Bernoulli **22**(3), 2016, 1448–1490
DOI: 10.3150/15-BEJ699

A stochastic volatility model with flexible extremal dependence structure

ANJA JANSSEN* and HOLGER DREES**

Department of Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany.
E-mail: *anja.janssen@math.uni-hamburg.de; **holger.drees@math.uni-hamburg.de.

Stochastic volatility processes with heavy-tailed innovations are a well-known model for financial time series. In these models, the extremes of the log returns are mainly driven by the extremes of the i.i.d. innovation sequence which leads to a very strong form of asymptotic independence, that is, the coefficient of tail dependence is equal to $1/2$ for all positive lags. We propose an alternative class of stochastic volatility models with heavy-tailed volatilities and examine their extreme value behavior. In particular, it is shown that, while lagged extreme observations are typically asymptotically independent, their coefficient of tail dependence can take on any value between $1/2$ (corresponding to exact independence) and 1 (related to asymptotic dependence). Hence, this class allows for a much more flexible extremal dependence between consecutive observations than classical SV models and can thus describe the observed clustering of financial returns more realistically.

The extremal dependence structure of lagged observations is analyzed in the framework of regular variation on the cone $(0,\infty)^d$. As two auxiliary results which are of interest on their own we derive a new Breiman-type theorem about regular variation on $(0,\infty)^d$ for products of a random matrix and a regularly varying random vector and a statement about the joint extremal behavior of products of i.i.d. regularly varying random variables.

Keywords: asymptotic independence; Breiman's lemma; coefficient of tail dependence; extremal dependence; financial time series; hidden regular variation; power products; regular variation on cones; stochastic volatility time series

1. Introduction

1.1. Extremal behavior of models for financial time series

Univariate time series of (log-)returns are usually described by multiplicative models of the form

$$X_t = \sigma_t \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{1.1}$$

where ε_t, $t \in \mathbb{Z}$, are i.i.d. innovations and $(\sigma_t)_{t\in\mathbb{Z}}$ is a non-negative stationary time series of so-called volatilities. The two most popular classes of multiplicative models vary in the way the volatilities are modeled. While σ_t is a function of past innovations ε_s, $s < t$, in GARCH-type models, stochastic volatility models (SV models, for short) assume in contrast that the volatilities are driven by a second time series $(\eta_t)_{t\in\mathbb{Z}}$ of innovations. More precisely, it is often assumed that

1350-7265

the log-volatilities are described by a Gaussian linear time series of the type

$$\log \sigma_t = \sum_{i=0}^{\infty} \alpha_i \eta_{t-i}, \qquad t \in \mathbb{Z}, \tag{1.2}$$

with i.i.d. normal innovations η_t, $t \in \mathbb{Z}$, independent of $(\varepsilon_t)_{t\in\mathbb{Z}}$ (although, sometimes, ε_t and η_{t+1} are assumed to be correlated to capture a leverage effect). Because returns are usually heavy-tailed and the volatilities are lognormal, in this modeling approach the innovations ε_t are often assumed to be regularly varying, that is,

$$\frac{P(\varepsilon_t > sx)}{P(|\varepsilon_t| > x)} \to ps^{-\alpha}, \qquad \frac{P(\varepsilon_t < -sx)}{P(|\varepsilon_t| > x)} \to (1-p)s^{-\alpha}, \qquad x \to \infty, \tag{1.3}$$

for all $s > 0$, some $\alpha > 0$ and $p \in [0, 1]$. By Breiman's lemma [see Breiman [6]] this implies that X_t is regularly varying as well.

While the (univariate) tails of X_t behave similarly in these SV and GARCH-type models, the extreme value dependence of consecutive returns differs between the two model classes. This can firstly be seen from the conditional probabilities of an extreme event at lag h, given an extreme event at time 0. For GARCH-type models,

$$\lim_{x\to\infty} P(X_h > x | X_0 > x) > 0, \qquad h \neq 0, \tag{1.4}$$

while for SV models as the one above

$$\lim_{x\to\infty} P(X_h > x | X_0 > x) = 0, \qquad h \neq 0 \tag{1.5}$$

[cf. Basrak, Davis and Mikosch [3] and Davis and Mikosch [10]]. We call the vector (X_0, X_h) asymptotically dependent in the first case and asymptotically independent in the second case.

The differences between (1.4) and (1.5) are mirrored by a different *asymptotic* cluster behavior of the related models. In case of asymptotic dependence exceedances over extreme thresholds typically occur in clusters, while in time series with asymptotically independent consecutive observations exceedances tend to be separated in time as the threshold increases. It is important to note, though, that for realistic sample sizes it will often be difficult to spot the difference with the naked eye, because sufficiently high thresholds are not exceeded. Figure 1 displays realizations of an SV time series [with AR(1) log-volatilities used in (1.2)] on the left-hand side and of a GARCH(1, 1) time series on the right-hand side. The parameters were obtained by fitting these models to a stretch of returns from the S&P 500 stock index. Although the left plot shows a realization from a model with asymptotically independent consecutive observations, it nevertheless exhibits a strong clustering of extremes, because the asymptotic independence becomes obvious only for exceedances over higher thresholds. Hence, the well-documented fact that large losses (and gains) often occur in clusters does not rule out SV models for the description of the extreme value dependence between consecutive returns. Indeed, the analysis of the extreme value dependence between consecutive returns from UBS stocks and Google stocks in Drees *et al.* [13] indicates that, despite the obvious clusters of large losses and gains, it is appropriate to assume asymptotic independence between observations of the same sign.

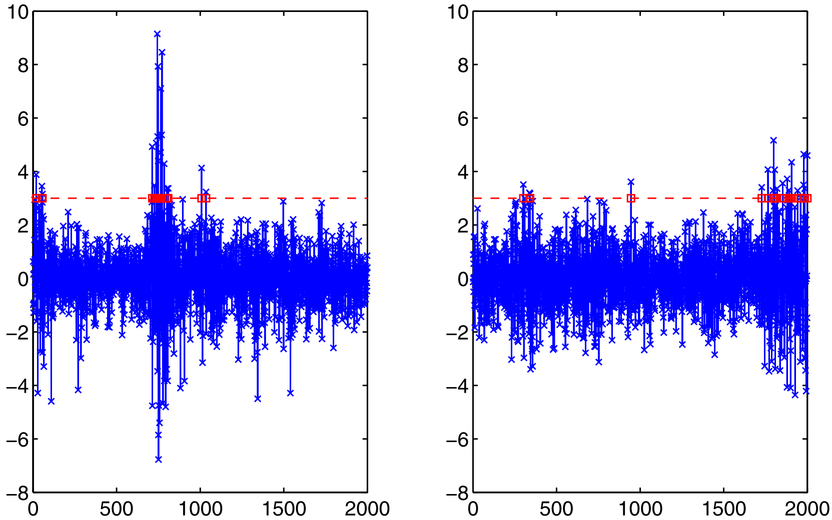

Figure 1. Large exceedances for 2000 consecutive observations of a simulated SV model with heavy-tailed (Student t) innovations (left) and a GARCH(1, 1) process with normal innovations (right). The model parameters for both models are estimated from the same data set, consisting of 2432 daily log returns (given in percent) of the S&P 500 stock index from 1999 to 2008, cf. Abanto-Valle *et al.* [1] for the SV parameters. The threshold 3 is approximately equal to the 99% quantile of the marginal distribution in both models.

In case of asymptotic independence, the strength of clustering of exceedances over high, but not extreme thresholds will depend on the rate at which

$$P(X_h > x | X_0 > x) = \frac{P(X_0 > x, X_h > x)}{P(X_0 > x)}, \qquad h \neq 0,$$

tend to 0. In the present paper, we thus focus on the asymptotics for the probabilities of joint exceedances $P(X_h > x, X_0 > x), h \neq 0$, as $x \to \infty$. Note that established models for financial time series are quite limited in their ability to reflect different behaviors of these probabilities. In particular, GARCH models imply that

$$\lim_{x\to\infty} \frac{P(X_0 > x, X_h > x)}{P(X_0 > x)} \in (0, \infty), \qquad h \neq 0,$$

while classical SV models lead to

$$\lim_{x\to\infty} \frac{P(X_0 > x, X_h > x)}{(P(X_0 > x))^2} \in (0, \infty), \qquad h \neq 0.$$

However, applying the methodology developed by Draisma *et al.* [12] to consecutive observations in time series of returns shows that often $P(X_0 > x, X_h > x)$ is apparently of the order $(P(X_0 > x))^{1/\eta_h}$ for some η_h strictly between $1/2$ and 1, indicating *asymptotic* independence, but a stronger dependence between exceedances over high thresholds than implied by classical SV models. We will therefore introduce an SV model which allows for a much more flexible modeling of the asymptotics of joint exceedance probabilities.

1.2. Regular variation on cones

In order to describe the extremal dependence structure of a time series, we will use the concept of regular variation of random vectors. An $\mathbb{R}^d$-valued random vector $\mathbf{Z}$ is said to be regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ if there exists a measure $\tilde{\nu} \neq 0$ on $\mathcal{B}([0,\infty)^d \setminus \{\mathbf{0}\})$ which is finite on $[0,\infty)^d \setminus [0,x]^d$ for all $x > 0$ such that

$$\frac{P(\mathbf{Z} \in xB)}{P(\|\mathbf{Z}\| > x)} \to \tilde{\nu}(B), \qquad x \to \infty, \tag{1.6}$$

for all Borel sets $B \subset [0,\infty)^d$ with $\tilde{\nu}(\partial B) = 0$ that are bounded away from the origin $\mathbf{0}$. (Here ∂B denotes the topological boundary of B and $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^d$.) By this notion, we follow the definition of $\mathbb{M}_0$-convergence of Hult and Lindskog [15] which can be shown to be equivalent to vague convergence on $[0,\infty]^d \setminus \{\mathbf{0}\}$. See, for example, Resnick [25], Section 6.1, for details on the definition of multivariate regular variation using vague convergence. The limit measure $\tilde{\nu}$ is necessarily homogeneous of some order $-\tilde{\alpha} < 0$, which is called the index of regular variation.

For GARCH time series, Basrak, Davis and Mikosch [3] show that under general conditions the vectors $(\sigma_{t_1},\dots,\sigma_{t_d})$ and $(X_{t_1},\dots,X_{t_d})$ are multivariate regularly varying for all $t_1 < \cdots < t_d$, and the same holds true for SV models with heavy-tailed innovations ε_t. However, while in the former case the limiting measure $\tilde{\nu}$ puts mass on $(0,\infty)^d$, it is concentrated on the axes for SV models (cf. Davis and Mikosch [9]). This asymptotic independence of lagged returns (and volatilities) renders convergence (1.6) rather uninformative. In particular, we can merely conclude that the probability $P(X_0 > x, X_h > x)$ of joint exceedances is of smaller order than $P(X_0 > x)$ for all $h \neq 0$, but we neither obtain its rate of convergence to 0, nor whether the probability can be standardized in a different way than in (1.6) to obtain a non-trivial limit. Therefore, in the case of asymptotic independence there is need for a refined analysis of the second-order extremal dependence behavior.

This can be most elegantly done in the framework of regular variation on the cone $\mathbb{E}^d := (0,\infty)^d$. In what follows, we use the abbreviation $\min(\mathbf{z}) := \min\{z_1,\dots,z_d\}$ for $\mathbf{z} = (z_1,\dots,z_d) \in \mathbb{R}^d$. Now, instead of (1.6), we assume that there exists a measure $\nu \neq 0$ on $\mathcal{B}(\mathbb{E}^d)$ which is finite on $(x,\infty)^d$ for all $x > 0$ such that

$$\frac{P(\mathbf{Z} \in xB)}{P(\min(\mathbf{Z}) > x)} \to \nu(B), \qquad x \to \infty, \tag{1.7}$$

for all Borel sets $B \subset \mathbb{E}^d$ with $\nu(\partial B) = 0$ that are bounded away from the topological boundary

$$\mathcal{O}^d := \partial(\mathbb{E}^d) = \{\mathbf{x} \in [0,\infty)^d \mid \min(\mathbf{x}) = 0\}$$

of the cone $\mathbb{E}^d$. Notice that here we consider events in which all components of $\mathbf{Z}$ are large, whereas in (1.6) just one coordinate needs to be extreme. Again ν is homogeneous of some order $-\alpha < 0$, which is called the index of regular variation on $\mathbb{E}^d$. If both (1.6) and (1.7) hold, then $\alpha \geq \tilde{\alpha}$. In the case of asymptotic independence (i.e., $\tilde{\nu}$ is concentrated on $[0, \infty)^d \setminus (0, \infty)^d$), the random vector $\mathbf{Z}$ is then said to exhibit hidden regular variation. See Resnick [24,26] and [25], Section 9.4, Das, Mitra and Resnick [7] and Lindskog, Resnick and Roy [21] for further details about hidden regular variation and regular variation on cones.

Kulik and Soulier [16] show that for SV models with heavy-tailed innovations the vector $(X_{t_1}, \ldots, X_{t_d})$ of lagged returns is regularly varying on $(0, \infty)^d$ and that the limit measure ν is the same as if the components of this vector were exactly independent. A similar result holds for the absolute returns. In particular, $P(X_0 > x, X_h > x)$ is of the same order as $(P(X_0 > x))^2$, which means that the coefficient of tail dependence introduced in Ledford and Tawn [19] and Ledford and Tawn [20] equals $\eta_h = 1/2$ for all lags $h > 0$. Recall that a bivariate random vector (Y_0, Y_1) with equal marginal quantile function $F^{\leftarrow}$ has a coefficient of tail dependence η if $x \mapsto P(Y_0 > F^{\leftarrow}(1-1/x), Y_1 > F^{\leftarrow}(1-1/x))$ is a regularly varying function with index $-1/\eta$. Thus, for independent Y_0, Y_1 the value of η is equal to $1/2$ and the higher the value of η, the stronger is the extremal dependence of Y_0 and Y_1.

Hence, the classical SV time series as described above show a very weak extremal dependence, which is barely influenced by the parameters of the model. It is the main aim of the present paper to propose and analyze a modified class of SV models which allows for a much more flexible and realistic extremal dependence than the classical version.

1.3. Outline

Our following analysis of different multiplicative models heavily relies on the fact that a product of two independent factors inherits both its tail behavior and extremal dependence from the factor with the heavier tail. This general heuristic principle was formalized by Breiman [6] for univariate random variables. A similar "Breiman-type" result on the first order dependence behavior for a product of a random matrix $\mathbf{A}$ and a random vector $\mathbf{Z}$ was proved by Basrak, Davis and Mikosch [3], who analyzed regular variation on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ and on $[0, \infty)^d \setminus \{\mathbf{0}\}$. In Section 2, we establish an analogous result for regular variation on $\mathbb{E}^d = (0, \infty)^d$, which is somewhat more involved, because one has to keep in mind that (1.7) only describes the asymptotic behavior of $\mathbf{Z}/x$ on sets B that are bounded away from the boundary $\mathcal{O}^d$ of $\mathbb{E}^d$, and this feature is not necessarily preserved under multiplication with a general matrix $\mathbf{A}$.

Our aim is to apply this Breiman result to SV models with a heavy-tailed volatility sequence and light-tailed innovations in order to show that for these models the second-order behavior of extremes is mainly determined by the volatility sequence. In Section 3, we introduce a particular class of such models with Gamma-type log-volatilities and derive their marginal tail behavior using results from Rootzén [27]. Moreover, the first-order extremal behavior is analyzed in terms of point processes of exceedances. Similar SV models which allow for both asymptotic dependence and asymptotic independence have been proposed in Mikosch and Rezapour [22], but their analysis is restricted to the in this case rather uninformative first-order extremal dependence of those models, while our focus is on the refined second-order behavior in the asymptotically independent case.

In our SV models, the volatilities are given as (generally infinite) products of powers of regularly varying i.i.d. random variables. Section 4 deals with the asymptotic behavior of joint exceedance probabilities of such products, which turns out to be intimately related to the solution of certain linear optimization problems. While the heuristics for this connection can easily be understood, the exact arguments are more delicate. We give a full result in the case of two products and the proof of a special case together with an outlook for the general case of more than two products. In Section 5, we discuss the consequences for our SV models with Gamma-type log-volatility. In particular, we show that for any finite number of given coefficients of tail dependence $\eta_h \in [1/2, 1]$ for the pairs (X_0, X_h), $1 \le h \le m$, one can find an SV model of the new type with exactly these characteristics. This result underpins the high flexibility of our approach to modeling the extremal dependence of consecutive returns. Most proofs are postponed to Section 6.

Throughout the paper, we use the following notation. We write $\mathcal{B}(M)$ for the Borel σ-algebra on the metric space M. Weak convergence is denoted by $\stackrel{w}{\to}$. The expression δ_x stands for the Dirac measure at x. Bold expressions like $\mathbf{0}, \mathbf{1}$, etc. stand for a vector of suitable dimension consisting of 0's, 1's, and so on. This convention extends to sets, for example, for a vector $\mathbf{x} = (x_1, \dots, x_d)$ we have $(\mathbf{x}, \infty) = \times_{i=1}^{d} (x_i, \infty)$. For a vector $\mathbf{x}$ or a matrix $\mathbf{A}$ we denote its transposed with $\mathbf{x}'$ and $\mathbf{A}'$, respectively. We denote the positive and negative part of $x \in \mathbb{R}$ by $x^+ := \max\{x, 0\}$ and $x^- := -\min\{x, 0\}$, respectively. The complement of a set A is denoted by A^c. The empty product, $\prod_{i\in\varnothing} X_i$, is by convention equal to 1. Finally, $f(x) \sim g(x)$ means that $\lim_{x\to\infty} f(x)/g(x) = 1$.

2. A Breiman-type result for regular variation on $(0, \infty)^d$

As explained in the Introduction, we will analyze the extremal dependence of the volatilities $(\sigma_{t_1}, \dots, \sigma_{t_d})$ and of the returns $(X_{t_1}, \dots, X_{t_d})$ of multiplicative time series as in (1.1) using the notion of regular variation on the cones $[0, \infty)^d \setminus \{\mathbf{0}\}$ and $\mathbb{E}^d = (0, \infty)^d$. Although we are mainly interested in the case $d = 2$, for the time being we allow for an arbitrary $d \in \mathbb{N}$.

Since in SV models we have

$$\begin{pmatrix} X_{t_1} \\ \vdots \\ X_{t_d} \end{pmatrix} = \begin{pmatrix} \sigma_{t_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{t_d} \end{pmatrix} \begin{pmatrix} \varepsilon_{t_1} \\ \vdots \\ \varepsilon_{t_d} \end{pmatrix} = \begin{pmatrix} \varepsilon_{t_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varepsilon_{t_d} \end{pmatrix} \begin{pmatrix} \sigma_{t_1} \\ \vdots \\ \sigma_{t_d} \end{pmatrix}, \tag{2.1}$$

multivariate Breiman-type results are very useful to establish regular variation of $(X_{t_1}, \dots, X_{t_d})$. For regular variation on $[0, \infty)^d \setminus \{\mathbf{0}\}$, such a result has been stated in Basrak, Davis and Mikosch [3], Proposition A.1. If $\mathbf{X}$ is regularly varying on $[0, \infty)^d \setminus \{\mathbf{0}\}$ with index $-\tilde{\alpha} < 0$ and $\mathbf{A}$ is a random matrix independent of $\mathbf{X}$ such that $E(\|\mathbf{A}\|_{\text{op}}^{\tilde{\alpha}+\varepsilon}) < \infty$ for some $\varepsilon > 0$, then $\mathbf{AX}$ is regularly varying on $[0, \infty)^d \setminus \{\mathbf{0}\}$ as well with the same index of regular variation. Here, $\|\cdot\|_{\text{op}}$ denotes the operator norm for matrices, which is defined by

$$\|\mathbf{A}\|_{\text{op}} = \sup_{\mathbf{x}\in\mathbb{R}^d : \|\mathbf{x}\|=1} \|\mathbf{Ax}\| = \sup_{\mathbf{x}\in\mathbb{R}^d : \|\mathbf{x}\|=1} d\big(\mathbf{Ax}, \{\mathbf{0}\}\big), \tag{2.2}$$

where $d(\mathbf{x}, B) := \inf_{\mathbf{y}\in B} \|\mathbf{x}-\mathbf{y}\|$ for $\mathbf{x} \in \mathbb{R}^d$, $B \subset \mathbb{R}^d$ denotes the usual distance function induced by the Euclidean norm on $\mathbb{R}^d$. More precisely, if $\mathbf{Z} = \mathbf{X}$ satisfies (1.6), then

$$\frac{P(\mathbf{AX} \in xB)}{P(\|\mathbf{X}\| > x)} \to E\big(\tilde{\nu}\big(\mathbf{A}^{-1}(B)\big)\big), \qquad x \to \infty,$$

for all $B \in \mathcal{B}([0,\infty)^d)$ bounded away from $\mathbf{0}$ with $E(\tilde{\nu}(\mathbf{A}^{-1}(\partial B))) = 0$. Therefore, if the vector $(\varepsilon_{t_1}, \dots, \varepsilon_{t_d})$ of i.i.d. innovations is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ with index $-\tilde{\alpha}$ and $E(\sigma_0^{\tilde{\alpha}+\varepsilon}) < \infty$, then $(X_{t_1}, \dots, X_{t_d})$ is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ as well. Likewise, one may draw this conclusion if the vector $(\sigma_{t_1}, \dots, \sigma_{t_d})$ is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ with index $-\tilde{\alpha}$ and $E(|\varepsilon_0|^{\tilde{\alpha}+\varepsilon}) < \infty$ for some $\varepsilon > 0$.

In the following, we will derive an analogous result for regular variation on the smaller cone $\mathbb{E}^d$. For the case of an SV model with heavy-tailed i.i.d. innovations which are regularly varying with index $-\alpha$ and light-tailed volatilities which satisfy $E(\sigma_0^{2d\alpha+\varepsilon}) < \infty$ for some $\varepsilon > 0$, Kulik and Soulier [16] have shown that the vector $(X_{t_1}, \dots, X_{t_d})$, $t_1 < \cdots < t_d$, shows regular variation on $\mathbb{E}^d$ with limiting measure

$$\nu^X\left(\bigtimes_{i=1}^d (s_i, \infty)\right) = \frac{E(\nu^\varepsilon((\mathbf{\Delta}(\sigma_{t_1}, \dots, \sigma_{t_d}))^{-1}(\bigtimes_{i=1}^d (s_i,\infty))))}{E(\nu^\varepsilon((\mathbf{\Delta}(\sigma_{t_1}, \dots, \sigma_{t_d}))^{-1}(\bigtimes_{i=1}^d (1,\infty))))} = \prod_{i=1}^d s_i^{-\alpha}, \tag{2.3}$$

for all $s_i > 0$, $1 \le i \le d$, where ν^ε with $\nu^\varepsilon(\bigtimes_{i=1}^d (s_i,\infty)) = \prod_{i=1}^d s_i^{-\alpha}$ denotes the limiting measure of $(\varepsilon_{t_1}, \dots, \varepsilon_{t_d})$ on $\mathbb{E}^d$ and $\mathbf{\Delta}(a_1, \dots, a_d)$ stands for a diagonal matrix with diagonal elements $a_1, \dots, a_d$.

We want to derive a general Breiman-type result for regular variation on the cone $\mathbb{E}^d$ for random products $\mathbf{AX}$. Because convergence (1.7) describes the behavior of random vectors exclusively on $\mathbb{E}^d$, we have to ensure that the random pre-image $\mathbf{A}^{-1}(B)$ of a set $B \subset \mathbb{E}^d$ is again a subset of $\mathbb{E}^d$. Moreover, as the limit measure ν in (1.7) may be infinite in a neighborhood of the boundary $\mathcal{O}^d$ of $\mathbb{E}^d$, for our Breiman-type result we must control the distance between $\mathbf{A}^{-1}(B)$ and $\mathcal{O}^d$.

To this end, set $\mathbb{F}^d := \mathbb{R}^d \setminus \mathbb{E}^d$ and introduce the notation

$$\tau(\mathbf{x}) := d\big(\mathbf{x}, \mathbb{F}^d\big) = \min\big(\mathbf{x}^+\big) \qquad \text{for } \mathbf{x} \in \mathbb{R}^d,$$

where $(x_1, \dots, x_d)^+ = (x_1^+, \dots, x_d^+)$. Furthermore, let

$$\mathcal{S}^d := \big\{\mathbf{x} \in \mathbb{E}^d \,|\, \tau(\mathbf{x}) = 1\big\}.$$

Observe that in the definition of regular variation on $\mathbb{E}^d$ the set $\mathcal{O}^d$ takes over the role which is played by the origin $\mathbf{0} \in \mathbb{R}^d$ in the definition of regular variation on $[0,\infty)^d \setminus \{\mathbf{0}\}$. Since $\tau(\mathbf{x}) = d(\mathbf{x}, \mathcal{O}^d)$ for all $\mathbf{x} \in \mathbb{E}^d$, one may consider $\mathcal{S}^d$ an analog to the unit sphere $\{\mathbf{x} \in \mathbb{R}^d \,|\, \|\mathbf{x}\| = d(\mathbf{x}, \mathbf{0}) = 1\}$ in the present setting. Cf. Das, Mitra and Resnick [7] for the use of $\mathcal{S}^d$ in the context of regular variation on cones.

Extend the definition of τ to matrices in analogy to (2.2) by setting

$$\tau(\mathbf{A}) := \tau_{\text{op}}(\mathbf{A}) := \sup_{\mathbf{x}\in\mathcal{S}^d} \tau(\mathbf{A}\mathbf{x}) \qquad \text{for } \mathbf{A} \in \mathbb{R}^{d\times d}, \tag{2.4}$$

with $\tau_{\text{op}}(\mathbf{A}) \in [0,\infty]$. For brevity, we will write $\tau(\mathbf{A})$ instead of $\tau_{\text{op}}(\mathbf{A})$ whenever it is clear that $\mathbf{A}$ denotes a matrix. It will turn out that one may prove a Breiman-type result for the regular variation on $\mathbb{E}^d$ if one replaces the moment condition $E(\|\mathbf{A}\|_{\text{op}}^{\tilde{\alpha}+\varepsilon}) < \infty$ used in Basrak, Davis and Mikosch [3] with the corresponding condition on $\tau(\mathbf{A})$. The next lemma provides some help for the interpretation of τ.

Lemma 2.1. *For a matrix $\mathbf{A} \in \mathbb{R}^{d\times d}$ the following three properties are equivalent*:

(i) $\mathbf{A}^{-1}(\mathbb{E}^d) \subset \mathbb{E}^d$;
(ii) $0 < \tau(\mathbf{A}) < \infty$;
(iii) $\mathbf{A}$ *is invertible and all entries of* $\mathbf{A}^{-1}$ *are non-negative.*

Furthermore, if $\mathbf{A}$ *possesses any of these properties, then* $\mathbf{A}^{-1}([0,\infty)^d) \subset [0,\infty)^d$.

For a d-dimensional diagonal matrix $\mathbf{\Delta}(\delta_1,\dots,\delta_d)$ with positive diagonal elements the value of τ may be directly derived from (2.4) as

$$\tau(\mathbf{\Delta}) = \max\{\delta_1,\dots,\delta_d\}.$$

For a general random matrix $\mathbf{A}$, the concrete value of $\tau(\mathbf{A}) \in (0,\infty)$ can be derived from the inverse matrix $\mathbf{A}^{-1}$.

Lemma 2.2. *Let* $\mathbf{A} \in \mathbb{R}^{d\times d}$ *satisfy* $0 < \tau(\mathbf{A}) < \infty$. *Then*

$$\tau(\mathbf{A}) = \frac{1}{\min(\mathbf{A}^{-1}\mathbf{1})}. \tag{2.5}$$

We are now ready to state our Breiman-type result for vectors that show regular variation on the cone $\mathbb{E}^d$.

Theorem 2.3. *Let* $\mathbf{Z} \in \mathbb{R}^d$ *be regularly varying on the cone* $\mathbb{E}^d$ *with index* $-\alpha$ *for some* $\alpha > 0$ *and limit measure* ν. *Moreover, let* $\mathbf{A} \in \mathbb{R}^{d\times d}$ *be a random matrix independent of* $\mathbf{Z}$ *which satisfies* $\tau(\mathbf{A}) > 0$ *almost surely and*

$$E\big(\tau(\mathbf{A})^{\alpha+\delta}\big) < \infty \tag{2.6}$$

for some $\delta > 0$. *Then*

$$\lim_{x\to\infty} \frac{P(\mathbf{A}\mathbf{Z} \in xB)}{P(\min(\mathbf{Z}) > x)} = E\big(\nu\big(\mathbf{A}^{-1}B\big)\big) < \infty \tag{2.7}$$

for all Borel sets $B \subset \mathbb{E}^d$ *bounded away from* $\mathcal{O}^d$ *with* $E(\nu(\partial(\mathbf{A}^{-1}B))) = 0$. *In particular,* $\mathbf{A}\mathbf{Z}$ *is regularly varying on* $\mathbb{E}^d$ *with index* $-\alpha$ *and limiting measure*

$$B \mapsto \frac{E(\nu(\mathbf{A}^{-1}B))}{E(\nu(\mathbf{A}^{-1}((1,\infty)^d)))}.$$

The above theorem is useful to derive the extremal behavior of stochastic volatility models.

Example 2.4. Let $(X_{1,t}, X_{2,t})_{t\in\mathbb{Z}}$ be a two-dimensional heavy-tailed stochastic volatility time series with negative dependence which satisfies

$$
\begin{aligned}
(X_{1,t}, X_{2,t})' &= \mathbf{L}_t \mathbf{G}_t (\varepsilon_{1,t}, \varepsilon_{2,t})', \qquad t \in \mathbb{Z},\\
\mathbf{L}_t &= \begin{pmatrix} 1 & 0 \\ -\exp(q_t/2) & 1 \end{pmatrix}, \qquad \mathbf{G}_t = \begin{pmatrix} \exp(h_{1,t}/2) & 0 \\ 0 & \exp(h_{2,t}/2) \end{pmatrix}, \qquad t \in \mathbb{Z},\\
q_{t+1} &= \alpha + \beta q_t + u_t, \qquad t \in \mathbb{Z}\\
h_{i,t+1} &= \mu_i + \phi_i h_{i,t} + \eta_{i,t}, \qquad i = 1,2, t \in \mathbb{Z},
\end{aligned}
$$

with $\beta, \phi_i \in (-1, 1)$ and i.i.d. $\varepsilon_{i,t}$ which have a Student-t-distribution with α degrees of freedom, $\eta_{i,t} \sim \mathcal{N}(0, \sigma^2_{\eta,i})$, $u_t \sim \mathcal{N}(0, \sigma^2_u)$, all mutually independent. Cf. Tsay [30], Section 9.6, and Asai, McAleer and Yu [2], Section 2.5.2, for similar models. Define independent random variables

$$
H_i \sim \mathcal{N}\bigl(\mu_i, \sigma^2_{\eta,i}/\bigl(1-\phi_i^2\bigr)\bigr), \qquad i = 1,2, \qquad Q \sim \mathcal{N}\bigl(\alpha, \sigma^2_u/\bigl(1-\beta^2\bigr)\bigr).
$$

The stationary distribution of $\mathbf{L}_t\mathbf{G}_t$ is equal to the distribution of $\mathbf{A}$ with

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} \exp(H_1/2) & 0 \\ -\exp\bigl((Q+H_1)/2\bigr) & \exp(H_2/2) \end{pmatrix},\\
\mathbf{A}^{-1} &= \begin{pmatrix} \exp(-H_1/2) & 0 \\ \exp\bigl((Q-H_2)/2\bigr) & \exp(-H_2/2) \end{pmatrix}.
\end{aligned}
$$

Since all entries of $\mathbf{A}^{-1}$ are a.s. non-negative, we may apply Theorem 2.3 to see that the stationary distribution of $(X_{1,t}, X_{2,t})$ is regularly varying on $\mathbb{E}^2$ with index -2α and limit measure

$$
\begin{aligned}
\nu_2^X\left(\bigtimes_{i=1}^{2} (s_i, \infty)\right) &= \frac{E((\exp(-H_1/2)s_1)^{-\alpha}(\exp((Q-H_2)/2)s_1 + \exp(-H_2/2)s_2)^{-\alpha})}{E((\exp(-H_1/2))^{-\alpha}(\exp((Q-H_2)/2) + \exp(-H_2/2))^{-\alpha})}\\
&= s_1^{-\alpha}\frac{E((\exp((Q-H_2)/2)s_1 + \exp(-H_2/2)s_2)^{-\alpha})}{E((\exp((Q-H_2)/2) + \exp(-H_2/2))^{-\alpha})}
\end{aligned}
$$

for $s_1, s_2 > 0$.

In the following, we are mainly interested in the behavior of lagged observations of univariate SV models. That all SV models with heavy-tailed i.i.d. innovations and light-tailed volatilities have the same limit measure (2.3) has been shown in Kulik and Soulier [16] and also follows from Theorem 2.3 (even under the weaker assumption that $E(\sigma_0^{d\alpha+\varepsilon}) < \infty$ instead of $E(\sigma_0^{2d\alpha+\varepsilon}) < \infty$

for some $\varepsilon > 0$). The following lemma shows that a much richer second-order structure evolves from SV models with heavy-tailed volatilities and light-tailed innovations.

Corollary 2.5. *Let $X_t = \sigma_t \varepsilon_t$, $t \in \mathbb{Z}$, where $(\sigma_t)_{t\in\mathbb{Z}}$ and $(\varepsilon_t)_{t\in\mathbb{Z}}$ are independent stationary time series with ε_t, $t \in \mathbb{Z}$, i.i.d. and $\sigma_t > 0$. Assume furthermore that $(\sigma_{t_1}, \dots, \sigma_{t_d})$ is regularly varying on $\mathbb{E}^d$ with index $-\alpha_d < 0$ and limit measure ν_d^σ. Let, in addition, $0 < E((\varepsilon_0^+)^{\alpha_d+\delta}) < \infty$ for some $\delta > 0$. Then $(X_{t_1}, \dots, X_{t_d})$ is regularly varying on $\mathbb{E}^d$ as well, with the same index $-\alpha_d$ and limit measure ν_d^X defined by*

$$\nu_d^X\left(\bigtimes_{i=1}^d (s_i, \infty)\right) = \frac{E(\nu_d^\sigma(\bigtimes_{i=1}^d (s_i(\varepsilon_{t_i}^+)^{-1}, \infty)))}{E(\nu_d^\sigma(\bigtimes_{i=1}^d ((\varepsilon_{t_i}^+)^{-1}, \infty)))} \tag{2.8}$$

for all $s_1, \dots, s_d > 0$.

Proof. Let $\varepsilon_{t_i}^*$, $1 \le i \le d$, be i.i.d. random variables independent of $(\sigma_{t_1}, \dots, \sigma_{t_d})$ with distribution $P^{\varepsilon_0|\varepsilon_0>0}$ and define $X_{t_i}^* := \varepsilon_{t_i}^* \sigma_{t_i}$, $i = 1, \dots, d$. Apply Theorem 2.3 to the second equation in (2.1) to conclude the regular variation of $(X_{t_1}^*, \dots, X_{t_d}^*)$ on $\mathbb{E}^d$ with limit measure ν_d^X, which is equivalent to the assertion. □

An analogous result holds true if X_t is replaced by $|X_t|$.

In view of Corollary 2.5, we are able to construct SV models with a flexible second-order extremal behavior of the lagged log returns as soon as we find a way to model the volatilities accordingly. A suitable model will be introduced in the next section.

3. SV models with Gamma-type log-volatilities

In order to allow for a richer second-order extremal dependence structure of SV models we modify the assumption of a normal distribution of the log-volatilities in a way which allows for heavy tails of the volatility process. In our construction, we rely on results from Rootzén [27] which guarantee the existence of stationary time series that meet our assumptions.

Definition 3.1. *Let*

$$X_t = \sigma_t \varepsilon_t, \qquad t \in \mathbb{Z}, \tag{3.1}$$

with ε_t, $t \in \mathbb{Z}$, i.i.d. such that $P(\varepsilon_0 > 0) > 0$ and $E(|\varepsilon_0|^{1+\delta}) < \infty$ holds for some $\delta > 0$. Furthermore, let

$$\log \sigma_t = \sum_{i=0}^\infty \alpha_i \eta_{t-i}, \qquad t \in \mathbb{Z}, \tag{3.2}$$

with:

(a) *coefficients $\alpha_i \in [0, 1]$, $i \in \mathbb{N}_0$, such that $\max_{i\in\mathbb{N}_0} \alpha_i = 1$ and $\alpha_i = \mathrm{O}(i^{-\theta})$ as $i \to \infty$ for some $\theta > 1$,*

(b) *i.i.d. innovations* $\eta_t, t \in \mathbb{Z}$, *which are independent of* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *with* $E(\eta_0^2) < \infty$ *and*

$$P(\eta_0 > z) \sim K z^{\beta} \mathrm{e}^{-z}, \qquad z \to \infty, \tag{3.3}$$

for a real constant $\beta \neq -1$ *and a positive constant* K.

We call $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *a* stochastic volatility (SV) model with Gamma-type log-volatility.

In accordance with common SV terminology, we will call the process $(\sigma_t)_{t\in\mathbb{Z}}$ the *volatility process*, $(\log \sigma_t)_{t\in\mathbb{Z}}$ the *log-volatility process* and $(\varepsilon_t)_{t\in\mathbb{Z}}$ the *innovation process*. The following theorem is based on Rootzén [27] and guarantees that a heavy-tailed stationary solution to our definition exists.

Theorem 3.2. *There exists a stationary solution* $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *to* (3.1) *and* (3.2) *as in Definition* 3.1 *and the marginal distributions of* $|X_0|$ *and* σ_0 *are regularly varying with index* -1. *Furthermore, the distribution of* X_0 *is tail balanced with*

$$\lim_{x\to\infty} \frac{P(X_0 > x)}{P(|X_0| > x)} = \frac{E(\varepsilon_0^+)}{E(|\varepsilon_0|)}, \qquad \lim_{x\to\infty} \frac{P(X_0 < -x)}{P(|X_0| > x)} = \frac{E(\varepsilon_0^-)}{E(|\varepsilon_0|)} \tag{3.4}$$

and

$$\lim_{x\to\infty} \frac{P(|X_0| > x)}{P(\sigma_0 > x)} = E\big(|\varepsilon_0|\big), \qquad \lim_{x\to\infty} \frac{P(X_0 > x)}{P(\sigma_0 > x)} = E\big(\varepsilon_0^+\big). \tag{3.5}$$

For normalizing constants $a_n := \hat{K} n (\log n)^{\hat{\beta}}, n \in \mathbb{N}$ [*see* (6.5) *and* (6.6) *for the definition of* $\hat{\beta}$ *and* $\hat{K}$] *and* $z > 0$ *we have*

$$\begin{aligned} &P(\sigma_0 \le a_n z)^n \to \exp(-1/z), \qquad P\big(|X_0| \le a_n z\big)^n \to \exp\big(-E\big(|\varepsilon_0|\big)/z\big), \\ &P(X_0 \le a_n z)^n \to \exp\big(-E\big(\varepsilon_0^+\big)/z\big), \qquad n \to \infty. \end{aligned} \tag{3.6}$$

Remark 3.3. (i) The assumption $\max_{i\in\mathbb{N}_0} \alpha_i = 1$ ensures that the index of regular variation equals -1. However, our model can easily be extended to an arbitrary (negative) index of regular variation. To this end, replace (3.2) with

$$\log \sigma_t = c \sum_{i=0}^{\infty} \alpha_i \eta_{t-i}, \qquad t \in \mathbb{Z},$$

for some $c > 0$. Together with the above assumptions this will lead to a solution of (3.2) which is regularly varying with index $-1/c$. If we assume that $E(|\varepsilon_0|^{1/c+\delta}) < \infty$ for some $\delta > 0$, then again by Breiman's lemma this implies a stationary solution to (3.1) which is regularly varying with the same index $-1/c$. For the sake of notational simplicity, we will stick to the original definition in the following analysis.

(ii) Stochastic volatility models often have an additional parameter specifying the mean of the log-volatilities (cf., e.g., Taylor [29]). In this case, equation (3.2) would read as $\log\sigma_t = \mu + \sum_{i=0}^{\infty}\alpha_i\eta_{t-i}$ for some $\mu \in \mathbb{R}$. Such an assumption is usually combined with the standardization of some moment of ε_0, for example, by setting $\operatorname{Var}(\varepsilon_0) = 1$. Otherwise, setting $\hat{\varepsilon}_i := \mathrm{e}^{\mu}\varepsilon_i$, $i \in \mathbb{Z}$, has the same effect as adding μ in the definition of $\log\sigma_t$. Since we make no assumptions about the particular form of (existing) moments of ε_0, we set $\mu = 0$ without loss of generality.

(iii) The case where $\beta = -1$ in (3.3) is a boundary case which is not treated in detail in Rootzén [27], therefore, it is excluded in Definition 3.1.

Example 3.4. An interesting special case is given by $\alpha_i := \alpha^i$, $i \in \mathbb{N}_0$, for some $\alpha \in (0, 1)$. This case corresponds to an AR(1) model for the log-volatilities, that is,

$$\log\sigma_t = \alpha\log\sigma_{t-1} + \eta_t, \qquad t \in \mathbb{Z}. \tag{3.7}$$

A similar model, with a modified assumption about the distribution of η_t, $t \in \mathbb{Z}$ (namely, that the distribution of $\exp(\eta_t)$ is regularly varying and that a stationary solution to (3.7) exists) has been analyzed with respect to its first-order extremal behavior in Mikosch and Rezapour [22].

Moreover, the conditional extreme value behavior of consecutive observations given that X_0 is large has been analyzed by Kulik and Soulier [17] in the case that the innovations η_t of the log-volatility series are double exponentially distributed. See Section 5 for a more detailed comparison with their results.

Next, we are interested in the extremal behavior of the *processes* $(X_t)_{t\in\mathbb{Z}}$ and $(\sigma_t)_{t\in\mathbb{Z}}$, particularly in their extremal dependence structure. Some information on their first-order extremal dependence behavior may readily be derived from the point process results in Rootzén [27] for the process of the log-volatilities. In the following, let $M_p(\mathbb{E})$ denote the set of Radon point measures on a topological space $\mathbb{E}$. For an introduction to point processes in the context of extreme values, see Resnick [25], Chapters 5 and 7.

Theorem 3.5. *Let $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ be an SV model with Gamma-type log-volatility. In the case that $\beta < -1$ assume additionally that $k := |\{n \in \mathbb{N}_0 : \alpha_n = 1\}| = 1$. With a_n, $n \in \mathbb{N}$, as in Theorem 3.2, let N_n^{σ}, $n \in \mathbb{N}$, and N_n^X, $n \in \mathbb{N}$, denote the point processes defined by*

$$N_n^{\sigma}(\cdot) := \sum_{i=1}^{n}\delta_{(i/n,\sigma_i/a_n)}(\cdot), \qquad N_n^{X}(\cdot) := \sum_{i=1}^{n}\delta_{(i/n,X_i/a_n)}(\cdot).$$

(i) *Then, as $n \to \infty$, $N_n^{\sigma} \xrightarrow{w} N^{\sigma}$ in $M_p([0,1]\times(0,\infty])$, where N^{σ} is a Poisson process with intensity measure $\mathrm{d}t \times z^{-2}\,\mathrm{d}z$.*

(ii) *Let $(t_{(i)}, z_{(i)})_{i\in\mathbb{N}}$ denote the points of the Poisson process N^{σ} and $(\varepsilon_{(i)})_{i\in\mathbb{N}}$ an i.i.d. sequence with $P^{\varepsilon_{(1)}} = P^{\varepsilon_0}$, independent of $(t_{(i)}, z_{(i)})_{i\in\mathbb{N}}$.*

Then $N_n^X(\cdot\cap D)\overset{w}{\to}N^X(\cdot\cap D)$ *in* $M_p(D)$, *as* $n\to\infty$, *where* $D:=[0,1]\times[-\infty,\infty]\setminus\{0\}$ *and* N^X *is a point process consisting of the points* $(t_{(i)},z_{(i)}\varepsilon_{(i)})_{i\in\mathbb{N}}$. *The restriction of* N^X *to* D *is a Poisson point process with intensity measure* $\mathrm{d}t\times z^{-2}[\mathbb{1}_{\{z<0\}}E(\varepsilon_0^-)+\mathbb{1}_{\{z>0\}}E(\varepsilon_0^+)]\,\mathrm{d}z$.

From part (i) of Theorem 3.5, we may conclude that

$$P\big(\max\{\sigma_1,\dots,\sigma_n\}\le a_nz\big)=P\big(N_n^\sigma\big([0,1]\times(z,\infty)\big)=0\big)\to\exp(-1/z)$$

for all $z>0$. A comparison with (3.6) thus yields that the extremal index of the stationary sequence $(\sigma_t)_{t\in\mathbb{Z}}$ exists and equals 1. The same holds true for the processes $(|X_t|)_{t\in\mathbb{Z}}$ and $(X_t)_{t\in\mathbb{Z}}$ by part (ii) of the theorem. Hence, the extremes in these processes do not cluster asymptotically, in the sense that the projections of the limiting point processes obtained in Theorem 3.5 on the time coordinate are simple (i.e., they do not have multiple points), cf. Leadbetter [18]. Moreover, the form of the limiting point process implies asymptotic independence, that is,

$$P(\sigma_h>x|\sigma_0>x)\to 0\quad\text{and}\quad P(X_h>x|X_0>x)\to 0$$

as $x\to\infty$ for all $h\in\mathbb{Z}\setminus\{0\}$.

Hence, in this respect, the SV model with Gamma-type log-volatility shows the same first-order extremal dependence behavior as classical SV time series. In Section 5, though, we will see that the second-order extremal behavior of these classes of processes is quite different. In what follows, we focus on the asymptotics for the probabilities $P(\sigma_0>s_0x,\sigma_h>s_hx)$ and $P(X_0>s_0x,X_h>s_hx)$ as $x\to\infty$ for different values of $h\in\mathbb{N}$ and $s_0,s_h>0$. For simplicity, we will restrict ourselves to the analysis of the upper tails of the process $(X_t)_{t\in\mathbb{Z}}$. However, the necessary changes to analyze both upper and lower tails become obvious by writing $P(X_0>s_0x,X_h<-s_hx)=P(\sigma_0\varepsilon_0^+>s_0x,\sigma_h\varepsilon_h^->s_hx)$ for $s_0,s_h,x>0$.

4. Joint extremal behavior of power products

In this section, we analyze the joint extremal behavior of products of the form $Y_i=\prod_{j=1}^\infty X_j^{\alpha_{ij}}$, $1\le i\le d$, for i.i.d. non-negative random variables X_j, $j\in\mathbb{N}$, which are regularly varying with index -1. The connection to SV models with Gamma-type log-volatilities becomes clear by writing

$$\sigma_t=\exp\Bigg(\sum_{i=0}^\infty\alpha_i\eta_{t-i}\Bigg)=\prod_{i=0}^\infty\big(\exp(\eta_{t-i})\big)^{\alpha_i},\qquad t\in\mathbb{Z}.$$

Most of this section deals with the case $d=2$ and we will write $\prod_{i=1}^\infty X_i^{\alpha_i}$ and $\prod_{i=1}^\infty X_i^{\beta_i}$ for notational convenience. The general case will briefly be discussed at the end of this section. We will show that the joint tail behavior of these two products is closely related to the following infinite-dimensional linear optimization problem:

$$\sum_{i=1}^\infty\kappa_i\to\min!$$

under the constraints

$$\sum_{i=1}^{\infty} \alpha_i \kappa_i \geq 1, \qquad \sum_{i=1}^{\infty} \beta_i \kappa_i \geq 1, \qquad \kappa_i \geq 0, \forall i \in \mathbb{N}.$$

This relation can be explained by the following heuristic argument for $\alpha_i, \beta_i \geq 0$. Suppose that $(\kappa_i)_{i\in\mathbb{N}}$ is a sequence that fulfills the constraints. Then the event $\{X_i > x^{\kappa_i}, i \in \mathbb{N}\}$ implies both

$$\prod_{i=1}^{\infty} X_i^{\alpha_i} > \prod_{i=1}^{\infty} x^{\alpha_i \kappa_i} \geq x$$

and

$$\prod_{i=1}^{\infty} X_i^{\beta_i} > \prod_{i=1}^{\infty} x^{\beta_i \kappa_i} \geq x$$

for $x \geq 1$. Now, $x \mapsto P(X_i > x^{\kappa_i}, i \in \mathbb{N})$ is regularly varying with index $-\sum_{i=1}^{\infty} \kappa_i$. Hence, if $\sum_{i=1}^{\infty} \kappa_i$ is minimized, the above event is, heuristically, the "most likely" combination of extremal events which lead to $\{\prod_{i=1}^{\infty} X_i^{\alpha_i} > x, \prod_{i=1}^{\infty} X_i^{\beta_i} > x\}$.

We will make frequent use of the so-called Potter bounds (Bingham *et al.* [5], Theorem 1.5.6) for functions $f : [0, \infty) \to (0, \infty)$ which are regularly varying with index $-\alpha$: for all $\varepsilon > 0$ there exists a constant $M = M(\varepsilon)$ such that if $\min\{x, sx\} > M$, then

$$(1 - \varepsilon) f(x) s^{-\alpha \mp \varepsilon} \leq f(sx) \leq (1 + \varepsilon) f(x) s^{-\alpha \pm \varepsilon}, \tag{4.1}$$

where, for abbreviation,

$$s^{\pm\varepsilon} := \max\{s^{\varepsilon}, s^{-\varepsilon}\}, \qquad s^{\mp\varepsilon} := \min\{s^{\varepsilon}, s^{-\varepsilon}\}. \tag{4.2}$$

Before we deal with infinite products, we first analyze products of two factors in the case that a unique solution to the above optimization problem exists with $\kappa_1, \kappa_2 > 0$.

Proposition 4.1. *Let $X_1, X_2 \geq 0$ be two independent random variables which are both regularly varying with index -1. For constants $\alpha_1, \alpha_2, \beta_1, \beta_2 \geq 0$, we assume that the linear optimization problem*

$$\begin{aligned} &\kappa_1 + \kappa_2 \to \min! \\ &\quad \textit{under the constraints } \kappa_1\alpha_1 + \kappa_2\alpha_2 \geq 1, \kappa_1\beta_1 + \kappa_2\beta_2 \geq 1, \kappa_1, \kappa_2 \geq 0 \end{aligned} \tag{4.3}$$

has a unique solution satisfying $\kappa_1, \kappa_2 > 0$. Let

$$C := \frac{|\alpha_1\beta_2 - \alpha_2\beta_1|}{(\alpha_1 - \alpha_2)(\beta_2 - \beta_1)}.$$

Then, to each $\varepsilon > 0$, *there exists a constant* $M = M(\varepsilon) > 0$ *such that*

$$(1-\varepsilon)Cz_1^{1\mp\varepsilon}z_2^{1\mp\varepsilon} \le \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \le (1+\varepsilon)Cz_1^{1\pm\varepsilon}z_2^{1\pm\varepsilon}$$

for all $x, z_1, z_2 > 0$ *satisfying* $\min\{x^{\kappa_1}, z_1^{-1}x^{\kappa_1}, x^{\kappa_2}, z_2^{-1}x^{\kappa_2}\} > M$.

The above result can be interpreted as a bivariate extension of the Potter bounds for random products and is essential for the proof of the following theorem.

Theorem 4.2. *Let* $X_i, i \in \mathbb{N}$, *be i.i.d. non-negative random variables which are regularly varying with index* -1. *Assume that* $\alpha_i, \beta_i \ge 0$, $i \in \mathbb{N}$, *are such that neither all* α_i *nor all* β_i *are equal to* 0 *and* $\sum_{i=1}^{\infty}\alpha_i < \infty$, $\sum_{i=1}^{\infty}\beta_i < \infty$. *Then*

$$Y_0 := \lim_{n\to\infty}\prod_{i=1}^{n} X_i^{\alpha_i}, \qquad Y_1 := \lim_{n\to\infty}\prod_{i=1}^{n} X_i^{\beta_i}$$

exist almost surely. We assume that $P(Y_i > 0) > 0, i = 0, 1$.

Then the optimization problem

$$\begin{aligned}&\sum_{i=1}^{\infty}\kappa_i \to \min!\\ &\text{under the constraints } \sum_{i=1}^{\infty}\alpha_i\kappa_i \ge 1, \sum_{i=1}^{\infty}\beta_i\kappa_i \ge 1, \kappa_i \ge 0, \forall i \in \mathbb{N},\end{aligned} \tag{4.4}$$

has a solution $(\kappa_i)_{i\in\mathbb{N}}$.

(i) *For all* $\varepsilon > 0$,

$$P(Y_0 > x, Y_1 > x) = \mathrm{o}\big(x^{\varepsilon - \sum_{i=1}^{\infty}\kappa_i}\big), \tag{4.5}$$

$$x^{-\varepsilon - \sum_{i=1}^{\infty}\kappa_i} = \mathrm{o}\big(P(Y_0 > x, Y_1 > x)\big) \tag{4.6}$$

as $x \to \infty$.

If the solution to (4.4) *is unique, then at most two of the* κ_i *are strictly positive.*

(ii) *If the solution is unique with* $\kappa_i, \kappa_j > 0$ *for some* $i \ne j$, *then for all* $s_0, s_1 > 0$

$$\lim_{x\to\infty}\frac{P(Y_0 > s_0x, Y_1 > s_1x)}{P(X_i > x^{\kappa_i})P(X_j > x^{\kappa_j})} = D\big(s_0^{\beta_i-\beta_j}s_1^{\alpha_j-\alpha_i}\big)^{1/(\alpha_i\beta_j-\alpha_j\beta_i)} \tag{4.7}$$

with

$$D := \frac{|\alpha_i\beta_j - \alpha_j\beta_i|}{(\alpha_i - \alpha_j)(\beta_j - \beta_i)} E\left(\prod_{m\in\mathbb{N}\setminus\{i,j\}} X_m^{(\alpha_m(\beta_j-\beta_i)+\beta_m(\alpha_i-\alpha_j))/(\alpha_i\beta_j-\alpha_j\beta_i)}\right) < \infty.$$

(iii) *If the solution is unique with $\kappa_i > 0$ for exactly one $i \in \mathbb{N}$ and $\kappa_j = 0$ else, then for all $s_0, s_1 > 0$*

$$\lim_{x\to\infty} \frac{P(Y_0 > s_0 x, Y_1 > s_1 x)}{P(X_i > x^{\kappa_i})}$$
$$= \begin{cases} E\left(\min\left(s_0^{-1} \prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\alpha_j}, s_1^{-1} \prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\beta_j}\right)^{1/\beta_i}\right), & \alpha_i = \beta_i, \\ E\left(\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\beta_j/\beta_i} \mathbb{1}_{\{X_j^{\alpha_j} > 0\}}\right) s_1^{-1/\beta_i}, & \alpha_i > \beta_i, \\ E\left(\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\alpha_j/\alpha_i} \mathbb{1}_{\{X_j^{\beta_j} > 0\}}\right) s_0^{-1/\alpha_i}, & \alpha_i < \beta_i, \end{cases} \tag{4.8}$$
$$< \infty.$$

Remark 4.3. (i) Under our assumptions about the summability of the α_i's and β_i's, the infinite-dimensional linear program (4.4) can always be boiled down to a finite-dimensional linear program, which can easily be solved numerically. Indeed, an optimal solution $(\kappa_i)_{i\in\mathbb{N}}$ must satisfy $\kappa_i = 0$ if $\max(\alpha_i, \beta_i) < 1/(2/\max_{j\in\mathbb{N}} \alpha_j + 2/\max_{j\in\mathbb{N}} \beta_j)$; see part 2 of the proof of Theorem 4.2 for details.

(ii) If $E(\log X_1)^- < \infty$, then Y_0 and Y_1 are almost surely strictly positive, because then $E(\sum_{i=1}^k (\alpha_i + \beta_i)\log X_i)$ converges in $\mathbb{R}$ as $k \to \infty$. In particular, it suffices to assume that $P(X_1 \le x) = o(|\log x|^{-(1+\varepsilon)})$ as $x \downarrow 0$ for some $\varepsilon > 0$.

(iii) The assumption that all X_j are identically distributed can be dropped if one of the following two conditions is fulfilled:

- there exists $n \in \mathbb{N}$ such that $\alpha_i = \beta_i = 0$ for all $i \ge n$,
- $\sup_{i\in\mathbb{N}} E(X_i^\varepsilon) < \infty$ for some $\varepsilon > 0$ (which ensures that the infinite products converge and a moment bound of similar type as (6.20) holds).

(iv) If the solution to the linear program is not unique, it is sometimes possible that a slight redefinition of the factors X_i leads to linear program with a unique solution. Think, for example, of $\alpha_i = \beta_i = 1, i = 1, 2$, $\sup_{i\ge 3}\{\alpha_i, \beta_i\} < 1$. If we define $\hat{X}_1 = X_1 X_2$, $\hat{X}_i = X_{i+1}, i \ge 2$, then $\hat{X}_1$ is regularly varying with index -1 by the corollary to Theorem 3 in Embrechts and Goldie [14]. The resulting linear program has a unique solution $\hat{\kappa}_1 = 1, \hat{\kappa}_i = 0, i \ge 2$, and $P(Y_0 > s_0 x, Y_1 > s_1 x) \sim E(\min\{s_0^{-1} \prod_{i\ge 3} X_i^{\alpha_i}, s_1^{-1} \prod_{i\ge 3} X_i^{\beta_i}\}) P(X_1 X_2 > x)$. However, the probability on the right-hand side cannot be easily expressed in terms of tail probabilities of X_1 and X_2. See Denisov and Zwart [11] and Embrechts and Goldie [14] for discussions of the distribution of the product of two factors with the same index of regular variation.

Remark 4.4. (i) In the situation of Theorem 4.2(ii) and (iii) $\min\{Y_0, Y_1\}$ is regularly varying with index $-\sum_{i=1}^{\infty} \kappa_i$.

(ii) Under the assumptions of Theorem 4.2(ii) the random vector (Y_0, Y_1) is regularly varying on the cone $\mathbb{E}^2 = (0, \infty)^2$ with limiting measure ν given by

$$\nu\big((s_0, \infty) \times (s_1, \infty)\big) = \big(s_0^{\beta_i - \beta_j} s_1^{\alpha_j - \alpha_i}\big)^{1/(\alpha_i \beta_j - \alpha_j \beta_i)}, s_0, \qquad s_1 > 0.$$

(iii) Under the assumptions of Theorem 4.2(iii) with $\alpha_i = \beta_i$ the random vector (Y_0, Y_1) is regularly varying on the cone $\mathbb{E}^2$ with limiting measure ν given by

$$\nu\big((s_0, \infty) \times (s_1, \infty)\big) = \frac{E(\min(s_0^{-1} \prod_{j \in \mathbb{N} \setminus \{i\}} X_j^{\alpha_j}, s_1^{-1} \prod_{j \in \mathbb{N} \setminus \{i\}} X_j^{\beta_j})^{1/\beta_i})}{E(\min(\prod_{j \in \mathbb{N} \setminus \{i\}} X_j^{\alpha_j}, \prod_{j \in \mathbb{N} \setminus \{i\}} X_j^{\beta_j})^{1/\beta_i})},$$

$s_0, s_1 > 0$.

(iv) If the assumptions of Theorem 4.2(iii) are fulfilled with $\alpha_i \neq \beta_i$, then there exists a measure ν on $\mathcal{B}((0, \infty]^2)$ such that convergence (1.7) holds for all ν-continuous Borel sets $B \subset (0, \infty]^2$ bounded away from $\mathcal{O}^2$, but ν is concentrated on $(\{\infty\} \times (0, \infty)) \cup ((0, \infty) \times \{\infty\})$. Thus, the restriction of ν to $\mathbb{E}^2$ equals the zero measure and (Y_0, Y_1) is not regularly varying on $\mathbb{E}^2$. Note that this case cannot occur if Y_0 and Y_1 have the same distribution, as it will be the case in the applications considered in the next section.

So far, we have analyzed the joint extremal behavior of only two products of powers of the random variables X_j. In the remainder of this section, we briefly discuss the joint extreme value behavior of an arbitrary number of such products. However, as a generalization of the Potter-type result established in Proposition 4.1 to arbitrary dimension is rather cumbersome, we focus on a special case in which one can use the one-dimensional Potter bounds instead.

In what follows, we consider products

$$Y_i := \prod_{j=1}^{n} X_j^{\alpha_{ij}}, \qquad 1 \leq i \leq d,$$

where we assume w.l.o.g. that $d \leq n$. (By the arguments given at the beginning of the proof of Theorem 4.2, one can easily deal with infinite products, too.) For the main result in this case, we do not assume any longer that all exponents α_{ij} are non-negative, but that the sub-matrix consisting of the most relevant exponents has a certain structure.

Theorem 4.5. *Assume that X_j, $1 \leq j \leq n$, are independent, non-negative random variables that are regularly varying with index -1 and bounded away from 0. If the optimization problem*

$$\sum_{j=1}^{n} \kappa_j \to \min!$$
$$\textit{under the constraints } \sum_{j=1}^{n} \alpha_{ij} \kappa_j \geq 1, \forall 1 \leq i \leq d, \kappa_j \geq 0, \forall 1 \leq j \leq n, \tag{4.9}$$

has a solution $(\kappa_j)_{1\le j\le n}$, *then for all* $\varepsilon > 0$,

$$P(Y_i > x\ \forall 1 \le i \le d) = \mathrm{o}\big(x^{\varepsilon - \sum_{j=1}^n \kappa_j}\big),$$
$$x^{-\varepsilon - \sum_{j=1}^n \kappa_j} = \mathrm{o}\big(P(Y_i > x\ \forall 1 \le i \le d)\big)$$

as $x \to \infty$.

Now suppose, in addition, that the solution is unique and that $J := \{j | \kappa_j > 0\}$ *has exactly* d *elements. Then the matrix* $\mathbf{A} := (\alpha_{ij})_{1\le i\le d, j\in J}$ *is invertible. If all entries* A^{-1}_{ji} *of its inverse* $\mathbf{A}^{-1}$ *are positive, then for all* $s_i > 0$, $1 \le i \le d$,

$$\lim_{x\to\infty} \frac{P(Y_i > s_i x\ \forall 1 \le i \le d)}{\prod_{j\in J} P(X_j > x^{\kappa_j})} = D \prod_{i=1}^d s_i^{-\sum_{j\in J} A^{-1}_{ji}} \tag{4.10}$$

with

$$D := \frac{1}{|\det \mathbf{A}|} \frac{\prod_{k\notin J} E(X_k^{\sum_{j\in J}\sum_{i=1}^d A^{-1}_{ji}\alpha_{ik}})}{\prod_{i=1}^d \sum_{j\in J} A^{-1}_{ji}}.$$

The assumption $A^{-1}_{ji} > 0$ for all $1 \le i \le d$, $j \in J$, ensures that all factors X_j, $j \in J$, must be large if all products $\prod_{j\in J} X_j^{\alpha_{ij}}$, $1 \le i \le d$, are large. If this condition is not satisfied, then a much more delicate analysis of the probability that at least one of the X_j is of smaller order than x^{κ_j} while still all products $\prod_{j\in J} X_j^{\alpha_{ij}}$, $1 \le i \le d$, exceed x is needed. We conjecture that (4.10) holds true under much more general conditions on $\mathbf{A}$ (in particular, that it holds for arbitrary non-negative matrices), but such a general result is left to a future publication.

5. Second-order behavior of SV models with Gamma-type log-volatility

The results from the two previous sections enable us to analyze the joint extremal behavior of two lagged observations from an SV model with Gamma-type log-volatility.

Theorem 5.1. *Let* $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *be an SV model with Gamma-type log-volatility as in Definition 3.1. Assume that for* $h \in \mathbb{N}$ *there exists a unique solution to the optimization problem*

$$\sum_{i=0}^\infty \kappa_i \to \min! \\ \text{under the constraints } \sum_{i=h}^\infty \alpha_{i-h}\kappa_i \ge 1, \sum_{i=0}^\infty \alpha_i\kappa_i \ge 1, \kappa_i \ge 0, \forall i \in \mathbb{N}_0. \tag{5.1}$$

In addition, suppose that $E((\varepsilon_0^+)^{\sum_{i=0}^\infty \kappa_i + \delta}) < \infty$ *for some* $\delta > 0$.

(i) *If* $\kappa_i, \kappa_j > 0$ *for some* $i \neq j$, *then for all* $s_0, s_h > 0$

$$\lim_{x\to\infty} \frac{P(\sigma_0 > s_0 x, \sigma_h > s_h x)}{P(\min\{\sigma_0, \sigma_h\} > x)} = \lim_{x\to\infty} \frac{P(X_0 > s_0 x, X_h > s_h x)}{P(\min\{X_0, X_h\} > x)} = \left(s_0^{\alpha_i - \alpha_j} s_h^{\alpha_{j-h} - \alpha_{i-h}}\right)^{1/(\alpha_j \alpha_{i-h} - \alpha_i \alpha_{j-h})},$$

where $\alpha_k := 0$ *for* $k < 0$.

(ii) *If* $\kappa_i > 0$ *for exactly one* $i \in \mathbb{N}_0$ *and* $\kappa_j = 0$ *else*, *then* $i \geq h$, $\alpha_i = \alpha_{i-h}$ *and for all* $s_0, s_h > 0$

$$\lim_{x\to\infty} \frac{P(\sigma_0 > s_0 x, \sigma_h > s_h x)}{P(\min\{\sigma_0, \sigma_h\} > x)} = \frac{E(\min(s_0^{-1} \prod_{j\geq h, j\neq i} \mathrm{e}^{\alpha_{j-h}\eta_{h-j}}, s_h^{-1} \prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\alpha_j \eta_{h-j}})^{1/\alpha_i})}{E(\min(\prod_{j\geq h, j\neq i} \mathrm{e}^{\alpha_{j-h}\eta_{h-j}}, \prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\alpha_j \eta_{h-j}})^{1/\alpha_i})},$$

and

$$\lim_{x\to\infty} \frac{P(X_0 > s_0 x, X_h > s_h x)}{P(\min\{X_0, X_h\} > x)} = \frac{E(\min(s_0^{-1} \varepsilon_0^+ \prod_{j\geq h, j\neq i} \mathrm{e}^{\alpha_{j-h}\eta_{h-j}}, s_h^{-1} \varepsilon_h^+ \prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\alpha_j \eta_{h-j}})^{1/\alpha_i})}{E(\min(\varepsilon_0^+ \prod_{j\geq h, j\neq i} \mathrm{e}^{\alpha_{j-h}\eta_{h-j}}, \varepsilon_h^+ \prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\alpha_j \eta_{h-j}})^{1/\alpha_i})}.$$

In both cases, (σ_0, σ_h) *and* (X_0, X_h) *are regularly varying on* $\mathbb{E}^2 = (0, \infty)^2$ *with index* $-\sum_{i=0}^{\infty} \kappa_i$.

Remark 5.2. (i) According to Definition 3.1(a) one has $\alpha_k = 1$ for some $k \in \mathbb{N}_0$. Because $\kappa_k = \kappa_{k+h} = 1$ and $\kappa_i = 0$ for all other $i \in \mathbb{N}_0$ defines a feasible solution of (5.1), the optimal solution satisfies $\sum_{i=0}^{\infty} \kappa_i \leq 2$. Hence, the moment condition on ε_0^+ is always fulfilled if $E((\varepsilon_0^+)^{2+\delta}) < \infty$ for some $\delta > 0$.

(ii) Again, a result similar to Theorem 5.1 holds for $(|X_0|, |X_h|)$ instead of (X_0, X_h) with ε_i^+ replaced by $|\varepsilon_i|$.

In general, the optimization problem (5.1) must be solved numerically. To this end, simple rules may help to find an optimal solution $(\kappa_i)_{i\in\mathbb{N}_0}$ more easily. For example, if $\kappa_i > 0$, then necessarily $\alpha_i + \alpha_{i-h} \geq 1$, because otherwise $\tilde{\kappa}_k := \kappa_k + \alpha_i \kappa_i$, $\tilde{\kappa}_{k+h} := \kappa_{k+h} + \alpha_{i-h}\kappa_i$, $\tilde{\kappa}_i := 0$ and $\tilde{\kappa}_j := \kappa_j$ for all $j \in \mathbb{N}_0 \setminus \{k, k+h, i\}$ with $\alpha_k = 1$ defines a feasible solution with a smaller total sum. This way, (5.1) is reduced to a finite-dimensional program. A special case are strictly decreasing coefficients, where the solution can be easily determined as the following corollary shows.

Corollary 5.3. *Let* $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *be an SV model with Gamma-type log-volatility as in Definition* 3.1 *with* α_i, $i \in \mathbb{N}_0$, *strictly decreasing* (*which implies* $\alpha_0 = 1$). *Then the unique solution* *to* (5.1) *is given by* $\kappa_0 = 1 - \alpha_h$, $\kappa_h = 1$ *and* $\kappa_i = 0$ *else. Furthermore, if* $E((\varepsilon_0^+)^{2-\alpha_h+\delta}) < \infty$ *for some* $\delta > 0$, *then*

$$\lim_{x\to\infty} \frac{P(\sigma_0 > s_0 x, \sigma_h > s_h x)}{P(\min\{\sigma_0, \sigma_h\} > x)} = \lim_{x\to\infty} \frac{P(X_0 > s_0 x, X_h > s_h x)}{P(\min\{X_0, X_h\} > x)} = s_0^{\alpha_h - 1} s_h^{-1}$$

for all $s_0, s_h > 0$.

Theorem 5.1 shows that under the stated assumptions the coefficient of tail dependence is the same for the vectors (σ_0, σ_h) and (X_0, X_h) and equal to $\eta_h = 1/\sum_{i=0}^{\infty} \kappa_i \in [1/2, 1]$. In the situation of Corollary 5.3, one has $\eta_h = 1/(2-\alpha_h)$. In particular, for the AR(1) model considered in Example 3.4 with $\alpha_h = \alpha^h$, $h \in \mathbb{N}_0$, for some $\alpha \in (0, 1)$, the coefficient of tail dependence of the lagged vectors is given by $1/(2 - \alpha^h)$.

If the sequence of coefficients α_h is decreasing, the coefficient of tail dependence is decreasing in h as well and converges to $1/2$ as $h \to \infty$. Thus, the extremal dependence gets weaker over time and its speed of convergence depends solely on the values of α_h, $h \in \mathbb{N}$ (resp., on $\alpha \in (0, 1)$ in the AR(1) model). The strictly monotonic decay of the coefficients of tail dependence seems a very reasonable assumption for asymptotically independent time series. Corollary 5.3 shows that SV models with Gamma-type log-volatility allow for all possible strictly monotonically decreasing functions $h \mapsto \eta_h \in [1/2, 1]$, provided $\sum_{h=1}^{\infty}(2 - 1/\eta_h) < \infty$. Moreover, it is also possible to reproduce arbitrary finite sequences of (not necessarily decreasing) coefficients of tail dependence η_h as long as they reflect a non-negative dependence (i.e., stay in the interval $[1/2, 1]$).

Theorem 5.4. *To each vector* $(\eta_1, \dots, \eta_m) \in [1/2, 1]^m$ $(m \in \mathbb{N})$ *there exists an SV model with Gamma-type log-volatility* $(X_t, \sigma_t)_{t\in\mathbb{Z}}$ *such that the coefficient of tail dependence of* (σ_0, σ_h) *and* (X_0, X_h) *equals* η_h *for all* $1 \le h \le m$.

Remark 5.5. The preceding theorem shows that SV models with Gamma-type log-volatility are also able to reflect $\eta_h = 1$ for $h > 0$. Remember that asymptotic dependence of the vector (X_0, X_h) implies $\eta_h = 1$ but not the other way round. In fact, it depends on the value of β in (3.3) whether our model allows for asymptotic dependence of lagged observations.

(i) If $\beta > -1$, then all vectors (σ_0, σ_h) and (X_0, X_h) show asymptotic independence by Theorem 3.5 and the following conclusions.

(ii) If $\beta < -1$ and $\eta_h = 1$ for some $h > 0$, then the vectors (σ_0, σ_h) and (X_0, X_h) show asymptotic dependence. See Section 6 for details.

We conclude this section with a comparison of our work and the results by Kulik and Soulier [17], Sections 3 and 4, who consider a similar class of SV models. They analyze the limit distributions

$$\lim_{x\to\infty} P\big(\sigma_0 \le s_0 x, \sigma_1 \le s_1 x^{\rho_1}, \dots, \sigma_h \le s_h x^{\rho_h} | \sigma_0 > x\big) \tag{5.2}$$

for a suitable choice of so-called “conditional scaling exponents” ρ_j, $j \in \mathbb{N}$, which lead to a non-degenerate limit. So while we examine the joint extremal behavior of consecutive volatilities or returns using regular variation on the cones $\mathbb{E}^d$, Kulik and Soulier [17] work in the framework of conditional extreme value models, which are discussed, for example, in Das and Resnick [8]. In the case of asymptotic dependence with $\rho_i = 1$ for all $i \in \mathbb{N}$, the resulting limit process is known as the tail process [cf. Basrak and Segers [4]]. Kulik and Soulier [17] consider an AR(1) model for the log-volatilities where the innovations $\eta_t, t \in \mathbb{Z}$, have a double exponential distribution, that is, they are symmetric with $P(\eta_t > x) = \exp(-\alpha x)/2$ for $x > 0$. This assumption fits into our model [we restrict our analysis to the case $\alpha = 1$ by standardization, cf. Remark 3.3(i)]. Furthermore, they deal with linear models of the form (1.2) with double exponentially distributed innovations under the additional assumptions that $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$ (i.e., they allow for a non-summable sequence of coefficients) and that $\alpha_0 = 1, \alpha_i < 1$ for all $i \geq 1$.

Kulik and Soulier show that for those models a non-degenerate limit in (5.2) exists if and only if the conditional scaling exponents ρ_j are chosen equal to α_j for all $j \in \mathbb{N}$. As $\alpha_j < 1$ for all $j \in \mathbb{N}$, this implies asymptotic independence of consecutive volatilities and of consecutive returns. Like Theorem 5.1, convergence (5.2) conveys refined information on their extremal dependence structure, but the focus of the approach by Kulik and Soulier is quite different from ours, and their mathematical techniques are in a sense considerably simpler than the ones employed in the present paper. Indeed, convergence (5.2) can be heuristically explained by the classical "Breiman's principle", according to which the tail behavior of a product is largely determined by the most heavy tailed factor. Under the condition $\alpha_j < \alpha_0$ for all $j \in \mathbb{N}$, a large value of $\sigma_0 = \prod_{j=0}^{\infty} e^{\alpha_j \eta_{-j}}$ is most likely caused by a large value of η_0. This in turn implies that, given an extreme event at time 0, the lagged volatility $\sigma_h = \prod_{j=0}^{\infty} e^{\alpha_j \eta_{h-j}}$ will be roughly of the order $e^{\alpha_h \eta_0}$, which yields $\rho_h = \alpha_h$ for all $h \in \mathbb{N}$.

In contrast, we consider events of the type that both σ_0 and σ_h exceed the same large threshold. The simple heuristic of above fails in this setting since our results show that a single extreme event at time 0 is not necessarily the most probable cause for the joint exceedance. Instead one has to find combinations of two factors e^{η_j} which are both sufficiently large (though potentially smaller than the single factor considered in the conditional extreme value approach) such that both products $\prod_{j=0}^{\infty} e^{\alpha_j \eta_{-j}}$ and $\prod_{j=0}^{\infty} e^{\alpha_j \eta_{h-j}}$ are large, which leads to the linear optimization problems investigated in the Sections 4 and 5. This clearly shows that in general one should neither expect a simple relationship between the coefficients of tail dependence obtained in this paper on the one hand and the conditional scaling exponents considered in Kulik and Soulier [17] on the other hand, nor between the respective limiting measures arising in both approaches. This fact somewhat qualifies the heuristic reasoning given in Section 1.5 of Kulik and Soulier [17].

6. Proofs

6.1. Proofs to Section 2

The following inequality will be useful. For all $\mathbf{x} \in \mathbb{E}^d$ and $\mathbf{A} \in \mathbb{R}^{d \times d}$, we have

$$\tau(\mathbf{A}\mathbf{x}) = \tau(\mathbf{x})\tau\left(\mathbf{A}\frac{\mathbf{x}}{\tau(\mathbf{x})}\right) \leq \tau(\mathbf{x})\tau(\mathbf{A}). \tag{6.1}$$

Proof of Lemma 2.1.

"(i) $\Rightarrow$ (ii)": By definition $\tau(\mathbf{A})=0$ if and only if $\mathbf{A}\mathcal{S}^d$ and $\mathbb{E}^d$ are disjoint, which in turn is equivalent to $\mathbf{A}^{-1}(\mathbb{E}^d)\cap\mathbb{E}^d=\varnothing$.

Now suppose that $\tau(\mathbf{A})=\infty$. Then there exists $\mathbf{x}\in\mathcal{S}^d$ such that $\min(\mathbf{Ax})>-\min(-\mathbf{A1})$. Thus $\mathbf{y}:=\mathbf{x}-\mathbf{1}\in\mathcal{O}^d\subset(\mathbb{E}^d)^c$ and $\min(\mathbf{Ay})\geq\min(\mathbf{Ax})+\min(-\mathbf{A1})>0$, which implies $\mathbf{Ay}\in\mathbb{E}^d$, and hence $\mathbf{A}^{-1}(\mathbb{E}^d)\not\subset\mathbb{E}^d$. Since this contradicts (i), we have shown that (i) $\Rightarrow$ (ii).

"(ii) $\Rightarrow$ (iii)": Note that $\mathbf{A}$ must be invertible if $\tau(\mathbf{A})\in(0,\infty)$. To see this, suppose that $\mathbf{A}$ were not invertible and choose some $\mathbf{y}\neq\mathbf{0}$ satisfying $\mathbf{Ay}=0$. Because $\tau(\mathbf{A})>0$, there exists a vector $\mathbf{x}\in\mathbb{E}^d$ such that $\tau(\mathbf{Ax})>0$. Moreover, for some $\lambda_0\in\mathbb{R}$ one has $\mathbf{x}+\lambda_0\mathbf{y}\in\mathcal{O}^d$ and $\mathbf{x}+\lambda\mathbf{y}\in\mathbb{E}^d$ for all λ between 0 and λ_0. By (6.1), we have for those λ

$$\tau(\mathbf{A})\geq\frac{\tau(\mathbf{A}(\mathbf{x}+\lambda\mathbf{y}))}{\tau(\mathbf{x}+\lambda\mathbf{y})}=\frac{\tau(\mathbf{Ax})}{\tau(\mathbf{x}+\lambda\mathbf{y})}\to\infty$$

as $\lambda\to\lambda_0$. Since this contradicts the assumption $\tau(\mathbf{A})\in(0,\infty)$, the matrix $\mathbf{A}$ must be invertible. In addition, if $\tau(\mathbf{A})\in(0,\infty)$ then all entries of $\mathbf{A}^{-1}$ must be non-negative. To see this, suppose that there exists a negative entry. Then there exists $\tilde{\mathbf{y}}\in\mathbb{E}^d$ with $\mathbf{y}:=\mathbf{A}^{-1}\tilde{\mathbf{y}}\in([0,\infty)^d)^c$. Furthermore, as above there exists $\mathbf{x}\in\mathbb{E}^d$ such that $\mathbf{Ax}\in\mathbb{E}^d$. Choose $\lambda_0>0$ such that $\mathbf{x}+\lambda_0\mathbf{y}\in\mathcal{O}^d$. Then, for all $\lambda\in[0,\lambda_0)$ we have

$$\tau(\mathbf{A})\geq\frac{\tau(\mathbf{A}(\mathbf{x}+\lambda\mathbf{y}))}{\tau(\mathbf{x}+\lambda\mathbf{y})}=\frac{\tau(\mathbf{Ax}+\lambda\tilde{\mathbf{y}})}{\tau(\mathbf{x}+\lambda\mathbf{y})}\geq\frac{\tau(\mathbf{Ax})}{\tau(\mathbf{x}+\lambda\mathbf{y})}\to\infty$$

by (6.1) as $\lambda\nearrow\lambda_0$. Since this contradicts the assumption, (iii) follows from (ii).

"(iii) $\Rightarrow$ (i)": If $\mathbf{A}$ is invertible with non-negative entries of $\mathbf{A}^{-1}$, then each row of $\mathbf{A}^{-1}$ has at least one strictly positive entry and thus $\mathbf{A}^{-1}\mathbf{x}\in\mathbb{E}^d$ for all $\mathbf{x}\in\mathbb{E}^d$. Therefore, (i) follows from (iii).

The last statement of the lemma follows from (i) by continuity of the linear mapping $\mathbf{x}\mapsto\mathbf{A}^{-1}\mathbf{x}$.

□

Proof of Lemma 2.2. Note first that we have $\{\mathbf{x}\in\mathbb{E}^d:\tau(\mathbf{x})=y\}=y\cdot\mathbf{1}+\mathcal{O}^d$ for all $y>0$. Now, since $0<\tau(\mathbf{A})<\infty$ by assumption and $\tau(t\mathbf{x})=t\tau(\mathbf{x})$ for all $\mathbf{x}\in\mathbb{E}^d$ and $t>0$, it follows from the definition of $\tau(\mathbf{A})$ that

$$\begin{aligned}1&=\inf_{\mathbf{x}\in\mathbb{E}^d:\,\tau(\mathbf{Ax})=\tau(\mathbf{A})}\tau(\mathbf{x})\\&=\inf_{\mathbf{x}:\,\mathbf{Ax}\in\tau(\mathbf{A})\mathbf{1}+\mathcal{O}^d}\tau(\mathbf{x})\\&=\inf_{\mathbf{y}\in\mathcal{O}^d}\tau\big(\mathbf{A}^{-1}\tau(\mathbf{A})\mathbf{1}+\mathbf{A}^{-1}\mathbf{y}\big)\\&=\tau\big(\mathbf{A}^{-1}\tau(\mathbf{A})\mathbf{1}\big)=\tau(\mathbf{A})\tau\big(\mathbf{A}^{-1}\mathbf{1}\big),\end{aligned}$$

where in the last but one equation we have used that $\mathbf{A}^{-1}\mathbf{y}\in[0,\infty)^d$ because all entries of $\mathbf{A}^{-1}$ are non-negative by Lemma 2.1. Thus, $\tau(\mathbf{A})=(\tau(\mathbf{A}^{-1}\mathbf{1}))^{-1}$ and the statement follows from the definition of τ. □

Proof of Theorem 2.3. For all Borel sets $B \subset \mathbb{E}^d$ bounded away from $\mathcal{O}^d$ there exists a constant $\delta_B > 0$ such that $\tau(\mathbf{x}) = \min(\mathbf{x}) > \delta_B$ for all $\mathbf{x} \in B$. Hence, for all $x > 0$, $M > 0$,

$$\begin{aligned} P\big(\mathbf{AZ} \in xB, \tau(\mathbf{A}) > M\big) &\le P\big(\tau(\mathbf{AZ}) > x\delta_B, \tau(\mathbf{A}) > M\big) \\ &\le P\big(\tau(\mathbf{A})\min(\mathbf{Z}) > x\delta_B, \tau(\mathbf{A}) > M\big) \\ &= P\big(\mathbb{1}_{\{\tau(\mathbf{A})>M\}}\tau(\mathbf{A})\min(\mathbf{Z}) > x\delta_B\big), \end{aligned}$$

where the second inequality follows from (6.1) and because $\mathbf{AZ} \in xB$ implies $\mathbf{Z} \in \mathbb{E}^d$ by Lemma 2.1. Since $\min(\mathbf{Z})$ is regularly varying with index $-\alpha$ and $\mathbf{A}$ and $\mathbf{Z}$ are assumed to be independent, the univariate version of Breiman's lemma in combination with (2.6) yields

$$\limsup_{x\to\infty} \frac{P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) > M)}{P(\min(\mathbf{Z}) > x)} \le \delta_B^{-\alpha} E\big(\tau(\mathbf{A})^\alpha \mathbb{1}_{\{\tau(\mathbf{A})>M\}}\big). \tag{6.2}$$

Since, by (6.1),

$$\tau\big(\mathbf{A}^{-1}\mathbf{x}\big) \ge \frac{\tau(\mathbf{AA}^{-1}\mathbf{x})}{\tau(\mathbf{A})} \ge \frac{\delta_B}{\tau(\mathbf{A})} > 0$$

a.s. for all $\mathbf{x} \in B$, the image of B under $\mathbf{A}^{-1}$ is again a.s. bounded away from $\mathcal{O}^d$. Hence,

$$\begin{aligned} \lim_{x\to\infty} \frac{P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) \le M)}{P(\min(\mathbf{Z}) > x)} &= \lim_{x\to\infty} \int_{\{\tau(\mathbf{a})\le M\}} \frac{P(\mathbf{Z} \in x\mathbf{a}^{-1}B)}{P(\min(\mathbf{Z}) > x)} \, \mathrm{d}P^{\mathbf{A}}(\mathbf{a}) && (6.3) \\ &= \int_{\{\tau(\mathbf{a})\le M\}} \nu\big(\mathbf{a}^{-1}B\big) \, \mathrm{d}P^{\mathbf{A}}(\mathbf{a}) \\ &= E\big(\nu\big(\mathbf{A}^{-1}B\big)\mathbb{1}_{\{\tau(\mathbf{A})\le M\}}\big), && (6.4) \end{aligned}$$

where $P^{\mathbf{A}}$ denotes the law of $\mathbf{A}$. For the second equation, we used that $E(\nu(\partial(\mathbf{A}^{-1}B))) = 0$ (i.e., $\nu(\partial(\mathbf{A}^{-1}B)) = 0$ a.s.) in combination with Pratt's lemma [cf. Pratt [23]], since the integrand in (6.3) is bounded by

$$\frac{P(\mathbf{aZ} \in xB)}{P(\min(\mathbf{Z}) > x)} \le \frac{P(\tau(\mathbf{a})\min(\mathbf{Z}) > x\delta_B)}{P(\min(\mathbf{Z}) > x)} \le \frac{P(M\min(\mathbf{Z}) > x\delta_B)}{P(\min(\mathbf{Z}) > x)} \to M^\alpha \delta_B^{-\alpha}.$$

Let $M \to \infty$ in (6.2) and (6.4) to obtain (2.7) by monotone convergence. Note that $E(\nu(\mathbf{A}^{-1}B))$ is finite since, by the definition of τ and homogeneity of ν,

$$\begin{aligned} E\big(\nu\big(\mathbf{A}^{-1}B\big)\big) &\le E\big(\nu\big(\mathbf{A}^{-1}[\delta_B, \infty)^d\big)\big) \le E\big(\nu\big(\big[\delta_B/\tau(\mathbf{A}), \infty\big)^d\big)\big) \\ &= \nu\big([\delta_B, \infty)^d\big) E\big(\tau(\mathbf{A})^\alpha\big) < \infty. \end{aligned}$$

□

6.2. Proofs to Section 3

Proof of Theorem 3.2. Let $\Lambda := \{n \in \mathbb{N}_0 : \alpha_n = 1\}$, $k := |\Lambda|$ (our assumptions guarantee that $k < \infty$),

$$\hat{\beta} := \begin{cases} k\beta + k - 1, & \beta > -1, \\ \beta, & \beta < -1 \end{cases} \tag{6.5}$$

and

$$\hat{K} := \begin{cases} K^k \dfrac{\Gamma(\beta+1)^k}{\Gamma(k(\beta+1))} E\Bigl(\exp\Bigl(\sum_{n \notin \Lambda} \alpha_n \eta_n\Bigr)\Bigr), & \beta > -1, \\ kK E\bigl(\exp(\eta_0)\bigr)^{k-1} E\Bigl(\exp\Bigl(\sum_{n \notin \Lambda} \alpha_n \eta_n\Bigr)\Bigr), & \beta < -1 \end{cases} \tag{6.6}$$

[cf. equations (7.8) and (7.9) in Rootzén [27]]. It follows from Lemma 7.2 in Rootzén [27] that

$$P(\log \sigma_t > z) = P\Biggl(\sum_{i=0}^{\infty} \alpha_i \eta_{t-i} > z\Biggr) \sim \hat{K} z^{\hat{\beta}} \mathrm{e}^{-z}, \qquad z \to \infty, t \in \mathbb{Z}. \tag{6.7}$$

Since the ε_t, $t \in \mathbb{Z}$, are assumed to be independent of the η_t, $t \in \mathbb{Z}$, the stationary solution to (3.2) implies the existence of a stationary solution $(X_t, \sigma_t)_{t \in \mathbb{Z}}$. For $s > 0$, one can conclude from (6.7) that

$$\lim_{x \to \infty} \frac{P(\sigma_0 > sx)}{P(\sigma_0 > x)} = \lim_{x \to \infty} \frac{(\log(sx))^{\hat{\beta}} (sx)^{-1}}{(\log x)^{\hat{\beta}} x^{-1}} = s^{-1},$$

which shows regular variation with index -1 of the marginal distributions of $(\sigma_t)_{t \in \mathbb{Z}}$.

The regular variation of the marginal distributions of $|X_0| = \sigma_0 |\varepsilon_0|$ and $X_0^+ = \sigma_0 \varepsilon_0^+$ follows by Breiman's lemma, which (under our moment assumptions on ε_0) gives

$$P\bigl(\sigma_0 |\varepsilon_0| > x\bigr) \sim E\bigl(|\varepsilon_0|\bigr) P(\sigma_0 > x) \quad \text{and} \quad P\bigl(\sigma_0 \varepsilon_0^+ > x\bigr) \sim E\bigl(\varepsilon_0^+\bigr) P(\sigma_0 > x).$$

Now the tail balance assertion (3.4) and relation (3.5) between the tails of X_0 and σ_0 are obvious.

The last statement follows from (3.5) in combination with Theorem 7.3 in Rootzén [27]. □

Proof of Theorem 3.5.

Proof of (i). It follows from Theorem 7.4 in Rootzén [27] that the point processes

$$N_n^{\log \sigma}(\cdot) := \sum_{i=1}^{n} \delta_{(i/n, \log \sigma_i - \log n - \hat{\beta} \log(\log n) - \log \hat{K})}(\cdot), \qquad n \in \mathbb{N},$$

converge weakly to a Poisson process $N^{\log \sigma}$ on $[0, 1] \times (-\infty, \infty]$ with intensity measure $\mathrm{d}t \times \mathrm{e}^{-x}\, \mathrm{d}x$. Now, $\exp(\cdot)$ is a continuous function such that the pre-image of a set $B \subset (0, \infty]$ is bounded away from $-\infty$ if B is bounded away from 0. We may thus apply Proposition 5.5 in Resnick [25] to derive that N_n^{σ} converges in $M_p([0, 1] \times (0, \infty])$ to a Poisson process with intensity $\mathrm{d}t \times z^{-2}\, \mathrm{d}z$ on $[0, 1] \times (0, \infty]$ which we denote by N^{σ}.

Proof of (ii). To derive the second assertion, for $n \in \mathbb{N}$, introduce the point process $N_n^{(\sigma,\varepsilon)}$ which consists of the points $(i/n, \sigma_i/a_n, \varepsilon_i)$, $i = 1, 2, \dots, n$. The first assertion implies that $N_n^{(\sigma,\varepsilon)}$ converges weakly to the Poisson process $N^{(\sigma,\varepsilon)}$ with points $(t_{(i)}, z_{(i)}, \varepsilon_{(i)})_{i\in\mathbb{N}}$ which has the intensity $\mathrm{d}t \times z^{-2}\,\mathrm{d}z \times \mathrm{d}P^{\varepsilon_0}$. One may now proceed similarly as in the proof of Proposition 7.5 in Resnick [25] to show that the point processes N_n^X, $n \in \mathbb{N}$, which consist of the points $(i/n, a_n^{-1}\sigma_i \cdot \varepsilon_i)$, $i = 1, 2, \dots, n$, converge to a Poisson point process N^X with points $(t_{(i)}, z_{(i)} \cdot \varepsilon_{(i)})_{i\in\mathbb{N}}$. Note that the additional first component $t_{(i)}$ of the points, the dependence between the σ_i, $i = 1, 2, \dots,$ and the possibly negative sign of the ε_i do not cause substantial changes in the course of the proof. The derivation of the stated intensity $\mathrm{d}t \times z^{-2}[\mathbb{1}_{\{z<0\}}E(\varepsilon_0^-) + \mathbb{1}_{\{z>0\}}E(\varepsilon_0^+)]\,\mathrm{d}z$ of the process N^X follows by an application of Proposition 5.2 in Resnick [25] to the continuous mapping

$$T : [0,1] \times (0,\infty] \times \big((-\infty,\infty) \setminus \{0\}\big) \to D, \qquad (t,x,y) \mapsto (t, x\cdot y)$$

in combination with a truncation argument like in step 4 of the proof of Proposition 7.5 in Resnick [25]. □

6.3. Proofs to Section 4

We start with a technical result on the tail behavior of a product of two factors.

Lemma 6.1. *Let $X, Y \geq 0$ be two independent random variables, such that both X and Y are regularly varying with index -1. Then, for α, β such that $0 \leq \beta < \min\{1, \alpha\}$ and all $\varepsilon > 0$ there exists an $M = M(\varepsilon) > 0$ such that*

$$(1-\varepsilon)\frac{\alpha}{\alpha-\beta}s^{-1/\alpha\mp\varepsilon} \leq \frac{P(X > x, X^\beta Y^\alpha > sx)}{P(X > x)P(Y^\alpha > x^{1-\beta})} \leq (1+\varepsilon)\frac{\alpha}{\alpha-\beta}s^{-1/\alpha\pm\varepsilon} \tag{6.8}$$

for all $s, x > 0$ such that $\min\{x^{1-\beta}, sx^{1-\beta}\} > M$.

Proof. If $\beta = 0$, then the statement follows from the independence of X and Y and by applying the Potter bounds to the function $x \mapsto P(Y^\alpha > x)$ which is regularly varying with index $-1/\alpha$.

In what follows, assume $\beta > 0$. For a start, check that

$$\begin{aligned}
&\frac{P(X > x, X^\beta Y^\alpha > sx)}{P(X > x)P(Y^\alpha > x^{1-\beta})} \\
&\quad= \frac{\int_0^\infty P(X > \max\{x, (sx/y)^{1/\beta}\})P^{Y^\alpha}(\mathrm{d}y)}{P(X > x)P(Y^\alpha > x^{1-\beta})} \\
&\quad= \int_0^{sx^{1-\beta}} \frac{P(X > (sx/y)^{1/\beta})}{P(X > x)P(Y^\alpha > x^{1-\beta})}P^{Y^\alpha}(\mathrm{d}y) \\
&\qquad+ \int_{sx^{1-\beta}}^\infty \frac{P(X > x)}{P(X > x)P(Y^\alpha > x^{1-\beta})}P^{Y^\alpha}(\mathrm{d}y).
\end{aligned} \tag{6.9}$$

The second summand equals

$$\frac{P(Y^\alpha > sx^{1-\beta})}{P(Y^\alpha > x^{1-\beta})} \in \bigl[(1-\varepsilon)s^{-1/\alpha\mp\varepsilon}, (1+\varepsilon)s^{-1/\alpha\pm\varepsilon}\bigr] \tag{6.10}$$

for $\varepsilon > 0$ if $\min\{sx^{1-\beta}, x^{1-\beta}\} > N$ for some $N = N(\varepsilon)$ by the Potter bounds. Again by the Potter bounds, to each $\varepsilon > 0$ there exists $N' = N'(\varepsilon)$ such that for all $x > N'$

$$\begin{aligned}
&\int_0^{sx^{1-\beta}} \frac{P(X > (sx/y)^{1/\beta})}{P(X > x)} P^{Y^\alpha}(\mathrm{d}y) \\
&\quad \le (1+\varepsilon) \int_0^{sx^{1-\beta}} \left(\frac{(sx/y)^{1/\beta}}{x}\right)^{-1+\varepsilon} P^{Y^\alpha}(\mathrm{d}y) \qquad (6.11)\\
&\quad = (1+\varepsilon)\bigl(sx^{1-\beta}\bigr)^{-(1-\varepsilon)/\beta} \int_0^{sx^{1-\beta}} y^{(1-\varepsilon)/\beta} P^{Y^\alpha}(\mathrm{d}y),
\end{aligned}$$

because $(sx/y)^{1/\beta} > x$. Since the distribution of Y^α is regularly varying with index $-1/\alpha$ and $(1-\varepsilon)/\beta > 1/\alpha$ for sufficiently small $\varepsilon > 0$ by assumption, a generalization of Karamata's theorem [Bingham *et al.* [5], Theorem 1.6.4] yields

$$\lim_{t\to\infty} \frac{\int_0^t y^{(1-\varepsilon)/\beta} P^{Y^\alpha}(\mathrm{d}y)}{t^{(1-\varepsilon)/\beta} P(Y^\alpha > t)} = \frac{1/\alpha}{(1-\varepsilon)/\beta - 1/\alpha} = \frac{\beta}{\alpha(1-\varepsilon) - \beta}.$$

Thus, for a suitable $N'' = N''(\varepsilon)$ and $x > N''$, the integral in (6.11) is bounded from above by

$$(1+\varepsilon)\frac{\beta}{\alpha(1-\varepsilon)-\beta}\bigl(sx^{1-\beta}\bigr)^{(1-\varepsilon)/\beta} P\bigl(Y^\alpha > sx^{1-\beta}\bigr).$$

Hence, by (6.10), the first summand in (6.9) is bounded by

$$(1+\varepsilon)^3 \frac{\beta}{\alpha(1-\varepsilon)-\beta} s^{-1/\alpha\pm\varepsilon} \tag{6.12}$$

for x large enough. It follows from (6.10) and (6.12) that for given $\varepsilon > 0$ one may find a suitable $\delta > 0$ and a corresponding constant $M(\delta)$ such that

$$\begin{aligned}
\frac{P(X > x, Y^\alpha X^\beta > sx)}{P(X > x)P(Y^\alpha > x^{1-\beta})} &\le (1+\delta)s^{-1/\alpha\pm\delta} + (1+\delta)^3 \frac{\beta}{\alpha(1-\delta)-\beta} s^{-1/\alpha\pm\delta} \\
&\le (1+\varepsilon)\frac{\alpha}{\alpha-\beta} s^{-1/\alpha\pm\varepsilon},
\end{aligned}$$

for $\min\{sx^{1-\beta}, x^{1-\beta}\} > M(\delta)$ which gives the upper bound in (6.8).

Using the lower Potter bound instead of the upper one and proceeding analogously, we arrive at

$$(1-\varepsilon)^3 s^{-1/\alpha\mp\varepsilon} \frac{\beta}{\alpha(1+\varepsilon)-\beta} \quad \text{and} \quad (1-\varepsilon)s^{-1/\alpha\mp\varepsilon}$$

as lower bounds for the first and second summand in (6.9), respectively, which leads to the lower bound in (6.8). □

Proof of Proposition 4.1. By our assumptions, the solution (κ_1, κ_2) to the linear program (4.3) is unique with $\kappa_1 > 0$ as well as $\kappa_2 > 0$. Therefore, it is impossible that $(\alpha_1 \geq \beta_1, \alpha_2 \geq \beta_2)$ or $(\alpha_1 \leq \beta_1, \alpha_2 \leq \beta_2)$, since this would imply a redundant restriction in (4.3) and thus multiple solutions or $\min\{\kappa_1, \kappa_2\} = 0$. Hence, one of the points (α_i, β_i), $i \in \{1, 2\}$, must lie above or on the main diagonal and one point below or on the main diagonal. W.l.o.g., we may assume that (α_1, β_1) lies below or on the main diagonal, that is, $\alpha_1 \geq \beta_1$. (Otherwise, interchange $(\alpha_1, \beta_1, X_1, \kappa_1)$ and $(\alpha_2, \beta_2, X_2, \kappa_2)$, which leaves the assertion unchanged.)

Next, note that $(\alpha_1 \geq \alpha_2, \beta_1 \geq \beta_2)$ or $(\alpha_1 \leq \alpha_2, \beta_1 \leq \beta_2)$ would imply that a solution with either $\kappa_1 = 0$ or $\kappa_2 = 0$ exists. Furthermore, $\alpha_1 = \alpha_2$ or $\beta_1 = \beta_2$ will always lead to multiple solutions or $\min\{\kappa_1, \kappa_2\} = 0$. Hence, we may conclude that $\min\{\alpha_1, \beta_2\} \geq \max\{\alpha_2, \beta_1\}$ and that $\alpha_1 > \alpha_2$ and $\beta_2 > \beta_1$.

Obviously, the equations

$$\kappa_1\alpha_1 + \kappa_2\alpha_2 = 1, \qquad \kappa_1\beta_1 + \kappa_2\beta_2 = 1 \tag{6.13}$$

must hold, so that

$$\kappa_1 = \frac{\beta_2 - \alpha_2}{\alpha_1\beta_2 - \alpha_2\beta_1}, \qquad \kappa_2 = \frac{\alpha_1 - \beta_1}{\alpha_1\beta_2 - \alpha_2\beta_1}.$$

By the Potter bounds,

$$\frac{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \in \left[(1-\varepsilon)z_1^{1\mp\varepsilon}z_2^{1\mp\varepsilon}, (1+\varepsilon)z_1^{1\pm\varepsilon}z_2^{1\pm\varepsilon}\right] \tag{6.14}$$

for all $\varepsilon > 0$ provided $\min\{x^{\kappa_1}, z_1^{-1}x^{\kappa_1}, x^{\kappa_2}, z_2^{-1}x^{\kappa_2}\} > N(\varepsilon)$ for sufficiently large $N(\varepsilon)$.

Set for abbreviation $\tilde{X}_i := z_iX_i$, $i = 1, 2$, and note that

$$\begin{aligned}
&\frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x)}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\
&\quad = \frac{P(\tilde{X}_1^{\alpha_1}\tilde{X}_2^{\alpha_2} > x, \tilde{X}_1^{\beta_1}\tilde{X}_2^{\beta_2} > x, \tilde{X}_1 \leq x^{\kappa_1}, \tilde{X}_2 \leq x^{\kappa_2})}{P(\tilde{X}_1 > x^{\kappa_1})P(\tilde{X}_2 > x^{\kappa_2})} \\
&\qquad - \frac{P(\tilde{X}_1^{\alpha_1}\tilde{X}_2^{\alpha_2} > x, \tilde{X}_1^{\beta_1}\tilde{X}_2^{\beta_2} > x, \tilde{X}_1 > x^{\kappa_1}, \tilde{X}_2 > x^{\kappa_2})}{P(\tilde{X}_1 > x^{\kappa_1})P(\tilde{X}_2 > x^{\kappa_2})} \\
&\qquad + \frac{P(\tilde{X}_1^{\alpha_1}\tilde{X}_2^{\alpha_2} > x, \tilde{X}_1^{\beta_1}\tilde{X}_2^{\beta_2} > x, \tilde{X}_1 > x^{\kappa_1})}{P(\tilde{X}_1 > x^{\kappa_1})P(\tilde{X}_2 > x^{\kappa_2})} \\
&\qquad + \frac{P(\tilde{X}_1^{\alpha_1}\tilde{X}_2^{\alpha_2} > x, \tilde{X}_1^{\beta_1}\tilde{X}_2^{\beta_2} > x, \tilde{X}_2 > x^{\kappa_2})}{P(\tilde{X}_1 > x^{\kappa_1})P(\tilde{X}_2 > x^{\kappa_2})}.
\end{aligned} \tag{6.15}$$

For $x \geq 1$ the first and second summand equal 0 and -1, respectively, because of (6.13).

Next, note that, again by (6.13),

$$
\begin{aligned}
&P\big(\tilde X_1^{\alpha_1}\tilde X_2^{\alpha_2}>x,\tilde X_1^{\beta_1}\tilde X_2^{\beta_2}>x,\tilde X_1>x^{\kappa_1}\big)\\
&=P\bigg(\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\alpha_1}\tilde X_2^{\alpha_2}>x^{1-\alpha_1\kappa_1},\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\beta_1}\tilde X_2^{\beta_2}>x^{1-\beta_1\kappa_1},\tilde X_1>x^{\kappa_1}\bigg)\\
&=P\bigg(\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\alpha_1}\tilde X_2^{\alpha_2}>x^{\alpha_2\kappa_2},\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\beta_1}\tilde X_2^{\beta_2}>x^{\beta_2\kappa_2},\tilde X_1>x^{\kappa_1}\bigg)\\
&=P\bigg(\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\alpha_1/\alpha_2}\tilde X_2>x^{\kappa_2},\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\beta_1/\beta_2}\tilde X_2>x^{\kappa_2},\tilde X_1>x^{\kappa_1}\bigg),
\end{aligned}
$$

where $(\tilde X_1/x^{\kappa_1})^{\alpha_1/\alpha_2}:=\infty$ if $\alpha_2=0$. According to the above discussion, we have $\beta_1/\beta_2<1<\alpha_1/\alpha_2$. Therefore, the last probability equals

$$
\begin{aligned}
P\bigg(\Big(\frac{\tilde X_1}{x^{\kappa_1}}\Big)^{\beta_1/\beta_2}\tilde X_2>x^{\kappa_2},\tilde X_1>x^{\kappa_1}\bigg)&=P\big(\tilde X_1^{\beta_1/\beta_2}\tilde X_2>x^{\kappa_2+\kappa_1\beta_1/\beta_2},\tilde X_1>x^{\kappa_1}\big)\\
&=P\big(\tilde X_1^{\beta_1\kappa_1}\tilde X_2^{\beta_2\kappa_1}>x^{\kappa_1},\tilde X_1>x^{\kappa_1}\big).
\end{aligned}
$$

Now, we substitute u for x^{κ_1}/z_1, use (6.13) and apply Lemma 6.1 and the Potter bounds to $x\mapsto P(X_2^{\beta_2\kappa_1}>x)$:

$$
\begin{aligned}
&\frac{P(\tilde X_1>x^{\kappa_1},\tilde X_1^{\beta_1\kappa_1}\tilde X_2^{\beta_2\kappa_1}>x^{\kappa_1})}{P(\tilde X_1>x^{\kappa_1})P(\tilde X_2>x^{\kappa_2})}\\
&=\frac{P(z_1X_1>x^{\kappa_1},(z_1X_1)^{\beta_1\kappa_1}(z_2X_2)^{\beta_2\kappa_1}>x^{\kappa_1})}{P(z_1X_1>x^{\kappa_1})P(z_2X_2>x^{\kappa_2})}\\
&=\frac{P(X_1>u,X_1^{\beta_1\kappa_1}X_2^{\beta_2\kappa_1}>z_1^{\beta_2\kappa_2}z_2^{-\beta_2\kappa_1}u)}{P(X_1>u)P(X_2^{\beta_2\kappa_1}>u^{1-\beta_1\kappa_1})}\cdot\frac{P(X_2^{\beta_2\kappa_1}>u^{1-\beta_1\kappa_1})}{P(X_2^{\beta_2\kappa_1}>z_1^{\beta_2\kappa_2}z_2^{-\beta_2\kappa_1}u^{1-\beta_1\kappa_1})}\\
&\in\bigg[(1-\eta)^2\frac{\beta_2}{\beta_2-\beta_1}\bigg(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\bigg)^{\mp2\eta},(1+\eta)^2\frac{\beta_2}{\beta_2-\beta_1}\bigg(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\bigg)^{\pm2\eta}\bigg]
\end{aligned}
$$

if both $u^{1-\beta_1\kappa_1}=(x^{\kappa_1}/z_1)^{\beta_2\kappa_2}$ and $z_1^{\beta_2\kappa_2}z_2^{-\beta_2\kappa_1}u^{1-\beta_1\kappa_1}=(x^{\kappa_2}/z_2)^{\beta_2\kappa_1}$ are larger than some $N(\eta)$, which is the case if both $x^{\kappa_1}/z_1>M(\eta)$ and $x^{\kappa_2}/z_2>M(\eta)$ for a suitably chosen constant $M(\eta)$. Adapting the value of η to each $\varepsilon>0$, we may hence find an $M'(\varepsilon)$ such that

$$
\begin{aligned}
&\frac{P(\tilde X_1^{\alpha_1}\tilde X_2^{\alpha_2}>x,\tilde X_1^{\beta_1}\tilde X_2^{\beta_2}>x,\tilde X_1>x^{\kappa_1})}{P(\tilde X_1>x^{\kappa_1})P(\tilde X_2>x^{\kappa_2})}\\
&\in\bigg[(1-\varepsilon)\frac{\beta_2}{\beta_2-\beta_1}z_1^{\mp\varepsilon}z_2^{\mp\varepsilon},(1+\varepsilon)\frac{\beta_2}{\beta_2-\beta_1}z_1^{\pm\varepsilon}z_2^{\pm\varepsilon}\bigg]
\end{aligned}
\tag{6.16}
$$

if both $x^{\kappa_1}/z_1 > M'(\varepsilon)$ and $x^{\kappa_2}/z_2 > M'(\varepsilon)$.

Analogously, one shows for the fourth summand in equation (6.15) that

$$\frac{P(\tilde{X}_1^{\alpha_1}\tilde{X}_2^{\alpha_2} > x, \tilde{X}_1^{\beta_1}\tilde{X}_2^{\beta_2} > x, \tilde{X}_2 > x^{\kappa_2})}{P(\tilde{X}_1 > x^{\kappa_1})P(\tilde{X}_2 > x^{\kappa_2})} \in \left[(1-\varepsilon)\frac{\alpha_1}{\alpha_1-\alpha_2}z_1^{\mp\varepsilon}z_2^{\mp\varepsilon}, (1+\varepsilon)\frac{\alpha_1}{\alpha_1-\alpha_2}z_1^{\pm\varepsilon}z_2^{\pm\varepsilon}\right] \tag{6.17}$$

if both $x^{\kappa_1}/z_1 > M''(\varepsilon)$ and $x^{\kappa_2}/z_2 > M''(\varepsilon)$ for a suitably chosen $M''(\varepsilon)$.

Finally, for $\varepsilon > 0$, combining (6.14)–(6.18) we arrive at

$$\frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \in \left[(1-\varepsilon')Cz_1^{1\mp\varepsilon'}z_2^{1\mp\varepsilon'}, (1+\varepsilon')Cz_1^{1\pm\varepsilon'}z_2^{1\pm\varepsilon'}\right]$$

for a suitable choice of ε' and a constant $N = N(\varepsilon)$ such that $\min\{x^{\kappa_1}, z_1^{-1}x^{\kappa_1}, x^{\kappa_2}, z_2^{-1}x^{\kappa_2}\} > N$. □

Proof of Theorem 4.2.

1. Almost sure convergence and joint positivity of infinite products

The almost sure convergence of $\prod_{i=1}^k X_i^{\alpha_i}$ (in $\mathbb{R}$) as $k \to \infty$ is equivalent to the almost sure convergence of $\sum_{i=1}^k \alpha_i(\log X_i)^+$ and the latter follows because the series is non-decreasing and $E(\sum_{i=1}^\infty \alpha_i(\log(X_i))^+) = \sum_{i=1}^\infty \alpha_i E(\log X_1)^+ < \infty$ by our assumptions. The almost sure convergence of $\prod_{i=1}^k X_i^{\beta_i}$ follows analogously.

Since we have assumed that $P(Y_i > 0) > 0$ for $i = 0, 1$, there is a positive probability for the event $\{X_i > 0 \text{ for all } i \text{ with } \alpha_i + \beta_i > 0\}$. Thus, the probability

$$P\left(\prod_{i=1}^\infty X_i^{\alpha_i+\beta_i} > 0 \,\middle|\, X_i > 0 \text{ for all } i \text{ with } \alpha_i + \beta_i > 0\right),$$

which can only be 0 or 1 by the 0–1-law, must be equal to 1. Hence, $P(Y_0Y_1 > 0) = P(Y_0 > 0, Y_1 > 0) > 0$ as well.

2. Reduction to a finite-dimensional linear program

In the following, we show that the infinite-dimensional linear program (4.4) is equivalent to the finite-dimensional one

$$\sum_{i=1}^{n^*}\kappa_i \to \min! \qquad \text{under the constraints } \sum_{i=1}^{n^*}\alpha_i\kappa_i \geq 1, \sum_{i=1}^{n^*}\beta_i\kappa_i \geq 1, \kappa_i \geq 0, \forall 1 \leq i \leq n^*, \tag{6.18}$$

for a suitably chosen n^*. To this end, choose $i^*, j^* \in \mathbb{N}_0$ such that $\alpha_{i^*} = \sup_{i\in\mathbb{N}} \alpha_i$ and $\beta_{j^*} = \sup_{j\in\mathbb{N}} \beta_j$. Choose n^* such that $\max\{\alpha_i, \beta_i\} < 1/(2/\alpha_{i^*} + 2/\beta_{j^*})$ for all $i > n^*$. Let $(\kappa_i)_{i\in\mathbb{N}}$ be any feasible solution to (4.4) and suppose that $\kappa_i > 0$ for some $i > n^*$. Then $\tilde{\kappa}_{i^*} := \kappa_{i^*} + \kappa_i\alpha_i/\alpha_{i^*}$, $\tilde{\kappa}_{j^*} := \kappa_{j^*} + \kappa_i\beta_i/\beta_{j^*}$ (resp., $\tilde{\kappa}_{i^*} := \kappa_{i^*} + \kappa_i \max\{\alpha_i/\alpha_{i^*}, \beta_i/\beta_{j^*}\}$ if $i^* = j^*$), $\tilde{\kappa}_i := 0$ and $\tilde{\kappa}_j := \kappa_j$ for all $j \in \mathbb{N} \setminus \{i^*, j^*, i\}$ defines a feasible solution to the constraints (4.4) with $\sum_{i=1}^{\infty} \tilde{\kappa}_i < \sum_{i=1}^{\infty} \kappa_i$. Hence, all optimal solutions $(\kappa_i)_{i\in\mathbb{N}}$ to (4.4) satisfy $\kappa_i = 0$ for all $i > n^*$. They can thus be identified with optimal solutions to the finite-dimensional problem (6.18), and vice versa. It will sometimes be useful in the following that by our definition of n^* we have

$$\max\{\alpha_i, \beta_i\} < \frac{1}{2((1/\alpha_{i^*}) + (1/\beta_{j^*}))} \le \frac{1}{2\sum_{j=1}^{\infty} \kappa_j} \qquad \text{for all } i > n^* \tag{6.19}$$

for any optimal solution $(\kappa_j)_{j\in\mathbb{N}}$, because $\bar{\kappa}_{i^*} := 1/\alpha_{i^*}$, $\bar{\kappa}_{j^*} := 1/\beta_{j^*}$ and $\bar{\kappa}_j = 0$ for all $j \in \mathbb{N} \setminus \{i^*, j^*\}$ (if $i^* \ne j^*$) defines a feasible solution.

A unique solution to (6.18) must be a basic feasible solution [Sierksma [28], Theorems 1.2 and 1.5], that is, $\kappa_l = 0$ for all but at most two indices $l \in \{1, \dots, n^*\}$. Thus, for a unique solution we may assume that $\kappa_1 > 0$ and $\kappa_i = 0$ for all $i \ge 3$ w.l.o.g.

3. A helpful moment bound

Let $m \in \mathbb{N}$, $\varepsilon \in (0, 1)$ and $c_i \in [0, 1 - \varepsilon)$, $i \ge m$, such that $\sum_{i=m}^{\infty} c_i < \infty$. Then

$$\begin{aligned} E\left(\prod_{i=m}^{\infty} X_i^{c_i}\right) &\le \liminf_{n\to\infty} \prod_{i=m}^{n} E\left(X_1^{c_i}\right) \\ &\le \liminf_{n\to\infty} \prod_{i=m}^{n} \left(E\left(X_1^{1-\varepsilon}\right)\right)^{c_i/(1-\varepsilon)} \\ &= \left(E\left(X_1^{1-\varepsilon}\right)\right)^{\sum_{i=m}^{\infty} c_i/(1-\varepsilon)} < \infty \end{aligned} \tag{6.20}$$

by Fatou's lemma and Jensen's inequality for independent and identically distributed $X_i \ge 0$ which are all regularly varying with index -1. [Note that this is the only instance in the proof where we use the assumption that all X_i are identically distributed, cf. Remark 4.3(iii).]

4. Proof of (i)

We prove the assertion by formalizing the heuristic arguments used to motivate the linear optimization problem. Let $(\kappa_i)_{i\in\mathbb{N}}$ be an optimal (not necessarily unique) solution to (4.4) and note that by part 2 of this proof we have $\kappa_i = 0$ for $i > n^*$. By our assumptions, there exists a $\delta > 0$ such that $P(\prod_{i=n^*+1}^{\infty} X_i^{\alpha_i} > \delta, \prod_{i=n^*+1}^{\infty} X_i^{\beta_i} > \delta) =: c > 0$, cf. part 1 of this proof. Fix $\varepsilon > 0$. The lower bound (4.6) follows from

$$\begin{aligned} &P(Y_0 > x, Y_1 > x) \\ &\quad \ge P\left(\prod_{i=n^*+1}^{\infty} X_i^{\alpha_i} > \delta, \prod_{i=n^*+1}^{\infty} X_i^{\beta_i} > \delta\right) P\left(X_i > \left(\delta^{-1}x\right)^{\kappa_i} \text{ for all } 1 \le i \le n^*\right) \\ &\quad \ge c \prod_{i=1}^{n^*} \left(\delta^{-1}x\right)^{-\kappa_i - \varepsilon/n^*} = c\delta^{\sum_{i=1}^{n^*} \kappa_i + \varepsilon} x^{-\sum_{i=1}^{n^*} \kappa_i - \varepsilon} = c\delta^{\sum_{i=1}^{\infty} \kappa_i + \varepsilon} x^{-\sum_{i=1}^{\infty} \kappa_i - \varepsilon} \end{aligned}$$

for sufficiently large x because of the regular variation of the independent random variables X_i.

To establish (4.5), write $Z := \max\{\prod_{i=n^*+1}^\infty X_i^{\alpha_i}, \prod_{i=n^*+1}^\infty X_i^{\beta_i}\}$ and fix $0 < \varepsilon < \sum_{i=1}^\infty \kappa_i$. Note that

$$\begin{aligned} &P(Y_0 > x, Y_1 > x)x^{\sum_{i=1}^\infty \kappa_i - \varepsilon} \\ &\quad \leq \int P\left(\prod_{i=1}^{n^*} X_i^{\alpha_i} > \frac{x}{z}, \prod_{i=1}^{n^*} X_i^{\beta_i} > \frac{x}{z}\right)\left(\frac{x}{z}\right)^{\sum_{i=1}^{n^*}\kappa_i - \varepsilon} z^{\sum_{i=1}^{n^*}\kappa_i - \varepsilon} P^Z(\mathrm{d}z). \end{aligned} \tag{6.21}$$

Because $(\kappa_i)_{1\leq i\leq n^*}$ is an optimal solution to (6.18) we have for all $y \geq 1$

$$\begin{aligned} &P\left(\prod_{i=1}^{n^*} X_i^{\alpha_i} > y, \prod_{i=1}^{n^*} X_i^{\beta_i} > y\right) \\ &\quad \leq P\left(\sum_{i=1}^{n^*} \alpha_i \frac{(\log X_i)^+}{\log y} \geq 1, \sum_{i=1}^{n^*} \beta_i \frac{(\log X_i)^+}{\log y} \geq 1\right) \\ &\quad \leq P\left(\sum_{i=1}^{n^*} \frac{(\log X_i)^+}{\log y} \geq \sum_{i=1}^{n^*} \kappa_i\right) \\ &\quad = P\left(\prod_{i=1}^{n^*} \max(X_i, 1) > y^{\sum_{i=1}^{n^*}\kappa_i}\right). \end{aligned}$$

This expression is a regularly varying function of y with index $-\sum_{i=1}^{n^*}\kappa_i$ by the corollary to Theorem 3 in Embrechts and Goldie [14]. Thus, the integrand in (6.21) tends to 0 for all $z > 0$ as $x \to \infty$. Furthermore, we have

$$E\left(Z^{\sum_{i=1}^{n^*}\kappa_i - \varepsilon}\right) \leq E\left(\prod_{i=n^*+1}^{\infty} X_i^{\alpha_i(\sum_{i=1}^{n^*}\kappa_i - \varepsilon)}\right) + E\left(\prod_{i=n^*+1}^{\infty} X_i^{\beta_i(\sum_{i=1}^{n^*}\kappa_i - \varepsilon)}\right) < \infty$$

by (6.19) and (6.20). Thus, $Mz^{\sum_{i=1}^{n^*}\kappa_i - \varepsilon}$ is an integrable majorant to the integrand in (6.21) for sufficiently large $M > 0$. Hence, $P(Y_0 > x, Y_1 > x) = o(x^{\varepsilon - \sum_{i=1}^\infty \kappa_i})$ follows by dominated convergence.

5. Proof of (ii)

In the following, we assume a unique solution to the optimization problem (4.4) with $\kappa_i > 0, \kappa_j > 0$. W.l.o.g. let us assume that $i = 1$, $j = 2$, thus $\kappa_1 > 0, \kappa_2 > 0$ and $\kappa_l = 0$, $l \geq 3$. Moreover, w.l.o.g. we may assume $\alpha_1 > \alpha_2$, $\beta_2 > \beta_1$ and $\min\{\alpha_1, \beta_2\} \geq \max\{\alpha_2, \beta_1\}$ (cf. the proof of Proposition 4.1).

As in the proof of Proposition 4.1, equation (6.13) holds and it follows that

$$\kappa_1 = \frac{\beta_2 - \alpha_2}{\alpha_1\beta_2 - \alpha_2\beta_1}, \qquad \kappa_2 = \frac{\alpha_1 - \beta_1}{\alpha_1\beta_2 - \alpha_2\beta_1}.$$

There exists a so-called dual problem to (6.18) [see Sierksma [28], Chapter 2] that is given by

$$\begin{aligned} &\hat{\kappa}_1 + \hat{\kappa}_2 \to \max! \\ &\quad \text{under the constraints } \alpha_l \hat{\kappa}_1 + \beta_l \hat{\kappa}_2 \le 1 \text{ for all } l = 1, \dots, n^*, \hat{\kappa}_1 \ge 0, \hat{\kappa}_2 \ge 0. \end{aligned} \tag{6.22}$$

Since we have assumed an optimal solution to the primal problem (6.18), there also exists an optimal solution to the dual problem [Sierksma [28], Theorem 2.2], denoted by $(\hat{\kappa}_1, \hat{\kappa}_2)$ with

$$\hat{\kappa}_1 + \hat{\kappa}_2 = \kappa_1 + \kappa_2. \tag{6.23}$$

Because of $\kappa_1, \kappa_2 > 0$, it follows by the complementary slackness theorem [Sierksma [28], Theorem 2.4] that

$$\alpha_1 \hat{\kappa}_1 + \beta_1 \hat{\kappa}_2 = 1 \quad \text{and} \quad \alpha_2 \hat{\kappa}_1 + \beta_2 \hat{\kappa}_2 = 1.$$

Therefore,

$$\hat{\kappa}_1 = \frac{\beta_2 - \beta_1}{\alpha_1 \beta_2 - \alpha_2 \beta_1} > 0, \qquad \hat{\kappa}_2 = \frac{\alpha_1 - \alpha_2}{\alpha_1 \beta_2 - \alpha_2 \beta_1} > 0. \tag{6.24}$$

If equality held in the constraints of (6.22) for some $l \in \{3, \dots, n^*\}$, then direct calculations show that

$$\tilde{\kappa}_1 = \frac{\beta_l - \alpha_l}{\alpha_1 \beta_l - \alpha_l \beta_1}, \qquad \tilde{\kappa}_l = \frac{\alpha_1 - \beta_1}{\alpha_1 \beta_l - \alpha_l \beta_1}, \qquad \tilde{\kappa}_i = 0, \forall i \in \{2, \dots, n^*\} \setminus \{l\}$$

would be another optimal solution to the primal problem in contradiction to our assumptions. Therefore, there exists an $\varepsilon > 0$ such that

$$\alpha_l \hat{\kappa}_1 + \beta_l \hat{\kappa}_2 < 1 - 2\varepsilon \qquad \text{for all } l \in \{3, \dots, n^*\}. \tag{6.25}$$

Let

$$Z_1(x_3, x_4, \dots) := \prod_{i=3}^{\infty} x_i^{(\alpha_i \beta_2 - \beta_i \alpha_2)/(\alpha_1 \beta_2 - \alpha_2 \beta_1)},$$

$$Z_2(x_3, x_4, \dots) := \prod_{i=3}^{\infty} x_i^{(\beta_i \alpha_1 - \alpha_i \beta_1)/(\alpha_1 \beta_2 - \alpha_2 \beta_1)}$$

if $x_i > 0$ for all $i \ge 3$ with $\max\{\alpha_i, \beta_i\} > 0$ and set $Z_1(x_3, x_4, \dots) = Z_2(x_3, x_4, \dots) = \infty$ else. These two quantities will play the role of z_1 and z_2 in the statement and proof of Proposition 4.1. For $\delta > 0$, define $M = M(\delta)$ as in Proposition 4.1,

$$\mathbb{M} = \mathbb{M}(x, M) := \{\mathbf{x} \in [0, \infty)^{\mathbb{N}} : M Z_1(\mathbf{x}) < x^{\kappa_1}, M Z_2(\mathbf{x}) < x^{\kappa_2}\}$$

and $\mathbf{X}_3^{\infty} := (X_3, X_4, \dots)$.

We start with showing (ii) for $s_0 = s_1 = 1$.

$$\begin{aligned}
&\frac{P(\prod_{i=1}^\infty X_i^{\alpha_i} > x, \prod_{i=1}^\infty X_i^{\beta_i} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\
&\quad = \int_{\mathbb{M}} \frac{P((Z_1(\mathbf{x})X_1)^{\alpha_1}(Z_2(\mathbf{x})X_2)^{\alpha_2} > x, (Z_1(\mathbf{x})X_1)^{\beta_1}(Z_2(\mathbf{x})X_2)^{\beta_2} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} P^{\mathbf{X}_3^\infty}(\mathrm{d}\mathbf{x}) \qquad (6.26)\\
&\qquad + \frac{P(\prod_{i=1}^\infty X_i^{\alpha_i} > x, \prod_{i=1}^\infty X_i^{\beta_i} > x, \mathbf{X}_3^\infty \in \mathbb{M}^c)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})}.
\end{aligned}$$

According to Proposition 4.1 and (6.24), the integrand of the first summand on the right-hand side of (6.26) is bounded from above by

$$(1+\delta)CZ_1(\mathbf{x})^{1\pm\delta}Z_2(\mathbf{x})^{1\pm\delta} \le (1+\delta)C\prod_{i=3}^\infty x_i^{(1\pm\delta)(\alpha_i\hat{\kappa}_1+\beta_i\hat{\kappa}_2)}$$

for $x > M$ with C as in Proposition 4.1. Thus, if we chose δ small enough then all the exponents in the last expression are less than $1-\varepsilon$ (by (6.25) for $3 \le i \le n^*$ and by (6.19) and (6.23) for $i > n^*$) and the upper bound is integrable w.r.t. $P^{\mathbf{X}_3^\infty}$ by (6.20). Therefore, by dominated convergence, we see that

$$\begin{aligned}
&\lim_{x\to\infty}\int_{\mathbb{M}} \frac{P((Z_1(\mathbf{x})X_1)^{\alpha_1}(Z_2(\mathbf{x})X_2)^{\alpha_2} > x, (Z_1(\mathbf{x})X_1)^{\beta_1}(Z_2(\mathbf{x})X_2)^{\beta_2} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} P^{\mathbf{X}_3^\infty}(\mathrm{d}\mathbf{x}) \\
&\quad = \int_{\mathbb{M}} \lim_{x\to\infty} \frac{P((Z_1(\mathbf{x})X_1)^{\alpha_1}(Z_2(\mathbf{x})X_2)^{\alpha_2} > x, (Z_1(\mathbf{x})X_1)^{\beta_1}(Z_2(\mathbf{x})X_2)^{\beta_2} > x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} P^{\mathbf{X}_3^\infty}(\mathrm{d}\mathbf{x}) \qquad (6.27)\\
&\quad = \int \prod_{i=3}^\infty CZ_1(\mathbf{x})Z_2(\mathbf{x})\mathbb{1}_{\{\max\{Z_1(\mathbf{x}),Z_2(\mathbf{x})\}<\infty\}} P^{\mathbf{X}_3^\infty}(\mathrm{d}\mathbf{x}) \\
&\quad = CE\left(\prod_{i=3}^\infty X_i^{\alpha_i\hat{\kappa}_1+\beta_i\hat{\kappa}_2}\right) = D \in (0,\infty),
\end{aligned}$$

where in the last step we have used that $\hat{\kappa}_1, \hat{\kappa}_2 > 0$ and thus the product vanishes if $Z_1(\mathbf{x})$ or $Z_2(\mathbf{x})$ is infinite.

It remains to be shown that

$$\begin{aligned}
&\frac{P(\prod_{i=1}^\infty X_i^{\alpha_i} > x, \prod_{i=1}^\infty X_i^{\beta_i} > x, \mathbf{X}_3^\infty \in \mathbb{M}^c)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\
&\quad \le \frac{P(\prod_{i=1}^\infty X_i^{\alpha_i} > x, \prod_{i=1}^\infty X_i^{\beta_i} > x, MZ_1(\mathbf{X}_3^\infty) \ge x^{\kappa_1})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \qquad (6.28)\\
&\qquad + \frac{P(\prod_{i=1}^\infty X_i^{\alpha_i} > x, \prod_{i=1}^\infty X_i^{\beta_i} > x, MZ_2(\mathbf{X}_3^\infty) \ge x^{\kappa_2})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})}
\end{aligned}$$

tends to 0. We will only examine the first term on the right-hand side, as the second term can be treated analogously.

The numerator of the first term can be bounded by

$$\begin{aligned} &P\left(\prod_{i=1}^{\infty} X_i^{\beta_i} > x,\, M^{\alpha_1\beta_2-\alpha_2\beta_1}\prod_{i=3}^{\infty} X_i^{\alpha_i\beta_2-\beta_i\alpha_2} \geq x^{\beta_2-\alpha_2}\right) \\ &\quad \leq P\left(\prod_{i=1}^{\infty} X_i^{\beta_i} > x,\, M^{\alpha_1\beta_2-\alpha_2\beta_1}\prod_{i=1}^{\infty} X_i^{\beta_i\alpha_2}\prod_{i=3}^{\infty} X_i^{\alpha_i\beta_2-\beta_i\alpha_2} \geq x^{\beta_2}\right) \qquad (6.29) \\ &\quad \leq P\left(\prod_{i=1}^{\infty} X_i^{\beta_i} > \min\{1, M^{(\alpha_2\beta_1-\alpha_1\beta_2)/\beta_2}\}x,\, X_1^{\alpha_2\beta_1/\beta_2}\prod_{i=2}^{\infty} X_i^{\alpha_i} \geq \min\{1, M^{(\alpha_2\beta_1-\alpha_1\beta_2)/\beta_2}\}x\right). \end{aligned}$$

By assertion (i), the asymptotic behavior of the above probability is related to the linear program

$$\begin{aligned} &\sum_{i=1}^{\infty}\tilde{\kappa}_i \to \min! \\ &\text{under the constraints } (\alpha_2\beta_1/\beta_2)\tilde{\kappa}_1 + \sum_{i=2}^{\infty}\alpha_i\tilde{\kappa}_i \geq 1,\, \sum_{i=1}^{\infty}\beta_i\tilde{\kappa}_i \geq 1,\, \tilde{\kappa}_i \geq 0,\, \forall i \in \mathbb{N}. \end{aligned} \qquad (6.30)$$

Since $\alpha_2\beta_1/\beta_2 < \alpha_1$, all optimal solutions $(\tilde{\kappa}_i)_{i\in\mathbb{N}}$ to (6.30) are also feasible solutions to (4.4), but they cannot be optimal solutions to (4.4) because the unique optimal solution $(\kappa_i)_{i\in\mathbb{N}}$ is not a feasible solution of (6.30). Hence, $\sum_{i=1}^{\infty}\tilde{\kappa}_i > \kappa_1 + \kappa_2 + 2\delta$ for sufficiently small $\delta > 0$ and, by assertion (i), the right-hand side of (6.29) is of smaller order than $x^{-(\kappa_1+\kappa_2+\delta)}$, while the denominator of the first term of the right-hand side of (6.28) is of larger order than $x^{-(\kappa_1+\kappa_2+\delta)}$. Thus, the first summand in (6.28) tends to 0. The second summand can be treated analogously, which concludes the proof in the case $s_0 = s_1 = 1$.

Finally, we prove the assertion (ii) for general $s_0, s_1 > 0$. Note that the above value of D in (6.27) does not depend on the distribution of X_1 or X_2. Therefore, we may replace X_1 and X_2 with z_1X_1 and z_2X_2, respectively, where

$$z_1 := s_1^{\alpha_2/(\alpha_1\beta_2-\alpha_2\beta_1)} s_0^{-\beta_2/(\alpha_1\beta_2-\alpha_2\beta_1)}, \qquad z_2 := s_0^{\beta_1/(\alpha_1\beta_2-\alpha_2\beta_1)} s_1^{-\alpha_1/(\alpha_1\beta_2-\alpha_2\beta_1)}, \qquad (6.31)$$

such that $z_1^{\alpha_1}z_2^{\alpha_2} = s_0^{-1}$ and $z_1^{\beta_1}z_2^{\beta_2} = s_1^{-1}$. As this substitution does not alter the solution to the linear program (4.4),

$$\begin{aligned} &\frac{P(Y_0 > s_0x, Y_1 > s_1x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\ &\quad = \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2}\prod_{i=3}^{\infty}X_i^{\alpha_i} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2}\prod_{i=3}^{\infty}X_i^{\beta_i} > x)}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\ &\qquad \times \frac{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \end{aligned}$$

converges to

$$Dz_1z_2 = Ds_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)}s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)}$$

as $x \to \infty$ by the first part of the proof and the regular variation of X_1 and X_2.

6. Proof of (iii)

In the following, we assume that a unique solution $(\kappa_i)_{i\in\mathbb{N}}$ to (4.4) exists with $\kappa_1 > 0$ and $\kappa_j = 0$ for all $j \geq 2$. Furthermore, assume w.l.o.g. that $\alpha_1 \geq \beta_1 > 0$, thus $\kappa_1 = 1/\beta_1$. In this case,

$$\begin{aligned} &P(Y_0 > s_0x, Y_1 > s_1x) \\ &= P\left(X_1^{\alpha_1}\prod_{i=2}^{\infty}X_i^{\alpha_i} > s_0x, X_1^{\beta_1}\prod_{i=2}^{\infty}X_i^{\beta_i} > s_1x\right) \\ &= P\left(X_1^{\beta_1}\left(s_0^{-1}\prod_{i=2}^{\infty}X_i^{\alpha_i}\right)^{\beta_1/\alpha_1}x^{1-\beta_1/\alpha_1} > x, X_1^{\beta_1}s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i} > x\right) \\ &= P\left(X_1^{\beta_1}\min\left\{\left(s_0^{-1}\prod_{i=2}^{\infty}X_i^{\alpha_i}\right)^{\beta_1/\alpha_1}x^{1-\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i}\right\} > x\right). \end{aligned} \tag{6.32}$$

Now, if $\alpha_1 = \beta_1$, then the second factor in the product in (6.32) equals $Z := \min\{s_0^{-1}\prod_{i=2}^{\infty}X_i^{\alpha_i}, s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i}\}$. Note that

$$\sum_{i=2}^{\infty}\kappa_i' > \kappa_1 = \frac{1}{\beta_1} \qquad \text{for all } \kappa_i' \geq 0, i \geq 2, \qquad \text{such that } \sum_{i=2}^{\infty}\alpha_i\kappa_i' \geq 1, \qquad \sum_{i=2}^{\infty}\beta_i\kappa_i' \geq 1$$

since we have assumed a unique solution to (4.4). Part (i) of the statement (applied to $\tilde{Y}_0 := \prod_{i=2}^{\infty}X_i^{\alpha_i}$ and $\tilde{Y}_1 := \prod_{i=2}^{\infty}X_i^{\beta_i}$) thus yields that $P(Z > x) = o(x^{-1/\beta_1-\varepsilon})$ for some $\varepsilon > 0$, which by Breiman's lemma in turn implies the first case in (4.8).

If $\alpha_1 > \beta_1$, then $\beta_i < \beta_1$ for all $i \geq 2$, since otherwise $\kappa_1 = 1/\beta_1, \kappa_i = 0, i \geq 2$, would not be a unique solution to (4.4). (If $\beta_i \geq \beta_1$, then $\kappa_1' = (1-\varepsilon)/\beta_1$, $\kappa_i' = \varepsilon/\beta_i$, for sufficiently small ε, and $\kappa_j' = 0$ for all $j \in \mathbb{N}\setminus\{1, i\}$ would satisfy (4.4) with $\kappa_1' + \kappa_i' \leq \kappa_1$.) For all $C > 0$, the second factor in the product in (6.32) is eventually (for sufficiently large x) bounded from below by $\min\{C\prod_{i=2}^{\infty}X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i}\}$. Thus,

$$\begin{aligned} &P\left(X_1^{\beta_1}\min\left\{C\prod_{i=2}^{\infty}X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i}\right\} > x\right) \\ &\quad\leq P(Y_0 > s_0x, Y_1 > s_1x) \leq P\left(X_1^{\beta_1}\mathbb{1}_{\{\prod_{i=2}^{\infty}X_i^{\alpha_i\beta_1/\alpha_1} > 0\}}s_1^{-1}\prod_{i=2}^{\infty}X_i^{\beta_i} > x\right) \end{aligned}$$

for x large enough. For both the left-hand side and the right-hand side of this inequality, the random variable $X_1^{\beta_1}$ is multiplied with an independent random variable bounded above by

$s_1^{-1}\prod_{i=2}^{\infty} X_i^{\beta_i}$. Because $\beta_i < \beta_1$ for all $i \geq 2$ we have $E((\prod_{i=2}^{\infty} X_i^{\beta_i})^{(1+\varepsilon)/\beta_1}) < \infty$ for sufficiently small $\varepsilon > 0$ by (6.20). We may thus again apply Breiman's lemma to get

$$\begin{aligned}
&E\left(\left(\min\left\{C\prod_{i=2}^{\infty} X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^{\infty} X_i^{\beta_i}\right\}\right)^{1/\beta_1}\right) \\
&\quad\leq \liminf_{x\to\infty} \frac{P(Y_1 > s_0 x, Y_1 > s_1 x)}{P(X_1^{\beta_1} > x)} \leq \limsup_{x\to\infty} \frac{P(Y_1 > s_0 x, Y_1 > s_1 x)}{P(X_1^{\beta_1} > x)} \\
&\quad\leq E\left(\left(\mathbb{1}_{\{\prod_{i=2}^{\infty} X_i^{\alpha_i\beta_1/\alpha_1} > 0\}} s_1^{-1}\prod_{i=2}^{\infty} X_i^{\beta_i}\right)^{1/\beta_1}\right) < \infty.
\end{aligned}$$

For $C \to \infty$, the lower bound converges to the upper bound, which yields the second case in (4.8). Analogous arguments with the role of α_1 and β_1 interchanged conclude the proof of part (iii) of the statement. □

Proof of Theorem 4.5. By assumption, $X_j \geq \delta$ for some $\delta \in (0, 1]$ and all $1 \leq j \leq n$, and thus $\alpha_{ij}\log X_j \leq \alpha_{ij}(\log X_j)^+ - \alpha_{ij}^-\log\delta$. Let $c := \max_{1\leq i\leq d}\sum_{j=1}^n \alpha_{ij}^-|\log\delta|$. Then, similarly as in the proof of part (i) of Theorem 4.2, the optimality of $(\kappa_j)_{1\leq j\leq n}$ implies

$$\begin{aligned}
P\left(\prod_{j=1}^n X_j^{\alpha_{ij}} > y, \forall 1 \leq i \leq d\right) &\leq P\left(\sum_{j=1}^n \alpha_{ij}\frac{(\log X_j)^+}{\log y - c} \geq 1, \forall 1 \leq i \leq d\right) \\
&= P\left(\prod_{j=1}^n \max(X_j, 1) \geq \left(y e^{-c}\right)^{\sum_{j=1}^n \kappa_j}\right) \\
&= \mathrm{o}\left(y^{\varepsilon - \sum_{j=1}^n \kappa_j}\right)
\end{aligned}$$

for all $\varepsilon > 0$ and $y > e^c$. The lower bound can be established in a similar way as in the proof of Theorem 4.2.

To prove the second assertion, first note that if a unique optimal solution of the given form exists, the equation $\mathbf{A}\mathbf{x} = \mathbf{1}$ must have $(\kappa_j)_{j\in J} = \mathbf{A}^{-1}\mathbf{1}$ as the unique solution. Let $\tilde{X}_j := X_j\prod_{i=1}^d s_i^{-A_{ji}^{-1}}$ for $j \in J$, $Z_i := \prod_{k\notin J} X_k^{\alpha_{ik}}$ for $1 \leq i \leq d$ and $\mathbb{M} := \{\mathbf{z} \in [0,\infty)^d \mid \prod_{i=1}^d z_i^{A_{ji}^{-1}} \leq M x^{\kappa_j}\ \forall j \in J\}$ for a sufficiently large M (to be specified later). Then

$$\begin{aligned}
&P(Y_i > s_i x\ \forall 1 \leq i \leq d) \\
&\quad= \int_{\mathbb{M}} P\left(\prod_{j\in J}\tilde{X}_j^{\alpha_{ij}} > x/z_i\ \forall 1 \leq i \leq d\right) P^{(Z_i)_{1\leq i\leq d}}(\mathrm{d}\mathbf{z}) \\
&\qquad+ P\left(Y_i > s_i x\ \forall 1 \leq i \leq d, (Z_i)_{1\leq i\leq d} \notin \mathbb{M}\right).
\end{aligned} \tag{6.33}$$

To show that the last summand is negligible, it suffices to prove that

$$P\left(Y_i > s_i x \ \forall 1 \le i \le d, \prod_{i=1}^{d} Z_i^{A_{li}^{-1}} > M x^{\kappa_l}\right) = \mathrm{o}\big(x^{-\sum_{j=1}^{n}\kappa_j - \delta}\big) \tag{6.34}$$

as $x \to \infty$ for some $\delta > 0$ and all $l \in J$. Now,

$$\begin{aligned} &P\left(Y_i > s_i x \ \forall 1 \le i \le d, \prod_{i=1}^{d} Z_i^{A_{li}^{-1}} > M x^{\kappa_l}\right) \\ &\quad = P\left(\prod_{j=1}^{n} X_j^{\alpha_{ij}} > s_i x \ \forall 1 \le i \le d, \prod_{k \notin J} X_k^{\sum_{i=1}^{d} \alpha_{ik} A_{li}^{-1}/\kappa_l} > M^{1/\kappa_l} x\right). \end{aligned}$$

Obviously, a feasible solution to the corresponding linear program

$$\begin{aligned} &\sum_{j=1}^{n} \rho_j \to \min! \\ &\quad \text{under the constraints } \sum_{j=1}^{n} \alpha_{ij}\rho_j \ge 1, \forall 1 \le i \le d, \\ &\qquad \sum_{k \notin J} \sum_{i=1}^{d} \frac{\alpha_{ik} A_{li}^{-1}}{\kappa_l} \rho_k \ge 1, \rho_j \ge 0, \forall 1 \le j \le n, \end{aligned} \tag{6.35}$$

is also a feasible solution of the original linear program. However, as the unique solution $(\kappa_j)_{1 \le j \le n}$ to the original linear program does not satisfy the second constraint in (6.35), an optimal solution $(\rho_j)_{1 \le j \le n}$ to the above linear program fulfills $\sum_{j=1}^{n} \rho_j > \sum_{j=1}^{n} \kappa_j$. Therefore, the first assertion (applied to the products $Y_1, \ldots, Y_d$ and $\prod_{i=1}^{d} Z_i^{A_{li}^{-1}/\kappa_l}$) implies (6.34).

The distribution of the random vector $(\tilde{X}_j)_{j \in J}$ is regularly varying on $\mathbb{E}^d$ in the sense of Lindskog, Resnick and Roy [21], Definition 3.2, w.r.t. the "multiplication" $(x, (t_j)_{j \in J}) \mapsto (x^{\kappa_j} t_j)_{j \in J}$, because

$$\frac{P((\tilde{X}_j)_{j \in J} \in \times_{j \in J} (x^{\kappa_j} t_j, \infty))}{P((\tilde{X}_j)_{j \in J} \in \times_{j \in J} (x^{\kappa_j}, \infty))} \to \prod_{j \in J} t_j^{-1} =: \nu\left(\mathop{\times}_{j \in J} (t_j, \infty)\right) \tag{6.36}$$

for all $t_j > 0$, $j \in J$ (cf. Example 3.1 of that paper).

By our assumptions, $\mathbf{A}^{-1}(\mathbb{E}^d) \subset \mathbb{E}^d$, so that

$$\exp\left(\mathbf{A}^{-1}\left(\mathop{\times}_{i=1}^{d} (-\log z_i, \infty)\right)\right) \subset \exp\big(\big(-\mathbf{A}^{-1}(\log z_i)_{1 \le i \le d}, \infty\big)\big) \tag{6.37}$$

which is bounded away from the boundary of $\mathbb{E}^d$ for fixed $(z_i)_{1\le i\le d}$. Hence, for the integrand of the first summand in (6.33) we obtain

$$\begin{aligned}
&\frac{P(\prod_{j\in J}\tilde{X}_j^{\alpha_{ij}} > x/z_i\ \forall 1\le i\le d)}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}\\
&\quad=\frac{P(\mathbf{A}(\log\tilde{X}_j)_{j\in J}\in \bigtimes_{i=1}^d(\log(x/z_i),\infty))}{P((\tilde{X}_j)_{j\in J}\in \bigtimes_{j\in J}(x^{\kappa_j},\infty))}\\
&\quad=\frac{P((\tilde{X}_j)_{j\in J}\in (x^{\kappa_j})_{j\in J}\exp(\mathbf{A}^{-1}(\bigtimes_{i=1}^d(-\log z_i,\infty))))}{P((\tilde{X}_j)_{j\in J}\in \bigtimes_{j\in J}(x^{\kappa_j},\infty))}\\
&\quad\to \nu\Bigg(\exp\Bigg(\mathbf{A}^{-1}\Bigg(\bigtimes_{i=1}^{d}(-\log z_i,\infty)\Bigg)\Bigg)\Bigg),
\end{aligned}\tag{6.38}$$

where by convention $(x^{\kappa_j})_{j\in J}B=\{(x^{\kappa_j}y_j)_{j\in J}|(y_j)_{j\in J}\in B\}$ for $B\subset\mathbb{E}^d$ and the measure ν is defined by (6.36). If we show that there exists an integrable majorant to (6.38), then by (6.33), (6.34) and dominated convergence

$$\frac{P(Y_i > s_i x\ \forall 1\le i\le d)}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}\to\int \nu\Bigg(\exp\Bigg(\mathbf{A}^{-1}\Bigg(\bigtimes_{i=1}^{d}(-\log z_i,\infty)\Bigg)\Bigg)\Bigg)P^{(Z_i)_{1\le i\le d}}(\mathrm{d}\mathbf{z}).$$

Direct calculations show that this limit equals

$$\frac{1}{|\det\mathbf{A}|}E\Bigg(\prod_{i=1}^{d}\frac{Z_i^{\sum_{j\in J}A_{ji}^{-1}}}{\sum_{j\in J}A_{ji}^{-1}}\Bigg)=D$$

and thus (4.10) follows with

$$\lim_{x\to\infty}\frac{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}{\prod_{j\in J}P(X_j > x^{\kappa_j})}=\prod_{i=1}^{d}s_i^{-\sum_{j\in J}A_{ji}^{-1}}.$$

It remains to define an integrable majorant to (6.38). As shown above [cf. (6.37)],

$$\begin{aligned}
&\frac{P(\prod_{j\in J}\tilde{X}_j^{\alpha_{ij}} > x/z_i\ \forall 1\le i\le d)}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}\\
&\quad=\frac{P((\tilde{X}_j)_{j\in J}\in\exp(\mathbf{A}^{-1}(\bigtimes_{i=1}^d(\log(x/z_i),\infty))))}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}\\
&\quad\le\frac{P((\tilde{X}_j)_{j\in J}\in (x^{\kappa_j})_{j\in J}\exp((\mathbf{A}^{-1}(-\log z_i)_{1\le i\le d},\infty)))}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})}.
\end{aligned}$$

Hence, for $\varepsilon > 0$ the Potter bounds yield

$$\frac{P(\prod_{j\in J}\tilde{X}_j^{\alpha_{ij}} > x/z_i\ \forall 1\le i\le d)}{\prod_{j\in J}P(\tilde{X}_j > x^{\kappa_j})} \le \prod_{j\in J}\frac{P(\tilde{X}_j > x^{\kappa_j}\prod_{i=1}^d z_i^{-A_{ji}^{-1}})}{P(\tilde{X}_j > x^{\kappa_j})}$$
$$\le (1+\varepsilon)\left(\prod_{j\in J}\prod_{i=1}^d z_i^{A_{ji}^{-1}}\right)^{1\pm\varepsilon}$$

for $\mathbf{z} \in \mathbb{M}$ if M is chosen sufficiently large. Since

$$\prod_{j\in J}\prod_{i=1}^d Z_i^{A_{ji}^{-1}} = \prod_{k\notin J} X_k^{\sum_{j\in J}\sum_{i=1}^d \alpha_{ik}A_{ji}^{-1}},$$

and the $\tilde{X}_k$ are independent and regularly varying with index -1, it suffices to show that $\sum_{j\in J}\sum_{i=1}^d \alpha_{ik}A_{ji}^{-1} < 1$ for all $k \notin J$. As in the proof of Theorem 4.2, the strong complementary slackness theorem implies that the corresponding dual problem

$$\sum_{i=1}^d \hat{\kappa}_i \to \max!$$

$$\text{under the constraints } \sum_{i=1}^d \alpha_{ij}\hat{\kappa}_i \le 1, \forall 1\le j\le n, \hat{\kappa}_i \ge 0, \forall 1\le i\le d,$$

has a solution that satisfies $\sum_{i=1}^d \alpha_{ij}\hat{\kappa}_i = 1$ for all $j \in J$, that is, $\hat{\kappa}_i = \sum_{j\in J} A_{ji}^{-1}$ for all $1 \le i \le d$, and $\sum_{j\in J}\sum_{i=1}^d \alpha_{ik}A_{ji}^{-1} = \sum_{i=1}^d \alpha_{ik}\hat{\kappa}_i < 1$ for all $k \notin J$, which concludes the proof. □

6.4. Proofs to Section 5

Proof of Theorem 5.1. Note that if a feasible solution $(\kappa_i)_{i\in\mathbb{N}_0}$ of (5.1) satisfies $\kappa_i > 0$ for exactly one index $i \in \mathbb{N}_0$ and $\alpha_i > \alpha_{i-h}$, then $\tilde{\kappa}_i := 1/\alpha_i$, $\tilde{\kappa}_{i+h} := (\alpha_i - \alpha_{i-h})/\alpha_i^2$ and $\tilde{\kappa}_j := 0$ for all other $j \in \mathbb{N}_0$ defines a feasible solution, too, with $\tilde{\kappa}_i + \tilde{\kappa}_{i+h} < 1/\alpha_{i-h} \le \kappa_i$. Hence, the first feasible solution cannot be optimal, which (together with an analogous argument in the case $\alpha_i < \alpha_{i-h}$) proves that for an optimal solution necessarily $\alpha_i = \alpha_{i-h}$ (and thus $i \ge h$), if $i \in \mathbb{N}_0$ is the only index with $\kappa_i > 0$.

Define coefficients $\hat{\alpha}_i := \alpha_{i-h}$ and $\hat{\beta}_i := \alpha_i$ (with the convention $\alpha_k = 0$ for $k < 0$) and let $\hat{X}_i := \mathrm{e}^{\eta_{h-i}}$, $i \in \mathbb{N}_0$, so that $(\sigma_0, \sigma_h) = (\prod_{i=0}^\infty \hat{X}_i^{\hat{\alpha}_i}, \prod_{i=0}^\infty \hat{X}_i^{\hat{\beta}_i})$. The random variables $\hat{X}_i$ are regularly varying with index -1, and the coefficients $\hat{\alpha}_i$ and $\hat{\beta}_i$ satisfy the conditions of Theorem 4.2 (where we start from the index $i = 0$ instead of $i = 1$). Moreover, by Remark 4.3(ii) and the assumption that $E(\eta_0^2) < \infty$, it follows that $\sigma_0, \sigma_h > 0$ almost surely. The statement about

(σ_0, σ_h) thus follows by an application of Theorem 4.2, cases (ii) and (iii), in combination with Remark 4.4(i)–(iii).

In particular, (σ_0, σ_h) is regularly varying on $\mathbb{E}^2$ with index $-\sum_{i=1}^{\infty}\kappa_i$. Hence, ε_0 fulfills the moment condition of Corollary 2.5, which yields the assertion on (X_0, X_h). □

Proof of Corollary 5.3. The stated unique form of the solution is immediate from the assumed strict monotonicity of the coefficients. The form of the limit measure then follows from Theorem 5.1(i) with $i = 0$, $j = h$ and $\alpha_0 = 1$. □

Proof of Theorem 5.4. Let $\alpha_{2m(i-1)} = 1$ and $\alpha_{2m(i-1)+i} = 2 - \eta_i^{-1} \in [0, 1]$ for $1 \le i \le m$ and $\alpha_j = 0$ else in Definition 3.1. This choice guarantees that for each $1 \le h \le m$ with $\eta_h > 1/2$ there exists exactly one $k(h) \in \mathbb{N}_0$ such that both $\alpha_{k(h)}$ and $\alpha_{k(h)+h}$ are positive; furthermore, for this $k(h)$ one has $\alpha_{k(h)} = 1$ and $\alpha_{k(h)+h} = 2 - \eta_h^{-1}$.

Fix $1 \le h \le m$. If $\eta_h = 1/2$, then there exists no i such that both α_i and α_{i+h} are positive and thus σ_0 and σ_h are independent and (σ_0, σ_h) has a coefficient of tail dependence equal to $1/2$.

If $\eta_h = 1$, then $\alpha_k = \alpha_{k+h} = 1$ for exactly one $k = k(h) \in \mathbb{N}_0$ and thus $\kappa_{k(h)+h} = 1$, $\kappa_j = 0$, $j \ne k(h) + h$, is the unique solution to the optimization problem (5.1). By Theorem 5.1(ii), the coefficient of tail dependence of (σ_0, σ_h) thus equals 1.

Finally, if $\eta_h \in (1/2, 1)$, write

$$\sigma_0 = \prod_{i\in\mathbb{Z}:\,\alpha_i>0} \mathrm{e}^{\alpha_i \eta_{-i}}, \qquad \sigma_h = Z_h \cdot \prod_{i\in\mathbb{Z}:\,\alpha_{i+h}\in(0,1)} \mathrm{e}^{\alpha_{i+h}\eta_{-i}} \tag{6.39}$$

with $\alpha_j := 0$ for $j < 0$, and $Z_h := \prod_{i\in\mathbb{Z}:\,\alpha_{i+h}=1} \mathrm{e}^{\eta_{-i}}$, which is independent of all other factors on the right-hand sides of (6.39), as $\alpha_{i+h} = 1$ implies $\alpha_i = 0$. Moreover, according to the corollary to Theorem 3 of Embrechts and Goldie [14], Z_h is regularly varying with index -1. Now, the joint behavior of σ_0 and σ_h can be derived by applying Theorem 4.2 to the representation (6.39). Observe that the unique optimal solution to the corresponding optimization problem to minimize $\tilde{\kappa} + \sum_{i\in\mathbb{Z}}\kappa_i$ under the constraints $\sum_{i\in\mathbb{Z}}\kappa_i\alpha_i \ge 1$, $\tilde{\kappa} + \sum_{i\in\mathbb{Z}:\,\alpha_{i+h}\in(0,1)}\kappa_i\alpha_{i+h} \ge 1$ and $\tilde{\kappa} \ge 0$, $\kappa_i \ge 0$, $i \in \mathbb{Z}$, is given by $\kappa_{k(h)} = 1$, $\tilde{\kappa} = 1 - \alpha_{k(h)+h} = \eta_h^{-1} - 1$ and $\kappa_j = 0$ for all other $j \in \mathbb{Z}$, because $\kappa_{k(h)}$ is the only value which contributes to both sums of the constraints and, at the same time, it is multiplied with the largest coefficient in the first sum. Therefore, according to Theorem 4.2,

$$P(\sigma_0 > x, \sigma_h > x) \sim c\,P\big(\mathrm{e}^{\eta_{-k(h)}} > x\big)\,P\big(Z_h > x^{\eta_h^{-1}-1}\big)$$

for some constant $c > 0$, and the coefficient of tail dependence of (σ_0, σ_h) equals $1/(\kappa_{k(h)} + \tilde{\kappa}) = \eta_h$.

By Corollary 2.5 and the following note, the same holds true for (X_0, X_h) (and $(|X_0|, |X_h|)$) for all $\eta_h \in [1/2, 1]$ if $E(|\varepsilon_0|^{2+\delta}) < \infty$ for some $\delta > 0$. □

***Details for Remark* 5.5(ii).** If $\eta_h = 1$ for some $h > 0$, then any optimal solution to (5.1) has to satisfy $\sum_{i=0}^{\infty}\kappa_i = 1$. (This even holds in the case that the solution is not unique since otherwise

Theorem 4.2(i) leads to a contradiction to $\eta_h = 1$.) But then $\kappa_i > 0$ can only hold if $\alpha_{i-h} = \alpha_i = 1$. Write

$$\sigma_0 = Z \prod_{i \ge h : \min\{\alpha_i, \alpha_{i-h}\} < 1} e^{\alpha_{i-h}\eta_{h-i}}, \qquad \sigma_h = Z \prod_{i \in \mathbb{N}_0 : \min\{\alpha_i, \alpha_{i-h}\} < 1} e^{\alpha_i \eta_{h-i}}$$

with $Z := \prod_{i \in \mathbb{N}_0 : \alpha_i = \alpha_{i-h} = 1} e^{\eta_{h-i}}$. Again, Z is regularly varying with index -1 by Embrechts and Goldie [14]. Then the corresponding linear program

$$\tilde{\kappa} + \sum_{i \in \mathbb{N}_0 : \min\{\alpha_i, \alpha_{i-h}\} < 1} \kappa_i \to \min!$$

under the constraints

$$\tilde{\kappa} + \sum_{i \ge h : \min\{\alpha_i, \alpha_{i-h}\} < 1} \alpha_{i-h}\kappa_i \ge 1, \qquad \tilde{\kappa} + \sum_{i \in \mathbb{N}_0 : \min\{\alpha_i, \alpha_{i-h}\} < 1} \alpha_i \kappa_i \ge 1,$$

$$\tilde{\kappa} \ge 0, \kappa_i \ge 0, \forall i \in \mathbb{N}_0,$$

has the unique solution $\tilde{\kappa} = 1, \kappa_i = 0$ for all $i \in \mathbb{N}_0$, and thus $P(\sigma_0 > x, \sigma_h > x) \sim cP(Z > x)$ for some constant $c > 0$ by Theorem 4.2. By Lemma 7.2 of Rootzén [27]

$$P(Z > x) = P\left(\sum_{i \in \mathbb{N}_0 : \alpha_i = \alpha_{i+h} = 1} \eta_{-i} > \log x \right) \sim \hat{K} (\log x)^{\beta} x^{-1},$$

for a constant $\hat{K} > 0$, where we have used equation (7.8) in Rootzén [27] and the fact that $\beta < -1$. On the other hand,

$$P(\sigma_0 > x) = P\left(\sum_{i=0}^{\infty} \alpha_i \eta_{-i} > \log x \right) \sim \hat{K}' (\log x)^{\beta} x^{-1},$$

for some constant $\hat{K}' > 0$ by the same arguments as above. Combining the above asymptotics, we arrive at

$$\lim_{x \to \infty} P(\sigma_h > x | \sigma_0 > x) = \lim_{x \to \infty} \frac{P(\sigma_0 > x, \sigma_h > x)}{P(\sigma_0 > x)} > 0$$

and thus asymptotic dependence of (σ_0, σ_h). The same holds true for (X_0, X_h).

In contrast, if $\beta > -1$, then one may conclude from Lemma 7.2 of Rootzén [27] that $P(\sigma_h > x | \sigma_0 > x) = O((\log x)^{-l(\beta+1)}) \to 0$ as $x \to \infty$ with $l := \sum_{i=0}^{\infty} \mathbb{1}_{\{\alpha_i = 1\}} - \sum_{i=0}^{\infty} \mathbb{1}_{\{\alpha_i = \alpha_{i+h} = 1\}} > 0$, which confirms the asymptotic independence of σ_0 and σ_h in this case.

Acknowledgments

The authors would like to thank the Associate Editor and two anonymous referees for their careful reading and many helpful comments. This project was supported by the German Research Foundation DFG, Grant no JA 2160/1.

References

[1] Abanto-Valle, C.A., Bandyopadhyay, D., Lachos, V.H. and Enriquez, I. (2010). Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Comput. Statist. Data Anal.* **54** 2883–2898. MR2727721

[2] Asai, M., McAleer, M. and Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Rev.* **25** 145–175. MR2256285

[3] Basrak, B., Davis, R.A. and Mikosch, T. (2002). Regular variation of GARCH processes. *Stochastic Process. Appl.* **99** 95–115. MR1894253

[4] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Process. Appl.* **119** 1055–1080. MR2508565

[5] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR0898871

[6] Breiman, L. (1965). On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10** 323–331.

[7] Das, B., Mitra, A. and Resnick, S. (2013). Living on the multidimensional edge: Seeking hidden risks using regular variation. *Adv. in Appl. Probab.* **45** 139–163. MR3077544

[8] Das, B. and Resnick, S.I. (2011). Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli* **17** 226–252. MR2797990

[9] Davis, R.A. and Mikosch, T. (2001). Point process convergence of stochastic volatility processes with application to sample autocorrelation: Probability, statistics and seismology. *J. Appl. Probab.* **38A** 93–104. MR1915537

[10] Davis, R.A. and Mikosch, T. (2009). Extremes of stochastic volatility models. In *Handbook of Financial Time Series* (T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch, eds.) 355–364. Berlin: Springer.

[11] Denisov, D. and Zwart, B. (2007). On a theorem of Breiman and a class of random difference equations. *J. Appl. Probab.* **44** 1031–1046. MR2382943

[12] Draisma, G., Drees, H., Ferreira, A. and de Haan, L. (2004). Bivariate tail estimation: Dependence in asymptotic independence. *Bernoulli* **10** 251–280. MR2046774

[13] Drees, H., Segers, J. and Warchoł, M. (2014). Statistics for tail processes of Markov chains. Preprint. Available at arXiv:1405.7721.

[14] Embrechts, P. and Goldie, C.M. (1980). On closure and factorization properties of subexponential and related distributions. *J. Austral. Math. Soc. Ser. A* **29** 243–256. MR0566289

[15] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910

[16] Kulik, R. and Soulier, P. (2013). Estimation of limiting conditional distributions for the heavy tailed long memory stochastic volatility process. *Extremes* **16** 203–239. MR3057196

[17] Kulik, R. and Soulier, P. (2013). Heavy tailed time series with extremal independence. Preprint.

[18] Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65** 291–306. MR0722133

[19] Ledford, A.W. and Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187. MR1399163

[20] Ledford, A.W. and Tawn, J.A. (2003). Diagnostics for dependence within time series extremes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **65** 521–543. MR1983762

[21] Lindskog, F., Resnick, S.I. and Roy, J. (2014). Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probab. Surv.* **11** 270–314. MR3271332

[22] Mikosch, T. and Rezapour, M. (2013). Stochastic volatility models with possible extremal clustering. *Bernoulli* **19** 1688–1713. MR3129030

[23] Pratt, J.W. (1960). On interchanging limits and integrals. *Ann. Math. Statist.* **31** 74–77. MR0123673
[24] Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5** 303–336 (2003). MR2002121
[25] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424
[26] Resnick, S.I. (2008). Multivariate regular variation on cones: Application to extreme values, hidden regular variation and conditioned limit laws. *Stochastics* **80** 269–298. MR2402168
[27] Rootzén, H. (1986). Extreme value theory for moving average processes. *Ann. Probab.* **14** 612–652. MR0832027
[28] Sierksma, G. (1996). *Linear and Integer Programming: Theory and Practice. Monographs and Textbooks in Pure and Applied Mathematics* **198**. New York: Dekker, Inc. MR1449398
[29] Taylor, S.J. (1986). *Modelling Financial Time Series*. Chichester: Wiley.
[30] Tsay, R.S. (2002). *Analysis of Financial Time Series*. Hoboken, NJ: Wiley.

Received October 2013 and revised August 2014