Extremes (2018) 21:261–284
https://doi.org/10.1007/s10687-017-0307-3

On the tail behavior of a class of multivariate conditionally heteroskedastic processes

Rasmus Søndergaard Pedersen[1] ·
Olivier Wintenberger[2,3]

Received: 31 January 2017 / Revised: 27 November 2017 / Accepted: 29 November 2017 /
Published online: 11 December 2017
© Springer Science+Business Media, LLC, part of Springer Nature 2017

Abstract Conditions for geometric ergodicity of multivariate autoregressive conditional heteroskedasticity (ARCH) processes, with the so-called BEKK (Baba, Engle, Kraft, and Kroner) parametrization, are considered. We show for a class of BEKK-ARCH processes that the invariant distribution is regularly varying. In order to account for the possibility of different tail indices of the marginals, we consider the notion of vector scaling regular variation (VSRV), closely related to non-standard regular variation. The characterization of the tail behavior of the processes is used for deriving the asymptotic properties of the sample covariance matrices.

Keywords Stochastic recurrence equations · Markov processes · Regular variation · Multivariate ARCH · Asymptotic properties · Geometric ergodicity

AMS 2000 Subject Classifications 60G70 · 60G10 · 60H25 · 39A50

✉ Rasmus Søndergaard Pedersen
rsp@econ.ku.dk

1 Department of Economics, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark

2 Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

3 Université Pierre et Marie Curie, LSTA, 4 Place Jussieu, 75005 Paris, France

1 Introduction

The aim of this paper is to investigate the tail behavior of a class of multivariate conditionally heteroskedastic processes. Specifically, we consider the BEKK-ARCH (or BEKK(1,0,l)) process, introduced by Engle and Kroner (1995), satisfying

$$X_t = H_t^{1/2} Z_t, \quad t \in \mathbb{N} \tag{1.1}$$

$$H_t = C + \sum_{i=1}^{l} A_i X_{t-1} X_{t-1}^{\intercal} A_i^{\intercal}, \tag{1.2}$$

with $(Z_t : t \in \mathbb{N})$ $i.i.d.$, $Z_t \sim N(0, I_d)$, C a $d \times d$ positive definite matrix, $A_1, ..., A_l \in M(d, \mathbb{R})$ (the set of $d \times d$ real matrices), and some initial value X_0. Due to the assumption that Z_t is Gaussian, it holds that X_t can be written as the stochastic recurrence equation (SRE)

$$X_t = \tilde{M}_t X_{t-1} + Q_t, \tag{1.3}$$

with

$$\tilde{M}_t = \sum_{i=1}^{l} m_{it} A_i \tag{1.4}$$

and $(m_{it} : t \in \mathbb{N})$ is an $i.i.d.$ process mutually independent of $(m_{jt} : t \in \mathbb{N})$ for $i \neq j$, with $m_{it} \sim N(0, 1)$. Moreover $(Q_t : t \in \mathbb{N})$ is an $i.i.d.$ process with $Q_t \sim N(0, C)$ mutually independent of $(m_{it} : t \in \mathbb{N})$ for all $i = 1, ..., l$.

To our knowledge, the representation in Eqs. 1.3-1.4 of the BEKK-ARCH process is new. Moreover, the representation will be crucial for studying the stochastic properties of the process. Firstly, we find a new sufficient condition in terms of the matrices $A_1, ..., A_l$ in order for $(X_t : t \geq 0)$ to be geometrically ergodic. In particular, for the case $l = 1$, we derive a condition directly related to the eigenvalues of A_1, in line with the strict stationarity condition found by Nelson (1990) for the univariate ARCH(1) process. This condition is milder compared to the conditions found in the existing body of literature on BEKK-type processes. Secondly, the representation is used to characterize the tails of the stationary solution to $(X_t : t \in \mathbb{N})$.

Whereas the tail behavior of univariate GARCH processes is well-established, see e.g. Basrak et al. (2002b), few results on the tail behavior of multivariate GARCH processes exist. Some exceptions are the multivariate constant conditional correlation (CCC) GARCH processes, see e.g. Stărică (1999), Pedersen (2016), and Matsui and Mikosch (2016), and a class of factor GARCH processes, see Basrak and Segers (2009). This existing body of literature relies on rewriting the (transformed) process on companion form that obeys a non-negative multivariate SRE. The characterization of the tails of the processes then follows by an application of Kesten's Theorem (Kesten 1973) for non-negative SREs. Such approach is not feasible when analyzing BEKK-ARCH processes, as these are stated in terms of an $\mathbb{R}^d$-valued SRE in Eq. 1.3. For some special cases of the BEKK-ARCH process, we apply existing results for $\mathbb{R}^d$-valued SREs in order to show that the stationary distribution for the BEKK-ARCH process is multivariate regularly varying. Specifically, when the matrix $\tilde{M}_t$ in Eq. 1.4 is invertible (almost surely) and has a law that is absolutely continuous with respect to the Lebesgue measure on $M(d, \mathbb{R})$ (denoted ID BEKK-ARCH) we argue that the classical results of Kesten (1973, Theorem 6), see also Alsmeyer and Mentemeier (2012), apply. Moreover, when $\tilde{M}_t$ is the product of a positive scalar and a random orthogonal matrix (denoted Similarity BEKK-ARCH) we show that the results of Buraczewski et al. (2009) apply. Importantly, we do also argue that the results of Alsmeyer and Mentemeier (2012) rely on rather restrictive conditions that can be shown not to hold for certain types of BEKK-ARCH processes, in particular the much applied process where $l = 1$ and A_1 is diagonal, denoted Diagonal BEKK-ARCH. Specifically, and as ruled out in Alsmeyer and Mentemeier (2012), we show that the Diagonal BEKK-ARCH process exhibits different marginal tail indices, i.e. $\mathbb{P}(\pm X_{t,i} > x)/c_i x^{-\alpha_i} \to 1$ as $x \to \infty$ for some constant $c_i > 0$, $i = 1, ..., d$ (denoted Condition **M**). In order to analyze this class of BEKK-ARCH processes, where the tail indices are allowed to differ among the elements of X_t, we introduce a new notion of vector scaling regular variation (VSRV) distributions, based on element-wise scaling of X_t instead of scaling by an arbitrary norm of X_t. We emphasize that the notion of VSRV is similar to the notion of non-standard regular variation (see Resnick (2007, Chapter 6)) under the additional Condition **M**. In addition, in the spirit of Basrak and Segers (2009), we introduce the notion of VSRV processes with particular attention to Markov chains and characterize their extremal behavior. We argue that the stationary distribution of the Diagonal BEKK-ARCH process is expected to be VSRV, which is supported in a simulation study. Proving that the VSRV property applies requires that new multivariate renewal theory is developed, and we leave such task for future research.

The rest of the paper is organized as follows. In Section 2, we state sufficient conditions for geometric ergodicity of the BEKK-ARCH process and introduce the notion of vector-scaling regular varying (VSRV) distributions. We show that the distribution of X_t satisfies this type of tail-behavior, under suitable conditions. In Section 3 we introduce the notion of VSRV processes and state that certain BEKK-ARCH processes satisfy this property. Moreover, we consider the extremal behavior of the process, in terms of the asymptotic behavior of maxima and extremal indices. Lastly, we consider the convergence of point processes based on VSRV processes. In Section 4, we consider the limiting distribution of the sample covariance matrix of X_t, which relies on point process convergence. Section 5 contains some concluding remarks on future research directions.

Notation: Let $GL(d, \mathbb{R})$ denote the set of $d \times d$ invertible real matrices. With $M(d, \mathbb{R})$ the set of $d \times d$ real matrices and $A \in M(d, \mathbb{R})$, let $\rho(A)$ denote the spectral radius of A. With $\otimes$ denoting the Kronecker product, for any real matrix A let $A^{\otimes p} = A \otimes A \otimes \cdots \otimes A$ (p factors). For two matrices, A and B, of the same dimension, $A \odot B$ denotes the elementwise product of A and B. Unless stated otherwise, $\|\cdot\|$ denotes an arbitrary matrix norm. Moreover, $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. For two matrices A and B of the same dimensions, $A \nleq B$ means that $A_{ij} > B_{ij}$ for some i, j.

For two positive functions f and g, $f(x) \sim g(x)$, if $\lim_{x\to\infty} f(x)/g(x) = 1$. Let $\mathcal{L}(X)$ denote the distribution of X. By default, the mode of convergence for distributions is weak convergence.

2 Stationary solution of the BEKK-ARCH model

2.1 Existence and geometric ergodicity

We start out by stating the following theorem that provides a sufficient condition for geometric ergodicity of the BEKK-ARCH process. To our knowledge, this result together with Proposition 2.3 below are new.

Theorem 2.1 *Let X_t satisfy* (1.1)-(1.2). *With $\tilde{M}_t$ defined in* Eq. 1.4, *suppose that*

$$\inf_{n\in\mathbb{N}} \left\{ \frac{1}{n} \mathbb{E}\left[\log\left(\left\| \prod_{t=1}^{n} \tilde{M}_t \right\| \right) \right] \right\} < 0. \tag{2.1}$$

Then $(X_t : t = 0, 1, ...)$ is geometrically ergodic, and for the associated stationary solution, $\mathbb{E}[\|X_t\|^s] < \infty$ for some $s > 0$.

The proof of the theorem follows by Alsmeyer (2003, Theorems 2.1-2.2, Example 2.6.d, and Theorem 3.2) and is hence omitted.

Remark 2.2 A sufficient condition for the existence of finite higher-order moments of X_t can be obtained from Theorem 5 of Feigin and Tweedie (1985). In particular, if $\rho(E[\tilde{M}_t^{\otimes 2n}]) < 1$ for some $n \in \mathbb{N}$, then, for the strictly stationary solution, $E[\|X_t\|^{2n}] < \infty$. For example, $\rho(\sum_{i=1}^{l} A_i^{\otimes 2}) < 1$ implies that $E[\|X_t\|^2] < \infty$. This result complements Theorem C.1 of Pedersen and Rahbek (2014) that contains conditions for finite higher-order moments for the case $l = 1$.

For the case where $\tilde{M}_t$ contains only one term, i.e. $l = 1$, the condition in Eq. 2.1 simplifies and a condition for geometric ergodicity can be stated explicitly in terms of the eigenvalues of the matrix A_1:

Proposition 2.3 *Let X_t satisfy* (1.1)-(1.2) *with $l = 1$ and let $A := A_1$. Then a necessary and sufficient condition for* Eq. 2.1 *is that*

$$\rho(A) < \exp\left\{ \frac{1}{2} \left[-\psi(1) + \log(2) \right] \right\} = 1.88736..., \tag{2.2}$$

where $\psi(\cdot)$ is the digamma function.

Proof The condition (2.1) holds if and only if there exists $n \in \mathbb{N}$ such that

$$\mathbb{E}\left[\log\left(\left\| \prod_{t=1}^{n} \tilde{M}_t \right\| \right) \right] < 0. \tag{2.3}$$

Let $m_t := m_{1t}$. It holds that

$$\begin{aligned}\mathbb{E}\left[\log\left(\left\|\prod_{t=1}^{n}\tilde{M}_t\right\|\right)\right] &= \mathbb{E}\left[\log\left(\left\|A^n\prod_{t=1}^{n}m_t\right\|\right)\right]\\ &= \log\left(\left\|A^n\right\|\right) - n\mathbb{E}\left[-\log(|m_t|)\right]\\ &= \log\left(\left\|A^n\right\|\right) - n\left\{\frac{1}{2}\left[-\psi(1)+\log(2)\right]\right\},\end{aligned}$$

and hence (2.3) is satisfied if

$$\log\left(\left\|A^n\right\|^{1/n}\right) < \frac{1}{2}\left[-\psi(1)+\log(2)\right].$$

The result now follows by observing that $\|A^n\|^{1/n} \to \rho(A)$ as $n \to \infty$. □

Remark 2.4 It holds that $\rho(A^{\otimes 2}) = (\rho(A))^2$. Hence the condition in Eq. 2.2 is equivalent to

$$\rho(A^{\otimes 2}) < \exp\{-\psi(1)+\log(2)\} = \frac{1}{2}\exp\left[-\psi\left(\frac{1}{2}\right)\right] = 3.56...,$$

which is similar to the strict stationary condition found for the ARCH coefficient of the univariate ARCH(1) process with Gaussian innovations; see Nelson (1990).

Boussama et al. (2011) derive sufficient conditions for geometric ergodicity of the GARCH-type BEKK process, where $H_t = C + \sum_{i=1}^{p} A_i X_{t-i} X_{t-i}^{\intercal} A_i^{\intercal} + \sum_{j=1}^{q} B_j H_{t-j} B_j^{\intercal}$, $A_i, B_j \in M(d,\mathbb{R})$, $i = 1, ..., p$, $j = 1, ..., q$. Specifically, they show that a sufficient condition is $\rho(\sum_{i=1}^{p} A_i^{\otimes 2} + \sum_{j=1}^{q} B_j^{\otimes 2}) < 1$. Setting $p = 1$ and $q = 0$, this condition simplifies to $\rho(A_1^{\otimes 2}) < 1$, which is stronger than the condition derived in Eq. 2.2.

Below, we provide some examples of BEKK-ARCH processes that are geometrically ergodic and that will be studied in detail throughout this paper.

Example 2.5 (**ID BEKK-ARCH**) Following Alsmeyer and Mentemeier (2012), we consider BEKK processes with corresponding SRE's satisfying certain irreducibility and density conditions (ID), that is conditions **(A4)-(A5)** in Appendix Section A1. Specifically, we consider the *bivariate* BEKK-ARCH process in Eqs. 1.1–1.2 with

$$H_t = C + \sum_{i=1}^{4} A_i X_{t-1} X_{t-1}^{\intercal} A_i^{\intercal},$$

where

$$A_1 = \begin{pmatrix} a_1 & 0 \\ 0 & 0 \end{pmatrix} \quad A_2 = \begin{pmatrix} 0 & 0 \\ a_2 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & a_3 \\ 0 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 0 & a_4 \end{pmatrix} \tag{2.4}$$

and

$$a_1, a_2, a_3, a_4 \neq 0. \tag{2.5}$$

Writing X_t as an SRE, we obtain

$$X_t = \tilde{M}_t X_{t-1} + Q_t, \tag{2.6}$$

with

$$\tilde{M}_t = \sum_{i=1}^{4} A_i m_{it} \tag{2.7}$$

where $(m_{1t}), (m_{2t}), (m_{3t}), (m_{4t})$ are mutually independent i.i.d. processes with $m_{it} \sim N(0, 1)$. Assuming that a_1, a_2, a_3, a_4 are such that the top Lyapunov exponent of $(\tilde{M}_t)$ is strictly negative, we have that the process is geometrically ergodic.

Notice that one could consider a more general d-dimensional process with the same structure as in Eqs. 2.4–2.7, but with $\tilde{M}_t$ containing d^2 terms such that $\tilde{M}_t$ has a Lebesgue density on $M(d, \mathbb{R})$, as clarified in Example 2.10 below. Moreover, one could include additional terms to $\tilde{M}_t$, say a term containing a full matrix A or an autoregressive term, as presented in Remark 2.8 below. We will focus on the simple bivariate process, but emphasize that our results apply to more general processes.

Example 2.6 (**Similarity BEKK-ARCH**) Consider the BEKK process in Eqs. 1.1–1.2 with $l = 1$ and $A := A_1 = aO$, where a is a positive scalar and O is an orthogonal matrix. This implies that the SRE (1.3) has $\tilde{M}_t = am_t O$. By definition, $\tilde{M}_t$ is a *similarity* with probability one, where we recall that a matrix is a similarity if it can be written as a product of a positive scalar and an orthogonal matrix. From Proposition 2.3, we have that if $a < \exp\left\{(1/2)\left[-\psi(1) + \log(2)\right]\right\} = 1.88736...$, then the process is geometrically ergodic. An important process satisfying the similarity property is the well-known scalar BEKK-ARCH process, where $H_t = C + aX_{t-1}X_{t-1}^{\intercal}$, $a > 0$. Here $A = \sqrt{a}I_d$, with I_d the identity matrix.

Example 2.7 (**Diagonal BEKK-ARCH**) Consider the BEKK-ARCH process in Eqs. 1.1-1.2 with $l = 1$ such that $A := A_1$ is diagonal. We refer to this process as the Diagonal BEKK-ARCH process. Relying on Proposition 2.3, the process is geometrically ergodic, if each diagonal element of A is less than $\exp\left\{(1/2)\left[-\psi(1) + \log(2)\right]\right\} = 1.88736...$ in modulus.

As discussed in Bauwens et al. (2006), diagonal BEKK models are typically used in practice, e.g. within empirical finance, due to their relatively simple parametrization. As will be shown below, even though the parametrization is simple, the tail behavior is rather rich in the sense that each marginal of X_t has different tail indices, in general.

Remark 2.8 As an extension to Eqs. 1.1-1.2, one may consider the autoregressive BEKK-ARCH (AR BEKK-ARCH) process

$$X_t = A_0 X_{t-1} + H_t^{1/2} Z_t, \quad t \in \mathbb{N}$$
$$H_t = C + \sum_{i=1}^{l} A_i X_{t-1} X_{t-1}^{\intercal} A_i^{\intercal},$$

with $A_0 \in M(\mathbb{R}, d)$. This process has recently been studied and applied by Nielsen and Rahbek (2014) for modelling the term structure of interest rates. Notice that the process has the SRE representation

$$X_t = \tilde{M}_t X_{t-1} + Q_t, \quad \tilde{M}_t = A_0 + \sum_{i=1}^{l} m_{it} A_i.$$

Following the arguments used for proving Theorem 2.1, it holds that the AR BEKK-ARCH process is geometrically ergodic if condition (2.1) is satisfied. Interestingly, as verified by simulations in Nielsen and Rahbek (2014) the Lyapunov condition may hold even if the autoregressive polynomial has unit roots, i.e. if $A_0 = I_d + \Pi$, where $\Pi \in M(\mathbb{R}, d)$ has reduced rank.

2.2 Multivariate regularly varying distributions

The stationary solution of the BEKK-ARCH process (see Theorem 2.1) can be written as

$$X_t = \sum_{i=0}^{\infty} \prod_{j=1}^{i} \tilde{M}_{t-j+1} Q_{t-i}, \quad t \in \mathbb{Z}. \tag{2.8}$$

Even if the random matrices $\tilde{M}_t$ are light-tailed under the Gaussian assumption, the maximum of the products $(\prod_{t=1}^{T} \tilde{M}_t)_{T \geq 0}$ may exhibit heavy tails when $T \to \infty$. More precisely, the tails of the stationary distribution are suspected to have an extremal behavior as a power law function: For any $u \in \mathbb{S}^{d-1}$,

$$\mathbb{P}(u^\intercal X_0 > x) \sim C(u) x^{-\alpha(u)}, \qquad x \to \infty, \tag{2.9}$$

with $\alpha(u) > 0$ and $C(u_0) > 0$ for some $u_0 \in \mathbb{S}^{d-1}$. The cases where $\alpha(u) = \alpha$ and $C(u) > 0$ for all $u \in \mathbb{S}^{d-1}$ are referred as Kesten's cases, because of the seminal paper (Kesten 1973), and are the subject of the monograph by Buraczewski et al. (2016). A class of multivariate distributions satisfying this property is the class of multivariate regularly varying distributions (de Haan and Resnick 1977):

Definition 2.9 Let $\bar{\mathbb{R}}_0^d := \bar{\mathbb{R}}^d \setminus \{0\}$, $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$, and $\bar{\mathcal{B}}_0^d$ be the Borel σ-field of $\bar{\mathbb{R}}_0^d$. For an $\mathbb{R}^d$-valued random variable X and some constant scalar $x > 0$, define $\mu_x(\cdot) := \mathbb{P}(x^{-1}X \in \cdot)/\mathbb{P}(\|X\| > x)$. Then X and its distribution are *multivariate regularly varying* if there exists a non-null Radon measure μ on $\bar{\mathcal{B}}_0^d$ which satisfies

$$\mu_x(\cdot) \to \mu(\cdot) \qquad \text{vaguely, as } x \to \infty. \tag{2.10}$$

For any μ-continuity set C and $t > 0$, $\mu(tC) = t^{-\alpha}\mu(C)$, and we refer to α as the index of regular variation.

We refer to de Haan and Resnick (1977) for the notion of vague convergence and additional details. Below, we provide two examples of multivariate regularly varying BEKK processes.

Example 2.10 (**ID BEKK-ARCH**, continued) Consider the ID BEKK-ARCH process (2.4)-(2.7) from Example 2.5. By verifying conditions **(A1)-(A7)** of Theorem 1.1 of Alsmeyer and Mentemeier (2012), stated in Appendix Section A.1 we establish that the process is multivariate regularly varying.

Since $(m_{1t}, m_{2t}, m_{3t}, m_{4t})$ and Q_t are Gaussian, we have that **(A1)-(A2)** hold. Moreover,

$$\tilde{M}_t = \begin{pmatrix} a_1 m_{1t} & a_3 m_{3t} \\ a_2 m_{2t} & a_4 m_{4t} \end{pmatrix} \tag{2.11}$$

is invertible with probability one, which ensures that **(A3)** is satisfied. From Eq. 2.11 we also notice that the distribution of $\tilde{M}_t$ has a Lebesgue density on $M(d, \mathbb{R})$ which is strictly positive in a neighborhood of I_2. This ensures that the irreducibility and density conditions **(A4)-(A5)** are satisfied. The fact that $Q_t \sim N(0, C)$ and independent of $\tilde{M}_t$ implies that condition **(A6)** holds. Lastly, condition **(A7)** holds by the fact that $(m_{1t}, m_{2t}, m_{3t}, m_{4t})$ and Q_t are Gaussian. By Theorem 1.1 of Alsmeyer and Mentemeier (2012) we have established the following proposition:

Proposition 2.11 *Let X_t satisfy (2.4)-(2.7) such that the top Lyapunov exponent of $(\tilde{M}_t)$ is strictly negative. Then for the stationary solution (X_t), there exists $\alpha > 0$ such that*

$$\lim_{t\to\infty} t^{\alpha}\mathbb{P}(x^{\intercal} X_0 > t) = K(x), \quad x \in \mathbb{S}^1, \tag{2.12}$$

for some finite, positive, and continuous function K on $\mathbb{S}^1$.

The proposition implies that each marginal of the distribution of X_0 is regularly varying of order α. By Theorem 1.1.(ii) of Basrak et al. (2002a), we conclude that X_0 is multivariate regularly varying whenever α is a non-integer. Moreover, since X_0 is symmetric, the multivariate regular variation does also hold if α is an odd integer, see Remark 4.4.17 in Buraczewski et al. (2016).

The proposition does also apply if $a_1 = 0$ or $a_4 = 0$. This can be seen by observing that $\prod_{k=1}^{n} \tilde{M}_k$ has a strictly positive density on $M(d, \mathbb{R})$ for n sufficiently large, which is sufficient for establishing conditions **(A4)-(A5)**.

Example 2.12 (**Similarity BEKK-ARCH**, continued) The Similarity BEKK-ARCH, introduced in Example 2.6, fits into the setting of Buraczewski et al. (2009), see also Section 4.4.10 of Buraczewski et al. (2016). Specifically, using the representation $\tilde{M}_t = a|m_t|\text{sign}(m_t)O$, we have that

(i) $\mathbb{E}[\log(|m_t a|)] < 0$ if $a < \exp\{(1/2)\left[-\psi(1) + \log(2)\right]\}$,
(ii) $\mathbb{P}(\tilde{M}_t x + Q_t = x) < 1$ for any $x \in \mathbb{R}^d$, and
(iii) $\log(|a m_t|)$ has a non-arithmetic distribution.

Then, due to Theorem 1.6 of Buraczewski et al. (2009), we have the following proposition:

Proposition 2.13 *Let X_t satisfy (1.1)–(1.2) with $l = 1$ such that $A := A_1 = aO$, where $a > 0$ and O is an orthogonal matrix. If $a < \exp\{(1/2)[-\psi(1) + \log(2)]\} = 1.88736...$, then the process has a unique strictly stationary solution (X_t) with X_t multivariate regularly varying with index $\alpha > 0$ satisfying $E[(|m_t|a)^\alpha] = 1$.*

In the following example, we clarify that the Diagonal BEKK-ARCH process, introduced in Example 2.7, does not satisfy the conditions of Theorem 1.1 of Alsmeyer and Mentemeier (2012). Moreover, we argue that the marginals may have different tail indices, which motivates the notion of *vector scaling regular variation*, introduced in the next section.

Example 2.14 (**Diagonal BEKK-ARCH**, continued) Consider the diagonal BEKK-ARCH process in Example 2.7, i.e. Eqs. 1.1–1.2 with $l = 1$ such that $A := A_1$ is diagonal, $m_t := m_{1t}$, and $M_t := \tilde{M}_t = m_t A$. For this process, the distribution of M_t is too restricted to apply the results by Alsmeyer and Mentemeier (2012), as in Example 2.10. Specifically, the irreducibility condition **(A4)** in Appendix A.1 can be shown not to hold, as clarified next. It holds that

$$\mathbb{P}\left\{\|x^\intercal \prod_{k=1}^{n} M_k\|^{-1}\left(x^\intercal \prod_{k=1}^{n} M_k\right) \in U\right\} = \mathbb{P}\left\{|\prod_{k=1}^{n} m_k|^{-1}\|x^\intercal A^n\|^{-1}\left(\prod_{k=1}^{n} m_k\right) x^\intercal A^n \in U\right\}$$
$$= \mathbb{P}\left\{\text{sign}\left(\prod_{k=1}^{n} m_k\right)\|x^\intercal A^n\|^{-1} x^\intercal A^n \in U\right\}.$$

Hence for any $x \in \mathbb{S}^{d-1}$ we can always find a non-empty open $U \subset \mathbb{S}^{d-1}$ such that

$$\max_{n\in\mathbb{N}} \mathbb{P}\left\{\text{sign}\left(\prod_{k=1}^{n} m_k\right)\|x^\intercal A^n\|^{-1} x^\intercal A^n \in U\right\} = 0. \tag{2.13}$$

As an example, for $d = 2$, choose $x = (1, 0)^\intercal$. Then $\|x^\intercal A^n\|^{-1} x^\intercal A^n \in \{(-1, 0)\} \cup \{(1, 0)\}$ for any $n \in \mathbb{N}$. We conclude that condition **(A4)** does not hold for the diagonal BEKK-ARCH process.

Note that, each element of $X_t = (X_{t,1}, ..., X_{t,d})^\intercal$ of the diagonal BEKK-ARCH process can be written as an SRE,

$$X_{t,i} = A_{ii} m_t X_{t-1,i} + Q_{t,i}, \qquad t \in Z, \qquad i = 1, \ldots, d.$$

By Theorem 4.1 of Goldie (1991), the stationary solution of the marginal equation exists if and only if $\mathbb{E}[\log(|A_{ii} m_0|)] < 0$. In that case there exists a unique $\alpha_i > 0$ such that $\mathbb{E}[|m_0|^{\alpha_i}] = |A_{ii}|^{-\alpha_i}$ and

$$\mathbb{P}(\pm X_{0,i} > x) \sim c_i x^{-\alpha_i} \qquad \text{where} \qquad c_i = \frac{\mathbb{E}[|X_{1,i}|^{\alpha_i} - |A_{ii} m_1 X_{0,i}|^{\alpha_i}]}{2\alpha_i \mathbb{E}[|A_{ii} m_1|^{\alpha_i} \log(|A_{ii} m_1|)]}.$$

Hence each marginal of X_0 may in general have different tail indices. More precisely, the tail indices are different if the diagonal elements of A, i.e. the A_{ii}s, are, and the heaviest marginal tail index α_{i_0} corresponds to the largest diagonal coefficient $A_{i_0 i_0}$. When i_0 is unique, i.e. $\alpha_{i_0} < \alpha_i$ for all $i = 1, ..., d$ except $i \neq i_0$, the distribution X_0

can be considered as multivariate regularly varying with index α_{i_0} and with a limit measure μ with degenerate marginals $i \neq i_0$.

2.3 Vector scaling regularly varying distributions

The previous Example 2.14 shows that the Diagonal BEKK-ARCH process fits into the case where $\alpha(u)$ in Eq. 2.9 is non-constant. Such cases have not attracted much attention in the existing body of literature. However, recent empirical studies, such as Matsui and Mikosch (2016), see also Damek et al. (2017), may suggest that it is more realistic to consider different marginal tail behaviors when modelling multidimensional financial observations. The idea is to use a vector scaling instead of the scaling $\mathbb{P}(\|X\| > x)$ in Definition 2.9 that reduced the regular variation properties of the vector X to the regular variation properties of the norm $\|X\|$ only. More precisely, let (X_t) be a stationary process in $\mathbb{R}^d$ and let $x = (x_1, \ldots, x_d)^\intercal \in \mathbb{R}^d$. Denote also $x^{-1} = (x_1^{-1}, \ldots, x_d^{-1})^\intercal$.

In our framework, we consider distributions satisfying the following condition:

Condition **M** Each marginal of X_0 is regularly varying of order $\alpha_i > 0, i = 1, ..., d$. The slowly varying functions $\ell_i(t) \to c_i > 0$ as $t \to \infty, i = 1, ..., d$.

Indeed, the Diagonal BEKK-ARCH process introduced in Example 2.14 satisfies Condition **M**. Moreover, any regularly varying distribution satisfying the Kesten property (2.9) satisfies Condition **M**. In particular, the ID and Similarity BEKK-ARCH processes, introduced in Examples 2.5 and 2.6 respectively, satisfy Condition **M**.

We introduce the notion of vector scaling regular variation as the nonstandard regular variation of the book of Resnick (2007) under Condition **M**, extended to negative components (Resnick 2007, Sections 6.5.5-6.5.6):

Definition 2.15 The distribution of the vector X_0 is vector scaling regularly varying (VSRV) if and only if it satisfies Condition **M** and it is non-standard regularly varying, i.e. there exists a normalizing sequence $x(t)$ and a Radon measure μ with non-null marginals such that

$$t\mathbb{P}(x(t)^{-1} \odot X_0 \in \cdot) \to \mu(\cdot), \qquad \text{vaguely.} \tag{2.14}$$

The usual way of analyzing non-standard regularly varying vectors is to consider a componentwise normalization that is standard regularly varying in the sense of Definition 2.9. Specifically, when $X_0 = (X_{0,1}, ..., X_{0,d})^\intercal$ satisfies Definition 2.15, $(c_1^{-1}(X_{0,1}/|X_{0,1}|)|X_{0,1}|^{\alpha_1}, ..., c_d^{-1}(X_{0,d}/|X_{0,d}|)|X_{0,d}|^{\alpha_d})^\intercal$ satisfies Definition 2.9 with index one. Throughout we find it helpful to focus on the non-normalized vector X_0 in order to preserve the multiplicative structure of the tail chain introduced in Section 3.2 below, which is used for analyzing the extremal properties of VSRV processes.

In the following proposition we state the VSRV vector X_0 has a polar decomposition. In the case where Condition **M** is not satisfied, note that the polar decomposition holds on a transformation of the original process. Under Condition **M**, the natural radius notion is $\|\cdot\|_\alpha$, where

$$\|x\|_\alpha := \max_{1\le i\le d} c_i^{-1}|x_i|^{\alpha_i}. \tag{2.15}$$

Notice that the homogeneity of $\|\cdot\|_\alpha$, due to Condition **M**, will be essential for the proof.

Proposition 2.16 *Suppose that the vector X_0 satisfies Condition **M**. Then X_0 is VSRV if and only if there exists a tail vector $Y_0 \in \mathbb{R}^d$ with non-degenerate marginals such that*

$$\mathcal{L}(((c_i t)^{-1/\alpha_i})_{1\le i\le d} \odot X_0 \mid \|X_0\|_\alpha > t) \to_{t\to\infty} \mathcal{L}(Y_0), \tag{2.16}$$

where $\|\cdot\|_\alpha$ is defined in (2.15). Moreover, $\|Y_0\|_\alpha$ is standard Pareto distributed.

Notice that a similar vector scaling argument has been introduced in Lindskog et al. (2014).

Proof Adapting Theorem 4 of de Haan and Resnick (1977), the definition of vector scaling regularly varying distribution of X_0 in Eq. 2.14 implies (2.16). Conversely, under Condition **M**, we have that $|X_{0,k}|^{\alpha_k}$ is regularly varying of order 1 for all $1 \le k \le d$ with slowly varying functions $\ell_i(t) \sim c_i$. Moreover $\|X_0\|_\alpha$ is regularly varying from the weak convergence in Eq. 2.16 applied on the Borel sets $\{\|X_0\|_\alpha > ty\}$, $y \ge 1$. Thus, $\|X_0\|_\alpha$ is regularly varying of order 1 with slowly varying function $\ell(t)$. One can rewrite (2.16) as

$$\ell(t)^{-1} t\mathbb{P}(x(t)^{-1} \odot X_0 \in \cdot, \|X_0\|_\alpha > t) \to \mathbb{P}(Y_0 \in \cdot).$$

Using the slowly varying property of ℓ, we obtain, for any $\epsilon > 0$,

$$\ell(t)^{-1} t\mathbb{P}(x(t)^{-1} \odot X_0 \in \cdot, \|X_0\|_\alpha > t\epsilon) \to \epsilon^{-1}\mathbb{P}(Y_0 \in \cdot).$$

Then by marginal homogeneity of $\|\cdot\|_\alpha$,

$$\ell(t)^{-1} t\mathbb{P}(x(t)^{-1} \odot X_0 \in \cdot, \|x(t)^{-1} \odot X_0\|_\alpha > \epsilon) \to \epsilon^{-1}\mathbb{P}(Y_0 \in \cdot).$$

Notice that $\ell(t)t^{-1} > 0$ is non-increasing as it is the tail of $\|X_0\|_\alpha$. So there exists a change of variable $t = h(t')$ so that $\ell(t)^{-1}t = t'$ and

$$t'\mathbb{P}(x(h(t'))^{-1} \odot X_0 \in \cdot, \|x(h(t'))^{-1} \odot X_0\|_\alpha > \epsilon) \to \epsilon^{-1}\mathbb{P}(Y_0 \in \cdot).$$

We obtain the existence of μ for $x' = x \circ h$ in Eq. 2.14 such that $\mu(\cdot, \|x\|_\alpha > \epsilon) = \mathbb{P}(\cdot)$, which is enough to characterize μ entirely, choosing $\epsilon > 0$ arbitrarily small. □

The spectral properties of VSRV X_0 can be expressed in terms of the tail vector Y_0. Notice that for any $u \in \{+1, 0, -1\}^d$, there exists $c_+(u) \ge 0$ satisfying

$$\lim_{t\to\infty} \mathbb{P}\left(\max_{1\le i\le d} c_i^{-1}(u_i X_{0,i})_+^{\alpha_i} > t \mid \|X_0\|_\alpha > t\right) = c_+(u).$$

Consider $c^{-1} \odot (u \odot X_0)_+^{\alpha}$, where $c^{-1} = (c_1^{-1}, \ldots, c_d^{-1})^{\intercal}$ and for $x \in \mathbb{R}^d$ and $\alpha = (\alpha_1, ..., \alpha_d)^{\intercal}$, $(x)_+^{\alpha} = ((x_1)_+^{\alpha_1}, ..., (x_d)_+^{\alpha_d})^{\intercal}$. If $c_+(u)$ is non-null, by a continuous mapping argument, $c^{-1} \odot (u \odot X_0)_+^{\alpha}$ satisfies

$$\mathcal{L}(t^{-1} c^{-1} \odot (u \odot X_0)_+^{\alpha} \mid \|(u \odot X_0)_+\|_{\alpha} > t) \to_{t\to\infty} \mathcal{L}(c_+(u)^{-1} (u \odot Y_0)_+^{\alpha}), \quad (2.17)$$

and $c^{-1} \odot (u \odot X_0)_+^{\alpha}$ is regularly varying of index 1. By homogeneity of the limiting measure in the multivariate regular variation (2.10), we may decompose the limit as a product

$$\frac{\mathbb{P}(\|(u \odot X_0)_+\|_{\alpha} > ty, c^{-1} \odot (u \odot X_0)_+^{\alpha} / \|(u \odot X_0)_+\|_{\alpha} \in \cdot)}{\mathbb{P}(\|(u \odot X_0)_+\|_{\alpha} > t)} \to y^{-\alpha} \mathbb{P}_{\Theta_u}(\cdot),$$

for any $y \geq 1$. Such limiting distribution is called a simple max-stable distribution, and $\P_{\Theta_u}$, supported by the positive orthant, is called the spectral measure of $c^{-1} \odot (u \odot X_0)_+^{\alpha}$, see de Haan and Resnick (1977) for more details. By identification of the two expressions of the same limit, we obtain the following proposition.

Proposition 2.17 *With Y_0 defined in Proposition 2.16, the distribution of $(u \odot Y_0)_+^{\alpha} / \|(u \odot Y_0)_+\|_{\alpha}$, if non-degenerate, is the spectral measure of $c^{-1} \odot (u \odot X_0)_+^{\alpha} \in [0, \infty)^d$. Moreover, it is independent of $\|(u \odot Y_0)_+\|_{\alpha}$, and $c_+(u)^{-1} \|(u \odot Y_0)_+\|_{\alpha}$ is standard Pareto distributed.*

Proof That $c_+(u)^{-1} \|(u \odot Y_0)_+\|_{\alpha}$ is standard Pareto distributed follows from the convergence in Eq. 2.17 associated with the regularly varying property, ensuring the homogeneity of the limiting measure. Then, using again the homogeneity in Eq. 2.17, it follows that $(u \odot Y_0)_+^{\alpha} / \|(u \odot Y_0)_+\|_{\alpha}$ and $c_+(u)^{-1} \|(u \odot Y_0)_+\|_{\alpha}$ are independent. □

Example 2.18 (**Diagonal BEKK-ARCH**, continued) We have not been able to establish the existence of Y_0 satisfying (2.16), except the case of the scalar BEKK-ARCH where the diagonal elements of A are identical. In this case the process is a special case of the Similarity BEKK-ARCH, see Example 2.6. Even in this case, the characterization of the spectral distribution is not an easy task because of the diagonality of A, ruling out Theorem 1.4 of Buraczewski et al. (2009). In Appendix Section A.2 we have included some estimates of the spectral measure of X_0 for the bivariate case. The plots suggest that the tails of the process are indeed dependent. We emphasize that new multivariate renewal theory should be developed in order to prove that the Diagonal-ARCH model is VSRV. We leave such task for future research.

3 Vector-scaling regularly varying time series and their extremal behavior

The existence of the tail vector in Proposition 2.16 allows us to extend the asymptotic results of Perfekt (1997) to VSRV vectors taking possibly negative values. In order to do so, we use the notion of tail chain from Basrak and Segers (2009) adapted to VSRV stationary sequences with eventually different tail indices.

3.1 Vector scaling regularly varying time series

We introduce a new notion of multivariate regularly varying time series based on VSRV of X_t.

Definition 3.1 The stationary process (X_t) is VSRV if and only if there exists a process $(Y_t)_{t\geq 0}$, with non-degenerate marginals for Y_0, such that

$$\mathcal{L}(((c_i t)^{-1/\alpha_i})_{1\leq i\leq d} \odot (X_0, X_1, \ldots, X_k) \mid \|X_0\|_\alpha > t) \to_{t\to\infty} \mathcal{L}(Y_0, \ldots, Y_k),$$

for all $k \geq 0$. The sequence $(Y_t)_{t\geq 0}$ is called the *tail process*.

Following Basrak and Segers (2009), we extend the notion of spectral measure to the one of *spectral processes* for any VSRV stationary process:

Definition 3.2 The VSRV stationary process (X_t) admits the spectral process (Θ_t) if and only if

$$\mathcal{L}(\|X_0\|_\alpha^{-1}(X_0, X_1, \ldots, X_k) \mid \|X_0\|_\alpha > t) \to_{t\to\infty} \mathcal{L}(\Theta_0, \ldots, \Theta_k),$$

for all $k \geq 0$.

By arguments similar to the ones in the proof of Proposition 2.17, it follows that the VSRV properties also characterize the spectral process of $(c^{-1} \odot (u \odot X_t)_+^\alpha)_{t\geq 0}$, with X_0 following the stationary distribution, which has the distribution of $((u \odot Y_t)_+^\alpha / \|(u \odot Y_0)_+\|_\alpha)_{t\geq 0}$. We have the following proposition.

Proposition 3.3 *For a VSRV stationary process (X_t), where Y_0 has non-degenerate marginals and $\|Y_0\|_\alpha$ is standard Pareto distributed, the spectral process of any non-degenerate $(c^{-1} \odot (u \odot X_t)_+^\alpha)_{t\geq 0}$ is distributed as $((u \odot Y_t)_+^\alpha / \|(u \odot Y_0)_+\|_\alpha)_{t\geq 0}$ and independent of $\|(u \odot Y_0)_+\|_\alpha$. Moreover $c_+(u)^{-1}\|(u \odot Y_0)_+\|_\alpha$ is standard Pareto distributed.*

3.2 The tail chain

In the following, we will focus on the dynamics of the tail process $(Y_t)_{t\geq 1}$ in Definition 3.1, given the existence of Y_0. We will restrict ourselves to the case where (X_t) is a Markov chain, which implies that (Y_t) is also a Markov chain called the *tail chain*; see Perfekt (1997). We have the following proposition.

Proposition 3.4 *Let (X_t) satisfy* (1.1)-(1.2) *be a VSRV stationary process. With $\tilde{M}_t$ defined in* Eq. 1.4*, the tail process (Y_t) admits the multiplicative form*

$$Y_{t+1} = \tilde{M}_{t+1} Y_t, \qquad t \geq 0. \tag{3.1}$$

Proof Following the approach of Janssen and Segers (2014), one first notices that the existence of the kernel of the tail chain does not depend on the marginal distribution. Thus the characterization of the kernel extends automatically from the usual multivariate regular variation setting to the vector scaling regular variation one. It is straightforward to check Condition 2.2 of Janssen and Segers (2014). We conclude that the tail chain has the multiplicative structure in Eq. 3.1. □

The tail chain for VSRV process satisfying (1.1)-(1.2) is the same no matter the values of the marginal tail indices; for the multivariate regularly varying case with common tail indices it coincides with the tail chain of Janssen and Segers (2014) under Condition **M**. Notice that we can extend the tail chain Y_t backward in time ($t < 0$) using Corollary 5.1 of Janssen and Segers (2014).

3.3 Asymptotic behavior of the maxima

From the previous section, we have that the tail chain (Y_t) quantifies the extremal behavior of (X_t) in Eqs. 1.1-1.2. Let us consider the asymptotic behavior of the component-wise maxima

$$\max(X_1, \dots, X_n) = \big(\max(X_{1,k}, \dots, X_{n,k})\big)_{1\le k\le d}\,.$$

Let $u = (1, \dots, 1) = \mathbf{1} \in \mathbb{R}^d$ and assume that $c_+(\mathbf{1}) = \lim_{t\to\infty} \mathbb{P}(X_0 \not\le x(t) \mid |X_0| \not\le x(t))$ is positive. Recall that for (X_t) *i.i.d.*, the suitably scaled maxima converge to the Fréchet distribution; see de Haan and Resnick (1977), i.e. for any $x = (x_1, \dots, x_d)^\intercal \in \mathbb{R}_+^d$, defining $u_n(x)$ such that $n\mathbb{P}(X_{0,i} > u_{n,i}(x)) \sim x_i^{-1}$, $1 \le i \le d$, we have

$$\mathbb{P}(\max(X_1, \dots, X_n) \le u_n(x)) \to \exp(-A^*(x)),$$

if and only if $(X_0)_+$ is vector scaling regularly varying. In such case, due to Condition **M**, we have the expression

$$A^*(x) = c_+(\mathbf{1})\mathbb{E}\left[\frac{1}{\|(Y_0)_+\|_\alpha} \max_{1\le i\le d} \frac{(Y_{0,i})_+^{\alpha_k}}{c_i x_i}\right]. \tag{3.2}$$

Let us assume the following Condition, slightly stronger than Eq. 2.1:

$$\text{There exists } p > 0 \text{ such that } \lim_{n\to\infty} \mathbb{E}[\|\tilde{M}_1 \cdots \tilde{M}_n\|^p]^{1/n} < 1. \tag{3.3}$$

Theorem 3.5 *Let X_t satisfy* (1.1)–(1.2). *With $\tilde{M}_t$ defined in* Eq. 1.4, *suppose that condition* (3.3) *holds. Suppose that the stationary distribution is VSRV. Assuming the existence of Y_0 in Definition 3.1, we have that*

$$\mathbb{P}(\max(X_m, \dots, X_n) \le u_n(x)) \to \exp(-A(x)),$$

where $A(x)$ admits the expression

$$c_+(\mathbf{1})\mathbb{E}\left[\max_{1\le i\le d} \frac{\max_{k\ge 0}\Big(\big(\prod_{1\le j\le k} \tilde{M}_{k-j} Y_0\big)_i\Big)_+^{\alpha_k}}{\|(Y_0)_+\|_\alpha c_i x_i} - \max_{1\le i\le d} \frac{\max_{k\ge 1}\Big(\big(\prod_{1\le j\le k} \tilde{M}_{k-j} Y_0\big)_i\Big)_+^{\alpha_k}}{\|(Y_0)_+\|_\alpha c_i x_i}\right]. \tag{3.4}$$

Proof We verify the conditions of Theorem 4.5 of Perfekt (1997). Condition B2 of Perfekt (1997) is satisfied under the more tractable Condition 2.2 of Janssen and Segers (2014). Indeed, the tail chain depends only on the Markov kernel and one can apply Lemma 2.1 of Janssen and Segers (2014), because it extends immediately to the vector scaling regularly varying setting. Condition $\mathrm{D(u_n)}$ of Perfekt (1997) holds by geometric ergodicity of the Markov chain for a sequence $u_n = C\log n$, with $C > 0$ sufficiently large. Lastly, the finite clustering condition,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}[\max(|X_m|,\dots,|X_{C\log n}|) \not\leq u_n(x) \mid |X_0| \not\leq u_n(x)] = 0, \tag{3.5}$$

holds for any $C > 0$ using the same reasoning as in the proof of Theorem 4.6 of Mikosch and Wintenberger (2013) under the drift condition (DC_p) for some $p < \alpha = \min\{\alpha_i : 1 \le i \le d\}$. As (X_t) is also standard α regularly varying, actually the drift condition holds thanks to Condition (3.3) on some sufficiently large iterations of the Markov kernel. Finally, as Eq. 3.5 is a special case of Condition $\mathrm{D^\infty(c\log n)}$ of Perfekt (1997), we obtain the desired result with the characterization given in Theorem 4.5 of Perfekt (1997),

$$A(x) = \int_{(0,\infty)^d\setminus(0,x)} \mathbb{P}\left(T_j \le x, k \ge 1 \mid T_0 = y\right) \nu(dy),$$

where $(T_k)_{k\ge 0}$ is the tail chain of the standardized Markov chain $(c_i^{-1}(X_{k,i})_+^{\alpha_i})_{1\le i\le d}$, $k \ge 0$. As μ restricted to $(0,\infty)^d \setminus (0,1)^d$ is the distribution of Y_0, we assume that $x_i \ge 1$ for all $1 \le i \le d$ so that we identify ν as the distribution of

$$(c_i^{-1}(Y_{0,i})_+^{\alpha_i})_{1\le i\le d} \qquad \text{under the constraint} \qquad \max_{1\le i\le d} c_i^{-1}(Y_{0,i})_+^{\alpha_i}/x_i > 1.$$

Thus we have

$$A(x) = \mathbb{P}\left(c_i^{-1}(Y_{k,i})_+^{\alpha_i}/x_i \le 1, k \ge 1, 1 \le i \le d, \max_{1\le i\le d} c_i^{-1}(Y_{0,i})_+^{\alpha_i}/x_i > 1\right).$$

To obtain an expression that is valid for any $x_i > 0$, we exploit the homogeneity property, and we obtain

$$\begin{aligned} A(x) &= \mathbb{P}\left(\max_{k\ge 0}\max_{1\le i\le d}(c_i x_i)^{-1} Y_{k,i}^{\alpha_i} > 1\right) - \mathbb{P}\left(\max_{k\ge 1}\max_{1\le i\le d}(c_i x_i)^{-1} Y_{k,i}^{\alpha_i} > 1\right) \\ &= c_+(\mathbf{1})\mathbb{E}\left[\frac{\max_{k\ge 0}\max_{1\le i\le d}(c_i x_i)^{-1}(Y_{k,i})_+^{\alpha_i}}{\|(Y_0)_+\|_\alpha} - \frac{\max_{k\ge 1}\max_{1\le i\le d}(c_i x_i)^{-1}(Y_{k,i})_+^{\alpha_i}}{\|(Y_0)_+\|_\alpha}\right] \end{aligned}$$

because $c_+(\mathbf{1})^{-1}\|(Y_0)_+\|_\alpha$ is standard Pareto distributed and independent of the spectral process $(Y_k)_+^\alpha/\|(Y_0)_+\|_\alpha$. This expression is homogeneous and extends to any possible x by homogeneity. □

3.4 Extremal indices

As the random coefficients $\tilde{M}_t$ in Eq. 1.4 may be large, consecutive values of X_t can be large. In the univariate case, one says that the extremal values appear in clusters. An indicator of the average length of the cluster is the inverse of the extremal index, an indicator of extremal dependence; see Leadbetter et al. (1983).

Thus, the natural extension of the extremal index is the function $\theta(x) = A(x)/A^*(x)$, with $A^*(x)$ and $A(x)$ defined in Eqs. 3.2 and 3.4, respectively. Notice that there is no reason why θ should not depend on x. When $x_i \geq c_+(\mathbf{1})$, for $1 \leq i \leq d$, we have the more explicit expression in terms of the spectral process,

$$\theta(x) = \mathbb{P}\left(Y_{k,i}^{\alpha_i} \leq c_i x_i, k \geq 1, 1 \leq i \leq d \mid Y_{0,i}^{\alpha_i} > c_i x_i, 1 \leq i \leq d\right). \quad (3.6)$$

However, the extremal index θ_i of the marginal $(X_{t,i})$ is still well-defined. It depends on the complete dependence structure of the multivariate Markov chain thanks to the following proposition:

Proposition 3.6 *Let X_t satisfy* (1.1)-(1.2). *With $\tilde{M}_t$ defined in* Eq. 1.4 *satisfying* (3.3) *and assuming the existence of Y_0 in Definition 3.1, the extremal index, θ, defined in* Eq. 3.6*, is a positive continuous function bounded from above by 1 that can be extended to* $(0, \infty]^d \setminus \{\infty, \ldots, \infty\}$. *The extremal indices of the marginals are*

$$\begin{aligned}\theta_i &= \theta(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty) \\ &= \frac{\mathbb{E}\left[\|(Y_0)_+\|_\alpha^{-1}\left(\max_{k\geq 0}\left(\left(\prod_{1\leq j\leq k}\tilde{M}_{k-j}Y_0\right)_i\right)_+^{\alpha_i} - \max_{k\geq 1}\left(\left(\prod_{1\leq j\leq k}\tilde{M}_{k-j}Y_0\right)_i\right)_+^{\alpha_i}\right)\right]}{\mathbb{E}\left[\|(Y_0)_+\|_\alpha^{-1}\left(Y_{0,i}\right)_+^{\alpha_i}\right]}.\end{aligned}$$

Proof Except for the positivity of the extremal index, the result follows by Proposition 2.5 in Perfekt (1997). The positivity is ensured by applying Corollary 2 in Segers (2005). □

Example 3.7 (**Diagonal BEKK-ARCH**, continued) Suppose that X_0 is VSRV as conjectured in Example 2.18. It follows from the tail chain approach of Janssen and Segers (2014) that the stationary Markov chain (X_t) is regularly varying. Thanks to the diagonal structure of the matrices $\tilde{M}_k = Am_k$, one can factorize $\|(Y_0)_+\|_\alpha^{-1}(Y_{0,i})^{\alpha_i}$ in the expression of θ_i provided in Proposition 3.6. Since $\|(Y_0)_+\|_\alpha^{-1}(Y_{0,i})^{\alpha_i}$ and m_k are independent for $k \geq 1$, we recover a similar expression as in the remarks after Theorem 2.1 in de Haan et al. (1989):

$$\theta_i = \mathbb{E}\left[\max_{k\geq 0}\left(A_{ii}^k \prod_{1\leq j\leq k} m_j\right)_+^{\alpha_i} - \max_{k\geq 1}\left(A_{ii}^k \prod_{1\leq j\leq k} m_j\right)_+^{\alpha_i}\right].$$

We did not manage to provide a link between the θ_i and the extremal index $\theta(x)$ of the (multivariate) stationary solution (X_t) of the Diagonal BEKK-ARCH. Due to the different normalising sequences in the asymptotic extremal result given in Theorem 3.5, the extremal index $\theta(x)$ depends on the constants $c_i, i = 1, ..., d$. For $x_i^* = c_+(\mathbf{1})$, $1 \leq i \leq d$, the expression (3.6) gets more simple because $c_+(\mathbf{1})^{-1}\|(Y_0)_+\|_\alpha$ is standard Pareto distributed and supported on $[1, \infty)$:

$$\theta(x^*) = \mathbb{P}\left(A_{ii}^k \prod_{1\leq j\leq k} m_j Y_{0,i} \leq (c_i c_+(\mathbf{1}))^{1/\alpha_i}, k \geq 1, 1 \leq i \leq d\right).$$

One can check that $\theta(x^*) \geq \theta_{i_0}$ where $1 \leq i_0 \leq d$ satisfies $A_{i_0 i_0} \geq A_{ii}$, $1 \leq i \leq d$ so that i_0 is the marginal with smallest tail and extremal indices. Thus the inverse of the extremal index of the multidimensional Diagonal BEKK-ARCH is not larger than the largest average length of the marginals clusters. It can be interpreted as the fact that the largest clusters are concentrated along the i_0 axis, following the interpretation of the multivariate extremal index given on p. 423 of Beirlant et al. (2006).

3.5 Convergence of point processes

Let us consider the vector scaling point process on $\mathbb{R}^d$

$$N_n(\cdot) = \sum_{t=1}^{n} \delta_{((c_i n)^{-1/\alpha_i})_{1\leq i\leq d} \odot X_t}(\cdot), \qquad n \geq 0. \tag{3.7}$$

We want to characterize the asymptotic distribution of the point process N_n when $n \to \infty$. We refer to Resnick (2007) for details on the convergence in distribution for random measures. In order to characterize the limit, we adapt the approach of Davis and Hsing (1995) to the multivariate VSRV case similar to Davis and Mikosch (1998). The limit distribution will be a cluster point process admitting the expression

$$N(\cdot) = \sum_{j=1}^{\infty} \sum_{t=1}^{\infty} \delta_{\left((c_i \Gamma_j)^{-1/\alpha_i}\right)_{1\leq i\leq d} \odot\, Q_{j,t}}(\cdot), \tag{3.8}$$

where Γ_j, $j = 1, 2, \ldots$, are arrival times of a standard Poisson process, and $(Q_{j,t})_{t\in\mathbb{Z}}$, $j = 1, 2, \ldots$, are mutually independent cluster processes. Following Basrak and Tafro (2016), we use the back and forth tail chain (Y_t) to describe the cluster process: Consider the process (Z_t), satisfying

$$\mathcal{L}\Big((Z_t)_{t\in\mathbb{Z}}\Big) = \mathcal{L}\Big((Y_t)_{t\in\mathbb{Z}} \mid \sup_{t\leq -1} \|Y_t\|_\alpha \leq 1\Big),$$

which is well defined when the anti-clustering condition (3.5) is satisfied. Then we have

$$\mathcal{L}\Big((Q_{j,t})_{t\in\mathbb{Z}}\Big) = \mathcal{L}\Big(L_Z^{-1}(Z_t)_{t\in\mathbb{Z}}\Big), \qquad j \geq 1,$$

with $L_Z = \sup_{t\in\mathbb{Z}} \|Z_t\|_\alpha$. Notice that the use of the pseudo-norm $\|\cdot\|_\alpha$ and the fact that $\|Y_0\|_\alpha$ is standard Pareto are crucial to mimic the arguments of Basrak and Tafro (2016). The limiting distribution of the point process N_n coincides with the one of N:

Theorem 3.8 *Let X_t satisfy* (1.1)-(1.2). *With $\tilde{M}_t$ defined in* Eq. 1.4, *suppose that* Eq. 3.3 *holds, and assume that Y_0 in Definition 3.1 exists. With N_n defined in* Eq. 3.7 *and N defined in* Eq. 3.8,

$$N_n \xrightarrow{d} N, \qquad n \to \infty.$$

Proof Let us denote $sign$ the operator $sign(x) = x/|x|$, $x \in \mathbb{R}$, applied coordinatewise to vectors in $\mathbb{R}^d$. We apply Theorem 2.8 of Davis and Mikosch (1998) to the transformed process $(c^{-1} \odot sign(X_t) \odot |X_t|^\alpha)_{t\in\mathbb{Z}}$ which is standard regularly varying of order 1. In order to do so, one has to check that the anti-clustering condition (3.5) is satisfied and that the cluster index of its max-norm is positive.

This follows from arguments developed in the proof of Theorem 3.5. The mixing condition of Davis and Mikosch (1998) is implied by the geometric ergodicity of (X_t). Thus, the limiting distribution of the point process $\sum_{t=1}^{n} \delta_{n^{-1}c^{-1}\odot sign(X_t)\odot|X_t|^{\alpha}}$ coincides with the one of the cluster point process $\sum_{j=1}^{\infty}\sum_{t=1}^{\infty}\delta_{\Gamma_j^{-1}\tilde{Q}_{j,t}}$ for some cluster process $(\tilde{Q}_{j,t})_{t\in\mathbb{Z}}$. A continuous mapping argument yields the convergence of N_n to $\sum_{j=1}^{\infty}\sum_{t=1}^{\infty}\delta_{((c_i\Gamma_j)^{-1/\alpha_i})_{1\le i\le d}\odot\, sign(\tilde{Q}_{j,t})\odot|\tilde{Q}_{j,t}|^{\alpha}}$. The limiting cluster process coincide with $Q_{j,t}$ in distribution thanks to the definition of VSRV processes. □

4 Sample covariances

In this section, we derive the limiting distribution of the sample covariances for certain BEKK-ARCH processes. Consider the sample covariance matrix,

$$\Gamma_{n,X} = \frac{1}{n}\sum_{t=1}^{n} X_t X_t^{\intercal}.$$

Let vech$(\cdot)$ denote the half-vectorization operator, i.e. for a $d\times d$ matrix $A=[a_{ij}]$, vech$(A) = (a_{11}, a_{21}, ..., a_{d1}, a_{22}, ..., a_{d2}, a_{33}, ..., a_{dd})^{\intercal}$ $(d(d+1)/2\times 1)$. The derivation of the limiting distribution of the sample covariance matrix relies on using the multidimensional regularly varying properties of the stationary process $(\text{vech}(X_t X_t^{\intercal}) : t\in\mathbb{Z})$. Let $\mathbf{a}_n^{-1}$ denote the normalization matrix,

$$\mathbf{a}_n^{-1} = \big(n^{-1/\alpha_i-1/\alpha_j} c_i^{-1/\alpha_i} c_j^{-1/\alpha_j}\big)_{1\le i,j\le d}.$$

Using Theorem 3.8 and adapting the continuous mapping argument of Proposition 3.1 of Davis and Mikosch (1998) yield the following result.

Proposition 4.1 *Let X_t satisfy* (1.1)-(1.2)*. With $\tilde{M}_t$ defined in* Eq. 1.4 *satisfying* (3.3) *and assuming the existence of Y_0 in Definition 3.1, we have*

$$\sum_{t=1}^{n}\delta_{\text{vech}(\boldsymbol{a}_n^{-1})\odot\text{vech}(X_tX_t^{\intercal})}\xrightarrow{d}\sum_{\ell=1}^{\infty}\sum_{t=1}^{\infty}\delta_{\text{vech}(\boldsymbol{P}_\ell)\odot\,\text{vech}(Q_{\ell,t}Q_{\ell,t}^{\intercal})},\qquad n\to\infty,$$

where

$$\boldsymbol{P}_\ell = \big(\Gamma_\ell^{-1/\alpha_i-1/\alpha_j} c_i^{-1/\alpha_i} c_j^{-1/\alpha_j}\big)_{1\le i,j\le d}.$$

Let us define $\alpha_{i,j} = \alpha_i\alpha_j/(\alpha_i+\alpha_j)$ and assume that $\alpha_{i,j}\neq 1$ and $\alpha_{i,j}\neq 2$ for all $1\le i\le j\le d$. Note that $\alpha_{i,j}$ is a candidate for the tail index of the cross product $X_{t,i}X_{t,j}$ and that $\alpha_{i,i}=\alpha_i/2$, $1\le i\le d$. Actually it is the case under some extra assumptions ensuring that the product $Y_{0,i}Y_{0,j}$ is non null, see Proposition 7.6 of Resnick (2007). In line with Theorem 3.5 of Davis and Mikosch (1998), we then get our main result on the asymptotic behavior of the empirical covariance matrix.

Theorem 4.2 *Let X_t satisfy* (1.1)-(1.2). *With $\tilde{M}_t$ defined in* Eq. 1.4, *suppose that* Eq. 3.3 *holds, and assume that Y_0 in Definition 3.1 exists. Moreover, for any (i, j) such that $1 < \alpha_{i,j} < 2$, suppose that*

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \mathbb{V}\mathrm{ar}\Big(n^{-1/\alpha_{i,j}} \sum_{t=1}^{n} X_{t,i} X_{t,j} \mathbf{1}_{|X_{t,i} X_{t,j}| \le n^{1/\alpha_{i,j}} \varepsilon}\Big) = 0. \tag{4.1}$$

Then

$$\Big(\sqrt{n} \wedge n^{1-1/\alpha_{i,j}} (\Gamma_{n,X} - \mathbb{E}[\Gamma_{n,X}] \mathbf{1}_{\alpha_{i,j} > 1})_{i,j}\Big)_{1 \le j \le i \le d} \xrightarrow{d} S, \qquad n \to \infty,$$

where $S_{i,j}$ is an $\alpha_{i,j} \wedge 2$-stable random variable for $1 \le i \le j \le d$ and non-degenerate for $i = j$.

When Theorem 4.2 applies, as $\alpha_{i,j} \ge (\alpha_i \wedge \alpha_j)/2$, the widest confidence interval on the covariance estimates is supported by the i_0th marginal satisfying $\alpha_{i_0} \le \alpha_i$ for all $1 \le i \le d$.

In order to apply Theorem 4.2, the main difficulty is to show that the condition (4.1) holds. However, notice that Theorem 4.2 applies simultaneously on the cross-products with $\alpha_{i,j} \notin [1, 2]$ with no extra assumption. Next, we apply Theorem 4.2 to the ongoing examples.

Example 4.3 (**Diagonal BEKK-ARCH**, continued) Consider the diagonal BEKK-ARCH process and the cross products $X_{t,i} X_{t,j}$ for some $i \le j$ and any $t \in \mathbb{Z}$. From Hölder's inequality (which turns out to be an equality in our case), we have

$$\mathbb{E}[|A_{ii} A_{jj} m_0^2|^{\alpha_{i,j}}] = \mathbb{E}[|A_{ii} m_0|^{\alpha_i}|]^{\alpha_{i,j}/\alpha_i} \mathbb{E}[|A_{jj} m_0|^{\alpha_j}]^{\alpha_{i,j}/\alpha_j} = 1.$$

Thus, $(X_{t,i} X_{t,j})$, which is a function of the Markov chain (X_t), satisfies the drift condition (DC_p) of Mikosch and Wintenberger (2013) for all $p < \alpha_{i,j}$. Then, one can show that Eq. 4.1 is satisfied using the same reasoning as in the proof of Theorem 4.6 of Mikosch and Wintenberger (2013).

Example 4.4 (**Similarity BEKK-ARCH**, continued) If $\alpha_{i,j} \notin [1, 2]$, the limiting distribution of the sample covariance matrix for the Similarity BEKK-ARCH follows directly from Theorem 4.2. If $\alpha_{i,j} \in (1, 2)$ the additional condition (4.1) has to be checked. Relying on the same arguments as in Example 4.3, one would have to verify that the condition (DC_p) of Mikosch and Wintenberger (2013) holds for the Similarity BEKK-ARCH process, which appears a difficult task as it requires to find a suitable multivariate Lyapunov function. We leave such task for future investigation. Consider the special case of the scalar BEKK-ARCH process introduced in Example 2.6. Here $A = \sqrt{a} I_d$, with I_d the identity matrix, such that $\tilde{M}_t$ is diagonal. In the case $\alpha_{i,j} \in (1, 2)$ for a least some pair (i, j), the limiting distribution of the sample covariance is derived along the lines of Example 4.3. Specifically, this relies on assuming that $a < \exp\{(1/2)\left[-\psi(1) + \log(2)\right]\}$ such that a stationary solution exists, and noting that the index of regular variation for each marginal of X_t is given by α satisfying $\mathbb{E}[|\sqrt{a} m_t|^{\alpha}] = 1$.

Example 4.5 (**ID BEKK-ARCH**, continued) Whenever $\alpha_{i,j} \notin [1, 2]$, the limiting distribution of the sample covariance matrix for the ID BEKK-ARCH follows directly from Theorem 4.2. Similar to Example 4.4 we leave for future investigation to show whether condition (4.1) holds.

The previous examples are important in relation to variance targeting estimation of the BEKK-ARCH model, as considered in Pedersen and Rahbek (2014). For the univariate GARCH process, Vaynman and Beare (2014) have shown that the limiting distribution of the (suitably scaled) variance targeting estimator follows a singular stable distribution when the tail index of the process lies in (2, 4). We expect a similar result to hold for the BEKK-ARCH process.

5 Concluding remarks

We have found a mild sufficient condition for geometric ergodicity of a class of BEKK-ARCH processes. By exploiting that the processes can be written as a multivaraite stochastic recurrence equation (SRE), we have investigated the tail behavior of the invariant distribution for different BEKK-ARCH processes. Specifically, we have demonstrated that existing Kesten-type results apply in certain cases, implying that each marginal of the invariant distribution has the same tail index. Moreover, we have shown for certain empirically relevant processes, existing renewal theory is not applicable. In particular, we show that the Diagonal BEKK-ARCH processes may have component-wise different tail indices. In light of this property, we introduce the notion of vector scaling regular varying (VSRV) distributions and processes. We study the extremal behavior of such processes and provide results for convergence of point processes based on VSRV processes. It is conjectured, and supported by simulations, that the Diagonal BEKK-ARCH process is VSRV. However, it remains an open task to verify formally that the property holds. Such task will require the development of new multivariate renewal theory.

Our results are expected to be important for future research related to the statistical analysis of the Diagonal BEKK-ARCH model. As recently shown by Avarucci et al. (2013), the (suitably scaled) maximum likelihood estimator for the general BEKK-ARCH model (with $l = 1$) does only have a Gaussian limiting distribution, if the second-order moments of X_t is finite. In order to obtain the limiting distribution in the presence of very heavy tails, i.e. when $\mathbb{E}[\|X_t\|^2] = \infty$, we believe that non-standard arguments are needed, and in particular the knowledge of the tail-behavior is expected to be crucial for the analysis. We leave additional considerations in this direction to future research.

Acknowledgments We are grateful for comments and suggestions from the editor-in-chief (Thomas Mikosch), an associate editor, and two referees, which have led to a much improved manuscript. Moreover, we thank Sebastian Mentemeier for valuable comments. Pedersen greatly acknowledges funding from the Carlsberg Foundation. Financial support by the ANR network AMERISKA ANR 14 CE20 0006 01 is gratefully acknowledged by Wintenberger.

A Appendix

A.1 Theorem 1.1 of Alsmeyer and Mentemeier (2012)

Consider the general SRE

$$Y_t = A_t Y_{t-1} + B_t \tag{4.2}$$

with (A_t, B_t) a sequence of i.i.d. random variables with generic copy (A, B) such that A is a $d \times d$ real matrix and B takes values in $\mathbb{R}^d$. Consider the following conditions of Alsmeyer and Mentemeier (2012):

- **(A1)** $\mathbb{E}[\log^+(\|A\|)] < \infty$, where $\|\cdot\|$ denotes the operator norm.
- **(A2)** $\mathbb{E}[\log^+(\|B\|)] < \infty$.
- **(A3)** $\mathbb{P}[A \in GL(d, \mathbb{R})] = 1$.
- **(A4)** $\max_{n\in\mathbb{N}} \mathbb{P}\left\{\|x^\top \prod_{i=1}^n A_i\|^{-1} \left(x^\top \prod_{i=1}^n A_i\right) \in U\right\} > 0$, for any $x \in \mathbb{S}^{d-1}$ and any non-empty open subset U of $\mathbb{S}^{d-1}$.
- **(A5)** Let $\mathcal{V}_\delta$ denote the open δ-ball in $GL(d, \mathbb{R})$ and let $\mathbb{LEB}$ denote the Lebesgue measure on $M(d, \mathbb{R})$. It holds that for any Borel set $A \in M(d, \mathbb{R})$, $\mathbb{P}(\prod_{i=1}^{n_0} A_i \in A) \geq \gamma_0 1_{\mathcal{V}_c(\Gamma_0)}(A)\mathbb{LEB}(A)$ for some $\Gamma_0 \in GL(d, \mathbb{R})$, $n_0 \in \mathbb{N}$, and $c, \gamma_0 > 0$.
- **(A6)** $\mathbb{P}(A_0 v + B_0 = v) < 1$ for any $v \in \mathbb{R}^d$.
- **(A7)** There exists $\kappa_0 > 0$ such that

$$\mathbb{E}[\inf_{x\in\mathbb{S}^{d-1}} \|x^\top A_0\|^{\kappa_0}] \geq 1, \quad \mathbb{E}[\|A_0\|^{\kappa_0} \log^+ \|A_0\|] < \infty, \quad \text{and} \quad 0 < \mathbb{E}[\|B_0\|^{\kappa_0}] < \infty.$$

Theorem A.1 (Alsmeyer and Mentemeier 2012, Theorem 1.1) *Consider the SRE in* Eq. 4.2 *suppose that* $\beta := \lim_{n\to\infty} n^{-1} \log(\|\prod_{i=1}^n A_i\|) < 0$ *and that* ***(A1)-(A7)*** *hold, then there exists a unique* $\kappa \in (0, \kappa_0]$ *such that*

$$\lim_{n\to\infty} n^{-1} \log(\|\prod_{i=1}^n A_i\|^\kappa) = 0.$$

Moreover, the SRE has a strictly stationary solution satisfying,

$$\lim_{t\to\infty} t^\kappa \mathbb{P}(x^\top Y_0 > t) = K(x) \quad forall\ x \in \mathbb{S}^{d-1},$$

where K *is a finite positive and continuous function on* $\mathbb{S}^{d-1}$.

A.2 Estimation of the spectral measure for the bivariate diagonal BEKK-ARCH process

In this section we consider the estimation of the spectral measure of the diagonal BEKK-ARCH process presented in Example 2.14. Specifically, we consider a special case of the BEKK-ARCH process in Eqs. 1.1-1.2, where $d = 2$:

$$X_t = m_t A X_{t-1} + Q_t,$$

with $\{Q_t : t \in \mathbb{N}\}$ an *i.i.d.* process with $Q_t \sim N(0, C)$ independent of $\{m_t : t \in \mathbb{N}\}$, and

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

Following the approach for *i.i.d.* sequences of vectors given in Einmahl et al. (2001), we consider the following estimator of the spectral measure of $X_t = (X_{t,1}, X_{t,2})^\intercal$:

$$\hat{\Phi}(\theta) = \frac{1}{k}\sum_{t=1}^{T} \mathbf{1}_{\{R_t^{(1)} \vee R_t^{(2)} \geq T+1-k, \arctan \frac{T+1-R_t^{(2)}}{T+1-R_t^{(1)}} \leq \theta\}}, \qquad \theta \in [0, \pi/2],$$

where $R_t^{(j)}$ denotes the rank of $X_{t,j}$ among $X_{1,j}, \ldots, X_{T,j}$, $j = 1, 2$, i.e.

$$R_t^{(j)} := \sum_{i=1}^{T} \mathbf{1}_{\{X_{i,j} \geq X_{t,j}\}}.$$

Here k is a sequence satisfying $k(T) \to \infty$ and $k(T) = o(T)$. Einmahl et al. (2001) showed that this estimator is consistent for i.i.d. series. We expect a similar result to hold for geometrically ergodic processes. The reason is that the asymptotic behavior of the empirical tail process used in Einmahl et al. (2001) has been extended to such cases in Kulik et al. (2015).

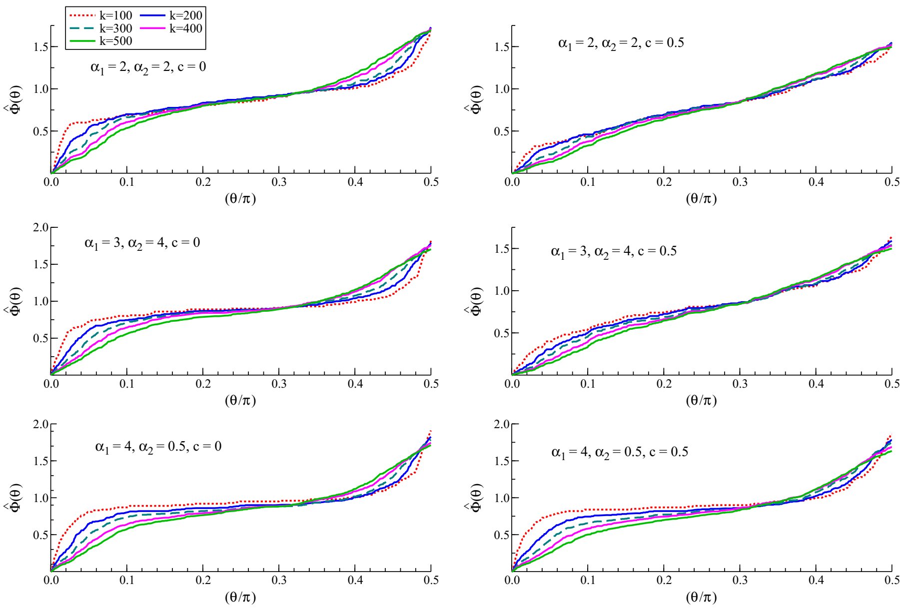

Fig. 1 Nonparametric estimates for $k = 100, 200, 300, 400, 500$ and for various choices of α_1, α_2, and c

We consider the estimation of the spectral measure for different values of C, A_{11}, and A_{22}. In particular, the matrix C is

$$C = 10^{-5} \begin{bmatrix} 1 & c \\ c & 1 \end{bmatrix}, \quad c \in \{0, 0.5\},$$

and the values A_{11} and A_{22} are determined according to choices of the tail indices of $X_{t,1}$ and $X_{t,2}$, respectively. I.e. A_{11} and A_{22} satisfy $\mathbb{E}[|m_t|^{\alpha_i}] = |A_{ii}|^{-\alpha_i}$ and are determined by analytical integration. Specifically, with $\phi(\cdot)$ the pdf of the standard normal distribution,

$$\begin{aligned} \alpha_i &= 0.5 \Rightarrow A_{ii} = (\int_{-\infty}^{\infty} |m|^{0.5} \phi(m) dm)^{-1/0.5} \approx 1.479 \\ \alpha_i &= 2.0 \Rightarrow A_{ii} = 1 \\ \alpha_i &= 3.0 \Rightarrow A_{ii} = (8/\pi)^{-1/6} \approx 0.8557 \\ \alpha_i &= 4.0 \Rightarrow A_{ii} = 3^{-1/4} \approx 0.7598 \end{aligned}$$

Figure 1 contains plots of the estimates of the spectral measure. The estimates $\hat{\Phi}(\theta)$ are based on one realization of the process with $T = 2{,}000$ and a burn-in period of 10,000 observations.

References

Alsmeyer, G.: On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results. J. Theor. Probab. **16**, 217–247 (2003)

Alsmeyer, G., Mentemeier, S.: Tail behaviour of stationary solutions of random difference equations: The case of regular matrices. J. Differ. Equ. Appl. **18**, 1305–1332 (2012)

Avarucci, M., Beutner, E., Zaffaroni, P.: On moment conditions for quasi-maximum likelihood estimation of multivariate ARCH models. Econ. Theory **29**, 545–566 (2013)

Basrak, B., Davis, R.A., Mikosch, T.: A characterization of multivariate regular variation. Ann. Appl. Probab. **12**, 908–920 (2002a)

Basrak, B., Davis, R.A., Mikosch, T.: Regular variation of GARCH processes. Stoch. Process. Appl. **99**, 95–115 (2002b)

Basrak, B., Segers, J.: Regularly varying multivariate time series. Stoch. Process. Appl. **119**, 1055–1080 (2009)

Basrak, B., Tafro, A.: A complete convergence theorem for stationary regularly varying multivariate time series. Extremes **19**, 549–560 (2016)

Bauwens, L., Laurent, S., Rombouts, J.V.K.: Multivariate GARCH models: A survey. J. Appl. Econ. **21**, 79–109 (2006)

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L.: Statistics of extremes: theory and applications. Wiley, NJ (2006)

Boussama, F., Fuchs, F., Stelzer, R.: Stationarity and geometric ergodicity of BEKK multivariate GARCH models. Stoch. Process. Appl. **121**, 2331–2360 (2011)

Buraczewski, D., Damek, E., Guivarc'h, Y., Hulanicki, A., Urban, R.: Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. Probab. Theory Relat. Fields **145**, 385–420 (2009)

Buraczewski, D., Damek, E., Mikosch, T.: Stochastic Models with Power-Law Tails: The Equation X = AX + B, Springer Series in Operations Research and Financial Engineering, Springer International Publishing (2016)

Damek, E., Matsui, M., Świątkowski, W.: Componentwise different tail solutions for bivariate stochastic recurrence equations, arXiv:1706.05800 (2017)

Davis, R.A., Hsing, T.: Point process and partial sum convergence for weakly dependent random variables with infinite variance. Ann. Probab. **23**, 879–917 (1995)

Davis, R.A., Mikosch, T.: The sample autocorrelations of heavy-tailed processes with applications to ARCH. Ann. Stat. **26**, 2049–2080 (1998)

de Haan, L., Resnick, S.I.: Limit theory for multivariate sample extremes. Z. Wahrscheinlichkeitstheorie Verwandte Geb. **40**, 317–337 (1977)

de Haan, L., Resnick, S.I., Rootzén, H., de Vries, C.G.: Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. Stoch. Process. Appl. **32**, 213–224 (1989)

Einmahl, J.H., de Haan, L., Piterbarg, V.I.: Nonparametric estimation of the spectral measure of an extreme value distribution. Ann. Stat. **29**, 1401–1423 (2001)

Engle, R.F., Kroner, K.F.: Multivariate simultaneous generalized ARCH. Econ. Theory **11**, 122–150 (1995)

Feigin, P., Tweedie, R.: Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. J. Time Ser. Anal. **6**, 1–14 (1985)

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

Janssen, A., Segers, J.: Markov tail chains. J. Appl. Probab. **51**, 1133–1153 (2014)

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

Kulik, R., Soulier, P., Wintenberger, O.: The tail empirical process of regularly varying functions of geometrically ergodic Markov chains. arXiv:1511.04903 (2015)

Leadbetter, M., Lindgren, G., Rootzén, H.: Extremes and related properties of random sequences and processes, Springer series in statistics. Springer-Verlag, Berlin (1983)

Lindskog, F., Resnick, S.I., Roy, J.: Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. Probab. Surv. **11**, 270–314 (2014)

Matsui, M., Mikosch, T.: The extremogram and the cross-extremogram for a bivariate GARCH(1,1) process. Adv. Appl. Probab. **48**, 217–233 (2016)

Mikosch, T., Wintenberger, O.: Precise large deviations for dependent regularly varying sequences. Probab. Theory Relat. Fields **156**, 851–887 (2013)

Nelson, D.B.: Stationarity and persistence in the GARCH(1,1) model. Econ. Theory **6**, 318–334 (1990)

Nielsen, H.B., Rahbek, A.: Unit root vector autoregression with volatility induced stationarity. J. Empir. Financ. **29**, 144–167 (2014)

Pedersen, R.S.: Targeting estimation of CCC-GARCH models with infinite fourth moments. Econ. Theory **32**, 498–531 (2016)

Pedersen, R.S., Rahbek, A.: Multivariate variance targeting in the BEKK-GARCH model. Econ. J. **17**, 24–55 (2014)

Perfekt, R.: Extreme value theory for a class of Markov chains with values in $\mathbb{R}^d$. Adv. Appl. Probab. **29**, 138–164 (1997)

Resnick, S.I.: Heavy-tail phenomena: probabilistic and statistical modeling, Springer Science & Business Media (2007)

Segers, J.: Generalized Pickands estimators for the extreme value index. J. Stat. Plan. Infer. **128**, 381–396 (2005)

Stărică, C.: Multivariate extremes for models with constant conditional correlations. J. Empir. Financ. **6**, 515–553 (1999)

Vaynman, I., Beare, B.K.: Stable limit theory for the variance targeting estimator. In: Essays in Honor of Peter C. B. Phillips, ed. by Y. Chang, T. B. Fomby, and J. Y. Park, Emerald Group Publishing Limited, vol. 33 of Advances in Econometrics, chap. 24, pp. 639–672. (2014)