Statistica Sinica **17**(2007), 161-175

A DOUBLE AR(p) MODEL: STRUCTURE AND ESTIMATION

Shiqing Ling

Hong Kong University of Science and Technology

Abstract: The paper considers the so-called double AR(p) model,

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \eta_t \sqrt{\omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2},$$

where $\eta_t \sim$ i.i.d. $N(0,1)$. It is shown that the necessary and sufficient condition for the existence of a strictly stationary solution to the model is that the top Lyapounov exponent γ, defined in the paper, be negative; the solution is then unique and geometrically ergodic. The necessary and sufficient condition for the existence of a strictly stationary solution to the model with $Ey_t^2 < \infty$ is also obtained. The maximum likelihood estimator of the parameters in the model is shown to be asymptotically normal. The condition for this is again only that γ is negative which includes the case with some roots of $1 - \sum_{i=1}^{p} \phi_i z^i = 0$ on or outside the unit circle, and the case with $Ey_t^2 = \infty$. The result is novel because all kinds of estimated ϕ_i's in these cases are not asymptotically normal in the classical AR(p) model with i.i.d. errors; it may provide new insights in this direction.

Key words and phrases: Asymptotic normality, double autoregressive model, maximum likelihood estimator, stationarity, geometric ergodicity.

1. Introduction

Consider the autoregressive (AR) model with conditional heteroscedasticity:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \eta_t \sqrt{\omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2}, \tag{1.1}$$

where $\omega, \alpha_i > 0$, $t \in \mathcal{N} \equiv \{-p, \ldots, 0, 1, 2, \ldots\}$, $\{\eta_t\}$ is an independent random sequence, $\eta_t \sim N(0,1)$, and y_s is independent of $\{\eta_t : t \geq 1\}$ for $s \leq 0$. Let $\mathcal{F}_t$ be the σ-field generated by $\{\eta_t, \ldots, \eta_1, y_0, \ldots, y_{-p}\}$, $t \in \mathcal{N}$. The conditional variance of y_t is $var(y_t|\mathcal{F}_{t-1}) = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2$. Model (1.1) is a special case of ARMA-ARCH models in Weiss (1986) and an example of weak ARMA models in Francq and Zakoïan (1998, 2000), but it differs from Engle's (1982) ARCH model if $\phi_i \neq 0$. We call (1.1) the pth-order double AR(DAR(p)) model. Motivation for the DAR(p) models can be found in Weiss (1984) and Ling (2004). Up to now, Engle's ARCH models have been well-understood. For some important results on Engle's ARCH models, we refer to Berkes, Horvath and Kokoszka (2003), Hall and Yao (2003), Francq and Zakoïan (2004) and Jensen and Rahbek (2004). However, we know relatively little bit about Weiss' ARCH-type models.

The first focus of this paper is to investigate the structure of (1.1). When $p = 1$, it was studied by Guégan and Diebolt (1994), Borkovec (2001) and Borkovec and Kluppelberg (2001). A general theory for this class of time series models was developed by Chen and An (1999) and Chen and Chen (2000). However, under which conditions the general DAR(p) model is stationary and ergodic remains an open problem. This paper solves this problem when the η_t are normal. In Section 2, we give the necessary and sufficient conditions for the strict stationarity and the weak stationarity of model (1.1). Furthermore, these conditions are sufficient for the ergodicity.

The second focus of this paper is to investigate the estimation of (1.1). Weiss (1986) first proved the asymptotic normality of the quasi-maximum likelihood estimator (MLE) of the parameters in (1.1). But he assumes that $Ey_t^4 < \infty$. The moment condition of y_t directly links the restriction to the parameters. The following table from Li, Ling and McAleer (2002) gives some moment conditions of with $\phi_1 = 0$ and $p = 1$, i.e., ARCH(1) model. It can be seen that $Ey_t^4 < \infty$ is

y_t	Strict stationarity	2nd moment	4th moment	8th moment
α_1	$(0, 3.5620\cdots)$	$(0, 1)$	$(0, 0.57\cdots)$	$(0, 0.3\cdots)$

a very strong condition. Up to date, the condition for $Ey_t^4 < \infty$ has not yet been established if $\phi_i \neq 0$. However, from the figure in Section 2, we can see that $Ey_t^2 < \infty$ is a much stronger condition than that for the strict stationarity and hence so is $Ey_t^4 < \infty$. Ling (2004) showed that the MLE of the parameters in (1.1) with $p = 1$ is consistent and asymptotically normal only under the strict stationarity condition. Its least absolute deviation estimator was studied by Chan and Peng (2005). In Section 3, we extend Ling's results to general DAR(p) models. All the proofs are given in Appendix A and Appendix B.

2. Structures of DAR(p) Models

We first let $\xi_t = (\xi_{1t}, \ldots, \xi_{pt})$ be an independent $p \times 1$ standard normal vector independent of $\{\eta_t\}$, and let A_t be the $p \times p$ random matrix

$$A_t = \begin{pmatrix} \phi_1 + \sqrt{\alpha_1}\xi_{1t} & \cdots & \phi_{p-1} + \sqrt{\alpha_{p-1}}\xi_{p-1,t} & \phi_p + \sqrt{\alpha_p}\xi_{pt} \\ & I_{p-1} & & O_{(p-1)\times 1} \end{pmatrix},$$

where I_r is the $r \times r$ identity matrix and $O_{r\times s}$ is the $r \times s$ zero matrix. Let $\|M\| = \sqrt{tr(MM')}$ for a vector or matrix, M. To state our result, we need the notion of top Lyapounov exponent, defined as

$$\gamma = \inf\{\frac{1}{n}E\ln\|A_1\cdots A_n\|,\ n \geq 1\}. \tag{2.1}$$

Since $\xi_{it} \sim N(0,1)$, we can show that $E\ln^+\|A_1\| < \infty$, where $\ln^+ x = \max\{\ln x, 0\}$. Thus, it follows from the subadditive ergodic theorem (see Kingman (1973, Theorem 6)) that, almost surely (a.s.),

$$\gamma = \lim_{n\to\infty}\frac{1}{n}\ln\|A_1\cdots A_n\|. \tag{2.2}$$

Given ϕ_i and α_i it is generally difficult to compute γ, but we can easily estimate γ by using simulation to sufficiently large n. Now, we can state the following theorem.

Theorem 2.1. *The necessary and sufficient condition for the existence of a strictly stationary solution to* (1.1) *is* $\gamma < 0$. *Furthermore, the solution* $\{y_t : t \in \mathcal{N}\}$ *is unique, geometrically ergodic, and* $E|y_t|^u < \infty$ *for some* $u > 0$.

Remark 2.1. The proof of Theorem 2.1 transforms (1.1) into a Markov chain. We then verify the regularity conditions in Tweedie (1983) and Tjϕstheim (1990) for sufficiency. This method has been commonly used to find sufficient conditions for stationarity of nonlinear time series models, see for example Tong (1990), Chen and An (1998, 1999), Carrasco and Chen (2002) and Ling and McAleer (2002). However, the verification of these conditions for (1.1) is nonstandard. Our main idea is to find a random coefficient AR (RCAR) model such that the model has the same transition probability as (1.1). The proof of necessity is to verify the conditions in Bougerol and Picard (1992).

Remark 2.2. The result in Theorem 2.1 holds for the following general case:

$$y_t = \sum_{i=1}^{p_1}\phi_i y_{t-i} + \eta_t\sqrt{\omega + \sum_{i=1}^{p_2}\alpha_i y_{t-i}^2},$$

by letting $p = \max\{p_1, p_2\}$ with $\phi_i = 0$ for $i > p_1$, and $\alpha_j = 0$ for $j > p_2$. Theorem 2.1 implies that the strictly stationary solution $\{y_t\}$ to (1.1) has an absolute finite moment $E|y_t|^u < \infty$ for some $u > 0$. This also implies that the strictly stationary solution to the ARCH (p) model (i.e., (1.1) with $\phi_1 = \cdots = \phi_p = 0$) has an finite absolute uth moment. A similar result for a GARCH(r, s) model was established in Ling (2007).

Let $\rho(A)$ be the moduli of the matrix A (i.e., $\rho(A) = \max_i |v_i|$, where v_i is the eigenvalues of A). We next give the result for the finite second moment of y_t.

Theorem 2.2. *The necessary and sufficient condition for the existence of a strictly stationary solution,* $\{y_t : t \in \mathcal{N}\}$ *with* $Ey_t^2 < \infty$, *to model* (1.1) *is that* $\rho[E(A_t \otimes A_t)] < 1$, *where* $\otimes$ *denotes the Kronecker product of matrices. The solution is unique and geometrically ergodic.*

Remark 2.3. From the proof, we can see that (1.1) is equivalent to an RCAR model in distribution. The equivalence of the two models in the second moment was given by Tsay (1987). The RCAR models were already studied in the literature by many authors, e.g., Pham (1986). However, the structures (or data generating mechanisms) of (1.1) and the RCAR model are not the same at all, although they have the same marginal distribution.

Remark 2.4. When $p = 1$, $\gamma = E\ln|\phi_1 + \sqrt{\alpha_1}\xi_t|$. In this case, the condition in Theorem 2.2 reduces to $\phi_1^2 + \alpha_1 < 1$. Figure 1 gives the regions of (ϕ_1, α_1) such that $\phi_1^2 + \alpha_1 < 1$ and $E\ln|\phi_1 + \sqrt{\alpha_1}\xi_t| < 0$. For the general case, note that $\|A_1 \cdots A_n\|^2 = C'[\prod_{t=1}^n (A_t \otimes A_t)]C$ with some constant vector C. By Jensen's Inequality, we see that $\rho[E(A_t \otimes A_t)] < 1$ implies $\gamma < 0$. From Figure 1, we can see that the condition $\gamma < 0$ allows the case with some roots of $\phi(z) \equiv 1 - \sum_{i=1}^p \phi_i z^i = 0$ on or outside the unit circle, and the case with $Ey_t^2 = \infty$.

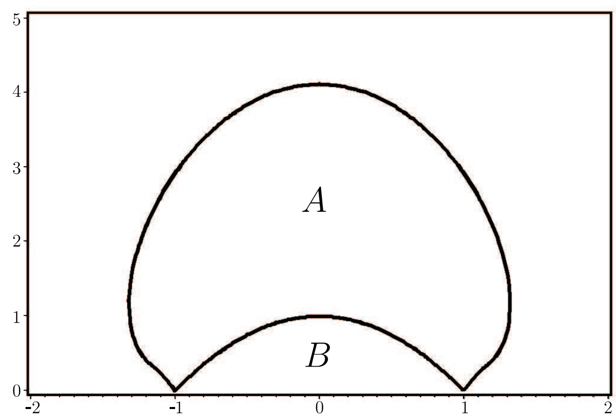

Figure 1. $E\ln|\phi_1 + \sqrt{\alpha_1}\xi_t| < 0$ as $(\phi_1, \alpha_1) \in A \cup B$ and $\phi_1^2 + \alpha_1 < 1$ as $(\phi_1, \alpha_1) \in B$.

Remark 2.5. Model (1.1) is a special case of the following general model,

$$y_t = f(y_{t-1}, \ldots, y_{t-p}) + \eta_t \sqrt{h(y_{t-1}, \ldots, y_{t-p})},$$

where $h(y_{t-1}, \ldots, y_{t-p}) > 0$ and $\{\eta_t\}$ is a sequence of i.i.d. random variables with zero mean and variable 1. Under this framework, stationarity and moment conditions have been extensively studied in the literature, e.g., Ango Nze (1992)

Masry and Tjϕtheim (1995), Lu (1998), Cline and Pu (1999) and Lu and Jiang (2001). Unlike our method, these papers obtain sufficient conditions for stationarity and the existence of moments by directly checking the regularity conditions in Tweedie (1983) or Tjϕstheim (1990). Applying their results to (1.1), the weakest sufficient conditions for stationarity is from Lu and Jiang (2001), which is

$$\sum_{i=1}^{p}|\phi_i| + E|\eta_t|\sum_{i=1}^{p}\alpha_i < 1,$$

where $E|\eta_t| = \sqrt{2/\pi}$ when η_t is normal. This condition is also sufficient for $E|y_t| < \infty$. Comparing this condition with that in Theorem 2.1, we can see that it is far from the necessary condition, see Remark 2.4. The weakest sufficient conditions for $Ey_t^2 < \infty$ is from Lu (1998), which is

$$\Big(\sum_{i=1}^{p}|\phi_i|\Big)^2 + \sum_{i=1}^{p}\alpha_i < 1.$$

Except for the case with $p = 1$, this condition is also stronger than that in Theorem 2.2. For example, when $\alpha_1 = \cdots = \alpha_p = 0$, the condition in Theorem 2.2 reduces to the necessary and sufficient one for stationarity of the usually AR(p) model, which is much weaker than $\sum_{i=1}^{p}|\phi_i| < 1$. In practice, we should first see if these simple conditions are satisfied. If not, we then check the condition in Theorem 2.1 or 2.2. Necessary and sufficient condtion for the DAR(p) model when $p > 1$ and η_t is not normal are still not known.

3. Maximum Likelihood Estimation

Let $\lambda = (\lambda_1', \lambda_2')'$ with $\lambda_1 = (\phi_1, \ldots, \phi_p)'$ and $\lambda_2 = (\omega, \alpha_1, \ldots, \alpha_p)'$. Suppose $\{y_{-p}, \ldots, y_n\}$ is generated by (1.1) with parameter $\lambda_0 = (\lambda_{10}', \lambda_{20}')'$. The conditional log-likelihood function (ignoring a constant) can be written as

$$L_n(\lambda) = \sum_{t=2}^{n} l_t(\lambda) \quad \text{and} \quad l_t(\lambda) = -\frac{1}{2}\ln(\lambda_2' Y_{2t-1}) - \frac{\varepsilon_t^2(\lambda)}{2(\lambda_2' Y_{2t-1})}, \qquad (3.1)$$

where $\varepsilon_t(\lambda) = y_t - \lambda_1' Y_{1t-1}$, $Y_{1t} = (y_t, \ldots, y_{t-p+1})'$, and $Y_{2t} = (1, y_t^2, \ldots, y_{t-p+1}^2)'$. Let Θ be the parameter space. The MLE of λ_0, denoted by $\hat\lambda_n$, is the maximizer of $L_n(\lambda)$ on Θ. Let $\underline{\omega}$, $\bar\omega$, $\underline{\alpha}$ and $\bar\alpha$ be some positive constants. Our condition is the following.

Assumption 3.1. *Θ is compact with $\underline{\omega} \le \omega \le \bar\omega$ and $\underline{\alpha} \le \alpha_i \le \bar\alpha$ ($i = 1, \ldots, p$), λ_0 is an interior point in Θ and $\gamma < 0$ for each $\lambda \in \Theta$.*

The following theorem gives the asymptotic properties of the MLE.

Theorem 3.1. *Suppose that* $\{y_t : t \in \mathcal{N}\}$ *is the strictly stationary and ergodic solution of* (1.1). *If Assumption* 3.1 *holds, then* $\sqrt{n}(\hat{\lambda}_n - \lambda_0) \longrightarrow_{\mathcal{L}} N\Big(0, \Omega^{-1}\Big)$, *as* $n \to \infty$, *where* $\to_{\mathcal{L}}$ *denote convergence in distribution, and*

$$\Omega = \mathrm{diag}\left\{E\left(\frac{Y_{1t}Y_{1t}'}{\lambda_{20}'Y_{2t})}\right),\ \frac{1}{2}E\left[\frac{Y_{2t}Y_{2t}'}{(\lambda_{20}'Y_{2t})^2}\right]\right\}.$$

Remark 3.1. When η_t is not normal, we still can use (3.1). In this case, $\hat{\lambda}_n$ is only the quasi-MLE of λ_0, and Theorem 3.1 holds with Ω^{-1} replaced by $\Omega^{-1}\Sigma\Omega^{-1}$ if $E\eta_t^4 < \infty$ and $J > 0$, where

$$\Sigma = E\left\{\left(\frac{Y_{1t}}{\sqrt{\lambda_{20}'Y_{2t}}}, \frac{Y_{2t}}{\lambda_{20}'Y_{2t}}\right) J \begin{pmatrix} \frac{Y_{1t}'}{\sqrt{\lambda_{20}'Y_{2t}}} \\ \frac{Y_{2t}'}{\lambda_{20}'Y_{2t}} \end{pmatrix}\right\} \ \textit{and}\ J = \begin{pmatrix} 1 & E\eta_t^3 \\ E\eta_t^3 & E\eta_t^4 - 1 \end{pmatrix}.$$

The proof for this is the same as that in the Appendix B, except that we need to modify Lemma B.5 with Ω replaced by Σ.

Remark 3.2. The condition $\underline{\alpha} > 0$ in Assumption 3.1 is used because $\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2$ can control the log-likelihood function, score function, and information matrix in such a way that they are bounded. We cannot obtain Theorem 3.1 without this condition. The MLE is an optimal estimator in LeCam's sense and its asymptotic normality implies that the log-likelihood function (3.1) is locally asymptotically normal. Theorem 3.1 shows that the MLE of λ_{10} may be asymptotically normal in the case with some roots of $\phi(z) = 0$ on or outside the unit circle, and in the case with $Ey_t^2 = \infty$. It is well known that all kinds of the estimated ϕ_i's in these cases are not asymptotically normal in the classical AR(p) model with i.i.d. errors, see Dickey and Fuller (1979), Chan and Tran (1989) and Davis, Knight and Liu (1992).

Appendix A. Proof of Theorems 2.1−2.2

Proof of Theorem 2.1. Let $Y_t = (y_t, y_{t-1}, \ldots, y_{t-p+1})'$, $\mathcal{B}_p$ be the class of Borel sets of R^p, and let ν_p be the Lebesgue measure on $(R^p, \mathcal{B}^p)$. Then $(R^p, \mathcal{B}^p, \nu_p)$ is the state space of the process $\{Y_t\}$. Let $m : R^p \to R$ be the projection map onto the first coordinate, i.e., $m(x) = x_1$ as $x \in R^p$. Then $\{Y_t\}$ is a homogeneous Markov chain with state space $(R^p, \mathcal{B}^p, \nu_p)$. It has the transition probability

$$P(x, A) = \int_{m(A)} \frac{1}{\sqrt{\lambda_2'\tilde{x}}} f\Big(\frac{z_1 - \lambda_1'x}{\sqrt{\lambda_2'\tilde{x}}}\Big) dz_1,\ x \in R^p \text{ and } A \in \mathcal{B}^p, \qquad \text{(A.1)}$$

where $x = (x_p, \ldots, x_1)$, $\tilde{x} = (1, x_p^2, \ldots, x_1^2)$, and $f(x) = (2\pi)^{-0.5} e^{-x^2/2}$.

We first check the ν_p-irreducibility of the Markov chain $\{Y_t\}$, i.e., whether $\sum_{n=1}^{\infty} P^n(x,A) > 0$ for every $x \in R^p$ whenever $\nu_p(A) > 0$, where

$$P^n(x,A) = \int_{R^p} P^{n-1}(y,A)P(x,dy), \ x \in R^p, A \in \mathcal{B}^p.$$

It is easy to see that the p-step transition probability of the Markov chain $\{Y_t\}$ is

$$P^{\,p}(x,A) = \int_A \prod_{i=1}^{p} \frac{1}{\sqrt{\lambda_2'\tilde{X}_i}} f\Big(\frac{z_i - \lambda_1' X_i}{\sqrt{\lambda_2'\tilde{X}_i}}\Big) dz_1 \cdots dz_p, \tag{A.2}$$

where $X_i = (z_i, \ldots, z_1, x_1, \ldots, x_{p-i})$ and $\tilde{X}_i = (1, z_i^2, \ldots, z_1^2, x_1^2, \ldots, x_{p-i}^2)$. Since the transition density kernel in (A.2) is positive, we know that $\{Y_t\}$ is ν_p-irreducible.

By (2.1), there exists an integer s such that $E \ln \|A_1 \cdots A_s\| < 0$. Let $\tilde{A}_t = A_t \cdots A_{t-s+1}$, and write $q(u) = E\|\tilde{A}_t\|^u$. Since $\eta_t \sim N(0,1)$, it is easy to show that $q(u)$ is differentiable on $[0,2)$ and $q'(u) = E[\|\tilde{A}_t\|^u \ln \|\tilde{A}_t\|]$. Note that $|\ln x^\delta| \le \max\{x^\delta, x^{-\delta}\} - 1$ for all $x > 0$ and $\delta \in [0, u/2]$. We can show that $E \sup_{u \in [0,1]} [\|\tilde{A}_t\|^u \ \ln \|\tilde{A}_t\|\Big|] < \infty$. By the Dominated Convergence Theorem, $\lim_{u \to 0} q'(u) = E \ln \|\tilde{A}_t\| < 0$. Thus, there exists a constant $\tilde{\delta} \in (0,2)$ such that $q(u)$ is strictly decreasing on $[0, \tilde{\delta}]$. Hence, there exists a constant $u \in (0,1)$ such that

$$E\|\tilde{A}_t\|^u < q(0) = 1. \tag{A.3}$$

Using (A.3), we next prove that the s-step Markov chain $\{Y_{ts}\}$ satisfies the drift condition of Theorem 4(ii) in Tweedie (1983), i.e., there exists a compact set K and a non-negative continuous function $g(x)$ such that $\nu_p(K) > 0$, $g(x) \ge 1$ on K, and

$$E(g(Y_{st})|Y_{(t-1)s} = x) \le (1-\epsilon)g(x), \ \ x \in K^c, \tag{A.4}$$

$$E(g(Y_{st})|Y_{(t-1)s} = x) \le M, \ \ x \in K, \tag{A.5}$$

for some $\epsilon > 0$. The key point is to find a function g such that (A.4)-(A.5) hold.

It is difficult to get g by a direct method. We first consider the RCAR(p) model

$$\tilde{Y}_t = A_t \tilde{Y}_{t-1} + \tilde{\eta}_t, \tag{A.6}$$

where $\tilde{\eta}_t = (\sqrt{\omega}\eta_t, 0, \ldots, 0)'$ and $\tilde{Y}_t$ is independent of $\{\tilde{\eta}_{t'} : t' < t\}$. It is easy to see that $\{\tilde{Y}_t\}$ is a homogeneous Markov chain with state space $(R^p, \mathcal{B}^p, \nu_p)$, and its transition probability is

$$P(x,A) = \int_{m(A)} \frac{1}{\sqrt{\lambda_2'\tilde{x}}} f\Big(\frac{z_1 - \lambda_1' x}{\sqrt{\lambda_2'\tilde{x}}}\Big) dz_1, \ x \in R^p \ and \ A \in \mathcal{B}^p. \tag{A.7}$$

We choose $g(x) = 1 + \|x\|^u$, where $x \in R^p$ and u is defined as in (A.3). For a fixed s such that (A.3) holds, we iterate (A.6) to obtain the following expansion:

$$\tilde{Y}_{ts} = \Big(\tilde{\eta}_{ts} + \sum_{j=1}^{s-1}\prod_{r=0}^{j-1} A_{ts-r}\tilde{\eta}_{ts-j}\Big) + \tilde{A}_{ts}\tilde{Y}_{(t-1)s}. \tag{A.8}$$

By (A.8), we have

$$\begin{aligned} E(g(\tilde{Y}_{st})|\tilde{Y}_{(t-1)s} = x) &\le 1 + E\|\tilde{\eta}_{ts} + \sum_{j=1}^{s-1}\prod_{r=0}^{j-1} A_{ts-r}\tilde{\eta}_{ts-j}\|^u + E\|\tilde{A}_{ts}\|^u\|x\|^u \\ &= E\|\tilde{A}_{ts}\|^u\|x\|^u + C, \end{aligned} \tag{A.9}$$

where C is some constant. Let $K = \{x : \|x\| \le L\}$ and L be a positive constant. It is easy to see that

$$E(g(\tilde{Y}_{st})|\tilde{Y}_{(t-1)s} = x) \le M, \;\; x \in K, \tag{A.10}$$

for some constant M. Note that $E\|\tilde{A}_{ts}\|^u = E\|\tilde{A}_t\|^u$. As L is large enough and $x \in K^c$, by (A.3) there exists $\epsilon > 0$ such that

$$\begin{aligned} E(g(\tilde{Y}_{st})|\tilde{Y}_{(t-1)s} = x) &= g(x) + (E\|\tilde{A}_{ts}\|^u - 1)\|x\|^u + C - 1 \\ &\le g(x)\{1 - [(1 - E\|\tilde{A}_{ts}\|^u) + \frac{C-1}{1+\|x\|^u}]\} \\ &\le (1-\epsilon)g(x). \end{aligned} \tag{A.11}$$

From (A.1) and (A.7), we know that $\{Y_t\}$ and $\{\tilde{Y}_t\}$ have the same transition kernel density. By (A.10)−(A.11), we know that (A.4)−(A.5) holds with the same $g(x)$ and K. For each bounded continuous function G on R^p, $E[G(Y_{st})|Y_{s(t-1)} = x]$ is continuous in x. Thus, $\{Y_{st}\}$ is a Feller chain. Furthermore, since Y_{st} is ν_p-irreducible, by Theorems 1−2 in Feigin and Tweedie (1985), we know that (i) Y_{st} is geometrically ergodic, which ensures that there exists a unique stationary distribution π for $\{Y_{st}\}$, and (ii)

$$\int \|Y_{st}\|^u d\pi \le \int_{R^p} g(x)\pi(dx) < \infty. \tag{A.12}$$

By Lemma 3.1 in Tjϕstheim (1990), Y_t is geometrically ergodic. Thus, Y_t has a unique stationary distribution π. Let Y_0 be initialized from the stationary distribution π. Then $\{y_t : t \in \mathcal{N}\}$ is the unique stationary solution to (1.1) and it is geometrically ergodic. Furthermore, by (A.12), $E|y_t|^u < \infty$.

We next consider the necessity. Assume $\{y_t : t \in \mathcal{N}\}$ is the strictly stationary solution to (1.1). Then the Markov chain $\{Y_t\}$ has a stationary distribution π.

Since $\{Y_t\}$ and $\{\tilde{Y}_t\}$ have the same transition kernel density, the Markov chain $\{\tilde{Y}_t\}$ has the stationary distribution π. Let $\tilde{Y}_0$ be initialized from the stationary distribution π. Then, $\{\tilde{Y}_t : t \in \mathcal{N}\}$ is the stationary solution to model (A.6).

By (A.8) with $s = p$, we have

$$\tilde{Y}_{tp} = B_{tp} + \tilde{A}_{tp}\tilde{Y}_{(t-1)p}, \tag{A.13}$$

where $B_{tp} = \tilde{\eta}_{tp} + \sum_{j=1}^{p-1}\prod_{r=0}^{j-1} A_{tp-r}\tilde{\eta}_{tp-j}$. Note that now $\tilde{A}_t = A_t \cdots A_{t-p+1}$. Since ξ_t is normal and $\ln(1+x) \le x$ as $x \ge 0$, we have that $E\ln^+\|\tilde{A}_{tp}\| = E\max\{\ln\|\tilde{A}_{tp}\|, 0\} = E[I\{\|\tilde{A}_{tp}\| \ge 1\}\ln\|\tilde{A}_{tp}\|] \le E[(\|\tilde{A}_{tp}\| - 1)I\{\|\tilde{A}_{tp}\| \ge 1\}] \le E\|\tilde{A}_{tp}\| + 1 < \infty$. Similarly, we can show that $E\ln^+\|B_{tp}\| < \infty$. By the assumption of model (A.6), $\{\tilde{Y}_{tp} : t \in \mathcal{N}\}$ is a nonanticipative solution to (A.13). Since $\{\tilde{Y}_{tp} : t \in \mathcal{N}\}$ is strictly stationary, it follows that

$$P(\tilde{Y}_{tp} \in A|\tilde{Y}_{(t-1)p} = x) = P(\tilde{Y}_p \in A|\tilde{Y}_0 = x) = P^p(x, A). \tag{A.14}$$

By (A.2) and (A.14), the density of $\tilde{Y}_{tp}$ given $\tilde{Y}_{(t-1)p} = x$ is positive. Let H be any affine invariant subspace of R^p under model (A.13) (i.e., $\{B_{tp} + \tilde{A}_{tp}x : x \in H\} \subseteq H$ a.s.). If $\nu_p(R^p - H) \ne 0$, then, for any $x \in H$,

$$\begin{aligned}P(B_{tp} + \tilde{A}_{tp}x \in H) &= P(\tilde{Y}_{tp} \in H|\tilde{Y}_{(t-1)p} = x)\\ &= P(\tilde{Y}_{tp} \in R^p|\tilde{Y}_{(t-1)p} = x) - P(\tilde{Y}_{tp} \in R^p - H|\tilde{Y}_{(t-1)p} = x) < 1.\end{aligned} \tag{A.15}$$

It is obvious that R^p is an affine invariant subspace. By (A.15), the affine invariant subspace is unique. Thus, model (A.13) is irreducible. By Theorem 2.5 in Bougerol and Picard (1992), the necessary condition for a nonanticipative strictly stationary solution $\tilde{Y}_{tp}$ to model (A.13) is that the top Lyapounov exponent

$$\tilde{\gamma} = \inf\left\{\frac{1}{t}E\ln\|\tilde{A}_p\tilde{A}_{2p}\cdots\tilde{A}_{tp}\|, t \ge 1\right\} < 0. \tag{A.16}$$

Note that $\tilde{A}_p\tilde{A}_{2p}\cdots\tilde{A}_{tp} = A_1\cdots A_{tp}$. By (A.16), there exists an s such that $E\ln\|A_1 \cdots A_{sp}\| < 0$. Write $\tilde{A}^*_t = A_t \cdots A_{t-sp+1}$. Then $E\ln\|\tilde{A}^*_{sp}\| = E\ln\|A_1\cdots A_{sp}\| < 0$. As for (A.3), there exists $u \in (0, 1)$ such that $E\|\tilde{A}^*_{sp}\|^u < 1$.

For any n, let $n = msp + r$, where $r = 0, \ldots, sp - 1$. We have $A_1\cdots A_n = \tilde{A}^*_{sp}\tilde{A}^*_{2sp}\cdots\tilde{A}^*_{msp} \cdot A_{msp+1}\cdots A_{msp+r}$, where $A_{msp+1}\cdots A_{msp+r} = 1$ when $r = 0$. Since $\{A_i\}$ is a sequence of i.i.d. random matrices, we have

$$\begin{aligned}E\|A_1A_2\cdots A_n\|^u &\le E\|\tilde{A}^*_{sp}\tilde{A}^*_{2sp}\cdots\tilde{A}^*_{msp}\|^u E\|A_{msp+1}\cdots A_{msp+r}\|^u\\ &\le [E\|\tilde{A}^*_{sp}\|^u]^m[E\|A_1\|^u]^r = O([E\|\tilde{A}^*_{sp}\|^u]^m) \to 0,\end{aligned}$$

as $n \to \infty$, which implies $m \to \infty$. Thus, there exists an n such that $E\ln\|A_1\cdots A_n\| \le u^{-1}\ln E\|A_1\cdots A_n\|^u < 0$, and hence $\gamma < 0$.

Proof of Theorem 2.2. Let $\{Y_t\}$ and $\{\tilde{Y}_t\}$ be defined as in the proof of Theorem 2.1. Since $\{Y_t\}$ and $\{\tilde{Y}_t\}$ have the same transition kernel density, similar to the proof of Theorem 2.1 we only need to show that $\rho[E(A_t \otimes A_t)] < 1$ is necessary and sufficient for the existence of a unique strictly stationary solution $\tilde{Y}_t$ to (A.6) with $E\|\tilde{Y}_t\|^2 < \infty$. The sufficiency and the uniqueness were given by Feigin and Tweedie (1985), while the necessity was given by Nicholls and Quinn (1982) and Ling (1999).

Appendix B. Proof of Theorem 3.1

The first lemma is Theorem 3.1 in Ling and McAleer (2003).

Lemma B.1. *Let $g(y, \theta)$ be a measurable function of y in Euclidean space for each $\theta \in \Theta$, a compact subset of R^p, and a continuous function of $\theta \in \Theta$ for each y. Suppose that $\{y_t\}$ is a sequence of strictly stationary and ergodic time series, such that $Eg(y_t, \theta) = 0$ and $E\sup_{\theta\in\Theta} |g(y_t, \theta)| < \infty$. Then*

$$\sup_{\theta\in\Theta}\Big|\frac{1}{n}\sum_{t=1}^{n} g(y_t, \theta)\Big| = o_p(1).$$

We now give the first- and second- derivatives of $l_t(\lambda)$ as follows.

$$\begin{aligned}
\frac{\partial l_t(\lambda)}{\partial \lambda_1} &= \frac{Y_{1t-1}\varepsilon_t(\lambda)}{\lambda_2' Y_{2t-1}},\\
\frac{\partial l_t(\lambda)}{\partial \lambda_2} &= -\frac{Y_{2t-1}}{2\lambda_2' Y_{2t-1}}\Big[1 - \frac{\varepsilon_t^2(\lambda)}{\lambda_2' Y_{2t-1}}\Big],\\
\frac{\partial^2 l_t(\lambda)}{\partial \lambda_1 \partial \lambda_1'} &= -\frac{Y_{1t-1}Y_{1t-1}'}{\lambda_2' Y_{2t-1}},\\
\frac{\partial^2 l_t(\lambda)}{\partial \lambda_2 \partial \lambda_2'} &= \frac{Y_{2t-1}Y_{2t-1}'}{2(\lambda_2' Y_{2t-1})^2}\Big[1 - \frac{2\varepsilon_t^2(\lambda)}{\lambda_2' Y_{2t-1}}\Big],\\
\frac{\partial^2 l_t(\lambda)}{\partial \lambda_1 \partial \lambda_2'} &= -\frac{Y_{1t-1}Y_{2t-1}'\varepsilon_t(\lambda)}{(\lambda_2' Y_{2t-1})^2}.
\end{aligned}$$

The following lemma is a basic result in our proof. It indicates why we need to bound ω and α_i in Θ and is used to obtain uniform convergence in Lemma B.3.

Lemma B.2. *If the assumption of Theorem* 3.1 *holds, then*

(i) $E\sup_{\lambda\in\Theta} |l_t(\lambda)| < \infty$,

(ii) $E\sup_{\lambda\in\Theta} \Big\|\dfrac{\partial^2 l_t(\lambda)}{\partial\lambda\partial\lambda'}\Big\| < \infty$.

Proof. (i) By Theorem 2.1, there exists a $u \in (0,1)$ such that $E|y_t|^u < \infty$. Let $\bar{\omega}^* = \max\{1, \bar{\omega}\}$. By Jensen's Inequality, we have

$$E\ln(\bar{\omega}^* + \sum_{i=1}^p \bar{\alpha}_i y_{t-i}^2) = \frac{2}{u} E\ln(\bar{\omega}^* + \sum_{i=1}^p \bar{\alpha}_i y_{t-i}^2)^{\frac{u}{2}}$$

$$\leq \frac{2}{u} \ln(\bar{\omega}^{*\frac{u}{2}} + \sum_{i=1}^p \bar{\alpha}_i^{\frac{u}{2}} E|y_{t-i}|^u) < \infty, \tag{B.1}$$

where the following elementary relation is used: $(\sum_{i=1}^p a_i)^s \leq \sum_{i=1}^p a_i^s$ for all $a_i > 0$ and $s \in (0,1)$. By (B.1), we have

$$E\sup_{\lambda\in\Theta} |\ln(\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2)| \leq E\sup_{\lambda\in\Theta} [I\{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 \geq 1\} \ln(\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2)]$$

$$+ E\sup_{\lambda\in\Theta} [-I\{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2 \leq 1\} \ln(\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2)]$$

$$\leq E\ln(\bar{\omega}^* + \sum_{i=1}^p \bar{\alpha}_i y_{t-i}^2) - I\{\underline{\omega} < 1\} \ln\underline{\omega} < \infty. \tag{B.2}$$

Furthermore, since $y_t - \sum_{i=1}^p \phi_i y_{t-i} = \varepsilon_t(\lambda_0) - \sum_{i=1}^p (\phi_i - \phi_{i0}) y_{t-i}$ and $\varepsilon_t(\lambda_0) = \eta_t \sqrt{\omega_0 + \sum_{i=1}^p \alpha_{i0} y_{t-i}^2}$, it follows that

$$E\sup_{\lambda\in\Theta} \Big[\frac{(y_t - \sum_{i=1}^p \phi_i y_{t-i})^2}{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2}\Big]$$

$$\leq 2E\sup_{\lambda\in\Theta} \Big[\frac{[\sum_{i=1}^p (\phi_i - \phi_{i0}) y_{t-1}]^2}{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2}\Big] + 2E\sup_{\lambda\in\Theta} \Big[\frac{\omega_0 + \sum_{i=1}^p \alpha_{i0} y_{t-i}^2}{\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2}\Big] \leq C, \tag{B.3}$$

where C is some finite constant. By (B.2)-(B.3), (i) holds.

(ii). As in (B.3), we can show that

$$E\sup_{\lambda\in\Theta} \Big[\frac{\partial l_t(\lambda)}{\partial \phi_i}\Big]^2 = E\sup_{\lambda\in\Theta} \frac{y_{t-i}^2 [\varepsilon_t(\lambda_0) - \sum_{i=1}^p (\phi_i - \phi_{i0}) y_{t-i}]^2}{(\omega + \sum_{i=1}^p \alpha_i y_{t-i}^2)^2}$$

$$\leq E\Big[\frac{C y_{t-i}^2 \sum_{j=1}^p y_{t-j}^2}{(\underline{\omega} + \sum_{i=1}^p \underline{\alpha}_i y_{t-i}^2)^2}\Big] + 2E\Big[\frac{y_{t-i}^2 \varepsilon_t^2(\lambda_0)}{(\underline{\omega} + \sum_{i=1}^p \underline{\alpha}_i y_{t-i}^2)^2}\Big]$$

$$\leq C_1 + 2E\Big[\frac{y_{t-i}^2 (\omega_0 + \sum_{i=1}^p \alpha_{i0} y_{t-i}^2)}{(\underline{\omega} + \sum_{i=1}^p \underline{\alpha}_i y_{t-i}^2)^2}\Big] \leq C,$$

where C and C_1 are some finite constants. Similarly, we can show that other terms in (ii) are finite. Thus, (ii) holds.

Before giving the proof of Theorem 2.1, we still need three lemmas. Lemmas B.3(i) and B.4 are used to obtain the consistency of $\hat{\lambda}_n$. Based on this consistency, we can use Lemmas B.3(ii) and B.5 to prove Theorem 2.1.

Lemma B.3. *If the assumption of Theorem* 3.1 *holds, then*

$$\text{(i)}\quad \sup_{\lambda\in\Theta}\Big|\frac{1}{n}L_n(\lambda)-El_t(\lambda)\Big| = o_p(1),$$

$$\text{(ii)}\quad \sup_{\lambda\in\Theta}\Big\|\frac{1}{n}\sum_{t=2}^{n}\Big\{\Big[\frac{\partial^2 l_t(\lambda)}{\partial\lambda\partial\lambda'}\Big]-E\Big[\frac{\partial^2 l_t(\lambda)}{\partial\lambda\partial\lambda'}\Big]\Big\}\Big\| = o_p(1).$$

Proof. This follows directly from Lemmas B.1 and B.2.

Lemma B.4. *Under the assumption of Theorem* 3.1*, $El_t(\lambda)$ has a unique maximum at λ_0.*

Proof. We first show that

$$c_1 = 0\ if\ c_1'Y_{1t}=0\ a.s. \quad\text{and}\quad c_2=0\ if\ c_2'Y_{2t}=0\ a.s., \tag{B.4}$$

where c_1 and c_2 are $p\times 1$ and $(p+1)\times 1$ constant vectors, respectively. If $c_1 \equiv (c_{11},\ldots,c_{1p})'$ is not a zero vector, for simplicity, we assume that $c_{11} = 1$. Then, $y_t = -\sum_{t=2}^{p} c_{1i}y_{t-i+1}$ a.s. because $c_1'Y_{1t} = 0$ a.s.. Thus, $E\eta_t^2 = E\{\eta_t[(-\lambda_{10}'Y_{1t-1} - \sum_{t=2}^{p} c_{1i}y_{t-i+1})/\sqrt{\lambda_{20}'Y_{2t-1}}]\} = 0$ since η_t is independent of $\mathcal{F}_{t-1}$, which is a contradiction with $\eta_t \sim N(0,1)$. Thus, $c_1 = 0$. Similarly, we can show that $c_2 = 0$.

As in (B.3), we can show that

$$\begin{aligned} El_t(\lambda) &= -\frac{1}{2}E\Big[\ln(\lambda_2'Y_{2t-1}) + \frac{(y_t-\lambda_1'Y_{1t-1})^2}{\lambda_2'Y_{2t-1}}\Big] \\ &= -\frac{1}{2}\Big\{E\ln(\lambda_2'Y_{2t-1}) + E\Big(\frac{\lambda_{20}'Y_{2t-1}}{\lambda_2'Y_{2t-1}}\Big)\Big\} - \frac{1}{2}E\Big(\frac{[(\lambda_1-\lambda_{10})'Y_{1t-1}]^2}{\lambda_2'Y_{2t-1}}\Big). \end{aligned} \tag{B.5}$$

The second term in (B.5) reaches its maximum at zero, and this occurs if and only if $(\lambda_1-\lambda_{10})'Y_{1t-1} = 0$ a.s., which holds if and only if $\lambda_1 = \lambda_{10}$ by (B.4). The first term in (B.5) is

$$-\frac{1}{2}\{-E\ln M_t + EM_t\} - \frac{1}{2}E\ln(\lambda_{20}'Y_{2t-1}), \tag{B.6}$$

where $M_t = \lambda_{20}'Y_{2t-1}/\lambda_2'Y_{2t-1}$. Note that, for any $x > 0$, $-f(x) \equiv -\ln x + x \geq 1$, and hence $-E\ln M_t + EM_t \geq 1$. When $M_t = 1$, we have $f(M_t) = f(1) = -1$. If $M_t \neq 1$, then $f(M_t) < f(1)$, so $Ef(M_t) \leq Ef(1)$ with equality only if $M_t = 1$ with probability one. Thus, (B.6) reaches its maximum $-1/2 - E\ln(\lambda_{20}'Y_{2t-1})/2$, if and only if $\lambda_2' Y_{2t-1} = \lambda_{20}' Y_{2t-1}$, which holds if and only if $\lambda_2 = \lambda_{20}$ by (B.4). Thus, $El_t(\lambda)$ is uniquely maximized at λ_0.

Lemma B.5. *If the assumption of Theorem* 3.1 *holds, then*

(i) Ω *is finite and positive definite,*

(ii) $\dfrac{1}{\sqrt{n}}\sum_{t=2}^{n}\dfrac{\partial l_t(\lambda_0)}{\partial\lambda} \longrightarrow_{\mathcal{L}} N(0,\Omega)$.

Proof. Since the density of η_t is symmetric, it is easy to show that

$$E\Big[\frac{\partial l_t(\lambda_0)}{\partial\lambda}\frac{\partial l_t(\lambda_0)}{\partial\lambda'}\Big] = \mathrm{diag}\Big\{E\Big(\frac{Y_{1t-1}Y_{1t-1}'}{\lambda_{20}'Y_{2t-1}}\Big), \frac{1}{2}E\Big[\frac{Y_{2t-1}Y_{2t-1}'}{(\lambda_{20}'Y_{2t-1})^2}\Big]\Big\} = \Omega.$$

By Assumption 3.1, we can see that $\|Y_{1t-1}\|^2/\lambda_{20}'Y_{2t-1}$ and $\|Y_{2t-1}\|^2/\lambda_{20}'Y_{2t-1}$ are bounded a.s. by some constant C. Thus, Ω is finite. If Ω is not positive definite, then there exists a nonzero $p\times 1$ constant vector c_1 such that $c_1'Y_{1t} = 0$ a.s. or a nonzero $(p+1)\times 1$ constant vector c_2 such that $c_2'Y_{2t} = 0$ a.s.. But this is impossible by (B.4). Thus, (i) holds. Using the Martingale Central Limit Theorem and the Crámer-Wold device, it is straightforward to show that (ii) holds.

Proof of Theorem 3.1. By Lemma B.3(i) and B.4, we have established all the conditions for consistency in Theorem 4.1.1 in Amemiya (1985), and hence $\hat{\lambda}_n \to \lambda_0$ in probability. By Lemma B.3(ii), we can obtain that $n^{-1}\sum_{t=2}^{n}[\partial^2 l_t(\lambda_n)/\partial\lambda'\partial\lambda]$ converges to Ω in probability for any sequence λ_n such that $\lambda_n \to \lambda_0$ in probability. Furthermore, by Lemma B.5, we have established all the conditions in Theorem 4.1.3 in Amemiya (1985), and hence the conclusion holds.

Acknowledgement

The author greatly appreciates two referees and the Co-Editor H.-C. Ho for their very helpful comments, and thanks the Hong Kong RGC's Competitive Earmarked Research Grants #HKUST602205/05P and #HKUST6428/06H for financial support.

References

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge.

Ango Nze, P. (1992). Criteria of ergodicity for some models with a Markovian representation. *C. R. Acad. Sci. Paris Ser. I Math.*, 1301-1304.

Borkovec, M. (2001). Asymptotic behavior of the sample autocovariance and autocorrelation function of the AR(1) process with ARCH(1). *Bernoulli* **6**, 847-872.

Borkovec, M. and Kluppelberg, C. (2001). The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Ann. Appl. Probab.* **11**, 1220-1241.

Berkes, I., Horvath L. and Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli* **9**, 201-207.

Bougerol, P. and Picard, N. (1992). Strict stationarity of generalized autoregressive processes. *Ann. Probab.* **20**, 1714-1730.

Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econom. Theory* **18**, 17-39.

Chan, N. H. and Tran, L. T. (1989). On the first-order autoregressive process with infinite variance. *Econom. Theory* **5**, 354-362.

Chan, N. H. and Peng, L. (2005). Weighted least absolute deviation estimation for an AR(1) process with ARCH(1) errors. *Biometrika* **92**, 477-484.

Chen, M. and An, H. Z. (1998). A note on the stationarity and the existence of moments of the GARCH model. *Statist. Sinica* **8**, 505-510.

Chen, M. and An, H. Z. (1999). The probabilistic properties of the nonlinear autoregressive model with conditional heteroskedasticity. *Acta Math. Appl. Sinica* (English Ser.) **15**, 9-17.

Chen, M. and Chen, G. (2000). Geometric ergodicity of nonlinear autoregressive models with changing conditional variances. *Canad. J. Statist.* **28**, 605-613.

Cline, D. B. H. and Pu, H. H. (1999). Geometric ergodicity of nonlinear time series. *Statist. Sinica* **9**, 1103-1118.

Davis, R. A., Knight, K. and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Process. Appl.* **40**, 145-180.

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.

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

Francq, C. and Zakoïan, J. M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference* **68**, 145-165.

Francq, C. and Zakoïan, J. M. (2000). Covariance matrix estimation for estimators of mixing weak ARMA models. *J. Statist. Plann. Inference* **83**, 369-394.

Francq, C. and Zakoïan, J. M. (2004). Maximum likelihood estimation of pure GSARCH and ARMA-GARCH processes. *Bernoulli* **10**, 605-637.

Feigin, P. D and Tweedie, R. L. (1985). Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. *J. Time Ser. Anal.* **6**, 1-14.

Guégan, D. and Diebolt, J. (1994). Probabilistic properties of the β-ARCH-model. *Statist. Sinica* **4**, 71-87.

Hall, P. and Yao, Q. W. (2003). Inference in ARCH and GARCH models. *Eonometrica* **71**, 285-317.

Jensen, S. T. and Rahbek, A. (2004). Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica* **72**, 641-646.

Kingman, J. F. C. (1973). Subadditive ergodic theory. *Ann. Probab.* **1**, 883-909.

Li, W. K., Ling, S. and McAleer, M. (2002). A survey of recent theoretical results for time series models with GARCH errors. *J. Economic Survey* **16**, 245-269.

Ling, S. (1999). On the stationarity and the existence of moments of conditional heteroskedastic ARMA models. *Statist. Sinica* **9**, 1119-1130.

Ling, S. (2004). Estimation and testing stationarity for double autoregressive models. *J. Roy. Statist. Soc. Ser. B* **66**, 63-78.

Ling, S. (2007). Self-weighted and local quasi-maximum likelihood estimator for ARMA-GARCH/IGARCH models. *J. Econometrics*. To appear.

Ling, S. and McAleer, M. (2002). Necessary and sufficient moment conditions for the GARCH and Asymmetric Power GARCH models. *Econometric Theory* **18** , 722-729.

Ling, S. and McAleer, M. (2003). Asymptotic theory for a new vector ARMA-GARCH model. *Econom. Theory* **19**, 280-310.

Lu, Z. D. (1998). On the geometric ergodicity of a non-linear autoregressive model with an autoregressive conditional heteroscedastic term. *Statist. Sinica* **8**, 1205-1217.

Lu, Z. D. and Jiang, Z. Y. (2001). L_1 geometric ergodicity of a multivariate nonlinear AR model with an ARCH term. *Statist. Probab. Lett.* **51**, 121-130.

Masry, E and Tjϕtheim, D. (1995). Nonparametric estimation and identification of nonlinear ARCH time series. *Econom. Theory* **11**, 258-289.

Nicholls, D. F. and Quinn, B. G. (1982). Random coefficient autoregressive models: an introduction. *Lecture Notes in Statistics* **11**. Springer-Verlag, New York-Berlin.

Pham, D. T. (1986). Bilinear Markovian representation and bilinear models. *Stochastic Process. Appl.* **20**, 295-306.

Tjϕstheim, D. (1990). Nonlinear time series and Markov chains. *Adv. Appl. Probab.* **22**, 587-611.

Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. The Clarendon Press, Oxford University Press, New York.

Tweedie, R. L. (1983). Criteria for rates of convergence of Markov chains, with application to queuing theory. In *Papers in Probability, Statistics and Analysis* (Edited by J. F. C. Kingman and G. E. H. Reuter). Cambridge University Press, Cambridge.

Tsay, R. S. (1987). Conditional heteroscedastic time series models. *J. Amer. Statist. Assoc.* **82**, 590-604.

Weiss, A. A. (1984). ARMA models with ARCH errors. *J. Time Ser. Anal.* **5**, 129-43.

Weiss, A. A. (1986). Asymptotic theory for ARCH models: estimation and testing. *Econom. Theory* **2**, 107-131.

Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong.

E-mail: maling@ust.hk

(Received August 2004; accepted April 2006)