Statistica Sinica 4(1994), 71-87

PROBABILISTIC PROPERTIES OF THE β-ARCH MODEL

Dominique Guégan and Jean Diebolt

Université Paris XIII and Université Paris VI

Abstract: In the present paper we consider the main probabilistic properties of the Markov chain $X_t = aX_{t-1} + [a_0 + (a_1^+(X_{t-1})^+ + a_1^-(X_{t-1})^-)^{2\beta}]^{1/2}\varepsilon_t$, that we call the β-ARCH model. We examine the invertibility, irreducibility, Harris recurrence, ergodicity, geometric ergodicity, α-mixing, existence and nonexistence of finite moments and exponential moments of some order and sharp upper bounds for the tails of the stationary density of the process $\{X_t\}$ in terms of the common density of the ε_t's.

Key words and phrases: Markov chain, invertibility, ergodicity, mixing, tail of the stationary density, ARCH model, nonlinear time series, autoregressive.

1. Introduction

The purpose of this article is to study the probabilistic properties of a particular case of the semi-parametric d-dimensional homogeneous Markov chain $\{\boldsymbol{X}_t; t \in \boldsymbol{N}\}$ defined by the recursive scheme:

$$\boldsymbol{X}_t = T(\boldsymbol{X}_{t-1}) + \sigma(\boldsymbol{X}_{t-1})\varepsilon_t, \tag{1.1}$$

where the $\boldsymbol{R}^d \to \boldsymbol{R}^d$ function $T(x)$ is such that $|T(x)| < \rho|x|$, $0 \le \rho < 1$ for all x in $\boldsymbol{R}^d$, $\sigma(x)$ is an unbounded function defined on $\boldsymbol{R}^d$ such that $|\sigma(x)| \to \infty$ as $|x| \to \infty$, $\sigma(x)$ has a differential at each point x such that $|x| \ge C$, for some positive C, and $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean 0 and variance $\boldsymbol{I}$ whose common distribution has a positive density $\mu(x)$ with respect to Lebesgue measure on $\boldsymbol{R}^d$. The class of models of time series defined by (1.1) embodies various nonlinear models. In this paper, we are particularly interested in the special case where $d = 1$ and $\{X_t\}$ is defined by the following equation:

$$X_t = T(X_{t-1}) + \sigma(X_{t-1})\varepsilon_t \tag{1.2}$$

with

$$T(x) = ax \tag{1.3}$$

and

$$\sigma(x) = \left[a_0 + (a_1^+ x^+ + a_1^- x^-)^{2\beta}\right]^{1/2}, \tag{1.4}$$

where $\alpha^+ = \max(\alpha, 0)$ and $\alpha^- = \max(-\alpha, 0)$, and β, a_0, a_1^+, a_1^- are positive constants.

We call this model, the β-ARCH model. When $\beta = 1$, the β-ARCH model reduces to the classical ARCH model introduced by Engle (1982) when $a = 0$, and by Weiss (1986) when $a \neq 0$. The models (1.2)-(1.4) are important for financial applications because they take into account the great variability of the data, allow for modelling of leptokurtik marginal distributions, avoid the limitation of the quadratic form of the ARCH model and allow the variance of the noise to follow the level of the process.

A possible relevant extension of the β-ARCH model is the β-GARCH model, which makes the conditional variance a weighted average of all past residuals. In this paper, we focus on the β-ARCH model, since it is much more tractable. However, a heuristic discussion of the β-GARCH model is included in the conclusion.

In this paper, we only investigate the case where $\sigma(x)$ is unbounded, since the case $\sigma(x)$ bounded has been extensively studied in the literature (see Section 2). We focus on results concerning the existence of polynomial and exponential moments, and on the asymptotic behavior of the stationary distribution of X_t, in terms of the common density of the ε_t's. Indeed, the existence of certain exponential moments and the knowledge of sharp upper bounds for the tails of the stationary distribution of (1.2) are of primary importance, since they lead to get general results on the fluctuations of the partial sums $S_n = X_1 + \cdots + X_n$ of the process $\{X_t\}$ and of the extreme values $X_{1,n}$ and $X_{n,n}$, where $X_{1,n} \leq \cdots \leq X_{n,n}$ denote the order statistics of the sample $\{X_1, \ldots, X_n\}$. In a companion paper we will consider the estimation of β using the results of this paper and the related properties of the extreme values $X_{1,n}$ and $X_{n,n}$, as $n \to \infty$.

The article is organized as follows: in Section 2, we give a brief survey of known probabilistic properties of (1.1) when $\sigma(x)$ is bounded and show how these properties can be applied to (1.2)-(1.4). In Section 3, we obtain conditions for the existence or non existence of exponential moments for (1.2) in the setup where $\sigma(x)$ is non bounded (Theorem 1). Then we derive results concerning sharp upper bounds for the tails of the stationary density of (1.2) in terms of the common density of the ε_t's. This is the object of Theorem 2. We present specific results concerning the ARCH model ($\beta = 1$ in (1.4)) in Theorem 3. The results that we obtain in this direction convey the idea that the classical AR model with ARCH noise is at the boundary of the set of ergodic models (1.2)–(1.4). Section 4 is devoted to discussion of some extensions of β-ARCH models. The proofs of

Theorems 1, 2 and 3 are postponed to Section 5.

2. Probabilistic Properties of the Markov Chain (1.2)–(1.4), with $\sigma(x)$ Bounded

In this section, we first consider the process defined by (1.1) in the case where $\sigma(x)$ is bounded. This Markov chain has been extensively investigated, and properties such as invertibility, irreducibility, aperiodicity, Harris-recurrence, ergodicity, geometric ergodicity and mixing have been studied in many papers. For example, Jones (1978) introduces a more restrictive class of models (scalar case, $\sigma(x)$ constant); Doukhan and Ghindès (1980) investigate such properties of (1.1) as irreducibility, ergodicity and geometric ergodicity for the case where $T(x)$ is continuous, $\sigma(x)$ is constant and the common distribution of the ε_t's is strongly symmetric; Chan and Tong (1985) present the first contribution to the case where $T(x)$ is continuous and $\sigma(x)$ is non-invertible, establishing the geometrical ergodicity of certain Markov chains (1.1) under the assumption of the existence, in a neighbourhood of 0, of a Lyapunov function of the discrete-time dynamical system generated by $T(x)$; Chan (1990) establishes the geometric ergodicity of the Markov chain defined by $X_t = f(X_{t-1}, \varepsilon_t)$ in connection with the stability of the deterministic difference equation $X_t = T(X_{t-1})$, Tjostheim (1990) investigates ergodicity, null recurrence and transience of (1.1) with $\sigma(x) =$ constant. Moreover, Mokkadem (1987) deals with a seemingly more general model, namely, $\boldsymbol{X}_t = T(\boldsymbol{X}_{t-1}) + \varepsilon(\boldsymbol{X}_{t-1}, t)$, using Tweedie's (1975, 1983) results. Finally, in a recent paper, Meyn and Guo (1990) study the geometric ergodicity of general Markov chains of the form (1.1) using an approach directly related to Control Theory. Diebolt and Guégan (1990) investigate most of the probabilistic properties of the Markov chain (1.1) in a thorough manner. It seems that the study of the Markov process (1.1) in the case where $\sigma(x)$ is bounded is almost complete (see Tjostheim (1990)).

We now consider the process defined by (1.2)–(1.4), with the following assumption: (H_1) The common distribution of the ε_t's has a density $\mu(x)$ with respect to the Lebesgue measure λ which is positive almost everywhere.

Under the assumption (H_1), if a_1^+, a_1^- and a_0 are positive constants, then the process (1.2)–(1.4) has the following properties.

(i) It is invertible.

(ii) It is an aperiodic, ν-irreducible and ν-Harris recurrent Markov chain for each ν such that $\lambda \ll \nu \ll \lambda$.

If in addition we suppose that $\beta < 1$, then:

(iii) A necessary and sufficient condition for geometric ergodicity is $|a| < 1$. In this case, (1.2)–(1.4) is α-mixing with a geometric rate of convergence.

(iv) If $|a| < 1$ and $E|\varepsilon|^s < \infty$ for an $s \geq 1$, then (1.2)–(1.4) has finite moments of order $\leq s$.

These properties are easily verified: Assertion (i) can be derived from the invertibility of (1.4). Assertion (ii) follows from the fact that the Markov probability transition kernel $P(x, B)$, with B in the Borel σ-field of $\boldsymbol{R}$ and $x \in \boldsymbol{R}$, has a positive density given by

$$p(x,y) = \sigma(x)^{-1}\mu\Big[\sigma(x)^{-1}\{y - ax\}\Big]. \tag{2.1}$$

To obtain Assertion (iii) we use Tweedie's (1975, p.394) result: If $|a| < 1$ and $\beta < 1$, then

$$\begin{aligned}&\Big\|ax + \Big(a_0 + (a_1^+x^+ + a_1^-x^-)^{2\beta}\Big)^{1/2}\varepsilon\Big\|_s \\ \leq\ &\Big\|ax + \Big((a_0 + a_1^+x^+ + a_1^-x^-)^{2\beta}\Big)^{1/2}\varepsilon\Big\|_s \leq |a||x|(1 + o(1)) \text{ as } |x| \to \infty,\end{aligned}$$

which implies that

$$\Big\|ax + \Big((a_0 + a_1^+x^+ + a_1^-x^-)^{2\beta}\Big)^{1/2}\varepsilon\Big\|_s \leq \rho|x| \text{ for } |x| \text{ large enough,}$$

where ρ is such that $0 \leq |a| < \rho < 1$. Here $\|Y\|_s = E^{1/s}(|Y|^s)$ denotes the L^s-norm ($1 \leq s \leq \infty$). If $|a| = 1$, then $EX_t = 0$ and $E(X_t^2) \geq a_0 t \to \infty$ as $t \to \infty$. If (2.1) had a stationary distribution ϕ^* and the distribution of X_0 was ϕ^*, then the distribution of X_t would be ϕ^* for all t, implying that $E(X_t^2)$ would be constant. Hence the contradiction. To obtain Assertion (iv) we can construct positive constants M, δ and $s \geq 1$, such that

$$E|ax + \sigma(x)\varepsilon|^s \leq |x|^s - \delta \text{ for all } x \text{ such that } |x| > M.$$

(See Tweedie (1983, p.192).)

Remark 1. For the case $\beta = 1$, (1.2)–(1.4) reduces to the classical ARCH model introduced by Engle (1982). Note that a sharp sufficient condition for geometric ergodicity then becomes $a^2 + a_1 < 1$. For $a = 0$ and ε Gaussian, Engle (1982) and Bollerslev (1986) have established sufficient conditions for the existence of the moments of some even integer order. Pantula (1988) has established the ergodicity of the ARCH model with $a = 0$ under the assumption $a_1 < 1$.

3. Existence or Nonexistence of Polynomial and Exponential Moments, with $\sigma(x)$ Non Bounded

In this section we are concerned with the process (1.2) under Assumption (H_1) and we assume that $\sigma(x)$ defined by (1.4) has a particular form. More precisely we assume that $\sigma(x)$ is differentiable, $|\sigma(x)| \to \infty$ as $|x| \to \infty$ and $\sigma(x) \sim b|x|^\beta$ as $|x| \to \infty$, with $b > 0$ and $\beta \geq 0$. We assume that $a = 0$ (so we are dealing with a pure β-ARCH model). Under the hypothesis that ε has all its polynomial moments, Theorem 1 states that the invariant density of (1.2) has all its polynomial moments if $0 \leq \beta < 1$. In contrast, (1.2) is not ergodic if $\beta > 1$. In Theorem 2, we address the problem of the existence of exponential moments if $\mu(x) \leq C\exp(-k|x|^\gamma)$ for $|x|$ large enough, where $k > 0$, $\gamma \geq 1$ and C are constants. We prove that the stationary distribution has all its exponential moments if $\beta < (\gamma-1)/\gamma$. If $\mu(x) \geq C\exp(-k|x|^\gamma)$ for $|x|$ large enough, then the stationary distribution has no exponential moment if $\beta > (\gamma-1)/\gamma$ or, in the special case where $\gamma = 1$, for all β, $0 < \beta < 1$. The proofs of the results of this section are postponed to Section 5.

Theorem 1. *Assume that* (H_1) *holds and* $a = 0$. *Assume in addition that*

$$\begin{aligned} &\sigma(x) \text{ is increasing on } [0,\infty),\\ &\sigma(x) \sim b|x|^\beta \text{ as } |x| \to \infty \text{ for some } b > 0 \text{ and } \beta \geq 0,\\ &\quad \sigma(x) \geq \sigma_0 > 0 \text{ for all } x,\\ &\sigma'(x) \sim \operatorname{sgn}(x)\beta b|x|^{\beta-1} (\text{ where } \operatorname{sgn}(x) = 1 \text{ if } x > 0 \text{ and } -1 \text{ if } x < 0),\\ &\quad \text{as } |x| \to \infty. \end{aligned} \tag{3.1}$$

(i) *If* $\beta < 1$ *and if* ε *has all its polynomial moments, then the invariant density of* (1.2) *has all its polynomial moments.*

(ii) *If* $\beta > 1$ *and* $b > 0$, *then* (1.2) *is not ergodic.*

Theorem 2. *Assume that* (H_1) *holds,* a_1^+, a_1^- *and* a_0 *are positive. Assume in addition that* $\beta < 1$ *and* $a = 0$.

(i) *If* ε *has all its polynomial moments with*

$$\mu(x) \leq C\exp(-k|x|^\gamma) \text{ for } |x| \text{ large enough,} \tag{3.2}$$

where $k > 0$, $\gamma \geq 1$ *and* C *are constants and if* $\beta < (\gamma-1)/\gamma$, *then* (1.2) *has a moment generating function defined in a neighbourhood of* 0.

(ii) *If* ε *has all its polynomial moments with*

$$\mu(x) \geq C\exp(-k|x|^\gamma) \text{ for } |x| \text{ large enough,} \tag{3.3}$$

then, if $\beta > (\gamma-1)/\gamma$, (1.2) *has no exponential moment. Moreover, if* $\gamma = 1$, *then* (1.2) *has no exponential moment whatever* β, $0 < \beta < 1$.

Remark 2. (a) Classical densities to which Theorems 1 and 2 can be applied include the Gaussian ($\gamma = 2$), the bilateral exponential (Laplace) ($\gamma = 1$), the Rayleigh ($\gamma = 2$) densities, the Generalized Error Distribution (GED) defined by $\mu(x) = C\exp(-k|x|^\gamma)$, where C and k only depend on the parameter γ (see, e.g., Nelson (1991)), the bilateral Weibull density defined by $\mu(x) = \frac{1}{2}\gamma\theta|x|^{\gamma-1}\exp\{-\theta|x|^\gamma\}$ for $\gamma \geq 1$, the bilateral Gamma density, $\mu(x) = \frac{k^p}{2\Gamma(p)}|x|^{p-1}$ $\exp\{-\theta|x|\}$ for $p \geq 1$ (in which case $\gamma = 1$) and the bilateral logistic, $\mu(x) = \frac{e^{-|x|}}{2(1+e^{-|x|})^2}$ (again with $\gamma = 1$).

(b) If we assume that for all $\gamma > 1$,

$$\mu(x) = o(\exp(-|x|^\gamma)) \text{ as } |x| \to \infty, \tag{3.4}$$

then the process defined by the assumptions of Theorem 1 has a moment generating function defined everywhere for all β, $0 \leq \beta < 1$. Note that the density $\mu(x) = C\exp(-e^{|x|})$ derived from the Gumbel density, satisfies (3.1).

(c) Broniatowski (1990) has studied estimation of the Weibull tail coefficient $\alpha > 0$ in the tail Weibull distribution defined by $1 - F(x) = \exp\{-x^{1/\alpha}L(x)\}$, $x \geq 0$ and $F(x) = \exp\{-|x|^{1/\alpha}L(x)\}$, $x \leq 0$, where the function $L(x)$ is slowly varying towards infinity. Our results can be applied to such distributions in case $L(x)$ has a derivative and $\alpha \geq 1$.

(d) In general, if $\beta = (\gamma - 1)/\gamma$, no conclusion can be drawn concerning the existence of exponential moments for processes (1.2) with $\sigma(x)$ defined by (3.1).

(e) Theorem 2 shows that, if $0 \leq \beta < 1$, $a = 0$ and the density $\mu(x)$ of the distribution of ε is assumed to have all its exponential moments, then the r.v.'s in (1.2) have all their exponential moments whenever β remains below $(\gamma-1)/\gamma$. On the contrary, the r.v.'s in (1.2) have no exponential moment for $\beta > (\gamma-1)/\gamma$. Since the parameter γ determines the tail behaviour of $\mu(x)$, our result highlights the strong relationship between the existence of exponential moments for (1.2) and the tail behavior of $\mu(x)$. If $\beta < (\gamma-1)/\gamma$, then rough versions of the large deviations inequality can be derived from the generalizations of Bernstein's inequality to ϕ-mixing processes (Bosq (1975)) and to α-mixing nonstationary processes (Carbon (1983)). An extension of Bosq and Carbon's results has been obtained in White and Wooldridge (1991), using truncation. The actual derivation of such results would be basic to obtaining approximate confidence intervals from the perspective of statistical inference in the context of (1.2). On the other hand, if $\beta > (\gamma - 1)/\gamma$, then the qualitative behavior of the process $\{X_t\}$ profoundly differs from the case $\beta < (\gamma-1)/\gamma$. This can be seen by examining the behaviour of the extremes $X_{1,n}$ and $X_{n,n}$ of the sample $\{X_1, \ldots, X_n\}$ as $n \to \infty$. Indeed, since $\{X_t\}$ is stationary and α-mixing, their limiting distribution only depends on the tail behaviour of the stationary distribution of $\{X_t\}$ (see, e.g., Leadbetter, Lindgren and Rootzén (1983)). Furthermore, from a modelling point of view, Theorem 2 shows that an erroneous choice for the distribution of ε may lead to some problems concerning the estimation of the parameter β. In conclusion, Theorem 2 provides a fresh insight into our newly introduced β-ARCH model defined by (1.2)–(1.4), which makes it interesting.

(f) In econometric contexts, several authors have considered non-Gaussian ε densities to model time series with heavy tails. In general they have focused on the case $\varepsilon > 0$ a.s. For instance, they have considered GED distributions. For a bibliographical survey, see Bollerslev, Chou and Kroner (1992).

We now establish new results concerning the existence of moments of the ARCH model defined as:

$$X_t = aX_{t-1} + (a_0 + a_1 X_{t-1}^2)^{1/2}\varepsilon_t, \tag{3.5}$$

where a_0 and a_1 are positive constants. This model has been introduced for the case $a = 0$ by Engle (1982) and by Weiss (1986) for $a \neq 0$; in both cases, ε has a standard Gaussian distribution. We assume the following hypotheses on the distribution of ε:

(H_2) There exists a bounded function $m : [0,\infty) \to [0,\infty)$, decreasing for large values of x, satisfying $\int_0^\infty r^{d-1}m(r)dr < \infty$ and $\mu(x) \le m(|x|)$ for x a.e. in $\boldsymbol{R}$.

(H_3) (i) For each $\alpha > \beta > 0$, $r^{d-1}m(\alpha r) = o\{m(\beta r)\}$ as $r \to \infty$.

(ii) For each $\alpha > \beta > 0$, $u^{-1}(r)m(\alpha r) = o\{m(\beta r)\}$ as $r \to \infty$,

where $u(r) = \frac{m(r)}{\int_t^\infty m(t)dt}$ $(r \ge 0)$ represents the hazard rate of the pseudo-density $m(r)$.

Theorem 3. *Let the process $\{X_t\}$ be defined by (3.5) and assume that $a = 0$ and $a_1^{1/2}\|\varepsilon\|_1 < 1$. Under the assumption ($H_1$)-($H_3$) with $\mu(x) = m(|x|)$, if there exists an $s > 1$ such that $a_1 > \frac{1}{\|\varepsilon\|_s^2}$, $a_0^{(s-1)/s} < \frac{1-a_1(E|\varepsilon|)^2}{a_1^{1/s}(E|\varepsilon|)^2}$ and $a_0^{(s-1)/s} < a_1^{(s-1)/s}(a_1^s\|\varepsilon\|_s^{2s} - 1)$, then (3.5) has no moment of order $\ge s$.*

Remark 3. (a) The assertions of Theorems 1, 2 and 3 cannot be derived from Tweedie's results, since their proofs make use of lower bounds for the transition kernel density.

(b) For the case $\beta = 1$ and $a = 0$, Theorem 3 shows the nonexistence of higher order polynomial moments of the invariant density. This is a basic property of AR processes with ARCH noise (i.e. $\beta = 1$). Furthermore, this property disappears when $\beta < 1$. It can be proved that if $\beta = 1$ and $b > 1/E|\varepsilon|$ then (3.5) is not ergodic, as for the (1.2)-(1.4) model when $\beta > 1$. This conveys the idea that the classical AR model with ARCH noise ($\beta = 1$) is located at the boundary of ergodic models of this kind, and, thus, appears as a limiting case. Note that Nelson (1991) proves, for the case $a = 0$, $a_1^+ = a_1^-$, $\beta = 1$, that X_t is ergodic if and only if $E[\log(a_1^2 \varepsilon_t^2)] < 0$, and that $E[|X|^{2p}] < \infty$ if and only if $E[(a_1^2 \varepsilon_t^2)^2] < 1$.

(c) Theorem 3 shows that for each $s > 1$, there exist values of a_0 and a_1 for which (3.5) has no moment of order s, whenever the density $\mu(x)$ satisfies (H_1)-(H_3). In particular, the stationary distribution of (3.5) can have infinite variance, even if the density $\mu(x)$ of ε decreases to 0 very quickly as $|x| \to \infty$. Note that, as far as we know, such a property has never been proved.

4. Conclusion

In general, obtaining stationary, ergodic and moment properties of ARCH models can be quite difficult. This paper illustrates several methods of deriving such properties in the context of a new ARCH model, as the β-ARCH model. The β-ARCH(1) model is basically an extension of the ARCH(1) model of Engle (1982), and shares a major limitation of that model, namely that it makes the conditional variance a function of only one lag of X_t.

Possible extensions of the ARCH(1) model are the ARCH(p) and the GARCH (p, q) models (Bollerslev (1986)). The latter makes the conditional variance a weighted average of all past residuals. In empirical applications (see e.g., Bollerslev, Chou and Kroner (1992)) the "GARCH" terms are almost always highly statistically significant.

Corresponding extensions of the β-ARCH(1) model discussed in this paper can be considered, namely,

— the β-ARCH(p) model, defined by

$$X_t = aX_{t-1} + \left[a_0 + \left(\sum_{i=1}^{p}(a_i^+ X_{t-i}^+ + a_i^- X_{t-i}^-)\right)^{2\beta}\right]^{1/2} \varepsilon_t \tag{4.1}$$

and

— the β-GARCH(p, q) model, defined by

$$X_t = \varepsilon_t h_t^{1/2} \tag{4.2}$$

with

$$h_t = \sum_{j=1}^{q} b_j h_{t-j} + \left[a_0 + \left(\sum_{i=1}^{p}(a_i^+ X_{t-i}^+ + a_i^- X_{t-i}^-)\right)^{2\beta}\right]^{1/2},$$

where ε_t is independent from $h_t, h_{t-1}, \ldots$. Unfortunately, since the techniques used in this paper can only deal with the Markovian models of the form (1.1) with invertible $\sigma(x)$, our results cannot be directly extended to the models (4.1)–(4.2) for $p \geq 2$.

For the case where $p = q = 1$, the β-GARCH(1,1) model can be rewritten as:

$$X_t = \varepsilon_t h_t^{1/2} \tag{4.3}$$

with,

$$h_t = b_1 h_{t-1} + \sigma^2 \left(\varepsilon_{t-1} h_{t-1}^{1/2}\right),$$

where $b_1 \geq 0$ and $\sigma(x)$ is defined by (1.4). Thus, in this case, $\{h_t\}$ is an autoregressive nonlinear Markovian process, but the recursion scheme (4.3) does not have the form (1.2) and appears to be much more difficult to handle than the recursion scheme (1.2)–(1.4).

However the ergodic and moment properties of autoregressive processes like (4.3) depend heavily on the response of the system to large values of $|\varepsilon_{t-1}|$ and h_{t-1}. But for such values of $|\varepsilon_{t-1}|$ and h_{t-1}, $\sigma^2(\varepsilon_{t-1} h_{t-1}^{1/2})$ can be approximated by $C|\varepsilon_{t-1}|^{2\beta} h_{t-1}^{\beta}$, for some positive constant C. Thus, (4.3) can be expected to exhibit the same kind of ergodic behaviour as a process of the form:

$$X_t = \varepsilon_t h_t^{1/2} \tag{4.4}$$

with

$$h_t = b_1 h_{t-1} + C h_{t-1}^{\beta} \eta_{t-1},$$

where $\eta_{t-1} = |\varepsilon_{t-1}|^{2\beta}$ is positive and indepedent of $h_{t-1}, h_{t-2}, \ldots$. Since (4.4) has a structure similar to that of (1.2)–(1.4), we conjecture that our results can be adapted to (4.3). The above discussion suggests that our results concerning several basic properties of the β-ARCH model capture the main features of the corresponding properties of the β-GARCH(1,1) model.

5. Proofs of the Theorem

Proof of Theorem 1

(i) Define, for $\alpha > 1$ and a positive constant A to be specified later,

$$C(\alpha, A) = \left\{\phi \in L^1(\boldsymbol{R}; \lambda) : \int_{-\infty}^{\infty} |\phi(x)| dx \leq 1 \text{ and } |\phi(x)| \leq |x|^{-\alpha} \right.$$
$$\left. \text{for a.e. } x \text{ such that } |x| \geq A \right\}.$$

Suppose for simplicity that $\phi(-x) = \phi(x)$ (without loss of generality, since $\mu(-x) = \mu(x)$), and choose $y \geq A$. Let H denote the linear operator of $L^1(\boldsymbol{R})$ defined by

$$H\phi(y) = \int_E p(x, y)\phi(x) dx.$$

Then, for $\phi \in C(\alpha, A)$, we have

$$H\phi(y) = I + II + III, \tag{5.1}$$

where

$$|I| = \left|2\int_0^1 \frac{1}{\sigma(x)}\mu\left\{\frac{y}{\sigma(x)}\right\}\phi(x)dx\right| \le 2\mu(y), \tag{5.2}$$

$$|II| = \left|2\int_1^A \frac{1}{\sigma(x)}\mu\left\{\frac{y}{\sigma(x)}\right\}\phi(x)dx\right| \le 2\sigma(1)^{-1}\mu\left\{\sigma(A)^{-1}y\right\} \tag{5.3}$$

and

$$|III| = \left|2\int_A^\infty \frac{1}{\sigma(x)}\mu\left\{\frac{y}{\sigma(x)}\right\}\phi(x)dx\right|. \tag{5.4}$$

Making the substitution $s = \frac{y}{\sigma(x)}$ in (5.4), we obtain

$$|III| = \left|2\int_0^{y/\sigma(A)} \frac{1}{\sigma'\{g(y/s)\}}\phi\left\{g\left(\frac{y}{s}\right)\right\}\frac{\mu(s)}{s}ds\right|,$$

where $g(x)$ denotes the inverse function of the restriction of σ to $[0,\infty)$. Since $g(x) \sim (\frac{x}{b})^{1/\beta}$ as $x \to \infty$ and $\phi \in C(\alpha, A)$, we have, for any $\varepsilon' > 0$,

$$|III| \le (1+\varepsilon')\beta^{-1}(b\|\varepsilon\|_{(\alpha-1)/\beta})^{(\alpha-1)/\beta}y^{-(\alpha+\beta-1)/\beta}. \tag{5.5}$$

Using (5.1)–(5.5), we obtain

$$|H\phi(y)| \le (1+\varepsilon')\beta^{-1}(\|\varepsilon\|_{(\alpha-1)/\beta})^{(\alpha-1)/\beta}y^{-(\alpha+\beta-1)/\beta} + o\left(\frac{1}{y^\alpha}\right). \tag{5.6}$$

Hence, if $\beta < 1$, then $\frac{\alpha+\beta-1}{\beta} > \alpha$, and from (5.6) it follows that $H\phi \in C(\alpha, A)$, which proves (i).

(ii) Following Tweedie (1976), it is enough to construct a nonnegative increasing function $g(x)$ defined on $[0,\infty)$ and a positive constant a, such that

$$(a) \quad 2\int_0^\infty \frac{1}{\sigma(x)}\mu\left\{\frac{y}{\sigma(x)}\right\}g(y)dy \ge g(x) \quad \text{for all} \quad x > a, \tag{5.7}$$

$$(b) \quad 2\int_0^\infty \frac{1}{\sigma(x)}\mu\left\{\frac{y}{\sigma(x)}\right\}|g(y)-g(x)|dy \le B \quad \text{for all} \quad x \in [0,\infty), \tag{5.8}$$

where B denotes some finite positive constant, and

$$(c) \quad g(x) > \sup_{0\le y\le a} g(y) \quad \text{for all} \quad x > a. \tag{5.9}$$

To this end, we define $g(x) = \lambda x$ for $0 \le x \le a$, where the positive constants a and λ will be precised in the course of the proof, and $g(x)$ increase from λa to $(1+\zeta)\lambda a$ as $x \to \infty$, $x > a$, for a given arbitrarily small constant $\zeta > 0$. Since $\sigma(x) \sim bx^{\beta}$ as $x \to \infty$, making the substitution $s = \frac{y}{\sigma(a)}$ in the integral in (5.7) yields

$$2\int_0^{\infty} \frac{1}{\sigma(x)}\mu\Big\{\frac{y}{\sigma(x)}\Big\}g(y)dy \ge \lambda b(1-\eta)\Big[2\int_0^a s\mu(s)ds\Big]x^{\beta} + \lambda a\Big[2\int_a^{\infty}\mu(s)ds\Big] \tag{5.10}$$

for all $x > a$, where $\eta > 0$ is some small constant. Now, the right-hand side of (5.10) is larger than $g(x)$ for all $x > a$ if a is chosen so large as to guarantee that this expression is larger than $(1+\zeta)\lambda a$, since $(1+\zeta)\lambda a = \sup_{x\in[0,\infty)} g(x)$. But this inequality holds whenever

$$\lambda b(1-\eta)\Big[2\int_0^a s\mu(s)ds\Big]a^{\beta} + \lambda a\Big[2\int_a^{\infty}\mu(s)ds\Big] \ge (1+\zeta)\lambda a. \tag{5.11}$$

Dividing both sides of (5.11) by λa yields

$$a^{\beta-1} \ge \frac{1+\zeta-2\displaystyle\int_a^{\infty}\mu(s)ds}{b(1-\eta)\Big[2\displaystyle\int_0^a s\mu(s)ds\Big]}. \tag{5.12}$$

Assume $\beta > 1$. Since $a^{\beta-1} \to \infty$ as $a \to \infty$, (5.12) holds for a large enough. Furthermore, conditions (b) and (c) are clearly satisfied with, e.g., $B = 2\lambda a$. This concludes the proof of (ii).

Proof of Theorem 2

In order to prove Theorem 2, we first need to establish a technical lemma.

Lemma 1. *Let*

$$J(z) = \int_0^z \exp\Big\{-\Big(s + \frac{\rho(z)}{s^{\mu}}\Big)\Big\}ds, \tag{5.13}$$

where $\mu > 0$, $\rho(z)$ *increases to* ∞ *as* $z \to \infty$ *and* $\rho(z) = o(z^{\mu+1})$ *as* $z \to \infty$. *Then, there exists* $z_0 > 0$ *such that*

$$J(z) \le \exp\Big\{-M\rho(z)^{1/(\mu+1)}\Big\} \quad \textit{for all} \quad z \ge z_0 \tag{5.14}$$

and

$$J(z) \ge C'\exp\Big\{-M'\rho(z)^{1/(\mu+1)}\Big\} \quad \textit{for all} \quad z \ge z_0 \tag{5.15}$$

for some positive constants C', M *and* M' *depending only on* μ.

Proof of Lemma 1. The proof basically relies on the substitution

$$t = \varphi_z(s) = s + \frac{\rho(z)}{s^\mu}, \quad 0 < s \leq z, \tag{5.16}$$

for all z large enough. For the sake of simplicity, we will denote $\varphi_z(s)$ more briefly by $\varphi(s)$ and, similarly, $\rho(z)$ by ρ, when no ambiguity can arise.

We define $s_{\min}(z) = s_{\min}$ as the unique value of s, $s > 0$, such that $\varphi'(s) = 0$. We have

$$s_{\min} = C_0 \rho^{1/(\mu+1)}, \tag{5.17}$$

where $C_0 = \mu^{1/(\mu+1)}$. The function $\varphi(s)$ takes its minimum value

$$\varphi_{\min} = \varphi(s_{\min}) = M_0 \rho^{1/(\mu+1)}, \tag{5.18}$$

where $M_0 = \mu^{1/(\mu+1)} + \mu^{-\mu/(\mu+1)}$, at $s = s_{\min}$. Moreover, the function $\varphi_1(s)$ defined by $\varphi_1(s) = \varphi(s)$ for $0 < s \leq s_{\min}$ is decreasing from ∞ to $\varphi_{\min}$, whereas the function $\varphi_2(s)$ defined by $\varphi_2(s) = \varphi(s)$ for $s_{\min} \leq s \leq z$ is increasing from $\varphi_{\min}$ to $\varphi(z)$. Thus, we will have to split the substitution (5.16) into two parts,

$$t = \varphi_1(s), \quad 0 < s \leq s_{\min}, \tag{5.19}$$

$$t = \varphi_2(s), \quad s_{\min} \leq s \leq z. \tag{5.20}$$

(i) First, we prove (5.14). Split $J(z)$ into $J(z) = J_1(z) + J_2(z)$, where

$$J_1(z) = \int_0^{s_{\min}} \exp\{-\varphi_1(s)\} ds,$$

$$J_2(z) = \int_{s_{\min}}^{z} \exp\{-\varphi_2(s)\} ds.$$

(a) We first seek upper bounds for $J_1(z) = I + II$, where

$$I = \int_0^{s_1} \exp\{-\varphi_1(s)\} ds$$

and

$$II = \int_{s_1}^{s_{\min}} \exp\{-\varphi_1(s)\} ds,$$

with s_1, $0 < s_1 < s_{\min}$, such that

$$|\varphi'(s)| \geq 1 \text{ for all } s \text{ in } (0, s_1] \text{ and } |\varphi'(s_1)| = 1. \tag{5.21}$$

Thus,

$$s_1 = C_1 \rho^{1/(\mu+1)}, \tag{5.22}$$

where $C_1 = (\frac{\mu}{2})^{1/(\mu+1)} < C_0$,

$$\varphi(s_1) = M_1 \rho^{1/(\mu+1)}, \tag{5.23}$$

and $M_1 > M_0$ is a constant depending only on μ. Making the substitution (5.19) in I and using (5.21)–(5.22) it follows that

$$I \leq \exp\Big(- M_1 \rho^{1/(\mu+1)}\Big). \tag{5.24}$$

Moreover,

$$II \leq (C_0 - C_1)\rho^{1/(\mu+1)} \exp\Big\{ - M_0 \rho^{1/(\mu+1)}\Big\}. \tag{5.25}$$

(b) We now turn to $J_2(z)$. Letting

$$s_2 = C_2 \rho^{1/(\mu+1)}, \tag{5.26}$$

where $C_2 = (2\mu)^{1/(\mu+1)} > C_0$, we have

$$\varphi(s_2) = M_2 \rho^{1/(\mu+1)}, \tag{5.27}$$

where $M_2 > M_0$ is a constant depending only on μ. As in (a), we split $J_2(z)$ into $J_2(z) = III + IV$, where

$$\begin{aligned} III &= \int_{s_{\min}}^{s_2} \exp\{-\varphi_2(s)\} ds \\ IV &= \int_{s_2}^{z} \exp\{-\varphi_2(s)\} ds. \end{aligned} \tag{5.28}$$

Proceeding as in (a), we obtain

$$III \leq (C_2 - C_0)\rho^{1/(\mu+1)} \exp\Big\{ - M_0 \rho^{1/(\mu+1)}\Big\} \tag{5.29}$$

$$IV \leq 2 \exp\Big\{ - M_2 \rho^{1/(\mu+1)}\Big\}. \tag{5.30}$$

Let M be any positive constant such that $M < M_0$. Collecting (5.24)–(5.30), it follows that $J(z) = o\{\exp(-M\rho^{1/(\mu+1)})\}$ as $z \to \infty$, which implies (5.14) in Lemma 1.

(ii) We now turn to the proof of (5.15). Proceeding along the same lines as in (i), we obtain

$$II + III \geq (C_2 - C_1)\rho^{1/(\mu+1)} \exp\Big\{ - M' \rho^{1/(\mu+1)}\Big\}, \tag{5.31}$$

where $M' = \max(M_1, M_2)$. Moreover,

$$I \geq \frac{\rho^{1/\mu}}{\mu} \int_{\varphi(s_1)}^{\infty} t^{-(\mu+1)/\mu} e^{-t} dt. \tag{5.32}$$

Using the classical result $\int_t^\infty t^\alpha e^{-t}dt \approx x^\alpha e^{-x}$ as $x \to \infty$ and taking into account that $\varphi(s_1) \to \infty$ as $z \to \infty$, we obtain from (5.32) that

$$I \geq \frac{1}{2\mu} M_1^{-(\mu+1)/\mu} \exp\Big\{ - M_1 \rho^{1/(\mu+1)} \Big\}, \tag{5.33}$$

for all z large enough.

(b) Finally,

$$IV \geq \frac{1}{2} \exp\Big\{ - M_2 \rho^{1/(\mu+1)} \Big\} \tag{5.34}$$

for all z large enough.

This completes the proof of Lemma 1.

Proof of Theorem 2

Proof of (i). This proof parallels that of Theorem 1 (i). We define

$$\begin{aligned} C(r, A) \;=\; & \Big\{ \phi \in L^1(\boldsymbol{R}; \lambda) : \int_{-\infty}^{\infty} |\phi(x)| dx \leq 1 \text{ and } |\phi(x)| \leq \exp(-r|x|) \\ & \text{for a.e. } \; x, |x| \geq A \Big\}, \end{aligned} \tag{5.35}$$

where r is a positive constant and $A > 0$ has to be chosen in the course of the proof. As in the proof of Theorem 1 (i), we suppose, for simplicity that $\phi(-x) = \phi(x)$ and $\sigma_0 = 1$, without loss of generality. For $\phi \in C(r, A)$, from (3.1), $|I|$ and $|II|$ have the same upper bounds as in (5.2) and (5.3), respectively. So we concentrate on $|III|$, as given by (5.4). By the substitution $s = \frac{y}{\sigma(x)}$ in (5.4), we obtain (5.5). Now the following upper bound for $|III|$ is obtained by replacing $\mu(x)$ by (3.2) and $\phi(g(\frac{y}{s}))$ by $\exp\{-rg(\frac{y}{s})\}$ (since $g(\frac{y}{s}) \geq A$ for all $0 < s \leq \frac{A}{\sigma(A)}$ and $y \geq A$). For any $\varepsilon' > 0$, we have

$$|III| \leq \frac{2C(1+\varepsilon')\mu}{b^{1/\beta}} k^{\mu(1-\beta)} y^{(1-\beta)/\beta} \int_0^z t^{-(1+\mu(1-\beta))} \exp\Big\{ - \Big(t + \frac{\rho(z)}{t^\mu} \Big) \Big\} \mathrm{d}t, \tag{5.36}$$

where $\mu = \frac{1}{\beta\gamma}$,

$$z = k \left[\frac{y}{\sigma(A)} \right]^\gamma \tag{5.37}$$

and

$$\rho(z) = \frac{r(1-\varepsilon)}{b^{1/\beta} k^\mu} \sigma(A)^{1/\beta} z^\mu, \tag{5.38}$$

with $z \geq k[\frac{A}{\sigma(A)}]^\gamma$ for $y \geq A$ provided that A is large enough. The integral in the righthand side of (5.36) has the same behaviour for $z \to \infty$ as the integral $J(z)$

defined in Lemma 1. Moreover, since $\rho(z) =$ Constant. $\sigma(A)^{1/\beta} z^{\mu} = o(z^{\mu+1})$ as $z \to \infty$, the assumptions of Lemma 1 are satisfied. Thus,

$$|III| = O\Big(y^{(1-\beta)/\beta} \exp\{-M\rho(z)^{1/(\mu+1)}\}\Big)$$

as $z \to \infty$. Now, from (5.37) and (5.38),

$$|III| = O\Big(y^{(1-\beta)/\beta} \exp\{-K y^{\mu\gamma/(\mu+1)}\}\Big)$$

as $y \to \infty$, where $K = M[\frac{r(1-\varepsilon)k^{\mu}}{b^{1/\beta}}]^{1/(\mu+1)}$ does not depend on A. Since $\beta < \frac{\gamma-1}{\gamma}$, it follows that $\mu\gamma/(\mu+1) > 1$, implying that

$$|III| = o(\exp(-qy)) \quad \text{for all} \quad q > 0,$$

as $y \to \infty$. Hence, if A is chosen large enough, $|III| \leq (1/3)\exp(-ry)$ for all $y \geq A$, whereas $|I|$ and $|II| \leq (1/3)\exp(-ry)$ for all $y \geq A$ from (5.2) and (5.3) respectively. Thus, it follows that for all $r > 0$, there exists A large enough such that $\phi \in C(r, A)$ implies $H\phi \in C(r, A)$. As in the proof of Theorem 2 (i), this is sufficient for proving that the stationary density ϕ^* of any process (1.1) satisfying the assumptions of Theorem 2 is in $C(r, A)$ for all $r > 0$ and A large enough. This achieves the proof of part (i) of Theorem 3.

Proof of (ii). It can be proved along the same lines that, if the assumption (3.3) is in force and $\beta > \frac{\gamma-1}{\gamma}$, then for all $r > 0$ and all densities ϕ such that $\phi(x) \geq \exp(-r|x|)$ for all $|x| \geq A$ and A large enough, $H\phi(x) \geq \exp(-r|x|)$ for all $|x| \geq A$. Since the sequence $H^n\phi(x)$ converges to $\phi^*(x)$ a.e. as $n \to \infty$, the proof is complete.

Proof of Theorem 3

From Högnäs (1986), it is sufficient to examine the case $a = 0$. Let $c > 0$ be a constant which will be chosen in the course of the proof. If $|x| \leq c$, then $\sigma_0 \leq \sigma(x) \leq \sigma_1$, where

$$\sigma_0 = (a_0)^{1/2} \quad \text{and} \quad \sigma_1 = (a_0 + a_1 c^2)^{1/2}. \tag{5.39}$$

If $|x| > c$, then $b_0\,|x| \leq \sigma(x) \leq b_1\,|x|$, where

$$b_0 = (a_1)^{1/2} \quad \text{and} \quad b_1 = (a_1)^{1/2}\left(1 + \frac{a_0}{a_1 c^2}\right)^{1/2} = b_0\left(1 + \frac{a_0}{a_1 c^2}\right)^{1/2}. \tag{5.40}$$

We need the condition $b_1 E|\varepsilon| < 1$ for the process to be ergodic, i.e.

$$b_0\left(1 + \frac{a_0}{a_1 c^2}\right)^{1/2} E\,|\varepsilon| < 1. \tag{5.41}$$

To show that $\int_{-\infty}^{+\infty} |x|^s \phi^*(x)dx = \infty$ where $\phi^*(x)$ denotes the stationary density of (3.5), we need to verify the following two conditions:

$$\left(1+\frac{a_0}{a_1 c^2}\right)^{-1/(2s)} (a_1)^{1/2} \|\varepsilon\|_s > 1 \tag{5.42}$$

and

$$\frac{b_1\sigma_0}{b_0\sigma_1}\left(\frac{\sigma_0}{b_0}\right)^s \geq c^s. \tag{5.43}$$

Condition (5.43) is satisfied if

$$\left(1+\frac{a_0}{a_1 a^2}\right)^{1/2}\left(\frac{a_0}{a_0+a_1 a^2}\right)^{1/2}\left(\frac{a_0}{a_1}\right)^{s/2} \geq a^s. \tag{5.44}$$

Moreover, the constant c has to be chosen so that

$$c > \frac{(a_0)^{1/2}}{(a_1)^{1/2}[(a_1)^s(\|\varepsilon\|_s)^{2s}-1]^{1/2}}, \tag{5.45}$$

$$c \leq \left(\frac{a_0}{a_1}\right)^{1/(2s)} \tag{5.46}$$

and a_1 and a_0 have to be chosen so that $a_1 > \frac{1}{(\|\varepsilon\|_s)^2}$,

$$(a_0)^{(s-1)/s} < \frac{1-a_1(E|\varepsilon|)^2}{(a_1)^{1/s}(E|\varepsilon|)^2}, \tag{5.47}$$

and

$$(a_0)^{(s-1)/s} < (a_1)^{(s-1)/s}\Big[(a_1)^s(\|\varepsilon\|_s)^{2s}-1\Big]. \tag{5.48}$$

Finally, for all $0 < a_0 \leq 1$ and $a_1 < 1$, there exists $s > 1$ large enough such that (5.47)–(5.48) are satisfied, which completes the proof.

References

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

Bollerslev, T., Chou, R. and Kroner, K. (1992). ARCH modelling in finance: A selective review of the theory and empirical evidence, with suggestions for future research. *J. Econom.* Forthcoming.

Bosq, D. (1975). Inégalité de Bernstein pour les processus stationnaires et mélangeants. *C.R.A.S.* **281** (I), 1095–1098.

Broniatowski, M. (1990). On the estimation of the Weibull tail coefficient. Technical Report n° **118**, L.S.T.A., Paris VI.

Carbon, M. (1983). Inégalité de Bernstein pour les processus mélangeants, non nécessairement stationnaires. Applications. *C.R.A.S.* **297** (I), 303–306.

Chan, K. S. (1990). Deterministic stability, stochastic stability and ergodicity. *Nonlinear Time Series: A Dynamical System Approach* (Edited by H. Tong), 448–466. Oxford Science Publications, Oxford.

Chan, K. S. and Tong, H. (1985). On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probab.* **17**, 666–678.

Diebolt, J. and Guégan, D. (1990). Probabilistic properties of the general nonlinear Markovian process of order one and applications to time series modelling. Technical Report n° **125**. L.S.T.A. Paris VI.

Doukhan, P. and Ghindès, M. (1980). Etude du processus $X_{n+1} = f(X_n) + \varepsilon_n$. *C.R.A.S.* **290** (A), 921–923.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1007.

Högnäs, G. (1986). Comparison of some non-linear autoregressive processes. *J. Time Ser. Anal.* **7**, 205–211.

Jones, D. A. (1978). Nonlinear autoregressive processes. *Proc. Roy. Soc. London Ser.A* **360**, 71–95.

Leadbetter, M. R., Lindgren G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York.

Meyn, S. P. and Guo, L. (1990). Geometric ergodicity of a bilinear time series model. Technical Report, University of Illinois at Urbana Champaign.

Mokkadem, A. (1987). Sur un modèle autorégressif non linéaire, ergodicité et ergodicité géométrique. *J. Time Ser. Anal.* **8**, 195–204.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset return: a new approach. *Econometrica* **59**, 347–370.

Pantula, S. G. (1988). Estimation of autoregressive models with ARCH errors. *Sankhyā Ser. B* **50**, 119–138.

Tjostheim, D. (1990). Non-linear time series and Markov chains. *Adv. Appl. Probab.* **22**, 587–611.

Tweedie, R. L. (1975). Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **3**, 385–403.

Tweedie, R. L. (1976). Criteria for classifying general Markov chains. *Adv. Appl. Probab.* **8**, 737–771.

Tweedie, R. L. (1983). The existence of moments for stationary Markov chains. *J. Appl. Probab.* **20**, 191–196.

Weiss, A. A. (1984). ARMA models with ARCH errors. *J. Time Ser. Anal.* **5**(2), 129–143.

White, H. and Wooldridge, J. M. (1991). Some results on Sieve estimation with dependent observations. *Nonparametric and Semiparametric Methods in Econometrics and Statistics* (Edited by Barnett, Powell and Tauchen), Cambridge University Press.

C.S.P., Department of Mathematics, University Paris XIII, Av. J. B. Clément, 93430, Villetaneuse, France.

L.S.T.A., University Paris VI, 4 place Jussieu, Tour 45-55, 75252 Paris cedex 05, France.

(Received December 1991; accepted May 1993)