# A stochastic volatility model with flexible extremal dependence structure

**Anja Janssen[*] and Holger Drees[**]**

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

**Abstract:** Since extreme observations of a stochastic volatility time series with heavy-tailed innovations are almost independent (with coefficient of tail dependence equal to $1/2$) for all positive lags, these models do not capture the well-known clustering of extreme losses which is often observed in financial returns. To overcome this drawback, 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 $(0,\infty)^d$. An important ingredient is a new Breiman-type result about regular variation on this cone for products of a random matrix and a regularly varying random vector, which is of interest of its own.

**Keywords and phrases:** asymptotic independence, Breiman's lemma, coefficient of tail dependence, extremal dependence, financial time series, hidden regular variation, regular variation on cones, power products, stochastic volatility time series.
Primary 60G70; secondary 60F99, 91B28, 91B84.

## 1. Introduction

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

$$X_t = \sigma_t \epsilon_t, \quad t \in \mathbb{Z}, \tag{1.1}$$

where $\epsilon_t$, $t \in \mathbb{Z}$, are i.i.d. innovations and $(\sigma_t)_{t\in\mathbb{Z}}$ is a stationary time series of so-called volatilities. The two most popular classes of multiplicative models vary in the way the volatilities are modeled. While $\sigma_t$ is a function of past innovations $\epsilon_s$, $s < t$, 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 the log-volatilities are described by a Gaussian linear time series of the type

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

with i.i.d. normal innovations $\eta_t$, $t \in \mathbb{Z}$, independent of $(\epsilon_t)_{t\in\mathbb{Z}}$ (although, sometimes, $\epsilon_t$ and $\eta_{t+1}$ 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 $\epsilon_t$ are often assumed to be regularly varying, i.e.

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

for all $s > 0$, some $\alpha > 0$ and $p \in [0, 1]$. By Breiman's lemma (see Breiman (1965)) 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 significantly between the two model classes. In the present paper, we focus on the concept of regular variation of random vectors to describe the extremal dependence structure of a time series. 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 $\mathbb{B}([0,\infty)^d \setminus \{\mathbf{0}\})$ which is finite on $[0,\infty)^d \setminus [0,x]^d$ for all $x > 0$ such that

$$\frac{P(\mathbf{Z} \in xB)}{P(\|\mathbf{Z}\| > x)} \ \to \ \tilde{\nu}(B), \quad x \to \infty, \tag{1.4}$$

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 $\partial B$ 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 (2006) which can be shown to be equivalent to vague convergence on $[0,\infty]^d \setminus \{\mathbf{0}\}$. See, for example, Resnick (2007), 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 et al. (2002) show that under general conditions the vectors $(\sigma_{t_1}, \ldots, \sigma_{t_d})$ and $(X_{t_1}, \ldots, X_{t_d})$ are multivariate regularly varying for all $t_1 < \ldots < t_d$, and the same holds true for SV models with heavy-tailed innovations $\epsilon_t$. 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 & Mikosch (2001)). This so-called asymptotic independence of lagged returns (and volatilities) renders convergence (1.4) rather uninformative. In particular, we can merely conclude that the probability $P(X_0 > s_0x, X_h > s_hx)$ of joint exceedances is of smaller order than $P(X_0 > x)$ for all $s_0, s_h > 0$, but we neither obtain its rate of convergence to 0, nor whether the probability can be standardized in a different way than in (1.4) 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 $\min(\mathbf{z}) := \min\{z_1, \ldots, z_d\}$ for $\mathbf{z} = (z_1, \ldots, z_d) \in \mathbb{R}^d$. Now, instead of (1.4), we assume that there exists a measure $\nu \neq 0$ on $\mathbb{B}(\mathbb{E}^d)$ which is finite on $(x,\infty)^d$ for all $x > 0$ such that

$$\frac{P(\mathbf{Z} \in xB)}{P(\min(\mathbf{Z}) > x)} \to \nu(B), \quad x \to \infty, \tag{1.5}$$

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 (1.4) just one coordinate needs to be extreme. Again $\nu$ is homogeneous of some order $-\alpha < 0$, which is called the index of regular variation on $\mathbb{E}^d$. If both (1.4) and (1.5) hold, then $\alpha \geq \tilde{\alpha}$. In the case of asymptotic independence (i.e. $\tilde{\nu}$ is concentrated on $[0,\infty)^d \backslash (0,\infty)^d$), the random vector $\mathbf{Z}$ is then said to exhibit hidden regular variation. See Resnick (2002), Resnick (2007), Section 9.4, Resnick (2008) and Das et al. (2011) for further details about hidden regular variation and regular variation on cones.

Corollary 2.5 below shows 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 $\nu$ is the same as if the components of this vector were exactly independent. A similar result holds for the absolute returns. In particular, $P(X_0 > s_0 x, X_h > s_h x)$ is of the same order as $(P(X_0 > x))^2$, which means that the coefficient of tail dependence introduced in Ledford & Tawn (1996) and Ledford and Tawn (2003) 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 $\eta$ 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 $-1/\eta$.

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. However, there is some empirical evidence (cf. Drees (2013)) that real consecutive returns exhibit a stronger extremal dependence with a coefficient of tail dependence $\eta_h$ strictly between 1/2 and 1 for small $h$, which means that although the dependence between exceedances over high thresholds vanishes asymptotically, it is nevertheless significantly stronger than modeled by classical SV time series. 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.

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 (1965) 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 et al. (2002), 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 (1.5) 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 preserved under multiplication with a general matrix $\mathbf{A}$.

According to these results, in the classical SV models the (absolute) returns inherit both their first and their second order dependence behavior (i.e., the regular variation on $[0,\infty)^d \backslash \{\mathbf{0}\}$ and on $\mathbb{E}^d$) from the i.i.d. innovations $\epsilon_t$ which are assumed to be more heavy tailed than the volatilities. To avoid the resulting very weak dependence while keeping a linear model for the log-volatilities with its nice probabilistic properties, one must therefore ensure that the volatility time series is more heavy tailed than the innovations. In Section 3 we introduce such a class of models with Gamma-type log-volatilities and derive their marginal tail behavior using results from Rootzén (1986). 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 (2012), but their analysis is restricted to the 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 two 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 seen, the exact arguments are more delicate and we split up the results into the cases of two, finitely many and infinitely many factors. 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 $(X_0, X_h), 1 \leq h \leq m$, 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: Weak convergence is denoted by $\overset{w}{\to}$. The expression $\delta_x$ stands for the Dirac measure at $x$. We denote the positive and negative part of $x \in \mathbb{R}$ by $x^+ := \max\{x, 0\}$ and $x^- := -\min\{x, 0\}$, respectively. For $x \in \mathbb{R}$ the expressions $\lceil x \rceil$ and $\lfloor x \rfloor$ denote the smallest integer greater than or equal to $x$ and the largest integer smaller than or equal to $x$, respectively. The complement of a set $A$ is denoted by $A^c$. The empty product, $\prod_{i\in\emptyset} X_i$, 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 $(\sigma_{t_1}, \ldots, \sigma_{t_d})$ and of the returns $(X_{t_1}, \ldots, X_{t_d})$ 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

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

multivariate Breiman-type results are very useful to establish regular variation of $(X_{t_1}, \ldots, X_{t_d})$. According to Basrak et al. (2002, Proposition A.1), the product $\mathbf{AX}$

of a random matrix $\mathbf{A}$ and a random vector $\mathbf{X}$ is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ if $\mathbf{X}$ is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ with index $-\tilde{\alpha} < 0$ and $E(\|\mathbf{A}\|_{\text{op}}^{\tilde{\alpha}+\epsilon}) < \infty$ for some $\epsilon > 0$. Here $\|\cdot\|_{\text{op}}$ denotes the operator norm for matrices, which is defined by

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

where $d(\mathbf{x}, A) := \inf_{\mathbf{y}\in A} \|\mathbf{x} - \mathbf{y}\|$ denotes the usual distance function induced by the Euclidean norm on $\mathbb{R}^d$. More precisely, if $\mathbf{X}$ satisfies (1.4), then

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

for all $B \in \mathbb{B}([0,\infty)^d)$ bounded away from $\mathbf{0}$ with $E(\tilde{\nu}(\mathbf{A}^{-1}(\partial B))) = 0$. Therefore, if the vector $(\epsilon_{t_1}, \dots, \epsilon_{t_d})$ of i.i.d. innovations is regularly varying on $[0,\infty)^d \setminus \{\mathbf{0}\}$ with index $-\tilde{\alpha}$ (which is obvious if $\epsilon_0$ is regularly varying) and $E(\sigma_0^{\tilde{\alpha}+\epsilon}) < \infty$, then one may conclude the regular variation of $(X_{t_1}, \dots, X_{t_d})$ on $[0,\infty)^d \setminus \{\mathbf{0}\}$. 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 $E(|\epsilon_0|^{\tilde{\alpha}+\epsilon}) < \infty$ for some $\epsilon > 0$.

We now want to derive a similar Breiman-type result for regular variation on $\mathbb{E}^d = (0,\infty)^d$. Because convergence (1.5) describes the behavior of random vectors exclusively on $\mathbb{E}^d$, we have to ensure that the preimage $\mathbf{A}^{-1}(B)$ of a set $B \subset \mathbb{E}^d$ is again a subset of $\mathbb{E}^d$. Moreover, as the limit measure $\nu$ in the definition (1.5) of regular variation on $\mathbb{E}^d$ 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, let

$$\mathcal{S}^d := \{\mathbf{x} \in \mathbb{E}^d \mid \min(\mathbf{x}) = 1\} = \{\mathbf{x} \in \mathbb{E}^d \mid d(\mathbf{x}, \mathbb{F}^d) = 1\},$$

where $\mathbb{F}^d := \mathbb{R}^d \setminus \mathbb{E}^d$. 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 $d(\mathbf{x}, \mathbb{F}^d) = d(\mathbf{x}, \mathcal{O}^d)$ for all $\mathbf{x} \in \mathbb{E}^d$, one may consider $\mathcal{S}^d$ an analog to the unit sphere $\{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x} - \mathbf{0}\| = 1\}$ in the present setting. Cf. Das et al. (2011) for the use of $\mathcal{S}^d$ in the context of regular variation on cones.

Now define $\tau : \mathbb{R}^{d\times d} \to [0,\infty]$ by

$$\tau(\mathbf{A}) := \sup_{\mathbf{x}\in\mathcal{S}^d} \min((\mathbf{A}\mathbf{x})^+) = \sup_{\mathbf{x}\in\mathcal{S}^d} d(\mathbf{A}\mathbf{x}, \mathbb{F}^d),$$

where $(x_1, \dots, x_d)^+ = (x_1^+, \dots, x_d^+)$. Note that, coming back to the above analog, $\tau$ bears similarity to the operator norm for matrices used by Basrak et al. (2002, Proposition A.1). Indeed, it will turn out that one may prove an analogous result for the regular variation on $\mathbb{E}^d$ if one replaces the moment condition $E(\|\mathbf{A}\|_{\text{op}}^{\tilde{\alpha}+\epsilon}) < \infty$ with the corresponding condition on $\tau(\mathbf{A})$. In order to ensure that $\mathbf{A}^{-1}(B) \subset \mathbb{E}^d$ for all $B \subset \mathbb{E}^d$ a much simpler condition on $\tau(\mathbf{A})$ suffices.

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

*(i)* $\mathbf{A}^{-1}(\mathbb{E}^d) \subset \mathbb{E}^d$;
*(ii)* $0 < \tau(\mathbf{A}) < \infty$.

*If* $\mathbf{A}$ *possesses these properties, then it is invertible and* $\mathbf{A}^{-1}([0,\infty)^d) \subset [0,\infty)^d$.

Sometimes it is helpful to express $\tau(\mathbf{A})$ in terms of the inverse matrix $\mathbf{A}^{-1}$.

**Lemma 2.2.** *Let* $\mathbf{A} \in \mathbb{R}^{d\times d}$ *satisfy* $0 < \tau(\mathbf{A}) < \infty$ *and denote the inverse of* $\mathbf{A}$ *by* $(A^{-1}_{ij})_{1\le i,j\le d}$. *Then*

$$\tau(\mathbf{A}) = \frac{1}{d(\mathbf{A}^{-1}\mathbf{1}, \mathbb{F}^d)} = \max_{i=1,\dots,d} \left\{ \left( \left( \sum_{j=1}^{d} A^{-1}_{ij} \right)^{+} \right)^{-1} \right\},$$

*where* $\mathbf{1} = (1,\dots,1)'$.

**Example 2.3.** *The function* $\tau$ *is particularly simple for a* $d$*-dimensional diagonal matrix* $\mathbf{\Delta}$ *with diagonal elements* $\delta_1,\dots,\delta_d > 0$ *where*

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

*by Lemma 2.2.*

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

**Theorem 2.4.** *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* $\nu$. *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*

$$E(\tau(\mathbf{A})^{\alpha+\delta}) < \infty \tag{2.2}$$

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

$$\lim_{x\to\infty} \frac{P(\mathbf{AZ} \in xB)}{P(\min(\mathbf{Z}) > x)} = E(\nu(\mathbf{A}^{-1}B)) < \infty \tag{2.3}$$

*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,* $\mathbf{AZ}$ *is regularly varying on* $\mathbb{E}^d$ *with index* $-\alpha$ *and limiting measure*

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

With the above theorem we are able to derive the second order extremal dependence behavior of the (absolute) returns of a classical SV model.

**Corollary 2.5.** *Let* $(X_t)_{t\in\mathbb{Z}}$ *satisfy* (1.1)–(1.3) *with normally distributed random variables* $\eta_t$. *Then* $(X_{t_1},\dots,X_{t_d})$ *is regularly varying on* $\mathbb{E}^d$ *with index* $-d\alpha$ *and limit measure* $\nu^X$ *given by*

$$\nu^X\left( \times_{i=1}^{d} (s_i,\infty) \right) = \prod_{i=1}^{d} s_i^{-\alpha} \tag{2.4}$$

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

**Proof**. Under the above conditions, $(\epsilon_{t_1}, \ldots, \epsilon_{t_d})$ is regularly varying on $\mathbb{E}^d$ with index $-d\alpha$ and limit measure $\nu^\epsilon$ given by

$$\nu^\epsilon\Big( \times_{i=1}^d (s_i, \infty)\Big) = \lim_{x\to\infty} \frac{P(\epsilon_{t_i} > s_i x\ \forall\, 1 \le i \le d)}{P(\epsilon_{t_i} > x\ \forall\, 1 \le i \le d)} = \prod_{i=1}^d s_i^{-\alpha}.$$

Denote by $\mathbf{A}$ the diagonal matrix with entries $\sigma_{t_1}, \ldots, \sigma_{t_d}$. As $\sigma_t$ is lognormally distributed, Example 2.3 shows that all moments of $\tau(\mathbf{A})$ are finite. We may thus apply Theorem 2.4 to the second representation of $(X_{t_1}, \ldots, X_{t_d})$ in (2.1) to see that this vector is regularly varying on $\mathbb{E}^d$ with limit measure (2.4). □

A similar result holds if we replace $X_t$ and $\epsilon_t$ by $|X_t|$ and $|\epsilon_t|$, respectively.

By the above corollary, in classical SV models the returns inherit the coefficient of tail dependence $\eta_h = 1/2$ of $(X_0, X_h)$ for all $h > 0$ from the independent innovations $\epsilon_t, t \in \mathbb{Z}$. In contrast, if the volatility sequence is more heavy tailed than the innovations $\epsilon_t$, the second order extremal dependence behavior of the return time series is determined by the volatility sequence.

**Corollary 2.6.** *Let $X_t = \sigma_t \epsilon_t$, $t \in \mathbb{Z}$, where $(\sigma_t)_{t\in\mathbb{Z}}$ and $(\epsilon_t)_{t\in\mathbb{Z}}$ are independent stationary time series with $\epsilon_t$, $t \in \mathbb{Z}$, i.i.d. and $\sigma_t > 0$. Assume furthermore that $(\sigma_{t_1}, \ldots, \sigma_{t_d})$ is regularly varying on $\mathbb{E}^d$ with index $-\alpha_d < 0$ and limit measure $\nu_d^\sigma$. Let in addition $0 < E((\epsilon_0^+)^{\alpha_d+\delta}) < \infty$ for some $\delta > 0$. Then $(X_{t_1}, \ldots, X_{t_d})$ is regularly varying on $\mathbb{E}^d$ as well, with the same index $-\alpha_d$ and limit measure $\nu_d^X$ defined by*

$$\nu_d^X\Big( \times_{i=1}^d (s_i, \infty)\Big) = \frac{E\Big(\nu_d^\sigma\Big( \times_{i=1}^d \Big(\frac{s_i}{\epsilon_{t_i}^+}, \infty\Big)\Big)\Big)}{E\Big(\nu_d^\sigma\Big( \times_{i=1}^d \Big(\frac{1}{\epsilon_{t_i}^+}, \infty\Big)\Big)\Big)} \tag{2.5}$$

*for all $s_1, \ldots, s_d > 0$.*

**Proof**. Let $\epsilon_{t_i}^*$, $1 \le i \le d$, be i.i.d. random variables independent of $(\sigma_{t_1}, \ldots, \sigma_{t_d})$ with distribution $P^{\epsilon_0|\epsilon_0>0}$ and define

$$\begin{pmatrix} X_{t_1}^* \\ \vdots \\ X_{t_d}^* \end{pmatrix} = \begin{pmatrix} \epsilon_{t_1}^* & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \epsilon_{t_d}^* \end{pmatrix} \begin{pmatrix} \sigma_{t_1} \\ \vdots \\ \sigma_{t_d} \end{pmatrix} =: \mathbf{A}\sigma. \tag{2.6}$$

Because of $E((\max_{1\le i\le d} \epsilon_{t_i}^*)^{\alpha_d+\delta}) \le dE((\epsilon_0^+)^{\alpha_d+\delta})/P(\epsilon_0 > 0) < \infty$ and Example 2.3, we have $E(\tau(\mathbf{A})^{\alpha_d+\delta}) < \infty$. Now apply Theorem 2.4 to conclude the regular variation of $(X_{t_1}^*, \ldots, X_{t_d}^*)$ on $\mathbb{E}^d$ with limit measure $\nu_d^X$, which is equivalent to the assertion. □

Again the above result can be extended to $(|X_{t_1}|, \ldots, |X_{t_d}|)$.

In view of Corollary 2.6, it is essential to analyze the extremal behavior of the volatility time series $(\sigma_t)_{t\in\mathbb{Z}}$. This will be done in Sections 4 and 5 for the SV model introduced in the next section.

## 3. SV models with Gamma-type log-volatilities

As shown in Corollary 2.5, in classical heavy-tailed SV time series the dependence between large observations is always very weak, irrespective of the specific model parameters. We will now introduce a new class of stochastic volatility models with a more flexible extremal dependence structure.

To this end, 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 (1986) which guarantee the existence of stationary time series that meet our assumptions.

**Definition 3.1.** *Let*

$$X_t = \sigma_t \epsilon_t, \quad t \in \mathbb{Z}, \tag{3.1}$$

*with* $\epsilon_t, t \in \mathbb{Z}$, *i.i.d. such that* $P(\epsilon_0 > 0) > 0$ *and* $E(|\epsilon_0|^{1+\delta}) < \infty$ *holds for some* $\delta > 0$. *Furthermore, let*

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

*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* $\alpha_i = O(i^{-\theta})$ *as* $i \to \infty$ *for some* $\theta > 1$,

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

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

*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 notations, 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 $(\epsilon_t)_{t \in \mathbb{Z}}$ the *innovation process*.

**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* $\sigma_0$ *are regularly varying with index* $-1$. *Furthermore, the distribution of* $X_0$ *is tail balanced with*

$$\lim_{x \to \infty} \frac{P(X_0 > x)}{P(|X_0| > x)} = \frac{E(\epsilon_0^+)}{E(|\epsilon_0|)}, \quad \lim_{x \to \infty} \frac{P(X_0 < -x)}{P(|X_0| > x)} = \frac{E(\epsilon_0^-)}{E(|\epsilon_0|)} \tag{3.4}$$

*and*

$$\lim_{x \to \infty} \frac{P(|X_0| > x)}{P(\sigma_0 > x)} = E(|\epsilon_0|), \quad \lim_{x \to \infty} \frac{P(X_0 > x)}{P(\sigma_0 > x)} = E(\epsilon_0^+). \tag{3.5}$$

**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*

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

*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 $E(|\epsilon_0|^{1/c+\delta}) < \infty$ 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., for example, Taylor (1986)). In this case, equation (3.2) would read as $\log(\sigma_t) = \mu + \sum_{i=0}^{\infty} \alpha_i \eta_{t-i}$ for a $\mu \in \mathbb{R}$. Such an assumption is usually combined with the standardization of some moment of $\epsilon_0$, for example by setting $Var(\epsilon_0) = 1$. Otherwise, setting $\hat{\epsilon}_i := \mathrm{e}^{\mu}\epsilon_i, i \in \mathbb{Z}$, has the same effect as adding $\mu$ in the definition of $\log(\sigma_t)$. Since we make no assumptions about the particular form of (existing) moments of $\epsilon_0$, we set $\mu = 0$ without loss of generality.*

**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, i.e.*

$$\log(\sigma_t) = \alpha \log(\sigma_{t-1}) + \eta_t, \quad t \in \mathbb{Z}. \tag{3.6}$$

*A similar model, with a modified assumption about the distribution of $\eta_t, t \in \mathbb{Z}$, (namely, that the distribution of $\exp(\eta_t)$ is regularly varying and that a stationary solution to (3.6) exists) has been analyzed with respect to its first order extremal behavior in Mikosch and Rezapour (2012).*

*Moreover, the conditional extreme value behavior of consecutive observations given that $X_0$ is large has been analyzed by Kulik and Soulier (2013) in the case that the innovations $\eta_t$ of the log-volatility series are double exponentially distributed. See Section 5 for a more detailed comparison with their results.*

The regular variation established in Theorem 3.2 implies that the distributions of both $\sigma_0$ and $X_0$ belong to the max-domain of attraction of the unit Fréchet distribution. More precisely, for normalizing constants $a_n := \hat{K} n (\log n)^{\hat{\beta}}, n \in \mathbb{N}$, (see (6.9) and (6.10) for the definition of $\hat{\beta}$ and $\hat{K}$) and $z > 0$ one has that

$$\begin{gathered} P(\sigma_0 \le a_n z)^n \to \exp(-1/z), \quad P(|X_0| \le a_n z)^n \to \exp(-E(|\epsilon_0|)/z), \\ P(X_0 \le a_n z)^n \to \exp(-E(\epsilon_0^+)/z) \end{gathered} \tag{3.7}$$

as $n \to \infty$, which follows from (3.5) in combination with Theorem 7.3 in Rootzén (1986).

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 (1986) 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 (2007), 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 sequence. In the case that $\beta < -1$ assume additionally that $k := |\{n \in \mathbb{N}_0 : \alpha_n = 1\}| = 1$. Let*

*$N_n^\sigma, n \in \mathbb{N}$, and $N_n^X, n \in \mathbb{N}$, denote the point processes defined by*

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

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

*(ii) Let $(t_{(i)}, z_{(i)})_{i\in\mathbb{N}}$ denote the points of the Poisson process $N^\sigma$ and $(\epsilon_{(i)})_{i\in\mathbb{N}}$ an i.i.d. sequence with $P^{\epsilon_{(1)}} = P^{\epsilon_0}$, independent of $(t_{(i)}, z_{(i)})_{i\in\mathbb{N}}$.*
*Then $N_n^X(\cdot \cap D) \xrightarrow{w} N^X(\cdot \cap D)$ 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 $(t_{(i)}, z_{(i)}\epsilon_{(i)})_{i\in\mathbb{N}}$. The restriction of $N^X$ to $D$ is a Poisson point process with intensity measure $dt \times z^{-2}\left[\mathbb{1}_{\{z<0\}}E(\epsilon_0^-) + \mathbb{1}_{\{z>0\}}E(\epsilon_0^+)\right]dz$.*

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

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

for all $z > 0$. A comparison with (3.7) 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 in the sense that the expected length of an extremal cluster is equal to 1; cf. Leadbetter (1983). This is also obvious from the fact 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). More formally speaking, the form of the limiting point process implies asymptotic independence, i.e.

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

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 $P(X_0 > s_0x, X_h < -s_hx) = P(\sigma_0\epsilon_0^+ > s_0x, \sigma_h\epsilon_h^- > s_hx)$ 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 $\prod_{i=1}^\infty X_i^{\alpha_i}$ and $\prod_{i=1}^\infty X_i^{\beta_i}$ for i.i.d. non-negative and regularly varying random variables $X_i, i \in \mathbb{N}$.

The connection to SV models with Gamma-type log-volatilities becomes clear by writing

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

We will show that the joint tail behavior of those products is closely related to the following infinite dimensional linear optimization problem:

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

under the constraints

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

This relation can be explained by the following heuristic argument. 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 leads 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. (1987), Theorem 1.5.6) for functions $f : [0,\infty) \to (0,\infty)$ which are regularly varying with index $-\alpha$: for all $\epsilon > 0$ there exists a constant $M = M(\epsilon)$ such that

$$(1-\epsilon) f(x) s^{-\alpha} \min\{s^{-\epsilon}, s^{\epsilon}\} \leq f(sx) \leq (1+\epsilon) f(x) s^{-\alpha} \max\{s^{-\epsilon}, s^{\epsilon}\} \tag{4.1}$$

for all $x, s$ such that $\min\{x, sx\} > M$. For brevity, introduce

$$\epsilon(s) = \begin{cases} \epsilon, & s \geq 1 \\ -\epsilon, & 0 < s < 1 \end{cases},$$

such that (4.1) reads as

$$(1-\epsilon) f(x) s^{-\alpha-\epsilon(s)} \leq f(sx) \leq (1+\epsilon) f(x) s^{-\alpha+\epsilon(s)}. \tag{4.2}$$

Before we deal with the general case, we first analyze products of two factors in the case that a unique solution to the above optimization problem exists.

**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*

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

*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 $\epsilon > 0$, there exists a constant $M = M(\epsilon) > 0$ such that*

$$\begin{aligned} &(1-\epsilon) C s_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)-\epsilon(s_0)} s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)-\epsilon(s_1)} \\ \leq\ &\frac{P(X_1^{\alpha_1}X_2^{\alpha_2} > s_0 x, X_1^{\beta_1}X_2^{\beta_2} > s_1 x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\ \leq\ &(1+\epsilon) C s_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)+\epsilon(s_0)} s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)+\epsilon(s_1)} \end{aligned}$$

*for all $s_0, s_1 > 0$ such that $\min\{x^{\kappa_1}, z_1^{-1}x^{\kappa_1}, x^{\kappa_2}, z_2^{-1}x^{\kappa_2}\} > M$ where*

$$z_1 := s_1^{\alpha_2/(\alpha_1\beta_2-\alpha_2\beta_1)} s_0^{-\beta_2/(\alpha_1\beta_2-\alpha_2\beta_1)}, \quad z_2 := s_0^{\beta_1/(\alpha_1\beta_2-\alpha_2\beta_1)} s_1^{-\alpha_1/(\alpha_1\beta_2-\alpha_2\beta_1)} \tag{4.4}$$

*(i.e., $s_0 = z_1^{-\alpha_1} z_2^{-\alpha_2}, s_1 = z_1^{-\beta_1} z_2^{-\beta_2}$).*

We are now ready to deal with products of finitely many random factors, where again the joint extremal behavior of these products is closely linked to the solution of a linear program.

**Theorem 4.2.** *Let $X_1, \ldots, X_n$ be independent non-negative random variables which are all regularly varying with index $-1$. Set $Y_0 = \prod_{i=1}^n X_i^{\alpha_i}, Y_1 = \prod_{i=1}^n X_i^{\beta_i}$ with $\alpha_1, \ldots, \alpha_n$, $\beta_1, \ldots, \beta_n \geq 0$ such that $\max_{1\leq i\leq n} \alpha_i > 0$ and $\max_{1\leq i\leq n} \beta_i > 0$. Then the optimization problem*

$$\sum_{i=1}^n \kappa_i \to \min! \tag{4.5}$$

*under the constraints*

$$\sum_{i=1}^n \alpha_i\kappa_i \geq 1, \quad \sum_{i=1}^n \beta_i\kappa_i \geq 1, \quad \kappa_i \geq 0 \ \ \forall\, 1 \leq i \leq n, \tag{4.6}$$

*has a solution $(\kappa_1, \ldots, \kappa_n)$. If the solution is unique, then at most two of the $\kappa_i$ are strictly positive and $\min\{Y_0, Y_1\}$ is regularly varying with index $-\sum_{i=1}^n \kappa_i$. Moreover,*

*(i) if the solution is unique and $\kappa_i, \kappa_j > 0$ for some $i \neq j$, then for all $s_0, s_1 > 0$*

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

*with*

$$D := \frac{|\alpha_i\beta_j - \alpha_j\beta_i|}{(\alpha_i-\alpha_j)(\beta_j-\beta_i)} E\left( \prod_{m\in\{1,\dots,n\}\setminus\{i,j\}} X_m^{(\alpha_m(\beta_j-\beta_i)+\beta_m(\alpha_i-\alpha_j))/(\alpha_i\beta_j-\alpha_j\beta_i)} \right);$$

*(ii) if the solution is unique, $\kappa_i > 0$ for exactly one $i \in \{1,\dots,n\}$ and $\kappa_j = 0$ else, then*

$$\lim_{x\to\infty} \frac{P(Y_0 > s_0 x, Y_1 > s_1 x)}{P(X_i > x^{\kappa_i})}$$

$$= \begin{cases} E\left(\min\left(s_0^{-1}\prod_{j\in J_i} X_j^{\alpha_j}, s_1^{-1}\prod_{j\in J_i} X_j^{\beta_j}\right)^{1/\beta_i}\right), & \alpha_i = \beta_i \\ E\left(\prod_{j\in J_i} X_j^{\beta_j/\beta_i} \mathbb{1}_{\{X_j^{\alpha_j}>0\}}\right) s_1^{-1/\beta_i}, & \alpha_i > \beta_i \\ E\left(\prod_{j\in J_i} X_j^{\alpha_j/\alpha_i} \mathbb{1}_{\{X_j^{\beta_j}>0\}}\right) s_0^{-1/\alpha_i}, & \alpha_i < \beta_i \end{cases} \tag{4.8}$$

*with $J_i := \{1,\dots,n\}\setminus\{i\}$;*

*(iii) in any case, for all $\epsilon > 0$,*

$$P(Y_0 > x, Y_1 > x) = o\left(x^{\epsilon - \sum_{i=1}^n \kappa_i}\right) \tag{4.9}$$

$$x^{-\epsilon-\sum_{i=1}^n \kappa_i} = o(P(Y_0 > x, Y_1 > x)) \tag{4.10}$$

*as $x \to \infty$.*

**Remark 4.3.** *(i) The linear program (4.5), (4.6) has multiple solutions if, e.g., the exponents $\beta_i$ of at least two factors $X_i^{\beta_i}$ equal $\beta^* := \max_{1\le j\le n}\beta_j$ and the corresponding exponents $\alpha_i$ are greater than or equal to $\beta^*$. (A simple example is $n = 2$ with $\alpha_i = \beta_i = 1$ for $i = 1, 2$.) In this case, there exists in general no simple relationship between the joint distribution (or indeed both marginal distributions) of $(Y_0, Y_1)$ and the distributions of the $X_i$; see Denisov and Zwart (2007) and Embrechts and Goldie (1980) for discussions of the distribution of the product of two factors with the same index of regular variation.*

*(ii) 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 a lower-dimensional linear program with a unique solution. Think, for example, of $n = 3, \alpha_i = \beta_i = 1, i = 2, 3, \max\{\alpha_1, \beta_1\} < 1$. If we define $\hat{X}_1 = X_1, \hat{X}_2 = X_2X_3$, then $\hat{X}_2$ is regularly varying with index $-1$ by the Corollary to Theorem 3 in Embrechts and Goldie (1980), the resulting linear program has a unique solution $\kappa_1 = 0, \kappa_2 = 1$ and $P(Y_0 > s_0x, Y_1 > s_1x) \sim E(\min\{s_0^{-1}X_1^{\alpha_1}, s_1^{-1}X_1^{\beta_1}\})P(X_2X_3 > x)$. However, the probability on the right hand side cannot be easily expressed in terms of tail probabilities of $X_2$ and $X_3$.*

*(iii) In the situation of part (ii) of Theorem 4.2 with $\alpha_i = \beta_i$ and in case (i), the random vector $(Y_0, Y_1)$ is regularly varying on the cone $(0,\infty)^2$ with limiting measure $\nu$ given by*

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

*in the situation of (i), and by*

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

*in (ii), for $s_0, s_1 > 0$.*
*If (ii) holds with $\alpha_i \neq \beta_i$, then the measure $\nu$ that is defined analogously by (1.5) would be concentrated on $(\{\infty\}\times(0,\infty))\cup((0,\infty)\times\{\infty\})$, which is not allowed in our definition of regular variation on $(0,\infty)^2$. 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.*

Under some additional assumptions, we can extend Theorem 4.2 to an infinite number of factors.

**Theorem 4.4.** *Let $X_i$, $i\in\mathbb{N}$, be i.i.d. non-negative random variables which are regularly varying with index $-1$. Assume that $\alpha_i, \beta_i \geq 0$, $i\in\mathbb{N}$, are such that $\sup_{i\in\mathbb{N}}\alpha_i > 0, \sup_{i\in\mathbb{N}}\beta_i > 0, \sum_{i=1}^\infty \alpha_i < \infty$ and $\sum_{i=1}^\infty \beta_i < \infty$. Then $\prod_{i=1}^n X_i^{\alpha_i}$ and $\prod_{i=1}^n X_i^{\beta_i}$ converge almost surely for $n\to\infty$. We assume that the limiting random variables $Y_0$ and $Y_1$, respectively, are strictly positive almost surely.*

*Then the optimization problem*

$$\sum_{i=1}^\infty \kappa_i \to \min! \tag{4.11}$$

*under the constraints*

$$\sum_{i=1}^\infty \alpha_i\kappa_i \geq 1, \quad \sum_{i=1}^\infty \beta_i\kappa_i \geq 1, \quad \kappa_i \geq 0 \;\; \forall\, i\in\mathbb{N}, \tag{4.12}$$

*has a solution $(\kappa_i)_{i\in\mathbb{N}}$. If the solution is unique, then at most two of the $\kappa_i$ are strictly positive and $\min\{Y_0, Y_1\}$ is regularly varying with index $-\sum_{i=1}^\infty \kappa_i$. Moreover,*

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

$$\lim_{x\to\infty}\frac{P(Y_0 > s_0 x, Y_1 > s_1 x)}{P(X_i > x^{\kappa_i})P(X_j > x^{\kappa_j})} = D s_0^{(\beta_i-\beta_j)/(\alpha_i\beta_j-\alpha_j\beta_i)} s_1^{(\alpha_j-\alpha_i)/(\alpha_i\beta_j-\alpha_j\beta_i)} \tag{4.13}$$

*with*

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

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

$$\lim_{x\to\infty} \frac{P(Y_0 > s_0 x, Y_1 > s_1 x)}{P(X_i > x^{\kappa_i})} = \begin{cases} E\left(\min\left(s_0^{-1}\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\alpha_j}, s_1^{-1}\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\beta_j}\right)^{1/\beta_i}\right), & \alpha_i = \beta_i \\ E\left(\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\beta_j/\beta_i}\right) s_1^{-1/\beta_i}, & \alpha_i > \beta_i \\ E\left(\prod_{j\in\mathbb{N}\setminus\{i\}} X_j^{\alpha_j/\alpha_i}\right) s_0^{-1/\alpha_i}, & \alpha_i < \beta_i; \end{cases} \tag{4.14}$$

*(iii)* *in any case, for all $\epsilon > 0$,*

$$\begin{aligned} P(Y_0 > x, Y_1 > x) &= o\big(x^{\epsilon - \sum_{i=1}^\infty \kappa_i}\big) \\ x^{-\epsilon-\sum_{i=1}^\infty \kappa_i} &= o(P(Y_0 > x, Y_1 > x)) \end{aligned}$$

*as $x \to \infty$.*

**Remark 4.5.** *(i) 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 $P(X_1 \le x) = o(|\log x|^{-(1+\epsilon)})$ as $x \downarrow 0$ for some $\epsilon > 0$.*

*(ii) Theorem 4.4 can be easily extended to independent, not necessarily identically distributed random variables $X_i$. To ensure the convergence of the products and a moment bound of similar type as (6.37) and (6.38), it suffices to assume that $\sup_{i\in\mathbb{N}} E(X_i^\epsilon) < \infty$ for some $\epsilon > 0$. The limiting random variables $Y_0$ and $Y_1$ are strictly positive almost surely if, in addition, one assumes that $\sup_{i\in\mathbb{N}} E(\log X_i)^- < \infty$.*

*(iii) Remarks 4.3 (ii) and (iii) to Theorem 4.2 apply to Theorem 4.4 analogously.*

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

The results from the two previous sections allow 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*

$$\sum_{i=0}^\infty \kappa_i \to \min! \tag{5.1}$$

*under the constraints*

$$\sum_{i=h}^\infty \alpha_{i-h}\kappa_i \ge 1, \quad \sum_{i=0}^\infty \alpha_i\kappa_i \ge 1, \quad \kappa_i \ge 0 \; \forall\, i \in \mathbb{N}_0. \tag{5.2}$$

*In addition, suppose that $E((\epsilon_0^+)^{\sum_{i=0}^\infty \kappa_i + \delta}) < \infty$ for some $\delta > 0$.*

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

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

*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 $s_0, s_h > 0$*

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

*and*

$$\lim_{x\to\infty} \frac{P(X_0 > s_0 x, X_h > s_h x)}{P(\min\{X_0, X_h\} > x)} = \frac{E\left(\min\left(s_0^{-1}\epsilon_0^+\prod_{j\geq h, j\neq i} \mathrm{e}^{\eta_{h-j}\alpha_{j-h}}, s_h^{-1}\epsilon_h^+\prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\eta_{h-j}\alpha_j}\right)^{1/\alpha_i}\right)}{E\left(\min\left(\epsilon_0^+\prod_{j\geq h, j\neq i} \mathrm{e}^{\eta_{h-j}\alpha_{j-h}}, \epsilon_h^+\prod_{j\in\mathbb{N}_0\setminus\{i\}} \mathrm{e}^{\eta_{h-j}\alpha_j}\right)^{1/\alpha_i}\right)}.$$

*In both cases, $(\sigma_0, \sigma_h)$ 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 (5.2), the optimal solution satisfies $\sum_{i=0}^\infty \kappa_i \leq 2$. Hence the moment condition on $\epsilon_0^+$ is always fulfilled if $E((\epsilon_0^+)^{2+\delta}) < \infty$ for some $\delta > 0$.*

*(ii) It is hardly possible to derive any specific properties of the (set of) solutions to the optimization problem (5.1), (5.2) for arbitrary sequences of coefficients $\alpha_i$, $i \in \mathbb{N}_0$. However, some 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.*

*(iii) Again, a similar result holds for $(|X_0|, |X_h|)$ instead of $(X_0, X_h)$ with $\epsilon_i^+$ replaced by $|\epsilon_i|$.*

Although in general the optimal solution of the linear program (5.1), (5.2) must be computed numerically, the solution is easily determined if the coefficients are strictly decreasing.

**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 $\alpha_i, i \in \mathbb{N}_0$, strictly decreasing (which implies $\alpha_0 = 1$). Then the unique solution to (5.1), (5.2) is given by $\kappa_0 = 1 - \alpha_h, \kappa_h = 1$ and $\kappa_i = 0$ else. Furthermore, if $E((\epsilon_0^+)^{2-\alpha_h+\delta}) < \infty$ 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 $(\sigma_0, \sigma_h)$ 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 $\alpha_h$ 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 $\alpha_h, h \in \mathbb{N}$, (respectively 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 $\eta_h$ 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 $(\sigma_0, \sigma_h)$ and $(X_0, X_h)$ equals $\eta_h$ 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 $\beta$ in (3.3) whether our model allows for asymptotic dependence of lagged observations.*

*(i) If $\beta > -1$, then all vectors $(\sigma_0, \sigma_h)$ 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 $(\sigma_0, \sigma_h)$ 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 (2013), Sections 3 and 4, who consider a similar class of SV models. They analyze the limit distributions

$$\lim_{x\to\infty} P(\sigma_0 \le s_0 x, \sigma_1 \le s_1 x^{\rho_1}, \dots, \sigma_h \le s_h x^{\rho_h} \mid \sigma_0 > x) \tag{5.3}$$

for a suitable choice of so-called "conditional scaling exponents" $\rho_j, j \in \mathbb{N}$, 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 (2013) work in the framework of conditional extreme value models, which are discussed, e.g., in Das and Resnick (2011). 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 (2009)). Kulik and Soulier (2013) consider an AR(1) model for the log-volatilities where the innovations $\eta_t, t \in \mathbb{Z}$, have a double exponential distribution, i.e. 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 long memory) and that $\alpha_0 = 1, \alpha_i < 1$ for all $i \geq 1$.

Kulik and Soulier show that for those models a nondegenerate limit in (5.3) exists if and only if the conditional scaling exponents $\rho_j$ are chosen equal to $\alpha_j$ 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 (5.3) 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 (5.3) 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 $\sigma_0 = \prod_{j=0}^{\infty} \mathrm{e}^{\alpha_j \eta_{-j}}$ is most likely caused by a large value of $\eta_0$. This in turn implies that, given the extreme event at time 0, the lagged volatility $\sigma_h = \prod_{j=0}^{\infty} \mathrm{e}^{\alpha_j \eta_{h-j}}$ will be roughly of the order $\mathrm{e}^{\alpha_h \eta_0}$, which yields $\rho_h = \alpha_h$ for all $h \in \mathbb{N}$.

In contrast, we consider events of the type that both $\sigma_0$ and $\sigma_h$ 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 the most probable cause for the joint exceedance. Instead one has to find combinations of two factors $\mathrm{e}^{\eta_j}$ which are both sufficiently large (though potentially smaller than the single factor considered in the conditional extreme value approach) such that both products $\prod_{j=0}^{\infty} \mathrm{e}^{\alpha_j \eta_{-j}}$ and $\prod_{j=0}^{\infty} \mathrm{e}^{\alpha_j \eta_{h-j}}$ 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 (2013) 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 (2013).

## 6. Proofs

### 6.1. Proofs to Section 2

The following inequality will be useful:

$$\begin{aligned} d(\mathbf{Ax}, \mathbb{F}^d) &= \min((\mathbf{Ax})^+) = \min(\mathbf{x}) \min\left(\left(\frac{\mathbf{Ax}}{\min(\mathbf{x})}\right)^+\right) \\ &\leq \min(\mathbf{x})\tau(\mathbf{A}) = d(\mathbf{x}, \mathbb{F}^d)\tau(\mathbf{A}) \end{aligned} \tag{6.1}$$

for all $\mathbf{x} \in \mathbb{E}^d$ and $\mathbf{A} \in \mathbb{R}^{d\times d}$.

**Proof of Lemma 2.1**. 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 $\mathbf{A}^{-1}(\mathbb{E}^d) \cap \mathbb{E}^d = \emptyset$.

Now suppose that $\tau(\mathbf{A}) = \infty$. Then there exists $\mathbf{x} \in \mathcal{S}^d$ such that $\min(\mathbf{Ax}) > -\min(-\mathbf{A1})$ where $\mathbf{1} = (1, \dots, 1)'$. Thus $y := \mathbf{x} - \mathbf{1} \in \mathcal{O}^d \subset (\mathbb{E}^d)^c$ 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) implies (ii).

For the converse implication "(ii) $\Rightarrow$ (i)", first note that $\mathbf{A}$ must be invertible if $\tau(\mathbf{A}) \in (0, \infty)$. To see this, suppose $\mathbf{A}$ 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 $\mathbf{Ax} \in \mathbb{E}^d$. 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 $\lambda$ between 0 and $\lambda_0$. As $(\mathbf{x} + \lambda\mathbf{y})/\min(\mathbf{x} + \lambda\mathbf{y}) \in \mathcal{S}^d$ for such values of $\lambda$, one may conclude

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

as $\lambda \to \lambda_0$. Since this contradicts the assumption $\tau(\mathbf{A}) \in (0, \infty)$, the matrix $\mathbf{A}$ must be invertible.

Next, recall that with $\tau(\mathbf{A}) > 0$ there exists at least one $\mathbf{x} \in \mathbb{E}^d$ such that $\mathbf{A}^{-1}\mathbf{x} \in \mathbb{E}^d$ and thus $d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) > 0$. Now, for all $\mathbf{x} \in \mathbb{E}^d$ with $d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) > 0$

$$1 = \frac{d((\mathbf{A} \cdot \mathbf{A}^{-1})\mathbf{x}, \mathbb{F}^d)}{d(\mathbf{x}, \mathbb{F}^d)} = \frac{d(\mathbf{A}(\mathbf{A}^{-1}\mathbf{x}), \mathbb{F}^d)}{d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d)} \frac{d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d)}{d(\mathbf{x}, \mathbb{F}^d)},$$

where, according to (6.1), the first factor is bounded by $\tau(\mathbf{A})$. Therefore,

$$d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) = 0 \quad \text{or} \quad d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) \geq \tau(\mathbf{A})^{-1}d(\mathbf{x}, \mathbb{F}^d) \tag{6.2}$$

for all $\mathbf{x} \in \mathbb{E}^d$. Now let $\delta > 0$. Since $d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) > 0$ for some $\mathbf{x} \in \mathbb{E}^d$, there also exists an $\mathbf{x}$ with $d(\mathbf{x}, \mathbb{F}^d) > \delta$ and $d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) > 0$ by linearity of $\mathbf{x} \mapsto \mathbf{A}^{-1}\mathbf{x}$. Since $\mathbf{x} \mapsto d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d)$ is a continuous function and $\tau(\mathbf{A})^{-1}d(\mathbf{x}, \mathbb{F}^d)$ is bounded away from 0 on the set $C_\delta := \left\{\mathbf{x} \in \mathbb{E}^d \mid d(\mathbf{x}, \mathbb{F}^d) > \delta\right\}$, it follows from (6.2) that the inequality in (6.2), which holds for some $\mathbf{x} \in C_\delta$, must be satisfied for all $\mathbf{x} \in C_\delta$. Let $\delta$ tend to 0 to see that

$$d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) \geq \frac{d(\mathbf{x}, \mathbb{F}^d)}{\tau(\mathbf{A})} \tag{6.3}$$

for all $\mathbf{x} \in \mathbb{E}^d$.

As the right hand side of this inequality is by assumption strictly positive for all $x \in \mathbb{E}^d$ we have $\mathbf{A}^{-1}(\mathbb{E}^d) \subset \mathbb{E}^d$. Continuity of the linear mapping $\mathbf{x} \mapsto \mathbf{A}^{-1}\mathbf{x}$ then gives $\mathbf{A}^{-1}([0,\infty)^d) \subset [0,\infty)^d$ as well, which concludes the proof. □

**Proof of Lemma 2.2.** Since $0 < \tau(\mathbf{A}) < \infty$ and $d(t\mathbf{x}, \mathbb{F}^d) = td(\mathbf{x}, \mathbb{F}^d)$ for all $\mathbf{x} \in \mathbb{E}^d$ and $t > 0$, it follows from the definition of $\tau(\mathbf{A})$ that

$$\inf_{\mathbf{x} \in [0,\infty)^d : d(\mathbf{A}\mathbf{x}, \mathbb{F}^d) = \tau(\mathbf{A})} d(\mathbf{x}, \mathbb{F}^d) = 1. \tag{6.4}$$

Recall that $\mathcal{O}^d = \{\mathbf{x} \in [0,\infty)^d : \min(\mathbf{x}) = 0\}$ denotes the boundary of $\mathbb{E}^d$ and note that

$$\{\mathbf{x} \in \mathbb{E}^d : d(\mathbf{x}, \mathbb{F}^d) = y\} = y \cdot \mathbf{1} + \mathcal{O}^d = \{y \cdot \mathbf{1} + \mathbf{x} : \mathbf{x} \in \mathcal{O}^d\}$$

for all $0 < y < \infty$. Now, for $\mathbf{x} \in [0,\infty)^d$,

$$d(\mathbf{A}\mathbf{x}, \mathbb{F}^d) = \tau(\mathbf{A}) \quad \Leftrightarrow \quad \mathbf{A}\mathbf{x} \in \tau(\mathbf{A}) \cdot \mathbf{1} + \mathcal{O}^d \quad \Leftrightarrow \quad \mathbf{x} \in \tau(\mathbf{A}) \cdot \mathbf{A}^{-1}\mathbf{1} + \mathbf{A}^{-1}\mathcal{O}^d.$$

Recall that $0 < \tau(\mathbf{A}) < \infty$ implies $\min(\mathbf{A}^{-1}\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{O}^d$ by the last statement of Lemma 2.1. Therefore, for points of the form $\tau(\mathbf{A}) \cdot \mathbf{A}^{-1}\mathbf{1} + \mathbf{A}^{-1}\mathbf{z}$ with $\mathbf{z} \in \mathcal{O}^d$ the minimum distance to $\mathbb{F}^d$ is attained for $\mathbf{z} = \mathbf{0}$. Thus, Equation (6.4) implies that

$$1 = \inf_{\mathbf{x} \in [0,\infty)^d : d(\mathbf{A}\mathbf{x}, \mathbb{F}^d) = \tau(\mathbf{A})} d(\mathbf{x}, \mathbb{F}^d) = d\big(\tau(\mathbf{A}) \cdot \mathbf{A}^{-1}\mathbf{1}, \mathbb{F}^d\big) = \tau(\mathbf{A}) d\big(\mathbf{A}^{-1}\mathbf{1}, \mathbb{F}^d\big),$$

which proves the first equation. The second equation follows from $d(\mathbf{x}, \mathbb{F}^d) = \min(\mathbf{x}^+)$. □

**Proof of Theorem 2.4.** Clearly,

$$P(\mathbf{AZ} \in xB) = P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) > M) + P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) \leq M) \tag{6.5}$$

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

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

where the second inequality follows from (6.1). 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 (2.2) yields

$$\limsup_{x \to \infty} \frac{P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) > M)}{P(\min(\mathbf{Z}) > x)} \leq \delta_B^{-\alpha} E\left(\tau(\mathbf{A})^\alpha \mathbb{1}_{\{\tau(\mathbf{A}) > M\}}\right). \tag{6.6}$$

As $d(B, \mathbb{F}^d) > \delta_B > 0$, the image of $B$ under $\mathbf{A}^{-1}$ is again a.s. bounded away from $\mathcal{O}^d$, since by (6.3) $d(\mathbf{A}^{-1}\mathbf{x}, \mathbb{F}^d) \geq \tau(\mathbf{A})^{-1}\delta_B$ for all $\mathbf{x} \in B$. Hence, for the second summand in (6.5) we obtain

$$\begin{aligned}
&\lim_{x\to\infty} \frac{P(\mathbf{AZ} \in xB, \tau(\mathbf{A}) \leq M)}{P(\min(\mathbf{Z}) > x)} \\
&= \lim_{x\to\infty} \int_{\{\tau(\mathbf{a})\leq M\}} \frac{P(\mathbf{Z} \in x\mathbf{a}^{-1}B)}{P(\min(\mathbf{Z}) > x)}\, dP^{\mathbf{A}}(\mathbf{a}) \qquad (6.7)\\
&= \int_{\{\tau(\mathbf{a})\leq M\}} \nu(\mathbf{a}^{-1}B)\, dP^{\mathbf{A}}(\mathbf{a}) \\
&= E(\nu(\mathbf{A}^{-1}B)\mathbb{1}_{\{\tau(\mathbf{A})\leq M\}}). \qquad (6.8)
\end{aligned}$$

For the second equation, we have used that $E(\nu(\partial\mathbf{A}^{-1}B)) = 0$ (i.e. $\nu(\partial\mathbf{A}^{-1}B) = 0$ a.s.) and Pratt's lemma (cf. Pratt (1960)), since the integrand in (6.7) is bounded by

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

Let $M \to \infty$ in (6.6) and (6.8) to obtain (2.3) by monotone convergence. Note that $E(\nu(\mathbf{A}^{-1}B))$ is finite since, by the definition of $\tau$,

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

□

### 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 $k < \infty$),

$$\hat{\beta} := \begin{cases} k\beta + k - 1, & \beta > -1 \\ \beta, & \beta < -1 \end{cases} \qquad (6.9)$$

and

$$\hat{K} := \begin{cases} K^k \frac{\Gamma(\beta+1)^k}{\Gamma(k(\beta+1))} E\left(\exp\left(\sum_{n\notin\Lambda} \alpha_n\eta_n\right)\right), & \beta > -1 \\ kKE(\exp(\eta_0))^{k-1} E\left(\exp\left(\sum_{n\notin\Lambda} \alpha_n\eta_n\right)\right), & \beta < -1 \end{cases} \qquad (6.10)$$

(cf. equations (7.8) and (7.9) in Rootzén (1986)). It follows from Lemma 7.2 in Rootzén (1986) that

$$P(\log(\sigma_t) > z) = P\left(\sum_{i=0}^{\infty} \alpha_i\eta_{t-i} > z\right) \sim \hat{K}z^{\hat{\beta}}\mathrm{e}^{-z}, \quad z \to \infty, \quad t \in \mathbb{Z}. \qquad (6.11)$$

Since the $\epsilon_t, t \in \mathbb{Z}$, are assumed to be independent of the $\eta_t, t \in \mathbb{Z}$, 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 (6.11) that

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

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 $|X_0| = \sigma_0|\epsilon_0|$ and $X_0^+ = \sigma_0\epsilon_0^+$ follows by Breiman's lemma, which (under our moment assumptions on $\epsilon_0$) gives

$$P(\sigma_0|\epsilon_0| > x) \sim E(|\epsilon_0|)P(\sigma_0 > x) \quad \text{and} \quad P(\sigma_0\epsilon_0^+ > x) \sim E(\epsilon_0^+)P(\sigma_0 > x).$$

Now the tail balance assertion (3.4) and relation (3.5) between the tails of $X_0$ and $\sigma_0$ are obvious. □

**Proof of Theorem 3.5.**

**Proof of (i)** It follows from Theorem 7.4 in Rootzén (1986) that the point processes

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

converge weakly to a Poisson process $N^{\log(\sigma)}$ on $[0,1] \times (-\infty, \infty]$ with intensity measure $dt \times \mathrm{e}^{-x}dx$. Now, $\exp(\cdot)$ is a continuous function such that the preimage of a set $B$ is bounded away from $-\infty$ if $B$ is bounded away from 0. We may thus apply Proposition 5.5 in Resnick (2007) to derive that $N_n^\sigma$ converges in $M_p([0,1] \times (0,\infty])$ to a Poisson process with intensity $dt \times z^{-2}dz$ on $[0,1] \times (0,\infty]$ which we denote by $N^\sigma$.

**Proof of (ii)** To derive the second assertion, for $n \in \mathbb{N}$, introduce the point process $N_n^{(\sigma,\epsilon)}$ which consists of the points $(i/n, \sigma_i/a_n, \epsilon_i), i = 1, 2, \ldots, n$. The first assertion implies that $N_n^{(\sigma,\epsilon)}$ converges weakly to the Poisson process $N^{(\sigma,\epsilon)}$ with points $(t_{(i)}, z_{(i)}, \epsilon_{(i)})_{i\in\mathbb{N}}$ which has the intensity $dt \times z^{-2}dz \times dP^{\epsilon_0}$. One may now proceed similarly as in the proof of Proposition 7.5 in Resnick (2007) to show that the point processes $N_n^X, n \in \mathbb{N}$, which consist of the points $(i/n, a_n^{-1}\sigma_i \cdot \epsilon_i), i = 1, 2, \ldots, n$, converge to a Poisson point process $N^X$ with points $(t_{(i)}, z_{(i)} \cdot \epsilon_{(i)})_{i\in\mathbb{N}}$: Note that the additional first component $t_{(i)}$ of the points, the dependence between the $\sigma_i, i = 1, 2, \ldots,$ and the possibly negative sign of the $\epsilon_i$ do not cause substantial changes in the course of the proof. The derivation of the stated intensity $dt \times z^{-2}\left[\mathbb{1}_{\{z<0\}}E(\epsilon_0^-) + \mathbb{1}_{\{z>0\}}E(\epsilon_0^+)\right]dz$ of the process $N^X$ follows by an application of Proposition 5.2 in Resnick (2007) to the continuous mapping

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

in combination with a truncation argument like in step 4 of the proof of Proposition 7.5 in Resnick (2007). □

### 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 $\alpha, \beta$ such that $0 \leq \beta < \min\{1, \alpha\}$,*

$$\lim_{x\to\infty} \frac{P(X > x, Y^\alpha X^\beta > x)}{P(X > x)P(Y^\alpha > x^{1-\beta})} = \frac{\alpha}{\alpha - \beta}. \tag{6.12}$$

*Furthermore, for each $\epsilon > 0$ there exists an $M = M(\epsilon) > 0$ such that*

$$(1-\epsilon)\frac{\alpha}{\alpha-\beta}s^{-\frac{1}{\alpha}-\epsilon(s)} \leq \frac{P(X > x, Y^\alpha X^\beta > sx)}{P(X > x)P(Y^\alpha > x^{1-\beta})} \leq (1+\epsilon)\frac{\alpha}{\alpha-\beta}s^{-\frac{1}{\alpha}+\epsilon(s)} \tag{6.13}$$

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

**Proof**. If $\beta = 0$ then the first statement follows directly from the independence of $X$ and $Y$. The second statement follows 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 $\beta > 0$. We only have to show (6.13), since then we conclude (6.12) by choosing $s = 1$ and $\epsilon \searrow 0$. To this end, check that

$$\begin{aligned}
&\frac{P(X > x, Y^\alpha X^\beta > sx)}{P(X > x)P\left(Y^\alpha > x^{1-\beta}\right)} \\
&\quad= \frac{\int_0^\infty P\left(X^\beta > \max\{sx/y, x^\beta\}\right)\, P^{Y^\alpha}(dy)}{P(X > x)P\left(Y^\alpha > x^{1-\beta}\right)} \\
&\quad= \int_0^{sx^{1-\beta}} \frac{P\left(X^\beta > sx/y\right)}{P(X > x)P\left(Y^\alpha > x^{1-\beta}\right)}\, P^{Y^\alpha}(dy) \\
&\qquad + \int_{sx^{1-\beta}}^\infty \frac{P\left(X^\beta > x^\beta\right)}{P(X > x)P\left(Y^\alpha > x^{1-\beta}\right)}\, P^{Y^\alpha}(dy).
\end{aligned} \tag{6.14}$$

The second summand equals

$$\frac{P(Y^\alpha > sx^{1-\beta})}{P(Y^\alpha > x^{1-\beta})} \in \left[(1-\epsilon)s^{-1/\alpha-\epsilon(s)}, (1+\epsilon)s^{-1/\alpha+\epsilon(s)}\right] \tag{6.15}$$

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

$$\begin{aligned}
&\int_0^{sx^{1-\beta}} \frac{P\left(X > (sx/y)^{1/\beta}\right)}{P(X > x)}\, P^{Y^\alpha}(dy) \\
&\quad\leq (1+\epsilon) \int_0^{sx^{1-\beta}} \left(\frac{(sx/y)^{1/\beta}}{x}\right)^{-1+\epsilon} P^{Y^\alpha}(dy)
\end{aligned}$$

$$= \quad (1+\epsilon)(sx^{1-\beta})^{-(1-\epsilon)/\beta} \int_0^{sx^{1-\beta}} y^{(1-\epsilon)/\beta} \, P^{Y^\alpha}(dy), \tag{6.16}$$

because $(sx/y)^{1/\beta} > x$. Since the distribution of $Y^\alpha$ is regularly varying with index $-1/\alpha$ and $(1-\epsilon)/\beta > 1/\alpha$ for sufficiently small $\epsilon > 0$ by assumption, a generalization of Karamata's theorem (Bingham et al. (1987), Theorem 1.6.4) yields

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

Thus, for a suitable $N'' = N''(\epsilon)$ and $x > N''$, the integral in (6.16) is bounded from above by

$$(1+\epsilon)\frac{\beta}{\alpha(1-\epsilon)-\beta}(sx^{1-\beta})^{(1-\epsilon)/\beta} P\left(Y^\alpha > sx^{1-\beta}\right).$$

Hence, by (6.15), the first summand in (6.14) is bounded by

$$(1+\epsilon)^3 \frac{\beta}{\alpha(1-\epsilon)-\beta} s^{-1/\alpha+\epsilon(s)} \tag{6.17}$$

for $x$ large enough. It follows from (6.15) and (6.17) that for given $\epsilon > 0$ one may find a suitable $\delta > 0$ and a corresponding constant $M(\delta)$ such that (with $\delta(s)$ defined analogously to $\epsilon(s)$)

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

for $\min\{sx^{1-\beta}, x^{1-\beta}\} > M(\delta)$ which gives the upper bound in (6.13).

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

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

as lower bounds for the first and second summand in (6.14), respectively, which leads to the lower bound in (6.13). □

**Remark 6.2.** *(i) Note that equation* (6.12) *of Lemma 6.1 follows immediately if we assume that $X$ has a unit Pareto distribution, since then*

$$\lim_{x\to\infty} \frac{P(X > x, Y^\alpha X^\beta > x)}{P(X > x)P(Y^\alpha > x^{1-\beta})} \quad = \quad \lim_{x\to\infty} \frac{P\left(Y^\alpha \left(\frac{X}{x}\right)^\beta > x^{1-\beta} \,|\, X > x\right)}{P(Y^\alpha > x^{1-\beta})}$$

$$= \lim_{x\to\infty} \frac{P\left(Y^\alpha X^\beta > x^{1-\beta}\right)}{P(Y^\alpha > x^{1-\beta})}$$
$$= E(X^{\beta/\alpha}) = \frac{\alpha}{\alpha-\beta}$$

*by Breiman's lemma.*

(ii) *The inequalities (6.13) can be interpreted as a special case of Potter type bounds for the function $(x,y) \mapsto P(X > x, Y^\alpha X^\beta > y)$ which is bivariate regularly varying on $(0,\infty)^2$ (cf. Resnick (1987), Equation (5.32)). We will use these bounds to extend our analysis to products of more than two random variables.*

**Proof of Proposition 4.1**. By our assumptions the solution $(\kappa_1, \kappa_2)$ 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 $(\alpha_i, \beta_i)$, $i \in \{1, 2\}$, 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 $(\alpha_1, \beta_1)$ lies below or on the main diagonal, i.e. $\alpha_1 \geq \beta_1$. (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

$$\kappa_1\alpha_1 + \kappa_2\alpha_2 = 1, \quad \kappa_1\beta_1 + \kappa_2\beta_2 = 1 \tag{6.18}$$

must hold, so that $\kappa_1$ and $\kappa_2$ are determined by

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

Now, write

$$\frac{P(X_1^{\alpha_1}X_2^{\alpha_2} > s_0x, X_1^{\beta_1}X_2^{\beta_2} > s_1x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} = \frac{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \times \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x)}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}, \tag{6.19}$$

with $z_1, z_2$ as in (4.4). Note that, by the Potter bounds, the first factor on the right hand side belongs to

$$\left[(1-\epsilon)z_1^{1-\epsilon(z_1)}z_2^{1-\epsilon(z_2)}, (1+\epsilon)z_1^{1+\epsilon(z_1)}z_2^{1+\epsilon(z_2)}\right] \tag{6.20}$$

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

The second factor in (6.19) is split up into four terms:

$$\frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x, z_1X_1 \leq x^{\kappa_1}, z_2X_2 \leq x^{\kappa_2})}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}$$

$$
\begin{aligned}
&- \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x, z_1X_1 > x^{\kappa_1}, z_2X_2 > x^{\kappa_2})}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\
&+ \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x, z_1X_1 > x^{\kappa_1})}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\
&+ \frac{P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x, z_2X_2 > x^{\kappa_2})}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}. \qquad (6.21)
\end{aligned}
$$

For $x \geq 1$ the first and second summand equal zero and $-1$, respectively, because of (6.18).

Next, note that, again by (6.18),

$$
\begin{aligned}
&P((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2} > x, z_1X_1 > x^{\kappa_1}) \\
&= P\left(\left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\alpha_1}(z_2X_2)^{\alpha_2} > x^{1-\alpha_1\kappa_1}, \left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\beta_1}(z_2X_2)^{\beta_2} > x^{1-\beta_1\kappa_1}, z_1X_1 > x^{\kappa_1}\right) \\
&= P\left(\left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\alpha_1}(z_2X_2)^{\alpha_2} > x^{\alpha_2\kappa_2}, \left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\beta_1}(z_2X_2)^{\beta_2} > x^{\beta_2\kappa_2}, z_1X_1 > x^{\kappa_1}\right) \\
&= P\left(\left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\alpha_1/\alpha_2} z_2X_2 > x^{\kappa_2}, \left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\beta_1/\beta_2} z_2X_2 > x^{\kappa_2}, z_1X_1 > x^{\kappa_1}\right),
\end{aligned}
$$

where $(z_1X_1/x^{\kappa_1})^{\alpha_1/\alpha_2} := \infty$ 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

$$
\begin{aligned}
&P\left(\left(\frac{z_1X_1}{x^{\kappa_1}}\right)^{\beta_1/\beta_2} z_2X_2 > x^{\kappa_2}, z_1X_1 > x^{\kappa_1}\right) \\
&= P\left((z_1X_1)^{\beta_1/\beta_2} z_2X_2 > x^{\kappa_2+\kappa_1\beta_1/\beta_2}, z_1X_1 > x^{\kappa_1}\right) \\
&= P\left((z_1X_1)^{\beta_1\kappa_1}(z_2X_2)^{\beta_2\kappa_1} > x^{\kappa_1}, z_1X_1 > x^{\kappa_1}\right).
\end{aligned}
$$

Now, we substitute $u$ for $x^{\kappa_1}/z_1$, use (6.18) and apply Lemma 6.1 and the Potter bounds to $x \mapsto P(X_2^{\beta_2\kappa_1} > x)$:

$$
\begin{aligned}
&\frac{P\left(z_1X_1 > x^{\kappa_1}, (z_1X_1)^{\beta_1\kappa_1}(z_2X_2)^{\beta_2\kappa_1} > x^{\kappa_1}\right)}{P(z_1X_1 > x^{\kappa_1})\,P(z_2X_2 > x^{\kappa_2})} \\
&= \frac{P\left(X_1 > u, X_1^{\beta_1\kappa_1}X_2^{\beta_2\kappa_1} > z_1^{\beta_2\kappa_2}z_2^{-\beta_2\kappa_1}u\right)}{P(X_1 > u)P(X_2^{\beta_2\kappa_1} > u^{1-\beta_1\kappa_1})} \frac{P(X_2^{\beta_2\kappa_1} > u^{1-\beta_1\kappa_1})}{P(X_2^{\beta_2\kappa_1} > z_1^{\beta_2\kappa_2}z_2^{-\beta_2\kappa_1}u^{1-\beta_1\kappa_1})} \\
&\in \Bigg[(1-\eta)^2\frac{\beta_2}{\beta_2-\beta_1}\min\left\{\left(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\right)^{2\eta}, \left(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\right)^{-2\eta}\right\}, \\
&\qquad (1+\eta)^2\frac{\beta_2}{\beta_2-\beta_1}\max\left\{\left(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\right)^{2\eta}, \left(\frac{z_1^{\beta_2\kappa_2}}{z_2^{\beta_2\kappa_1}}\right)^{-2\eta}\right\}\Bigg]
\end{aligned}
$$

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 $\eta$ to each $\epsilon > 0$, we may hence find an $M'(\epsilon)$ such that

$$\begin{aligned} &\frac{P\left(z_1X_1 > x^{\kappa_1}, (z_1X_1)^{\beta_1\kappa_1}(z_2X_2)^{\beta_2\kappa_1} > x^{\kappa_1}\right)}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\ &\in \left[(1-\epsilon)\frac{\beta_2}{\beta_2-\beta_1} z_1^{-\epsilon(z_1)} z_2^{-\epsilon(z_2)}, (1+\epsilon)\frac{\beta_2}{\beta_2-\beta_1} z_1^{\epsilon(z_1)} z_2^{\epsilon(z_2)}\right] \end{aligned} \tag{6.22}$$

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

Analogously, the fourth summand in equation (6.21) equals

$$\begin{aligned} &\frac{P\left(\left(\frac{z_2X_2}{x^{\kappa_2}}\right)^{\alpha_2/\alpha_1} z_1X_1 > x^{\kappa_1}, z_2X_2 > x^{\kappa_2}\right)}{P(z_2X_2 > x^{\kappa_2})P(z_1X_1 > x^{\kappa_1})} \\ &= \frac{P\left(z_2X_2 > x^{\kappa_2}, (z_2X_2)^{\alpha_2\kappa_2}(z_1X_1)^{\alpha_1\kappa_2} > x^{\kappa_2}\right)}{P(z_2X_2 > x^{\kappa_2})P((z_1X_1)^{\alpha_1\kappa_2} > x^{\alpha_1\kappa_1\kappa_2})}. \end{aligned}$$

Again, a substitution (set $u = x^{\kappa_2}/z_2$) and Lemma (6.1) lead to

$$\begin{aligned} &\frac{P\left(z_2X_2 > x^{\kappa_2}, (z_2X_2)^{\alpha_2\kappa_2}(z_1X_1)^{\alpha_1\kappa_2} > x^{\kappa_2}\right)}{P(z_2X_2 > x^{\kappa_2})P((z_1X_1)^{\alpha_1\kappa_2} > x^{\alpha_1\kappa_1\kappa_2})} \\ &= \frac{P\left(X_2 > u, X_2^{\alpha_2\kappa_2}X_1^{\alpha_1\kappa_2} > z_2^{\alpha_1\kappa_1} z_1^{-\alpha_1\kappa_2} u\right)}{P(X_2 > u)P(X_1^{\alpha_1\kappa_2} > u^{1-\alpha_2\kappa_2})} \frac{P(X_1^{\alpha_1\kappa_2} > u^{1-\alpha_2\kappa_2})}{P(X_1^{\alpha_1\kappa_2} > z_2^{\alpha_1\kappa_1} z_1^{-\alpha_1\kappa_2} u^{1-\alpha_2\kappa_2})} \\ &\in \left[(1-\epsilon)\frac{\alpha_1}{\alpha_1-\alpha_2} z_1^{-\epsilon(z_1)} z_2^{-\epsilon(z_2)}, (1+\epsilon)\frac{\alpha_1}{\alpha_1-\alpha_2} z_1^{\epsilon(z_1)} z_2^{\epsilon(z_2)}\right] \end{aligned} \tag{6.23}$$

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

Finally, for $\epsilon > 0$, combining (6.19)-(6.23) we arrive at

$$\begin{aligned} &\frac{P(X_1^{\alpha_1}X_2^{\alpha_2} > s_0x, X_1^{\beta_1}X_2^{\beta_2} > s_1x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\ &\in \left[(1-\epsilon')Cz_1z_2z_1^{-\epsilon'(z_1)}z_2^{-\epsilon'(z_2)}, (1+\epsilon')Cz_1z_2z_1^{+\epsilon'(z_1)}z_2^{+\epsilon'(z_2)}\right] \\ &\subset \Big[(1-\epsilon)Cs_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)-\epsilon(s_0)} s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)-\epsilon(s_1)}, \\ &\qquad (1+\epsilon)Cs_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)+\epsilon(s_0)} s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)+\epsilon(s_1)}\Big] \end{aligned}$$

for a suitable choice of $\epsilon'$ and a constant $N = N(\epsilon)$ 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.**
**Proof of (iii)** We prove the assertion by formalizing the heuristic arguments used to motivate the linear optimization problem. Let $(\kappa_1, \ldots, \kappa_n)$ be a solution to the optimization problem. The lower bound (4.10) follows from

$$P(Y_0 > x, Y_1 > x) \geq P(X_i > x^{\kappa_i} \text{ for all } 1 \leq i \leq n) \geq \prod_{i=1}^{n} x^{-\kappa_i - \epsilon/n}$$

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

To establish (4.9), let $c := (2n)/\epsilon$ and, for $x > 1$, $Z_{i,x} := \max\left\{\lceil c \log X_i / \log x \rceil, 0\right\}$. Then $\log X_i \in (Z_{i,x} - 1, Z_{i,x}] \log x / c$ if $Z_{i,x} > 0$. Since $(\kappa_1, \ldots, \kappa_n)$ is an optimal solution of the linear program (4.5) and (4.6), it follows

$$\begin{aligned}
&P(Y_0 > x, Y_1 > x) \\
&\leq P\Big(\sum_{i=1}^{n} \alpha_i Z_{i,x}/c \geq 1, \sum_{i=1}^{n} \beta_i Z_{i,x}/c \geq 1\Big) \\
&\leq P\Big(\sum_{i=1}^{n} Z_{i,x}/c \geq \sum_{i=1}^{n} \kappa_i\Big) \\
&\leq \sum_{i=1}^{n} P\Big(Z_{i,x} > c \sum_{j=1}^{n} \kappa_j\Big) + \sum_{(j_1, \ldots, j_n)} P(Z_{i,x} = j_i \text{ for all } 1 \leq i \leq n)
\end{aligned}$$

where the second sum runs over all $j_1, \ldots, j_n \in \{0, \ldots, \lfloor c \sum_{i=1}^{n} \kappa_i \rfloor\}$ such that $\sum_{i=1}^{n} j_i \geq c \sum_{i=1}^{n} \kappa_i$. As these are only finitely many and both

$$\begin{aligned}
P(Z_{i,x} = j_i \text{ for all } 1 \leq i \leq n) &\leq \prod_{i=1}^{n} (P(X_i > x^{(j_i - 1)/c} \mathbb{1}_{\{j_i > 0\}} + \mathbb{1}_{\{j_i = 0\}}) \\
&= o(x^{-\sum_{i=1}^{n} j_i / c + \epsilon}) \\
&= o(x^{-\sum_{i=1}^{n} \kappa_i + \epsilon}),
\end{aligned}$$

and

$$P\Big(Z_{i,x} > c \sum_{j=1}^{n} \kappa_j\Big) \leq P\Big(X_i > x^{\sum_{j=1}^{n} \kappa_j - 1/c}\Big) = o(x^{-\sum_{j=1}^{n} \kappa_j + \epsilon}),$$

we obtain the upper bound (4.9) which concludes the proof of (iii).

Since we have assumed a unique solution to (4.5) and (4.6), it must be a basic feasible solution (Sierksma (1996), Theorems 1.2 and 1.5), i.e., $\kappa_l = 0$ for all but at most two indices $l \in \{1, \ldots, n\}$.

**Proof of (ii)**
W.l.o.g. assume that $\kappa_1 > 0$ and $\kappa_j = 0$ for all $j \in \{2, \ldots, n\}$ and that $\alpha_1 \geq \beta_1$, thus $\kappa_1 = 1/\beta_1$. In this case,

$$P(Y_0 > s_0 x, Y_1 > s_1 x)$$

$$= P\left(X_1^{\alpha_1}\prod_{i=2}^n X_i^{\alpha_i} > s_0x, X_1^{\beta_1}\prod_{i=2}^n X_i^{\beta_i} > s_1x\right)$$
$$= P\left(X_1^{\beta_1}\left(s_0^{-1}\prod_{i=2}^n X_i^{\alpha_i}\right)^{\beta_1/\alpha_1} x^{1-\beta_1/\alpha_1} > x, X_1^{\beta_1}s_1^{-1}\prod_{i=2}^n X_i^{\beta_i} > x\right)$$
$$= P\left(X_1^{\beta_1}\min\left\{\left(s_0^{-1}\prod_{i=2}^n X_i^{\alpha_i}\right)^{\beta_1/\alpha_1} x^{1-\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^n X_i^{\beta_i}\right\} > x\right). \tag{6.24}$$

(Remember that in the case $n = 1$ the empty product is by our convention equal to 1.) Now, if $\alpha_1 = \beta_1$, then the second factor in the product in (6.24) equals $Z := \min\left\{s_0^{-1}\prod_{i=2}^n X_i^{\alpha_i}, s_1^{-1}\prod_{i=2}^n X_i^{\beta_i}\right\}$. If $n = 1$ then (4.8) follows directly from (6.24). Note that for $n \geq 2$

$$\sum_{i=2}^n \kappa_i' > \kappa_1 = \frac{1}{\beta_1} \text{ for all } \kappa_2', \dots, \kappa_n' \geq 0 \text{ such that } \sum_{i=2}^n \alpha_i\kappa_i' \geq 1, \ \sum_{i=2}^n \beta_i\kappa_i' \geq 1$$

since we have assumed a unique solution to (4.5), (4.6). Part (iii) of the statement (applied to $\tilde{Y}_0 := \prod_{i=2}^n X_i^{\alpha_i}$ and $\tilde{Y}_1 := \prod_{i=2}^n X_i^{\beta_i}$) thus yields that $P(Z > x) = o(x^{-1/\beta_1-\epsilon})$ for some $\epsilon > 0$, which by Breiman's lemma in turn implies (4.8).

If $\alpha_1 > \beta_1$, then $\beta_i < \beta_1$ for all $i = 2, \dots, n$, since otherwise $\kappa_1 = 1/\beta_1, \kappa_i = 0, i = 2, \dots, n$, would not be a unique solution to (4.5) and (4.6). (If $\beta_i \geq \beta_1$, then $\kappa_1' = (1-\epsilon)/\beta_1$, $\kappa_i' = \epsilon/\beta_i$, for sufficiently small $\epsilon$, and $\kappa_j' = 0$ for all $j \in \{2, \dots, n\}\setminus\{i\}$ would satisfy (4.6) with $\kappa_1' + \kappa_i' \leq \kappa_1$.) For all $C > 0$, the second factor in the product in (6.24) is eventually (for sufficiently large $x$) bounded from below by $\min\left\{C\prod_{i=2}^n X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^n X_i^{\beta_i}\right\}$. Thus,

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

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 $s_1^{-1}\prod_{i=2}^n X_i^{\beta_i}$ and we may thus again apply Breiman's lemma to get

$$E\left(\left(\min\left\{C\prod_{i=2}^n X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^n X_i^{\beta_i}\right\}\right)^{1/\beta_1}\right)$$
$$\leq \liminf_{x\to\infty}\frac{P(Y_1 > s_0x, Y_1 > s_1x)}{P(X_1^{\beta_1} > x)} \leq \limsup_{x\to\infty}\frac{P(Y_1 > s_0x, Y_1 > s_1x)}{P(X_1^{\beta_1} > x)}$$

$$\leq \quad E\left(\left(\mathbb{1}_{\{\prod_{i=2}^n X_i^{\alpha_i\beta_1/\alpha_1}>0\}} s_1^{-1} \prod_{i=2}^n X_i^{\beta_i}\right)^{1/\beta_1}\right) < \infty,$$

since $\beta_i < \beta_1$ for all $i \in \{2, \ldots, n\}$. For $C \to \infty$, the lower bound converges to the upper bound, which concludes the proof for part (ii) of the statement.

**Proof of (i) for $\mathbf{s_0 = s_1 = 1}$**

W.l.o.g. let us assume that $i = 1, j = 2$, thus $\kappa_1 > 0, \kappa_2 > 0$ and $\kappa_l = 0, l \in \{3, \ldots, n\}$. 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 (6.18) holds and it follows that

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

There exists a so-called dual problem to (4.5) and (4.6) (see Sierksma (1996), Chapter 2) that is given by

$$\hat{\kappa}_1 + \hat{\kappa}_2 \to \max! \tag{6.25}$$

under the constraints

$$\alpha_l\hat{\kappa}_1 + \beta_l\hat{\kappa}_2 \leq 1 \quad \text{for all } l = 1, \ldots, n. \tag{6.26}$$

Since we have assumed an optimal solution to the primal problem, there also exists an optimal solution to the dual problem (Sierksma (1996), Theorem 2.2), denoted by $(\hat{\kappa}_1, \hat{\kappa}_2)$. Because of $\kappa_1, \kappa_2 > 0$, it follows by the complementary slackness theorem (Sierksma (1996), Theorem 2.4) that

$$\alpha_1\hat{\kappa}_1 + \beta_1\hat{\kappa}_2 = 1 \text{ and } \alpha_2\hat{\kappa}_1 + \beta_2\hat{\kappa}_2 = 1. \tag{6.27}$$

Therefore

$$\hat{\kappa}_1 = \frac{\beta_2 - \beta_1}{\alpha_1\beta_2 - \alpha_2\beta_1} > 0, \quad \hat{\kappa}_2 = \frac{\alpha_1 - \alpha_2}{\alpha_1\beta_2 - \alpha_2\beta_1} > 0. \tag{6.28}$$

If equality held in (6.26) for some $l \geq 3$, then direct calculations show that

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

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

$$\alpha_l\hat{\kappa}_1 + \beta_l\hat{\kappa}_2 < 1 - \epsilon \text{ for all } l \in \{3, \ldots, n\}. \tag{6.29}$$

Let

$$Z_1(x_3, \ldots, x_n) \quad := \quad \left(\prod_{i=3}^n x_i^{\beta_i}\right)^{-\alpha_2/(\alpha_1\beta_2 - \alpha_2\beta_1)} \left(\prod_{i=3}^n x_i^{\alpha_i}\right)^{\beta_2/(\alpha_1\beta_2 - \alpha_2\beta_1)}$$

$$= \prod_{i=3}^{n} x_i^{(\alpha_i\beta_2-\beta_i\alpha_2)/(\alpha_1\beta_2-\alpha_2\beta_1)}$$

and

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

if $x_i > 0$ for all $i \in \{3,\dots,n\}$ with $\max\{\alpha_i,\beta_i\} > 0$ and set $Z_1(x_3,\dots,x_n) = Z_2(x_3,\dots,x_n) = \infty$ 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,

$$\mathbb{M} = \mathbb{M}(x, M) := \left\{\mathbf{x} \in [0,\infty)^{n-2} \;:\; MZ_1(\mathbf{x}) < x^{\kappa_1}, MZ_2(\mathbf{x}) < x^{\kappa_2}\right\}$$

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

Then,

$$\begin{aligned}
&\frac{P\left(\prod_{i=1}^n X_i^{\alpha_i} > x, \prod_{i=1}^n X_i^{\beta_i} > x\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\
&= \int \frac{P\left(X_1^{\alpha_1}X_2^{\alpha_2} > \prod_{i=3}^n x_i^{-\alpha_i}x, X_1^{\beta_1}X_2^{\beta_2} > \prod_{i=3}^n x_i^{-\beta_i}x\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \mathbb{1}_{\mathbb{M}}(\mathbf{x}_3^n)\, P^{\mathbf{X}_3^n}(d\mathbf{x}_3^n) \\
&\quad + \frac{P\left(\prod_{i=1}^n X_i^{\alpha_i} > x, \prod_{i=1}^n X_i^{\beta_i} > x, \mathbf{X}_3^n \in \mathbb{M}^c\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})}.
\end{aligned} \tag{6.30}$$

Let $s_0(\mathbf{x}) := \prod_{i=3}^n x_i^{-\alpha_i}$ and $s_1(\mathbf{x}) := \prod_{i=3}^n x_i^{-\beta_i}$. According to Proposition 4.1 and (6.28), the integrand of the first summand on the right hand side of (6.30) is bounded from above by

$$\begin{aligned}
&(1+\delta)Cs_0(\mathbf{x})^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)+\delta(s_0(\mathbf{x}))} s_1(\mathbf{x})^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)+\delta(s_1(\mathbf{x}))} \\
&= (1+\delta)C \prod_{i=3}^n x_i^{\alpha_i\hat\kappa_1+\beta_i\hat\kappa_2-\alpha_i\delta(s_0(\mathbf{x}))-\beta_i\delta(s_1(\mathbf{x}))}
\end{aligned}$$

for $x > M$ with $C$ as in Proposition 4.1. In view of (6.29), $\delta$ may be chosen so small such that all the exponents in the last expression are strictly less than 1, and hence the upper bound is integrable w.r.t. $P^{\mathbf{X}_3^n}$. Therefore, by dominated convergence, we see that

$$\begin{aligned}
&\lim_{x\to\infty} \int \frac{P\left(X_1^{\alpha_1}X_2^{\alpha_2} > \prod_{i=3}^n x_i^{-\alpha_i}x, X_1^{\beta_1}X_2^{\beta_2} > \prod_{i=3}^n x_i^{-\beta_i}x\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \mathbb{1}_{\mathbb{M}}(\mathbf{x}_3^n)\, P^{\mathbf{X}_3^n}(d\mathbf{x}_3^n) \\
&= \int \lim_{x\to\infty} \frac{P\left(X_1^{\alpha_1}X_2^{\alpha_2} > \prod_{i=3}^n x_i^{-\alpha_i}x, X_1^{\beta_1}X_2^{\beta_2} > \prod_{i=3}^n x_i^{-\beta_i}x\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \mathbb{1}_{\mathbb{M}}(\mathbf{x}_3^n)\, P^{\mathbf{X}_3^n}(d\mathbf{x}_3^n) \\
&= \int \prod_{i=3}^n C x_i^{\alpha_i\hat\kappa_1+\beta_i\hat\kappa_2} \mathbb{1}_{\{\max\{Z_1(\mathbf{x}),Z_2(\mathbf{x})\}<\infty\}}\, P^{\mathbf{X}_3^n}(d\mathbf{x}_3^n)
\end{aligned}$$

$$= \quad CE\left(\prod_{i=3}^{n} X_i^{\alpha_i\hat\kappa_1+\beta_i\hat\kappa_2}\right) = D \in (0,\infty), \tag{6.31}$$

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

$$\begin{aligned} &\frac{P\left(\prod_{i=1}^n X_i^{\alpha_i} > x, \prod_{i=1}^n X_i^{\beta_i} > x, \mathbf{X}_3^n \in \mathbb{M}^c\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\ &\quad\leq \quad \frac{P\left(\prod_{i=1}^n X_i^{\alpha_i} > x, \prod_{i=1}^n X_i^{\beta_i} > x, MZ_1(X_3,\dots,X_n) \geq x^{\kappa_1}\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\ &\qquad\quad + \frac{P\left(\prod_{i=1}^n X_i^{\alpha_i} > x, \prod_{i=1}^n X_i^{\beta_i} > x, MZ_2(X_3,\dots,X_n) \geq x^{\kappa_2}\right)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \end{aligned} \tag{6.32}$$

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

$$\begin{aligned} &P\left(\prod_{i=1}^n X_i^{\beta_i} > x, M^{\alpha_1\beta_2-\alpha_2\beta_1}\prod_{i=3}^n X_i^{\alpha_i\beta_2-\beta_i\alpha_2} \geq x^{\beta_2-\alpha_2}\right) \\ &\leq \quad P\left(\prod_{i=1}^n X_i^{\beta_i} > x, M^{\alpha_1\beta_2-\alpha_2\beta_1}\prod_{i=1}^n X_i^{\beta_i\alpha_2}\prod_{i=3}^n X_i^{\alpha_i\beta_2-\beta_i\alpha_2} \geq x^{\beta_2}\right) \\ &\leq \quad P\left(\prod_{i=1}^n X_i^{\beta_i} > \min\{1, M^{\alpha_2\beta_1-\alpha_1\beta_2}\}x, X_1^{\alpha_2\beta_1/\beta_2}\prod_{i=2}^n X_i^{\alpha_i} \geq \min\{1, M^{\alpha_2\beta_1-\alpha_1\beta_2}\}x\right). \end{aligned} \tag{6.33}$$

By part (iii) of the assertion the asymptotic behavior of the above probability is related to the linear program

$$\sum_{i=1}^n \tilde\kappa_i \to \min! \tag{6.34}$$

under the constraints

$$(\alpha_2\beta_1/\beta_2)\tilde\kappa_1 + \sum_{i=2}^n \alpha_i\tilde\kappa_i \geq 1, \quad \sum_{i=1}^n \beta_i\tilde\kappa_i \geq 1, \quad \tilde\kappa_i \geq 0 \;\forall\, 1 \leq i \leq n. \tag{6.35}$$

Since $\alpha_2\beta_1/\beta_2 < \alpha_1$, all optimal solutions $(\tilde\kappa_i)_{1\leq i\leq n}$ to (6.34), (6.35) also fulfill (4.6), but they cannot be optimal solutions to (4.5), (4.6), because the unique optimal solution $(\kappa_i)_{1\leq i\leq n}$ does not fulfill (6.35). Hence $\sum_{i=1}^n \tilde\kappa_i > \kappa_1 + \kappa_2 + 2\delta$ for sufficiently small $\delta > 0$ and, by assertion (iii), the right hand side of (6.33) is of smaller order than $x^{-(\kappa_1+\kappa_2+\delta)}$, while by the potter bounds the denominator of the first term of the right hand side of (6.32) is of larger order than $x^{-(\kappa_1+\kappa_2+\delta)}$. Thus the first summand in (6.32) 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 (i) for general $s_0, s_1 > 0$. Note that the above value of $D$ does not depend on the distribution of $X_1$ or $X_2$. Therefore, we may replace $X_1$ and $X_2$ with $z_1 X_1$ and $z_2 X_2$ respectively, where $z_1$ and $z_2$ are as in (4.4), i.e. $z_1^{\alpha_1} z_2^{\alpha_2} = s_0^{-1}$ and $z_1^{\beta_1} z_2^{\beta_2} = s_1^{-1}$. As this substitution does not alter the solution to the linear program (4.5), (4.6),

$$\begin{aligned}
&\frac{P(Y_0 > s_0 x, Y_1 > s_1 x)}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})} \\
&= \quad \frac{P\left((z_1X_1)^{\alpha_1}(z_2X_2)^{\alpha_2}\prod_{i=3}^n X_i^{\alpha_i} > x, (z_1X_1)^{\beta_1}(z_2X_2)^{\beta_2}\prod_{i=3}^n X_i^{\beta_i} > x\right)}{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})} \\
&\qquad \times \frac{P(z_1X_1 > x^{\kappa_1})P(z_2X_2 > x^{\kappa_2})}{P(X_1 > x^{\kappa_1})P(X_2 > x^{\kappa_2})}
\end{aligned}$$

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. □

**Proof of Theorem 4.4**. We first show that the infinite dimensional optimization problem (4.11), (4.12) is equivalent to a finite dimensional one of type (4.5), (4.6) for sufficiently large $n$. To this end, choose $i^*, j^* \in \mathbb{N}_0$ such that $\alpha_{i^*} = \sup_{i\in\mathbb{N}} \alpha_i$ and $\beta_{j^*} = \sup_{j\in\mathbb{N}} \beta_j$. There exists a natural number $n$ such that $\max\{\alpha_i, \beta_i\} \le 1/(2/\alpha_{i^*} + 2/\beta_{j^*})$ for all $i > n$. Let $(\kappa_i)_{i\in\mathbb{N}}$ be any feasible solution which satisfies the constraints (4.12) 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 (4.12) 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 (4.11), (4.12) satisfy $\kappa_i = 0$ for all $i > n$. They can thus be identified with optimal solutions to the finite dimensional problem (4.5), (4.6), and vice versa. In particular, $\kappa_i > 0$ for at most two indices $i, j \in \{1, \dots, n\}$ (cf. Theorem 4.2); w.l.o.g., we may assume that $\kappa_1 > 0$ and $\kappa_i = 0$ for all $i \ge 3$.

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 in the same way.

Now,

$$\begin{pmatrix} Y_0 \\ Y_1 \end{pmatrix} = \begin{pmatrix} \prod_{i=n+1}^\infty X_i^{\alpha_i} & 0 \\ 0 & \prod_{i=n+1}^\infty X_i^{\beta_i} \end{pmatrix} \begin{pmatrix} \prod_{i=1}^n X_i^{\alpha_i} \\ \prod_{i=1}^n X_i^{\beta_i} \end{pmatrix} \tag{6.36}$$

where the vector and the matrix on the right hand side are independent. We will use the Breiman type Theorem 2.4 to establish the assertion in the cases (ii) with $\alpha_i = \beta_i$ and (i).

It follows from Theorem 4.2 that in these cases the vector on the right hand side is regularly varying on $(0,\infty)^2$ with index $-(\kappa_1+\kappa_2)$ and limiting measure $\nu$ defined in Remark 4.3 (iii).

According to Example 2.3 we have

$$
\begin{aligned}
&E\left(\tau\left(\begin{pmatrix}\prod_{i=n+1}^\infty X_i^{\alpha_i} & 0\\ 0 & \prod_{i=n+1}^\infty X_i^{\beta_i}\end{pmatrix}\right)^{3(\kappa_1+\kappa_2)/2}\right)\\
&\quad= E\left(\left(\max\left\{\prod_{i=n+1}^\infty X_i^{\alpha_i}, \prod_{i=n+1}^\infty X_i^{\beta_i}\right\}\right)^{3(\kappa_1+\kappa_2)/2}\right)\\
&\quad\leq E\left(\left(\prod_{i=n+1}^\infty X_i^{\alpha_i}\right)^{3(\kappa_1+\kappa_2)/2}\right)+E\left(\left(\prod_{i=n+1}^\infty X_i^{\beta_i}\right)^{3(\kappa_1+\kappa_2)/2}\right).
\end{aligned}
$$

Since $2\alpha_i(\kappa_1+\kappa_2)\leq 2\alpha_i(1/\alpha_{i^*}+1/\beta_{j^*})\leq 1$ for $i>n$, Fatou's lemma and Jensen's inequality yield

$$
\begin{aligned}
E\left(\left(\prod_{i=n+1}^\infty X_i^{\alpha_i}\right)^{3(\kappa_1+\kappa_2)/2}\right) &\leq \liminf_{m\to\infty}\prod_{i=n+1}^m E\left(X_1^{3\alpha_i(\kappa_1+\kappa_2)/2}\right)\\
&\leq \liminf_{m\to\infty}\prod_{i=n+1}^m \left(EX_1^{3/4}\right)^{2\alpha_i(\kappa_1+\kappa_2)}\\
&= \left(EX_1^{3/4}\right)^{2(\kappa_1+\kappa_2)\sum_{i=n+1}^\infty\alpha_i}\\
&< \infty. \qquad (6.37)
\end{aligned}
$$

Together with analogous calculations for $E((\prod_{i=n+1}^\infty X_i^{\beta_i})^{3(\kappa_1+\kappa_2)/2})$, we have verified condition (2.2) for the diagonal matrix of (6.36), $\alpha=\kappa_1+\kappa_2$ and $\epsilon=(\kappa_1+\kappa_2)/2$. By Theorem 2.4

$$
\begin{aligned}
&\lim_{x\to\infty}\frac{P(Y_0>s_0x,Y_1>s_1x)}{P\Big(\min\left(\prod_{i=1}^n X_i^{\alpha_i},\prod_{i=1}^n X_i^{\beta_i}\right)>x\Big)}\\
&\quad= E\left(\nu\left(\left(s_0\prod_{i=n+1}^\infty X_i^{-\alpha_i},\infty\right)\times\left(s_1\prod_{i=n+1}^\infty X_i^{-\beta_i},\infty\right)\right)\right).
\end{aligned}
$$

Using Theorem 4.2 and Remark 4.3 (iii), we conclude in the case (i) that

$$
\begin{aligned}
&\lim_{x\to\infty}\frac{P(Y_0>s_0x,Y_1>s_1x)}{P(X_1>x^{\kappa_1})P(X_2>x^{\kappa_2})}\\
&\quad= E\left(\left(s_0\prod_{i=n+1}^\infty X_i^{-\alpha_i}\right)^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)}\left(s_1\prod_{i=n+1}^\infty X_i^{-\beta_i}\right)^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)}\right)
\end{aligned}
$$

$$
\begin{aligned}
& \times \frac{|\alpha_1\beta_2 - \alpha_2\beta_1|}{(\alpha_1-\alpha_2)(\beta_2-\beta_1)} E\left(\prod_{m=3}^{n} X_m^{(\alpha_m(\beta_2-\beta_1)+\beta_m(\alpha_1-\alpha_2))/(\alpha_1\beta_2-\alpha_2\beta_1)}\right) \\
= \; & s_0^{(\beta_1-\beta_2)/(\alpha_1\beta_2-\alpha_2\beta_1)} s_1^{(\alpha_2-\alpha_1)/(\alpha_1\beta_2-\alpha_2\beta_1)} \\
& \times \frac{|\alpha_1\beta_2 - \alpha_2\beta_1|}{(\alpha_1-\alpha_2)(\beta_2-\beta_1)} E\left(\prod_{m=3}^{\infty} X_m^{(\alpha_m(\beta_2-\beta_1)+\beta_m(\alpha_1-\alpha_2))/(\alpha_1\beta_2-\alpha_2\beta_1)}\right).
\end{aligned}
$$

In the case (ii) with $\alpha_1 = \beta_1$, the assertion can be derived by the same arguments.

In the situation (ii) with $\alpha_1 > \beta_1$, the assertion follows by similar arguments as in the proof of Theorem 4.2. Since $Y_0 > 0$ almost surely, for all $C > 0$, one has for sufficiently large $x$

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

Recall from the proof of Theorem 4.2 that in the present case $\varrho := \left(\beta_1/\sup_{i\geq 2}\beta_i\right)^{1/2} > 1$. Therefore we have again by Fatou's lemma and Jensen's inequality that

$$
\begin{aligned}
E\left(\left(\prod_{i=2}^{\infty} X_i^{\beta_i}\right)^{\varrho/\beta_1}\right) & \leq \liminf_{m\to\infty} \prod_{i=2}^{m} E\left(X_1^{\varrho\beta_i/\beta_1}\right) \\
& \leq \liminf_{m\to\infty} \prod_{i=2}^{m} E\left(X_1^{1/\varrho}\right)^{\varrho^2\beta_i/\beta_1} \\
& = E\left(X_1^{1/\varrho}\right)^{\varrho^2 \sum_{i=2}^{\infty}\beta_i/\beta_1} < \infty.
\end{aligned}
$$

Thus we may apply Breiman's lemma to the lower and upper bound to obtain

$$
\begin{aligned}
& E\left(\left(\min\left\{C\prod_{i=2}^{\infty} X_i^{\alpha_i\beta_1/\alpha_1}, s_1^{-1}\prod_{i=2}^{\infty} X_i^{\beta_i}\right\}\right)^{1/\beta_1}\right) \\
\leq \; & \liminf_{x\to\infty} \frac{P(Y_1 > s_0 x, Y_1 > s_1 x)}{P(X_1^{\beta_1} > x)} \leq \limsup_{x\to\infty} \frac{P(Y_1 > s_0 x, Y_1 > s_1 x)}{P(X_1^{\beta_1} > x)} \\
\leq \; & E\left(\left(s_1^{-1}\prod_{i=2}^{\infty} X_i^{\beta_i}\right)^{1/\beta_1}\right) < \infty. \qquad (6.38)
\end{aligned}
$$

Let $C$ tend to $\infty$ to conclude the proof in this case. The case $\alpha_1 < \beta_1$ can be treated similarly.

Finally, we prove assertion (iii). The lower bound on $P(Y_0 > x, Y_1 > x)$ is an immediate consequence of the lower bound for $P(\prod_{i=1}^n X_i^{\alpha_i} > x/\delta, \prod_{i=1}^n X_i^{\beta_i} > x/\delta)$ from Theorem 4.2 (iii) and of the fact that $P(\prod_{i=n+1}^\infty X_i^{\alpha_i} > \delta, \prod_{i=n+1}^\infty X_i^{\beta_i} > \delta) > 0$ for some $\delta > 0$, because $Y_0, Y_1 > 0$ almost surely.

Moreover, with $Z := \max\{\prod_{i=n+1}^\infty X_i^{\alpha_i}, \prod_{i=n+1}^\infty X_i^{\beta_i}\}$ and $0 < \epsilon \le \kappa_1 + \kappa_2$

$$\begin{aligned}&P(Y_0 > x, Y_1 > x) x^{\sum_{i=1}^\infty \kappa_i - \epsilon}\\ &\le \int P\Big(\prod_{i=1}^n X_i^{\alpha_i} > \frac{x}{z}, \prod_{i=1}^n X_i^{\beta_i} > \frac{x}{z}\Big)\Big(\frac{x}{z}\Big)^{\kappa_1+\kappa_2-\epsilon} z^{\kappa_1+\kappa_2-\epsilon} P^Z(dz).\end{aligned}$$

By Theorem 4.2 (iii), the integrand tends to 0 as $x \to \infty$ for all $z > 0$ and, in view of (6.37), $Mz^{\kappa_1+\kappa_2-\epsilon}$ is an integrable majorant to the integrand for sufficiently large $M > 0$. Hence $P(Y_0 > x, Y_1 > x) = o(x^{\epsilon - \sum_{i=1}^\infty \kappa_i})$ follows by dominated convergence. □

### *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 (5.2) 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 4.4 (where we start from the index $i = 0$ instead of $i = 1$). Moreover, by Remark 4.5 (i) and the assumption that $E(\eta_0^2) < \infty$, it follows that $\sigma_0, \sigma_h > 0$ almost surely. The statement about $(\sigma_0, \sigma_h)$ thus follows by an application of Theorem 4.4, cases (i) and (ii), in combination with Remark 4.5 (iii).

In particular, $(\sigma_0, \sigma_h)$ is regularly varying on $\mathbb{E}^2$ with index $-\sum_{i=1}^\infty \kappa_i$. Hence, $\epsilon_0$ fulfills the moment condition of Corollary 2.6, 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 $\alpha_i$ and $\alpha_{i+h}$ are positive and thus $\sigma_0$ and $\sigma_h$ are independent and $(\sigma_0, \sigma_h)$ 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 $\kappa_{k(h)+h} = 1, \kappa_j = 0, j \neq k(h) + h$, is the unique solution to the optimization problem (5.1), (5.2). By Theorem 5.1 (ii) the coefficient of tail dependence of $(\sigma_0, \sigma_h)$ thus equals 1.

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

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

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 Corollary to Theorem 3 of Embrechts and Goldie (1980), $Z_h$ is regularly varying with index $-1$. Now, the joint behavior of $\sigma_0$ and $\sigma_h$ 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 $\sum_{i\in\mathbb{Z}} \kappa_i\alpha_i \geq 1, \tilde{\kappa} + \sum_{i\in\mathbb{Z}:\alpha_{i+h}\in(0,1)} \kappa_i\alpha_{i+h} \geq 1$ and $\tilde{\kappa} \geq 0, \kappa_i \geq 0, i \in \mathbb{Z}$, 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,

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

for some constant $c > 0$, and the coefficient of tail dependence of $(\sigma_0, \sigma_h)$ equals $1/(\kappa_{k(h)} + \tilde{\kappa}) = \eta_h$.

By Corollary 2.6 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 $E(|\epsilon_0|^{2+\delta}) < \infty$ 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), (5.2) 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 4.4 (iii) leads to a contradiction to $\eta_h = 1$.) But then $\kappa_i > 0$ can only hold if $\alpha_{i-h} = \alpha_i = 1$. Write

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

with $Z := \prod_{i\in\mathbb{N}_0:\alpha_i=\alpha_{i-h}=1} \mathrm{e}^{\eta_{h-i}}$. Again, $Z$ is regularly varying with index $-1$ by Embrechts and Goldie (1980). 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

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

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

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 4.4. By Lemma 7.2 of Rootzén (1986)

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

for a constant $\hat{K} > 0$, where we have used equation (7.8) in Rootzén (1986) 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

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

and thus asymptotic dependence of $(\sigma_0, \sigma_h)$. The same holds true for $(X_0, X_h)$.

In contrast, if $\beta > -1$, then one may conclude from Lemma 7.2 of Rootzén (1986) that $P(\sigma_h > x \mid \sigma_0 > x) = O(\log(x)^{-l(\beta+1)}) \to 0$ 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 $\sigma_0$ and $\sigma_h$ in this case.

**Acknowledgment:** This project was supported by the German Research Foundation DFG, grant no JA 2160/1.

## References

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

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

Bingham, N. H., Goldie, C. M. and Teugels, J. L.: Regular Variation. Cambridge University Press, Cambridge (1987).

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

Das, B., Mitra, A. and Resnick, S. I.: Living on the multi-dimensional edge: seeking hidden risks using regular variation. Arxiv Preprint, http://arxiv.org/abs/1108.5560 (2011).

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

Davis, R.A. and Mikosch, T.: Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab.* **38**, 93—104 (2001).

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

Drees, H.: Estimating the coefficient of tail dependence for stationary time series. Preprint, University of Hamburg (2013).

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

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

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

Leadbetter, M.R.: Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65**, 291—306 (1983).

Ledford, A.W. and Tawn, J.A.: Statistics for near independence in multivariate extremes. *Biometrika* **83**, 169–187 (1996).

Ledford, A.W. and Tawn, J.A.: Diagnostics for dependence within time series extremes. *J. R. Statist. Soc. B* **65**, 521–543 (2003).

Mikosch, T. and Rezapour, M.: Stochastic volatility models with possible extremal clustering. *Bernoulli (to appear)* (2012).

Pratt, J. W.: On interchanging limits and integrals. *Ann. Math. Statist.* **31**, 74–77 (1960)

Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer-Verlag, New York (1987).

Resnick, S.I.: Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5**, 303–336 (2002).

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

Resnick, S.I.: Multivariate regular variation on cones: application to extreme values, hidden regular variation and conditioned limit laws. *Stochastics* **80**, 269–298 (2008).

Rootzén, H.: Extreme value theory for moving average processes. *Ann. Probab.* **14**, 612–652 (1986).

Sierksma, G.: Linear and Integer Programming: Theory and Practice. Dekker, New York (1996).

Taylor, S.J.: Modelling Financial Time Series. Chichester, Wiley (1986).