

INSTITUT DE STATISTIQUE  
BIOSTATISTIQUE ET  
SCIENCES ACTUARIELLES  
(ISBA)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN



DISCUSSION  
PAPER

2013/17

Markov Tail Chains

JANSSEN, A. and J. SEGERS

30 April 2013

**MARKOV TAIL CHAINS**A. JANSSEN,\* *University of Hamburg*J. SEGERS,\*\* *Université catholique de Louvain***Abstract**

The extremes of a univariate Markov chain with regularly varying stationary marginal distribution and asymptotically linear behavior are known to exhibit a multiplicative random walk structure called the tail chain. In this paper, we extend this fact to Markov chains with multivariate regularly varying marginal distribution in  $\mathbb{R}^d$ . We analyze both the forward and the backward tail process and show that they mutually determine each other through a kind of adjoint relation. In a broader setting, it will be seen that even for non-Markovian underlying processes a Markovian forward tail chain always implies that the backward tail chain is Markovian as well. We analyze the resulting class of limiting processes in detail. Applications of the theory yield the asymptotic distribution of both the past and the future of univariate and multivariate stochastic difference equations conditioned on an extreme event.

2010 Mathematics Subject Classification: Primary 60G70; 60J05

Secondary 60G10; 60H25; 62P05

*Keywords:* autoregressive conditional heteroskedasticity; extreme value distribution; (multivariate) Markov chain; multivariate regular variation; random walk; stochastic difference equation; tail chain; tail-switching potential

---

\* Postal address: University of Hamburg, Department of Mathematics, Bundesstr. 55, 20146 Hamburg, Germany

\* Email address: anja.janssen@math.uni-hamburg.de

\*\* Postal address: Université catholique de Louvain, Institut de statistique, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium

\*\* Email address: johan.segers@uclouvain.be

## 1. Introduction

Consider a discrete-time,  $\mathbb{R}^d$ -valued random process  $\{X_t : t = 0, 1, 2, \dots\}$  defined by the recursive equation

$$X_t = \Phi(X_{t-1}, \varepsilon_t), \quad t = 1, 2, \dots, \quad (1.1)$$

where

- (i)  $\varepsilon_1, \varepsilon_2, \dots$  are independent and identically distributed random elements of a measurable space  $(\mathbb{E}, \mathcal{E})$  and independent of  $X_0$ ; (1.2)
- (ii)  $\Phi$  is a measurable function from  $\mathbb{R}^d \times \mathbb{E}$  to  $\mathbb{R}^d$ .

If the process  $\{X_t\}$  happens to be stationary, it will be assumed to be defined for all integer  $t$ . The distribution of  $X_0$  is assumed to be multivariate regularly varying.

The aim of the paper is to analyze the special structure of weak limits of the finite-dimensional distributions of the process conditionally on  $\|X_0\|$  being large, where  $\|\cdot\|$  denotes the Euclidean norm. More precisely, we will investigate the weak limits, called the forward tail chain, of vectors of the form  $(X_0, \dots, X_t)$  given that  $\|X_0\|$  exceeds a high threshold. If in addition the process is stationary we will extend this to find the so-called back-and-forth tail chain, which corresponds to the weak limits of vectors of the form  $(X_{-s}, \dots, X_t)$  given that  $\|X_0\|$  is large. A close relation of these processes to multivariate regular variation of the whole process has been analyzed in Basrak and Segers (2009). In this article, we are interested in the special form of the processes, in particular the Markovian structure of both the forward and the backward process and how they necessarily determine each other.

The process  $\{X_t\}$  is obviously a discrete-time homogeneous Markov chain. On the other hand, every homogeneous discrete-time Markov chain  $\{X_t\}$  on a complete separable metric space can be represented as in (1.1)–(1.2) (Kifer, 1986). Of course, for a given Markov chain  $\{X_t\}$  the above representation is not unique. Still, in examples, the way in which Markov chains are defined is often through a recursive equation; all examples in Goldie (1991, pp. 126–127), for instance, are of this type. The chain is stationary if and only if the random vectors  $X_1 = \Phi(X_0, \varepsilon_1)$  and  $X_0$  are equal in law.

In Smith (1992) and Perfekt (1994), excursions of a univariate Markov chain over a high threshold following an extreme event are shown to behave asymptotically and

under quite general conditions as a (multiplicative) random walk. The theory has been extended to multivariate Markov chains in Perfekt (1997) and to higher-order Markov chains in Yun (1998, 2000). More recently, Resnick and Zeber (2013) have analyzed the topic with a special view towards the convergence of Markov kernels and added a criterion to distinguish between extreme and non-extreme states of a Markov chain as the threshold rises. The random-walk representation is useful from a statistical perspective because it gives a handle on how to model the extremes of certain time series (Bortot and Coles (2000); Coles et al. (1997); Smith et al. (1997)). A useful, well-investigated class of processes for which the random walk structure is quite revealing are the stationary solutions to certain stochastic difference equations, including squared (generalized) autoregressive conditionally heteroskedastic (ARCH/GARCH) processes as a special case (Basrak et al. (2002b); Gomes et al. (2004); de Haan et al. (1989)).

A limitation of the theory of Smith (1992), Perfekt (1994) and Resnick and Zeber (2013) is that it is specialized to univariate, nonnegative Markov chains. Similarly, Perfekt (1997) only considers the upper extremes of a multivariate Markov chain. When extending the theory to real-valued and higher dimensional chains, one has to keep in mind that extremes may be both positive or negative and that extreme values of  $X_t$  may depend not only on  $\|X_{t-1}\|$  but also on  $X_t/\|X_{t-1}\|$ . The simplest case of the extension on which we will focus deals with real-valued univariate Markov chains, where an extreme value of  $X_t$  may depend on the sign of  $X_{t-1}$  as can be observed for instance in time series of logreturns of prices of financial securities in periods of high volatility. The observation of this so-called leverage effect has led to the formulation of asymmetric extensions of GARCH models (cf., for example, Zivot (2009)). For such Markov chains with tail switching potential, the random walk representation of excursions over high thresholds breaks down in the sense that the distribution of the multiplicative increment now depends in general on the sign of the chain on the previous step. In Bortot and Coles (2003), a more general representation is postulated, involving in fact four transition mechanisms rather than one, corresponding to the four cases of transitions from and to upper or lower extreme states.

The novelty of this paper is two-fold: first, to explicitly state the random walk representation in the general  $\mathbb{R}^d$ -valued case; second, in the stationary case, to study the joint distribution of the forward and backward tail chain, coined the *back-and-*

*forth tail chain.* Throughout, some remarkable simplifications in the (univariate) real-valued case will be studied in more detail. In particular, in the univariate case the backward tail chain is again a random walk which is in some sense dual the forward tail chain. Besides the assumption that the distribution of  $X_0$  is regularly varying, the only condition is a relatively easy-to-check statement on the asymptotic behaviour of  $\Phi(x, \cdot)$  for large  $\|x\|$ .

The outline of the paper is as follows. The forward tail chain of a possibly non-stationary  $\mathbb{R}^d$ -valued Markov chain is studied in section 2. For stationary Markov chains, the tail chain can be extended to the past of the process, the backward tail chain, see section 3. Section 4 describes a kind of adjoint relation between distributions which is motivated by a general property of tail processes of stationary processes. In section 5, we show that a certain class of processes, coined back-and-forth tail chains, which are derived from this adjoint distribution, form exactly the class of tail processes which arise in our Markovian setting. Finally, section 6 provides some examples to the theory, including an application to stationary solutions of (multivariate) stochastic difference equations.

To conclude this section, let us fix some notations. We write  $(x)_+ = \max(x, 0)$  for the positive part of  $x \in \mathbb{R}$  and  $(x)_- = \min(x, 0)$  for the negative part. The transpose of a matrix  $A$  is denoted by  $A'$ . The law of a random vector  $X$  is denoted by  $\mathcal{L}(X)$ ; weak convergence of probability measures is denoted by  $\rightsquigarrow$ . The probability measure degenerate at a point  $x$  is denoted by  $\delta_x$ , and  $\text{Unif}(E)$  denotes the uniform distribution on a compact set  $E$ . The indicator of an event  $A$  is denoted by  $\mathbf{1}_A(\cdot)$ . We write  $\overline{\mathbb{R}}$  for  $\mathbb{R} \cup \{-\infty, \infty\}$ ,  $\mathbb{S}^{d-1}$  for  $\{x \in \mathbb{R}^d : \|x\| = 1\}$  and  $\mathbf{0}$  for a vector (of suitable dimension) which consists of all zeros. Let  $\mathbb{Z}$  be the set of integers and  $\mathbb{N}_0$  be the set of nonnegative integers.

## 2. Forward tail chains

Let  $X_0, X_1, X_2, \dots$  be a homogeneous Markov chain as in (1.1) and (1.2), not necessarily stationary. The focus of this section is on the weak limits of the finite-dimensional distributions of the process conditionally on  $\|X_0\|$  being large (Theorem 2.1). Two conditions are required: Condition 2.1 on the tails of  $X_0$ , and Condition 2.2 on the

asymptotics of  $x \mapsto \Phi(x, e)$  for large  $\|x\|$ . See for instance Resnick (2007) for details on multivariate regular variation.

**Condition 2.1.** *The distribution of  $X_0$  is multivariate regularly varying on  $\overline{\mathbb{R}^d} \setminus \{0\}$ , that is, there exists a non-degenerate probability measure  $\Upsilon$  on  $\mathbb{S}^{d-1}$  (called the spectral measure) and an  $\alpha > 0$  such that*

$$\lim_{x \rightarrow \infty} \mathbb{P} \left( \|X_0\| > ux, \frac{X_0}{\|X_0\|} \in S \mid \|X_0\| > x \right) = u^{-\alpha} \Upsilon(S) \quad (2.1)$$

for all Borel sets  $S \subset \mathbb{S}^{d-1}$  which satisfy  $\Upsilon(\partial S) = 0$  and  $u \geq 1$ .

The second condition states that the function  $\Phi$  in (1.1) is asymptotically homogeneous in  $x$  for large values of  $\|x\|$ .

**Condition 2.2.** *There exists a measurable map  $\phi : \mathbb{S}^{d-1} \times \mathbb{E} \mapsto \mathbb{R}^d$  such that, for all  $e \in \mathbb{E}$ ,*

$$\lim_{x \rightarrow \infty} x^{-1} \Phi(xs(x), e) = \phi(s, e) \quad (2.2)$$

whenever  $s(x) \rightarrow s$  in  $\mathbb{S}^{d-1}$ .

Moreover, if  $\mathbb{P}(\phi(s, \varepsilon_1) = 0) > 0$  for some  $s \in \mathbb{S}^{d-1}$ , then also  $\mathbb{P}(\varepsilon_1 \in \mathbb{W}) = 1$ , where  $\mathbb{W}$  is a measurable subset of  $\mathbb{E}$  such that for all  $e \in \mathbb{W}$ ,

$$\sup_{\|y\| \leq x} |\Phi(y, e)| = O(x), \quad x \rightarrow \infty. \quad (2.3)$$

We extend the domain of the limit function  $\phi$  in (2.2) to  $\mathbb{R}^d \times \mathbb{E}$  by setting

$$\phi(v, e) = \begin{cases} \|v\| \phi(v/\|v\|, e) & \text{if } v \neq 0, \\ 0 & \text{if } v = 0. \end{cases} \quad (2.4)$$

**Lemma 2.1.** *If Condition 2.2 holds, then*

$$\lim_{x \rightarrow \infty} x^{-1} \Phi(xv(x), e) = \phi(v, e) \quad (2.5)$$

whenever  $v(x) \rightarrow v \in \mathbb{R}^d \setminus 0$  and  $e \in \mathbb{E}$ . If  $\mathbb{P}(\phi(s, \varepsilon_1) = 0) > 0$  for some  $s \in \mathbb{S}^{d-1}$ , then (2.5) also holds for  $v(x) \rightarrow v = 0$  and  $e \in \mathbb{W}$ .

*Proof.* If  $v(x) \rightarrow v \in \mathbb{R}^d \setminus 0$ , then both  $\|v(x)\| \rightarrow \|v\|$  and  $v(x)/\|v(x)\| \rightarrow v/\|v\|$ .

Thus

$$\lim_{x \rightarrow \infty} \frac{\Phi(xv(x), e)}{x} = \lim_{x \rightarrow \infty} \|v(x)\| \frac{\Phi(x\|v(x)\| (v(x)/\|v(x)\|), e)}{x\|v(x)\|} = \|v\| \phi(v/\|v\|, e)$$

which, by (2.4), gives (2.5). The case  $v(x) \rightarrow 0$  follows from (2.3).

**Theorem 2.1.** *Let  $\{X_t : t \in \mathbb{N}_0\}$  be given by (1.1)–(1.2). If Conditions 2.1 and 2.2 hold, then for every integer  $t \geq 0$ , as  $x \rightarrow \infty$ ,*

$$\mathcal{L}\left(\frac{\|X_0\|}{x}, \frac{X_0}{\|X_0\|}, \frac{X_1}{\|X_0\|}, \dots, \frac{X_t}{\|X_0\|} \middle| \|X_0\| > x\right) \rightsquigarrow \mathcal{L}(Y, M_0, M_1, \dots, M_t) \quad (2.6)$$

with

$$M_j = \phi(M_{j-1}, \varepsilon_j), \quad j = 1, 2, \dots, \quad (2.7)$$

and

- (i)  $Y, M_0, \varepsilon_1, \varepsilon_2, \dots$  are independent with  $\varepsilon_t$  as in (1.2)(i);
  - (ii)  $P(Y > y) = y^{-\alpha}$  for  $y \geq 1$ ;
  - (iii)  $\mathcal{L}(M_0) = \Upsilon$ .
- (2.8)

*Proof.* The argument is by induction on  $t$ . The case  $t = 0$  is a straightforward consequence of Condition 2.1. So let  $t$  be a positive integer and let  $f : \mathbb{R} \times (\mathbb{R}^d)^{t+1} \rightarrow \mathbb{R}$  be bounded and continuous. We have to show that

$$\lim_{x \rightarrow \infty} \mathbb{E}\left[f\left(\frac{\|X_0\|}{x}, \frac{X_0}{\|X_0\|}, \dots, \frac{X_t}{\|X_0\|}\right) \middle| \|X_0\| > x\right] = \mathbb{E}[f(Y, M_0, \dots, M_t)]. \quad (2.9)$$

By (1.1), if  $X_0 \neq 0$ ,

$$\frac{X_t}{\|X_0\|} = \frac{\Phi(X_{t-1}, \varepsilon_t)}{\|X_0\|} = \frac{\Phi\left(x \frac{\|X_0\|}{x} \frac{X_{t-1}}{\|X_0\|}, \varepsilon_t\right)}{x \frac{\|X_0\|}{x}}.$$

Hence,

$$\begin{aligned} & \mathbb{E}\left[f\left(\frac{\|X_0\|}{x}, \frac{X_0}{\|X_0\|}, \dots, \frac{X_t}{\|X_0\|}\right) \middle| \|X_0\| > x\right] \\ &= \mathbb{E}\left[g_x\left(\frac{\|X_0\|}{x}, \frac{X_0}{\|X_0\|}, \dots, \frac{X_{t-1}}{\|X_0\|}\right) \middle| \|X_0\| > x\right] \end{aligned} \quad (2.10)$$

where

$$g_x(y, x_0, \dots, x_{t-1}) = \mathbb{E}\left[f\left(y, x_0, \dots, x_{t-1}, \frac{\Phi(xy x_{t-1}, \varepsilon_t)}{xy}\right)\right]. \quad (2.11)$$

Define

$$g(y, x_0, \dots, x_{t-1}) = \mathbb{E}[f(y, x_0, \dots, x_{t-1}, \phi(x_{t-1}, \varepsilon_t))]. \quad (2.12)$$

By (2.7),

$$\mathbb{E}[f(Y, M_0, \dots, M_t)] = \mathbb{E}[g(Y, M_0, \dots, M_{t-1})]. \quad (2.13)$$

In view of the identities (2.10) and (2.13), the limit relation in (2.9) will follow if we can show that

$$\mathbb{E} \left[ g_x \left( \frac{\|X_0\|}{x}, \frac{X_0}{\|X_0\|}, \dots, \frac{X_{t-1}}{\|X_0\|} \right) \middle| \|X_0\| > x \right] \rightarrow \mathbb{E}[g(Y, M_0, \dots, M_{t-1})] \quad (2.14)$$

as  $x \rightarrow \infty$ . In turn, (2.14) will follow from the induction hypothesis and an extension of the continuous mapping theorem (van der Vaart, 1998, Theorem 18.11) provided

$$\lim_{x \rightarrow \infty} g_x(y(x), x_0(x), \dots, x_{t-1}(x)) = g(y, x_0, \dots, x_{t-1}) \quad (2.15)$$

whenever  $y(x) \rightarrow y$  and  $x_i(x) \rightarrow x_i$  as  $x \rightarrow \infty$  with  $(y, x_0, \dots, x_{t-1})$  ranging over a set  $E \subset \mathbb{R} \times (\mathbb{R}^d)^t$  with  $\mathbb{P}((Y, M_0, \dots, M_{t-1}) \in E) = 1$ . From the definitions of  $g_x$  and  $g$  in (2.11) and (2.12), respectively, equation (2.15) is implied by

$$\lim_{x \rightarrow \infty} \frac{\Phi(xw(x), v)}{x} = \phi(w, v) \quad (2.16)$$

whenever  $\lim_{x \rightarrow \infty} w(x) = w$  and where  $w$  and  $v$  range over sets that receive probability one by the distributions of  $M_{t-1}$  and  $\varepsilon_1$ , respectively. Since (2.16) is ensured by Condition 2.2 and Lemma 2.1, the statement follows.

### 3. Backward tail chains

From now on, the process  $\{X_t\}$  in (1.1) and (1.2) is assumed to be strictly stationary. A necessary and sufficient condition for stationarity is that

$$\mathcal{L}(\Phi(X_0, \varepsilon_1)) = \mathcal{L}(X_0). \quad (3.1)$$

It may be highly non-trivial to find the law for  $X_0$  that solves (3.1). But even when the stationary distribution does not admit an explicit expression, its tails may in many cases be found by the theory developed originally in Kesten (1973), Letac (1986) and Goldie (1991). For recent results on specific models, see for instance Klüppelberg and Pergamenchtchikov (2003, 2004), De Saporta et al. (2004), Mirek (2011), Buraczweski et al. (2012), and Collamore and Vidyashankar (2013).

If the process  $\{X_t\}$  is stationary, then by Kolmogorov's extension theorem and changing the probability space if necessary, the range of  $t$  can without loss of generality be assumed to be the set of all integers,  $\mathbb{Z}$ ; recall that we are interested in distributional properties only, not in almost sure properties, for instance.



Our aim is to extend Theorem 2.1 and find the asymptotic distribution of the random vector  $(X_{-s}, \dots, X_t)$  conditionally on  $\|X_0\| > x$  as  $\|x\| \rightarrow \infty$ , for all integer  $s$  and  $t$  (Corollary 5.1). According to Basrak and Segers (2009, Theorem 2.1), if the underlying process is stationary, the existence of a forward tail process ( $t \in \mathbb{N}_0$ ) is enough to guarantee the existence of the tail process as a whole ( $t \in \mathbb{Z}$ ).

**Proposition 3.1.** *Let  $\{X_t : t \in \mathbb{Z}\}$  be a stationary Markov chain with distribution determined by (1.1), (1.2) and (3.1). If Conditions 2.1 and 2.2 hold, then there exists a process  $\{M_t : t \in \mathbb{Z}\}$  such that*

$$\mathcal{L}\left(\frac{X_{-s}}{\|X_0\|}, \dots, \frac{X_0}{\|X_0\|}, \dots, \frac{X_t}{\|X_0\|} \middle| \|X_0\| > x\right) \rightsquigarrow \mathcal{L}(M_{-s}, \dots, M_0, \dots, M_t) \quad (3.2)$$

for all integer  $s, t \geq 0$ .

*Proof.* This follows from our Theorem 2.1 and Theorem 2.1 in Basrak and Segers (2009), combined with a continuous mapping argument.

We call the process  $\{M_t : t \in \mathbb{Z}\}$  the *spectral (tail) process* of  $\{X_t : t \in \mathbb{Z}\}$ , in accordance with the definition of the process  $\{\Theta_t : t \in \mathbb{Z}\}$  in Basrak and Segers (2009).

Basrak and Segers (2009) also state an important property of the limiting process.

**Proposition 3.2.** *Let  $\{X_t : t \in \mathbb{Z}\}$  be a stationary Markov chain with distribution determined by (1.1), (1.2) and (3.1) and spectral process  $\{M_t : t \in \mathbb{Z}\}$ . Then for all  $s, t \geq 0$  and for all bounded and measurable  $f : (\mathbb{R}^d)^{s+t+1} \rightarrow \mathbb{R}$  satisfying  $f(y_{-s}, \dots, y_t) = 0$  whenever  $y_{-s} = 0$ ,*

$$E[f(M_{-s}, \dots, M_t)] = E\left[f\left(\frac{M_0}{\|M_s\|}, \dots, \frac{M_{s+t}}{\|M_s\|}\right) \|M_s\|^\alpha \mathbf{1}_{\{M_s \neq 0\}}\right]. \quad (3.3)$$

*Proof.* It follows directly from our Proposition 3.1 and Theorem 3.1 in Basrak and Segers (2009) that

$$E[f(M_{-s-i}, \dots, M_{t-i})] = E\left[f\left(\frac{M_{-s}}{\|M_i\|}, \dots, \frac{M_t}{\|M_i\|}\right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}}\right]. \quad (3.4)$$

holds for all bounded and continuous  $f : (\mathbb{R}^d)^{t+s+1} \rightarrow \mathbb{R}$  satisfying  $f(y_{-s}, \dots, y_t) = 0$  whenever  $y_0 = 0$  (instead of  $y_{-s} = 0$ ) and all  $i \in \mathbb{Z}$ . We have added the indicator function on the right-hand side for greater clarity. Let  $s, t$  and  $f$  be as in the statement

of the Proposition. Apply (3.4) to the indices  $(\underline{s}, \underline{t}, \underline{i}) = (0, t + s, s)$  to arrive at (3.3); note that  $\underline{s} + 1 + \underline{t} = s + 1 + t$  and that  $f(x_{-\underline{s}}, \dots, x_{\underline{t}}) = 0$  as soon as  $x_0 = 0$ . Thus, for functions  $f$  which are additionally assumed to be continuous, the statement follows directly.

For the general case, set for abbreviation  $\mathbb{A}^* := (\mathbb{R}^d)^{s+t+1} \setminus (\{0\} \times (\mathbb{R}^d)^{s+t})$ . Furthermore, let  $\mu$  denote the restriction of the law of  $(M_{-s}, \dots, M_t)$  to  $\mathbb{A}^*$  and let  $\nu$  denote the measure on  $\mathbb{A}^*$  defined by

$$\nu(f) = E \left[ f \left( \frac{M_{-s}}{\|M_{-s}\|}, \dots, \frac{M_t}{\|M_t\|} \right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}} \right]$$

for all bounded and continuous  $f$  on  $\mathbb{A}^*$ . In order to show (3.3) for general bounded and measurable  $f$  with  $f(y_{-s}, \dots, y_t) = 0$  if  $y_{-s} = 0$  it suffices to show that  $\mu$  and  $\nu$  coincide. The closed sets of  $(\mathbb{R}^d)^{s+t+1}$  which are bounded away from  $\{0\} \times (\mathbb{R}^d)^{s+t}$  are a  $\pi$ -system generating  $\mathbb{B}(\mathbb{A}^*)$ . Indicator functions of closed sets  $A$  can be written as pointwise limits of continuous functions with values in  $[0, 1]$ . If  $A$  is bounded away from  $\{0\} \times (\mathbb{R}^d)^{s+t}$  we can choose these approximating continuous functions in such a way that they vanish on  $\{0\} \times (\mathbb{R}^d)^{s+t}$ . Thus, by dominated convergence  $\mu(A) = \nu(A)$  for all sets  $A$  of a generating  $\pi$ -system and therefore  $\mu = \nu$  on the Borel sets of  $\mathbb{A}^*$  (Billingsley, 1968, Theorem 2.2), which finishes the proof.

By Lemma 2.2 in Basrak and Segers (2009) it follows that the distribution of  $\{M_t : t \in \mathbb{Z}\}$  is uniquely determined by the distribution of  $\{M_t : t \in \mathbb{N}_0\}$  (and  $\alpha > 0$ ). We will use (3.3) to analyze the structure of the spectral process with a special focus on the backward process  $\{M_{-t} : t \in \mathbb{N}_0\}$ . At the heart of the connection between the forward and backward processes is an adjoint relation between the laws of  $(M_0, M_1)$  and  $(M_0, M_{-1})$ , studied next.

#### 4. An adjoint relation between distributions

A special case of the equality (3.3) is

$$E[f(M_{-1}, M_0)] = E \left[ f \left( \frac{M_0}{\|M_1\|}, \frac{M_1}{\|M_1\|} \right) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}} \right] \quad (4.1)$$

for all  $f : (\mathbb{R}^d)^2 \rightarrow \mathbb{R}$  satisfying  $f(y_0, y_1) = 0$  whenever  $y_0 = 0$ . Starting from a given distribution of  $(M_0, M_1)$  we will in the following characterize the distributions of

$(M_{-1}, M_0)$  which satisfy (4.1). For such an adjoint distribution to exist, the distribution  $(M_0, M_1)$  cannot be chosen arbitrarily from the distributions on  $\mathbb{S}^{d-1} \times \mathbb{R}^d$ . We therefore introduce the following set of “admissible” distributions.

**Definition 4.1.** For  $\alpha \in (0, \infty)$ , let  $\mathcal{M}_\alpha = \mathcal{M}_{\alpha,d}$  be the set of all probability measures  $P$  on  $\mathbb{S}^{d-1} \times \mathbb{R}^d$  such that

$$\int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} \mathbf{1}_S(m/\|m\|) \|m\|^\alpha P(ds, dm) \leq P(S \times \mathbb{R}^d) \quad (4.2)$$

for every Borel set  $S \subset \mathbb{S}^{d-1}$ . We call  $\mathcal{M}_\alpha$  the set of *admissible distributions* for  $\alpha > 0$ .

Note that for  $P \in \mathcal{M}_\alpha$  we have

$$\int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} \|m\|^\alpha P(ds, dm) \leq 1.$$

We now make the already mentioned notion of an “adjoint” distribution more concise.

**Definition 4.2.** For  $P \in \mathcal{M}_\alpha$ , define a signed Borel measure  $P^*$  on  $\mathbb{S}^{d-1} \times \mathbb{R}^d$  by

$$P^*(S \times \{0\}) = P(S \times \mathbb{R}^d) - \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} \mathbf{1}_S(m/\|m\|) \|m\|^\alpha P(ds, dm), \quad (4.3)$$

$$P^*(E) = \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} \mathbf{1}_E(m/\|m\|, s/\|m\|) \|m\|^\alpha P(ds, dm), \quad (4.4)$$

for Borel sets  $S \subset \mathbb{S}^{d-1}$  and  $E \subset \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . We call  $P^*$  the *adjoint measure* of  $P$  in  $\mathcal{M}_\alpha$ .

**Lemma 4.1.** *Let  $P \in \mathcal{M}_\alpha$  and let  $P^*$  be as in Definition 4.2.*

(i)  *$P^*$  is a probability measure and the marginal distributions induced by  $P$  and  $P^*$  on  $\mathbb{S}^{d-1}$  are the same.*

(ii) *For every measurable function  $f : \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\}) \rightarrow \mathbb{R}$ ,*

$$\begin{aligned} \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(s^*, m^*) P^*(ds^*, dm^*) \\ = \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(m/\|m\|, s/\|m\|) \|m\|^\alpha P(ds, dm) \end{aligned} \quad (4.5)$$

*in the sense that if one integral exists, then so does the other, and they are the same.*

(iii)  *$P^* \in \mathcal{M}_\alpha$ .*

(iv)  $(P^*)^* = P$ .

*Proof.* (i) By (4.2),  $P^*$  is a nonnegative Borel measure. Let  $S$  be a Borel subset of  $\mathbb{S}^{d-1}$ . We have

$$P^*(S \times \mathbb{R}^d) = P^*(S \times \{0\}) + P^*(S \times (\mathbb{R}^d \setminus \{0\})).$$

Applying (4.3) to the first term on the right-hand side and applying (4.4) with  $E = S \times (\mathbb{R}^d \setminus \{0\})$  to the second term on the right-hand side yields

$$P^*(S \times \mathbb{R}^d) = P(S \times \mathbb{R}^d).$$

It follows that  $P^*$  is a probability measure (take  $S = \mathbb{S}^{d-1}$ ) on  $\mathbb{S}^{d-1} \times \mathbb{R}^d$  inducing the same marginal distribution on  $\mathbb{S}^{d-1}$  as  $P$ .

(ii) By (4.4), equation (4.5) holds for indicator functions  $\mathbf{1}_E$  of Borel subsets  $E$  of  $\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . The extension to general bounded, measurable functions follows from the definition of the integral.

(iii) Let  $S$  be a Borel subset of  $\mathbb{S}^{d-1}$ . We will apply (4.5) to the function

$$f(s, m) = \mathbf{1}_S(m/\|m\|) \|m\|^\alpha \quad \text{for } (s, m) \in \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\}).$$

We find

$$\begin{aligned} & \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} \mathbf{1}_S(m^*/\|m^*\|) \|m^*\|^\alpha P^*(ds^*, dm^*) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(s^*, m^*) P^*(ds^*, dm^*) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(m/\|m\|, s/\|m\|) \|m\|^\alpha P(ds, dm) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} \mathbf{1}_S\left(\frac{s/\|m\|}{\|(s/\|m\|)\|}\right) \|(s/\|m\|)\|^\alpha \|m\|^\alpha P(ds, dm) \\ &= P(S \times (\mathbb{R}^d \setminus \{0\})) \\ &\leq P(S \times \mathbb{R}^d) = P^*(S \times \mathbb{R}^d), \end{aligned}$$

where we applied (i) in the last step.

(iv) Let  $Q = (P^*)^*$ . We already know that  $Q$  is a probability measure on  $\mathbb{S}^{d-1} \times \mathbb{R}^d$ , that  $Q \in \mathcal{M}_\alpha$ , and that the marginal induced by  $Q$  on  $\mathbb{S}^{d-1}$  coincides with the one of  $P^*$  and thus with the one of  $P$ . Let  $f$  be a nonnegative, measurable function on

$\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . Define the nonnegative, measurable function  $g$  on  $\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$  by

$$g(s, m) = f(m/\|m\|, s/\|m\|) \|m\|^\alpha, \quad \text{for } (s, m) \in \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\}).$$

We have

$$\begin{aligned} & g(m/\|m\|, s/\|m\|) \|m\|^\alpha \\ &= f\left(\frac{s/\|m\|}{\|(s/\|m\|)\|}, \frac{m/\|m\|}{\|(m/\|m\|)\|}\right) \|(s/\|m\|)\|^\alpha \|m\|^\alpha = f(s, m). \end{aligned} \quad (4.6)$$

By (4.5) applied first to  $Q$  and  $f$  and then to  $P^*$  and  $g$ , we have

$$\begin{aligned} & \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(s, m) Q(ds, dm) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(m^*/\|m^*\|, s^*/\|m^*\|) \|m^*\|^\alpha P^*(ds^*, dm^*) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} g(s^*, m^*) P^*(ds^*, dm^*) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} g(m/\|m\|, s/\|m\|) \|m\|^\alpha P(ds, dm) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})} f(s, m) P(ds, dm), \end{aligned}$$

where we used (4.6) in the last step. It follows that  $Q$  and  $P$  coincide on  $\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . As  $Q$  and  $P$  also induce the same marginal distributions on  $\mathbb{S}^{d-1}$ , it follows that they must also coincide on  $\mathbb{S}^{d-1} \times \{0\}$ . As a consequence,  $Q$  is equal to  $P$ .

The next lemma shows that the class  $\mathcal{M}_\alpha$  and the adjoint relation on it arise naturally in the context of regularly varying Markov chains.

**Lemma 4.2.** *Let  $\{X_t : t \in \mathbb{Z}\}$  be a stationary Markov chain with distribution determined by (1.1), (1.2) and (3.1). If Conditions 2.1 and 2.2 hold, then  $\mathcal{L}(M_0, M_1)$  belongs to  $\mathcal{M}_\alpha$  and its adjoint is equal to  $\mathcal{L}(M_0, M_{-1})$ .*

*Proof.* To prove admissibility, we have to show that

$$\mathbb{E}[\mathbf{1}_S(M_1/\|M_1\|) \|M_1\|^\alpha] \leq \mathbb{P}(M_0 \in S) \quad (4.7)$$

for every Borel set  $S \subset \mathbb{S}^{d-1}$ . Let  $f$  be a bounded, nonnegative and continuous function on  $\mathbb{S}^{d-1}$ . We will show that

$$\mathbb{E}[f(M_1/\|M_1\|) \|M_1\|^\alpha] \leq \mathbb{E}[f(M_0)]. \quad (4.8)$$

Equation (4.8) implies (4.7) for closed sets  $S$  because the indicator function of a closed set  $S$  can be written as the pointwise limit of a decreasing sequence of continuous functions taking values in the interval  $[0, 1]$ . From this we arrive at (4.7) for an arbitrary Borel set  $S$  by invoking an increasing sequence of closed sets  $S_n$  contained in  $S$  such that  $\mathbb{E}[\mathbf{1}_{S_n}(M_1/\|M_1\|) \|M_1\|^\alpha]$  and  $\mathbb{P}(M_0 \in S_n)$  converge to  $\mathbb{E}[\mathbf{1}_S(M_1/\|M_1\|) \|M_1\|^\alpha]$  and  $\mathbb{P}(M_0 \in S)$  respectively; see for instance Theorem 1.1 on p. 7 in Billingsley (1968).

Let  $\delta > 0$ . By stationarity of  $\{X_t : t \in \mathbb{Z}\}$  and by definition of the spectral process  $\{M_t : t \in \mathbb{Z}\}$ , we have

$$\begin{aligned} \mathbb{E}[f(M_0)] &= \lim_{x \rightarrow \infty} \mathbb{E}[f(X_1/\|X_1\|) \mid \|X_1\| > x] \\ &\geq \limsup_{x \rightarrow \infty} \mathbb{E}[\mathbf{1}_{\{\|X_0\| > \delta x\}} f(X_1/\|X_1\|) \mid \|X_1\| > x] \\ &= \limsup_{x \rightarrow \infty} \frac{\mathbb{P}[\|X_0\| > \delta x]}{\mathbb{P}[\|X_1\| > x]} \mathbb{E}[f(X_1/\|X_1\|) \mathbf{1}_{\{\|X_1\| > x\}} \mid \|X_0\| > \delta x] \\ &= \delta^{-\alpha} \mathbb{E}[f(M_1/\|M_1\|) \mathbf{1}_{\{Y\|M_1\| > \delta^{-1}\}}]. \end{aligned}$$

In the last line,  $Y$  is a Pareto( $\alpha$ ) random variable, independent of  $M_1$ . As  $\mathbb{P}(Y\|M_1\| = \delta^{-1}) = 0$  by continuity of the law of  $Y$ , the last equality in the above display follows from the continuous mapping theorem.

Since the distribution of  $Y^{-\alpha}$  is uniform on the interval  $(0, 1)$ , we have

$$\begin{aligned} \delta^{-\alpha} \mathbb{E}[f(M_1/\|M_1\|) \mathbf{1}_{\{Y\|M_1\| > \delta^{-1}\}}] &= \delta^{-\alpha} \mathbb{E}[f(M_1/\|M_1\|) \mathbf{1}_{\{\delta^\alpha \|M_1\|^\alpha > Y^{-\alpha}\}}] \\ &= \delta^{-\alpha} \mathbb{E}[\mathbb{E}[f(M_1/\|M_1\|) \mathbf{1}_{\{\delta^\alpha \|M_1\|^\alpha > Y^{-\alpha}\}} \mid M_1]] \\ &= \delta^{-\alpha} \mathbb{E}[f(M_1/\|M_1\|) \min(\delta^\alpha \|M_1\|^\alpha, 1)] \\ &= \mathbb{E}[f(M_1/\|M_1\|) \min(\|M_1\|^\alpha, \delta^{-\alpha})]. \end{aligned}$$

We obtain that for every  $\delta > 0$ ,

$$\mathbb{E}[f(M_0)] \geq \mathbb{E}[f(M_1/\|M_1\|) \min(\|M_1\|^\alpha, \delta^{-\alpha})].$$

Take the limit as  $\delta \rightarrow 0$  and apply the monotone convergence theorem to obtain (4.8).

Next we show that the adjoint of  $\mathcal{L}(M_0, M_1)$  is equal to  $\mathcal{L}(M_0, M_{-1})$ . We have to check the two equations

$$\begin{aligned} \mathbb{P}((M_0, M_{-1}) \in S \times \{0\}) &= \mathbb{P}(M_0 \in S) - \mathbb{E}[\mathbf{1}_{\mathbb{R}^d \setminus \{0\}}(M_1) \mathbf{1}_S(M_1/\|M_1\|) \|M_1\|^\alpha], \\ \mathbb{P}((M_0, M_{-1}) \in E) &= \mathbb{E}[\mathbf{1}_{\mathbb{R}^d \setminus \{0\}}(M_1) \mathbf{1}_E(M_1/\|M_1\|, M_0/\|M_1\|) \|M_1\|^\alpha], \quad (4.9) \end{aligned}$$

for all Borel sets  $S \subset \mathbb{S}^{d-1}$  and  $E \subset \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . Since the first component  $M_0$  is common to both laws, it is sufficient to check only the second equation, (4.9).

Set  $f(m_{-1}, m_0) = \mathbf{1}_E(m_0, m_{-1})$  on  $\mathbb{R} \times \mathbb{S}^{d-1}$ . Note that  $f(0, m_0) = 0$ . Apply equation (4.1) to  $f$ :

$$\begin{aligned} \mathbb{P}((M_0, M_{-1}) \in E) &= \mathbb{E}[f(M_{-1}, M_0)] \\ &= \mathbb{E}[f(M_0/\|M_1\|, M_{-1}/\|M_1\|)\|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}] \\ &= \mathbb{E}[\mathbf{1}_{\mathbb{R}^d \setminus \{0\}}(M_1) \mathbf{1}_E(M_{-1}/\|M_1\|, M_0/\|M_1\|)\|M_1\|^\alpha], \end{aligned}$$

which gives (4.9), as required.

**Remark 4.1.** The determination of the adjoint measure is particularly simple for probability measures  $P$  such that

$$\int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} \|m\|^\alpha P(ds, dm) = 1, \quad (4.10)$$

since in this case  $P^*(\mathbb{S}^{d-1} \times \{0\}) = 0$  by (4.3) and  $P^*$  is completely described by (4.4).

Definition 4.2 and Lemma 4.1 generalize Proposition 3.1 in Segers (2007) to the multivariate case. Examples 3.2–3.4 in the latter reference illustrate the adjoint relation for laws on  $\{-1, +1\} \times \mathbb{R}$ . We conclude the section with a multivariate example.

**Example 4.1.** Let  $\alpha > 0$  and let  $P$  be the law of  $(I, RQI)$  with  $I$ ,  $R$  and  $Q$  independent,  $I$  taking values in  $\mathbb{S}^{d-1}$ ,  $R$  a positive random variable with  $\mathbb{E}[R^\alpha] = 1$ , and  $Q$  a random orthogonal  $d \times d$  matrix, that is  $Q' = Q^{-1}$  a.s.; also assume that the laws of  $I$  and  $QI$  are the same (cf. also Example 6.1). One verifies easily that  $P \in \mathcal{M}_\alpha$  and that (4.10) holds, so that the adjoint law  $P^*$  is concentrated on  $\mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ . It may thus be derived from (4.4) that for Borel sets  $S \subset \mathbb{S}^{d-1}$  and  $T \subset \mathbb{R}^d \setminus \{0\}$ ,

$$P^*(S \times T) = \mathbb{E}[\mathbf{1}_S(QI) \mathbf{1}_T(I/R) R^\alpha]. \quad (4.11)$$

It follows that  $P^*$  is the law of  $(I^*, R^*Q^*I^*)$ , with  $I^*$ ,  $R^*$  and  $Q^*$  independent,  $\mathcal{L}(I^*) = \mathcal{L}(I)$ ,  $\mathcal{L}(Q^*) = \mathcal{L}(Q')$ , and the law of  $R^* > 0$  given by  $\mathbb{E}[f(R^*)] = \mathbb{E}[f(1/R) R^\alpha]$  for measurable functions  $f$  on  $(0, \infty)$ .

### 5. Back-and-forth tail chains and the spectral process

In this section, we will analyze a certain class of discrete-time processes which are constructed from a pair of adjoint distributions. We will see that this class of processes fulfills equation (3.3) for all  $i, s, t \in \mathbb{Z}$  with  $s \leq 0 \leq t$ .

**Definition 5.1.** A  $d$ -dimensional discrete-time process  $\{M_t : t \in \mathbb{Z}\}$  is called a *back-and-forth tail chain* with index  $\alpha > 0$ , notation  $\text{BFTC}(\alpha)$ , if the following properties hold:

- (i)  $\mathcal{L}(M_0, M_1)$  and  $\mathcal{L}(M_0, M_{-1})$  belong to  $\mathcal{M}_\alpha$  and are adjoint;
- (ii) the forward process  $\{M_t : t \in \mathbb{N}_0\}$  is a Markov chain with respect to the filtration  $\sigma(M_s, -\infty < s \leq t)$ ,  $t \geq 0$ , and the Markov kernel satisfies

$$\begin{aligned} & \mathbb{P}(M_t \in \cdot \mid M_{t-1} = x_{t-1}) \\ &= \begin{cases} \delta_0(\cdot) & \text{if } x_{t-1} = 0, \\ \mathbb{P}(\|x_{t-1}\| M_1 \in \cdot \mid M_0 = x_{t-1}/\|x_{t-1}\|) & \text{if } x_{t-1} \neq 0; \end{cases} \end{aligned}$$

- (iii) the backward process  $\{M_t : t \in \mathbb{N}_0\}$  is a Markov chain with respect to the filtration  $\sigma(M_{-s}, -\infty < s \leq t)$ ,  $t \geq 0$ , and the Markov kernel satisfies

$$\begin{aligned} & \mathbb{P}(M_{-t} \in \cdot \mid M_{-t+1} = x_{-t+1}) \\ &= \begin{cases} \delta_0(\cdot) & \text{if } x_{-t+1} = 0, \\ \mathbb{P}(\|x_{-t+1}\| M_{-1} \in \cdot \mid M_0 = x_{-t+1}/\|x_{-t+1}\|) & \text{if } x_{-t+1} \neq 0. \end{cases} \end{aligned}$$

Clearly,  $\{M_t : t \in \mathbb{Z}\}$  is a  $\text{BFTC}(\alpha)$  if and only if  $\{M_{-t} : t \in \mathbb{Z}\}$  is a  $\text{BFTC}(\alpha)$ . The distribution of a  $\text{BFTC}(\alpha)$  is completely determined by an admissible law of  $(M_0, M_1)$  (and  $\alpha > 0$ ).

The fact that the distributions  $P = \mathcal{L}(M_0, M_1)$  and  $P^* = \mathcal{L}(M_0, M_{-1})$  are adjoint in  $\mathcal{M}_\alpha$  implies that for every measurable function  $f : \mathbb{R}^d \times \mathbb{S}^{d-1} \rightarrow \mathbb{R}$  such that  $f(0, s) = 0$  for all  $s \in \mathbb{S}^{d-1}$ , we have

$$\begin{aligned} \mathbb{E}[f(M_{-1}, M_0)] &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \times \{0\})} f(m, s) P^*(ds, dm) \\ &= \int_{\mathbb{S}^{d-1} \times (\mathbb{R}^d \times \{0\})} f(s/\|m\|, m/\|m\|) \|m\|^\alpha P(ds, dm) \end{aligned}$$



$$= \mathbb{E} \left[ f \left( \frac{M_0}{\|M_1\|}, \frac{M_1}{\|M_1\|} \right) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}} \right], \quad (5.1)$$

in the sense that if one expectation exists, then so does the other, the two expectations being equal. This corresponds to equation (4.1) which originally motivated the definition of an adjoint distribution. The above formula is the special case  $s = 1$  and  $t = 0$  of the following result.

**Proposition 5.1.** *Let  $\{M_t : t \in \mathbb{Z}\}$  be a BFTC( $\alpha$ ). For all integer  $s, t \geq 0$  and for all measurable functions  $f : (\mathbb{R}^d)^{s+1+t} \rightarrow \mathbb{R}$  vanishing on  $\{0\} \times (\mathbb{R}^d)^{s+t}$ , the  $s+1$  numbers*

$$\mathbb{E} \left[ f \left( \frac{M_{-s+i}}{\|M_i\|}, \dots, \frac{M_{t+i}}{\|M_i\|} \right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}} \right], \quad i = 0, \dots, s, \quad (5.2)$$

are all the same, in the sense that if one integral exists, then they all exist and they are equal.

*Proof.* For  $s = 0$  there is nothing to prove, so assume that  $s \geq 1$ . By definition of the integral, it is sufficient to consider the case where  $f$  is nonnegative, in which case the expectations in (5.2) are always well-defined, possibly equal to infinity.

*Reduction to the case  $i \in \{0, 1\}$ .* Suppose first that we can show that the numbers corresponding to  $i = 0$  and  $i = 1$  in (5.2) are equal, that is (note that  $\|M_0\| = 1$ ),

$$\mathbb{E}[f(M_{-s}, \dots, M_t)] = \mathbb{E} \left[ f \left( \frac{M_{-s+1}}{\|M_1\|}, \dots, \frac{M_{t+1}}{\|M_1\|} \right) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}} \right]. \quad (5.3)$$

Take arbitrary  $i = 0, \dots, s-1$ . Note that

$$\mathbb{E} \left[ f \left( \frac{M_{-s+i}}{\|M_i\|}, \dots, \frac{M_{t+i}}{\|M_i\|} \right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}} \right] = \mathbb{E}[g(M_{-s+i}, \dots, M_{t+i})]$$

for a measurable function  $g : (\mathbb{R}^d)^{s+1+t} \rightarrow \mathbb{R}$  with that vanishes as soon as its first  $d$ -tuple of arguments is zero. By (5.3) applied to  $\tilde{s} = s - i$  and  $\tilde{t} = t + i$ , we find

$$\mathbb{E}[g(M_{-s+i}, \dots, M_{t+i})] = \mathbb{E} \left[ g \left( \frac{M_{-s+i+1}}{\|M_1\|}, \dots, \frac{M_{t+i+1}}{\|M_1\|} \right) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}} \right].$$

By definition of  $g$ , if  $M_1 \neq 0$ , then

$$\begin{aligned} & g \left( \frac{M_{-s+i+1}}{\|M_1\|}, \dots, \frac{M_{t+i+1}}{\|M_1\|} \right) \\ &= f \left( \frac{M_{-s+i+1}/\|M_1\|}{\|(M_{i+1}/\|M_1\|)\|}, \dots, \frac{M_{t+i+1}/\|M_1\|}{\|(M_{i+1}/\|M_1\|)\|} \right) \|(M_{i+1}/\|M_1\|)\|^\alpha \mathbf{1}_{\{M_{i+1} \neq 0\}} \\ &= f \left( \frac{M_{-s+i+1}}{\|M_{i+1}\|}, \dots, \frac{M_{t+i+1}}{\|M_{i+1}\|} \right) \frac{\|M_{i+1}\|^\alpha}{\|M_1\|^\alpha} \mathbf{1}_{\{M_{i+1} \neq 0\}}. \end{aligned}$$

Combine the previous three displays to see that

$$\begin{aligned} & \mathbb{E} \left[ f \left( \frac{M_{-s+i}}{\|M_i\|}, \dots, \frac{M_{t+i}}{\|M_i\|} \right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}} \right] \\ &= \mathbb{E} \left[ f \left( \frac{M_{-s+i+1}}{\|M_{i+1}\|}, \dots, \frac{M_{t+i+1}}{\|M_{i+1}\|} \right) \|M_{i+1}\|^\alpha \mathbf{1}_{\{M_1 \neq 0, M_{i+1} \neq 0\}} \right]. \end{aligned}$$

By definition of the forward chain  $(M_t)_{t \geq 0}$ , we have  $M_{i+1} = 0$  as soon as  $M_1 = 0$ . As a consequence, we may suppress the event  $\{M_1 \neq 0\}$  in the indicator function on the right-hand side, and thus

$$\begin{aligned} & \mathbb{E} \left[ f \left( \frac{M_{-s+i}}{\|M_i\|}, \dots, \frac{M_{t+i}}{\|M_i\|} \right) \|M_i\|^\alpha \mathbf{1}_{\{M_i \neq 0\}} \right] \\ &= \mathbb{E} \left[ f \left( \frac{M_{-s+i+1}}{\|M_{i+1}\|}, \dots, \frac{M_{t+i+1}}{\|M_{i+1}\|} \right) \|M_{i+1}\|^\alpha \mathbf{1}_{\{M_{i+1} \neq 0\}} \right]. \end{aligned}$$

We conclude that in order to show (5.2), it is enough to show (5.3). We will show (5.3) by induction on  $s \geq 1$ .

*Proof of (5.3) if  $s = 1$ .* We have to show that

$$\mathbb{E}[f(M_{-1}, \dots, M_t)] = \mathbb{E} \left[ f \left( \frac{M_0}{\|M_1\|}, \dots, \frac{M_{t+1}}{\|M_1\|} \right) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}} \right]. \quad (5.4)$$

We will proceed by induction on  $t \geq 0$ .

The case  $t = 0$  is nothing more than the adjoint relation between the laws of  $(M_0, M_1)$  and  $(M_0, M_{-1})$ , see (5.1).

Let  $t \geq 1$  and let (5.4) be fulfilled for  $t - 1$ . By the Markov property,

$$\mathbb{E}[f(M_{-1}, \dots, M_t)] = \mathbb{E}[g(M_{-1}, \dots, M_{t-1})]$$

with

$$g(m_{-1}, \dots, m_{t-1}) = \mathbb{E}\{f(m_{-1}, \dots, m_{t-1}, M_t) \mid M_{t-1} = m_{t-1}\}$$

As  $g(0, m_0, \dots, m_{t-1}) = 0$ , we can apply the induction hypothesis, yielding

$$\mathbb{E}[g(M_{-1}, \dots, M_{t-1})] = \mathbb{E}[g(M_0/\|M_1\|, \dots, M_t/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}].$$

The defining property of a BFTC implies that for every  $c > 0$ , for every integer  $r \geq 1$  and for every nonnegative, measurable function  $h$  on  $\mathbb{R}^d$ ,

$$\mathbb{E}[h(cM_r) \mid M_{r-1} = m/c] = \begin{cases} h(0) & \text{if } m = 0, \\ \mathbb{E}[h(\|m\|M_1) \mid M_0 = m/\|m\|] & \text{if } m \neq 0, \end{cases} \quad (5.5)$$

the right-hand side not depending on the scaling constant  $c$  nor on the time index  $r$ .

It follows that if  $m_1 \neq 0$ ,

$$\begin{aligned} & g(m_0/\|m_1\|, \dots, m_t/\|m_1\|) \\ &= \mathbb{E}[f(m_0/\|m_1\|, \dots, m_t/\|m_1\|, M_t) \mid M_{t-1} = m_t/\|m_1\|] \\ &= \mathbb{E}[f(m_0/\|m_1\|, \dots, m_t/\|m_1\|, M_{t+1}/\|m_1\|) \mid M_t = m_t]. \end{aligned}$$

We find that, on the event  $\{M_1 \neq 0\}$ , by the Markov property,

$$\begin{aligned} & g(M_0/\|M_1\|, \dots, M_t/\|M_1\|) \\ &= \mathbb{E}[f(M_0/\|M_1\|, \dots, M_t/\|M_1\|, M_{t+1}/\|M_1\|) \mid M_0, \dots, M_t]. \end{aligned}$$

We can conclude that

$$\begin{aligned} & \mathbb{E}[f(M_{-1}, \dots, M_t)] \\ &= \mathbb{E}[g(M_{-1}, \dots, M_{t-1})] \\ &= \mathbb{E}[g(M_0/\|M_1\|, \dots, M_t/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}] \\ &= \mathbb{E}[f(M_0/\|M_1\|, \dots, M_t/\|M_1\|, M_{t+1}/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}], \end{aligned}$$

as required.

*Proof of (5.3) for general  $s \geq 1$ .* The case  $s = 1$  was treated above. So let  $s \geq 2$ .

By the Markov property, we have

$$\mathbb{E}[f(M_{-s}, \dots, M_t)] = \mathbb{E}[g(M_{-s+1}, \dots, M_t)]$$

with  $g : (\mathbb{R}^d)^{s+t} \rightarrow \mathbb{R}$  a nonnegative, measurable function defined by

$$g(m_{-s+1}, \dots, m_t) = \mathbb{E}\{f(M_{-s}, m_{-s+1}, \dots, m_t) \mid M_{-s+1} = m_{-s+1}\}.$$

Conditionally on  $M_{-s+1} = 0$ , we have  $M_{-s} = 0$ , and thus  $f(M_{-s}, \dots) = 0$  too. It follows that  $g(0, m_{-s+2}, \dots, m_t) = 0$ . By the induction hypothesis, we therefore have

$$\mathbb{E}[g(M_{-s+1}, \dots, M_t)] = \mathbb{E}[g(M_{-s+2}/\|M_1\|, \dots, M_{t+1}/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}].$$

As for the forward chain in (5.5), we have for every nonnegative, measurable function  $h$  on  $\mathbb{R}^d$  and every  $c > 0$ ,

$$\mathbb{E}[h(cM_{-r}) \mid M_{-r+1} = m/c] = \begin{cases} h(0) & \text{if } m = 0, \\ \mathbb{E}[h(\|m\|M_{-1}) \mid M_0 = m/\|m\|] & \text{if } m \neq 0, \end{cases} \quad (5.6)$$

the right-hand side not depending on the scaling constant  $c > 0$  nor on the time index  $r = 1, 2, \dots$ . It follows that for  $m_1 \neq 0$ , we have

$$\begin{aligned} & g(m_{-s+2}/\|m_1\|, \dots, m_{t+1}/\|m_1\|) \\ &= \mathbb{E}[f(M_{-s}, m_{-s+2}/\|m_1\|, \dots, m_{t+1}/\|m_1\|) \mid M_{-s+1} = m_{-s+2}/\|m_1\|] \\ &= \mathbb{E}[f(M_{-s+1}/\|m_1\|, m_{-s+2}/\|m_1\|, \dots, m_{t+1}/\|m_1\|) \mid M_{-s+2} = m_{-s+2}]. \end{aligned}$$

Invoking the Markov property again, we conclude that

$$\begin{aligned} \mathbb{E}[f(M_{-s}, \dots, M_t)] &= \mathbb{E}[g(M_{-s+1}, \dots, M_t)] \\ &= \mathbb{E}[g(M_{-s+2}/\|M_1\|, \dots, M_{t+1}/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}] \\ &= \mathbb{E}[f(M_{-s+1}/\|M_1\|, \dots, M_{t+1}/\|M_1\|) \|M_1\|^\alpha \mathbf{1}_{\{M_1 \neq 0\}}], \end{aligned}$$

as required. This concludes the proof of Proposition 5.1.

The following proposition connects BFTCs and spectral processes.

**Proposition 5.2.** *Let  $\{Y_t : t \in \mathbb{Z}\}$  be an  $\mathbb{R}^d$ -valued process and let  $\{M_t : t \in \mathbb{Z}\}$  be an  $\mathbb{R}^d$ -valued BFTC( $\alpha$ ). If*

$$\mathcal{L}(Y_0, \dots, Y_t) = \mathcal{L}(M_0, \dots, M_t) \quad (5.7)$$

for all  $t \geq 0$  and if

$$\mathbb{E} \left[ f(Y_{-s}, \dots, Y_t) \right] = \mathbb{E} \left[ f \left( \frac{Y_0}{\|Y_s\|}, \dots, \frac{Y_{s+t}}{\|Y_s\|} \right) \|Y_s\|^\alpha \mathbf{1}_{\{Y_s \neq 0\}} \right] \quad (5.8)$$

for all  $s, t \geq 0$  and for all bounded and measurable  $f : (\mathbb{R}^d)^{s+t+1} \rightarrow \mathbb{R}$  satisfying  $f(y_{-s}, \dots, y_t) = 0$  whenever  $y_{-s} = 0$ , then

$$\mathcal{L}(Y_{-s}, \dots, Y_t) = \mathcal{L}(M_{-s}, \dots, M_t) \quad (5.9)$$

for all  $s, t \geq 0$ .

*Proof.* The proof relies on the fact that both the process  $\{Y_t : t \in \mathbb{Z}\}$  which satisfies (5.8) and the BFTC( $\alpha$ ) are uniquely determined by their forward process. Our proof is by induction on  $s$ . For  $s = 0$ , equation (5.9) is equal to the assumption (5.7) for all  $t \geq 0$ . For the induction step, assume that (5.9) holds for a fixed value of  $\tilde{s} = s - 1 \geq 0$  and all  $t \geq 0$ . Let  $f : (\mathbb{R}^d)^{s+t+1} \rightarrow \mathbb{R}$  be a bounded continuous function. Write

$$f(y_{-s}, \dots, y_t) = f_1(y_{-s}, \dots, y_t) + f_2(y_{-s}, \dots, y_t)$$

with

$$\begin{aligned} f_1(y_{-s}, \dots, y_t) &= f(0, y_{-s+1}, \dots, y_t), \\ f_2(y_{-s}, \dots, y_t) &= f(y_{-s}, y_{-s+1}, \dots, y_t) - f(0, y_{-s+1}, \dots, y_t). \end{aligned}$$

and note that  $f_2(0, y_{-s+1}, \dots, y_t) = 0$ , while the value of  $f_1$  does not depend on the first coordinate of the argument. Then

$$\begin{aligned} & \mathbb{E}[f(Y_{-s}, \dots, Y_t)] \\ &= \mathbb{E}[f_1(Y_{-s}, \dots, Y_t)] + \mathbb{E}[f_2(Y_{-s}, \dots, Y_t)] \\ &= \mathbb{E}[f_1(Y_{-s}, \dots, Y_t)] + \mathbb{E}\left[f_2\left(\frac{Y_0}{\|Y_s\|}, \dots, \frac{Y_{s+t}}{\|Y_s\|}\right) \|Y_s\|^\alpha \mathbf{1}_{\{Y_s \neq 0\}}\right] \\ &= \mathbb{E}[f_1(M_{-s}, \dots, M_t)] + \mathbb{E}\left[f_2\left(\frac{M_0}{\|M_s\|}, \dots, \frac{M_{s+t}}{\|M_s\|}\right) \|M_s\|^\alpha \mathbf{1}_{\{M_s \neq 0\}}\right], \end{aligned}$$

where both the induction hypothesis and equations (5.8) and (5.9) have been used. Since  $\{M_t : t \in \mathbb{Z}\}$  is a BFTC( $\alpha$ ), we may apply Proposition 5.1 for  $i = s$  and  $i = 0$  (note that  $\|M_0\| = 1$ ), so that the above expression is equal to

$$\mathbb{E}[f_1(M_{-s}, \dots, M_t)] + \mathbb{E}[f_2(M_{-s}, \dots, M_t)] = \mathbb{E}[f(M_{-s}, \dots, M_t)],$$

which finishes the induction step and the proof.

**Remark 5.1.** Proposition 5.2 can be read in the following way: Every spectral process  $\{M_t : t \in \mathbb{Z}\}$  with a forward process (meaning:  $\{M_t : t \in \mathbb{N}_0\}$ ) which has a BFTC( $\alpha$ ) structure, automatically has a BFTC( $\alpha$ )-backward-distribution as well. This means that a Markovian structure in the forward spectral process (which may also arise in settings where the underlying process is non-Markovian) is enough to secure a Markovian structure of the backward spectral process as well.

**Corollary 5.1.** *Let  $\{X_t : t \in \mathbb{Z}\}$  be a stationary Markov chain with distribution determined by (1.1), (1.2) and (3.1). Then the corresponding spectral process  $\{M_t : t \in \mathbb{Z}\}$  is a BFTC( $\alpha$ ).*

*Proof.* The existence of a corresponding spectral process follows from Proposition 3.1. Furthermore, it follows from Theorem 2.1 that the forward process  $\{M_t : t \in \mathbb{N}_0\}$  is equal in law to the forward process of a BFTC( $\alpha$ ). By Proposition 5.2 the statement follows.

In the univariate case, BFTCs have an additional structure which generalizes a multiplicative random walk in that the distribution of the increment depends on the sign of the process in its current state (Segers, 2007). The random walk structure of the forward tail chain was first observed in Smith (1992) for one-sided extremes and extended to allow for both positive and negative extremes in Bortot and Coles (2003).

## 6. Examples for BFTCs

We conclude the paper with some examples of BFTCs for multivariate Markov processes. For univariate examples, see Segers (2007, section 7).

**Example 6.1.** Let  $(A_t, B_t), t \in \mathbb{Z}$ , be i.i.d. with  $A_t \in \mathbb{R}^{d \times d}$  and  $B_t \in \mathbb{R}^d$ . The stationary distribution and asymptotic behavior of the corresponding random difference equation

$$X_t = A_t X_{t-1} + B_t, \quad t \in \mathbb{Z}, \quad (6.1)$$

have been studied initially in the seminal work by Kesten (1973). Let us assume that the distribution of  $(A_t, B_t)$  satisfies the technical, but mild assumptions of Theorems A and B (nonnegative case, i.e. all components of  $X_t, t \in \mathbb{N}_0$ , are nonnegative almost surely) or Theorem 6 (general case) in Kesten (1973). Together with results in Boman and Lindskog (2009) this implies that the stationary distribution of  $X_t$  for (6.1) is multivariate regularly varying in the nonnegative case. In the general case, multivariate regular variation follows if  $\kappa_1 > 0$  in Kesten (1973), Equation (4.8), is not an integer, cf. Basrak et al. (2002). Let  $\Upsilon$  denote the spectral measure and  $\alpha > 0$  the index of regular variation of the stationary distribution of  $X_t$ . In this case, it can be shown that

$$\mathbb{E} \left[ f \left( \frac{AI}{\|AI\|} \right) \|AI\|^\alpha \right] = \mathbb{E}[f(I)] \quad (6.2)$$

for all bounded, continuous functions  $f$  on  $\mathbb{S}^{d-1}$ , where  $I \in \mathbb{S}^{d-1}$  has distribution  $\Upsilon$  and  $A \in \mathbb{R}^{d \times d}$  is independent of  $I$  with  $\mathcal{L}(A) = \mathcal{L}(A_1)$ , cf. Basrak and Segers (2009).

Due to the linear structure of (6.1), Theorem 2.1 applies with  $P(Y > y) = y^{-\alpha}, y > 1, \mathcal{L}(M_0) = \Upsilon$  and  $\phi(M_{j-1}, \epsilon_j) = \epsilon_j M_{j-1}$  where the  $\epsilon_j \in \mathbb{R}^{d \times d}, j = 1, 2, \dots$ , are i.i.d. with  $\mathcal{L}(\epsilon_j) = \mathcal{L}(A_1)$ . In order to find the distribution of the backward tail chain note that Remark 4.1 applies to this example by equation (6.2). So the law  $P^*$  of  $(M_0, M_{-1})$

is given by

$$P^*(E) = \mathbb{E} \left[ \mathbf{1}_E \left( \frac{AI}{\|AI\|}, \frac{I}{\|AI\|} \right) \|AI\|^\alpha \right] \quad (6.3)$$

for all Borel sets  $E \subset \mathbb{S}^{d-1} \times \mathbb{R}^d$ .

Additional assumptions about  $\mathcal{L}(A)$  allow us to simplify this characterization: Let us assume that  $A$  has a multiplicative form like in Example 4.1, i.e.  $A = RQ$  for a positive random variable  $R$  with  $\mathbb{E}[R^\alpha] = 1$  and  $Q$  is an orthogonal matrix independent of  $R$ . We may additionally assume that both  $R$  and  $Q$  have a density on  $\mathbb{R}_+$  and  $\mathbb{S}^{d-1}$  respectively and that  $\mathbb{E}[\|B_1\|^\alpha] < \infty$  which guarantees that all assumptions of Theorem 6 in Kesten (1973) are met. The spectral measure  $\Upsilon$  is then the uniform distribution on  $\mathbb{S}^{d-1}$  and  $\alpha > 0$  is the index of regular variation since

$$\mathbb{E} \left[ f \left( \frac{AI}{\|AI\|} \right) \|AI\|^\alpha \right] = \mathbb{E}[f(QI)R^\alpha] = \mathbb{E}[f(QI)]\mathbb{E}[R^\alpha] = \mathbb{E}[f(I)]$$

holds for all bounded, continuous functions  $f$  on  $\mathbb{S}^{d-1}$  if  $I \sim \text{Unif}(\mathbb{S}^{d-1})$ . Since  $\mathcal{L}(I) = \mathcal{L}(QI)$ , all assumptions of Example 4.1 are met and the adjoint measure  $P^*$  is determined by (4.11) and equal to the law of  $(I^*, R^*Q^*I^*)$  with  $R^*, Q^*, I^*$  independent,  $\mathcal{L}(I^*) = \text{Unif}(\mathbb{S}^{d-1})$ ,  $\mathcal{L}(Q^*) = \mathcal{L}(Q')$  and  $R^*$  has density  $f_{R^*}(y) = f_R(y^{-1})y^{-(2+\alpha)}$ ,  $y > 0$ , where  $f_R$  denotes the density of  $R$ . Thus, both the forward and the backward tail chain have a simple multiplicative structure:

$$M_t = M_0 A_1 \cdot \dots \cdot A_t, \quad M_{-t} = M_0 A_{-1} \cdot \dots \cdot A_{-t}, \quad t \geq 1,$$

with  $A_1, A_2, \dots$  as above and  $A_{-1}, A_{-2}, \dots$  i.i.d. with the same distribution as  $R^*Q^*$ , all independent of each other and of  $M_0 \sim \text{Unif}(\mathbb{S}^{d-1})$ .

**Example 6.2.** While the preceding example dealt with random difference equations where the random increment  $B_t$  has a relatively light tail [Kesten (1973) assumes that  $\mathbb{E}(\|B_1\|^\alpha) < \infty$ ], the following example deals with AR(1) processes where the innovations themselves are regularly varying. Let

$$X_t = AX_{t-1} + B_t, \quad t \in \mathbb{Z}, \quad (6.4)$$

where  $A$  is a deterministic  $\mathbb{R}^{d \times d}$ -matrix and  $B_t \in \mathbb{R}^d$ ,  $t \in \mathbb{Z}$ , are i.i.d. and multivariate regularly varying with index  $\alpha > 0$  and spectral measure  $\lambda$  on  $\mathbb{S}^{d-1}$ . For extensions to random but light-tailed random matrices  $A_t$ , see for instance Hult and Samorodnitsky (2008).

If  $\sup_{x \in \mathbb{S}^{d-1}} \|A^m x\| < 1$  for some positive integer  $m$ , then (6.4) has the stationary solution

$$X_t = \sum_{n=0}^{\infty} A^n B_{t-n}, \quad t \in \mathbb{Z}.$$

It has been shown in Meinguet and Segers (2010) that in this case the stationary distribution of  $X_t$  is multivariate regularly varying as well, with the same index  $\alpha$  and spectral measure  $\Upsilon = \sum_{n=0}^{\infty} p_n \lambda_n$ , where

$$p_n := \frac{c_n}{\sum_{k=0}^{\infty} c_k} \quad \text{with} \quad c_n := \int_{\mathbb{S}^{d-1}} \|A^n \theta\| \lambda(d\theta), \quad n \in \mathbb{N}_0,$$

and where  $\lambda_n$  is the spectral measure of  $A^n B$ , provided  $c_n > 0$ , i.e.

$$\lambda_n(f) := \frac{1}{c_n} \int_{\mathbb{S}^{d-1}} f \left( \frac{A^n s}{\|A^n s\|} \right) \|A^n s\|^\alpha \lambda(ds), \quad n \in \mathbb{N}_0, \quad \text{if } c_n > 0,$$

for all bounded, continuous functions  $f$  on  $\mathbb{S}^{d-1}$  (Meinguet and Segers, 2010, Example 9.3). The spectral process  $\{M_t : t \in \mathbb{Z}\}$  in Proposition 3.1 is of the form

$$M_{-N+t} = \begin{cases} A^t \Theta, & t = 0, 1, 2, \dots, \\ 0, & t = -1, -2, \dots \end{cases} \quad (6.5)$$

for a random integer  $N$  with  $P(N = n) = p_n$ ,  $n \in \mathbb{N}_0$ , and a random vector  $\Theta$  with distribution

$$P(\Theta \in E \mid N = n) = \frac{1}{c_n} \int_{\mathbb{S}^{d-1}} \mathbf{1}_E(s/\|A^n s\|) \|A^n s\|^\alpha \lambda(ds)$$

for  $n \in \mathbb{N}_0$  and Borel sets  $E \in \mathbb{R}^d$ . Here, the forward tail chain has a deterministic multiplicative structure with  $M_0 \sim \Upsilon$  and  $M_n = A M_{n-1}$  for  $n \geq 1$ . The backward process is Markovian as well, by Corollary 5.1. This is also clear if one looks at (6.5) and notices that  $M_{-(n+h)} = 0$  if  $M_{-n} = 0$  for all  $h \geq 1, n \geq 1$ . Furthermore, if  $M_{-n} \neq 0$  then  $(M_{-n+1}, \dots, M_0) = (A M_{-n}, \dots, A^n M_{-n})$  contains no more information about  $M_{-(n+1)}$  than  $M_{-n}$  does.

The distribution of  $(M_0, M_{-1})$  is adjoint to the one of  $(M_0, M_1) = (M_0, A M_0)$ . By (4.4) and since  $M_0 \sim \Upsilon$ , we find, for every Borel set  $E \subset \mathbb{S}^{d-1} \times (\mathbb{R}^d \setminus \{0\})$ ,

$$\begin{aligned} P((M_0, M_{-1}) \in E) &= E \left[ \mathbf{1}_E \left( \frac{M_1}{\|M_1\|}, \frac{M_0}{\|M_1\|} \right) \|M_1\|^\alpha \right] \\ &= \frac{1}{\sum_{k=0}^{\infty} c_k} \sum_{n \geq 0} \int_{\mathbb{S}^{d-1}} \mathbf{1}_E \left( \frac{A^{n+1} s}{\|A^{n+1} s\|}, \frac{A^n s}{\|A^{n+1} s\|} \right) \|A^{n+1} s\|^\alpha \lambda(ds). \end{aligned}$$



Choosing  $E = S \times (\mathbb{R}^d \setminus \{0\})$  for a Borel set  $S \subset \mathbb{S}^{d-1}$  yields, upon taking complements with respect to  $\{M_0 \in S\}$  and noting that  $\|s\| = 1$  for  $s \in \mathbb{S}^{d-1}$ ,

$$P(M_0 \in S, M_{-1} = 0) = \frac{1}{\sum_{k=0}^{\infty} c_k} \lambda(S). \quad (6.6)$$

In particular,  $P(M_{-1} = 0) = p_0 = P(N = 0)$ . The backward tail chain now follows from Definition 5.1(iii) together with the distribution of  $(M_0, M_{-1})$ .

In the special case that  $A$  is invertible, we find from (6.5) that  $M_{-(t+1)}$  is equal to either  $A^{-1}M_{-t}$  or to 0 with conditional probabilities depending on  $M_{-t}/\|M_{-t}\|$ : if  $M_{-t} = 0$ , then  $M_{-(t+1)} = 0$  too, while if  $M_{-t} = x \neq 0$ , then

$$M_{-(t+1)} = \begin{cases} A^{-1}x & \text{with probability } 1 - P(M_{-1} = 0 \mid M_0 = x/\|x\|), \\ 0 & \text{with probability } P(M_{-1} = 0 \mid M_0 = x/\|x\|). \end{cases}$$

To derive a concrete form of the backward Markov kernel, let us assume that  $\lambda$  has a Lebesgue density  $f_\lambda$  on  $\mathbb{S}^{d-1}$ . Then all measures  $\lambda_n$  and thus  $\Upsilon$  have Lebesgue densities as well and (6.6) gives us

$$P(M_{-1} = 0 \mid M_0 = s) = \frac{1}{\sum_{k=0}^{\infty} c_k} \frac{f_\lambda(s)}{f_\Upsilon(s)}$$

for all  $s \in \mathbb{S}^{d-1}$  such that  $f_\Upsilon(s) > 0$ .

### Acknowledgement

The authors thank Richard Davis and Holger Drees for helpful discussions. Furthermore, they wish to thank the organisers, especially Paul Doukhan, of the workshop ‘‘Extremes and risk management’’ which took place during September 2012 at the university of Cergy-Pontoise. Anja Janßen was supported by DFG (DFG project JA 2160/1-1). Johan Segers was supported by contract ‘‘Projet d’Actions de Recherche Concertées’’ No. 12/17-045 of the ‘‘Communauté française de Belgique’’ and by IAP research network grant No. P7/06 of the Belgian government (Belgian Science Policy).

### References

BASRAK, B., DAVIS, R. A. and MIKOSCH, T. (2002). A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

- BASRAK, B., DAVIS, R. A. and MIKOSCH, T. (2002). Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–115.
- BASRAK, B. and SEGERS, J. (2009). Regularly varying multivariate time series. *Stoch. Proc. Appl.* **119**, 1055–1080.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*, Wiley, New York.
- BOMAN, J. and LINDSKOG, F. (2009). Support theorems for the Radon transform and Cramér–Wold theorems, *J. Theoret. Probab.* **22**, 683–710.
- BORTOT, P. and COLES, S. G. (2000). A sufficiency property arising from the characterization of extremes of Markov chains. *Bernoulli* **6**, 183–190.
- BORTOT, P. and COLES, S. G. (2003). Extremes of Markov chains with tail switching potential. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65**, 851–867.
- BURACZEWSKI, D., DAMEK, E. and MIREK, M. (2012). Asymptotics of stationary solutions of multivariate stochastic recursions with heavy tailed inputs and related limit theorems. *Stoch. Proc. Appl.* **122**, 42–67.
- COLLAMORE, J. F. and VIDYASHANKAR, A. N. (2013). Tail estimates for stochastic fixed point equations via nonlinear renewal theory. *Stoch. Proc. Appl.*, forthcoming. [arXiv:1103.2317](https://arxiv.org/abs/1103.2317)
- COLES, S. G., SMITH, R. L., and TAWN, J. A. (1997). A seasonal Markov model for extremely low temperatures. *Environmetrics* **5**, 221–239.
- EMBRECHTS, P., KLÜPPELBERG, C., and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin.
- GOLDIE, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.
- GOMES, I., DE HAAN, L. and PESTANA, D. (2004). Joint exceedances of the ARCH process. *J. Appl. Prob.* **41**, 919–926.

- DE HAAN, L., RESNICK, S. I., ROOTZÉN, H. and de VRIES, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213–224.
- HULT, H. and SAMORODNITSKY, G. (2008). Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli* **14**, 838–864.
- KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.
- KIFER, Y. (1986). *Ergodic Theory of Random Transformations*, Birkhäuser, Boston.
- KLÜPPELBERG, C. and PERGAMENCHTCHIKOV, S. (2003). Renewal theory for functionals of a Markov chain with compact state space. *Ann. Probab.* **31**, 2270–2300.
- KLÜPPELBERG, C. and PERGAMENCHTCHIKOV, S. (2004). The tail of the stationary distribution of a random coefficient AR( $q$ ) model. *Ann. Appl. Probab.* **14**, 971–1005.
- LETAC, G. (1986). A contraction principle for certain Markov chains and its applications. *Random matrices and their applications: Proceedings* (Brunswick, Maine, 1984), 263–273. *Contemp. Math.*, **50**, Amer. Math. Soc., Providence, RI.
- MEINGUET, T. and SEGERS, J. (2010). Regularly varying time series in Banach spaces. *Arxiv eprint*, <http://arxiv.org/abs/1001.3262>
- MIREK, M. (2011). Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. *Probab. Theory Relat. Fields* **151**, 705–734.
- PERFEKT, R. (1994). Extremal behaviour of stationary Markov chains with applications. *Ann. Appl. Probab.* **4**, 529–548.
- PERFEKT, R. (1997). Extreme value theory for a class of Markov chains with values in  $\mathbb{R}^d$ . *Adv. Appl. Prob.* **29**, 138–164.
- RESNICK, S. I. (2007). *Heavy-Tail Phenomena. Probabilistic and Statistical Modeling*, Springer, New York.

- RESNICK, S. I. and ZEBER, D. (2013). Asymptotics of Markov kernels and the tail chain. *Adv. Appl. Prob.*, **45**, 186–213.
- DE SAPORTA, B., GUIVARC'H, Y., and LE PAGE, E. (2004). On the multidimensional stochastic equation  $Y(n+1) = A(n)Y(n) + B(n)$ . *C. R. Acad. Sci.* **339**, 499–502.
- SEGERS, J. (2007). Multivariate regular variation of heavy-tailed Markov chains. Institut de statistique, Université catholique de Louvain, Discussion Paper 0703, arXiv:0701411.
- SMITH, R. L. (1992). The extremal index for a Markov chain. *J. Appl. Prob.* **29**, 37–45.
- SMITH, R. L., TAWN, J. A., and COLES, S. G. (1997). Markov chain models for threshold exceedances. *Biometrika* **84**, 249–268.
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge.
- YUN, S. (1998). The extremal index of a higher-order stationary Markov chain. *Ann. Appl. Probab.* **8**, 408–437.
- YUN, S. (2000). The distributions of cluster functionals of extreme events in a  $d$ th-order Markov chain. *J. Appl. Prob.* **37**, 29–44.
- ZIVOT, E. (2009). Practical issues in the analysis of univariate GARCH models. In *Handbook of Financial Time Series*, ed. T. G. Andersen, R. A. Davis, J.-P. Kreiß, T. Mikosch. Springer, Berlin, pp. 113–155.