Journal of the
Nigerian Mathematical Society

Vol. 36, pp. 307-317, 2017
©Nigerian Mathematical Society

ON THE GEOMETRIC ERGODICITY OF THE MIXTURE AUTOREGRESSIVE MODEL

M. I. AKINYEMI[1] AND G. N. BOSHNAKOV

ABSTRACT. Geometric ergodicity is very useful in establishing mixing conditions and central limit results for parameter estimates of a model. It also justifies the use of laws of large numbers and forms part of the basis for exploring the asymptotic theory of a model.
The class of mixture autoregressive (MAR) models provides a flexible way to model various features of time series data and is well suited for density forecasting. The MAR models are able to capture many stylised properties of real data, such as multimodality, asymmetry and heterogeneity. We show here that the MAR model is geometrically ergodic and by implication satisfies the absolutely regular and strong mixing conditions.

Keywords and phrases: mixture autoregressive model, MAR model, geometric ergodicity, mixing conditions, drift condition
2010 Mathematical Subject Classification: 37A25

1. Introduction

Geometric ergodicity is very useful in establishing mixing conditions and central limit results for parameter estimators of a model. It also justifies the use of laws of large numbers and forms a basis for exploring the asymptotic theory of the model. This further translates into examining the consistency and asymptotic normality of the parameter estimates of the model [22]. Detailed discussions on geometric ergodicity and mixing conditions are given by [16], [21], [23], [5]. Furthermore, [20] provide criteria for judging the strong ergodicity of regime-switching diffusion processes. They considered processes in one dimensional space and in multidimensional space separately.[12] stated sufficient conditions for simultaneous geometric ergodicity of Markov chain classes. In particular, they deal with non asymptotic computable bounds for the geometric convergence rate of homogeneous ergodic Markov processes. [19] derived sufficient conditions for geometric ergodicity of a general class of asymmetric nonparametric stochastic processes with stochastic volatility models with skewness driven by the hidden Markov Chain with

Received by the editors May 28, 2015; Revised: January 10, 2017; Accepted: March 16, 2017
www.nigerianmathematicalsociety.org

[1]Corresponding author

switching.[9] showed that the Gibbs sampler of the Poisson change-point model is geometrically ergodic. They prove that the transition kernel is a trace-class operator, which implies geometric ergodicity of the sampler and apply the sampler to a model for the quarterly driver fatality counts for the state of Victoria, Australia. [18] showed the geometric ergodicity of the block Gibbs Markov chain. They showed under mild conditions that when a Bayesian version of the general linear mixed model is created by adopting a conditionally conjugate prior distribution, a simple block Gibbs sampler can be employed to explore the resulting intractable posterior density. [14] provided explicit connections between the V - geometric ergodicity of Markov kernel P on a measurable space X and that of finite-rank non-negative sub-Markov kernels $\hat{P}_k$ approximating P.

Mixture models have over the years played a significant role in data modelling especially in capturing the dynamics of financial time series [13] studied the class of mixed normal conditional heteroskedastic (MixN- GARCH) models, which couples a mixed normal distributional structure with GARCH-type dynamics.They introduced two different flexible time-varying weight model structures. [6] proposed mixtures of stable Paretian distributions for univariate asset returns. The model lends itself to use in a multivariate context for portfolio selection. They applied the model to out-of-sample risk forecasting exercise for seven major FX and equity indices.[15] proposed the large margin mixture of AR (LMMAR) models. They proposed methods which are applied on the simulated time series data, electrocardiogram data, speech data for E-set in English alphabet and electroencephalogram time series data. [17] considered a class of mixtures of structured autoregressive (AR) models and methods for sequential estimation within the said class. [26] proposed the hierarchical Bayesian information criterion (HBIC) for model selection in finite mixture models.

In this paper we show that the mixture autoregressive (MAR) model is geometrically ergodic and, as a consequence, satisfies the absolutely regular and strong mixing conditions.

The rest of this paper is structured as follows. In Section we describe the class of MAR models and give some assumptions associated with it. In Section we discuss the concepts of geometric ergodicity and mixing, as well as the relationship between them. Finally, we show the geometric ergodicity of the MAR model in Section .

2. Mixture Autoregressive (MAR) Model

The mixture autoregressive model of [24] is defined as follows.

Definition 1: A process $\{y_t\}$ is said to be a mixture autoregressive (MAR) process if the conditional distribution function of y_t given past information is given by

$$F_{t|t-1}(x) = \sum_{k=1}^{g} \pi_k F_k \left(\frac{x - \phi_{k,0} - \sum_{i=1}^{p_k} \phi_{k,i} y_{t-i}}{\sigma_k} \right), \tag{0.1}$$

where g is a positive integer representing the number of components in the model and the kth component of the model, for $k = 1, \dots, g$, is specified by its mixing proportion $\pi_k > 0$, scale parameter $\sigma_k > 0$, autoregressive order p_k, intercept $\phi_{k,0}$, autoregressive coefficients $\phi_{k,i}$, $i = 1, \dots, p_k$, and cumulative distribution function $F_k(\cdot)$. The mixing proportions π_k define a discrete distribution $\boldsymbol{\pi}$, so $\sum_{k=1}^{g} \pi_k = 1$.

We denote by MAR$(g; p_1, \dots, p_g)$ a g-component MAR model whose components are of orders $p_1, \dots, p_g$. The noise distribution functions F_k, $k = 1, \dots, g$, are typically taken to be standard Gaussian [24] or (standardised) Student-t [25]. We will denote by $f_k(\cdot)$ the corresponding probability density functions. It is also convenient to set $p = \max_{1 \le k \le g} p_k$ and $\phi_{k,i} = 0$ for $i > p_k$.

We do not discuss estimation theory in this paper but it can be developed under relatively mild conditions, usually met in practice. The noise probability densities, $f_k(.)$, need to be continuous and positive everywhere, non-periodic and bounded on compacts sets for all k. Detailed study of the asymptotic theory is given by [1].

A useful interpretation of the MAR model is that at each time t one of g autoregressive-like equations is picked at random to generate y_t. This can be formalised by writing the MAR model as a random coefficient autoregressive model [4]. Let $\{z_t\}$ be an i.i.d. sequence of random variables with distribution $\boldsymbol{\pi}$ (see Definition), such that $\Pr\{z_t = k\} = \pi_k$, for $k = 1, \dots, g$. Then y_t can be written as

$$y_t = \phi_{z_t,0} + \sum_{i=1}^{p} \phi_{z_t,i} y_{t-i} + \sigma_{z_t} \epsilon_{z_t}(t) \tag{0.2}$$

$$= \mu_{z_t}(y_t) + \sigma_{z_t} \epsilon_{z_t}(t), \tag{0.3}$$

where the noise variables $\{\epsilon_k\}$ are jointly independent and independent of past ys, $\mu_{z_t}(y_t) = \phi_{z_t,0} + \sum_{i=1}^{p} \phi_{z_t,i} y_{t-i}$, and the probability density of $\epsilon_k(t)$ is $f_k(.)$, see [4] for further details see.

The conditional density of y_t given z_t and the past values of both y_t and z_t is

$$f_\theta(y_t \mid z_t, z_s, y_s, s \le t-1) = \frac{1}{\sigma_{z_t}} f_{z_t} \left(\frac{y_t - \phi_{z_t,0} - \sum_{i=1}^{p_{z_t}} \phi_{z_t,i} y_{t-i}}{\sigma_{z_t}} \right). \tag{0.4}$$

Define a vector $Z_t = [Z_{t,1}, \dots, Z_{t,g}]'$, such that $Z_{t,k}$ takes the value one when $z_t = k$ and zero otherwise. Let also $Y_t = (y_t, \dots, y_{t-p+1})'$.

Our results hold under the following somewhat more general assumptions about the process Z_t [1].

Assumptions A

(1) *For each* $k \in \{1, \dots, g\}$, $\{Z_{t,k} : t \geq 0\}$ *is an irreducible, aperiodic Markov chain on a finite space* S *with probability distribution* $\pi_1, \dots, \pi_g$ *and transition probability matrix* $A = (a_{ij})$. Hence, $Z_{t,k}$ inherits the properties of $\{Z_t\}$.

(2) *The chain* $\{Z_t\}$ *is independent of the* ϵ_t *and for all* i, j, $P(z_t = j \mid z_{t-1} = i, \mathcal{F}_{t-1}) = P(z_t = j \mid z_{t-1} = i)$, *where* $\mathcal{F}_{t-1} = \sigma\{Y_r,\ r \leq t-1\}$. This assumption means that the hidden process Z_t is independent of the past observations given its own past, i.e. $Z_{t,1}$.

Assumptions A(1) and A(2) are general conditions for hidden Markov models (see [10] or [11] for examples). Furthermore, we assume for the MAR model defined in Equation (0.2) that:

(3) *The noise* $\{\epsilon_t\}$ *has probability density that is continuous and positive everywhere.*

(4) $f_k(y)$ *is non-periodic and bounded on all compact sets for all* k

Under Assumption A, it is possible to define an aperiodic $S \times \mathbb{R}^p$-valued Markov chain

$$Q_t = (Z_t, Y_t). \tag{0.5}$$

The process $\{Z_t, t > 0\}$ in Equation (0.5) is a simple case of a hidden Markov chain on a finite state space $S = \{0, 1\}$ with stationary k-step transition probability matrix. $\{Z_t, t > 0\}$ drives the dynamics of Y_t.

3. Geometric Ergodicity

The chain $\{Y_t\}$ is called *geometrically ergodic* if there exists a positive constant $\rho < 1$ such that

$$\lim_{t \to +\infty} \rho^{-t} \|p^t(y, \cdot) - \pi(\cdot)\| = 0,\ \forall y \in S. \tag{0.6}$$

Recurrence, existence of an invariant probability measure and φ-irreducibility properties are not generally easily verified for all models. [22] and [16] suggested exploring the use of the *drift condition* for proving geometric ergodicity. The following theorem is by [22, page 591] and [16, page 368].

Theorem 1: (**Geometric ergodicity**) Suppose that the Markov process Y_t is aperiodic and φ-irreducible. Suppose also that there exists a petite set A, positive constants $0 < \rho < 1$, $\varepsilon > 0$, $M < \infty$ and a non-negative measurable function $V \geq 1$ such that:

$$E[V(Y_t) \mid y_{t-1} = y] \leq \begin{cases} \rho V(y) - \varepsilon, & \text{if } y \in A^c, \\ M, & \text{if } y \in A. \end{cases} \tag{0.7}$$

Then Y_t is geometrically ergodic.

The function V is said to be a ***drift criterion*** and is also referred to as a *test function* [23].

3.1. Geometric Ergodicity and β-mixing conditions

[7] and [5] showed that for an ergodic Markov chain Y_t, of invariant probability measure π,

$$\beta_Y(t) = \int \|P^t(y,.) - \pi\|\pi(dy). \tag{0.8}$$

The rate ρ in Equation (0.6) can be chosen independently of the initial point. If Equation (0.6) holds then it follows that $\beta_Y(t) = O(\rho^t)$. Then $\{Y_t\}$ is stationary and geometrically ergodic and hence β-mixing. Thus, β-mixing is a major consequence of geometric ergodicity.

Furthermore, β-mixing implies α-mixing [8]. Therefore, if the Markov chain Y_t is geometrically ergodic, then it is both β-mixing (absolutely regular) and α-mixing (strong mixing) at a geometric rate.

4. Geometric Ergodicity of the MAR model

Let y_t be a MAR process as defined in Definition , We show here that $Y_t = (y_t, \dots, y_{t-p+1})'$ is geometrically ergodic and by implication β-mixing as well as α-mixing, at a geometric rate.

The following assumptions are made in addition to Assumption A in Section .

Assumptions B

(1) *For each $z \in S$, there exist $c_i(z), d_i(z) \in \mathbb{R}^p$, $c_i(z) \geq 0$, $d_i(z) \geq 0$, $i = 1, \dots, p$, such that for $y = (y_1, \dots, y_p)$ the following inequalities hold:*
   (a) $|\mu_{z_t}(y)| \leq \sum_{i=1}^p c_{i(z)}|y_i| + o(\|y\|)$ *as* $\|y\| \to \infty$ *and*
   (b) $\sigma^2_{z_t}(y) \leq \sum_{i=1}^p d_{i(z)}|y_i^2| + o(\|y\|^2)$ *as* $\|y\| \to \infty$.

(2) *The Foster-Lyapunov drift condition* [22, 16] *holds, that is, there exists a real valued measure function $V \geq 1$, such that for some constants $\varepsilon > 0$, $0 < \rho < 1$, M_1 and a small set $A = \{y \in \mathbb{R} : \|y\| \leq M_1\}$, the following holds:*

$$E[V(Q_t) \mid Q_{t-1} = q] \leq \rho V(q) \qquad \text{for } y \in A^c, \tag{0.9}$$

$$\sup_{x \in A} E[V(Q_t) \mid Q_{t-1} = q] < \infty \qquad \text{for } y \in A. \tag{0.10}$$

The following result is useful for the proof of our claim.

Lemma 1:[[16, Theorem 5.5.7]] For an aperiodic, φ-irreducible Markov chain, all petite sets are small sets.

We now prove the following result for the chain $Q_t = (Z_t, Y_t)$ defined in Equation (0.5).

Proposition 1: The Markov chain $Q_t = (Z_t, Y_t)$ is aperiodic and φ-irreducible. Furthermore, for every compact set $C \in \mathbb{R}^p$, $S \times C$ is a small set.

Proof. For ease of notation, we will write the conditional probability $P(Y_t \mid Z_1 = z_1, \dots, Z_t = z_t)$ as $P(Y_t^{(z)})$.

Let $g(q, y \mid z)$ denote the joint density of Q_t, Y_t given Y_t is in state z and let $z = (z_1, \ldots, z_t) \in S'$ and $y = (y_1, \ldots, y_t) \in S$.

Denote the transition probabilities of moving between alternate states of z_t by $p_{z_0 z_1}, \ldots, p_{z_{t-1} z_t} > 0$.

For any compact set $C \subset \mathbb{R}^p$ and A such that $\varphi(A) > 0$, we have by [2, Lemma 1] and [21]

$$\int_A g(q, y \mid z) d\varphi(q) > 0 \qquad \text{and} \qquad \inf_{y \in C} \int_A g(q, y \mid z) d\varphi(q) > 0. \quad (0.11)$$

Hence, it is possible to write

$$P(Y_t^{(z)} \in A \mid Y_{t-p} = q) = \int_A g(q, y \mid z_{t-p+1}, \ldots, z_t) d\varphi(y) > 0, \ \forall q.$$

It then follows that

$$P(Y_t^{(z)} \in A \mid Y_0 = y) > 0 \qquad \text{and} \qquad \inf_{y \in C} P(Y_t^{(z)} \in A \mid Y_0 = y) > 0,$$

which shows that Q_t is φ-irreducible. Furthermore, for any compact set $C \in \mathbb{R}^p$,

$$\inf_{(z_0, y) \in S \times C} \sum_{n=1}^{j} P^n((z_0, y), S' \times A) > 0, \quad (0.12)$$

implying that every compact subset of $S \times C$ is a small set, which completes the proof. □

To verify the geometric ergodicity of the MAR model, we need to:

(1) Prove that the process $Q_t = (Z_t, Y_t)$ is φ-irreducible and aperiodic.
(2) Show the existence of a test function $V(Q_t)$ satisfying the drift condition (Equation (0.9) above).

The two steps are summarized in the following theorem.

Theorem 2: Consider the aperiodic Markov Chain $Q_t = (Z_t, Y_t)$. Let A be a small set and $\{Y_t = (y_t, \ldots, y_{t-p+1})'; t \geq 0\}$ be an aperiodic, φ-irreducible process such that each y_t is a MAR process defined by Equation (0.3). Suppose that Assumption A and Assumption B are satisfied and

$$\sup_z E\left[\sum_{j=1}^{p} c_i(Z_t) c_j(Z_t) + E(\epsilon_{z_t}^2) d_i(Z_t) \mid Z_{t-1} = z\right] < 1. \quad (0.13)$$

Denote by π the unique invariant distribution of Y_t and let $\pi_y(A) = \pi(S \times A \times \mathbb{R}^{p-1})$, $A \in B(\mathbb{R})$.

Then

(1) $\{Y_t; t \geq 0\}$ is geometrically ergodic with $V(y) = 1 + \|y\|^2$,
(2) $\{Y_t; t \geq 0\}$ has a stationary distribution with finite second moments, i.e. $E_{\pi_y}[y_t^2] < \infty$,

(3) $\{Y_t; t \geq 0\}$ is β-mixing and hence strong mixing at geometric rate.

Proof. We start by showing that the drift condition (Equation (0.9)) is satisfied.

Let us choose $\delta > 0$, such that $\sum_{i=1}^{p} \xi_i + \delta = 1$, where

$$\xi_i = \sup_z E[\sum_{j=1}^{p} c_{i(z)} c_{j(z)} \\ + E(\epsilon_t^2) d_{i(z)} \mid Z_{t-1} = z] < 1.$$

Choose $V(z, y) = 1 + \|y^2\|$ to be the test function, $V : S \times R^p \to R$. Since $\xi_i > 0$ and $\sum_{i=1}^{p} \xi_i + \delta = \sum_{i=1}^{p}(\xi_i + \delta/p)$, we also have $\xi_i \leq (1 - \frac{\delta}{p})$ for $1 \leq i \leq p$.

Hence, by Equation (0.9) we have for $y = (y_1, \ldots, y_p)$,

$$\begin{aligned} E[V(Q_t) \mid Q_{t-1} = q] &= E[V(Q_t) \mid Q_{t-1} = (z, y)] \\ &= E[(\mu_{z_t}(y) + \sigma_{z_t}\epsilon_t)^2 \mid Z_{t-1} = z] + 1 \\ &\leq E_z[(\mu_{z_t}(y) + \sigma_{z_t}\epsilon_t)^2] + \sum_{i=2}^{p} y_{i-1}^2 + 1. \end{aligned}$$

Let $\tau_i(z) = \sum_{j=1}^{p} c_{i(z)} c_{j(z)}$. From the last inequality and Assumption B we get

$$\begin{aligned} E[V(Q_t) \mid Q_{t-1} = q] \leq &\sum_{i=1}^{p} E_z[\tau_{i(z)} + E\epsilon_{z_t}^2 d_{i(z)}] y_i^2 + \sum_{i=2}^{p} y_{i-1}^2 \\ &+ E_z[(2o(\|y\| \sum_{i=1}^{p} c_{iz})|y_i|) + (o(\|y\|))^2 + E(\epsilon_t^2)(o(\|y\|^2))^2] + 1. \end{aligned}$$

This can be written more concisely as

$$E[V(Q_t) \mid Q_{t-1} = q] \leq \sum_{i=1}^{p} \xi_i y_1^2 + \sum_{i=2}^{p} y_{i-1}^2 + E_z[L_{z_t}(y)] + 1,$$

where

$$L_{z_t}(y) = (2o(\|y\| \sum_{i=1}^{p} c_{i(z)})|y_i|) + (o(\|y\|))^2 + E(\epsilon_t^2)(o(\|y\|^2))^2. \quad (0.14)$$

Furthermore, for $\delta > 0$, it follows that

$$\begin{aligned}
E[V(Q_t) \mid Q_{t-1} = q] &\leq y_1(1 - \frac{\delta}{p}) + \sum_{i=2}^{p} y_{i-1}^2 + E_z[L_{z_t}(y)] + 1 \\
&\leq \sum_{i=1}^{p} y_i^2 - \frac{\delta}{p} \sum_{i=1}^{p} y_i^2 + E_z[L_{z_t}(y)] + 1 \\
&\leq \sum_{i=1}^{p} y_i^2 - \frac{\delta}{p} \sum_{i=1}^{p} y_i^2 + E_z[L_{z_t}(y)] + 1 + \frac{\delta}{p} - \frac{\delta}{p} \\
&= (1 + \sum_{i=1}^{p} y_i^2) - \frac{\delta}{p}(1 + \sum_{i=1}^{p} y_i^2) + E_z[L_{z_t}(y)] + \frac{\delta}{p} \\
&= V(z, y) - \frac{\delta}{p}(V(z, y)) + E_z[L_{z_t}(y)] + \frac{\delta}{p} \\
&= V(z, y) \left[1 - \frac{\delta}{p} + \frac{1}{V(z, y)} \left[E_z\left[L_{z_t}(y)\right] + \frac{\delta}{p}\right]\right].
\end{aligned}$$

However, $\frac{E_z[L_{z_t}(y)]}{V(z,y)} \to 0$, $\frac{\delta/p}{V(z,y)} \to 0$ as $\|y\| \to \infty$, so that we have

$$E[V(Q_t) \mid Q_{t-1} = q] \leq V(z, y)[1 - \frac{\delta}{p} + \frac{\delta/p}{V(z, y)}] = V(z, y)(1 - \frac{\delta}{p}).$$

Now suppose that $y \in A^c$ and there exists $M_1 > 1$ such that $\|y\| > M_1$ so that $\frac{\delta}{p} < \varepsilon < 1$, ε is a strictly positive constant. Choose δ, p and ρ such that $1 - \frac{\delta}{p} < \rho < 1$. It follows that the first part of Equation (0.9) holds. Furthermore, since $\mu_{z_t}(y)$ is locally bounded for $y \in A$, the second part of Equation (0.9) holds, which concludes the proof of part (1) of Theorem . Part (3) follows from Section .

We prove part (2) of Theorem as follows. Using Assumptions B we get

$$
\begin{aligned}
y_t^2 &\leq \left[\sum_{i=1}^{p} c_{i(z)}|y_{i-1}| + o(\|y\|) + (\sum_{i=1}^{p} d_{i(z)}|y_{i-1}^2| + o\|y\|^2)^{\frac{1}{2}}\epsilon_{z_t}\right]^2 \\
&= \left(\sum_{i=1}^{p} c_{i(z)}|y_{i-1}| + o(\|y\|)\right)^2 + \sum_{i=1}^{p} d_{i(z)}|y_{i-1}^2| + o\|y\|^2)\epsilon_{z_t}^2 \\
&+ 2\mu_{z_t(y)}\sigma_{z_t}\epsilon_{z_t} \\
&= \sum_{i=1}^{p} c_{i(z)}|y_{i-1}|c_{j(z)}|y_{j-1}| + 2\sum_{i=1}^{p} c_{i(z)}y_{i-1}o(\|y\|) \\
&\quad + (o\|y\|)^2 \sum_{i=1}^{p} d_{i(z)}|y^2| + o\|y\|^2)\epsilon_{z_t}^2] + 2\mu_{z_t(y)}\sigma_{z_t}\epsilon_{z_t} \\
&= \sum_{i=1}^{p}(\tau_{i(z)} + \epsilon_{z_t}^2 d_{i(z)} y_{i-1}^2) \\
&+ 2\sum_{i=1}^{p} c_{i(z)}y_{i-1}o(\|y\|) + (o\|y\|)^2 + o\|y\|^2)\epsilon_{z_t}^2] \\
&\quad + 2\mu_{z_t(y)}\sigma_{z_t}\epsilon_{z_t}.
\end{aligned}
$$

Since both y_{t-1} and z_t are independent of ϵ_{z_t}, after taking expectation and setting $L_{z_t}(y)$ the same as in Equation (0.14), we get

$$
Ey_t^2 \leq \sum_{i=1}^{p}(\tau_{i(z)} + E\epsilon_{z_t}^2 d_{i(z)})Ey_{i-1}^2 + L_{z_t}(y).
$$

Simplifying by pulling the y_ts together and solving for Ey_t^2 we have

$$
Ey_t^2 \leq \frac{L_{z_t}(y)}{1 - [\sum_{i=1}^{p}(\tau_{i(z)} + E\epsilon_{z_t}^2 d_{i(z)})]}.
$$

The proof of Theorem (1) and (3) indicates that the Foster criterion F1 and F2 of [23] holds. Hence, by [23, Theorem 2 and 1(iii)], there exists a finite invariant measure π such that

$$
\frac{L_{z_t}(y)}{1 - [\sum_{i=1}^{p}(\tau_{i(z)} + E\epsilon_{z_t}^2 d_{i(z)})]} < \infty.
$$

Therefore, $E_\pi(y_t^2) < \infty$, as required. □

4. CONCLUDING REMARKS

In this paper, we established the geometric ergodicity of the MAR model and by implication have shown that it satisfies the absolutely regular (β-mixing) and strong mixing (α-mixing) conditions. In addition, we have shown that the process $\{y_t\}$ has a stationary distribution with finite second moments.

ACKNOWLEDGEMENTS

The author would like to thank the anonymous referee whose comments improved the original version of this manuscript.

REFERENCES

[5] R. C. Bradley, *Basic properties of strong mixing conditions. a survey and some open questions*, Probability Surveys **2** (2), 107-144, 2005.
[16] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer, Verlag, London, 1993.
[20] J. Shao and F.Xi, *Strong ergodicity of the regime-switching diffusion processes*, Stochastic Processes and their Applications **123** (11), 3903-3918, 2013.
[21] J.P. Stockis,J. Franke, J.T. Kamgaing, *On geometric ergodicity of charme models*, Journal of Time Series Analysis **31** (3), 141-152, 2010.
[22] D. Tjostheim, *Non-linear time series and Markov chains*, Advances in Applied Probability **22** (3), 587-611, 1990.
[23] R.L. Tweedie, *Invariant measures for Markov chains with no irreducibility assumptions* Journal of Applied Probability 25, pp. 275-285, 1988.
[1] M.I. Akinyemi, *Mixture autoregressive models: asymptotic properties and application to financial risk*, Ph.D. thesis, Mathematical Finance, Probability and Statistics Group, School of Mathematics, University of Manchester, 2013.
[2] R.Bhattacharya and C.Lee, *On geometric ergodicity of nonlinear autoregressive models*, Statistics and Probability Letters **41** (4) 439-440, 1999
[3] G.N. Boshnakov, *Analytic expressions for predictive distributions in mixture autoregressive models*, Statistics and Probability Letters **79** (15), 1704-1709, 1999.
[4] G.N. Boshnakov, *On first and second order stationarity of random coefficient models*, Linear Algebra Applications, **434** (2), 415-423, 2011.
[5] R.C. Bradley, *Basic properties of strong mixing conditions. a survey and some open questions*, Probability Surveys **2** (2), 107-144, 2005.
[6] S.A. Broda, M. Haas, J.Krause, M.S.Paolella, S.C.Steude, *Stable mixture GARCH models*. Journal of Econometrics **172** (2), 292-306, 2013.
[7] Y.A. Davydov, *Mixing conditions for Markov chains. Theory Probability and its Applications* **18** (2), 321-338, 1973.
[8] J. Dedecker, *Weak Dependence: With Examples and Applications*, Springer, New York, 2007.
[9] M. Fitzpatrick, *Geometric ergodicity of the Gibbs sampler for the poisson changepoint model*, Statistics and Probability Letters **91**(0), 55-61, 2014.
[10] C. Francq and M. Roussignol, *Ergodicity of autoregressive processes with Markov-switching and consistency of the maximum-likelihood estimator*, Statistics **32**(2), 151-173, 1998.
[11] C. Francq and J.M. Zakoian, *Garch Models: Structure, Statistical Inference and Financial Applications*, Wiley, 2010.
[12] L. Galtchouk and S. Pergamenshchikov, *Geometric ergodicity for classes of homogeneous Markov chains. Stochastic Processes and their Applications* **124** (10), 3362-3391, 2014.
[13] M.Haas, J. Krause, M.S. Paolella, S.C. Steude, *Time-varying mixture GARCH models and asymmetric volatility*, The North American Journal of Economics and Finance **26** (0), 602-623, 2013.

[14] L. Herve and J. Ledoux, *Approximating Markov chains and geometric ergodicity via weak perturbation theory*, Stochastic Processes and their Applications **124**(1), 613-638, 2014.

[15] B.V. Kini and C.C. Sekhar, *Large margin mixture of AR models for time series classification. Applied Soft Computing* **13** (1), 361-371, 2013.

[16] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*, Springer, Verlag, London, 1993.

[17] R. Prado, *Sequential estimation of mixtures of structured autoregressive models*, Computational Statistics and Data Analysis **58** (0), 58-70, 2013.

[18] J.C. Roman and J.P. Hobert, *Geometric ergodicity of Gibbs samplers for bayesian general linear mixed models with proper priors*, Linear Algebra and its Applications-special issue on Statistics **473** (0), 54-77, 2015. .

[19] J.P.Rydlewski and M. Snarska, *On geometric ergodicity of skewed-svcharme models*, Statistics and Probability Letters **84** (0), 192-197 2014.

[20] J. Shao and F.Xi, *Strong ergodicity of the regime-switching diffusion processes*, Stochastic Processes and their Applications **123** (11), 3903-3918, 2013.

[21] J.P. Stockis,J. Franke, J.T. Kamgaing, *On geometric ergodicity of charme models*, Journal of Time Series Analysis **31** (3), 141-152, 2010.

[22] D. Tjostheim, *Non-linear time series and Markov chains*, Advances in Applied Probability **22** (3), 587-611, 1990.

[23] R.L. Tweedie, *Invariant measures for Markov chains with no irreducibility assumptions* Journal of Applied Probability 25, pp. 275-285, 1988.

[24] C. Wong and W. Li, *On a mixture autoregressive model*, Journal of the Royal Statistical Society. Series B: Statistical Methodology **62** (1), 95-115, 2000.

[25] C.S. Wong, W.S. Chan, P.L. Kam, *A student t-mixture autoregressive model with applications to heavy-tailed financial data*, Biometrika **96** (3), 751-760, 2009.

[26] J. Zhao, L. Jin, L.Shi, *Mixture model selection via hierarchical BIC*, Computational Statistics and Data Analysis **88** (0), 139-153, 2015.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LAGOS, LAGOS, NIGERIA

E-mail address: `makinyemi@unilag.edu.ng`

SCHOOL OF MATHEMATICS, UNIVERSITY OF MANCHESTER, MANCHESTER M13 9PL, UK

E-mail address: `georgi.boshnakov@manchester.ac.uk`