

That is, Assumption A.4 holds. By Theorem 2.11 in Wooldridge and White (1988), $\{\sum_{t=1}^{[n\tau]} \xi_t/\sqrt{n}\}$ is tight in D. When $\sigma^2 = 0$, by C4, $E[\sum_{t=1}^{n} \lambda'\xi_t/\sqrt{n}]^2 \to 0$ and hence, we can obtain $E[\sum_{t=1}^{[nt]} \lambda'\xi_t/\sqrt{n}]^2 \to 0$, where $0 \le \tau \le 1$. It is easy to show that all the finite distributions of $\{\sum_{t=1}^{[n\tau]} \xi_t/\sqrt{n}\}$ converge to zero in distribution. Thus when $\sigma^2 = 0$, (3.9) holds. When $\sigma^2 \neq 0$, since $[n\tau]/n \to \tau$, by C4,

$$\frac{1}{n\sigma_n^2} E\left(\sum_{t=1}^{[n\tau]} \xi_t\right)^2 = \frac{[n\tau]}{n\sigma_n^2}\left[\frac{1}{[n\tau]} E\left(\sum_{t=1}^{[n\tau]} \xi_t\right)^2\right] \to \tau,$$

as $n \to \infty$, where $0 \le \tau \le 1$. That is, Assumption A.5 holds. By Theorem 2.11 in Wooldridge and White (1988), (3.9) holds. This completes the proof. □

The remaining part of this section will be devoted to invariance principles of some random variables involved in Section 4. We first introduce three lemmas. The proofs of Lemmas 3.1 and 3.2 can be found in Ling and Li (1996); see also Bai (1993) and Ling and Li (1997a). The proof of Lemma 3.3 can be found in Appendix B.

LEMMA 3.1. (i) *Under Assumption* 2, $\psi^{-1}(z)$ *has power series expansion*,

$$\psi^{-1}(z) = \sum_{k=0}^{\infty} v_\varepsilon(k) z^k, \qquad |z| \le 1, \tag{3.10}$$

and $v_\varepsilon(k) = O(\rho^k)$ *with* $0 < \rho < 1$.

(ii) *Under Assumption* 3, $\alpha(z)\beta^{-1}(z)$ *has power series expansion*,

$$\alpha(z)\beta^{-1}(z) = \sum_{k=1}^{\infty} v_h(k) z^k, \qquad |z| \le 1, \tag{3.11}$$

and $v_h(k) = O(\rho^k)$ *with* $0 < \rho < 1$.

LEMMA 3.2. *Suppose that the process* $\{\varepsilon_t\}$ *is defined by* (1.2) *and Assumption* 3 *holds. Then* ε_t *is strictly stationary and ergodic and* ε_t^2 *has the following causal representations*,

$$\varepsilon_t^2 = \gamma'\xi_t + \sum_{j=1}^{\infty} \gamma' \prod_{i=0}^{j-1} A_{t-i}\xi_{t-j}, \tag{3.12}$$

where $\xi_t = (\alpha_0\eta_t^2, 0, \ldots, 0, \alpha_0, \ldots, 0)'_{(r+s)\times 1}$ *with the first component* $\alpha_0\eta_t^2$ *and the* $(r+1)$*th component* α_0, $\gamma = (1, 0, \ldots, 0)'_{(r+s)\times 1}$ *and* A_t *is defined in Assumption* 5.

LEMMA 3.3. *Let* $G_{t-m}^{t+m} = \sigma\{\eta_{t+m}, \ldots, \eta_{t-m}\}$, *where* $m = 0, 1, \ldots$. *Under Assumptions* 3 *and* 5, *there exists a constant* ρ, $0 < \rho < 1$, *such that*, *for*

$k = 0, 1, \ldots, m$,

(a) $$E\left|\varepsilon_{t-k}^2 - E\left(\varepsilon_{t-k}^2 | G_{t-m}^{t+m}\right)\right|^2 = O(\rho^{m-k});$$

(b) $$E\left|h_{t-k} - E\left(h_{t-k} | G_{t-m}^{t+m}\right)\right|^2 = O(\rho^{m-k});$$

(c) $$E\left|\frac{1}{h_t} - E\left(\frac{1}{h_t} | G_{t-m}^{t+m}\right)\right|^2 = O(\rho^{m});$$

(d) $$E\left|\sqrt{h_{t-k}} - E\left(\sqrt{h_{t-k}} | G_{t-m}^{t+m}\right)\right|^2 = O(\rho^{m-k});$$

(e) $$E\left|\varepsilon_{t-k} - E\left(\varepsilon_{t-k} | G_{t-m}^{t+m}\right)\right|^2 = O(\rho^{m-k});$$

(f) *if* ξ_t *is one of the following random variables*:

$$\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}, \quad \sum_{k=1}^{t-1} (-1)^{t-k} v_h(k)\varepsilon_{t-k}, \quad \sum_{k=1}^{t-1} v_\beta(k)\tilde{\varepsilon}_t,$$

$$\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\sin k\theta, \quad \sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\cos k\theta,$$

$$\psi^{-1}(B)\varepsilon_{t-i}, i = 1, \ldots, p, \quad \phi^{*-1}(B)\varepsilon_{t-i}, j = 1, \ldots, p^*,$$

$$\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\psi^{-1}(B)\varepsilon_{t-k}, \quad \sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\phi^{*-1}(B)\varepsilon_{t-k},$$

then $E|\xi_t - E(\xi_t | G_{t-m}^{t+m})|^2 = O(\rho^m)$, *where* $(1 - \beta_1 z - \cdots - \beta_s z^s)^{-1} = \sum_{k=1}^{\infty} v_\beta(k) z^k$ *and* $\tilde{\varepsilon}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-r}^2, h_{t-1}, \ldots, h_{t-s})'$.

Now denote

$$A_t = \left(\varepsilon_t, \frac{\varepsilon_t}{h_t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\right)',$$

$$B_t = \left((-1)^t\varepsilon_t, (-1)^t\frac{\varepsilon_t}{h_t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\sum_{k=1}^{t-1} (-1)^{t-k} v_h(k)\varepsilon_{t-k}\right)',$$

$$C_{1t}(\theta) = \sqrt{2}\left(\varepsilon_t \sin t\theta, \frac{\varepsilon_t}{h_t}\sin t\theta - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right) \times \sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\sin(t-k)\theta\right)',$$

$$C_{2t}(\theta) = \sqrt{2}\left(\varepsilon_t \cos t\theta, \frac{\varepsilon_t}{h_t}\cos t\theta - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\cos(t-k)\theta\right)',$$

$$Z_t = (\partial l_{zt}(\lambda^*)/\partial m^{*\prime}, \partial l_{zt}(\lambda^*)/\partial\delta')',$$

where l_{zt} is defined as in Section 2.

LEMMA 3.4. *Suppose that* g_t *is one of* A_t, B_t, $C_{1t}(\theta)$, $C_{2t}(\theta)$ *and* Z_t. *Then there exists a constant* ρ, $0 \le \rho < 1$, *such that* $E\|g_t - E(g_t|G_{t-m}^{t+m})\|^2 = O(\rho^m)$.

PROOF. By Lemma 3.1(ii) and Lemma 3.3, it is not difficult to verify that the conclusion holds. This completes the proof. □

THEOREM 3.3. *Let*

$$S_t = (A_t', B_t', C_{1t}'(\theta_1), C_{2t}'(\theta_1), \ldots, C_{1t}'(\theta_l), C_{2t}'(\theta_l), Z_t')',$$

where $\theta_i \in (0, \pi)$, $\theta_i \ne \theta_j$ *if* $i \ne j$, $i, j = 1, \ldots, l$. *Then*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{[n\tau]} S_t \to_{\mathscr{L}} W(\tau) \quad in \quad D^{4(l+1)+p^*+q+r+s+1},$$

where $W(\tau) = (W_1'(\tau), W_2'(\tau), \ldots, W_{2l+1}'(\tau), W_{2l+2}'(\tau), N'(\tau))'$, *the* $W_i(\tau)$*'s are sequences of i.i.d. bivariate Brownian motions with mean zero and covariance*

$$\tau\Omega = \tau \begin{pmatrix} Eh_t & 1 \\ 1 & E(1/h_t) + \kappa \sum_{k=1}^{\infty} v_h^2(k) E\left(\varepsilon_{t-k}^2/h_t^2\right) \end{pmatrix},$$

$\kappa = E\eta_t^4 - 1$, $N(\tau)$ *is a* $(p^* + q + r + s + 1)$*-dimension Brownian motion, which is independent of* $W_i(\tau)$, $i = 1, \ldots, 2l + 2$, *and has mean zero and covariance* $\tau\Sigma$, *where* Σ *is defined in Theorem* 2.1.

REMARK. When h_t is a constant, $W(\tau)$ is a singular multidimensional Brownian motion but its components are still usual Brownian motions. This theorem serves a similar purpose as Theorem 2.2 in Chan and Wei (1988). The elements of S_t will be basic processes corresponding to nonstationary componentwise arguments in Section 4 and the stationary componentwise argument. The theorem actually is a special extension of Theorem 2.2 in Chan and Wei (1998). When their assumption that $E(\varepsilon_t^2|\mathscr{F}_t)$ = a constant is replaced by the assumption that ε_t is a GARCH process defined by (1.2), by Theorem 2.2 in Kurtz and Protter (1991) and Theorem 3.3, we can show that, under Assumptions 1–5, Chan and Wei's (1988) results still hold. Similarly, in this case, Jeganathan's (1991) results also hold.

PROOF. Let λ be a $(4l + p^* + q + r + s + 5)$-dimension constant vector with $\lambda'\lambda \ne 0$, $\xi_t = \lambda' S_t$ and $\Omega^* = \text{diag}(I \otimes \Omega, \Sigma)$, where I is a $(l + 1) \times (l + 1)$ identity matrix. In the following, we verify that ξ_t satisfies conditions C1–C4 in Theorem 3.2.

(a) By Assumption 5, the fourth moment of ε_t is finite and further, by Lemma 3.1(ii), it is easy to verify that $\sup_t E\xi_t^2 < \infty$ and $E\xi_t = 0$, $t = 1, 2, \ldots$.

(b) By Lemma 3.1(ii) and Lemma 3.4, we can show that

$$E\left|\xi_t - E\left(\xi_t|G_{t-m}^{t+m}\right)\right|^2 = O(m^{-2v})$$

for some $v > 1/2$.

(c) By direct verification, condition C3 is satisfied.

(d) Note that, for $\theta, \theta^* \in [0, 2\pi]$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \cos k\theta \sin k\theta^* = 0$$

and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \cos k\theta \cos k\theta^* = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sin k\theta \sin k\theta^* = 0 \quad \text{if } \theta \neq \theta^*.$$

Thus it is not difficult to show that, for any two different vectors, ξ_{1t} and ξ_{2t}, chosen among A_t, B_t, $C_{1t}(\theta_k)$ and $C_{2t}(\theta_k)$, $k = 1, \ldots, l$,

$$\lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{t=1}^{n} \xi_{1t}\right)\left(\sum_{t=1}^{n} \xi_{2t}'\right)\right] = 0. \tag{3.13}$$

Since η_t has a symmetrical distribution, by straightforward calculation, we can also obtain

$$\lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{t=1}^{n} \xi_{1t}\right)\left(\sum_{t=1}^{n} Z_t'\right)\right] = 0. \tag{3.14}$$

Again, since η_t has a symmetrical distribution,

$$\begin{aligned} &\frac{1}{n} E\left[\left(\sum_{t=1}^{n} A_t\right)\left(\sum_{t=1}^{n} A_t'\right)\right] \\ &\quad = \frac{1}{n} \sum_{t=1}^{n} \begin{pmatrix} Eh_t & 1 \\ 1 & E\left(h_t^{-1}\right) + \kappa \sum_{k=1}^{t-1} \left[v_h^2(k) E\left(\varepsilon_{t-k}^2 h_t^{-2}\right)\right] \end{pmatrix}, \end{aligned} \tag{3.15}$$

$$\begin{aligned} &\frac{1}{n} \sum_{t=1}^{n} E\left[\sum_{k=1}^{t-1} \frac{v_h^2(k)\varepsilon_{t-k}^2}{h_t^2}\right] \\ &\quad = \frac{n}{n} E\left[\sum_{k=1}^{\infty} \frac{v_h^2(k)\varepsilon_{t-k}^2}{h_t^2}\right] - \frac{1}{n} \sum_{k=1}^{n} E\left[\sum_{k=t}^{\infty} \frac{v_h^2(k)\varepsilon_{t-k}^2}{h_t^2}\right], \end{aligned} \tag{3.16}$$

$$\frac{1}{n} \sum_{t=1}^{n} E\left[\sum_{k=t}^{\infty} \frac{v_h^2(k)\varepsilon_{t-k}^2}{h_t^2}\right] \leq E\left(\frac{\varepsilon_{t-i}^2}{\alpha_0^2}\right) \frac{1}{n} \sum_{t=1}^{n} O(\rho^{2t}) \to 0, \tag{3.17}$$

by Lemma 3.1(ii). By (3.15)–(3.17), we have

$$\lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{t=1}^{n} A_t\right)\left(\sum_{t=1}^{n} A_t'\right)\right] = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \Omega = \Omega. \tag{3.18}$$

Similarly, we can show that

$$\lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{t=1}^{n} B_t\right)\left(\sum_{t=1}^{n} B_t'\right)\right] = \Omega, \tag{3.19}$$

$$\lim_{n\to\infty} \frac{1}{n} E\left[\left(\sum_{t=1}^{n} C_{it}(\theta_k)\right)\left(\sum_{t=1}^{n} C_{it}'(\theta_k)\right)\right] = \Omega \quad \text{for } i = 1, 2,\ k = 1, \ldots, l. \tag{3.20}$$

By Assumptions 1–5, the process $\{Z_t\}$ is strictly stationary and ergodic with finite fourth moments [cf. Ling and Li (1997a), Theorems 3.1–3.2 and Weiss (1986), Theorems 3.2–3.3]. Thus,

$$\lim_{n\to\infty}\frac{1}{n}E\left[\left(\sum_{t=1}^{n}Z_t\right)\left(\sum_{t=1}^{n}Z_t'\right)\right]=\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}E(Z_tZ_t')=\Sigma. \tag{3.21}$$

From (3.13)–(3.14) and (3.18)–(3.21), we know that

$$\sigma_n^2=\frac{1}{n}E\left(\sum_{t=1}^{n}\xi_t\right)^2\to\sigma^2=\lambda'\Omega^*\lambda. \tag{3.22}$$

Combining (a)–(d), we have already shown that conditions C1–C4 in Theorem 3.2 are satisfied and hence,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}\lambda'S_t\to_{\mathscr{L}}\lambda'W(\tau)\quad\text{in } D, \tag{3.23}$$

where W is a $(4l+p^*+q+r+s+5)$-dimension Brownian motion with mean zero and covariance Ω^*. Finally, by Proposition 4.1 of Wooldridge and White (1988), we complete the proof. □

THEOREM 3.4. *Suppose that g_t is one of the following types of random variables*:

$$\frac{\varepsilon_t^2}{h_t^3}\left[\sum_{k=1}^{t-1}v_h(k)\varepsilon_{t-k}\right]^2,\quad\frac{\varepsilon_t^2}{h_t^3}\left[\sum_{k=1}^{t-1}(-1)^k v_h(k)\varepsilon_{t-k}\right]^2,$$

$$\frac{\varepsilon_t^2}{h_t^3}\left[\sum_{k=1}^{t-1}v_h(k)\varepsilon_{t-k}\sin k\theta\right]^2,\quad\frac{1}{h_t},\quad\frac{\varepsilon_t^2}{h_t^3}\left[\sum_{k=1}^{t-1}v_h(k)\varepsilon_{t-k}\cos k\theta\right]^2,$$

$$\frac{\varepsilon_t^2}{h_t^3}\left[\sum_{k=1}^{t-1}v_h(k)\varepsilon_{t-k}\sin k\theta\right]\left[\sum_{k=1}^{t-1}v_h(k)\varepsilon_{t-k}\cos k\theta\right].$$

Then

(a) $$n^{-1}\sum_{t=1}^{n}|g_t-Eg_t|=O_p(1);$$

(b) $$\max_{1\le t\le n}|g_t-Eg_t|/\sqrt{n}=o_p(1);$$

(c) $$\frac{1}{n}E\left[\sum_{t=1}^{n}(g_t-Eg_t)\right]^2\to\sigma_0^2,$$
$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}(g_t-Eg_t)\to_{\mathscr{L}}\sigma_0\omega_0(\tau)\quad in\ D;$$

(d)
$$\frac{1}{n}E\left[\sum_{t=1}^{n}(g_t - Eg_t)\sin t\theta^*\right]^2 \to \sigma_1^2,$$
$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}(g_t - Eg_t)\sin t\theta^* \to_{\mathscr{D}} \sigma_1\omega_1(\tau) \quad \text{in } D;$$

(e)
$$\frac{1}{n}E\left[\sum_{t=1}^{n}(g_t - Eg_t)\cos t\theta^*\right]^2 \to \sigma_2^2,$$
$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}(g_t - Eg_t)\cos t\theta^* \to_{\mathscr{D}} \sigma_2\omega_2(\tau) \quad \text{in } D,$$

where σ_0, σ_1 *and* σ_2 *are constants,* $\theta^* \neq 0$ *and* ω_i, $i = 0, 1, 2$, *are standard Brownian motions.*

PROOF. We only consider the case with $g_t = (\varepsilon_t^2/h_t^3)[\sum_{k=1}^{t-1} v_h(k)\varepsilon_{t-k}\sin k\theta]^2$; other cases are similar.

(a) Denote $\tilde{g}_t = (\varepsilon_t^2/h_t^3)[\sum_{k=1}^{\infty} v_h(k)\varepsilon_{t-k}]^2$. By Lemma 3.2 and Assumption 5, $\{\tilde{g}_t\}$ is strictly stationary and ergodic with finite variance. Since $|g_t| \le \tilde{g}_t$, by the ergodic theorem, we know that (a) holds.

(b) It is clear that $\max_{1\le t\le n}|g_t - Eg_t|/\sqrt{n} \le \max_{1\le t\le n}\tilde{g}_t/\sqrt{n} + \max_{1\le t\le n} Eg_t/\sqrt{n}$. It is easy to show that the second term converges to zero. Note that $\{\tilde{g}_t\}$ have a common distribution and $E\tilde{g}_t^2 < \infty$. By Chung (1968), page 93], $\max_{1\le t\le n}\tilde{g}_t/\sqrt{n} = o_p(1)$. Thus (b) holds.

(c) Denote $g_t^* = (\varepsilon_t^2/h_t^3)[\sum_{k=1}^{\infty} v_h(k)\varepsilon_{t-k}\sin k\theta]^2$. Then, by Lemma 3.2 and Assumption 5, we can show that $\{g_t^*\}$ is strictly stationary and ergodic with finite variance. Let $\xi_t = g_t - Eg_t$ and $\xi_t^* = g_t^* - Eg_t^*$. In the following, we first prove the fact that, for $k = 0, 1, \ldots,$

$$E(\xi_0^*\xi_k^*) = O(\rho^k). \tag{3.24}$$

Let $m = [k/3]$. Then, for k large enough, $E(\xi_0^*|G_{-m}^m)$ and $E(\xi_k^*|G_{k-m}^{k+m})$ are two independent random variables and hence,

$$E\left[E(\xi_0^*|G_{-m}^m)E\left(\xi_k^*|G_{k-m}^{k+m}\right)\right] = [E\xi_0^*][E\xi_k^*] = 0, \tag{3.25}$$

where G_{k-m}^{k+m} is defined as in Lemma 3.3. On the other hand, by Lemma 3.3, it is easy to obtain that, for $m = 0, 1, \ldots$, $E|\xi_k^* - E(\xi_k^*|G_{k-m}^{k+m})|^2 = O(\rho_1^m)$ with $\rho_1 \in (0, 1)$. Thus, by stationarity of ξ_t^* and the Cauchy–Schwarz inequality, we have

$$\begin{aligned}|E(\xi_0^*\xi_k^*)| &\le \left|E\left[E(\xi_0^*|G_{-m}^m)E\left(\xi_k^*|G_{k-m}^{k+m}\right)\right]\right| \\ &\quad + 2\left\{E\left[E(\xi_0^*|G_{-m}^m)\right]^2\right\}^{1/2}\left\{E\left[\xi_k^* - E\left(\xi_k^*|G_{k-m}^{k+m}\right)\right]^2\right\}^{1/2} \\ &\quad + E\left[\xi_k^* - E\left(\xi_k^*|G_{k-m}^{k+m}\right)\right]^2 \\ &= O(\rho_1^m) = O(\rho^k),\end{aligned}$$

where $\rho \in (0, 1)$ is large enough such that $\rho_1^m \le C\rho^k$ for some constant C. That is, (3.24) holds.

Now, by (3.24), we know $\sum_{k=1}^{n-1} E(\xi_0^*\xi_k^*)$ and $\sum_{k=1}^{n-1} kE(\xi_0^*\xi_k^*)$ converge absolutely. Therefore,

$$\begin{aligned}\frac{1}{n}E\left(\sum_{t=1}^{n}\xi_t^*\right)^2 &= E\xi_0^{*2} + 2\sum_{k=1}^{n-1}\frac{n-k}{n}E(\xi_0^*\xi_k^*)\\ &\to E\xi_0^{*2} + 2\sum_{k=1}^{\infty}E(\xi_0^*\xi_k^*),\end{aligned} \tag{3.26}$$

as $n \to \infty$. Note that

$$\begin{aligned}\frac{1}{n}\left(\sum_{t=1}^{n}\xi_t\right)^2 &= \frac{1}{n}\left(\sum_{t=1}^{n}\xi_t^*\right)^2 + \frac{2}{n}\left[\sum_{t=1}^{n}(\xi_t - \xi_t^*)\right]\left(\sum_{t=1}^{n}\xi_t^*\right)\\ &\quad + \frac{1}{n}\left[\sum_{t=1}^{n}(\xi_t - \xi_t^*)\right]^2.\end{aligned} \tag{3.27}$$

By Lemma 3.1(ii), it is easy to obtain that $E(\xi_t - \xi_t^*)^2 = O(\rho^t)$ with $\rho \in (0,1)$ and hence, by applying Minkowski's inequality, we know that $E[\sum_{t=1}^{n}(\xi_t - \xi_t^*)]^2/n \to 0$. Further, by (3.26), we can show that the second term in (3.27) also converges to zero in probability. Thus,

$$\frac{1}{n}E\left(\sum_{t=1}^{n}\xi_t\right)^2 \to E\xi_0^{*2} + 2\sum_{k=1}^{\infty}E(\xi_0^*\xi_k^*) = \text{a constant } \sigma_0^2, \tag{3.28}$$

as $n \to \infty$. By (3.28), Assumption 5, which implies $E\varepsilon_t^4 < \infty$ [see Ling (1995), Theorem 6.2] and Lemma 3.3, we can verify that conditions C1–C4 in Theorem 3.2 are satisfied. Thus (c) holds.

The proof of (d) can be found in Ling and Li (1996). The proof of (e) is similar to (d) and hence is omitted. This completes the proof. □

4. The asymptotic behaviors of componentwise arguments. As shown in Section 2, the general model (1.1)–(1.2) can be transformed into various componentwise arguments corresponding to the location of their roots. In this section, we will discuss the asymptotic behaviors of these component arguments according to the different locations of unit roots. These results will be used to prove Theorem 2.1. All of the limiting results obtained in this section are jointly convergent by Theorem 2.3 of Chan and Wei (1988) and Theorem 3.3. We will no longer give special statements.

4.1. *Roots equal to* 1 *and* -1. In this section, first we consider the model

$$(1-B)^a u_t = \psi(B)\varepsilon_t, \tag{4.1.1}$$

where $\psi(B)$ and ε_t are defined as in (1.1)–(1.2) and the initial value $\mathbf{u}_0 = (u_0, \ldots, u_{-a+1})' = \mathbf{0}$. Define $\mathbf{u}_t = (u_t, \ldots, u_{t-a+1})'$, $u_t(k) = (1-B)^{a-k}u_t$, $k =$

$0, 1, \ldots, a$ and $U_t = (u_t(a), \ldots, u_t(1))'$. Then

$$u_t(0) = \psi(B)\varepsilon_t, \tag{4.1.2}$$

$$u_t(k+1) = \sum_{i=1}^{t} u_i(k) \quad \text{for } k = 0, 1, \ldots, a-1, \tag{4.1.3}$$

and, as shown by Chan and Wei (1988), there exists a nonsingular $a \times a$ matrix M such that $M\mathbf{u}_t = U_t$. Denote $J_n = N_n^{-1}M$ and $N_n = \text{diag}(n^a, n^{a-1}, \ldots, n)$. Let

$$F_0(t) = B_1(t), \qquad F_j(t) = \int_0^t F_{j-1}(s)\, ds, \tag{4.1.4}$$

$$\xi = \left(\int_0^1 F_{a-1}(t)\, dB_2(t), \ldots, \int_0^1 F_0(t)\, dB_2(t) \right)', \tag{4.1.5}$$

$$F = (\sigma_{ij})_{a\times a} \quad \text{and} \quad \sigma_{ij} = \int_0^1 F_i(t)F_j(t)\, dt, \tag{4.1.6}$$

where $i, j = 0, 1, \ldots, a-1$ and $W_1(t) = (B_1(t), B_2(t))'$ is a bivariate Brownian motion with mean zero and covariance $t\Omega$ defined as in Theorem 3.3. For the process $\{\mathbf{u}_t\}$, we have the following theorem.

THEOREM 4.1.

(a) $$J_n \sum_{t=1}^{n} A_{1t} \to_{\mathscr{L}} \xi;$$

(b) $$\sum_{t=1}^{n} J_n B_{1t} J_n' \to_{\mathscr{L}} KF,$$

where

$$A_{1t} = \frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i)\mathbf{u}_{t-i-1} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \sum_{i=1}^{t-1} \sum_{j=0}^{t-1} v_h(i) v_\varepsilon(j) \varepsilon_{t-i} \mathbf{u}_{t-i-j-1},$$

$$B_{1t} = \frac{1}{h_t} \sum_{i=0}^{t-1} \sum_{j=0}^{t-1} v_\varepsilon(i) v_\varepsilon(j) \mathbf{u}_{t-i-1}\mathbf{u}_{t-j-1}'$$

$$+ \frac{2\varepsilon_t^2}{h_t^3} \sum_{i_1, i_2=1}^{t-1} \sum_{j_1, j_2=0}^{t-1} v_h(i_1) v_h(i_2) v_\varepsilon(j_1) v_\varepsilon(j_2) \tag{4.1.7}$$

$$\times \varepsilon_{t-i_1}\varepsilon_{t-i_2} \mathbf{u}_{t-i_1-j_1-1}\mathbf{u}_{t-i_2-j_2-1}',$$

$$K = E\left(\frac{1}{h_t} \right) + 2 \sum_{i=1}^{\infty} v_h^2(i) E\left(\frac{\varepsilon_{t-i}^2}{h_t^2} \right),$$

and $v_h(i)$ *and* $v_\varepsilon(i)$ *are defined as in Lemma* 3.1.

Before giving the proof of Theorem 4.1, we first present two lemmas.

LEMMA 4.1.

(a) $$\sqrt{n}\,N_n^{-1}U_{[nt]} \to_{\mathscr{L}} \psi(1)\,\xi^*(t);$$

(b) $$N_n^{-1}\sum_{t=1}^{n}\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]U_{t-1} \to_{\mathscr{L}} \psi(1)\,\xi,$$

(c) $$N_n^{-1}\sum_{t=1}^{n}U_{t-1}U_{t-1}'N_n^{-1} \to_{\mathscr{L}} \psi^2(1)F,$$

where $\xi^*(t) = (\int_0^t F_{a-1}(s)\,ds, \ldots, \int_0^t F_0(s)\,ds)'$.

PROOF. For (a),

$$\begin{aligned} u_t(1) &= \sum_{i=1}^{t} u_i(0) = \sum_{i=1}^{t}\psi(B)\,\varepsilon_i \\ &= \psi(1)\sum_{i=1}^{t}\varepsilon_i - \sum_{r=1}^{q}\psi_r\sum_{i=0}^{r-1}\varepsilon_{t-i} \qquad (4.1.8)\\ &= \psi(1)\sum_{i=1}^{t}\varepsilon_i + R_t(1), \end{aligned}$$

where $ER_t^2(1) \le M$ (a constant) uniformly on t. By (4.1.8),

$$\begin{aligned} u_t(k) &= \sum_{i_1}^{t}\cdots\sum_{i_{k-1}=1}^{i_{k-2}} R_t(1) + \psi(1)u_t^*(k) \qquad (4.1.9)\\ &= R_t(k) + \psi(1)u_t^*(k), \qquad k = 1, \ldots, a, \end{aligned}$$

where $ER_t^2(k) \le Mt^{2(k-1)}$, $(1-B)^k u_t^*(k) = \varepsilon_t$ and thus,

$$u_t^*(k+1) = \sum_{i=1}^{t} u_i^*(k), \qquad k = 0, 1, \ldots, a-1. \qquad (4.1.10)$$

Using Theorem 3.3, we have

$$\frac{1}{\sqrt{n}}u_{[nt]}^*(1) = \frac{1}{\sqrt{n}}\sum_{i=1}^{[nt]}\varepsilon_i \to_{\mathscr{L}} B_1(t) \quad \text{in } D. \qquad (4.1.11)$$

By (4.1.10)–(4.1.11) and repeatedly applying the continuous mapping theorem [Billingsley (1968), Theorem 5.1],

$$n^{-k+1/2}u_{[nt]}^*(k) \to_{\mathscr{L}} \int_0^t F_{k-1}(s)\,ds \quad \text{in } D, \qquad k = 1, \ldots, a. \qquad (4.1.12)$$

Further, by (4.1.9) and (4.1.12), we can claim that (a) holds.

For (b), by (4.1.9), the kth element of (b) can be written as

$$
\begin{aligned}
&n^{-k}\sum_{t=1}^{n}\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]u_{t-1}(k)\\
&\quad=n^{-k}\sum_{t=1}^{n}\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]R_{t-1}(k)\\
&\qquad+\psi(1)n^{-k}\sum_{t=1}^{n}\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]u_{t-1}^{*}(k)\\
&\quad=I_1+\psi(1)I_2,\quad\text{say.}
\end{aligned}
\tag{4.1.13}
$$

Note that, by (1.2) and (3.11), $\varepsilon_{t-i}^2/h_t<1/v_h(i)$ a.s. and further, by $1/\sqrt{h_t}\le 1/\sqrt{\alpha_0}$ a.s., we know that, almost surely,

$$
\begin{aligned}
&\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]^2\\
&\qquad\le c\left[|\eta_t|+\left(\eta_t^2+1\right)\sum_{i=1}^{t-1}|v_h(i)|\sqrt{\frac{\varepsilon_{t-i}^2}{h_t}}\right]^2\\
&\qquad\le c\left(|\eta_t|+\eta_t^2+1\right)^2,
\end{aligned}
$$

where c is a constant not depending on t. Since $R_{t-1}(k)$ is $\mathscr{F}_{t-1}$-measurable, we have

$$
\begin{aligned}
EI_1^2&=n^{-2k}\sum_{t=1}^{n}E\left\{\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]^2R_{t-1}^2(k)\right\}\\
&\le c_1n^{-2k}\sum_{i=1}^{n}ER_{t-1}^2(k)=n^{-2k}O(n^{2k-1})=o(1),
\end{aligned}
\tag{4.1.14}
$$

where $c_1=cE(|\eta_t|+\eta_t^2+1)^2$.

Let

$$
Y_{nt}=\left(1/\sqrt{n}\right)\left[\varepsilon_t/h_t-(1/h_t)\left(\varepsilon_t^2/h_t-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right].
$$

Then $\{Y_{nt}\}$ is a $\mathscr{F}_t$-measurable Martingale difference. It is easy to verify that $\sup_n\sum_{t=1}^{n}|Y_{nt}|^2<\infty$. Further, by (4.1.12), Theorem 2.2 in Kurtz and Protter (1991) and Theorem 3.3,

$$
\begin{aligned}
I_2&=\sum_{t=1}^{n}\left[n^{-k+1/2}u_{t-1}^{*}(k)\right]\left\{\frac{1}{\sqrt{n}}\left[\frac{\varepsilon_t}{h_t}-\frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t}-1\right)\sum_{i=1}^{t-1}v_h(i)\,\varepsilon_{t-i}\right]\right\}\\
&\to_{\mathscr{L}}\int_0^1F_k(t)\,dB_2(t)\quad\text{for } k=0,\ldots,a-1.
\end{aligned}
\tag{4.1.15}
$$

By (4.1.13)–(4.1.15), we can show that (b) holds.

For (c), by (4.1.9), Lemma 4.2(a) below and Fuller's (1976) Lemma 5.1.4, the (k, j)th element of (c) can be written as

$$
\begin{aligned}
n^{-k-j}\sum_{t=1}^{n} u_{t-1}(k)u_{t-1}(j) &= n^{-k-j}\sum_{t=1}^{n} O_p(t^{k+j-3/2}) \\
&\quad + \psi^2(1)n^{-k-j}\sum_{t=1}^{n} u^*_{t-1}(k)u^*_{t-1}(j).
\end{aligned} \tag{4.1.16}
$$

In the last equation, the first term converges to zero and, by (4.1.12) and the continuous mapping theorem, the second term converges in distribution to $\psi^2(1)\sigma_{kj}$, $k, j = 1, \ldots, a$. Thus (c) holds. This completes the proof. □

LEMMA 4.2.

(a) $$E\big(u_t^2(k)\big) = O(t^{2(k-1)+1}), \qquad k = 1, \ldots, a;$$

(b) $$N_n^{-1}\sum_{t=1}^{n}\left[\frac{1}{h_t} - E\left(\frac{1}{h_t}\right)\right]U_{t-1}U'_{t-1}N_n^{-1} = o_p(1);$$

(c) $$N_n^{-1}\sum_{t=1}^{n}\left\{\left[\frac{\varepsilon_t^2}{h_t^3}\left(\sum_{i=1}^{t-1} v_h(i)\,\varepsilon_{t-i}\right)^2\right] - E\left[\frac{\varepsilon_t^2}{h_t^3}\left(\sum_{i=1}^{t-1} v_h(i)\,\varepsilon_{t-i}\right)^2\right]\right\} \times U_{t-1}U'_{t-1}N_n^{-1} = o_p(1).$$

PROOF. Similar to Lemma 3.3.5 of Chan and Wei (1988), (a) can be established, and the detail is in Ling and Li (1996). (b) and (c) hold by Theorem 3.1, Theorem 3.4 and Lemma 4.1(b)–(c). This completes the proof. □

PROOF OF THEOREM 4.1. For (a), consider the kth element of $J_n\sum_{t=1}^{n} A_{1t}$,

$$
\begin{aligned}
&n^{-k}\sum_{t=1}^{n}\left[\frac{\varepsilon_t}{h_t}\sum_{i=0}^{t-1} v_\varepsilon(i)u_{t-i-1}(k)\right. \\
&\qquad \left. - \frac{1}{h_t}\left(\frac{\varepsilon_t}{h_t} - 1\right)\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i)v_\varepsilon(j)\varepsilon_{t-i}u_{t-i-j-1}(k)\right] \\
&\quad = I_1 - I_2.
\end{aligned} \tag{4.1.17}
$$

Note that

$$
\begin{aligned}
u_{t-i-1}(k) &= \sum_{j=1}^{t-i-1} u_j(k-1) \\
&= \sum_{j=1}^{t-1} u_j(k-1) - \sum_{j=t-i}^{t-1} u_j(k-1) \\
&= u_{t-1}(k) - \sum_{j=1}^{i} u_{t-j}(k-1), \qquad k = 1, \ldots, a,
\end{aligned} \tag{4.1.18}
$$

where the last term is defined as zero if $i = 0$. Now,

$$
\begin{aligned}
I_1 &= n^{-k} \sum_{t=1}^{n} \frac{\varepsilon_t}{h_t} \left[\sum_{i=0}^{t-1} v_\varepsilon(i) u_{t-i-1}(k) \right] \\
&= n^{-k} \sum_{t=1}^{n} \left[\frac{\varepsilon_t}{h_t} \left(\sum_{i=0}^{t-1} v_\varepsilon(i) \right) u_{t-1}(k) \right] \\
&\quad - n^{-k} \sum_{t=1}^{n} \left[\frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i) \sum_{j=1}^{i} u_{t-j}(k-1) \right] \\
&= n^{-k} \psi^{-1}(1) \sum_{t=1}^{n} \frac{\varepsilon_t}{h_t} u_{t-1}(k) \\
&\quad - n^{-k} \sum_{t=1}^{n} \frac{\varepsilon_t}{h_t} \left(\sum_{i=t}^{\infty} v_\varepsilon(i) \right) u_{t-1}(k) \\
&\quad - n^{-k} \sum_{t=1}^{n} \left[\frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i) \sum_{j=1}^{i} u_{t-j}(k-1) \right].
\end{aligned} \tag{4.1.19}
$$

By Lemma 3.1(i), Lemma 4.2(a) and the Cauchy–Schwarz inequality, the expectation of the absolute value of the second term above is less than

$$
n^{-k} \sum_{t=1}^{n} \sqrt{E\left(\frac{\varepsilon_t}{h_t} \right)^2} \rho^t O(n^{k-1/2}) = o(1), \tag{4.1.20}
$$

and the expectation of the square of the last term is given by

$$
\begin{aligned}
& n^{-2k} \sum_{t=1}^{n} E\left[\frac{1}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i) \sum_{j=1}^{i} u_{t-j}(k-1) \right]^2 \\
& \le n^{-2k} \sum_{t=1}^{n} \left[\sum_{i=0}^{t-1} |v_\varepsilon(i)| \sum_{j=1}^{i} \sqrt{E u_{t-j}^2(k-1)} \right]^2 \\
& = \begin{cases} O(n^{-1}), & \text{if } k = 1 \\ n^{-2k} \sum_{t=1}^{n} O\left(\left(\sum_{i=0}^{t-1} i \rho^i n^{k-3/2} \right)^2 \right) = O(n^{-2}), & \text{if } k > 1 \end{cases} \\
& = o(1).
\end{aligned} \tag{4.1.21}
$$

By (4.1.19)–(4.1.21),

$$
I_1 = \psi^{-1}(1) n^{-k} \sum_{t=1}^{n} \frac{\varepsilon_t}{h_t} u_{t-1}(k) + o_p(1). \tag{4.1.22}
$$

Similar to (4.1.22), by Lemma 3.1, we can show that

$$I_2 = \psi^{-1}(1) n^{-k} \sum_{t=1}^{n} \frac{1}{h_t}\left(\frac{\varepsilon_t}{h_t} - 1\right)\left(\sum_{i=1}^{t-1} v_h(i)\,\varepsilon_{t-1}\right) u_{t-1}(k) + o_p(1). \tag{4.1.23}$$

By (4.1.17), (4.1.22)–(4.1.23) and Lemma 4.1(b), we complete the proof of (a).

For (b), consider the (k, j)th element of $\sum_{t=1}^{n} J_n B_{1t} J_n'$,

$$\begin{aligned} & n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t} \sum_{i_1=0}^{t-1} \sum_{i_2=0}^{t-1} v_\varepsilon(i_1) v_\varepsilon(i_2) u_{t-i_1-1}(k) u_{t-i_2-1}(j) \\ & + 2n^{-k-j} \sum_{t=1}^{n} \frac{\varepsilon_t^2}{h_t^3}\Bigg[\sum_{i_1,\, i_2=1}^{t-1} \sum_{j_1,\, j_2=0}^{t-1} v_h(i) v_h(i_2) v_\varepsilon(j_1) v_\varepsilon(j_2) \\ & \qquad \times \varepsilon_{t-i_1}\varepsilon_{t-i_2} u_{t-i_1-j_1-1}(k) u_{t-i_2-j_2-1}(j)\Bigg] \\ & = I_1 + I_2, \end{aligned} \tag{4.1.24}$$

say. By (4.1.18),

$$\begin{aligned} I_1 &= n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t} \sum_{i_1=0}^{t-1} \sum_{i_2=0}^{t-1} v_\varepsilon(i_1) v_\varepsilon(i_2) u_{t-i_1-1}(k) u_{t-i_2-1}(j) \\ &= n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t}\left[\sum_{i_1=0}^{t-1} \sum_{i_2=0}^{t-1} v_\varepsilon(i_1) v_\varepsilon(i_2)\right] u_{t-1}(k) u_{t-1}(j) \\ &\quad - n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t}\left[\sum_{i_1=0}^{t-1} v_\varepsilon(i_1) \sum_{i_2=1}^{i_1} u_{t-i_2}(k-1)\right] \\ &\qquad \times \left[\sum_{i_2=0}^{t-1} v_\varepsilon(i_2) u_{t-i_2-1}(j)\right] \\ &\quad - n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t}\left[\sum_{i_1=0}^{t-1} v_\varepsilon(i_1) u_{t-1}(k)\right] \\ &\qquad \times \left[\sum_{j_1=0}^{t-1} v_\varepsilon(j_1) \sum_{i_2=1}^{j_1} u_{t-i_2}(j-1)\right]. \end{aligned} \tag{4.1.25}$$

Denote the last summation above by $T_1 + T_2 + T_3$. Then

$$\begin{aligned} E|T_2| \le n^{-k-j} \sum_{t=1}^{n} \Bigg(& E\left[\sum_{i_1=0}^{t-1} v_\varepsilon(i_1) \sum_{j_1=1}^{i_1} u_{t-j_1}(k-1)\right]^2 \\ & \times E\left[\sum_{i_2=0}^{t-1} v_\varepsilon(i_2) u_{t-i_2-1}(j)\right]^2\Bigg)^{1/2} \end{aligned} \tag{4.1.26}$$

$$\leq \begin{cases} n^{-1-j}\sum_{t=1}^{n} O(t^{j-1/2}), & \text{if } k=1, \\ n^{-k-j}\sum_{t=1}^{n}\left[O(t^{2(k-2)+1})O(t^{2(j-1)+1})\right]^{1/2}, & \text{if } k>1, \end{cases}$$

$$= \left\{ \begin{cases} O(n^{-1/2}), & \text{if } k=1, \\ n^{-k-j}\sum_{t=1}^{n} O(t^{k+j-2}) = O(n^{-1}), & \text{if } k>1 \end{cases} \right\} = o(1).$$

Similarly, we can show that $E|T_3| = o(1)$. Thus, we have

$$I_1 = n^{-k-j}\psi^{-2}(1)\sum_{t=1}^{n}\frac{1}{h_t}u_{t-1}(k)u_{t-1}(j) - 2n^{-k-j}\psi^{-1}(1)\sum_{t=1}^{n}\frac{1}{h_t}\left(\sum_{j=t}^{\infty} v_\varepsilon(i)\right)$$
$$\times u_{t-1}(k)u_{t-1}(j) + n^{-k-j}\sum_{t=1}^{n}\frac{1}{h_t}\left(\sum_{i=t}^{\infty} v_\varepsilon(i)\right)^2 u_{t-1}(k)u_{t-1}(j) + o_p(1).$$

By Lemma 3.1(i) and Lemma 4.2(a), it is easy to show that the second and third terms converge to zero in probability. Thus,

$$\begin{aligned} I_1 &= n^{-k-j}\psi^{-2}(1)E\left(\frac{1}{h_t}\right)\sum_{t=1}^{n} u_{t-1}(k)u_{t-1}(j) \\ &\quad + n^{-k-j}\psi^{-2}(1)\sum_{t=1}^{n}\left[\frac{1}{h_t} - E\left(\frac{1}{h_t}\right)\right]u_{t-1}(k)u_{t-1}(j) + o_p(1) \\ &= n^{-k-j}\psi^{-2}(1)E\left(\frac{1}{h_t}\right)\sum_{t=1}^{n} u_{t-1}(k)u_{t-1}(j) + o_p(1), \end{aligned} \tag{4.1.27}$$

where $k = 1, \ldots, a$ and the last equation holds since the second term in the first equation converges to zero by Lemma 4.2(b). Further, we have

$$I_2 = 2n^{-k-j}\psi^{-2}(1)\sum_{i=1}^{\infty} v_h^2(i)E\left(\frac{\varepsilon_{t-i}^2}{h_t^2}\right)\sum_{t=1}^{n} u_{t-1}(k)u_{t-1}(j) + o_p(1). \tag{4.1.28}$$

The proof of (4.1.28) can be found in Ling and Li (1996). By (4.1.24), (4.1.27), (4.1.28) and Lemma 4.1(c), we complete the proof. □

For the case with unit root -1, we consider the model

$$(1+B)^b v_t = \psi(B)\varepsilon_t, \tag{4.1.29}$$

where $\psi(B)$ and ε_t are defined as in (1.1)–(1.2) and the initial value $\mathbf{v}_0 = (v_0, \ldots, v_{-b+1})' = \mathbf{0}$. Similar to the process $\{\mathbf{u}_t\}$, we define $\mathbf{v}_t = (v_t, \ldots, v_{t-b+1})'$, $v_t(k) = (1+B)^{b-k}v_t$, $k = 0, 1, \ldots, b$, and $V_t = (v_t(b), \ldots, v_t(1))'$. Then $v_t(0) = \psi(B)\varepsilon_t$, $(-1)^t v_t(k+1) = \sum_{i=1}^{t}(-1)^i v_i(k)$, for $k = 0, 1, \ldots, b-1$, and there exists a nonsingular $b \times b$ matrix $\tilde{M}$ such that $\tilde{M}\mathbf{v}_t = V_t$. Denote

$\tilde{J}_n = \tilde{N}_n^{-1}\tilde{M}$ and $\tilde{N}_n = \operatorname{diag}(n^b, n^{b-1}, \ldots, n)$. Let

$$\tilde{\xi} = -\left(\int_0^1 \tilde{F}_{b-1}(t)\, dB_2(t), \ldots, \int_0^1 \tilde{F}_0(t)\, dB_2(t)\right)', \tag{4.1.30}$$

and $\tilde{F} = (\tilde{\sigma}_{ij})_{b\times b}$, where $\tilde{F}_j(t)$ and $\tilde{\sigma}_{ij}$, $i, j = 0, 1, \ldots, b-1$, are defined similarly as (4.1.4) and (4.1.6), respectively. For the process $\{\mathbf{v}_t\}$, we have the following theorem.

THEOREM 4.2.

(a) $$\tilde{J}_n \sum_{t=1}^{n} A_{2t} \to_{\mathscr{L}} \tilde{\xi};$$

(b) $$\sum_{t=1}^{n} \tilde{J}_n B_{2t} \tilde{J}_n' \to_{\mathscr{L}} K\tilde{F},$$

where

$$A_{2t} = \frac{\varepsilon_t}{h_t}\sum_{i=0}^{t-1} v_\varepsilon(i)\mathbf{v}_{t-i-1} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i)v_\varepsilon(j)\varepsilon_{t-i}\mathbf{v}_{t-i-j-1},$$

$$B_{2t} = \frac{1}{h_t}\sum_{i=0}^{t-1}\sum_{j=0}^{t-1} v_\varepsilon(i)v_\varepsilon(j)\mathbf{v}_{t-i-1}\mathbf{v}_{t-j-1}' + \frac{2\varepsilon_t^2}{h_t^3}\sum_{i_1,i_2=1}^{t-1}\sum_{j_1,j_2=0}^{t-1} v_h(i_1)v_h(i_2)v_\varepsilon(j_1)v_\varepsilon(j_2) \times \varepsilon_{t-i_1}\varepsilon_{t-i_2}\mathbf{v}_{t-i_1-j_1-1}\mathbf{v}_{t-i_2-j_2-1}',$$

K is defined as in Theorem 4.1 *and $v_h(i)$ and $v_\varepsilon(i)$ are defined in Lemma* 3.1.

The proof of Theorem 4.2 is similar to that of Theorem 4.1 and can be found in Ling and Li (1996).

4.2. *Roots equal to $e^{i\theta}$ and $e^{i\theta}$.* In this section, we consider the model

$$(1 - 2\cos\theta B + B^2)^d x_t = \psi(B)\varepsilon_t, \tag{4.2.1}$$

where $\psi(B)$ and ε_t are defined in (1.1)–(1.2) and the initial value $\mathbf{x}_0 = (x_0, \ldots, x_{-2d+1}) = 0$. Define $x_t(j) = (1 - 2\cos\theta B + B^2)^{d-j}x_t$ for $j = 0, 1, \ldots, d$. $\mathbf{x}_t = (x_t, \ldots, x_{t-2d+1})'$ and $X_t = (x_t(1), x_{t-1}(1), \ldots, x_t(d), x_{t-1}(d))'$. Then, as in Chan and Wei (1988), there exists a nonsingular $2d \times 2d$ matrix C such that $C\mathbf{x}_t = X_t$. Note that

$$(1 - 2\cos\theta B + B^2)x_t(j+1) = x_t(j) \quad \text{for } j = 0, 1, \ldots, d-1.$$

Since $\mathbf{x}_0 = 0$ implies $X_0 = 0$, we have

$$x_t(j+1) = \frac{1}{\sin\theta}\sum_{k=1}^{t} \sin(t-k+1)\theta x_k(j) \quad \text{for } j = 0, 1, \ldots, d-1. \tag{4.2.2}$$

Let

$$S_t(j) = \sum_{k=1}^{t} \cos k\theta x_k(j) \quad \text{and} \quad T_t(j) = \sum_{k=1}^{t} \sin k\theta x_k(j). \tag{4.2.3}$$

Similar to those of Chan and Wei (1988), we have the following identities:

$$\sin \theta x_t(j) = S_t(j-1)\sin(t+1)\theta - T_t(j-1)\cos(t+1)\theta, \tag{4.2.4}$$

$$\begin{aligned} &2\sin\theta S_t(j) \\ &= \sum_{k=1}^{t} \big[\sin\theta S_k(j-1) - \cos\theta T_k(j-1) \\ &\qquad + \sin(2k+1)\theta S_k(j-1) - \cos(2k+1)\theta T_k(j-1)\big], \end{aligned} \tag{4.2.5}$$

$$\begin{aligned} &2\sin\theta T_t(j) \\ &= \sum_{k=1}^{t} \big[\cos\theta S_k(j-1) + \sin\theta T_k(j-1) \\ &\qquad - \cos(2k+1)\theta S_k(j-1) - \sin(2k+1)\theta T_k(j-1)\big]. \end{aligned} \tag{4.2.6}$$

In the following, we first introduce some notations:

$$\zeta = (\xi_1, \ldots, \xi_{2d})', \qquad H = (\sigma_{ij})_{2d\times 2d},$$

$$\xi_{2j-1} = \frac{1}{2\sin\theta}\left(\int_0^1 f_{j-1}(s)\, dB_2(s) - \int_0^1 g_{j-1}(s)\, d\tilde{B}_2(s)\right),$$

$$\begin{aligned} \xi_{2j} = \frac{1}{2\sin\theta}\Bigg\{&\cos\theta\left[\int_0^1 f_{j-1}(s)\, dB_2(s) - \int_0^1 g_{j-1}(s)\, d\tilde{B}_2(s)\right] \\ &-\sin\theta\left[\int_0^1 f_{j-1}(s)\, d\tilde{B}_2(s) + \int_0^1 g_{j-1}(s)\, dB_2(s)\right]\Bigg\}, \end{aligned}$$

$$\begin{aligned} \sigma_{2k-1,2j-1} &= \sigma_{2k,2j} \\ &= \frac{1}{4\sin^2\theta}\left(\int_0^1 f_{k-1}(s) f_{j-1}(s)\, ds + \int_0^1 g_{k-1}(s) g_{j-1}(s)\, ds\right), \end{aligned}$$

$$\begin{aligned} \sigma_{2k-1,2j} &= \sigma_{2j,2k-1} \\ &= \frac{1}{4\sin^2\theta}\Bigg\{\cos\theta\left[\int_0^1 f_{k-1}(s) f_{j-1}(s)\, ds + \int_0^1 g_{k-1}(s) g_{j-1}(s)\, ds\right] \\ &\qquad -\sin\theta\left[\int_0^1 f_{j-1}(s) g_{k-1}(s)\, ds - \int_0^1 g_{j-1}(s) f_{k-1}(s)\, ds\right]\Bigg\}, \end{aligned}$$

$$f_j(t) = \frac{1}{2\sin\theta}\left(\sin\theta\int_0^t f_{j-1}(s)\, ds - \cos\theta\int_0^t g_{j-1}(s)\, ds\right),$$

$$g_j(t) = \frac{1}{2\sin\theta}\left(\cos\theta\int_0^t f_{j-1}(s)\, ds + \sin\theta\int_0^t g_{j-1}(s)\, ds\right),$$

$$f_0(t) = \tilde{B}_1(t) \quad \text{and} \quad g_0(t) = B_1(t),$$

where $(\tilde{B}_1(t), \tilde{B}_2(t))$ and $(B_1(t), B_2(t))$ are two independent bivariate Brownian motions with mean zero and covariance $t\Omega$ as in Theorem 3.3. Further denote $L_n = N_n^{-1}C$ and $N_n = \operatorname{diag}(nI_{2\times 2}, \ldots, n^d I_{2\times 2})$.

THEOREM 4.3.

(a) $$L_n \sum_{t=1}^{n} A_{3t} \to_{\mathscr{L}} \zeta;$$

(b) $$\sum_{t=1}^{n} L_n B_{3t} L_n' \to_{\mathscr{L}} KH,$$

where

$$A_{3t} = \frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i)\mathbf{x}_{t-i-1} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right) \sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i) v_\varepsilon(j) \varepsilon_{t-i}\mathbf{x}_{t-i-j-1},$$

$$B_{3t} = \frac{1}{h_t} \sum_{i=0}^{t-1}\sum_{j=0}^{t-1} v_\varepsilon(i) v_\varepsilon(j)\mathbf{x}_{t-i-1}\mathbf{x}_{t-j-1}' + \frac{2\varepsilon_t^2}{h_t^3} \sum_{i_1, i_2=1}^{t-1} \sum_{j_1, j_2=0}^{t-1} v_h(i_1) v_h(i_2) v_\varepsilon(j_1) v_\varepsilon(j_2) \times \varepsilon_{t-i_1}\varepsilon_{t-i_2}\mathbf{x}_{t-i_1-j_1-1}\mathbf{x}_{t-i_2-j_2-1}',$$

K is defined as in Theorem 4.1 and $v_h(i)$ and $v_\varepsilon(i)$ are defined as in Lemma 3.1. The following are two auxiliary lemmas.

LEMMA 4.3. *For* $d \geq j \geq 0$,

$$\sqrt{2}\, n^{-j-1/2}\big(S_{[nt]}(j), T_{[nt]}(j)\big)' \to_{\mathscr{L}} \psi(R)\big(f_j(t), g_j(t)\big)' \quad \text{in } D^2, \tag{4.2.7}$$

where $R = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$ *and* $\psi(R)$ *is the matrix polynomial* $\sum_{i=1}^{q} \psi_i R^i$.

PROOF. We prove the lemma by induction. By the definitions of $S_t(0)$ and $T_t(0)$ in (4.2.3),

$$\begin{aligned} S_t(0) &= \sum_{k=1}^{t} \cos k\theta\, \psi(B)\varepsilon_k \\ &= \sum_{i=0}^{q} \psi_i \sum_{k=1}^{t} \cos k\theta \varepsilon_{k-i} \\ &= \sum_{i=0}^{q} \psi_i \sum_{k=1}^{t-i} \cos(k+i)\theta\varepsilon_k + O_p(1) \\ &= \sum_{i=0}^{q} \psi_i \sum_{k=1}^{t} \cos(k+i)\theta\varepsilon_k - \sum_{i=0}^{q} \psi_i \sum_{k=t-i}^{t} \cos(k+i)\theta\varepsilon_k + O_p(1) \end{aligned} \tag{4.2.8}$$

$$= \sum_{i=0}^{q} \psi_i \left(\cos i\theta \sum_{k=1}^{t} \cos k\theta \varepsilon_k - \sin i\theta \sum_{k=1}^{t} \sin k\theta \varepsilon_k \right) + O_p(1)$$

$$= \sum_{i=0}^{q} \psi_i (\cos i\theta S_t^* - \sin i\theta T_t^*) + O_p(1),$$

where $\psi_0 = 1$, $S_t^* = \sum_{k=1}^{t} \cos k\theta \varepsilon_k$ and $T_t^* = \sum_{k=1}^{t} \sin k\theta \varepsilon_k$. Similarly, we have

$$T_t(0) = \sum_{i=0}^{q} \psi_i (\sin i\theta S_t^* + \cos i\theta T_t^*) + O_p(1). \tag{4.2.9}$$

Writing (4.2.8)–(4.2.9) in the vector form,

$$\begin{aligned} \begin{pmatrix} S_t(0) \\ T_t(0) \end{pmatrix} &= \sum_{i=1}^{q} \psi_i \begin{pmatrix} \cos i\theta & -\sin i\theta \\ \sin i\theta & \cos i\theta \end{pmatrix} \begin{pmatrix} S_t^* \\ T_t^* \end{pmatrix} + O_p(1) \\ &= \sum_{i=1}^{q} \psi_i \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}^i \begin{pmatrix} S_t^* \\ T_t^* \end{pmatrix} + O_p(1) \qquad (4.2.10) \\ &= \psi(R) \begin{pmatrix} S_t^* \\ T_t^* \end{pmatrix} + O_p(1). \end{aligned}$$

By Theorem 3.3 and (4.2.10), we know that

$$\sqrt{2}\, n^{-1/2} \begin{pmatrix} S_{[nt]}(0) \\ T_{[nt]}(0) \end{pmatrix} \to_{\mathscr{L}} \psi(R) \begin{pmatrix} \tilde{B}_1(t) \\ B_1(t) \end{pmatrix} \quad \text{in } D^2. \tag{4.2.11}$$

That is, (4.2.7) holds for $j = 0$. Now suppose that (4.2.7) holds for $j - 1$, that is,

$$\sqrt{2}\, n^{-j+1/2} \begin{pmatrix} S_{[nt]}(j-1) \\ T_{[nt]}(j-1) \end{pmatrix} \to_{\mathscr{L}} \psi(R) \begin{pmatrix} f_{j-1}(t) \\ g_{j-1}(t) \end{pmatrix} \quad \text{in } D^2. \tag{4.2.12}$$

By Proposition 8 in Jeganathan (1991), we have that

$$\sup_{0 \le t \le n} \left| n^{-j-1/2} \sum_{k=1}^{t} \sin(2k+1)\theta S_k(j-1) \right| = o_p(1), \tag{4.2.13}$$

$$\sup_{0 \le t \le n} \left| n^{-j-1/2} \sum_{k=1}^{t} \cos(2k+1)\theta S_k(j-1) \right| = o_p(1), \tag{4.2.14}$$

$$\sup_{0 \le t \le n} \left| n^{-j-1/2} \sum_{k=1}^{t} \sin(2k+1)\theta T_k(j-1) \right| = o_p(1), \tag{4.2.15}$$

$$\sup_{0 \le t \le n} \left| n^{-j-1/2} \sum_{k=1}^{t} \cos(2k+1)\theta T_k(j-1) \right| = o_p(1). \tag{4.2.16}$$

By (4.2.5), (4.2.6) and (4.2.13)–(4.2.16),

$$
\begin{aligned}
&\sqrt{2}\,n^{-j-1/2}\begin{pmatrix} S_{[nt]}(j) \\ T_{[nt]}(j) \end{pmatrix} \\
&\quad = \frac{\sqrt{2}\,n^{-j-1/2}}{2\sin\theta}\sum_{k=1}^{[nt]}\begin{pmatrix} \sin\theta & -\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}\begin{pmatrix} S_k(j-1) \\ T_k(j-1) \end{pmatrix} + o_p(1) \\
&\quad = \frac{1}{2\sin\theta}\begin{pmatrix} \sin\theta & -\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}\left[\frac{1}{n}\sum_{k=1}^{[nt]}\sqrt{2}\,n^{-(j-1)-1/2}\begin{pmatrix} S_k(j-1) \\ T_k(j-1) \end{pmatrix}\right] \\
&\qquad + o_p(1) \qquad (4.2.17) \\
&\quad \to_{\mathscr{L}} \frac{1}{2\sin\theta}\begin{pmatrix} \sin\theta & -\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}\psi(R)\begin{pmatrix} f_{j-1}(t) \\ g_{j-1}(t) \end{pmatrix} \quad \text{in } D^2,
\end{aligned}
$$

where the last step holds by (4.2.12) and the continuous mapping theorem. Since $\psi(R)$ can be written as the form $\psi(R) = \left(\begin{smallmatrix} c_1 & -c_2 \\ c_2 & c_1 \end{smallmatrix}\right)$, by straightforward calculation, the above limiting distribution can be written as the distribution of

$$
\psi(R)\left[\frac{1}{2\sin\theta}\begin{pmatrix} \sin\theta & -\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}\right]\begin{pmatrix} f_{j-1}(t) \\ g_{j-1}(t) \end{pmatrix} = \psi(R)\begin{pmatrix} f_j(t) \\ g_j(t) \end{pmatrix}.
$$

Thus (4.2.7) holds for all $j = 0, 1, \ldots, d-1$. This completes the proof. □

LEMMA 4.4. *For $k, j = 1, \ldots, d$,*

(a) $$E\big(S_t^2(k-1)\big) = O(t^{2(k-1)+1}), \qquad E\big(T_t^2(k-1)\big) = O(t^{2(k-1)+1}),$$

(b) $$E\big(x_t^2(k)\big) = O(t^{2(k-1)+1}),$$

(c)
$$
\begin{aligned}
&n^{-k-j}\sum_{t=1}^{n}\left[\frac{1}{h_t} - E\left(\frac{1}{h_t}\right)\right](S_{t-1}(k-1), T_{t-1}(k-1)) \\
&\qquad \times\big[\psi^{-1}(R)\big]' c_{1t}c_{1t}'\big[\psi^{-1}(R)\big]\begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix} = o_p(1),
\end{aligned}
$$

(d)
$$
\begin{aligned}
&n^{-k-j}\sum_{t=1}^{n}(S_{t-1}(k-1), T_{t-1}(k-1))\big[\psi^{-1}(R)\big]' \\
&\qquad \times\left[\sum_{i_1=1}^{t-1}\sum_{i_2=1}^{t-1} v_h(i_1)v_h(i_2)R^{i_1'}c_{1t}c_{1t}'\omega_t(i_1,i_2)R^{i_2}\right] \\
&\qquad \times\psi^{-1}(R)\begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix} = o_p(1),
\end{aligned}
$$

where $c_{1t} = (\sin t\theta, -\cos t\theta)'$ *and*

$$\frac{\omega_t(i_1, i_2) = \varepsilon_t^2 \varepsilon_{t-i_1}\varepsilon_{t-i_2}}{h_t^3 - E\left(\varepsilon_t^2 \varepsilon_{t-i_1}\varepsilon_{t-i_2}/h_t^3\right)}.$$

PROOF. Similar to Lemma 3.3.5 of Chan and Wei (1988), (a) can be obtained and the details are omitted, (b) follows from (a) and 4.2.4. Since $[1/h_t - E(1/h_t)]c_{1t}c_{1t}'$ and $\sum_{t=1}^{n}[\sum_{i_1=1}^{t-1}\sum_{i_2=1}^{t-1} v_h(i_1)v_h(i_2)R^{i_1'}c_{1t}c_{1t}'\omega_t(i_1, i_2)R^{i_2}]$ are composed of the types of random variables as in Theorem 3.4, (c) and (d) are immediately obtained by Lemma 4.3, Theorem 3.4 and Theorem 3.1. This completes the proof. □

PROOF OF THEOREM 4.3. By (4.2.3),

$$\begin{aligned} S_{t-i-1}(j) &= \sum_{k=1}^{t-i-1} \cos k\theta x_k(j) \\ &= \sum_{k=1}^{t-1} \cos k\theta x_k(j) - \sum_{k=t-i}^{t-1} \cos k\theta x_k(j) \\ &= S_{t-1}(j) - \sum_{k=1}^{i} \cos(t-k)\theta x_{t-k}(j). \end{aligned} \tag{4.2.18}$$

Similarly,

$$T_{t-i-1}(j) = T_{t-1}(j) - \sum_{k=1}^{i} \sin(t-k)\theta x_{t-k}(j). \tag{4.2.19}$$

By (4.2.4), (4.2.18) and (4.2.19),

$$\begin{aligned} x_{t-i-1}(j) &= \frac{1}{\sin\theta}\big[\sin(t-i)\theta S_{t-i-1}(j-1) \\ &\qquad -\cos(t-i)\theta T_{t-i-1}(j-1)\big] \\ &= \frac{1}{\sin\theta}\big[\sin(t-i)\theta S_{t-1}(j-1) \\ &\qquad -\cos(t-i)\theta T_{t-1}(j-1)\big] + R_{1t} \\ &= \frac{1}{\sin\theta}(\sin t\theta, -\cos t\theta)R^i\begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix} + R_{1t}, \end{aligned} \tag{4.2.20}$$

where $R_{1t} = -(1/\sin\theta)[\sin(t-i)\theta\sum_{k=1}^{i}\cos(t-k)\theta x_{t-k}(j) - \cos(t-i)\theta\sum_{k=1}^{i}\sin(t-k)\theta x_{t-k}(j)]$ and R is defined in Lemma 4.3.

The $(2k-1)$th element of $\sum_{t=1}^{n} L_n A_{3t}$ is

$$
\begin{aligned}
& n^{-k} \sum_{t=1}^{n} \left[\frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i)\, x_{t-i-1}(k) \right. \\
& \qquad \left. - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \sum_{i=1}^{t-1} \sum_{j=0}^{t-1} v_h(i)\, v_\varepsilon(j)\, x_{t-i-j-1}(k) \right] \\
& \quad = \frac{n^{-k}}{2 \sin\theta} \sum_{t=1}^{n} c'_{1t} \left\{ \frac{\varepsilon_t}{h_t} \sum_{i=0}^{t-1} v_\varepsilon(i) R^i - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \right. \\
& \qquad \left. \times \left[\sum_{i=1}^{t-1} v_h(i)\, \varepsilon_{t-i} R^i \right] \left[\sum_{j=0}^{t-1} v_\varepsilon(j) R^j \right] \right\} \begin{pmatrix} S_{t-1}(k-1) \\ T_{t-1}(k-1) \end{pmatrix} + R_{2n} \qquad (4.2.21) \\
& \quad = \frac{n^{-k}}{2 \sin\theta} \sum_{t=1}^{n} c'_{1t} \left[\frac{\varepsilon_t}{h_t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \left(\sum_{i=1}^{t-1} v_h(i)\, \varepsilon_{t-i} R^i \right) \right] \\
& \qquad \times \psi^{-1}(R) \begin{pmatrix} S_{t-1}(k-1) \\ T_{t-1}(k-1) \end{pmatrix} \\
& \qquad - \frac{n^{-k}}{2 \sin\theta} \sum_{t=1}^{n} c'_{1t} \left[\frac{\varepsilon_t}{h_t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \left(\sum_{i=1}^{t-1} v_h(i)\, \varepsilon_{t-i} R^i \right) \right] \\
& \qquad \cdot \left(\sum_{j=t}^{\infty} v_\varepsilon(j) R^j \right) \begin{pmatrix} S_{t-1}(k-1) \\ T_{t-1}(k-1) \end{pmatrix} + R_{2n},
\end{aligned}
$$

where c_{1t} is defined as in Lemma 4.4. Denote the last summation above by $C_{1n} + C_{2n} + R_{2n}$. By Lemma 3.1(i) and Lemma 4.4(a), similar to (4.1.21), it is easy to show that the term C_{2n} converges to zero in probability. Note that R_{2n} is a function of R_{1t}. Using Lemma 4.4(a) and (b), similar to (4.1.21), we can show that R_{2n} converges to zero in probability. By Lemma 4.3, Theorem 3.3 and the continuous mapping theorem, we have

$$
\begin{aligned}
C_{1n} &= \frac{n^{-k}}{2 \sin\theta} \sum_{t=1}^{n} \left(S_{t-1}(k), T_{t-1}(k) \right) \left[\psi^{-1}(R) \right]' \\
&\qquad \times \left\{ \frac{\varepsilon_t}{h_t} c_{1t} - \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1 \right) \left[\sum_{i=1}^{t-1} v_h(i)\, \varepsilon_{t-i} c_{1,t-i} \right] \right\} \\
&\to_{\mathscr{L}} \xi_{2k-1} \quad \text{in } D.
\end{aligned}
$$

Hence, the $(2k-1)$th element of $L_n \sum_{t=1}^{n} A_{3t} \to_{\mathscr{L}} \xi_{2k-1}$ in D. Similarly, we can show that the $2k$th element of $L_n \sum_{t=1}^{n} A_{3t} \to_{\mathscr{L}} \xi_{2k}$ in D. This completes the proof of (a).

For (b), we consider the $(2k, 2j)$th element of $\sum_{t=1}^{n} L_n B_{3t} L_n'$,

$$
\begin{aligned}
& n^{-k-j} \sum_{t=1}^{n} \frac{1}{h_t} \sum_{i_1=0}^{t-1} \sum_{i_2=0}^{t-1} v_\varepsilon(i_1) v_\varepsilon(i_2)\, x_{t-i_1-1}(k)\, x_{t-i_2-1}(j) \\
& \quad + 2n^{-k-j} \sum_{t=1}^{n} \frac{\varepsilon_t^2}{h_t^3} \Bigg[\sum_{i_1, i_2=1}^{t-1} \sum_{j_1, j_2=0}^{t-1} v_h(i_1) v_h(i_2) v_\varepsilon(j_1) v_\varepsilon(j_2) \\
& \qquad \times \varepsilon_{t-i_1} \varepsilon_{t-i_2} x_{t-i_1-j_1-1}(k)\, x_{t-i_2-j_2-1}(j) \Bigg] \\
& \quad = I_1 + 2I_2,
\end{aligned} \tag{4.2.22}
$$

say. By (4.2.20), Lemma 3.1 and Lemma 4.4, similar to the proofs of (4.1.25)–(4.1.27), I_1 can be written as

$$
\begin{aligned}
I_1 = {} & \frac{n^{-k-j}}{2\sin^2\theta} E\left(\frac{1}{h_t}\right) \sum_{t=1}^{n} \left(S_{t-1}(k-1), T_{t-1}(k-1)\right)\left[\psi^{-1}(R)\right]' \\
& \qquad \times c_{1t} c_{1t}'\left[\psi^{-1}(R)\right] \begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix} \\
& + \frac{n^{-k-j}}{2\sin^2\theta} \sum_{t=1}^{n} \left[\frac{1}{h_t} - E\left(\frac{1}{h_t}\right)\right] \left(S_{t-1}(k-1), T_{t-1}(k-1)\right)\left[\psi^{-1}(R)\right]' \\
& \qquad \times c_{1t} c_{1t}'\left[\psi^{-1}(R)\right] \begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix} + o_p(1).
\end{aligned}
$$

By Lemma 4.4(c), the second term converges to zero in probability. Note that

$$
\begin{aligned}
c_{1t} c_{1t}' &= \begin{pmatrix} \sin t\theta \\ -\cos t\theta \end{pmatrix} (\sin t\theta, \cos t\theta) \\
&= \begin{pmatrix} \sin^2 t\theta & -\sin t\theta \cos t\theta \\ \sin t\theta \cos t\theta & \cos^2 t\theta \end{pmatrix} \\
&= \tfrac{1}{2} \begin{pmatrix} 1 - \cos 2t\theta & -\sin 2t\theta \\ \sin 2t\theta & 1 + \cos 2t\theta \end{pmatrix} \\
&= \tfrac{1}{2} I_{2\times 2} + \tfrac{1}{2} \begin{pmatrix} -\cos 2t\theta & -\sin 2t\theta \\ \sin 2t\theta & \cos 2t\theta \end{pmatrix}.
\end{aligned} \tag{4.2.23}
$$

I_1 can be further written as

$$
\begin{aligned}
I_1 = {} & \frac{n^{-k-j}}{4\sin^2\theta} E\left(\frac{1}{h_t}\right) \sum_{t=1}^{n} \left(S_{t-1}(k-1), T_{t-1}(k-1)\right) \\
& \qquad \times \left[\psi^{-1}(R)\right]' \left[\psi^{-1}(R)\right] \begin{pmatrix} S_{t-1}(j-1) \\ T_{t-1}(j-1) \end{pmatrix}
\end{aligned}
$$

$$+\frac{n^{-k-j}}{4\sin^2\theta}E\left(\frac{1}{h_t}\right)\sum_{t=1}^{n}\left(S_{t-1}(k-1),T_{t-1}(k-1)\right)\left[\psi^{-1}(R)\right]'$$
$$\times\begin{pmatrix}-\cos 2t\theta & -\sin 2t\theta\\ \sin 2t\theta & \cos 2t\theta\end{pmatrix}\psi^{-1}(R)\begin{pmatrix}S_{t-1}(j-1)\\ T_{t-1}(j-1)\end{pmatrix}$$
$$+o_p(1).$$

By Proposition 8 in Jeganathan (1991) and Lemma 4.3, we know that the second term converges to zero in probability. Further, by Lemma 4.3 and the continuous mapping theorem,

$$I_1 \to_{\mathscr{L}} \frac{1}{4\sin^2\theta}E\left(\frac{1}{h_t}\right)\left(\int_0^1 f_{k-1}(s)f_{j-1}(s)\,ds+\int_0^1 g_{k-1}(s)g_{j-1}(s)\,ds\right). \tag{4.2.24}$$

Further, we have

$$I_2 \to_{\mathscr{L}} \frac{1}{4\sin^2\theta}\sum_{i_1=1}^{\infty}v_h^2(i_1)E\left[\frac{\varepsilon_{t-i_1}^2}{h_t^2}\right]$$
$$\times\left[\int_0^1 f_{k-1}(s)f_{j-1}(s)\,ds+\int_0^1 g_{k-1}(s)g_{j-1}(s)\,ds\right]. \tag{4.2.25}$$

The proof of (4.2.25) can be found in Ling and Li (1996). By (4.2.22), (4.2.24) and (4.2.25), the $(2k,2j)$th element of $\sum_{t=1}^{n}L_nB_{3t}L_n'$ converges to $K\sigma_{2k,2j}$, $k,j=1,\ldots,d$. Similarly, we can show that the $(2k-1,2j)$th element of $\sum_{t=1}^{n}L_nB_{3t}L_n'$ converges to $K\sigma_{2k-1,2j}$, $k,j=1,\ldots,d$. This completes the proof. □

5. Proof of the main result. Before giving the proof of Theorem 2.1, we first state two lemmas. First, note that by (2.2) and (2.4)–(2.5), QD_t can be decomposed as

$$QD_t=\left(D_{u,t}',D_{v,t}',D_{x_1,t}',\ldots,D_{x_l,t}',Z_t\right)',$$

where $D_{u,t}$, $D_{v,t}$ and $D_{x_k,t}$, $k=1,\ldots,l$, correspond to nonstationary componentwise arguments $\mathbf{u}_t$, $\mathbf{v}_t$ and $\mathbf{x}_{k,t}$, $k=1,\ldots,l$, respectively. Here $Z_t=(D_{\phi^*,t}',D_{\psi,t}',D_{\delta,t}')'$ corresponds to stationary componentwise arguments, which are defined as in Theorem 3.3. Similarly, QI_tQ' can also be decomposed as $(l+3)\times(l+3)$ block matrices. They are denoted as $I_{uu,t}$, $I_{vv,t}$, $I_{uv,t}$, etc., which are respectively the information blocks in terms of componentwise arguments $\mathbf{u}_t$, $\mathbf{v}_t$, the product of $\mathbf{u}_t$ and $\mathbf{v}_t$, etc. For these subvectors and block-matrices, we have the following lemmas.

LEMMA 5.1.

(a) $$\sum_{t=1}^{n}J_nD_{u,t}\to_{\mathscr{L}}\xi;$$

(b) $$\sum_{t=1}^{n}\tilde{J}_nD_{v,t}\to_{\mathscr{L}}\tilde{\xi};$$

(c) $$\sum_{t=1}^{n} L_{k,n} D_{x_k,t} \to_{\mathscr{L}} \zeta_k, \qquad k = 1, \ldots, l;$$

(d) $$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} Z_t \to_{\mathscr{L}} N(O, \Sigma);$$

(e) $$\sum_{t=1}^{n} J_n I_{uu,t} J_n' \to_{\mathscr{L}} -KF;$$

(f) $$\sum_{t=1}^{n} \tilde{J}_n I_{vv,t} \tilde{J}_n' \to_{\mathscr{L}} -K\tilde{F};$$

(g) $$\sum_{t=1}^{n} L_{k,n} I_{x_k x_k,t} L_{k,n}' \to_{\mathscr{L}} -KH_k, \qquad k = 1, \ldots, l;$$

(h) $$\frac{1}{n} \sum_{t=1}^{n} I_{z,t} \to_{\mathscr{L}} -\Sigma^*,$$

where all notations are defined as in Theorem 2.1.

PROOF. To simplify notation, in this proof, ε_{0t} and h_{0t} are still written as ε_t and h_t. By (A.2)–(A.4) in Appendix A, (2.2) and (2.4)–(2.5), $D_{u,t} = A_{1t}$, $D_{v,t} = A_{2t}$, $D_{x_k,t} = A_{3,k,t}$, $k = 1, \ldots, l$, where A_{1t} and A_{2t} are defined exactly as in Section 4.1 and $A_{3,k,t}$ are the same types of random vectors as A_{3t} defined in Section 4.2. By Theorems 4.1(a)–4.3(a), respectively, we know that (a)–(c) hold.

Since Z_t and $I_{z,t}$ are stationary and ergodic, similar to Ling and Li (1997a), we can show that (d) and (h) hold.

For (e)–(g), by (A.7) in Appendix A, (2.2) and (2.4), we have

$$I_{uu,t} = -B_{1t} + R_{1t}, \qquad I_{vv,t} = -B_{2t} + R_{2t},$$
$$I_{x_k x_k,t} = -B_{3,k,t} + R_{3,k,t}, \qquad k = 1, \ldots, l,$$

where B_{1t} and B_{2t} are defined exactly as in Section 4.1, $B_{3,k,t}$ are the same types of random matrices as B_{3t} defined in Section 4.2,

$$\begin{aligned} R_{1t} = & -\frac{2}{h_t^2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left[\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i) v_\varepsilon(j) \mathbf{u}_{t-i-j-1} \varepsilon_{t-i}\right] \\ & \times \left[\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i) v_\varepsilon(j) \mathbf{u}_{t-i-j-1} \varepsilon_{t-i}\right]' \\ & + \frac{1}{h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left[\sum_{i=1}^{t-1}\sum_{j=0}^{t-1}\sum_{j_1=0}^{t-1} v_h(i) v_\varepsilon(j) v_\varepsilon(j_1) \mathbf{u}_{t-i-j-1} \mathbf{u}_{t-i-j_1-1}'\right] \\ & + \frac{4\varepsilon_t}{h_t^2}\left[\sum_{j=0}^{t-1} v_\varepsilon(j) \mathbf{u}_{t-j-1}\right]\left[\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i) v_\varepsilon(j) \varepsilon_{t-i} \mathbf{u}_{t-i-j-1}\right]', \end{aligned} \tag{5.1}$$

and similarly define R_{2t} and $R_{3,k,t}$.

By Theorems 4.1(b)–4.3(b), it is sufficient for (e)–(g) to hold if

$$\sum_{t=1}^{n} J_n R_{1t} J_n' = o_p(1), \qquad \sum_{t=1}^{n} \tilde{J}_n R_{2t} \tilde{J}_n' = o_p(1), \tag{5.2}$$

$$\sum_{t=1}^{n} L_{k,n} R_{3,k,t} L_{k,n}' = o_p(1), \qquad k = 1, \ldots, l. \tag{5.3}$$

The proofs of (5.2)–(5.3) can be found in Ling and Li (1996). □

LEMMA 5.2.

(a) $$J_n \sum_{t=1}^{n} I_{uv,t} \tilde{J}_n' \to_p 0;$$

(b) $$J_n \sum_{t=1}^{n} I_{ux_k,t} L_{k,n}' \to_p 0, \qquad 1 \le k \le l;$$

(c) $$\tilde{J}_n \sum_{t=1}^{n} I_{vx_k,t} L_{k,n}' \to_p 0, \qquad 1 \le k \le l;$$

(d) $$L_{j,n} \sum_{t=1}^{n} I_{x_j x_k,t} L_{k,n}' \to_p 0, \qquad 1 \le j \ne k \le l;$$

(e) $$J_n \sum_{t=1}^{n} I_{uz,t}/\sqrt{n} \to_p 0;$$

(f) $$\tilde{J}_n \sum_{t=1}^{n} I_{vz,t}/\sqrt{n} \to_p 0;$$

(g) $$L_{kn} \sum_{t=1}^{n} I_{x_k z,t}/\sqrt{n} \to_p 0, \qquad 1 \le k \le l.$$

LEMMA 5.3. *Provided* $\|(1/\sqrt{n})(Q'G_n')^{-1}(\lambda - \lambda_0)\| < 1$,

$$\sum_{t=1}^{n} G_n Q[I_t(\lambda) - I_t] Q' G_n' = O_p\left(\left\| \frac{1}{\sqrt{n}} (Q'G_n')^{-1}(\lambda - \lambda_0) \right\|\right).$$

The proofs of Lemmas 5.2 and 5.3 can be found in Ling and Li (1996).

PROOF OF THEOREM 2.1. Multiply $(\lambda - \lambda_0)'/n$ to (2.3); we have

$$\begin{aligned} &\frac{1}{n}(\lambda - \lambda_0)' \frac{\partial Li_n(\lambda)}{\partial \lambda} \\ &\quad = \left[\frac{1}{\sqrt{n}} (Q'G_n')^{-1}(\lambda - \lambda_0)\right]' \left[\frac{1}{\sqrt{n}} \sum_{t=1}^{n} G_n Q D_t\right] \\ &\qquad + \left[\frac{1}{\sqrt{n}} (Q'G_n')^{-1}(\lambda - \lambda_0)\right]' \left[\sum_{t=1}^{n} G_n Q I_t Q' G_n' + R_n(\lambda)\right] \\ &\qquad \times \left[\frac{1}{\sqrt{n}} (Q'G_n')^{-1}(\lambda - \lambda_0)\right], \end{aligned} \tag{5.4}$$

where $R_n(\lambda) = \sum_{t=1}^{n} G_n Q[I_t(\lambda_n^*) - I_t] Q' G_n'$. Denote the last term by Π.

Let ν and ε be two given and sufficiently small positive numbers. Let $V_n(\varepsilon) = \{\lambda\colon \|(1/\sqrt{n})(Q'G_n')^{-1}(\lambda - \lambda_0)\| = \varepsilon\}$. Note that F, $\tilde{F}$ and H_k, $k = 1, \ldots, l$, are the same as those given by Chan and Wei (1988) and, hence, these information blocks are negative definite in probability. By Lemmas 5.1 and 5.2, there is a constant c_1 and an integer N_1 such that, as $n > N_1$,

$$P\left\{ \sum_{t=1}^{n} G_n Q I_t Q' G_n' < -c_1 I_{m\times m} \right\} > 1 - \nu,$$

where $m = p + q + r + s + 1$. By Lemma 5.3, there exists a constant c such that, for small enough ε, as $n > N_1$ and $\lambda \in V_n(\varepsilon)$,

$$P\left\{ \sum_{t=1}^{n} \left[G_n Q I_t Q' G_n' + R_n(\lambda) \right] < -c I_{m\times m} \right\} > 1 - \nu.$$

Hence, as $n > N_1$ and $\lambda \in V_n(\varepsilon)$,

$$P\{\Pi < -c\varepsilon^2\} > 1 - \nu. \tag{5.5}$$

By Lemma 5.1(a)–(d), we know that $\sum_{t=1}^{n} G_n Q D_t = O_p(1)$. Hence, there exists an integer N_2 such that, as $n > N_2$, $P\{n^{-1/2}\|\sum_{t=1}^{n} G_n Q D_t\| < c\varepsilon/2\} > 1 - \nu$. Thus, as $n > N_2$ and $\lambda \in V_n(\varepsilon)$,

$$P\left\{ \left[\frac{1}{\sqrt{n}} (Q'G_n')^{-1}(\lambda - \lambda_0) \right]' \left[\frac{1}{\sqrt{n}} \sum_{t=1}^{n} G_n Q D_t \right] < \frac{c}{2}\varepsilon^2 \right\} > 1 - \nu. \tag{5.6}$$

Thus by (5.4), (5.5) and (5.6), as $n > \max\{N_1, N_2\}$ and $\lambda \in V_n(\varepsilon)$, with at least probability $1 - \nu$,

$$\frac{1}{n}(\lambda - \lambda_0)' \frac{\partial Li_n(\lambda)}{\partial \lambda} < -c\varepsilon^2 + \frac{c}{2}\varepsilon^2 < 0. \tag{5.7}$$

Let $\Upsilon = (1/\sqrt{n})(Q'G_n')^{-1}(\lambda - \lambda_0)/\varepsilon$ and $g(\lambda) = G_n Q(\partial Li_n(\lambda)/\partial\lambda)$. Then, by (5.7),

$$\|\Upsilon\| = 1 \quad \text{and} \quad \Upsilon' g\left(\sqrt{n}\,\varepsilon Q'G_n'\Upsilon + \lambda_0\right) < 0.$$

Since $\partial Li_n(\lambda)/\partial\lambda$ is continuous and g is also continuous on Υ, by the fixed point theorem [Aitchison and Silvey (1958)], there is a solution $\hat{\lambda}_n$ satisfying $g(\sqrt{n}\,\varepsilon Q'G_n'\Upsilon + \lambda_0) = 0$, that is, $\partial Li_n(\hat{\lambda}_n)/\partial\lambda = 0$ and $\|(1/\sqrt{n})(Q'G_n')^{-1}(\hat{\lambda}_n - \lambda_0)\| < \varepsilon$. Consequently, the proof of part (a) is complete. For such a sequence of $\hat{\lambda}_n$, we have

$$\begin{aligned} &(Q'G_n')^{-1}\left(\hat{\lambda}_n - \lambda_0\right) \\ &\quad = -\left[\sum_{t=1}^{n} G_n Q I_t Q' G_n' \right. \\ &\qquad \left. + O_p\left(\left\| \frac{1}{\sqrt{n}} (Q'G_n')^{-1}\left(\hat{\lambda}_n - \lambda_0\right) \right\| \right) \right]^{-1} \left[\sum_{t=1}^{n} G_n Q D_t \right]. \end{aligned} \tag{5.8}$$

By (a), $(1/\sqrt{n})(Q'G_n')^{-1}(\hat{\lambda}_n - \lambda_0)$ converges to zero in probability. By Lemmas 5.1 and 5.2, Theorem 2.3 of Chan and Wei (1988) and Theorem 3.3 in Section 3, all random variables in $\sum_{t=1}^{n} G_n Q D_t$ and $\sum_{t=1}^{n} G_n Q I_t Q' G_n'$ converge

jointly. Again by Lemmas 5.1 and 5.2 and (5.8), we complete the proof of part (b). □

APPENDIX A

Denote $m = (\phi', \psi')'$, $\varepsilon_t = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \psi_i \varepsilon_{t-i}$ and $\tilde{\varepsilon}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-r}^2, h_{t-1}, \ldots, h_{t-s})'$. The following are some first- and second-partial derivatives of the equations (2.1):

$$\frac{\partial Li_n(\lambda)}{\partial \lambda} = \sum_{t=1}^{n} \frac{\partial l_t(\lambda)}{\partial \lambda}, \qquad \frac{\partial l_t(\lambda)}{\partial \delta} = \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial h_t}{\partial \delta}, \tag{A.1}$$

$$\frac{\partial l_t(\lambda)}{\partial m} = \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial h_t}{\partial m} - \frac{\varepsilon_t}{h_t}\frac{\partial \varepsilon_t}{\partial m}, \tag{A.2}$$

$$\frac{\partial \varepsilon_t}{\partial \phi} = -\psi^{-1}(B)\mathbf{y}_{t-1} = -\sum_{i=0}^{t-1} v_\varepsilon(i)\mathbf{y}_{t-i-1} \quad (\text{if } y_t = 0 \text{ for } t \le 0), \tag{A.3}$$

$$\begin{aligned}\frac{\partial h_t}{\partial \phi} &= 2\sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}\frac{\partial \varepsilon_{t-i}}{\partial \phi} + \sum_{i=1}^{s} \beta_i \frac{\partial h_{t-i}}{\partial \phi} \\ &= 2\sum_{i=1}^{t-1}\sum_{j=0}^{t-1} v_h(i)v_\varepsilon(j)\mathbf{y}_{t-i-j-1}\varepsilon_{t-i} \quad (\text{if } y_t = 0 \text{ for } t \le 0),\end{aligned} \tag{A.4}$$

$$\frac{\partial \varepsilon_t}{\partial \psi_j} = -\psi^{-1}(B)\varepsilon_{t-j}, \qquad j = 1, \ldots, q, \tag{A.5}$$

$$\begin{aligned}\frac{\partial h_t}{\partial \psi} &= 2\sum_{i=1}^{r} \alpha_i \varepsilon_{t-i}\lambda\frac{\partial \varepsilon_{t-i}}{\partial \psi} + \sum_{i=1}^{s} \beta_i \frac{\partial h_{t-i}}{\partial \psi}, \\ \frac{\partial h_t}{\partial \delta} &= \tilde{\varepsilon}_t + \sum_{i=1}^{s} \beta_i \frac{\partial h_{t-i}}{\partial \delta},\end{aligned} \tag{A.6}$$

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \phi\, \partial \phi'} &= -\frac{1}{h_t}\frac{\partial \varepsilon_t}{\partial \phi}\frac{\partial \varepsilon_t}{\partial \phi'} - \frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial \phi}\frac{\partial h_t}{\partial \phi'} \\ &\quad + \left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \phi}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial \phi'}\right] + \frac{2\varepsilon_t}{h_t^2}\frac{\partial \varepsilon_t}{\partial \phi}\frac{\partial h_t}{\partial \phi'} - \frac{\varepsilon_t}{h_t}\frac{\partial^2 \varepsilon_t}{\partial \phi\, \partial \phi'},\end{aligned} \tag{A.7}$$

$$\begin{aligned}\frac{\partial^2 l_t}{\partial \phi\, \partial \psi'} &= -\frac{1}{h_t}\frac{\partial \varepsilon_t}{\partial \phi}\frac{\partial \varepsilon_t}{\partial \psi'} - \frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial \phi}\frac{\partial h_t}{\partial \psi'} \\ &\quad + \left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial \phi}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial \psi'}\right] + \frac{2\varepsilon_t}{h_t^2}\frac{\partial \varepsilon_t}{\partial \phi}\frac{\partial h_t}{\partial \psi'} - \frac{\varepsilon_t}{h_t}\frac{\partial^2 \varepsilon_t}{\partial \phi\, \partial \psi'},\end{aligned} \tag{A.8}$$

$$\frac{\partial^2 l_t}{\partial\phi\,\partial\delta'} = \frac{\varepsilon_t}{h_t^2}\frac{\partial\varepsilon_t}{\partial\phi}\frac{\partial h_t}{\partial\delta'} + \frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial^2 h_t}{\partial\phi\,\partial\delta'} + \frac{1}{2h_t^2}\left(1 - \frac{2\varepsilon_t^2}{h_t}\right)\frac{\partial h_t}{\partial\phi}\frac{\partial h_t}{\partial\delta'}, \tag{A.9}$$

$$\frac{\partial^2 l_t}{\partial\psi\,\partial\psi'} = -\frac{1}{h_t}\frac{\partial\varepsilon_t}{\partial\psi}\frac{\partial\varepsilon_t}{\partial\psi'} - \frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial\psi}\frac{\partial h_t}{\partial\psi'} + \left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial\psi}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial\psi'}\right] + \frac{2\varepsilon_t}{h_t}\frac{\partial\varepsilon_t}{\partial\psi}\frac{\partial h_t}{\partial\psi'} - \frac{\varepsilon_t}{h_t}\frac{\partial^2\varepsilon_t}{\partial\psi\,\partial\psi'}, \tag{A.10}$$

$$\frac{\partial^2 l_t}{\partial\delta\,\partial\delta'} = -\frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial\delta}\frac{\partial h_t}{\partial\delta'} + \left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{\partial}{\partial\delta}\left[\frac{1}{2h_t}\frac{\partial h_t}{\partial\delta'}\right]. \tag{A.11}$$

APPENDIX B

PROOF OF LEMMA 3.3. Let

$$\varepsilon_{t-k,\,m-k}^2 = \gamma'\xi_{t-k} + \sum_{j=1}^{m-k}\gamma'\prod_{i=0}^{j-1}A_{t-k-i}\,\xi_{t-k-j},$$

where γ, ξ_t and A_t are defined in Lemma 3.2. Then $\varepsilon_{t-k,\,m-k}^2$ is G_{t-m}^{t+m}-measurable. Thus,

$$\begin{aligned} &E\left|\varepsilon_{t-k}^2 - E\left(\varepsilon_{t-k}^2|G_{t-m}^{t+m}\right)\right|^2 \\ &\quad\le 2E|\varepsilon_{t-k}^2 - \varepsilon_{t-k,\,m-k}^2|^2 + 2E\left|E\left[\left(\varepsilon_{t-k}^2 - \varepsilon_{t-k,\,m-k}^2\right)|G_{t-m}^{t+m}\right]\right|^2. \end{aligned} \tag{B.1}$$

Note that, since $\{\eta_t\}$ are i.i.d. random variables, $\{A_t\}$ are i.i.d. random matrices and, further, we have

$$\begin{aligned} &E\left[\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\right)\otimes\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\right)\right] \\ &\quad= E\left(\prod_{i=0}^{m-k-1}A_{t-k-i}\otimes\prod_{i=0}^{m-k-1}A_{t-k-i}\right) \\ &\qquad\times E\left[\left(\sum_{j=m-k+1}^{\infty}\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\otimes\left(\sum_{j=m-k+1}^{\infty}\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\right] \\ &\quad= \left[E(A_t\otimes A_t)\right]^{m-k}\left\{\sum_{j=m-k+1}^{\infty}E\left[\left(\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\otimes\left(\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\right]\right. \end{aligned} \tag{B.2}$$

$$
\begin{aligned}
&+\sum_{r=1}^{\infty}\sum_{j=m-k+1}^{\infty}\left(E\left[\left(\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\right.\right.\\
&\qquad\left.\otimes\left(\prod_{i=m-k}^{j-1+r}A_{t-k-i}\right)\right]\\
&\left.\left.+E\left[\left(\prod_{i=m-k}^{j-1+r}A_{t-k-i}\right)\otimes\left(\prod_{i=m-k}^{j-1}A_{t-k-i}\right)\right]\right)\right\}\\
&=\left[E(A_t\otimes A_t)\right]^{m-k}\left\{\sum_{j=m-k+1}^{\infty}\left[E(A_t\otimes A_t)\right]^{j-m+k}\right.\\
&\qquad+\sum_{r=1}^{\infty}\sum_{j=m-k+1}^{\infty}\left[E(A_t\otimes A_t)\right]^{j-m+k}\\
&\qquad\left.\times\left\{I\otimes\left[E(A_t)\right]^r+\left[E(A_t)\right]^r\otimes I\right\}\right\}\\
&=O(\rho^{m-k}),
\end{aligned}
$$

where $0\le\rho<1$, the last equation holds by $\rho(E(A_t))<1$, which is equivalent to Assumption 3 [Ling (1995)] and Assumption 5. Since $\{(A_t,\xi_t)\}$ are i.i.d. random variables, by (B.2), we have

$$
\begin{aligned}
&E|\varepsilon_{t-k}^2-\varepsilon_{t-k,m-k}^2|^2\\
&\quad=E\left|\sum_{j=m-k+1}^{\infty}\gamma'\prod_{i=0}^{j-1}A_{t-k-i}\xi_{t-k-i}\right|^2\\
&\quad=(\gamma'\otimes\gamma')\operatorname{vec}E\left[\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\xi_{t-k-j}\right)\right.\\
&\qquad\left.\times\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\xi_{t-k-j}\right)'\right] \qquad \text{(B.3)}\\
&\quad=(\gamma'\otimes\gamma')E\left[\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\right)\right.\\
&\qquad\left.\otimes\left(\sum_{j=m-k+1}^{\infty}\prod_{i=0}^{j-1}A_{t-k-i}\right)\right]\operatorname{vec}\left[E(\xi_t\xi_t')\right]\\
&\quad=O(\rho^{m-k}),
\end{aligned}
$$

$$
\begin{aligned}
&E\left[\left(\varepsilon_{t-k}^2 - \varepsilon_{t-k,m-k}^2\right)|G_{t-m}^{t+m}\right] \\
&\quad = E\left[\sum_{j=m-k+1}^{\infty} \gamma' \prod_{i=0}^{j-1} A_{t-k-i}\,\xi_{t-k-i}|G_{t-m}^{t+m}\right] \\
&\quad = \gamma' \prod_{i=0}^{m-k} A_{t-k-i} \cdot \left[I_{(r+s)\times(r+s)} + \sum_{j=m-k+2}^{\infty} E\left(\prod_{i=m-k+1}^{j-1} A_{t-k-i}\right)\right] \\
&\qquad \times E\left(\xi_{t-k-j}\right) \qquad \text{(B.4)} \\
&\quad = \gamma' \prod_{i=0}^{m-k} A_{t-k-i} c,
\end{aligned}
$$

where $c = (I_{(r+s)\times(r+s)} + \sum_{j=m-k+2}^{\infty}[E(A_t)]^{j-m+k-1})E\xi_t < \infty$, since $\rho(E(A_t)) < 1$. Similar to (B.3), we can show that

$$
E\left|\gamma' \prod_{i=0}^{m-k} A_{t-k-i} c\right|^2 = O\left(\rho^{m-k}\right). \qquad \text{(B.5)}
$$

By (B.1), (B.3)–(B.5), we know that (a) holds.

Using (a), we can show that (b) holds, and the detail is in Ling and Li (1996). By (b), it is easy to verify that (c) and (d) hold. (e) comes directly from (d). The proof of (f) is similar to that of (b) and hence is omitted. This completes the proof. □

Acknowledgment. The authors would like to thank Professor N. H. Chan, two referees, an Associate Editor and the Editor for helpful comments which led to improvements of this paper.

REFERENCES

AITCHISON, J. and SILVEY, S. D. (1958). Maximum-likelihood estimation of parameters subject to restraints. *Ann. Math. Statist.* **29** 813–828.

BAI, J. (1993). On the partial sums of residuals in autoregressive and moving average models. *J. Time Ser. Anal.* **14** 247–260.

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

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327.

BOLLERSLEV, T., CHOU, R. Y. and KRONER, K. F. (1992). ARCH modeling in finance. *J. Econometrics* **52** 5–59.

BOLLERSLEV, T., ENGLE, R. F. and NELSON, D. B. (1994). ARCH model. In *Handbook of Econometrics IV* (R. F. Engle and D. C. McFadden, eds.) 2959–3038. North-Holland, Amsterdam.

CHAN, N. H. and WEI, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15** 1050–1063.

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

CHUNG, K. L. (1968). *A Course in Probability Theory*. Harcourt, Brace and World, New York.

CHUNG, K. L. and WILLIAMS, R. J. (1990). *Introduction of Stochastic Integration*, 2nd ed. Birkhäuser, Boston.

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.

ENGLE, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica* **50** 987–1008.

FULLER, W. A. (1976). *Introduction to Statistical Time Series*. Wiley, New York.

HASZA, D. P. and FULLER, W. A. (1979). Estimation for autoregressive processes with unit roots. *Ann. Statist.* **7** 1106–1120.

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.

KIM, K. and SCHMIDT, P. (1993). Unit root tests with conditional heteroskedasticity. *J. Econometrics* **59** 287–300.

KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorems to stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** 1035–1070.

LEE, S.-W. and HANSEN, B. E. (1994). Asymptotic theory for GARCH $(1, 1)$ quasi-maximum likelihood estimator. *Econometric Theory* **10** 29–52.

LING, S. (1995). On probability properties of a double threshold ARMA conditional heteroskedasticity models. Technical report, Dept. Statistics, Univ. Hong Kong.

LING, S. and LI, W. K. (1996). Limiting distributions of maximum likelihood estimators for unstable ARMA time series with GARCH errors. Technical report, Dept. Statistics, Univ. Hong Kong.

LING, S. and LI, W. K. (1997a). On fractionally integrated autoregressive moving-average time series models with conditional heteroscedasticity. *J. Amer. Statist. Assoc.* To appear.

LING, S. and LI, W. K. (1997b). Estimating and testing for unit root processes with GARCH $(1, 1)$ errors. Technical report, Dept. Statistics, Univ. Hong Kong.

NELSON, D. B. (1990). Stationarity and persistence in the GARCH $(1, 1)$ model. *Econometric Theory* **6** 318–334.

PANTULA, S. G. (1989). Estimation of autoregressive models with ARCH errors. *Sankhyā Ser. B* **50** 119–138.

PANTULA, S. G. and HALL, A. (1991). Testing for unit roots in autoregressive moving average models. *J. Econometrics* **48** 325–353.

PETERS, T. A. and VELOCE, W. (1988). Robustness of unit root tests in ARMA models with generalized ARCH errors. Unpublished manuscript.

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

PHILLIPS, P. C. B. and DURLAUF, S. N. (1986). Multiple time series regression with integrated processes. *Rev. Econom. Stud.* **LIII** 473–495.

TSAY, R. S. and TIAO, G. C. (1984). Consistent estimates of autoregressive parameters and the extended sample autocorrelation function of nonstationary and stationary ARMA models. *J. Amer. Statist. Assoc.* **79** 84–96.

TSAY, R. S. and TIAO, G. C. (1990). Asymptotic properties of multivariate nonstationary processes with applications to autoregressions. *Ann. Statist.* **18** 220–250.

WEISS, A. A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Econometric Theory* **2** 107–131.

WOOLDRIDGE, J. M. and WHITE, H. (1988). Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory* **4** 210–230.

YAP, S. F. and REINSEL, G. C. (1995a). Estimation and testing for unit roots in a partially nonstationary vector autoregressive moving average model. *J. Amer. Statist. Assoc.* **90** 253–267.

YAP, S. F. and REINSEL, G. C. (1995b). Results on estimation and testing for unit roots in the nonstationary autoregressive moving average model. *J. Time Ser. Anal.* **16** 339–353.

DEPARTMENT OF STATISTICS
UNIVERSITY OF HONG KONG
POKFULAM ROAD
HONG KONG
E-MAIL: statist@hkucc.hku.hk