The Annals of Applied Probability
2001, Vol. 00, No. 00, 000–000

THE TAIL OF THE STATIONARY DISTRIBUTION OF AN AUTOREGRESSIVE PROCESS WITH ARCH(1) ERRORS

BY MILAN BORKOVEC AND CLAUDIA KLÜPPELBERG

Munich University of Technology

We consider the class of autoregressive processes with ARCH(1) errors given by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\,\varepsilon_n\,, \quad n \in \mathbb{N}\,,$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. random variables. Under general and tractable assumptions we show the existence and uniqueness of a stationary distribution. We prove that the stationary distribution has a Pareto-like tail with a well-specified tail index which depends on α, λ and the distribution of the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. This paper generalizes results for the ARCH(1) process (the case $\alpha = 0$) proved by Kesten (1973), Vervaat (1979) and Goldie (1991). The generalization requires a new method of proof and we invoke a Tauberian theorem.

1. Introduction. Recently there has been considerable interest in nonlinear time series models (see e.g. the books by Priestley (1988), Tong (1990) and Taylor (1995)). Many of these models were introduced to allow the conditional variance of a time series model to depend on past information (conditional heteroscedasticity). It has turned out that such models fit very well to many types of financial data. Early empirical work (see e.g. Mandelbrot (1963), Fama (1965)) has shown that large changes in equity returns and exchange rates, with high sampling frequency, tend to be followed by large changes settling down after some time to a more normal behavior. This observation leads to models of the form

$$X_n = \sigma_n \varepsilon_n\,, \quad n \in \mathbb{N}, \tag{1.1}$$

where the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. symmetric random variables with mean zero, and the volatility σ_n describes the change of (conditional) variance.

The autoregressive conditionally heteroscedastic (ARCH) models are one of the specifications of (1.1). In this case the conditional variance σ_n^2 is a linear function of the squared past observations. ARCH(p) models introduced by Engle (1982) are

Received

AMS 1991 subject classifications. Primary 60H25; secondary 60G10, 60J05.

Key words and phrases. ARCH model, autoregressive process, geometric ergodicity, heavy tail, heteroscedastic model, Markov process, recurrent Harris chain, regular variation, Tauberian theorem.

defined by

$$(1.2)\ \sigma_n^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j X_{n-j}^2 \,, \quad \alpha_0 > 0\,,\, \alpha_1, \dots, \alpha_{p-1} \geq 0,\, \alpha_p > 0\,, \quad n \in \mathbb{N},$$

where p is the order of the ARCH process.

In a series of papers, the ARCH model has been analyzed, generalized and used to test for time-varying risk premia in the financial market. We refer for instance to the survey article by Bollerslev, Chou and Kroner (1992). The most famous generalization to so-called generalized ARCH (GARCH) processes was proposed in Bollerslev (1986). The volatility σ_n is now a linear function of X_{n-1}, $X_{n-2}, \dots$ and σ_{n-1}, $\sigma_{n-2}, \dots$. ARCH and GARCH models are widely used to model financial time series since they capture certain empirical observations in financial data, namely the tendency for volatility clustering and the fact that unconditional price and return distributions tend to have fatter tails than the normal distribution.

The class of autoregressive (AR) models with ARCH errors introduced by Weiss (1984) are another extension. These models are also called SETAR-ARCH models (self-exciting autoregressive). They are defined by

$$(1.3) \qquad X_n = f(X_{n-1}, \dots, X_{n-k}) + \sigma_n \varepsilon_n, \quad n \geq k\,,$$

where f is again a linear function in its arguments and σ_n is given by (1.2). This model combines the advantages of an AR model which targets more on the conditional mean of X_n (given the past) and an ARCH model which concentrates on the conditional variance of X_n (given the past).

The class of models defined by (1.3) embodies various nonlinear models. In this paper we focus on the AR(1) process with ARCH(1) errors, i.e. $f(X_{n-1}, \dots, X_{n-k}) = \alpha X_{n-1}$ for some $\alpha \in \mathbb{R}$ and σ_n is given in (1.2) with $p = 1$. This Markovian model is analytically tractable and may serve as a prototype for the larger class of models (1.3).

The purpose of this article is to investigate the tail of the stationary distribution of the AR(1) process with ARCH(1) errors $(X_n)_{n\in\mathbb{N}}$. The model has also been considered by Diebolt and Guégan (1990) and Maercker (1997). For $\lambda = 0$ the process is an AR(1) process whose stationary distribution is determined by the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$, for ε_n normal it is a Gaussian process. In the ARCH(1) case (the case when $\alpha = 0$) the tail is known (see e.g. Goldie (1991) or Embrechts, Klüppelberg and Mikosch (1997), Section 8.4). The result was obtained by considering the square ARCH(1) process which leads to a stochastic difference equation which fits in the setting of Kesten (1973) and Vervaat (1979). This approach is, however, in general not possible or at least not obvious for $\alpha \neq 0$. Nevertheless for ε_n normal, provided a stationary distribution exists, a characteristic function argument transforms the model such that the results by Kesten (1973), Vervaat (1979) and Goldie (1991) may be applied. We refer to Remark 10 for further details.

For the general case we present another technique for evaluating the tail of the stationary distribution using the Drasin-Shea Tauberian theorem which can be found for instance in Bingham, Goldie and Teugels (1987). In contrast to Kesten (1973) and Goldie (1991), this approach has the drawback that it gives no information on the slowly varying function present in the tail of the stationary distribution. However, on the other side, the method also applies to processes which do not fit in the framework of Kesten (1973) or Goldie (1991). Furthermore, the Tauberian approach does not depend on additional assumptions which are often very hard to check (e.g. the existence of certain moments of the stationary distribution). See also the discussion in the introduction of Section 4. Combining our method with results in Goldie (1991), we finally specify the slowly varying function of the tail of the stationary distribution of $(X_n)_{n\in\mathbb{N}}$. Note that Goldie's results cannot be applied in the general case without the Tauberian approach. The Tauberian approach guarantees that the assumptions in Goldie (1991) are satisfied. The results in the present paper can be applied to study the behavior of the extremes and of the sample autocovariance and autocorrelation function of $(X_n)_{n\in\mathbb{N}}$; see Borkovec (2000) and Borkovec (2001).

The organization of this paper is as follows. In Section 2 we present the model and introduce the required assumptions on the innovations $(\varepsilon_n)_{n\in\mathbb{N}}$. We distinguish between the so-called general conditions and the technical conditions $(D.1)-(D.3)$. They are assumed to hold throughout this paper if it is not stated otherwise. In Section 3 we determine the parameter set of stationarity for our model and the tail of the stationary distribution. In Theorem 3 we summarize some probabilistic properties of $(X_n)_{n\in\mathbb{N}}$, in particular the existence and uniqueness of a stationary distribution. Section 4 investigates the tail of the stationary distribution. Theorem 8 is the main theorem in this section. We show that the stationary distribution has a Pareto-like tail with a well-specified tail index. For $\alpha = 0$ our result coincides with the corresponding result in Goldie (1991) whereas for $\alpha \neq 0$ the tail index is determined by the autoregressive coefficient α and the ARCH(1) parameter λ. The proof of this result will be an application of a modification of the Drasin-Shea Tauberian theorem.

2. Assumptions on the model. We consider throughout this paper an autoregressive model of order 1 with autoregressive conditionally heteroscedastic errors of order 1 (AR(1) model with ARCH(1) errors) which is defined by the stochastic difference equation

$$X_n = \alpha X_{n-1} + \sqrt{\beta + \lambda X_{n-1}^2}\varepsilon_n\,, \quad n \in \mathbb{N}\,, \tag{2.1}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are i.i.d. symmetric random variables, $\alpha \in \mathbb{R}$, β, $\lambda > 0$ and X_0 is independent of $(\varepsilon_n)_{n\in\mathbb{N}}$.

Let ε be a generic random variable with the same distribution function H as ε_n. In what follows, we assume without loss of generality $\alpha \geq 0$ (for a justification see Remark 4 below) and that the following *general conditions* for ε are in force:

$$\begin{aligned} \varepsilon \text{ has full support } \mathbb{R}\,,\\ \varepsilon \text{ is symmetric with continuous Lebesgue density } p\,,\\ \text{the second moment of } \varepsilon \text{ exists}\,. \end{aligned} \tag{2.2}$$

Note that the process is evidently a homogeneous Markov chain with state space $\mathbb{R}$ equipped with the Borel σ-algebra. The transition kernel density is given by

$$P(X_1 \in dy \,|\, X_0 = x) = \frac{1}{\sqrt{\beta + \lambda x^2}}\, p(\frac{y - \alpha x}{\sqrt{\beta + \lambda x^2}})dy\,, \quad x \in \mathbb{R}. \tag{2.3}$$

Under appropriate conditions on α and λ, Theorem 3 in Section 2 guarantees the existence and uniqueness of a stationary distribution π of $(X_n)_{n\in\mathbb{N}}$. By F we denote the distribution function of π and X is a random variable with distribution function F. From the stochastic difference equation (2.1) it is straightforward that X satisfies the fixpoint equation

$$X \stackrel{d}{=} \alpha X + \sqrt{\beta + \lambda X^2}\,\varepsilon\,, \tag{2.4}$$

where ε is independent of X. In order to determine the tail of the stationary distribution function F we need some additional technical assumptions on p and $\overline{H} = 1 - H$, the density and the distribution tail of ε:

$(D.1)$ $p(x) \geq p(x')$ for every $0 \leq x < x'$.
$(D.2)$ The lower and upper Matuszewska indices of $\overline{H}$ are equal, i.e.

$$\begin{aligned} -\infty \leq \gamma := \lim_{\nu\to\infty} \frac{\log \limsup_{x\to\infty} \overline{H}(\nu x)/\overline{H}(x)}{\log \nu} \\ = \lim_{\nu\to\infty} \frac{\log \liminf_{x\to\infty} \overline{H}(\nu x)/\overline{H}(x)}{\log \nu} \leq 0\,. \end{aligned}$$

$(D.3)$ If $\gamma = -\infty$ then for all $\delta > 0$ there exist constants $q \in (0,1)$ and $x_0 > 0$ such that for all $x > x_0$ and $t > x^q$

$$p(\frac{x \pm \alpha t}{\sqrt{\lambda t^2}}) \geq (1 - \delta)\, p(\frac{x \pm \alpha t}{\sqrt{\beta + \lambda t^2}})\,. \tag{2.5}$$

If $\gamma > -\infty$ then for all $\delta > 0$ there exist constants $x_0 > 0$ and $T > 0$ such that for all $x > x_0$ and $t > T$ the inequality (2.5) holds.

The definition of the lower and upper Matuszewska indices can be found e.g. in Bingham et al. (1987), p. 68; for the above representation we used Theorem 2.1.5 and Corollary 2.1.6. The case $\gamma = -\infty$ corresponds to a tail which is exponentially decreasing. For $\gamma \in (-\infty, 0]$ condition $(D.2)$ is equivalent to the existence of constants $0 \leq c \leq C < \infty$ such that for all $\Lambda > 1$, uniformly in $\nu \in [1, \Lambda]$,

$$c(1+o(1))\nu^{\gamma} \leq \frac{\overline{H}(\nu x)}{\overline{H}(x)} \leq C(1+o(1))\nu^{\gamma}\,, \quad x \to \infty\,. \tag{2.6}$$

In particular, a distribution with a regularly varying tail satisfies $(D.2)$; the value γ is then the tail index. Due to the equality of the Matuszewska indices and the monotonicity of p we obtain easily some asymptotic properties of $\overline{H}$ and of p, respectively.

Proposition 1. *Suppose the general conditions (2.2) and $(D.1) - (D.3)$ hold. Then the following holds:*
(a) $\lim_{x\to\infty} x^m \overline{H}(x) = 0$ *and* $E(|\varepsilon|^m) < \infty$ *for all* $m < -\gamma$.
(b) $\lim_{x\to\infty} x^m \overline{H}(x) = \infty$ *and* $E(|\varepsilon|^m) = \infty$ *for all* $m > -\gamma$.
(c) $\lim_{x\to\infty} x^{m+1} p(x) = 0$ *for all* $m < -\gamma$.
(d) If $\gamma > -\infty$*, there exist constants* $0 < c \leq C < \infty$ *such that*

$$c \leq \liminf_{x\to\infty} \frac{x\,p(x)}{\overline{H}(x)} \leq \limsup_{x\to\infty} \frac{x\,p(x)}{\overline{H}(x)} \leq C\,.$$

Moreover, there exist constants $0 \leq d \leq D < \infty$ *such that for all* $\Lambda > 1$*, uniformly in* $\nu \in [1, \Lambda]$,

$$d(1+o(1))\nu^{\gamma-1} \leq \frac{p(\nu x)}{p(x)} \leq D(1+o(1))\nu^{\gamma-1}\,, \quad x \to \infty\,. \tag{2.7}$$

Furthermore, in this case (2.7) is equivalent to (2.6) or $(D.2)$.

Proof. Statements (a) and (b) are immediate consequences of Theorem 2.2.2 of Bingham et al. (1987). (c) follows from (a) and the monotonicity of p. Applying (2.6) and using again the monotonicity of p yields (d).

The general conditions (2.2) and assumption $(D.1)$ are fairly general and can be checked easily, whereas $(D.2)$ and in particular $(D.3)$ seem to be quite technical and intractable. Nevertheless, numerous densities satisfy these assumptions.

Example 1. We give two different classes of densities, which satisfy the general conditions (2.2) and $(D.1) - (D.3)$.
(a) $p_{\rho,\theta}(x) \propto \exp(-\frac{|x|^{\rho}}{\theta})$, $x \in \mathbb{R}$, for parameters $\rho, \theta > 0$.
Note that this family of densities includes the Laplace (double exponential) density

($\rho = 1$) and the normal density with mean 0 ($\rho = 2$).
It is straightforward that the general conditions and $(D.1), (D.2)$ with $\gamma = -\infty$ hold. In order to show $(D.3)$, choose $q \in (\rho/(\rho+2), 1)$. Then for every $x > 0$ and $t > x^q$,

$$\frac{p_{\rho,\theta}\left(\dfrac{x \pm \alpha t}{\sqrt{\lambda t^2}}\right)}{p_{\rho,\theta}\left(\dfrac{x \pm \alpha t}{\sqrt{\beta + \lambda t^2}}\right)} = \frac{p_{\rho,\theta}\left(\dfrac{x}{\sqrt{\lambda} t} \pm \dfrac{\alpha}{\sqrt{\lambda}}\right)}{p_{\rho,\theta}\left(\left(\dfrac{x}{\sqrt{\lambda} t} \pm \dfrac{\alpha}{\sqrt{\lambda}}\right)\left(1 + \dfrac{\beta}{\lambda t^2}\right)^{-1/2}\right)}$$
$$= \exp\left(-\frac{1}{\theta}\left|\frac{x}{\sqrt{\lambda} t} \pm \frac{\alpha}{\sqrt{\lambda}}\right|^{\rho}\left(1 - \left|1 + \frac{\beta}{\lambda t^2}\right|^{-\rho/2}\right)\right)$$
$$\geq \exp\left(-\frac{c_\rho \rho \beta}{2\theta\lambda^{1+\rho/2}}|x|^{\rho - \rho q - 2q} - \frac{c_\rho \rho \beta \alpha^{\rho}}{2\theta\lambda^{1+\rho/2}}|x|^{-2q}\right),$$

where $c_\rho = \max(1, 2^\rho)$. The rhs is arbitrary close to 1 for x sufficiently large and therefore $(D.3)$ holds.
(b) $p_{a,\rho,\theta}(x) \propto \big(1 + \dfrac{x^2}{\theta}\big)^{-(\rho+1)/2}\big(1 + a \sin\big(2\pi \log(1 + \dfrac{x^2}{\theta})\big)\big)$, $x \in \mathbb{R}$,
for parameters $\rho > 2$, $\theta > 0$ and $a \in \big[0, \dfrac{\rho+1}{\rho+1+4\pi}\big)$.
This family of densities includes e.g. the Student's distribution density with parameter ρ (set $a = 0$ and $\theta = \rho$).
One can easily see that the general conditions hold. $(D.1)$ is satisfied because of the choice of a. Furthermore, for all $\Lambda > 1$, uniformly in $\nu \in [1, \Lambda]$,

$$\frac{1-a}{1+a}(1+o(1))\nu^{-(\rho+1)} \leq \frac{p(\nu x)}{p(x)} \leq \frac{1+a}{1-a}(1+o(1))\nu^{-(\rho+1)}, \quad x \to \infty.$$

In particular, $p_{a,\rho,\theta}$ is regularly varying if and only if $a = 0$. By Proposition 1(d), condition $(D.2)$ is satisfied with $\gamma = -\rho$. It remains to show $(D.3)$. Let $\delta > 0$ be arbitrary and choose T such that

$$\left(1 + \frac{\beta}{\lambda T^2}\right)^{-(\rho+1)/2}\left(1 - \frac{2\pi\, a\, \beta}{(1-a)\,\lambda T^2}\right) \geq 1 - \delta. \tag{2.8}$$

Next note that for every $x > 0$, setting $b(t) = 1 + \beta/(\lambda t^2)$, $t \geq 0$, we obtain

$$\left|\frac{1 + a \sin\left(2\pi \log\left(1 + y^2 b(T)/\theta\right)\right)}{1 + a \sin\left(2\pi \log\left(1 + y^2/\theta\right)\right)} - 1\right| \leq \frac{2\pi\, a\, \beta}{(1-a)\,\lambda T^2}. \tag{2.9}$$

Using (2.8) and (2.9), we have for every $t \geq T$, $x > 0$ and $y = \left(\dfrac{x}{\sqrt{\lambda}t} \pm \dfrac{\alpha}{\sqrt{\lambda}}\right) / \sqrt{b(T)}$,

$$
\begin{aligned}
\frac{p_{a,\rho,\theta}\left(\dfrac{x \pm \alpha t}{\sqrt{\lambda t^2}}\right)}{p_{a,\rho,\theta}\left(\dfrac{x \pm \alpha t}{\sqrt{\beta + \lambda t^2}}\right)} &= \frac{p_{a,\rho,\theta}\left(y\sqrt{b(t)}\right)}{p_{a,\rho,\theta}(y)} \\
&\geq \frac{p_{a,\rho,\theta}\left(y\sqrt{b(T)}\right)}{p_{a,\rho,\theta}(y)} \\
&\geq b(T)^{-(\rho+1)/2} \frac{1 + a\sin\left(2\pi \log\left(1 + y^2 b(T)/\theta\right)\right)}{1 + a\sin\left(2\pi\log\left(1 + y^2/\theta\right)\right)} \\
&\geq b(T)^{-(\rho+1)/2}\left(1 - \frac{2\pi\, a\,\beta}{(1-a)\,\lambda T^2}\right) \\
&\geq 1 - \delta\,.
\end{aligned}
$$

3. Existence and uniqueness of a stationary distribution. In this section we summarize in Theorem 3 some properties of the process $(X_n)_{n\in\mathbb{N}}$. In particular, the geometric ergodicity guarantees the existence and uniqueness of a stationary distribution. For an introduction to Markov chain terminology we refer to Tweedie (1976) or Meyn and Tweedie (1993).

The next proposition follows easily from well-known properties of moment generating functions (one can follow the proof of the case $\alpha = 0$; see e.g. Lemma 8.4.6 of Embrechts et al. (1997)).

PROPOSITION 2. *Let ε be a random variable with probability density p satisfying the general conditions (2.2). Define $h_{\alpha,\lambda} : [0,\infty) \to [0,\infty]$ for $\alpha \in \mathbb{R}$ and $\lambda > 0$ by*

$$h_{\alpha,\lambda}(u) := E(|\alpha + \sqrt{\lambda}\,\varepsilon|^u)\,, \quad u \geq 0\,. \tag{3.1}$$

(a) The function $h_{\alpha,\lambda}(\cdot)$ is strictly convex in $[0,T)$, where

$$T := \inf\{u \geq 0 \,|\, E(|\sqrt{\lambda}\,\varepsilon|^u) = \infty\}\,.$$

(b) If furthermore the parameters α and λ are chosen such that

$$h'_{\alpha,\lambda}(0) = E(\log|\alpha + \sqrt{\lambda}\,\varepsilon|) < 0\,, \tag{3.2}$$

then there exists a unique solution $\kappa = \kappa(\alpha, \lambda) > 0$ *to the equation* $h_{\alpha,\lambda}(u) = 1$. *Moreover, under* $h'_{\alpha,\lambda}(0) < 0$,

$$\kappa(\alpha, \lambda) \begin{cases} > 2, & \alpha^2 + \lambda E(\varepsilon^2) < 1, \\ = 2, & \alpha^2 + \lambda E(\varepsilon^2) = 1, \\ < 2, & \alpha^2 + \lambda E(\varepsilon^2) > 1. \end{cases} \tag{3.3}$$

REMARK 2. (a) By Jensen's inequality $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ implies $h'_{\alpha,\lambda}(0) < 0$.
(b) Proposition 2 holds in particular for a standard normal random variable ε. In this case $T = \infty$.
(c) In general, it is not possible to determine explicitly which parameters α and λ satisfy (3.2). If $\alpha = 0$ (i.e. in the ARCH(1)-case) and $\varepsilon \sim N(0, 1)$ (3.2) is fulfilled if and only if $\lambda \in (0, 2e^{\gamma})$, where γ is Euler's constant (see e.g. Embrechts et al. (1997), Section 8.4). For $\alpha \neq 0$, Tables 1 and 2 show numerical domains of α and λ for $\varepsilon \sim N(0, 1)$. See also Kiefersbeck (1999) for numerical results in some non-normal cases.
(d) Note that κ is a function of α and λ. Since ε is symmetric κ does not depend on the sign of α. For $\varepsilon \sim N(0, 1)$ we can show: for fixed λ, κ is decreasing in $|\alpha|$. See also Table 3.

PROOF Let $\varphi(\cdot \,|\, \mu, \sigma^2)$ denote the normal density with mean μ and variance σ^2. Then, by symmetry of φ,

$$\begin{aligned} \frac{\partial h_{\alpha,\lambda}(u)}{\partial \alpha} &= \frac{1}{\lambda} \int_{-\infty}^{\infty} |y|^u (y - \alpha) \varphi(y|\alpha, \lambda) dy \\ &= \frac{1}{\lambda} \left(\int_{-\infty}^{0} (-y)^u (y - \alpha) \varphi(y|\alpha, \lambda) dy + \int_0^{\infty} y^u (y - \alpha) \varphi(y|\alpha, \lambda) dy \right) \\ &= u \int_0^{\infty} y^{u-1} \left(\varphi(y|\alpha, \lambda) - \varphi(y| - \alpha, \lambda) \right) dy > 0, \quad u \geq 0, \end{aligned}$$

where the last line follows by integration by parts with respect to y. We may therefore conclude that, if $\alpha' > \alpha$ then $h_{\alpha,\lambda}(u) < h_{\alpha',\lambda}(u)$ for any λ, u. Assume $\kappa(\alpha) \leq \kappa(\alpha')$. Then we have by Proposition 2(b) and Hölder's inequality that

$$1 = h_{\alpha,\lambda}(\kappa(\alpha)) < h_{\alpha',\lambda}(\kappa(\alpha)) \leq h_{\alpha',\lambda}(\kappa(\alpha'))^{\kappa(\alpha)/\kappa(\alpha')} = 1,$$

which is a contradiction.

We are now ready to state the following theorem.

THEOREM 3. *Consider the process* $(X_n)_{n \in \mathbb{N}}$ *in (2.1) with* $(\varepsilon_n)_{n \in \mathbb{N}}$ *satisfying the general conditions (2.2) and with parameters* α *and* λ *satisfying (3.2). Then the following assertions hold:*

$\lvert\alpha\rvert$	0	0.1	0.2	0.3	0.4	0.5	0.6
λ	(0,3.56]	(0,3.55]	(0,3.52]	(0,3.47]	(0,3.39]	(0,3.30]	(0,3.18]
$\lvert\alpha\rvert$	0.8	0.9	1	1.1	1.2	1.25	1.27
λ	(0,2.87]	(0,2.66]	(0,2.42]	(0.17,2.11]	(0.38,1.69]	(0.58,1.38]	(0.75,1.19]

TABLE 1

Numerical domain of λ dependent on $\lvert\alpha\rvert$ such that $h'_{\alpha,\lambda}(0) < 0$ in the case $\varepsilon \sim N(0,1)$.

λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\lvert\alpha\rvert$	1.05	1.11	1.16	1.20	1.23	1.25	1.26	1.27	1.28
λ	1	1.1	1.2	1.5	2	2.5	3	3.5	3.56
$\lvert\alpha\rvert$	1.28	1.27	1.27	1.23	1.13	0.97	0.72	0.24	0.04

TABLE 2

Numerical supremum of $\lvert\alpha\rvert$ dependent on λ such that $h'_{\alpha,\lambda}(0) < 0$ in the case $\varepsilon \sim N(0,1)$.

$\lvert\alpha\rvert$ \ λ	0.2	0.4	0.6	0.8	1.0	1.2	1.5	2.0	2.5	3.0	3.5
0	12.85	6.09	3.82	2.67	1.99	1.54	1.07	0.61	0.33	0.15	0.01
0.2	11.00	5.49	3.52	2.51	1.89	1.46	1.03	0.59	0.32	0.13	0.01
0.4	8.12	4.28	2.87	2.10	1.61	1.26	0.90	0.51	0.27	0.10	-
0.6	5.41	3.03	2.12	1.60	1.25	0.99	0.71	0.39	0.19	0.05	-
0.8	3.00	1.85	1.37	1.07	0.85	0.68	0.48	0.25	0.09	-	-
1.0	0.96	0.83	0.70	0.57	0.47	0.37	0.25	0.09	-	-	-
1.2	-	0.01	0.01	0.01	0.01	0.01	0.01	-	-	-	-

TABLE 3

Numerical solution of $h_{\alpha,\lambda}(\kappa) = 1$ for $\kappa = \kappa(\alpha,\lambda)$ dependent on α and λ in the case $\varepsilon \sim N(0,1)$. For $\alpha = 0$ a similar table can be found in de Haan et al. (1989).

(a) *Let ν be the normalized Lebesgue-measure $\nu(\cdot) := \lambda(\cdot \cap [-M, M])/\lambda([-M, M])$. Then $(X_n)_{n\in\mathbb{N}}$ is an aperiodic positive ν-recurrent Harris chain with regeneration set $[-M, M]$ for M large enough.*

(b) *$(X_n)_{n\in\mathbb{N}}$ is geometric ergodic. In particular, $(X_n)_{n\in\mathbb{N}}$ has a unique stationary distribution and satisfies the strong mixing condition with geometric rate of convergence. The stationary distribution is continuous and symmetric.*

(c) *If $\alpha^2 + \lambda E(\varepsilon^2) < 1$, then the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ has finite second moment.*

REMARK 4. (a) Statements (a) and (b) are basically a collection of results of Diebolt and Guégan (1990) and Maercker (1997). They assume $\alpha^2 + \lambda E(\varepsilon^2) < 1$ and hence only cover the finite variance case. The model fits also into the more general framework of "iterated random Lipschitz functions"; see Alsmeyer (2000).
(b) When we study the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ we may w.l.o.g. assume that $\alpha \geq 0$. For a justification, consider the process $(\widetilde{X}_n)_{n\in\mathbb{N}} = ((-1)^n X_n)_{n\in\mathbb{N}}$ which satisfies to the stochastic difference equation

$$\widetilde{X}_n = -\alpha \widetilde{X}_{n-1} + \sqrt{\beta + \lambda \widetilde{X}_{n-1}^2}\, \varepsilon_n\,, \quad n \in \mathbb{N},$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are the same random variables as in (2.1) and $\widetilde{X}_0 = X_0$. If $\alpha < 0$, because of the symmetry of the stationary distribution, we may hence study the new process $(\widetilde{X}_n)_{n\in\mathbb{N}}$.
(c) By statement (c), the assumption $\alpha^2 + \lambda E(\varepsilon^2) < 1$ is sufficient for the existence of the second moment. We will see in Remark 9(c) that it is also necessary.

PROOF OF THEOREM 3.3. Because of the strict positivity and continuity of the transition density the process $(X_n)_{n\in\mathbb{N}}$ is a ν-irreducible Feller chain. By Feigin and Tweedie (1984), p. 3, this implies that every compact set of the state space with positive Lebesgue measure is small and thus $[-M, M]$ is small for arbitrary $M > 0$. Finally, by Proposition 5.3 of Tweedie (1976), $[-M, M]$ is a status set for $(X_n)_{n\in\mathbb{N}}$.

(a) Because of Proposition 2, for $\alpha \in \mathbb{R}$ and $\lambda > 0$ such that $h'_{\alpha,\lambda}(0) < 0$ there exists a $\kappa > 0$ such that $h_{\alpha,\lambda}(u) < 1$ for every $u \in (0, \kappa)$ and $h_{\alpha,\lambda}(0) = h_{\alpha,\lambda}(\kappa) = 1$. Now choose $\eta \in (0, \min(\kappa, 2))$ and $\delta \in (0, 1 - h_{\alpha,\lambda}(\eta))$ arbitrary. For any such η and δ there exists a constant $C = C(\eta, \delta) \in (0, 1)$ such that

$$h_{\alpha,\lambda}(\eta) + \delta \leq 1 - 2C\,. \tag{3.4}$$

Define $g(x) := 1 + |x|^\eta \geq 1$ for every $x \in \mathbb{R}$. For M large enough and $|x| > M$ we have by continuity of $h_{\alpha,\lambda}$ in α

$$\left| h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) - h_{\alpha,\lambda}(\eta) \right| < \delta \tag{3.5}$$

and

$$C\, g(x) \geq 1 + (h_{\alpha,\lambda}(\eta) \pm \delta)(-1 + O(|x|^{\eta-2}))\,, \tag{3.6}$$

since $\eta < 2$, $h_{\alpha,\lambda}(\eta) - \delta$ is independent of x and g increases to ∞. From (2.3) we obtain for $x \to \infty$

$$\begin{aligned}
&\int_{(-\infty,\infty)} g(y) P(X_1 \in dy \,|\, X_0 = x) \\
&= 1 + (\beta + \lambda x^2)^{\eta/2} E(|\frac{\alpha x}{\sqrt{\lambda x^2 + \beta}} + \varepsilon|^\eta) \\
&= 1 + (\frac{\beta}{\lambda} + x^2)^{\eta/2} h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + (1 + O(x^{-2}))|x|^\eta\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + O(|x|^{\eta-2})\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) + |x|^\eta\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) \\
&= 1 + \left(-1 + O(|x|^{\eta-2})\right)\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta) + g(x)\, h_{\alpha x/\sqrt{x^2+\beta/\lambda},\lambda}(\eta)\,,
\end{aligned}$$

where the third line follows from Taylor expansion. Together with (3.4)-(3.6), we obtain for every $x \in \mathbb{R}$ with $|x| > M$,

$$\begin{aligned}
\int_{(-\infty,\infty)} g(y) P(X_1 \in dy \,|\, X_0 = x) &\leq C\, g(x) + (1 - 2\,C) g(x) \\
&= (1 - C) g(x)\,.
\end{aligned} \tag{3.7}$$

Define

$$\tau_{[-M,M]} := \inf\{n \geq 1 \,|\, X_n \in [-M, M]\}$$

and let $x \in \mathbb{R}$ be arbitrary. Then we have

$$\begin{aligned}
E(\tau_{[-M,M]} \,|\, X_0 = x) &= E(1_{\{X_1 \in [-M,M]\}} E(\tau_{[-M,M]}|X_1)|X_0 = x) \\
&\qquad + E(1_{\{X_1 \in [-M,M]^c\}} E(\tau_{[-M,M]}|X_1)|X_0 = x) \\
&\leq 1 + E(1_{\{X_1 \in [-M,M]^c\}} E(\tau_{[-M,M]}|X_1)|X_0 = x) \\
&\leq 1 + \int_{[-M,M]^c} E(\tau_{[-M,M]}|X_1 = y) P(X_1 \in dy|X_0 = x)\,.
\end{aligned}$$

By (3.7), Theorem 3 of Tweedie (1983a) holds and we obtain for all $x \in \mathbb{R}$,

$$\begin{aligned} E(\tau_{[-M,M]}|X_0 = x) &\leq 1 + \int_{[-M,M]^c} \frac{g(y)}{C} P(X_1 \in dy|X_0 = x) \\ &\leq 1 + \frac{1}{C} + E\left(\left|\alpha x + \sqrt{\lambda x^2 + \beta}\varepsilon\right|^{\eta}\right) < \infty \end{aligned} \tag{3.8}$$

and thus $[-M, M]$ is Harris recurrent. Since the transition density of $(X_n)_{n\in\mathbb{N}}$ is strictly positive on $[-M, M]$ we know from Asmussen (1987), p. 151, that there exists some constant $\widetilde{C} \in (0, 1)$ such that

$$P(X_1 \in B \,|\, X_0 = x) \geq \widetilde{C}\, \nu(B) \tag{3.9}$$

for every $x \in [-M, M]$ and any Borel-measurable set B, i.e. $(X_n)_{n\in\mathbb{N}}$ is a Harris chain with regeneration set $[-M, M]$. Finally, by Theorem 9.1 of Tweedie (1976), (3.7) and the fact that $[-M, M]$ is a status set, $(X_n)_{n\in\mathbb{N}}$ is positive Harris ν-recurrent.

(b) Note that

$$\begin{aligned} &\sup_{x\in[-M,M]} \int_{\mathbb{R}} g(y)P(X_1 \in dy|X_0 = x) \\ &= 1 + \sup_{x\in[-M,M]} E\left(\left|\alpha x + \sqrt{\lambda x^2 + \beta}\varepsilon\right|^{\eta}\right) \;<\; \infty\,. \end{aligned} \tag{3.10}$$

Thus the geometric ergodicity follows from Theorem 4 of Tweedie (1983a) and the same arguments as in the proof of statement (a) of this theorem. The process is therefore strongly mixing with a geometric rate. The symmetry of the stationary distribution follows from the ergodicity and the fact that the processes $(X_n)_{n\in\mathbb{N}}$ and $(-X_n)_{n\in\mathbb{N}}$ have the same transition probabilities, hence the same unique stationary distribution. Finally, because of the continuity of the transition probabilities, the stationary distribution function is continuous as well.

(c) Define now the small set

$$A := \left\{x \in \mathbb{R} \,|\, x^2 \leq \max\{1, \frac{\beta E(\varepsilon^2)}{(1-2\delta) - (\alpha^2 + \lambda\, E(\varepsilon^2))}\}\right\}$$

with $\delta > 0$ such that $(1 - 2\,\delta) - (\alpha^2 + \lambda\, E(\varepsilon^2)) > 0$. Choose $g(x) = 1 + x^2$. Note that for every $x \in A^c$,

$$\begin{aligned} \int_{\mathbb{R}} g(y)P(X_1 \in dy \,|\, X_0 = x) &\leq 1 + x^2 \left(\alpha^2 + \lambda\, E(\varepsilon^2) + \frac{\beta\, E(\varepsilon^2)}{x^2}\right) \\ &\leq 1 + x^2\,(1 - 2\delta) \\ &= 1 - x^2\delta + x^2(1 - \delta) \\ &\leq 1 - \delta + x^2(1 - \delta) \;=\; g(x)\,(1 - \delta)\,. \end{aligned}$$

This together with (3.10) for $\eta = 2$ and A instead of $[-M, M]$, Theorem 3 of Tweedie (1983b) holds and the second moment of the stationary distribution is finite.

Even if the building blocks $(\varepsilon_n)_{n\in\mathbb{N}}$ have moments of all orders, i.e. $\gamma = -\infty$, not all moments of the stationary distribution are finite.

PROPOSITION 3. *Suppose $(X_n)_{n\in\mathbb{N}}$ is given by equation (2.1) with $(\varepsilon_n)_{n\in\mathbb{N}}$ satisfying the general conditions (2.2) and with parameters α and λ satisfying (3.2). Let X be the stationary limit variable of $(X_n)_{n\in\mathbb{N}}$. Choose $N > 0$ such that*

$$E(|\sqrt{\lambda}\varepsilon|^N) > 2\,. \tag{3.11}$$

Then

$$E(|X|^N) = \infty\,.$$

PROOF. Assume that the N-th moment is finite. As a consequence of (2.4) (recall that w.l.o.g. $\alpha \geq 0$)

$$\begin{aligned}
E(|X|^N) &= E(|\alpha X + \sqrt{\beta + \lambda X^2}\,\varepsilon|^N) \\
&= E(1_{\{X<0\}}|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}\,(-\varepsilon)|^N) \\
&\qquad + E(1_{\{X>0\}}|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}\,\varepsilon|^N) \\
&= E(|X|^N|\alpha + \sqrt{\frac{\beta}{X^2} + \lambda}\,\varepsilon|^N) \\
&\geq E(|X|^N)E(1_{\{\varepsilon>0\}}|\sqrt{\lambda}\,\varepsilon|^N) \\
&> E(|X|^N)\,,
\end{aligned}$$

where we used in the third and fourth line that X and ε are independent. The last line is a consequence of (3.11) and the symmetry of ε.

REMARK 5. (a) Note that $N > 2$ if $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$ since the second moment exists by Theorem 3(c).
(b) Condition (3.11) can be replaced by $E(1_{\{\varepsilon>0\}}|\alpha + \sqrt{\lambda}\varepsilon|^N) > 1$ for $\alpha \geq 0$ and $E(1_{\{\varepsilon<0\}}|\alpha + \sqrt{\lambda}\varepsilon|^N) > 1$ for $\alpha < 0$, respectively. These alternative conditions may enable us to find a smaller N.

Because of Proposition 3 we know that the distribution of X is heavy-tailed in the sense that not all moments exist. The following section considers the precise asymptotic behavior of its tail.

4. The tail of the stationary distribution. Estimating the (heavy) tail of a stationary distribution of a Markov process is in general a non-trivial problem and few explicit results are known in the literature. There are basically two articles which refer to this topic and which are somewhat related to our problem. Kesten (1973) investigates the tail of the limit distribution of the solution of a linear difference equation, and Goldie (1991) proves and extends Kesten's results in the one-dimensional case by applying a renewal type argument.

Unfortunately, both approaches are not directly applicable for the AR(1) process with ARCH(1) errors, since $(X_n)_{n\in\mathbb{N}}$ does not fit in their framework. Consider instead the process $(Y_n)_{n\in\mathbb{N}}$ given by the stochastic difference equation

$$Y_n = \left|\alpha Y_{n-1} + \sqrt{\beta + \lambda Y_{n-1}^2}\varepsilon_n\right|, \quad n \geq 1, \tag{4.1}$$

where $(\varepsilon_n)_{n\in\mathbb{N}}$ are the same i.i.d. random variables as in Theorem 3, the constants are the same as for the process $(X_n)_{n\in\mathbb{N}}$ and Y_0 equals $|X_0|$ a.s. It can be seen easily that $(Y_n)_{n\in\mathbb{N}} \stackrel{d}{=} (|X_n|)_{n\in\mathbb{N}}$ if $X_0 \sim \pi$. Hence $(Y_n)_{n\in\mathbb{N}}$ and $(|X_n|)_{n\in\mathbb{N}}$ have the same stationary distribution and $P(X > x) = 1/2\, P(Y > x)$, $x \in \mathbb{R}$. Setting $M := |\alpha + \sqrt{\lambda}\,\varepsilon|$ and κ as in Lemma 3.1, the conditions of Corollary 2.4 of Goldie (1991) on M are satisfied. Thus, under the additional assumption that

$$E\left(\left|(|\alpha Y + \sqrt{\beta + \lambda Y^2}\,\varepsilon|)^\kappa - (|\alpha + \sqrt{\lambda}\,\varepsilon|Y)^\kappa\right|\right) < \infty, \tag{4.2}$$

the tail of the stationary distribution of $(Y_n)_{n\in\mathbb{N}}$ is Pareto, i.e.

$$P(Y > x) \sim c\, x^{-\kappa}, \quad x \to \infty, \tag{4.3}$$

where c is a well-specified non-negative constant. Note that a sufficient condition for (4.2) is $E(Y^{\kappa-1}) < \infty$.

The above procedure using Goldie's result seems to be at first sight very simple. However, in spite of the strength and elegance of the results in Goldie (1991), additional conditions such as (4.2) are hard to check. Since the knowledge of the existence of moments is in some way equivalent to the knowledge of the (unknown) tail distribution (or at least the tail index of the stationary distribution), we consider directly the tail of the stationary distribution of the process $(X_n)_{n\in\mathbb{N}}$. The tail is derived by applying a Tauberian theorem which, as far as we know, is a new approach. This method may also be applied to other processes given by random recurrence equations which do not fit in the framework of Kesten (1973) or Goldie (1991), or which simply do not fulfill all the conditions in the two referred articles. Note that our approach gives no information on the slowly varying function present in the tail of the heavy-tailed stationary distribution. In the case of the AR(1) process with ARCH(1) errors we determine the tail index of the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ with our new approach and draw then the conclusion that the slowly varying function is a well-specified constant.

In order to present our method we need the notion of O-regular variation; see Bingham et al. (1987), Chapter 2, for relevant definitions and results.

PROPOSITION 4. *Let $\overline{F}(x) := P(X > x)$, $x \geq 0$, be the tail of the stationary solution of the process $(X_n)_{n\in\mathbb{N}}$ given by (2.1). Then $\overline{F}$ is O-regularly varying. In particular, if $\overline{H} := 1 - H$ denotes the tail of the distribution function of ε, for every $\Lambda \geq 1$,*

$$\frac{\overline{F}(\Lambda x)}{\overline{F}(x)} \geq \overline{H}\left(\max(0, \frac{\Lambda - \alpha}{\sqrt{\lambda}})\right) \quad \textit{for all } x \geq 0\,. \tag{4.4}$$

PROOF. Let $\Lambda \geq 1$ be arbitrary. Since X is symmetric and $(|X_n|)$ and (Y_n) have the same law when $X_0 \sim \pi$, we have for every $x \geq 0$,

$$\begin{aligned}
\frac{P(X > \Lambda x)}{P(X > x)} &= \frac{P(Y > \Lambda x)}{P(Y > x)} \\
&\geq \frac{P(\alpha Y + \sqrt{\beta + \lambda Y^2}\varepsilon > \Lambda x, \varepsilon > 0)}{P(Y > x)} \\
&\geq \frac{P(Y > \Lambda x/(\alpha + \sqrt{\lambda}\varepsilon), \varepsilon > 0)}{P(Y > x)} \\
&\geq \int_{\max\left(0,(\Lambda-\alpha)/\sqrt{\lambda}\right)}^{\infty} \frac{P(Y > \Lambda x/(\alpha + \sqrt{\lambda}t))}{P(Y > x)} p(t)dt
\end{aligned}$$

By monotonicity, the integrand is bounded from below by 1. Therefore, (4.4) holds. Note that the rhs of (4.4) does not depend on x. Letting $x \to \infty$ and applying Corollary 2.0.6 of Bingham et al. (1987) shows that $\overline{F}$ is O-regularly varying.

REMARK 6. Since $\overline{F}$ is O-regularly varying, its lower Matuszewska index $\gamma > -\infty$. Therefore, by Theorem 2.2.2 of Bingham et al. (1987), for every $\tau \in (-\gamma, \infty)$ there exist $C > 0$ and $x_0 > 0$ such that $x^\tau\, \overline{F}(x) \geq C$ for all $x \geq x_0$.

It turns out that the following modification of the Drasin-Shea Theorem (Bingham et al. (1987), Theorem 5.2.3, p. 273) is the key to our result.

THEOREM 7. *Let $k : [0,\infty) \to [0,\infty)$ be an integrable function and let (a,b) be the maximal open interval (where $a < 0$) such that*

$$\check{k}(z) = \int_{(0,\infty)} t^{-z}\frac{k(t)}{t}dt < \infty\,, \quad \textit{for } z \in (a,b)\,.$$

If $a > -\infty$, *assume* $\lim_{\delta\downarrow 0} \check{k}(a+\delta) = \infty$, *if* $b < \infty$, *assume* $\lim_{\delta\downarrow 0} \check{k}(b-\delta) = \infty$. *Let* $h : [0,\infty) \to [0,\infty)$ *be locally bounded. Assume* h *has bounded increase. If*

$$\lim_{x\to\infty} \frac{\int_{(0,\infty)} k(x/t)h(t)dt/t}{h(x)} = c > 0\,, \tag{4.5}$$

then

$$c = \check{k}(\rho) \text{ for some } \rho \in (a,b) \text{ and } h(x) \sim x^{\rho} l(x)\,, \quad x \to \infty\,,$$

where l *is some slowly varying function.*

We will identify h with the tail $\overline{F}$ of the distribution of X. The following is our main theorem.

THEOREM 8. *Suppose* $(X_n)_{n\in\mathbb{N}}$ *is given by equation (2.1) with* $(\varepsilon_n)_{n\in\mathbb{N}}$ *satisfying the general conditions (2.2) and* $(D.1)-(D.3)$ *and with parameters* α *and* λ *satisfying (3.2). Let* $\overline{F}(x) = P(X > x)$, $x \geq 0$, *be the right tail of the stationary distribution function. Then*

$$\overline{F}(x) \sim c\,x^{-\kappa}\,, \quad x \to \infty\,,$$

where

$$c = \frac{1}{2\kappa} \frac{E\left(\left|\alpha|X| + \sqrt{\beta + \lambda X^2}\varepsilon\right|^{\kappa} - \left|(\alpha + \sqrt{\lambda}\varepsilon)|X|\right|^{\kappa}\right)}{E\left(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa} \log|\alpha + \sqrt{\lambda}\varepsilon|\right)} \tag{4.6}$$

and κ *is given as the unique positive solution to*

$$E(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa}) = 1\,. \tag{4.7}$$

REMARK 9. (a) For the ARCH(1) process (i.e. the case $\alpha = 0$) this result is well-known (see Goldie (1991) or Embrechts et al. (1997), Section 8.4).
(b) Let $E(|\alpha + \sqrt{\lambda}\varepsilon|^{\kappa}) = h_{\alpha,\lambda}(\kappa)$ be as in Lemma 2. Recall that for $\varepsilon \sim N(0,1)$ and fixed λ, the exponent κ is decreasing in $|\alpha|$. This means that the distribution of X gets heavier tails. In particular, our new model has for $\alpha \neq 0$ heavier tails than the ARCH(1) process (see also Table 3).
(c) Theorem 8 together with Lemma 2 implies that the second moment of the stationary distribution exists if and only if $\alpha^2 + \lambda\, E(\varepsilon^2) < 1$.

The proof of Theorem 8 will be an application of Theorem 7. Proposition 5 presents an implicit formula for the right tail $\overline{F} = 1 - F$ of the distribution of X. We shall need the formula to show that assumption (4.5) is fulfilled. In the following all assumptions of Theorem 8 hold. Recall that w.l.o.g. $\alpha \geq 0$.

PROPOSITION 5.

$$1 = \frac{\overline{H}(x/\sqrt{\beta})}{\overline{F}(x)} + \int_0^\infty f(x,t)dt + \int_0^\infty h(x,t)dt\,, \quad x>0\,, \tag{4.8}$$

where $\overline{H}(x) = P(\varepsilon > x)$, $x > 0$, *and for every* $x > 0$, $t > 0$,

$$f(x,t) := \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}}\frac{\overline{F}(t)}{\overline{F}(x)}\frac{1}{t} \geq 0\,,$$

$$h(x,t) := \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) - p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{\alpha\beta t}{(\beta+\lambda t^2)^{3/2}}\frac{\overline{F}(t)}{\overline{F}(x)}\frac{1}{t} \geq 0\,.$$

PROOF. By (2.4) and the symmetry of X, we have

$$\begin{aligned}
\overline{F}(x) &= \int_{-\infty}^\infty P(\alpha X + \sqrt{\beta+\lambda X^2}\,\varepsilon > x \mid X = t)dF(t)\\
&= -\int_0^\infty P(-\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x)dF(-t) + \int_0^\infty P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x)dF(t)\\
&= -\int_0^\infty P(-\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x)d\overline{F}(t) + \int_0^\infty P(\alpha t + \sqrt{\beta+\lambda t^2}\varepsilon > x)dF(t)\\
&= -\int_0^\infty \Big(\overline{H}(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}}) + \overline{H}(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)d\overline{F}(t)\,.
\end{aligned}$$

Integration by parts (see e.g. Theorem 18.4 in Billingsley (1995)) and again symmetry yields

$$\begin{aligned}
\overline{F}(x) &= \overline{H}(\frac{x}{\sqrt{\beta}}) - \int_0^\infty \Big(p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\frac{\alpha(\beta+\lambda t^2)-(x+\alpha t)\lambda t}{(\beta+\lambda t^2)^{3/2}}\\
&\qquad\qquad + p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}})\frac{-\alpha(\beta+\lambda t^2)-(x-\alpha t)\lambda t}{(\beta+\lambda t^2)^{3/2}}\Big)\overline{F}(t)dt\\
&= \overline{H}(\frac{x}{\sqrt{\beta}}) + \int_0^\infty \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) + p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{x\lambda t^2}{(\beta+\lambda t^2)^{3/2}}\overline{F}(t)\frac{dt}{t}\\
&\qquad + \int_0^\infty \Big(p(\frac{x-\alpha t}{\sqrt{\beta+\lambda t^2}}) - p(\frac{x+\alpha t}{\sqrt{\beta+\lambda t^2}})\Big)\frac{\alpha\beta t}{(\beta+\lambda t^2)^{3/2}}\overline{F}(t)\frac{dt}{t}\,.
\end{aligned}$$

Finally, $h(x,t) \geq 0$ for every $x > 0$, $t > 0$ because of $(D.1)$ and the symmetry of p. This finishes the proof.

We investigate now (4.8). Using Proposition 3, Proposition 4 and Remark 6 we derive some technical results in the next three lemmata. These results will be crucial in applying Theorem 7.

LEMMA 1. *For every $a \geq 0$ and $b > 0$,*

$$\lim_{x\to\infty} \frac{\overline{H}\left((x-a)/b\right)}{\overline{F}(x)} = 0\,.$$

PROOF. Assume first that $\gamma = -\infty$. Because of Proposition 1(a) and Remark 6 the statement follows immediately.
Now consider the case where $\gamma > -\infty$. Let $N := \inf\{n > 0 : E(|\varepsilon|^n) > 2\}$ and choose $m \in (N, -\gamma)$. This is possible because of Proposition 1(a). Similarly as in Proposition 4 we derive that

$$\begin{aligned}
\frac{\overline{F}(x)}{\overline{H}\left((x-a)/b\right)} &= \frac{1}{2}\frac{P(|\alpha Y + \sqrt{\beta+\lambda Y^2}\,\varepsilon| > x)}{\overline{H}\left((x-a)/b\right)} \\
&\geq \frac{1}{2}\int_0^\infty \frac{\overline{H}((x-\alpha\, t)/\sqrt{\beta+\lambda t^2})}{\overline{H}\left((x-a)/b\right)} dF_Y(t) \\
&\geq \frac{1}{2}\int_{\max\{2a/\alpha, b/\sqrt{\lambda}\}}^\infty \frac{\overline{H}\left((x-a)/\sqrt{\lambda}t\right)}{\overline{H}\left((x-a)/b\right)} dF_Y(t)\,.
\end{aligned}$$

Applying the Lemma of Fatou and Proposition 2.2.1(a) of Bingham et al. (1987) yields

$$\begin{aligned}
\liminf_{x\to\infty} \frac{\overline{F}(x)}{\overline{H}\left((x-a)/b\right)} &\geq \frac{1}{2}\int_{\max\{2a/\alpha, b/\sqrt{\lambda}\}}^\infty \liminf_{x\to\infty} \frac{\overline{H}\left((x-a)/\sqrt{\lambda}t\right)}{\overline{H}\left((x-a)/b\right)} dF_Y(t) \\
&\geq const \int_{\max\{2a/\alpha, b/\sqrt{\lambda}\}}^\infty t^m dF_Y(t) \\
&= const\; E(|X|^m 1_{\{|X| > \max\{2a/\alpha, b/\sqrt{\lambda}\}\}})\,.
\end{aligned}$$

Since $m > N$ and $E(|X|^N) = \infty$, the statement follows by Proposition 3.

LEMMA 2. *For every $T > 0$,*

$$\lim_{x\to\infty}\int_0^\infty f(x,t)dt = \lim_{x\to\infty}\int_T^\infty f(x,t)dt = 1\,.$$

Moreover, if the lower Matuszewska index $\gamma = -\infty$, then for every $q \in (0,1)$,

$$\lim_{x\to\infty}\int_{x^q}^\infty f(x,t)dt = 1\,.$$

PROOF. Note that

$$0 \leq h(x,t) \leq \frac{\alpha\beta}{x\lambda} f(x,t), \quad \text{for every } t \geq 1 \text{ and } x > 0\,. \tag{4.9}$$

Thus, for every $x > 0$,

$$0 \leq \int_1^\infty h(x,t)dt \leq \frac{\alpha\beta}{x\lambda}\int_1^\infty f(x,t)dt\,. \tag{4.10}$$

Next choose $T \geq 0$ arbitrary. By $(D.1)$, for every $t \in [0,T]$ and x large enough

$$0 \leq \max\{f(x,t),\, h(x,t)\} \leq \frac{\max\{2\lambda T, \alpha\beta\}}{\beta^{3/2}}\, p(\frac{x-\alpha T}{\sqrt{\beta+\lambda T^2}}) \frac{x}{\overline{F}(x)}\,,$$

and therefore distinguishing again between $\gamma = -\infty$ and $\gamma > -\infty$ (in the first case use Remark 6 otherwise Lemma 1 and Proposition 1(d)) we get

$$\lim_{x\to\infty} f(x,t) = 0 \quad \text{and} \quad \lim_{x\to\infty} h(x,t) = 0\,, \qquad \text{for every } t \in [0,T]\,.$$

Thus, by the dominated convergence theorem,

$$\lim_{x\to\infty} \int_0^T f(x,t)dt = 0 \quad \text{and} \quad \lim_{x\to\infty} \int_0^T h(x,t)dt = 0\,. \tag{4.11}$$

Combining the result in Lemma 1 with (4.10) and (4.11) the first statement follows.

Finally, by $(D.1)$, Remark 6 and Proposition 1(c), supposing that $\gamma = -\infty$ and x large enough,

$$\begin{aligned}\int_T^{x^q} f(x,t)dt &\leq 2\, p(\frac{x-\alpha x^q}{\sqrt{\beta+\lambda x^{2q}}}) \frac{x^{q+1}\lambda}{(\beta+\lambda T^2)^{3/2}} \frac{1}{\overline{F}(x)} x^q \\ &\leq const(T)\, p(\frac{x^{1-q}-\alpha}{\sqrt{\beta/x^{2q}+\lambda}})(x^{1-q})^{(2q+1+\tau)/(1-q)} \;\to\; 0\,, \quad x\to\infty\,.\end{aligned}$$

This completes the proof.

LEMMA 3. *Define for $x > 0$, $t > 0$*

$$g(x,t) := \left(p(\frac{x-\alpha t}{\sqrt{\lambda}t}) + p(\frac{x+\alpha t}{\sqrt{\lambda}t})\right) \frac{x\lambda t^2}{(\lambda t^2)^{3/2}} \frac{\overline{F}(t)}{\overline{F}(x)} \frac{1}{t}\,,$$

then $\lim_{x\to\infty} \int_0^\infty g(x,t)dt = 1$.

PROOF. Note first that integration by parts and Lemma 1 yield for every $T > 0$

$$\begin{aligned} 0 &\leq \limsup_{x\to\infty} \int_0^T g(x,t)dt \\ &\leq \limsup_{x\to\infty} \overline{F}(T) \left(\frac{\overline{H}\Big((x-\alpha T)/\sqrt{\lambda}T\Big)}{\overline{F}(x)} + \frac{\overline{H}\Big((x+\alpha T)/\sqrt{\lambda}T\Big)}{\overline{F}(x)} \right) \\ &\quad + \limsup_{x\to\infty} \int_0^T \frac{\Big(\overline{H}\Big((x-\alpha t)/\sqrt{\lambda}t\Big) + \overline{H}\Big((x+\alpha t)/\sqrt{\lambda}t\Big)\Big)}{\overline{F}(x)} dF(t) \qquad (4.12) \\ &\leq 4 \limsup_{x\to\infty} \frac{\overline{H}\Big((x-\alpha T)/\sqrt{\lambda}T\Big)}{\overline{F}(x)} = 0\,. \end{aligned}$$

Furthermore, by the general conditions (2.2) and assumption $(D.1)$, for every $x > 0$, $t \geq 0$

$$p(\frac{x \pm \alpha t}{\sqrt{\lambda}\, t}) \leq p(\frac{x \pm \alpha t}{\sqrt{\beta + \lambda t^2}})$$

and hence with Lemma 2 and (4.12) we get

$$\begin{aligned}
\limsup_{x\to\infty} \int_0^\infty g(x,t)dt &= \limsup_{x\to\infty} \int_T^\infty g(x,t)dt \\
&\leq \left(\frac{\beta}{\lambda T^2} + 1\right)^{3/2} \limsup_{x\to\infty} \int_T^\infty f(x,t)dt \\
&= \left(\frac{\beta}{\lambda T^2} + 1\right)^{3/2} .
\end{aligned}$$

Letting $T \to \infty$ we conclude that

$$\limsup_{x\to\infty} \int_0^\infty g(x,t)dt \leq 1 .$$

It remains to show that the converse inequality holds for the limes inferior. We restrict ourselves to $\gamma = -\infty$ (for $\gamma > -\infty$ replace in what follows the lower integration limit x^q with T). Choose $\delta > 0$ arbitrary and let q be the constant in $(D.3)$. By assumption $(D.3)$ and Lemma 2

$$\begin{aligned}
\liminf_{x\to\infty} \int_0^\infty g(x,t)dt &\geq \liminf_{x\to\infty} \int_{x^q}^\infty g(x,t)dt \\
&\geq (1-\delta) \liminf_{x\to\infty} \int_{x^q}^\infty f(x,t)\frac{(\beta+\lambda t^2)^{3/2}}{(\lambda t^2)^{3/2}}dt \\
&\geq (1-\delta) \liminf_{x\to\infty} \int_{x^q}^\infty f(x,t)dt \\
&= 1-\delta .
\end{aligned}$$

Since $\delta > 0$ was arbitrary the statement follows.

We are now ready to prove Theorem 8.

Proof of Theorem 8. The proof is an application of Theorem 7. Choose

$$k(x) = \frac{x}{\sqrt{\lambda}}\Big(p(\frac{x-\alpha}{\sqrt{\lambda}}) + p(\frac{x+\alpha}{\sqrt{\lambda}})\Big), \quad x > 0, \tag{4.13}$$

and

$$h(x) = \overline{F}(x), \quad x > 0 . \tag{4.14}$$

One can readily see that k is non-negative, h is non-negative, locally bounded and of bounded increase since it is non-increasing. Note that for every $z \in (-\infty, \infty)$

$$\begin{aligned}\check{k}(z) &= \int_0^\infty t^{-z} \frac{k(t)}{t} dt \\ &= \int_0^\infty t^{-z} \frac{1}{\sqrt{\lambda}} p(\frac{t-\alpha}{\sqrt{\lambda}}) dt + \int_{-\infty}^0 (-t)^{-z} \frac{1}{\sqrt{\lambda}} p(\frac{t-\alpha}{\sqrt{\lambda}}) dt \\ &= E(|\alpha + \sqrt{\lambda}\varepsilon|^{-z}) .\end{aligned}$$

Let (a,b) be the maximal open interval such that

$$\check{k}(z) < \infty \quad \text{for } z \in (a,b) .$$

Note that $a = -T = -\inf\{u \geq 0 \,|\, h_{\alpha,\lambda}(u) = \infty\} < 0$ and $b = 1$ because of Proposition 2 and the fact that for $z \geq 0$

$$\int_1^\infty t^{-z} \frac{k(t)}{t} dt \leq \int_1^\infty \frac{1}{\sqrt{\lambda}} \Big(p(\frac{t-\alpha}{\sqrt{\lambda}}) + p(\frac{t+\alpha}{\sqrt{\lambda}}) \Big) dt < \infty$$

and

$$\int_0^1 t^{-z} \frac{k(t)}{t} dt \leq const \int_0^1 t^{-z} dt = \begin{cases} < \infty\,, & z < 1\,, \\ = \infty\,, & z \geq 1\,. \end{cases}$$

Furthermore, by the dominated and monotone convergence theorem, respectively,

$$\begin{aligned}\lim_{\delta \downarrow 0} \check{k}(a+\delta) &= \lim_{\delta \downarrow 0} E\left(1_{\{|\alpha+\sqrt{\lambda}\,\varepsilon| \leq 1\}} |\alpha + \sqrt{\lambda}\,\varepsilon|^{-(a+\delta)} \right) \\ &\qquad + \lim_{\delta \downarrow 0} E\left(1_{\{|\alpha+\sqrt{\lambda}\,\varepsilon| > 1\}} |\alpha + \sqrt{\lambda}\,\varepsilon|^{-(a+\delta)} \right) \\ &= E\left(1_{\{|\alpha+\sqrt{\lambda}\,\varepsilon| \leq 1\}} |\alpha + \sqrt{\lambda}\,\varepsilon|^{T} \right) + E\left(1_{\{|\alpha+\sqrt{\lambda}\,\varepsilon| > 1\}} |\alpha + \sqrt{\lambda}\,\varepsilon|^{T} \right) \\ &= h_{\alpha,\lambda}(T) = \infty\end{aligned}$$

and

$$\begin{aligned}\lim_{\delta \downarrow 0} \check{k}(b-\delta) &= \lim_{\delta \downarrow 0} \int_0^\infty t^{-(1-\delta)} \frac{1}{\sqrt{\lambda}} \Big(p(\frac{t-\alpha}{\sqrt{\lambda}}) + p(\frac{t+\alpha}{\sqrt{\lambda}}) \Big) dt \\ &\geq const \lim_{\delta \downarrow 0} \int_0^1 t^{-(1+\delta)} dt = const \lim_{\delta \downarrow 0} \frac{1}{\delta} = \infty .\end{aligned}$$

Finally, by Lemma 3, we have

$$\lim_{x \to \infty} \frac{\int_0^\infty k(x/t) \overline{F}(t) dt/t}{\overline{F}(x)} = \lim_{x \to \infty} \int_0^\infty g(x,t) dt = 1$$

and hence condition (4.5) is fulfilled with $c = 1$. Therefore all assumptions of Theorem 7 are satisfied and we conclude (setting $\kappa = -\rho$)

$$\overline{F}(x) \sim x^{-\kappa} l(x) , \quad x \to \infty , \tag{4.15}$$

where l is some slowly varying function and κ is determined by the equation

$$E(|\alpha+\sqrt{\lambda}\varepsilon|^{\kappa})=1\,,\quad \text{for some } \kappa\in(-1,T)\,. \tag{4.16}$$

Since the tail of the stationary distribution function is decreasing, the solution κ in (4.16) has to be strictly positive and hence by Theorem 7 there exists a solution $\kappa\in(0,T)$ in (4.16) which is unique because of Lemma 2. Finally, with the proceeding described in the introduction of Section 4 it follows that the slowly varying function l is the constant c given in Theorem 8.

REMARK 10. The approach proposed in this paper for evaluating the tail of the stationary distribution of $(X_n)_{n\in\mathbb{N}}$ is quite lengthy and technical and requires the unpleasant conditions $(D.2)$ and $(D.3)$. Unfortunately, as already mentioned in the introduction of Section 4, there does not exist any obvious simpler derivation for general ε. However, in the case $\varepsilon\sim N(0,1)$, the result in Theorem 8 can be obtained much more easily using the special structure of the characteristic function of the normal distribution.

Recall that the random variable X which has the stationary distribution function is characterized by the fixpoint equation

$$X\stackrel{d}{=}\alpha X+\sqrt{\beta+\lambda X^2}\,\varepsilon\,. \tag{4.17}$$

Now note that for every $t\in\mathbb{R}$

$$\begin{aligned} E(e^{itX}) &= E(e^{it\alpha X}E(e^{it\sqrt{\beta+\lambda X^2}\,\varepsilon}\mid X))\\ &= e^{-\beta t^2/2}E(e^{it\alpha X-t^2\lambda X^2/2})\\ &= E(e^{it\sqrt{\beta}N_1})\,E(e^{it(\alpha X+\sqrt{\lambda}X\,N_2)})\,, \end{aligned} \tag{4.18}$$

where N_1 and N_2 are independent standard normal random variables, independent of X. From (4.18) we obtain the fixpoint equation

$$X\stackrel{d}{=}\sqrt{\beta}N_1+(\alpha+\sqrt{\lambda}N_2)X\,.$$

Hence X is limit variable of the ergodic process $(\widetilde{X}_n)_{n\in\mathbb{N}}$ given by the stochastic difference equation

$$\widetilde{X}_n=\sqrt{\beta}N_{1,n}+(\alpha+\sqrt{\lambda}N_{2,n})\widetilde{X}_{n-1}\,, \tag{4.19}$$

where $(N_{1,n},N_{2,n})_{n\in\mathbb{N}}$ is an iid sequence of random variables with same distribution as (N_1,N_2). The stationary distribution of the process $(\widetilde{X}_n)_{n\in\mathbb{N}}$ follows from Goldie (1991, Theorem 4.1), see also Embrechts et al. (1997), Section 8.4.

Acknowledgement. We take pleasure in thanking Charles Goldie for many discussions, helpful comments and his hospitality in Brighton. We are also grateful to Thomas Mikosch for useful comments and Hans-Jürgen Schuh for making us aware of one of Tweedie's papers. Finally, we would like to thank an unknown referee whose insightful comments led to a substantial improvement of the paper.

REFERENCES

ALSMEYER, G. (2000) On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. Preprint. Institut für Mathematische Statistik, University of Münster.

ASMUSSEN, S. (1987) *Applied Probability and Queues.* Wiley, Chichester.

BILLINGSLEY, P. (1995) *Probability and Measure.* Wiley, New York.

BINGHAM, N.H., GOLDIE, C.M. AND TEUGELS, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroscedasticity. *Journ. Econometrics* **31**, 307–327.

BOLLERSLEV, T., CHOU, R.Y. AND KRONER, K.F. (1992) ARCH modelling in finance: a review of the theory and empirical evidence. *Journ. Econometrics* **52**, 5-59.

BORKOVEC, M. (2000) Extremal behavior of the autoregressive process with ARCH(1) errors. *Stoch. Proc. Appl.* **85**, 189-207.

BORKOVEC, M. (2001) Asymptotic behavior of the sample autocovariance and autocorrelation function of an AR(1) process with ARCH(1) errors. Submitted for publication. *http://www.ma.tum.de/stat/*

DIEBOLT, J. AND GUÉGAN, D. (1990) Probabilistic properties of the general nonlinear Markovian process of order one and applications to time series modelling. *Rapp. Tech.* **No.125**, L.S.T.A., Paris 6.

EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Heidelberg.

ENGLE, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation. *Econometrica* **50**, 987–1007.

FAMA, E.F. (1965) The behavior of stock market prices. *Journ. Business* **38**, 34–105.

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

GOLDIE, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Prob.* **1**, 126–166.

JOHNSON, N.L., KOTZ, S. AND BALAKRISHNAN, N. (1995) *Continuous Univariate Distributions, Vol.2.* Wiley, New York.

KESTEN, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

KIEFERSBECK, N. (1999) *Stationarität und Tailindex von zeitdiskreten Volatilitätsmodellen*, Diploma Thesis.
http://www-m4.mathematik.tu-muenchen.de/m4/Diplarb/

MAERCKER, G. (1997) *Statistical Inference in Conditional Heteroskedastic Autoregressive Models.* Shaker, Aachen. (Dissertation, University of Braunschweig)

MANDELBROT, B. (1963) The variation of certain speculative prices. *Journ. Business* **36**, 394–419.

MEYN, S.P. AND TWEEDIE, R.L. (1993) *Markov Chains and Stochastic Stability.* Springer, London.

PRIESTLEY, M.B. (1988) *Non-linear and Non-stationary Time Series.* Academic Press, New York.

TAYLOR, S. (1995) *Modelling financial time series.* Wiley, Chichester.

TONG, H. (1990) *Non-linear Time Series – A Dynamical System Approach.* Oxford University Press, Oxford.

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

TWEEDIE, R.L. (1983a) Criteria for rates of convergence of Markov chains, with application to queueing and storage theory. In J.F.C. Kingman and G.E.H. Reuter, editors, *Probability, Statistics and Analysis.* London Mathematics Society Lecture Note Series, Cambridge University Press, Cambridge.

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

VERVAAT, W. (1979) On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.

WEISS, A.A. (1984) ARMA models with ARCH errors. *J. Time Series Analysis,* **3**, 129–143.

MILAN BORKOVEC AND CLAUDIA KÜPPELBERG
CENTER OF MATHEMATICAL SCIENCES
MUNICH UNIVERSITY OF TECHNOLOGY
D-80290 MUNICH, GERMANY
Email: cklu@ma.tum.de
http://www.ma.tum.de/stat/