



# Asymptotic Independence ex machina - Extreme Value Theory for the Diagonal SRE Model

Sebastian Mentemeier, Olivier Wintenberger

## ► To cite this version:

Sebastian Mentemeier, Olivier Wintenberger. Asymptotic Independence ex machina - Extreme Value Theory for the Diagonal SRE Model. *Journal of Time Series Analysis*, In press, 43 (5), pp.750-780. 10.1111/jtsa.12637 . hal-02591409v3

**HAL Id: hal-02591409**

**<https://hal.science/hal-02591409v3>**

Submitted on 4 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Asymptotic Independence *ex machina* - Extreme Value Theory for the Diagonal SRE Model

Sebastian Mentemeier\*, Olivier Wintenberger†

January 4, 2022

## Abstract

We consider multivariate stationary processes  $(\mathbf{X}_t)$  satisfying a stochastic recurrence equation of the form

$$\mathbf{X}_t = \mathbb{M}_t \mathbf{X}_{t-1} + \mathbf{Q}_t,$$

where  $(\mathbf{Q}_t)$  are iid random vectors and

$$\mathbb{M}_t = \text{Diag}(b_1 + c_1 M_t, \dots, b_d + c_d M_t)$$

are iid diagonal matrices and  $(M_t)$  are iid random variables. We obtain a full characterization of the Vector Scaling Regular Variation properties of  $(\mathbf{X}_t)$ , proving that some coordinates  $X_{t,i}$  and  $X_{t,j}$  are asymptotically independent even though all coordinates rely on the same random input  $(M_t)$ . We prove the asynchrony of extreme clusters among marginals with different tail indices. Our results are applied to some multivariate autoregressive conditional heteroskedastic (BEKK-ARCH and CCC-GARCH) processes and to log-returns. Angular measure inference shows evidences of asymptotic independence among marginals of diagonal SRE with different tail indices.

*AMS 2010 subject classifications:* 60G70, 60G10

*Keywords:* Stochastic recurrence equations, multivariate ARCH, multivariate regular variation, non-standard regular variation

## 1 Introduction

We consider multivariate stationary processes  $(\mathbf{X}_t)$ , satisfying a diagonal stochastic recurrence equation (SRE) of the form

$$\mathbf{X}_t = \mathbb{M}_t \mathbf{X}_{t-1} + \mathbf{Q}_t, \quad t \in \mathbb{Z}, \quad (1.1)$$

---

\*Universität Hildesheim, Germany. [mentemeier@uni-hildesheim.de](mailto:mentemeier@uni-hildesheim.de)

†Sorbonne Université Paris, France. [olivier.wintenberger@upmc.fr](mailto:olivier.wintenberger@upmc.fr)

where  $(\mathbb{M}_t)$  is an iid sequence of matrices such that for non-negative coefficients  $b_i, c_i, 1 \leq i \leq d$ ,

$$\mathbb{M}_t = \text{Diag}(b_1 + c_1 M_t, \dots, b_d + c_d M_t), \quad t \in \mathbb{Z}, \quad (1.2)$$

and  $(\mathbf{Q}_t)$  is an iid sequence of  $\mathbb{R}^d$  random vectors with marginals  $Q_{t,i}, 1 \leq i \leq d$ , independent of the iid real random variables  $(M_t)$ . Stationary solutions of SRE have attracted a lot of research in the past few years, see [5] and references therein. However, in the present setting of diagonal matrices, only marginal tail behavior has been investigated so far using the result of the seminal paper of [15]. Under Assumptions (A1) – (A6) that are introduced in Section 2.3, applying the Kesten-Goldie-Theorem [15, 19] we get that

$$\mathbb{P}(X_{0,i} > x) \sim a_i x^{-\alpha_i}, \quad x \rightarrow \infty, \quad (1.3)$$

where  $a_i$  is a positive constant and  $\alpha_i > 0$  is the unique solution of the equation  $\mathbb{E}[|b_i + c_i M_0|^{\alpha_i}] = 1$ . Here and below,  $f(x) \sim g(x)$  means that  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$ .

Heavy-tails as in (1.3) favor the appearance of extreme values. On the contrary with the iid case, these values tend to appear consecutively in time due to the dependency in the diagonal SRE model (1.1). As discussed in [8], this extremal clustering is an important phenomenon to take into account in risk analysis. It is usually described via the notion of *regular variation for stationary time series* defined in [1].

**Definition 1.1.** The stationary time series  $(\mathbf{X}_t)$  is regularly varying if and only if  $\|\mathbf{X}_0\|$  is regularly varying and for all  $t \geq 0$  there exist weak limits

$$\lim_{x \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_0\|^{-1}(\mathbf{X}_0, \dots, \mathbf{X}_t) \in \cdot | \|\mathbf{X}_0\| > x) = \mathbb{P}((\boldsymbol{\Theta}_0, \dots, \boldsymbol{\Theta}_t) \in \cdot).$$

By stationarity and using Kolmogorov consistency theorem one can extend the trajectories  $(\boldsymbol{\Theta}_0, \dots, \boldsymbol{\Theta}_t)$  into a process  $(\boldsymbol{\Theta}_t)$  called the *spectral tail process*. Consecutive big values of the spectral tail process around  $\boldsymbol{\Theta}_0$  characterize the extreme clustering due to an extremal event  $\{\|\mathbf{X}_0\| > x\}$  for  $x$  sufficiently large. Note that serial extremal dependence of the marginal sequences  $(X_{t,i})_{t \in \mathbb{Z}}$  for any  $1 \leq i \leq d$  is well known since the pioneer work of [12].

For diagonal SRE (1.1) we will show that  $\alpha_i \neq \alpha_j$  in many situations, so that the marginals of  $\mathbf{X}_0$  are *not tail equivalent*. If  $\alpha_i < \max_j \alpha_j$ , then

$$\mathbb{P}(|X_{0,i}| > x) = o(\mathbb{P}(\|\mathbf{X}_0\| > x)), \quad x \rightarrow \infty.$$

In this case, the notion of regular variation of [1] is not suitable. By an application of the results in [18], the corresponding marginal of the spectral tail process is degenerated, *i.e.*,  $\boldsymbol{\Theta}_{t,i} = 0$  a.s. for all  $t \in \mathbb{Z}$ . Hence,

information about extreme clustering in this coordinate is lost. The notion of *non-standard regular variation* was introduced to circumvent this issue in full generality (see [26] and references therein). For Pareto equivalent marginal tails satisfying (1.3), such as coordinates of stationary solutions to the diagonal SRE (1.1), *Vector Scaling Regular Variation (VSRV)* in the sense of [24] is a simple alternative of the non-standard regular variation of [26], see preliminary **Section 2** for a formal definition. The extremal behaviour of stationary VSRV time series is described by the spectral tail VSRV process  $(\tilde{\Theta}_t)$ , extending the notion of spectral tail process to SRE with different marginal tail indices.

As an illustration of our approach, consider the bivariate case  $(X_{0,1}, X_{0,2})$  such that  $c_2/c_1 > b_2/b_1 \geq 1$  (with the convention  $0/0 = 1$ ) and  $c_1 > 0$  and both  $M = M_0$  as well as the components  $Q_1 = Q_{0,1}$  and  $Q_2 = Q_{0,2}$  are positive. Then  $\alpha_1 > \alpha_2$  and we prove in **Section 3** that  $X_{0,1}$  and  $X_{0,2}$  are *asymptotically independent* in the sense that

$$\lim_{x \rightarrow \infty} \mathbb{P}(X_{0,1} > x^{\frac{1}{\alpha_1}} \mid X_{0,2} > x^{\frac{1}{\alpha_2}}) = \lim_{x \rightarrow \infty} \mathbb{P}(X_{0,2} > x^{\frac{1}{\alpha_2}} \mid X_{0,1} > x^{\frac{1}{\alpha_1}}) = 0. \quad (1.4)$$

This result remains true also when  $Q_1 = Q_2$ . Thus, even though  $X_{0,1}$  and  $X_{0,2}$  are perfectly dependent in the sense that all their randomness comes from the same random variables, extremes never occur simultaneously. This asynchrony is due to different typical time scales until the appearance of an extremal event in coordinates with different tail indices. Note that the different vector scaling  $(x^{1/\alpha_1}, x^{1/\alpha_2})$  in (1.4) depends on the power-law index in (1.3). This asymptotical independence result (1.4) is responsible for the VSRV properties of the stationary time series  $(X_{t,1}, X_{t,2})$ . We refer to **Section 3** for the general result in  $\mathbb{R}^d$ . We prove asymptotic independence between blocks with different tail indices, and asymptotic dependence within blocks. Clusters of extremes of coordinates with different tail indices in the diagonal SRE are necessarily asynchronous. We describe this phenomenon as *asynchronous extreme clustering*.

The diagonal SRE (1.1) is a very simple model that may coincide with some classical multivariate GARCH ones. For the specific case  $(M_t)$  are iid  $\mathcal{N}(0, 1)$ ,  $b_i = 0$  for all  $1 \leq i \leq d$  and  $(Q_t)$  are iid  $\mathcal{N}(0, \Sigma)$  the diagonal SRE coincides with the diagonal BEKK-ARCH(1) model as in [24]. For  $(Q_t)$  degenerated to a constant the diagonal SRE model coincides with the volatility process of some CCC-GARCH model. Such diagonal SRE models are very interesting as they generate potentially different marginal tail indices  $\alpha_i > 0$ . This freedom is not offered by the general CCC-GARCH model whom marginals have the same tail index, as discussed in [5] and [27]. This feature is important for modeling: Heavy tailed data, such as in finance, may exhibit different tail indices indicating different magnitude in the responses during financial crisis.

We illustrate in **Section 4** via simulations of the CCC-GARCH model the consequence of different marginal tail indices on extreme clustering. We show the asynchrony of the clusters of extreme values generated by multivariate GARCH models with different tail indices and asymptotic independence of bivariate processes. Even if the conditional covariance dependence is total, i.e. the multiplicative random term is common to all GARCH models, the clustering of extremes does not propagate between blocks of different tail indices. We believe that this asynchronous extreme clustering phenomenon is shared by general models that do not fit into the setting of our work. We apply our approach on real data. We fit the more general and realistic DCC-GARCH model of [9] on log-returns of IBM and GOOGLE. Despite strong *correlation clustering*<sup>1</sup> and close tail indices among the marginals of the model, the fitted DCC-GARCH model exhibits some evidences of asymptotic independence on simulations despite the model does not satisfy our assumptions. These evidences are not shared by the real data. Asynchrony of extreme clusters constitute, in our opinion, a strong and intrinsic limitation of multivariate GARCH models with different marginal tail indices. The model captures extreme clustering with different magnitude only when they are due to different financial crisis. The existence of a multivariate GARCH model exhibiting extremal clustering among marginals with different tail indices is left as an open question.

## Structure of the paper

**Section 2** contains preliminaries such as notation, VSRV notion and assumptions required to formulate the results contained in **Section 3**. Illustrations of our results are provided in **Section 4** both on simulated and real data. The proofs are collected in **Section 5** using a crucial lemma on the different typical time scales until the first appearance of an extreme value in different coordinates. The proof of this lemma is contained in the **Appendix** together with a result on the non-standard regular variation properties of VSRV random vectors with positive coordinates.

## 2 Preliminaries

### 2.1 Notation

The max-norm on  $\mathbb{R}^d$  is denoted  $\|\cdot\|$  and the euclidean norm  $\|\cdot\|_2$ . For vectors, we use bold notation  $\mathbf{x} = (x_1, \dots, x_d)$ . Operations between vectors or scalar and vector are interpreted coordinate wise, *e.g.*,  $x^{-1/\alpha} = (x^{-1/\alpha_1}, \dots, x^{-1/\alpha_d})$  for positive  $x$  and  $\mathbf{ab} = (a_i b_i)_{1 \leq i \leq d}$ . A notation that

---

<sup>1</sup>Correlation clustering is a stylized fact captured by the DCC-GARCH model, see the V-Lab documentation <https://vlab.stern.nyu.edu/docs/correlation/GARCH-DCC>.

will be used frequently is vector scaling of a sequence of  $\mathbb{R}^d$ -valued random variables, e.g.

$$\begin{aligned} x^{-1/\alpha}(\mathbf{X}_0, \dots, \mathbf{X}_t) &= (x^{-1/\alpha}\mathbf{X}_0, \dots, x^{-1/\alpha}\mathbf{X}_t) \\ &= ((x^{-1/\alpha_i}X_{0,i})_{1 \leq i \leq d}, \dots, (x^{-1/\alpha_i}X_{t,i})_{1 \leq i \leq d}). \end{aligned}$$

For some potentially distinct  $\alpha_1, \dots, \alpha_d$  we define the following notion of vector scaling distance:

$$\|\mathbf{x}\|_{\alpha} = \max_{1 \leq i \leq d} |x_i|^{\alpha_i} = \|\mathbf{x}^{\alpha}\|, \quad \mathbf{x} = (x_i)_{1 \leq i \leq d} \in \mathbb{R}^d.$$

Here  $\mathbf{x}^{\alpha}$  denotes the vector  $(\text{sign}(x_i)|x_i|^{\alpha_i})_{1 \leq i \leq d}$  in  $\mathbb{R}^d$ . We want to stress that  $\|\mathbf{x}\|_{\alpha}$  is neither homogeneous nor does it satisfy the triangle inequality for general values of  $\alpha_1, \dots, \alpha_d$ . Thus, it is not a (pseudo-)norm but it will provide a meaningful scaling function. Note that  $\mathbf{x} \mapsto \|\mathbf{x}\|_{\alpha}$  is a continuous function and is  $1/\alpha$ -homogeneous in the following sense:

$$\|\lambda^{1/\alpha}\mathbf{X}_0\|_{\alpha} = \max_{1 \leq i \leq d} |\lambda^{1/\alpha_i}X_{0,i}|^{\alpha_i} = \lambda\|\mathbf{X}_0\|_{\alpha}$$

The components of the vector

$$\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha}\mathbf{X}_t = (\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha_i}X_{t,i})_{1 \leq i \leq d}$$

have  $\|\cdot\|_{\alpha}$  and max-norm equal to one when  $t = 0$  thus it belongs to  $\mathcal{S}_{\infty}^{d-1} = \{\mathbf{x} \in \mathbb{R}^d; \|\mathbf{x}\|_{\alpha} = 1\}$  the max-norm-unit sphere.

## 2.2 Vector Scaling Regular Variation

In order to treat the temporal dependence of the stationary solution  $(\mathbf{X}_t)$ , we will use the notion of Vector Scaling Regular Variation (VSRV) introduced in [24] as follows:

**Definition 2.1** (VSRV). A stationary time series  $(\mathbf{X}_t)$  is VSRV of order  $\alpha = (\alpha_1, \dots, \alpha_d)$  if for all  $1 \leq i \leq d$ ,

$$\mathbb{P}(|X_{0,i}| > x) \sim a_i x^{\alpha_i} \text{ with } a_i > 0, \quad (2.1)$$

$\|\mathbf{X}_0\|_{\alpha}$  is regularly varying and there exists weak limits

$$\lim_{x \rightarrow \infty} \mathbb{P}\left(\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha}(\mathbf{X}_0, \dots, \mathbf{X}_t) \in \cdot \mid \|\mathbf{X}_0\|_{\alpha} > x\right) = \mathbb{P}((\tilde{\Theta}_0, \dots, \tilde{\Theta}_t) \in \cdot), \quad (2.2)$$

for any  $t \geq 0$ .

One can extend the trajectories  $(\tilde{\Theta}_0, \dots, \tilde{\Theta}_t)$  into a process  $(\tilde{\Theta}_t)$  called the *spectral tail VSRV process*. Note that a VSRV time series  $(\mathbf{X}_t)$  with indices  $\alpha_1, \dots, \alpha_d$  is such that  $(\mathbf{X}_t^{\alpha})$  is regularly varying as in Definition 1.1 with tail index 1. When  $\alpha_1 = \dots = \alpha_d$  then  $(\mathbf{X}_t)$  is regularly varying and  $(\tilde{\Theta}_t)$  coincides with the spectral tail process  $(\Theta_t)$ . It is one advantage of considering VSRV as it extends the regular variation of time series in [1] to SRE solutions with different marginal tail indices.

### 2.3 The univariate marginal SRE and the assumptions

Due to the diagonal multiplicative term in (1.1), the marginals of  $\mathbf{X}_t = (X_{t,1}, \dots, X_{t,d})^\top$  are satisfying the univariate marginal SREs

$$X_{t,i} = (b_i + c_i M_t) X_{t-1,i} + Q_{t,i}, \quad t \in \mathbb{Z}, \quad \text{for } 1 \leq i \leq d.$$

We work under the following set of assumptions that implies the ones of [15] on the marginal SREs. Note that this includes the assumption that the compoments of  $\mathbf{Q}$  are (relatively) light tailed, in order to avoid that the tails of the stationary solution are dominated by  $\mathbf{Q}$ ; see assumption A3 below. We impose this assumption because we are interested in the effect of the multiplicative part  $\mathbb{M}$  on the tail properties of the solution. See [17] for a discussion of the case when  $\mathbf{Q}$  dominates. Denoting by  $(M, \mathbf{Q})$  a generic copy of  $(M_t, \mathbf{Q}_t)$ , we assume that for all  $1 \leq i \leq d$ ,

$$\mathbb{E}[\log |b_i + c_i M|] < 0. \quad (\text{A1})$$

This implies that the top Lyapunov coefficient of the product  $\mathbb{M}_1 \cdots \mathbb{M}_k$  is negative, hence the Markov chain  $(\mathbf{X}_t)$  has a unique stationary distribution (see [19]). It is given by the law of the random vector

$$\mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} := \sum_{k=1}^{\infty} \mathbb{M}_1 \cdots \mathbb{M}_{k-1} \mathbf{Q}_k. \quad (2.3)$$

We further assume that there exist positive constants  $\alpha_1, \dots, \alpha_d$  such that for  $1 \leq i \leq d$

$$\mathbb{E}[|b_i + c_i M|^{\alpha_i}] = 1. \quad (\text{A2})$$

Given these  $\alpha_1, \dots, \alpha_d$ , we assume for  $1 \leq i \leq d$

$$\mathbb{E}[|M|^{\alpha_i + \epsilon}] < \infty, \quad \mathbb{E}[\|\mathbf{Q}\|^{\alpha_i + \epsilon}] < \infty \quad \text{for some } \epsilon > 0. \quad (\text{A3})$$

Of course, it suffices to check this condition for the maximal  $\alpha_i$ . We also need the technical assumption that

$$\text{the distributions of } \log |b_i + c_i M| \text{ are non-arithmetic for all } 1 \leq i \leq d. \quad (\text{A4})$$

Finally, to avoid degeneracy, we require for  $1 \leq i \leq d$  that

$$\mathbb{P}((b_i + c_i M)x + Q_i = x) < 1 \quad \text{for all } x \in \mathbb{R}, \text{ and } \mathbb{P}(Q_i > 0) > 0. \quad (\text{A5})$$

Given (A1)–(A5), an application of the Kesten-Goldie-Theorem of [15, 19] yields the existence of a Pareto tail equivalent stationary distribution, i.e. the equivalence in (1.3) is met, namely  $\mathbb{P}(X_{0,i} > x) \sim a_i x^{-\alpha_i}$  as  $x \rightarrow \infty$ .

The positivity of  $a_i$  can be deduced by non trivial classical arguments, as follows.

If  $P(b_i + c_i M < 0) > 0$ , then positivity of  $a_i$  is proved in [15, Theorem 4.1]. If  $b_i + c_i M > 0$  a.s., then additional arguments are needed: Assumptions (A1) and (A2) together imply that there are  $m, m'$  in the support of  $M$ , such that  $b_i + c_i m < 1$ ,  $b_i + c_i m' > 1$ . By (A5), there is  $q_i > 0$  in the support of  $Q_i$ . Since  $Q_i$  and  $M$  are assumed to be independent, we have that  $(b_i + c_i m, q_i)$  and  $(b_i + c_i m', q_i)$  are in the support of  $(b_i + c_i M, Q_i)$ . Then [5, Proposition 2.5.4] yields that the support of  $X_i$  is unbounded at  $+\infty$ , which together with [5, Theorem 2.4.6] implies that  $a_i$  is positive.

We note in addition that the stationarity condition (A1) can be deduced from (A2) as soon as  $M$  is not constant a.s. (which is implied by (A4)), see the comments after Theorem 2.4.4 in [5].

## 2.4 Assumptions specific to the multivariate SRE

We finish with assumptions concerning the interplay of different marginals; these are specific to our multivariate diagonal setting and not required for the study of marginal SREs.

For all pairs  $1 \leq i, j \leq d$  such that  $\alpha_i > \alpha_j$ , we will require that

$$\lim_{u \rightarrow \infty} \log(u) \mathbb{P}\left(\frac{|Q_j|}{|Q_i|} > u^\epsilon\right) = 0 \quad \text{for all } \epsilon > 0. \quad (\text{A6})$$

Considering (2.3), we see that the ratio  $X_i/X_j$  may be affected not only by the ratios  $(b_i + c_i M_k)/(b_j + c_j M_k)$  of the multiplicative part, but also by the ratios  $Q_{i,k}/Q_{j,k}$ . Assumption (A6) ensures that the effect of the latter ratio does not dominate. It is easy to check in examples (see [7] for a discussion); it holds e.g. if  $\mathbf{Q} \sim \mathcal{N}(0, \Sigma)$ , or whenever the ratio  $Q_j/Q_i$  is bounded or has finite expectation.

We suppose that coordinates are chosen in such a way that  $\alpha_i$  decreases with  $i$ . We partition  $\{1, \dots, d\} = I_1 \cup I_2 \cup \dots \cup I_r$  such that  $b_i = b_j$  and  $c_i = c_j$  (hence  $\alpha_i = \alpha_j$ ) if and only if  $i, j \in I_\ell$  for some  $1 \leq \ell \leq r$ .

To be able to compare the effects of the multiplicative part  $b_i + c_i M$  in different marginals, we assume that one of the two following cases holds for all  $i, j$  with  $i \in I_\ell, j \in I_{\ell'}, \ell < \ell'$  (hence  $i < j$ ):

$$\begin{aligned} b_i = b_j = 0, \quad c_j > c_i > 0, \quad M_t \text{ is } \mathbb{R}\text{-valued} \quad (\text{Case I}) \\ b_j \geq b_i > 0, \quad c_j > c_i > 0, \quad \frac{c_j}{c_i} \geq \frac{b_j}{b_i}, \quad M_t > 0 \text{ a.s.} \quad (\text{Case II}) \end{aligned}$$

That is, we allow the diagonal entries to be linear or affine transformations of  $M_t$ . These cases cover a variety of multivariate diagonal GARCH models, see the examples in Section 4. In the affine case Case II, we require positivity to ensure that one entry dominates the other in modulus almost surely; i.e., we always know, which multiplicative component is the largest one. Recalling



the definition of  $\alpha_i$  by the property

$$\mathbb{E}\left[|b_i + c_i M|^{\alpha_i}\right] = 1,$$

we note that in both cases  $|b_j + c_j M| > |b_i + c_i M|$  a.s., which implies  $\alpha_i > \alpha_j$ .

We further denote by

$$\mathbb{R}^{|I_\ell|} = \{x \in \mathbb{R}^d; x_i = 0 \text{ for } i \notin I_\ell\}$$

the (embedded) subspace corresponding with coordinates indexed by  $I_\ell$  and by

$$\mathbb{S}_\infty^{|I_\ell|-1} = \{x \in \mathbb{R}^d; \max_{i \in I_\ell} |x_i| = 1 \text{ and } x_i = 0 \text{ for } i \notin I_\ell\}$$

its max-norm-unit sphere. Note that if  $I_\ell = \{i\}$  is a singleton, then  $\mathbb{S}_\infty^{|I_\ell|-1} = \{e_i, -e_i\}$  where  $e_i$  denotes the corresponding standard basis vectors.

### 3 Main Result

In this section we provide the vector scaling regular variation properties of the diagonal SRE in full generality.

**Theorem 3.1.** *Let  $(\mathbf{X}_t)$  a stationary process satisfying the diagonal SRE (1.1) and assume that (A1)–(A6) hold. Assume that (Case I) or (Case II) holds, then  $(\mathbf{X}_t)$  is a VSRV process satisfying*

$$\tilde{\Theta}_t = \mathbb{M}_t \tilde{\Theta}_{t-1}, \quad t \geq 1. \quad (3.1)$$

and there exist probability measures  $\xi_\ell$  with support contained in  $\mathbb{S}_\infty^{|I_\ell|-1}$ ,  $1 \leq \ell \leq r$ , such that

$$\mathbb{P}(\tilde{\Theta}_0 \in \cdot) = \lim_{x \rightarrow \infty} \mathbb{P}\left(\|\mathbf{X}_0\|_\alpha^{-1/\alpha} \mathbf{X}_0 \in \cdot \mid \|\mathbf{X}_0\|_\alpha > x\right) = \frac{1}{c} \sum_{1 \leq \ell \leq r} c_\ell \xi_\ell(\cdot),$$

for positive constants  $c, c_1, \dots, c_r$  with  $c_1 + \dots + c_r = c$ . In particular,

$$\text{Supp}(\tilde{\Theta}_0) \subset \cup_{1 \leq \ell \leq r} \mathbb{S}_\infty^{|I_\ell|-1}. \quad (3.2)$$

*Proof.* The proof of the main and auxiliary results is the content of Section 5, the actual proof of Theorem 3.1 is in Subsection 5.4.  $\square$

With probability 1, the components of  $\tilde{\Theta}_0$  are equal to zero except for the entries within exactly one block  $I_\ell$ . Thus, we have asymptotic independence between blocks  $I_\ell$ : If  $i \in I_\ell, j \in I_{\ell'}$  with  $\ell \neq \ell'$ , then

$$\begin{aligned} & \lim_{x \rightarrow \infty} \mathbb{P}\left(\|\mathbf{X}_0\|_\alpha^{-1/\alpha(\ell)} X_{0,i} > 0, \|\mathbf{X}_0\|_\alpha^{-1/\alpha(\ell')} X_{0,j} > 0 \mid \|\mathbf{X}_0\|_\alpha > x\right) \\ &= \mathbb{P}(\tilde{\Theta}_{0,i} > 0, \tilde{\Theta}_{0,j} > 0) = 0. \end{aligned}$$

A natural question to ask is in which cases equality holds in (3.2), or, more precisely, in which cases is the support of  $\xi_\ell$  equal to  $\mathbb{S}_\infty^{|I_\ell|-1}$ ? The following result gives sufficient conditions for equality. For a vector  $\mathbf{x}$  in  $\mathbb{R}^d$  we write  $\mathbf{x}_\ell = (x_i)_{i \in I_\ell}$  and denote by  $\text{span}(M)$  the linear space generated by the set of vectors  $M$ .

**Lemma 3.2.** *Fix  $1 \leq \ell \leq r$  and let  $m = |I_\ell|$ . Under the assumptions of Theorem 3.1,*

$$\text{supp}(\tilde{\Theta}_{0,\ell}) \subset \text{span}(\text{supp}(\mathbf{Q}_\ell)) \cap S_\infty^{m-1}. \quad (3.3)$$

*In addition, the following implications hold:*

(a) *In (Case I), if  $\text{supp}(M)$  is dense in  $\mathbb{R}$ , then*

$$\text{supp}(\tilde{\Theta}_{0,\ell}) = \text{span}(\text{supp}(\mathbf{Q}_\ell)) \cap S_\infty^{m-1}.$$

(b) *In (Case II), if  $\text{supp}(M)$  is dense in  $\mathbb{R}_+$ , then*

$$\text{supp}(\tilde{\Theta}_{0,\ell}) = \{a_1 \mathbf{q}_1 + \dots + a_n \mathbf{q}_n : n \in \mathbb{N}, a_i > 0, \mathbf{q}_i \in \text{supp}(\mathbf{Q}_\ell)\} \cap S_\infty^{m-1},$$

*i.e. it equals the convex cone generated by  $\text{supp}(\mathbf{Q}_\ell)$  intersected by the unit sphere.*

(c) *If  $\text{supp}(\mathbf{Q}_\ell)$  is dense in  $\mathbb{R}^m$ , then  $\text{supp}(\tilde{\Theta}_{0,\ell}) = S_\infty^{m-1}$ .*

*Proof.* This result is proved in Lemma 5.5 below.  $\square$

We can summarize the previous result as follows. If  $M$  has full support and  $\mathbf{Q}_\ell$  is not confined to a linear subspace, then  $\tilde{\Theta}_{0,\ell}$  charges the whole unit sphere  $S_\infty^{|I_\ell|-1}$ , see Remark 5.6 for a precise statement.

## 4 Application to multivariate GARCH models and real data

In this section, we show how our result applies to different multivariate GARCH models.

### 4.1 Diagonal BEKK-ARCH(1) model

We consider  $(\mathbf{X}_t)$  the solution of the diagonal BEKK-ARCH(1) model defined as in [24] by the system

$$\begin{cases} \mathbf{X}_t &= \mathbf{H}_t^{1/2} \mathbf{Z}_t, & t \in \mathbb{Z}, \\ \mathbf{H}_t &= \Sigma + \text{Diag}(c_1, \dots, c_d) \mathbf{X}_{t-1} \mathbf{X}_{t-1}^\top \text{Diag}(c_1, \dots, c_d), \end{cases}$$

where  $(\mathbf{Z}_t)$  is an iid sequence of gaussian random vectors  $\mathcal{N}_d(0, I)$  and  $\mathbf{\Sigma}$  is a variance matrix. Due to the assumption that  $\mathbf{Z}_t$  is gaussian, we can identify the gaussian transition kernel of the diagonal BEKK-ARCH(1) model. It coincides with the one of the diagonal SRE model (1.1) with

$$(\mathbf{Q}_t) \text{ are iid } \mathcal{N}(0, \mathbf{\Sigma}),$$

$$\mathbb{M}_t = \text{Diag}(c_1 M_t, \dots, c_d M_t)$$

and  $(M_t)$  iid  $\mathcal{N}(0, 1)$ , see [24]. Thus a version of  $(\mathbf{X}_t)$  satisfies the recursion (1.1). Under the top-Lyapunov condition

$$c_i^2 < 2e^\gamma, \quad 1 \leq i \leq d, \quad (4.1)$$

where  $\gamma \approx 0.5772$  is the Euler constant, it exists a stationary solution  $(\mathbf{X}_t)$ ; see e.g. [23]. Its multivariate extremal behaviour is given by the following corollary which follows from an application of Theorem 3.1 in (Case I):

**Corollary 4.1.** *If the stationarity assumption (4.1) is satisfied, then the stationary solution  $(\mathbf{X}_t)$  of the diagonal BEKK-ARCH(1) model is a VSRV process satisfying*

$$\text{Supp}(\tilde{\mathbf{\Theta}}_0) = \cup_{1 \leq \ell \leq r} \mathbb{S}_\infty^{|I_\ell|-1}$$

and

$$\tilde{\mathbf{\Theta}}_t = M_t \text{Diag}(c_1, \dots, c_d) \tilde{\mathbf{\Theta}}_{t-1}, \quad t \geq 1. \quad (4.2)$$

*Proof.* We have to check the assumptions of Theorem 3.1. This is readily done for (A1)-(A5), see [24] for details. Considering (A6), let  $\sigma_i^2 = \text{Var}(Q_i)$  and  $\rho_{ij}$  be the correlation coefficient of  $Q_i$  and  $Q_j$ ;  $\mathbb{E}Q_i = \mathbb{E}Q_j = 0$ . Then the ratio  $Q_i/Q_j$  has a Cauchy distribution with location parameter  $a = \rho_{ij} \frac{\sigma_i}{\sigma_j}$  and scale parameter  $b = \frac{\sigma_i}{\sigma_j} \sqrt{1 - \rho_{ij}^2}$ ; see e.g. [7, (3.3)]. The Cauchy distributions are 1-stable, hence

$$\mathbb{P}\left(\frac{|Q_i|}{|Q_j|} > u\right) = O(u)$$

and (A6) follows if  $I, J$  are singletons. To compare  $Q_I^* = \max_{i \in I} |Q_i|$  with  $Q_J^* = \max_{j \in J} |Q_j|$  we use the simple bound (fix any  $j \in J$ )

$$\left\{ \frac{Q_I^*}{Q_J^*} > u \right\} \subset \bigcup_{i \in I} \left\{ \frac{|Q_i|}{|Q_j|} > u \right\}$$

to conclude that the probability of this event still decays as  $O(u)$ . Thus (A6) also holds in this case.

It remains to show that  $\text{supp}(\tilde{\mathbf{\Theta}}_0)$  is equal to  $\cup_{1 \leq \ell \leq r} \mathbb{S}_\infty^{|I_\ell|-1}$ . Therefore, we can focus on a particular block  $I$  and show that the spectral measure of the restriction  $(X_{0,i})_{i \in I}$  has full support  $\mathbb{S}_\infty^{|I|-1}$ .

If  $I$  is a singleton, then this means nothing but that left and right tails are regularly varying with the same index; which already follows from the Goldie-Kesten theorem applied to  $M$  with  $\mathbb{P}(M < 0) > 0$ . If  $|I| > 1$  then the result follows from Lemma 3.2 (a), since  $M$  and  $(Q_i)_{i \in I}$  are independent Gaussians, and  $\text{span}(\text{supp}((Q_i)_{i \in I})) = \mathbb{R}^{|I|}$  since  $C$ , the variance of  $\mathbf{Q}$ , has full rank.  $\square$

The multivariate regular variation properties of the diagonal BEKK-ARCH(1) process is quite simple as the support is preserved by the multiplicative form of the spectral tail VSRV process: The spectral tail VSRV process is a mixture of multiplicative random walks with distinct supports. Each support corresponds to the span of the diagonal coefficients of the multiplicative matrix that are equal. From a risk analysis point of view, it means that the extremal risks are dependent only in the directions of equal diagonal coefficients. Thus the model is relevant only if the each group of dependent extremal risks are due to different financial crisis.

Under more restrictive assumptions, it is also possible to provide second-order results, see the arXiv version of the present paper [22]. See also [21] for a recent analysis of the general BEKK-ARCH model.

## 4.2 CCC-GARCH(1,1) model

The Constant Conditional Correlation CCC-GARCH(1,1) model has been introduced by [2] such as the stationary solution of the system

$$\begin{cases} \mathbf{R}_t &= \boldsymbol{\Sigma}_t N_t, & t \in \mathbb{Z}, \\ \boldsymbol{\Sigma}_t &= \text{Diag}(\sigma_{t,1}, \dots, \sigma_{t,d}), \\ \sigma_{t,i}^2 &= a_i + b_i \sigma_{t-1,i}^2 + c_i R_{t-1,i}^2, \end{cases} \quad (4.3)$$

where  $N_t$  iid is distributed as  $\mathcal{N}_d(0, \mathbf{C})$  for  $\mathbf{C}$  a correlation matrix and the coefficients  $a_i$ s,  $b_i$ s and  $c_i$ s are positive. The general CCC-GARCH(1,1) model of [27] is defined by the same system with extra cross terms  $\sum_{j \neq i} b_{i,j} \sigma_{t-1,j}^2$  and  $\sum_{j \neq i} c_{i,j} R_{t-1,j}^2$  in the second equation. By a direct application of Theorem of [19], the volatility  $\mathbf{X}_t = (\sigma_{t,1}^2, \dots, \sigma_{t,d}^2)^\top \in \mathbb{R}_+^d$  of a general CCC-GARCH(1,1) is regularly varying when  $b_{i,j} + c_{i,j} > 0$  for all  $i, j$ , see [27] for more details.

Back to the original CCC-GARCH model (4.3) we consider the degenerate case  $N_t = (1, \dots, 1)^\top Z_t$  with  $Z_t \sim \mathcal{N}(0, 1)$ . Then the original CCC-GARCH(1,1) volatility  $\mathbf{X}_t = (\sigma_{t,i}^2)_{1 \leq i \leq d}$  satisfies the diagonal SRE (1.1) with  $\mathbf{Q}_t \equiv \mathbf{q} = (a_1, \dots, a_d)^\top$  and

$$\mathbb{M}_t = \text{Diag}(b_1 + c_1 Z_t^2, \dots, b_d + c_d Z_t^2).$$

The existence of a stationary solution is then ensured under the top-Lyapunov conditions

$$\mathbb{E}[\log(b_i + c_i Z_0^2)] < 0, \quad 1 \leq i \leq d. \quad (4.4)$$

which is equivalent to (A1) in this case.

We get from another application of Theorem 3.1 in (Case II):

**Corollary 4.2.** *If the stationarity assumption (4.4) is satisfied and  $c_j/c_i \geq b_j/b_i$  whenever  $\alpha_i > \alpha_j$ , then the volatility  $(\mathbf{X}_t)$  of the stationary solution of the CCC-GARCH model (4.3) with  $N_t = (1, \dots, 1)^\top Z_t$  is a VSRV process satisfying*

$$\text{Supp}(\tilde{\Theta}_0) = \cup_{1 \leq \ell \leq r} \{\bar{\mathbf{q}}_\ell\}$$

where  $\bar{\mathbf{q}}_\ell$  is given by

$$(\bar{\mathbf{q}}_\ell)_i = \begin{cases} 0 & i \notin I_\ell, \\ \mathbf{q}_i / \mathbf{q}_\ell^* & i \in I_\ell \end{cases}$$

with  $\mathbf{q}_\ell^* := \max_{j \in I_\ell} \mathbf{q}_j$ . Moreover

$$\tilde{\Theta}_t = \text{Diag}(b_1 + c_1 Z_t^2, \dots, b_d + c_d Z_t^2) \tilde{\Theta}_{t-1}, \quad t \geq 1.$$

*Proof.* To apply Theorem 3.1 in (Case II), we have to check conditions (A1)–(A6) for  $1 \leq i \leq d$ . The stationarity assumption (4.4) is (A1) and implies moreover, together with the fact that  $M = Z_0^2$  has a  $\chi^2$ -distribution, that all moments of  $b_i + c_i M$  exist and  $\lim_{s \rightarrow \infty} \mathbb{E}[|b_i + c_i M|^s] = \infty$ , while  $\mathbb{E}[|b_i + c_i M|^\epsilon] < 1$  for sufficiently small  $\epsilon > 0$ . Hence, for any  $1 \leq i \leq d$  there exists  $\alpha_i > 0$  satisfying

$$\mathbb{E}[(b_i + c_i Z_0^2)^{\alpha_i}] = 1. \quad (4.5)$$

The further assumptions are readily checked using that  $b_i + c_i M$  has a density on  $[b, \infty)$  and that  $\mathbf{Q}_t \equiv \mathbf{q}$  is deterministic. We conclude that

$$\text{supp}(\tilde{\Theta}_0) \subset \cup_{1 \leq \ell \leq r} \mathbb{S}_\infty^{|I_\ell|-1}$$

Finally, by Lemma 3.2 (b), for each  $I_\ell$ -block,

$$\text{supp}(\tilde{\Theta}_0) \cap \mathbb{S}_\infty^{|I_\ell|-1} = \{\bar{\mathbf{q}}_\ell\}. \quad \square$$

Thus, despite the constant correlation matrix  $\Sigma$  being totally correlated, the volatility process exhibits asymptotic independence among marginals. On the other hand, if the components of the CCC-GARCH(1,1) model are independent, which is the case when  $\mathbf{C} = \text{Diag}(1, \dots, 1)$ , they are asymptotically independent as well.

We conjecture asymptotic independence to be true for all choices of  $N_t$  in the original CCC-GARCH(1,1) model, since the one we considered is the most dependent one. A recent preprint [10] gives a partial positive answer to this open question.

### 4.3 A simulation study

We provide some empirical evidences on the results of Theorem 3.1 and Lemma 3.2. We simulate a trajectory of length  $10^8$  of the bidimensional diagonal SRE

$$\mathbf{X}_t = \begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = \begin{pmatrix} b_1 + c_1 M_t & 0 \\ 0 & b_2 + c_2 M_t \end{pmatrix} \begin{pmatrix} X_{t-1,1} \\ X_{t-1,2} \end{pmatrix} + \mathbf{Q}_t.$$

Following (2.2), we consider the ratios  $(\|\mathbf{X}_t\|_\alpha^{-1/\alpha} \mathbf{X}_t)$  of the exceedances satisfying  $\|\mathbf{X}_t\|_\alpha > x$  as an approximation of the spectral component  $\tilde{\Theta}_0$ . We fix the threshold  $x$  as the empirical  $(1 - 10^{-5})$ -percentile of the simulated values  $(\|\mathbf{X}_t\|_\alpha)$ . We estimate empirically the angular measure  $\arctan(\tilde{\Theta}_{0,1}/\tilde{\Theta}_{0,2})$  by an histogram.

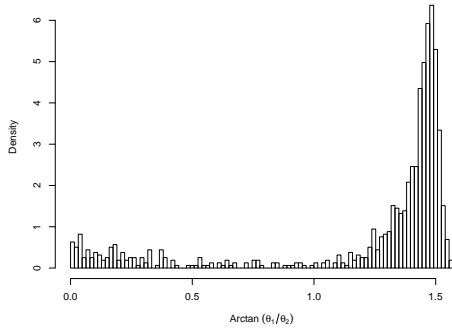


Figure 1:  $\alpha_1 = 2$  and  $\alpha_2 = 4$

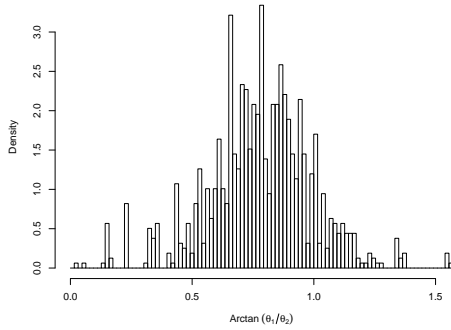


Figure 2:  $\alpha_1 = \alpha_2 = 2$

We first consider Case I with  $M_t \sim \mathcal{N}(0, 1)$  and  $\mathbf{Q}_t$  a bivariate standard gaussian vector with correlation 0.9. This corresponds to the diagonal BEKK-ARCH model. To simplify the graphical presentation, we consider the spectral measure for the absolute values of  $\mathbf{X}_t$ . Figure 1 corresponds to different coefficients  $c_1 = 1$  and  $c_2 = (1/3)^{1/4}$  whereas Figure 2 corresponds to equal coefficients  $c_1 = c_2 = 1$ . In accordance with Theorem 3.1, the angular measure in Figure 1 is concentrated around 0 and  $\pi/2$ . Following Lemma 3.2 (c) the support of the angular measure should be  $[0, \pi/2)$  which is not in contradiction with Figure 2.

Next we consider Case II with  $M_t = Z_t^2$ ,  $Z_t \sim \mathcal{N}(0, 1)$ ,  $b_1 = b_2 = 0.1$  and constant  $\mathbf{Q}_t = (0.2, 0.1)^\top$ . This corresponds to the totally correlated CCC-GARCH model. Again the angular measure is concentrated around 0 and  $\pi/2$  in Figure 3 for different coefficients  $c_1 = 0.9$  and  $c_2$  the positive root of  $3c_2^2 + 0.2c_2 + 0.01 = 1$ . For equal coefficients  $c_1 = c_2 = 0.9$  in Figure 4, the support of the angular measure of  $\tilde{\Theta}_0$  is concentrated around  $\arctan(Q_{t,1}/Q_{t,2}) = \arctan(2)$  as expected from Lemma 3.2 (b).

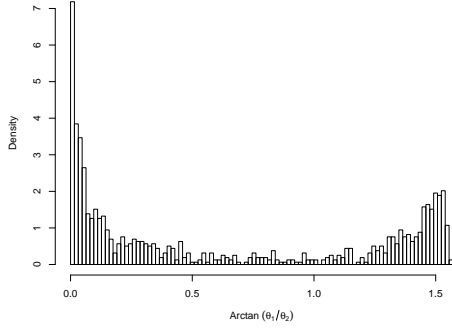


Figure 3:  $\alpha_1 = 1$  and  $\alpha_2 = 2$

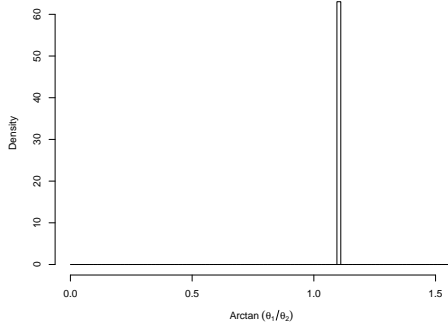


Figure 4:  $\alpha_1 = \alpha_2 = 1$

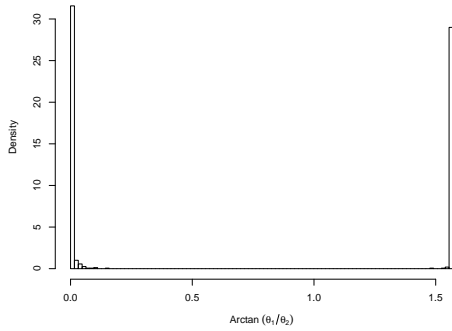


Figure 5:  $\alpha_1 = 1$  and  $\alpha_2 = 2$

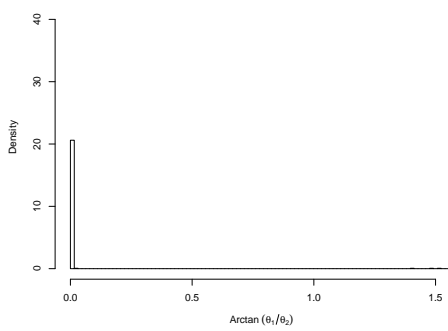


Figure 6:  $\alpha_1 = \alpha_2 = 1$

Figures 3 and 4 correspond to the totally correlated CCC-GARCH. In order to investigate our conjecture we consider the case where the covariance matrix  $\Sigma$  is standardized with correlation 0.5. In Figures 5 and 6 the parameters are the same as in Figures 3 and 4. Both angular measures are concentrated around 0 and  $\pi/2$ , supporting our conjecture and encouraging us to extend it: The asymptotic independence might be the rule for any CCC-GARCH model except when the CCC-GARCH model is totally correlated with equal marginal tail index.

#### 4.4 Real data application

We illustrate our approach by fitting a DCC-GARCH model on real data. The DCC-GARCH model was introduced in [9] as a more realistic generalization of the CCC-GARCH model satisfying the system of recursive equations

$$\begin{cases} \mathbf{R}_t &= \boldsymbol{\Sigma}_t N_t, & t \in \mathbb{Z}, \\ \boldsymbol{\Sigma}_t &= \text{Diag}(\sigma_{t,1}, \dots, \sigma_{t,d}), \\ \sigma_{t,i}^2 &= a_i + b_i \sigma_{t-1,i}^2 + c_i R_{t-1,i}^2, \end{cases} \quad (4.6)$$

where the correlation matrices  $(\mathbf{C}_t)$  of  $N_t$  are no longer constant but satisfies the dynamic model

$$\mathbf{C}_t = (1 - a - b)\mathbf{C} + a\mathbf{R}_t\boldsymbol{\Sigma}_t^{-2}\mathbf{R}_t^\top + b\mathbf{C}_{t-1}, \quad t \in \mathbb{Z},$$

for some  $a, b > 0$  such that  $a + b < 1$  and  $\mathbf{C}$  being an unconditional correlation matrix. Fitted<sup>2</sup> on IBM and GOOGLE log returns  $\mathbf{R}_t = (\mathbf{R}_{t,1}, \mathbf{R}_{t,2})$  from 2007-01-03 to 2021-08-30 (3690 observations), we found a large unconditional correlation of  $\mathbf{C}_{1,2} = 0.37$  together with some dynamical effect  $(a, b) = (0.03, 0.85)$ . The DCC-GARCH model captures the financial stylized fact called correlation clustering: The conditional correlation between log-returns is more likely to be high at time  $t$  if it was high at time  $t - 1$ . The marginal parameters of the GARCH(1,1) models are close  $(b_1, c_1) = (0.11, 0.82)$  and  $(b_2, c_2) = (0.08, 0.89)$  and solving the corresponding unit-root equations (4.5), we found very close tail indices  $(\alpha_1, \alpha_2) = (1.46, 1.41)$  for the squared log-ratios  $\mathbf{X}_t = \mathbf{R}_t^2$ . We consider again the ratios  $(\|\mathbf{X}_t\|_\alpha^{-1/\alpha} \mathbf{X}_t)$  of the exceedances satisfying  $\|\mathbf{X}_t\|_\alpha > x$ , for  $x$  the empirical  $(1 - 10^{-2})$ -percentile, as an approximation of the spectral component  $\tilde{\Theta}_0$ . We estimate empirically the angular measure  $\arctan(\tilde{\Theta}_{0,1}/\tilde{\Theta}_{0,2})$  by an histogram in Figure 7. We do not find clear evidence of asymptotic independence. Next, simulating a trajectory of length 3690 following the fitted DCC-GARCH model (4.6), we estimate similarly the angular measure using exceedances above the empirical  $(1 - 10^{-2})$ -percentile. We find in Figure 8 evidences of asymptotic independence despite the DCC-GARCH model does not satisfy the conditions of our paper.

This illustration may lead to several discussions; first, it seems, as conjectured at the end of Section 4.2, that the asynchrony in multivariate GARCH models with different tail indices is found beyond the conditions of our study and applies to DCC-GARCH models. Second, correlation clustering is not extremal clustering. When the values are extreme, the DCC-GARCH model does not capture any clustering among marginals with different tail indices. Third, it seems realistic to use GARCH models with different marginal tail indices capturing extreme clustering among them. However, up to our knowledge, such a model does not exist yet.

---

<sup>2</sup>We used the package `rmgarch` of the R-CRAN software [14].



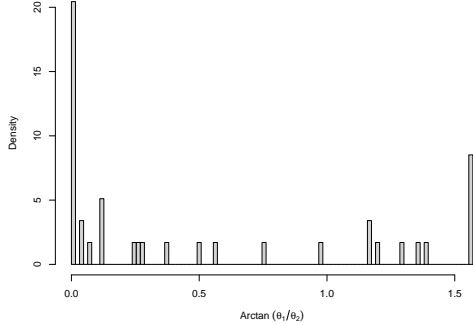


Figure 7: Log-returns

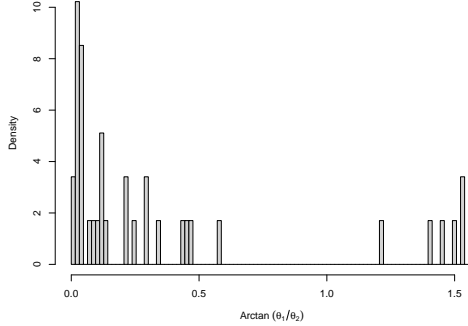


Figure 8: Fitted DCC-GARCH model

## 5 Proofs

To prove our main result, we proceed as follows. As the first step, in Subsection 5.1, we consider a bivariate diagonal SRE with distinct coefficients  $b_1, b_2$  and  $c_1, c_2$ ; for we can always reduce the study of asymptotic independence between blocks  $I_\ell$  to the comparison of two components with distinct coefficients.

As the second step, in Subsection 5.2, we consider a multivariate diagonal SRE where all coefficients are equal. There, we study properties of the spectral vector *within* one block  $I_\ell$ . This part also contains the proof of Lemma 3.2.

As the third step, in Subsection 5.3, we provide further properties of Vector Scaling Regular Variation (VSRV), in order to study the serial dependence structure provided by  $(\tilde{\Theta}_0, \dots, \tilde{\Theta}_t)$ .

We conclude with the proof of Theorem 3.1 in Subsection 5.4, using the findings of the previous steps.

### 5.1 The diagonal SRE with distinct coefficients

In this section we will show that the marginals of the diagonal SRE with distinct coefficients are asymptotically independent. A standard argument reduces the discussion to the bivariate case. We consider the bivariate random recursive process  $\mathbf{X}_t = \mathbb{M}_t \mathbf{X}_{t-1} + \mathbf{Q}_t$ , defined by  $\mathbf{X}_0 = 0$  and

$$\begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = \begin{pmatrix} b_1 + c_1 M_t & 0 \\ 0 & b_2 + c_2 M_t \end{pmatrix} \begin{pmatrix} X_{t-1,1} \\ X_{t-1,2} \end{pmatrix} + \mathbf{Q}_t. \quad (5.1)$$

We assume that  $(M_t)_{t \in \mathbb{N}}$  are iid random variables,  $(\mathbf{Q}_t)_{t \in \mathbb{N}}$  are iid  $\mathbb{R}^2$ -valued random vectors independent of  $(M_t)$  and that (A1)-(A6) are satisfied. We

assume as well that either (Case I) or (Case II) holds, which read in the bivariate setting as follows.

$$\begin{aligned} b_2 = b_1 = 0, \quad c_2 > c_1 > 0, \quad M_t \text{ is } \mathbb{R}\text{-valued} \quad (\text{Case I}) \\ b_2 \geq b_1 > 0, \quad c_2 > c_1 > 0, \quad \frac{c_2}{c_1} \geq \frac{b_2}{b_1}, \quad M_t > 0 \text{ a.s.} \quad (\text{Case II}) \end{aligned}$$

As before, by the definition of  $\alpha_i$  (see (A2)) it holds  $\alpha_1 > \alpha_2$ .

Under our assumptions, by the Kesten-Goldie-Theorem of [15, 19] applied to multiplicative factors with  $b_i + c_i M$ ,  $i = 1, 2$ , we have for the random variables  $X_i$ , defined by (2.3)

$$\lim_{u \rightarrow \infty} u^{\alpha_1} \mathbb{P}(X_1 > u) = a_1, \quad \lim_{u \rightarrow \infty} u^{\alpha_2} \mathbb{P}(X_2 > u) = a_2 \quad (5.2)$$

for constants  $a_1, a_2$  which are positive, see Section 2.3 for details. We are going to prove that

$$\lim_{u \rightarrow \infty} u \mathbb{P}(X_2 > u^{1/\alpha_2}, X_1 > u^{1/\alpha_1}) = 0. \quad (5.3)$$

which by (5.2) is equivalent to the asymptotic independence

$$\lim_{u \rightarrow \infty} \mathbb{P}(X_2 > u^{1/\alpha_2} \mid X_1 > u^{1/\alpha_1}) = 0.$$

of the extremes.

### 5.1.1 Reduction to the case of nonnegative $M$ and $Q_i$

From Definition (2.3), it is obvious that we can bound  $X_i$  by the following sums over nonnegative random variables:

$$X_i \leq \sum_{k=1}^{\infty} \prod_{\ell=1}^{k-1} |b_i + c_i M_\ell| |Q_{k,i}| =: X_i^*$$

We notice that  $X_i^*$  satisfies the fixed point equation, in distribution,

$$X_i^* \stackrel{\text{law}}{=} |b_i + c_i M| X_i^* + |Q_i|, \quad i = 1, 2,$$

(where  $\stackrel{\text{law}}{=}$  denotes equality in law between random variables on both sides). In particular, thanks to (A1)–(A4), the Kesten-Goldie theorem, now used in the case of positive coefficients, applies and yields

$$\lim_{u \rightarrow \infty} u \mathbb{P}(X_2^* > u^{1/\alpha_2}) = a_2^* > 0 \quad \lim_{u \rightarrow \infty} u \mathbb{P}(X_1^* > u^{1/\alpha_1}) = a_1^* > 0. \quad (5.4)$$

Note that the tail indices  $\alpha_1, \alpha_2$  remain unchanged thanks to their definition in (A2). Since  $|X_i| \leq X_i^*$ ,  $i = 1, 2$ , the result (5.3) will follow from the relation

$$\lim_{u \rightarrow \infty} u \mathbb{P}(X_2^* > u^{1/\alpha_2}, X_1^* > u^{1/\alpha_1}) = 0.$$

### 5.1.2 Asymptotic independent diagonal SRE

By the previous discussion, it is enough to consider the following cases

$$\begin{aligned} b_2 = b_1 = 0, \quad c_2 > c_1 > 0, \quad M_t, Q_i > 0 \text{ a.s. (Case I')} \\ b_2 \geq b_1 > 0, \quad c_2 > c_1 > 0, \quad \frac{c_2}{c_1} > \frac{b_2}{b_1}, \quad M_t, Q_i > 0 \text{ a.s.} \\ \text{(Case II')} \end{aligned}$$

We summarize these two cases under the condition  $c_2/c_1 > b_2/b_1 \geq 1$  and  $c_1 > 0$  (with the convention  $0/0 = 1$ ). We are going to prove the following result.

**Theorem 5.1.** *Assume (A1)–(A6) for  $i = 1, 2$  with  $c_2/c_1 > b_2/b_1 \geq 1$  and  $c_1 > 0$ . Then we have*

$$\lim_{u \rightarrow \infty} u \mathbb{P}(X_2 > u^{1/\alpha_2}, X_1 > u^{1/\alpha_1}) = 0,$$

*i.e.,  $X_1$  and  $X_2$  are asymptotically independent.*

The basic tool in the proof is to analyze the behavior of  $X_2$  under an exponential change of measure that favors large values for  $X_1$ . Namely, we consider the probability measure  $\mathbb{P}^{\alpha_1}$ , under which  $(M_n)$  is still an iid sequence, but with the new law

$$\mathbb{P}^{\alpha_1}(M \in \cdot) := \mathbb{E}[(b_1 + c_1 M)^{\alpha_1} 1(M \in \cdot)].$$

The law of the sequences  $(Q_{n,i})$  remains unchanged and independent of  $(M_n)$  under  $\mathbb{P}^{\alpha_1}$ .

Considering the random variables

$$W_1 := \log(b_1 + c_1 M) \quad \text{and} \quad W_2 := \log(b_2 + c_2 M),$$

with associated iid sequences  $W_{n,i} := \log(b_i + c_i M_n)$ , we denote their respective  $\mathbb{P}^{\alpha_1}$ -drift by

$$\mu_{j|1} := \mathbb{E}[\log(b_j + c_j M)(b_1 + c_1 M)^{\alpha_1}] = \mathbb{E}^{\alpha_1}[W_j], \quad j = 1, 2.$$

We have the following result.

**Lemma 5.2.** *In both (Case I') and (Case II'), it holds that*

$$\alpha_2 \mu_{2|1} < \alpha_1 \mu_{1|1}. \tag{5.5}$$

*Proof.* Using Jensen's inequality under the change of measure, we obtain

$$\begin{aligned}
\alpha_2 \mu_{2|1} - \alpha_1 \mu_{1|1} &= \mathbb{E} \left[ \log \left( \frac{(b_2 + c_2 M)^{\alpha_2}}{(b_1 + c_1 M)^{\alpha_1}} \right) (b_1 + c_1 M)^{\alpha_1} \right] \\
&= \mathbb{E}^{\alpha_1} \left[ \log \left( \frac{(b_2 + c_2 M)^{\alpha_2}}{(b_1 + c_1 M)^{\alpha_1}} \right) \right] \\
&< \log \mathbb{E}^{\alpha_1} \left[ \left( \frac{(b_2 + c_2 M)^{\alpha_2}}{(b_1 + c_1 M)^{\alpha_1}} \right) \right] \\
&= \log \mathbb{E} \left[ \frac{(b_2 + c_2 M)^{\alpha_2}}{(b_1 + c_1 M)^{\alpha_1}} (b_1 + c_1 M)^{\alpha_1} \right] = 0
\end{aligned}$$

The strict inequality holds since  $\log$  is strictly convex and the random variable  $(b_2 + c_2 M)^{\alpha_2} / (b_1 + c_1 M)^{\alpha_1}$  is not constant a.s., due to the different exponents and condition (A4) which implies that  $M$  is not degenerate.  $\square$

*Proof of Theorem 5.1.* We are going to study partial sums converging to the random variables  $X_1, X_2$  given by (2.3), namely

$$X_{j:m,i} := \sum_{k=j+1}^m \prod_{l=1}^{k-1} (b_i + c_i M_l) Q_{k,i}, \quad i = 1, 2. \quad (5.6)$$

We write  $X_{n,i} := X_{0:n,i}$  and observe that  $X_i = \lim_{n \rightarrow \infty} X_{n,i} = \sup_{n \geq 0} X_{n,i}$  a.s. Note the distinction between the Markov chain  $(X_{t,i})$  (the *forward process*) and the almost surely convergent series  $(X_{n,i})$  defined above (the *backward process*); see [20].

**Step 1.** We gain additional control by introducing the first exit time for  $(X_{n,1})$ ,

$$T_u := \inf \{n \in \mathbb{N} : X_{n,1} > u^{1/\alpha_1}\}.$$

As  $X_i = \sup_{n \geq 0} X_{n,i}$  for  $i = 1, 2$  we have  $\{X_1 > u^{1/\alpha_1}\} = \{T_u < \infty\}$ . By (5.4) we have

$$\lim_{u \rightarrow \infty} u \cdot \mathbb{P}(T_u < \infty) > 0. \quad (5.7)$$

Thus, the desired result will follow from the relation

$$\lim_{u \rightarrow \infty} \mathbb{P}(X_2 > u^{1/\alpha_2} \mid T_u < \infty) = 0. \quad (5.8)$$

On the set  $\{T_u < \infty\}$ , it holds

$$X_2 = X_{T_u,2} + \prod_{l=1}^{T_u} (b_2 + c_2 M_l) X_{T_u:\infty,2}. \quad (5.9)$$

The simple inclusion

$$\{X_2 > s\} \subset \underbrace{\{X_{T_u,2} > u^{1/\alpha_2}/2\}}_{=: A_u} \cup \underbrace{\left\{\prod_{l=1}^{T_u} (b_2 + c_2 M_l) X_{T_u:\infty,2} > u^{1/\alpha_2}/2\right\}}_{=: B_u}$$

allows us to consider the contributions in (5.9) separately. The following lemma, to be proved subsequently, provides stronger control and is the crucial ingredient for evaluating the contributions of  $A_u$  and  $B_u$ . The proof of this lemma is deferred to the appendix.

**Lemma 5.3.** *For any  $\epsilon > 0$ , define the set  $C_u(\epsilon)$  as the intersection*

$$\begin{aligned} & \left\{T_u \leq L_u\right\} \cap \left\{X_{T_u,1} \leq u^{\frac{1+\epsilon}{\alpha_1}}\right\} \cap \left\{\max_{1 \leq k \leq L_u} \frac{Q_{k,2}}{Q_{k,1}} \leq u^{\epsilon/\alpha_1}\right\} \\ & \cap \left\{\sum_{l=1}^{T_u} (W_{l,2} - W_{l,1}) - T_u(\mu_{2|1} - \mu_{1|1}) \leq \epsilon T_u\right\} \cap \left\{\sum_{l=1}^{L_u} W_{l,1} \leq \frac{1+\epsilon}{\alpha_1} \log u\right\} \end{aligned}$$

where  $L_u := \log(u)/(\mu_{1|1}\alpha_1) + Cf(u)$ ,  $f(u) := \sqrt{\log(u) \cdot \log(\log(u))}$  and  $C$  is a (suitably large) constant that can be chosen indepently of  $\epsilon$ . Then it holds that

$$\lim_{u \rightarrow \infty} \mathbb{P}(\{X_2 > u^{1/\alpha_2}\} \cap C_u(\epsilon) \mid T_u < \infty) = \lim_{u \rightarrow \infty} \mathbb{P}(X_2 > u^{1/\alpha_2} \mid T_u < \infty)$$

if either of the limits exists.

**Step 2.** Considering the event  $A_u$ , we have, using  $b_1 \leq b_2$  and  $c_1 < c_2$  and the controls provided by  $C_u(\epsilon)$ , that

$$\begin{aligned} X_{T_u,2} &= \sum_{k=1}^{T_u} \prod_{l=1}^{k-1} (b_2 + c_2 M_l) Q_{k,2} \\ &\leq \left(\max_{1 \leq k \leq T_u} \frac{Q_{k,2}}{Q_{k,1}}\right) \sum_{k=1}^{T_u} \prod_{l=1}^{k-1} \frac{b_2 + c_2 M_l}{b_1 + c_1 M_l} (b_1 + c_1 M_l) Q_{k,1} \\ &\leq \left(\max_{1 \leq k \leq L_u} \frac{Q_{k,2}}{Q_{k,1}}\right) \left(\prod_{l=1}^{T_u-1} \frac{b_2 + c_2 M_l}{b_1 + c_1 M_l}\right) \sum_{k=1}^{T_u} \prod_{l=1}^{k-1} (b_1 + c_1 M_l) Q_{k,1} \\ &\leq \left(\prod_{l=1}^{T_u-1} \frac{b_2 + c_2 M_l}{b_1 + c_1 M_l}\right) \left(\max_{1 \leq k \leq L_u} \frac{Q_{k,2}}{Q_{k,1}}\right) X_{T_u,1} \\ &\leq \left(\prod_{l=1}^{T_u} \frac{b_2 + c_2 M_l}{b_1 + c_1 M_l}\right) u^{\epsilon/\alpha_1} u^{(1+\epsilon)/\alpha_1} \\ &\leq e^{\sum_{l=1}^{T_u} W_{l,2} - W_{l,1}} u^{(1+2\epsilon)/\alpha_1}. \end{aligned} \tag{5.10}$$

Now we use that on  $C_u(\varepsilon)$  we have the relation

$$\sum_{l=1}^{T_u} (W_{l,2} - W_{l,1}) \leq T_u(\mu_{2|1} - \mu_{1|1}) + \epsilon T_u \leq L_u(\mu_{2|1} - \mu_{1|1} + \epsilon)$$

so that (5.10) yields

$$\begin{aligned} \frac{\log X_{T_u,2}}{\log u} &\leq \frac{\mu_{2|1} - \mu_{1|1} + \epsilon}{\mu_{1|1}\alpha_1} + \frac{1 + 3\epsilon}{\alpha_1} \\ &= \frac{\mu_{2|1} + \epsilon(1 + 3\mu_{1|1})}{\mu_{1|1}\alpha_1} = \frac{1}{\alpha_2} \frac{\alpha_2\mu_{2|1} + \epsilon\alpha_2(1 + 3\mu_{1|1})}{\alpha_1\mu_{1|1}}. \end{aligned} \quad (5.11)$$

Here we have used that

$$\exp(\sqrt{\log u}) \log u = \exp(\log u / \sqrt{\log u}) \log u = u^{1/\sqrt{\log u}} \log u \leq u^{\epsilon/\alpha_1}$$

for any fixed  $\epsilon > 0$ , as soon as  $u$  is large enough.

Under the condition (5.5) it is always possible to find  $\epsilon$  so small that

$$\eta := \frac{1}{\alpha_2} \frac{\alpha_2\mu_{2|1} + \epsilon\alpha_2(1 + 3\mu_{1|1})}{\alpha_1\mu_{1|1}} \leq \frac{1}{\alpha_2} - \epsilon$$

and hence by (5.11),

$$\{X_{T_u,2} > u^{1/\alpha_2}/2\} \cap C_u(\epsilon) \subset \{u^\eta \geq X_{T_u,2} > u^{1/\alpha_2}/2\} = \emptyset$$

for  $u$  sufficiently large. It follows that the first term  $A_u$  in (5.9) does not contribute on  $C_u(\varepsilon)$ .

**Step 3.** Turning to  $B_u$ , we start by bounding the multiplicative factor on  $C_u(\epsilon)$ . By Lemma 5.3,

$$\begin{aligned} \prod_{l=1}^{T_u} (b_2 + c_2 M_l) &= \exp\left(\sum_{l=1}^{T_u} (W_{l,2} - W_{l,1})\right) \exp\left(\sum_{l=1}^{T_u} W_{l,1}\right) \\ &\leq e^{L_u(\mu_{2|1} - \mu_{1|1} + \epsilon)} u^{(1+\epsilon)/\alpha_1} \leq u^\eta \end{aligned}$$

where we used the same calculations as the ones leading to (5.11). Hence

$$\begin{aligned} &\mathbb{P}\left(\left\{\prod_{l=1}^{T_u} (b_2 + c_2 M_l) X_{T_u:\infty,2} > \frac{1}{2} u^{1/\alpha_2}\right\} \cap C_u(\epsilon) \mid T_u < \infty\right) \\ &\leq \mathbb{P}(X_{T_u:\infty,2} > u^{1/\alpha_2 - \eta}/2 \mid T_u < \infty) = \mathbb{P}(X_2 > u^{1/\alpha_2 - \eta}/2). \end{aligned}$$

since  $X_{T_u:\infty,2}$  is independent of  $\{T_u < \infty\}$ . Since  $1/\alpha_2 > \eta$ , the last probability tends to zero.

Combining the two previous steps, we have proved that

$$\lim_{u \rightarrow \infty} \mathbb{P}\left(\left\{X_2 > u^{1/\alpha_2}\right\} \cap C_u(\epsilon) \mid T_u < \infty\right) = 0$$

which by Lemma 5.3 is enough to conclude (5.8) and thus the desired result.  $\square$

## 5.2 The diagonal SRE with equal coefficients

In this section we focus on the case where  $b_i = b \geq 0$  and  $c_i = c > 0$  for any  $1 \leq i \leq d$  so that

$$\mathbf{X}_t = (b + cM_t)\mathbf{X}_{t-1} + \mathbf{Q}_t, \quad t \in \mathbb{Z}.$$

We can interpret the multiplicative factor  $(b + cM_t)$  as multiplication with the random similarity matrix  $(b + cM_t)I_d$ , thus we are in the framework of [4]. From there, we obtain the following result:

**Theorem 5.4.** *Assume (A1)–(A5) for all  $1 \leq i \leq d$ . Let  $\mathbf{X}_0$  have the stationary distribution. Then  $\mathbf{X}_0$  is VSRV and  $(\mathbf{X}_t)_{t \geq 0}$  is a VSRV process of order  $\alpha = (\alpha, \dots, \alpha)$ , and its spectral tail process satisfies the relation*

$$\tilde{\Theta}_t = (b + cM_t)\tilde{\Theta}_{t-1}, \quad t \geq 1.$$

*Proof.* By [4, Theorem 1.6], there is a non-null Radon measure  $\mu$  on  $[-\infty, \infty]^d \setminus \{0\}$  such that

$$x^\alpha \mathbb{P}(x^{-1}\mathbf{X}_0 \in \cdot) \xrightarrow{v} \mu, \quad x \rightarrow \infty.$$

[See [5, Theorem 4.4.21] for a reformulation of the quoted result which is more consistent with our notation.] Hence,  $\mathbf{X}_0$  is (standard) regularly varying and also VSRV of order  $\alpha = (\alpha, \dots, \alpha)$  since  $\Theta_0$  and  $\tilde{\Theta}_0$  coincide then.

The remaining assertions follow from a direct application of Proposition 5.8.  $\square$

In order to determine whether the components of  $\mathbf{X}_0$  are asymptotically independent or dependent, we are interested in information about the support of  $\mathbb{P}(\tilde{\Theta}_0 \in \cdot)$ . By (2.2), two or more components of  $\mathbf{X}_0$  can be large simultaneously if and only if two or more components of  $\tilde{\Theta}_0$  can be nonzero simultaneously; i.e.,  $\tilde{\Theta}_0$  is not concentrated on the standard basis vectors  $e_i$ .

We write  $\text{supp}(\mathbf{Q})$  for the support of the law of  $\mathbf{Q}$  and  $\text{span}(E)$  for the linear space spanned by set  $E \subset \mathbb{R}^d$ . Let  $S_\infty^{d-1}$  denote the unit sphere in  $\mathbb{R}^d$  with respect to  $\|\cdot\|_\alpha$  which coincides with the unit sphere for the max-norm whatever is  $\alpha$ .

**Lemma 5.5.** *Under the assumptions of Theorem 5.4,*

$$\text{supp}(\tilde{\Theta}_0) \subset \text{span}(\text{supp}(\mathbf{Q})) \cap S_\infty^{d-1}. \quad (5.12)$$

*In addition, the following implications hold:*

(a) *If  $b = 0$ ,  $c > 0$  and  $\text{supp}(M)$  is dense in  $\mathbb{R}$ , then*

$$\text{supp}(\tilde{\Theta}_0) = \text{span}(\text{supp}(\mathbf{Q})) \cap S_\infty^{d-1}.$$

(b) If  $b > 0$ ,  $c > 0$  and  $\text{supp}(M)$  is dense in  $\mathbb{R}_+$ , then

$$\text{supp}(\tilde{\Theta}_0) = \{a_1 \mathbf{q}_1 + \cdots + a_n \mathbf{q}_n : n \in \mathbb{N}, a_i > 0, \mathbf{q}_i \in \text{supp}(\mathbf{Q})\} \cap S_\infty^{d-1},$$

i.e. it equals the convex cone generated by  $\text{supp}(\mathbf{Q})$  intersected by the unit sphere.

(c) If  $\text{supp}(\mathbf{Q})$  is dense in  $\mathbb{R}^d$ , then  $\text{supp}(\tilde{\Theta}_0) = S_\infty^{d-1}$ .

*Proof of Lemma 5.5.* The proof is based on [4, Remark 1.9], which gives that the support of the spectral measure  $\sigma_\infty$  with respect to the Euclidean norm is given by the directions (subsets of the unit sphere  $S^{d-1}$ ) in which the support of  $\mathbf{X}_0$  is unbounded. More precisely, consider the measures

$$\sigma_t(A) := \mathbb{P}\left(\|\mathbf{X}_0\|_2 > t, \frac{\mathbf{X}_0}{\|\mathbf{X}_0\|_2} \in A\right)$$

Then  $\text{supp}(\sigma_\infty) = \bigcap_{t>0} \text{supp}(\sigma_t)$ . The surprising part of this result is that all directions, in which the support of  $\mathbf{X}_0$  is unbounded, do matter. One does not need a lower bound on the decay of mass at infinity. But if we know that the support of the spectral measure w.r.t. the Euclidean norm is the intersection of a particular subspace with the unit sphere, we immediately deduce the same for the spectral measure w.r.t the max-norm, i.e., for  $\mathbb{P}(\Theta_0 \in \cdot)$ , as well as for  $\mathbb{P}(\tilde{\Theta} \in \cdot)$ .

Thus, to proceed, we have to study the support of  $\mathbf{X}_0$ . For simplicity we write, for the remainder of the proof,  $(m, \mathbf{q})$  for a realization of the random variables  $(b + cM, \mathbf{Q})$ . We identify a pair  $(m, \mathbf{q})$  with the affine mapping  $h(\mathbf{x}) = m\mathbf{x} + \mathbf{q}$ , we say that  $h \in \text{supp}((b + cM, \mathbf{Q}))$  if  $(m, \mathbf{q}) \in \text{supp}((b + cM, \mathbf{Q}))$ . We consider the semigroup generated by mappings in  $\text{supp}((b + cM, \mathbf{Q}))$ ,

$$\mathcal{G} := \left\{ h_1 \cdots h_n : h_i \in \text{supp}((M, \mathbf{Q})), 1 \leq i \leq n, n \geq 1 \right\}.$$

Then, by [4, Lemma 2.7]

$$\text{supp}(\mathbf{X}_0) = \text{closure of } \left\{ \frac{1}{1-m} \mathbf{q} : (m, \mathbf{q}) \in \mathcal{G}, |m| < 1 \right\}.$$

This is obviously a subset of  $\text{span}(\mathbf{Q})$ , hence (5.12) holds. [Again, see [5, Proposition 4.3.1] for a reformulation of the quoted result which is more consistent with our notation.]

Since  $M$  and  $\mathbf{Q}$  are independent,  $\text{supp}((b + cM, \mathbf{Q})) = \text{supp}(b + cM) \times \text{supp}(\mathbf{Q})$  and a general element in  $\mathcal{G}$  is of the form

$$h(\mathbf{x}) = m_1 \cdots m_n \mathbf{x} + \left( \mathbf{q}_1 + \sum_{k=2}^n m_1 \cdots m_{k-1} \mathbf{q}_k \right)$$



with  $m_i \in \text{supp}(b + cM)$ ,  $\mathbf{q}_i \in \text{supp}(\mathbf{Q})$ . Thus, a generic point in  $\text{supp}(\mathbf{X}_0)$  is of the form

$$\frac{1}{1 - m_1 \cdots m_n} \left( \mathbf{q}_1 + \sum_{k=2}^n m_1 \cdots m_{k-1} \mathbf{q}_k \right), \quad (5.13)$$

with

$$m_i \in \text{supp}(b + cM), \quad \mathbf{q}_i \in \text{supp}(\mathbf{Q}), \quad |m_1 \cdots m_n| < 1.$$

The prefactor in (5.13) is scalar, while the bracket term represents a linear combination of  $\mathbf{q}_k \in \text{supp}(\mathbf{Q})$ . Now we can prove the two implications.

Concerning (a), if  $\text{supp}(M)$  is dense in  $\mathbb{R}$ , then the bracket term in (5.13) can be chosen such that its direction approximates any direction of  $y \in \text{span}(\text{supp}(\mathbf{Q}))$ . Then, given  $t > 0$ ,  $m_n$  can be chosen arbitrarily small, such that  $|m_1 \cdots m_n| < 1$  and moreover, the norm of (5.13) exceeds  $t$ . It follows that  $\text{supp}(\sigma_t) = \text{span}(\text{supp}(\mathbf{Q})) \cap S^{d-1}$  for all  $t$ , which yields the assertion since  $\text{supp}(\sigma_\infty) = \bigcap_{t>0} \text{supp}(\sigma_t)$ .

Concerning (b), note that  $m \in \text{supp}(b + cM)$  is bounded from below by  $b > 0$ , with  $b < 1$  due to assumption (A1). If  $\text{supp}(M)$  is dense in  $\mathbb{R}_+$ , then the bracket term in (5.13) can be chosen such that its direction approximates any direction of  $y \in \text{span}(\text{supp}(\mathbf{Q}))_+$  given that its norm is suitably large. Similarly as for (a) above, the desired assertion follows.

Concerning (c), if  $\text{supp}(\mathbf{Q})$  is dense in  $\mathbb{R}^d$ , then the bracket term can be chosen such that it approximates an arbitrary element of  $\mathbb{R}^d$  and its modulus is larger than  $t$ , while (A1) entails that there are  $m_i \in \text{supp}(M)$  such that  $|m_1 \cdots m_n| < 1$ .  $\square$

**Remark 5.6.** We conclude for the diagonal SRE with equal coefficients, under the assumptions of Lemma 5.5: As soon as  $\mathbf{Q}$  is not confined to a linear subspace of  $\mathbb{R}^d$ , we have that  $\mathbf{X}_0$  is multivariate regularly varying and its components are asymptotically dependent. In fact, the spectral measure charges the whole unit sphere.

### 5.3 Stationary VSRV Markov chains

We adapt the work of [18] to our framework. We consider a Markov chain  $(\mathbf{X}_t)_{t \geq 0}$  with values in  $\mathbb{R}^d$  satisfying the recursive equation

$$\mathbf{X}_t = \Phi(\mathbf{X}_{t-1}, Z_t), \quad t \geq 0, \quad (5.14)$$

where  $\Phi : \mathbb{R}^d \times \mathcal{E} \mapsto \mathbb{R}^d$  is measurable and  $(Z_t)$  is an iid sequence taking values in a Polish space  $\mathcal{E}$ . We work under the following assumption, which is the vector scaling adaptation of [18, Condition 2.2]. As above, we fix in advance the positive indices  $\alpha_1, \dots, \alpha_d$ .

**VS Condition for Markov chains:** *There exists a measurable function  $\phi : S_\infty^{d-1} \times \mathcal{E} \mapsto \mathbb{R}^d$  such that, for all  $e \in \mathcal{E}$ ,*

$$\lim_{x \rightarrow \infty} x^{-1/\alpha} \Phi(x^{1/\alpha} \mathbf{s}(x), e) \rightarrow \phi(\mathbf{s}, e),$$

*whenever  $\mathbf{s}(x) \rightarrow \mathbf{s}$  in  $S_\infty^{d-1}$ . Moreover, if  $\mathbb{P}(\phi(\mathbf{s}, Z_0) = 0) > 0$  for some  $\mathbf{s} \in S_\infty^{d-1}$  then  $Z_0 \in \mathcal{W}$  a.s. for a subset  $\mathcal{W} \subset \mathcal{E}$  such that, for all  $e \in \mathcal{W}$ ,*

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

We extend  $\phi$  over  $\mathbb{R}^d \times \mathcal{E}$  thanks to the relation

$$\phi(\mathbf{v}, e) = \begin{cases} \|\mathbf{v}\|_\alpha^{1/\alpha} \phi\left(\|\mathbf{v}\|_\alpha^{-1/\alpha} \mathbf{v}, e\right) & \text{if } \mathbf{v} \neq 0, \\ 0 & \text{if } \mathbf{v} = 0. \end{cases}$$

We have the following result which extends Theorem 2.1 of [18]

**Theorem 5.7.** *If the Markov chain  $(\mathbf{X}_t)$  satisfies the recursion (5.14) with  $\Phi$  satisfying the VS condition and if the vector  $\mathbf{X}_0$  is VSRV with positive indices  $\alpha_1, \dots, \alpha_d$  then  $(\mathbf{X}_t)_{t \geq 0}$  is a VSRV process and its spectral tail process satisfies the relation*

$$\tilde{\Theta}_t = \phi(\tilde{\Theta}_{t-1}, Z_t), \quad t \geq 0.$$

*started from  $\tilde{\Theta}_0$ , the spectral component of  $\mathbf{X}_0$ .*

*Proof.* The result follows by an application of Theorem 2.1 in [18] to the Markov chain  $(\mathbf{Y}_t)_{t \geq 0} = (\mathbf{X}_t^\alpha)_{t \geq 0}$ . We have  $\mathbf{Y}_0$  regularly varying since  $\mathbf{X}_0^\alpha$  is VSRV. Moreover

$$\mathbf{Y}_t = \tilde{\Phi}(\mathbf{Y}_{t-1}, Z_t), \quad t \geq 0,$$

with  $\tilde{\Phi}(x, z) = (\Phi(x^{1/\alpha}, z))^\alpha$ . As the VS condition for Markov chain is the vector scaling version of the condition 2.2. of [18] on  $\tilde{\Phi}$  associated to the limit  $\tilde{\phi}((x, z)) = \phi((x^{1/\alpha}, z))^\alpha$ , i.e.

$$\lim_{x \rightarrow \infty} x^{-1} \tilde{\Phi}(x \mathbf{s}(x), e) \rightarrow \tilde{\phi}(\mathbf{s}, e)$$

whenever  $\mathbf{s}(x) \rightarrow \mathbf{s}$  in  $S_\infty^{d-1}$ . We obtain that the spectral tail process of  $(\mathbf{Y}_t)_{t \geq 0}$  satisfies the recursion

$$\Theta_t^Y = \tilde{\phi}(\Theta_{t-1}^Y, Z_t), \quad t \geq 1.$$

The desired result follows as  $\tilde{\phi}((x, z)) = \tilde{\phi}((x^{1/\alpha}, z))^\alpha$  and  $\tilde{\Theta}_t^\alpha = \Theta_t^Y$ ,  $t \geq 0$ .  $\square$

We are specially interested in Stochastic Recurrence Equations (SRE) corresponding to the Markov chains

$$\mathbf{X}_t = \Phi(\mathbf{X}_{t-1}, (\mathbf{M}, \mathbf{Q})_t) = \mathbf{M}_t \mathbf{X}_{t-1} + \mathbf{Q}_t, \quad t \geq 0.$$

In this setting  $(\mathbf{M}_t)$  are iid random  $d \times d$  matrices and  $(\mathbf{Q}_t)$  iid random vectors in  $\mathbb{R}^d$ . We have

**Proposition 5.8.** *The SRE Markov chain  $(\mathbf{X}_t)_{t \geq 0}$  satisfies Condition VS for positive indices  $\alpha_1, \dots, \alpha_d$  if and only if  $M_{ij} = 0$  a.s. for any  $(i, j)$  so that  $\alpha_i > \alpha_j$ . Then*

$$\phi(\mathbf{s}, (\mathbf{M}, \mathbf{Q})) = \left( \sum_{j=1}^d M_{ij} \mathbf{1}_{\alpha_i = \alpha_j} s_j \right)_{1 \leq i \leq d}.$$

*Proof.* As  $x \rightarrow \infty$  and  $\mathbf{s}(x) \rightarrow \mathbf{s}$ , we have

$$\begin{aligned} \lim_{x \rightarrow \infty} x^{-1/\alpha} \Phi((x^{1/\alpha} \mathbf{s}(x), (\mathbf{M}, \mathbf{Q}))) &= \lim_{x \rightarrow \infty} x^{-1/\alpha} (\mathbf{M}(x^{1/\alpha} \mathbf{s}(x)) + \mathbf{Q}) \\ &= \lim_{x \rightarrow \infty} \left( \sum_{j=1}^d M_{ij} s(x)_j x^{1/\alpha_j - 1/\alpha_i} \right)_{1 \leq i \leq d}. \end{aligned}$$

Each coordinate converges to  $\sum_{j=1}^d M_{ij} \mathbf{1}_{\alpha_i = \alpha_j} s_j$  for any  $\mathbf{s} \in S_\infty^{d-1}$  if and only if  $M_{ij} = 0$  a.s. for any  $(i, j)$  so that  $\alpha_i > \alpha_j$ .  $\square$

**Remark 5.9.** In case of distinct  $\alpha_i$ 's, it means that the dynamic tail process depends only on the diagonal elements of  $\mathbf{M}$ . In general, specifying  $\mathbf{M}_t$  to be diagonal, we ensure that if  $\mathbf{X}_0$  is VSRV then the SRE process is VSRV with

$$\tilde{\Theta}_t = \mathbf{M}_t \tilde{\Theta}_{t-1}, \quad t \geq 1,$$

whatever are the positive indices  $\alpha_1, \dots, \alpha_d$ .

## 5.4 Proof of the Main Result

*Proof of Theorem 3.1.* We start by proving that  $(\mathbf{X}_t)$  is a VSRV process. According to Proposition 5.8 and Remark 5.9, it suffices to prove that  $\mathbf{X}_0$  is VSRV, then (3.1) and the VSRV of  $(\mathbf{X}_t)$  follow.

We use the following short-hand notation: For  $\mathbf{x} \in \mathbb{R}^d$ , let  $\mathbf{x}_\ell = (x_i)_{i \in I_\ell}$ ,  $\|\mathbf{x}\|_\ell := \max_{i \in I_\ell} |x_i|$  and  $\alpha(\ell)$  is the common tail index of all coordinates in  $I_\ell$ .

Let  $\epsilon > 0$ ,  $\ell \neq k$ . By (5.3) of Theorem 5.1, it holds that

$$\begin{aligned} & \lim_{x \rightarrow \infty} x \cdot \mathbb{P} \left( \|\mathbf{X}_0\|_\ell > \epsilon \|\mathbf{X}_0\|_\alpha^{1/\alpha(\ell)}, \|\mathbf{X}_0\|_k > \epsilon \|\mathbf{X}_0\|_\alpha^{1/\alpha(k)}, \|\mathbf{X}_0\|_\alpha > x \right) \\ & \leq \lim_{x \rightarrow \infty} x \cdot \mathbb{P} \left( \|\mathbf{X}_0\|_\ell > \epsilon x^{1/\alpha(\ell)}, \|\mathbf{X}_0\|_k > \epsilon x^{1/\alpha(k)} \right) \\ & \leq \sum_{i \in I_\ell, j \in I_k} \lim_{x \rightarrow \infty} x \cdot \mathbb{P} \left( |X_{0,i}| > \epsilon x^{1/\alpha_i}, |X_{0,j}| > \epsilon x^{1/\alpha_j} \right) = 0 \end{aligned} \quad (5.15)$$

We note from the results of Section 5.2 that there are positive constants  $c_\ell$  and probability measures  $\tilde{\xi}_\ell$  on the  $|I_\ell|$ -dimensional unit sphere (w.r.t. the max-norm), such that for all  $1 \leq \ell \leq r$

$$\lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_\ell > x^{1/\alpha(\ell)}, \|\mathbf{X}_0\|_\ell^{-1} \mathbf{X}_{0,\ell} \in \cdot\right) = c_\ell \tilde{\xi}_\ell(\cdot).$$

Applying the inclusion-exclusion principle, we have

$$\begin{aligned} \lim_{x \rightarrow \infty} x \cdot \mathbb{P}(\|\mathbf{X}_0\|_\alpha > x) &= \lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\bigvee_{1 \leq \ell \leq r} \|\mathbf{X}_0\|_\ell > x^{1/\alpha(\ell)}\right) \\ &= \sum_{1 \leq \ell \leq r} \lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_\ell > x^{1/\alpha(\ell)}\right) \\ &\quad - \sum_{1 \leq \ell < k \leq r} \lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_\ell > x^{1/\alpha(\ell)}, \|\mathbf{X}_0\|_k > x^{1/\alpha(k)}\right) + \dots \\ &= c_1 + \dots + c_r =: c, \end{aligned} \tag{5.16}$$

since all intersection terms vanish asymptotically due to 5.15 (with  $\epsilon = 1$ ).

Thus we have shown that  $\|\mathbf{X}_0\|_\alpha$  is regularly varying. We claim that

$$\lim_{x \rightarrow \infty} \mathbb{P}\left(\|\mathbf{X}_0\|_\alpha^{-1/\alpha} \mathbf{X}_0 \in \cdot \mid \|\mathbf{X}_0\|_\alpha > x\right) = \frac{1}{c} \sum_{1 \leq \ell \leq r} c_\ell \tilde{\xi}_\ell(\cdot),$$

where  $\xi_\ell$  is the extension of  $\tilde{\xi}_\ell$  to a measure on the unit sphere  $S_\infty^{d-1}$  in  $\mathbb{R}^d$  by putting unit mass in the origin of the additional coordinates. Hence, its support is contained in  $\mathbb{S}_\infty^{|I_\ell|-1}$ . In particular, (3.2) follows once this claim is proved.

By the Portmanteau lemma, it suffices to study closed sets. Note that for any closed set  $B \subset S_\infty^{d-1}$ , it holds that

$$B_{\ell,\epsilon} := \{\mathbf{x}_\ell : \mathbf{x} \in B, |x_j| < \epsilon \text{ for } j \notin I_\ell\} \rightarrow \{\mathbf{x}_\ell : \mathbf{x} \in B \cap \mathbb{S}^{|I_\ell|-1}\} =: B_\ell$$

as  $\epsilon \rightarrow 0$ . Using (5.15) and the inclusion-exclusion principle, we obtain

$$\begin{aligned}
& \limsup_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha} \mathbf{X}_0 \in B, \|\mathbf{X}_0\|_{\alpha} > x\right) \\
&= \limsup_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha} \mathbf{X}_0 \in B, \bigvee_{1 \leq k \leq r} \|\mathbf{X}_0\|_k > x^{1/\alpha(k)}\right) \\
&= \sum_{1 \leq \ell \leq r} \limsup_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha} \mathbf{X}_0 \in B, \|\mathbf{X}_0\|_{\ell} > x^{1/\alpha(\ell)}, \right. \\
&\quad \left. \bigwedge_{k \neq \ell} \|\mathbf{X}_0\|_k \leq \epsilon \|\mathbf{X}_0\|_{\alpha}^{1/\alpha(k)}\right) \\
&\leq \sum_{1 \leq \ell \leq r} \lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\ell}^{-1} \mathbf{X}_{0,\ell} \in B_{\ell,\epsilon}, \|\mathbf{X}_0\|_{\ell} > x^{1/\alpha(\ell)}, \right. \\
&\quad \left. \bigwedge_{k \neq \ell} \|\mathbf{X}_0\|_k \leq \epsilon x^{1/\alpha(k)}\right) \\
&= \sum_{1 \leq \ell \leq r} \lim_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\ell}^{-1} \mathbf{X}_{0,\ell} \in B_{\ell,\epsilon}, \|\mathbf{X}_0\|_{\ell} > x^{1/\alpha(\ell)}\right) \\
&= \sum_{1 \leq \ell \leq r} c_{\ell} \tilde{\xi}_{\ell}(B_{\ell,\epsilon})
\end{aligned}$$

This holds for all  $\epsilon > 0$ . Since the sequence  $B_{\ell,\epsilon}$  is decreasing, we conclude by the continuity of  $\tilde{\xi}_{\ell}$  that

$$\begin{aligned}
\limsup_{x \rightarrow \infty} x \cdot \mathbb{P}\left(\|\mathbf{X}_0\|_{\alpha}^{-1/\alpha} \mathbf{X}_0 \in B, \|\mathbf{X}_0\|_{\alpha} > x\right) &\leq \sum_{1 \leq \ell \leq r} c_{\ell} \tilde{\xi}_{\ell}(B_{\ell}) \\
&= \sum_{1 \leq \ell \leq r} c_{\ell} \xi_{\ell}(B).
\end{aligned}$$

Combined with (5.16), this proves the weak convergence by an application of the Portmanteau lemma.  $\square$

## Appendix

### A VSRV and non-standard regular variation

In this section we show that indeed any VSRV random vector with positive coordinates  $\mathbf{X}_0 \in (0, \infty)^d$  is also non-standard regularly varying; which is defined in [26] as follows.

Assume that marginals are positive and (one-dimensional) regularly varying with possibly different tail indices  $\alpha_i$  and cdf  $F_i$ ,  $1 \leq i \leq d$ . Then non-standard regular variation holds if and only if

$$\lim_{x \rightarrow \infty} x \cdot \mathbb{P}(x^{-1} \widetilde{\mathbf{X}}_0 \in \cdot)$$

exists in the vague sense, where the standardized vector  $\widetilde{\mathbf{X}}_0$  is defined as

$$\widetilde{\mathbf{X}}_0 = (1/(1 - F_i(X_{0,i})))_{1 \leq i \leq d}.$$

Following [11, Theorem 4], we note that  $\widetilde{\mathbf{X}}_0$  is regularly varying in the classical sense, i.e.  $\|\widetilde{\mathbf{X}}_0\|$  is regularly varying with tail index 1 and there exists an angular measure which is the weak limit of

$$\lim_{x \rightarrow \infty} \mathbb{P}\left(\|\widetilde{\mathbf{X}}_0\|^{-1} \widetilde{\mathbf{X}}_0 \in \cdot \mid \|\widetilde{\mathbf{X}}_0\| > x\right).$$

Note that the standardization is made so that all coordinates of  $\widetilde{\mathbf{X}}_0$  are tail equivalent

$$\mathbb{P}(\widetilde{X}_{0,i} > x) \sim x^{-1}, \quad x \rightarrow \infty, \quad 1 \leq i \leq d.$$

**Proposition A.1.** *Let  $\mathbf{X}_0$  be a VSRV random vector with positive coordinates of order  $\alpha = (\alpha_1, \dots, \alpha_d)$ . Then  $\mathbf{X}_0$  is non-standard regularly varying and the angular measure is given by*

$$\frac{\mathbb{E}[\|\mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha\| \mathbf{1}(\|\mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha\|^{-1} \mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha \in \cdot)]}{\mathbb{E}[\|\mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha\|]},$$

where  $\mathbf{a} = (a_1, \dots, a_d)$  is the vector of standardization coefficients given by 2.1.

We remark that the angular measure of  $\mathbf{X}_0$  is completely determined by the spectral tail process  $(\widetilde{\Theta}_t)$ . However its expression is intricate because of the different marginal standardizations  $\mathbf{a}$  whereas we will derive explicit expressions of  $(\widetilde{\Theta}_t)$  for many Markov chains in Section 5.3. We emphasize that this simplicity is the main motivation for introducing the notion of VSRV rather than using the more general notion of non-standard regular variation.

*Proof.* The standardized vector

$$\widetilde{\mathbf{X}}_0' = \mathbf{a}^{-1} \mathbf{X}_0^\alpha$$

has marginal tails equivalent to the standard Pareto marginally distributed vector  $\widetilde{\mathbf{X}}_0 = (1/(1 - F_i(X_{0,i})))_{1 \leq i \leq d}$ . Moreover  $\|\widetilde{\mathbf{X}}_0'\|_\alpha$  tail is Pareto equivalent with tail index 1 since an union bound yields

$$\mathbb{P}(|X_{0,1}|^{\alpha_1} > x) \leq \mathbb{P}(\|\mathbf{X}_0\|_\alpha > x) \leq \sum_{i=1}^d \mathbb{P}(|X_{0,i}|^{\alpha_i} > x)$$

and a sandwich argument concludes. Thus  $\|\widetilde{\mathbf{X}}'_0\|$  is also regularly varying because, denoting  $a_* = \min_{1 \leq i \leq d} a_i$  and  $a_0 = \lim x^{-1} \mathbb{P}(\|\widetilde{\mathbf{X}}'_0\|_\alpha > x)$ , we have

$$\begin{aligned}
\mathbb{P}(\|\widetilde{\mathbf{X}}'_0\| > x) &= \mathbb{P}(\|\mathbf{a}^{-1} \mathbf{X}_0^\alpha\| > x, a_*^{-1} \|\mathbf{X}_0^\alpha\| > x) \\
&= \mathbb{P}(\|\mathbf{a}^{-1} \mathbf{X}_0^\alpha\| > x \mid \|\mathbf{X}_0^\alpha\| > xa_*) \mathbb{P}(\|\mathbf{X}_0^\alpha\| > xa_*) \\
&\sim \mathbb{P}\left(\frac{\|\mathbf{a}^{-1} \mathbf{X}_0^\alpha\|}{\|\mathbf{X}_0^\alpha\|} > \frac{x}{\|\mathbf{X}_0^\alpha\|} \mid \|\mathbf{X}_0^\alpha\| > xa_*\right) a_0 a_*^{-1} x^{-1} \\
&\sim \mathbb{P}\left(\left\|\mathbf{a}^{-1} \left(\frac{\mathbf{X}_0}{\|\mathbf{X}_0^\alpha\|^{1/\alpha}}\right)^\alpha\right\| > \frac{x}{\|\mathbf{X}_0^\alpha\|} \mid \|\mathbf{X}_0^\alpha\| > xa_*\right) a_0 a_*^{-1} x^{-1} \\
&\sim \mathbb{P}(\|\mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha\| > a_* Y^{-1}) a_0 a_* x^{-1} \\
&\sim \mathbb{E}[\|\mathbf{a}^{-1} \widetilde{\Theta}_0^\alpha\|] a_0 x^{-1}.
\end{aligned}$$

We conclude that  $\mathbf{X}_0$  is non-standard regularly varying and the angular measure is the limit, as  $x \rightarrow \infty$ , of the ratio

$$\begin{aligned}
\mathbb{P}(\|\widetilde{\mathbf{X}}_0\|^{-1} \widetilde{\mathbf{X}}_0 \in \cdot \mid \|\widetilde{\mathbf{X}}_0\| > x) &= \frac{\mathbb{P}(\|\widetilde{\mathbf{X}}_0\|^{-1} \widetilde{\mathbf{X}}_0 \in \cdot, \|\widetilde{\mathbf{X}}_0\| > x)}{\mathbb{P}(\|\widetilde{\mathbf{X}}_0\| > x)} \\
&= \frac{\mathbb{P}(\|\widetilde{\mathbf{X}}'_0\|^{-1} \widetilde{\mathbf{X}}'_0 \in \cdot, \|\widetilde{\mathbf{X}}'_0\| > x)}{\mathbb{P}(\|\widetilde{\mathbf{X}}'_0\| > x)} \\
&= \frac{\mathbb{P}(\|\widetilde{\mathbf{X}}'_0\|^{-1} \widetilde{\mathbf{X}}'_0 \in \cdot, \|\widetilde{\mathbf{X}}'_0\| > x \mid \|\mathbf{X}_0^\alpha\| > xa_*)}{\mathbb{P}(\|\widetilde{\mathbf{X}}'_0\| > x \mid \|\mathbf{X}_0^\alpha\| > xa_*)}
\end{aligned}$$

and the desired result follows by definition of  $\widetilde{\Theta}_0$ .  $\square$

## B Proof of Lemma 5.3

The fundamental ingredient in the proof is a large deviation result for  $T_u$  by [3] (see also [6]). It gives a very precise bound on the typical range of  $T_u$ , which allows us to deduce properties of the relevant random variables at time  $T_u$ , by replacing the random time by a deterministic bound.

**Step 1.** Fix  $\epsilon > 0$  and write  $C_u = C_u(\epsilon)$ . It is enough to show that  $\lim_{u \rightarrow \infty} \mathbb{P}(C_u^c \mid T_u < \infty) = 0$ . Indeed, we can sandwich the conditional probabilities as follows

$$\begin{aligned}
&\mathbb{P}(X_2 > u^{1/\alpha_2} \mid T_u < \infty) \\
&\geq \mathbb{P}(\{X_2 > u^{1/\alpha_2}\} \cap C_u \mid T_u < \infty) \\
&= \mathbb{P}(X_2 > u^{1/\alpha_2} \mid T_u < \infty) - \mathbb{P}(\{X_2 > u^{1/\alpha_2}\} \cap C_u^c \mid T_u < \infty) \\
&\geq \mathbb{P}(X_2 > u^{1/\alpha_2} \mid T_u < \infty) - \mathbb{P}(C_u^c \mid T_u < \infty).
\end{aligned}$$

Then the desired result follows by letting  $u \rightarrow \infty$ . We will consider each of the contributions to  $C_u^c$  separately:

$$\begin{aligned} C_u^c &= \left\{ T_u > L_u \right\} \cup \left\{ X_{T_u,1} > u^{(1+\epsilon)/\alpha_1} \right\} \cup \left\{ \max_{1 \leq k \leq L_u} \frac{Q_{k,2}}{Q_{k,1}} > u^{\epsilon/\alpha_1} \right\} \\ &\cup \left\{ \sum_{l=1}^{T_u} (W_{2,l} - W_{1,l}) - T_u(\mu_{2|1} - \mu_{1|1}) > \epsilon T_u \right\} \cup \left\{ \sum_{l=1}^{L_u} W_{l,1} > \frac{1+\epsilon}{\alpha_1} \log u \right\} \\ &= A \cup B \cup D \cup E \cup F. \end{aligned}$$

By (5.7), the required assertion  $\lim_{u \rightarrow \infty} P(B|T_u < \infty) = 0$  will as well follow from

$$\lim_{u \rightarrow \infty} u \cdot P(B \cap \{T_u < \infty\}) \leq \lim_{u \rightarrow \infty} u \cdot P(B) = 0.$$

**Step 2.** The negligibility of  $A$  is a direct consequence of [3, Lemma 4.3]) which provides that for a sufficiently large constant  $C$ ,

$$\lim_{u \rightarrow \infty} \mathbb{P} \left( \left| T_u - \frac{\log u}{\mu_1 \alpha_1} \right| \geq C f(u), \left| T_u < \infty \right. \right) = 0,$$

where  $f(u) = \sqrt{\log(u) \cdot \log(\log(u))}$ .

**Step 3.** Negligibility of  $B$  and  $F$ : Considering  $B$ , we have by (5.4) that  $\lim_{u \rightarrow \infty} u \mathbb{P}(X_1 > u^{(1+\epsilon)/\alpha_1}) = 0$  implying that

$$\lim_{u \rightarrow \infty} u \mathbb{P}(X_{T_u,1} > u^{(1+\epsilon)/\alpha_1}) = 0,$$

since  $X_1 = \sup_n X_{n,1}$ .

By the classical Cramér estimate for the random walk  $W_{n,1}$  (see [13, XII.(5.13)] it holds

$$\lim_{u \rightarrow \infty} u^{1/\alpha_1} \mathbb{P} \left( \sup_{n \in \mathbb{N}} \left( \sum_{l=1}^n W_{l,1} \right) > \log u \right) = c^* \in (0, \infty)$$

and hence in particular

$$\lim_{u \rightarrow \infty} u \mathbb{P} \left( \sum_{l=1}^{L_u} W_{l,1} > \frac{1+\epsilon}{\alpha_1} \log u \right) = 0.$$

**Step 4.** Now we turn to  $D$ . A union bound yields

$$\mathbb{P} \left( \max_{1 \leq k \leq L_u} \frac{Q_{k,2}}{Q_{k,1}} > u^{\epsilon/\alpha_1}, T_u < \infty \right) \leq \sum_{k=1}^{L_u} \mathbb{P} \left( u^{\epsilon/\alpha_1} Q_{k,1} < Q_{k,2}, T_u < \infty \right).$$



We decompose for any  $k \geq 0$

$$\begin{aligned}
& \mathbb{P}(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, T_u < \infty) \\
&= \mathbb{P}(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, X_1 > u^{1/\alpha_1}) \\
&\leq \mathbb{P}\left(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, \right. \\
&\quad \left. \sum_{j \neq k} \prod_{\ell=1}^{j-1} (b_1 + c_1 M_\ell) Q_{j,1} + \prod_{\ell=1}^{k-1} (b_1 + c_1 M_\ell) Q_{k,1} > u^{1/\alpha_1}\right).
\end{aligned}$$

We bound this probability by the sum of two terms

$$\begin{aligned}
& \mathbb{P}\left(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, \sum_{j \neq k} \prod_{\ell=1}^{j-1} (b_1 + c_1 M_\ell) Q_{j,1} > \frac{1}{2} u^{1/\alpha_1}\right) \\
&+ \mathbb{P}\left(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, \prod_{\ell=1}^{k-1} (b_1 + c_1 M_\ell) Q_{k,1} > \frac{1}{2} u^{1/\alpha_1}\right) \quad (\text{B.1})
\end{aligned}$$

and have to show that both contributions, when summed over  $k = 0, \dots, L_u$ , are of order  $o(u^{-1})$ .

We estimate the second term in (B.1) thanks to Markov's inequality of order  $\alpha_1/(1 + \varepsilon) < \kappa < \alpha_1$ :

$$\begin{aligned}
& \mathbb{P}\left(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, \prod_{\ell=1}^{k-1} (b_1 + c_1 M_\ell) Q_{k,1} > \frac{1}{2} u^{1/\alpha_1}\right) \\
&\leq \mathbb{P}\left(\prod_{\ell=1}^{k-1} (b_1 + c_1 M_\ell) Q_{k,2} > \frac{1}{2} u^{(1+\varepsilon)/\alpha_1}\right) \\
&\leq \frac{2^\kappa (\mathbb{E}[(b_1 + c_1 M)^\kappa])^k \mathbb{E}[\|\mathbf{Q}\|^\kappa]}{u^{\kappa((1+\varepsilon)/\alpha_1)}}.
\end{aligned}$$

As  $\alpha_1/(1 + \varepsilon\alpha_1) < \kappa < \alpha_1$  we have that  $q := \mathbb{E}[(b_1 + c_1 M)^\kappa] < 1$  and conclude

$$\begin{aligned}
& \sum_{k=0}^{\infty} \mathbb{P}\left(u^{\varepsilon/\alpha_1} Q_{k,1} < Q_{k,2}, \prod_{\ell=1}^{k-1} (b_1 + c_1 M_\ell) Q_{k,1} > \frac{1}{2} u^{1/\alpha_1}\right) \\
&\leq \frac{2^\kappa \mathbb{E}[\|\mathbf{Q}\|^\kappa]}{1 - q} \frac{1}{u^{\kappa((1+\varepsilon)/\alpha_1)}} = o(u^{-1}).
\end{aligned}$$

**Step 5.** Finally we turn to  $E$ . Note that  $E = \emptyset$  in Case I', since then

$$W_{l,2} - W_{l,1} = \log\left(\frac{c_2 M_l}{c_1 M_l}\right) = \log c_2 - \log c_1 = \mu_{2|1} - \mu_{1|1} \quad \text{a.s.}$$

Hence, we consider only Case II'. We have to prove the conditional large deviation result

$$\lim_{u \rightarrow \infty} \mathbb{P} \left( \underbrace{\sum_{k=1}^{T_u} (W_{k,2} - W_{k,1}) - T_u(\mu_{2|1} - \mu_{1|1})}_{=: D_u} > \epsilon T_u \mid T_u < \infty \right) = 0$$

Under  $\mathbb{P}^{\alpha_1}$  we have that  $S_n = \sum_{k=1}^n W_{k,1}$  constitutes a random walk with positive drift  $\mu_{1|1} = \mathbb{E}^{\alpha_1}[\log(b_1 + c_1 M)] > 0$ . Thus under the change of measure  $S_n \rightarrow \infty$ ,  $T_u < \infty$  a.s. Note here that  $T_u$  is not a stopping time for the random walk, but for the sequence  $X_{1,n}$ . However, divergence of  $S_n$  implies divergence of  $X_{1,n}$ , see e.g. [16, Theorem 2.1]. Hence we have the identity

$$\mathbb{P}(D_u \mid T_u < \infty) = \frac{\mathbb{E}^{\alpha_1}[e^{-\alpha_1 S_{T_u}} 1_{D_u}]}{\mathbb{P}(T_u < \infty)}.$$

Since  $u\mathbb{P}(T_u < \infty) \rightarrow c > 0$  as  $u \rightarrow \infty$ , it is enough to show that

$$\lim_{u \rightarrow \infty} u \mathbb{E}^{\alpha_1}[e^{-\alpha_1 S_{T_u}} 1_{D_u}] = 0.$$

We have

$$\begin{aligned} u \mathbb{E}^{\alpha_1}[e^{-\alpha_1 S_{T_u}} 1_{D_u}] &= \mathbb{E}^{\alpha_1} \left[ \left( \frac{u^{1/\alpha_1}}{X_{T_u,1}} \right)^{\alpha_1} \left( \frac{X_{T_u,1}}{e^{S_{T_u}}} \right)^{\alpha_1} 1_{D_u} \right] \\ &\leq \mathbb{E}^{\alpha_1} \left[ \left( \frac{X_{T_u,1}}{e^{S_{T_u}}} \right)^{\alpha_1} 1_{D_u} \right] \end{aligned}$$

by definition of  $T_u$  implying that  $X_{T_u,1} > u^{1/\alpha_1}$ . Then, using Chernoff's device, we achieve for any  $\lambda > 0$  the following upper-bound

$$\mathbb{E}^{\alpha_1} \left[ \left( \frac{X_{T_u,1}}{e^{S_{T_u}}} \right)^{\alpha_1} 1_{D_u} \right] \leq \mathbb{E}^{\alpha_1} \left[ \left( \frac{X_{T_u,1}}{e^{S_{T_u}}} \right)^{\alpha_1} e^{\lambda(\sum_{k=1}^{T_u} (W_{k,2} - W_{k,1}) - T_u(\mu_{2|1} - \mu_{1|1}) - \epsilon T_u)} \right].$$

We will show that

$$Z_n(\epsilon, \lambda) := \left( \frac{X_{n,1}}{e^{S_n}} \right)^{\alpha_1} \frac{e^{\lambda \sum_{k=1}^n (W_{k,2} - W_{k,1})}}{(e^{\lambda((\mu_2 - \mu_1) + \epsilon)})^n}$$

is an integrable stochastic process with  $\sup_{n \geq 1} \mathbb{E}^{\alpha_1}[Z_n(\epsilon, \lambda)] < \infty$  a.s. for any  $\epsilon > 0$  small by choosing  $\lambda > 0$  accordingly. Then  $\mathbb{E}^{\alpha_1}[Z_n(\epsilon', \lambda)] \rightarrow 0$  for any  $\epsilon' > \epsilon$  which is the desired result.

In order to show the uniform bound for  $\mathbb{E}^{\alpha_1}[Z_n]$  one has to introduce the sequence

$$V_n = \frac{X_{n,1}}{e^{S_n}} = \frac{\sum_{k=1}^n \prod_{l=1}^{k-1} A_l Q_{k,1}}{\prod_{l=1}^n A_l} = \frac{1}{A_n} V_{n-1} + \frac{Q_{n,1}}{A_n}$$

with  $A_n = b_1 + c_1 M_n$ . We abbreviate  $\Delta W_k := W_{k,2} - W_{k,1}$  and  $\Delta\mu := \mu_{2|1} - \mu_{1|1}$  and note that  $\mathbb{E}^{\alpha_1}(\Delta W_1) = \Delta\mu$  and that, due to the assumptions of (Case II')

$$\begin{aligned} d_* &:= \log(b_2) - \log(b_1) \leq \Delta W_1 = \log\left(\frac{b_2 + c_2 M_1}{b_1 + c_1 M_1}\right) \\ &\leq \log(c_2) - \log(c_1) =: d^*. \end{aligned}$$

We will use the recursive formula

$$\begin{aligned} Z_n(\epsilon, \lambda) &= V_n^{\alpha_1} \frac{e^{\lambda \sum_{k=1}^n \Delta W_k}}{(e^{\lambda \Delta\mu + \epsilon})^n} \\ &= \left( \frac{1}{A_n} V_{n-1} + \frac{Q_{n,1}}{A_n} \right)^{\alpha_1} \frac{e^{\lambda \sum_{k=1}^{n-1} \Delta W_k}}{(e^{\lambda \Delta\mu + \epsilon})^{n-1}} \cdot \frac{e^{\lambda \Delta W_n}}{e^{\lambda(\Delta\mu + \epsilon)}}. \end{aligned}$$

Then

$$\begin{aligned} &\mathbb{E}^{\alpha_1} [Z_n(\epsilon, \lambda) \mid \mathcal{F}_{n-1}] \\ &\leq \mathbb{E}^{\alpha_1} \left[ \frac{1}{A_n^{\alpha_1}} \frac{e^{\lambda(\Delta W_n)}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] Z_{n-1} + \mathbb{E}^{\alpha_1} \left[ \left( \frac{Q_{n,1}}{A_n} \right)^{\alpha_1} \frac{e^{\lambda \Delta W_n}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] \frac{e^{\lambda \sum_{k=1}^{n-1} \Delta W_k}}{(e^{\lambda \Delta\mu + \epsilon})^{n-1}} \end{aligned} \tag{B.2}$$

and we are going to prove that both the factors

$$\begin{aligned} \mathbf{m}(\lambda) &:= \mathbb{E}^{\alpha_1} \left[ \frac{1}{A_n^{\alpha_1}} \frac{e^{\lambda \Delta W_n}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] = \mathbb{E} \left[ \frac{e^{\lambda \Delta W_1}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] \\ \mathbf{c}(\lambda) &:= \mathbb{E}^{\alpha_1} \left[ \frac{e^{\lambda \Delta W_1}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] \end{aligned}$$

are less than one for suitably small  $\lambda$ . Upon taking  $\mathbb{P}^{\alpha_1}$ -expectations in (B.2), we infer

$$\mathbb{E}^{\alpha_1} [Z_n(\epsilon, \lambda)] \leq \mathbf{m}(\lambda) \mathbb{E}^{\alpha_1} [Z_{n-1}(\epsilon, \lambda)] + \mathbf{c}(\lambda)^{n-1} C_Q$$

with

$$C_Q = \mathbb{E}^{\alpha_1} \left[ \left( \frac{Q_{n,1}}{A_n} \right)^{\alpha_1} \frac{e^{\lambda \Delta W_n}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] = \mathbb{E} \left[ Q_{n,1}^{\alpha_1} \frac{e^{\lambda \Delta W_n}}{e^{\lambda(\Delta\mu + \epsilon)}} \right] \leq \mathbb{E} [Q_{n,1}^{\alpha_1}] \frac{e^{\lambda d^*}}{e^{\lambda(\Delta\mu + \epsilon)}} < \infty.$$

Using again the boundedness of  $\Delta W_1$ , an application of Hoeffding's lemma yields that

$$\mathbb{E}^{\alpha_1} [e^{\lambda \Delta W_1}] \leq \exp \left( \lambda(\Delta\mu) + \frac{\lambda^2}{8} (d^* - d_*) \right)$$

and hence

$$\mathbf{c}(\lambda) \leq \exp \left( \frac{\lambda^2}{8} (d^* - d_*) - \lambda\epsilon \right).$$

Thus  $\mathfrak{c}(\lambda) < 1$  for suitably small  $\lambda > 0$ .

Turning to  $\mathfrak{m}(\lambda)$ , we again use Hoeffding's lemma to get that

$$\mathfrak{c}(\lambda) \leq \exp \left( \lambda \mathbb{E}[\Delta W_1] + \frac{\lambda^2}{8} (d^* - d_*) - \lambda \mathbb{E}^{\alpha_1}[\Delta W_1] - \lambda \epsilon \right)$$

Hence, it suffices to show that difference of the expectations is nonpositive.

We write

$$\Delta W_1 = \log \left( \frac{b_2 + c_2 M_1}{b_1 + c_1 M_1} \right) = \log f(A_1)$$

for  $A_1 = b_1 + c_1 M_1$  and

$$f(a) = \frac{b_2 + c_2 \left( \frac{a - b_1}{c_1} \right)}{a} = -\frac{b_1}{a} \left( \frac{c_2}{c_1} - \frac{b_2}{b_1} \right) + \frac{c_2}{c_1}, \quad a \geq b_1.$$

The function  $f$  is increasing and  $\log f(a) \geq 0$  for  $a \geq b_1$ . Then

$$\begin{aligned} \mathbb{E}[\Delta W_1] - \mathbb{E}^{\alpha_1}[\Delta W_1] &= \mathbb{E} \left[ \log f(A_1) (1 - A_1^{\alpha_1}) \right] \\ &= \mathbb{E} \left[ \log f(A_1) (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A \leq 1\}} \right] + \mathbb{E} \left[ \log f(A_1) (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A > 1\}} \right] \end{aligned}$$

Recall that  $\mathbb{E}[A_1^{\alpha_1}] = 1$ , thus

$$\begin{aligned} -\mathbb{E} \left[ (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A > 1\}} \right] &= \mathbb{E} \left[ (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A \leq 1\}} \right] \\ -\mathbb{E} \left[ \log f(A_1) (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A > 1\}} \right] &\geq \mathbb{E} \left[ \log f(A_1) (1 - A_1^{\alpha_1}) \mathbf{1}_{\{A \leq 1\}} \right] \end{aligned}$$

where we have used that the function  $f$  is increasing and nonnegative.

## Acknowledgement

Olivier Wintenberger would like to acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR20-CE40-0025-01 (T-REX project). He also would like to thank Christian Francq and Jean-Michel Zakoïan for useful discussions on the topic.

Sebastian Mentemeier was partly supported by Deutscher Akademischer Austauschdienst (DAAD) under the grant no. 57388888.

## Data Availability Statement

Data sharing not applicable to this article as no datasets were generated during the current study.

## References

- [1] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Process. Appl.*, 119(4):1055–1080.
- [2] Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Rev. Econom. Statist.*, 72(3):498–505.
- [3] Buraczewski, D., Collamore, J. F., Damek, E., and Zienkiewicz, J. (2016). Large deviation estimates for exceedance times of perpetuity sequences and their dual processes. *Ann. Probab.*, 44(6):3688–3739.
- [4] Buraczewski, D., Damek, E., Guivarc’h, Y., Hulanicki, A., and Urban, R. (2009). Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. *Probab. Theory Relat. Fields*, 145(3-4):385–420
- [5] 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.
- [6] Buraczewski, D., Damek, E., and Zienkiewicz, J. (2018). Pointwise estimates for first passage times of perpetuity sequences. *Stochastic Process. Appl.*, 128(9):2923–2951.
- [7] Curtiss, J. H. (1941). On the distribution of the quotient of two chance variables. *Ann. Math. Statistics*, 12:409–421.
- [8] Embrechts, P., Klüppelberg, C., and Mikosch, T. (2013). *Modelling extremal events: for insurance and finance* Springer Science & Business Media.
- [9] Engle, R. F. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models. *Journal of Business and Economic Statistics*, 20(3), 339–350.
- [10] Damek, E. (2021). Diagonal stochastic recurrence equation–multivariate regular variation. *preprint arXiv:2106.11243*.
- [11] de Haan, L., and Resnick, S. I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. Verw. Gebiete*, 40(4), 317–337.
- [12] De Haan, L., Resnick, S., Rootzén, H., and de Vries, C. (1989). Extremal behavior of solutions to a stochastic difference equation, with applications to ARCH processes. *Stochastic Process. Appl.*, 32(2):213–224.

- [13] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. Vol. II. Second edition. Wiley, New York.
- [14] Ghalanos, A. (2019). *rmgarch: Multivariate GARCH models*. R package version 1.3-7.
- [15] Goldie, C. M. (1991). Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.*, 1(1):126–166.
- [16] Goldie, C. M. and Maller, R. A. (2000). Stability of perpetuities. *Ann. Prob.*, 28(3):1195–1218.
- [17] Grey, D. R. (1994). Regular variation in the tail behaviour of solutions of random difference equations. *Ann. Appl. Probab.*, 4(1):169–183.
- [18] Janssen, A. and Segers, J. (2014). Markov tail chains. *J. Appl. Probab.*, 51(4):1133–1153.
- [19] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.*, 131:207–248.
- [20] Letac, G. (1986). A contraction principle for certain markov chains and its applications. *Contemp. Math*, 50:263–273.
- [21] Matsui, M. and Pedersen, R. S. (2021). Characterization of the tail behavior of a class of BEKK processes: A stochastic recurrence equation approach. *Econometric Theory*, 1-34. doi:10.1017/S0266466620000584
- [22] Mentemeier, S. and Wintenberger, O. (2019). Asymptotic Independence ex machina – Extreme Value Theory for the Diagonal BEKK-ARCH(1) Model. *preprint arXiv:1907.10379*.
- [23] Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, 6:318–334.
- [24] Pedersen, R. S. and Wintenberger, O. (2018). On the tail behavior of a class of multivariate conditionally heteroskedastic processes. *Extremes*, 21:261–284.
- [25] Perfekt, R. (1997). Extreme value theory for a class of markov chains with values in  $\mathbb{R}^d$ . *Adv. in Appl. Probab.*, 29(1):138–164.
- [26] Resnick, S. I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media.
- [27] Stărică, C. (1999). Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance*, 6(5), 515–553.