Geometric Ergodicity of a Vector Non-linear AR Model Mixed with an ARCH Term

Zudi Lu

Institut de Statistique
Universite Catholique de Louvain
Voie du Roman Pays 20
B-1348, Louvain-La-Neuve
Belgium

Abstract: In this note, the sufficient condition for the geometric ergodicity of a vector-valued nonlinear AR model mixed with an ARCH term (also called CHARN model by Härdle and his coauthors (e.g. Härdle, Tsybakov and Yang (1996)) is investigated. The main result obtained, to some large degree, relaxes the restrictive limitation on the order of the model and improves the result in Ango Nze (1992) for this model, which was adopted by Härdle and his coauthors in a series of papers for the statistical inferences. As an application of the main result, the geometric ergodicity of an additive AR model mixed with multiplicative ARCH is also discussed. The results of this paper ensure that the geometric α–mixing stationary solution of the model exists under some mild conditions and hence expand the wide applicability of Härdle's (and others') papers.

Keywords and phrases: autoregression, conditional heteroscedasticity, geometric ergodicity, Markov chain, vector nonlinear AR model with ARCH term.

Short running title: Geometric ergodicity of vector AR with ARCH.

1. Introduction

Consider the vector non-linear autoregressive (AR) model mixed with autoregressive conditional heteroscedastic (ARCH) term (denoted by AR-ARCH):

$$X_t = f(X_{t-1}, \cdots, X_{t-p}) + [H(X_{t-1}, \cdots, X_{t-q})]\varepsilon_t, \tag{1.1}$$

where $f : R^{mp} \to R^m$ is a Borel measurable vector-valued function on R^{mp} and $H : R^{mq} \to R^m \times R^m$ is a positive-definite Borel measurable matrix-valued function on R^{mq}, and where $\{X_t\}$ and $\{\varepsilon_t\}$ are two sequences of m-dimensional random vectors defined on the same probability space $(\Omega, \mathcal{F}, \mathcal{P})$, and $\{\varepsilon_t\}$ consists of i.i.d. random variables with $E\varepsilon_t = 0 \in R^m$, $E|\varepsilon_t| < \infty$ for which ε_t is independent of $\{X_s,\ s < t\}$, where throughout the paper, we all use $|\cdot|$ denote the absolute value for number, the norms for vector and matrix, respectively, defined by $|U| = \sum_{i=1}^{p} |u_i|$ for $U = (u_1, \cdots, u_p)' \in R^p$ and $|V| = \sum_{i=1}^{m} \sum_{j=1}^{p} |v_{ij}|$ for $V = (v_{ij})_{m\times p} \in R^m \times R^p$. This is also called conditional heteroscedastic autoregressive nonlinear model (denoted CHARN) by a series of papers of Härdle and his coauthors (e.g., Härdle and Tsybakov (1996), Härdle, Tsybakov and Yang (1996), Härdle and Yang (1996), Yang and Härdle (1996)). Clearly this model is quite general; it not only includes the usual non-linear autoregressive model (Tong (1990)) but also the general ARCH model proposed first by Engle (1982). It has the advantages of both the AR model and the ARCH model: the conditional mean f describes the prediction value and the conditional variance H measures the risk (volatility) of this prediction based on the past information; hence it is extensively applied in Econometrics, particularly for financial time series (see, for example, Bossaerts, Härdle and Hafner (1996)).

So far, there have been many works concerned with model (1.1). On the nonparametric statistical inferences of the functions f and H, Auestad and Tjϕstheim (1990), Mckeague and Zhang (1994), Masry and Tjϕstheim (1995), a series of papers of Härdle and his coauthors (see references stated above), Lu and Cheng (1995), Lu (1996) and the references therein contain a lot of detailed discussions for the different nonparametric methods. In those works, the asymptotic properties, such as (strong) consistency and asymptotic normality of the nonparametric estimators of the functions f and H have been studied based on a preliminary probabilistic property — geometric ergodicity (see (1.3) for the definition) of model (1.1). Hence the (geometric) ergodicity is very important for the statistical inferences of the model.

Concerning the (geometric) ergodicity of model (1.1), there have also been some publications in the literature. For the scalar case ($m = 1$), if $H \equiv \sigma$ (positive constant), that is (1.1) is an AR model, see Chan and Tong (1985), Tjϕstheim (1990, 1994) and An and Huang (1994) which gave many sufficient conditions; if $f \equiv 0$, that is (1.1) is an ARCH, see Guegan and Diebolt (1994), Lu (1994, 1995a) which considered a lot of ARCH type models and obtained general theorems that cover many ARCH type models which have appeared in the econometric literature; if $p = q = 1$, H is bounded away from zero and ∞, Bhattacharya and Lee (1995) conducted a detailed study of the geometric ergodicity for every case of the limit of $f(x)/x$ as $x \to \pm\infty$; for the general p and q, if $E\varepsilon_t^2 < \infty$, Lu (1995b) rectified the vacuous condition of Masry and Tjϕstheim (1995) and did obtain a non-vacuous and mild condition. Of course, these models considered are all the scalar cases of model (1.1). For the general vector case, Ango Nze (1992), assuming $E|\varepsilon_t| < \infty$ (first absolute moment), studied the geometric ergodicity of (1.1), but his condition seems quite limited, which we specify in a bit detail as follows.

Without loss of generality, we may assume that $p = q$ in (1.1) in the following (otherwise, take $\max(p, q)$ as the order of the model). Denote $x = (x_1', \cdots, x_p')' \in R^{mp}$, $x_j \in R^m$. Throughout the paper, y' denotes the transpose of a vector or matrix, y.

(A1) The error variables ε_t, are i.i.d., $E|\varepsilon_t| < \infty$. The density $\psi(\cdot)$ of ε_t exists and satisfies $\inf_{x_1 \in \mathcal{K}} \psi(x_1) > 0$ for any compact $\mathcal{K} \subset R^m$.

(A2) There exist constants $C_1 > 0$, $C_2 > 0$ such that

$$|f(x)| \leq C_1(1 + \sum_{j=1}^{p} |x_j|), \qquad |H(x)| \leq C_2(1 + \sum_{j=1}^{p} |x_j|). \tag{1.2}$$

(A3) The matrix function $H(x)$ is symmetric for any $x \in R^{mp}$, and satisfies $\inf_{x \in \mathcal{K}} \lambda_{\min} \{H(x)\} > \lambda_{\mathcal{K}} > 0$ for any compact $\mathcal{K} \subset R^{mp}$, where $\lambda_{\min}\{H\}$ denotes the minimal eigenvalue of a real symmetric matrix H.

(A4) $C_1 + C_2 E|\varepsilon_1| < 1/p$.

Theorem 0 ($p = q$ in (1.1))(i.e. Proposition 3 of Ango Nze (1992)) *Under the conditions (A1) to (A4) the Markov chain $\{Y_t\}$ is geometrically ergodic, i.e. there exists a probability measure π on $(R^{mp}, \mathcal{B}^{mp})$, a positive constant $\rho < 1$, and a π-integrable non-negative measurable function $J(x)$ such that as $n \to \infty$*

$$||P^n(x, \cdot) - \pi(\cdot)||_\tau \leq \rho^n J(x), \qquad x \in R^{mp}, \tag{1.3}$$

where $Y_t = (X_t', X_{t-1}', \cdots, X_{t-p+1}')'$, $P^n(\cdot, \cdot)$ *is the n–step transition probability of* $\{Y_t\}$ *and* $||\cdot||_\tau$ *denotes the total variation norm.*

We remark here that so far as the model is considered alone, this theorem considered the vector case and hence is general, but the conditions (A2) and (A4) are clearly quite limited. If the order p of the model is large (of course the number of the variables involved in (1.1) may be small; see Tjϕstheim (1994), Tjϕstheim and Auestad (1994)), then the right-hand side of the inequality in (A4) may be close to 0, and hence so may two constants C_1 and C_2. Thus by (1.2), $f(x)$ and $H(x)$ in (1.1) are also close to 0, which is obviously unsatisfactory.

In a series of papers of Härdle and his coauthors, the conditions (A1) to (A4), and this theorem were adopted frequently to ensure that the α-mixing stationary solution of model (1.1) exists with an exponential mixing rate. Hence it is clear that the conditions (A2) and A(4) confine the wide applicability of the inference results of Härdle and his coauthors. This motivates me to consider the geometric ergodicity of model (1.1) once again.

In this paper, we will follow the fashion of Lu (1995b), but since, if ε_t has second order moment, the result of Theorem 1 of Lu (1995b) is easily extended to the vector case, hence we mainly consider the case that only the first absolute moment of ε_t is assumed, which is the case Ango Nze considered. In Section 2, we'll state the main result, in which conditions (A2) and (A4) are relaxed, to some large degree, to the following two mild conditions:

(A2') There exist constants $a_{ij} \geq 0$, $b_{ij} \geq 0$, $1 \leq i \leq m$, $1 \leq j \leq p$, such that, as $|x| \to \infty$,

$$|f(x)| \leq \sum_{i=1}^{m} \sum_{j=1}^{p} a_{ij}|x_{ij}| + o(|x|), \qquad |H(x)| \leq \sum_{i=1}^{m} \sum_{j=1}^{p} b_{ij}|x_{ij}| + o(|x|). \quad (1.4)$$

(A4') $\max_{1\leq i \leq m}\{\sum_{j=1}^{p} a_{ij} + E|\varepsilon_1| \sum_{j=1}^{p} b_{ij}\} < 1$.

Also, condition (A3) may be relaxed a little as shown in Lu (1995b).

(A3') The matrix function $H(x)$ is symmetric for any $x \in R^{mp}$, and satisfies either (1) $\inf_{x\in\mathcal{K}} \lambda_{\min}\{H(x)\} > \lambda_{\mathcal{K}} > 0$ for any compact $\mathcal{K} \subset R^{mp}$, or (2) $\lambda_{\min}\{H(x)\} > 0$ for any $x \in R^{mp}$ and $H(x)$ is continuous on R^{mp}.

Section 3 will give some applications and example, in which the geometric ergodicity of an additive AR model mixed with multiplicative ARCH will be particularly discussed. The proofs of the results will be deferred in Section 4.

2. Markovian representation and main result

For model (1.1) ($p = q$), let

$$Y_t = (X_t', X_{t-1}', \cdots, X_{t-p+1}')';$$

$$T_1(Y_{t-1}, \varepsilon_t) = f(Y_{t-1}) + (H(Y_{t-1}))\varepsilon_t;$$

$$T_i(Y_{t-1}, \varepsilon_t) = X_{t-i+1}, i = 2, 3, \cdots, p,$$

$$T(Y_{t-1}, \varepsilon_t) = ((T_1(Y_{t-1}, \varepsilon_t))', (T_2(Y_{t-1}, \varepsilon_t))', \cdots, (T_p(Y_{t-1}, \varepsilon_t))')',$$

then model (1.1) can be expressed in the state equation as

$$X_t = \gamma Y_t, \tag{2.1a}$$

$$Y_t = T(Y_{t-1}, \varepsilon_t), \tag{2.1b}$$

where $\gamma = (I_m, 0, \cdots, 0) \in R^m \times R^{mp}$, I_m is an identity matrix of order m. Since ε_t is independent of Y_{t-1}, obviously, it follows from (2.1b) that $\{Y_t\}$ is a Markov chain whose state space is R^{mp}, the mp-dimensional Euclidean space.

Under conditions (A1) and (A3'), we may prove that the Markov chain $\{Y_t\}$ is irreducible with respect to the Lebesgue measure μ_{mp} of R^{mp} and aperiodic. (For the definitions of irreducibility and aperiodicity of Markov chain, see Chan (1990), Nummelin (1984)).

Lemma 1. Under conditions (A1) and (A3'), the Markov chain $\{Y_t\}$ defined by (2.1b) is μ_{mp}-irreducible and aperiodic.

The following property is also necessary for the proof of Theorem 1 below. We may prove that for the Markov chain $\{Y_t\}$, any bounded set K in $\mathcal{B}^{mp}$ with a positive μ_{mp}-measure is a small set. (For the definition of the small set, also see Chan (1990) or Nummelin (1984).) That is the following

Lemma 2. If the conditions of Lemma 1 hold, then any bounded set K in R^{mp} with positive Lebesgue measure is a small set of the Markov chain $\{Y_t\}$.

The main purpose of this paper is to study the geometric ergodicity of model (1.1) under conditions (A1), (A2') to (A4'). The following theorem is just what we need, which to some large degree improves the result of Ango Nze (1992).

Theorem 1. *Under conditions (A1) and (A2') to (A4'), the Markov chain $\{Y_t\}$ in (2.1b) defined by (1.1) ($p = q$) is geometrically ergodic (hence so is model (1.1)) and $E_\pi|X_t| < \infty$, where π is the limiting probability measure (see (1.3)).*

At the end of this section we point out that in Theorem 1, only the first absolute moment of ε_t is assumed, and hence $E_\pi|X_t| < \infty$ is obtained; similarly, if $E\varepsilon_{ti}^2 < \infty$, $i = 1, \cdots, m$, then the geometric ergodicity of model (1.1) and the second order moment of X_t with respect to the limiting probability measure π may be obtained under the analogous conditions of Theorem 1 of Lu (1995b), that is a vector version of Theorem 1 of Lu (1995b) (which is easily written out and hence omitted here).

3. Some applications and example

In this section we will give some corollaries as an application of Theorem 1 and make some simulation to the α–mixing stationary solution of an example to illustrate the realization of the solution explicitely.

In Theorem 1, if $a_{ij} = a_j$, $i = 1, \cdots, m$; $j = 1, \cdots, p$, and similarly for b_{ij}, then Theorem 1 becomes the following

Corollary 1. *Under the conditions of Theorem 1 with (A2') and (A4') replaced by conditions (A2") and (A4") as follows, respectively,*

(A2") There exist constants $a_j \geq 0$, $b_j \geq 0$, $1 \leq j \leq p$, such that, as $|x| \to \infty$,

$$|f(x)| \leq \sum_{j=1}^{p} a_j|x_j| + o(|x|), \qquad |H(x)| \leq \sum_{j=1}^{p} b_j|x_j| + o(|x|), \tag{3.1}$$

(A4") $\sum_{j=1}^{p} a_j + E|\varepsilon_1| \sum_{j=1}^{p} b_j < 1$,

the conclusion of Theorem 1 holds.

Furthermore, if all a_j's are equal to each other and similarly for b_j's in Corollary 1, we have

Corollary 2. *Under the conditions of Theorem 1 with (A2') and (A4') replaced by conditions (A2"') and (A4"') as follows, respectively,*

(A2"') There exist constants $a \geq 0$, $b \geq 0$, such that, as $|x| \to \infty$,

$$|f(x)| \leq a(\sum_{j=1}^{p} |x_j|) + o(|x|), \qquad |H(x)| \leq b(\sum_{j=1}^{p} |x_j|) + o(|x|), \tag{3.2}$$

(A4"') $a + bE|\varepsilon_1| < 1/p$,

the conclusion of Theorem 1 holds.

Remark 1. Clearly, Proposition 3 of Ango Nze (i.e. Theorem 0 above) may be easily derived from Corollary 2, which shows that our Theorem 1 extends the result of Ango Nze (1992) largely. Also, the result of Ango Nze is limited to apply for large p (see introduction above). However, just in such case (p large), the additive model for $f(x)$ and the multiplicative model for $H(x)$ are of especial interests to avoid "the curse of dimensionarity" for nonparametric inferences, which was stressed by Härdle and his coauthors (e.g. Yang and Härdle (1996), in which the conditions (A1) to (A4) were adopted, but in fact only the geometric ergodicity is needed). Hence the results of this paper and Lu (1995b) would expand the applicability of a series of papers by Härdle and his coauthors.

To more precisely specify the usefulness of Theorem 1 to the additive and multiplicative model, we discuss the geometric ergodicity of the model more specifically as follows.

Consider the following model (for simplicity we only consider the scalar case, $m = 1$):

$$X_t = f_1(X_{t-1}) + \cdots + f_p(X_{t-p}) + \varepsilon_t[h_1(X_{t-1}) \cdots h_q(X_{t-q})], \tag{3.3}$$

where $f_i : R^1 \rightarrow R^1$ is a Borel measurable function on R^1 and $h_j : R^1 \rightarrow R^1$ is a positive Borel measurable function on R^1, and the others are defined analogously as in (1.1).

We are interested in the question of under what form for the f_i's and h_j's the model (3.1) is geometrically ergodic. For the f_i's, it is natural to consider the linearity form at infinity as usual, that is

$$f_i(x_i) = c_i x_i + o(|x_i|) \quad (|x_i| \rightarrow \infty), \tag{3.4}$$

where c_i is a real number; for the ARCH part, we consider the form discussed in Guegan and Diebolt (1994), that is $h_i(x_i) \sim d_i|x_i|^{\beta_i}$ $(|x_i| \rightarrow \infty)$ with $d_i > 0$ and $\beta_i \geq 0$, which may be expressed in a slightly extensive form

$$h_i(x_i) \leq d_i|x_i|^{\beta_i} + o(|x_i|^{\beta_i}), \quad (|x_i| \rightarrow \infty) \tag{3.5}$$

where $d_i > 0$ and $\beta_i \geq 0$ are two real constants. The following theorem gives some sufficient conditions for the geometric ergodicity of model (3.3).

Theorem 2. ($p = q$ in (3.3)) *Under the conditions (A1) and (A3'), if (3.4) and (3.5) hold, and either one of the following two conditions is also satisfied:*

(1) $\sum_{i=1}^{p} \beta_i < 1$, and all the roots of the polynomial $z^p - c_1 z^{p-1} - \cdots - c_p$ lie in the unit circle;

(2) $\sum_{i=1}^{p} \beta_i = 1$, and $\sum_{i=1}^{p} |c_i| + \prod_{j=1}^{q} d_j E|\varepsilon_t| < 1$,

then the result of Theorem 1 holds.

Remark 2. When $\sum_{i=1}^{p} \beta_i > 1$, if $p = 1$, Guegan and Diebolt (1994) pointed out that the model is not geometrically ergodic. Hence the cases (1) and (2) in Theorem 2 are of particular interests for statistical inferences. The vector version of Theorem 2 is easily written out and is omitted.

At the end of this section, we consider an explicit example of nonlinear time series model as an application of Theorem 2.

Example 1. Let $\{X_t\}$ satisfy

$$X_t = \frac{1}{2}X_{t-1} + \sin(X_{t-2}) + \cos(X_{t-10}) + \varepsilon_t(1 + |\sin(X_{t-3})|)(1 + \alpha|X_{t-5}|^{\beta}), \tag{3.6}$$

where $\alpha > 0$ and $\beta \geq 0$ are two constants, and ε_t is defined as in (A1). This is an additive AR with multiplicative ARCH model. If $\beta < 1$, by (1) of Theorem 2, the model (3.6) is geometrically ergodic; if $\beta = 1$, and $1/2 + 2\alpha E|\varepsilon_t| < 1$, that is $\alpha E|\varepsilon_t| < 1/4$, then the geometric ergodicity of the model also follows from (2) of Theorem 2. Clearly, since $p = 10$ in (3.6) and $1/2 > 1/10$, it is impossible to deduce the geometric ergodicity by the result of Ango Nze (1992) (i.e. Theorem 0 above). The geometric α-mixing stationary solutions to the model (3.6) are shown in Figures 1 and 2 ($\varepsilon_t \sim N(0, 1)$), from which the high changing volatility could be clearly seen.

4. Proofs

Lemmas 1 and 2 may be proved analougously as Chan and Tong (1985), Ango Nze (1992) and Lu (1995b) did and hence the proofs are omitted.

The following lemma is crucial in the proof of the main result, Theorem 1, which is usually called the drift criterion for geometric ergodicity.

Lemma 3. Let $\{Y_t\}$ be an aperiodic and irreducible Markov chain. Suppose that there exists a small set C, a non-negative measurable function g, and constants $0 < r < 1$, $\gamma > 0$, and $B > 0$ such that

$$E(g(Y_{t+1})|Y_t = y) < rg(y) - \gamma, \ y \notin C; \tag{4.1}$$

$$E(g(Y_{t+1})|Y_t = y) < B, \ y \in C. \tag{4.2}$$

Then $\{Y_t\}$ is geometrically ergodic. And if g is also bounded away from 0 and ∞ on C, then $E_\pi g(Y_t) < \infty$, where E_π means the expectation with respect to the limiting probability measure π in (1.3).

Proof. For the proof of this lemma, see Numellin (1984) and Tweedie (1975, 1976, 1983).

Proof of Theorem 1. By Lemma 1, $\{Y_t\}$ is irreducible and aperiodic, so, to prove the geometric ergodicity, we only need to find a non-negative test function $g(x)$ for which the conditions of Lemma 3 hold.

For $y = (y_1', \cdots, y_p')' \in R^{mp}$, $y_j = (y_{1j}, \cdots, y_{mj})' \in R^m$, $j = 1, \cdots, p$, take the test function $g(y)$ as

$$g(y) = 1 + \sum_{i=1}^{m} |y_{i1}| + \sum_{i=1}^{m}\sum_{j=1}^{p-1} r_{ij}|y_{i,j+1}|$$

where r_{ij}, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, p-1$, are positive constants and to be specified later.

By (1.1) and (1.4), as $|y| \to \infty$,

$$\begin{aligned}
E(g(Y_{t+1})|Y_t = y) \quad &= 1 + E(|X_{t+1}||Y_t = y) + \textstyle\sum_{i=1}^{m}\sum_{j=1}^{p-1} r_{ij}|y_{ij}| \\
&= 1 + |f(y)| + |H(y)|E|\varepsilon_t| + \textstyle\sum_{i=1}^{m}\sum_{j=1}^{p-1} r_{ij}|y_{ij}| \\
&\leq 1 + (\textstyle\sum_{i=1}^{m}\sum_{j=1}^{p} a_{ij}|y_{ij}| + E|\varepsilon_t|\sum_{i=1}^{m}\sum_{j=1}^{p} b_{ij}|y_{ij}|)+ \\
&\quad + \textstyle\sum_{i=1}^{m}\sum_{j=1}^{p-1} r_{ij}|y_{ij}| + o(|y|) \\
&= 1 + \textstyle\sum_{i=1}^{m}\sum_{j=1}^{p-1}(a_{ij} + E|\varepsilon_t|b_{ij} + r_{ij})|y_{ij}| \qquad\qquad (4.3) \\
&\qquad + \textstyle\sum_{i=1}^{m}(a_{ip} + E|\varepsilon_t|b_{ip}+)|y_{ip}| + o(|y|) \\
&= 1 + \textstyle\sum_{i=1}^{m}(a_{i1} + E|\varepsilon_t|b_{i1} + r_{i1})|y_{i1}| \\
&\qquad + \textstyle\sum_{i=1}^{m}\sum_{j=2}^{p-1} \frac{a_{ij}+E|\varepsilon_t|b_{ij}+r_{ij}}{r_{i,j-1}} r_{i,j-1}|y_{ij}| \\
&\qquad + \textstyle\sum_{i=1}^{m} \frac{a_{ip}+E|\varepsilon_t|b_{ip}}{r_{i,p-1}} r_{i,p-1}|y_{ip}| + o(|y|)
\end{aligned}$$

If we choose the positive constants r_{ij}'s satisfying

$$\begin{aligned}
&a_{i1} + b_{i1}E|\varepsilon_t| + r_{i1} < 1, \\
&\tfrac{a_{ij}+b_{ij}E|\varepsilon_t|+r_{ij}}{r_{i,j-1}} < 1,\ j = 2, 3, \cdots, p-1, \\
&\tfrac{a_{ip}+b_{ip}E|\varepsilon_t|}{r_{i,p-1}} < 1,\ i = 1, 2, \cdots, m,
\end{aligned}$$

that is,

$$\begin{aligned}
&a_{ip} + b_{ip}E|\varepsilon_t| < r_{i,p-1} < 1 - \textstyle\sum_{j=1}^{p-1}(a_{ij} + b_{ij}E|\varepsilon_t|), \\
&a_{i,j+1} + b_{i,j+1}E|\varepsilon_t| + r_{i,j+1} < r_{i,j} < 1 - \textstyle\sum_{k=1}^{i}(a_{ik} + b_{ik}E|\varepsilon_t|), \\
&\qquad j = 1, 2, \cdots, p-2,\ i = 1, 2, \cdots, m,
\end{aligned}$$

(Since (A4') holds, such choosing r_{ij}'s is possible.) and set

$$\begin{aligned}\rho &= \max\{a_{i1}+b_{i1}E|\varepsilon_t|+r_{i1},\ \tfrac{a_{ij}+b_{ij}E|\varepsilon_t|+r_{ij}}{r_{i,j-1}},\ j=2,\,3,\,\cdots,\,p-1,\\ &\quad \tfrac{a_{ip}+b_{ip}E|\varepsilon_t|}{r_{i,p-1}},\ i=1,\,2,\,\cdots,\,m\}\end{aligned}$$

then $0<\rho<1$,

$$\begin{aligned}E(g(Y_{t+1}|Y_t=y) \quad &\le \rho(1+\textstyle\sum_{i=1}^m|y_{i1}|+\sum_{i=1}^m\sum_{j=1}^{p-1}r_{ij}|y_{i,j+1}|)+o(|y|)\\ &= (\rho+o(1))[1+\textstyle\sum_{i=1}^m|y_{i1}|+\sum_{i=1}^m\sum_{j=1}^{p-1}r_{ij}|y_{i,j+1}|]-1,\end{aligned} \tag{4.4}$$

where $o(1)\to 0$ as $|y|\to\infty$. Choose $c>0$ large enough, such that

$$\rho+o(1)<r_0<1, \quad as\ |y|>c.$$

Set $C=\{|y|\le c\}$, then C is the bounded set with $\mu_{mp}(C)>0$, which shows from lemma 2 that C is the small set of $\{Y_t\}$. Then (4.1), (4.2) obviously hold. Hence by Leema 3, $\{Y_t\}$ is geometrically ergodic (so is $\{X_t\}$) and $E_\pi|X_t|<\infty$.

Proof of Corollaries 1 and 2. By Theorem 1, Corollaries 1 and 2 are obvious.

Proof of Theorem 2. If $\sum_{i=1}^p\beta_i<1$, it is obvious that

$$h_1(x_1)\cdots h_q(x_q)=o(|x|), \quad (|x|\to\infty),$$

by which (1) of Theorem 2 easily follows from the non-vacuous part of Lemma 3.1 of Masry and Tjϕstheim (1995) and Theorem 2 of Lu (1995b).

If $\sum_{i=1}^p\beta_i=1$, then by (3.4),

$$h_1(x_1)\cdots h_q(x_q)\le\prod_{i=1}^q(d_i|x_i|^{\beta_i}+o(|x_i|^{\beta_i}))\le(\prod_{i=1}^q d_i)(\sum_{i=1}^q\beta_i|x_i|)+o(|x|), \quad (|x|\to\infty),$$

where in the last inequality, the following inequality is used:$\prod_{i=1}^q|x_i|^{\beta_i}\le\sum_{i=1}^q\beta_i|x_i|$, since $\sum_{i=1}^p\beta_i=1$. Thus, (2) of Theorem 2 also easily follows from Theorem 1 above.

Acknowledgements: I am grateful very much to Prof. L. Simar for his encouraging comments and pointing out some errors. The support of Institut de Statistique, Universite Catholique de Louvain, is greatly acknowledged.

REFERENCES

[1] An, H. Z. and Huang, F. C. (1994) The geometric ergodicity of nolinear autoregressive model. Technical Report, Institute of Applied Mathematics, Academia Sinica.

[2] Ango Nze, P. (1992) Critéres d'ergodicité de quelques modéles à représentation markovenne, *C. R. Acad. Sci. Paris,* 315, sér 1, 1301–1304.

[3] Auestad, B. and Tjϕstheim, D. (1990) Identification of nonlinear time series: First order characterization and order determination. *Biometrika*, 77, 669–687.

[4] Bhattacharya, R. N. and Lee, C. (1995) Ergodicity of nonlinear first order autoregressive models. *Journal of Theoretical Probability*, 8, 207–219.

[5] Bossaerts, P., Härdle, W. and Hafner, C. (1995) A new method for volatility estimation with application in foreign exchange series, discussion paper, Institut für Statistik und Öconometrie, Wirtschaftswissenschaftliche Fakultät, Humboldt Universität zu Berlin.

[6] Chan, K. S. (1990) Deterministic stability, stochastic stability, and ergodicity. ed. Tong H., 1990, *Nonlinear Time Series:a dynamical system approach*, 448-466, Oxford University Press, Oxford.

[7] Chan, K. S. and Tong, H. (1985) On the use of the deterministic Lyapunov fuction for the ergodicity of stochastic difference equation. *Advances of Applied Probability*, 17, 666–678.

[8] Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of United Kingdom inflation. *Econometrica*, 50, 987-1007.

[9] Guegan, D. and Diebolt, J. (1994) Probabilistic properties of the β–ARCH model. *Statistica Sinica*, 4, 71-87.

[10] Härdle, W. and Tsybakov, A. (1996) Local polynomial estimation of the volatility function in nonparametric autoregression, to appear in *Journal of Econometrics*.

[11] Härdle, W., Tsybakov, A. and Yang, L. (1996) Nonparametric vector autoregression, to appear in *Journal of Statistical Planning and Inference*.

[12] Härdle, W. and Yang, L. (1996) Nonparametric time series model selection, discussion paper, Institut für Statistik und Öconometrie, Wirtschaftswissenschaftliche Fakultät, Humboldt Universität zu Berlin.

[13] Lu, Z. D. (1994) A note on geometric ergodicity of autoregressive conditional heteroscedasticity (ARCH) model. To appear in *Statistics and Probability Letters*.

[14] Lu, Z. D. (1995a) Two general theorems on geometric ergodicity of a general ARCH type model with applications to some typical models. Submitted.

[15] Lu, Z. D. (1995b) On geometric ergodicity of a non-linear autoregressive (AR) model with an autoregressive conditional heteroscedastic (ARCH) term. Submitted.

[16] Lu, Z. D. and Cheng P. (1995) Distribution–free Strong Consistency for Nonparametric Kernel Regression Involving Nonlinear Time Series. Submitted.

[17] Lu, Z. D. (1996) Some topics for nonlinear time series—Geometric ergodicity and nonparametric identification, Ph.D dissertation, Institute of Systems Science, Academia Sinica, Beijing.

[18] Masry, E. and Tjϕstheim (1995) Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality. *Econometric Theory*, 11, 258–289.

[19] Mckeague, I. W. and Zhang, M. J. (1994) Identification of nonlinear time series from first order cumulative characteristics. *Annal of Statistics*, 22, 495–514.

[20] Nummelin, E. (1984) *General Irreducible Markov Chains and Non-negative Operations*. Cambridge University Press, Cambridge.

[21] Tjϕstheim, D. (1990) Nonlinear time series and Markov chains. *Advances in Applied Probability*, 22, 587–611.

[22] Tjϕstheim, D. (1994) Nonlinear time series: a selective review. *Scandinavian Journal of Statistics*, 21, 97–130.

[23] Tjϕstheim, D. and Auestad, B. (1994) Non-parametric identification of nonlinear time series: selecting significant lags, *Journal of American Statistical Association*, 89, 1410–1419.

[24] Tong, H. (1994) *Nonlinear time series: a dynamical system approach.* Oxford University Press, Oxford.

[25] Tweedie, R. L. (1975) Sufficient conditions for ergodicity and recurrence of Markov chain on a general state space. *Stochastic Processes and Their Applications*, 3, 385–403.

[26] Tweedie, R. L. (1976) Criteria for rates of convergence of Markov chains with applications to queueing theory. *Papers in Probability,Statistics and Analysis*, J. K. C. Kingman and G.E.H. Reuter (eds.), 260–276, Cambridge University Press, Cambridge.

[27] Tweedie, R. L. (1983) The existence of moments for stationary Markov chains. *Journal of Applied Probability*, 29, 191-196.

[28] Yang, L. and Härdle, W. (1996) Nonparametric autoregression with multiplicative volatility and additive mean, discussion paper, Institut für Statistik und Öconometrie, Wirtschaftswissenschaftliche Fakultät, Humboldt Universität zu Berlin.