3.1 Introduction

Many theoretical and practical examples allege the assumption of asymptotic dependence for multivariate extremes. Some multivariate random variables have a strength of dependence that weakens at high or extreme values. For instance, Sibuya et al., 1960 show that all bivariate normal variables with correlations less than one are asymptotically independent such that their pairwise extremal correlations are zeros. Moreover, practical examples such as those in Souvairan, 2022 and Ledford et al., 1996 reveal that the asymptotic dependence is not a realistic assumption for a financial dataset of the "S&P 500" stocks and a climatic dataset of "wind and rain" provided by Anderson et al., 1993, respectively. To this extent, Coles, 2001 asserts that the most extreme events are near-independent. Hence, the conventional methods assuming the asymptotic dependence for multivariate extremes can lead to misleading results for these events.

As discussed in Section (2.5), the extremal tree model, provided by Engelke and Volgushev, 2020, is a graphical model under the assumption of asymptotic dependence with respect to a connected graph which is a tree. This assumption is unrealistic, restrictive, and inherited from trees' intrinsic connectedness. In this thesis, we suggest the extremal forest models. From Section (2.2), recall that a forest is an acyclic graph which can be either a connected graph that consists of a tree or an unconnected graph consisting of several trees as its connected components (the connected components of a forest are connected and acyclic.). Therefore, the advantage of modeling under both regimes, i.e., asymptotic dependence and asymptotic independence, places the extremal forest models ahead of extremal tree models. In the following, we represent the extremal forest models with two approaches.

3.2 Notations

In this section, we repeat some notations in EVT and graph theory that we use in the current and following chapters.
We denote a multivariate Pareto distribution, defined as in Theorem (2.1.27), by $\mathbf{Y}$. The support of $\mathbf{Y}$ is an L-shaped space denoted by $\mathcal{L} = \{\mathbf{x} \in \mathcal{E} : \|\mathbf{x}\|_\infty > 1\}$ where $\mathcal{E} := [0,\infty)^d \setminus \{\mathbf{0}\}$ is a cone on $\mathbb{R}^d$. We denote faces of $\mathcal{E}$ by $\mathcal{E}^I := \{\mathbf{x} \in \mathcal{E} : \mathbf{x}_I > \mathbf{0}, \mathbf{x}_{\setminus I} = \mathbf{0}\}$, where $\emptyset \neq I \subseteq V = \{1,...,d\}$ and $\mathcal{E} = \bigcup_{\emptyset \neq I \subseteq V} \mathcal{E}^I$. We also denote $\mathcal{E}_I := [0,\infty)^{|I|} \setminus \{\mathbf{0}\}$, which is a cone on the subspace of $\mathcal{E}$. For any non-empty $J \subseteq I$, we denote $\mathcal{E}_I^J := \{\mathbf{x} \in \mathcal{E}_I : \mathbf{x}_J > \mathbf{0}, \mathbf{x}_{\setminus J} = \mathbf{0}\}$. The random vector $\mathbf{Y}^m$ is defined as $\mathbf{Y}$ conditioned on the event that $\{Y_m > 1\}$. Therefore, $\mathbf{Y}^m$ is supported on the product space $\mathcal{L}^m = \{\mathbf{x} \in \mathcal{L} : x_m > 1\}$.

A forest is denoted by $\mathcal{F} = (V, E_{\mathcal{F}})$. $\mathfrak{C}(\mathcal{F})$ denotes the set of connected components of $\mathcal{F}$. These connected components are acyclic and connected, so they are trees. We denote these trees by $\mathcal{T}_1, ..., \mathcal{T}_{c \geq 2}$ where $c := |\mathfrak{C}(\mathcal{F})|$ and for all $i \in \{1,...,c\}$, $\mathcal{T}_i = (V_i, E_i)$ such that E_i is restricted to the pairs in V_i. The $\mathcal{T}_i^m = (V, E_i^m)$ denotes a tree rooted at an arbitrary but fixed node $m \in V_i$.

3.3 First representation of extremal forest models

In what follows, we introduce a helpful lemma that can be useful in our proofs related to two representations of the extremal forest models; later, we suggest the first representation of extremal forest models.

Lemma 3.3.1. *Let $\mathbf{Y}$ be a multivariate Pareto distribution, as in Theorem (2.1.27), that is defined on $\mathcal{L} = \{\mathbf{x} \in \mathcal{E} : \|\mathbf{x}\|_\infty > 1\}$, where $\mathcal{E} := [0,\infty)^d \setminus \{\mathbf{0}\}$. Then for any Borel set $A \subseteq \mathcal{L}$*

$$\mathbb{P}(\mathbf{Y} \in A) = \frac{\Lambda(A)}{\Lambda(\mathbf{1})}, \tag{3.3.1}$$

where Λ is the exponent measure defined in Theorem (2.1.22).

Proof. Proposition 6(ii) in Engelke and Volgushev, 2020 shows that the limiting measure ν in Theorem 2(e) of Segers, 2020 satisfies $\mathbb{P}(\mathbf{Y} \in A) = \nu(A)/\nu(\mathcal{L})$ for all Borel sets $A \subseteq \mathcal{L}$. By paragraph 3 from page 10 of Segers, 2020, for any $I \subseteq \{1,...,d\}$, a measure ν is any Borel measure on $\mathcal{E}_{0,I} = \{\mathbf{x} \in \mathcal{E} : \max(\mathbf{x}_I) > 0\}$ with the property that $\nu(B)$ is finite for every Borel set B of $\mathcal{E}_{0,I}$ that is contained in a set of the form $\{\mathbf{x} \in \mathcal{E} : \max(\mathbf{x}_I) > \varepsilon\}$ for some $\varepsilon > 0$. The exponent measure Λ has the latter property (for more details, see the explanation below the Equation (8.3) on the concentration of the exponent measure in Beirlant et al., 2004). Therefore, for any Borel set $A \subseteq \mathcal{L}$ we have that,

$$\mathbb{P}(\mathbf{Y} \in A) = \frac{\Lambda(A)}{\Lambda(\mathcal{L})}.$$

Recall that, $\Lambda(\mathbf{z})$ is a shorthand for $\Lambda(\mathcal{E} \setminus [\mathbf{0}, \mathbf{z}])$. Therefore, $\Lambda(\mathbf{1})$ is a shorthand notation for $\Lambda(\mathcal{L}) = \Lambda(\mathcal{E} \setminus [\mathbf{0}, \mathbf{1}])$. Consequently, the above equation leads to Equation (3.3.1). □

Proposition 3.3.2. *Let $\mathbf{Y}$ be a multivariate Pareto distribution that is an extremal graphical model with respect to the forest $\mathcal{F} = (V, E_{\mathcal{F}})$, where $\mathbf{Y}$ and $\mathcal{F}$ are defined as Section (3.2). Let P be a standard Pareto distribution, independent of $\{W_e : e \in E_i^m\}$ (for any edge $e \in E_{\mathcal{T}}^m$ incident with nodes p and q, we have $W_e = W_q^p$) for all $i \in \{1,...,c\}$ and an arbitrary $m \in V_i$. Then, we have the joint stochastic representation for $\mathbf{Y}^m$ on $\mathcal{L}^m$*

$$Y_j^m \stackrel{d}{=} \begin{cases} P \times \prod_{e \in ph(mj)} W_e, & \text{for } j \in V_i \\ 0, & \text{for } j \in V \setminus V_i \end{cases} \tag{3.3.2}$$

where $ph(mj)$ denotes the set of edges on the unique path from node m to node j on the tree $\mathcal{T}_i^m$, and by convention, $ph(mm) = \emptyset$ and $\prod_{e \in \emptyset} W_e = 1$.

Proof. For any $i \in \{1,...,c\}$, the sets V_i and $V \setminus V_i$ are disjoint sets. The random vector $\mathbf{Y}$ is an extremal graphical model with respect to the forest $\mathcal{F}$; hence, by Definition (2.5.1), we have that

$$\mathbf{Y}_{V_i} \perp_e \mathbf{Y}_{V \setminus V_i}.$$

By Lemma (2.4.2)(3), the independence of $\mathbf{Y}_{V_i}$ and $\mathbf{Y}_{V \setminus V_i}$ is equivalent to

$$\Lambda_I(\mathcal{E}_I^I) = 0,$$

for any I, where $I \cap V_i \neq \emptyset$ and $I \cap V \setminus V_i \neq \emptyset$. Moreover, Lemma (3.3.1) implies that

$$\mathbb{P}\left(\mathbf{Y} \in \mathcal{E}_I^I\right) = 0,$$

for any I, where $I \cap V_i \neq \emptyset$ and $I \cap V \setminus V_i \neq \emptyset$. Therefore, $\mathbf{Y}$ must lie either on the subface $\mathcal{E}^{V \setminus V_i}$ or on the subface $\mathcal{E}^{V_i}$. Conditioning on $\{Y_m > 1\}$ for any arbitrary $m \in V_i$ implies that $\mathbf{Y}^m$ lies on the subface $\mathcal{E}^{V_i}$. Thus,

$$Y_j^m \stackrel{d}{=} 0,$$

for any $j \in V \setminus V_i$.

Moreover, the marginal $\mathbf{Y}_{V_i} = (Y_j)_{j \in V_i}$ of $\mathbf{Y}$ conditioned on $\{\|\mathbf{Y}_{V_i}\|_\infty > 1\}$ is defined on $\mathcal{L}_{V_i} = \{\mathbf{x}_{V_i} \in [0, \infty)^{|V_i|} \setminus \{\mathbf{0}\} : \|\mathbf{x}_{V_i}\|_\infty > 1\}$, and it is a graphical model with respect to the tree $\mathcal{T}_i$. Consequently, by Theorem (2.5.4), we have that

$$Y_j^m \stackrel{d}{=} P \times \prod_{e \in ph(mj)} W_e,$$

for any $j \in V_i$. □

Corollary (3.3.3) follows Equation (2.8.1)

Corollary 3.3.3. *Let* $\mathbf{Y}$ *be a multivariate Pareto distribution that is an extremal graphical model with respect to the forest* $\mathcal{F} = (V, E_{\mathcal{F}})$*, where* $\mathbf{Y}$ *and* $\mathcal{F}$ *are defined as Section (3.2), and all marginals say* $\mathbf{Y}_{V_i}$ *for* $i \in \{1, ..., c\}$*, follow a Hüsler–Reiss distribution that is an extremal graphical model with respect to the tree* $\mathcal{T}_i$*. Let* P *be a standard Pareto distribution, independent of* $\{W_e : e \in E_i^m\}$ *(for any edge* $e \in E_{\mathcal{T}}^m$ *incident with nodes* p *and* q*, we have* $W_e = W_q^p$*) for all* $i \in \{1, ..., c\}$ *and an arbitrary* $m \in V_i$*. Then, we have the joint stochastic representation for* $\mathbf{Y}^m$ *on* $\mathcal{L}^m$

$$Y_j^m \overset{d}{=} \begin{cases} P \exp\left(\boldsymbol{U}^m - \Gamma_{\cdot m}/2\right) & \text{for } j \in V_i \\ 0, & \text{for } j \in V \setminus V_i \end{cases}$$

where $ph(mj)$ *denotes the set of edges on the unique path from node* m *to node* j *on the tree* $\mathcal{T}_i^m$*, and by convention,* $ph(mm) = \emptyset$ *and* $\prod_{e \in \emptyset} W_e = 1$.

3.4 Second representation of extremal forest models

In what follows, we introduce some helpful Lemma for providing the second representation of extremal forest models; subsequently, we propose the second representation of extremal forest models. For the following lemma and proposition, the definition of max-domain of attraction is crucially required for the case that Λ places mass on the subfaces. As Remark 1 in Engelke and Hitz, 2020 states, the assumption of multivariate regular variation provided in Proposition 5.15(a) S. Resnick, 2008 implies the asymptotic dependence between all components. We exploit the notion of multivariate regular variation in Theorem 6.1 and Lemma 6.1 of S. I. Resnick, 2007 and Proposition 3.3.1 of Mariko, 2015 that is define multivariate regular variation under asymptotic independence.

Lemma 3.4.1. *Let* $\mathbf{X}_1, ..., \mathbf{X}_{c \geq 2}$ *be independent multivariate distributions with standard Pareto marginals, defined on* $\mathcal{L}_{V_1}, ..., \mathcal{L}_{V_c}$ *where* $\mathcal{L}_{V_i} = \{\mathbf{x}_i \in \mathcal{E}_{V_i} : \|\mathbf{x}_i\|_\infty > 1\}$ *and* $\mathcal{E}_{V_i} = [0, \infty)^{|V_i|} \setminus \{\mathbf{0}\}$*, and let* $\mathbf{X}_1, ..., \mathbf{X}_c$ *be in the max-domain of attraction of multivariate Pareto distributions* $\mathbf{Y}_1, ..., \mathbf{Y}_c$*, respectively; such that* $\mathbf{X} := (\mathbf{X}_1, ..., \mathbf{X}_c)$ *be in the max-domain of attraction of* $\mathbf{Y}$*. Then* $\mathbb{P}(\mathbf{Y} > \mathbf{1}) = 0$*.*

Proof. The random vector $\mathbf{X}$ is in the max-domain of attraction of $\mathbf{Y}$. Multivariate Pareto distributions are the only possible limits for

$$\lim_{u \to \infty} \mathbb{P}\left(\frac{\mathbf{X}}{u} > \mathbf{1} \middle| \|\mathbf{X}\|_\infty > u\right),$$

therefore, the random vector $\mathbf{Y}$ follows a multivariate Pareto distribution. Hence, by Theorem (2.1.27), $\mathbf{Y} \overset{d}{=} \mathbf{X}/u$; conditioned on that $\{\|\mathbf{X}\|_\infty > u\}$ where u goes to ∞. Therefore,

$$\begin{aligned}
\mathbb{P}(\mathbf{Y} > \mathbf{1}) &= \lim_{u \to \infty} \mathbb{P}\left(\frac{\mathbf{X}}{u} > \mathbf{1} \middle| \|\mathbf{X}\|_\infty > u\right) \\
&= \lim_{u \to \infty} \mathbb{P}\left(\mathbf{X} > u\mathbf{1} \middle| \|\mathbf{X}\|_\infty > u\right) \\
&= \lim_{u \to \infty} \mathbb{P}\left(\forall i \in \{1, ..., c\} : \mathbf{X}_i > u\mathbf{1} \middle| \exists i \in \{1, ..., c\} : \max(\mathbf{X}_i) > u\right) \\
&= \lim_{u \to \infty} \frac{\mathbb{P}\left(\{\forall i \in \{1, ..., c\} : \mathbf{X}_i > u\mathbf{1}\} \cap \{\exists i \in \{1, ..., c\} : \max(\mathbf{X}_i) > u\}\right)}{\mathbb{P}\left(\exists i \in \{1, ..., c\} : \max(\mathbf{X}_i) > u\right)} \\
&= \lim_{u \to \infty} \frac{\mathbb{P}\left(\forall i \in \{1, ..., c\} : \mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(\exists i \in \{1, ..., c\} : \max(\mathbf{X}_i) > u\right)} \\
&= \lim_{u \to \infty} \frac{\prod_{i=1}^c \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(\exists i \in \{1, ..., c\} : \max(\mathbf{X}_i) > u\right)}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{u\to\infty} \frac{\prod_{i=1}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left((\max(\mathbf{X}_1) > u) \vee ... \vee (\max(\mathbf{X}_c) > u)\right)} \\
&\leq \lim_{u\to\infty} \frac{\prod_{i=1}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(\max(\mathbf{X}_1) > u\right)} \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left(\mathbf{X}_1 > u\mathbf{1}\right) \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(\max(\mathbf{X}_1) > u\right)} \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left(\forall j \in \{1,...,b\} : X_{1,j} > u\right) \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(\exists j \in \{1,...,b\} : X_{1,j} > u\right)} \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left((X_{1,1} > u) \wedge ... \wedge (X_{1,b} > u)\right) \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left((X_{1,1} > u) \vee ... \vee (X_{1,b} > u)\right)} \\
&\leq \lim_{u\to\infty} \frac{\mathbb{P}\left(X_{1,1} > u\right) \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left((X_{1,1} > u) \vee ... \vee (X_{1,b} > u)\right)} \\
&\leq \lim_{u\to\infty} \frac{\mathbb{P}\left(X_{1,1} > u\right) \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right)}{\mathbb{P}\left(X_{1,1} > u\right)} \\
&= \lim_{u\to\infty} 1 \cdot \prod_{i=2}^{c} \mathbb{P}\left(\mathbf{X}_i > u\mathbf{1}\right) \\
&= 0
\end{aligned}
$$

where $\mathbf{X_1} = (X_{1,1}, ..., X_{1,b})$. For the inequalities, we use the facts:

- Since $A \subseteq A \cup B$, for the probability measure $\mathbb{P}$, we have $\mathbb{P}(A) \leq \mathbb{P}(A \cup B)$.
- Since $A \cap B \subseteq A$, for the probability measure $\mathbb{P}$, we have $\mathbb{P}(A \cap B) \leq \mathbb{P}(A)$.

$\square$

Proposition 3.4.2. *Let $\mathbf{X}_1, ..., \mathbf{X}_{c\geq 2}$ be independent multivariate distributions with standard Pareto marginals, defined on $\mathcal{L}_{V_1}, ..., \mathcal{L}_{V_c}$ where $\mathcal{L}_{V_i} = \{\mathbf{x}_i \in \mathcal{E}_{V_i} : \|\mathbf{x}_i\|_\infty > 1\}$ and $\mathcal{E}_{V_i} = [0,\infty)^{|V_i|} \setminus \{\mathbf{0}\}$, and let $\mathbf{X}_1, ..., \mathbf{X}_c$ be in the max-domain of attraction of multivariate Pareto distributions $\mathbf{Y}_1, ..., \mathbf{Y}_c$, that are graphical models with respect to the trees $\mathcal{T}_1, ..., \mathcal{T}_c$, respectively (with the notation in Section (3.2)); such that $\mathbf{X} \stackrel{d}{=} (\mathbf{X}_1, ..., \mathbf{X}_c)$ be in the max-domain of attraction of $\mathbf{Y}$. Then, the limiting distribution $\mathbf{Y}$ is a multivariate Pareto distribution which is a graphical model with respect to the forest $\mathcal{F} = (V, E_{\mathcal{F}})$ where the set of connected components of $\mathcal{F}$ consists of the trees $\mathcal{T}_1, ..., \mathcal{T}_c$.*

Proof. The random vector $\mathbf{X}$ is in the max-domain of attraction of $\mathbf{Y}$. Multivariate Pareto distributions are the only possible limits for

$$
\lim_{u\to\infty} \mathbb{P}\left(\frac{\mathbf{X}}{u} > \mathbf{1} \,\middle|\, \|\mathbf{X}\|_\infty > u\right),
$$

therefore, the random vector $\mathbf{Y}$ follows a multivariate Pareto distribution (for more details, see Rootzén et al., 2006).

First (i), we need to prove that the marginals $\mathbf{Y}_1, ..., \mathbf{Y}_c$ are asymptotically independent or equivalently, the random vector $\mathbf{Y}$ lies on the subfaces $\mathcal{E}^{V_1}, ..., \mathcal{E}^{V_c}$, where $\mathcal{E}^{V_i} := \{\mathbf{x} \in \mathcal{E} : \mathbf{x}_i > \mathbf{0}, \mathbf{x}_{\setminus i} = \mathbf{0}\}$. Secondly (ii), we must show that every marginal in $\mathbf{Y}_1, ..., \mathbf{Y}_c$, say $\mathbf{Y}_i$, is a graphical model with respect to the tree $\mathcal{T}_i$. Eventually, as a consequence of these two, we state that $\mathbf{Y}$ is a multivariate Pareto distribution, which is a graphical model with respect to the forest $\mathcal{F}$ where the set of connected components of $\mathcal{F}$ consists of the trees $\mathcal{T}_1, ..., \mathcal{T}_c$.

(i). By Lemma (3.4.1), we have that:

$$
\mathbb{P}(\mathbf{Y} > \mathbf{1}) = 0 \tag{3.4.1}
$$

Recall that in Subsubsection (2.1.9), we discussed that a multivariate Pareto distribution is characterized by the homogeneity of order -1, which is as follows.

$$\mathbb{P}(\mathbf{Y} \in tB) = t^{-1}\mathbb{P}(\mathbf{Y} \in B), \quad t \geq 1,$$

where for any Borel subset $B \subseteq \mathcal{L}$, we define $tB = \{t\mathbf{x} : \mathbf{x} \in B\}$. Now we show that $\mathbf{Y}$ lies on the subfaces. To do so, we must show that for any set $\mathcal{S}$ that $\mathcal{S} \subseteq \mathcal{E} \setminus \left(\bigcup_{i=1,\dots,c} \mathcal{E}^{V_i}\right)$ $\mathbb{P}(\mathbf{Y} \in \mathcal{S}) = 0$. By contradiction, suppose that there is $\varepsilon > 0$, where $A := \{\mathbf{y} : y_i > \varepsilon \text{ for all } i \in \{1, \dots, d\}\}$ and $\mathbb{P}(\mathbf{Y} \in A) > 0$. Now let $t := \varepsilon^{-1} > 0$. Then by Equation (2.1.26):

$$\mathbb{P}(\mathbf{Y} > \mathbf{1}) = t^{-1}\mathbb{P}(\mathbf{Y} \in A) > 0,$$

which is in contradiction with the result of Lemma (3.4.1), i.e., $\mathbb{P}(\mathbf{Y} > \mathbf{1}) = 0$ [see Figure (3.1)]

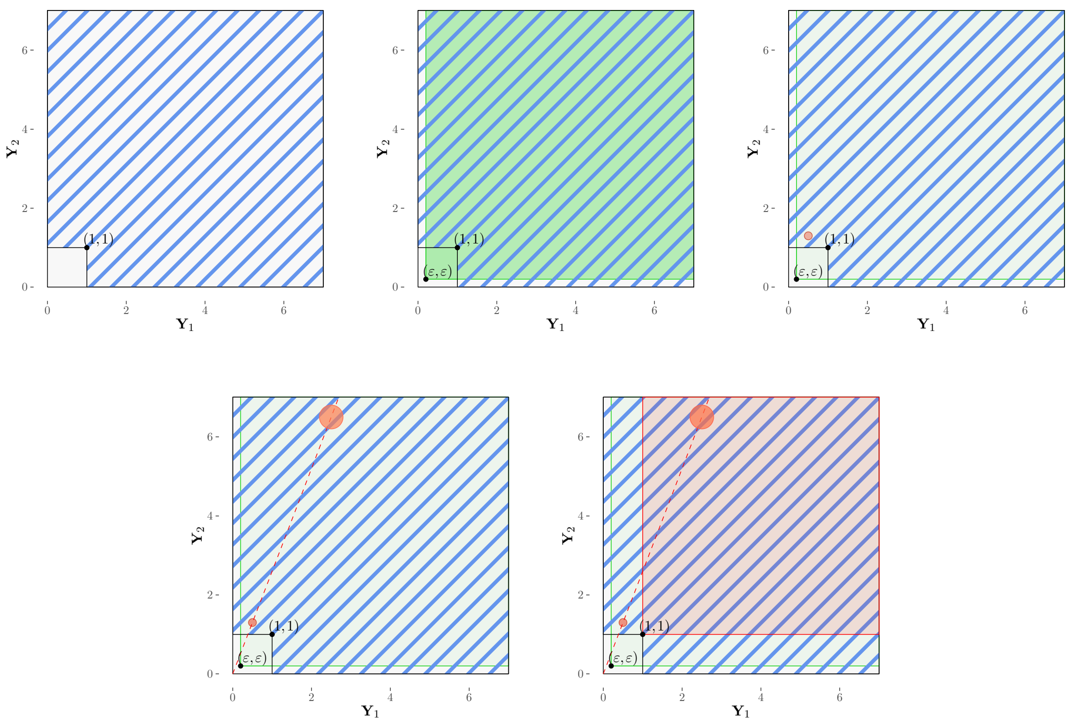

Figure 3.1: Figures illustrate that $\mathbb{P}(\mathbf{Y} \in A) > 0$ for any $A := \{\mathbf{y} : y_i > \varepsilon \text{ for all } i \in \{1, \dots, d\}\}$ is in contradiction with the $\mathbb{P}(\mathbf{Y} > \mathbf{1}) = 0$. : the support of $\mathbf{Y}$, i.e., $\mathcal{L}$. : The possible area for choosing A such that every component of A is larger than ε. : A (small circle) and tA (large circle). : The area corresponding to $\mathbf{Y} > \mathbf{1}$.

Hence, for any set $\mathcal{S} \subseteq \mathcal{E} \setminus \left(\bigcup_{i=1,\dots,c} \mathcal{E}^{V_i}\right)$, $\mathbb{P}(\mathbf{Y} \in \mathcal{S}) = 0$. By Lemma (3.3.1) $\mathbb{P}(\mathbf{Y} \in \mathcal{S}) = \dfrac{\Lambda(\mathcal{S})}{\Lambda(\mathbf{1})} = 0$, thus $\Lambda(\mathcal{S}) = 0$ for any set $\mathcal{S} \subseteq \mathcal{E} \setminus \left(\bigcup_{i=1,\dots,c} \mathcal{E}^{V_i}\right)$. Consequently, from Lemma (2.4.2)(3), $\mathbf{Y}_1, \dots, \mathbf{Y}_c$ are asymptotically independent. Therefore, by Definition (2.5.1), $\mathbf{Y}_1, \dots, \mathbf{Y}_c$ are extremal graphical models with respect to connected components on $V_1, \dots, V_c$.

(ii). Now, we must show that every marginal in $\mathbf{Y}_1, \dots, \mathbf{Y}_c$, say $\mathbf{Y}_i$, is a graphical model with respect to the tree $\mathcal{T}_i$. As a consequence of the first part of the proof, i.e.,(i), we know that $\mathbf{Y}$ lies on the subfaces and we have that,

$$\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) = \lim_{u \to \infty} \mathbb{P}\left(\frac{\mathbf{X}}{u} \leq \mathbf{y} \,\middle|\, \|\mathbf{X}\|_\infty > u\right), \quad \mathbf{y} \in \bigcup_{i=1}^{c} \mathcal{E}^{V_i}.$$

For $\mathbf{y} \in \bigcup_{i=1}^{c} \mathcal{E}^{V_i}$, we have that

$$
\begin{aligned}
\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) &= \lim_{u\to\infty} \mathbb{P}\left(\frac{\mathbf{X}}{u} \leq \mathbf{y} \middle| \bigvee_{i=1}^{c} \|\mathbf{X}_i\|_\infty > u\right) \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left(\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\bigvee_{i=1}^{c}\{\|\mathbf{X}_i\|_\infty > u\}\}\right)}{\mathbb{P}\left(\bigvee_{i=1}^{c}\{\|\mathbf{X}_i\|_\infty > u\}\right)} \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left(\bigvee_{i=1}^{c}\{\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\}\right)}{\mathbb{P}\left(\bigvee_{i=1}^{m}\{\|\mathbf{X}_i\|_\infty > u\}\right)}.
\end{aligned}
\tag{3.4.2}
$$

Moreover, for any $i, j \in \{1, ..., c\}$

$$
\begin{aligned}
\mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u \wedge \|\mathbf{X}_j\|_\infty > u\right) &\leq \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right) \\
&= \mathbb{P}\left(\bigvee_{k=1}^{p} \{X_{i,k} > u\}\right) \\
&\leq \sum_{k=1}^{p} \mathbb{P}\left(X_{i,k} > u\right)
\end{aligned}
\tag{3.4.3}
$$

where $\mathbf{X}_i = (X_{i,1}, ..., X_{i,p})$. Since for any $k \in \{1, ..., p\}$ for the standard pareto marginals $X_{i,k}$ we have that

$$
\lim_{u\to\infty} \mathbb{P}\left(X_{i,k} > u\right) = 0. \tag{3.4.4}
$$

Therefore, by Equations (3.4.3) and (3.4.4),

$$
\lim_{u\to\infty} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u \wedge \|\mathbf{X}_j\|_\infty > u\right) = 0. \tag{3.4.5}
$$

By Equation (3.4.5), we conclude in both the numerator and denominator of Equation (3.4.2), the probability measure applied on the sets, which are the union of pairwise disjoint sets. Because the probability of the union of pairwise disjoint sets is equal to the sum of the probabilities on each of these sets, we deduce that

$$
\begin{aligned}
\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) &= \lim_{u\to\infty} \frac{\mathbb{P}\left(\bigvee_{i=1}^{c}\{\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\}\right)}{\mathbb{P}\left(\bigvee_{i=1}^{c}\{\|\mathbf{X}_i\|_\infty > u\}\right)} \\
&= \lim_{u\to\infty} \frac{\sum_{i=1}^{c} \mathbb{P}\left(\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\right)}{\sum_{i=1}^{c} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)}.
\end{aligned}
\tag{3.4.6}
$$

From the first part of the proof, (i), we know that $\mathbf{Y}$ lies on the subfaces of $\mathcal{E}$. When u goes to ∞, $\{\|\mathbf{X}_i\|_\infty > u\}$ for any $i \in \{1, ..., c\}$, results in $\mathbf{X}_{\setminus i} = \mathbf{0}$.

$$
\lim_{u\to\infty} \mathbb{P}\left(\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\right) = \lim_{u\to\infty} \mathbb{P}\left(\{\mathbf{X}_i \leq u\mathbf{y}_i\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\right)
$$

Therefore, Equation (3.4.6) implies that

$$
\begin{aligned}
\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) &= \lim_{u\to\infty} \frac{\sum_{i=1}^{c} \mathbb{P}\left(\{\mathbf{X} \leq u\mathbf{y}\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\right)}{\sum_{i=1}^{c} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)} \\
&= \lim_{u\to\infty} \frac{\sum_{i=1}^{c} \mathbb{P}\left(\{\mathbf{X}_i \leq u\mathbf{y}_i\} \wedge \{\|\mathbf{X}_i\|_\infty > u\}\right)}{\sum_{i=1}^{c} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)} \\
&= \lim_{u\to\infty} \frac{\sum_{i=1}^{c} \mathbb{P}\left(\mathbf{X}_i \leq u\mathbf{y}_i \middle| \|\mathbf{X}_i\|_\infty > u\right) \cdot \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)}{\sum_{i=1}^{c} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)}.
\end{aligned}
$$

Consequently,

$$
\mathbb{P}(\mathbf{Y} \leq \mathbf{y}) = \sum_{i=1}^{c} \mathbb{P}\left(\mathbf{Y}_i \leq \mathbf{y}_i\right) \cdot \lim_{u\to\infty} \frac{\mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)}{\sum_{i=1}^{c} \lim_{u\to\infty} \mathbb{P}\left(\|\mathbf{X}_i\|_\infty > u\right)}, \qquad \mathbf{y} \in \bigcup_{i=1}^{c} \mathcal{E}^{V_i}. \tag{3.4.7}
$$

Equation (3.4.7) shows that the cumulative probability of $\mathbf{Y}$ in each subface, say $\mathcal{E}^{V_i}$, corresponds to the cumulative probability of the marginal $\mathbf{Y}_i$ on that subface, which is a graphical model with respect to the tree $\mathcal{T}_i$. Hence, the probability distribution functions on the subfaces factorize on the corresponding trees.

As a consequence of parts (i) and (ii), we conclude that the limiting distribution of $\mathbf{X}/u$ conditioned on that $\{\|\mathbf{X}\|_\infty > u\}$ is a graphical model with respect to the forest $\mathcal{F} = (V, E_{\mathcal{F}})$ consisting of the trees $\mathcal{T}_1, ..., \mathcal{T}_c$. □

Corollary 3.4.3. *Let $\mathbf{X}_1, ..., \mathbf{X}_{c\geq 2}$ be independent multivariate Hüsler–Reiss distributions, defined on $\mathcal{L}_{V_1}, ..., \mathcal{L}_{V_c}$ where $\mathcal{L}_{V_i} = \{\mathbf{x}_i \in \mathcal{E}_{V_i} : \|\mathbf{x}_i\|_\infty > 1\}$ and $\mathcal{E}_{V_i} = [0, \infty)^{|V_i|} \setminus \{\mathbf{0}\}$, and let $\mathbf{X}_1, ..., \mathbf{X}_c$ be in the max-domain of attraction of multivariate Pareto distributions $\mathbf{Y}_1, ..., \mathbf{Y}_c$, that are graphical models with respect to the trees $\mathcal{T}_1, ..., \mathcal{T}_c$, respectively (with the notation in Section (3.2)); such that $\mathbf{X} \overset{d}{=} (\mathbf{X}_1, ..., \mathbf{X}_c)$ be in the max-domain of attraction of $\mathbf{Y}$. Then, the limiting distribution $\mathbf{Y}$ is a multivariate Pareto distribution which is a graphical model with respect to the forest $\mathcal{F} = (V, E_{\mathcal{F}})$ where the set of connected components of $\mathcal{F}$ consists of the trees $\mathcal{T}_1, ..., \mathcal{T}_c$.*

Propositions (3.3.2) and (3.4.2) provide elegant approaches for the simulations of samples from a multivariate Pareto distribution, denoted by $\mathbf{Y}$, which is an extremal graphical model with respect to a forest $\mathcal{F}$, consisting of the trees $\mathcal{T}_1, ..., \mathcal{T}_c$. To simulate samples of $\mathbf{Y}$ by Proposition (3.3.2), one can generate samples from the Pareto marginals that are extremal graphical models with respect to the trees and embed them with enough zeros. The sample generation from a Pareto distribution which is a graphical model with respect to a tree can be done using the function `rmpareto_tree` of the package in `R`, called `graphicalExtremes`. Moreover, to simulate samples of $\mathbf{Y}$ by Proposition (3.4.2), one can generate samples from a max-stable marginal for which the limiting distribution is an extremal graphical model with respect to a tree and embed these samples with the samples from other max-stable marginals corresponding to other trees of the forest. The sample generation from Pareto distributions, which are graphical models with respect to the trees, can be done using the function `rmstable_tree` of package `graphicalExtremes`. Eventually, one can make a sample of Pareto distribution corresponding to a forest using the function `data2mpareto` on the latter sample. (for more details on the simulation method and functions, see Engelke and Hitz, 2020; Engelke, Hitz, et al., 2022, respectively).

3.5 Simulation studies

In this section, we introduce the methodology of simulation of data for an arbitrary underlying forest. We will illustrate the convergence of first and second representations provided in the previous sections, using simulations.

3.5.1 Methodology

In the first step, we should provide a method to generate an arbitrary forest $\mathcal{F} = (V, E_{\mathcal{F}})$. This can be done using the notion of the Prüfer sequence of a labeled tree which is a unique sequence associated with the tree (for more details, see Lewis, 1999) Now, the arbitrary generated forest $\mathcal{F}$ is consisting of trees $\mathcal{T}_1, ..., \mathcal{T}_c$. One can sample max-stable Hüsler–Reiss observations on these trees using the function `rmstable_tree` from Engelke, Hitz, et al., 2022. By Corollary (3.4.3), the joint observations of Hüsler–Reiss max-stable over the extremal trees conditioned on the event that at least one of these observations is larger than an increasing threshold u, when u tends to ∞, is an observation of Hüsler–Reiss Pareto distribution over the extremal forest consisting of the trees. One can simulate data from an underlying extremal forest structure for any pre-asymptotic u, using this approach. In the following, instead of $u \to \infty$, we increase