Multivariate Markov-switching ARMA processes with regularly varying noise

Robert Stelzer*

23rd June 2006

Abstract

The tail behaviour of stationary $\mathbb{R}^d$-valued Markov-Switching ARMA processes driven by a regularly varying noise is analysed. It is shown that under appropriate summability conditions the MS-ARMA process is again regularly varying as a sequence. Moreover, the feasible stationarity condition given in Stelzer (2006) is extended to a criterion for regular variation. Our results complement in particular those of Saporta (2005) where regularly varying tails of one-dimensional MS-AR(1) processes coming from consecutive large parameters were studied.

AMS Subject Classification 2000: 60G10, 60G70, 60J10

Keywords:
heavy tails, regular variation, non-linear time series models, stochastic difference equation

1 Introduction

Markov-switching ARMA (MS-ARMA) processes are a modification of the well-known ARMA processes by allowing for time-dependent ARMA coefficients, which are modelled as a Markov chain. These processes are particularly popular in econometric modelling since the seminal paper by Hamilton (1989). In this paper we study the tail behaviour of multivariate MS-ARMA processes which are driven by a regularly varying i.i.d. noise sequence. In our analysis we allow the driving parameter chain to have a general state space as in Stelzer (2006), instead

*Graduate Programme Applied Algorithmic Mathematics, Centre for Mathematical Sciences, Munich University of Technology, Boltzmannstraße 3, D-85747 Garching, Germany. *Email:* `rstelzer@ma.tum.de`, `www.ma.tum.de/stat`

of assuming only finitely many regimes as usual (see e.g. Francq & Zakoïan (2001) or Krolzig (1997)).

Under appropriate summability conditions on the coefficients, we establish that the MS-ARMA process is (multivariate) regularly varying with the same index of regular variation as the driving noise sequence. Moreover, the spectral measure of regular variation is determined by the spectral measure of the noise. Extending a result of Stelzer (2006) we see that the summability conditions are satisfied (for all indices of regular variation), if in almost all regimes the sum of some norm of the autoregressive coefficients is strictly less than one.

Recently Saporta (2005) studied one-dimensional MS-AR(1) processes with finitely many regimes and obtained that the possible appearance of consecutive large AR(1) coefficients (explosive regimes) implies that the tail of the stationary distribution follows a power law under some technical conditions. For random coefficient autoregressive processes (i.e. the AR coefficients are i.i.d.) similar results are given in Kesten (1973) and Klüppelberg & Pergamenchtchikov (2004).

The paper is organized as follows. In Section 2 we briefly recall the details of the MS-ARMA model and in Section 3 the details of multivariate regular variation. Thereafter, we analyse MS-ARMA processes with a regularly varying noise in Section 4 and conclude with some examples in Section 5.

2 Markov-switching ARMA processes

Below (stationary) multivariate Markov-switching ARMA processes are briefly reviewed. For more details we refer to Stelzer (2006).

In defining MS-ARMA processes, one starts from a (multivariate) ARMA equation (see e.g. Brockwell & Davis (1991)) with drift and allows for random coefficients which are modelled as a Markov chain. We denote the real $d \times d$ ($m \times n$) matrices by $M_d(\mathbb{R})$ ($M_{m,n}(\mathbb{R})$).

Definition 2.1 (MS-ARMA(p,q) process)**.** *Let $p, q \in \mathbb{N}_0$, $p + q \geq 1$ and $\Delta = (\Sigma_t, \Phi_{1t}, \ldots, \Phi_{pt}, \Theta_{1t}, \ldots, \Theta_{qt})_{t\in\mathbb{Z}}$ be a stationary and ergodic Markov chain with some (measurable) subset S of $M_d(\mathbb{R})^{1+p+q}$ as state space. Moreover, let $\epsilon = (\epsilon_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence of $\mathbb{R}^d$-valued random variables independent of Δ and set $Z_t := \Sigma_t \epsilon_t \in \mathbb{R}^d$. A stationary process $(X_t)_{t\in\mathbb{Z}}$ in $\mathbb{R}^d$ is called* MS-ARMA(p, q, Δ, ϵ) process*, if it satisfies*

$$X_t - \Phi_{1t}X_{t-1} - \cdots - \Phi_{pt}X_{t-p} = Z_t + \Theta_{1t}Z_{t-1} + \cdots + \Theta_{qt}Z_{t-q} \tag{2.1}$$

for all $t \in \mathbb{Z}$. (2.1) is referred to as the MS-ARMA(p, q, Δ, ϵ) equation.

Furthermore, a stationary process $(X_t)_{t\in\mathbb{Z}}$ is said to be an MS-ARMA(p,q) process, if it is an MS-ARMA(p, q, Δ, ϵ) *process for some Δ and ϵ satisfying the above conditions.*

The elements of S are called "regimes", and "ergodic" is to be understood in its general measure theoretic meaning.

Remark 2.2. Compared to Stelzer (2006) we do not include an intercept (mean) μ_t in the parameter chain Δ and the defining equation (2.1), as this makes the following results notationally easier. Note, however, that the results of this paper can be immediately applied to the case with a general μ_t under an appropriate condition ensuring relative light-tailedness of $\sum_{k=0}^{\infty} \mathbf{A}_0\mathbf{A}_{-1}\cdots\mathbf{A}_{-k+1}\mathbf{m}_{-k}$ using Basrak (2000, Remark 2.1.20) (see also Mikosch (1999, Remarks 1.3.5, 1.5.11)) with $\mathbf{m}_t := (\mu^\intercal, 0^\intercal, \ldots, 0^\intercal)^\intercal \in \mathbb{R}^{d(p+q)}$. □

Given some i.i.d. noise (ϵ_t) and parameter chain (Δ_t), the natural question arising is, whether there exists a stationary (always understood in the strict sense) solution to (2.1). Below, the zeros appearing denote zeros in $M_{m,n}(\mathbb{R})$ or $\mathbb{R}^d$ with the appropriate dimensions m, n and d being obvious from the context.

Proposition 2.3 (State Space Representation, Stelzer (2006, Prop. 2.3))**.** *Define*

$$\begin{aligned}
\mathbf{X}_t &= (X_t^\intercal, X_{t-1}^\intercal, \ldots, X_{t-p+1}^\intercal, Z_t^\intercal, \ldots, Z_{t-q+1}^\intercal)^\intercal \in \mathbb{R}^{d(p+q)}, \qquad (2.2)\\
\mathbf{\Sigma}_t &= (\Sigma_t^\intercal, \underbrace{0^\intercal, \ldots, 0^\intercal}_{p-1}, \Sigma_t^\intercal, \underbrace{0^\intercal, \ldots, 0^\intercal}_{q-1})^\intercal \in M_{d(p+q),d}(\mathbb{R}), \ \ \mathbf{C}_t = \mathbf{\Sigma}_t\epsilon_t,\\
\mathbf{\Phi}_t &= \begin{pmatrix} \Phi_{1t} & \cdots & \Phi_{(p-1)t} & \Phi_{pt} \\ I_d & 0\cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & \cdots 0 & I_d & 0 \end{pmatrix} \in M_{dp}(\mathbb{R}), \ \ \mathbf{J} = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ I_d & 0 & \cdots & 0 \\ 0 & \ddots & 0\cdots & \vdots \\ 0 & \cdots 0 & I_d & 0 \end{pmatrix} \in M_{dq}(\mathbb{R}),\\
\mathbf{\Theta}_t &= \begin{pmatrix} \Theta_{1t} & \cdots & \Theta_{(q-1)t} & \Theta_{qt} \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \in M_{dp,dq}(\mathbb{R}),
\end{aligned}$$

$$\mathbf{A}_t = \begin{pmatrix} \mathbf{\Phi}_t & \mathbf{\Theta}_t \\ 0 & \mathbf{J} \end{pmatrix} \in M_{d(p+q)}(\mathbb{R}). \qquad (2.3)$$

Then (2.1) has a stationary and ergodic solution, if and only if

$$\mathbf{X}_t = \mathbf{A}_t\mathbf{X}_{t-1} + \mathbf{C}_t \qquad (2.4)$$

has one.

The process $\mathbf{X}$ as defined above is called the state space representation of the MS-ARMA process.

In order to avoid degeneracies in the state space representation, we presume without loss of generality $p \geq 1$ from now on. Moreover, in the case of a purely autoregressive MS-ARMA equation, i.e. $q = 0$, it is implicitly understood that $\mathbf{J}_t$ and $\mathbf{\Theta}_t$ vanish, $\mathbf{X}_t = \left(X_t^\intercal, X_{t-1}^\intercal, \ldots, X_{t-p+1}^\intercal\right)^\intercal$, $\mathbf{\Sigma}_t = \left(\Sigma_t^\intercal, 0^\intercal, \ldots, 0^\intercal\right)^\intercal$ and $\mathbf{A}_t = \mathbf{\Phi}_t$.

Regarding notation, $\|\cdot\|$ shall denote any norm on $\mathbb{R}^{d(p+q)}$ as well as the induced operator norm and $\xrightarrow{\mathscr{D}}$ convergence in distribution. If $k = 0$, the product $\mathbf{A}_t\mathbf{A}_{t-1}\cdots\mathbf{A}_{t-k+1}$ below is understood to be identical to the identity $I_{d(p+q)}$ on $\mathbb{R}^{d(p+q)}$, a convention to be used throughout for products of this structure.

Theorem 2.4 (Stelzer (2006, Th. 2.5 a)))**.** *Equation (2.4) and the MS-ARMA(p, q, Δ, ϵ) equation (2.1) have a unique stationary and ergodic solution, if $E(\log^+\|\mathbf{A}_0\|)$ and $E(\log^+\|\mathbf{C}_0\|)$ are finite and the Lyapunov exponent $\gamma := \inf_{t\in\mathbb{N}_0}\left(\frac{1}{t+1}E\left(\log\|\mathbf{A}_0\mathbf{A}_{-1}\cdots\mathbf{A}_{-t}\|\right)\right)$ is strictly negative. The unique stationary solution $\mathbf{X} = (\mathbf{X}_t)_{t\in\mathbb{Z}}$ of (2.4) is given by*

$$\mathbf{X}_t = \sum_{k=0}^{\infty}\mathbf{A}_t\mathbf{A}_{t-1}\cdots\mathbf{A}_{t-k+1}\mathbf{C}_{t-k} \tag{2.5}$$

and this series converges absolutely a.s.

3 Multivariate regular variation

As we are dealing with processes in $\mathbb{R}^d$ and also shall consider the state space representation of an MS-ARMA process, we recall first some results on multivariate regular variation in this section. Comprehensive references on this topic are Resnick (1987, Section 5.4.2; 2004), Mikosch (2003), and for univariate regular variation Bingham, Goldie & Teugels (1989).

Let $\|\cdot\|$ denote an arbitrary, fixed norm on $\mathbb{R}^d$ and $\mathbb{S}^{d-1}$ the unit sphere in $\mathbb{R}^d$, i.e. $\mathbb{S}^{d-1} = \partial B_1(0)$, with respect to this norm $\|\cdot\|$. Moreover, let $\xrightarrow{v}$ denote vague convergence, $M_+(E)$ the set of Radon measures over some space E, $\mathcal{B}(E)$ the Borel sets over E and $\mathscr{B}_\mu$ the μ-boundaryless sets for some measure μ, i.e. all sets B with $\mu(\partial B) = 0$. We define multivariate regular variation as follows.

Definition 3.1 (Regular variation on $\mathbb{R}^d$)**. a)** *Let X be an $\mathbb{R}^d$-valued random variable. If there exists an $\mathbb{S}^{d-1}$-valued random variable θ such that for some $\alpha > 0$ and every $u > 0$*

$$\frac{P\left(\|X\| > tu, \frac{X}{\|X\|} \in \cdot\right)}{P(\|X\| > t)} \xrightarrow{v} u^{-\alpha}P(\theta \in \cdot)$$

in $M_+(\mathbb{S}^{d-1})$ for $t \to \infty$, then X is said to be (multivariate) regularly varying *and we write $X \in \mathcal{R}_\alpha$.*

The parameter α is called the index of regular variation *and $P(\theta \in \cdot) \in M_+(\mathbb{S}^{d-1})$ the* spectral measure of regular variation *of X*

b) *A random sequence $(X_n)_{n\in\mathbb{Z}}$ in $\mathbb{R}^d$ is called* regularly varying (as a sequence)*, if all its finite dimensional distributions are regularly varying.*

For the necessary background on vague convergence of Radon measures on locally compact Polish spaces see, for example, Resnick (1987) or Bauer (1992).

Several equivalent definitions for multivariate regular variation exist, confer e.g. Basrak (2000), Lindskog (2004) or Resnick (2004) for detailed discussions. We employ the following characterization, as it makes transformations straightforward.

Theorem 3.2. *Let X be an $\mathbb{R}^d$-valued random variable. Then the following are equivalent:*

(i) *X is regularly varying.*

(ii) *There exists a positive sequence $(a_n)_{n\in\mathbb{N}}$, $a_n \to \infty$ as $n \to \infty$, and a non-zero $\nu_X \in M_+\left(\overline{\mathbb{R}^d}\backslash\{0\}\right)$ with $\nu_X\left(\overline{\mathbb{R}^d}\backslash\mathbb{R}^d\right) = 0$ such that*

$$nP(X \in a_n\cdot) \xrightarrow{v} \nu_X(\cdot)$$

in $M_+\left(\overline{\mathbb{R}^d}\backslash\{0\}\right)$ for $n \to \infty$.

If (ii) holds, then there exists an $\alpha > 0$ such that $\nu_X(tA) = t^{-\alpha}\nu_X(A)$ for all Borel sets A and $\partial B_\delta(0) \in \mathscr{B}_{\nu_X}$ for all $\delta > 0$. In particular, ν_X has no atoms.

ν_X is referred to as the measure of regular variation *of X.*

Remark 3.3. a) (*One point uncompactification*) $\overline{\mathbb{R}^d}\backslash\{0\}$ is called the one point uncompactification of $\mathbb{R}^d$. For $d = 1$ this is obtained as follows: Take the space $\mathbb{R}$ with the usual topology and form the two point compactification by setting $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ and adding the neighbourhoods of $\pm\infty$, i.e. the sets $[-\infty, a)$ and $(a, \infty]$ with $a \in \mathbb{R}$, to the basic open sets. Then take $\overline{\mathbb{R}}\backslash\{0\}$ and remove the open neighbourhoods of 0 from the topology. For the d-dimensional case one takes the compactification $\overline{\mathbb{R}^d}$, which is simply the d-fold product of $\overline{\mathbb{R}}$, with the product topology. Then one removes the point 0 from $\overline{\mathbb{R}^d}$ and the open neighbourhoods of 0 from the topology.

One can interpret this procedure as interchanging the roles of zero and infinity. In $\overline{\mathbb{R}^d}\backslash\{0\}$ compact sets can by characterized by being closed (in the usual sense) and bounded away from zero. By this procedure we obtain a locally compact Polish space, a possible metric on $\overline{\mathbb{R}}\backslash\{0\}$ is given by $d(x, y) := |x^{-1} - y^{-1}|$ (cf. Resnick (1987, p. 225f)). For the construction of a possible metric on $\overline{\mathbb{R}^d}\backslash\{0\}$ see Lindskog (2004, Theorem 1.5), for instance.

b) (*ii*) is norm-free and thus it does not matter, which norm is used in the definition. Therefore the results of this paper do not depend on the particular norm. However, the spectral measure is different for different norms, see also Hult & Lindskog (2002).

c) ν_X is non-degenerate, if and only if $\nu_X((a, \infty]\mathbb{S}^{d-1}) > 0$ for one and hence all $a > 0$ (note $(a, \infty]S^{d-1} := \{xz : x \in (0, \infty], z \in \mathbb{S}^{d-1}\}$).

Moreover, we need a precise notion of L^r-spaces of multivariate random variables.

Definition 3.4. *Denote by $L^r_{\mathbb{R}}$ with $r \in (0, \infty]$ the usual space of r-times integrable real-valued random variables and be $\mathbb{R}^d$ (or $M_d(\mathbb{R})$) equipped with a norm $\|\cdot\|$. Then $L^r_{\mathbb{R}^d}$ (or $L^r_{M_d(\mathbb{R})}$) is defined as the space of all $\mathbb{R}^d$- (or $M_d(\mathbb{R})$-) valued random variables X with $\|X\| \in L^r_{\mathbb{R}}$. For short we often omit the space subscript and write L^r.*

Moreover, $\|\cdot\|_{L^r} : L^r \to \mathbb{R}^+_0$, $X \mapsto E(\|X\|^r)^{1/r}$ defines (up to a.s. identity) a norm on L^r for $r \geq 1$ and $d_{L^r}(\cdot,\cdot) : L^r \times L^r \to \mathbb{R}^+_0$, $(X, Y) \mapsto E(\|X - Y\|^r)$ a metric on L^r for $0 < r < 1$.

The L^r spaces are independent of the norm $\|\cdot\|$ used on $\mathbb{R}^d$ (or $M_d(\mathbb{R})$), when viewed solely as sets. This is, however, not true for the norms $\|\cdot\|_{L^r}$ and metrics $d_{L^r}(\cdot,\cdot)$. Yet, due to the equivalence of all norms on $\mathbb{R}^d$ (or $M_d(\mathbb{R})$) it is immediate to see that for different norms $\|\cdot\|$ the induced norms and metrics on L^r are equivalent. All results from the well-known theory of the $L^r_{\mathbb{R}}$ spaces extend immediately to the multidimensional L^r spaces (see e.g. the overview in Stelzer (2005, Section 2.4)).

Remark 3.5. Note that for a regularly varying random variable X with index α one has that $X \in L^\beta \ \forall\, 0 < \beta < \alpha$ and $X \notin L^\beta \ \forall\, \beta > \alpha$ □

The next theorem provides the basis for our analysis of MS-ARMA processes with regularly varying noise. For some matrix A we denote by A^{-1} the pre-image under A.

Theorem 3.6. *Let $\epsilon = (\epsilon_k)_{k\in\mathbb{N}_0}$ be an i.i.d. sequence of $\mathbb{R}^d$-valued random variables in $\mathcal{R}_\alpha$ and ν, $(a_n)_{n\in\mathbb{N}}$ be the measure and normalizing sequence associated to ϵ_k in Theorem 3.2 (ii). Assume, moreover, that $A = (A_k)_{k\in\mathbb{N}_0}$ is a sequence of $M_{qd}(\mathbb{R})$-valued random variables independent of ϵ.*
If $\alpha < 1$, assume that there is a $0 < \eta < \alpha$ with $\alpha + \eta < 1$ such that $A_k \in L^{\alpha+\eta}$ for all $k \in \mathbb{N}_0$ and

$$\sum_{k=0}^{\infty} E\left(\|A_k\|^{\alpha+\eta}\right) < \infty \text{ and } \sum_{k=0}^{\infty} E\left(\|A_k\|^{\alpha-\eta}\right) < \infty. \tag{3.1}$$

If $\alpha \geq 1$, assume that there is a $0 < \eta < \alpha$ such that $A_k \in L^{\alpha+\eta}$ for all $k \in \mathbb{N}_0$ and

$$\sum_{k=0}^{\infty} E\left(\|A_k\|^{\alpha+\eta}\right)^{1/(\alpha+\eta)} < \infty \text{ and } \sum_{k=0}^{\infty} E\left(\|A_k\|^{\alpha-\eta}\right)^{1/(\alpha+\eta)} < \infty. \tag{3.2}$$

Then the tail behaviour of $Y = \sum_{k=0}^\infty A_k\epsilon_k$ is given by

$$nP\left(\sum_{k=0}^{\infty} A_k\epsilon_k \in a_n\cdot\right) \xrightarrow{v} \tilde{\nu}(\cdot) := \sum_{k=0}^{\infty} E\left(\nu \circ A_k^{-1}(\cdot)\right) \tag{3.3}$$

in $M_+\left(\overline{\mathbb{R}^q}\backslash\{0\}\right)$.

In particular, $Y = \sum_{k=0}^\infty A_k\epsilon_k$ is in $\mathcal{R}_\alpha$ with associated measure $\tilde{\nu}$ and normalizing sequence $(a_n)_{n\in\mathbb{N}}$, provided there is a relatively compact $K \in \mathcal{B}\left(\overline{\mathbb{R}^q}\backslash\{0\}\right)$ and an index $j \in \mathbb{N}_0$ such that $E\left(\nu\left(A_j^{-1}(K)\right)\right) > 0$.

This theorem is a straightforward generalization of Resnick & Willekens (1991, Th. 2.1), who consider random vectors and matrices with positive entries. We omit giving a proof, since an inspection of Resnick & Willekens (1991) shows that all their arguments carry through to our set-up (see also Stelzer (2005, Th. 3.19)).

Remark 3.7. Condition (3.1) or (3.2), respectively, is independent of the norm used and motivated mainly by the proof. If $A_k \in L^\beta$ for some $\beta > \alpha$ and all $k \in \mathbb{N}_0$ and

$$\limsup_{k\to\infty} E\left(\|A_k\|^\beta\right)^{1/k} < 1,$$

then (3.1) or (3.2), respectively, is satisfied for all admissible η with $\eta \leq \beta - \alpha$, as the root criterion from standard analysis shows. $\square$

4 MS-ARMA processes driven by regularly varying noise

Returning back to MS-ARMA processes we are now equipped with the necessary tools to study the effects of a regularly varying noise sequence ϵ.

Theorem 4.1. *Let $(\epsilon_t)_{t\in\mathbb{Z}}$ be an i.i.d. sequence of $\mathbb{R}^d$-valued random variables in $\mathcal{R}_\alpha$ and ν, $(a_n)_{n\in\mathbb{N}}$ the associated measure and normalizing sequence of Theorem 3.2 (ii). Assume further that $E(\log^+\|\mathbf{A}_0\|) < \infty$ and $\gamma < 0$.*

If $\alpha < 1$, assume there is an η with $0 < \eta < \alpha$ and $\alpha+\eta < 1$ such that $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k} \in L^{\alpha+\eta}$ for all $k \in \mathbb{N}_0$ and that

$$\sum_{k=0}^{\infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k}\|^{\alpha+\eta}\right) < \infty, \quad \sum_{k=0}^{\infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k}\|^{\alpha-\eta}\right) < \infty. \tag{4.1}$$

If $\alpha \geq 1$, assume that there is an η with $0 < \eta < \alpha$ such that $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k} \in L^{\alpha+\eta}$ for all $k \in \mathbb{N}_0$ and that

$$\sum_{k=0}^{\infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k}\|^{\alpha+\eta}\right)^{\frac{1}{\alpha+\eta}} < \infty, \quad \sum_{k=0}^{\infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k}\|^{\alpha-\eta}\right)^{\frac{1}{\alpha+\eta}} < \infty. \tag{4.2}$$

Then the following hold:

a) *There is a unique stationary and ergodic solution $X = (X_t)_{t\in\mathbb{Z}}$ to the MS-ARMA equation (2.1) given by Theorem 2.4.*

b) *The tail behaviour of $\mathbf{X}_0$, the state space representation of the stationary solution, is given by*

$$nP(\mathbf{X}_0 \in a_n \cdot) \xrightarrow{v} \tilde{\nu}(\cdot) = \sum_{k=0}^{\infty} E\left(\nu \circ (\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k})^{-1}(\cdot)\right). \tag{4.3}$$

c) *For the stationary solution X_0 the tail behaviour is described by*

$$nP(X_0 \in a_n \cdot) \xrightarrow{v} \bar{\nu}(\cdot) = \sum_{k=0}^{\infty} E\left(\nu \circ (\mathbf{P}\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k})^{-1}(\cdot)\right), \tag{4.4}$$

where $\mathbf{P} := (I_d, 0, \ldots, 0) \in M_{d,(p+q)d}(\mathbb{R})$ with I_d being the identity on $\mathbb{R}^d$.

d) *Provided there is a relatively compact $K \in \overline{\mathbb{R}^d}\backslash\{0\}$ with $E\left(\nu \circ \Sigma_0^{-1}(K)\right) > 0$, $\mathbf{X}_0$ and X_0 are in $\mathcal{R}_\alpha$ with normalizing sequence (a_n) and measures $\tilde{\nu}$ and $\bar{\nu}$, respectively.*

e) *Finally, if $\epsilon_0 \in L^\alpha$, then $\mathbf{X}_0$ and X_0 are in L^α.*

Proof: From $\mathbf{\Sigma}_0 \in L^{\alpha+\eta}$ and $\epsilon_0 \in \mathcal{R}_\alpha$ one gets $\mathbf{C}_0 = \mathbf{\Sigma}_0\epsilon_0 \in L^\beta \;\forall\, 0 < \beta < \alpha$ and, hence, $E(\log^+\|\mathbf{C}_0\|) < \infty$. Therefore a) is Theorem 2.4. Parts b) and c) follow from Theorem 3.6 using the series representation of $\mathbf{X}_0$ given in Theorem 2.4 and $X_t = \sum_{k=0}^{\infty} \mathbf{P}\mathbf{A}_t \cdots \mathbf{A}_{t-k+1}\mathbf{\Sigma}_{t-k}\epsilon_{t-k}$, noting that $\mathbf{P}$ is the projection on the first d coordinates.

Turning to d) we observe that

$$E\left(\nu \circ \mathbf{\Sigma}_0^{-1}(A_1 \times A_2 \times \cdots \times A_{p+q})\right) = E(\nu \circ \Sigma_0^{-1}(A_1 \cap A_{p+1})) \times \delta_0(A_2 \times \cdots \times A_{p-1} \times A_{p+2} \times \cdots A_{p+q}) \tag{4.5}$$

for $A_i \in \mathcal{B}(\mathbb{R}^d)$, where δ_0 denotes the Dirac measure with respect to 0 in $\mathbb{R}^{d(p+q-2)}$. So, setting $\tilde{K} = K \times 0_{\mathbb{R}^{d(p-1)}} \times K \times 0_{\mathbb{R}^{d(q-1)}}$ gives a relatively compact set with $E(\nu \circ \mathbf{\Sigma}_0^{-1}(\tilde{K})) > 0$. Furthermore, $E(\nu \circ (\mathbf{P\Sigma}_0)^{-1}(K)) = E(\nu \circ \mathbf{\Sigma}_0^{-1}(K \times \mathbb{R}^{d(p+q-1)})) \overset{(4.5)}{=} E(\nu \circ \Sigma_0^{-1}(K)) > 0$ and thus $\tilde{\nu}$ and $\bar{\nu}$ are non-degenerate, which proves d).

$\epsilon_0 \in L^\alpha$ and (4.1) or (4.2), respectively, ensure that the conditions (4.1) or (4.2) of Stelzer (2006, Th. 4.2) hold with $r = \alpha$, as ϵ and Δ are independent. Thus e) follows from this Theorem. □

Remark 4.2. a) From the above results the extremal domain of attraction of the stationary marginal distribution of the MS-ARMA process can be immediately deduced using e.g. Resnick (1987, Corollary 5.18). In the case $d = 1$, we have tail equivalence of the stationary distribution and the driving noise and, in particular, that the distributions of ϵ_0 and X_0 both belong to the maximum domain of attraction of the Fréchet distribution Φ_α (cf. e.g. Resnick (1987) or Embrechts, Klüppelberg & Mikosch (1997)), provided the upper tails are non-degenerate.

b) For the non-degeneracy condition $E(\nu \circ \Sigma^{-1}(K)) > 0$ in d) it suffices that Σ_0 has a strictly positive probability of being invertible. (If Σ_0 is invertible, $\Sigma_0^{-1}(B_0(1)) \subseteq \|\Sigma_0^{-1}\| B_0(1)$, hence $\Sigma_0^{-1}\left((1,\infty]\mathbb{S}^{d-1}\right) = \left(\Sigma_0^{-1}(B_0(1))\right)^c \supseteq (\|\Sigma_0^{-1}\|, \infty]\mathbb{S}^{d-1}$ and thus

$$\nu \circ \mathbf{\Sigma}_0^{-1}\left((1,\infty]\mathbb{S}^{d-1} \times 0_{\mathbb{R}^{d(p-1)}} \times (1,\infty]\mathbb{S}^{d-1} \times 0_{\mathbb{R}^{d(q-1)}}\right) = \nu \circ \Sigma_0^{-1}\left((1,\infty]\mathbb{S}^{d-1}\right) > 0$$

due to the non-degeneracy of ν.)

d) For the one-dimensional stochastic difference equation $X_t = A_t X_{t-1} + C_t$ with i.i.d. (A_t, C_t) similar results are to be found in Grey (1994) or Konstantinides & Mikosch (2005) and for one-dimensional positive valued random coefficient autoregressive models in Resnick & Willekens (1991). For the multivariate general stochastic difference equation $\mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{C}_t$ with ergodic $(\mathbf{A}_t)$ and i.i.d. $(\mathbf{C}_t)$ independent of $(\mathbf{A}_t)$ see Stelzer (2005, Section 4.3). □

The regular variation results can be strengthened further.

Theorem 4.3. *If all conditions of Theorem 4.1 including the existence of a relatively compact* $K \in \overline{\mathbb{R}^d}\backslash\{0\}$ *with* $E\left(\nu \circ \Sigma_0^{-1}(K)\right) > 0$ *are satisfied, then* $\mathbf{X} = (\mathbf{X}_t)_{t\in\mathbb{Z}}$ *as well as* $X = (X_t)_{t\in\mathbb{Z}}$ *are regularly varying as a sequence with index* α.

Proof: It remains to show that all finite dimensional distributions of $\mathbf{X} = (\mathbf{X}_t)_{t\in\mathbb{Z}}$ are regularly varying. We restrict ourselves to showing that the two-dimensional marginals are again regularly varying. It is obvious that the very same arguments can be used for all higher dimensional marginals.

W.l.o.g. we only consider the joint distribution of $\mathbf{X}_0$ and $\mathbf{X}_h$ for $h \in \mathbb{N}$. From the series representations of $\mathbf{X}_0$ and $\mathbf{X}_h$ we construct a series representation of $(\mathbf{X}_0^\mathsf{T}, \mathbf{X}_h^\mathsf{T})^\mathsf{T}$ as follows. Set

$$\mathbb{A}_h = \begin{pmatrix} 0_{M_{d(p+q),d}(\mathbb{R})} \\ \mathbf{\Sigma}_h \end{pmatrix}, \; \mathbb{A}_{h-k} = \begin{pmatrix} 0_{M_{d(p+q),d}(\mathbb{R})} \\ \mathbf{A}_h \mathbf{A}_{h-1} \cdots \mathbf{A}_{h-k+1} \mathbf{\Sigma}_{h-k} \end{pmatrix} \text{ for } k = 1, 2, \ldots, h-1$$

$$\mathbb{A}_0 = \begin{pmatrix} \mathbf{\Sigma}_0 \\ \mathbf{A}_h \mathbf{A}_{h-1} \cdots \mathbf{A}_1 \mathbf{\Sigma}_0 \end{pmatrix},$$

$$\mathbb{A}_{h-k} = \begin{pmatrix} \mathbf{A}_0 \mathbf{A}_{-1} \cdots \mathbf{A}_{h-k+1} \mathbf{\Sigma}_{h-k} \\ \mathbf{A}_h \mathbf{A}_{h-1} \cdots \mathbf{A}_{h-k+1} \mathbf{\Sigma}_{h-k} \end{pmatrix} \text{ for } k = h+1, h+2, \ldots,$$

then $(\mathbf{X}_0^\mathsf{T}, \mathbf{X}_h^\mathsf{T})^\mathsf{T} = \sum_{k=0}^\infty \mathbb{A}_{h-k} \epsilon_{h-k}$ and the sequences $(\mathbb{A}_{h-k})_{k \in \mathbb{N}_0}$ and $(\epsilon_{h-k})_{k \in \mathbb{N}_0}$ are mutually independent. On $\mathbb{R}^{2d(p+q)}$ consider the norm $\|\cdot\|_*$ defined via the norm $\|\cdot\|$ used on $\mathbb{R}^{d(p+q)}$ by $\|(x_1^\mathsf{T}, x_2^\mathsf{T})^\mathsf{T}\|_* = \max\{\|x_1\|, \|x_2\|\}$. For any matrix $A \in M_{2d(p+q),d}(\mathbb{R})$ with $A = (A_1^\mathsf{T}, A_2^\mathsf{T})^\mathsf{T}$, where $A_1, A_2 \in M_{d(p+q),d}(\mathbb{R})$, it holds that $\|A\|_* \leq \max\{\|A_1\|, \|A_2\|\} \leq \|A_1\| + \|A_2\|$. Using (4.1) or (4.2), respectively, the triangle inequalities in $L^{\alpha \pm \eta}$ and the elementary inequality $|a+b|^r \leq |a|^r + |b|^r$ for $0 < r \leq 1$ and all $a, b \in \mathbb{R}$, we thus obtain from the definition of $\mathbb{A}_{h-i}$ that $\mathbb{A}_{h-i} \in L^{\alpha+\eta}$ for all $i \in \mathbb{N}_0$ and $\sum_{k=0}^\infty E\left(\|\mathbb{A}_{h-k}\|_*^{\alpha+\eta}\right) < \infty$, $\sum_{k=0}^\infty E\left(\|\mathbb{A}_{h-k}\|_*^{\alpha-\eta}\right) < \infty$, if $\alpha < 1$, or $\sum_{k=0}^\infty E\left(\|\mathbb{A}_{h-k}\|_*^{\alpha+\eta}\right)^{1/(\alpha+\eta)} < \infty$, $\sum_{k=0}^\infty E\left(\|\mathbb{A}_{h-k}\|_*^{\alpha-\eta}\right)^{1/(\alpha+\eta)} < \infty$, if $\alpha \geq 1$. So Theorem 3.6 gives $nP\left((\mathbf{X}_0^\mathsf{T}, \mathbf{X}_h^\mathsf{T})^\mathsf{T} \in a_n \cdot\right) \xrightarrow{v} \hat{\nu}(\cdot) := \sum_{k=0}^\infty E\left(\nu \circ \mathbb{A}_{h-k}^{-1}(\cdot)\right)$ as $n \to \infty$. Since $\mathbb{A}_h^{-1}\left(0_{\mathbb{R}^{d(p+q)}} \times K \times 0_{\mathbb{R}^{d(p-1)}} \times K \times 0_{\mathbb{R}^{d(q-1)}}\right) = \Sigma_h^{-1}(K)$, the measure $\hat{\nu}$ is non-degenerate under the non-degeneracy condition of Theorem 4.1 d) and so $(\mathbf{X}_0^\mathsf{T}, \mathbf{X}_h^\mathsf{T})^\mathsf{T}$ is multivariate regularly varying with index α, measure $\hat{\nu}$ and normalizing sequence (a_n). To obtain the result for the marginal distribution of the MS-ARMA process, i.e. for (X_0, X_h), one again simply needs to employ a projection onto the first and $(p+q+1)$th d-dimensional coordinate. □

Using Remark 3.7 a) and the Jensen's inequality to obtain $\gamma < 0$ one gets some asymptotic criteria replacing the summability conditions.

Lemma 4.4. *Let $(\epsilon_t)_{t \in \mathbb{Z}}$ be a sequence of i.i.d. $\mathbb{R}^d$-valued random variables in $\mathcal{R}_\alpha$ and ν, $(a_n)_{n \in \mathbb{N}}$ the associated measure and normalizing sequence of Theorem 3.2 (ii). Assume that there is a $\beta > \alpha$ such that $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \mathbf{\Sigma}_{-k} \in L^\beta$ and $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \in L^\beta$ for all $k \in \mathbb{N}_0$ and that*

$$\limsup_{n \to \infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-n+1} \mathbf{\Sigma}_{-n}\|^\beta\right)^{1/(n+1)} < 1, \quad \limsup_{n \to \infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-n+1}\|^\beta\right)^{1/n} < 1. \quad (4.6)$$

Then the conditions of Theorem 4.1 are satisfied.

Under an independence condition this simplifies further.

Corollary 4.5. *Assume $\mathbf{\Sigma}_{-k}$ is independent of $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}$ $\forall\, k \in \mathbb{N}_0$. Then $\mathbf{\Sigma}_0 \in L^\beta$ and $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \in L^\beta$ $\forall\, k \in \mathbb{N}_0$ are sufficient for $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \mathbf{\Sigma}_{-k} \in L^\beta$ $\forall k \in \mathbb{N}_0$ and $\limsup_{n \to \infty} E\left(\|\mathbf{A}_0 \cdots \mathbf{A}_{-n+1}\|^\beta\right)^{1/n} < 1$ implies already that (4.6) is satisfied.*

In order to obtain a condition that can be verified easily, we use the following Theorem from Stelzer (2006) and thereby extend the feasible stationarity condition given in Corollary 3.4 of that paper to ensure that the conditions for a regularly varying noise to determine the tail-behaviour of the stationary MS-ARMA process are satisfied.

Theorem 4.6. *Let $d, p \in \mathbb{N}$, $q \in \mathbb{N}_0$ and $\mathcal{A} \subset M_{d(p+q)}(\mathbb{R})$ be a set of matrices such that for each $A \in \mathcal{A}$ there are matrices $A_1(A), \ldots, A_p(A), B_1(A), \ldots, B_q(A) \in M_d(\mathbb{R})$ such that*

$$A = \begin{pmatrix} A_1(A) & \cdots & A_{p-1}(A) & A_p(A) & B_1(A) & \cdots & B_{q-1}(A) & B_q(A) \\ I_d & 0\cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & 0 & 0 & \cdots & \cdots & \vdots \\ 0 & \cdots 0 & I_d & 0 & 0 & \cdots & \cdots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & I_d & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots & 0 & \ddots & 0\cdots & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots 0 & I_d & 0 \end{pmatrix}.$$

Assume, moreover, that there is a norm $\|\cdot\|_d$ on $\mathbb{R}^d$ and $c < 1$ such that $\sup\limits_{A \in \mathcal{A}} \sum_{i=1}^p \|A_i(A)\|_d < c$ and $\sup\limits_{A \in \mathcal{A}} \sum_{i=1}^q \|B_i(A)\|_d < \infty$ hold for the induced operator norm.

Then there is a norm $\|\cdot\|$ on $\mathbb{R}^{d(p+q)}$ and $c' < 1$ such that $\sup_{A \in \mathcal{A}} \|A\| < c'$ in the induced operator norm. Especially, $\|x_0 x_1 \cdots x_k\| < (c')^{k+1}$ for any $k \in \mathbb{N}$ and sequence $(x_n)_{n \in \mathbb{N}_0}$ with elements in $\mathcal{A}$.

Lemma 4.7. *Assume that there are $c < 1$, $C, M \in \mathbb{R}^+$ and a norm $\|\cdot\|$ on $\mathbb{R}^d$ such that $\sum_{i=1}^p \|\Phi_{i0}\| \leq c$, $\sum_{i=1}^q \|\Theta_{i0}\| \leq M$ and $\|\Sigma_0\| \leq C$ a.s. Then $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k} \in L^\beta$, $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \in L^\beta$ for all $k \in \mathbb{N}_0$ and (4.6) is satisfied for all $\beta > \alpha$.*

Proof: Define the subset $\mathcal{A} = \{\mathbf{A}_0 : \sum_{i=1}^p \|\Phi_{i0}\|_d \leq \bar{c}\}$ of the state space of $\mathbf{\Phi}_t$. Then the conditions of this Lemma imply that the process $(\mathbf{A}_t)_{t \in \mathbb{Z}}$ a.s. takes only values in $\mathcal{A}$ at all times $t \in \mathbb{Z}$. From Theorem 4.6 we thus obtain obtain an operator norm $\|\cdot\|$ which ensures $\|\mathbf{A}_0\mathbf{A}_{-1} \cdots \mathbf{A}_{-k+1}\| < (c')^k$ a.s. for some $c' < 1$ and all $k \in \mathbb{N}_0$. Thus, $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1} \in L^\beta$ for all $\beta > \alpha$ and the second part of (4.6) is satisfied. Furthermore, $\|\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k}\| \leq C(c')^k$ implies $\mathbf{A}_0 \cdots \mathbf{A}_{-k+1}\mathbf{\Sigma}_{-k} \in L^\beta$ for all $k \in \mathbb{N}_0$ and $\beta > \alpha$ and that the first part of (4.6) is satisfied. □

Note that in Stelzer (2006) it was shown that under similar conditions an MS-ARMA process is not only stationary, but also geometrically ergodic/strong mixing and has finite moments of at least as many orders as the driving noise ϵ.

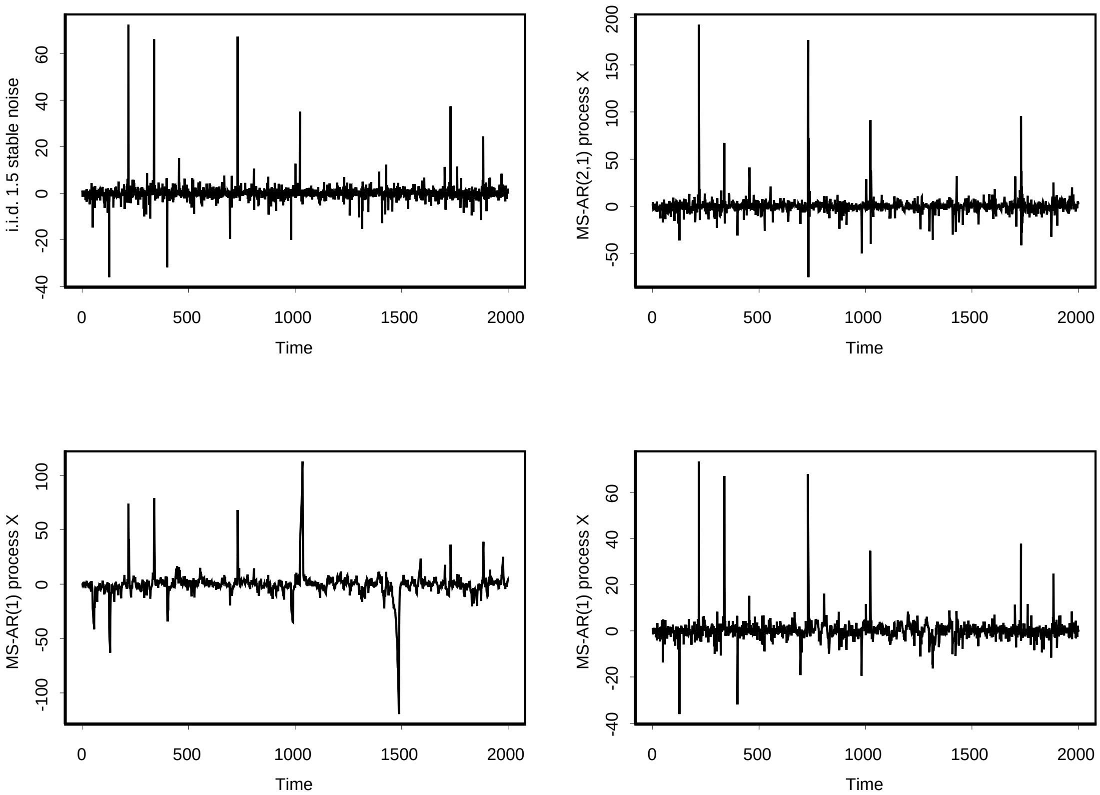

Figure 1: Simulations of an i.i.d. symmetric 1.5-stable noise sequence (upper left), the MS-ARMA(2,1) process from Example 7.1 (upper right) and the MS-AR(1) processes from Examples 7.2 (lower left) and 7.3 (lower right)

5 Some illustrative examples

Finally, we consider some examples and simulate sample models. We shall look at real-valued MS-ARMA(p,q) processes with $\Sigma_t = 1$, i.e. $X_t = \Phi_{1t}X_{t-1} + \ldots + \Phi_{pt}X_{t-p} + \epsilon_t + \Theta_{1t}\epsilon_{t-1} + \ldots + \Theta_{qt}\epsilon_{t-q}$. As noise we take an i.i.d. sequence ϵ_t with symmetric 1.5-stable distribution, cf. Figure 1 (upper left) for a simulation. In particular, this noise is non-degenerately regularly varying in both tails with index 1.5. The results of Stelzer (2006) give that all examples below are geometrically ergodic. Thus, we use arbitrary starting values for the MS-ARMA processes and show the simulated values after an appropriate burn-in period only.

In the first two examples we presume that there are only two possible states of Δ given by $\Delta^{(1)}$ and $\Delta^{(2)}$ and that the transition matrix of the Markov parameter chain Δ is

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} \bar{p} & 1-\bar{p} \\ 1-\bar{p} & \bar{p} \end{pmatrix}$$

for some $\bar{p} \in (0,1)$. Thus, the stationary distribution is $(\pi^{(1)}, \pi^{(2)}) = (1/2, 1/2)$ and Δ is aperiodic and irreducible.

Example 5.1: Take $\bar{p} = 3/4$ and let us consider an MS-ARMA(2,1) process with the two regimes given by the equations

$$X_t = 0.6X_{t-1} - 0.3X_{t-2} + \epsilon_t + 2\epsilon_{t-1} \text{ and } X_t = -0.5X_{t-1} + 0.2X_{t-2} + \epsilon_t + 0.5\epsilon_{t-1}.$$

Obviously the conditions of Lemma 4.7 are satisfied and so Theorem 4.1 in combination with Lemma 4.4 shows that the MS-ARMA process is stationary and regularly varying with index 1.5. The simulation in Figure 1 (upper right) illustrates this, in particular, that the stationary distribution is tail equivalent to the noise.

Example 5.2: Take $\bar{p} = 3/4$ and consider a real valued MS-AR(1) process with two regimes given by the AR(1) coefficients $\mathbf{\Phi}_{(1)} = 1/2$ and $\mathbf{\Phi}_{(2)} = 11/10$, respectively. Although the second regime is explosive and thus Lemma 4.7 is not applicable, we can still show stationarity and regular variation. Regarding the conditions of Lemma 4.4, the only problem is (4.6), but as we have only finitely many regimes, we can use the tools of Francq & Zakoïan (2001). Define

$$Q = \begin{pmatrix} p_{11}{\mathbf{\Phi}_{(1)}}^2 & p_{21}{\mathbf{\Phi}_{(1)}}^2 \\ p_{12}{\mathbf{\Phi}_{(2)}}^2 & p_{22}{\mathbf{\Phi}_{(2)}}^2 \end{pmatrix}.$$

Then one calculates $\rho(Q) = (219 + \sqrt{23761})/400 \approx 0.9328650868$ for the spectral radius of Q. Tedious but elementary calculations along similar lines as the arguments in Francq & Zakoïan (2001) (cf. Stelzer (2005, Theorem 5.23)) show that $\rho(Q) < 1$ implies that (4.6) holds with $\beta = 2$. Thus, all conditions of Theorems 4.1 and 4.3 are satisfied. Again, the simulation in Figure 1 (lower left) illustrates the fact that the MS-AR process is regularly varying (with index 1.5) and the stationary distribution is tail equivalent to the noise.

Of particular interest is the downwards going spike at about time 1500, which obviously is not caused by a large shock in the noise sequence ϵ. In fact, it comes from the autoregressive coefficient being 1.1 over a rather long period. MS-AR(1) processes with finitely many regimes where the tails of the stationary distribution are determined by such events were studied in Saporta (2005). Our theoretic results show that the tails are, however, determined by the noise in this example and it is also easy to see that this does not change, if we take any other noise ϵ which is regularly varying with index less than two. However, using Saporta (2005) and calculating the λ as defined in Theorem 1.1 (1) of this article numerically as $\lambda = 2.8875$ one obtains that the MS-AR(1) process given above with a different noise ϵ has a stationary distribution that is regularly varying with index 2.8875 provided $\epsilon_0 \in L^r$ for some $r > 2.8875$. In this case not the noise but the possible occurrence of explosive regimes determines the tail behaviour. The regularly varying tail behaviour can be seen in Figure 2 where the process is simulated with a standard normal i.i.d. noise ϵ. Note in particular how the spikes build up due to consecutive occurrences of the explosive regime.

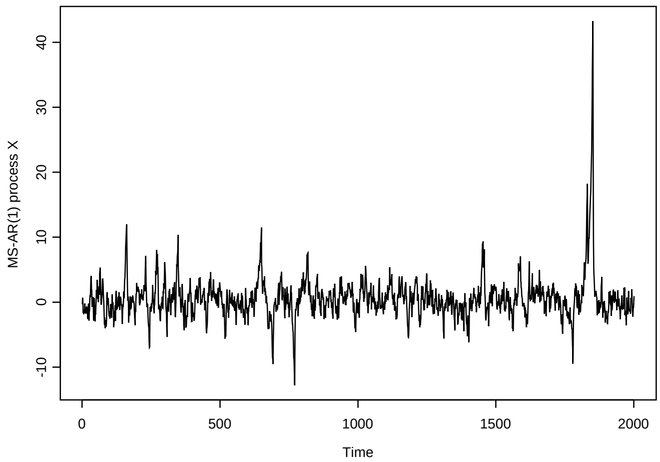

Figure 2: Simulation of the MS-AR(1) process in Example 5.2 with a standard normal noise ϵ

Observe also that we have obtained regular variation as a sequence for the MS-ARMA processes with a regularly varying noise above, whereas the results of Saporta (2005) only give that the stationary distribution is regularly varying in the tails. What happens for noise sequences ϵ which are regularly varying with index two or greater and do not satisfy $\epsilon_0 \in L^r$ for some $r > 2.8875$ requires a detailed analysis which is beyond the scope of the present paper.

Example 5.3: Finally, we consider an MS-AR(1) process with an uncountable state space for the parameter Φ_{1t}. Take a, b, c such that $-1 < a < b < 1$ and $c > 0$ and an i.i.d. sequence (u_t) uniformly distributed on the interval $[-1, 1]$. Then the evolution of the autoregressive coefficient shall be given by $\Phi_{1t} = \max\left(\min\left(\Phi_{1,t-1} + cu_t, b\right), a\right)$, i.e. we choose the new parameter uniformly from the neighbourhood with radius c of the old one, but do not allow it to leave the interval $[a; b]$. Using Lemma 4.7, Lemma 4.4 and Theorem 4.1 one sees that the MS-AR(1) process is stationary and regularly varying with index 1.5. The simulation in Figure 1 (lower right) with $a = -0.9$, $b = 0.9$ and $c = 0.05$ again illustrates in particular the regular variation and tail equivalence.

Observe that in the above examples one deducts immediately from (4.3) that both tails of the stationary distribution of the MS-AR(1) process are non-degenerately regularly varying, since this holds for the noise ϵ_t.

Acknowledgments

The research leading to this paper has originally been undertaken for the author's diploma thesis. The author would like to thank his supervisors Claudia Klüppelberg and Alexander

Lindner for all their advice. Furthermore financial support by the Deutsche Forschungsgemeinschaft through the graduate programme "Angewandte Algorithmische Mathematik" at the Munich University of Technology is gratefully acknowledged.

References

Basrak, B. (2000). *The Sample Autocorrelation Function of Non-Linear Time Series*, PhD-Thesis, Rijksuniversiteit Groningen, The Netherlands.

Bauer, H. (1992). *Maß- und Integrationstheorie*, 2nd edn, Walter de Gruyter, Berlin.

Bingham, G. H., Goldie, C. M. & Teugels, J. L. (1989). *Regular Variation*, 1st paperback edn, Cambridge University Press, Cambridge.

Brockwell, P. J. & Davis, R. A. (1991). *Time Series: Theory and Methods*, 2nd edn, Springer, New York.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

Francq, C. & Zakoïan, J.-M. (2001). Stationarity of multivariate Markov-switching ARMA models, *J. Econometrics* **102**: 339–364.

Grey, D. R. (1994). Regular variation in the tail behaviour of random difference equations, *Ann. Appl. Probab.* **4**: 169–183.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* **57**: 357–384.

Hult, H. & Lindskog, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions, *Adv. in Appl. Probab.* **34**: 587–608.

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

Klüppelberg, C. & Pergamenchtchikov, S. (2004). The tail of the stationary distribution of a random coefficient AR(q) model, *Ann. Appl. Probab.* **14**: 971–1005.

Konstantinides, D. G. & Mikosch, T. (2005). Large deviations and ruin probabilities for solutions to stochastic recurrence equations with heavy-tailed innovations, *Ann. Probab.* **33**: 1992–2035.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions*, Vol. 454 of *Lecture Notes in Economics and Mathematical Systems*, Springer, Berlin.

Lindskog, F. (2004). *Multivariate Extremes and Regular Variation for Stochastic Processes*, Dissertation, Department of Mathematics, ETH Zürich, Zürich, Switzerland. available from: `e-collection.ethbib.ethz.ch/diss/`.

Mikosch, T. (1999). Regular variation, subexponentiatility and their applications in probability theory, *Lecture notes for the workshop "Heavy Tails and Queues"*, EURANDOM Institute, Eindhoven, The Netherlands. available from `www.math.ku.dk/~mikosch`.

Mikosch, T. (2003). Modeling dependence and tails of financial time series, *in* B. Finkenstädt & H. Rootzén (eds), *Extreme Values in Finance, Telecommunications, and the Environment*, Chapman and Hall, London, pp. 185–286.

Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*, Springer, New York.

Resnick, S. I. (2004). On the foundations of multivariate heavy tail analysis, *J. Appl. Probab.* **41A**: 191–212.

Resnick, S. I. & Willekens, E. (1991). Moving averages with random coefficients and random coefficient autoregressive models, *Comm. Statist. Stochastic Models* **7**: 511–525.

Saporta, B. d. (2005). 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**: 1954–1978.

Stelzer, R. (2006). On Markov-switching ARMA processes – stationarity, existence of moments and geometric ergodicity, *Submitted for publication*, Centre for Mathematical Sciences, Munich University of Technology, Garching bei München. available from: `www.ma.tum.de/stat`.

Stelzer, R. J. (2005). *On Markov-Switching Models – Stationarity and Tail Behaviour*, Diploma Thesis, Centre for Mathematical Sciences, Munich University of Technology, D-85747 Garching bei München, Germany. available from: `www.ma.tum.de/stat`.