Characterization of the tail behavior of a class of BEKK processes: A stochastic recurrence equation approach

Pedersen, Rasmus Søndergaard; Matsui, Muneya

Published in:
Econometric Theory

Publication date:
2021

Document version
Peer reviewed version

Characterization of the tail behavior of a class of BEKK processes: A stochastic recurrence equation approach

Muneya Matsui[*] and Rasmus Søndergaard Pedersen[†]

December 1, 2020

Abstract

We consider conditions for strict stationarity and ergodicity of a class of multivariate BEKK processes $(X_t : t = 1, 2, \ldots)$ and study the tail behavior of the associated stationary distributions. Specifically, we consider a class of BEKK-ARCH processes where the innovations are assumed to be Gaussian and a finite number of lagged X_t's may load into the conditional covariance matrix of X_t. By exploiting that the processes have multivariate stochastic recurrence equation representations, we show the existence of strictly stationary solutions under mild conditions where only a fractional moment of X_t may be finite. Moreover, we show that each component of the BEKK processes is regularly varying with some tail index. In general, the tail index differs along the components, which contrasts with most of the existing literature on the tail behavior of multivariate GARCH processes. Lastly, in an empirical illustration of our theoretical results, we quantify the model-implied tail index of the daily returns on two cryptocurrencies.

Keywords: Regular variation, GARCH, BEKK, stochastic recurrence equation.
JEL: C32 and C58.

---

[*]Department of Business Administration, Nanzan University, 18 Yamazato-cho Showa-ku Nagoya, 466-8673, Japan. `mmuneya@nanzan-u.ac.jp`. [†]Department of Economics, University of Copenhagen. Øster Farimagsgade 5, DK-1353 Copenhagen K, Denmark. `rsp@econ.ku.dk`
We thank Peter C. B. Phillips (editor), Robert Taylor (co-editor), and two referees for helpful comments and suggestions that have led to a much improved paper. Moreover, we thank Ewa Damek, Rustam Ibragimov, Anders Rahbek, and participants at the 2019 CFE Conference for valuable suggestions. Matsui's research is partly supported by the JSPS Grant-in-Aid for Young Scientists B (16k16023) and for Scientific Research C (19K11868). Pedersen is grateful for support from the Carlsberg Foundation (CF16-0909, "Robust methods for volatility modelling").

1 Introduction

In this paper we present novel results about the tail properties for the stationary solution to a class of multivariate conditionally heteroskedastic BEKK processes. Specifically, with $X_t \in \mathbb{R}^d$ we consider BEKK-ARCH (BEKK$(q,0,l)$) processes of the form

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

$$H_t = C + \sum_{i=1}^{q} \sum_{j=1}^{l} A_{ij} X_{t-i} X_{t-i}' A_{ij}', \tag{1.2}$$

where $(Z_t : t \in \mathbb{N})$ is i.i.d., $Z_t \stackrel{d}{=} N(0, I_d)$ (i.e. Z_t is $N(0, I_d)$-distributed), with I_d the $d \times d$ identity matrix, C is a $d \times d$ positive definite matrix, $A_{ij} \in M(d, \mathbb{R})$ (the set of $d \times d$ real matrices) for $i = 1, ..., q$ and $j = 1, ..., l$, and $X_0, ..., X_{-(q-1)} \in \mathbb{R}^d$ are some initial values. This class of processes was originally introduced by Engle and Kroner (1995). By relying on results for stochastic recurrence equations (SREs), we find a new, mild condition for the existence of an almost surely unique stationary solution to the process in (1.1)-(1.2). Specifically, the condition is stated in terms of the top Lyapunov exponent of random coefficient matrices associated with the SREs. In the case where $l = q = 1$ a sufficient stationarity condition is given explicitly in terms of the spectral radius of the matrix A_{11}, similar to the stationarity condition found by Nelson (1990) for one-dimensional ARCH processes. The conditions are mild in the sense that they impose that only fractional moments of X_t may be finite, whereas existing results impose that second-order moments of X_t are finite. Next, again relying on results for SREs, we demonstrate that for various specifications of the matrices A_{ij} and various values of q and l that each component (of the stationary solution) to (1.1)-(1.2) is regularly varying with some index of regular variation, or tail index, $\alpha_i > 0$, $i = 1, ..., d$. Importantly, we show that the tail indexes may in general be different, which contrasts with most of the existing body of literature on regularly varying solutions to multivariate GARCH processes, where the tail indexes are assumed to be the same along the components of X_t, see e.g. Stărică (1999), Basrak and Segers (2009), and Pedersen (2016). Cases of component-wise different tail indexes in the context of multivariate GARCH-type processes are considered in recent articles by Matsui and Mikosch (2016), for constant conditional correlation (CCC) GARCH processes, and Pedersen and Wintenberger (2018) for the process in (1.1)-(1.2) with $q = l = 1$ and A_{11} diagonal (i.e. Diagonal BEKK-ARCH processes). The results in the present paper extend the theory in Pedersen and Wintenberger (2018) in several directions: for $q = 1$ and $l \geq 1$ we consider the component-wise tail behavior of X_t for cases where the matrices $A_{11}, \ldots, A_{1l}$ are simultaneous diagonalizable or simultaneous triangularizable. These cases include several special cases such as triangular A_{1j} and cases where X_t stacks univariate ARCH(1) processes. For $q \geq 1$ we rely on recent results by Guivarc'h and Le Page (2016) in order to characterize the tail behavior of X_t.

In a vast amount of applications within quantitative economics and finance, it is well-documented that certain time series exhibit power law tails, see e.g. Loretan and Phillips (1994) and Gabaix (2009). Classic examples of such time series are the series of daily returns on publicly traded shares of stocks; e.g. Cont (2001) and Ibragimov et al. (2015). In addition to exhibiting extreme values, such return series typically exhibit conditional heteroskedasticity. The latter has led to an entire research area on univariate and multivariate GARCH models, and it is by now well-known that certain GARCH random variables are heavy tailed, see e.g. Davis and Mikosch (2009) for a discussion on regular variation of univariate GARCH variables and Pedersen and Wintenberger (2018) for references on heavy tailed multivariate GARCH variables.

In addition to providing new results about the properties of a class of BEKK-ARCH processes in (1.1)-(1.2), we conjecture that our results are important for obtaining a better understanding of the properties of the quasi-maximum likelihood (QML) estimators for the BEKK class of models. In particular, Avarucci et al. (2013) have shown that for a particular class of BEKK-ARCH models (with $q = l = 1$), as considered in the present paper, the log-likelihood score contribution has a finite variance if and only if the second-order moments of X_t are finite. Hence, standard arguments used to prove asymptotic normality of QML estimators rely on the assumption that X_t has finite variances. Such a condition may not necessarily be satisfied in practice. For instance, Ibragimov et al. (2015, Section 3.2) document that daily returns on certain emerging market foreign exchange rates may have tail index less than two, and hence infinite variance. Likewise, as argued in Pedersen and Rahbek (2014), the much applied two-step covariance targeting estimator, that relies on computing the sample unconditional covariance matrix of X_t, does only seem to obey a Gaussian limiting distribution (at the usual $\sqrt{T}$-rate) provided that at least the fourth-order moments of X_t are finite. In order to derive the limiting distributions of the aforementioned estimators in the case where the moment restrictions on X_t are not satisfied, it appears essential to have results for the tail behavior of X_t, as done by Pedersen (2016) who consider stable limit theory for the variance targeting estimator for multivariate constant conditional correlation (CCC) GARCH models.

The remainder of the paper is organized as follows. Section 2 provides a brief overview of recent results on the tail behavior of BEKK-ARCH processes, and we outline our main contributions. In Section 3 we state a new, mild condition for strict stationarity and ergodicity of BEKK processes. In Sections 4 and 5 we present results on tail behavior of BEKK-ARCH processes of order $q = 1$ for the cases where the collection of matrices $\{A_{11}, \ldots, A_{1l}\}$ is simultaneously diagonalizable and simultaneous triangularizable, respectively. In Section 6 we present theory for BEKK-ARCH processes of arbitrary order $q \geq 1$. In Section 7 we provide a small empirical illustration of our results in relation to the tail behavior of the returns on cryptocurrencies. We provide concluding remarks in Section 8. In the appendix we provide auxiliary results for regularly varying random variables and SREs as well as technical arguments.

We end this section by providing some definitions and notation used throughout the paper. For any column vector $x \in \mathbb{R}^n$, let $|x|$ denote any vector norm of x. When $x \in \mathbb{R}$, $|x|$ denotes the absolute value of x. For any real matrix A, let $||A||$ denote the operator norm $||A|| = \sup_{x:|x|=1} |Ax|$. We let $\mathbb{S}^{n-1}$ denote the unit sphere in $\mathbb{R}^n$, i.e. $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. For $x \in \mathbb{R}$, $x_+ = \max\{x, 0\}$ and $x_- = \max\{-x, 0\}$. For functions $f, g : \mathbb{R} \to \mathbb{R}$, $f(x) \sim g(x)$ means that $\lim_{x\to\infty} f(x)/g(x) \to 1$. We refer to Appendix A for a definition of regular variation.

2 Existing results and our contributions

We observe that the process in (1.1)-(1.2) can be stated as an SRE: For $i = 1, ..., q$ and $j = 1, ..., l$, let $(m_{i,j,t} : t \in \mathbb{Z})$ be an i.i.d. process with $m_{i,j,t}$ univariate standard normal, $m_{i,j,t} \overset{d}{=} N(0, 1)$, and let $(m_{i,j,t} : t \in \mathbb{Z})$ and $(m_{r,s,t} : t \in \mathbb{Z})$ be mutually independent for all $i \neq r$ or $j \neq s$. Let $(B_t : t \in \mathbb{Z})$ be an i.i.d. process with $B_t \overset{d}{=} N(0, C)$ and mutually independent of $(m_{i,j,t} : t \in \mathbb{Z})$ for all i, j. With $Y_t = (X_t', ..., X_{t-(q-1)}')'$, noting that Z_t is Gaussian, it holds that

$$Y_t = M_t Y_{t-1} + Q_t, \tag{2.1}$$

where

$$M_t = \begin{pmatrix} M_{1,t} & M_{2,t} & \dots & M_{q,t} \\ I_d & & & 0 \\ & \ddots & & \vdots \\ & & I_d & 0 \end{pmatrix}, \tag{2.2}$$

$Q_t = (B_t', 0', \dots, 0')'$, $M_{i,t} = \sum_{j=1}^{l} m_{i,j,t} A_{ij}$ for $i = 1, ..., q$, with $M_t = M_{1,t}$ if $q = 1$.

We use this representation to find new conditions for stationarity and ergodicity and in order to consider the (component-wise) regular variation of various BEKK-ARCH processes of the form (1.1)-(1.2). Recently, Pedersen and Wintenberger (2018) considered the tail behavior of X_t for $q = 1$ under the following conditions for M_t defined in (2.2):

(a) M_t is invertible (almost surely) and has a positive Lebesgue density on $M(d, \mathbb{R})$.

(b) M_t is a similarity (almost surely). Specifically, they consider the case where $l = 1$ and $A_{11} = aO$ with a a positive constant and O an orthogonal matrix. This includes the scalar BEKK process, by setting $O = I_d$.

(c) $l = 1$ and A_{11} is diagonal such that M_t is diagonal (Diagonal BEKK).

For the cases (a) and (b) they show that (under suitable conditions) X_t is multivariate regularly varying with each component having the same tail index; we refer the reader to the monograph by Resnick (2007) for more details on multivariate regular variation. For case (c) the tail indexes of the components of X_t differ whenever the diagonal elements of A_{11} differ in modulus. In order to understand this property, we note that for the diagonal case with $l = 1$, with $\tilde{A}_{ii}$ denoting the ith diagonal element of A_{11},

$$\begin{pmatrix} X_{1,t} \\ \vdots \\ X_{d,t} \end{pmatrix} = \begin{pmatrix} \tilde{A}_{11} m_{1,1,t} X_{1,t-1} + Q_{1,t} \\ \cdots \\ \tilde{A}_{dd} m_{1,1,t} X_{d,t-1} + Q_{d,t} \end{pmatrix}. \tag{2.3}$$

Hence each entry of X_t obeys a one-dimensional SRE, and the component-wise tail indexes are determined by the Kesten-Goldie theorem, see Lemma A.3 in the appendix.

We consider the tail behavior for larger classes of BEKK processes:

(1) $q = 1$, $l \geq 1$ and the matrices $A_{11}, ..., A_{1l}$ are simultaneously diagonalizable. This includes the important special case where $l = 1$ and A_{11} is full or triangular and diagonalizable. Another special case is when X_t stacks d (potentially independent) one-dimensional ARCH(1) processes.

(2) $q = 1$, $l \geq 1$ and the matrices $A_{11}, ..., A_{1l}$ are simultaneously triangularizable. This includes the special case where $l = 2$ and A_{11} and A_{12} are triangular but not simultaneously diagonalizable.

(3) $q \geq 1$ and the distribution of the matrix M_t satisfies certain irreducibility and contraction conditions in the spirit of Guivarc'h and Le Page (2016).

For case (1), considered in Section 4, the strategy is to consider a suitable transformation of X_t. To fix ideas, in the case $l = 1$, $PA_{11}P^{-1} =: D$ is diagonal, such that $Y_t = Dm_{1,t}Y_{t-1} + PQ_t$ with $Y_t = PX_t$. Here Y_t is of the form (2.3), such that each $Y_{i,t}$ forms an SRE. This allows us to study the component-wise regular variation of Y_t, and hence of $X_t = P^{-1}Y_t$, by carefully applying results for sums of regularly varying random variables.

For case (2), considered in Section 5, suppose that $d = 2$ and

$$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} M_{11,t} & M_{12,t} \\ 0 & M_{22,t} \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \begin{pmatrix} Q_{1,t} \\ Q_{2,t} \end{pmatrix}, \tag{2.4}$$

with $M_{11,t}, M_{12,t}, M_{22,t}$ non-degenerate. We see that $X_{2,t}$ obeys an SRE and is hence regularly varying under suitable conditions. However, $X_{1,t}$ does not obey an SRE. In this case, depending on the properties of $M_{11,t}$, $X_{1,t}$ may inherit the tail shape of $X_{2,t}$ or it has fatter tails than $X_{2,t}$. The characterization of the tail behavior of $X_{1,t}$ is non-trivial and requires new technical arguments, extending the recent results by Damek et al. (2019) who study the component-wise tail behavior of $\mathbb{R}^2_+$-valued SREs of the form (2.4).

For case (3), studied in Section 6, we show that (under suitable conditions) the BEKK-ARCH(q) process satisfies some irreducibility and contraction conditions, recently considered by Guivarc'h and Le Page (2016). In particular, one may note that for $q > 1$ the distribution of M_t is singular with respect to the Lebesgue measure on $M(dq, \mathbb{R})$, and the approach used in Pedersen and Wintenberger (2018) (case (a) above) cannot be applied.

The results for the cases (1) and (2) rely on the condition that $q = 1$ such that some entry of X_t satisfies a univariate SRE, which allows us to characterize its tail behavior. For general $q > 2$ and arbitrary structure of the matrices A_{ij}, such an approach is not feasible, and, hence, other arguments are needed, such as the ones applied for case (3) that imposes additional structure on the matrices A_{ij}. It is left as an open task to characterize the tail behavior of X_t for arbitrary q and arbitrary structure on A_{ij}.

3 Strict stationarity and ergodicity of BEKK processes

In this section we use the SRE representation of the BEKK process in order to state new results on strict stationarity and ergodicity. We make the following assumption.

Assumption 3.1. *With M_t defined in (2.2), let γ denote the top Lyapunov exponent associated with the process in (2.1), i.e.*

$$\gamma = \inf_{n \in \mathbb{N}} n^{-1} \mathbb{E}[\log \|M_1 \cdots M_n\|].$$

It holds that $\gamma < 0$.

Under Assumption 3.1, noting that $\mathbb{E}[(\log \|M_t\|)_+] < \infty$ and $\mathbb{E}[(\log |Q_t|)_+] < \infty$, we obtain the following result by an application of Theorem 4.1.4 of Buraczewski et al. (2016) (BDM henceforth):

Theorem 3.2. *Suppose that Assumption 3.1 holds. Then there exists an almost surely unique strictly stationary ergodic causal solution to the SRE in (2.1). In particular, with Z_t Gaussian there exists a strictly stationary ergodic solution to the BEKK process in (1.1)-(1.2).*

Remark 3.3. *Note that for the case $d = q = l = 1$ the BEKK process in (1.1)-(1.2) is a univariate ARCH(1) process, i.e. $X_t = (C + A_{11}^2 X_{t-1}^2)^{1/2} Z_t$ where $C > 0$ and $A_{11}^2 \geq 0$ are scalars, and Z_t is i.i.d.$N(0,1)$. Nelson (1990) showed that a necessary and sufficient condition for the existence of a stationary solution to such a process is that $\mathbb{E}[\log(A_{11}^2 Z_t^2)] < 0$, i.e. that $A_{11}^2 < \exp(-\psi(1) + \log 2) = 3.56\ldots$, where ψ denotes the digamma function. As recently noticed by Pedersen and Wintenberger (2018), one can show that for the case $q = l = 1$ a sufficient condition for the existence of a stationary solution to (1.1)-(1.2) is that $\rho(A_{11} \otimes A_{11}) < 3.56...$, where ρ denotes the spectral radius. In this case, the process in (1.1)-(1.2) (which is a Markov chain for $q = 1$) is geometrically ergodic. The condition is mild in the sense that only fractional moments of $|X_t|$ may be finite, which contrasts with existing results on stationarity conditions that impose finite second-order moments of $|X_t|$ (see e.g. Francq and Zakoïan (2010, Section 11.3) and Boussama et al. (2011) for sufficient conditions for stationarity of BEKK-GARCH processes). Moreover, as noted by Nicholls and Quinn (1982, Corollary 2.1.1), a sufficient condition for stationarity, stronger than Assumption 3.1, is that $\rho(\mathbb{E}[M_t \otimes M_t]) < 1$, which implies that $\mathbb{E}[|X_t|^2] < \infty$.*

Remark 3.4. *The BEKK process in* (1.1)-(1.2) *could be extended with an autoregressive term such that $X_t = \Phi X_{t-1} + H_t^{1/2} Z_t$ with H_t given by* (1.2). *Such a process, which one may denote a vector double autoregressive (DAR) process, has been studied by Nielsen and Rahbek (2014), see also Ling and Li (2008) and the references therein for details on one-dimensional DAR processes. The vector DAR process has an SRE representation of the form* (2.1)-(2.2) *with $M_{1,t} = \Phi + \sum_{j=1}^{l} m_{1,j,t} A_{1j}$. Note that the process may have a strictly stationary solution even if the matrix $\Phi - I$ has reduced rank, in contrast to standard vector autoregressive processes of order one. In the remainder of this paper we focus on the BEKK processes of the form* (1.1)-(1.2), *i.e. with $\Phi = 0$, but emphasize that the results in the following sections are straightforward to adapt to certain vector DAR processes. Hence, we note that certain vector DAR processes are indeed heavy-tailed, as conjectured by Nielsen and Rahbek (2014, Remarks 5 and 6).*

Remark 3.5. *Another potential extension of the process is to allow for vector autoregressive processes with BEKK errors. Specifically, consider the d-dimensional AR(p)-BEKK(q) process given by*

$$X_t = \sum_{i=1}^{p} \Phi_i X_{t-i} + \epsilon_t, \quad \epsilon_t = H_t^{1/2} Z_t, \quad H_t = C + \sum_{i=1}^{q} \sum_{j=1}^{l} A_{ij} \epsilon_{t-i} \epsilon_{t-i}' A_{ij}',$$

for some $(d \times d)$ *matrices* $\Phi_1, \ldots, \Phi_p$. *Observe that the process has the representation*

$$X_t = \sum_{i=1}^{p} \Phi_i X_{t-i} + \sum_{i=1}^{q} M_{i,t}\epsilon_{t-i} + B_t,$$

with $M_{i,t}$ *and* B_t *defined as in Section 2. Let* $r = \max(p, q)$ *and make the convention that* $\Phi_i = 0$ *for* $i > p$ *and* $M_{i,t} = 0$ *for* $i > q$. *Define*

$$\tilde{M}_{i,t} = \begin{cases} 0 & \text{for} \quad i = 1 \\ -\sum_{j,k=1:j+k=i}^{r} M_{k,t}\Phi_j & \text{for} \quad i = 2, \ldots, 2r \end{cases},$$

$$M^{\dagger}_{i,t} = \begin{cases} \Phi_i + M_{i,t} + \tilde{M}_{i,t} & \text{for} \quad i = 1, \ldots, r \\ \tilde{M}_{i,t} & \text{for} \quad i = r+1, \ldots, 2r \end{cases}.$$

Then, noting that $\epsilon_t = X_t - \sum_{i=1}^{p} \Phi_i X_{t-i}$, *it holds that*

$$\begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ \\ X_{t-(2r-1)} \end{pmatrix} = \begin{pmatrix} M^{\dagger}_{1,t} & M^{\dagger}_{2,t} & \cdots & M^{\dagger}_{2r-1,t} & M^{\dagger}_{2r,t} \\ I_d & & & & 0 \\ & \ddots & & & \vdots \\ & & \ddots & & \\ & & & I_d & 0 \end{pmatrix} \begin{pmatrix} X_{t-1} \\ X_{t-2} \\ \vdots \\ \\ X_{t-2r} \end{pmatrix} + \begin{pmatrix} B_t \\ 0 \\ \vdots \\ \\ 0 \end{pmatrix}.$$

Hence, the process has an SRE representation. The first d *rows of this SRE yield a vector random coefficient autoregression of order* $2r$,

$$X_t = \sum_{i=1}^{2r} M^{\dagger}_{i,t} X_{t-i} + B_t.$$

This representation generalizes the random coefficient representation found by Tsay (1987, Theorem 1) for univariate AR(p)-ARCH(q) processes to arbitrary dimension of X_t.

Having shown that there exists a strictly stationary solution to the class of BEKK processes in (1.1)-(1.2), we turn to characterizing the tail properties of the associated stationary law of the processes.

4 Simultaneous diagonalization

We now consider the BEKK process in (1.1)-(1.2) for $q = 1$, which we denote the BEKK-ARCH(1) process. Specifically, for $t \in \mathbb{Z}$,

$$X_t = H_t^{1/2} Z_t, \quad Z_t \text{ is } i.i.d. N(0, I_d), \quad H_t = C + \sum_{i=1}^{l} A_i X_{t-1} X_{t-1}' A_i', \tag{4.1}$$

and we note that the process has the SRE representation,

$$X_t \;=\; M_t X_{t-1} + Q_t, \quad M_t = \sum_{i=1}^{l} m_{i,t} A_i, \tag{4.2}$$

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

We consider BEKK-ARCH(1) processes satisfying that the collection $\{A_i : i = 1, ..., l\}$ is simultaneously diagonalizable, i.e. the collection satisfies that there exists a real non-singular matrix P such that $D_i = P A_i P^{-1}$ is diagonal for $i = 1, ..., l$.[1] Note that if $l = 1$, we simply have that the matrix A_1 should be diagonalizable. Note that a sufficient, but indeed not necessary, condition for A_1 being diagonalizable is that all its eigenvalues are distinct. Since the collection $\{A_i : i = 1, ..., l\}$ is simultaneously diagonalizable, we may, using (4.2), define $Y_t = P X_t$ such that

$$Y_t = \widetilde{M}_t Y_{t-1} + \widetilde{Q}_t, \quad \widetilde{M}_t = \sum_{j=1}^{l} m_{j,t} D_j, \quad \widetilde{Q}_t = P Q_t. \tag{4.3}$$

With $D_{ii,j}$ the ith diagonal element of D_j, we have that the ith component of Y_t, $Y_{i,t}$, can be written as an SRE,

$$Y_{i,t} = \left(\sum_{j=1}^{l} m_{j,t} D_{ii,j} \right) Y_{i,t-1} + \widetilde{Q}_{i,t}, \quad i = 1, ..., d. \tag{4.4}$$

The idea is then to apply Lemma A.3 to each component $Y_{i,t}$. Specifically, under certain conditions stated in Theorem 4.3 below, there exist constants $c_{i,+} > 0$ and $\alpha_i^{(Y)} > 0$ such that $\mathbb{P}(Y_{i,t} > x) \sim c_{i,+} x^{-\alpha_i^{(Y)}}$ and $\mathbb{P}(Y_{i,t} < -x) \sim c_{i,+} x^{-\alpha_i^{(Y)}}$, i.e. $Y_{i,t}$ is regularly varying with tail index $\alpha_i^{(Y)}$. We have that $X_t = P^{-1} Y_t$, such that with P^{ij} denoting element (i, j) of P^{-1}, $X_{i,t} = \sum_{j=1}^{d} P^{ij} Y_{j,t}$. The tail index of $X_{i,t}$ is then obtained by careful investigation of

[1] From Theorem 1.3.21 of Horn and Johnson (2013) we have that a set of diagonalizable matrices $\{A_i : i = 1, ..., l\}$, $l \geq 2$, is simultaneously diagonalizable if and only if every pair of the set commutes.

the sum of the regularly varying variables $Y_{j,t}$. To provide a bit of intuition, suppose that $d = l = 2$, $\alpha_1^{(Y)} \neq \alpha_2^{(Y)}$, and that

$$P = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \theta \in [0, 2\pi].$$

It then holds that

$$\begin{aligned} X_{1,t} &= \cos(\theta)Y_{1,t} - \sin(\theta)Y_{2,t}, \\ X_{2,t} &= \sin(\theta)Y_{1,t} + \cos(\theta)Y_{2,t}. \end{aligned}$$

In the case where $\cos(\theta), \sin(\theta) \neq 0$, we may exploit that a sum of regularly varying variables inherits the smallest tail index of the summands (see Lemma A.2), such that the tail index of $X_{1,t}$ and $X_{2,t}$ is $\min(\alpha_1^{(Y)}, \alpha_2^{(Y)})$. We note that this case corresponds to a situation where at least one of the matrices A_1 and A_2 is a proper full matrix. Hence, intuitively, the fact that at least one of the matrices is full, implies that the tail index does not differ along the elements of X_t. On the other hand, if e.g. $\theta = 0$ such that $\cos(\theta) = 1$ and $\sin(\theta) = 0$, $X_{i,t}$ has tail index $\alpha_i^{(Y)}$, $i = 1, 2$. In this case, A_1 and A_2 are diagonal matrices, and we have that the tail indexes differ. Essentially, the cases with non-full matrices allow for the possibility of different tail indexes along the components of X_t. Such cases have not been given much attention in the existing body of literature on the tail behavior of multivariate GARCH processes. For instance, in Pedersen (2016) and Basrak and Segers (2009, Example 5.3), who consider respectively constant conditional correlation GARCH and factor GARCH processes, the coefficient matrices are assumed to be full, implying (under suitable conditions) that all components of the observed variable have the same tail index.

Note that, since the matrices $D_1, \ldots, D_l$ in (4.3) are diagonal, the elements of Y_t are univariate ARCH(1) processes. Hence, if X_t is a BEKK-ARCH process of the form (4.1) with simultaneously diagonalizable matrices $A_1, \ldots, A_l$, then there exists a transformation PX_t with marginals obeying univariate ARCH(1) processes.

We make the following assumptions that imply strict stationarity of the BEKK-ARCH(1) process and regular variation of $Y_{i,t}$.

Assumption 4.1. *Let $(X_t : t \in \mathbb{Z})$ be the BEKK-ARCH(1) process given in* (4.1). *The collection $\{A_i : i = 1, ..., l\}$ is simultaneously diagonalizable, such that there exist a non-singular $P \in M(d, \mathbb{R})$ and diagonal matrices $D_1, \ldots, D_l \in M(d, \mathbb{R})$ such that $D_j = PA_jP^{-1}$ for $j = 1, \ldots, l$. Let $D_{ii,j}$ denote the ith diagonal element of matrix D_j. For $i = 1 \ldots, d$ there exists $\alpha_i^{(Y)} > 0$ such that $\mathbb{E}[|\sum_{j=1}^{l} D_{ii,j}m_{j,t}|^{\alpha_i^{(Y)}}] = 1$.*

Remark 4.2. *Noting that $m_{j,t}$ and $m_{i,t}$ are independent for $i \neq j$, we have that $\mathbb{E}[|\sum_{j=1}^{l} D_{ii,j} m_{j,t}|^{\alpha_i^{(Y)}}] = \mathbb{E}[|(\sum_{j=1}^{l} D_{ii,j}^2)^{1/2} z|^{\alpha_i^{(Y)}}]$ with z a standard normal random variable. Hence it is straightforward to check (numerically) if $\alpha_i^{(Y)} > 0$ in Assumption 4.1 exists.*

We obtain the following theorem.

Theorem 4.3. *Let $(X_t : t \in \mathbb{Z})$ be the BEKK-ARCH(1) process given in (4.1) with Z_t Gaussian. Suppose that Assumption 4.1 holds. Then the process has an almost surely unique strictly stationary and ergodic solution, $X_t = (X_{1,t}, \ldots, X_{d,t})'$. Let P^{ij} denote element (i, j) of P^{-1} and define the collection $\mathcal{A}_i = \{\alpha = \alpha_j^{(Y)} : j = 1, ..., d \text{ and } P^{ij} \neq 0\}$. Suppose that $\alpha_i = \min \mathcal{A}_i$ has multiplicity one, i.e. α_i appears only once in $\mathcal{A}_i$. Then $X_{i,t}$ is regularly varying with index α_i.*

Proof. We start out by showing that $(X_t : t \in \mathbb{Z})$ has an almost surely unique strictly stationary and ergodic solution. Since $X_t = P^{-1} Y_t$, with Y_t given by (4.3), it suffices to show that $(Y_t : t \in \mathbb{Z})$ has an almost surely unique strictly stationary and ergodic solution. By Theorem 3.2, this is the case if the top Lyapunov exponent $\gamma^Y := \inf_{n \in \mathbb{N}} n^{-1} \mathbb{E}[\log \|\widetilde{M}_1 \cdots \widetilde{M}_n\|] < 0$. Let $\gamma_i := \mathbb{E}[\log |\sum_{j=1}^{l} D_{ii,j} m_{j,t}|]$. By Theorem 1.1 of Gerencsér et al. (2008), $\gamma^Y = \max_{i=1,...,d} \gamma_i$. Moreover, by Assumption 4.1 and Jensen's inequality, we have that $\gamma_i < 0$ for all $i = 1, \ldots, d$, and we conclude that $(Y_t : t \in \mathbb{Z})$ has an almost surely unique strictly stationary and ergodic solution.

Next, it is straightforward to show that each $Y_{i,t}$ given by (4.4) satisfies Lemma A.3 under Assumption 4.1. Specifically, one may conclude that $Y_{i,t}$ is regularly varying with index $\alpha_i^{(Y)} > 0$.

It remains to characterize the tail behavior of $X_{i,t} = \sum_{j=1}^{d} P^{ij} Y_{j,t}$ for $i = 1, ..., d$. Let $\mathcal{K}_i = \{j = 1, ..., d : P^{ij} \neq 0\}$ and define the collection of component-wise tail indexes of Y_t that are relevant for $X_{i,t}$, $\mathcal{A}_i = \{\alpha = \alpha_j^{(Y)} : j \in \mathcal{K}_i\}$. When $\alpha_i := \min \mathcal{A}_i$ has multiplicity one, we may without loss of generality assume that $\alpha_i = \alpha_1^{(Y)}$, i.e. $\{1\} = \{j = 1, ..., d : P^{ij} \neq 0, \alpha_j^{(Y)} = \alpha_i\}$. Then $X_{i,t} = P^{i1} Y_{1,t} + \sum_{j \in \mathcal{K}_i \setminus \{1\}} P^{ij} Y_{j,t}$. Using that each component of Y_t has a symmetric distribution, and by repeated use of Lemma A.2, we conclude that $P^{i1} Y_{1,t}$ has a lower tail index than $\sum_{j \in \mathcal{K}_i \setminus \{1\}} P^{ij} Y_{j,t}$ such that $\mathbb{P}(X_{i,t} > x) \sim c_i x^{-\alpha_i}$ and $\mathbb{P}(X_{i,t} < -x) \sim c_i x^{-\alpha_i}$ for some constant $c_i > 0$. □

We next consider some applications of Theorem 4.3, where it is (implicitly) assumed that Assumption 4.1 holds.

Example 4.4. *Let e_i denote a d-dimensional column vector satisfying, with $e_{j,i}$ denoting its jth entry, $e_{i,i} = 1$ and $e_{j,i} = 0$ for $j \neq i$. Let $q = 1$, $l = d$ and $A_i := A_{1,i} = e_i e_i' a_i$ for some non-zero constant a_i, $i = 1, ..., d$. Here X_t stacks d univariate ARCH(1) processes, potentially correlated. (The processes are mutually independent, if C is diagonal.) The* *matrix A_i is diagonal, and hence the collection $\{A_i : i = 1, ..., d\}$ is simultaneously diagonalizable, choosing $P = I_d$. The tail index, α_i, of $X_{i,t}$ satisfies $\mathbb{E}[|a_i m_{i,t}|^{\alpha_i}] = 1$.*

Example 4.5. *Consider the case $l = d = 2$ where,*

$$A_1 = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \quad a \neq b.$$

We have that A_1 and A_2 are simultaneous diagonalizable such that $D_1 = PA_1P^{-1}$ and $D_2 = PA_2P^{-1}$ with

$$D_1 = \begin{pmatrix} a-b & 0 \\ 0 & a+b \end{pmatrix}, \quad D_2 = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \quad P = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad \text{and} \quad P^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

With z a standard normal random variable, it holds that $\mathbb{E}[|((a-b)^2 + c^2)^{1/2} z|^{\alpha_1^{(Y)}}] = \mathbb{E}[|((a+b)^2 + c^2)^{1/2} z|^{\alpha_2^{(Y)}}] = 1$ where $\alpha_1^{(Y)} \neq \alpha_2^{(Y)}$, since $a \neq b$. In particular, $\alpha_1^{(Y)} < \alpha_2^{(Y)}$ ($\alpha_1^{(Y)} > \alpha_2^{(Y)}$) if $|a-b| > |a+b|$ ($|a-b| < |a+b|$). Hence, $X_{1,t}$ and $X_{2,t}$ have tail index $\alpha_1^{(Y)} \wedge \alpha_2^{(Y)}$.

Example 4.6. *In contrast to the previous example, we may for $l = 1$ have that A_1 has some non-distinct eigenvalues. Suppose that $d = 3$, and that with $a, b, c \neq 0$ and $a \neq b$,*

$$A_1 = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & c & b \end{pmatrix}.$$

Then $D_1 = PA_1P^{-1}$ with

$$D_1 = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}, \quad P = \begin{pmatrix} 0 & \frac{c}{a-b} & 0 \\ 1 & 0 & 0 \\ 0 & -\frac{c}{a-b} & 1 \end{pmatrix}, \quad \text{and} \quad P^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{a-b}{c} & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

In this case, $\alpha_1^{(Y)} = \alpha_2^{(Y)}$, satisfying $\mathbb{E}[|am_{1,t}|^{\alpha_1^{(Y)}}] = 1$ and $\alpha_3^{(Y)}$ satisfies $\mathbb{E}[|bm_{1,t}|^{\alpha_3^{(Y)}}] = 1$. Due to the structure of P^{-1}, we have that $X_{1,t}$ and $X_{2,t}$ have tail index $\alpha_1^{(Y)}$, whereas $X_{3,t}$ has index $\alpha_1^{(Y)} \wedge \alpha_3^{(Y)}$, since $a \neq b$.

The above example motivates the following theorem.

Theorem 4.7. *Let $(X_t : t \in \mathbb{Z})$ be a BEKK-ARCH(1) process given by (4.1) with Z_t Gaussian. Suppose that Assumption 4.1 holds such that the process is strictly stationary. Let $\mathcal{A}_i$ be defined as in Theorem 4.3, $\alpha_i = \min \mathcal{A}_i$, and $\mathcal{G}_i = \{j = 1, ..., d : P^{ij} \neq 0, \alpha_j^{(Y)} =$*

$\alpha_i\}$. For a given $s \in \{1, \ldots, l\}$ with $D_{jj,s}$ the jth diagonal element of $D_s = PA_sP^{-1}$, suppose that $D_{jj,s} = D_{kk,s}$ for all $j, k \in \mathcal{G}_i$. Then $X_{i,t}$ is regularly varying with index α_i.

Proof. With $Y_t = PX_t$, we have that $Y_{i,t}$ has tail index $\alpha_i^{(Y)} > 0$, satisfying $\mathbb{E}[|\sum_{j=1}^{l} D_{ii,j}m_{j,t}|^{\alpha_i^{(Y)}}] = 1$. Let $\mathcal{G}_i^\dagger = \{j = 1, ..., d : P^{ij} \neq 0\} \backslash \mathcal{G}_i$. It holds that $X_{i,t} = \sum_{j\in\mathcal{G}_i} P^{ij}Y_{j,t} + \sum_{j\in\mathcal{G}_i^\dagger} P^{ij}Y_{j,t}$. For a given $k \in \{1, \ldots, l\}$, with $j \in \mathcal{G}_i$ let $\lambda_k = D_{jj,k}$. Then $\sum_{j\in\mathcal{G}_i} P^{ij}Y_{j,t} = \sum_{j\in\mathcal{G}_i} P^{ij}(\sum_{k=1}^{l} D_{jj,k}m_{k,t})Y_{j,t-1} + \sum_{j\in\mathcal{G}_i} P^{ij}\widetilde{Q}_{j,t} = (\sum_{k=1}^{l} \lambda_k m_{k,t}) \sum_{j\in\mathcal{G}_i} P^{ij}Y_{j,t-1} + \sum_{j\in\mathcal{G}_i} P^{ij}\widetilde{Q}_{j,t}$, where $\widetilde{Q}_t = PQ_t$. Hence $\sum_{j\in\mathcal{G}_i} P^{ij}Y_{j,t}$ obeys an SRE, and Lemma A.3 implies that $\sum_{j\in\mathcal{G}_i} P^{ij}Y_{j,t}$ has tail index α_i. By repeated use of Lemma A.2 we conclude that the tail index of $X_{i,t}$ is α_i. □

Example 4.8. *For $l = 1$ and $d = 3$, suppose that for $a, b \neq 0$ and $a \neq b$,*

$$A_1 = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}.$$

Then $D_1 = PA_1P^{-1}$ with

$$D_1 = \begin{pmatrix} a-b & 0 & 0 \\ 0 & a-b & 0 \\ 0 & 0 & a+2b \end{pmatrix}, \; P = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \; \text{and } P^{-1} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

In this case, $\alpha_1^{(Y)} = \alpha_2^{(Y)}$, satisfying $\mathbb{E}[|(a-b)m_{1,t}|^{\alpha_1^{(Y)}}] = 1$ and $\alpha_3^{(Y)}$ satisfies $\mathbb{E}[|(a+2b)m_{1,t}|^{\alpha_3^{(Y)}}] = 1$. We note that $\alpha_1^{(Y)} = \alpha_2^{(Y)} = \alpha_3^{(Y)}$ if and only if $a = -b/2$. In light of Theorem 4.7, we have that each component of X_t has tail index $\alpha_1^{(Y)}$ if $\alpha_1^{(Y)} < \alpha_3^{(Y)}$ (i.e. if $|a-b| > |a+2b|$) and index $\alpha_3^{(Y)}$ if $\alpha_3^{(Y)} < \alpha_1^{(Y)}$ (i.e. if $|a-b| < |a+2b|$).

5 Simultaneous triangularization

In this section, we consider the BEKK-ARCH(1) process in (4.1) for $d = 2$, and $l \geq 1$, where the matrices $A_1, \ldots, A_l$ are simultaneous triangularizable in the sense that there exists a nonsingular $P \in M(2, \mathbb{R})$ such that $U_i = PA_iP^{-1}$ is upper triangular for all $i = 1, \ldots, l$.[2] A special case is that $A_i = U_i$ such that $P = I_2$. Defining $Y_t = PX_t$, we have the SRE representation of the form (4.3) where $\widetilde{M}_t = \sum_{i=1}^{l} m_{it}U_i$ and $\widetilde{Q}_t = PQ_t$. Recall that the original process X_t is easily recovered by $X_t = P^{-1}Y_t$. We hence study the special case

[2]By Horn and Johnson (2013, Theorem 2.4.8.7 and the comments thereafter), we have that a set of square matrices $\{A_i : i = 1, \ldots, l\}$, $l \geq 2$ is simultaneously triangularizable by a unitary matrix if every pair of the set commutes.

$A_i = U_i$ and $P = I_2$, so that we work on (2.4) with X_t replaced by Y_t,

$$Y_t = M_t Y_{t-1} + Q_t, \tag{5.1}$$

which we may write as

$$\begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = \begin{pmatrix} M_{11,t} & M_{12,t} \\ 0 & M_{22,t} \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} + \begin{pmatrix} Q_{1,t} \\ Q_{2,t} \end{pmatrix}. \tag{5.2}$$

Note that this SRE has the coordinate-wise representation,

$$Y_{1,t} = M_{11,t} Y_{1,t-1} + D_t, \tag{5.3}$$

$$Y_{2,t} = M_{22,t} Y_{2,t-1} + Q_{2,t}, \tag{5.4}$$

where $D_t = M_{12,t} Y_{2,t-1} + Q_{1,t}$. For later use, we note that $Y_{1,t}$ has a decomposition given by $Y_{1,t} = \widetilde{Y}_{1,t} + \widehat{Y}_{1,t}$, where $\widetilde{Y}_{1,t}$ and $\widehat{Y}_{1,t}$ satisfy, respectively, the SREs

$$\widehat{Y}_{1,t} = M_{11,t} \widehat{Y}_{1,t-1} + Q_{1,t}, \tag{5.5}$$

$$\widetilde{Y}_{1,t} = M_{11,t} \widetilde{Y}_{1,t-1} + \widetilde{D}_t, \quad \widetilde{D}_t = M_{12,t} Y_{2,t-1}. \tag{5.6}$$

We occasionally, for notationally convenience, omit the subscript 0 in $M_{ij,0}$, $Q_{i,0}$ and just write M_{ij} and Q_i. Moreover, we define for $t \in \mathbb{Z}$,

$$\Pi^{(i)}_{t,s} = \prod_{j=s}^{t} M_{ii,j},\ t \geq s,\ i = 1, 2, \quad \text{and} \quad \Pi^{(i)}_{t,s} = 1,\ t < s, \quad \text{and} \quad \Pi^{(i)}_{t} = \Pi^{(i)}_{t,1}.$$

Since M_t is triangular, it holds that the stationarity condition in Assumption 3.1 can be simplified. Specifically, let

$$\gamma_i = \inf_{n \in \mathbb{N}} n^{-1} \mathbb{E}[\log |\Pi^{(i)}_n|] = \mathbb{E}[\log |M_{ii}|]. \tag{5.7}$$

Then by Theorem 1.1 of Gerencsér et al. (2008), Assumption 3.1 holds if and only if

$$\max_{i=1,2} \gamma_i < 0. \tag{5.8}$$

If $l = 1$, this condition for $A_{ii,1}$, $i = 1, 2$ reduces to those stated in Remark 3.3 for the case $q = l = 1$. In line with Lemma A.3 we assume that there exist $\alpha_1 > 0$ and $\alpha_2 > 0$ such that

$$\mathbb{E}[|M_{11}|^{\alpha_1}] = 1 \qquad \text{and} \qquad \mathbb{E}[|M_{22}|^{\alpha_2}] = 1. \tag{5.9}$$

By an application of Jensen's inequality, it is easily concluded that condition (5.9) implies the stationarity condition (5.8); see also Proposition 2.1 of Damek et al. (2019) and the proof of Theorem 4.3. Notice that the conditions in (5.9) and the fact that $\mathbb{E}[|M_{ii}|^0] = 1$ and $\mathbb{E}[\log|M_{ii}|] < 0$ imply by convexity that $\mathbb{E}[|M_{ii}|^\beta] < 1$ for $0 < \beta < \alpha_i$, $i = 1, 2$.

We now turn to the tail behavior of each component (5.3) and (5.4). By Gaussianity it is easy to see that

$$\mathbb{E}[|M_{ii}|^{\alpha_i}(\log|M_{ii}|)_+] < \infty \qquad \text{and} \qquad \mathbb{E}[|Q_i|^{\alpha_i}] < \infty, \quad i = 1, 2. \tag{5.10}$$

Noting that conditions (1) and (3) of Lemma A.3 are immediately satisfied, the tail behavior of component Y_2 follows from Lemma A.3. On the other hand, the tail behavior of Y_1, given by the SRE (5.3), is far from trivial to obtain as the random sequence $((M_{11,t}, D_t) : t \in \mathbb{Z})$ is stationary but not i.i.d. and, moreover, $Y_{1,t-1}$ and D_t are dependent. Indeed, standard Kesten-Goldie-type theory (i.e. Lemma A.3) is not applicable for this type of SRE, as recently pointed out by Damek et al. (2019), who consider the case of $\mathbb{R}^2_+$-valued SREs. The following theorem states the tail properties of Y_1 and Y_2.

Theorem 5.1. *Let* $k_i = \mathbb{E}[|M_{ii}|^{\alpha_i}\log|M_{ii}|]$, $i = 1, 2$. *Consider the bivariate SRE* (5.2) *such that* (5.9) *holds and such that* $\mathbb{P}(M_{12} > 0) > 0$. *Then there exists a stationary solution to the SRE,* Y_0*, that satisfies*

$$\mathbb{P}(Y_{1,0} > x) \sim \mathbb{P}(Y_{1,0} < -x) \sim \begin{cases} \overline{c}_1 x^{-\alpha_1} & \text{if } \alpha_1 < \alpha_2 \\ \tilde{c}_1 x^{-\alpha_2} & \text{if } \alpha_1 > \alpha_2, \end{cases}$$

and

$$\mathbb{P}(Y_{2,0} > x) \sim \mathbb{P}(Y_{2,0} < -x) \sim c_2 x^{-\alpha_2}, \qquad x \to \infty, \tag{5.11}$$

where $k_i > 0$ *and*

$$\begin{aligned}
\overline{c}_1 &= \frac{1}{2\alpha_1 k_1}\mathbb{E}[|M_{11,0}Y_{1,-1} + D_0|^{\alpha_1} - |M_{11,0}Y_{1,-1}|^{\alpha_1}], \\
\tilde{c}_1 &= c_2 \lim_{s\to\infty} \mathbb{E}[|\sum_{i=1}^{s} \Pi^{(1)}_{0,2-i}\Pi^{(2)}_{-i,1-s} M_{12,1-i}|^{\alpha_1}] > 0, \\
c_2 &= \frac{1}{2\alpha_2 k_2}\mathbb{E}[|M_{22,0}Y_{2,-1} + Q_{2,0}|^{\alpha_2} - |M_{11,0}Y_{2,-1}|^{\alpha_2}] > 0.
\end{aligned}$$

Let $\widehat{Y}_1$ *and* $\widetilde{Y}_1$ *be solutions, respectively, to the SREs* (5.5) *and* (5.6). *Representations of such solutions are given, respectively, by* (B.3) *and* (B.4) *in Appendix B. With* $0 <$

$\mathbb{E}[|\widehat{Y}_{1,0}|^{\alpha_1} - |M_{11,0}\widehat{Y}_{1,-1}|^{\alpha_1}]$, *if*

$$\mathbb{E}[|\widehat{Y}_{1,0}|^{\alpha_1} - |M_{11,0}\widehat{Y}_{1,-1}|^{\alpha_1}] \neq \mathbb{E}[|\widetilde{Y}_{1,0}|^{\alpha_1} - |M_{11,0}\widetilde{Y}_{1,-1}|^{\alpha_1}], \tag{5.12}$$

then $\overline{c}_1 > 0$.

Remark 5.2. *Since* Z_t *is Gaussian, the condition* $\mathbb{P}(M_{12} > 0) > 0$ *is equivalent to* $\mathbb{P}(M_{12} = 0) = 0$. *This holds if and only if at least one of the matrices* $A_1, \ldots, A_l$ *has a non-zero upper off-diagonal element. If this condition does not hold, then necessarily all matrices* $A_1, \ldots, A_l$ *are diagonal, and we are in the setting of Section 4, and* Y_1 *has tail index* α_1.

Remark 5.3. (a) *In order to prove that* $\overline{c}_1 > 0$ *we note that it is not possible to follow the approach used by Goldie (1991, proof of Theorem 4.1), since* (D_t) *is a dependent process. Specifically, the key argument made by Goldie is to apply an extended version of Lévy's inequality (Goldie (1991, Proposition 4.2)), which is not applicable in our setting.*
(b) *The relation in* (5.12)*, ensuring that* $\overline{c}_1 > 0$*, is quite weak. Indeed, suppose that the relation does not hold for a given process. Then we may consider a new process of the form* (5.2) *with* $M_{12,t}$ *replaced by* $aM_{12,t}$ *with* $a \in \mathbb{R}$. *For the new process, the right-hand side of* (5.12) *becomes* $|a|^{\alpha_1}\mathbb{E}[|\widetilde{Y}_{1,0}|^{\alpha_1} - |M_{11,0}\widetilde{Y}_{1,-1}|^{\alpha_1}]$. *Hence, the new process satisfies the relation* (5.12) *for* $|a| \neq 1$. *We note that* $\overline{c}_1 > 0$ *may hold even if* (5.12) *is violated, but we leave it as an open task to prove this.*
(c) *Note that the positivity of* $\mathbb{E}[|\widehat{Y}_{1,0}|^{\alpha_1} - |M_{11,0}\widehat{Y}_{1,-1}|^{\alpha_1}]$ *follows by Lemma A.3.*
(d) *The above theorem is easily extended to general bivariate SREs of the form* (5.2) *with* M_t *and* Q_t *non-Gaussian, at the expense of additional technical assumptions, see also Damek and Matsui (2019).*

In order to prove the theorem, in light of the above discussion, it suffices to establish the tail behavior for Y_1, which is done by extending the theory recently developed by Damek et al. (2019) to $\mathbb{R}^2$-valued SREs. This extension requires lengthy technical arguments given in Appendix B.

Example 5.4. *Consider the SRE in* (5.1) *with* $l = 2$ *such that*

$$A_1 = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}, \quad A_2 = \begin{pmatrix} c & 0 \\ 0 & \tilde{c} \end{pmatrix}, \quad a, b, c, \tilde{c} \neq 0,\ |c| \neq |\tilde{c}|,\ a + c, a + \tilde{c} \neq 0.$$

Noting that A_1 *is non-diagonalizable,* A_1 *and* A_2 *are not simultaneously diagonalizable, but trivially simultaneously triangularizable with* $P = I_2$. *Suppose that there exist* $\alpha_1 > 0$ *and* $\alpha_2 > 0$ *such that* $\mathbb{E}[|am_{1,t} + cm_{2,t}|^{\alpha_1}] = \mathbb{E}[|am_{1,t} + \tilde{c}m_{2,t}|^{\alpha_2}] = 1$ *and such that* (5.12) *holds. Due to Theorem 5.1,* $X_{1,0}$ *has tail index* $\alpha_1 \wedge \alpha_2$ *while* $X_{2,0}$ *has index* α_2.

In the next example, we consider the case where $l = 2$ and A_1 and A_2 are non-triangular, but simultaneously triangularizable.

Example 5.5. *Let $l = 2$ and consider the SRE in* (5.1) *where*

$$A_1 = \begin{pmatrix} a & \frac{b-a}{2} \\ \frac{a-b}{2} & b \end{pmatrix}, \quad A_2 = \begin{pmatrix} a & c \\ a-b+c & b \end{pmatrix}, \quad |a| \neq |b|,\ a, b, c \neq 0,\ c \neq \frac{b-a}{2}, -a, b.$$

Note that A_1 and A_2 are not commutable (and hence not simultaneously diagonalizable) since

$$[A_1A_2]_{12} = ac + \frac{b^2 - ab}{2} \neq cb + \frac{ab - a^2}{2} = [A_2A_1]_{12},$$

where $[\cdot]_{ij}$ is the ij element of the matrix in the bracket. However, they are simultaneously triangularizable: $U_1 = PA_1P^{-1}$ and $U_2 = PA_2P^{-1}$ with

$$U_1 = \begin{pmatrix} \frac{a+b}{2} & b-a \\ 0 & \frac{a+b}{2} \end{pmatrix}, \quad U_2 = \begin{pmatrix} a+c & b-a \\ 0 & b-c \end{pmatrix}, \quad P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Let $d := \sqrt{(a+b)^2/4 + (a+c)^2}$ and $e := \sqrt{(a+b)^2/4 + (b-c)^2}$, and suppose that for a standard normal random variable z, $\mathbb{E}[|dz|^{\beta_1}] = \mathbb{E}[|ez|^{\beta_2}] = 1$ with $\beta_1 \neq \beta_2$. Moreover, assume that a and b are chosen such that (5.12) *holds. Then according to Theorem 5.1, $\alpha_1^{(Y)} = \beta_1 \wedge \beta_2$ and $\alpha_2^{(Y)} = \beta_2$. Since $X = P^{-1}Y$ by Lemma A.2, $\alpha_1^{(X)} = \alpha_2^{(X)} = \beta_1$ when $\beta_1 < \beta_2$. However, for $\beta_2 < \beta_1$, so that $\alpha_1^{(Y)} = \alpha_2^{(Y)} = \beta_2$, it is not clear how to determine the tail behavior of X_0, since $Y_{1,0}$ and $Y_{2,0}$ are dependent.*

Remark 5.6. (a) *Consider the simple SRE in* (5.1) *where $l = 1$ and A_1 non-diagonalizable,*

$$A_1 = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}, \quad a \neq 0.$$

In this case, (5.9) *implies that $\alpha_1 = \alpha_2$, and obtaining the tail properties of Y_1 is not straightforward. A similar case has recently been studied by Damek and Zienkiewicz (2018) who consider an SRE of the type* (5.1) *with M_{11} and M_{22} non-negative almost surely. However, an extension of their results to the case where M_{11} and M_{22} are allowed to take negative values (with positive probability) is non-trivial, see also Damek and Matsui (2019), and we leave the case $\alpha_1 = \alpha_2$ for future research.*
(b) *Our results may be extended to the d-dimensional case by extending recent results for $\mathbb{R}_+^d$-valued SREs by Matsui and Świątkowski (2018).*

6 Tail properties of BEKK-ARCH(q)

In this section we consider the tail properties of the BEKK-ARCH process of order $q \geq 1$. Recall that this process has the SRE representation given by (2.1)-(2.2), and the main idea is to show that the SRE satisfies certain irreducibility and contraction conditions recently considered by Guivarc'h and Le Page (2016), see also Section 4.4.8 of BDM.

With M_t defined in (2.2), let $\mathbb{P}_M$ denote its distribution. Define

$$G_M = \{s \in M(d, \mathbb{R}) : s = a_1 \cdots a_n,\ a_i \in \mathrm{supp}\mathbb{P}_M,\ i = 1, \dots, n,\ n \in \mathbb{N}\}, \tag{6.1}$$

where $\mathrm{supp}\mathbb{P}_M$ denotes the support of $\mathbb{P}_M$. We initially make the following high-level assumptions (BDM, p.189):

Assumption 6.1. *(a) With G_M defined in (6.1), there exists no finite union $\mathcal{W} = \bigcup_{i=1}^{n} W_i$ of proper subspaces $W_i \subsetneq \mathbb{R}^{dq}$ such that for any $v \in G_M$, $v\mathcal{W} = \mathcal{W}$. (b) G_M contains a matrix that has a unique largest eigenvalue in modulus with multiplicity one.*

Assumption 6.1(a) is an irreducibility condition, and Assumption 6.1(b) is a contraction condition stating that G_M contains a proximal matrix. In Lemmas 6.3-6.5 below we state more primitive sufficient conditions for Assumption 6.1. The following theorem states that the stationary solution to the SRE in (2.1)-(2.2) is multivariate regularly varying; see e.g. Resnick (2007).

Theorem 6.2. *For the BEKK-ARCH process of order $q \geq 1$ with Z_t Gaussian and with SRE representation given by (2.1)-(2.2), suppose that Assumptions 3.1 and 6.1 hold, that $\mathbb{P}[\det(M_{q,t}) = 0] = 0$, and that there exists $\alpha > 0$ such that $\inf_{n\in\mathbb{N}}(\mathbb{E}[\|M_1 \cdots M_n\|^{\alpha}])^{1/n} = 1$. Then the stationary solution, Y_t, to the SRE is multivariate regularly varying with index α, i.e. there exists a probability measure $\mathbb{P}_\Theta$ on $\mathbb{S}^{dq-1}$ such that*

$$\frac{\mathbb{P}(|Y_t| > sx, \widetilde{Y}_t \in \cdot)}{\mathbb{P}(|Y_t| > x)} \xrightarrow{w} s^{-\alpha}\mathbb{P}_\Theta(\cdot), \quad \text{as } x \to \infty, \quad s > 0, \quad \widetilde{Y}_t = Y_t/|Y_t|, \tag{6.2}$$

where $\xrightarrow{w}$ denotes weak convergence.

The multivariate regular variation in (6.2) implies that for any $y \in \mathbb{S}^{dq-1}$, $P(y'Y_t > x) \sim c(y)x^{-\alpha}$ as $x \to \infty$, where $c(y)$ may depend on y and $c(\tilde{y}) > 0$ for some $\tilde{y} \in \mathbb{S}^{dq-1}$. Moreover, $|Y_t|$ is regularly varying with index α.

Proof. The theorem is proved by verifying the conditions of Theorem 5.2 of Guivarc'h and Le Page (2016); see also Theorem 4.4.18 in BDM. It suffices to show that (i) M_t is invertible almost surely, (ii) for all $x \in \mathbb{R}^{dq}$, $\mathbb{P}(M_t x + Q_t = x) < 1$, and (iii) $\mathbb{E}[\|M_t\|^{\alpha+\delta}] < \infty$,

$\mathbb{E}[||M_t||^\alpha ||M_t^{-1}||^\delta] < \infty$, and $\mathbb{E}[|Q_t|^{\alpha+\delta}] < \infty$ for some $\delta > 0$. Condition (i) is clearly satisfied as $\det(M_t) = \det(M_{q,t}) \neq 0$ almost surely. Condition (ii) is immediate as M_t and Q_t are independent and Q_t is non-degenerate. Condition (iii) holds by noting that the elements of M_t and Q_t are Gaussian and an application of Hölder's inequality, choosing $\delta > 0$ sufficiently small. □

The following lemmas give sufficient conditions for Assumption 6.1(a).

Lemma 6.3. *With $M_{1,t}, \dots, M_{q,t}$ the random matrices in* (2.2)*, let $M_t^{(1,q)}$ denote the $d \times dq$ matrix given by*

$$M_t^{(1,q)} = (M_{1,t}, \dots, M_{q,t}). \tag{6.3}$$

Suppose that for any non-zero $x \in \mathbb{R}^{dq}$ the distribution of $M_t^{(1,q)} x$ has a density with respect to the Lebesgue measure strictly positive on $\mathbb{R}^d$. Then Assumption 6.1(a) holds.

Proof. The proof extends the arguments in Section 4.4.9 of BDM to arbitrary dimension d. The strategy is to show that the only space that satisfies $v\mathcal{W} = \mathcal{W}$ for all $v \in G_M$ is $\mathcal{W} = \mathbb{R}^{dq}$. We show this by contradiction by assuming that the space $\mathcal{W}$ is not equal to $\mathbb{R}^{dq}$. Specifically, $\mathcal{W} = \bigcup_{i=1}^n W_i$ for proper subspaces $W_i \subsetneq \mathbb{R}^{dq}$. Let x be some non-zero vector from one of the subspaces, and consider the partition $x = (x_1', \dots, x_q')'$, $x_i \in \mathbb{R}^d$. Let $M_{(1)}^{(1,q)}, \dots, M_{(q)}^{(1,q)}$ denote q independent copies of $M_t^{(1,q)}$, and likewise let $M_{(1)}, \dots, M_{(q)}$ denote q independent copies of M_t, where the first d rows of $M_{(i)}$ are given by $M_{(i)}^{(1,q)}$. Then

$$M_{(1)} x = \begin{pmatrix} M_{(1)}^{(1,q)} x \\ x_1 \\ \vdots \\ x_{q-1} \end{pmatrix}.$$

Since $M_{(1)}^{(1,q)} x$ has a Lebesgue density strictly positive on $\mathbb{R}^d$, necessarily there must be a subspace W_{i_1} satisfying $\mathcal{V}_1 := \{(z_1', x_1', \dots, x_{q-1}')' \colon z_1 \in \mathbb{R}^d\} \subset W_{i_1}$. Next, the action $M_{(2)}$ on $\mathcal{V}_1$ yields,

$$M_{(2)} v_1 = \begin{pmatrix} M_{(2)}^{(1,q)} v_1 \\ z_1 \\ x_1 \\ \vdots \\ x_{q-2} \end{pmatrix}, \quad v_1 = (z_1', x_1', \dots, x_{q-1}')' \in \mathcal{V}_1.$$

Using again that $M_{(2)}^{(1,q)} v_1$ has a Lebesgue density strictly positive on $\mathbb{R}^d$ (for $v_1 \neq 0$), there must exist a subspace W_{i_2} such that $\mathcal{V}_2 := \{(z_1', z_2', x_1', \dots, x_{q-2}')' \colon z_1, z_2 \in \mathbb{R}^d\} \subset W_{i_2}$. By repeating these arguments we conclude that one of the subspaces W_i equals $\mathbb{R}^{dq}$. □

For $q = 1$ with $d = l = 2$ the following lemma is useful. The lemma is also applicable for checking the irreducibility condition in Alsmeyer and Mentemeier (2012); see Alsmeyer and Mentemeier (2012, Condition (A4)) and Theorem 4.4.15 of BDM.

Lemma 6.4. *Let $M_t\,(= M_{1,t} = M_t^{(1,1)}) = m_{1,t}A_1 + m_{2,t}A_2$ with A_1, $A_2 \in M(2,\mathbb{R})$ where $m_{i,t} \overset{d}{=} N(0,1)$, $i = 1,2$ are independent, namely SRE* (2.1) *reduces to SRE* (4.2). *With $x = (x_1, x_2)' \in \mathbb{R}^2$ we write $x^{(n)} = \Pi_{i=1}^n M_i x$. For any $x \neq 0$ assume that there exists $n \in \mathbb{N}$ such that almost surely*

$$A_1 x^{(n)},\ A_2 x^{(n)} \neq 0, \quad \text{and} \quad A_1 x^{(n)} \neq k A_2 x^{(n)} \quad \text{for any}\, k \in \mathbb{R},$$

i.e. the vectors $A_1 x^{(n)}$ and $A_2 x^{(n)}$ are not parallel. Then Assumption 6.1(a) holds.

Proof. It suffices to observe that since $m_{i,n+1} \overset{d}{=} N(0,1)$, $i = 1,2$ are independent,

$$M_{n+1} x^{(n)} = m_{1,n+1} A_1 x^{(n)} + m_{2,n+1} A_2 x^{(n)}$$

may take any value in $\mathbb{R}^2$. Thus Assumption 6.1(a) follows. □

The following lemma gives a sufficient condition for Assumption 6.1(b).

Lemma 6.5. *With M_t given in* (2.2)*, suppose that for any $i = 1, \dots, q$, $M_{i,t}$ has a density with respect to the Lebesgue measure on $M(d,\mathbb{R})$ that is strictly positive on a neighborhood around zero. Then Assumption 6.1(b) holds.*

Proof. The result is immediate by noting that $M_{i,t}$ and $M_{j,t}$ are independent for $i \neq j$. □

Remark 6.6. *As mentioned, Lemmas 6.3-6.4 state sufficient conditions for Assumption 6.1(a), and Lemma 6.5 states a sufficient condition for Assumption 6.1(b). It is possible to relax these conditions: For instance, under the conditions stated in Lemma 6.4 where $q = 1$, $M_t^{(1,1)}x$ may not have a Lebesgue density on $\mathbb{R}^2$ for some $x \neq 0$, as required in Lemma 6.3. As an example, consider the case*

$$A_1 = \begin{pmatrix} a & b \\ b & 0 \end{pmatrix}, \quad \text{and} \quad A_2 = \begin{pmatrix} a & a \\ a & 0 \end{pmatrix}, \quad a, b \neq 0, \quad a \neq b.$$

Here $M_t^{(1,1)}x = (m_{1,t}A_1 + m_{2,t}A_2)x$ does not have a Lebesgue density on $\mathbb{R}^2$ if the first entry of x is zero. On the other hand it is easy to check that the conditions stated in Lemma

6.4 hold for $n = 1$. For the case $q > 1$ the conditions stated in Lemma 6.3 appear more difficult to relax, and, in light of the fact that the entries of the matrix M_t are Gaussian, the conditions essentially correspond to the ones obtained for one-dimensional processes in BDM (Section 4.4.9). In terms of Assumption 6.1(b), we note that the conditions stated in Lemma 6.5 are not necessary. For instance, let $q = d = l = 2$, and

$$M_t = \begin{pmatrix} m_{1,1,t}a & 0 & m_{2,1,t}c & 0 \\ 0 & m_{1,2,t}b & 0 & m_{2,2,t}d \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

with (potentially complex) eigenvalues $bm_{1,2,t}/2 \pm (b^2 m_{1,2,t}^2 + 4dm_{2,2,t})^{1/2}/2$ and $am_{1,1,t}/2 \pm (a^2 m_{1,1,t}^2 + 4cm_{2,1,t})^{1/2}/2$. Clearly, the conditions of Lemma 6.5 are not satisfied, but using that $m_{i,j,t}$ is standard normal and independent across both i and j ($i, j = 1, 2$), we note that G_M contains a matrix with eigenvalues 0 (with multiplicity two) and $am_{1,1,t}/2 \pm |am_{1,1,t}|/2$, and, hence, for $a \neq 0$, we have that Assumption 6.1(b) holds.

The BEKK-ARCH process in the following example satisfies Assumption 6.1.

Example 6.7. *Consider the case $d = q = 2$ and $l = 4$ where*

$$A_{i1} = \begin{pmatrix} a_{i1} & 0 \\ 0 & 0 \end{pmatrix}, A_{i2} = \begin{pmatrix} 0 & 0 \\ a_{i2} & 0 \end{pmatrix}, A_{i3} = \begin{pmatrix} 0 & a_{i3} \\ 0 & 0 \end{pmatrix}, A_{i4} = \begin{pmatrix} 0 & 0 \\ 0 & a_{i4} \end{pmatrix}, \; i = 1, 2,$$

for some non-zero a_{ij}. Since all elements of the matrices $M_{1,t}$ and $M_{2,t}$ are independent and Gaussian, we have that $M_{1,t}$ and $M_{2,t}$ have densities strictly positive on $M(d, \mathbb{R})$. Moreover, for any non-zero $x \in \mathbb{R}^4$ the distribution of $M_t^{(1,2)} x$ has a density that is strictly positive on $\mathbb{R}^2$. By Lemmas 6.3 and 6.5, the process satisfies Assumption 6.1. We may also note that $\mathbb{P}[\det(M_{2,t}) = 0] = 0$. For suitable values of constants a_{ij} Assumption 3.1 holds and there exists $\alpha > 0$ such that $\inf_{n \in \mathbb{N}} (\mathbb{E}[\|M_1 \cdots M_n\|^\alpha])^{1/n} = 1$. Hence, under these conditions, Theorem 6.2 applies.

In the following example we consider the case where $q = 1$ and $d = l = 2$ and the matrices A_1 and A_2 are neither simultaneously diagonalizable nor simultaneously triangularizable.

Example 6.8. *Suppose that $q = 1$, $d = l = 2$ and consider (4.2) with $M_t = m_{1,t}A_1 + m_{2,t}A_2$ where*

$$A_1 = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad A_2 = \begin{pmatrix} a & b \\ -b & -a \end{pmatrix}, \quad |a| > |b|, \; a, b \neq 0.$$

Since the eigenvalues of A_1 *and* A_2 *are respectively* $a \pm b$ *and* $\pm\sqrt{a^2-b^2}$*,* A_1 *and* A_2 *are diagonalizable. However, due to non-commutability, they are not simultaneously diagonalizable. We check the conditions of Theorem 6.2. Define a set of vectors* $\mathcal{V} = \{x = (x_1, x_2)' \in \mathbb{R}^2 \mid x = (\pm b, \mp a)', (\pm a, \mp b)'\}$*. For non-zero* $x \in \mathbb{R}^2 \setminus \mathcal{V}$*,* $A_1 x$ *and* $A_2 x$ *are linearly independent, while for* $x \in \mathcal{V}$*,* $M_1 x$ *is proportional to either* $(0,1)' \notin \mathcal{V}$ *or* $(1,0)' \notin \mathcal{V}$*. Thus via Lemma 6.4, Assumption 6.1(a) holds. For Assumption 6.1(b) take* $n = 1$ *in* G_M *and observe that* $\det(M_{1,t} - \lambda I) = 0 \Leftrightarrow \lambda^2 - 2am_{1,t}\lambda + (a^2-b^2)(m_{1,t}^2 - m_{2,t}^2) = 0$*, so that the eigenvalues may differ. Note that* $\det(M_t) = (a^2-b^2)(m_{1,t}^2 + m_{2,t}^2) \neq 0$ *almost surely. In order to assure the existence of* $\alpha > 0$ *such that* $\inf_{n\in\mathbb{N}}(\mathbb{E}\|M_1 \cdots M_n\|^{\alpha})^{1/n} = 1$*, as in Remark 4.4.16 of BDM, it is enough to assume that for some* $p > 0$*,* $\mathbb{E}[(\lambda_{\min}(M_t M_t'))^{p/2}] \geq 1$ *where* $\lambda_{\min}(M_t M_t')$ *is the smallest eigenvalue of* $M_t M_t'$*. However, in view of the characteristic equation of* $M_t M_t'$*:*

$$\lambda^2 - 2(a^2+b^2)(m_{1,t}^2 + m_{2,t}^2)\lambda + (a^2-b^2)(m_{1,t}^2 - m_{2,t}^2) = 0,$$

non-negative eigenvalues of $M_t M_t'$ *are proportional to* (a,b) *and we can choose appropriate values. In a similar manner, we can adjust* (a,b) *so that Assumption 3.1 holds.*

7 Empirical illustration

We consider the daily continuously compounded returns on the Bitcoin/US dollar and XRP/US dollar cryptocurrencies from August 5, 2013 to January 7, 2017, see Figure 1.[3]

[Figure 1 about here.]

We seek to estimate BEKK models of the form considered in the previous sections based on the cryptocurrency returns and compare the estimated model-implied tail indexes to tail index estimates obtained from a conventional, non-parametric estimation approach. Specifically, we estimate various submodels of the two-dimensional BEKK-ARCH model of order $q = 1$ and with $l = 4$ as in (4.1). With $R_{t,\text{Bitcoin}}$ and $R_{t,\text{XRP}}$ the continuously compounded returns on Bitcoin/US Dollar and XRP/US Dollar, respectively, let $X_t = (R_{t,\text{Bitcoin}}, R_{t,\text{XRP}})'$. Moreover, we let

$$C = \tilde{C}\tilde{C}', \quad \tilde{C} = \begin{pmatrix} C_1 & 0 \\ C_2 & C_3 \end{pmatrix}, \quad C_1, C_3 > 0, \quad C_2 \in \mathbb{R},$$

[3] The daily exchange rates are daily volume-weighted average closing prices across various exchanges and obtained from https://www.coingecko.com/en.

where the inequality constraints ensure that C is positive definite, and

$$A_1 = \begin{pmatrix} A_{1,11} & A_{1,12} \\ 0 & A_{1,22} \end{pmatrix}, \quad A_2 = \begin{pmatrix} A_{2,11} & 0 \\ 0 & A_{2,22} \end{pmatrix}, \quad A_3 = \begin{pmatrix} A_{3,11} & 0 \\ 0 & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 \\ 0 & A_{4,22} \end{pmatrix}.$$

We consider four submodels given, respectively, by the restrictions

$$A_{1,11}, A_{1,22} \in \mathbb{R}, \quad A_{1,12} > 0, \quad A_{2,11}, A_{2,22} \geq 0, \quad A_{3,11} = A_{4,22} = 0, \tag{7.1}$$

$$A_{1,11}, A_{1,22} \in \mathbb{R}, \quad A_{1,12} > 0, \quad A_{2,11} = A_{2,22} = 0, \quad A_{3,11}, A_{4,22} \geq 0, \tag{7.2}$$

$$A_{1,11} > 0, \quad A_{1,22}, A_{1,12} \in \mathbb{R}, \quad A_{2,11} = A_{2,22} = A_{3,11} = 0, \quad A_{4,22} \geq 0, \tag{7.3}$$

$$A_{1,11} = A_{1,22} = A_{1,12} = 0, \quad A_{2,11} > 0, \quad A_{2,22} \in \mathbb{R}, \quad A_{3,11} = 0, \quad A_{4,22} \geq 0. \tag{7.4}$$

The inequality constraints are imposed in order to ensure identification; we refer to Engle and Kroner (1995, pp.127–130) for additional details on identification in BEKK models. Since the models are not closed with respect to the ordering of the elements of X_t, we also consider models with specifications (7.1)-(7.4) with X_t replaced by $X_t^\star = (R_{t,\text{XRP}}, R_{t,\text{Bitcoin}})'$, which we label (7.1*)-(7.4*). The model parameters are estimated by maximum likelihood.[4,5] Table 1 contains point estimates of the model parameters as well as the associated log-likelihood values (Lik) and Akaike (AIC) and Bayesian (BIC) information criteria. Both information criteria suggest that model (7.3) is to be preferred. Note that the nested models (7.3) and (7.4) have very similar log-likelihood values, and based on a likelihood ratio test with a conventional critical value, one cannot reject the hypothesis that $A_{1,12} = 0$. For the remainder of the analysis we hence focus on the parameter estimates for the models (7.3) and (7.4).

[Table 1 about here.]

Next, we estimate the model-implied tail index of the two cryptocurrency returns based on the parameter estimates for model (7.3) and (7.4). In terms of model (7.3), the matrices $A_1, \ldots, A_4$ are upper triangular (and, hence, trivially simultaneously triangularizable) such that we may apply Theorem 5.1 to determine the tail indexes. Specifically, given the ordering of the returns in X_t, the tail index of $R_{t,\text{XRP}}$ corresponds to α_2 in Theorem 5.1, such that $\mathbb{E}[|(A_{1,22}^2 + A_{4,22}^2)^{1/2} z|^{\alpha_2}] = 1$ with $z \overset{d}{=} N(0,1)$. The estimate of α_2 is then found by solving the equation numerically using the point estimates of $A_{1,22}$ and $A_{4,22}$. From Theorem 5.1 we have the tail index of $R_{t,\text{Bitcoin}}$ can be either α_2 or α_1, where $\mathbb{E}[|A_{1,11} z|^{\alpha_1}] =$

[4]We notice that the maximum likelihood estimator is consistent (under suitable conditions) even if only a fractional moment of $|X_t|$ is finite. We refer to Hafner and Preminger (2009) and Avarucci et al. (2013) for additional details.

[5]The estimation is carried out in OxMetrics 8.0. The log-likelihood function is maximized using the `MaxSQP` algorithm, imposing the parameter constraints.

1 with $z \overset{d}{=} N(0,1)$. In particular, in light of Remark 5.2, if $\alpha_2 < \alpha_1$ and $A_{1,12} \neq 0$, then the tail index is α_2. Based on the estimates of the model parameters we have that the estimates of α_1 and α_2 are, respectively, 3.87 and 3.03, suggesting that the tail index estimate for both returns should be 3.03. However, the aforementioned likelihood ratio test suggests that $A_{1,12} = 0$ such that tail index estimate for $R_{t,\text{Bitcoin}}$, based on model (7.3), should be 3.87. We report the latter estimate in the second column of Table 2. In the third column of the table, we report the tail index estimates based on model (7.4) that has diagonal matrices $A_1, \ldots, A_4$. The tail index estimates for this model are determined along the lines of Theorem 4.3 and Remark 4.2 using numerical methods. We note that the estimates are very similar to the ones obtained from model (7.3). For comparison, the fourth and fifth columns state the tail index estimates obtained from log-log-rank-size regressions, as considered in Gabaix and Ibragimov (2011), based, respectively, on the 5 and 10 percent largest observations of $|R_{t,\text{XRP}}|$ and $|R_{t,\text{Bitcoin}}|$. These estimates are smaller than the model-implied estimates but are subject to rather large estimation uncertainty, as indicated by the confidence intervals. In particular, we note that for both returns the model-implied tail index estimate lies within the confidence interval based on $\hat{\alpha}_{\text{RS},5\%}$, i.e. the estimate using the 5 percent largest observations.

[Table 2 about here.]

8 Concluding remarks

We conclude by stating some important directions for future research. For the cases considered in Sections 4 and 5, we focused on the *component-wise* tail behavior of X_t. Ideally, one would also be interested in obtaining results for the tail dependence structure of X_t, as this can be used for establishing stable limit theory for X_t, see e.g. Section 4.5 of BDM and Pedersen and Wintenberger (2018, Section 4). As the components, or marginals, of X_t have different indexes of regular variation, it seems appealing to find conditions such that X_t is *non-standard regularly varying* in the sense of Resnick (2007, Section 6.5.6) or *vector scaling regularly varying* as introduced in Pedersen and Wintenberger (2018). Finding such conditions for general multivariate SREs is a tremendous task and an active area of research.

The SRE representation for the BEKK-ARCH process in (2.1)-(2.2) relies crucially on the assumption that the noise variable Z_t is Gaussian and that the process is of the ARCH-type of finite order, i.e. H_t does not include lagged values of itself. Characterizing the tail behavior of general GARCH-type BEKK processes with non-Gaussian noise is indeed an interesting open issue that inherently seems to require another approach than relying on SRE representations of the processes.

A Appendix: Auxiliary results on regular variation and one-dimensional SREs

The following definitions and results can be found in the recent monograph by BDM. A positive measurable function f on $(0, \infty)$ is said to be regularly varying with index $\kappa \in \mathbb{R}$, if for any constant $c > 0$, $\lim_{x\to\infty} f(cx)/f(x) = c^{\kappa}$. We say that an $\mathbb{R}$-valued random variable X is regularly varying with index $\alpha \geq 0$ if the function $f(x) = \mathbb{P}(|X| > x)$ is regularly varying with index $-\alpha$ and there exist constants $p, q \geq 0$ such that $p + q = 1$ and

$$\lim_{x\to\infty} \frac{\mathbb{P}(X > x)}{\mathbb{P}(|X| > x)} = p \quad \text{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(X \leq -x)}{\mathbb{P}(|X| > x)} = q. \tag{A.1}$$

Note that if X is regularly varying with index $\alpha > 0$, then $\mathbb{E}[|X|^{\delta}] < \infty$ for any $0 \leq \delta < \alpha$, and $\mathbb{E}[|X|^{\tilde{\delta}}] = \infty$ for any $\tilde{\delta} > \alpha$.

The following result is a generalization of Breiman's (1965) lemma, and is useful for characterizing the product of a regularly varying random variable and a lighter-tailed random variable.

Lemma A.1. *Let X and Y be independent random variables. Assume that X is regularly varying with index $\alpha > 0$, and that there exists an $\varepsilon > 0$ such that $\mathbb{E}|Y|^{\alpha+\varepsilon} < \infty$. Then XY is regularly varying with index α. In particular,*

$$\lim_{x\to\infty} \frac{\mathbb{P}(XY > x)}{\mathbb{P}(|X| > x)} = p\mathbb{E}Y_+^{\alpha} + q\mathbb{E}Y_-^{\alpha}, \quad \textit{and} \quad \lim_{x\to\infty} \frac{\mathbb{P}(XY < -x)}{\mathbb{P}(|X| > x)} = p\mathbb{E}Y_-^{\alpha} + q\mathbb{E}Y_+^{\alpha}, \tag{A.2}$$

where the constants p and q are given by (A.1).

Proof. Note that $X = X_+ - X_-$ and $Y = Y_+ - Y_-$. Hence for $x > 0$ we have

$$\mathbb{P}(XY > x) = \mathbb{P}((X_+ - X_-)(Y_+ - Y_-) > x) = \mathbb{P}(X_+Y_+ > x) + \mathbb{P}(X_-Y_- > x), \tag{A.3}$$

and the first part of (A.2) follows by an application of Breiman's lemma (c.f. Lemma B.5.1 in BDM) to each term in (A.3). The second part of (A.2) follows by a similar argument. □

The following result states that regular variation is closed under convolution. A proof is given in Section B.6 of BDM.

Lemma A.2. *Let X and Y be random variables, and assume that X_+ is regularly varying with index $\alpha > 0$ such that $\mathbb{P}(|Y| > x) = o(\mathbb{P}(|X| > x))$ as $x \to \infty$. Then*

$$\mathbb{P}(X + Y > x)/\mathbb{P}(X > x) \to 1 \quad \textit{as } x \to \infty. \tag{A.4}$$

Lastly, we state the following lemma about the strictly stationary solution to one-dimensional SREs. The first result on strict stationarity is given in Theorem 2.1.3 of BDM, but has been stated elsewhere in the literature under similar assumptions, see e.g. Bougerol and Picard (1992). The second part on regular variation is given in Theorem 2.4.7 of BDM and was originally proved by Goldie (1991).

Lemma A.3. *Let X_t be an $\mathbb{R}$-valued random variable satisfying the SRE*

$$X_t = A_t X_{t-1} + B_t, \quad t \in \mathbb{Z}, \tag{A.5}$$

with $((A_t, B_t) : t \in \mathbb{Z})$ an $\mathbb{R}^2$-valued i.i.d. sequence.

Suppose that $\mathbb{P}(A_t = 0) = 0$, $-\infty \leq \mathbb{E}[\log|A|] < 0$, and $\mathbb{E}[(\log|B|)_+] < \infty$. Then there exists an almost surely unique causal ergodic strictly stationary solution to the SRE in (A.5). Let $\mathbb{P}_0$ denote the distribution of the strictly stationary solution.

Suppose in addition that (1) $\mathbb{P}(A_t < 0) > 0$ and the conditional distribution of $\log|A_t|$ given $A_t \neq 0$ is non-arithmetic, (2) there exists an $\alpha > 0$ such that $\mathbb{E}[|A_t|^\alpha] = 1$, $\mathbb{E}[|B_t|^\alpha] < \infty$, and $\mathbb{E}[|A_t|^\alpha(\log|A_t|)_+] < \infty$, and (3) $\mathbb{P}(A_t x + B_t = x) < 1$ for all $x \in \mathbb{R}$.

Let (A, B) have the same distribution as (A_t, B_t). Then the stochastic fixed point equation

$$X \stackrel{d}{=} AX + B \tag{A.6}$$

has a solution X which is independent of (A, B) and that has distribution $\mathbb{P}_0$. Moreover, there exists a constant $c_+ > 0$ such that

$$\mathbb{P}_0(X > x) \sim c_+ x^{-\alpha} \quad \textit{and} \quad \mathbb{P}_0(X < -x) \sim c_+ x^{-\alpha}, \quad \textit{as } x \to \infty, \tag{A.7}$$

where

$$c_+ = \frac{1}{2\alpha m_\alpha}\mathbb{E}[|AX + B|^\alpha - |AX|^\alpha] \quad \textit{and} \quad m_\alpha = \mathbb{E}[|A|^\alpha \log|A|] > 0. \tag{A.8}$$

B Appendix: Proof of Theorem 5.1

Throughout the proof, we apply the following component-wise series representations of the unique stationary solution to (5.2), $Y_t = (Y_{1,t}, Y_{2,t})$, which are given by

$$Y_{1,t} = \sum_{i=1}^{\infty} \Pi^{(1)}_{t,t+2-i} D_{t+1-i}, \quad \text{where} \quad D_t = M_{12,t} Y_{2,t-1} + Q_{1,t}, \tag{B.1}$$

$$Y_{2,t} = \sum_{i=1}^{\infty} \Pi^{(2)}_{t,t+2-i} Q_{2,t+1-i}. \tag{B.2}$$

We start out by verifying that these representations are well-defined. The expression (B.2) follows easily from (5.4), and the series converges absolutely almost surely (see e.g. proof of Theorem 2.1.3 in BDM). Turning to (B.1), consider the SRE (5.3). Note that the random element $Y_{2,t}$ is measurable w.r.t. the σ-field generated by $(M_{t-s}, Q_{t-s})_{s\in\mathbb{Z}_-}$ (see Section 2.6 of Straumann, 2005), and so is $(M_{11,t}, Q_{1,t})$. Thus $(M_{11,t}, D_t)$ with $D_t = M_{12,t}Y_{2,t-1} + Q_{1,t}$ is also measurable, where we notice that component-wise measurability is equivalent to the measurability of a vector. Then due to e.g. Proposition 4.3 of Krengel (2011) , $(M_{11,t}, D_t)_{t\in\mathbb{Z}}$ is a stationary and ergodic sequence. Using that $\mathbb{E}[\log|M_{11}|] < 0$ and $\mathbb{E}[(\log|Y_2|)_+] < \infty$, it follows by Theorem 1 of Brandt (1986) that (B.1) is the unique stationary solution to (5.3) and that the series converges absolutely almost surely. Since (5.2) has a unique solution, we conclude that the solution, $(Y_{1,t}, Y_{2,t})$ of (B.1) and (B.2) is the component-wise series representation.

We start out by considering the solutions to the decomposed SREs in (5.5) and (5.6). By the same reasoning as above, these SREs have unique solutions, respectively,

$$\widehat{Y}_{1,0} = \sum_{i=1}^{\infty} \Pi^{(1)}_{0,2-i} Q_{1,1-i}, \tag{B.3}$$

and

$$\widetilde{Y}_{1,0} = \sum_{i=1}^{\infty} \Pi^{(1)}_{0,2-i} M_{12,1-i} Y_{2,-i}, \tag{B.4}$$

where the series converge absolutely almost surely. Thus we have that

$$Y_{1,0} = \widetilde{Y}_{1,0} + \widehat{Y}_{1,0}. \tag{B.5}$$

Proof. Throughout c denotes a generic positive constant.
(i) Case $\alpha_1 > \alpha_2$. Our strategy is that we further decompose $\widetilde{Y}_{1,0}$ into several parts. By comparing their tail behaviors we specify the dominant term, which determines the tail behavior of Y_1. First we show the general scheme. The detailed tail asymptotics of the dominant and negligible terms are given later. Without loss of generality, we consider the upper tail $\mathbb{P}(Y_1 > x)$. Observe that in (B.5), $\widehat{Y}_1$ is regularly varying with index α_1, i.e.

$$\mathbb{P}(\widehat{Y}_{1,0} > x) \sim \hat{c}_1 x^{-\alpha_1}$$

for a positive constant $\hat{c}_1$, as the SRE (5.5) satisfies the conditions of Lemma A.3. We next turn to the tail properties of $\widetilde{Y}_1$. We decompose $\widetilde{Y}_1$ into three parts,

$$\widetilde{Y}_{1,0} = \Big(\underbrace{\sum_{i=1}^{s}}_{\widetilde{Z}_s} + \underbrace{\sum_{i=s+1}^{\infty}}_{\widetilde{Y}^s} \Big) \Pi^{(1)}_{0,2-i} M_{12,1-i} Y_{2,-i} =: \underbrace{\widetilde{Y}_{s,1} + \widetilde{Y}_{s,2}}_{\widetilde{Z}_s} + \widetilde{Y}^s, \tag{B.6}$$

where in $\widetilde{Z}_s$ we apply the iteration of the SRE for $Y_{2,-i}$ until time $-s < -i$,

$$Y_{2,-i} = \Pi^{(2)}_{-i,1-s} Y_{2,-s} + \sum_{k=0}^{s-i-1} \Pi^{(2)}_{-1,1-i-k} Q_{2,-i-k},$$

and substitute this into $\widetilde{Z}_s$, so that

$$\widetilde{Z}_s = \underbrace{\sum_{i=1}^{s} \Pi^{(1)}_{0,2-i} M_{12,1-i} \Pi^{(2)}_{-i,1-s} Y_{2,-s}}_{\widetilde{Y}_{s,1}} + \underbrace{\sum_{i=1}^{s} \Pi^{(1)}_{0,2-i} M_{12,1-i} \sum_{k=0}^{s-i-1} \Pi^{(2)}_{-i,1-i-k} Q_{2,-i-k}}_{\widetilde{Y}_{s,2}}. \tag{B.7}$$

The idea is then to study the tail behavior of each term in (B.6). Specifically, we later show that there are constants $C > 0$, $0 < \eta < 1$ such that for every s

$$\mathbb{P}(|\widetilde{Y}^s| > x) \le C\eta^s x^{-\alpha_2}. \tag{B.8}$$

Moreover, for a fixed (but arbitrary) s

$$\lim_{x\to\infty} \mathbb{P}(|\widetilde{Y}_{s,2}| > x) x^{\alpha_2} = 0 \tag{B.9}$$

and

$$\lim_{x\to\infty} \mathbb{P}(\widetilde{Y}_{s,1} > x) x^{\alpha_2} = c_2 w_s, \tag{B.10}$$

where c_2 is that in (5.11) and

$$w_s = \mathbb{E}\Big| \sum_{i=1}^{s} \Pi^{(1)}_{0,2-i} \Pi^{(2)}_{-i,1-s} M_{12,1-i} \Big|^{\alpha_2}$$

with

$$\sup_{s\in\mathbb{N}} w_s < \infty. \tag{B.11}$$

Hence we note that the term $\widetilde{Y}_{s,1}$ is the dominating term in (B.6). Now, using (B.6), we have that

$$\mathbb{P}(Y_1 > x) \le \mathbb{P}(\widetilde{Y}_{s,1} > (1-3\varepsilon)x) + \mathbb{P}(\widehat{Y}_1 > \varepsilon x) + \mathbb{P}(\widetilde{Y}_{s,2} > \varepsilon x) + \mathbb{P}(\widetilde{Y}^s > \varepsilon x),$$
$$\mathbb{P}(Y_1 > x) \ge \mathbb{P}(\widetilde{Y}_{s,1} > (1+3\varepsilon)x) - \mathbb{P}(\widehat{Y}_1 < -\varepsilon x) - \mathbb{P}(\widetilde{Y}_{s,2} < -\varepsilon x) - \mathbb{P}(\widetilde{Y}^s < -\varepsilon x).$$

Then after multiplying x^{α_2} to both sides of inequalities, we make the limit operation of $x \to \infty$ and obtain

$$\begin{aligned}(1+3\varepsilon)^{-\alpha_2} c_2 w_s - C\eta^s &\le \liminf_{x\to\infty} x^{\alpha_2}\mathbb{P}(Y_1 > x) \\ &\le \limsup_{x\to\infty} x^{\alpha_2}\mathbb{P}(Y_1 > x) \\ &\le (1-3\varepsilon)^{-\alpha_2} c_2 w_s + C\eta^s,\end{aligned} \tag{B.12}$$

where C may depend on ε. In the upper and lower bounds, we take a converging subsequence w_{s_k} of w_s, which is possible by (B.11), and then $\varepsilon \downarrow 0$. Note here that it is not possible to change the order of the limits. Due to (B.12), the limit for $k \to \infty$ satisfies

$$c_2 \lim_{k\to\infty} w_{s_k} = \lim_{x\to\infty} x^{\alpha_2}\mathbb{P}(Y_1 > x).$$

Since every converging subsequence converges to the same limit, since there is no oscillation, we have

$$\lim_{x\to\infty} x^{\alpha_2}\mathbb{P}(Y_1 > x) = c_2 \lim_{s\to\infty} w_s.$$

It remains to prove (B.8)–(B.11). We begin with (B.11) and recall from the argument below (5.9) that $\mathbb{E}|M_{11}|^{\alpha_2} < 1$. If $\alpha_2 \le 1$ then

$$w_s \le \sum_{i=1}^{\infty} (\mathbb{E}|M_{11}|^{\alpha_2})^{(i-1)} \mathbb{E}|M_{12}|^{\alpha_2} < \infty$$

and if $\alpha_2 > 1$ then

$$w_s^{1/\alpha_2} \le \sum_{i=1}^{\infty} (\mathbb{E}|M_{11}|^{\alpha_2})^{(i-1)/\alpha_2} (\mathbb{E}|M_{12}|^{\alpha_2})^{1/\alpha_2} < \infty.$$

Since the bounds above do not depend on s, (B.11) follows. Concerning (B.8) we use the Markov inequality and obtain

$$\begin{aligned}x^{\alpha_2}\mathbb{P}(|\widetilde{Y}^s| > x) &= x^{\alpha_2}\mathbb{P}\Big(\Big|\sum_{i=s+1}^{\infty} \Pi^{(1)}_{0,2-i} M_{12,1-i} Y_{2,-i}\Big| > x\Big) \\ &\le x^{\alpha_2}\mathbb{P}\Big(\sum_{i=1}^{\infty} |\Pi^{(1)}_{0,2-(s+i)} M_{12,1-(s+i)} Y_{2,-(s+i)}| > x\Big)\end{aligned}$$

$$\le \sum_{i=1}^{\infty} x^{\alpha_2}\mathbb{P}(|\Pi^{(1)}_{0,2-(s+i)} M_{12,1-(s+i)} Y_{2,-(s+i)}| > x\, i^{-\mu}\, /\zeta(\mu))$$
$$= \sum_{i=1}^{\infty} \mathbb{E}[x^{\alpha_2}\mathbb{P}(G_i|Y_{2,-(s+i)}| > x \mid G_i)],$$

where $G_i = \zeta(\mu)\, i^{\mu}\, |\Pi^{(1)}_{0,2-(i+s)} M_{12,1-(s+i)}|$ with $\mu > 1$ and $\zeta(\cdot)$ is zeta function. The integrability above follows from the fact $\sum_{i=1}^{\infty} i^{-\mu} = \zeta(\mu)$. Notice that G_i and $|Y_{2,-(i+s)}|$ are independent, $\mathbb{E}G_i^{\alpha} < \infty$ and there is a constant c such that for every $x > 0$,

$$\mathbb{P}(|Y_{2,-(i+s)}| > x) \le c x^{-\alpha_2}.$$

Hence it follows from

$$\mathbb{E}[x^{\alpha_2}\mathbb{P}(G_i|Y_{2,-(s+i)}| > x \mid G_i)] \le c\mathbb{E}G_i^{\alpha_2}$$

that

$$\begin{aligned}
x^{\alpha_2}\mathbb{P}(|\widetilde{Y}^s| > x) &\le c \sum_{i=1}^{\infty} \mathbb{E}G_i^{\alpha_2} \\
&= c \sum_{i=1}^{\infty} \mathbb{E}|\Pi_{0,2-(s+i)}|^{\alpha_2}\mathbb{E}|M_{12,1-(s+i)}|^{\alpha_2}\zeta(\mu)^{\alpha_2}\, i^{\alpha_2\mu} \\
&\le \underbrace{c\mathbb{E}|M_{12}|^{\alpha_2}\zeta(\mu)^{\alpha_2} \sum_{i=1}^{\infty}(\mathbb{E}|M_{11}|^{\alpha_2})^{i-1} i^{\alpha_2\mu}}_{C} \cdot(\mathbb{E}|M_{11}|^{\alpha_2})^{s},
\end{aligned}$$

where $\sum_{i=1}^{\infty}(\mathbb{E}|M_{11}|^{\alpha_2})^{i-1} i^{\alpha_2\mu} < \infty$ since $\mathbb{E}|M_{11}|^{\alpha_2} < 1$. Putting $\eta = \mathbb{E}|M_{11}|^{\alpha_2}$, we obtain (B.8). We prove (B.9) by showing that $\mathbb{E}|\widetilde{Y}_{s,2}|^{\alpha_2} < \infty$ for any fixed s. We work on the expression in (B.7). Recall that (M_t, Q_t) are i.i.d. so that $\Pi^{(1)}_{0,2-i} M_{12,1-i}$ and $\sum_{k=0}^{s-i-1} \Pi^{(2)}_{-i,1-i-k} Q_{2,-i-k}$, $i = 1, 2, \dots, s$ are independent. We further recall that $\mathbb{E}|M_{22}|^{\alpha_2} = 1$, $\mathbb{E}|M_{11}|^{\alpha_2} < 1$, $\mathbb{E}|M_{12}|^{\alpha_2} < \infty$ and $\mathbb{E}|Q_2|^{\alpha_i} < \infty$. For $\alpha_2 > 1$, by Minkowski's inequality,

$$\begin{aligned}
\mathbb{E}|\widetilde{Y}_{s,2}|^{\alpha_2} &= \mathbb{E}\Big| \sum_{i=1}^{s} \Pi^{(1)}_{0,2-i} M_{12,-i} \sum_{k=0}^{s-i-1} \Pi^{(2)}_{-i,1-i-k} Q_{2,-i-k}\Big|^{\alpha_2} \\
&\le \Big[\sum_{i=1}^{s} \Big\{(\mathbb{E}|M_{11}|^{\alpha_2})^{i-1}\mathbb{E}|M_{12}|^{\alpha_2}\mathbb{E}|Q_2|^{\alpha_2}(s-i)^{\alpha_2}\Big\}^{1/\alpha_2}\Big]^{\alpha_2} < \infty
\end{aligned}$$

and for $\alpha_2 \le 1$, by sub-additivity,

$$\mathbb{E}|\widetilde{Y}_{s,2}|^{\alpha_2} \le \sum_{i=1}^{s} \mathbb{E}\Big|\Pi^{(1)}_{0,2-i} M_{12,-i} \sum_{k=0}^{s-i-1} \Pi^{(2)}_{-i,1-i-k} Q_{2,-i-k}\Big|^{\alpha_2}$$

$$\leq \sum_{i=1}^{s} (\mathbb{E}|M_{11}|^{\alpha_2})^{i-1}\mathbb{E}|M_{12}|^{\alpha_2} \sum_{k=0}^{s-i-1} (\mathbb{E}|M_{22}|^{\alpha_2})^k \mathbb{E}|Q_2|^{\alpha_2}$$
$$= \mathbb{E}|M_{12}|^{\alpha_2}\mathbb{E}|Q_2|^{\alpha_2} \sum_{i=1}^{s} (\mathbb{E}|M_{11}|^{\alpha_2})^{i-1}(s-i) < \infty.$$

Hence for fixed s we have (B.9). Finally we observe that

$$\widetilde{Y}_{s,1} = R_s Y_{2,-s},$$

where $R_s := \sum_{i=1}^{s} \Pi^{(1)}_{0,2-i} M_{12,-i} \Pi^{(2)}_{-i,1-s}$ and $Y_{2,-s}$ are independent. Hence Breiman's lemma (Lemma A.1) yields

$$\lim_{x\to\infty} x^{\alpha_2}\mathbb{P}(\widetilde{Y}_{s,1} > x) = c_2 w_s,$$

which is (B.10).

Next, we verify that $\tilde{c}_1 > 0$. This has been established in Damek and Matsui (2019) in a more general setting, and we here provide a brief explanation. Since $\mathbb{P}(M_{22} = 0) = 0$ by Gaussianity, we may write

$$|R_s| = |\Pi^{(2)}_{0,1-s}| \Big| \sum_{i=1}^{s} \Pi^{(1)}_{0,2-i}(\Pi^{(2)}_{0,2-i})^{-1} M_{12,1-i} M^{-1}_{22,1-i} \Big| = |\Pi^{(2)}_{0,1-s}| \Big| \sum_{i=1}^{s} \widetilde{A}_0 \cdots \widetilde{A}_{2-i} \widetilde{B}_{1-i} \Big|,$$

where

$$\widetilde{A}_i = M_{11,i} M^{-1}_{22,i}, \quad \widetilde{B}_i = M_{12,i} M^{-1}_{22,i} \quad \text{and} \quad \widetilde{A}_0 \widetilde{A}_1 = 1.$$

Then we apply a change of measure: let $\mathcal{F}_n$ be the filtration defined by the matrix sequence $(\mathbf{A}_i, \mathbf{B}_i)$; $\mathcal{F}_n = \sigma((\mathbf{A}_i, \mathbf{B}_i)_{-n\leq i\leq 0})$. The expectation $\mathbb{E}_0$ so that the new probability measure $\mathbb{P}_0$ is defined by $\mathbb{E}_0[Z] = \mathbb{E}[|\Pi_{0,-n}|^{\alpha_2} Z]$ where Z is measurable w.r.t. $\mathcal{F}_n$. Since

$$\mathbb{E}_0|\widetilde{A}_0|^{\alpha_1} = \mathbb{E}|M_{11,0}|^{\alpha_2} < 1 \quad \text{and} \quad \mathbb{E}_0|\widetilde{B}_0|^{\alpha_2} = \mathbb{E}|M_{12,0}|^{\alpha_2} < \infty \tag{B.13}$$

by the same reasoning as that for (B.11), the series $R = \sum_{i=1}^{\infty} \widetilde{A}_0 \cdots \widetilde{A}_{2-i} \widetilde{B}_{1-i}$ exists *a.s.* and moreover

$$\lim_{s\to\infty} w_s = \lim_{s\to\infty} \mathbb{E}_0 \Big| \sum_{i=1}^{s} \widetilde{A}_0 \cdots \widetilde{A}_{2-i} \widetilde{B}_{1-i} \Big|^{\alpha_2} = \mathbb{E}_0 \Big| \sum_{i=1}^{\infty} \widetilde{A}_0 \cdots \widetilde{A}_{2-i} \widetilde{B}_{1-i} \Big|^{\alpha_2} =: w.$$

Thus for positivity of $\tilde{c}_1$, it suffices to prove that

$$w = \mathbb{E}_0|R|^{\alpha_2} \neq 0, \quad i.e. \quad \mathbb{P}(R \neq 0) > 0.$$

Observe that

$$R_t = \sum_{i=1}^{\infty} \widetilde{A}_t \cdots \widetilde{A}_{t+2-i} \widetilde{B}_{t+1-i}$$

is a stationary solution of the SRE: $X_t = \widetilde{A}_t X_{t-1} + \widetilde{B}_t$. Indeed, by Jensen's inequality, the moment conditions (B.13) imply

$$\mathbb{E}_0[\log |\widetilde{A}_0|] < 0 \quad \text{and} \quad \mathbb{E}_0[(\log |\widetilde{B}_0|)_+] < \infty,$$

which assures that R_t is the unique stationary solution. Hence if $\mathbb{P}(R_t = 0) = 1$ it must hold that $\mathbb{P}(\widetilde{B}_t = 0) = 1$, which is impossible since $\mathbb{P}(M_{12} > 0) > 0$ (by assumption). This finishes the first part of the proof.

(ii) Case $\alpha_1 < \alpha_2$. By stationarity we have from SRE (5.3) that

$$Y_{1,0} = D_0 + M_{11,0} Y_{1,-1},$$

where $Y_{1,-1}$ has the same law as $Y_{1,0}$ and independent of $M_{11,0}$. We apply Theorem 2.3 Case 2 of Goldie (1991) that states that if

$$I_+ = \int_0^\infty \left| \mathbb{P}(Y_{1,-1} > x) - \mathbb{P}(M_{11,0} Y_{1,-1} > x) \right| x^{\alpha_1 - 1} dx < \infty$$

and

$$I_- = \int_0^\infty \left| \mathbb{P}(Y_{1,-1} < -x) - \mathbb{P}(M_{11,0} Y_{1,-1} < -x) \right| x^{\alpha_1 - 1} dx < \infty,$$

then

$$\begin{aligned} \lim_{x\to\infty} \mathbb{P}(Y_{1,0} > x) x^{\alpha_1} &= \lim_{x\to\infty} \mathbb{P}(Y_{1,0} < -x) x^{\alpha_1} \\ &= \frac{1}{2m_1} \int_0^\infty (\mathbb{P}(|Y_1| > x) - \mathbb{P}(|M_{11} Y_{1,-1}| > x)) x^{\alpha_1 - 1} dx. \end{aligned} \tag{B.14}$$

Due to Goldie (1991, Lemma 9.4), if I_+, $I_- < \infty$ then the right-hand side in (B.14) equals $\overline{c}_1$, given in Theorem 5.1, where we notice that $|x|^{\alpha_1} = x_+^{\alpha_1} + x_-^{\alpha_1}$ for any $x \in \mathbb{R}$. We focus on showing that $I_+ < \infty$ since the proof of $I_- < \infty$ follows by similar arguments replacing $Y_{1,-1}$ by $(-Y_{1,-1})$ in I_+. In view of Goldie (1991, Lemma 9.4), we have that

$$\begin{aligned} I_+ &= \int_0^\infty \left| \mathbb{P}(Y_{1,-1} > x) - \mathbb{P}(M_{11,0} Y_{1,-1} > x) \right| x^{\alpha_1 - 1} dx \\ &= \int_0^\infty \left| \mathbb{P}(D_0 + M_{11,0} Y_{1,-1} > x) - \mathbb{P}(M_{11,0} Y_{1,-1} > x) \right| x^{\alpha_1 - 1} dx \end{aligned}$$

$$= \frac{1}{\alpha_1}\mathbb{E}\big[\big|(D_0 + M_{11,0}Y_{1,-1})_+^{\alpha_1} - (M_{11,0}Y_{1,-1})_+^{\alpha_1}\big|\big],$$

which holds regardless of whether I_+ is finite or infinite. By elementary inequalities we observe that for $\alpha_1 \leq 1$,

$$\big|(D_0 + M_{11,0}Y_{1,-1})_+^{\alpha_1} - (M_{11,0}Y_{1,-1})_+^{\alpha_1}\big| \leq |D_0|^{\alpha_1}.$$

Using that $\alpha_1 < \alpha_2$, we have that $E[|D_0|^{\alpha_1}] < \infty$ such that $I_+ < \infty$. It remains to consider the case $\alpha_1 > 1$, where we note that

$$\begin{aligned}
&\big|(D_0 + M_{11,0}Y_{1,-1})_+^{\alpha_1} - (M_{11,0}Y_{1,-1})_+^{\alpha_1}\big| \\
&\quad \leq \alpha_1(D_0 + M_{11,0}Y_{1,-1})_+^{\alpha_1-1}\{(D_0 + M_{11,0}Y_{1,-1})_+ - (M_{11,0}Y_{1,-1})_+\}I_{\{D_0>0\}} \\
&\qquad + \alpha_1(M_{11,0}Y_{1,-1})_+^{\alpha_1-1}\{(M_{11,0}Y_{1,-1})_+ - (D_0 + M_{11,0}Y_{1,-1})_+\}I_{\{D_0<0\}} \\
&\quad \leq \alpha_1(|D_0| + |M_{11,0}Y_{1,-1}|)^{\alpha_1-1}|D_0|.
\end{aligned}$$

Thus we have

$$I_+ \leq c(\mathbb{E}|D_0|^{\alpha_1} + \mathbb{E}|M_{11,0}Y_{1,-1}|^{\alpha_1-1}|D_0|),$$

where we use Minkowski's inequality and sub-additivity of concave functions depending on whether $\alpha_1 > 2$ or $1 < \alpha_1 \leq 2$. We need to prove that $\mathbb{E}[|M_{11,0}Y_{1,-1}|^{\alpha_1-1}|D_0|] < \infty$. Note that

$$\begin{aligned}
&\mathbb{E}[|M_{11,0}Y_{1,-1}|^{\alpha_1-1}|D_0|] \\
&\ \leq \mathbb{E}[|M_{11,0}Y_{1,-1}|^{\alpha_1-1}(|Q_{1,0}| + |M_{12,0}Y_{2,-1}|)] \\
&\ \leq \mathbb{E}[|M_{11,0}|^{\alpha_1-1}|Q_{1,0}|]\mathbb{E}[|Y_{1,-1}|^{\alpha_1-1}] + \mathbb{E}[|M_{11,0}|^{\alpha_1-1}|M_{12,0}|]\,\mathbb{E}[|Y_{1,-1}|^{\alpha_1-1}|Y_{2,-1}|].
\end{aligned}$$

By Hölder's inequality $\mathbb{E}|M_{11,0}|^{\alpha_1-1}|Q_{1,0}|$ and $\mathbb{E}|M_{11,0}|^{\alpha_1-1}|M_{12,0}|$ are finite, since all quantities included have finite moments of any (finite) order. We now show that $\mathbb{E}|Y_{1,-1}|^{\alpha_1-1}|Y_{2,-1}|$ is finite, and note that $\mathbb{E}|Y_{1,-1}|^{\alpha_1-1} < \infty$ follows by a similar argument. Choose some small $\varepsilon > 0$ such that $p := (\alpha_1-\varepsilon)/(\alpha_1-1) > 0$ and $q := p/(p-1) < \alpha_2$. By Hölder's inequality,

$$\mathbb{E}|Y_{1,0}|^{\alpha_1-1}|Y_{2,0}| \leq \big(\mathbb{E}|Y_{1,0}|^{p(\alpha_1-1)}\big)^{1/p}\big(\mathbb{E}|Y_{2,0}|^q\big)^{1/q}.$$

With $\beta := \alpha_1 - \varepsilon > 0$ by applying Minkowski's inequality to $Y_{1,0} = \sum_{i=0}^{\infty}\Pi_{0,1-i}^{(1)}D_{-i}$, we obtain

$$\big(\mathbb{E}|Y_{1,0}|^\beta\big)^{1/\beta} \leq \sum_{i=0}^{\infty}\big(\mathbb{E}|\Pi_{0,1-i}^{(1)}D_{-i}|^\beta\big)^{1/\beta} = \sum_{i=0}^{\infty}\big(\mathbb{E}|M_{11,0}|^\beta\big)^{i/\beta}\big(\mathbb{E}|D_0|^\beta\big)^{1/\beta} < \infty,$$

since $\mathbb{E}|M_{11,0}|^\beta < 1$ by convexity and $\mathbb{E}|D_0|^\beta < \infty$. We conclude that $I_+ < \infty$ for $\alpha_1 > 1$.

Lastly, we show that $\overline{c}_1 > 0$. We notice that the solution $\widehat{Y}_1$ of the SRE (5.5) satisfies $\mathbb{P}(\pm\widehat{Y}_1 > x) \sim \hat{c}_1 x^{-\alpha_1}$ with $\hat{c}_1 > 0$ from Lemma A.3. Similarly $\widetilde{Y}_1$ of SRE (5.6) has $\mathbb{P}(\pm\widetilde{Y}_1 > x) \sim \tilde{c}_1' x^{-\alpha_1}$ possibly $\tilde{c}_1' = 0$. Here $\hat{c}_1$ and $\tilde{c}_1'$ are given by $1/(2\alpha_1 k_1)$ times the left and right of (5.12), respectively. The condition (5.12) is equivalent to $\hat{c}_1 \neq \tilde{c}_1'$, and without loss of generality we assume $\hat{c}_1 > \tilde{c}_1'$. Then for arbitrary $\varepsilon > 0$, it follows that

$$\begin{aligned}
\mathbb{P}(Y_1 > x) &= \mathbb{P}(\widehat{Y}_1 + \widetilde{Y}_2 > x) \\
&\geq \mathbb{P}(\widehat{Y}_1 > (1+\varepsilon)x,\ \widetilde{Y}_1 \geq -\varepsilon x) \\
&\geq \mathbb{P}(\widehat{Y}_1(1+\varepsilon)x) - \mathbb{P}(\widetilde{Y}_1 < -\varepsilon x) \\
&\sim \frac{1}{\varepsilon^{\alpha_1}}\Big(\hat{c}_1\Big(\frac{\varepsilon}{1+\varepsilon}\Big)^{\alpha_1} - \tilde{c}_1'\Big)x^{-\alpha_1}.
\end{aligned}$$

Therefore, for sufficiently large ε, $\mathbb{P}(Y_1 > x) \geq c x^{-\alpha_1}$ with some $c > 0$. In a similar manner we see $\mathbb{P}(-Y_1 > x) \geq c' x^{-\alpha_1}$ with some $c' > 0$. This finishes the proof. □

References

[1] Alsmeyer, G. and Mentemeier, S. (2012) Tail behavior of stationary solutions of random difference equations: the case of regular matrices. *Journal of Differential Equations and Applications*, **18**, 1305–1332.

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

[3] Basrak, B. and Segers, J. (2009) Regularly varying multivariate time series. *Stochastic Processes and their Applications*, **119**, 1055–1080.

[4] Bougerol, P. and Picard N. (1992) Strict stationarity of generalized autoregressive processes. *Annals of Probability*, **20**, 1714–1730.

[5] Brandt, A. (1986) The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability*, **18**, 211–220.

[6] Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory of Probability and its Applications*, **10**, 323–331.

[7] Boussama, F., Fuchs, F., and Stelzer, R. (2011) Stationarity and geometric ergodicity of BEKK multivariate GARCH models. *Stochastic Processes and their Applications*, **121**, 2331–2360.

[8] Buraczewski, D., Damek, E., Guivarc'h, Y., Hulanicki, A. and Urban, R. (2009) Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. *Probability Theory and Related Fields*, **145**, 385–420.

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

[10] Cont, R. (2001) Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, **1**, 223-236.

[11] Damek, E., Matsui, M. and Świątkowski, W. (2019) Componentwise different tail solutions for bivariate stochastic recurrence equations with application to GARCH(1,1) processes. *Colloquium Mathematicum*, **155**, 227–254.

[12] Damek, E. and Matsui, M. (2019) Tails of bivariate stochastic recurrence equation. *in preparation*.

[13] Damek, E. and Zienkiewicz, J. (2018) Affine stochastic equation with triangular matrices. *Journal of Differential Equations and Applications*, **24**, 520–542.

[14] Davis, R.A. and Mikosch, T. (2009) Extreme value theory for GARCH processes. In Andersen, T.G., Davis, R.A., Kreiß, J.-P. and Mikosch T. (Eds.), *Handbook of Financial Time Series*, 186-200. Springer, New York, NY.

[15] Engle, R.F and Kroner K.F. (1995) Multivariate simultaneous generalized ARCH. *Econometric Theory*, **11**, 122–150.

[16] Francq, C. and Zakoïan (2010) *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons, Chichester.

[17] Gabaix, X. (2009) Power laws in economics and finance. *The Annual Review of Economics*, **1**, 255–293.

[18] Gabaix, X. and Ibragimov, R. (2011) Rank - 1/2: A simple way to improve the OLS estimation of tail exponents. *Journal of Business & Economic Statistics*, **29**, 24–39.

[19] Gerencsér, L., Michaletzky, G. and Orlovits, Z. (2008) Stability of block-triangular stationary random matrices. *Systems & Control Letters*, **57**, 620–625.

[20] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Annals of Applied Probability*, **1**, 126–166.

[21] Guivarc'h, Y. and Le Page, É. (2016) Spectral gap properties for linear random walks and Pareto's asymptotics for affine stochastic recursions. *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, **52**, 503–574.

[22] Hafner, C. and Preminger, A. (2009) On asymptotic theory for multivariate GARCH models. *Journal of Multivariate Anlysis*, **100**, 2044–2054.

[23] Horn, R.A. and Johnson, C.R. (2013) *Matrix Analysis*, 2nd edition, Cambridge University Press.

[24] Ibragimov, M., Ibragimov, R. and Walden, J. (2015) *Heavy-Tailed Distributions and Robustness in Economics and Finance. Lecture Notes in Statistics* **214**. Springer International Publishing, Switzerland.

[25] Krengel, U. (2011) *Ergodic theorems*. Vol. 6. Walter de Gruyter, Berlin.

[26] Ling, S. and Li, D. (2008) Asymptotic inference for a nonstationary double AR(1) model. *Biometrika*, **95**, 257–263.

[27] Loretan, M. and Phillips, P.C.B. (1994) Testing for covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance*, **1**, 211–248.

[28] Matsui, M. and Mikosch, T. (2016) The extremogram and the cross-extremogram for a bivariate GARCH(1,1) process. *Advances in Applied Probability*, **48A**, 217–233.

[29] Matsui, M. and Świątkowski, W. (2020) Tail indices for AX+B recursion with triangular matrices. *Journal of Theoretical Probability* (online first).

[30] Nelson, D.B. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, **6**, 318–334.

[31] Nicholls, D.F. and Quinn, B.G. (1982) *Random Coefficient Autoregressive Models: An Introduction. Lecture Notes in Statistics* **11**. Springer, New York, NY.

[32] Nielsen, H.B. and Rahbek, A. (2014) Unit root vector autoregression with volatility induced stationarity. *Journal of Empirical Finance*, **29**, 144–167.

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

[34] Pedersen R.S. and Rahbek, A. (2014) Multivariate variance targeting in the BEKK-GARCH model. *The Econometrics Journal*, **17**, 24–55.

[35] Pedersen R.S. and Wintenberger, O. (2018) On the tail behavior of a class of multivariate conditionally heteroskedastic processes. *Extremes*, **21**, 261–284.

[36] Resnick, S.I. (2007) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York, NY.

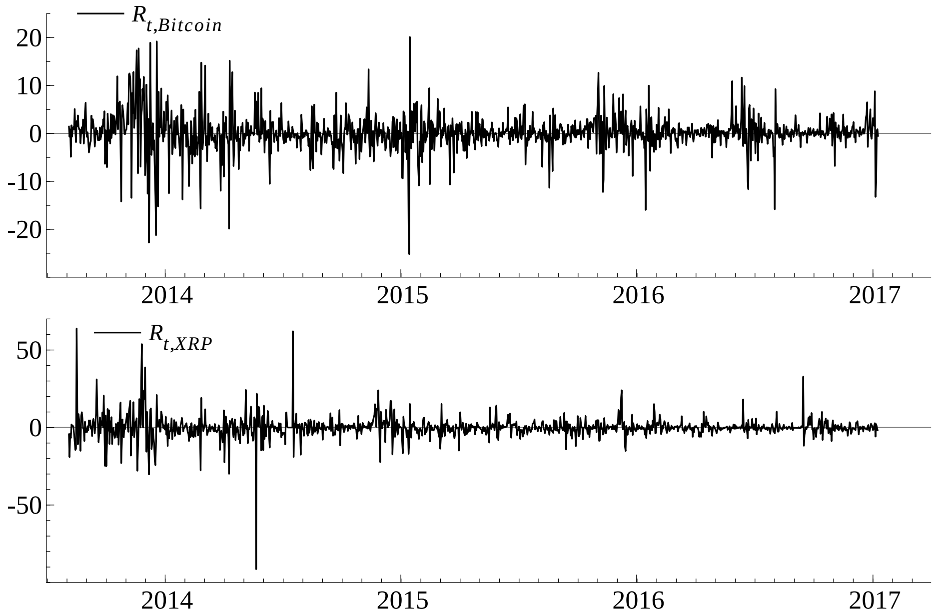
[37] Stărică, C. (1999) Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance*, **6**, 515–553.

[38] Straumann, D. (2005) *Estimation in conditionally heteroscedastic time series models.* Lecture Notes in Statistics **181**, Springer-Verlag, Berlin.

[39] Sun, P. and Zhou C. (2014) Diagnosing the distribution of GARCH innovations. *Journal of Empirical Finance*, **29**, 287–303.

[40] Tsay, R. S. (1987) Conditional heteroscedastic time series models. *Journal of the American Statistical Association*, **82**, 590–604.

Figure 1: Daily returns on Bitcoin/USD and XRP/USD



Note: Daily continuously compounded returns (multiplied by 100) on Bitcoin/US Dollar and XRP/US Dollar exchange rates from August 5, 2013 - January 7, 2017.

Table 1: Estimation results

	(7.1)	(7.2)	(7.3)	(7.4)	(7.1*)	(7.2*)	(7.3*)	(7.4*)
C_1	2.80	2.74	2.81	2.81	5.18	5.06	5.13	5.29
C_2	0.614	0.600	0.879	0.882	0.234	0.265	0.417	0.469
C_3	5.23	5.23	5.21	5.21	2.75	2.74	2.77	2.77
$A_{1,11}$	0.162	0.249	0.769	.	0.0799	0.0798	0.834	.
$A_{1,12}$	0.0772	0.203	0.0174	.	0.442	0.545	0.298	.
$A_{1,22}$	0.870	0.473	0.216	.	0.823	0.650	0.231	.
$A_{2,11}$	0.742	.	.	0.778	0.822	.	.	0.856
$A_{2,22}$	0.00000	.	.	0.210	0.00000	.	.	0.191
$A_{3,11}$	.	0.630	.	.	.	0.821	.	.
$A_{4,22}$	.	0.729	0.828	0.829	.	0.505	0.749	0.754
Lik	-5107.9	-5104.9	-5111.7	-5112.2	-5097.9	-5094.0	-5105.4	-5112.2
AIC	10231.9	10225.9	10237.6	10236.4	10211.9	10204.0	10224.9	10236.4
BIC	10272.9	10267.0	10273.5	10267.2	10252.9	10245.1	10260.8	10267.2

Table 2: Estimated tail indexes

	$\hat{\alpha}$, model (7.3)	$\hat{\alpha}$, model (7.4)	$\hat{\alpha}_{\mathrm{RS},5\%}$	$\hat{\alpha}_{\mathrm{RS},10\%}$
$R_{t,\mathrm{Bitcoin}}$	3.87	3.77	3.68 (2.40,4.97)	2.67 (2.01,3.32)
$R_{t,\mathrm{XRP}}$	3.03	2.99	2.40 (1.56,3.23)	2.24 (1.69,2.80)