Electronic Journal of Probability

Electron. J. Probab. **18** (2013), no. 90, 1–12.
ISSN: 1083-6489 DOI: 10.1214/EJP.v18-1871

Geometric ergodicity of asymmetric volatility models with stochastic parameters

Neelabh Rohan* T. V. Ramanathan†

Abstract

In this paper, we consider a general family of asymmetric volatility models with stationary and ergodic coefficients. This family can nest several non-linear asymmetric GARCH models with stochastic parameters into its ambit. It also generalizes Markov-switching GARCH and GJR models. The geometric ergodicity of the proposed process is established. Sufficient conditions for stationarity and existence of moments have also been investigated. Geometric ergodicity of various volatility models with stochastic parameters has been discussed as special cases.

Keywords: Asymmetric volatility models ; geometric ergodicity ; irreducibility ; stationarity ; stochastic parameter GARCH model.
AMS MSC 2010: 60J05 ; 60G10.
Submitted to EJP on March 12, 2012, final version accepted on October 18, 2013.

1 Introduction

The failure of GARCH models to account for the asymmetry and overestimation of persistence of variance has led to plenty of research in conditional heteroscedastic models. The findings of Black [4] about the asymmetric behavior of the market with respect to news resulted in number of asymmetric volatility models, such as EGARCH, GJR, AGARCH, TGARCH etc., see Engle and Ng [11] for details. Lamoureux and Lastrapes [26] and recently, Hwang and Pereira [22] have claimed that the persistence in variance may be overestimated because of the existence and failure to account for structural shifts in the volatility processes. Mikosch and Starica [30] showed that structural breaks in the unconditional variance of the GARCH volatility process can cause spurious high persistence. Moreover, these break points may be responsible for the asymmetry and long range dependence in the data set.

The immense popularity of the problem of structural breaks has resulted in several approaches of modeling the volatility with stochastically varying parameters, especially in the form of Markov-switching ARCH and GARCH models, see for example [7], [18], [16], [24], [23] and [17]. Ardia [1] modified Haas' [17] model to a Markov-switching asymmetric GJR model and discussed its Bayesian estimation procedures. Bauwens et

*University of Pune, INDIA. E-mail: neelabhrohan@gmail.com
†University of Pune, INDIA. E-mail: ram@stats.unipune.ac.in

al. [2] have established the stability of a regime switching GARCH model. Also see [12], [20] and [32] and references therein for various other approaches. In all the above papers, authors consider only GARCH models with parameter variation restricted to a finite state space.

Given the changing pace of the world economy at the time of various financial crises, several authors have considered the continuous variation in the parameters of GARCH models. Dahlhaus and Subba Rao [10], Čížek and Spokoiny [9] and Rohan and Ramanathan [33] developed inference procedures for the time varying ARCH (∞) and GARCH models. Recently, Fryzlewicz and Subba Rao [14] proved the mixing properties of a time varying ARCH model. However, the variation considered in the parameters of time varying ARCH and GARCH models is deterministic. Asymmetric volatility models address the limitations of a simple GARCH model, such as persistence and volatility clustering, and if we allow the coefficients of such models to vary stochastically according to time, then they can be better candidates for modeling volatility. In fact, Granger [15] advocated the use of stochastically varying parameter models by commenting 'Most of the time varying parameters change stochastically rather than deterministically'. In this paper, we introduce a general asymmetric volatility model with stochastically varying parameters and study its probabilistic properties. Sufficient conditions for the geometric ergodicity and stationarity of the chain associated with the model are given. The results discussed here provide a basis for a similar investigation on various other asymmetric volatility models with stochastic parameters. An important implication of the geometric ergodicity is that the law of large numbers and central limit theorems can be applied and therefore it becomes possible to develop a rigorous asymptotic estimation theory.

Establishing the probabilistic properties of GARCH type models has always been a dynamic topic of research due to the their vast applicability to financial data. Carrasco and Chen [8] provided sufficient conditions for the β-mixing and existence of higher order moments for various GARCH models as an application of a generalized random coefficient autoregressive model. Later, Francq and Zakoïan [13] proved the geometric ergodicity of a GARCH (1,1) process with rather mild conditions. Saïdi and Zakoïan [34] introduced a class of non-linear ARCH models and studied its probabilistic aspects. Using the similar ideas as in Carrasco and Chen [8], Meitz and Saikkonen [27] established the mixing properties of a class of Markov models and applied their results to the GARCH (1,1) model. However, all these papers study the probabilistic properties of standard GARCH models with constant coefficients. This article has a different contribution as it introduces various GARCH models with stationary and ergodic coefficients and establishes their geometric ergodicity.

The paper is organized as follows. A family of asymmetric volatility models with stochastic parameters is introduced in Section 2. In Section 3, we discuss the geometric ergodicity of the process and provide sufficient conditions for the strict and covariance stationarity. Various stochastic parameter volatility models and their probabilistic properties are also considered in Section 3.

2 A family of asymmetric volatility models with stochastic parameters

To define a family of asymmetric volatility models, we start from a general family of GARCH type equation for the volatility process with coefficients modeled as a Markov chain.

Definition 2.1. *Let $\theta_{\mathbf{t}} = (\omega_t, \alpha_t, \beta_t, b_t, c_t, \lambda_t, \nu_t)$ be a stationary and ergodic Markov chain with state space $\mathcal{D}$, a measurable subset of $\mathcal{R}_+^2 \times (0,1) \times \mathcal{R} \times [-1,1] \times \mathcal{R}_+^2$, where*

$\mathcal{R}_+ = (0, \infty)$. Let $\{\epsilon_t\}$ be a process such that $E(\epsilon_t|\mathcal{F}_{t-1}) = 0$ and $E(\epsilon_t^2|\mathcal{F}_{t-1}) = \sigma_t^2$, where $\mathcal{F}_{t-1} = \sigma\{\epsilon_{t-1}, \epsilon_{t-2}, \ldots\}$. Suppose $\{v_t\}$ is a sequence of real valued independent and identically distributed (i.i.d.) random variables independent of $\{\theta_t\}$, having mean 0 and variance 1. The family of asymmetric volatility models with stochastic parameters is defined as

$$\begin{aligned} \epsilon_t &= \sigma_t v_t, \\ \sigma_{t+1}^{\lambda_t} &= \omega_t + \{\alpha_t \lambda_t [|v_t - b_t| - c_t(v_t - b_t)]^{\nu_t} + \beta_t\} \sigma_t^{\lambda_{t-1}} \end{aligned} \tag{2.1}$$

for all $t \in \mathcal{Z}$, the set of all integers.

We call the elements of $\mathcal{D}$ as *regimes*. The state space $\mathcal{D}$ is assumed to have a metric inherited from some norm on $\mathcal{R}_+^2 \times (0,1) \times \mathcal{R} \times [-1,1] \times \mathcal{R}_+^2$ and the Borel σ-algebra $\mathcal{B}(\mathcal{D})$. This state space ensures the positivity of the variance process in every regime. The function $\Delta(\theta, D), \theta \in \mathcal{D}, D \in \mathcal{B}(\mathcal{D})$ denotes the one-step Markov transition function of the chain $\{\theta_\mathbf{t}\}$.

The model defined in (2.1) may be considered as a generalization of the family of volatility models given by Hentschel [21], which can nest almost all the volatility models for different parameter choices. This definition of the process also generalizes other Markov-switching models in two ways. First, it allows infinitely many regimes as opposed to only finite number of regimes in the other Markov-switching GARCH models. Second, this general model also allows for a shift in the parameters of asymmetry of both types i.e., the asymmetry due to shift, denoted by b, and that due to rotation, denoted by c, in the news impact curve, see [21]. It can handle the situations, where there is asymmetry in one regime and symmetry in the other or one type of asymmetry in one regime and another type in the other. Hence this model can capture different dynamics in low and high volatility periods efficiently as it can take different asymmetric volatility models into account.

3 Stationarity and geometric ergodicity

Due to the stochastic nature of the parameters, it would be interesting to establish the existence of a stationary solution of the process defined in (2.1). Towards this, first we make the following assumption.

$\mathbf{A_1}$: *Let the parameter chain $\{\theta_t\}$ be such that $E\ [log(\alpha_0 \lambda_0 (|v_0 - b_0| - c_0(v_0 - b_0))^{\nu_0} + \beta_0)] < 0$ and $E(log\omega_0)^+ < \infty$, where, $(log\omega_0)^+ = \max(0, log\omega_0)$.*

Theorem 3.1. *The family of asymmetric volatility process defined in (2.1) has a unique stationary solution, if A_1 is satisfied. For a given sequence $\{\theta_\mathbf{t}\}$, the unique stationary solution is*

$$\bar{\sigma}_{t+1}^{\lambda_t} = \omega_t + \sum_{i=1}^{\infty} \left(\prod_{k=t-i+1}^{t} (\alpha_k \lambda_k [|v_k - b_k| - c_k(v_k - b_k)]^{\nu_k} + \beta_k) \right) \omega_{t-i} \tag{3.1}$$

such that $|\sigma_{t+1}^{\lambda_t} - \bar{\sigma}_{t+1}^{\lambda_t}| \to 0$ almost surely for $t \geq 1$ with some well defined initial value at $t = 0$, where $\sigma_{t+1}^{\lambda_t}$ satisfies (2.1). The series inside the bracket on the right hand side of (3.1) converges absolutely almost surely. Moreover, $\inf_t \omega_t/(1 - \inf_t \beta_t) \leq \sigma_{t+1}^{\lambda_t} < \infty$ almost surely.

Remark 3.2. *1. The top Lyapounov exponent associated with the model (2.1), defined as*

$$\gamma_0 = \inf \left\{ \left(\frac{1}{t} \sum_{i=1}^{t} log(\alpha_i \lambda_i [|v_i - b_i| - c_i(v_i - b_i)]^{\nu_i} + \beta_i) \right) \right\},$$

is strictly negative. In fact, it is equal to $E\ [log(\alpha_0 \lambda_0 [|v_0 - b_0| - c_0(v_0 - b_0)]^{\nu_0} + \beta_0)]$ (see [5], [25] and [35]).

2. *If the random sequence $\{\theta_t\}$ is i.i.d., then the assumption A_1 is also necessary for the existence of a unique stationary solution (see [5]).*
3. *Under the i.i.d. assumption of the parameter chain $\{\theta_t\}$, $\sigma_t^{\lambda_{t-1}}$ and (A_t, B_t) are independent, where $A_t = (\alpha_t\lambda_t[|v_t - b_t| - c_t(v_t - b_t)]^{\nu_t} + \beta_t)$ and $B_t = \omega_t$. Also, $\sigma_1^{\lambda_0}$ is equal in distribution to $\left[\omega_0 + \{\alpha_0\lambda_0[|v_0 - b_0| - c_0(v_0 - b_0)]^{\nu_0} + \beta_0\}\sigma_1^{\lambda_0}\right]$ ([6]).*

Next, we consider general order moments of $\sigma_{t+1}^{\lambda_t}$. The following theorem puts an upper and lower bound on the moments of $\sigma_{t+1}^{\lambda_t}$ and finds conditions for their existence.

Theorem 3.3. *Suppose that the assumption A_1 is satisfied. Then,*

$$\left[\{E(\omega_t^r)\}^{1/r} + \sum_{i=1}^{\infty}\{E(\eta_{i,t}\omega_{t-i})^r\}^{1/r}\right]^r \leq E(\sigma_{t+1}^{r\lambda_t}) \leq E(\omega_t^r) + \sum_{i=1}^{\infty} E(\eta_{i,t}\omega_{t-i})^r \qquad (3.2)$$

for $0 < r \leq 1$, where $\eta_{i,t} = \prod_{k=t-i+1}^{t}[\alpha_k\lambda_k(|v_k - b_k| - c_k(v_k - b_k))^{\nu_k} + \beta_k]$. When $1 \leq r < \infty$, then both the inequalities in (3.2) are reversed, that is,

$$E(\omega_t^r) + \sum_{i=1}^{\infty} E(\eta_{i,t}\omega_{t-i})^r \leq E(\sigma_{t+1}^{r\lambda_t}) \leq \left[\{E(\omega_t^r)\}^{1/r} + \sum_{i=1}^{\infty}\{E(\eta_{i,t}\omega_{t-i})^r\}^{1/r}\right]^r. \qquad (3.3)$$

Also, the series on the right hand sides of (3.2) (for $0 < r \leq 1$) and (3.3) (for $1 < r < \infty$) converge if

$$\limsup_{n\to\infty}\left[E\left(\prod_{k=1}^{n} A_{t-k+1}^r\,\omega_{t-n}^r\right)\right]^{1/n} < 1 \text{ and } E(\omega_t)^r < \infty. \qquad (3.4)$$

where $A_t = (\alpha_t\lambda_t[|v_t - b_t| - c_t(v_t - b_t)]^{\nu_t} + \beta_t)$. Further, $E(\sigma_{t+1}^{\lambda_t})^r < \infty$ if (3.4) holds.

Remark 3.4.
1. *If any of the expectations $E(\omega_t^r)$ and $E(\eta_{i,t}\omega_{t-i})^r$ in (3.2) does not exist, then $E(\sigma_{t+1}^{r\lambda_t}) = +\infty$.*
2. *If the chain $\{\theta_t\}$ is i.i.d., then the sufficient condition for the existence of r^{th} moment (3.4) reduces to $E(A_t^r) < 1$ and $E(\omega_t^r) < \infty$, which is equivalent to the condition assumed by Nelson [31] for a constant parameter stationary GARCH models.*
3. *If the chain $\{\theta_t\}$ is degenerate at some constant θ, then the condition (3.4) reduces to $E\left(\alpha\lambda[|v_t - b| - c(v_t - b)]^{\nu} + \beta\right)^r < 1$, which is a sufficient condition for the existence of r^{th} moment of the Hentschel [21] family of models with constant parameters. This condition further reduces to $E(2\alpha v_t^2 + \beta)^r < 1$ for a simple GARCH model, which is same as the Nelson's [31] condition for existence of moments.*

To establish the geometric ergodicity of the variance process defined in (2.1), we prove the same for $z_t = (h_t, \theta_t)^\top$, where $h_t = \sigma_{t+1}^{\lambda_t}$. Notice that v_t is independent of z_{t-1}. Suppose that the state space of the process $\{z_t\}$ is given by $\mathcal{X} \subset \mathcal{R}_+ \times \mathcal{D}$, where this subset $\mathcal{X}$ is such that all the sets in it with $h_t < \inf_t \omega_t/(1 - \inf_t \beta_t)$ have probability zero. This restriction is due to the lower limit on h_t established in Theorem 3.1. Since $\{\theta_t\}$ is a vector Markov chain, there exists a measurable function $F_\theta : \mathcal{D} \times \mathcal{R}^7 \to \mathcal{D}$ such that $\theta_t = F_\theta(\theta_{t-1}, u_t)$, where the error term $u_t \in \mathcal{R}^7$ is i.i.d. and independent of θ_{t-1} and v_t (according to the definition of θ_t). Let $\zeta_t = (u_t, v_t)$. Then,

$$z_t = \begin{pmatrix} h_t \\ \theta_t \end{pmatrix} = \begin{pmatrix} \omega_t + \{\alpha_t\lambda_t f_t^{\nu_t}(v_t) + \beta_t\}h_{t-1} \\ \theta_t \end{pmatrix} = F(z_{t-1}, \zeta_t),$$

where $F : \mathcal{X} \times \mathcal{R}^8 \to \mathcal{X}$ and $f_t(v_t) = |v_t - b_t| - c_t(v_t - b_t)$. Here, independence of v_t with z_{t-1} ensures that $\{z_t\}$ is a Markov chain. Hence, although the variance process alone in model (2.1) is non-Markovian, the process $z_t = (h_t, \theta_t)^\top$ is Markovian, which simplifies our derivation. Now consider the following assumptions:

$\mathbf{A_2}$: *The random variable v_t has a density $g(\cdot)$ with respect to Lebesgue measure on real line, which is positive and continuous everywhere.*

$\mathbf{A_3}$: $E(|v_t|^\vartheta) < \infty,\ E(\omega_t)^\vartheta < \infty$ *for some $\vartheta > 0$.*

The assumption A_2 is required for the irreducibility of z_t. A_3 is used to prove its geometric ergodicity.

Suppose $\mathcal{B}(\mathcal{X})$ denotes the Borel sigma field of $\mathcal{X}$ and λ^* is the measure associated with the chain, such that $\lambda^*(A)$ is well defined for every $A \in \mathcal{B}(\mathcal{X})$. Here $\lambda^* = \nu_L \otimes \phi$ is the product measure on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, where ν_L denotes the Lebesgue measure and ϕ is an irreducibility measure on $\mathcal{B}(\mathcal{D})$. Denote $p^m(z_0, B) = P(z_t \in B | z_{t-m} = z_0)$, $B \in \mathcal{B}(\mathcal{X}), z_0 \in \mathcal{X}$, the probability that starting with z_0, the process moves to B in m steps.

The first major step towards establishing the geometric ergodicity of the process $\{z_t\}$ is to prove the irreducibility. We deal with this in the following theorem.

Theorem 3.5. *Let the chain $\{\theta_t\}$ be $\phi-$irreducible and aperiodic. Suppose that the assumption A_2 regarding $\{v_t\}$ holds. Then the chain $\{z_t\}$ is λ^*-irreducible and aperiodic. Further, there exists a probability measure Ψ on $\mathcal{B}(\mathcal{X})$, called the maximal irreducibility measure, such that $\{z_t\}$ is Ψ-irreducible.*

The following corollary establishes that if transitions of the chain $\{\theta_t\}$ are continuous, then the chain $\{z_t\}$ is weak Feller. We also prove that the process $\{z_t\}$ is a T-chain. The definitions of small sets, petite sets, Feller chain and T-chain, which are used in proving the geometric ergodicity, can be found in [29].

Corollary 3.6. *(i) Suppose there exists a measurable function F_θ such that $\theta_t = F_\theta(\theta_{t-1}, u_t)$, where $\{u_t\}$ is an i.i.d. sequence independent of θ_{t-1}, assuming values in the measurable space $(D, \mathcal{B}(D))$ and $F_\theta(\cdot, u)$ is continuous for any fixed u. Then, the process $\{z_t\}$ is weak Feller.*
(ii) Further, if the support of ϕ has a non-empty interior, that is, the measurable space $(D, \mathcal{B}(D))$ has at least one set with positive $\phi-$measure, then the process $\{z_t\}$ is a T-chain. Moreover, $\mathcal{X}$ can be written as the union of open small sets and every compact set in $\mathcal{X}$ is petite.

Remark 3.7. *The proof of the weak Feller property of $\{z_t\}$ is similar to the proof of the Proposition 6.1.2 of Meyn and Tweedie [29], in which the weak Feller property of nonlinear state space (NSS) model is proved. Compared to NSS model, which requires the assumption of smoothness of F, our assumption is weaker because we do not impose any differentiability restriction on F. In fact, F is not differentiable in our case.*

Towards proving the geometric ergodicity of $\{z_t\}$, we state the following lemma first. The proof of the lemma is based on the idea that under the assumption A_1, we can find a drift function (see [29]) to show that the process (θ_t, v_t) is geometrically ergodic.

Lemma 3.8. *Let the assumptions of Theorem 3.5 and Corollary 3.6 hold. Then, the process $\{(\theta_t, v_t)\}$ is geometrically ergodic if the assumptions A_1 to A_3 are satisfied.*

Now we prove the geometric ergodicity of the process z_t with the help of an extra assumption ensuring the existence of a small order moment of the process (2.1). Notice that condition (3.5) in Theorem 3.9 below is sufficient for the existence of moment of order ϑ, $E(h_t)^\vartheta < \infty$ for some small $\vartheta > 0$.

Theorem 3.9. *Suppose that the assumptions A_1 to A_3 along with those of Theorem 3.5 and Corollary 3.6 are satisfied. Then, the process $\{z_t\}$ is geometrically ergodic if the following condition holds.*

$$\limsup_{n\to\infty} \left[E\left(\prod_{k=1}^{n} A_{t-k+1}^{\vartheta}\ \omega_{t-n}^{\vartheta} \right) \right]^{1/n} < 1 \ \ for\ some\ \vartheta > 0 \tag{3.5}$$

where A_t is as defined in Theorem 3.3.

Corollary 3.10. *Let all the assumptions of Theorem 3.9 be satisfied. Then the process $\{z_t\}$ is strictly stationary and strong or $\alpha-$mixing if its initial distribution is same as the stationary distribution.*

As mentioned in Section 1, the model (2.1) nests several GARCH models with stochastic parameters (SP hereafter) and hence provides the basis for investigation of their probabilistic properties. In Table 1, we provide the constraints on the parameters of model (2.1) which lead to a specific volatility model with stochastic parameters. In this table, by equating some of the components of θ_t to a constant, we mean that, they are degenerate or take a constant value with probability one.

The results of Theorems 1-4 and their corollaries are also applicable to the various GARCH models with stochastic parameters mentioned in Table 1. That is, each of the models is stationary and geometrically ergodic under the corresponding specific assumptions subject to the constraints of Table 1. The stationary solution and conditions for existence of moments for all the models can also be directly obtained using (3.1) and (3.4) under the constraints mentioned in Table 1.

Table 1. Constraints on model (2.1) leading to specific stochastic parameter (SP) volatility models

Model	Constraints on model (2.1)
SPGARCH	$\lambda_t = \nu_t = 2,\ b_t = c_t = 0$
SPGJR	$\lambda_t = \nu_t = 2,\ b_t = 0$
SPTGARCH	$\lambda_t = \nu_t = 1,\ b_t = 0$
SPNAGARCH	$\lambda_t = \nu_t = 2,\ c_t = 0$
SPNARCH	$\lambda_t = \nu_t,\ b_t = c_t = 0$
SPAPARCH	$\lambda_t = \nu_t,\ b_t = 0$
SPAGARCH	$\lambda_t = \nu_t = 1$
SPEGARCH	$\lambda_t \to 0, \nu_t = 1, b_t = 0$

Remark 3.11. *SPGARCH and SPGJR models (Table 1) can be considered as the generalization of Markov-switching GARCH and GJR models with only finitely many regimes of Haas et al. [17] and Ardia [1] respectively.*

Remark 3.12. *Meitz and Saikkkonen [28] studied the stability of a simple AR-GARCH model. However, they require stringent smoothness assumptions and hence their results are not applicable to the threshold type non-linear volatility models. Our approach differs from theirs in two ways. First, it considers several GARCH models with stochastic parameters as discussed above. Second, due to moderate assumptions, it is possible to take into account the threshold type non-linear models such as GJR and TGARCH with stochastic parameters.*

Appendix: Proofs

Proof of Theorem 3.1. The variance process (2.1) can be written as,

$$Y_{t+1} = A_t Y_t + B_t,$$

where $Y_{t+1} = \sigma_{t+1}^{\lambda_t}$ and A_t, B_t are as defined in Remark 3.2. Here $(\theta_t,\ v_t)$ and hence (A_t, B_t) is a stationary ergodic sequence. Therefore, the result immediately follows from Brandt ([6], Theorem 1). Now from (3.1), we can write

$$\bar{\sigma}_{t+1}^{\lambda_t} \geq \inf_t \omega_t \left(1 + \sum_{i=1}^{\infty} \prod_{k=t-i+1}^{t} \inf_t \beta_t \right).$$

Hence, the lower limit for $\sigma_{t+1}^{\lambda_t}$ follows. □

Proof of Theorem 3.3. Let $0 < r \leq 1$. From (3.1) and Minkowski's integral inequality (Hardy et al. [19], Theorem 199),

$$E(\sigma_{t+1}^{r\lambda_t}) \leq E(\omega_t^r) + \sum_{i=1}^{\infty} E\left[\left(\prod_{k=t-i+1}^{t} (\alpha_k \lambda_k [|v_k - b_k| - c_k(v_k - b_k)]^{\nu_k} + \beta_k)\right)\omega_{t-i}\right]^r. \quad (3.6)$$

Now, using the Cauchy root criterion and assumption (3.4), the series on the right hand side of (3.6) converges. We obtain the lower bound in (3.2) using Theorem 198 of Hardy et al. [19] for $0 < r \leq 1$,

$$\{E(\sigma_{t+1}^{r\lambda_t})\}^{1/r} \geq \{E(\omega_t^r)\}^{1/r} + \sum_{i=1}^{\infty} \left\{E\left[\left(\prod_{k=t-i+1}^{t} (\alpha_k \lambda_k [|v_k - b_k| - c_k(v_k - b_k)]^{\nu_k} + \beta_k)\right)\omega_{t-i}\right]^r\right\}^{1/r}.$$

Similarly, for $1 < r < \infty$, using Minkowski's integral inequality (Hardy et al. [19], Theorem 198) and (3.1),

$$\{E(\sigma_{t+1}^{r\lambda_t})\}^{1/r} \leq \{E(\omega_t^r)\}^{1/r} + \sum_{i=1}^{\infty} \left\{E\left[\left(\prod_{k=t-i+1}^{t} (\alpha_k \lambda_k [|v_k - b_k| - c_k(v_k - b_k)]^{\nu_k} + \beta_k)\right)\omega_{t-i}\right]^r\right\}^{1/r}.$$

Here as in (3.6), the series on the right hand side converges using Cauchy root criterion and assumption (3.4). Also, using Theorem 199 of Hardy et al. [19] for $r > 1$,

$$E(\sigma_{t+1}^{r\lambda_t}) \geq E(\omega_t^r) + \sum_{i=1}^{\infty} E\left[\left(\prod_{k=t-i+1}^{t} (\alpha_k \lambda_k [|v_k - b_k| - c_k(v_k - b_k)]^{\nu_k} + \beta_k)\right)\omega_{t-i}\right]^r.$$

Hence the theorem follows.

□

Proof of Theorem 3.5. To prove the λ^*-irreducibility of the process, we need to show that $p^k(z_0, B) > 0$ for some $k \geq 1$ and $\forall\ z_0 \in \mathcal{X}$, where B is any measurable set such that $B \in \mathcal{B}(\mathcal{X})$ with $\lambda^*(B) > 0$. This is done by showing that starting with an arbitrary $(h_0, \theta_0) \in \mathcal{X}$, any point $(h, \theta) \in B$ can be reached in finite number of steps with a positive probability. Denote the component sets of B as B^h and B^θ which are Borel sets such that $B \equiv B^h \times B^\theta$. The $\phi-$irreducibility of the process $\{\theta_t\}$ ensures that if $\phi(B^\theta) > 0$, then there exists an $m \geq 0$ (depending on θ_0 and B^θ) such that $\Delta^{m+1}(\theta_0, B^\theta) > 0$, where $\Delta^{m+1}(\theta_0, B^\theta)$ represents the $(m+1)$-step transition probability of $\{\theta_t\}$ to the set B^θ, starting with θ_0. Thus there exist intermediate Borel sets $B_i^\theta \in \mathcal{B}(\mathcal{D}),\ i = 1, 2, \ldots, m$ with $\phi(B_i^\theta) > 0$ and $\Delta(\theta_{i-1}, B_i^\theta) > 0,\ \theta_{i-1} \in B_{i-1}^\theta$, through which the chain $\{\theta_t\}$ transits to B starting with θ_0. Corresponding to such sets, identify the sets B_i^h such that starting with z_0, the chain $\{z_t\}$ can reach to z through the Borel sets $B_i \equiv B_i^h \times B_i^\theta$ in $m+1$ steps with a positive probability. The sets B_i^h are defined as

$$\begin{aligned} B_1^h = &\ \left\{h : h = \omega_1 + [\alpha_1 \lambda_1 (|v_1 - b_1| - c_1(v_1 - b_1))^{\nu_1} + \beta_1] h_0,\ \theta_1 \in B_1^\theta, v_1 \in \mathcal{R}\right\}, \\ B_i^h = &\ \left\{h : h = \omega_i + [\alpha_i \lambda_i (|v_i - b_i| - c_i(v_i - b_i))^{\nu_i} + \beta_i] h_{i-1},\ \theta_i \in B_i^\theta, v_i \in \mathcal{R}, h_{i-1} \in B_{i-1}^h\right\}, \\ &\ i = 2, 3, \ldots, m. \end{aligned}$$

Therefore, the chain $\{z_t\}$ can transit through the following path

$$(h_0, \theta_0) \to (h_1, \theta_1) \to \cdots \to (h_m, \theta_m) \to (h, \theta).$$

Here each $(h_i, \theta_i) \in B_i,\ i = 1, 2, \ldots, m$. By the transition $(h_{i-1}, \theta_{i-1}) \to (h_i, \theta_i)$, we mean that the transition is from $(h_{i-1}, \theta_{i-1}) \in B_{i-1}$ to B_i. The probability of reaching $z = (h, \theta) \in B$ from $z_0 = (h_0, \theta_0)$ in the m intermediate steps through sets $B_i,\ i = 1, 2, \ldots, m$ is,

$$\begin{aligned} p^{m+1}(z_0, B) &\geq \prod_{i=1}^{m+1} P(z_i \in B_i | z_{i-1}), \text{ where } z_{i-1} \in B_{i-1} \\ &= \prod_{i=1}^{m+1} P(h_i \in B_i^h | z_{i-1}, \theta_i \in B_i^\theta) P(\theta_i \in B_i^\theta | z_{i-1}) \\ &= \prod_{i=1}^{m+1} P\left((\omega_i + [\alpha_i \lambda_i f_i^{\nu_i}(v_i) + \beta_i] h_{i-1}) \in B_i^h | z_{i-1}, \theta_i \in B_i^\theta\right)\ \Delta(\theta_{i-1}, B_i^\theta) > 0 \end{aligned}$$

where $f_i(v_i) = |v_i - b_i| - c_i(v_i - b_i)$, $B_{m+1} = B$ and $z_{m+1} = z$. Here, the first probability is strictly greater than zero under the assumption A_2. In this way, each $z \in B$ with $\lambda^*(B) > 0$ can be reached in finite number of steps with a positive probability, i.e., $p^{m+1}(z_0, B) > 0$. Hence the process z_t is λ^*-irreducible.

For $B \in \mathcal{B}(\mathcal{X})$, define the probability measure Ψ as,

$$\Psi(B) = \int_{\mathcal{X}} \lambda^*(dy) \sum_{n=0}^{\infty} p^n(y, B) 2^{-(n+1)}.$$

Hence the Proposition 4.2.2 of Meyn and Tweedie ([29], page 90) implies that $\{z_t\}$ is Ψ-irreducible where Ψ is a maximal irreducibility measure. Aperiodicity of the chain is obvious as $\{\theta_t\}$ is aperiodic and corresponding to every B_i^θ there exist sets and B_i^h as shown above.

□

Proof of Corollary 3.6. (i) This proof is similar to the proof of the Proposition 6.1.2 of Meyn and Tweedie [29]. Denote the class of all bounded continuous functions from $\mathcal{X}$ to $\mathcal{R}$ by $\mathcal{C}(\mathcal{X})$. Using the similar notations as in Meyn and Tweedie ([29], Section 6.1), the transition probability kernel P acts on the bounded functions through the following mapping

$$Ph(z) = \int P(z, dy) h(y).$$

where $h \in \mathcal{C}(\mathcal{X})$ is a bounded continuous function. Then by the definition of the mapping $F(z, \zeta)$ and the continuity of F_θ, it is possible to claim that $h(F(z, \zeta))$ is bounded and continuous. Thus it follows that (Meyn and Tweedie [29], Section 3.5.5)

$$Ph(z) = E(h(z_{t+1}) | z_t = z) = E(h(F(z_t, \zeta_{t+1})) | z_t = z) = E[h(F(z, \zeta))]$$

is a continuous function of $z \in \mathcal{X}$. Therefore, by the definition, $\{z_t\}$ is a weak Feller chain.

(ii) Let $z \in \mathcal{X}$ be such that it can be reached from any $z_0 \in \mathcal{X}$ in a finite number of steps. First, we show that z is *reachable*, that is, every neighborhood of z is reachable with a positive probability (see [29], Section 6.1.2). Let $B \in \mathcal{B}(\mathcal{X})$ be any neighborhood of z. It is shown in Theorem 3.5 that the chain $\{z_t\}$ is λ^*-irreducible. Also, the

λ^*-measure of the boundary points is zero and support of ϕ has a non-empty interior (which implies that $\{\theta_t\}$ is a T-chain). Hence, there exists an open ball $B' \subset B$ with $\lambda^*(B') > 0$, such that all $z \in B'$ can be reached in a finite number of steps, starting with any z_0. This implies that for any open set B containing z, we have $\lambda^*(B) \geq \lambda^*(B') > 0$ and hence $\Psi(B) \geq \Psi(B') > 0$. Therefore, the point z is *reachable* and hence support of Ψ has a non empty interior (Meyn and Tweedie [29], Lemma 6.1.4). We have already shown that z_t is a $\Psi-$irreducible Feller chain. Thus, $\{z_t\}$ is a T-chain (Meyn and Tweedie [29], Theorem 6.2.9) and hence the state space can be written as a union of open small sets (Meyn and Tweedie [29], Theorem 6.0.1).

□

Proof of Lemma 3.8. With the given assumptions, it can be easily proved that the chain $\{(\theta_t, v_t)\}$ is $\Psi-$irreducible, aperiodic and ergodic. To prove the geometric ergodicity, we use the $t-$step drift criterion of Tjostheim [36] and Meyn and Tweedie [29]. Define a function

$$g_1(\theta, v) = \omega + [\alpha\lambda[|v-b| - c(v-b)]^{\nu} + \beta],$$

and an associated drift function $V(\theta, v) : \mathcal{D} \times \mathcal{R} \to [1, \infty)$ as

$$V(\theta, v) = 1 + [g_1(\theta, v)]^{k_0},$$

for some $0 < k_0 < \vartheta$. Define the test set as a compact set $C_{\theta,v} = \{(\theta, v) : [g_1(\theta, v)]^{k_0} \leq c^*\}$, where c^* is a finite constant. Notice that under ergodicity assumption of $\{\theta_t\}$, the difference between the $t-$step ahead transition probability kernel of θ_t (given θ_0) and its unique invariant stationary probability is negligible for a sufficiently large t (Meyn and Tweedie [29], Chapter 13). Therefore, for a sufficiently large t,

$$E(V(\theta_t, v_t)|\theta_0, v_0) = E(V(\theta_0, v_0)) \leq 1 + E(\omega_0)^{k_0} + E[\alpha_0\lambda_0[|v_0 - b_0| - c_0(v_0 - b_0)]^{\nu_0} + \beta_0]^{k_0}.$$

where the inequality on the right follows using the C_r inequality. Here the unconditional expectation is taken with respect to the stationary distribution of $\{(\theta_t, v_t)\}$. Now, using similar arguments as in Remark 2.9 of Basrak et al. [3], we can show that $E[\alpha_0\lambda_0[|v_0 - b_0| - c_0(v_0 - b_0)]^{\nu_0} + \beta_0]^{k_0} < 1$ for some $0 < k_0 < \vartheta$, when the assumptions A_1 and A_3 ($E|v_0|^{\vartheta} < \infty$) are satisfied. Therefore,

$$E(V(\theta_t, v_t)|\theta_0, v_0) \leq c_1^*$$

for a sufficiently large t, where $c_1^* = 1 + E(\omega_0)^{k_0} + c_2^*$ and $0 < c_2^* < 1$. Now let $\delta_{\theta_0} = ([g_1(\theta_0, v_0)]^{k_0} - c^*)/(1 + g_1(\theta_0, v_0)^{k_0})$, where $c^* = (c_2^* + E(\omega_0)^{k_0})$. Therefore, $\delta_{\theta_0} > 0$ for $\theta_0 \in \bar{C}_{\theta,v}$ (complement of $C_{\theta,v}$) and hence,

$$E(V(\theta_t, v_t)|\theta_0, v_0) \leq (1 - \delta_{\theta_0})V(\theta_0, v_0) \quad \text{if } \theta_0 \in \bar{C}_{\theta,v}$$

Clearly, if $\theta_0 \in C_{\theta,v}$, then $E(V(\theta_t, v_t)|\theta_0, v_0)$ is bounded in view of the assumptions A_3 ($E(\omega_0)^{\vartheta} < \infty$) and the fact that $E[\alpha_0\lambda_0[|v_0 - b_0| - c_0(v_0 - b_0)]^{\nu_0} + \beta_0]^{k_0} < 1$. Therefore using Meyn and Tweedie ([29], Theorem 15.0.1) along with Theorem 3.5, it follows that $\{(\theta_t, v_t)\}$ is geometrically ergodic.

□

Proof of Theorem 3.9. We have shown in Lemma 3.8 that the assumptions A_1 to A_3 along with those of Theorem 3.5 and Corollary 3.6 are sufficient for $\{(\theta_t, v_t)\}$ to be geometrically ergodic. Now, we prove the geometric ergodicity of $\{z_t\}$. By recursively substituting for h_t, the variance process (2.1) can be written as

$$h_t = \prod_{i=1}^{t}(\alpha_i\lambda_i f_i^{\nu_i}(v_i) + \beta_i)h_0 + \left(\omega_t + \sum_{j=1}^{t-1}\prod_{k=t-j+1}^{t}(\alpha_k\lambda_k f_k^{\nu_k}(v_k) + \beta_k)\omega_{t-j}\right). \quad (3.7)$$

Now, by the strong law of large numbers (SLLN), as $t \to \infty$

$$\prod_{i=1}^{t}(\alpha_i\lambda_i f_i^{\nu_i}(v_i)+\beta_i)^{1/t} = exp\left(\frac{1}{t}\sum_{i=1}^{t} log(\alpha_i\lambda_i f_i^{\nu_i}(v_i)+\beta_i)\right) \to exp\,(\gamma^*) < 1 \;\; a.s.,$$

where $\gamma^* = E\,[\,log(\alpha_0\lambda_0[|v_0-b_0|-c_0(v_0-b_0)]^{\nu_0}+\beta_0)] < 0$ using A_1. Therefore, there exists an N, sufficiently large such that $\forall\; t > N$,

$$\left(\prod_{i=1}^{t}(\alpha_i\lambda_i f_i^{\nu_i}(v_i)+\beta_i)^{1/N}\right) = exp\left(\frac{t}{N}\frac{1}{t}\sum_{i=1}^{t} log(\alpha_i\lambda_i f_i^{\nu_i}(v_i)+\beta_i)\right) \le \rho < 1. \qquad (3.8)$$

Using C_r inequality in (3.7), consider for some $t > N$,

$$E(h_t^{1/N}|z_0=(h_0,\theta_0)) \le E\left(\prod_{i=1}^{t}(\alpha_i\lambda_i f_i^{\nu_i}(v_i)+\beta_i)^{1/N}|\theta_0\right) h_0^{1/N} + M, \qquad (3.9)$$

where $M = E\left[\left(\omega_t + \sum_{j=1}^{t-1}\prod_{k=t-j+1}^{t}(\alpha_k\lambda_k f_k^{\nu_k}(v_k)+\beta_k)\omega_{t-j}\right)|\theta_0\right]^{1/N}$

$$\le E(\omega_t^{1/N}|\theta_0) + \sum_{j=1}^{t-1} E\left[\left(\prod_{k=t-j+1}^{t}(\alpha_k\lambda_k f_k^{\nu_k}(v_k)+\beta_k)^{1/N}\omega_{t-j}^{1/N}\right)|\theta_0\right]$$

$$\le E(\omega_t^{1/N}|\theta_0) + \sum_{j=1}^{\infty} E\left[\left(\prod_{k=t-j+1}^{t}(\alpha_k\lambda_k f_k^{\nu_k}(v_k)+\beta_k)^{1/N}\omega_{t-j}^{1/N}\right)|\theta_0\right]$$

Now for $N \to \infty$, we have for $t > N$, $E(\omega_t^{1/N}|\theta_0) = E(\omega_t^{1/N})$ due to ergodicity of $\{\theta_t\}$. Also, using assumption (3.5) and the Cauchy root criterion, $M < \infty$. Now, from (3.8) and (3.9), we can write,

$$E(h_t^{1/N}|z_0) \le \rho h_0^{1/N} + M,$$

where, $0 \le \rho < 1$. Define a drift function $V : \mathcal{X} \to [1,\infty)$ as

$$V(z) = 1 + h^{\vartheta}.$$

We define the test set as the compact set $C = \{(h,\theta) \in \mathcal{X} : h^{\vartheta} \le c\}$, where c is a constant. The drift function V is bounded on the set C, which is a petite set from Corollary 3.6 and can serve as a test set for the drift criterion. Now,

$$E(V(z_t)|z_0) \le 1 + M + \rho h_0^{\vartheta}$$

Define $c = (1+M-\delta)/(\delta-\rho)$, where δ is a constant satisfying $\rho < \delta < 1$. Then, if $h_0^{\vartheta} > c$ ($z_0 \in \bar{C}$, the complement of C),

$$E(V(z_t)|z_0) \le \delta(1+h_0^{\vartheta}) = \delta V(z_0), \;\; \text{for } z_0 \in \bar{C}.$$

Now if $z_0 \in C$, then the $E(V(z_t)|z_0)$ is clearly bounded. Hence there exist constants $b < \infty,\; 0 < \delta < 1$, a petite set C and the function V as defined above satisfying

$$E(V(z_t)|z_0) \le \delta V(z_0) + bI_{z_0\in C}, \;\; z_0 \in \mathcal{X}.$$

Hence the drift criterion is satisfied which together with Theorem 3.5 and Corollary 3.6 implies that $\{z_t\}$ is geometrically ergodic (Meyn and Tweedie [29], Theorem 15.0.1). □

References

[1] Ardia, D.: Bayesian estimation of a Markov-switching threshold asymmetric GARCH model with Student-t innovations. *The Econometrics Journal* **12**, (2009), 105–126. MR-2500199

[2] Bauwens, L., Preminger A., and Rombouts, J.: Theory and inference for a Markov-switching GARCH model. *The Econometrics Journal* **13**, (2010), 218–244. MR-2722883

[3] Basrak, B., Davis, R. A. and Mikosch, T.: Regular variation of GARCH processes. *Stochastic Processes and their Applications* **99**, (2002), 95–115. MR-1894253

[4] Black, F.: Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association*, (1976), 177–181.

[5] Bougerol, P. and Picard, N.: Strict stationarity of generalized autoregressive processes. *Annals of Probability* **20**, (1992), 1714–1730. MR-1188039

[6] Brandt, A.: The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability* **18**, (1986), 211–220. MR-0827336

[7] Cai, J.: A Markov model of switching regime ARCH. *Journal of Business and Economic Statistics* **12**, (1994), 309–316.

[8] Carrasco, M. and Chen, X.: Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18**, (2002), 17–39. MR-1885348

[9] Čížek, P. and Spokoiny, V.: Varying coefficient GARCH models. In Handbook of Financial Time Series edited by T. G. Andersen, R. A. Davis, J. P. Kreiβ and T. Mikosch. *Springer-Verlag*, Berlin Heidelberg, 2009, 169–183.

[10] Dahlhaus, R. and Subba Rao, S.: Statistical inference for time varying ARCH processes. *The Annals of Statistics* **34**, (2006), 1075–1114. MR-2278352

[11] Engle, R. F. and Ng, V. K.: Measuring and testing the impact of news on volatility. *Journal of Finance* **48**, (1993), 1749–1778.

[12] Francq, C. and Zakoïan, J. M.: The L^2-structures of standard and switching-regime GARCH models. *Stochastic Processes and their Applications* **115**, (2005), 1557–1582. MR-2158020

[13] Francq, C. and Zakoïan, J. M.: Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* **22**, (2006), 815–834. MR-2291218

[14] Fryzlewicz, P. and Subba Rao, S.: On some mixing properties of ARCH and time-varying ARCH processes. *Bernoulli* **17**, (2011), 320–346. MR-2797994

[15] Granger, C. W. J.: Non-linear models: where do we go next-time varying parameter models?. *Studies in Nonlinear Dynamics & Econometrics* **12**, (2008), 1–9. MR-2443320

[16] Gray, S. F.: Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics* **42**, (1996), 27–62.

[17] Haas, M., Mittnik, S. and Paolella, M.: A new approach to Markov switching GARCH models. *Journal of Financial Econometrics* **2**, (2004), 493–530.

[18] Hamilton, J. D. and Susmel, R.: Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics* **64**, (1994), 307–337.

[19] Hardy, G. H., Littlewood, J.E. and Polya, G.: Inequalities. Second edition, *Cambridge University Press*, Cambridge, 1988. MR-0944909

[20] He, Z. and Maheu, J. M.: Real time detection of structural breaks in GARCH models. *Computational Statistics and Data Analysis* **54**, (2010), 2628–2640. MR-2720465

[21] Hentschel, L.: All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* **39**, (1995), 71–104.

[22] Hwang, S. and Pereira, P. L. V.: The effects of structural breaks in ARCH and GARCH parameters on persistence of GARCH models. *Communications in Statistics - Simulation and Computation* **37**, (2009), 571–578. MR-2749970

[23] Klaassen, F.: Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics* **27**, (2002), 363–394.

[24] Kokoszka, P. and Leipus, R.: Change point estimation in ARCH models. *Bernoulli* **6**, (2000), 513–539. MR-1762558

[25] Kristensen, D.: On stationarity and ergodicity of the bilinear model with applications to GARCH models. *Journal of Time Series Analysis* **30**, (2005), 125–144. MR-2488638

[26] Lamoureux, C. G. and Lastrapes W. D.: Persistence in variance, structural change, and the GARCH model. *Journal of Business and Economic Statistics* **8**, (1990), 225–234.

[27] Meitz, M. and Saikkonen, P.: Ergodicity, mixing, and existence of moments of a class of Markov models with applications to GARCH and ACD models. *Econometric Theory* **24**, (2008a), 1291–1320. MR-2440741

[28] Meitz, M. and Saikkonen, P.: Stability of nonlinear AR-GARCH models. *Journal of Time Series Analysis* **29**, (2008b), 453–475. MR-2410184

[29] Meyn, S. P. and Tweedie R. L.: Markov Chains and Stochastic Stability *Springer Verlag*, London, 1993. MR-1287609

[30] Mikosch, T. and Starica, C.: Non-stationarities in financial time series, the long-range dependence and the IGARCH effects. *Review of Economics and Statistics* **86**, (2004), 378–390.

[31] Nelson, D. B.: Stationarity and persistence in the GARCH (1,1) model. *Econometric Theory* **6**, (1990), 318–334. MR-1085577

[32] Rohan, N. and Ramanathan, T. V.: Asymmetric volatility models with structural breaks. *Communications in Statistics- Simulations and Computations* **41**, (2012), 1519–1543. MR-2924000

[33] Rohan, N. and Ramanathan, T. V.: Non-parametric estimation of a time-varying GARCH model. *Journal of Non-parametric Statistics* **25**, (2013), 33–52. MR-3039969

[34] Saïdi, I. and Zakoïan, J. M.: Stationarity and geometric ergodicity of a class of non-linear ARCH models. *Annals of Applied Probability* **16**, (2006), 2256–2271. MR-2288721

[35] Stelzer, R.: On Markov-switching ARMA processes-stationarity, existence of moments and geometric ergodicity. *Econometric Theory* **25**, (2009), 43–62. MR-2472046

[36] Tjostheim, D.: Non-linear time series and Markov chains. *Advances in Applied Probability* **22**, (1990), 587–611. MR-1066965

Acknowledgments. Authors are thankful to the anonymous referee and the associate editor for some useful comments