Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 117 (2007) 840–861
stochastic processes and their applications
www.elsevier.com/locate/spa

Regular variation of order 1 nonlinear AR-ARCH models

Daren B.H. Cline*

Department of Statistics, Texas A&M University, College Station, TX 77843-3143, United States

Received 8 August 2005; received in revised form 24 July 2006; accepted 17 October 2006
Available online 7 November 2006

Abstract

We prove both geometric ergodicity and regular variation of the stationary distribution for a class of nonlinear stochastic recursions that includes nonlinear AR-ARCH models of order 1. The Lyapounov exponent for the model, the index of regular variation and the spectral measure for the regular variation all are characterized by a simple two-state Markov chain.

MSC: primary 60G10; 60J05; secondary 60G70; 62M10; 91B84

Keywords: ARCH; Ergodicity; Regular variation; Stationary distribution; Stochastic recursion

1. Introduction

1.1. Overview

Several papers have been devoted to bounding and/or characterizing the probability tails of the stationary distribution for a (generalized) autoregressive conditional heteroscedastic ((G)ARCH) model [15,19,30,3,23]. In each of these, the conditional variances can be characterized as linear in the squared components of the "state vector" and the model can be embedded in a random (matrix) coefficients model, with iid coefficients. This puts it within the stochastic recursion framework of Kesten [22] and Goldie [17] who used renewal theory arguments to identify the tail behavior. Unfortunately, this framework does not allow for extended models such as a combined

* Tel.: +1 979 8451443; fax: +1 979 8453144.
E-mail address: dcline@stat.tamu.edu.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2006.10.009

AR-(G)ARCH model or a threshold (G)ARCH model. Any attempt to embed these models in random coefficients models leads to "coefficients" that are no longer independent and, indeed, not known *a priori* even to be stationary.

Recent papers that have capitalized on regular variation of (G)ARCH models to study the sample autocovariance function include Davis and Mikosch [14], Mikosch and Stărică [25] and Borkovec [6]. Papers that deal with extremal behavior include Borkovec [5], Hult and Lindskog [20] and Hult, Lindskog, Mikosch and Samorodnitsky [21].

In this paper we will provide conditions for, and characterize, both the ergodicity and the tail behavior of a general one-dimensional stochastic recursion model that includes standard nonlinear ARCH and AR-ARCH models. The results here are precise, as opposed to the stronger ergodicity condition and bounds given in Diebolt and Guégan [15] and Guégan and Diebolt [19]. Our approach will avoid a random coefficient embedding and therefore may have more promise for other nonlinear models. Instead, we use the *piggyback* method of Cline and Pu [13] to show ergodicity and we verify and solve an invariance equation to determine regular variation. Like Borkovec and Klüppelberg [7], who studied an order 1 AR-ARCH model, our approach is essentially Tauberian in nature but it applies more generally to nonlinear models.

Specifically, we consider the Markov chain on $\mathbb{R}$ given by

$$\xi_t = a(\xi_{t-1}, e_t) \stackrel{\text{def}}{=} b(\xi_{t-1}/|\xi_{t-1}|, e_t)|\xi_{t-1}| + c(\xi_{t-1}, e_t) \tag{1.1}$$

where $\{e_t\}$ is an iid sequence, $|b(x/|x|, u)| \leq \bar{b}(1 + |u|)$ and $|c(x, u)| \leq \bar{c}(1 + |u|)$ for finite $\bar{b}, \bar{c}$. The point to be made here is that the first term on the right is homogeneous in ξ_{t-1} while the second is bounded in ξ_{t-1}. Such a decomposition is possible for any first order AR-ARCH model and for first order threshold AR-ARCH models. For example, suppose

$$\xi_t = a(\xi_{t-1}, e_1) = \begin{cases} a_{10} + a_{11}\xi_{t-1} + (b_{10} + b_{11}\xi_{t-1}^2)^{1/2} e_t, & \text{if } \xi_{t-1} < x_1, \\ a_{20} + a_{21}\xi_{t-1} + (b_{20} + b_{21}\xi_{t-1}^2)^{1/2} e_t, & \text{if } x_1 \leq \xi_{t-1} \leq x_2, \\ a_{30} + a_{31}\xi_{t-1} + (b_{30} + b_{31}\xi_{t-1}^2)^{1/2} e_t, & \text{if } \xi_{t-1} > x_2, \end{cases} \tag{1.2}$$

with each $b_{ij} \geq 0$. Then we may set $b(-1, u) = -a_{11} + b_{11}^{1/2} u$, $b(1, u) = a_{31} + b_{31}^{1/2} u$ and $c(x, u) = a(x, u) - b(x/|x|, u)|x|$.

A similar decomposition holds for models with smooth transitions and for certain random switching models (see Section 3).

1.2. Assumptions

Throughout we assume the following.

Assumption A.1. The error sequence $\{e_t\}$ is iid and $E(|e_t|^\beta) < \infty$ for all $\beta > 0$.

Assumption A.2. There exist $\bar{b} < \infty$, $\tilde{b}_1 > 0$, $\tilde{b}_2 \geq 0$ and $\bar{c} < \infty$ such that

(i) $\max(\tilde{b}_1|u| - \tilde{b}_2, 0) \leq |b(\theta, u)| \leq \bar{b}(1 + |u|)$ for all $u \in \mathbb{R}$, $\theta \in \{-1, 1\}$, and
(ii) $|c(x, u)| \leq \bar{c}(1 + |u|)$ for all $u \in \mathbb{R}$, $x \in \mathbb{R}$.

Note the lower bound on $b(\theta, u)$ as well as the upper bound. This is the generalized ARCH-like behavior and it also applies to random coefficient and bilinear models.

Assumption A.3. For each $\theta \in \{-1, 1\}$, $b(\theta, e_1)$ has absolutely continuous distribution, $0 < P(b(\theta, e_1) > 0) < 1$, $E(|\log(|b(\theta, e_1)|)|) < \infty$, and either

$$\Delta_{-1} \overset{\text{def}}{=} \min_{\theta=\pm 1} \liminf_{w\to\infty} \frac{P(b(\theta, e_1) < -w)}{P(|b(\theta, e_1)| > w)} > 0$$

or

$$\Delta_1 \overset{\text{def}}{=} \min_{\theta=\pm 1} \liminf_{w\to\infty} \frac{P(b(\theta, e_1) > w)}{P(|b(\theta, e_1)| > w)} > 0.$$

In the time series literature, one often sees the assumption that e_t has a positive density. In such a case, Assumption A.3 simply requires some regularity on the functions $b(-1, \cdot)$ and $b(1, \cdot)$. However, even in a nonlinear time series setting, the assumption typically applies.

Assumption A.4. $\{\xi_t\}$ is an aperiodic, Lebesgue irreducible T-chain.

The reader is asked to refer to standard texts on Markov processes (such as [24]) for the definition of these terms, as well as the terms "ergodic" and "transient". The T-chain property is a generalization of the Feller property and is needed here because, as is common with threshold models, the transition probabilities may not be continuous in the current state.

We are making the last assumption outright, as the primary focus of this paper is on the regular variation of the tails of the stationary distribution rather than on the ergodicity of the process, though we do identify a critical condition for ergodicity. Assumption A.4 will be valid, however, if the following hold (cf. [10]).

(i) The distribution of e_t has Lebesgue density f on $\mathbb{R}$ which is bounded and locally bounded away from 0, and
(ii) for each $x \in \mathbb{R}$, $a(x, \cdot) = b(x/|x|, \cdot)|x| + c(x, \cdot)$ is strictly increasing, with a derivative that is locally bounded and locally bounded away from 0, locally uniformly in x.

In particular, (1.2) satisfies Assumptions A.2–A.4 if (i) holds and each $b_{i0} > 0$, $i = 1, 2, 3$, and each $b_{i1} > 0$, $i = 1, 3$. These assumptions are likewise easily checked for each of the examples in Section 3.

1.3. Objectives

Our objectives are two-fold.

First, we establish a sufficient condition for $\{\xi_t\}$ to be geometrically ergodic, meaning that

$$\lim_{n\to\infty} r^n \sup_A |P(\xi_n \in A \mid \xi_0 = x) - \Pi(A)| < \infty$$

for some $r > 1$, some probability distribution Π and every $x \in \mathbb{R}$ [24, Ch. 15]. Simply stated, the condition is that the (largest) Lyapounov exponent of the process,

$$\liminf_{n\to\infty} \limsup_{|x|\to\infty} \frac{1}{n} E\left(\log\left(\frac{1 + |\xi_n|}{1 + |\xi_0|}\right) \,\middle|\, \xi_0 = x\right), \tag{1.3}$$

is negative, meaning ξ_t tends to contract when very large in magnitude.

In a random coefficients setting, Bougerol and Picard [8,9] define the Lyapounov exponent in terms of the asymptotic behavior of the sequential product of random coefficients. Its value is easily seen to equal a limiting behavior of the process itself, such as the limit above. Indeed, as will become clear in the next section, the Lyapounov exponent in our context also may be interpreted in terms of a sequential product of random variables. (See [13], also.) We point out, however, that our definition is not to be confused with the Lyapounov exponent of a noisy chaos.

The key result is that the value of this exponent may be expressed in terms of the stationary distribution of a simpler process ((1.6) below). We actually will verify geometric ergodicity through the Foster–Lyapounov drift condition method, thereby endowing the process with mixing, strong laws, etc. (cf. [24]).

The second, and greater, objective is to verify that if $\{\xi_t\}$ satisfies an appropriate drift condition then its stationary distribution Π has regularly varying tails with some index $-\kappa < 0$. That is, under stationarity,

$$\lim_{r\to\infty} \frac{P(\xi_t < -\lambda r)}{P(|\xi_t| > r)} = \mu_{-1}\lambda^{-\kappa} \quad \text{and} \quad \lim_{r\to\infty} \frac{P(\xi_t > \lambda r)}{P(|\xi_t| > r)} = \mu_1\lambda^{-\kappa}, \quad \text{all } \lambda > 0. \tag{1.4}$$

Knowing that Π has regularly varying tails helps to establish the existence of moments (none are of order greater than κ) and limit theorems for statistics such as the sample autocovariance and autocorrelation functions (see the references in Section 1.1).

Let $(R_t, \tilde{\theta}_t) = (|\xi_t|, \xi_t/|\xi_t|)$ and define

$$w(\theta, u) = |b(\theta, u)|, \qquad \eta(\theta, u) = b(\theta, u)/|b(\theta, u)|, \quad \text{for } \theta \in \{-1, 1\},\ u \in \mathbb{R}.$$

A related (though inherently non-ergodic) process is the homogeneous form of (1.1):

$$\xi_t^* = b(\xi_{t-1}^*/|\xi_{t-1}^*|, e_t)|\xi_{t-1}^*|. \tag{1.5}$$

This can be collapsed to a two-state Markov chain on $\{-1, 1\}$:

$$\theta_t^* \overset{\text{def}}{=} \xi_t^*/|\xi_t^*| = \eta(\theta_{t-1}^*, e_t). \tag{1.6}$$

Also, let $W_t^* = w(\theta_{t-1}^*, e_t)$. The "collapsed" process is Markov and ergodic. Its behavior (and more specifically, the behavior of W_t^*) determines both the ergodicity and the distribution tails of the original process $\{\xi_t\}$.

2. Main results

2.1. The collapsed process

We first describe the principal properties of the process $\{\theta_t^*\}$ which will, in turn, inform the behavior of $\{\xi_t\}$. Let

$$p_{ij} = P(\theta_1^* = j \mid \theta_0^* = i) = P(\eta(i, e_1) = j), \quad i, j \in \{-1, 1\}. \tag{2.1}$$

Then, clearly, $\{\theta_t^*\}$ has stationary distribution given by

$$\pi_1 = 1 - \pi_{-1} = \frac{p_{-1,1}}{p_{1,-1} + p_{-1,1}}.$$

To establish the ergodicity criterion (in the proof of Theorem 2.2), we will require a function $\nu : \{-1, 1\} \to \mathbb{R}$ and a constant γ which solve the equilibrium (Poisson) equation

$$E(\nu(\theta_1^*) - \nu(\theta_0^*) + \log W_1^* \mid \theta_0^* = i) = \gamma, \quad i = \pm 1. \tag{2.2}$$

The solution is easily seen to be

$$\nu(\pm 1) = \pm \frac{E(\log W_1^* \mid \theta_0^* = 1) - E(\log W_1^* \mid \theta_0^* = -1)}{2(p_{1,-1} + p_{-1,1})} \tag{2.3}$$

with

$$\begin{aligned}\gamma &= \pi_{-1} E(\log W_1^* \mid \theta_0^* = -1) + \pi_1 E(\log W_1^* \mid \theta_0^* = 1) \\ &= \pi_{-1} E(\log |b(-1, e_1)|) + \pi_1 E(\log |b(1, e_1)|),\end{aligned} \tag{2.4}$$

the expectation of $\log W_1^*$ under the stationary distribution π. Since the collapsed process is ergodic, it is clear that

$$\gamma = \lim_{n\to\infty} \frac{1}{n} E(\log(W_1^* \cdots W_n^*)) = \lim_{n\to\infty} \frac{1}{n} \log(W_1^* \cdots W_n^*) \quad \text{a.s.}$$

Ergodicity of $\{\xi_t\}$ depends on the value of γ. The regular variation, however, relies on a different set of characters from the collapsed process. These are given in the following lemma.

Lemma 2.1. *Suppose the value of γ in (2.4) is negative. Then there exist unique $\kappa > 0$ and probability measure μ on $\{-1, 1\}$ such that μ is invariant for the (transition) matrix M_κ with elements*

$$m_{\kappa ij} \stackrel{\text{def}}{=} E((W_1^*)^\kappa 1_{\theta_1^* = j} \mid \theta_0^* = i) = E(|b(i, e_1)|^\kappa 1_{\eta(i,e_1)=j}), \quad i, j \in \{-1, 1\}. \tag{2.5}$$

For this κ, M_κ has maximal eigenvalue 1 and μ is the corresponding left eigenvector with

$$\mu_1 = 1 - \mu_{-1} = \frac{m_{\kappa,-1,1}}{1 - m_{\kappa,1,1} + m_{\kappa,-1,1}} = \frac{1 - m_{\kappa,-1,-1}}{1 - m_{\kappa,-1,-1} + m_{\kappa,1,-1}}. \tag{2.6}$$

Actually evaluating the κ in Lemma 2.1 seems to be a non-trivial task. Since M_κ is a 2×2 matrix, we can say that the solution must satisfy

$$m_{\kappa,-1,-1} < 1, \qquad m_{\kappa,1,1} < 1 \quad \text{and} \quad (1 - m_{\kappa,-1,-1})(1 - m_{\kappa,1,1}) = m_{\kappa,-1,1} m_{\kappa,1,-1}, \tag{2.7}$$

or, equivalently,

$$m_{\kappa,-1,-1} + m_{\kappa,1,1} + \sqrt{(m_{\kappa,-1,-1} - m_{\kappa,1,1})^2 + 4 m_{\kappa,-1,1} m_{\kappa,1,-1}} = 2.$$

2.2. *Geometric ergodicity*

The now quite standard argument for ergodicity of a nonlinear time series, and for Markov chains in general, includes demonstrating a Foster–Lyapounov drift condition. Ours is no exception. The basic idea of the piggyback method is that a Foster–Lyapounov test function may be computed from the equilibrium equation (2.2).

Indeed, the value γ from the equilibrium equation (2.2) holds the key to ergodicity. The following is taken from Cline and Pu [13]. We will demonstrate it here as well, however, partly because the (one-dimensional) model here is more general and partly because the earlier arguments were specifically designed for a multidimensional Markov model.

Theorem 2.2. *Let γ be as in (2.2) and (2.4).*

(i) *The Lyapounov exponent for $\{\xi_t\}$ (see (1.3)) is γ. Indeed,*

$$\lim_{n\to\infty} \limsup_{|x|\to\infty} \left| \frac{1}{n} E(\log(|\xi_n|/|\xi_0|) \mid \xi_0 = x) - \gamma \right| = 0. \tag{2.8}$$

(ii) *Suppose $\gamma < 0$ and let κ be as in Lemma 2.1. For any $0 < \zeta < \kappa$, there exists a function $V : \mathbb{R} \to \mathbb{R}_+$ satisfying*

(a) *there exist finite, positive d_1, d_2 such that*

$$d_1|x|^\zeta \le V(x) \le d_2(1 + |x|^\zeta), \tag{2.9}$$

and

(b) *there exist finite M_0, K_0, and $\rho < 1$ such that*

$$E(V(\xi_1) \mid \xi_0 = x) \le \rho V(x) 1_{V(x)>M_0} + K_0 1_{V(x)\le M_0}, \quad \textit{for all } x \in \mathbb{R}. \tag{2.10}$$

(iii) *If $\gamma < 0$ then $\{\xi_t\}$ is geometrically ergodic, but if $\gamma > 0$ then $\{\xi_t\}$ is transient.*

When $\gamma < 0$, we let Π be the stationary distribution for $\{\xi_t\}$.

2.3. Regular variation

We now describe the tail behavior for the stationary distribution Π. For our argument, it will be advantageous to think of Π as the stationary distribution of $(R_t, \tilde{\theta}_t) = (|\xi_t|, \xi_t/|\xi_t|)$ and to define the measure Q_v on $\mathbb{R}_+ \times \{-1, 1\}$ by

$$Q_v((r, \infty) \times \{i\}) = \frac{\Pi((rv, \infty) \times \{i\})}{\Pi((v, \infty) \times \{-1, 1\})}, \quad \text{for } r > 0, i \in \{-1, 1\}. \tag{2.11}$$

Regular variation of Π (recall (1.4)) is equivalent to $Q_v \overset{\mathrm{v}}{\to} Q$ (vague convergence) as $v \to \infty$, for some measure Q with $Q((1, \infty) \times \{-1, 1\}) = 1$. If this occurs then necessarily [26, p. 277]

$$Q((r, \infty) \times \{i\}) = r^{-\kappa} \mu(\{i\}),$$

with some index of regular variation $\kappa > 0$ and some spectral probability measure μ on $\{-1, 1\}$. In fact, we can identify κ and μ from the collapsed process.

Theorem 2.3. *Suppose the Lyapounov exponent γ is negative and $\{\xi_t\}$ has stationary distribution Π. Let κ and μ be as in Lemma 2.1. Then Π has regularly varying tails with index of regular variation κ and spectral probability measure μ. That is, (1.4) holds.*

We note that our assumptions of irreducibility and $0 < P(B_1 > 0) < 1$ ensure that *both* probability tails are regularly varying. A one-sided result holds as well but arguing it would require specialization in the proof of the theorem and of Lemma 4.2 below, and we leave this to the reader. See [17] for one-sided examples under continuity assumptions.

In proving regular variation, we will first verify that the probability tails of R_t are dominated varying, under stationarity. This will entail consideration of the Matuszewska indices (cf. [4, Ch. 2]), defined as follows.

Definition 2.4. Let $p(v)$ be a positive function on $(0, \infty)$.

(i) The upper Matuszewska index for p is the infimum of those α such that

$$\inf_{c>1} \limsup_{v\to\infty} \sup_{1\le\lambda\le c} \frac{\lambda^{-\alpha} p(\lambda v)}{p(v)} < \infty.$$

(ii) The lower Matuszewska index for p is the supremum of those β such that

$$\sup_{c>1} \liminf_{v\to\infty} \inf_{1\le\lambda\le c} \frac{\lambda^{-\beta} p(\lambda v)}{p(v)} > 0.$$

Since probability tails are nonincreasing, the indices will be nonpositive. More importantly, we will need to verify that they are finite, negative and equal. Although equality of the Matuszewska indices generally does not imply regular variation, it will in fact suffice for us.

3. Examples

3.1. Random coefficients model

Goldie [17] analyzes the tail behavior for the stationary distributions of models of the form

$$\xi_t = B_t \xi_{t-1} + c(\xi_{t-1}, B_t, C_t), \tag{3.1}$$

where $e_t \stackrel{\text{def}}{=} (B_t, C_t)$ is an iid sequence in $\mathbb{R}^2$, $c(\cdot, B, C)$ is *continuous* for each (B, C) and $|c(x, B, C)| \le \bar{c}(1 + |B| + |C|)$ for some finite $\bar{c}$. An important special case, studied by Kesten [22] and also by de Saporta [27], is the one-dimensional random coefficients model

$$\xi_t = B_t \xi_{t-1} + C_t.$$

Model (3.1) is a special case of (1.1) with $b(x, B, C) = \operatorname{sgn}(x)B$. There is no loss in allowing e_t to be multidimensional as long as our other assumptions are met. Those assumptions are not automatic, however. For example, $C_t = m(1 - B_t)$ almost surely for some constant m leads to a degenerate stationary distribution for the random coefficients model (cf. [17]), but the model is not irreducible. (See also [12].)

From (2.4), $\gamma = E(\log|b(\pm 1, e_1)|) = E(\log|B_1|)$. Verwaat [29] and Grincevičius [18] (for example) showed that $\gamma < 0$ suffices for ergodicity. Likewise, from Lemma 2.1, the parameter κ satisfies $E(|B_t|^\kappa) = 1$ and $\mu_1 = \mu_{-1} = \frac{1}{2}$ since $m_{\kappa,-1,1} = 1 - m_{\kappa,1,1} = E(|B_t|^\kappa 1_{B_t<0})$, in agreement with Goldie (under the assumption $0 < P(B_1 > 0) < 1$).

3.2. AR-ARCH model

The AR-ARCH model of order 1 is

$$\xi_t = a_0 + a_1 \xi_{t-1} + (b_0 + b_1 \xi_{t-1}^2)^{1/2} e_t.$$

This is the model examined by Borkovec and Klüppelberg [7], under the additional assumption that e_t has a distribution symmetric about 0. The ordinary ARCH(1) model is a special case with $a_1 = a_0 = 0$. If $a_1 \ne 0$, however, the combination of an autoregression term with the ARCH term precludes the possibility of embedding it in a random coefficients model. We have $b(i, e_1) = ia_1 + b_1^{1/2} e_1$, $i = \pm 1$, so that

$$p_{-1,1} = P(-a_1 + b_1^{1/2} e_1 > 0) \quad \text{and} \quad p_{1,-1} = P(a_1 + b_1^{1/2} e_1 < 0).$$

From (2.4), the Lyapounov exponent is

$$\gamma = \frac{p_{1,-1} E(\log|a_1 - b_1^{1/2} e_1|) + p_{-1,1} E(\log|a_1 + b_1^{1/2} e_1|)}{p_{1,-1} + p_{-1,1}}.$$

The index of regular variation, κ, solves (2.7) with

$$m_{\kappa ij} = E(|ia_1 + b_1^{1/2}e_1|^\kappa 1_{j(ia_1+b_1^{1/2}e_1)>0}), \quad i, j \in \{-1, 1\},$$

and the tail weights are given by

$$\mu_1 = 1 - \mu_{-1} = \frac{E(|a_1 - b_1^{1/2}e_1|^\kappa 1_{a_1 - b_1^{1/2}e_1<0})}{1 - E(|a_1 - b_1^{1/2}e_1|^\kappa 1_{a_1-b_1^{1/2}e_1<0}) + E(|a_1 + b_1^{1/2}e_1|^\kappa 1_{a_1+b_1^{1/2}e_1>0})}.$$

When e_1 is assumed to have a symmetric distribution, the results simplify considerably. In this case, $|a_1 - b_1^{1/2}e_1| \overset{\mathrm{D}}{=} |a_1 + b_1^{1/2}e_1|$ so that $\gamma = E(\log|a_1 + b_1^{1/2}e_1|)$, $\mu_{-1} = \mu_1 = 1/2$ and κ solves $E(|a_1 + b_1^{1/2}e_1|^\kappa) = 1$.

When $a_0 = a_1 = 0$, we of course have the standard ARCH model. Here, $\gamma = \log b_1^{1/2} + E(\log|e_1|)$, κ satisfies $b_1^{\kappa/2}E(|e_1|^\kappa) = 1$ and $\mu_1 = b_1^{\kappa/2}E(|e_1|^\kappa 1_{e_1>0})$. Note that ξ_t^2 satisfies a random coefficients model. Goldie's results would only determine the tail properties of $|\xi_t|$, whereas we also identify the tail weights.

3.3. Threshold AR-ARCH model

The results for the threshold model (1.2) are only slightly more involved. Here, we have

$$p_{-1,1} = P(-a_{11} + b_{11}^{1/2}e_1 > 0) \quad \text{and} \quad p_{1,-1} = P(a_{31} + b_{31}^{1/2}e_1 < 0).$$

The Lyapounov exponent is

$$\gamma = \frac{p_{1,-1}E(\log|a_{11} - b_{11}^{1/2}e_1|) + p_{-1,1}E(\log|a_{31} + b_{31}^{1/2}e_1|)}{p_{1,-1} + p_{-1,1}}$$

and κ solves (2.7) with

$$m_{\kappa,-1,j} = E(|a_{11} - b_{11}^{1/2}e_1|^\kappa 1_{j(a_{11}-b_{11}^{1/2}e_1)<0}), \quad j \in \{-1, 1\},$$

and

$$m_{\kappa,1,j} = E(|a_{13} + b_{13}^{1/2}e_1|^\kappa 1_{j(a_{31}+b_{31}^{1/2}e_1)>0}), \quad j \in \{-1, 1\}.$$

Again, these quantities are used in (2.6) to compute μ_1 and μ_{-1}.

For a threshold ARCH model (without the autoregression term), $a_{11} = a_{31} = 0$. Consequently,

$$\gamma = p\log b_{11}^{1/2} + (1-p)\log b_{31}^{1/2} + E(\log|e_1|),$$

where $p = P(e_1 < 0)$. Also, $\mu_1 = E\left(|e_1|^\kappa 1_{e_1>0}\right)/E\left(|e_1|^\kappa\right)$ and κ solves

$$b_{11}^{\kappa/2}E(|e_1|^\kappa 1_{e_1<0}) + b_{31}^{\kappa/2}E(|e_1|^\kappa 1_{e_1>0}) = 1.$$

Smooth transition models also fall within the framework here. Suppose G is a continuous probability distribution function on $\mathbb{R}$, with $\sup_{x\in\mathbb{R}}|x|G(x)(1-G(x)) < \infty$, and

$$\begin{aligned}\xi_t &= (a_{10} + a_{11}\xi_{t-1} + (b_{10} + b_{11}\xi_{t-1}^2)^{1/2}e_t)(1 - G(\xi_{t-1}))\\ &\quad + (a_{30} + a_{31}\xi_{t-1} + (b_{30} + b_{31}\xi_{t-1}^2)^{1/2}e_t)G(\xi_{t-1}).\end{aligned}$$

Then the above conclusions hold exactly as stated.

3.4. Random switching AR-ARCH model

Our results allow for some nonlinearity in the errors. For example, regime switching could be signaled by the value (or sign) of the errors rather than by the time series itself. A simple example that satisfies our assumptions is

$$\xi_t = a_0 + a_1\xi_{t-1} + (b_0 + b_1\xi_{t-1}^2)^{1/2} e_t G(e_t) - (d_0 + d_1\xi_{t-1}^2)^{1/2} e_t G(-e_t),$$

where again G is a continuous probability distribution function on $\mathbb{R}$. Now $b(i, e_1) = ia_1 + b_1^{1/2} e_t G(e_t) - d_1^{1/2} e_t G(-e_t)$, $i = \pm 1$, and γ, κ and μ can be computed accordingly from (2.4) and Lemma 2.1.

4. Proofs

4.1. Showing ergodicity

Here we show that γ is in fact the Lyapounov exponent for $\{\xi_t\}$ and that $\gamma < 0$ implies $\{\xi_t\}$ is geometrically ergodic. This argument is actually a much simpler version of the piggyback argument in Cline and Pu [13] where we dealt primarily with higher order AR-ARCH models.

Lemma 4.1. *Let ν and γ be as in (2.3) and (2.4), respectively. Extend ν to $\mathbb{R}$ by $\nu(x) = \nu(x/|x|)$ if $x \neq 0$ and $\nu(0) = 0$. Then*

$$\lim_{|x|\to\infty} E\left(\nu(\xi_1) - \nu(\xi_0) + \log\left(\frac{1+|\xi_1|}{1+|\xi_0|}\right) \middle|\ \xi_0 = x\right) = \gamma. \tag{4.1}$$

Proof. By the definitions of ξ_t and θ_t^*, if $x = i|x|$, $i = \pm 1$, then

$$\begin{aligned} |E(\nu(\theta_1^*) \mid \theta_0^* = i) - E(\nu(\xi_1) \mid \xi_0 = x)| &\le |\nu(1)| P(\eta(i, e_1) \neq a(x, e_1)/|a(x, e_1)|) \\ &\le |\nu(1)| P(|c(x, e_1)| > |b(x/|x|, e_1)|\,|x|) \\ &\le |\nu(1)| P(\bar{c}(1 + |e_1|) > |b(i, e_1)|\,|x|). \end{aligned} \tag{4.2}$$

Obviously, therefore,

$$\lim_{|x|\to\infty,\, x/|x|=i} |E(\nu(\theta_1^*) \mid \theta_0^* = i) - E(\nu(\xi_1) \mid \xi_0 = x)| = 0. \tag{4.3}$$

By Assumption A.3, $E(|\log W_1^*| \mid \theta_0^* = i) = E(|\log(|b(i, e_1)|)|) < \infty$. Also, Assumptions A.1 and A.2 imply

$$E\left(\log\left(\frac{1 + |b(i, e_1)|x| + c(x, e_1)|}{1 + |x|}\right)\right) \le E\left(\log\left(1 + |b(i, e_1)| + |c(x, e_1)|\right)\right) < \infty,$$

$$E\left(\log\left(\frac{1 + |b(i, e_1)x|/2}{1 + |x|}\right)\right) \ge E\left(\log\left(|b(i, e_1)|/2\right) 1_{|b(i,e_1)x|<2}\right) > -\infty$$

and

$$\begin{aligned} &E\left(\log\left(\frac{1 + |b(i, e_1)|x| + c(x, e_1)|}{1 + |b(i, e_1)||x|/2}\right)\right) \\ &\quad \ge E(-\log\left(1 + |b(i, e_1)x|/2\right) 1_{|c(x,e_1)|>|b(i,e_1)x|/2}) \\ &\quad \ge E(-\log\left(1 + |c(x, e_1)|\right)) > -\infty. \end{aligned}$$

Then easily by dominated convergence,

$$\lim_{|x|\to\infty,\, x/|x|=i} E\left(\log\left(\frac{1+|\xi_1|}{1+|\xi_0|}\right)\,\middle|\, \xi_0 = x\right)$$
$$= \lim_{|x|\to\infty,\, x/|x|=i} E\left(\log\left(\frac{1+|b(i,e_1)|x|+c(x,e_1)|}{1+|x|}\right)\right)$$
$$= E(\log b(i,e_1)) = E(\log W_1^* \mid \theta_0^* = i). \tag{4.4}$$

The conclusion (4.1) follows from (2.2), (4.3) and (4.4). □

Proof of Theorem 2.2. (i) Fix $L < \infty$ arbitrarily. Observe that

$$\limsup_{|x|\to\infty} P(|\xi_1| \le L \mid \xi_0 = x) = \limsup_{|x|\to\infty} P(|a(x,e_1)| \le L) = 0.$$

Let $\epsilon > 0$ and choose L_0 such that $\sup_{|x|>L_0} P(|\xi_1| \le L \mid \xi_0 = x) < \epsilon$. Thus,

$$\limsup_{|x|\to\infty} P(|\xi_t| \le L \mid \xi_0 = x)$$
$$\le \limsup_{|x|\to\infty} E(P(|\xi_t| \le L \mid \xi_{t-1}) 1_{|\xi_{t-1}|>L_0} \mid \xi_0 = x) + \limsup_{|x|\to\infty} P(|\xi_{t-1}| \le L_0 \mid \xi_0 = x)$$
$$\le \epsilon + \limsup_{|x|\to\infty} P(|\xi_{t-1}| \le L_0 \mid \xi_0 = x).$$

Hence, inductively, for any $L < \infty$,

$$\limsup_{|x|\to\infty} P(|\xi_t| \le L \mid \xi_0 = x) = 0, \quad \text{each } t \ge 1. \tag{4.5}$$

Now let $\bar{B}_t = \nu(\xi_t) - \nu(\xi_{t-1}) + \log(\frac{1+|\xi_t|}{1+|\xi_{t-1}|})$ for $t \ge 1$. Fix $\epsilon > 0$. From Lemma 4.1 we may choose L_1 such that

$$\sup_{|x|>L_1} |E(\bar{B}_1 \mid \xi_0 = x) - \gamma| < \epsilon.$$

Also, let

$$L_2 = \sup_{|x|\le L_1} |E(\bar{B}_1 \mid \xi_0 = x) - \gamma|.$$

Then, using (4.5),

$$\begin{aligned}\limsup_{|x|\to\infty} |E(\bar{B}_t \mid \xi_0 = x) - \gamma| &\le \limsup_{|x|\to\infty} E(|E(\bar{B}_t \mid \xi_{t-1}) - \gamma| \mid \xi_0 = x) \\ &\le \limsup_{|x|\to\infty} E(\epsilon 1_{|\xi_{t-1}|>L_1} + L_2 1_{|\xi_{t-1}|\le L_1} \mid \xi_0 = x) \\ &\le \epsilon + L_2 \limsup_{|x|\to\infty} P(|\xi_{t-1}| \le L_1 \mid \xi_0 = x) \le \epsilon.\end{aligned}$$

Therefore, since ϵ is arbitrary,

$$\limsup_{|x|\to\infty} |E(\bar{B}_t \mid \xi_0 = x) - \gamma| = 0, \quad \text{each } t \ge 1. \tag{4.6}$$

From (4.6) we thus have

$$\limsup_{|x|\to\infty} \left|\frac{1}{n} E\left(\nu(\xi_n) - \nu(\xi_0) + \log\left(\frac{1+|\xi_n|}{1+|\xi_0|}\right)\middle|\, \xi_0 = x\right) - \gamma\right|$$

$$\leq \frac{1}{n} \sum_{t=1}^{n} \limsup_{|x|\to\infty} |E(\bar{B}_t \mid \xi_0 = x) - \gamma| = 0$$

and conclude

$$\lim_{n\to\infty} \limsup_{|x|\to\infty} \left| \frac{1}{n} E\left(\log\left(\frac{1+|\xi_n|}{1+|\xi_0|} \right) \middle| \xi_0 = x \right) - \gamma \right| = 0,$$

which is (2.8).

(ii) This is similar to the proof of Lemma 4.1. For $\zeta < \kappa$, define M_ζ to be the matrix with positive elements

$$m_{\zeta ij} \stackrel{\text{def}}{=} E((W_1^*)^\zeta 1_{\theta_1^*=j} \mid \theta_0^* = i), \quad i, j \in \{-1, 1\}.$$

Let ρ_ζ be the maximal eigenvalue with corresponding *right* eigenvector ϕ, which we interpret as a function on $\{-1, 1\}$. Note that ϕ is nonnegative. We will demonstrate in the proof of Lemma 2.1 below that $\rho_\zeta < 1$. Define

$$V(x) = \begin{cases} 1 + \phi(x/|x|)|x|^\zeta, & x \neq 0, \\ 1, & x = 0. \end{cases}$$

This function satisfies (2.9).

Recall (4.2) and (4.3), which say in essence that

$$\limsup_{|x|\to\infty} P(\eta(i, e_1) \neq a(x, e_1)/|a(x, e_1)|) = 0.$$

We similarly have

$$\limsup_{|x|\to\infty} E(|b(1, e_1)|^\zeta 1_{\eta(i,e_1)\neq a(x,e_1)/|a(x,e_1)|}) = 0.$$

Thus,

$$\begin{aligned} &\limsup_{x\to\infty} E\left(\frac{V(\xi_1)}{V(\xi_0)} \middle| \xi_0 = x \right) \\ &\quad = \limsup_{x\to\infty} E\left(\frac{\phi(a(x, e_1)/|a(x, e_1)|)(|b(1, e_1)\, |x| + c(x, e_1)|^\zeta)}{\phi(1)|x|^\zeta} \right) \\ &\quad = \frac{\phi(-1)}{\phi(1)} E(|b(1, e_1)|^\zeta 1_{\eta(1,e_1)<0}) + E(|b(1, e_1)|^\zeta 1_{\eta(1,e_1)>0}) \\ &\quad = \frac{\phi(-1)}{\phi(1)} m_{\zeta,1,-1} + m_{\zeta,1,1} = \rho_\zeta < 1. \end{aligned} \tag{4.7}$$

Likewise,

$$\limsup_{x\to-\infty} E\left(\frac{V(\xi_1)}{V(\xi_0)} \middle| \xi_0 = x \right) = \rho_\zeta < 1. \tag{4.8}$$

Since $E(V(\xi_1) \mid \xi_0 = x)$ is locally bounded as a function of x, (4.7) and (4.8) suffice to prove (2.10) with $\rho \in (\rho_\zeta, 1)$.

(iii) From (2.8) and [11], $\gamma > 0$ implies $|\xi_t| \to \infty$ in probability and thus that the process is transient.

Suppose instead that $\gamma < 0$. By Assumption A.4, $\{\xi_t\}$ is irreducible with Lebesgue measure as a maximal irreducibility measure, is aperiodic and is a T-chain. Consequently, by Thm. 15.0.1 of [24], (2.9) and (2.10) are sufficient to ensure $\{\xi_t\}$ is geometrically ergodic. □

4.2. Showing regular variation

In this, the final and longest, subsection we assume $\{\xi_t\}$ is stationary with negative Lyapounov exponent γ and distribution Π. Here, we will verify the regular variation of its probability tails.

We start by proving Lemma 2.1.

Proof of Lemma 2.1. Since, for any $\kappa > 0$, all the elements of M_κ are positive, it has a nonnegative maximal eigenvalue ρ_κ and a unique *left* eigenvector μ_κ ($\mu'_\kappa M_\kappa = \rho_\kappa \mu'_\kappa$) such that μ_κ is a probability measure on $\{-1, 1\}$. We want to show that there is a unique $\kappa > 0$ such that $\rho_\kappa = 1$.

We first show that $\rho_\kappa^{1/\kappa}$ is strictly increasing in κ. To this end, let $\zeta > \kappa > 0$ and define the matrix M_ζ accordingly. Let ρ_ζ be the maximal eigenvalue for M_ζ. Define p_{ij} as in (2.1) and

$$d_{\kappa ij} = \frac{m_{\kappa ij}}{p_{ij}}, \qquad d_{\zeta ij} = \frac{m_{\zeta ij}}{p_{ij}}.$$

Thus, there exists $c < 1$ such that

$$\begin{aligned} d_{\kappa ij} &= E((W_1^*)^\kappa \mid \theta_0^* = i, \theta_1^* = j) \\ &< c(E((W_1^*)^\zeta \mid \theta_0^* = i, \theta_1^* = j))^{\kappa/\zeta} = c\left(d_{\zeta ij}\right)^{\kappa/\zeta}, \quad i, j \in \{-1, 1\}. \end{aligned}$$

It follows that, for any probability measure μ and vector $1 = \binom{1}{1}$,

$$\begin{aligned} \mu' M_\kappa^n 1 &= \sum_{\substack{i_l = \pm 1 \\ l = 0, \ldots, n}} \mu_{i_0} \prod_{l=1}^n p_{i_{l-1} i_l} \prod_{l=1}^n d_{\kappa i_{l-1} i_l} \\ &< c^n \sum_{\substack{i_l = \pm 1 \\ l = 0, \ldots, n}} \mu_{i_0} \prod_{l=1}^n p_{i_{l-1} i_l} \prod_{l=1}^n (d_{\zeta i_{l-1} i_l})^{\kappa/\zeta} \\ &< c^n \left(\sum_{\substack{i_l = \pm 1 \\ l = 0, \ldots, n}} \mu_{i_0} \prod_{l=1}^n p_{i_{l-1} i_l} \prod_{l=1}^n d_{\zeta i_{l-1} i_l} \right)^{\kappa/\zeta} \\ &= c^n (\mu' M_\zeta^n 1)^{\kappa/\zeta}. \end{aligned}$$

Hence,

$$\rho_\kappa = \lim_{n\to\infty} (\mu' M_\kappa^n 1)^{1/n} \le c \lim_{n\to\infty} ((\mu' M_\zeta^n 1)^{\kappa/\zeta})^{1/n} = c\left(\rho_\zeta\right)^{\kappa/\zeta},$$

showing the strict monotonicity as desired.

Now let $\tilde{E}$ be the matrix with elements

$$e_{ij} = \exp(E((\log W_1^*) 1_{\theta_1^* = j} \mid \theta_0^* = i)), \quad i, j \in \{-1, 1\}.$$

and note that

$$\lim_{\kappa \downarrow 0} d_{\kappa ij}^{1/\kappa} = \exp(E(\log W_1^* \mid \theta_0^* = i, \theta_1^* = j)).$$

So, by an argument similar to the above, $\lim_{\kappa\downarrow 0} \rho_\kappa^{1/\kappa}$ is the maximal eigenvalue of $\tilde{E}$. Since $\binom{\exp(\nu(-1))}{\exp(\nu(1))}$ is a nonnegative eigenvector for $\tilde{E}$ with corresponding eigenvalue e^γ, by (2.2), it must be that e^{γ} is the maximal eigenvalue of $\tilde{E}$. Therefore, since $\gamma < 0$, ρ_κ must be less than 1 for small enough κ. Also, ρ_κ clearly is continuous in κ and, by Assumptions A.1 and A.3, $\rho_\kappa > 1$ for large enough κ. From all this, it follows that there is a unique positive κ for which $\rho_\kappa = 1$. □

Lemma 4.2. *Let κ and μ_j, $j = \pm 1$, be the solution in Lemma 2.1 and set*

$$T_{ij}(w) = P(W_1^* \le w,\, \theta_1^* = j \mid \theta_0^* = i).$$

Suppose q_{-1}, q_1 are nonnegative measurable functions on $\mathbb{R}_+$ such that

$$\sup_{0<r\le 1} r^{\kappa+\delta} q_j(r) < \infty \quad \text{and} \quad \sup_{r\ge 1} r^{\kappa-\delta} q_j(r) < \infty, \qquad j = \pm 1,$$

for every $\delta > 0$. Suppose also that they solve the system of equations

$$q_j(r) = \sum_{i=\pm 1} \int_0^\infty q_i(r/w) T_{ij}(\mathrm{d}w). \tag{4.9}$$

and $q_{-1}(1) + q_1(1) = 1$. Then $q_j(r) = \mu_j r^{-\kappa}$.

Proof. Define $g_j(x) = e^{\kappa x} q_j(e^x)$. By Assumption A.3, each T_{ij} is absolutely continuous with density, say, t_{ij}. Define $\tau_{ij}(x) = e^{\kappa x} t_{ij}(e^x)$. Then (4.9) becomes

$$g_j(x) = \sum_{i=\pm 1} \int_{-\infty}^{\infty} g_i(y) \tau_{ij}(x - y)\, \mathrm{d}y, \qquad j = \pm 1, \tag{4.10}$$

namely, a linear system of integral equations with a convolution kernel, subject to $e^{-\delta|x|} g_j(x)$, is bounded, $j = \pm 1$. By Assumptions A.1 and A.2, we also deduce that $\int_{-\infty}^{\infty} e^{\zeta|x|} \tau_{ij}(x)\, \mathrm{d}x < \infty$ for all ζ, i, j. Expressing (4.10) more simply,

$$g_{-1} = g_{-1} * \tau_{-1,-1} + g_1 * \tau_{1,-1} \quad \text{and} \quad g_1 = g_{-1} * \tau_{-1,1} + g_1 * \tau_{1,1}.$$

We are thus justified in computing

$$\begin{aligned} g_1 * (1 - \tau_{-1,-1}) &= g_{-1} * (1 - \tau_{-1,-1}) * \tau_{-1,1} + g_1 * \tau_{1,1} * (1 - \tau_{-1,-1}) \\ &= g_1 * \tau_{1,-1} * \tau_{-1,1} + g_1 * \tau_{1,1} * (1 - \tau_{-1,-1}), \end{aligned} \tag{4.11}$$

or, equivalently, $g_1 = g_1 * \sigma$, where

$$\sigma = \tau_{-1,-1} + \tau_{1,1} - \tau_{-1,-1} * \tau_{1,1} + \tau_{-1,1} * \tau_{1,-1}.$$

Similarly, $g_{-1} = g_{-1} * \sigma$. Let $\hat{\tau} = \{\hat{\tau}_{ij}\}$ be the matrix of Fourier transforms for the τ_{ij}'s. So

$$\hat{\tau}_{ij}(\alpha) = E((W_1^*)^{\kappa + i\alpha} 1_{\theta_1^* = j} \mid \theta_0^* = i) \quad \text{and} \quad \hat{\sigma}(\alpha) = \det(I - \hat{\tau}(\alpha)). \tag{4.12}$$

From classical results (e.g., Sec. 11.2 of [28]), the solutions to (4.11) are linear combinations of $e^{i\alpha_k x} P_{jk}(x)$, where α_k is a root of $\hat{\sigma}(\alpha) = \det(I - \hat{\tau}(\alpha)) = 0$ in the strip $\mathcal{I}m(\alpha) < \delta$ and P_{jk} is a polynomial of degree one less than the multiplicity of α_k. Note that $\alpha = 0$ is root, by (4.12) and Lemma 2.1, and it has multiplicity 1 because $\hat{\tau}(0) = M_\kappa$ has a simple eigenvalue equal to 1. Also, δ may be chosen arbitrarily small. Hence, the only nonnegative solutions to (4.10) are constant functions which thus satisfy

$$g_j = \sum_{i=\pm 1} g_i \int_{-\infty}^{\infty} \tau_{ij}(y)\, \mathrm{d}y = \sum_{i=\pm 1} g_i m_{\kappa ij}.$$

By the conclusion of Lemma 2.1, and since $g_{-1} + g_1 = 1$, g_{-1} and g_1 must be equal to the elements of μ.

We conclude, then, that $q_j(r) = \mu_j r^{-\kappa}$ gives the unique nonnegative solution to (4.9) subject to $q_{-1}(1) + q_1(1) = 1$. □

The significance of the above result is in the next one, which essentially identifies the unique invariant measure for the *transient* process $\{\xi_t^*\}$ defined in (1.5). Observe that

$$\begin{aligned} &P(|\xi_t^*| > r, \xi_t^*/|\xi_t^*| = i \mid |\xi_{t-1}^*| = s, \xi_{t-1}^*/|\xi_{t-1}^*| = \theta) \\ &\quad = P(W_1^* > r/s, \theta_1^* = i \mid \theta_0^* = \theta) \\ &\quad = P(w(\theta, e_1) > r/s, \eta(\theta, e_1) = i). \end{aligned} \tag{4.13}$$

Corollary 4.3. *Let κ and μ_j, $j = \pm 1$, be the solution in Lemma 2.1. Suppose Q is a measure on $\mathbb{R}_+ \times \{-1, 1\}$ satisfying $Q((1, \infty) \times \{-1, 1\}) = 1$,*

$$\sup_{r \le 1} r^{\kappa+\delta} Q((r, \infty) \times \{i\}) < \infty \quad \textit{and} \quad \sup_{r \ge 1} r^{\kappa-\delta} Q((r, \infty) \times \{i\}) < \infty,$$

for every $\delta > 0$, and

$$\begin{aligned} Q((r, \infty) \times \{i\}) = \int_{\mathbb{R}_+ \times \{-1,1\}} & P(w(\theta, e_1) > r/s, \eta(\theta, e_1) = i) Q(\mathrm{d}s\mathrm{d}\theta), \\ & r > 0,\ i \in \{-1, 1\}. \end{aligned} \tag{4.14}$$

Then $Q((r, \infty) \times \{i\}) = \mu_i r^{-\kappa}$.

Proof. Let $q_i(r) = Q((r, \infty) \times \{i\})$. Then by (4.13) and a simple integration by parts, (4.14) is exactly the same as (4.9). □

We now turn to the tail behavior of the stationary distribution Π. It actually will be convenient to think of Π as the stationary distribution of $\{(R_t, \tilde{\theta}_t)\} = \{(|\xi_t|, \xi_t/|\xi_t|)\}$.

A helpful alternative to Definition 2.4 is given by the following result (cf. [4], Thm. 2.2.2, or [1]).

Theorem 4.4 (*Aljančić and Arandelović*). *Let $p(v)$ be a positive function on $(0, \infty)$.*

(i) *The upper Matuszewska index for p is the infimum of those α such that there exist finite K and v_0 with*

$$\frac{p(\lambda v)}{p(v)} \le K\lambda^\alpha, \quad \textit{for } \lambda v \ge v \ge v_0.$$

(ii) *The lower Matuszewska index for p is the supremum of those β such that there exist finite K and v_0 with*

$$\frac{p(\lambda v)}{p(v)} \le K\lambda^\beta, \quad \textit{for } v \ge \lambda v \ge v_0.$$

Lemma 4.5. *Suppose $\{\xi_t\}$ is stationary.*

(i) *Then $P(R_t > v)$ is of dominated variation: its Matuszewska indices are finite.*

(ii) *Let* $-\kappa_L$ *be the lower Matuszewska index for* $P(R_t > v)$. *Then for any* $\beta > \kappa_L$ *there exists* $K_1 < \infty$ *and* $v_0 < \infty$ *such that*

$$\frac{P(R_1 > \lambda v)}{P(R_1 > v)} \le K_1 \lambda^{-\beta}, \quad \text{for } v > \lambda v \ge v_0. \tag{4.15}$$

(iii) *Additionally,*

$$\lim_{v\to\infty} \sup_{x\in\mathbb{R}} \frac{P(|c(x, e_1)| > \epsilon v)}{P(R_1 > v)} = 0, \quad \text{for all } \epsilon > 0. \tag{4.16}$$

Proof. Recall Assumption A.2. We may assume $\tilde{b}_2 \ge \max(\tilde{b}_1/2, 2)$ without any loss. Note that

$$R_1 = |b(\tilde{\theta}_0, e_1)R_0 + c(\tilde{\theta}_0 R_0, e_1)| \ge (\tilde{b}_1|e_1| - \tilde{b}_2)R_0 - \bar{c}(1 + |e_1|).$$

Let $M_1 = 8\bar{c}/\tilde{b}_1$. Then $R_0 > M_1$ and $|e_1| \ge 2\tilde{b}_2/\tilde{b}_1 \ge 1$ imply

$$R_1 \ge (\tilde{b}_1|e_1| - \tilde{b}_2)R_0 - \bar{c}(1 + |e_1|) \ge R_0(\tilde{b}_1/2 - 2\bar{c}/M_1)|e_1| \ge R_0\tilde{b}_1|e_1|/(2\tilde{b}_2).$$

Let $0 < \delta < 1$. Given $R_1 \overset{\mathrm{D}}{=} R_0$ and $v > M_1/\delta$, we have

$$P(R_1 > v) \ge P(R_0 > \delta v, \tilde{b}_1|e_1| > 2\tilde{b}_2/\delta) = P(R_1 > \delta v)P(\tilde{b}_1|e_1| > 2\tilde{b}_2/\delta).$$

Hence,

$$\sup_{v>M_1/\delta} \frac{P(R_1 > \delta v)}{P(R_1 > v)} \le K_2 \overset{\text{def}}{=} \frac{1}{P(\tilde{b}_1|e_1| > 2\tilde{b}_2/\delta)} < \infty, \tag{4.17}$$

showing that R_1 has dominated varying probability tail (cf. [4]).

In particular, this means the probability tail has a finite (and nonpositive) lower Matuszewska index, say $-\kappa_L$. From Theorem 4.4(ii) we find that for each $\beta > \kappa_L$, (4.15) must hold with some finite K_1. In particular (take $\lambda v = v_0$ in (4.15)), $P(R_1 > v) > \delta_0 v^{-\beta}$ for some $\delta_0 > 0$. Taken with Assumption A.2 and the fact $E(|e_1|^\beta) < \infty$, this implies (4.16). □

Lemma 4.5 shows that the Matuszewska indices are finite. Lemma 4.7 below will show that they are in fact negative. Ultimately, they will turn out to be equal to each other.

Lemma 4.6. *Assume as in Lemma* 4.5. *For any* $r > 0$ *and* $\epsilon > 0$, *there exists* $\delta > 0$ *and* $M_2 < \infty$ *such that*

$$\frac{P(R_1 > rv, R_0 < \delta v)}{P(R_0 > v)} < \epsilon, \quad \text{for all } v > M_2.$$

Proof. Let F_1 be the distribution of $1 + |e_1|$. Suppose $0 < \delta \le 1$ and let $\beta > \kappa_L$. By Assumption A.1 and Lemma 4.5, we know

$$\lim_{v\to\infty} \frac{P(|e_1| > cv)}{P(R_1 > v)} = 0, \quad \text{for any } c > 0. \tag{4.18}$$

Choose v_0 as in (4.15) with $v_0 \ge \bar{c}/\bar{b}$. If $v > v_0/\delta$ then, using (4.15),

$$\begin{aligned}
&P(R_1 > rv, R_0 < \delta v)\\
&\quad \le P((\bar{b}R_0 + \bar{c})(1 + |e_1|) > rv, R_0 < \delta v)\\
&\quad \le P(\bar{b}R_0(1 + |e_1|) > rv/2, r/(2\delta\bar{b}) \le 1 + |e_1| \le rv/(2\bar{b}v_0))
\end{aligned}$$

$$
\begin{aligned}
&+ P(1+|e_1| > rv/(2\bar{b}v_0)) \\
&= \int_{r/(2\delta\bar{b})}^{rv/(2\bar{b}v_0)} P\left(R_0 > rv/(2\bar{b}u)\right) F_1(\mathrm{d}u) + P(1+|e_1| > rv/(2\bar{b}r_0)) \\
&\le K_1 \int_{r/(2\delta\bar{b})}^{rv/(2\bar{b}v_0)} \left(\frac{2\bar{b}u}{r}\right)^{\beta} F_1(\mathrm{d}u) P(R_0 > v) + P(1+|e_1| > rv/(2\bar{b}v_0)). \qquad (4.19)
\end{aligned}
$$

We may choose $\delta > 0$ to make $K_1 \int_{r/(2\delta\bar{b})}^{\infty} (\frac{2\bar{b}u}{r})^{\beta} F_1(\mathrm{d}u) < \epsilon/2$ and, by (4.18), we may choose $M_2 > v_0/\delta$ so that

$$\frac{P(1+|e_1| > rv/(2\bar{b}v_0))}{P(R_0 > v)} < \epsilon/2, \quad \text{for all } v > M_2.$$

Combining these with (4.19) gives the result. □

Lemma 4.7. *Suppose $\{\xi_t\}$ is stationary with (2.9) and (2.10) holding. Let the stationary distribution be Π.*

(i) *The upper Matuszewska index for $P(R_t > v)$ is no bigger than $-\zeta$.*
(ii) *For each $k \ge 0$, the measures $\bar{Q}_v^k$, $v \ge 1$, given by*

$$\bar{Q}_v^k((r,\infty)\times\{i\}) = \frac{\Pi((\max(r, 2^{-k})v, \infty)\times\{i\})}{\Pi((v,\infty)\times\{-1,1\})}, \quad r > 0,\ i \in \{-1,1\}, \qquad (4.20)$$

are tight on $\mathbb{R}_+ \times \{-1,1\}$.

Proof. First, suppose $k = 0$. Note that $\bar{Q}_v^0$ is the conditional distribution of $(R_t/v, \tilde{\theta}_t)$, given $R_t > v$, under stationarity.

By (2.9) and (2.10) and Assumption A.2,

$$V(\xi_1) \le d_2(|b(\tilde{\theta}_1, e_1)\xi_0| + |c(\xi_0, e_1)|)^{\zeta} + d_2 \le d_2(\bar{b}|\xi_0| + \bar{c})^{\zeta}(1+|e_1|)^{\zeta} + d_2.$$

This implies the existence of finite, positive d_3, d_4 such that

$$V(\xi_1) \le (d_3 V(\xi_0) + d_4)(1+|e_1|)^{\zeta} + d_2.$$

Let $K_3 = E((1+|e_1|)^{\zeta})$. Then, if $r > M_0$,

$$
\begin{aligned}
E(V(\xi_1)1_{V(\xi_1)>r}) &= E(V(\xi_1)1_{V(\xi_1)>r, V(\xi_0)>r}) + E(V(\xi_1)1_{V(\xi_1)>r, V(\xi_0)\le r}) \\
&\le E(E(V(\xi_1) \mid \xi_0)1_{V(\xi_0)>r}) \\
&\quad + E(((d_3 r + d_4)(1+|e_1|)^{\zeta} + d_2)1_{V(\xi_1)>r}) \\
&\le \rho E(V(\xi_0)1_{V(\xi_0)>r}) + ((d_3 r + d_4)K_3 + d_2)P(V(\xi_1) > r).
\end{aligned}
$$

Under stationarity, $V(\xi_1) \overset{\mathrm{D}}{=} V(\xi_0)$, and $E(V(\xi_0)) < \infty$ by Meyn and Tweedie [24, Thm. 14.0.1]. Hence

$$
\begin{aligned}
\frac{1}{r}E(V(\xi_1) \mid V(\xi_1) > r) &= \frac{E(V(\xi_1)1_{V(\xi_1)>r})}{rP(V(\xi_1) > r)} \\
&\le \rho \frac{E(V(\xi_1)1_{V(\xi_1)>r})}{rP(V(\xi_1) > r)} + d_3 K_3 + (d_4 K_3 + d_2)/M_0. \qquad (4.21)
\end{aligned}
$$

It follows from (4.21) that

$$\sup_{r>M_0} \frac{1}{r} E(V(\xi_1) \mid V(\xi_1) > r) \le K_4 \stackrel{\text{def}}{=} \frac{d_3 K_3 + (d_4 K_3 + d_2)/M_0}{1-\rho} < \infty. \tag{4.22}$$

Furthermore, we have $d_1 R_0^\zeta \le V(\xi_0) \le d_2(1 + R_0^\zeta)$. Thus,

$$E(R_0^\zeta 1_{R_0^\zeta > r}) \le \frac{1}{d_1} E(V(\xi_0) 1_{V(\xi_0) > d_1 r}) \tag{4.23}$$

and, if $d_1 r > 2d_2$,

$$P(R_0^\zeta > d_1 r/(2d_2)) \ge P(V(\xi_0) > d_1 r). \tag{4.24}$$

Let $\delta = (d_1/(2d_2))^{1/\zeta}$ and obtain M_1, K_2 from the proof of Lemma 4.5. Set

$$r_0 = \max(M_0, 2d_2, 2d_2 M_1^\zeta)/d_1.$$

Then, by (4.17) and (4.22)–(4.24), we obtain

$$\begin{aligned}
\sup_{v > r_0^{1/\zeta}} \int_1^\infty \int_{\{-1,1\}} s^\zeta \bar{Q}_v^0(\mathrm{d}s\mathrm{d}\theta) &= \sup_{r>r_0} \frac{1}{r} E(R_0^\zeta \mid R_0^\zeta > r) \\
&= \sup_{r>r_0} \frac{P(R_0^\zeta > d_1 r/(2d_2))}{P(R_0^\zeta > r)} \frac{E(R_0^\zeta 1_{R_0^\zeta > r})}{r P(R_0^\zeta > d_3 r/(2d_4))} \\
&\le K_2 \sup_{r>r_0} \frac{1}{d_1 r} E(V(\xi_1) \mid V(\xi_1) > d_1 r) \\
&\le K_2 K_4 < \infty.
\end{aligned} \tag{4.25}$$

This is sufficient for the probability measures $\{\bar{Q}_v^0\}_{v\ge1}$ to be tight on $\mathbb{R}_+ \times \{-1, 1\}$.

Indeed, from (4.25), we easily determine that

$$\frac{P(R_0 > \lambda v)}{P(R_0 > v)} \le K_2 K_4 \lambda^{-\zeta}, \quad \lambda v \ge v \ge r_0^{1/\zeta}.$$

Hence, the upper Matuszewska index is no more than $-\zeta$.

Let $\epsilon_1 > 0$. By the above we can choose $M_3 \in [1, \infty)$ so that $\bar{Q}_v^k((M_3, \infty) \times \{-1, 1\}) = \bar{Q}_v^0((M_3, \infty) \times \{-1, 1\}) < \epsilon_1$, for all $v \ge r_0^{1/\zeta}$, $k \ge 1$. This proves the tightness of $\{\bar{Q}_v^k\}_{v\ge1}$ on $\mathbb{R}_+ \times \{-1, 1\}$ for each k. □

In fact, assuming $\gamma < 0$, we may choose any $\zeta < \kappa$, by Theorem 2.2(ii). This implies that the upper Matuszewska index is no more than $-\kappa$, but is still some way from saying Π is regularly varying or even that the two indices are equal.

Next is the lemma that is at the heart of our proof. Recall the definition of Q_v in (2.11).

Lemma 4.8. *Assume as in Lemma 4.7. For any sequence $\tilde{v}_n \to \infty$, there exists a subsequence $v_n \to \infty$ and a continuous measure Q on $\mathbb{R}_+ \times \{-1, 1\}$ such that $Q_{v_n} \stackrel{\text{v}}{\to} Q$, $Q((1, \infty) \times \{-1, 1\}) = 1$ and*

$$Q((r, \infty) \times \{i\}) = \int_{\mathbb{R}_+ \times \{-1,1\}} P(w(\theta, e_1) > r/s, \eta(\theta, e_1) = i) Q(\mathrm{d}s\mathrm{d}\theta), \tag{4.26}$$

for $i \in \{-1, 1\}$, $r > 0$.

Proof. Let $\bar{Q}_v^k$, $v \geq 1$, be as in (4.20), namely the restriction of Q_v to $(2^{-k}, \infty) \times \{-1, 1\}$. Note that, by (4.17) for each k, the measures $\bar{Q}_v^k$ are uniformly bounded. By Lemma 4.7, the probability measures $\bar{Q}_v^0$ are tight on $\mathbb{R}_+ \times \{-1, 1\}$. Given any sequence $\tilde{v}_n \to \infty$, there exists a subsequence $v_n^0 \to \infty$ and a measure $\bar{Q}^0$ such that $\bar{Q}_{v_n^0}^0 \xrightarrow{\text{v}} \bar{Q}^0$ and, of course, $\bar{Q}^0((1, \infty) \times \{-1, 1\}) = 1$. Iteratively we may find a further subsequence $v_n^k \to \infty$ and a measure $\bar{Q}^k$ such that $\bar{Q}_{v_n^k}^k \xrightarrow{\text{v}} \bar{Q}^k$ and $\bar{Q}^k$ agrees with $\bar{Q}^{k-1}$ on $(2^{1-k}, \infty) \times \{-1, 1\}$. Letting $v_n = v_n^n$, we have $Q_{v_n} \xrightarrow{\text{v}} Q$ where Q is a measure that agrees with $\bar{Q}^k$ on $(2^{-k}, \infty) \times \{-1, 1\}$ for each k. At this point we do not know that Q is continuous.

Note that $w(\tilde{\theta}_0, e_1) R_0 > (1 + \epsilon/2) v$ implies either $R_1 > v$ or $|c(\tilde{\theta}_0 R_0, e_1)| > \epsilon v/2$. Thus,

$$\begin{aligned} & P(R_1 > v, \tilde{\theta}_1 = i \mid R_0 = s, \tilde{\theta}_0 = \theta) \\ & \quad \geq P(w(\theta, e_1) > (1 + \epsilon/2) v/s, \eta(\theta, e_1) = i) - P(|c(\theta s, e_1)| \geq \epsilon v/2). \end{aligned} \tag{4.27}$$

Using (4.27) and (4.16),

$$\begin{aligned} \liminf_{n \to \infty} Q_{v_n}((r, \infty) \times \{i\}) &= \liminf_{n \to \infty} \frac{P(R_1 > r v_n, \tilde{\theta}_1 = i)}{P(R_0 > v_n)} \\ &\geq \liminf_{n \to \infty} \int_{\mathbb{R}_+ \times \{-1,1\}} P(R_1 > r v_n, \tilde{\theta}_1 = i \mid R_0 = s, \tilde{\theta}_0 = \theta) \\ &\quad \times 1_{(2^{-k} v_n, \infty)}(s) \frac{\Pi(\mathrm{d}s \mathrm{d}\theta)}{\Pi((v_n, \infty) \times \{-1, 1\})} \\ &= \liminf_{n \to \infty} \int_{\mathbb{R}_+ \times \{-1,1\}} P(R_1 > r v_n, \tilde{\theta}_1 = i \mid R_0 = v_n s, \tilde{\theta}_0 = \theta) \\ &\quad \times \bar{Q}_{v_n}^k(\mathrm{d}s \mathrm{d}\theta) \\ &\geq \liminf_{n \to \infty} \int_{\mathbb{R}_+ \times \{-1,1\}} P(w(\theta, e_1) > (1 + \epsilon/2) r/s, \eta(\theta, e_1) = i) \\ &\quad \times \bar{Q}_{v_n}^k(\mathrm{d}s \mathrm{d}\theta). \end{aligned}$$

By Assumption A.3, $P(w(\theta, e_1) > \cdot, \eta(\theta, e_1) = \cdot)$ is continuous in $\mathbb{R}_+ \times \{-1, 1\}$. It follows by standard theory for vague convergence (e.g., [2, Thm. 4.5.1]) that

$$\begin{aligned} & \liminf_{n \to \infty} Q_{v_n}((r, \infty) \times \{i\}) \\ & \quad \geq \int_{\mathbb{R}_+ \times \{-1,1\}} P(w(\theta, e_1) > (1 + \epsilon/2) r/s, \eta(\theta, e_1) = i) \bar{Q}^k(\mathrm{d}s \mathrm{d}\theta). \end{aligned}$$

By monotone convergence, as $k \uparrow \infty$ and $\epsilon \downarrow 0$,

$$\liminf_{n \to \infty} Q_{v_n}((r, \infty) \times \{i\}) \geq \int_{\mathbb{R}_+ \times \{-1,1\}} P(w(\theta, e_1) > r/s, \eta(\theta, e_1) = i) Q(\mathrm{d}s \mathrm{d}\theta). \tag{4.28}$$

Fix $m \geq 0$. Let $\epsilon_1 > 0$ be chosen arbitrarily. By Lemma 4.6, with $r = 2^{-m}$, we may choose $\delta = 2^{-k}$ to make

$$\limsup_{n \to \infty} \frac{P(R_1 \geq 2^{-m} v_n, R_0 < 2^{-k} v_n)}{\Pi((v_n, \infty) \times \{-1, 1\})} < \epsilon_1.$$

Then, again using (4.16) and [2, Thm. 4.5.1],

$$\begin{aligned}
&\limsup_{n\to\infty} Q_{v_n}([2^{-m},\infty)\times\{-1,1\}) \\
&\quad\le \limsup_{n\to\infty}\int_{\mathbb{R}_+\times\{-1,1\}} P(R_1\ge 2^{-m}v_n \mid R_0=s,\tilde\theta_0=\theta)1_{[2^{-k}v_n,\infty)}(s) \\
&\qquad\times\frac{\Pi(\mathrm{d}s\mathrm{d}\theta)}{\Pi((v_n,\infty)\times\{-1,1\})} \\
&\qquad+\limsup_{n\to\infty}\frac{P(R_1\ge 2^{-m}v_n, R_0<2^{-k}v_n)}{\Pi((v_n,\infty)\times\{-1,1\})} \\
&\quad\le \limsup_{n\to\infty}\int_{\mathbb{R}_+\times\{-1,1\}} P(R_1\ge 2^{-m}v_n \mid R_0=v_ns,\tilde\theta_0=\theta)\bar{Q}^k_{v_n}(\mathrm{d}s\mathrm{d}\theta)+\epsilon_1 \\
&\quad\le \limsup_{n\to\infty}\int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)>(1-\epsilon/2)2^{-m}/s)\bar{Q}^k_{v_n}(\mathrm{d}s\mathrm{d}\theta)+\epsilon_1 \\
&\quad= \int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)>(1-\epsilon/2)2^{-m}/s)\bar{Q}^k(\mathrm{d}s\mathrm{d}\theta)+\epsilon_1.
\end{aligned}$$

Dominated convergence as $\epsilon\downarrow 0$ and monotone convergence as $k\uparrow\infty$ yields

$$\limsup_{n\to\infty} Q_{v_n}([2^{-m},\infty)\times\{-1,1\})\le\int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)\ge 2^{-m}/s)Q(\mathrm{d}s\mathrm{d}\theta)+\epsilon_1.$$

The arbitrary choice of ϵ_1 finally gives

$$\limsup_{n\to\infty} Q_{v_n}([2^{-m},\infty)\times\{-1,1\})\le\int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)\ge 2^{-m}/s)Q(\mathrm{d}s\mathrm{d}\theta). \tag{4.29}$$

Since $P(w(\theta,e_1)=2^{-m}/s)=0$ for all $s>0$, (4.29) combined with (4.28) for $r=2^{-m}$ implies

$$\lim_{n\to\infty} Q_{v_n}([2^{-m},\infty)\times\{-1,1\})=\int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)\ge 2^{-m}/s)Q(\mathrm{d}s\mathrm{d}\theta). \tag{4.30}$$

From (4.28) and (4.30), we can now conclude that the measures $\bar{Q}^m_{v_n}$ converge vaguely, once again by Ash [2, Thm. 4.5.1]. Therefore, in fact $Q_{v_n}\xrightarrow{\mathrm{v}}\tilde{Q}$, where $\tilde{Q}$ is continuous and defined by

$$\tilde{Q}((r,\infty)\times\{i\})=\int_{\mathbb{R}_+\times\{-1,1\}} P(w(\theta,e_1)>r/s,\eta(\theta,e_1)=i)Q(\mathrm{d}s\mathrm{d}\theta).$$

Finally, since also $Q_{v_n}\xrightarrow{\mathrm{v}} Q$, it must be that $\tilde{Q}=Q$ and (4.26) holds. $\square$

Lemma 4.9. *Let Q be a vague subsequential limit of Q_v, as in Lemma 4.8. Then*

$$\textit{either}\quad \inf_{r>0}\frac{Q((r,\infty)\times\{-1\})}{Q((r,\infty)\times\{-1,1\})}>0\quad\textit{or}\quad \inf_{r>0}\frac{Q((r,\infty)\times\{1\})}{Q((r,\infty)\times\{-1,1\})}>0. \tag{4.31}$$

Proof. Let Δ_{-1} and Δ_1 be as in Assumption A.3 and assume $\Delta_1>0$. Choose r_0 such that $P(b(\theta,e_1)>r)\ge\frac{\Delta_1}{2}P(|b(\theta,e_1)|>r)$ for all $r\ge r_0$. Let $\delta=\min(\frac{\Delta_1}{2},P(b(\theta,e_1)>r_0))$. Then

$$P(b(\theta,e_1)>r/s)\ge\delta P(|b(\theta,e_1)|>r/s),\quad\text{for all } r>0,\ s\le r/r_0,$$

and

$$P(b(\theta, e_1) > r/s) \geq P(b(\theta, e_1) > r_0) \geq \delta P(|b(\theta, e_1)| > r/s), \quad \text{for all } r > 0,\ s \geq r/r_0.$$

Note that $b(\theta, e_1) > r$ iff $w(\theta, e_1) > r$ and $\eta(\theta, e_1) = 1$. Thus, from (4.26),

$$\begin{aligned} Q((r, \infty) \times \{1\}) &= \int_{\mathbb{R}_+ \times \{-1,1\}} P(b(\theta, e_1) > r/s) Q(\mathrm{d}s\mathrm{d}\theta) \\ &\geq \delta \int_{\mathbb{R}_+ \times \{-1,1\}} P(|b(\theta, e_1)| > r/s) Q(\mathrm{d}s\mathrm{d}\theta) = \delta Q((r, \infty) \times \{-1, 1\}). \end{aligned}$$

This shows that the second inequality in (4.31) holds if $\Delta_1 > 0$. A similar argument applies to show that the first inequality in (4.31) holds if $\Delta_{-1} > 0$. □

Finally, we are ready to prove our principal result.

Proof of Theorem 2.3. Let $-\kappa_L \leq -\kappa_U$ be the lower and upper Matuszewska indices, respectively, for the function $p(r) = P(R_t > r)$ under stationarity. From Lemmas 4.5 and 4.7 we know they are finite and negative. Before we can proceed further, we need to show that these indices are both equal to κ. This will require several steps. First, let Q_{v_n} be a sequence converging vaguely to Q, as in Lemma 4.8. Let $\alpha > -\kappa_U$. By Theorem 4.4,

$$\frac{Q((\lambda r, \infty) \times \{-1, 1\})}{Q((r, \infty) \times \{-1, 1\})} = \lim_{n\to\infty} \frac{P(R_t > \lambda r v_n)}{P(R_t > r v_n)} \leq K\lambda^\alpha, \tag{4.32}$$

for some finite K and all $\lambda \geq 1$. Consequently, the upper Matuszewska index for $q(r) = Q((r, \infty) \times \{-1, 1\})$ is no more than $-\kappa_U$. Likewise, the lower Matuszewska index for $q(r)$ is no less than $-\kappa_L$. This is true for *any* vague subsequential limit Q.

Next, we apply the Pólya peak theorem (Thm. 2.5.2 in [4], from [16]): there exists a sequence $\tilde{v}_n \to \infty$ such that

$$\limsup_{n\to\infty} Q_{\tilde{v}_n}((r, \infty) \times \{-1, 1\}) = \limsup_{n\to\infty} \frac{P(R_t > r\tilde{v}_n)}{P(R_t > \tilde{v}_n)} \leq r^{-\kappa_L}, \quad \text{for all } r > 0. \tag{4.33}$$

From Lemma 4.8, there is a subsequence $v_n \to \infty$ and a continuous measure Q such that $Q_{v_n} \xrightarrow{\mathrm{v}} Q$. Define the upper and lower orders for $q(r) = Q((r, \infty) \times \{-1, 1\})$ by

$$\omega_L = \liminf_{r\to\infty} \frac{\log q(r)}{\log r} \quad \text{and} \quad \omega_U = \limsup_{r\to\infty} \frac{\log q(r)}{\log r}.$$

Hence, $\omega_L \leq \omega_U \leq -\kappa_L$, by (4.33). From Prop. 2.2.5 of [4] and our comment above about the lower Matuszewska index for $q(r)$, $\omega_L \geq -\kappa_L$ also. Thus, $\omega_L = -\kappa_L$. Therefore, by Thm. 2.3.11 (note the misprint) of [4] and the continuity of Q, there exists a function $g(r)$ which is regularly varying on $\mathbb{R}_+$ with index $-\kappa_L$ such that $\liminf_{r\to\infty} Q((r, \infty) \times \{-1, 1\})/g(r) = 1$. Define

$$q_i = \liminf_{r\to\infty} \frac{Q((r, \infty) \times \{i\})}{g(r)}, \quad i = \pm 1. \tag{4.34}$$

Both q_{-1} and q_1 are finite and at least one is positive by Lemma 4.9 (critical points that have compelled us to this intricate, nested argument of indices and orders).

Recall that Q satisfies (4.26). Letting $T_{ij}(w) = P(W_1^* \le w,\ \theta_1^* = j \mid \theta_0^* = i)$, (4.26) may be reexpressed as

$$Q((r,\infty)\times\{j\}) = \sum_{i=\pm1}\int_0^\infty Q((r/w,\infty)\times\{i\})\,T_{ij}(\mathrm{d}w). \tag{4.35}$$

From (4.34) and (4.35) and the regular variation of $g(r)$ we get

$$\begin{aligned} q_j &\ge \sum_{i=\pm1}\int_0^\infty \liminf_{r\to\infty}\frac{Q((r/w,\infty)\times\{i\})}{g(r/w)}\frac{g(r/w)}{g(r)}\,T_{ij}(\mathrm{d}w) \\ &= \sum_{i=\pm1}\int_0^\infty q_i w^{\kappa_L} T_{ij}(\mathrm{d}w) = \sum_{i=\pm1} q_i m_{\kappa_L ij}, \end{aligned}$$

where $m_{\kappa_L ij}$ is defined by (2.5). This means the maximal eigenvalue of the matrix $M_{\kappa_L} = [m_{\kappa_L ij}]_{ij}$ is no larger than 1. That, in turn, implies $\kappa_L \le \kappa$ by the proof of Lemma 2.1. A similar argument (or the comment following the proof of Lemma 4.7) confirms $\kappa_U \ge \kappa$. But $\kappa_L \ge \kappa_U$. Therefore, $\kappa_L = \kappa_U = \kappa$.

The point to all this is that we may now claim that for all $\delta > 0$ (see (4.32) with $r = 1$ and $\alpha = -\kappa + \delta$),

$$\sup_{\lambda\ge1}\lambda^{\kappa-\delta}Q((\lambda,\infty)\times\{j\}) < \infty$$

for any vague subsequential limit Q. Likewise, Theorem 4.4 also yields

$$\sup_{\lambda\le1}\lambda^{\kappa+\delta}Q((\lambda,\infty)\times\{j\}) < \infty,$$

for all $\delta > 0$. We therefore have the conditions of Corollary 4.3 fulfilled so that the unique solution to (4.35) (and thus to (4.26)), subject to $Q((1,\infty)\times\{-1,1\}) = 1$, is given by $Q((r,\infty)\times\{i\}) = \mu_i r^{-\kappa}$. Since this true for any vague subsequential limit, we conclude that $Q_v \xrightarrow{\mathrm{v}} Q$, and therefore Π has regularly varying tails. $\square$

References

[1] S. Aljančić, D. Arandelović, O-regularly varying functions, Publications de l'Institut Mathématique, Belgrade (nouvelle série) 22 (36) (1977) 5–22.

[2] R.B. Ash, Real Analysis and Probability, Academic Press, 1972.

[3] B. Basrak, R. Davis, T. Mikosch, Regular variation of GARCH processes, Stochastic Process Appl. 99 (2002) 95–115.

[4] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, 1987.

[5] M. Borkovec, Extremal behavior of the autoregressive process with ARCH(1) errors, Stochastic Process Appl. 85 (2000) 189–207.

[6] M. Borkovec, Asymptotic behavior of the sample autocovariance and autocorrelation function of the AR(1) process with ARCH(1) errors, Bernoulli 7 (6) (2001) 847–872.

[7] M. Borkovec, C. Klüppelberg, The tail of the stationary distribution of an autoregressive process with ARCH(1) errors, Ann. Appl. Probab. 11 (2001) 1220–1241.

[8] P. Bougerol, N. Picard, Strict stationarity of generalized autoregressive processes, Ann. Probab. 20 (1992) 1714–1730.

[9] P. Bougerol, N. Picard, Stationarity of GARCH processes and some nonnegative time series, J. Econom. 52 (1992) 115–127.

[10] D.B.H. Cline, H.H. Pu, Verifying irreducibility and continuity of a nonlinear time series, Statist. Probab. Lett. 40 (1998) 139–148.

[11] D.B.H. Cline, H.H. Pu, Geometric transience of nonlinear time series, Statist. Sinica 11 (2001) 273–287.

[12] D.B.H. Cline, H.H. Pu, A note on a simple Markov bilinear stochastic process, Statist. Probab. Lett. 56 (2002) 283–288.
[13] D.B.H. Cline, H.H. Pu, Stability and the Lyaponouv exponent of threshold AR-ARCH models, Ann. Appl. Probab. 14 (2004) 1920–1949.
[14] R.A. Davis, T. Mikosch, The sample autocorrelations of heavy-tailed processes with applications to ARCH, Ann. Statist. 26 (1998) 2049–2080.
[15] J. Diebolt, D. Guégan, Tail behaviour of the stationary density of general non-linear autoregressive processes of order 1, J. Appl. Probab. 30 (1993) 315–329.
[16] D. Drasin, D.F. Shea, Pólya peaks and the oscillation of positive functions, Proc. Amer. Math. Soc. 34 (1972) 403–411.
[17] C.M. Goldie, Implicit renewal theory and tails of solutions of random equations, Ann. Appl. Probab. 1 (1991) 126–166.
[18] A.K. Grincevičius, A random difference equation, Lithuanian Math. J. 20 (1981) 279–282.
[19] D. Guégan, J. Diebolt, Probabilistic properties of the β-ARCH model, Statist. Sinica 4 (1994) 71–87.
[20] H. Hult, F. Lindskog, Extremal behavior of regularly varying stochastic processes, Stochastic Process Appl. 115 (2005) 249–274.
[21] H. Hult, F. Lindskog, T. Mikosch, G. Samorodnitsky, Functional large deviations for multivariate regularly varying random walks, Ann. Appl. Probab. 15 (2006) 2651–2680.
[22] H. Kesten, Random difference equations and renewal theory for products of random matrices, Acta Math. 131 (1973) 207–248.
[23] C. Klüppelberg, P. Pergamenchtchikov, The tail of the stationary distribution of a random coefficient AR(q) model, Ann. Appl. Probab. 14 (2004) 971–1005.
[24] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Springer-Verlag, London, 1993.
[25] T. Mikosch, C. Stărică, Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process, Ann. Statist. 28 (2000) 1427–1451.
[26] S.I. Resnick, Extreme Values, Regular Variation and Point Processes, Springer-Verlag, 1987.
[27] B. de Saporta, Tail of the stationary solution of the stochastic equation $Y_{n+1} = a_n Y_n + b_n$ with Markovian coefficients, Stochastic Process Appl. 115 (2005) 1954–1978.
[28] E.C. Titchmarsh, Introduction to the Theory of Fourier Integrals, 2nd ed., Oxford University Press, 1948.
[29] W. Verwaat, On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, Adv. Appl. Prob. 11 (1979) 750–783.
[30] Z. Zhang, H. Tong, A note on stochastic difference equations and its application to GARCH models, Chinese J. Appl. Probab. Statist. 20 (2004) 259–269.