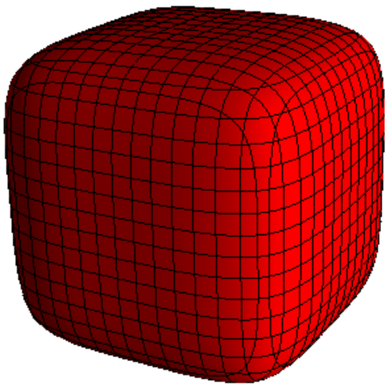
Lehrstuhl für Statistik
Institut für Mathematik
Universität Würzburg

# Characterization of the $D$-Norm Corresponding to a Multivariate Extreme Value Distribution



Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades
der Bayerischen Julius–Maximilians–Universität Würzburg

vorgelegt von


Daniel Hofmann

aus
Coburg


August 2009

# Contents

CHAPTER 1

# Introduction

Extreme value theory is an active research area, mainly due to its applications in various fields like the investigation of water levels of rivers (Michel (2006, [**26**])), corrosion of materials (Rivas et. al. (2008, [**34**])), wind speeds (de Haan and Ronde (1998, [**9**])) or insurance data (Reiss and Thomas (2007, [**33**])), only to name a few.
In the last years multivariate extreme value theory became particularly more interesting since there is the demand of practitioners for statistical tools not only to analyze data from independent rare events, but also from rare events that are not independent from each other. This occurs for instance when rare events are analyzed at several sites that are far away from each other so that they are not completely dependent but still not far away enough to be completely independent.

The main part of this thesis deals with a representation of multivariate extreme value distributions in arbitrary dimension. It is well-known known that a $d$-dimensional extreme value distribution (EVD) $G$ with negative exponential margins can be represented as $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|_D)$, $\mathbf{x} \le \mathbf{0}$, where $\|\cdot\|_D$ is the so called <u>$D$-norm</u>. This $D$-norm can be expressed in terms of the <u>Pickands dependence function</u> $D$ via

$$\|\mathbf{x}\|_D = \|\mathbf{x}\|_1 D\left(|x_1| / \|\mathbf{x}\|_1, \dots, |x_{d-1}| / \|\mathbf{x}\|_1\right),$$

where $\|\mathbf{x}\|_1 = \sum_{j\le d} |x_j|$ denotes the usual $L_1$-norm of $\mathbf{x} \in \mathbb{R}^d$. We refer to Section 4.3 in Falk et al. (2004, [**13**]) for more details.
As shown in Falk (2006, Remark 1), there are norms $\|\cdot\|$ on $\mathbb{R}^d$ that are <u>not</u> $D$-norms, i.e. there are norms $\|\cdot\|$ such that $\exp(-\|\mathbf{x}\|)$, $\mathbf{x} \le \mathbf{0}$, does not define a distribution function (df). For the bivariate case $d = 2$, Falk (2006, [**12**]) states a necessary and sufficient condition for a norm to obtain a distribution function. But this condition is not sufficient in the case of dimension $d \ge 3$, see Section 6.2 in Hofmann (2006, [**22**]) and Lemma 5.3.1 below.

Therefore a characterization of the $D$-Norm and hence also for the Pickands dependence function is still an open issue which this thesis aims to settle.

Chapter 2 gives an introduction to the theory of multivariate extreme value distributions and multivariate generalized Pareto distributions. In Theorem 2.2.2 we answer the open question whether there are $1 \leq \lambda < \infty$ for which $W_\lambda(\mathbf{x}) := 1 - \|\mathbf{x}\|_\lambda$, $\|\mathbf{x}\|_\lambda \leq 1$ defines a distribution function over its entire support in dimension 3 and higher, which turns out to be not the case. By $\|\mathbf{x}\|_\lambda = (\sum_{j \leq d} |x_j|^\lambda)^{1/\lambda}$ for $\lambda \in [0, \infty)$ and $\|\mathbf{x}\|_\lambda = \max_{i \in \{1,\dots,d\}} |x_i|$ for $\lambda = \infty$ we denote the usual $\lambda$-norm on $\mathbb{R}^d$.

In Chapter 3 we will state a necessary and sufficient condition for a norm in $\mathbb{R}^d$, such that $G(x) := \exp(-\|\mathbf{x}\|)$, $\mathbf{x} \leq \mathbf{0}$, defines a distribution function. In this case, $G$ is obviously an EVD with negative exponential margins $G_i(x_i) = \exp(x_i)$, $x_i \leq 0$, $i \leq d$. This is the Main Theorem of this thesis. Thus the Main Theorem provides a characterization of the $D$-Norm. There are already other one-to-one representations of multivariate extreme value distributions as the exponent measure (see Balkema and Resnick (1977, [**2**])) and the angular measure (see de Haan and Resnick (1977, [**11**])). More details about the exponent measure and the angular measure are provided in Section 2.1.

Molchanov (2008, [**29**]) developed a completely different approach to this problem in terms of convex geometry and the theory of random sets. His access to this topic will be presented in Chapter 4. In fact using his results it is possible to give an alternative proof of the Main Theorem. Since our proof of the Main Theorem uses only results from measure theory we think that it is easier accessible.

Applications of the Main Theorem are given in Chapter 5. The bivariate case is examined in Section 5.1. In Section 5.2 the Main Theorem is carried over to the Pickands dependence function and a necessary and sufficient condition for a function to be a Pickands dependence function is given.
Section 5.3 uses the Main Theorem to show that a condition for the nested logistic model, which is known to be sufficient, is also necessary.
The last two sections 5.4 and 5.5 in this chapter introduce ways to construct new

norms that define an EVD using the Main Theorem.

Chapter 6 is based on a theorem from a yet unpublished paper by Aulbach, Bayer and Falk ([**1**]), that goes back to Buishand et al. (2008, [**7**]). As a first consequence we can use this theorem to specify the left neighborhood in the definition of a GPD. Theorem 6.2.1 shows that this left neighborhood can be chosen to be $\left[-\frac{1}{d}, 0\right]^d$.
Furthermore we introduce the GPD-Flow. The theorem from Aulbach, Bayer and Falk can be used to obtain a GPD as a function of a copula. Since the GPD has again an underlying copula, this step can be iterated over and over again, which will be called the GPD-Flow. Simulations indicate, that the GPD-Flow converges against the copula of complete dependence. Nevertheless the convergence of the GPD-Flow is not yet proven, but in Theorem 6.5.4 we see that if it converges it must be the copula of complete dependence.

Chapter 7 deals with the simulation of random vectors following a GPD. The Shi-Transformation for generating random vectors that follow a GPD from the logistic type introduced by Michel (2006, [**26**]) is generalized in dimension 3 to generate random vectors from a GPD of the nested logistic model.

The restriction of an EVD to have negative exponential margins is not a real constraint since a transformation of the one dimensional margins to these margins can always be achieved. Further information will be provided in Section 2.1.

Throughout this thesis all vectors are denoted in bold letters and, if not explicitly stated otherwise, the components of a vector $\mathbf{x}$ are given by $x_1, \ldots, x_d$. Furthermore, all operations on vectors such as $\mathbf{x} + \mathbf{y}$, $\max(\mathbf{x}, \mathbf{y})$ and $\mathbf{x} \leq \mathbf{y}$ etc. are meant component wise. We define $0^0 := 1$ and $\infty^0 := 1$. By the symbol $\subsetneq$ we denote a real subset, i.e. $A \subsetneq B$ means that $a \in A \Rightarrow a \in B$ but $A \neq B$.
By $^\complement$ we denote the complement of a set, i.e. for a universe $U$ and a subset $A \subset U$ the complement of $A$ in $U$ is denoted by $A^\complement$ and given by $A^\complement = U \backslash A$.
Furthermore we set $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$ and $\mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}$.
We denote the $i$-th value of the data $x_1, \ldots, x_n$ (in non-decreasing order) by $x_{i:n}$.
Finally $I(x \leq y)$ denotes the indicator function with $I(x \leq y) = 1$ of $y \leq x$ and 0, otherwise.

CHAPTER 2

# Multivariate Extreme Value and Generalized Pareto Distributions

In this chapter an introduction to multivariate extreme value and generalized Pareto distributions is given. We assume that basic concepts of the univariate theory are known. Reiss and Thomas (2007, [**33**]) give in Section 1.3 and 1.4 an overview of these distributions.

## 2.1. Multivariate Extreme Value Distributions

We start this section with the definition of a multivariate extreme value distribution as given in Section 12.1 in Reiss and Thomas (2007, [**33**]).

DEFINITION 2.1.1. *We call a d-variate distribution function $G$ an <u>extreme value distribution (EVD)</u> if and only if $G$ is <u>max-stable</u> that is for certain vectors $\mathbf{b}_n$ and $\mathbf{a}_n > \mathbf{0}$ it is*

$$G^n\left(\mathbf{b}_n + \mathbf{a}_n\mathbf{x}\right) = G\left(\mathbf{x}\right).$$

Since the univariate marginal distributions of an EVD are univariate EVD the Theorem of Fisher-Tippett (see Fisher and Tippet (1928, [**15**])) implies that the univariate margins is either a Gumbel, Fréchet or Weibull distribution.
A multivariate distribution function consists of univariate marginal distributions and the dependence among them. In general multivariate distribution theory a common approach to model the dependence is the concept of a copula. We refer to Nelsen (2006, [**31**]) for an introduction to copulas and the 9th issue of the journal <u>Extremes</u> in 2006 for a controversial discussion about the usefulness of copulas.
However in extreme value theory other concepts of modeling dependence are common. Next we will give the definition of the Pickands dependence function for which we will establish new results in Section 5.2.

THEOREM 2.1.2. *A $d$-variate extreme value distribution $G$ with negative exponential univariate margins can be written as*

$$G(\mathbf{x}) = \exp\left(\int_{S_d} \min_{j \le d}(u_j x_j) \,\mathrm{d}\,\mu(\mathbf{u})\right), \mathbf{x} < 0,$$

*where $\mu$ is a finite measure on the $d$-variate unit simplex*

$$S_d = \left\{\mathbf{u} : \sum_{j \le d} u_j = 1, u_j \ge 0\right\}$$

*with*

$$\int_S u_j \,\mathrm{d}\,\mu(\mathbf{u}) = 1, j \le d,$$

*see Theorem 4.3.1 Falk et al. (2004, [**13**]).*

COROLLARY 2.1.3. *The Pickands dependence function $D$, $D : \overline{R}_{d-1} \to [0, \infty)$ is defined by*

$$D(t_1, \ldots, t_{d-1}) := \int_{S_d} \max\left(u_1 t_1, \ldots, u_{d-1} t_{d-1}, u_d \left(1 - \sum_{i \le d-1} t_i\right)\right) \mathrm{d}\,\mu(\mathbf{u}),$$

*where*

$$\overline{R}_d := \left\{(t_1, \ldots, t_d) \in [0,1]^d : \sum_{i \le d} t_i \le 1\right\}.$$

*Therefore $G$ can be written as*

$$G(\mathbf{x}) = \exp\left((x_1 + \cdots + x_d)\, D\left(\frac{x_1}{x_1 + \cdots + x_d}, \ldots, \frac{x_d}{x_1 + \cdots + x_d}\right)\right).$$

We now list some properties of the Pickands dependence function $D$ taken from Section 4.3 in Falk et al. (2004, [**13**]).

(i) $D$ is continuous.

(ii) We have

$$D(\mathbf{0}) = D(\mathbf{e}_i) = 1, \quad 1 \le i \le d-1,$$

where $\mathbf{e}_i$ denotes the $i$-th unit vector in $\mathbb{R}^{d-1}$.

(iii) It is $D(\mathbf{t}) \le 1$ for all $\mathbf{t} \in \overline{R}_{d-1}$.

(iv) $D$ is a convex function, i.e. for $\mathbf{s}, \mathbf{t} \in \overline{R}_{d-1}$ and $\lambda \in [0,1]$ we have

$$D(\lambda \mathbf{s} + (1-\lambda)\mathbf{t}) \le \lambda D(\mathbf{s}) + (1-\lambda) D(\mathbf{t}).$$

(v) For any $\mathbf{t} \in \overline{R}_{d-1}$ we have

$$D(\mathbf{t}) \geq \max\left(t_1, \dots, t_{d-1}, 1 - \sum_{j=1}^{d-1} t_i\right) \geq \frac{1}{d}.$$

(vi) The convex combination of two Pickands dependence functions $D_1$ and $D_2$ is again a Pickands dependence function, i.e. for every $\lambda \in [0,1]$ $D(\mathbf{t}) = (1-\lambda) D_1(\mathbf{t}) + \lambda D_2(\mathbf{t})$ is a Pickands dependence function.

(vii) If we set $\|\mathbf{x}\|_D := \|\mathbf{x}\|_1 D\left(\frac{|x_1|}{\|\mathbf{x}\|_1}, \dots, \frac{|x_{d-1}|}{\|\mathbf{x}\|_1}\right)$ for $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\|\mathbf{0}\|_D := 0$, then $\|\cdot\|_D$ defines a norm on $\mathbb{R}^d$, the so called $D$-Norm.

Besides the Pickands dependence function and the norm, we deal with in our Main Theorem, there are other representations of multivariate extreme value distributions, all having advantages and disadvantages. A $d$-variate EVD $G$ with negative exponential margins can be represented as

$$\begin{aligned} G(x_1, \dots, x_d) &= \exp\left(-\mu\left((([-\infty, x_1] \times \cdots \times [-\infty, x_d])^{\complement}\right)\right) \\ &= \exp\left(\int_{S_E} \min_{i \leq d}(a_i x_i) \,\mathrm{d}\,\nu(\mathbf{a})\right) \\ &= \exp\left(-l(x_1, \dots, x_d)\right) \\ &= \mathcal{D}(\exp(x_1), \dots, \exp(x_d)) \\ &= \Omega(\exp(x_1), \dots, \exp(x_d)) \exp(x_1 + \cdots + x_d), \quad x_1, \dots, x_d < 0, \end{aligned}$$

where $S_E$ denotes the unit-sphere pertaining to the norm that underlies $\nu$.
$\mu$ is called exponent measure (see Balkema and Resnick (1977, [**2**])), $\nu$ is the angular measure (see de Haan and Resnick (1977, [**11**])), $l(\cdot)$ is the stable tail dependence function (see Huang (1992, [**23**])), $k$ is the dependence function of Tiago de Oliveira (see Tiago de Oliveira (1966, [**43**]), $\mathcal{D}$ is the dependence function of Galambos (see Definition 5.2.1 of Galambos (1978, [**19**]); note that Galambos denotes his dependence function by $D$, but due to the danger of confusion with the Pickands dependence function, we rename it by $\mathcal{D}$). In fact, the dependence function of Galambos is actually a copula and those copulas are also called extreme value copulas (see Nelsen (2006, [**31**])). Finally $\Omega$ is the dependence function of Sibuya (see Sibuya (1960, [**39**])). There are several approaches to estimate those dependency structures, see for instance the book by de Haan and Ferreira (2006, [**10**]).

All the dependency representations above are given due to negative exponential margins. With an easy transformation the margins of any EVD can be transformed to be negative exponential (see Lemma 5.4.7 in Falk et. al(2004, [**13**])). The nondegenerate univariate EVDs can be parametrized by one parameter $\alpha \in \mathbb{R}$ with

$$G_\alpha(x) := \begin{cases} \exp\left(-(-x)^\alpha\right), & x \le 0 \\ 1, & x > 0 \end{cases} \text{ for } \alpha > 0,$$
$$G_0(x) := \exp\left(-\exp(-x)\right), x \in \mathbb{R}$$

and

$$G_\alpha(x) := \begin{cases} 0, & x \le 0 \\ \exp\left(-x^\alpha\right), & x > 0 \end{cases} \text{ for } \alpha < 0.$$

For $\alpha > 0$ we have the family of Weibull, for $\alpha < 0$ the family of Fréchet and for $\alpha = 0$ the Gumbel distribution. In this terms the negative exponential distribution is the $G_1$ distribution.
For any $\alpha_i \in \mathbb{R}$ we define the function $\psi_{\alpha_i} : \{x \in \mathbb{R} : 0 < G_{\alpha_i}(x) < 1\} \mapsto \mathbb{R}$ by

$$\begin{aligned} \psi_{\alpha_i}(x) &:= \log\left(G_{\alpha_i}(x)\right) \\ &= \begin{cases} -(-x)^{\alpha_i}, & x < 0, \text{ if } \alpha_i > 0 \\ -\exp(-x), & x \in \mathbb{R}, \text{ if } \alpha_i = 0 \\ -x^{\alpha_i}, & x > 0, \text{ if } \alpha_i < 0 \end{cases}. \end{aligned}$$

With $\psi_{\alpha_i}$ it is possible to transform any EVD to negative exponential margins. Let $G_{(\alpha_1,\dots,\alpha_d)}$ be a multivariate EVD whose $i$-th margin is an EVD with parameter $\alpha_i$. Then we have

$$G_{(\alpha_1,\dots,\alpha_d)}(x_1,\dots,x_d) = G_{(1,\dots,1)}\left(\psi_{\alpha_1}(x_1),\dots,\psi_{\alpha_d}(x_d)\right),$$

see Lemma 5.4.7 in Falk et. al (2004, [**13**]).

Finally we will study in more detail the exponent measure.

Definition 2.1.4. *A $\sigma$-finit measure $\mu$ on $[-\infty,\infty)^d$ is called <u>exponent measure</u> of the distribution function $F(\mathbf{x}) := \exp\left(-\mu\left([-\boldsymbol{\infty},\mathbf{x}]^{\complement}\right)\right)$.*

The concept of a max-stable distribution function will be extended to max-infinite divisible distribution functions.

DEFINITION 2.1.5. *A distribution function $F$ will be called max-infinitely divisible (max-id) if for every nonnegative integer $n$ there is a distribution function $F_n$ such that*

$$F_n^n = F.$$

Obviously a max-stable distribution function is also max-infinitely divisible (set $F_n(\mathbf{x}) = F(\mathbf{b}_n + \mathbf{a}_n\mathbf{x})$). In the univariate case every distribution function is max-id since $F_n = F^{\frac{1}{n}}$ is a distribution function.

The following characterization shows that max-id distribution function are in a one-to-one relationship with the exponent measure.

THEOREM 2.1.6 (Balkema and Resnick). *A distribution function $F$ on $\mathbb{R}^d$ is max-id if and only if it has an exponent measure.*

PROOF. The bivariate case is Theorem 3 in Balkema and Resnick (1977, [**2**]). □

## 2.2. Multivariate Generalized Pareto Distributions

In the univariate extreme value theory the limit distribution of peaks over a threshold is given by a generalized Pareto distribution $W$ which turns out to be in a simple relation to the EVD, namely

$$W(x) = 1 + \log(G(x), \text{ if } \log G(x) > -1.$$

However there is no natural generalization of limiting distributions for multivariate peaks over threshold. There are three different approaches closely related to each other. First Kaufmann and Reiss (1995, [**25**]) introduced a definition for the bivariate case which is generalized to arbitrary dimension in Section 5.1 of Falk et al. (2004, [**13**]). In this manuscript we will stick to this definition. Other definitions are given by Tajvidi (1996, [**41**]) and Beirlant et al. (2004, [**5**], Section 8.3) which is more investigated in Rootzén and Tajvidi (2006, [**35**]). But in the region of interest, namely $\{(x, y) : u < x \leq 0, v < y \leq 0\}$ with $(u, v) < \mathbf{0}$ and close enough to the origin, these definitions are all identical, see Section 13.1 in

Reiss and Thomas (2007, [**33**]) and Section 8.3 of Beirlant et al. (2004, [**5**]).
We start with the definition of a generalized Pareto distribution we use within this manuscript. As mentioned above this is the one in Section 5.1 of Falk et al. (2004, [**13**]).

Definition 2.2.1. *A d-variate distribution function W will be called a multivariate generalized Pareto distribution function (GPD) if there is some EVD G that*

$$W(\mathbf{x}) = 1 + \log G(\mathbf{x}),$$

*where* $\mathbf{x} \leq \mathbf{0}$ *and* $\mathbf{x}$ *is in a left neighborhood of* $\omega(G) = (\omega(G_1), \ldots, \omega(G_d))$, $\omega(G_i) = \sup\{x \in \mathbb{R} : G_i(x) < 1\}$.
*Furthermore we call the function*

$$W(\mathbf{x}) = 1 + \log G(\mathbf{x}), \log G(\mathbf{x}) \geq -1,$$

*a generalized Pareto function (GPF).*

Obviously we can restrict ourselves to GPDs coming from an EVD with negative exponential margins since we can apply the transformation $\psi_{\alpha_i}$ on the coordinates as in the case of the EVD, see page 12. For more details see Corollary 5.4.8 in Falk et. al (2004, [**13**]).

In the bivariate case the GP function is actually a distribution function, see Kaufmann and Reiss (1995, [**25**]) or Lemma 5.1.1 in Falk et al. (2004, [**13**]), but in dimension 3 and higher this is no longer valid.
An example is given in Section 5.1 in Falk et al. (2004, [**13**]). There it is shown that for $W(x_1, x_2, x_3) = \max(1 + x_1 + x_2 + x_3, 0)$ the cube $\left(-\frac{1}{2}, 0\right]^3$ would get probability $-\frac{1}{2}$ and therefore $W$ cannot be a distribution function. In Theorem 2.3.12 in Michel (2006, [**26**]) this example was extended to dimension greater than 3. By using the continuity of the $\lambda$-norms Michel also showed that there exists a $\lambda_0 > 1$ that for all $\lambda \in [1, \lambda_0)$ the GP function $W_\lambda(\mathbf{x}) := \max(1 - \|\mathbf{x}\|_\lambda, 0)$ does not define a distribution function.
We will extend this example one step further by showing that for dimension 3 or higher and arbitrary and finite $\lambda \geq 1$ the GP function $W_\lambda(x) = \max(1 - \|\mathbf{x}\|_\lambda, 0)$ does not define a distribution function. This will be established by showing that a certain cube would get a negative probability. Since in the case of $\lambda = \infty$ the GP function $W_\infty(\mathbf{x}) := \max(1 - \|\mathbf{x}\|_\infty, 0)$ is a distribution function this cube

must depend on $\lambda$ and as $\lambda$ converges to $\infty$ this cube must converge to a cube having probability 0. Otherwise the continuity of the $\lambda$-norms implies that also for $\lambda = \infty$ the cube would have negative probability which cannot be true.

THEOREM 2.2.2. *For any $\lambda \in [1, \infty)$, the GP function*

$$W_\lambda(\mathbf{x}) = \max\left(1 - \|\mathbf{x}\|_\lambda, 0\right) =: \left(1 - \|\mathbf{x}\|_\lambda\right)^+, \mathbf{x} \le \mathbf{0},$$

*does not define a distribution function for $d \ge 3$.*

PROOF. Let $d = 3$. Assume that $W_\lambda$ does define a distribution function. We define the two points

$$\mathbf{a}_{x,\lambda} := \left(0, 0, -(1-2x)^{\frac{1}{\lambda}}\right)^T$$

and

$$\mathbf{b}_{x,\lambda} := \left(-x^{\frac{1}{\lambda}}, -x^{\frac{1}{\lambda}}, -(1-x)^{\frac{1}{\lambda}}\right)^T$$

for $x \in \left[0, \frac{1}{2}\right]$.
Obviously these two points satisfy $\mathbf{b}_{x,\lambda} \le \mathbf{a}_{x,\lambda} \le 0$. Now we calculate the probability of the set $K_{x,\lambda} := (\mathbf{b}_{x,\lambda}, \mathbf{a}_{x,\lambda}]$:

$$\begin{aligned}
h_\lambda(x) :=& \mathrm{P}\left(K_{x,\lambda}\right) \\
=& W_\lambda\left((a_1, a_2, a_3)^T\right) \\
& - W_\lambda\left((b_1, a_2, a_3)^T\right) - W_\lambda\left((a_1, b_2, a_3)^T\right) - W_\lambda\left((a_1, a_2, b_3)^T\right) \\
& + W_\lambda\left((a_1, b_2, b_3)^T\right) + W_\lambda\left((b_1, a_2, b_3)^T\right) + W_\lambda\left((b_1, b_2, a_3)^T\right) \\
& - W_\lambda\left((b_1, b_2, b_3)^T\right) \\
=& \left(1 - (1-2x)^{\frac{1}{\lambda}}\right)^+ - 3\left(1 - (1-x)^{\frac{1}{\lambda}}\right)^+ + 3(1-1)^+ - \left(1 - (1+x)^{\frac{1}{\lambda}}\right)^+ \\
=& 3(1-x)^{\frac{1}{\lambda}} - (1-2x)^{\frac{1}{\lambda}} - 2.
\end{aligned}$$

Evaluation of $h_\lambda$ at 0 shows $h_\lambda(0) = 0$. The function $h_\lambda$ is differentiable in the interior of its domain and we obtain

$$h'_\lambda(x) = \frac{1}{\lambda}\left(2(1-2x)^{\frac{1-\lambda}{\lambda}} - 3(1-x)^{\frac{1-\lambda}{\lambda}}\right).$$

Thus

$$\lim_{x \to 0} h'_\lambda(x) = -\frac{1}{\lambda}.$$

Hence for $\epsilon_2 := \frac{1}{2\lambda}$ there exists a $\delta_2 > 0$ such that $h'_\lambda(x) \in \left(-\frac{1}{\lambda} - \epsilon_2, -\frac{1}{\lambda} + \epsilon_2\right) = \left(-\frac{3}{2\lambda}, -\frac{1}{2\lambda}\right)$ for $x \in (0, \delta_2]$.
The continuity of $h_\lambda$ implies, that for $\epsilon_1 := \delta_2 \frac{1}{16\lambda}$ there exists a $\delta_1 > 0$ such that $h_\lambda(x) \in (-\epsilon_1, \epsilon_1)$ for $x \in [0, \delta_1]$.
With $\delta := \min\left(\delta_1, \frac{3}{4}\delta_2\right)$ and $\xi \in [\delta, \delta_2]$ we use Taylor's Theorem and obtain

$$\begin{aligned} h_\lambda(\delta_2) &= h_\lambda(\delta) + (\delta_2 - \delta) h'_\lambda(\xi) \\ &< \epsilon_1 + \frac{\delta_2}{4}\left(-\frac{1}{\lambda} + \epsilon_2\right) \\ &= \delta_2 \frac{1}{16\lambda} - \frac{\delta_2}{8\lambda} \\ &= -\frac{\delta_2}{16\lambda} \\ &< 0. \end{aligned}$$

Since $K_{x,\lambda}$ would have a negative probability this is a contradiction to the assumption of a distribution function and hence for $d = 3$ the GP function is not a distribution function.
Now suppose that the GP function $W_\lambda$ is a distribution function for $d > 3$. Then the marginal distribution of the first three components is the GP function $W_\lambda$ of dimension 3. But this is not a distribution function and we have to reject the assumption that the $d$-dimensional GP function is a distribution function.

□

A useful approach in dealing with GPD is the decomposition of the coordinates in an angular and a radial component using the so called Pickands coordinates.

Definition 2.2.3.
*For $d \in \mathbb{N}$, $d \geq 2$, define the transformation $T_P : (-\infty, 0]^d \setminus \{0\}^d \to \bar{R}_{d-1} \times (-\infty, 0]$ by*

$$T_P(\mathbf{x}) := \left(\frac{x_1}{x_1 + \cdots + x_d}, \ldots, \frac{x_{d-1}}{x_1 + \cdots + x_d}, x_1 + \cdots + x_d\right) =: (z_1, \ldots, z_{d-1}, c)$$

*with*

$$\overline{R}_d := \left\{ \mathbf{x} \in (-\infty, 0]^d : \sum_{i=1}^{d} x_i \leq 1 \right\}.$$

*$T_P$ is called transformation to (standard) Pickands coordinates* $\mathbf{z} := (z_1, \ldots, z_{d-1})$*, $c$. At this $z$ is called angular component and $c$ is called radial component.*

The Pickands coordinates are similar to the polar coordinates which consists also of an angular and a radial component but the polar coordinates use the euclidean norm $\|\cdot\|_2$ for the components contrary to the sum norm $\|\cdot\|_1$ used by the Pickands coordinates. For more information on $T_P$ we refer to Falk and Reiss (2005, [**14**]) and Section 5.4 of Falk et al. (2004, [**13**]).

Next we introduce the so called Pickands density.

Definition 2.2.4. *For a GPD $W$ that has, in a left neighborhood of $\mathbf{0}$, continuous partial derivatives of order $d$, the function*

$$\phi(\mathbf{z}) := |c|^{d-1} \left( \frac{\partial^d}{\partial x_1 \ldots \partial x_d} W \right) \left( T_P^{-1}(z, c) \right), \mathbf{z} \in R_{d-1}$$

*with*

$$R_d := \left\{ \mathbf{x} \in (0, \infty)^d : \sum_{i=1}^{d} x_i < 1 \right\}$$

*is called the Pickands density.*

Note that the Pickands density does not depend on $c$ (see Theorem 5.4.2 in Falk et al. (2004, [**13**])). Furthermore assume that $(X_1, \ldots, X_d)$ follows a differentiable GPD $W$. Let $C := X_1 + \cdots + X_d$ and $Z := \left( \frac{X_1}{C}, \ldots, \frac{X_{d-1}}{C} \right)$ be the standard random Pickands coordinates. Then conditional on $C > c_0$ for $c_0 < 0$ close to 0 the Pickand coordinate $Z$ has the density

$$f(\mathbf{z}) = \frac{\phi(\mathbf{z})}{\int_{R_{d-1}} \phi(\mathbf{v})\, \mathrm{d}\, \mathbf{v}},$$

see also Theorem 5.4.2 in Falk et al. (2004, [**13**]).
For more details see Section 5.4 in Falk et al. (2004, [**13**]) or Section 2.1 in Michel (2006, [**26**]).

As mentioned in the previous section one of the dependence structures for an EVD is the angular measure $\nu$ (see page 2.1). Therefore a multivariate GPD can also be written in a left neighborhood of $\mathbf{0}$ in terms of the angular measure namely

$$W(x_1, \ldots, x_d) := \\ 1 - \int_{\bar{R}_{d-1}} \min\left(u_1 x_1, \ldots, u_{d-1} x_{d-1}, \left(1 - \sum_{i=1}^{d-1} u_i\right) x_d\right) \nu(\mathrm{d}\, u).$$

If the angular measure $\nu$ restricted to $R_{d-1}$ possesses a density $l$ we call it the <u>angular density</u>.

CHAPTER 3

# Main theorem

## 3.1. The Main Theorem

The following theorem is the main result of this thesis which gives a one-to-one characterization for multivariate extreme value distribution with negative exponential margins in terms of a norm.

THEOREM 3.1.1 (Main Theorem). *For any norm* $\|\cdot\|$ *on* $\mathbb{R}^d$ *the following assertions are equivalent:*

(i) *the function* $G(\mathbf{x}) := \exp(-\|\mathbf{x}\|)$, $\mathbf{x} \leq \mathbf{0}$, *defines a multivariate extreme value distribution function*

(ii) *there exists a measure* $\mu$ *on* $[-\infty, \infty) \setminus \{-\infty\}$ *with*

$$\mu\left([-\infty, \mathbf{x}]^{\complement}\right) = \begin{cases} \|\mathbf{x}\|, & \text{for } \mathbf{x} \leq \mathbf{0} \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

(iii) *the norm satisfies*

$$\sum_{\substack{\mathbf{m} \in \{0,1\}^d \\ m_i = 1, i \in K}} (-1)^{d+1-\sum_{j \leq d} m_j} \left\| \left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right) \right\| \geq 0 \tag{3.2}$$

*for every* $K \subsetneq \{1, \ldots, d\}$ *and* $-\infty < a_j \leq b_j \leq 0$, $1 \leq j \leq d$.

*A norm which fulfills one (and therefore all) conditions from above will be called a D-Norm.*

If $G$ is a multivariate distribution function, then we have with the laws for the exponential function and the homogeneity of the norm

$$G^n\left(\frac{1}{n}\mathbf{x}\right) = \exp\left(-\left\|\frac{1}{n}\mathbf{x}\right\|\right)^n$$

$$
\begin{aligned}
&= \exp\left(-n\frac{1}{n}\|\mathbf{x}\|\right) \\
&= G(\mathbf{x}).
\end{aligned}
$$

Thus $G$ is max-stable (see Definition 2.1.1) and hence $G$ is an extreme value distribution.
Because a $d$-dimensional distribution function $G$ is max-id if and only if it has an exponent measure (see Theorem 2.1.6), the assertion (i) is equivalent to (ii). Therefore we only have to proof that (ii) is equivalent to (iii).
Since the proof is rather long it will be splitted into several lemmas and corollaries given in the next sections.

## 3.2. Proof of the sufficiency

First we give an outline of the proof of the sufficiency.
We start by showing that condition (3.2) in the Main Theorem is equivalent to another condition. Then we construct in a similar way to the Lebesgue measure (see Section 1.4 and 1.6 in Bauer (1972, [**3**])) several measures on $(-\infty, 0] \times \cdots \times \{-\infty\} \times \cdots \subset [-\infty, 0]^d$, i.e. the spaces where in certain components we have the nonpositive real numbers (with $-\infty$ excluded) and in the other components we fix the point $\{-\infty\}$. There the condition (3.2) from the Main Theorem 3.1.1 implies that we can really construct these measure since the condition guarantees that the measures obtain nonnegative values.
Using these measures we construct a new measure $\mu^*$ on $[-\boldsymbol{\infty}, \mathbf{0}] \setminus \{-\boldsymbol{\infty}\}$.
Then we prove that if a certain condition is fulfilled then $\mu^*$ obtains certain values for certain sets.
Furthermore we show that $\mu^*$ has the desired property that $\mu^*\left([-\boldsymbol{\infty}, \mathbf{x}]^{\complement}\right) = \|\mathbf{x}\|$.
Finally the proof will be finished by showing that the certain condition named above follows from the condition of the Main Theorem.

LEMMA 3.2.1. *Condition* (3.2) *in the Main Theorem is satisfied if and only if for every* $K \subsetneq \{1, \ldots, d\}$, *for every* $L \subseteq K$ *and* $-\infty < a_j \le b_j \le 0$, $1 \le j \le d$ *it is*

$$
\lim_{t \to -\infty} \sum_{\substack{\mathbf{m} \in \{0,1\}^d \\ m_i = 1, i \in K \\ b_i = t, i \in L}} (-1)^{d+1-\sum_{j \le d} m_j} \left\| \left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right) \right\| \in [0, \infty) \tag{3.3}
$$

Proof. Condition (3.3) implies condition (3.2) by just setting $L = \emptyset$.
Now we prove the converse implication.
We define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K \\ b_i=x, i\in L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\|$$

and with this new function condition (3.3) can be reformulated as $\lim_{x\to-\infty} f(x) \in [0,\infty)$. First we show that the limit exists. We begin this proof by showing that $f$ is bounded from below. In the sequel we use the inequality $\big| \|\mathbf{x}\| - \|\mathbf{y}\| \big| \le \|\mathbf{x}-\mathbf{y}\|$ derived from the triangle-inequality for norms.
Choose an index $q \in \{1,\dots,d\} \setminus K$. We have

$$\begin{aligned}
f(x) &= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K \\ b_i=x, i\in L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K \\ b_i=x, i\in L \\ m_q=1}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&\quad + \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K \\ b_i=x, i\in L \\ m_q=0}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K \\ b_i=x, i\in L \\ m_q=1}} (-1)^{d+1-\sum_{j\le d} m_j} \Bigg( \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, b_q, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&\quad - \left\| \left( b_1^{m_1} a_1^{1-m_1}, \dots, a_q, \dots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \Bigg) \\
&\in \left[ 2^{d-|K|-1} (a_q - b_q), 2^{d-|K|-1} (b_q - a_q) \right].
\end{aligned}$$

Therefore $f$ is bounded from below.

Next we show that $f$ is increasing in $x$. Without loss of generality it suffices to prove that $f$ is increasing if $L = \{1\}$ (if $|L| = 1$ it can be proved in a completely analogous way and the case that $|L| > 1$ follows from the fact that $f$ is increasing in every argument).

Let $c_1 \in \mathbb{R}$ with $a_1 < c_1 < b_1$.
We have

$$
\begin{aligned}
f(b_1) &= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( c_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&\quad \underbrace{+ \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K\setminus\{1\}}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_1^{m_1} c_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d} \right) \right\|}_{\geq 0} \\
&\geq \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( c_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d} \right) \right\| \\
&= f(c_1).
\end{aligned}
$$

Thus we have shown that $f$ is increasing and together with the boundness from below we obtain the convergence of $f$.
Finally we need to prove that $f$ is nonnegative. Condition (3.2) implies that for all $x \in (-\infty, 0]$ we have $f(x) \geq 0$.
Now assume that $\lim_{x\to-\infty} f(x) = y < 0$. Because of the definition of the limit, there exists an $S \in \mathbb{R}$ such that $|f(x) - y| < y$ for all $x < S$. As a consequence $f(x) < 2y < 0$ for all $x < S$. But this is a contradiction to condition (3.2) and we have $\lim_{x\to-\infty} f(x) \geq 0$. Hence everything is proved.

$\square$

Let $\mathcal{K} := \{K : K \subsetneq \{1, \ldots, d\}\}$. We define for every $K \in \mathcal{K}$

$$
\begin{aligned}
\mathcal{I}_K := \big\{ \times_{k=1}^d I_k : I_k = \{-\infty\} \text{ for } k \in K, I_k = [a_k, b_k) \\
\text{with } -\infty < a_k \leq b_k \leq 0 \text{ for } k \notin K \big\}.
\end{aligned}
$$

Then $\mathcal{I}_K$ is a semiring (see Definition A.1 in the appendix) in $[-\infty, 0)^d$ and the sets $\mathcal{I}_K, K \in \mathcal{K}$ are pairwise disjoint.
In the following we assume that condition (3.2) always holds and because of Lemma 3.2.1 also condition (3.3).

Define furthermore $\mu_K$ on $\mathcal{I}_k$ by

$$\mu_K(I) := \lim_{t\to-\infty} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right) \right\| \tag{3.4}$$

with $b_{(i)} = t$, if $i \in K$, $b_{(i)} = b_i, a_{(i)} = a_i$ otherwise. Condition (3.3) implies that $\mu_K(I) \ge 0$ for $I \in \mathcal{I}$.

Similar to the Lebesgue measure (see Section 1.4 and 1.6 in Bauer (1972, [**3**])) we will show that $\mu_K$ defines a measure. For this purpose we proof that $\mu_K$ is additiv (Lemma 3.2.2) and then that it is even $\sigma$-finite (Corollary 3.2.7). With the $\mu_K$ we define a new measure $\mu$ on a ring $\mathcal{F}$ which will then be extended to a measure $\mu^*$ on the $\sigma$-algebra generated by $\mathcal{F}$ (Theorem 3.2.10). Finally we proof that $\mu^*$ is an exponent measure with the desired property $\mu^*\left([-\boldsymbol{\infty}, \mathbf{x}]^{\complement}\right) = \|\mathbf{x}\|$ (see Lemma 3.2.13).

Lemma 3.2.2. *$\mu_K$ is additiv, i.e. for $n \in \mathbb{N}$ and pairwise disjoint $I_k \in \mathcal{I}_K$ with $\cup_{i=1}^n I_i = I \in \mathcal{I}_K$ we have $\mu_K(\cup_{i=1}^n I_i) = \sum_{i=1}^n \mu_K(I_i)$.*

Proof. For $I \in \mathcal{I}_K$ choose an index $k \in \{1, \ldots, d\} \setminus K$ and $c \in (a_k, b_k)$. The hyperplane $H = \{\mathbf{z} \le 0 : z_k = c\}$ separates the set $I$ into two disjoint sets $I_1, I_2 \in \mathcal{I}_K$. We have (using that the limit of a sum is the sum of the limits if those exists)

$$\begin{aligned}
&\mu_K(I_1) + \mu_K(I_2) \\
&= \lim_{t\to-\infty} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, y_k^{m_k} c^{1-m_k}, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\| \\
&\quad + \lim_{t\to-\infty} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, c^{m_k} x_k^{1-m_k}, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\| \\
&= \lim_{t\to-\infty} \left( \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K\cup\{k\}}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, y_k, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\| \right. \\
&\quad - \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K\cup\{k\}}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, c, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\| \\
&\quad + \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K\cup\{k\}}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, c, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\|
\end{aligned}$$

$$
\left. - \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K\cup\{k\}}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, x_k, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\| \right)
$$

$$
= \lim_{t\to-\infty} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_{(1)}^{m_1} x_{(1)}^{1-m_1}, \ldots, y_k^{m_k} x_k^{1-m_k}, \ldots, y_{(d)}^{m_d} x_{(d)}^{1-m_d} \right) \right\|
$$

$$
= \mu_K\left(I\right).
$$

Decomposing an $I \in \mathcal{I}_K$ by a finite number of hyperplanes, in the same way as above, results in that for pairwise disjoint $I_1, \ldots, I_n \in \mathcal{I}_K$ we get

$$
\mu_K\left(I\right) = \sum_{i=1}^{n} \mu_K\left(I_i\right).
$$

Finally we show that for pairwise disjoint $I_1, \ldots, I_n \in \mathcal{I}_K$ with $I = \cup_{i=1}^n I_i \in \mathcal{I}_K$ it is $\mu_K\left(I\right) = \sum_{i=1}^n \mu_K\left(I_i\right)$. Without loss of generality we can assume that every $I_i$ is nonempty. So there exists $-\infty < a_{i,j} < b_{i,j} \le 0$, $j \in \{1, \ldots, d\} \setminus K$, with $I_i = \times_{j=1}^d X_{i,j}$ where $X_{i,j} = \{-\infty\}$ for $j \in K$ and otherwise $X_{i,j} = [a_{i,j}, b_{i,j})$. If we split $I_0$ by hyperplanes of the form $\xi_i = a_{i,j}$ or $\xi_i = b_{i,j}$ the set $I_0$ decomposes in pairwise disjoint sets $I'_1, \ldots, I'_m$. Each of the $I_1, \ldots, I_n$ decomposes into certain $I'_k$, $k \in \{1, \ldots, m\}$. Applying $(n+1)$ times the case from above we obtain the equality.

□

Lemma 3.2.3. *For pairwise disjoint $I_1, \ldots, I_n \in \mathcal{I}_K$ with $\cup_{i=1}^n I_i \subseteq I \in \mathcal{I}_K$ it is*

$$
\sum_{i=1}^{n} \mu_K\left(I_i\right) \le \mu_K\left(I\right).
$$

Proof. There exists pairwise disjoint intervals $J_1, \ldots, J_m$ that are all also disjoint with $\cup_{i=1}^n I_i$ so that we have $I = \bigcup_{i=1}^n I_i \cup \bigcup_{j=1}^m J_j$. From the additivity of $\mu_k$ proven in Lemma 3.2.2 it follows that $\mu_K\left(I\right) = \sum_{i=1}^n \mu_K\left(I_i\right) + \sum_{j=1}^m \mu_K\left(J_j\right)$. Since $\mu_K \ge 0$ the proposition is proved.

□

Lemma 3.2.4. *If $\cup_{i=1}^{n} I_i = \cup_{j=1}^{m} J_j$, $I_i, J_j \in \mathcal{I}_K$ with $I_i$ pairwise disjoint then we have*

$$\sum_{i=1}^{n} \mu_K\left(I_i\right) \leq \sum_{j=1}^{m} \mu_K\left(J_j\right).$$

Proof. We get

$$\begin{aligned}\sum_{j=1}^{m} \mu_K\left(J_j\right) &= \sum_{j=1}^{m} \mu_K\left(\cup_{i=1}^{n}\left(J_j \cap I_i\right)\right)\\ &= \sum_{j=1}^{m}\sum_{i=1}^{n} \mu_K\left(\underbrace{J_j \cap I_i}_{\in \mathcal{I}_K}\right).\end{aligned}$$

The sets $I_i \cap J_j \in \mathcal{I}_K$, $i \leq m$, $j \leq d$ can only be constructed by disjoint $B_k \in \mathcal{I}_K$, $k = 1, \ldots, N$, i.e. $I_i \cap J_j = B_1^{(i,j)} \cup \cdots \cup B_{k(i,j)}^{(i,j)}$, $i \leq m$, $j \leq d$ and $B_r^{(i,j)} \in \{B_1, \ldots, B_N\}$. Then we can conclude from the additivity of $\mu_K$

$$\begin{aligned}\sum_{i \leq m}\sum_{j \leq n} \mu_K\left(I_i \cap J_j\right) &= \sum_{i \leq m}\sum_{j \leq n}\sum_{r \leq k(i,j)} \mu_K\left(B_r^{(i,j)}\right)\\ &\geq \mu_K\left(B_1\right) + \cdots + \mu_K\left(B_N\right),\end{aligned}$$

because every $B_k$ appears at least one time in the sum or otherwise it can be omitted. Since every $I_i$ can be represented as a union of certain $B_i$ and the $I_i$ are pairwise disjoint we must have $\mu_K\left(B_1\right) + \cdots + \mu_K\left(B_N\right) = \sum_{i \leq n} \mu_K\left(I_i\right)$ and the assertion is proved.

□

Lemma 3.2.5. *For $I \subseteq \cup_{i=1}^{n} I_i$ with $I, I_i \in \mathcal{I}_K$, $i \in \{1, \ldots, n\}$ we have*

$$\mu_K\left(I\right) \leq \sum_{i=1}^{n} \mu_K\left(I_i\right).$$

Proof. There exists pairwise disjoint $B_1, \ldots, B_N$ that

$$I_1 \cup \cdots \cup I_n = B_1 \cup \cdots \cup B_N.$$

Since $I = \cup_{j \leq n}\left(I \cap I_j\right) = \cup_{k \leq N}\left(I \cap B_k\right)$ it follows from Lemma 3.2.4 and the additivity of $\mu_K$ (Lemma 3.2.2)

$$\mu_K\left(I\right) = \mu_K\left(\cup_{k \leq N}\left(I \cap B_k\right)\right)$$

$$= \sum_{k \le N} \mu_K (I \cap B_k)$$
$$\le \sum_{j \le n} \mu_K (I \cap I_j)$$
$$\le \sum_{j \le n} \mu_K (I_j) .$$

The last inequality follows from the monotony of $\mu_K$ which is included in Lemma 3.2.3.

□

LEMMA 3.2.6. *For $I \subseteq \cup_{i \in \mathbb{N}} I_i$ with $I, I_i \in \mathcal{I}_K$, $i \in \mathbb{N}$ we have*

$$\mu_K (I) \le \sum_{i \in \mathbb{N}} \mu_K (I_i) .$$

PROOF. Without loss of generality let $K = \{1, \dots, k\}$.
Set $L = \{1, \dots, d\} \backslash K$ and let $I = \times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_i, y_i)$ and $I_n = \times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_{n,i}, y_{n,i})$.
For $\epsilon > 0$ it is

$$\times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_i, y_i - \epsilon] \subseteq \cup_{n \in \mathbb{N}} \times_{i \in K} \{-\infty\} \times \times_{i \in L} \left( x_{n,i} - \frac{\epsilon}{2^n}, y_{n,i} \right) .$$

We can conclude from the Theorem of Heine–Borel (see Theorem 15 in Cairns (1961, [**8**])) that there exists an $n_o \in \mathbb{N}$ with

$$\times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_i, y_i - \epsilon) \subseteq \times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_i, y_i - \epsilon]$$
$$\subseteq \cup_{n \le n_0} \times_{i \in K} \{-\infty\} \times \times_{i \in L} \left( x_{n,i} - \frac{\epsilon}{2^n}, y_{n,i} \right) .$$

From Lemma 3.2.5 we conclude

$$\mu_K \left( \times_{i \in K} \{-\infty\} \times \times_{i \in L} [x_i, y_i - \epsilon) \right) \le$$
$$\sum_{n \le n_0} \mu_K \left( \times_{i \in K} \{-\infty\} \times \times_{i \in L} \left( x_{n,i} - \frac{\epsilon}{2^n}, y_{n,i} \right) \right) ,$$

whereas by (3.4) and the continuity of a norm

$$\mu_K \left( \times_{i \in K} \{-\infty\} \times \times_{i \in L} \left( x_{n,i} - \tfrac{\epsilon}{2^n}, y_{n,i} \right) \right) =$$

$$\lim_{t\to-\infty} \sum_{\substack{m\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq m} m_j}$$
$$\left\| t,\dots,t, y_{n,k+1}^{m_{k+1}} \left( x_{n,k+1} - \frac{\epsilon}{2^n} \right)^{1-m_{k+1}}, \dots, y_{n,d}^{m_d} \left( x_{n,d} - \frac{\epsilon}{2^n} \right)^{1-m_d} \right\|.$$

From the triangle inequality of a norm we obtain $\big| \|\mathbf{x}\| - \|\mathbf{y}\| \big| \leq \|\mathbf{x}-\mathbf{y}\|$ and therefore

$$\begin{aligned}
&\left| \mu_K \left( \times_{i\in K} \{-\infty\} \times \times_{i\in L} \left( x_{n,i} - \frac{\epsilon}{2^n}, y_{n,i} \right) \right) - \mu_K (I_n) \right| \\
&\leq \lim_{t\to-\infty} \sum_{\substack{m\in\{0,1\}^d \\ m_i=1, i\in K}} \Bigg| \left\| t,\dots,t, y_{n,k+1}^{m_{k+1}} \left( x_{n,k+1} - \frac{\epsilon}{2^n} \right)^{1-m_{k+1}}, \dots, y_{n,d}^{m_d} \left( x_{n,d} - \frac{\epsilon}{2^d} \right)^{1-m_d} \right\| \\
&\quad - \left\| t,\dots,t, y_{n,k+1}^{m_{k+1}} x_{n,k+1}^{1-m_{k+1}}, \dots, y_{n,d}^{m_d} x_{n,d}^{1-m_d} \right\| \Bigg| \\
&\leq \sum_{\substack{m\in\{0,1\}^d \\ m_i=1, i\in K}} \left\| 0^{m_1} \left( -\frac{\epsilon}{2^n} \right)^{1-m_1}, \dots, 0^{m_d} \left( -\frac{\epsilon}{2^n} \right)^{1-m_d} \right\| \\
&\leq \text{const} \cdot \frac{\epsilon}{2^n}, \quad n \in \mathbb{N},
\end{aligned}$$

whereas const $> 0$ is independent from $\epsilon$ and $n$.
It follows

$$\begin{aligned}
&\mu_K \left( \times_{i\in K} \{-\infty\} \times \times_{i\in L} [x_i, y_i - \epsilon) \right) \\
&\quad \leq \sum_{n\leq n_0} \mu_K \left( \times_{i\in K} \{-\infty\} \times \times_{i\in L} \left( x_{n,i} - \frac{\epsilon}{2^n}, y_{n,i} \right) \right) \\
&\quad \leq \sum_{n\leq n_o} \mu_K (I_n) + \text{const} \cdot \epsilon \\
&\quad \leq \sum_{n\in\mathbb{N}} \mu_K (I_n) + \text{const} \cdot \epsilon.
\end{aligned}$$

Since $\epsilon$ can be arbitrary small the continuity of the norm implies

$$\mu_K (I) \leq \sum_{n\in\mathbb{N}} \mu_K (I_n).$$

□

Corollary 3.2.7. *$\mu_K$ is $\sigma$-additiv.*

Proof. This follows directly from Lemma 3.2.6 and Lemma 3.2.3.

□

Now we extend our definition of $\mu_K$ on
$\mathcal{F}_K := \{F : F = \cup_{i=1}^n I_i, I_i \in \mathcal{I}_K$ and pairwise disjoint$\}$ by

$$\bar{\mu}_K(F) := \sum_{j \leq n} \mu_K(I_j)$$

with $F = \cup_{i=1}^n I_i$. We will show that this is well-defined.
Suppose that $F \in \mathcal{F}_K$ can be represented as two different unions of pairwise disjoint elements from $\mathcal{I}_K$, i.e. $F = \cup_{i=1}^n I_i = \cup_{j=1}^m J_j$. Hence we have $I_i = I_i \cap F = \cup_{j=1}^m (I_i \cap J_j)$ and in the same way $J_j = F \cap J_j = \cup_{i=1}^n (I_i \cap J_j)$. The finite additivity of $\mu_K$ (see Lemma 3.2.2) implies

$$\mu_K(I_i) = \sum_{j=1}^m \mu_K(I_i \cap J_j)$$

and

$$\mu_K(I_j) = \sum_{i=1}^n \mu_K(I_i \cap J_j)$$

and therefore

$$\sum_{i=1}^n \mu_K(I_i) = \sum_{j=1}^m \mu_K(J_j).$$

Thus the definition is independent of the decomposition of $F$.

Corollary 3.2.8. *$\bar{\mu}_K : \mathcal{F}_K \to [0, \infty)$ is $\sigma$-additiv.*

Proof. Choose pairwise disjoint $F_i$ in $\mathcal{F}_K$, $i \in \mathbb{N}$ with $\cup_{i \in \mathbb{N}} F_i \in \mathcal{F}_K$, i.e. $\cup_{i \in \mathbb{N}} F_i = I_1^* \cup \cdots \cup I_m^*$ with $I_j^* \in \mathcal{I}_K$, $j \leq m$, pairwise disjoint and $F_i = I_{1,i} \cup \cdots \cup I_{m_i,i}$ is a union of disjoint intervals from $\mathcal{I}_K$, $i \in \mathbb{N}$.
From the definition of $\bar{\mu}_K$ follows

$$\bar{\mu}_K(\cup_{i \in \mathbb{N}} F_i) = \mu_K(I_1^* \cup \cdots \cup I_m^*) = \mu_K(I_1^*) + \cdots + \mu_K(I_m^*).$$

Furthermore from the $\sigma$-additivity of $\mu_K$ it follows

$$\begin{aligned}
\sum_{k\leq m} \mu_K\left(I_k^*\right) &= \sum_{k\leq m} \mu_K\left(\cup_{i\in\mathbb{N}}\left(F_i\cap I_k^*\right)\right)\\
&= \sum_{k\leq m} \mu_K\left(\cup_{i\in\mathbb{N}}\cup_{j\leq m_i}\left(\underbrace{I_{j,i}\cap I_k^*}_{\text{pairwise disjoint}}\right)\right)\\
&= \sum_{k\leq m}\sum_{i\in\mathbb{N}}\sum_{j\leq m_i} \mu_K\left(I_{j,i}\cap I_k^*\right)\\
&= \sum_{i\in\mathbb{N}}\sum_{j\leq m_i}\left(\underbrace{\sum_{k\leq m}\mu_K\left(I_{j,i}\cap I_k^*\right)}_{\mu(I_{j,i})}\right)\\
&= \sum_{i\in\mathbb{N}}\left(\underbrace{\sum_{j\leq m_i}\mu_K\left(I_{j,i}\right)}_{=\bar{\mu}_K(F_i)}\right)\\
&= \sum_{i\in\mathbb{N}}\bar{\mu}_K\left(F_i\right),
\end{aligned}$$

i.e. we obtain the assertion.

□

Corollary 3.2.9. *Since $\bar{\mu}_K$ is a ($\sigma$-finite) measure on the ring $\mathcal{F}_K$ there exists an unique measure $\mu_K^*$ on the $\sigma$-ring $\sigma\left(\mathcal{F}_K\right) = \{\sigma$-algebra that is generated by $\mathcal{F}_K\}$ , which coincide with $\bar{\mu}_K$ on $\mathcal{F}_K$.*

Proof. See Theorem A, §13 in Halmos (1973).

□

THEOREM 3.2.10. *With*

$$\mathcal{F} := \{\cup_{i=1}^{n} F_i : F_i \in \cup_{K \in \mathcal{K}} \mathcal{I}_K,\ F_i \textit{ are pairwise disjoint}\},$$

*and*

$$\bar{\mathbb{R}}_K := \times_{i=1}^{d} M_i,\ \textit{with } M_i = \{-\infty\} \textit{ for } i \in K \textit{ and } M_i = (-\infty, 0) \textit{ for } i \notin K$$

*there exists a measure $\mu$ on $\sigma(\mathcal{F})$ with*

$$\mu(I) := \sum_{K \in \mathcal{K}} \mu_K\left(I \cap \bar{\mathbb{R}}_K\right) \textit{ for } I \in \cup_{K \in \mathcal{K}} \mathcal{I}_K.$$

PROOF. Note that $I \cap \bar{\mathbb{R}}_K = \emptyset$ if $I \notin \mathcal{I}_K$.
The finite additivity of the $\mu_K$ implies the finite additivity of $\mu$. For $I_0, I_1, \ldots, I_n \in \mathcal{I}_0$ with $I_0 = \cup_{i=1}^{n} I_i$ we have

$$\begin{aligned}
\mu(I_0) &= \sum_{K \in \mathcal{K}} \mu_K\left(I_0 \cap \bar{\mathbb{R}}_K\right) \\
&= \sum_{K \in \mathcal{K}} \mu_K\left(\left(\cup_{i=1}^{n} I_i\right) \cap \bar{\mathbb{R}}_K\right) \\
&= \sum_{K \in \mathcal{K}} \mu_K\left(\cup_{i=1}^{n}\left(I_i \cap \bar{\mathbb{R}}_K\right)\right) \\
&= \sum_{K \in \mathcal{K}} \sum_{i=1}^{n} \mu_K\left(I_i \cap \bar{\mathbb{R}}_K\right) \\
&= \sum_{i=1}^{n} \sum_{K \in \mathcal{K}} \mu_K\left(I_i \cap \bar{\mathbb{R}}_K\right) \\
&= \sum_{i=1}^{n} \mu(I_i).
\end{aligned}$$

If $F \in \mathcal{F}$ can be represented as two different unions of pairwise disjoint elements from $\cup_{K \in \mathcal{K}} \mathcal{I}_K$, i.e. $F = \cup_{i=1}^{n} I_i = \cup_{j=1}^{m} J_j$, then we have $I_i = I_i \cap F = \cup_{j=1}^{m}(I_i \cap J_j)$ and in the same way $J_j = F \cap J_j = \cup_{i=1}^{n}(I_i \cap J_j)$. The finite additivity of $\mu$ implies

$$\mu(I_i) = \sum_{j=1}^{m} \mu(I_i \cap J_j)$$

and

$$\mu\left(I_j\right) = \sum_{i=1}^{n} \mu\left(I_i \cap J_j\right).$$

Therefore we get

$$\sum_{i=1}^{n} \mu\left(I_i\right) = \sum_{j=1}^{m} \mu\left(J_j\right).$$

So we can define $\bar{\mu}$ on $\mathcal{F}$ by

$$\bar{\mu}\left(\cup_{i=1}^{m} I_i\right) = \sum_{i=1}^{m} \mu\left(I_i\right). \tag{3.5}$$

From the considerations above we see that (3.5) is well defined.
Let $F_n$, $n \in \mathbb{N}$ be pairwise disjoint sets from $\mathcal{F}$ with $\cup_{n\in\mathbb{N}} F_n \in \mathcal{F}$, i.e. there exists pairwise disjoint $I_1, \ldots, I_m \in \cup_{K\in\mathcal{K}} \mathcal{I}_K$ with $\cup_{n\in\mathbb{N}} F_n = I_1 \cup \cdots \cup I_m$. This implies $\cup_{n\in\mathbb{N}} F_n = \cup_{j\leq m} \underbrace{\cup_{n\in\mathbb{N}} \left(F_n \cap I_j\right)}_{=I_j}$ and therefore we obtain

$$\begin{aligned}
\bar{\mu}\left(\cup_{n\in\mathbb{N}} F_n\right) &= \sum_{j\leq m} \mu\left(I_j\right) \\
&= \sum_{j\leq m} \mu\left(\underbrace{\cup_{n\in\mathbb{N}} \left(F_n \cap I_j\right)}_{=I_j}\right) \\
&= \sum_{j\leq m} \sum_{n\in\mathbb{N}} \mu\left(F_n \cap I_j\right) \\
&= \sum_{n\in\mathbb{N}} \sum_{j\leq m} \mu\left(F_n \cap I_j\right) \\
&= \sum_{n\in\mathbb{N}} \bar{\mu}\left(\underbrace{\cup_{j\leq m} \left(F_n \cap I_j\right)}_{=F_n}\right) \\
&= \sum_{n\in\mathbb{N}} \bar{\mu}\left(F_n\right).
\end{aligned}$$

The third equality sign follows from the fact that $\mu_K$ defines a measure on $\mathcal{I}_k$.
Therefore $\bar{\mu}$ can be extended to a measure $\mu^*$ on $\sigma\left(\mathcal{F}\right)$.

□

LEMMA 3.2.11. *If for every $K, L \subsetneq \{1, \ldots, d\}$ with $K \cap L = \emptyset$ and $K \cup L \neq \{1, \ldots, d\}$*

$$\lim_{s\to-\infty}\lim_{t\to-\infty}\left(\left\|\sum_{i\in K} t\mathbf{e}_i + \sum_{i\in L} s\mathbf{e}_i + \sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\| - \left\|\sum_{i\in K} t\mathbf{e}_i + \sum_{i\in L} s\mathbf{e}_i + \sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right)$$
$$= \lim_{t\to-\infty}\left(\left\|\sum_{i\in K\cup L} t\mathbf{e}_i + \sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\| - \left\|\sum_{i\in K\cup L} t\mathbf{e}_i + \sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right) \in \mathbb{R} \tag{3.6}$$

*then we have for* $[\mathbf{a}, \mathbf{b}) \subseteq [-\boldsymbol{\infty}, \mathbf{0}]^d \setminus \left\{(-\infty)^d\right\}$

$$\mu^*([\mathbf{a}, \mathbf{b})) = \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right)^T\right\|,$$

*with* $K = \{i : a_i = -\infty\}$.

PROOF. We set

$$M(\mathbf{a}, \mathbf{b}, L, K, i) := \begin{cases} \{-\infty\}, & i \in L \\ (-\infty, b_i), & i \in K\backslash L \\ [a_i, b_i), & i \notin K \end{cases}$$

and

$$\tilde{M}(\mathbf{a}, \mathbf{b}, L, K, i, s) := \begin{cases} \{-\infty\}, & i \in L \\ [s, b_i), & i \in K\backslash L \\ [a_i, b_i), & i \notin K \end{cases}.$$

With this notation we have the disjoint decomposition

$$[\mathbf{a}, \mathbf{b}) = \cup_{L\subseteq K}\left(\times_{i=1}^d M(\mathbf{a}, \mathbf{b}, L, K, i)\right)$$

and therefore

$$\begin{aligned}
&\mu^*([\mathbf{a}, \mathbf{b})) \\
&\quad = \sum_{L\subseteq K}\lim_{s\to-\infty}\mu^*\left(\times_{i=1}^d \tilde{M}(\mathbf{a}, \mathbf{b}, L, K, i, s)\right) \\
&\quad = \lim_{s\to-\infty}\sum_{L\subseteq K}\mu^*\left(\times_{i=1}^d \tilde{M}(\mathbf{a}, \mathbf{b}, L, K, i, s)\right)
\end{aligned}$$

$$= \lim_{s\to-\infty} \sum_{L\subseteq K} \lim_{t\to-\infty} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in K\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\|$$

$$= \lim_{s\to-\infty} \lim_{t\to-\infty} \sum_{L\subseteq K} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in K\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\|,$$

where $a_{(i)} = a_i$, $b_{(i)} = b_i$ for $i \notin K$; $a_{(i)}$ need not be defined for $i \in L$ and can be set to 0 for instance.

Choose an index $k \in K$. We split the sum into two parts, whether $k \in L$ or not. With the notation $\tilde{K} := K \setminus \{k\}$ we obtain

$$\sum_{L\subseteq K} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in K\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\|$$

$$= \sum_{L\subseteq \tilde{K}} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in K\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\|$$

$$+ \sum_{L\subseteq \tilde{K}} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k\} \\ b_{(i)}=t, i\in L\cup\{k\} \\ a_{(i)}=s, i\in \tilde{K}\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\|$$

$$= \sum_{L\subseteq \tilde{K}} \left( \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k\} \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in \tilde{K}\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left( \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, b_{(k)} = b_k, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \right.\right.$$
$$\left.\left. - \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \ldots, a_{(k)} = s, \ldots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \right) \right.$$

$$
\left. + \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k\} \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in\tilde{K}\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, b_{(k)} = t, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \right)
$$

$$
\begin{aligned}
&= \sum_{L\subseteq\tilde{K}} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k\} \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in\tilde{K}\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \\
&\quad + \sum_{L\subseteq\tilde{K}} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k\} \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in\tilde{K}\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \Bigg( \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, b_{(k)} = t, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \\
&\qquad\qquad - \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, a_{(k)} = s, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \Bigg) \\
&=: A + B
\end{aligned}
$$

Choose an index $r \in \{1, \dots, d\} \setminus K$ and decompose the sum $B$ according to $m_r$. Then we obtain

$$
\begin{aligned}
B = \sum_{L\subseteq\tilde{K}} \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in L\cup\{k,r\} \\ b_{(i)}=t, i\in L \\ a_{(i)}=s, i\in K\setminus L}} (-1)^{d+1-\sum_{j\le d} m_j} \Bigg( \\
&\left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, a_{(k)} = s, \dots, a_{(r)} = a_r, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \\
&- \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, a_{(k)} = s, \dots, b_{(r)} = b_r, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \\
&- \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, b_{(k)} = t, \dots, a_{(r)} = a_r, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \\
&+ \left\| b_{(1)}^{m_1} a_{(1)}^{1-m_1}, \dots, b_{(k)} = t, \dots, b_{(r)} = b_r, \dots, b_{(d)}^{m_d} a_{(d)}^{1-m_d} \right\| \Bigg) .
\end{aligned}
$$

From condition (3.6) it follows $\lim_{s\to-\infty} \lim_{t\to-\infty} B = 0$. In the sum $A$ we now have $m_k = 1$, $b_{(k)} = b_k$ and therefore we can iterate those steps above, i.e. chose $\tilde{k} \in \tilde{K}$ and decompose the set $L \subseteq \tilde{K}$ into two disjoint one-to-one subsets. This iteration can be repeated until we have $\tilde{K} = \emptyset$. Thus we have proven the assertion.

□

In the following example we will see that condition (3.6) from Lemma 3.2.11 is not satisfied by all norms.

Example 3.2.12. For $d \geq 3$ let

$$\|\mathbf{x}\| := \left( \mathbf{x}^T \underbrace{\begin{pmatrix} 1 & \delta & \dots & \delta \\ \delta & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \delta \\ \delta & \dots & \delta & 1 \end{pmatrix}}_{=:\mathbf{B}} \mathbf{x} \right)^{\frac{1}{2}},$$

with $\delta \in (0,1)$.

Since $\mathbf{B}$ is a symmetric positive definite matrix it is well known that we really defined a norm.
If we set $K = 1$ and $L = 2$ and choose $\mathbf{0} \leq \mathbf{x}, \mathbf{y}$ with $\sum_{i=3}^{d} (x_i - y_i) \neq 0$ then we obtain with $a^2 - b^2 = (a-b)(a+b)$:

$$\left\| t\mathbf{e}_1 + s\mathbf{e}_2 + \sum_{i=3}^{d} x_i \mathbf{e}_i \right\| - \left\| t\mathbf{e}_1 + s\mathbf{e}_2 + \sum_{i=3}^{d} y_i \mathbf{e}_i \right\| =: \frac{z(t,s)}{n(t,s)},$$

where

$$\begin{aligned} z(t,s) =& 2t \sum_{i=3}^{d} (x_i - y_i)\, \mathbf{e}_1^T \mathbf{B} \mathbf{e}_i + 2s \sum_{i=3}^{d} (x_i - y_i)\, \mathbf{e}_1^T \mathbf{B} \mathbf{e}_i \\ &+ \left( \sum_{i=3}^{d} x_i \mathbf{e}_i \right)^T \mathbf{B} \left( \sum_{i=3}^{d} x_i \mathbf{e}_i \right) - \left( \sum_{i=3}^{d} y_i \mathbf{e}_i \right)^T \mathbf{B} \left( \sum_{i=3}^{d} y_i \mathbf{e}_i \right) \end{aligned}$$

and

$$\begin{aligned} n(t,s) =& t \left( \left( \mathbf{e}_1 + \frac{s}{t}\mathbf{e}_2 + \frac{1}{t} \sum_{i=3}^{d} x_i \mathbf{e}_i \right)^T \mathbf{B} \left( \mathbf{e}_1 + \frac{s}{t}\mathbf{e}_2 + \frac{1}{t} \sum_{i=3}^{d} x_i \mathbf{e}_i \right) \right)^{\frac{1}{2}} \\ &+ t \left( \left( \mathbf{e}_1 + \frac{s}{t}\mathbf{e}_2 + \frac{1}{t} \sum_{i=3}^{d} y_i \mathbf{e}_i \right)^T \mathbf{B} \left( \mathbf{e}_1 + \frac{s}{t}\mathbf{e}_2 + \frac{1}{t} \sum_{i=3}^{d} y_i \mathbf{e}_i \right) \right)^{\frac{1}{2}}. \end{aligned}$$

Therefore we have for $s \neq t$

$$\lim_{t\to\infty} \frac{z(t,s)}{n(t,s)} = \frac{\sum_{i=3}^{d} (y_i - x_i)\, \mathbf{e}_1^T \mathbf{B} \mathbf{e}_i}{\left(\mathbf{e}_1^T \mathbf{B} \mathbf{e}_1\right)^{\frac{1}{2}}} = \delta \sum_{i=3}^{d} (y_i - x_i) \tag{3.7}$$

and $s = t$

$$\begin{aligned}\lim_{t\to\infty} \frac{z(t,t)}{n(t,t)} &= \frac{\sum_{i=3}^{d} (y_i - x_i)\, \mathbf{e}_1^T \mathbf{B} \mathbf{e}_i + \sum_{i=3}^{d} (y_i - x_i)\, \mathbf{e}_2^T \mathbf{B} \mathbf{e}_i}{\left( (\mathbf{e}_1 + \mathbf{e}_2)^T \mathbf{B} (\mathbf{e}_1 + \mathbf{e}_2) \right)^{\frac{1}{2}}} \\ &= \frac{\sqrt{2}\delta}{\sqrt{1+\delta}} \sum_{i=3}^{d} (y_i - x_i)\,. \end{aligned} \tag{3.8}$$

Since $\sum_{i=3}^{d} (x_i - y_i) \neq 0$ the limits (3.7) and (3.8) are different. As a consequence the condition (3.6) from Lemma 3.2.11 does not hold.

Lemma 3.2.13. *We have*

$$\mu^* \left( [-\boldsymbol{\infty}, \mathbf{x}]^{\complement} \right) = \|\mathbf{x}\|\,.$$

Proof. For $\mathbf{x} \leq \mathbf{0}$, $i \in \{1, \ldots, d\}$ and $y_i \leq x_i$ we can deduce from Lemma 3.2.11 and the continuity of a norm

$$\mu^* \left( \{\mathbf{z} : \mathbf{z} \leq \mathbf{x}, z_i > y_i\} \right) = \left\| (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_d)^T \right\| - \|\mathbf{x}\|\,.$$

Therefore we obtain

$$\begin{aligned}\mu^* \left( [-\boldsymbol{\infty}, \mathbf{x}]^{\complement} \right) &= \mu^* \left( \{\mathbf{z} : \mathbf{z} \leq \mathbf{0}, z_i > x_i \text{ for at least one } i \in \{1, \ldots, d\}\} \right) \\ &= \mu^* \left( \cup_{i=1}^{d} \{\mathbf{z} : \mathbf{z} \leq \mathbf{0}, z_i > x_i\} \right) \\ &= \mu^* \left( \dot\cup_{i=1}^{d} \{\mathbf{z} : \mathbf{z} \leq \mathbf{0}, z_i > x_i, z_j \leq x_j \text{ for } j \in \{1, \ldots, i-1\}\} \right) \\ &= \sum_{i=1}^{d} \mu \left( \{\mathbf{z} : \mathbf{z} \leq \mathbf{0}, z_i > x_i, z_j \leq x_j \text{ for } j \in \{1, \ldots, i-1\}\} \right) \\ &= \sum_{i=1}^{d} \left( \left\| (x_1, \ldots, x_i, 0 \ldots, 0)^T \right\| - \left\| (x_1, \ldots, x_{i-1}, 0, \ldots, 0)^T \right\| \right) \end{aligned}$$

$$= \sum_{i=1}^{d-1} \left( \left\| (x_1, \ldots, x_i, 0 \ldots, 0)^T \right\| - \left\| (x_1, \ldots, x_i, 0, \ldots, 0)^T \right\| \right) \\ - \|\mathbf{0}\| + \|\mathbf{x}\|$$

$$= \|\mathbf{x}\| .$$

$\square$

Finally we want to show that condition (3.6) in Lemma 3.2.11 already follows from condition (3.2) in Lemma 3.2.1.
In order to show this we will need the following lemma.

LEMMA 3.2.14. *Let $f : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}_0^+$ be an in both arguments monotone decreasing function, i.e for all $\epsilon > 0$ it is*

$$f(t, s) \geq f(t + \epsilon, s)$$

*and*

$$f(t, s) \geq f(t, s + \epsilon),$$

*then we have*

$$\lim_{s\to\infty} \lim_{t\to\infty} f(t, s) = \lim_{t\to\infty} f(t, t).$$

PROOF. Since $f$ is bounded from below by 0 and monotone decreasing in each argument, both limits exist. Furthermore we have for any $s \in \mathbb{R}^+$

$$\lim_{t\to\infty} f(t, s) =: c(s)$$
$\Leftrightarrow$ for all $\epsilon{>}0$ exists an $t_0(s){>}0$ that for all $t{\geq}t_0(s)$ it is $|f(t, s) - c(s)| < \epsilon$
$\Leftrightarrow$ for all $\epsilon{>}0$ exists an $t_0(s){>}0$ that for all $t{\geq}t_0(s)$ it is

$$c(s) - \epsilon < f(t, s) < c(s) + \epsilon.$$

In the same way we have

$$\lim_{t\to\infty} f(t, t) =: d$$
$\Leftrightarrow$ for all $\epsilon{>}0$ exists an $\tilde{t}_0{>}0$ that for all $t{\geq}\tilde{t}_0$ it is $|f(t, t) - d| < \epsilon$

$\Leftrightarrow$ for all $\epsilon > 0$ exists an $\tilde{t_0} > 0$ that for all $t \geq \tilde{t_0}$ it is

$$d - \epsilon < f(t,t) < d + \epsilon.$$

For any $s \in \mathbb{R}^+$ we therefore have for all $\epsilon > 0$ and for all $t \geq \max\left\{t_0(s), \tilde{t_0}, s\right\}$

$$c(s) - \epsilon < f(t,s) < c(s) + \epsilon$$

and

$$d - \epsilon < f(t,t) < d + \epsilon.$$

Using the fact that $f$ is decreasing in both arguments we have

$$\begin{aligned}
&c(s) + \epsilon > f(t,s) \geq f(t,t) > d - \epsilon \\
\Rightarrow &c(s) > d - 2\epsilon \\
\Rightarrow &c(s) \geq d
\end{aligned}$$

and

$$\begin{aligned}
&f(s,s) \geq f(t,s) > c(s) - \epsilon \\
\Rightarrow &f(s,s) > c(s) - \epsilon \\
\Rightarrow &f(s,s) \geq c(s),
\end{aligned}$$

since $\epsilon$ can be arbitrary small.
Using the inequalities from above we get

$$\lim_{s\to\infty} \lim_{t\to\infty} f(t,s) = \lim_{s\to\infty} c(s) \leq \lim_{s\to\infty} f(s,s) = \lim_{t\to\infty} f(t,t)$$

and

$$\lim_{s\to\infty} \lim_{t\to\infty} f(t,s) = \lim_{s\to\infty} c(s) \geq \lim_{s\to\infty} d = d = \lim_{t\to\infty} f(t,t).$$

Thus we obtain

$$\lim_{s\to\infty} \lim_{t\to\infty} f(t,s) = \lim_{t\to\infty} f(t,t).$$

$\square$

Lemma 3.2.15. *Condition (3.2) from Lemma 3.2.1 implies condition (3.6) from Lemma 3.2.11.*

Proof. First we prove the existence of the limit. We start by assuming that $\mathbf{0} \leq \mathbf{y} \leq \mathbf{x}$. Using condition (3.2) we have for any $k \in \{1, \ldots, d\}$, $k \neq l \in \{1, \ldots, d\}$ and $t_1 \geq t_0 > 0$

$$\begin{aligned}
&\left(\left\| t_0\mathbf{e}_k + x_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + y_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\|\right) \\
&- \left(\left\| t_1\mathbf{e}_k + x_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\| - \left\| t_1\mathbf{e}_k + y_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\|\right) \geq 0 \\
\Leftrightarrow &\left(\left\| t_0\mathbf{e}_k + x_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + y_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\|\right) \geq \\
&\left(\left\| t_1\mathbf{e}_k + x_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\| - \left\| t_1\mathbf{e}_k + y_l\mathbf{e}_l + \sum_{i\in\{1,\ldots,d\}\setminus\{k,l\}} x_i\mathbf{e}_i \right\|\right).
\end{aligned}$$

Using the above inequality we obtain furthermore

$$\begin{aligned}
&\left(\left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} y_i\mathbf{e}_i \right\|\right) \\
= &\left(\left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} y_i\mathbf{e}_i \right\|\right) \\
&+ \sum_{j=1}^{d-1} \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,j\}\setminus\{k\}} y_i\mathbf{e}_i + \sum_{i\in\{j+1,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| \\
&- \sum_{j=1}^{d-1} \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,j\}\setminus\{k\}} y_i\mathbf{e}_i + \sum_{i\in\{j+1,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| \\
= &\left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + \sum_{i\in\{1\}\setminus\{k\}} y_i\mathbf{e}_i + \sum_{i\in\{2,\ldots,d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| \\
&+ \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d-1\}\setminus\{k\}} y_i\mathbf{e}_i + \sum_{i\in\{d\}\setminus\{k\}} x_i\mathbf{e}_i \right\| - \left\| t_0\mathbf{e}_k + \sum_{i\in\{1,\ldots,d\}\setminus\{k\}} y_i\mathbf{e}_i \right\|
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{j=1}^{d-2}\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+1,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\\
&-\sum_{j=2}^{d-1}\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+1,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\\
=&\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|-\left\|t_0\mathbf{e}_k+\sum_{i\in\{1\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\\
&+\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,d-1\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|-\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}y_i\mathbf{e}_i\right\|\\
&+\sum_{j=1}^{d-2}\left(\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i+\sum_{i\in\{j+1\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\right.\\
&\left.-\left\|t_0\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i+\sum_{i\in\{j+1\}\setminus\{k\}}y_i\mathbf{e}_i\right\|\right)\\
\geq&\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|-\left\|t_1\mathbf{e}_k+\sum_{i\in\{1\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\\
&+\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,d-1\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|-\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}y_i\mathbf{e}_i\right\|\\
&+\sum_{j=1}^{d-2}\left(\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i+\sum_{i\in\{j+1\}\setminus\{k\}}x_i\mathbf{e}_i\right\|\right.\\
&\left.-\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,j\}\setminus\{k\}}y_i\mathbf{e}_i+\sum_{i\in\{j+2,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i+\sum_{i\in\{j+1\}\setminus\{k\}}y_i\mathbf{e}_i\right\|\right)\\
=&\left(\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}x_i\mathbf{e}_i\right\|-\left\|t_1\mathbf{e}_k+\sum_{i\in\{1,\dots,d\}\setminus\{k\}}y_i\mathbf{e}_i\right\|\right).
\end{aligned}
$$

By applying the above inequality to several choices of $k$ we obtain that for any arbitrary $K \subseteq \{1, \dots, d\}$

$$\begin{aligned}
&\left( \left\| \sum_{k \in K} t_0 \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t_0 \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right) \\
\geq &\left( \left\| \sum_{k \in K} t_1 \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t_1 \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right). \qquad (3.9)
\end{aligned}$$

The differences in (3.9) are descending in $t$. Applying condition (3.2) from the Main Theorem $d$ times using all subsets $K$ with $|K| = d-1$, yields that those differences are always nonnegative. Therefore the limit exists, i.e.

$$\lim_{t \to \infty} \left( \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right) \in \mathbb{R}_0^+. \qquad (3.10)$$

Now we drop the assumption that $\mathbf{y} \leq \mathbf{x}$. Then we obtain

$$\begin{aligned}
&\lim_{t \to \infty} \left( \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right) \\
&= \lim_{t \to \infty} \left( \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} \max\{x_i, y_i\} \mathbf{e}_i \right\| \right. \\
&\quad \left. + \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} \max\{x_i, y_i\} \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right) \\
&= \lim_{t \to \infty} \left( \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} x_i \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} \max\{x_i, y_i\} \mathbf{e}_i \right\| \right) \\
&\quad + \lim_{t \to \infty} \left( \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} \max\{x_i, y_i\} \mathbf{e}_i \right\| - \left\| \sum_{k \in K} t \mathbf{e}_k + \sum_{i \in \{1,\dots,d\} \setminus K} y_i \mathbf{e}_i \right\| \right),
\end{aligned}$$

where the last equal sign holds since both limits exist by (3.10). Hence the limit exists for arbitrary $\mathbf{x}, \mathbf{y} \geq \mathbf{0}$.

Next we prove that the limit does not change if we make the two step limit. We first assume again that $\mathbf{0} \leq \mathbf{y} \leq \mathbf{x}$. Above we have shown that the difference is decreasing in every component that is equal in both norms of the difference.

Therefore we can use Lemma 3.2.14 and obtain

$$\lim_{s\to\infty}\lim_{t\to\infty}\left(\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right)$$
$$=\lim_{t\to\infty}\left(\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|-\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right).$$

With a similar argumentation as above we obtain for arbitrary $\mathbf{x},\mathbf{y}\geq\mathbf{0}$

$$\lim_{s\to\infty}\lim_{t\to\infty}\left(\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right)$$
$$\begin{aligned}=\lim_{s\to\infty}\lim_{t\to\infty}\Bigg(&\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|\\
&-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|\\
&+\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|\\
&-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\Bigg)\\
=\lim_{s\to\infty}\lim_{t\to\infty}\Bigg(&\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|\\
&-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|\Bigg)\\
+\lim_{s\to\infty}\lim_{t\to\infty}\Bigg(&\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|\\
&-\left\|\sum_{i\in K} t\mathbf{e_i}+\sum_{i\in L} s\mathbf{e}_i+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\Bigg)\\
=\lim_{t\to\infty}&\left(\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} x_i\mathbf{e}_i\right\|-\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|\right)\\
+\lim_{t\to\infty}&\left(\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} \max\{x_i,y_i\}\,\mathbf{e}_i\right\|-\left\|\sum_{i\in K\cup L} t\mathbf{e_i}+\sum_{i\notin K\cup L} y_i\mathbf{e}_i\right\|\right)\end{aligned}$$

$$= \lim_{t\to\infty} \left( \left\| \sum_{i\in K\cup L} t\mathbf{e_i} + \sum_{i\notin K\cup L} x_i \mathbf{e}_i \right\| - \left\| \sum_{i\in K\cup L} t\mathbf{e_i} + \sum_{i\notin K\cup L} y_i \mathbf{e}_i \right\| \right),$$

where again the limit of the sum is the sum of the limits since these limits exist.

□

All the considerations above can be summarized in the following theorem.

THEOREM 3.2.16. *Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^d$. If*

$$\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left(b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d}\right) \right\| \ge 0$$

*holds for every $K \subsetneq \{1,\dots,d\}$ and $-\infty < a_j \le b_j \le 0,\ 1 \le j \le d$, then there exists a measure $\mu$ on $[-\infty,0]^d \setminus \{-\infty\}^d$ with*

$$\mu\left([-\infty,\mathbf{x}]^{\complement}\right) = \|\mathbf{x}\|.$$

## 3.3. Proof of the necessity

LEMMA 3.3.1. *Condition* (3.2) *is necessary to define an extreme value distribution with the norm $\|\cdot\|$.*

PROOF. Let $\nu$ be the exponent measure belonging to the norm $\|\cdot\|$, i.e.

$$\nu\left([-\infty,\mathbf{x}]^{\complement}\right) = \begin{cases} \|\mathbf{x}\|, & \text{for } \mathbf{x} \le \mathbf{0} \\ 0, & \text{otherwise.} \end{cases}$$

The continuity of a norm implies furthermore

$$\nu\left([-\infty,\mathbf{x})^{\complement}\right) = \begin{cases} \|\mathbf{x}\|, & \text{for } \mathbf{x} \le \mathbf{0} \\ 0, & \text{otherwise.} \end{cases} \tag{3.11}$$

For any $K \subsetneq \{1,\dots,d\}$ and $-\boldsymbol{\infty} < \mathbf{x} \le \mathbf{y} \le \mathbf{0}$ we define

$$A_K(\mathbf{x},\mathbf{y}) := \{\mathbf{z} : z_i < y_i, z_j \ge x_j, j \in \{1,\dots,d\} \setminus K\}.$$

For $l \notin K$ and $j \in \{1,\dots,d\}$ we have

$$\begin{aligned} A_{\{1,\dots,d\}}[\mathbf{x},\mathbf{y}) &= [-\boldsymbol{\infty},\mathbf{y}), \\ A_{\{1,\dots,d\}\setminus\{j\}}[\mathbf{x},\mathbf{y}) &= \{\mathbf{z} : z_i < y_i \text{ for } 1 \le i \le d, z_j \ge x_j\} \\ &= \{\mathbf{z} : z_i < y_i \text{ for } 1 \le i \le d\} \setminus \{\mathbf{z} : z_i < y_i \text{ for } 1 \le i \le d, z_j < x_j\} \end{aligned}$$

$$
\begin{aligned}
&= A_{\{1,\ldots,d\}}\left(\mathbf{x},\mathbf{y}\right)\backslash A_{\{1,\ldots,d\}}\left(\mathbf{x},\left(y_1,\ldots,y_{j-1},x_j,y_{j+1},\ldots,y_d\right)\right)\\
&= \left[-\boldsymbol{\infty},\mathbf{y}\right)\setminus\left[-\boldsymbol{\infty},\left(y_1,\ldots,y_{j-1},x_j,y_{j+1},\ldots,y_d\right)\right)\\
&= \left[-\boldsymbol{\infty},\left(y_1,\ldots,y_{j-1},x_j,y_{j+1},\ldots,y_d\right)\right)^{\complement}\setminus\left[-\boldsymbol{\infty},\mathbf{y}\right)^{\complement} \qquad (3.12)
\end{aligned}
$$

and, more general,

$$
\begin{aligned}
A_K\left[\mathbf{x},\mathbf{y}\right) &= \left\{\mathbf{z}: z_i<y_i \text{ for } 1\le i\le d, z_j\ge x_j \text{ for } j\notin K\right\}\\
&= \left\{\mathbf{z}: z_i<y_i \text{ for } 1\le i\le d, z_j\ge x_j \text{ for } j\notin K\backslash\left\{l\right\}\right\}\cap\\
&\qquad \left\{\mathbf{z}: z_i<y_i \text{ for } 1\le i\le d, z_l<x_l, z_j\ge x_j \text{ for } j\notin K\backslash\left\{l\right\}\right\}^{\complement}\\
&= A_{K\cup\{l\}}\left[\mathbf{x},\mathbf{y}\right)\backslash A_{K\cup\{l\}}\left[\mathbf{x},\left(y_1,\ldots,y_{l-1},x_l,y_{l+1},\ldots,y_d\right)\right). \qquad (3.13)
\end{aligned}
$$

We will show by induction over $k:=d-|K|$ that for any $K\subsetneq\{1,\ldots,d\}$ the $\sigma$-additivity implies that

$$
\nu\left(A_K\left(\mathbf{x},\mathbf{y}\right)\right)=\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\in K}}(-1)^{\left(d+1-\sum_{j\le d}m_j\right)}\left\|\left(y_1^{m_1}x_1^{1-m_1},\ldots,y_d^{m_d}x_d^{1-m_d}\right)\right\|. \qquad (3.14)
$$

Set $k=1$ and therefore take a subset $K=\{1,\ldots,d\}\setminus\{l\}$, $1\le l\le d$. Equations (3.12) and (3.11) imply

$$
\begin{aligned}
\nu\left(A_K\left(\mathbf{x},\mathbf{y}\right)\right) &= \nu\left(\left[-\boldsymbol{\infty},\left(y_1,\ldots,y_{l-1},x_l,y_{l+1},\ldots,y_d\right)\right)^{\complement}\setminus\left[-\boldsymbol{\infty},\mathbf{y}\right)^{\complement}\right)\\
&= \left\|\left(y_1,\ldots,y_{l-1},x_l,y_{l+1},\ldots,y_d\right)\right\|-\left\|\mathbf{y}\right\|\\
&= (-1)^{d+1-(d-1)}\left\|\left(y_1,\ldots,y_{l-1},x_l,y_{l+1},\ldots,y_d\right)\right\|+(-1)^{d+1-d}\left\|\mathbf{y}\right\|\\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\neq l}}(-1)^{\left(d+1-\sum_{j\le d}m_j\right)}\left\|\left(y_1^{m_1}x_1^{1-m_1},\ldots,y_d^{m_d}x_d^{1-m_d}\right)\right\|.
\end{aligned}
$$

Assume now that we have established the assertion for any $k$ with $k<d$. By using the induction assumption and equations (3.13) and (3.11) we obtain for any subset $K$ with $k=d-|K|$ and $l\notin K$

$$
\begin{aligned}
\nu\left(A_K\left(\mathbf{x},\mathbf{y}\right)\right) &= \mu\left(A_{K\cup\{l\}}\left(\mathbf{x},\mathbf{y}\right)\backslash A_{K\cup\{l\}}\left(\mathbf{x},\left(y_1,\ldots,y_{l-1},x_l,y_{l+1},\ldots,y_d\right)\right)\right)\\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\in K\cup\{l\}}}(-1)^{\left(d+1-\sum_{j\le d}m_j\right)}\left\|\left(y_1^{m_1}x_1^{1-m_1},\ldots,y_d^{m_d}x_d^{1-m_d}\right)\right\|
\end{aligned}
$$

$$+ \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_j=1, j\in K, m_l=0}} (-1)^{\left(d+1-\sum_{j\leq d} m_j\right)} \left\| \left(y_1^{m_1} x_1^{1-m_1}, \ldots, y_d^{m_d} x_d^{1-m_d}\right) \right\|$$

$$= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_j=1, j\in K}} (-1)^{\left(d+1-\sum_{j\leq d} m_j\right)} \left\| \left(y_1^{m_1} x_1^{1-m_1}, \ldots, y_d^{m_d} x_d^{1-m_d}\right) \right\|.$$

So we have established equation (3.14) and together with the nonnegativity of the measure $\nu$ we obtain condition (3.2).

□

CHAPTER 4

# Approach via convex geometry

In this chapter we state another access to the Main Theorem using results from convex geometry and the theory of random sets based on an article by Molchanov (2007, [**29**]). Note that Molchanov uses Fréchet margins and we adapted this to Weibull margins, since this fits better with the representation of an EVD using norms. At the end of the chapter we can give another proof of the Main Theorem based on the work of Molchanov.
For instance this approach is used in Molchanov (2007, [**28**]) and Molchanov and Schmutz (2008, [**30**]) to link economic statements to geometric properties and conversely.
We assume that the reader is familiar with basic knowledge of convex geometry and the theory of random sets. Some important definitions are given in the Appendix. For an introduction to convex geometry we refer to Schneider (1993, [**36**]) and for the theory of random sets the book by Molchanov (2005, [**27**]).
The notation in this chapter is mostly in accordance with the one used by Molchanov (2007, [**29**]).
By $\mathbf{E}$ we denote the selection expectation (also called Aumann expectation) (the definition is given in the Appendix in Definition A.4; for more information see Section 2.1 in Molchanov (2005, [**27**])) and for $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{R}^d$ we set

$$\Delta_{\mathbf{a}} := \operatorname{conv}\left(\{\mathbf{0}, a_1\mathbf{e}_1, \dots, a_d\mathbf{e}_d\}\right),$$

where $\operatorname{conv}(\cdot)$ denotes the convex hull of the corresponding set.
In convex geometry there is an important correspondence between a convex set and a certain function, the so called support function. We will deal with special convex sets and therefore we will need support functions later on.

Definition 4.1. *The support function $h$ of a set $M \subseteq \mathbb{R}^d$ is defined as*

$$h(M, \mathbf{x}) = \sup\left\{\langle \mathbf{z}, \mathbf{x}\rangle : \mathbf{z} \in M\right\}$$

*where $\langle \mathbf{z}, \mathbf{x}\rangle$ is the scalar product in $\mathbb{R}^d$.*

In the definition of the support function a convex set yields a function. The next lemmas shows that every sublinear function is a support function of a (unique) convex set. Since support functions are sublinear functions the correspondence between sublinear functions and convex sets are one-to-one.

LEMMA 4.2. *If $f : [0,\infty)^d \to \mathbb{R}$ is a sublinear function, i.e.*

$$f(\lambda x) = \lambda f(x) \text{ for all } \lambda \geq 0 \text{ and all } \mathbf{x} \in [0,\infty)^d$$

*and*

$$f(\mathbf{x}+\mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y}) \text{ for all } \mathbf{x},\mathbf{y} \in [0,\infty)^d,$$

*then there exists a unique convex body $\mathbb{K} \in \mathcal{K}^d$ with support function $f$, where $\mathcal{K}^d$ denotes the set of all non empty, compact and convex subset of $[0,\infty)^d$.*

PROOF. This is Theorem 1.7.1 in Schneider (1993, [**36**]) and we refer to the proofs (three different are given) stated there.

□

A max-zonoid which is defined next is the crucial concept of this approach.

DEFINITION 4.3. *The set $\mathbb{K} = c\mathbf{E}\Delta_\eta$ where $c > 0$ and $\eta$ is a random vector on $\mathbb{S}_+ := \left\{\mathbf{x} \in [0,\infty)^d : \|\mathbf{x}\| = 1\right\}$ (with respect to any chosen norm) is said to be a max-zonoid. If $\hat{\sigma}$ is the distribution of $\eta$, then $\sigma = c\hat{\sigma}$ is the spectral measure of $\mathbb{K}$. If $c\mathbf{E}\eta = (1,\ldots,1)$, then the max-zonoid of $\mathbb{K}$ is called the dependency set associated with the spectral measure $\sigma$.*

The next theorem connects the max-zonoids with EVDs with Weibull margins.

THEOREM 4.4. *A convex set $\mathbb{K}$ is a max-zonoid if and only if there exists a random vector $\xi$ with cumulative distribution function $F(\mathbf{x}) = \exp(-h(\mathbb{K},\mathbf{x}))$ for all $\mathbf{x} \in (-\infty,0]^d$.*

PROOF. This is a version of Proposition 1 in Molchanov (2007, [**29**]) using Weibull margins instead of Fréchet margins.

□

Next we introduce max-completely alternating functions. Later on we will see that the condition of being a max-completely alternating function is equivalent to the condition of the Main Theorem.

DEFINITION 4.5. *Consider a function* $f : [0,\infty)^d \to \mathbb{R}$. *For* $n \geq 1$ *and* $\mathbf{x}$, $\mathbf{x}_1, \ldots, \mathbf{x}_n \in [0,\infty)^d$ *we define the following successive differences*

$$\Delta_{\mathbf{x}_1} f(\mathbf{x}) := f(\mathbf{x}) - f(\max\{\mathbf{x}, \mathbf{x}_1\}),$$
$$\Delta_{\mathbf{x}_n} \ldots \Delta_{\mathbf{x}_1} f(\mathbf{x}) = \Delta_{\mathbf{x}_{n-1}} \ldots \Delta_{\mathbf{x}_1} f(\mathbf{x}) - \Delta_{\mathbf{x}_{n-1}} \ldots \Delta_{\mathbf{x}_1} f(\max\{\mathbf{x}, \mathbf{x}_n\}),$$

*where the maximum of the vectors is meant component wise as usual.*
*The function* $f$ *is said to be* <u>*max-completely alternating*</u> *if all successive differences are non positive.*

For more information on max-completely alternating functions, max-completely monotone functions (functions where all differences are nonnegative), completely alternating functions and completely monotone functions (these are generalizations of the max-completely alternating/monotone functions) we refer to Section 6 in Molchanov (2008, [**29**]), Section I.1.2 in Molchanov (2005, [**27**]) and Section 4.6 in Berg et. al. (1984, [**6**]).

We will see that we can restrict to a certain set if we want to check whether a function is max-completely alternating.

LEMMA 4.6. *Set* $G := \{\lambda \mathbf{e}_i : 1 \leq i \leq d, \lambda \in [0,\infty)\}$. *A function* $f : [0,\infty)^d \to \mathbb{R}$ *is max-completely alternating if and only if*

$$\Delta_{\mathbf{x}_n} \ldots \Delta_{\mathbf{x}_1} f(\mathbf{x}) \leq 0 \textit{ for } \mathbf{x} \in [0,1)^d \textit{ and } \mathbf{x}_1, \ldots, \mathbf{x}_n \in G, n \geq 1.$$

PROOF. $G$ is a generator set of the semigroup $\left([0,\infty)^d, \wedge\right)$, where the operation $\wedge$ is the component wise maximum. Therefore this is a special case of Proposition 6.6 in Berg et. al. (1984, [**6**]).

□

The next theorem is crucial to link this approach to our Main Theorem. It connects max-zonoids to max-completely alternating functions.

THEOREM 4.7. *A convex set* $\mathbb{K} \subseteq [0,\infty)^d$ *is a max-zonoid if and only if* $h(\mathbb{K}, \mathbf{x})$ *is a max-completely alternating function of* $\mathbf{x}$.

PROOF. This is Theorem 7 in Molchanov (2007, [**29**]) and we refer to this paper for the proof.

□

As mentioned above we will now prove that the condition of being max-completely alternating is equivalent to the condition given in the Main Theorem.

LEMMA 4.8. *An arbitrary norm* $\|\cdot\|$ *restricted to* $[0,\infty)^d$ *is a max-completely alternating function if and only if condition* (3.2) *from the Main Theorem 3.1.1 is fulfilled.*

PROOF. Without loss of generalization we can assume that in the condition for max-completely alternating functions we can restrict ourselves to the case where $\mathbf{x} \leq \mathbf{x}_i$, $i = \{1, \ldots, n\}$. Otherwise the vector $\mathbf{x}_i$ can be replaced by $\max(\mathbf{x}, \mathbf{x}_i)$. For $\mathbf{a} \leq \mathbf{b} \leq \mathbf{0}$ we set $\mathbf{c}_i := (b_1, \ldots, b_{i-1}a_i, b_{i+1}, \ldots, b_d)$, $1 \leq i \leq d$ and $\mathbf{c}_0 := \mathbf{b}$. Without loss of generalization we assume that $K = \{k, \ldots, d\}$, $2 \leq k \leq d$ (otherwise apply an permutation on the elements of $K$). We prove via induction over $k$ the equality

$$\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\| \left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right) \right\| = -\Delta_{-\mathbf{c}_{k-1}} \ldots \Delta_{-\mathbf{c}_1} \|-\mathbf{c}_0\|$$

and thus the terms in the conditions are the same and have opposed signs.
So let $K = \{2, \ldots, d\}$. Then we have (note that $\mathbf{a} \leq \mathbf{b} \leq 0$ and thus $0 \leq -\mathbf{b} \leq -\mathbf{a}$)

$$\begin{aligned}
&\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\| \left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right) \right\| \\
&= \sum_{m_1\in\{0,1\}} (-1)^{2-m_1} \left\| \left(b_1^{m_1} a_1^{1-m_1}, b_2, \ldots, b_d\right) \right\| \\
&= (-1)^1 \|\mathbf{c}_0\| + (-1)^2 \|\mathbf{c}_1\| \\
&= (-1)\left(\|-\mathbf{c}_0\| - \|-\mathbf{c}_1\|\right)
\end{aligned}$$

$$= -\Delta_{-\mathbf{c}_1} \|-\mathbf{c}_0\|$$

and therefore the base case of the induction is valid.
Now assume that we have already proven the assumption for $k < d$. Then we obtain for $k+1$

$$\begin{aligned}
&\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k+1}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right)\right\| \\
&= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k+1 \\ m_k=0}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right)\right\| \\
&\qquad + \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k+1 \\ m_k=1}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right)\right\| \\
&= -\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|-\left(b_1^{m_1} a_1^{1-m_1}, \ldots, a_k, \ldots, b_d^{m_d} a_d^{1-m_d}\right)\right\| \\
&\qquad + \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\geq k}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|-\left(b_1^{m_1} a_1^{1-m_1}, \ldots, b_d^{m_d} a_d^{1-m_d}\right)\right\| \\
&= -\left(\Delta_{-\mathbf{c}_{k-1}} \ldots \Delta_{-\mathbf{c}_1} \|-\mathbf{c}_0\| - \Delta_{-\mathbf{c}_{k-1}} \ldots \Delta_{-\mathbf{c}_1} \|-\mathbf{c}_k\|\right) \\
&= -\Delta_{-\mathbf{c}_k} \ldots \Delta_{-\mathbf{c}_1} \|-\mathbf{c}_0\| .
\end{aligned}$$

Thus the assertion is shown.
Furthermore we have

$$\max\{-\mathbf{c}_0, -\mathbf{c}_i\} = \max\{-\mathbf{c}_0, -a_i\mathbf{e}_i\}$$

and thus

$$\begin{aligned}
\Delta_{-\mathbf{c}_i} f(-\mathbf{c}_0) &= f(\mathbf{c}_0) - f(\max\{-\mathbf{c}_0, -\mathbf{c}_i\}) \\
&= f(\mathbf{c}_0) - f(\max\{-\mathbf{c}_0, -a_i\mathbf{e}_i\}) \\
&= \Delta_{-a_i\mathbf{e}_i} f(-\mathbf{x}_0) .
\end{aligned}$$

Iterating the step above we obtain

$$-\Delta_{-\mathbf{c}_k} \ldots \Delta_{-\mathbf{c}_1} \|-\mathbf{c}_0\| = -\Delta_{-a_k\mathbf{e}_k} \ldots \Delta_{-a_1\mathbf{e}_1} \|-\mathbf{c}_0\| .$$

Altogether we have

$$\begin{aligned}&\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_i=1, i\geq k}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1},\ldots,b_d^{m_d} a_d^{1-m_d}\right)\right\|\\ &= -\Delta_{-a_k\mathbf{e}_k}\ldots\Delta_{-a_i\mathbf{e}_1}\left\|-\mathbf{c}_0\right\|.\end{aligned}$$

Therefore if the norm $\|\cdot\|$ is max-completely alternating then the condition of the Main Theorem is fulfilled.

On the other hand if the condition from the Main Theorem is fulfilled and thus for the vectors from the set $G$ as defined in Lemma 4.6 the successive differences are non positive. Then Lemma 4.6 implies that the norm $\|\cdot\|$ is max-completely alternating.

$\square$

The previous results from this chapter leads to a new proof of the Main Theorem. Since the proof does not need the exponent measure we will leave it out in this formulation of the Main Theorem.

THEOREM 4.9. *For any norm $\|\cdot\|$ on $\mathbb{R}^d$ the following assertions are equivalent*

(i) *the function $G(\mathbf{x}) := \exp(-\|\mathbf{x}\|)$, $\mathbf{x}\leq\mathbf{0}$, defines a multivariate extreme value distribution function*

(ii) *the norm satisfies*

$$\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(b_1^{m_1} a_1^{1-m_1},\ldots,b_d^{m_d} a_d^{1-m_d}\right)\right\| \geq 0 \tag{4.1}$$

*for every $K\subsetneq\{1,\ldots,d\}$ and $-\infty<a_j\leq b_j\leq 0,\ 1\leq j\leq d$.*

PROOF. First assume that $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|)$, $\mathbf{x}\leq\mathbf{0}$, defines a distribution function having Weibull margins.

According to Lemma 4.2 there is a convex body $\mathbb{K}$ with support function $\|\cdot\|$ and Theorem 4.4 implies that $\mathbb{K}$ is a max-zonoid.

With Theorem 4.7 we conclude that $\|\cdot\|$ is a max-completely alternating function and because of Lemma 4.8 the condition in the Main Theorem is fulfilled.

Now let the condition from the Main Theorem be fulfilled and because of Lemma 4.8 the norm $\|\cdot\|$ is a max-completely alternating function.
According to Lemma 4.2 the norm $\|\cdot\|$ is support function of a convex body $\mathbb{K}$. Therefore Theorem 4.7 implies that $\mathbb{K}$ is a max-zonoid. Using Theorem 4.4 we obtain that $G(\mathbf{x}) = \exp(-\|\mathbf{x}\|)$, $\mathbf{x} \leq \mathbf{0}$, defines a distribution function.

$\square$

CHAPTER 5

# Applications

## 5.1. The bivariate case

The bivariate case turns out to be more simple than the case of dimension 3 or higher. In a way this is due to the fact that the Pickands dependence function is in the bivariate case a one-dimensional function. Falk (2006, [**12**]) proved that with a norm $\|\cdot\|$ a bivariate extreme value distribution can be defined if and only if the norm satisfies $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq \|\mathbf{x}\|_1$ for $\mathbf{x} \geq \mathbf{0}$. This section ends with an alternative proof for this fact using our Main Theorem 3.1.1. But first of all we start with two definitions of special properties a norm can have, because they will be useful in the following. Then we will show that these two properties are related to the corresponding one used by Falk (2006, [**12**]).

Definition 5.1.1. *A norm $\|\cdot\|$ on $\mathbb{R}^d$ is called monotone if for any vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ with $\mathbf{0} \leq \mathbf{a} \leq \mathbf{b}$ the norm of the vectors is ordered in the same way, i.e.*

$$\|\mathbf{a}\| \leq \|\mathbf{b}\|.$$

Definition 5.1.2. *A norm $\|\cdot\|$ on $\mathbb{R}^d$ is called standardized if every standard basis vector $\mathbf{e}_i$ has norm 1, i.e. $\|\mathbf{e}_i\| = 1$, $i = 1, \ldots, d$.*

Now we relate these two properties to the one used in Falk (2006, [**12**]).

Lemma 5.1.3. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. If $\|\cdot\|$ is monotone and standardized then we have for $\mathbf{0} \leq \mathbf{x} \in \mathbb{R}^d$*

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq \|\mathbf{x}\|_1.$$

*For $d = 2$ the converse statement is also true.*

Proof. Let $\mathbf{0} \leq \mathbf{x} = (x_1, \ldots, x_d)^T \in \mathbb{R}^d$. Since the norm is standardized we have

$$\left\|(x_1, \ldots, x_d)^T\right\| \leq \left\|(x_1, 0, \ldots, 0)^T\right\| + \cdots + \left\|(0, \ldots, 0, x_d)^T\right\|$$

$$\begin{aligned}
&= x_1 + \cdots + x_d \\
&= \left\| (x_1, \ldots, x_d)^T \right\|_1 .
\end{aligned}$$

Furthermore we obtain for all $i \in \{1, \ldots, d\}$

$$\begin{aligned}
\left\| (x_1, \ldots, x_d)^T \right\| &\geq \left\| (0, \ldots, 0, x_i, 0 \ldots, 0)^T \right\| \\
&= x_i \left\| \mathbf{e}_i \right\| \\
&= x_i.
\end{aligned}$$

Therefore it is $\left\| (x_1, \ldots, x_d)^T \right\| \geq \max \{x_1, \ldots, x_d\} = \left\| (x_1, \ldots, x_d)^T \right\|_\infty$. Altogether we have

$$\left\| \mathbf{x} \right\|_\infty \leq \left\| \mathbf{x} \right\| \leq \left\| \mathbf{x} \right\|_1 .$$

Now let $d = 2$ and the norm satisfies $\left\| \mathbf{x} \right\|_\infty \leq \left\| \mathbf{x} \right\| \leq \left\| \mathbf{x} \right\|_1$ for $\mathbf{0} \leq \mathbf{x}$. Then we have for the standard basis vectors

$$1 = \left\| \mathbf{e}_i \right\|_\infty \leq \left\| \mathbf{e}_i \right\| \leq \left\| \mathbf{e}_i \right\|_1 = 1$$

and thus the norm is standardized.
Take $\mathbf{a} = (a_1, a_2)^T \in \mathbb{R}^2$ and $\mathbf{b} = (b_1, b_2)^T \in \mathbb{R}^2$ with $\mathbf{0} \leq \mathbf{a} \leq \mathbf{b}$ and $\mathbf{0} < \mathbf{b}$. The condition $\left\| \mathbf{x} \right\|_\infty \leq \left\| \mathbf{x} \right\|$ implies that $b_i \leq \max \{b_1, b_2\} = \left\| \mathbf{b} \right\|_\infty \leq \left\| \mathbf{b} \right\|$ for $i = 1, 2$. From the triangle inequality we obtain

$$\begin{aligned}
\left\| (a_1, b_2)^T \right\| &= \left\| \frac{b_1 - a_1}{b_1} (0, b_2)^T + \frac{a_1}{b_1} (b_1, b_2)^T \right\| \\
&\leq \frac{b_1 - a_1}{b_1} \underbrace{\left\| (0, b_2)^T \right\|}_{= b_2 \leq \left\| \mathbf{b} \right\|} + \frac{a_1}{b_1} \left\| (b_1, b_2)^T \right\| \\
&\leq \left( \frac{b_1 - a_1}{b_1} + \frac{a_1}{b_1} \right) \left\| (b_1, b_2)^T \right\| \\
&= \left\| \mathbf{b} \right\|
\end{aligned}$$

and

$$\begin{aligned}
\left\| \mathbf{a} \right\| &= \left\| (a_1, a_2)^T \right\| \\
&= \left\| \frac{b_2 - a_2}{b_2} (a_1, 0)^T + \frac{a_2}{b_2} (a_1, b_2)^T \right\|
\end{aligned}$$

$$\leq \frac{b_2 - a_2}{b_2} \underbrace{\left\| (a_1, 0)^T \right\|}_{= a_1 \leq b_1 \leq \|\mathbf{b}\|} + \frac{a_2}{b_2} \underbrace{\left\| (a_1, b_2)^T \right\|}_{\leq \|\mathbf{b}\|, \text{see above}}$$

$$\leq \left( \frac{b_2 - a_2}{b_2} + \frac{a_2}{b_2} \right) \|\mathbf{b}\|$$

$$= \|\mathbf{b}\| .$$

Therefore the norm is monotone.

□

REMARK 5.1.4. In Example 2.19 in Hofmann (2006, [**22**]) a 3-dimensional norm is given that is between the maximum norm and the sum norm but which is not monotone. The unit sphere of this norm is pictured in Figure 5.1. Therefore the equivalence in Lemma 5.1.3 is only true in dimension 2.

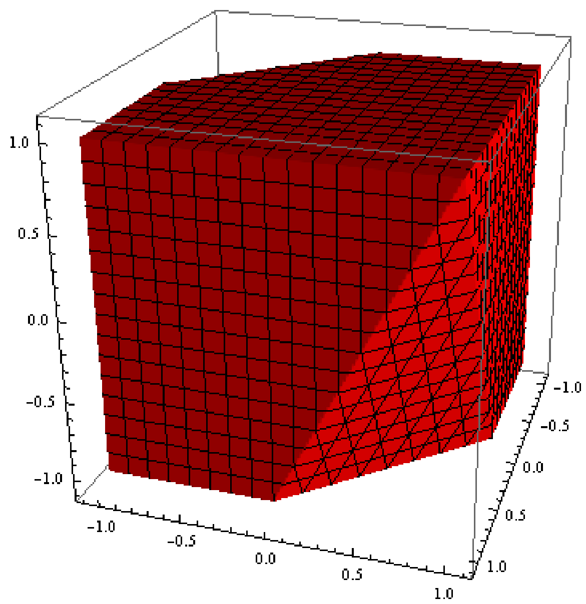



FIGURE 5.1. The unit sphere of the norm stated in Remark 5.1.4

In the next lemma we will see that the bivariate case turns out to be in particular simple. In this case it is necessary and sufficient for the norm to be monotone.

LEMMA 5.1.5. *Take an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^2$. Then*

$$G(\mathbf{x}) := \exp(-\|\mathbf{x}\|), \ \mathbf{x} \leq \mathbf{0}$$

*defines a bivariate distribution function if and only if the norm is monotone .*

PROOF. We have to check equation (3.2) for any subset $N \neq \emptyset$ of $\{1,2\}$. If $|N| = 1$ then (3.2) holds if and only if the norm is monotone.
For $N = \{1,2\}$ the equation holds for every norm on $\mathbb{R}^2$, because for $\mathbf{0} \leq \mathbf{a} \leq \mathbf{b}$, $\mathbf{a} \neq \mathbf{b}$, $\mathbf{0} < \mathbf{b}$ and $\alpha := \frac{b_1a_2 - a_1a_2}{b_1b_2 - a_1a_2}$, $\beta := \frac{b_1a_2 - a_1a_2}{b_1b_2 - a_1a_2}$, $\gamma := \frac{b_1b_2 - b_1a_2}{b_1b_2 - a_1a_2}$ and $\delta := \frac{b_1b_2 - a_1b_2}{b_1b_2 - a_1a_2}$.
We have

$$\begin{aligned}
\alpha, \beta, \gamma, \delta &\in \mathbb{R}_+, \\
\alpha + \gamma &= 1, \\
\beta + \delta &= 1, \\
\mathbf{a} &= \alpha (a_1, b_2)^T + \beta (b_1, a_2)^T \\
\mathbf{b} &= \gamma (a_1, b_2)^T + \delta (b_1, a_2)^T
\end{aligned}$$

and hence

$$\begin{aligned}
\|\mathbf{a}\| + \|\mathbf{b}\| &= \left\|\alpha (a_1, b_2)^T + \beta (b_1, a_2)^T\right\| + \left\|\gamma (a_1, b_2)^T + \delta (b_1, a_2)^T\right\| \\
&\leq \alpha \left\|(a_1, b_2)^T\right\| + \beta \left\|(b_1, a_2)^T\right\| + \gamma \left\|(a_1, b_2)^T\right\| + \delta \left\|(b_1, a_2)^T\right\| \\
&= \left\|(a_1, b_2)^T\right\| + \left\|(b_1, a_2)^T\right\|.
\end{aligned}$$

□

Usually we restrict ourselves to EVD with negative exponential margins. In this case a necessary and sufficient condition is that the norm is between the sum and the maximum norm as proved by Falk (2006, [**12**]). Using the results from above we can give an alternative proof.

Corollary 5.1.6. *Take an arbitrary norm* $\|\cdot\|$ *on* $\mathbb{R}^2$. *Then*

$$G(\mathbf{x}) := \exp\left(-\|\mathbf{x}\|\right), \ \mathbf{x} \le \mathbf{0}$$

*defines a bivariate distribution function with negative exponential margins if and only if the norm satisfies for every* $\mathbf{0} \le \mathbf{x} \in \mathbb{R}^d$

$$\|\mathbf{x}\|_\infty \le \|\mathbf{x}\| \le \|\mathbf{x}\|_1 . \tag{5.1}$$

Proof. According to Lemma 5.1.5 the norm is monotone if $G$ is a bivariate distribution function. Since all margins are negative exponential the norm must be standardized and therefore Lemma 5.1.3 implies the if-part.
On the other hand from condition (5.1) we get by Lemma 5.1.3 that the norm is monotone and standardized. From the monotonicity we conclude with Lemma 5.1.5 that $G$ is a distribution function. Since the norm is standardized we obtain with the homogeneity of the norm that $G$ has negative exponential margins.

$\square$

## 5.2. The Pickands dependence function

Now we investigate the Pickands dependence function. We start with a convex function and give a necessary and sufficient condition such that we can define a norm with that function. After that we can use our Main Theorem to establish a necessary and sufficient condition for a convex function to be a Pickands dependence function.

Lemma 5.2.1. *Let* $D : \overline{R}_{d-1} \to (0,\infty)$ *be a convex function, where*

$$\overline{R}_d := \left\{ (t_1,\dots,t_d) \in [0,1]^d : \sum_{j\le d} t_j \le 1 \right\}.$$

*Put for* $\mathbf{x} \in \mathbb{R}^d$, $\mathbf{x} \ne \mathbf{0}$

$$\|\mathbf{x}\|_D := \|\mathbf{x}\|_1 \, D\left( \frac{|x_1|}{\sum_{j\le d} |x_j|}, \dots, \frac{|x_{d-1}|}{\sum_{j\le d} |x_j|} \right) \tag{5.2}$$

*and* $\|\mathbf{0}\| = 0$, *where* $\|\mathbf{x}\|_1 = \sum_{j\le d} |x_j|$ *denotes the usual* $L_1$*-norm in* $\mathbb{R}^d$. *Then* $\|\mathbf{x}\|_D$ *defines a norm on* $\mathbb{R}^d$ *iff for* $\mathbf{0} \le \mathbf{x} \le \mathbf{y}$, $\mathbf{x} \ne \mathbf{0}$ *we have*

$$D\left(\frac{x_1}{\sum_{j\le d} x_j}, \dots, \frac{x_{d-1}}{\sum_{j\le d} x_j}\right) \le \frac{\sum_{j\le d} y_j}{\sum_{j\le d} x_j} D\left(\frac{y_1}{\sum_{j\le d} y_j}, \dots, \frac{y_{d-1}}{\sum_{j\le d} y_j}\right). \tag{5.3}$$

Proof. First assume that (5.3) holds. We have, obviously, $\|\lambda\mathbf{x}\|_D = |\lambda| \, \|\mathbf{x}\|_D$, $\lambda \in \mathbb{R}$, as well as $\|\mathbf{x}\|_D \ge 0$ and $\|\mathbf{x}\|_D = 0 \iff \mathbf{x} = \mathbf{0}$. The triangle inequality follows from the convexity of D, the triangle inequality of the absolute value and equation (5.3):

$$\begin{aligned}
&\|\mathbf{x}+\mathbf{y}\|_D \\
&= \|\mathbf{x}+\mathbf{y}\|_1 \, D\left(\frac{|x_1+y_1|}{\sum_{j\le d}(|x_j+y_j|)}, \dots, \frac{|x_{d-1}+y_{d-1}|}{\sum_{j\le d}(|x_j+y_j|)}\right) \\
&\le_{(5.3)} (\|\mathbf{x}\|_1 + \|\mathbf{y}\|_1) \, D\left(\frac{|x_1|+|y_1|}{\sum_{j\le d}(|x_j|+|y_j|)}, \dots, \frac{|x_{d-1}|+|y_{d-1}|}{\sum_{j\le d}(|x_j|+|y_j|)}\right) \\
&= (\|\mathbf{x}\|_1 + \|\mathbf{y}\|_1) \, D\left(\frac{\sum_{j\le d}|x_j|}{\sum_{j\le d}(|x_j|+|y_j|)} \left(\frac{|x_1|}{\sum_{j\le d}|x_j|}, \dots, \frac{|x_{d-1}|}{\sum_{j\le d}|x_j|}\right)\right. \\
&\qquad\qquad \left. + \frac{\sum_{j\le d}|y_j|}{\sum_{j\le d}(|x_j|+|y_j|)} \left(\frac{|y_1|}{\sum_{j\le d}|y_j|}, \dots, \frac{|y_{d-1}|}{\sum_{j\le d}|y_j|}\right)\right) \\
&\le (\|\mathbf{x}\|_1 + \|\mathbf{y}\|_1) \left(\frac{\sum_{j\le d}|x_j|}{\sum_{j\le d}(|x_j|+|y_j|)} D\left(\frac{|x_1|}{\sum_{j\le d}|x_j|}, \dots, \frac{|x_{d-1}|}{\sum_{j\le d}|x_j|}\right)\right. \\
&\qquad\qquad \left. + \frac{\sum_{j\le d}|y_j|}{\sum_{j\le d}(|x_j|+|y_j|)} \, D\left(\frac{|y_1|}{\sum_{j\le d}|y_j|}, \dots, \frac{|y_{d-1}|}{\sum_{j\le d}|y_j|}\right)\right) \\
&= \|\mathbf{x}\|_D + \|\mathbf{y}\|_D \, .
\end{aligned}$$

So we have established the if-part.

Now assume that (5.2) defines a norm on $\mathbb{R}^d$. It is sufficient to prove equation (5.3) for $\mathbf{0} \le \mathbf{x} \le \mathbf{y}$, where $\mathbf{x}$ and $\mathbf{y}$ differ only in the $k$-th component, i.e. $x_i = y_i$, $1 \le i \le d$ and $i \ne k$. By iterating this step with every component one gets the general equation.

Set $\tilde{\mathbf{y}} := \mathbf{y} - 2y_k\mathbf{e}_k$. The vector $\tilde{\mathbf{y}}$ differs from $\mathbf{y}$ only in the $k$-th component and the absolute value of this component is equal. Therefore $\|\mathbf{y}\|_D = \|\tilde{\mathbf{y}}\|_D$ and with the convexity of the $D$-norm we obtain

$$
\begin{aligned}
D\left(\frac{x_1}{\sum_{j\le d} x_j},\dots,\frac{x_{d-1}}{\sum_{j\le d} x_j}\right) &= \frac{1}{\|\mathbf{x}\|_1}\|\mathbf{x}\|_D \\
&= \frac{1}{\|\mathbf{x}\|_1}\left\|\frac{x_k+y_k}{2y_k}\mathbf{y}+\frac{y_k-x_k}{2y_k}\tilde{\mathbf{y}}\right\|_D \\
&\le \frac{1}{\|\mathbf{x}\|_1}\left(\frac{x_k+y_k}{2y_k}\|\mathbf{y}\|_D+\frac{y_k-x_k}{2y_k}\|\tilde{\mathbf{y}}\|_D\right) \\
&= \frac{1}{\|\mathbf{x}\|_1}\|\mathbf{y}\|_D \\
&= \frac{\sum_{j\le d} y_j}{\sum_{j\le d} x_j} D\left(\frac{y_1}{\sum_{j\le d} y_j},\dots,\frac{y_{d-1}}{\sum_{j\le d} y_j}\right).
\end{aligned}
$$

$\square$

Theorem 5.2.2. *Let $D$ be a positive and convex function on $\overline{R}_{d-1}$. Then $D$ is a Pickands dependence function, i.e.*

$$
G(\mathbf{x}) := \exp\left(\left(\sum_{j\le d} x_j\right) D\left(\frac{x_1}{\sum_{j\le d} x_j},\dots,\frac{x_{d-1}}{\sum_{j\le d} x_j}\right)\right), \quad \mathbf{x}\le \mathbf{0},
$$

*defines a $d$-dimensional EVD with standard exponential margins,*
*if and only if*

$$
\sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_j=1, j\in E}} \left[(-1)^{d+1-\sum_{j\le d} m_j}\left(\sum_{j\le d}\left(-y_j^{m_j}x_j^{1-m_j}\right)\right) D\left(\frac{y_1^{m_1}x_1^{1-m_1}}{\sum_{j\le d}\left(y_j^{m_j}x_j^{1-m_j}\right)},\dots,\frac{y_{d-1}^{m_{d-1}}x_{d-1}^{1-m_{d-1}}}{\sum_{j\le d}\left(y_j^{m_j}x_j^{1-m_j}\right)}\right)\right] \ge 0 \quad (5.4)
$$

*for any $\mathbf{x}\le\mathbf{y}\le\mathbf{0}$ and any subset $E\subset\{1,\dots,d\}$, $E\ne\{1,\dots,d\}$ and $D$ satisfies $D(\tilde{\mathbf{e}}_i)=D(\mathbf{0})=1$, $i\le d$, where $\tilde{\mathbf{e}}_i$ denotes the $i$-th unit vector in $\mathbb{R}^{d-1}$.*

Proof. Condition (5.4) implies condition (5.3) from Lemma 5.2.1 in the following way.
For any $m\in\{1,\dots,d\}$ we use condition (5.4) with $E=\{1,\dots,d\}\setminus\{m\}$ on the

vectors $\sum_{i=1}^{m} x_i \mathbf{e}_i + \sum_{i=m+1}^{d} y_i \mathbf{e}_i$ and $\sum_{i=1}^{m-1} x_i \mathbf{e}_i + \sum_{i=m}^{d} y_i \mathbf{e}_i$. Thus we obtain with $\alpha_m := \sum_{i=1}^{m-1} x_i + \sum_{i=m}^{d} y_i$

$$
\begin{aligned}
&- (-\alpha_m)\, D\left(\frac{x_1}{\alpha_m}, \ldots, \frac{x_{m-1}}{\alpha_m}, \frac{y_m}{\alpha_m}, \ldots, \frac{y_{d-1}}{\alpha_m}\right) \\
&\qquad\qquad + (-\alpha_{m+1})\, D\left(\frac{x_1}{\alpha_{m+1}}, \ldots, \frac{x_m}{\alpha_{m+1}}, \frac{y_{m+1}}{\alpha_{m+1}}, \ldots, \frac{y_{d-1}}{\alpha_{m+1}}\right) \geq 0.
\end{aligned}
$$

Furthermore set $\tau_m := D\left(\frac{x_1}{\alpha_m}, \ldots, \frac{x_{m-1}}{\alpha_m}, \frac{y_m}{\alpha_m}, \ldots, \frac{y_d}{\alpha_m}\right)$. Summation over $m$ from 1 to $d$ yields

$$
\begin{aligned}
0 &\leq \sum_{m=1}^{d} \left( \underbrace{\alpha_m \tau_m - \alpha_{m+1}\tau_{m+1}}_{\geq 0} \right) \\
&= \sum_{m=1}^{d} \alpha_m \tau_m - \sum_{m=1}^{d} \alpha_{m+1}\tau_{m+1} \\
&= \sum_{m=1}^{d} \alpha_m \tau_m - \sum_{m=2}^{d+1} \alpha_m \tau_m \\
&= \alpha_1 \tau_1 - \alpha_{d+1}\tau_{d+1} \\
&= \left(\sum_{j=1}^{d} y_j\right) D\left(\frac{y_1}{\sum_{j=1}^{d} y_j}, ..., \frac{y_{d-1}}{\sum_{j=1}^{d} y_j}\right) - \left(\sum_{j=1}^{d} x_j\right) D\left(\frac{x_1}{\sum_{j=1}^{d} x_j}, ..., \frac{x_{d-1}}{\sum_{j=1}^{d} x_j}\right) \\
\Leftrightarrow & D\left(\frac{x_1}{\sum_{j=1}^{d} x_j}, \ldots, \frac{x_{d-1}}{\sum_{j=1}^{d} x_j}\right) \leq \frac{\sum_{j=1}^{d} y_j}{\sum_{j=1}^{d} x_j} D\left(\frac{y_1}{\sum_{j=1}^{d} y_j}, \ldots, \frac{y_{d-1}}{\sum_{j=1}^{d} y_j}\right).
\end{aligned}
$$

Hence we can define with $D$ a norm as described in Lemma 5.2.1. Using the definition of the $D$-norm we get

$$
\begin{aligned}
0 \leq \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_j = 1, j\in E}} & \Bigg[ (-1)^{d+1-\sum_{j\leq d} m_j} \left( \sum_{j\leq d} \left( -y_j^{m_j} x_j^{1-m_j} \right) \right) \\
& \qquad D\left( \frac{y_1^{m_1} x_1^{1-m_1}}{\sum_{j\leq d} \left( y_j^{m_j} x_j^{1-m_j} \right)}, \ldots, \frac{y_{d-1}^{m_{d-1}} x_{d-1}^{1-m_{d-1}}}{\sum_{j\leq d} \left( y_j^{m_j} x_j^{1-m_j} \right)} \right) \Bigg]
\end{aligned}
$$

$$= \sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1, j\in E}} \left[ (-1)^{d+1-\sum_{j\le d} m_j} \left\| \left( y_1^{m_1} x_1^{1-m_1}, \dots, y_{d-1}^{m_{d-1}} x_{d-1}^{1-m_{d-1}} \right)^T \right\|_D \right].$$

By applying the Main Theorem 3.1.1 we get the assertion.

□

The considerations from this and the preceding section can be utilized to characterize a Pickands dependence function in the bivariate case.

THEOREM 5.2.3. *Consider an arbitrary function* $D : [0,1] \to (0,\infty)$ *and put* $\|(x,y)\|_D := (|x|+|y|)D(|x|/(|x|+|y|))$ *for* $x, y \in \mathbb{R}$ *with the convention* $\|\mathbf{0}\|_D = 0$*. Then the following statements are equivalent.*

(i) $\|\cdot\|_D$ *is a monotone and standardized norm.*
(ii) $\|\cdot\|_D$ *is a norm that satisfies* $\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_D \le \|\mathbf{x}\|_1$, $\mathbf{0} \le \mathbf{x}$.
(iii) $G(x,y) := \exp((x+y)D(x/(x+y)))$, $x, y \le 0$, *defines a bivariate EVD with standard reverse exponential margins.*
(iv) *The function* $D$ *is convex and satisfies* $\max(t, 1-t) \le D(t) \le 1$, $t \in [0,1]$.
(v) *The function* $D$ *is convex and satisfies* $\|\mathbf{x}\|_D \le \|\mathbf{y}\|_D$ *for* $\mathbf{0} \le \mathbf{x} \le \mathbf{y}$ *as well as* $D(0) = D(1) = 1$.

PROOF. The equivalence of (i), (ii) and (iii) is a consequence of Lemma 5.1.3 and Lemma 5.1.5. Next we show that (ii) and (iv) are equivalent. Suppose condition (ii) holds and choose $\lambda, t_1, t_2 \in [0,1]$. The triangle inequality implies

$$\begin{aligned} D(\lambda t_1 + (1-\lambda)t_2) &= \|\lambda(t_1, 1-t_1) + (1-\lambda)(t_2, 1-t_2)\|_D \\ &\le \lambda \|(t_1, 1-t_1)\|_D + (1-\lambda)\|(t_2, 1-t_2)\|_D \\ &= \lambda D(t_1) + (1-\lambda)D(t_2), \end{aligned}$$

i.e., $D$ is a convex function. Moreover we have for $t \in [0,1]$

$$\max(t, 1-t) = \|(t, 1-t)\|_\infty \le \|(t, 1-t)\|_D = D(t) \le \|(t, 1-t)\|_1 = 1,$$

which is (iv).
In what follows we show that (iv) implies (ii). The inequalities $\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_D \le \|\mathbf{x}\|_1$, $\mathbf{0} \le \mathbf{x} = (x,y)$, are obvious by putting $t = x/(x+y)$ in (iv). We also obtain $D(t) \ge 1/2$, $t \in [0,1]$, and, thus, $\|\mathbf{x}\|_D = 0$ if and only if $\mathbf{x} = \mathbf{0}$ as well as

$\|\lambda\mathbf{x}\|_D = |\lambda| \, \|\mathbf{x}\|_D$, $\lambda \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^d$. The triangular inequality will follow from the subsequent considerations. The inequality $\max(t, 1-t) \le D(t) \le 1$, $t \in [0,1]$, implies for $a, b \ge 0$, $a+b>0$,

$$D\left(\frac{a}{a+b}\right) \ge \frac{b}{a+b} = \frac{b}{a+b}D(0) = \frac{b}{a+b}D(1)$$

as well as

$$D\left(\frac{a}{a+b}\right) \ge \frac{a}{a+b} = \frac{a}{a+b}D(0) = \frac{a}{a+b}D(1).$$

Hence we obtain for $0 \le (x_1, x_2) \le (y_1, y_2)$ with $x_1 + x_2 > 0$, $y_i > 0$, $i = 1, 2$,

$$\begin{aligned}
D\left(\frac{x_1}{x_1+y_2}\right) &= D\left(\left(\frac{(y_1-x_1)y_2}{y_1(x_1+y_2)}\right)\cdot 0 + \left(\frac{(y_1+y_2)x_1}{y_1(x_1+y_2)}\right)\frac{y_1}{y_1+y_2}\right)\\
&\le \frac{(y_1-x_1)y_2}{y_1(x_1+y_2)}D(0) + \frac{(y_1+y_2)x_1}{y_1(x_1+y_2)}D\left(\frac{y_1}{y_1+y_2}\right)\\
&\le \frac{(y_1-x_1)(y_1+y_2)}{y_1(x_1+y_2)}D\left(\frac{y_1}{y_1+y_2}\right) + \frac{(y_1+y_2)x_1}{y_1(x_1+y_2)}D\left(\frac{y_1}{y_1+y_2}\right)\\
&= \frac{y_1+y_2}{x_1+y_2}D\left(\frac{y_1}{y_1+y_2}\right).
\end{aligned}$$

Summarizing the preceding inequalities we obtain

$$\begin{aligned}
D\left(\frac{x_1}{x_1+x_2}\right) &= D\left(\left(\frac{(y_2-x_2)x_1}{y_2(x_1+x_2)}\right)\cdot 1 + \left(\frac{(x_1+y_2)x_2}{y_2(x_1+x_2)}\right)\frac{x_1}{x_1+y_2}\right)\\
&\le \frac{y_2-x_2}{y_2(x_1+x_2)}x_1 D(1) + \frac{x_2}{y_2}\frac{x_1+y_2}{x_1+x_2}D\left(\frac{x_1}{x_1+y_2}\right)\\
&\le \frac{y_2-x_2}{y_2(x_1+x_2)}y_1 D(1) + \frac{x_2}{y_2}\frac{y_1+y_2}{x_1+x_2}D\left(\frac{y_1}{y_1+y_2}\right)\\
&\le \frac{y_2-x_2}{y_2}\frac{y_1+y_2}{x_1+x_2}D\left(\frac{y_1}{y_1+y_2}\right) + \frac{x_2}{y_2}\frac{y_1+y_2}{x_1+x_2}D\left(\frac{y_1}{y_1+y_2}\right)\\
&= \frac{y_1+y_2}{x_1+x_2}D\left(\frac{y_1}{y_1+y_2}\right).
\end{aligned}$$

The monotonicity $\|\mathbf{x}\|_D \le \|\mathbf{y}\|_D$, $\mathbf{0} \le \mathbf{x} \le \mathbf{y}$, established above together with the convexity of $D$ implies that $\|\cdot\|_D$ satisfies the triangular inequality for arbitrary $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$:

$$\begin{aligned}
&\|\mathbf{x}+\mathbf{y}\|_D\\
&= (|x_1+y_1| + |x_2+y_2|)\, D\left(\frac{|x_1+y_1|}{|x_1+y_1|+|x_2+y_2|}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \|(|x_1 + y_1|, |x_2 + y_2|)\|_D \\
&\le \|(|x_1| + |y_1|, |x_2| + |y_2|)\|_D \\
&= (|x_1| + |x_2| + |y_1| + |y_2|) \\
&\quad \times D\left(\frac{|x_1| + |x_2|}{|x_1| + |x_2| + |y_1| + |y_2|}\frac{|x_1|}{|x_1| + |x_2|} + \frac{|y_1| + |y_2|}{|x_1| + |x_2| + |y_1| + |y_2|}\frac{|y_1|}{|y_1| + |y_2|}\right) \\
&\le (|x_1| + |x_2|)\, D\left(\frac{|x_1|}{|x_1| + |x_2|}\right) + (|y_1| + |y_2|)\, D\left(\frac{|y_1|}{|y_1| + |y_2|}\right) \\
&= \|\mathbf{x}\|_D + \|\mathbf{y}\|_D .
\end{aligned}
$$

Next we show that (iv) and (v) are equivalent. Suppose condition (iv) is valid. Then, obviously, $D(0) = D(1) = 1$. The monotonicity of $\|\cdot\|_D$ was established in the proof of the implication (iv) $\implies$ (ii). Therefore it remains to show that (v) implies (iv). The convexity of $D$ implies

$$D(t) = D((1-t)\cdot 0 + t\cdot 1) \le (1-t)D(0) + tD(1) = t, \qquad t \in [0,1].$$

The monotonicity of $\|\cdot\|_D$ implies

$$(x_1 + x_2)D\left(\frac{x_1}{x_1 + x_2}\right) \le (y_1 + y_2)D\left(\frac{y_1}{y_1 + y_2}\right), \qquad \mathbf{0} \le \mathbf{x} \le \mathbf{y}.$$

Choosing $x_1 \in [0,1]$ and putting $x_2 = 0$, $y_1 = x_1$, $y_2 = 1 - x_1$, we obtain from the above inequality

$$x_1 D(1) = x_1 \le D(x_1).$$

Choosing $x_2 \in [0,1]$ and putting $x_1 = 0$, $y_1 = 1 - x_2$, $y_2 = x_2$, we obtain

$$x_2 D(0) = x_2 \le D\left(1 - \frac{y_2}{y_1 + y_2}\right) = D(1 - x_2),$$

i.e. we have established (iv).

□

### 5.3. Nested Logistic Model

The nested logistic model is a nonexchangeable model, i.e. for any permutation $(i_1, \ldots, i_d)$ of $(1, \ldots, d)$ the distributions of $(X_1, \ldots, X_d)$ and $(X_{i_1}, \ldots, X_{i_d})$ don't coincide. It was first described in Section 3 of Joe (1994, [**24**]) and is derived from the logistic model in a recursive way. In each recursion step a lower dimensional nested logistic model is nested in a bivariate logistic model. Joe gives only a sufficient condition for the logistic models to obtain a distribution function, namely that the parameter from the bivariate logistic model in the recursion step is not greater than any parameters used in the preceding recursion steps. In the next lemma we also give the proof that this condition is necessary by finding vectors for which the condition of the Main Theorem 3.1.1 is not fulfilled. Regarding dimension 3 we see that similar to the situation in Theorem 2.2.2 that the sum in the condition of the Main Theorem must get arbitrary small as one parameters tends to be equal to the other one since in the case of equality we are in the logistic model.

Lemma 5.3.1. *For $\lambda_1, \ldots, \lambda_{d-1} \geq 1$ we define recursive a norm in $\mathbb{R}^d$, $d \geq 3$, by*

$$\left\| (x_1, \ldots, x_d)^T \right\|_{\lambda_1, \ldots, \lambda_{d-1}} := \left\| \left( \left\| (x_1, \ldots, x_{d-1})^T \right\|_{\lambda_1, \ldots, \lambda_{d-2}}, |x_d| \right)^T \right\|_{\lambda_{d-1}},$$

*where $\|\cdot\|_\lambda$ is the usual $\lambda$-norm.*
*Then*

$$F(\mathbf{x}) := \exp\left( - \|\mathbf{x}\|_{\lambda_1, \ldots, \lambda_{d-1}} \right), \; \mathbf{x} \leq \mathbf{0}$$

*is a distribution function (and hence an extreme value distribution function) if and only if it is*

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{d-1} \geq 1.$$

Proof. The recursive definition of the norm as stated above yields really a norm. A prove is given in Lemma 2.22 in Hofmann (2006, [**22**]). More details concerning the nesting of norms are given in the remark following this proof.
In Section 5 of Joe (1994, [**24**]) the sufficiency of the condition is proved by showing that the density function is nonnegative. But the results in this paper are not expressed in terms of norms.

We proof now the necessity.
First let $d = 3$ and $\lambda_1 < \lambda_2$.

We consider the two points $\mathbf{a}_x = (0,0,1-x)^T$ and $\mathbf{b}_x = \left(x^{\frac{1}{\lambda_2}}, x^{\frac{1}{\lambda_2}}, 1\right)^T$, $x \in [0,1]$. Define for $x \in [0,1]$

$$\begin{aligned}
s(x) :=& -\left\|(0,0,1-x)^T\right\|_{\lambda_1,\lambda_2} \\
&+ \left\|\left(x^{\frac{1}{\lambda_2}}, 0, 1-x\right)^T\right\|_{\lambda_1,\lambda_2} + \left\|\left(0, x^{\frac{1}{\lambda_2}}, 1-x\right)^T\right\|_{\lambda_1,\lambda_2} + \left\|(0,0,1)^T\right\|_{\lambda_1,\lambda_2} \\
&- \left\|\left(0, x^{\frac{1}{\lambda_2}}, 1\right)^T\right\|_{\lambda_1,\lambda_2} - \left\|\left(x^{\frac{1}{\lambda_2}}, 0, 1\right)^T\right\|_{\lambda_1,\lambda_2} - \left\|\left(x^{\frac{1}{\lambda_2}}, x^{\frac{1}{\lambda_2}}, 1-x\right)^T\right\|_{\lambda_1,\lambda_2} \\
&+ \left\|\left(x^{\frac{1}{\lambda_2}}, x^{\frac{1}{\lambda_2}}, 1\right)^T\right\|_{\lambda_1,\lambda_2} \\
=& \left(1 + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{\frac{1}{\lambda_2}} + 2\left((1-x)^{\lambda_2} + x\right)^{\frac{1}{\lambda_2}} + 1 \\
& \qquad - 2(1+x)^{\frac{1}{\lambda_2}} - \left((1-x)^{\lambda_2} + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{\frac{1}{\lambda_2}} - (1-x).
\end{aligned}$$

The evaluation of the function in zero shows $s(0) = 0$. On the interval $(0,1)$ the function is infinitely often continuous differentiable and we obtain for $x \in (0,1)$

$$\begin{aligned}
s'(x) = 1 + \frac{1}{\lambda_2}\Bigg(& -2(1+x)^{\frac{1-\lambda_2}{\lambda_2}} + 2^{\frac{\lambda_2}{\lambda_1}}\left(1 + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{\frac{1-\lambda_2}{\lambda_2}} \\
& +2\left((1-x)^{\lambda_2} + x\right)^{\frac{1-\lambda_2}{\lambda_2}}\left(1 - (1-x)^{\lambda_2 - 1}\lambda_2\right) \\
& - \left((1-x)^{\lambda_2} + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{\frac{1-\lambda_2}{\lambda_2}}\left(2^{\frac{\lambda_2}{\lambda_1}} - (1-x)^{\lambda_2 - 1}\lambda_2\right)\Bigg)
\end{aligned}$$

and hence it is $\lim_{x \downarrow 0} s'(x) = 0$.
Furthermore we obtain for $x \in (0,1)$:

$$\begin{aligned}
s''(x) = & -(1-x)^{-2+\lambda_2}\left((1-x)^{\lambda_2} + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{-1+\frac{1}{\lambda_2}}(-1+\lambda_2) \\
& - \frac{2(1+x)^{-2+\frac{1}{\lambda_2}}\left(-1 + \frac{1}{\lambda_2}\right)}{\lambda_2} + \frac{2^{\frac{2\lambda_2}{\lambda_1}}\left(1 + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{-2+\frac{1}{\lambda_2}}\left(-1 + \frac{1}{\lambda_2}\right)}{\lambda_2} \\
& - \frac{1}{\lambda_2}\left((1-x)^{\lambda_2} + 2^{\frac{\lambda_2}{\lambda_1}} x\right)^{-2+\frac{1}{\lambda_2}}\left(-1 + \frac{1}{\lambda_2}\right)\left(2^{\frac{\lambda_2}{\lambda_1}} - (1-x)^{-1+\lambda_2}\lambda_2\right)^2 \\
& + 2\Bigg((1-x)^{-2+\lambda_2}\left((1-x)^{\lambda_2} + x\right)^{-1+\frac{1}{\lambda_2}}(-1+\lambda_2) +
\end{aligned}$$

$$\frac{1}{\lambda_2}\left((1-x)^{\lambda_2}+x\right)^{-2+\frac{1}{\lambda_2}}\left(-1+\frac{1}{\lambda_2}\right)\left(1-(1-x)^{-1+\lambda_2}\lambda_2\right)^2\bigg).$$

Consequently

$$\lim_{x\downarrow 0} s''(x) = 2\frac{\lambda_2-1}{\lambda_2}\left(2-2^{\frac{\lambda_2}{\lambda_1}}\right) =: c < 0,$$

i.e. for $\epsilon_3 := -\frac{c}{2}$ exists a $\delta_3 > 0$ that for all $x \in (0, \delta_3]$ it is $s''(x) \in \left[\frac{3c}{2}, \frac{c}{2}\right]$.
Because $s$ is a continuous function (as a composition of continuous functions), there exists for $\epsilon_1 := -c\frac{\delta_3^2}{256}$ a $\delta_1 > 0$ such that for all $x \in (0, \delta_1]$ it is $s(x) \in [-\epsilon_1, \epsilon_1]$.
For the first derivative we showed that $\lim_{x\downarrow 0} s'(x) = 0$, i.e. for $\epsilon_2 := -c\frac{\delta_3}{256}$ there exists a $\delta_2 > 0$ such that for all $x \in (0, \delta_2]$ it is $s'(x) \in [-\epsilon_2, \epsilon_2]$.
Now we set $\delta := \min\left(\delta_1, \delta_2, \frac{3\delta_3}{4}\right)$. We make a Taylor's expansion in the point $\delta_3$ around the point $\delta$ and obtain with $\xi \in [\delta, \delta_3]$

$$\begin{aligned}
s(\delta_3) &= s(\delta) + (\delta_3-\delta)\, s'(\delta) + \frac{1}{2}(\delta_3-\delta)^2 s''(\xi) \\
&\leq \epsilon_1 + \underbrace{(\delta_3-\delta)}_{\leq \delta_3}\epsilon_2 + \frac{1}{2}(\delta_3-\delta)^2\frac{c}{2} \\
&\leq -c\frac{\delta_3^2}{256} - \delta_3 c\frac{\delta_3}{256} + \frac{1}{4}\left(\frac{\delta_3}{4}\right)^2 c \\
&= -c\delta_3^2\left(\frac{1}{256}+\frac{1}{256}-\frac{1}{64}\right) \\
&= c\delta_3^2\frac{1}{128} \\
&< 0.
\end{aligned}$$

Hence there exists an $x \in [0,1]$ with $s(x) < 0$. But according to the Main Theorem 3.1.1 this is a necessary condition to the norm and thus for $d = 3$ and $\lambda_1 < \lambda_2$ the recursive defined norm does not yield a distribution function.

Now regard the case of arbitrary dimension $d > 3$.
Using the case from above we obtain that for the trivariate marginal distribution of the $i$, $j$, $k$ component, $1 \leq i < j < k \leq d-1$, it is necessary that $\lambda_i \geq \lambda_j \geq \lambda_k$.
Hence we obtain that it is necessary that $\lambda_1 \geq \cdots \geq \lambda_{d-2} \geq \lambda_{d-1} \geq 1$.

□

REMARK 5.3.2. Nesting of norms can be done in more ways than in the nested logistic model. For a monotone norm $\|\cdot\|_o$ on the $\mathbb{R}^k$ and for $i \in \{1, \dots, k\}$ let $\|\cdot\|_i$ be a norm on the $\mathbb{R}^{d_i}$ with $d_i \in \mathbb{N}\setminus\{0\}$ and $d = \sum_{i=1}^{d} d_i$. Then

$$\left\|(x_1, \dots, x_d)^T\right\| :=$$
$$\left\|\left(\left\|(x_1, ..., x_{d_1})^T\right\|_{d_1}, \left\|(x_{d_1+1}, ..., x_{d_1+d_2})^T\right\|_{d_2}, ..., \left\|(x_{d-d_k+1}, ..., x_d)^T\right\|_{d_k}\right)^T\right\|_o$$

is a norm on $\mathbb{R}^d$ (see Lemma 2.22 in Hofmann (2006, [**22**])).
The example of the nested logistic model shows that nesting norms that fulfill the condition of the Main Theorem 3.1.1 does not necessarily yield a norm that fulfills also this condition too.
The $\lambda$-norms are natural candidates for nesting due to their simplicity and publicity and therefore we restrict ourselves to them. In dimension 3 there is no other possibility than the usual nested logistic model. In the 4th dimension however there is a further possibility, namely

$$\left\|(x_1, x_2, x_3, x_4)^T\right\| := \left\|\left(\left\|(x_1, x_2)^T\right\|_{\lambda_1}, \left\|(x_3, x_4)^T\right\|_{\lambda_2}\right)^T\right\|_{\lambda_3},$$

with $\lambda_1, \lambda_2, \lambda_3 \geq 1$.
Looking at the 3 dimensional margins we obtain the necessary condition $\lambda_1 \geq \lambda_3$ and $\lambda_2 \geq \lambda_3$. By regarding the possible density of the possible EVD we see that $\lambda_1 \geq \lambda_3$ and $\lambda_2 \geq \lambda_3$ is also a sufficient condition.

We close this section with the calculation of the density of the GPD of nested logistic type in dimension 3, because this result will be necessary in Chapter 7. Since there is no general formula for the density in arbitrary dimension we restrict ourselves to dimension 3.

LEMMA 5.3.3. *The GPD of the nested logistic model in dimension* 3 *with parameters* $\lambda_1 \geq \lambda_2 \geq 1$ *has density*

$$w_{\lambda_1,\lambda_2}(x_1, x_2, x_3) =$$
$$(\lambda_2 - 1)(-x_1)^{\lambda_1-1}(-x_2)^{\lambda_1-1}(-x_3)^{\lambda_2-1}$$
$$\left((-x_1)^{\lambda_1} + (-x_2)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}-2}\left(\left((-x_1)^{\lambda_1} + (-x_2)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}} + (-x_3)^{\lambda_2}\right)^{\frac{1}{\lambda_2}-3}$$

$$\left((-x_3)^{\lambda_2}(\lambda_1-\lambda_2)+\left((-x_1)^{\lambda_1}+(-x_2)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}(\lambda_1+\lambda_2-1)\right)$$

PROOF. We have

$$w_{\lambda_1,\lambda_2}(x_1,x_2,x_3)=\frac{\partial^3}{\partial x_1\partial x_2\partial x_3}\left(1-\|x_1,x_2,x_3\|_{\lambda_1,\lambda_2}\right).$$

Straightforward computation yields the result.

□

## 5.4. The A-Norm

In this section we will obtain a $D$-Norm from another $D$-Norm by multiplying the vector with a certain Matrix $A$ and taking the norm of that vector. The Main Theorem 3.1.1 yields an sufficient condition for $A$.

DEFINITION AND THEOREM 5.4.1. *Let $\|\cdot\|$ be a standardized norm on $\mathbb{R}^d$ that fulfills the condition of the Main Theorem* (3.1.1) *and let $\mathbf{A}=(a_{i,j})\in\mathbb{R}^{d\times d}$ a regular matrix with only nonnegative entries.*
*With*

$$\tilde{\mathbf{A}}:=(\tilde{a}_{i,j}):=\begin{pmatrix}\frac{a_{1,1}}{\|\mathbf{Ae}_1\|}, & \dots, & \frac{a_{1,d}}{\|\mathbf{Ae}_d\|}\\ \vdots & \ddots & \vdots\\ \frac{a_{d,1}}{\|\mathbf{Ae}_1\|}, & \dots, & \frac{a_{d,d}}{\|\mathbf{Ae}_d\|}\end{pmatrix}$$

*we define a new norm*

$$\|\mathbf{x}\|_{\mathbf{A}}:=\left\|\tilde{\mathbf{A}}\mathbf{x}\right\|$$

*and for this norm*

$$G(\mathbf{x}):=\exp\left(-\|\mathbf{x}\|_{\mathbf{A}}\right),\quad \mathbf{x}\le\mathbf{0}$$

*defines a multivariate extreme value distribution function with negative exponential margins.*

PROOF. It is easy to check that $\|\cdot\|_{\mathbf{A}}$ defines a norm . Of course $\|\cdot\|_{\mathbf{A}}$ is nonnegative and since $\mathbf{A}$ is regular $\tilde{\mathbf{A}}$ is also regular (since $\det\tilde{\mathbf{A}}=\prod_{i=1}^d\frac{1}{\|\mathbf{Ae}_i\|}\det\mathbf{A}\neq 0$, see Harville (1997, [**21**]) Lemma 13.2.2) and therefore $\|\mathbf{x}\|_{\mathbf{A}}=0\Leftrightarrow\mathbf{x}=\mathbf{0}$.

Furthermore we have for $\lambda \in \mathbb{R}$

$$\begin{aligned}\|\lambda \mathbf{x}\|_{\mathbf{A}} &= \left\|\tilde{\mathbf{A}}\left(\lambda \mathbf{x}\right)\right\| \\ &= \left\|\lambda \tilde{\mathbf{A}}\mathbf{x}\right\| \\ &= |\lambda| \left\|\tilde{\mathbf{A}}\mathbf{x}\right\| \\ &= |\lambda| \|\mathbf{x}\|_{\mathbf{A}}\end{aligned}$$

and therefore we obtain the homogeneity.
Finally we show the triangle inequality, i.e. for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ we have

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|_{\mathbf{A}} &= \left\|\tilde{\mathbf{A}}\left(\mathbf{x}+\mathbf{y}\right)\right\| \\ &= \left\|\tilde{\mathbf{A}}\mathbf{x}+\tilde{\mathbf{A}}\mathbf{y}\right\| \\ &\leq \left\|\tilde{\mathbf{A}}\mathbf{x}\right\| + \left\|\tilde{\mathbf{A}}\mathbf{y}\right\| \\ &= \|\mathbf{x}\|_{\mathbf{A}} + \|\mathbf{y}\|_{\mathbf{A}}.\end{aligned}$$

Thus $\|\cdot\|_{\mathbf{A}}$ really defines a norm.
For $\mathbf{0} \leq \mathbf{x} \leq \mathbf{y} \in \mathbb{R}^d$ we have

$$\tilde{\mathbf{x}} := \tilde{\mathbf{A}}\mathbf{x} = \left(\sum_{i=1}^{d}\left(\sum_{j=1}^{d}\tilde{a}_{i,j}x_j\right)\mathbf{e_i}\right) \leq \left(\sum_{i=1}^{d}\left(\sum_{j=1}^{d}\tilde{a}_{i,j}y_j\right)\mathbf{e_i}\right) = \tilde{\mathbf{A}}\mathbf{y} =: \tilde{\mathbf{y}}.$$

From the definition of $\|\cdot\|_{\mathbf{A}}$ we see that $\|\mathbf{x}\|_{\mathbf{A}} = \|\tilde{\mathbf{x}}\|$ and $\|\mathbf{y}\|_{\mathbf{A}} = \|\tilde{\mathbf{y}}\|$. Since $\|\cdot\|$ fulfills condition (3.2) from the Main Theorem 3.1.1 we have for every $K \subsetneq \{1, \ldots, d\}$

$$\begin{aligned}0 &\leq \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(\tilde{y}_1^{m_1}\tilde{x}_1^{1-m_1}, \ldots, \tilde{y}_d^{m_d}\tilde{x}_d^{1-m_d}\right)\right\| \\ &= \sum_{\substack{\mathbf{m}\in\{0,1\}^d \\ m_i=1, i\in K}} (-1)^{d+1-\sum_{j\leq d} m_j} \left\|\left(y_1^{m_1}x_1^{1-m_1}, \ldots, y_d^{m_d}x_d^{1-m_d}\right)\right\|_{\mathbf{A}}.\end{aligned}$$

We can see that the norm $\|\cdot\|_{\mathbf{A}}$ fulfills condition (3.2) from the Main Theorem 3.1.1 and therefore $G(\mathbf{x}) := \exp\left(-\|\mathbf{x}\|_{\mathbf{A}}\right), \mathbf{x} \leq \mathbf{0}$ defines a distribution function. The matrix $\tilde{\mathbf{A}}$ is constructed in a way such that the norm is standardized, because we have

$$\|\mathbf{e}_i\|_{\mathbf{A}} = \left\|\tilde{\mathbf{A}}\mathbf{e}_i\right\|$$

$$\begin{aligned}
&= \left\| \sum_{j=1}^{d} \tilde{a}_{j,i} \mathbf{e}_j \right\| \\
&= \left\| \frac{1}{\|\mathbf{A}\mathbf{e}_i\|} \sum_{j=1}^{d} a_{j,i} \mathbf{e}_j \right\| \\
&= \frac{1}{\|\mathbf{A}\mathbf{e}_i\|} \|\mathbf{A}\mathbf{e}_i\| \\
&= 1
\end{aligned}$$

Therefore the margins are negative exponential.

$\square$

REMARK 5.4.2. If not all entries of the matrix $A$ are nonnegative no conclusion can be drawn. For instance, set $d = 2$, $\|\mathbf{x}\| = \|\mathbf{x}\|_5 = \left(|x_1|^5 + |x_2|^5\right)^{\frac{1}{5}}$. Then for

$$\mathbf{A} := \begin{pmatrix} \frac{1}{\sqrt{2}}, & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{2}} \end{pmatrix}$$

we can define an extreme value distribution with the norm $\|\cdot\|_{\mathbf{A}}$ and for

$$\mathbf{B} := \begin{pmatrix} \frac{\sqrt{3}}{2}, & -\frac{1}{2} \\ \frac{1}{2}, & \frac{\sqrt{3}}{2} \end{pmatrix}$$

we can <u>not</u> define an extreme value distribution with the norm $\|\cdot\|_{\mathbf{B}}$.
Without loss of generalization we assume that $x_1 \geq x_2 \geq 0$. Then we have

$$\tilde{\mathbf{A}} = \begin{pmatrix} \frac{1}{\sqrt[5]{2}}, & -\frac{1}{\sqrt[5]{2}} \\ \frac{1}{\sqrt[5]{2}}, & \frac{1}{\sqrt[5]{2}} \end{pmatrix}$$

and thus for $\mathbf{x} := (x_1, x_2)^T$

$$\begin{aligned}
\|\mathbf{x}\|_{\mathbf{A}} &= \left\| \tilde{\mathbf{A}}\mathbf{x} \right\| \\
&= \sqrt[5]{\frac{1}{2}\left((x_1 - x_2)^5 + (x_1 + x_2)^5\right)} \\
&= \sqrt[5]{\frac{1}{2}\left(2x_1^5 + 20x_1^3x_2^3 + 10x_1x_2^4\right)} \\
&\geq \sqrt[5]{x_1^5}
\end{aligned}$$

$$
\begin{aligned}
&= x_1 \\
&= \max(x_1, x_2) \\
&= \|\mathbf{x}\|_\infty .
\end{aligned}
$$

Since the norm $\|\cdot\|_{\mathbf{A}}$ is constructed in a way such that it is standardized, the norm $\|\cdot\|_{\mathbf{A}}$ is embedded between the sum and the maximum norm. By Corollary 5.1.6 we can define an extreme value distribution function with this norm.
We have

$$
\tilde{\mathbf{B}} = \begin{pmatrix} \frac{\sqrt{3}}{\left(1+9\sqrt{3}\right)^{1/5}} & -\frac{1}{\left(-1+9\sqrt{3}\right)^{1/5}} \\ \frac{1}{\left(1+9\sqrt{3}\right)^{1/5}} & \frac{\sqrt{3}}{\left(-1+9\sqrt{3}\right)^{1/5}} \end{pmatrix}
$$

and thus

$$
\begin{aligned}
\|\mathbf{e}_1\|_{\mathbf{B}} &= 1 \\
&> \left(\left(\frac{\sqrt{3}}{5(-1+9\sqrt{3})^{1/5}} + \frac{1}{1+9\sqrt{3}^{1/5}}\right)^5 + \left(\frac{\sqrt{3}}{(1+9\sqrt{3})^{1/5}} - \frac{1}{5(-1+9\sqrt{3})^{1/5}}\right)^5\right)^{1/5} \\
&= \|\mathbf{e}_1 + 0.2\mathbf{e}_2\|_{\mathbf{B}} \\
&\approx 0.95.
\end{aligned}
$$

Therefore $\|\cdot\|_{\mathbf{B}}$ is not monotone and by Lemma 5.1.5 we can not define a distribution function with the norm $\|\cdot\|_{\mathbf{B}}$.

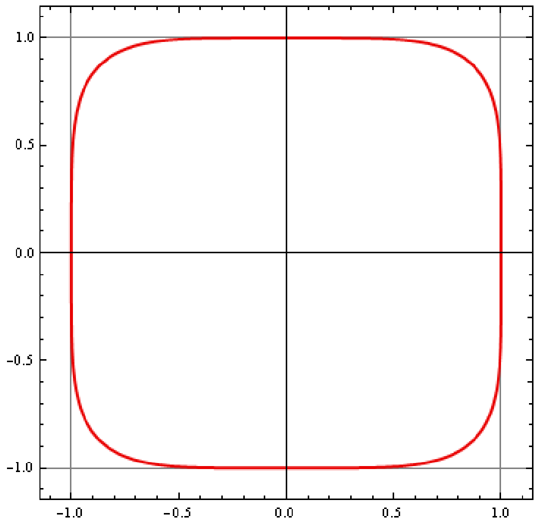



FIGURE 5.2. The unit sphere of the norm $\|\cdot\|_3$

Note that **A** is a rotation matrix with angle $\alpha = 45°$ and **B** is a rotation matrix with angle $\alpha = 30°$. Figures 5.2, 5.3 and 5.4 show the unit spheres of these norms.

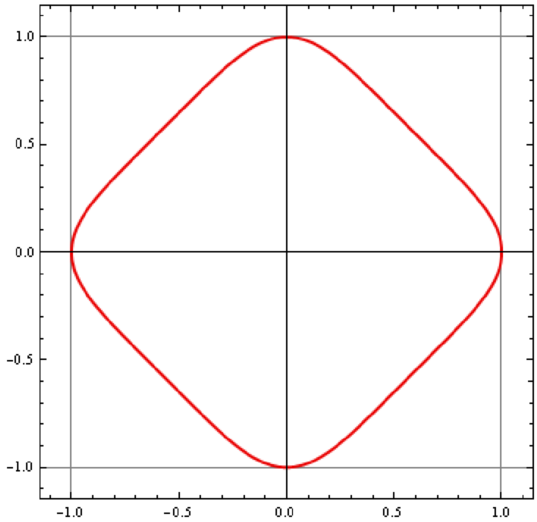



FIGURE 5.3. The unit sphere of the norm $\|\cdot\|_{\mathbf{A}}$

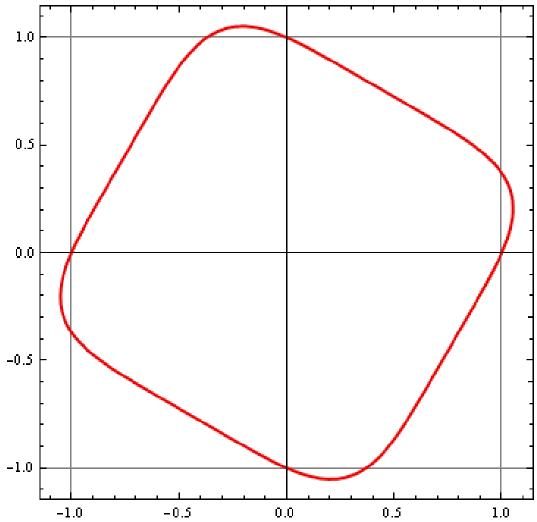



FIGURE 5.4. The unit sphere of the norm $\|\cdot\|_{\mathbf{B}}$

## 5.5. The Generalized Pareto Function of generalized asymmetric type

The family of asymmetric logistic distributions was first introduced in Tawn (1990, [**42**]) for the extreme value case and in Michel (2006, [**26**]) for the generalized pareto distributions. We now give a generalization of these models that is a little bit more readable though the side conditions to obtain the usual asymmetric logistic distributions are not easy to write down.

Definition and Theorem 5.5.1. *For* $n \in \mathbb{N}$ *choose* $n$ *norms on* $\mathbb{R}^d$ $\|\cdot\|_{\{1\}}, \dots, \|\cdot\|_{\{n\}}$ *that each fulfill condition* (3.2) *of the Main Theorem 3.1.1. Let* $0 \le \psi_{i,j}$ *for* $i = 1, \dots, n$ *and* $j = 1, \dots, d$ *with* $\sum_{i=1}^n \psi_{i,j} = 1$ *for every* $j = 1, \dots, d$.
*Then*

$$\|x_1, \dots, x_d\| := \sum_{i=1}^n \|\psi_{i,1}x_1, \dots, \psi_{i,d}x_d\|_{\{i\}}$$

*defines a norm that we call* <u>*generalized asymmetric norm*</u> *on* $\mathbb{R}^d$ *and*

$$G(x_1, \dots, x_d) := \exp(-\|x_1, \dots, x_d\|)$$

*defines an EVD that is called the* <u>*generalized asymmetric distribution*</u>*. Furthermore*

$$W(x_1, \dots, x_d) := 1 - \|x_1, \dots, x_d\|$$

*defines a GPF that is called the* <u>*generalized Pareto function of generalized asymmetric type*</u>*.*

Proof. We first verify that we have obtained a norm and thus we check the norm conditions.
As a sum of norms the new norm is nonnegative. It is zero iff every summand is zero and because every summand is a norm this is the case iff $\psi_{i,j}x_j = 0$. The side condition $\sum_{i=1}^n \psi_{i,j} = 1$ then implies that this is the case iff $x_j = 0$.
The homogeneity follows directly from the homogeneity of the original norms.

We have

$$\begin{aligned}\|\lambda x_1,\ldots,\lambda x_d\| &= \sum_{i=1}^{n}\|\psi_{i,1}\lambda x_1,\ldots,\psi_{i,d}\lambda x_d\|_{\{i\}}\\ &= \sum_{i=1}^{n}|\lambda|\,\|\psi_{i,1}x_1,\ldots,\psi_{i,d}x_d\|_{\{i\}}\\ &= |\lambda|\sum_{i=1}^{n}\|\psi_{i,1}x_1,\ldots,\psi_{i,d}x_d\|_{\{i\}}\\ &= |\lambda|\,\|\lambda x_1,\ldots,\lambda x_d\|.\end{aligned}$$

Finally we prove the triangle equation.

$$\begin{aligned}&\|x_1+y_1,\ldots,x_d+y_d\|\\ &\qquad= \sum_{i=1}^{n}\|\psi_{i,1}(x_1+y_1),\ldots,\psi_{i,d}(x_d+y_d)\|_{\{i\}}\\ &\qquad\leq \sum_{i=1}^{n}\left(\|\psi_{i,1}x_1,\ldots,\psi_{i,d}x_d\|_{\{i\}}+\|\psi_{i,1}y_1,\ldots,\psi_{i,d}y_d\|_{\{i\}}\right)\\ &\qquad= \|x_1,\ldots,x_d\|+\|y_1,\ldots,y_d\|.\end{aligned}$$

So we have really defined a new norm.
Now we check the condition of the Main Theorem 3.1.1.
For any two vectors $\mathbf{x},\mathbf{y}\in(-\infty,0]^d$ with $\mathbf{x}\leq\mathbf{y}$ and any subset $E\subsetneq\{1,\ldots,d\}$ we have

$$\begin{aligned}&\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\in E}}(-1)^{d+1-\sum_{j=1}^d m_j}\left\|y_1^{m_1}x_1^{1-m_1},\ldots,y_d^{m_d}x_d^{1-m_d}\right\|\\ &=\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\in E}}(-1)^{d+1-\sum_{j=1}^d m_j}\sum_{i=1}^{n}\left\|\psi_{i,1}y_1^{m_1}x_1^{1-m_1},\ldots,\psi_{i,d}y_d^{m_d}x_d^{1-m_d}\right\|_{\{i\}}\\ &=\sum_{i=1}^{n}\underbrace{\sum_{\substack{\mathbf{m}\in\{0,1\}^d\\ m_j=1,j\in E}}(-1)^{d+1-\sum_{j=1}^d m_j}\left\|(\psi_{i,1}y_1)^{m_1}(\psi_{i,1}x_1)^{1-m_1},\ldots,(\psi_{i,d}y_d)^{m_d}(\psi_{i,d}x_d)^{1-m_d}\right\|_{\{i\}}}_{\geq 0,\text{ since }\|\cdot\|_{\{i\}}\text{ fulfills the condition of the Main Theorem}}\\ &\geq 0.\end{aligned}$$

With our new norm we can define an EVD and therefore also a GPD.

□

Remark 5.5.2. As already indicated by its name the asymmetric logistic distribution is a special case of the generalized asymmetric distribution. Set $n = 2^d - d$ and the first norm as the sum-norm. All other norms are chosen to be $\lambda$-norms and some coefficient has to be zero. For the case of $d = 2$ no more condition is needed. We obtain the trivariate asymmetric logistic distribution by putting $\psi_{2,3} = 0$, $\psi_{3,2} = 0$, $\psi_{4,1} = 0$.
In contrary to the original asymmetric logistic distribution the generalization is no longer parameterizable due to the nonparametrizability of the $n$ norms. In the following we will overcome this by restricting ourselves to $\lambda$-norms.

Lemma 5.5.3. *Let $\|\cdot\|$ be a generalized asymmetric norm as defined in Definition 5.5.1 with the $n$ norms chosen to be $\lambda$-norms, i.e. for $\lambda_i \geq 1$, $i = 1, \ldots, n$, it is $\|\cdot\|_{\{i\}} = \|\cdot\|_{\lambda_i}$ and let $W$ be the corresponding generalized Pareto function. Then $W$ has the density $w$ given by*

$$w(x_1, \ldots, x_d)$$
$$= \sum_{i=1}^{n} \prod_{j=1}^{d} \psi_{i,j} \prod_{j=1}^{d-1} (j\lambda_i - 1) \prod_{j=1}^{d} (-\psi_{i,j} x_j)^{\lambda_i - 1} \left\| (\psi_{i,1} x_1, \ldots, \psi_{i,d} x_d)^T \right\|_{\lambda_i}^{1-d\lambda_i},$$

$\mathbf{x} < 0$ *and close to the origin.*

Proof. We know from Lemma 2.3.6 in Michel (2006, [**26**]) that the generalized Pareto distribution of logistic type $W_\lambda$ has the density

$$\begin{aligned} w_\lambda(x_1, \ldots, x_d) &:= \frac{\partial^d}{\partial x_1 \ldots \partial x_d} W_\lambda(x_1, \ldots, x_d) \\ &= \prod_{i=1}^{d-1} (i\lambda - 1) \prod_{i=1}^{d} (-x_i)^{\lambda - 1} \|\mathbf{x}\|_\lambda^{1-d\lambda}. \end{aligned}$$

So we obtain

$$w(x_1, \ldots, x_d)$$
$$:= \frac{\partial^d}{\partial x_1 \ldots \partial x_d} W(x_1, \ldots, x_d)$$

$$
\begin{aligned}
&= \frac{\partial^d}{\partial x_1 \dots \partial x_d} (1 - \|\mathbf{x}\|) \\
&= \frac{\partial^d}{\partial x_1 \dots \partial x_d} \left( 1 - \sum_{i=1}^{n} \|\psi_{i,1} x_1, \dots, \psi_{i,d} x_d\|_{\lambda_i} \right) \\
&= \frac{\partial^d}{\partial x_1 \dots \partial x_d} \left( (1-n) + \sum_{i=1}^{n} \left( 1 - \|\psi_{i,1} x_1, \dots, \psi_{i,d} x_d\|_{\lambda_i} \right) \right) \\
&= \frac{\partial^d}{\partial x_1 \dots \partial x_d} \left( (1-n) + \sum_{i=1}^{n} W_{\lambda_i} \left( \psi_{i,1} x_1, \dots, \psi_{i,d} x_d \right) \right) \\
&= \sum_{i=1}^{n} \left( \prod_{j=1}^{d} \psi_{i,j} \right) w_{\lambda_i} \left( \psi_{i,1} x_1, \dots, \psi_{i,d} x_d \right) \\
&= \sum_{i=1}^{n} \left( \prod_{j=1}^{d} \psi_{i,j} \prod_{j=1}^{d-1} (j\lambda_i - 1) \prod_{j=1}^{d} (-\psi_{i,j} x_j)^{\lambda_i - 1} \left\| (\psi_{i,1} x_1, \dots, \psi_{i,d} x_d)^T \right\|_{\lambda_i}^{1-d\lambda_i} \right).
\end{aligned}
$$

□

Lemma 5.5.4. *Under the same conditions as in Lemma 5.5.3, the Pickands density is given by*

$$
\begin{aligned}
\phi(z_1, \dots, z_{d-1}) = \sum_{i=1}^{n} \Bigg( & \prod_{j=1}^{d} \psi_{i,j} \prod_{j=1}^{d-1} (j\lambda_i - 1) \\
& \prod_{j=1}^{d-1} (\psi_{i,j} z_j)^{\lambda_i - 1} \left( \psi_{i,d} \left( 1 - \sum_{k=1}^{d-1} z_k \right) \right)^{\lambda_i - 1} \\
& \left\| \psi_{i,1} z_1, \dots, \psi_{i,d-1} z_{d-1}, \psi_{i,d} \left( 1 - \sum_{k=1}^{d-1} z_k \right) \right\|_{\lambda_i}^{-d\lambda_i + 1} \Bigg).
\end{aligned}
$$

Proof. Remind the definition of the Pickands density $\phi$ given in Definition 2.2.4, i.e.

$$\phi(z_1, \dots, z_{d-1}) = |c|^{d-1} w \left( T_P^{-1} (z_1, \dots, z_{d-1}, c) \right),$$

where $T_P^{-1}(z_1, \dots, d_{d-1}, c) := c \left( z_1, \dots, z_{d-1}, 1 - \sum_{i=1}^{d-1} z_i \right)^T$, $0 \le z_i$, $\sum_{i=1}^{d-1} z_i \le 1$, is the inverse of the transformation to standard Pickands coordinates. Then we obtain

$$\begin{aligned}
&\phi\left(z_1,\ldots,z_{d-1}\right)\\
&\quad= |c|^{d-1}\, w\left(T_P^{-1}\left(z_1,\ldots,z_{d-1},c\right)\right)\\
&\quad= |c|^{d-1}\sum_{i=1}^{n}\left[\prod_{j=1}^{d}\psi_{i,j}\prod_{j=1}^{d-1}\left(j\lambda_i-1\right)\right.\\
&\qquad\quad \prod_{j=1}^{d-1}\left(-c\psi_{i,j}z_j\right)^{\lambda_i-1}\left(-c\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^{\lambda_i-1}\\
&\qquad\quad \left.\left\|\left(c\psi_{i,1}z_1,\ldots,c\psi_{i,d-1}z_{d-1},c\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^T\right\|_{\lambda_i}^{1-d\lambda_i}\right]\\
&\quad= |c|^{d-1}\sum_{i=1}^{n}\left[\prod_{j=1}^{d}\psi_{i,j}\prod_{j=1}^{d-1}\left(j\lambda_i-1\right)\right.\\
&\qquad\quad (-c)^{(\lambda_i-1)(d-1)}\prod_{j=1}^{d-1}\left(\psi_{i,j}z_j\right)^{\lambda_i-1}(-c)^{(\lambda_i-1)}\left(\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^{\lambda_i-1}\\
&\qquad\quad \left.|c|^{1-d\lambda_i}\left\|\left(\psi_{i,1}z_1,\ldots,\psi_{i,d-1}z_{d-1},\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^T\right\|_{\lambda_i}^{1-d\lambda_i}\right]\\
&\quad= \sum_{i=1}^{n}\left[\prod_{j=1}^{d}\psi_{i,j}\prod_{j=1}^{d-1}\left(j\lambda_i-1\right)\prod_{j=1}^{d-1}\left(\psi_{i,j}z_j\right)^{\lambda_i-1}\left(\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^{\lambda_i-1}\right.\\
&\qquad\quad \left.\left\|\left(\psi_{i,1}z_1,\ldots,\psi_{i,d-1}z_{d-1},\psi_{i,d}\left(1-\sum_{j=1}^{d-1}z_j\right)\right)^T\right\|_{\lambda_i}^{1-d\lambda_i}\right].
\end{aligned}$$

□

REMARK 5.5.5. Looking at Lemma 5.5.4 we see that the $i$-th summand of the Pickands density vanishes for every $\mathbf{z}$ if and only if $\lambda_i = 1$ or $\prod_{j=1}^{d} \psi_{i,j} = 0$, i.e. there is at least one index $j$ with $\psi_{i,j} = 0$.
So the Pickands density of the asymmetric model reduces to the last summand and therefore only depends on $d+1$ parameters, namely $\lambda_d$ and $\psi_{i,1}, \ldots, \psi_{i,d}$.

LEMMA 5.5.6. *Under the conditions of Lemma 5.5.4 the angular density is given by*

$$
\begin{aligned}
l\left(z_1, \ldots, z_{d-1}\right) = \sum_{i=1}^{n} \Bigg( & \prod_{j=1}^{d} \psi_{i,j} \prod_{j=1}^{d-1} \left(j\lambda_i - 1\right) \\
& \prod_{j=1}^{d-1} \left(-\psi_{i,j}\frac{1}{z_j}\right)^{\lambda_i - 1} \left(-\psi_{i,d}\frac{1}{1-\sum_{j=1}^{d-1} z_j}\right)^{\lambda_d - 1} \\
& \left\| \left(\psi_{i,1}\frac{1}{z_1}, \ldots, \psi_{i,d-1}\frac{1}{z_{d-1}}, \psi_{i,d}\frac{1}{1-\sum_{j=1}^{d-1} z_j}\right)^T \right\|_{\lambda_i}^{1-d\lambda_i} \Bigg).
\end{aligned}
$$

PROOF. According to Theorem 2.2.4 in Michel (2006, [**26**]) a generalized Pareto distribution $W$, that is continuously differentiable of order $d$, the angular density $l$ fulfills

$$
l\left(\frac{\frac{1}{x_1}}{\sum_{i=1}^{d} \frac{1}{x_i}}, \ldots, \frac{\frac{1}{x_{d-1}}}{\sum_{i=1}^{d} \frac{1}{x_i}}\right) = \frac{x_1^2 \ldots x_d^2}{\left(-\sum_{i=1}^{d} \frac{1}{x_i}\right)^{-(d+1)}} \frac{\partial^d}{\partial x_1 \ldots \partial x_d} W\left(x_1, \ldots, x_d\right).
$$

For $0 < z_i < 1$, $1 \le i \le d-1$, we set $x_i := \frac{1}{z_i}$, $1 \le i \le d-1$, and $x_d := \frac{1}{1-\sum_{j=1}^{d-1} z_j}$. Then we have $z_i = \frac{1}{x_i}$, for $1 \le i \le d-1$, $\frac{1}{x_d} = 1 - \sum_{j=1}^{d-1} \frac{1}{x_j}$ and hence $\sum_{i=1}^{d} \frac{1}{x_i} = 1$. Using the representation of $w$ from Lemma 5.5.3 we obtain:

$$
\begin{aligned}
& l\left(z_1, \ldots, z_{d-1}\right) \\
&= (-1)^{d+1} \prod_{i=1}^{d-1} z_i^{-2} \left(1 - \sum_{i=1}^{d-1} z_i\right)^{-2} w\left(\frac{1}{z_1}, \ldots, \frac{1}{z_{d-1}}, \frac{1}{1-\sum_{i=1}^{d-1} z_i}\right) \\
&= (-1)^{d+1} \prod_{i=1}^{d-1} z_i^{-2} \left(1 - \sum_{i=1}^{d-1} z_i\right)^{-2} \\
&\quad \sum_{i=1}^{n} \Bigg( \prod_{j=1}^{d} \psi_{i,j} \prod_{j=1}^{d-1} \left(j\lambda_i - 1\right) \prod_{j=1}^{d-1} \left(-\psi_{i,j}\frac{1}{z_j}\right)^{\lambda_i - 1} \left(-\psi_{i,d}\frac{1}{1-\sum_{j=1}^{d-1} z_j}\right)^{\lambda_d - 1}
\end{aligned}
$$

$$\left\| \left( \psi_{i,1} \frac{1}{z_1}, \ldots, \psi_{i,d-1} \frac{1}{z_{d-1}}, \psi_{i,d} \frac{1}{1 - \sum_{j=1}^{d-1} z_j} \right)^T \right\|_{\lambda_i}^{1-d\lambda_i} \Bigg).$$

□

Remark 5.5.7. We have the same situation as in Remark 5.5.5. From Lemma 5.5.4 we obtain that the $i$th summand of the angular density vanishes for every $\mathbf{z}$ if and only if $\lambda_i = 1$ or $\prod_{j=1}^d \psi_{i,j} = 0$, i.e. there is at least one index $j$ with $\psi_{i,j} = 0$.
So the angular density of the asymmetric model reduces to the last summand and therefore only depends on $d + 1$ parameters, namely $\lambda_d$ and $\psi_{i,1}, \ldots, \psi_{i,d}$.

CHAPTER 6

# The GPD-Flow

## 6.1. Introduction

Aulbach et al. extended in an unpublished paper ([**1**]) a method for the simulation of random variables of GPDs that was introduced in the bivariate case by Buishand et al. (2008, [**7**]). The according theorem is the basis of this chapter

THEOREM 6.1.1. (i) *Let $W$ be a multivariate GPD with uniform margins in a left neighborhood of $\mathbf{0} \in \mathbb{R}^d$. Then there exists a random vector $\mathbf{Z} = (Z_1, \ldots, Z_d)$ with $Z_i \in [0, d]$ and $\mathrm{E}(Z_i) = 1$, $i \le d$, and a vector $\mathbf{x}_0 < \mathbf{0}$ such that*

$$W(\mathbf{x}) = \mathrm{P}\left(-U\left(\frac{1}{Z_1}, \ldots, \frac{1}{Z_d}\right) \le \mathbf{x}\right), \quad \mathbf{x}_0 \le \mathbf{x} \le \mathbf{0},$$

*where the random variable $U$ is uniformly on $(0,1)$ distributed and independent of $\mathbf{Z}$.*

(ii) *The random vector $-U(1/Z_1, \ldots, 1/Z_d)$ follows a GPD with uniform margins in a left neighborhood of $\mathbf{0} \in \mathbb{R}^d$ if $U$ is independent of $\mathbf{Z} = (Z_1, \ldots, Z_d)$ and $0 \le Z_i \le c_i$ a.s. with $\mathrm{E}(Z_i) = 1$, $i \le d$, for some $c_1, \ldots, c_d \ge 1$.*

In the following we will repeat the proof by Aulbach, Bayer and Falk ([**1**]).
PROOF. First we establish part (i). Recall that a multivariate GPD $W$ with uniform margins $1 - W_i(x) = \text{const}\, x$, $i \le d$, in a left neighborhood of $\mathbf{0} \in \mathbb{R}^d$ can be represented as

$$\begin{aligned} W(\mathbf{x}) &= 1 + \text{const}\left(\sum_{j \le d} x_j\right) \int_{S_d} \max_{i \le d}\left(\frac{x_i}{\sum_{j \le d} x_j} t_i\right) \mu(\mathrm{d}\,\mathbf{t}) \\ &= 1 + \text{const}\left(\sum_{j \le d} x_j\right) D\left(\frac{x_1}{\sum_{j \le d} x_j}, \ldots, \frac{x_{d-1}}{\sum_{j \le d} x_j}\right) \end{aligned}$$

with const $> 0$ and some measure $\mu$ on $S_d$ such that $\mu(S_d) = d$ and $\int_{S_d} t_i \mu(d\mathbf{t}) = 1$, $i \le m$. For the sake of simplicity we assume in the following that const $= 1$. Now $\tilde{\mu}(\cdot) = \mu(\cdot)/d$ defines a probability measure on $S_d$. Let $\mathbf{T} = (T_1, \dots, T_d)$ be a random variable with values in $S_d$ that has a distribution $\tilde{\mu}$ and put $\mathbf{Z} := d\mathbf{T}$. Then $\mathbf{Z} \in [0, d]^d$ and $\mathrm{E}(Z_i) = \int_{S_d} t_i \mu(d\mathbf{t}) = 1$, $i \le d$. We have, further, for $\mathbf{x} \le \mathbf{0} \in \mathbb{R}^d$ with $x_j \ge -1/d$, $j \le d$,

$$
\begin{aligned}
&\mathrm{P}\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \mathbf{x}\right) \\
&= \mathrm{P}\left(-U\left(\frac{1}{T_1}, \dots, \frac{1}{T_d}\right) \le d\mathbf{x}\right) \\
&= \int_{S_d} \mathrm{P}\left(-U\left(\frac{1}{t_1}, \dots, \frac{1}{t_d}\right) \le d\mathbf{x} \middle| \mathbf{T} = \mathbf{t}\right) (P * \mathbf{T})(\mathrm{d}\,\mathbf{t}) \\
&= \int_{S_d} \mathrm{P}\left(-U\left(\frac{1}{t_1}, \dots, \frac{1}{t_d}\right) \le d\mathbf{x}\right) \tilde{\mu}(\mathrm{d}\,\mathbf{t}) \\
&= \frac{1}{d} \int_{S_d} \mathrm{P}\left(-U\left(\frac{1}{t_1}, \dots, \frac{1}{t_d}\right) \le d\mathbf{x}\right) \mu(\mathrm{d}\,\mathbf{t}) \\
&= \frac{1}{d} \int_{S_d} \mathrm{P}\left(U \ge d \max_{i \le d}(-x_i t_i)\right) \mu(\mathrm{d}\,\mathbf{t}) \\
&= \frac{1}{d} \int_{S_d} \mathrm{P}\left(U \ge -d\left(\sum_{j \le d} x_j\right) \max_{i \le d}\left(\frac{x_i}{\sum_{j \le d} x_j} t_i\right)\right) \mu(\mathrm{d}\,\mathbf{t}) \\
&= \frac{1}{d} \int_{S_d} 1 + d\left(\sum_{j \le d} x_j\right) \max_{i \le d}\left(\frac{x_i}{\sum_{j \le d} x_j} t_i\right) \mu(\mathrm{d}\,\mathbf{t}) \\
&= 1 + \left(\sum_{j \le d} x_j\right) \int_{S_d} \max_{i \le d}\left(\frac{x_i}{\sum_{j \le d} x_j} t_i\right) \mu(\mathrm{d}\,\mathbf{t}).
\end{aligned}
$$

This implies part (i) of the Proposition.
On the other hand we have for $\mathbf{x} \le \mathbf{0}$ and large $s > 0$

$$
\begin{aligned}
&\mathrm{P}\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \frac{1}{s}\mathbf{x}\right)^s \\
&= \left(\int_{[\mathbf{0}, \mathbf{x}]} \mathrm{P}\left(U \ge \frac{1}{s} \max_{i \ge d}(-x_i z_i)\right) (P * \mathbf{Z})(\mathrm{d}\,\mathbf{z})\right)^s
\end{aligned}
$$

$$
\begin{aligned}
&= \left(1 - \frac{1}{s}\int_{[\mathbf{0},\mathbf{x}]} \max_{i\geq d}\left(-x_i z_i\right)\left(P * \mathbf{Z}\right)\left(\mathrm{d}\,\mathbf{z}\right)\right)^s \\
&\to_{s\to\infty} \exp\left(-\int_{[\mathbf{0},\mathbf{x}]} \max_{i\geq d}\left(-x_i z_i\right)\left(P * \mathbf{Z}\right)\left(\mathrm{d}\,\mathbf{z}\right)\right) \\
&=: G\left(\mathbf{x}\right)
\end{aligned}
$$

with $\mathbf{c} := (c_1, \ldots, c_d)$.
Lemma 7.2.1 in Reiss (1989, [**32**]) now implies that $G$ is a distribution function which is obviously max-stable: $G^s\left(s^{-1}\mathbf{x}\right) = G\left(\mathbf{x}\right)$, $s > 0$, i.e., $G$ is a multivariate EVD and has negative standard exponential margins $G_i\left(x\right) = \exp\left(x\,\mathrm{E}\left(Z_i\right)\right) = \exp\left(x\right)$, $x \leq 0$. As a consequence, $1 + \log\left(G\left(\mathbf{x}\right)\right)$ is a GP function with

$$
\begin{aligned}
1 + \log\left(G\left(\mathbf{x}\right)\right) &= 1 - \int_{[\mathbf{0},\mathbf{c}]} \max_{i\leq d}\left(-x_i z_i\right)\left(P * \mathbf{Z}\right)\left(d\mathbf{z}\right) \\
&= \mathrm{P}\left(-U\left(\frac{1}{Z_1}, \ldots, \frac{1}{Z_d}\right) \leq \mathbf{x}\right)
\end{aligned}
$$

for $\mathbf{x}_0 \leq \mathbf{x} \leq \mathbf{0}$ and some $\mathbf{x}_0 < \mathbf{0}$.

□

Let $C$ be a copula and $\mathbf{V}$ is a random vector having the distribution function $C$. Then $\mathbf{Z} := 2\mathbf{V}$ is a proper choice in part (ii) of Theorem 6.1.1. As a consequence we can create random vectors that follow a GPD if we can create random vectors of the corresponding copula.
First we use the theorem to specify the neighborhood in which a GPD is a distribution function. Then we want to find a representation for the $D$-Norm depending on the copula in the bivariate case. Next we generalize this result to arbitrary dimension. Starting with a copula we get a multivariate distribution and therefore we get a new copula. With this copula we can again use part (ii) of Theorem 6.1.1 and this yields a new multivariate GPD. Apparently this step can be iterated over and over again which we call the "GPD-Flow". It will be explored in Section 6.5.

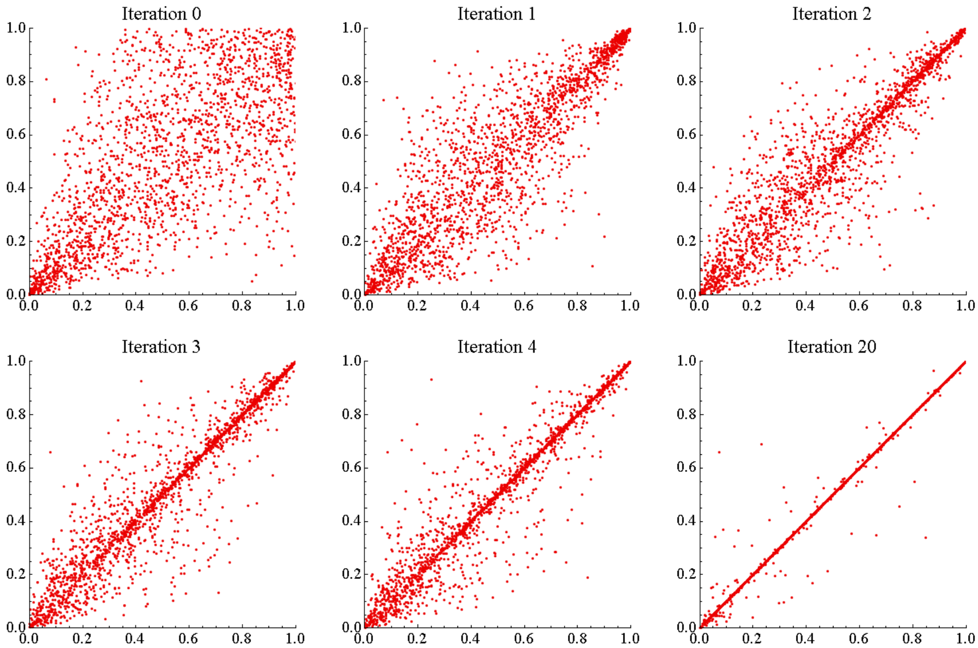



FIGURE 6.1. 2000 realizations of the GPD-Flow starting with the Clayton-Copula with $\theta = 2$

Figure 6.1 shows realizations of iterations of the GPD-Flow starting with the Clayton-Copula with $\theta = 2$. The first iteration will be analytical regarded in Example 6.3.1. In this example we can see that the GPD-Flow tends to become a line and hence the underlying copula tends to become the copula of complete dependence. Figure 6.2 displays the estimated unit sphere of the underlying norm in the positive quadrant. The norm is estimated using the estimator given in Section 10.2 of Reiss and Thomas (2007, [**33**]) namely

$$\hat{D}_n(z) = 1 - \frac{1}{nc} \sum_{i=1}^{n} I\left(x_i > -c(1-z), y_i > -cz\right), 0 < c \leq 1,$$

where $(x_i, y_i)$ are the underlying data. Furthermore we used the homogeneity of the norm $\hat{\eta}(z)$ since we estimate the norm for $(z, 1-z)$, $z \in [0,1]$, and the estimated point on the unit sphere is $\left(\frac{z}{\hat{\eta}(z)}, \frac{1-z}{\hat{\eta}(z)}\right)$. This figure also hints that the GPD-Flow converges to complete dependency since the unit spheres approaches the unit sphere of the maximum norm.

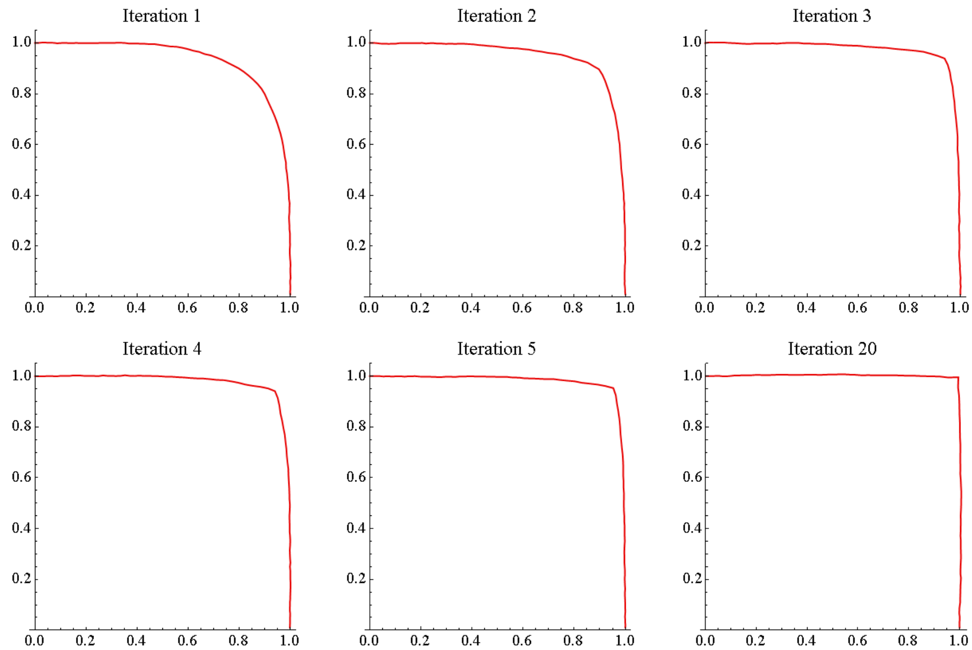



FIGURE 6.2. estimated $D$-Norm of the GPD-Flow started with the Clayton Copula with $\theta = 2$, $n$=100000 and $c = 0.5$

Furthermore we also simulated the GPD-Flow starting with the data that underlies the GPD of tail-independence, thus $W(x, y) = \max(1 + x + y, 0)$, $x, y \leq 0$. The copula $C$ that pertains to $W$ is given by $C(u, v) = \max(u + v - 1, 0)$. This copula is the Fréchet-Hoeffding lower bound (see Section 2.2 in Nelsen (2006, [**31**])). Figure 6.3 shows realizations of that GPD-Flow and Figure 6.4 shows the estimated unit sphere of the $D$-Norm, where the $D$-Norm is estimated as mentioned above.
It seems that the GPD-Flow also converges, as seen in the example of the Clayton Copula, against the GPD of total tail-dependence. In a sense the GPD of tail-independence is the antipode of the GPD of total tail-dependence. Thus although if we start "farthest" away from the case of total tail-dependence the GPD-Flow converges to that case. This also hints that the GPD-Flow converges.

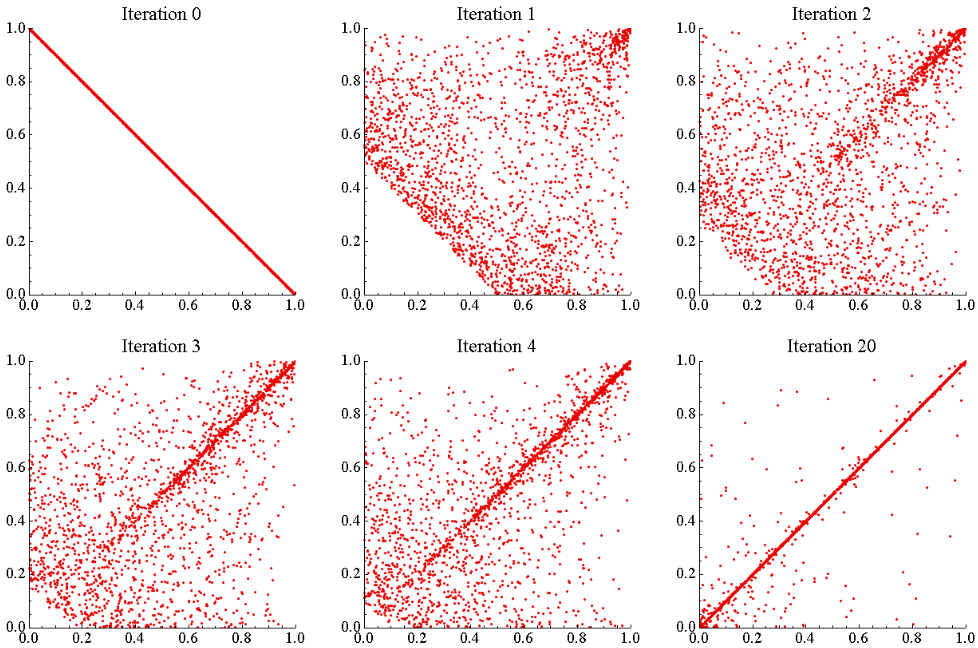



FIGURE 6.3. 2000 realizations of the GPD-Flow started with GPD of tail-independence

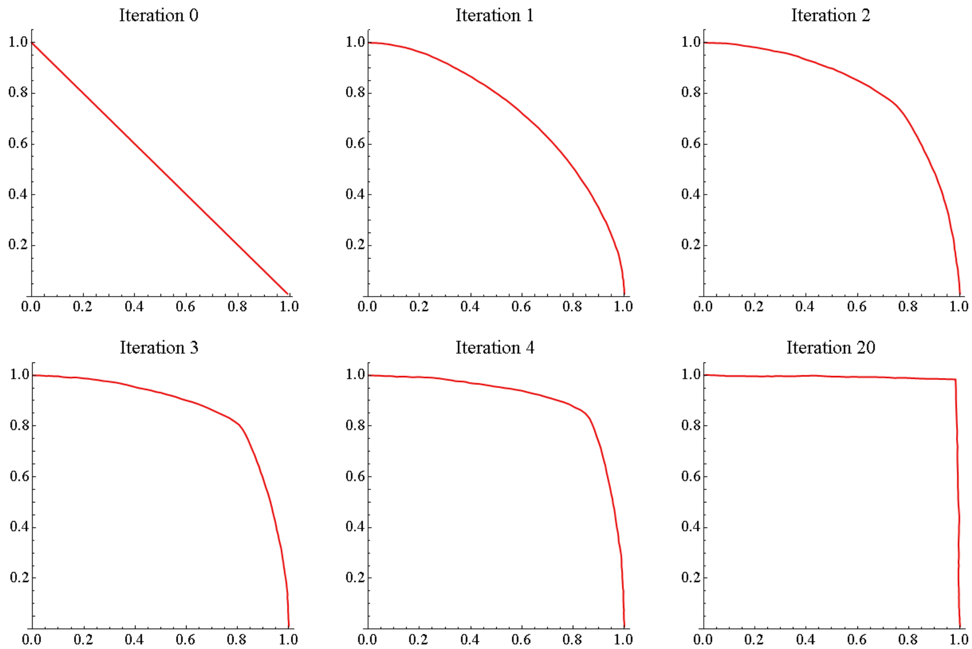



FIGURE 6.4. estimated $D$-Norm of the GPD-Flow started with GPD of tail-independence with $n$=100000 and $c = 0.5$

Finally Figure 6.5 shows realizations of the 3 dimensional GPD-Flow started with the Clayton Copula with parameter $\theta = 2$. In Example 6.4.1 we will explore the first iteration step. Figure 6.6 shows the estimated unit spheres of the underlying norms. The estimation procedure is done in a similar way as in the 2 dimensional case as stated above. Thereby we use

$$1 - F(\mathbf{x}) = \sum_{j \le d} (-1)^{j+1} \sum_{|K|=j} \overline{F}_K(\mathbf{x})$$

where $\overline{F}_K(\mathbf{x})$ denotes the margins of the survivor function $\overline{F}$ and is defined by

$$\overline{F}_K(\mathbf{x}) := \mathrm{P}(X_k > x_k, k \in K).$$

This is equation (8.3) in Reiss and Thomas (2007, [**33**]).
Hence we have for a GPD $W$ with uniform margins in a left neighborhood of $\mathbf{0}$

$$\begin{aligned}
\|\mathbf{z}\| &= 1 - W(\mathbf{z}) \\
&= \overline{W}_{\{1\}}(\mathbf{z}) + \overline{W}_{\{2\}}(\mathbf{z}) + \overline{W}_{\{3\}}(\mathbf{z}) \\
&\qquad\qquad - \overline{W}_{\{1,2\}}(\mathbf{z}) - \overline{W}_{\{1,3\}}(\mathbf{z}) - \overline{W}_{\{2,3\}}(\mathbf{z}) + \overline{W}(\mathbf{z}) \\
&= z_1 + z_2 + z_3 - \overline{W}_{\{1,2\}}(\mathbf{z}) - \overline{W}_{\{1,3\}}(\mathbf{z}) - \overline{W}_{\{2,3\}}(\mathbf{z}) + \overline{W}(\mathbf{z}).
\end{aligned}$$

Therefore an estimator $\hat{\eta}$ for the $D$-Norm is given by

$$\begin{aligned}
\hat{\eta}(\mathbf{z}) = \|\mathbf{z}\|_1 &- \frac{1}{n}\sum_{i=1}^{n} I(u_i > z_1)\, I(v_i > z_2) - \frac{1}{n}\sum_{i=1}^{n} I(u_i > z_1)\, I(w_i > z_3) \\
&- \frac{1}{n}\sum_{i=1}^{n} I(v_i > z_2)\, I(w_i > z_3) + \frac{1}{n}\sum_{i=1}^{n} I(u_i > z_1)\, I(v_i > z_2)\, I(w_i > z_3)
\end{aligned}$$

where $(u, v, w)_i$, $1 \le i \le n$, is governed by a GPD $W$ and $\mathbf{z}$ is in left neighborhood of $\mathbf{0}$.

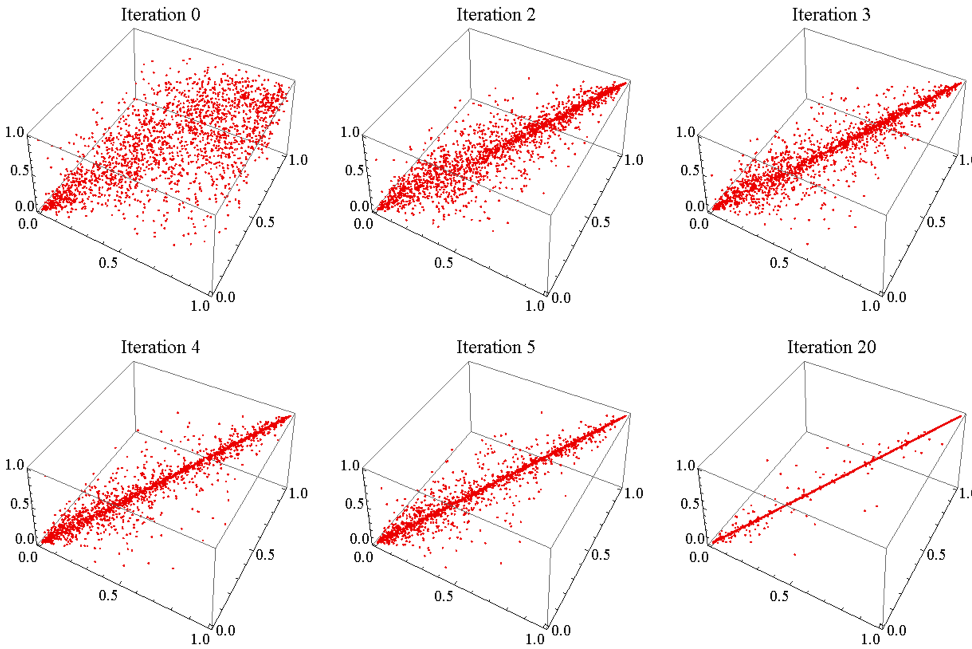



FIGURE 6.5. 2000 realizations of the GPD-Flow started with the Clayton Copula with $\theta = 2$

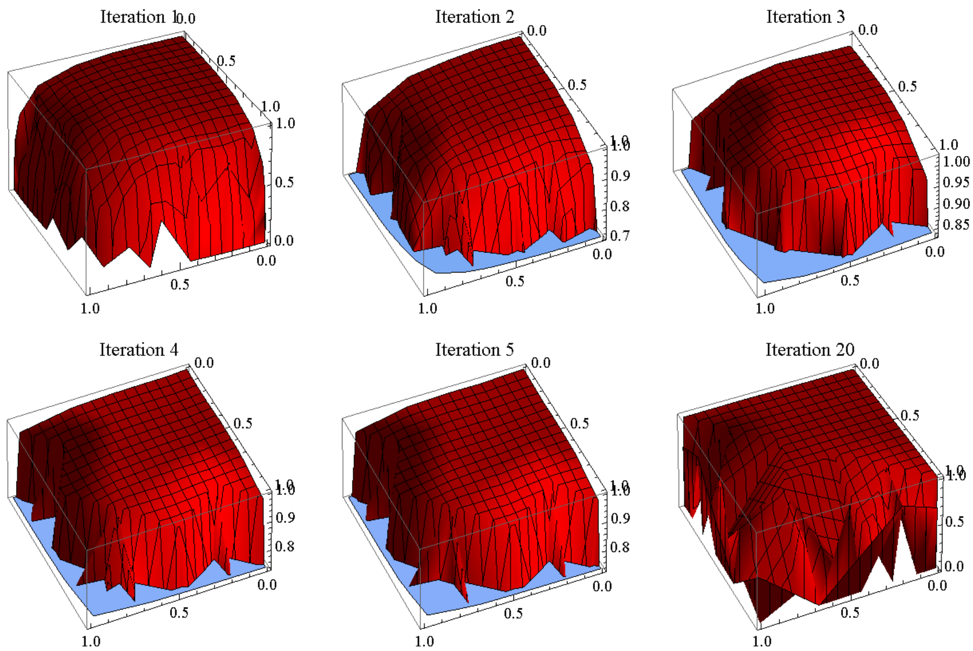



FIGURE 6.6. estimated $D$-Norm of the 3 dimensional GPD-Flow started with the Clayton Copula with $\theta = 2$, $n$=50000 and $c = \frac{1}{3}$

## 6.2. The domain of a GPD

This section contains an important application of Theorem 6.1.1. In the definition of a GPD there is the statement that there exists an $x_0 < 0$ that the GPD is a distribution function on $[x_0, 0]^d$. In the next theorem we see that this can be concretized namely $x_0$ can be chosen to be $x_0 = -\frac{1}{d}$.

THEOREM 6.2.1. *Let $\|.\|_D$ be a d-dimensional D-Norm. Than there exists a d-variate distribution function $W$ with*

$$W(\mathbf{x}) = 1 - \|\mathbf{x}\|_D, \quad \text{for } \mathbf{x} \in \left[-\frac{1}{d}, 0\right]^d.$$

PROOF. Let $\tilde{W}$ be a GPD with associated $D$-Norm $\|.\|_D$. According to part (i) of Theorem 6.1.1 there exists a random vector $\mathbf{Z} = (Z_1, \dots, Z_d)$ with $Z_i \in [0, d]$ and an $\mathbf{x}_0 < 0$ with

$$\tilde{W}(\mathbf{x}) = 1 - \|\mathbf{x}\|_D = \mathrm{P}\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \mathbf{x}\right), \quad \mathbf{x}_0 \le \mathbf{x} \le \mathbf{0}.$$

By $F$ we denote the distribution function of $\mathbf{Z}$ and by $1 \le \gamma_i \le d$ we denote the upper bound of $Z_i$, i.e. $Z_i \in [0, \gamma_i]$ with probability 1. We set $\mathbf{X} := -U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right)$ and let $W$ be the distribution function of $\mathbf{X}$. Now we calculate the probability $\mathrm{P}\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \mathbf{x}\right)$ for $-\frac{1}{\gamma_i} \le x_i < 0$, $1 \le i \le d$.
Without loss of generalization we assume that $\gamma_d x_d \le \dots \le \gamma_1 x_1 < \gamma_0 x_0 := 0$. Then we obtain

$$\begin{aligned}
&\mathrm{P}\left(-U\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \mathbf{x}\right) \\
&= \int_0^1 \mathrm{P}\left(-u\left(\frac{1}{Z_1}, \dots, \frac{1}{Z_d}\right) \le \mathbf{x} \,\middle|\, U = u\right) \mathrm{d}\,u \\
&= \int_0^1 F\left(-\frac{u}{x_1}, \dots, -\frac{u}{x_d}\right) \mathrm{d}\,u \\
&= \int_0^{-\gamma_1 x_1} F\left(-\frac{u}{x_1}, \dots, -\frac{u}{x_d}\right) \mathrm{d}\,u + \int_{-\gamma_1 x_1}^{-\gamma_2 x_2} F\left(\gamma_1, -\frac{u}{x_2}, \dots, -\frac{u}{x_d}\right) + \dots + \\
&\quad \int_{-\gamma_{d-1} x_{d-1}}^{-\gamma_d x_d} F\left(\gamma_1, \dots, \gamma_{d-1}, -\frac{u}{x_d}\right) + \int_{-\gamma_d x_d}^{1} 1 \,\mathrm{d}\,u \\
&= (1 + \gamma_d x_d) + \sum_{i=1}^{d} \int_{-\gamma_{i-1} x_{i-1}}^{-\gamma_i x_i} F\left(\sum_{j=1}^{i-1} \gamma_j \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{u}{x_j}\right)\right) \mathrm{d}\,u
\end{aligned}$$

$$= 1 - \left( -\gamma_d x_d - \sum_{i=1}^{d} \left(\gamma_{i-1} x_{i-1} - \gamma_i x_i\right) \right.$$
$$\left. \int_0^1 F\left( \sum_{j=1}^{i-1} \gamma_j \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left( -\frac{\left(\gamma_{i-1} x_{i-1} - \gamma_i x_i\right) u - \gamma_{i-1} x_{i-1}}{x_j} \right) \right) \mathrm{d}\, u \right).$$

For $\mathbf{x} < \mathbf{0}$ set

$$h\left(\mathbf{x}\right) := -\left(\gamma x\right)_{1:d} - \sum_{i=1}^{d} \left( \left(\gamma x\right)_{d+2-i:d} - \left(\gamma x\right)_{d+1-i:d} \right)$$
$$\int_0^1 F\left( \sum_{j=1}^{i-1} \gamma_{(d+1-j,d)} \mathbf{e}_{(d+1-j,d)} + \sum_{j=i}^{d} \mathbf{e}_{(d+1-j,d)} \right.$$
$$\left. \left( -\frac{\left( \left(\gamma x\right)_{d+2-i:d} - \left(\gamma x\right)_{d+1-i:d} \right) u + \left(\gamma x\right)_{d+2-i:d}}{x_{(d+1-j,d)}} \right) \right) \mathrm{d}\, u,$$

where $\left(\gamma x\right)_{i:d}$ denotes the $i$-th greatest value of the $\gamma_j x_j$. Furthermore the ordering of $x_{(i,d)}$ $\gamma_{(i,d)}$ and $\mathbf{e}_{(i,d)}$ is meant also according to the $\gamma_j x_j$.
We see that $h$ is homogeneous, i.e. for all $\mathbf{x} < \mathbf{0}$ and $\lambda > 0$ we have $h\left(\lambda \mathbf{x}\right) = \lambda h\left(\mathbf{x}\right)$. Furthermore $h$ coincides with the $D$-Norm for $\mathbf{x} \in [\mathbf{x}_0, \mathbf{0})$ and thus, since $h$ and the $D$-Norm are homogeneous both coincides in the negative orthant.
For $\mathbf{x} \in \times_{i=1}^{d} \left[-\frac{1}{\gamma_i}, 0\right]$ and thus in particular for $\mathbf{x} \in \left[-\frac{1}{d}, 0\right)^d \subseteq \times_{i=1}^{d} \left[-\frac{1}{\gamma_i}, 0\right]$ we have

$$\begin{aligned} W\left(\mathbf{x}\right) &= \mathrm{P}\left(\mathbf{X} \leq \mathbf{x}\right) \\ &= \mathrm{P}\left( -U\left( \frac{1}{Z_1}, \dots, \frac{1}{Z_d} \right) \leq \mathbf{x} \right) \\ &= 1 - h\left(\mathbf{x}\right) \\ &= 1 - \left\|\mathbf{x}\right\|_D \end{aligned}$$

□

## 6.3. A bivariate GPD as a function of a copula

Let $(Z_1, Z_2)$ be distributed according to a copula $C$, i.e.

$$P(Z_1 \leq x, Z_2 \leq y) := C(x, y)$$

and independent of $U$ and $U$ is uniform distributed on $(0, 1)$.
Then we have for $x, y < 0$

$$\begin{aligned}
\mathrm{P}\Bigg( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \leq (x,y) \Bigg| U = u\Bigg) \\
&= \mathrm{P}\left(\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \geq (Z_1, Z_2)\right) \\
&= C\left(-\frac{u}{2x}, -\frac{u}{2y}\right)
\end{aligned}$$

Without loss of generalization let $y \leq x < 0$ ($\Leftrightarrow 0 < -\frac{u}{2y} \leq -\frac{u}{2x}$). We get

$$\begin{aligned}
&\mathrm{P}\Bigg( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \leq (x,y)\Bigg) \\
&= \int_0^1 \mathrm{P}\Bigg( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \leq (x,y) \Bigg| U = u\Bigg) \,\mathrm{d}\, u \\
&= \int_0^1 C\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \mathrm{d}\, u. \qquad (6.1)
\end{aligned}$$

Since the GPD is only defined close to the origin we first regard the case that $y \geq -\frac{1}{2}$ and obtain using the formula deduced in the proof of Theorem 6.2.1

$$\begin{aligned}
&\mathrm{P}\Bigg( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \leq (x,y)\Bigg) \\
&= 1 - \Bigg(-2y + 2x \int_0^1 C\left(\frac{1}{2}\left(-\frac{-2xu}{x}\right), \frac{1}{2}\left(-\frac{-2xu}{y}\right)\right) \mathrm{d}\, u \\
&\qquad - (2x - 2y) \int_0^1 C\left(\frac{1}{2}2, \frac{1}{2}\left(\frac{(2x-2y)\,u - 2x}{y}\right)\right) \mathrm{d}\, u\Bigg) \\
&= 1 - \left(-2y + 2x \int_0^1 C\left(u, u\frac{x}{y}\right) \mathrm{d}\, u + (2y - 2x) \int_0^1 \frac{(x-y)\,u - x}{y} \,\mathrm{d}\, u\right) \\
&= 1 - \left(-2y + 2x \int_0^1 C\left(u, u\frac{x}{y}\right) \mathrm{d}\, u + (2y - 2x) \frac{\frac{1}{2}(x-y) - x}{y} \,\mathrm{d}\, u\right)
\end{aligned}$$

$$= 1 - \left( -2y + 2x \int_0^1 C\left(u, u\frac{x}{y}\right) \mathrm{d}\, u + \frac{y^2 - x^2}{y} \mathrm{d}\, u \right)$$

$$= 1 - \left( 2x \int_0^1 C\left(u, u\frac{x}{y}\right) \mathrm{d}\, u - \frac{x^2 + y^2}{y} \right)$$

$$= 1 - \|x, y\|_D \,.$$

So we have obtained a representation of the GPD in dependence of the copula $C$. Next we want to look at the case that $y < -\frac{1}{2} \le x$. We have

$$\mathrm{P}\left( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \le (x, y) \right)$$

$$= \int_0^{-2x} C\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \mathrm{d}\, u + \int_{-2x}^{1} -\frac{u}{2y} \mathrm{d}\, u$$

$$= -2x \int_0^1 C\left(u, u\frac{x}{y}\right) \mathrm{d}\, u + \frac{x^2 - \frac{1}{4}}{y}.$$

Finally let $x < -\frac{1}{2}$ and we get

$$\mathrm{P}\left( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \le (x, y) \right)$$

$$= \int_0^1 C\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \mathrm{d}\, u.$$

Hence certain $D$-Norms can be written as a function of copula by

$$\|x, y\|_D = \begin{cases} -2\,|x| \int_0^1 C\left(u, u\left|\frac{x}{y}\right|\right) \mathrm{d}\, u + \frac{x^2+y^2}{|y|}, & 0 < |x| \le |y| \\ -2\,|y| \int_0^1 C\left(u\left|\frac{x}{y}\right|, u\right) \mathrm{d}\, u + \frac{x^2+y^2}{|x|}, & 0 < |y| < |x| \\ |x| + |y|\,, & x = 0 \text{ oder } y = 0 \end{cases} .$$

We start with an easy example namely the Clayton-Copula with parameter $\theta = 2$. The Clayton-Copula is a special case of the Archimedean Copula.

Example 6.3.1 (Example of a bivariate Clayton-Copula with parameter $\theta = 2$). We consider the Clayton-Copula $C$ with parameter 2, i.e. the generator of $C$ is

$\psi_2(x) = (1+2x)^{-\frac{1}{2}}$. It is $\psi_2^{-1}(x) = \frac{1}{2}\left(\frac{1}{x^2} - 1\right)$ and therefore the copula

$$\begin{aligned} C(u,v) &= \psi_2\left(\psi_2^{-1}(u) + \psi_2^{-1}(v)\right) \\ &= \left(\frac{1}{u^2} + \frac{1}{v^2} - 1\right)^{-\frac{1}{2}},\ u,v \in (0,1]. \end{aligned}$$

We have

$$\int C\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \mathrm{d}\, u = -\sqrt{4\left(x^2+y^2\right) - u^2}$$

and hence (note that $y < 0$)

$$\begin{aligned} \int_0^{-2x} C\left(-\frac{u}{2x}, -\frac{u}{2y}\right) \mathrm{d}\, u &= -\sqrt{4\left(x^2+y^2\right) - (-2x)^2} + \sqrt{4\left(x^2+y^2\right)} \\ &= 2y + 2\sqrt{x^2+y^2}. \end{aligned}$$

Altogether we obtain with the formula from above for $-\frac{1}{2} \le y \le x < 0$

$$\begin{aligned} &\mathrm{P}\left( \left(-\frac{U}{2Z_1}, -\frac{U}{2Z_2}\right) \le (x,y)\right) \\ &= 1 - \left(-\frac{x^2+3y^2}{y} - 2\sqrt{x^2+y^2}\right). \end{aligned}$$

Hence

$$\|x,y\|_D = \begin{cases} \frac{x^2+3y^2}{|y|} - 2\sqrt{x^2+y^2}, & \text{for } 0 < |x| \le |y| \\ \frac{y^2+3x^2}{|x|} - 2\sqrt{x^2+y^2}, & \text{for } \ |x| > |y| > 0 \\ |x| + |y|, & \text{for } \ |x| = 0 \text{ or } \ |y| = 0 \end{cases}.$$

### 6.4. A multivariate GPD as a function of a copula

Analog to the bivariate case let $(Z_1, \dots, Z_d)$ be distributed according to a copula $C$, i.e.

$$P(Z_1 \le x_1, \dots, Z_d \le x_d) := C(x_1, \dots, x_d)$$

and independent of $U$ and $U$ is uniform distributed on $(0,1)$.
The formula deduced in the proof of Theorem 6.2.1 then implies for $-\frac{1}{2} \le x_d \le \dots \le x_1 < x_0 := 0$

$$\mathrm{P}\left( \left(-\frac{U}{2Z_1}, \dots, -\frac{U}{2Z_d}\right) \le (x_1, \dots, x_d)\right)$$

$$
\begin{aligned}
&= 1 - \Bigg( -2x_d - \sum_{i=1}^{d} 2\,(x_{i-1} - x_i) \\
&\qquad \int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{(x_{i-1} - x_i)\,u - x_{i-1}}{x_j}\right)\right) \mathrm{d}\,u \Bigg) \\
&= 1 - \Bigg( -2x_d - 2\,(x_{d-1} - x_d) \int_0^1 -\frac{(x_{d-1} - x_d)\,u - x_{d-1}}{x_d}\,\mathrm{d}\,u \\
&\quad - \sum_{i=1}^{d-1} 2\,(x_{i-1} - x_i) \int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{(x_{i-1} - x_i)\,u - x_{i-1}}{x_j}\right)\right) \mathrm{d}\,u \Bigg) \\
&= 1 - \Bigg( -2x_d + 2\,(x_{d-1} - x_d)\, \frac{\frac{1}{2}\,(x_{d-1} - x_d) - x_{d-1}}{x_d} \\
&\quad - \sum_{i=1}^{d-1} 2\,(x_{i-1} - x_i) \int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{(x_{i-1} - x_i)\,u + x_{i-1}}{x_j}\right)\right) \mathrm{d}\,u \Bigg) \\
&= 1 - \Bigg( -2x_d - \frac{x_{d-1}^2 - x_d^2}{x_d} \\
&\quad - \sum_{i=1}^{d-1} 2\,(x_{i-1} - x_i) \int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{(x_{i-1} - x_i)\,u + x_{i-1}}{x_j}\right)\right) \mathrm{d}\,u \Bigg) \\
&= 1 - \Bigg( -\frac{x_{d-1}^2 + x_d^2}{x_d} \\
&\quad - \sum_{i=1}^{d-1} 2\,(x_{i-1} - x_i) \int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{(x_{i-1} - x_i)\,u + x_{i-1}}{x_j}\right)\right) \mathrm{d}\,u \Bigg).
\end{aligned}
$$

We obtain again a representation of a $D$-Norm that is deduced from a copula. Set $y_i := |x_i|$, $1 \le i \le d$, and $\delta := \max\{i : y_{i:d} = 0\}$. Then we have for $\delta < d - 1$

$$
\begin{aligned}
\|x_1, \dots, x_d\|_D &= +\frac{y_{d-1:d}^2 + y_{d:d}^2}{y_{d:d}} + \sum_{i=1}^{\delta-1} 2\,(y_{i-1:d} - y_{i:d}) \\
&\int_0^1 C\left(\sum_{j=1}^{i-1} \mathbf{e}_{(j),d} + \sum_{j=i}^{\delta} \mathbf{e}_{(j),d} \left(-\frac{(y_{i-1:d} - y_{i:d})\,u + y_{i-1:d}}{y_{j:d}}\right) + \sum_{j=\delta+1}^{d} \mathbf{e}_{(j),d}\right) \mathrm{d}\,u
\end{aligned}
$$

and for $\delta \geq d-1$

$\|x_1, \ldots, x_d\|_D = y_{d:d}$.

Example 6.4.1 (The Clayton-Copula with parameter 2 in arbitrary dimension). Once again we consider the Clayton-Copula with parameter 2 in dimension $d$. As above the generator is $\psi_2(x) = (1+2x)^{-\frac{1}{2}}$ and hence $\psi_2^{-1}(x) = \frac{1}{2}\left(\frac{1}{x^2} - 1\right)$ and therefore the copula

$$\begin{aligned} C(u_1, \ldots, u_d) &= \psi_2\left(\sum_{i=1}^{d} \psi_2^{-1}(u_i)\right) \\ &= \left(\sum_{i=1}^{d} \frac{1}{u_i^2} + 1 - d\right)^{-\frac{1}{2}}, \ u_1, \ldots, u_d \in (0,1]. \end{aligned}$$

We have

$$\int \left(\frac{\alpha}{u^2} + \beta\right)^{-\frac{1}{2}} \mathrm{d}\, u = \frac{1}{\beta}\sqrt{\alpha + \beta u^2}.$$

Using the formula from the preceding section we obtain (with $-\frac{1}{2} \leq x_d \leq \cdots \leq x_1 < x_0 := 0$)

$$\begin{aligned} &\mathrm{P}\left( \left(-\frac{U}{2Z_1}, \ldots, -\frac{U}{2Z_d}\right) \leq (x_1, \ldots, x_d)\right) \\ &= 1 - \left(-\sum_{i=1}^{d-1} \int_{-2x_{i-1}}^{-2x_i} C\left(\sum_{j=i}^{i-1} \mathbf{e}_j + \sum_{j=i}^{d} \mathbf{e}_j \left(-\frac{u}{2x_j}\right)\right) \mathrm{d}\, u - \frac{x_{d-1}^2 + x_d^2}{x_d}\right) \\ &= 1 - \left(-\sum_{i=1}^{d-1} \int_{-2x_{i-1}}^{-2x_i} \left(\frac{\sum_{j=i}^{d} 4x_j^2}{u^2} + (i-d)\right)^{-\frac{1}{2}} \mathrm{d}\, u - \frac{x_{d-1}^2 + x_d^2}{x_d}\right) \\ &= 1 - \left(-\sum_{i=1}^{d-1} \left[\frac{2}{i-d}\sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)\left(\frac{u}{2}\right)^2}\,\right]_{-2x_{i-1}}^{-2x_i} - \frac{x_{d-1}^2 + x_d^2}{x_d}\right) \\ &= 1 - \left(-\sum_{i=1}^{d-1} \left(\frac{2}{i-d}\left(\sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_i^2}\right.\right.\right. \\ &\qquad\qquad \left.\left.\left. - \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2}\right)\right) - \frac{x_{d-1}^2 + x_d^2}{x_d}\right) \end{aligned}$$

$$
\begin{aligned}
&= 1 - \left( - \sum_{i=1}^{d-1} \left( \frac{2}{i-d} \left( \sqrt{\sum_{j=i+1}^{d} x_j^2 + (1+i-d)x_i^2} \right.\right.\right. \\
&\qquad\qquad\qquad \left.\left.\left. - \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right) \right) - \frac{x_{d-1}^2 + x_d^2}{x_d} \right) \\
&= 1 - \left( - \sum_{i=2}^{d} \left( \frac{2}{i-1-d} \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right) \right. \\
&\qquad\qquad \left. + \sum_{i=1}^{d-1} \left( \frac{2}{i-d} \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right) - \frac{x_{d-1}^2 + x_d^2}{x_d} \right) \\
&= 1 - \left( +2x_d - \sum_{i=2}^{d-1} \frac{2}{i-1-d} \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right. \\
&\qquad \left. + \sum_{i=2}^{d-1} \frac{2}{i-d} \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} - \frac{2}{d-1} \sqrt{\sum_{i=1}^{d} x_i^2} - \frac{x_{d-1}^2 + x_d^2}{x_d} \right) \\
&= 1 - \left( - \frac{2}{d-1} \sqrt{\sum_{i=1}^{d} x_i^2} - 2x_d - \frac{x_{d-1}^2 + x_d^2}{x_d} \right. \\
&\qquad\qquad \left. + \sum_{i=2}^{d-1} \left( \frac{2}{i-d} - \frac{2}{i-1-d} \right) \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right) \\
&= 1 - \left( \frac{x_{d-1}^2 + 3x_d^2}{-x_d} - \frac{2}{d-1} \sqrt{\sum_{i=1}^{d} x_i^2} \right. \\
&\qquad\qquad \left. - \sum_{i=2}^{d-1} \frac{2}{(d-i)\,(d+1-i)} \sqrt{\sum_{j=i}^{d} x_j^2 + (i-d)x_{i-1}^2} \right)
\end{aligned}
$$

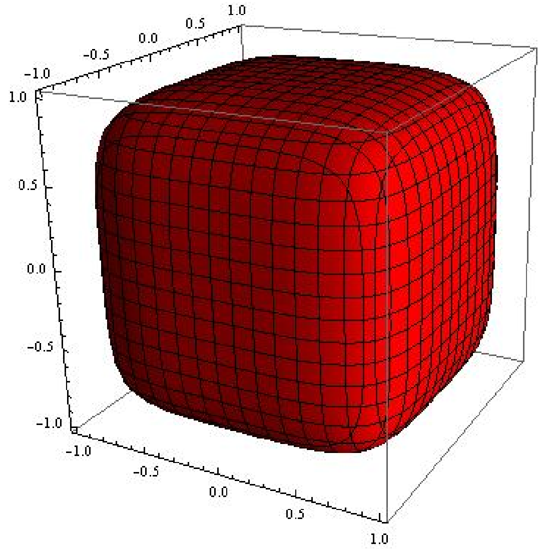


Figure 6.7. The unit sphere coming from the Clayton-Copula with $\theta = 2$ in dimension 3

Figure 6.7 shows the unit sphere of the norm of the GPD associated with the Clayton-Copula with $\theta = 2$ in dimension 3.

## 6.5. The GPD-Flow

In the preceding section we have seen that starting with a copula $C$ we obtain a multivariate GPD and thus a copula $C^{(1)}$, a so called GPD-Copula. This GPD-Copula $C^{(1)}$ can then be used to iterate the step and yields another GPD-Copula $C^{(2)}$. This procedure can be iterated over and over again. This is called a GPD-Flow.

Though a natural question arises: does the GPD-Flow converges and if it does, which is the limit copula?

The convergence of the GPD-Flow is still an open problem. But we will show that if it converges then it has to be to the copula of complete dependence, as already indicated by the simulations displayed in Figures 6.2 and 6.4.

Let $U_n$, $n \in \mathbb{N} \cup \{0\}$ be independent and uniform distributed random variables that are independent from the random vector $(Z_1, \dots, Z_d)$ which is distributed according a copula $C^{(0)}$. Then we define $\mathbf{V}_0 := \left(-\frac{U}{2Z_1}, \dots, -\frac{U}{2Z_d}\right)$ and

$W_0(x_1, \ldots, x_d) := \mathrm{P}(\mathbf{V}_0 \leq (x_1, \ldots, x_d))$.
The $j$-th marginal distribution of $\mathbf{V}_0$ is given by

$$\mathrm{P}\left(\mathbf{V}_0^T \mathbf{e}_j \leq x\right) = \begin{cases} 1 + x, & \text{if } -\frac{1}{2} \leq x \leq 0, \\ -\frac{1}{4x}, & \text{if } x < -\frac{1}{2}, \end{cases}$$
$$=: H(x), x \leq 0,$$

see the unpublished paper by Aulbach, Bayer and Falk ([**1**]) on page 6.
We have

$$H^{-1}(x) = \begin{cases} x - 1, & \text{if } x \in \left[\frac{1}{2}, 1\right] \\ -\frac{1}{4x}, & \text{if } x \in \left(0, \frac{1}{2}\right) \end{cases}$$

and observe that $H^{-1}$ is strictly increasing.
Furthermore we get

$$\begin{aligned} & C^{(n)}(u_1, \ldots, u_d) \\ & := \mathrm{P}\left(H\left(\mathbf{V}_{n-1}^T \mathbf{e}_1\right) \leq u_1, \ldots, H\left(\mathbf{V}_{n-1}^T \mathbf{e}_d\right) \leq u_d\right) \\ & = \mathrm{P}\left(\mathbf{V}_{n-1}^T \mathbf{e}_1 \leq H^{-1}(u), \ldots, \mathbf{V}_{n-1}^T \mathbf{e}_d \leq H^{-1}(u_d)\right) \\ & = W_{n-1}\left(H^{-1}(u_1), \ldots, H^{-1}(u_d)\right) \end{aligned}$$

and furthermore this simplifies in the bivariate case to

$$\begin{aligned} & C^{(n)}(u, v) \\ & = \begin{cases} W_{n-1}(u - 1, v - 1), & \text{if } u, v \in \left[\frac{1}{2}, 1\right], \\ W_{n-1}\left(u - 1, -\frac{1}{4v}\right), & \text{if } u \in \left[\frac{1}{2}, 1\right], v \in \left(0, \frac{1}{2}\right), \\ W_{n-1}\left(-\frac{1}{4u}, v - 1\right), & \text{if } u \in \left(0, \frac{1}{2}\right), v \in \left[\frac{1}{2}, 1\right], \\ W_{n-1}\left(-\frac{1}{4u}, -\frac{1}{4v}\right), & \text{if } u, v \in \left(0, \frac{1}{2}\right) \end{cases} \end{aligned}$$

$$= \begin{cases} 1 + \int\limits_0^{-2u+2} C^{(n-1)}\left(-\frac{w}{2u-2}, -\frac{w}{2v-2}\right) \mathrm{d}\, w + \frac{(u-1)^2+(v-1)^2}{v-1}, & \text{if } \frac{1}{2} \le v \le u \le 1 \\ 1 + \int\limits_0^{-2v+2} C^{(n-1)}\left(-\frac{w}{2u-2}, -\frac{w}{2v-2}\right) \mathrm{d}\, w + \frac{(u-1)^2+(v-1)^2}{u-1}, & \text{if } \frac{1}{2} \le u \le v \le 1 \\ \int\limits_0^{-2u+2} C^{(n-1)} \left(-\frac{w}{2u-2}, 2vw\right) \mathrm{d}\, w + v - 4v(u-1)^2, & \text{if } 0<v<\frac{1}{2}\le u \le 1 \\ \int\limits_0^{-2v+2} C^{(n-1)} \left(2uw, -\frac{w}{2v-2}\right) \mathrm{d}\, w + u - 4v(v-1)^2, & \text{if } 0<u<\frac{1}{2}\le v \le 1 \\ \int\limits_0^{1} C^{(n-1)} \left(2uw, 2vw\right) \mathrm{d}\, w, & \text{if } u, v \in \left(0, \frac{1}{2}\right) \end{cases} \tag{6.2}$$

where

$$\mathbf{V}_n := \left(-\frac{U_n}{2\mathbf{V}_{n-1}^T \mathbf{e}_1}, \ldots, -\frac{U_n}{2\mathbf{V}_{n-1}^T \mathbf{e}_d}\right)$$

and

$$W_n\left(x_1, \ldots, x_d\right) := \mathrm{P}\left(\mathbf{V}_n^T \mathbf{e}_1 \le x_1, \ldots, \mathbf{V}_n^T \mathbf{e}_d \le x_d\right).$$

Since $H$ is the distribution function of the margins, $C^{(n)}$ is a copula.
Finally by $\|x_1, \ldots, , x_d\|_n$ we denote the corresponding norm to the GPD of $\mathbf{V}_n$, i.e. for small $x_1, \ldots, x_d < 0$ it is $W_n\left(x_1, \ldots, x_d\right) =: 1 - \|(x_1, \ldots, x_d)\|_n$.

THEOREM 6.5.1. *There exists no bivariate copula different from the complete dependence copula that remains fixed under the iteration step as stated above.*
*Or equivalent:*
*If $C$ is a bivariate copula that remains fixed under the iteration step then $C$ is the copula of complete dependence.*

PROOF. Setting $u = v$ in equation (6.2) yields

$$C^{(1)}(u, u) = \begin{cases} -1 + \int\limits_0^{-2u+2} C^{(0)}\left(-\frac{v}{2u-2}, -\frac{v}{2u-2}\right) \mathrm{d}\, v + 2u, & \text{if } u \in \left[\frac{1}{2}, 1\right], \\ \int\limits_0^{1} C^{(0)}\left(2uv, 2uv\right) \mathrm{d}\, v, & \text{if } u \in \left(0, \frac{1}{2}\right) \end{cases}$$

and then using integration by substitution we get

$$= \begin{cases} -1 + 2u + 2(1-u) \int\limits_0^1 C^{(0)}(v,v)\,\mathrm{d}\,v, & \text{if } u \in \left[\frac{1}{2}, 1\right], \\ \frac{1}{2u} \int\limits_0^{2u} C^{(0)}(v,v)\,\mathrm{d}\,v, & \text{if } u \in \left(0, \frac{1}{2}\right) \end{cases}. \tag{6.3}$$

So we have for all $u \in \left[\frac{1}{2}, 1\right]$

$$C^{(1)}(u,u) = 2\left(1 - \int\limits_0^1 C^{(0)}(v,v)\,\mathrm{d}\,v\right) u + 2\int\limits_0^1 C^{(0)}(v,v)\,dv - 1.$$

Consider the special case that $C^{(1)} = C^{(0)} =: C$ and set $\alpha := \int\limits_0^1 C^{(0)}(v,v)\,\mathrm{d}\,v$. We will prove that for $u \in (0,1]$ it is $C(u,u) = m \cdot u + t$. By induction over $n \in \mathbb{N}_0$ we will show that $C(u,u) = m \cdot u + t$ for $u \in [2^{-1-n}, 2^{-n}]$. The initial step of the induction is already proven above with $m := 2(1-\alpha)$ and $t := 2\alpha - 1$.
Now assume that we have already shown that $C(u,u) = m \cdot u + t$ for $u \in [2^{-n-1}, 2^{-n}]$.
Take $u \in [2^{-n-2}, 2^{-n-1}]$. From formula (6.3) we obtain

$$2^{-n} C\left(2^{-n-1}, 2^{-n-1}\right) = 2^{-n} \frac{1}{2^{-n}} \int\limits_0^{2 \cdot 2^{-n-1}} C(v,v)\,dv = \int\limits_0^{2^{-n}} C(v,v)\,\mathrm{d}\,v$$

and therefore

$$\begin{aligned} C(u,u) &= \frac{1}{2u} \int\limits_0^{2u} C(v,v)\,\mathrm{d}\,v \\ &= \frac{1}{2u} \left( \int\limits_0^{2^{-n}} C(v,v)\,\mathrm{d}\,v - \int\limits_{2u}^{2^{-n}} C(v,v)\,\mathrm{d}\,v \right) \\ &= \frac{1}{2u} \left( 2^{-n} C\left(2^{-n-1}, 2^{-n-1}\right) - \left[\frac{1}{2} m v^2 + t v\right]_{2u}^{2^{-n}} \right) \\ &= \frac{1}{2u} \left( 2^{-n} C\left(2^{-n-1}, 2^{-n-1}\right) - 2^{-n} \underbrace{\left(m \cdot 2^{-n-1} + t\right)}_{= C(2^{-n-1}, 2^{-n-1})} + 2u\,(m \cdot u + t) \right) \\ &= \frac{1}{2u} \left( 2^{-n} C\left(2^{-n-1}, 2^{-n-1}\right) - 2^{-n} C\left(2^{-n-1}, 2^{-n-1}\right) + 2u\,(m \cdot u + t) \right) \end{aligned}$$

$$= m \cdot u + t.$$

Since $C(u,u)$ is continuous from the right we even have $C(u,u) = m \cdot u + t$ for $u \in [0,1]$. Furthermore $C(0,0) = 0$ implies $t = 0$ and $C(1,1) = 1$ implies $m = 1$. So we have $C(u,u) = u$.
Therefore $C$ is the copula of complete dependence (see Example 3.17 (a) in Nelsen (2006, [**31**])).

□

Corollary 6.5.2. *In dimension 2 or higher the only copula that remains fix under the GPD-Flow is the copula of complete dependence.*

Proof. If $C$ is a copula that remains fix under the GPD-Flow then according to Theorem 6.5.1 all bivariate margins are complete dependent and therefore $C$ is the copula of complete dependence.

□

Next we will see that the GPD-Flow is continuous.

Lemma 6.5.3. *Let $M_n$ be the set of all copulas of dimension $n$. We define a metric $d$ on $M_n$ by*

$$d(C_1, C_2) := \sup_{\mathbf{x} \in \mathbb{R}^n} |C_1(\mathbf{x}) - C_2(\mathbf{x})|, \qquad C_1, C_2 \in M_n$$

*By $f : M_n \mapsto M_n$ we denote the function that represents the iteration step of the GPD-Flow, i.e. $C^{(i)} = f\left(C^{(i-1)}\right)$ where $C^{(i)}$ is the copula from the i-th iteration of the GPD-Flow.*

*Then $f$ is continuous.*

Proof. We have for $C \in M_n$, $(x_1, \dots, x_n) \in \mathbb{R}^n$ and $H^{-1}$ as defined on page 100

$$f(C(x_1, \dots, x_n)) = \int_0^1 C\left(-\frac{u}{2H^{-1}(x_1)}, \dots, -\frac{u}{2H^{-1}(x_n)}\right) \mathrm{d}\, u.$$

For all $C \in M_n$ and all $\epsilon > 0$ and $\delta := \epsilon$ we have for all $D \in M_n$ that satisfy $d(C,D) < \delta$

$$d(f(C), f(D)) = \sup_{(x_1, \dots, x_n) \in \mathbb{R}^n} \left| \int_0^1 C\left(-\frac{u}{2H^{-1}(x_1)}, \dots, -\frac{u}{2H^{-1}(x_n)}\right) \mathrm{d}\, u \right.$$

$$
\begin{aligned}
&\qquad - \int_0^1 D\left(-\frac{u}{2H^{-1}(x_1)}, \dots, -\frac{u}{2H^{-1}(x_n)}\right) \mathrm{d}\, u\Bigg| \\
&\leq \sup_{(x_1,\dots,x_n)\in\mathbb{R}^n} \int_0^1 \Bigg| C\left(-\frac{u}{2H^{-1}(x_1)}, \dots, -\frac{u}{2H^{-1}(x_n)}\right) \\
&\qquad\qquad - D\left(-\frac{u}{2H^{-1}(x_1)}, \dots, -\frac{u}{2H^{-1}(x_n)}\right)\Bigg| \mathrm{d}\, u \\
&< \delta \\
&= \epsilon.
\end{aligned}
$$

Thus $f$ is continuous.

□

THEOREM 6.5.4. *If the GPD-Flow converges to a copula then this is the copula of complete dependence.*

PROOF. By $C^{(i)} \in M_n$ we denote the copula of the $i$-th iteration of the GPD-Flow and by $f$ we denote the function of the iteration step of the GPD-Flow. Then we have

$$
C^{(i+1)} = f\left(C^{(i)}\right), \qquad i \in \mathbb{N}.
$$

If $C^{(i)}$ converges to a copula $C$, i.e. $C = \lim_{i\to\infty} C^{(i)}$ then we obtain using the continuity of $f$ (see Lemma 6.5.3):

$$
\begin{aligned}
f(C) &= f\left(\lim_{i\to\infty} C^{(i)}\right) \\
&= \lim_{i\to\infty} f\left(C^{(i)}\right) \\
&= \lim_{i\to\infty} C^{(i+1)} \\
&= C.
\end{aligned}
$$

Hence $C$ remains fix under the iteration step of the GPD-Flow and according to Corollary 6.5.2 $C$ is the copula of complete dependence.

□

Simulations (see Figure 6.2 and Figure 6.4) indicate that for every $n \in \mathbb{N}_0$ we have $\|\mathbf{x}\|_n \geq \|\mathbf{x}\|_{n+1}$, $\mathbf{x} \geq \mathbf{0}$, i.e. the unit sphere of the norm of the $n$-th iteration is completely contained in the unit sphere of the norm of the $(n+1)$-th iteration. But this is actually not true as we will see in the example below.
We will construct the copula leading to the corresponding norms using the so called Diagonal Copula introduced by Fredericks and Nelsen (1997, [**17**]).

Definition 6.5.5. *A function $\delta : [0,1] \mapsto [0,1]$ will be called a* <u>*diagonal*</u> *if it satisfies*

(i) $\delta(1) = 1$,
(ii) $\delta(t) \leq t$ *for all* $t \in [0,1]$ *and*
(iii) $0 \leq \delta(t_2) - \delta(t_1) \leq 2(t_2 - t_1)$ *for all* $t_1, t_2 \in [0,1]$ *with* $t_1 \leq t_2$.

Definition and Theorem 6.5.6. *Let $\delta$ be any diagonal and set*

$$C(u,v) := \min\left(u, v, \frac{1}{2}(\delta(u) + \delta(v))\right).$$

*Then $C$ is a copula whose diagonal section is $\delta$, i.e. $C(t,t) = \delta(t)$.*
*Copulas that can be written in this form are called* <u>*diagonal copulas*</u>.

Proof. For the proof we refer to Fredericks and Nelsen (1997, [**17**]).

□

Example 6.5.7. Now we are ready to construct the copula as mentioned above. Define for $\tau \in \left[0, \frac{1}{2}\right]$ the function $\delta_\tau : [0,1] \to [0,1]$ by

$$\delta_\tau(t) := \begin{cases} 0, & t \in [0,\tau) \\ 2(t-\tau), & t \in [\tau, 2\tau) \\ t, & t \in [2\tau, 1] \end{cases}.$$

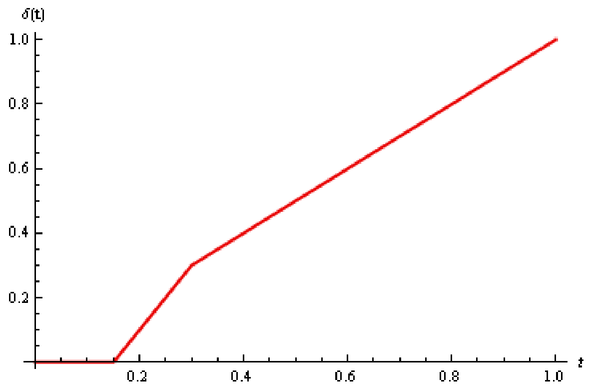



FIGURE 6.8. The diagonal function $\delta_{\frac{3}{20}}$

Obviously $\delta_\tau$ is for $\tau \in \left[0, \frac{1}{2}\right]$ a diagonal and therefore

$$C^{(0)}(u,v) := \min\left(u, v, \frac{1}{2}\left(\delta_{\frac{3}{20}}(u) + \delta_{\frac{3}{20}}(v)\right)\right)$$

is a Diagonal Copula.
We have

$$\begin{aligned}
&\int_0^1 \delta_\tau(t)\,\mathrm{d}\,t \\
&= \int_0^\tau \delta_\tau(t)\,\mathrm{d}\,t + \int_\tau^{2\tau} \delta_\tau(t)\,\mathrm{d}\,t + \int_{2\tau}^1 \delta_\tau(t)\,\mathrm{d}\,t \\
&= 0 + \tau^2 + \frac{1}{2} - 2\tau^2 \\
&= \frac{1}{2} - \tau^2
\end{aligned}$$

and furthermore

$$\begin{aligned}
&\int_0^1\int_0^1 \delta_\tau(uv)\,\mathrm{d}\,v\,\mathrm{d}\,u \\
&= \int_0^\tau\int_0^1 \underbrace{\delta_\tau(uv)}_{=0}\,\mathrm{d}\,v\,\mathrm{d}\,u + \int_\tau^1\int_0^{\frac{\tau}{u}} \underbrace{\delta_\tau(uv)}_{=0}\,\mathrm{d}\,v\,\mathrm{d}\,u + \int_\tau^{2\tau}\int_{\frac{\tau}{u}}^1 \underbrace{\delta_\tau(uv)}_{=2(uv-\tau)}\,\mathrm{d}\,v\,\mathrm{d}\,u \\
&\quad + \int_{2\tau}^1\int_{\frac{\tau}{u}}^{\frac{2\tau}{u}} \underbrace{\delta_\tau(uv)}_{=2(uv-\tau)}\,dvdu + \int_{2\tau}^1\int_{\frac{2\tau}{u}}^1 \underbrace{\delta_\tau(uv)}_{=uv}\,\mathrm{d}\,v\,\mathrm{d}\,u \\
&= 0 + 0 + \frac{1}{2}\tau^2(2\log 2 - 1) - \tau^2\log(2\tau) + \frac{1}{4} - \tau^2 + 2\tau^2\log(2\tau)
\end{aligned}$$

$$= \frac{1}{4} + \tau^2 \left( 2\log 2 - \frac{3}{2} \right) + \tau^2 \log \tau.$$

Using integration by substitution we have for $x \in \left[-\frac{1}{2}, 0\right]$

$$\begin{aligned}
W_1(x,x) &= (-2x) \int_0^1 C^{(1)}(u,u)\,\mathrm{d}\,u + 2x + 1 \\
&= (-2x) \left( \int_0^{\frac{1}{2}} C^{(1)}(u,u)\,\mathrm{d}\,u + \int_{\frac{1}{2}}^1 C^{(1)}(u,u)\,du \right) + 2x + 1 \\
&= (-2x) \left( \int_0^{\frac{1}{2}} \int_0^1 C^{(0)}(2uv, 2uv)\,\mathrm{d}\,v\,\mathrm{d}\,u \right. \\
&\quad \left. + \int_{\frac{1}{2}}^1 \left( \int_0^{-2u+2} C^{(0)} \left( -\frac{v}{2u-2}, -\frac{v}{2u-2} \right) dv + 2u - 1 \right) \mathrm{d}\,u \right) + 2x + 1 \\
&= (-2x) \left( \frac{1}{2} \int_0^1 \int_0^1 C^{(0)}(uv, uv)\,\mathrm{d}\,v\,\mathrm{d}\,u \right. \\
&\quad \left. + \int_{\frac{1}{2}}^1 (-2u+2) \int_0^1 C^{(0)}(v,v)\,\mathrm{d}\,v\,\mathrm{d}\,u + \left[u^2 - u\right]_{\frac{1}{2}}^1 \right) + 2x + 1 \\
&= (-2x) \left( \frac{1}{2} \int_0^1 \int_0^1 C^{(0)}(uv, uv)\,\mathrm{d}\,v\,\mathrm{d}\,u \right. \\
&\quad \left. + \int_0^1 C^{(0)}(v,v)\,\mathrm{d}\,v \int_{\frac{1}{2}}^1 (-2u+2)\,\mathrm{d}\,u + \frac{1}{4} \right) + 2x + 1 \\
&= 1 + (-2x) \left( \frac{1}{2} \int_0^1 \int_0^1 C^{(0)}(uv, uv)\,\mathrm{d}\,v\,\mathrm{d}\,u + \frac{1}{4} \int_0^1 C^{(0)}(u,u)\,\mathrm{d}\,u - \frac{3}{4} \right).
\end{aligned}$$

Using the formula from the first section and again integration by substitution we obtain still with $x \in \left[-\frac{1}{2}, 0\right)$

$$\|x,x\|_0 = (-2x)\left(1 - \int\limits_0^1 C^{(0)}(u,u)\,\mathrm{d}\,u\right)$$

and from the considerations above we obtain

$$\begin{aligned}\|x,x\|_1 &= (-2x)\left(\frac{3}{4} - \frac{1}{2}\int\limits_0^1\int\limits_0^1 C^{(0)}(uv,uv)\,\mathrm{d}\,v\,\mathrm{d}\,u - \frac{1}{4}\int\limits_0^1 C^{(0)}(u,u)\,\mathrm{d}\,u\right)\\ &= \|x,x\|_0\\ &\qquad + (-2x)\left(-\frac{1}{4} - \frac{1}{2}\int\limits_0^1\int\limits_0^1 C^{(0)}(uv,uv)\,\mathrm{d}\,v\,\mathrm{d}\,u + \frac{3}{4}\int\limits_0^1 C^{(0)}(u,u)\,\mathrm{d}\,u\right)\end{aligned}$$

Furthermore we have for the Diagonal Copula $C^{(0)}$

$$\begin{aligned}&-1 - 2\int_0^1\int_0^1 C^{(0)}(uv,uv)\,\mathrm{d}\,v\,\mathrm{d}\,u + 3\int_0^1 C^{(0)}(u,u)\,\mathrm{d}\,u\\ &\quad= -1 - 2\int_0^1\int_0^1 \delta_{\frac{3}{20}}(uv)\,\mathrm{d}\,v\,\mathrm{d}\,u + 3\int_0^1 \delta_{\frac{3}{20}}(u)\,\mathrm{d}\,u\\ &\quad= -1 - \frac{1}{2} - 2\tau^2\left(2\log 2 - \frac{3}{2}\right) - 2\tau^2\log\tau + 3\frac{1}{2} - 3\tau^2\\ &\quad= -2\tau^2\log(4\tau)\\ &\quad= -2\left(\frac{3}{20}\right)^2\underbrace{\log\left(\frac{3}{5}\right)}_{<0}\\ &\quad\approx 0.02299\\ &\quad> 0\end{aligned}$$

and thus

$$\|x,x\|_1 > \|x,x\|_0\,.$$

CHAPTER 7

# Simulation via the Shi Transformation

The need for realizations of random variables of a certain distribution arises in several fields. Often technical systems are so complex that they couldn't be described in a deterministic way. A common approach to cope with this problem is to use Monte-Carlo simulations to get a statistical distribution of the behavior of the system. Input parameters are distributed according to a certain distribution so there is the need for the realization of random variables from those distributions.
Another field with the need for realizations of random variables is testing the behavior of statistics methods. There is practically never exact knowledge of the underlying distribution function of "real data". In order to check the behavior of a method it is therefore useful to test it with data you know exactly, like the underlying distribution function or whether there is independence.

In his PhD thesis Michel (2006, [**26**]) established an algorithm that creates random vectors that follow a GPD of logistic type using the so called Shi Transformation which was first introduced by Shi (1995, [**37**]).
In this chapter we generalize this Transformation to the nested logistic model of arbitrary dimension $d \geq 2$. However things get more complicated in this case. Therefore we are able to deduce an algorithm only in dimension 3, which creates random vectors from a GPD from the nested logistic model.

Definition 7.1. *For $\lambda_1, \dots, \lambda_{d-1} \geq 1$ we call the transformation $ST_{\lambda_1,\dots,\lambda_{d-1}}$ : $(0, \infty) \times \left(0, \frac{\pi}{2}\right)^{d-1} \to (-\infty, 0)^d$, with*

$$(c, \psi_1, \dots, \psi_{d-1}) \mapsto$$
$$-c\left(\prod_{i=1}^{d-1} \sin^{\frac{2}{\lambda_i}} \psi_i, \cos^{\frac{2}{\lambda_1}} \psi_1 \prod_{i=2}^{d-1} \sin^{\frac{2}{\lambda_i}} \psi_i, \cos^{\frac{2}{\lambda_2}} \psi_2 \prod_{i=3}^{d-1} \sin^{\frac{2}{\lambda_i}} \psi_i, \dots, \cos^{\frac{2}{\lambda_{d-1}}} \psi_{d-1}\right)$$

*the Shi transformation for the nested logistic model. This transformation is one-to-one and infinitely often differentiable.*
*Let $(x_1, \ldots, x_d) \in (-\infty, 0)^d$. The components of the vector*

$$(c, \psi_1, \ldots, \psi_{d-1}) := ST^{-1}_{\lambda_1,\ldots,\lambda_{d-1}} (x_1, \ldots, x_d)$$

*are the Shi coordinates of $(x_1, \ldots, x_d)$. $c$ is called the radial component and $\psi := (\psi_1, \ldots, \psi_{d-1})$ is called the angular component.*
*By $(C, \Psi_1, \ldots, \Psi_{d-1}) = ST^{-1}_{\lambda_1,\ldots,\ldots,\lambda_{d-1}} (X_1, \ldots, X_d)$ we denote the Shi coordinates of a random vector $(X_1, \ldots, X_d) \in (-\infty, 0)^d$.*

Remark 7.2. The Shi transformation can also be defined in a recursive way:

$$ST_{\lambda_1,\ldots,\lambda_k} (c, \psi_1, \ldots, \psi_k)$$

$$:= -c \left( -\sin^{\frac{2}{\lambda_k}} \psi_k \mathbf{A}_k ST_{\lambda_1,\ldots,\lambda_{k-1}} (1, \psi_1, \ldots, \psi_{k-1}) + e_{k+1} \cos^{\frac{2}{\lambda_k}} \psi_k \right)$$

and

$$ST_{\lambda_1} (c, \psi_1) := -c \left( \sin^{\frac{2}{\lambda_1}} \psi_1, \cos^{\frac{2}{\lambda_1}} \psi_1 \right)$$

where

$$\mathbf{A}_k := \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 \\ 0 & \ldots & \ldots & 0 \end{pmatrix} \in \mathbb{R}^{k+1 \times k}.$$

To simplify the notation we set

$$\mathbf{z}_k := -ST_{\lambda_1,\ldots,\lambda_{k-1}} (1, \psi_1, \ldots, \psi_{k-1})$$

and with it we can abbreviate the recursion from above

$$ST_{\lambda_1,\ldots,\lambda_k} (c, \psi_1, \ldots, \psi_k) := -c \left( \sin^{\frac{2}{\lambda_k}} \psi_k \mathbf{A}_k \mathbf{z}_k + e_{k+1} \cos^{\frac{2}{\lambda_k}} \psi_k \right).$$

where $\mathbf{A}_k$ is defined as above.

LEMMA 7.3. *For* $(c, \psi_1, \ldots, \psi_{d-1}) \in (0,\infty) \times \left(0, \frac{\pi}{2}\right)^{d-1} \to (-\infty, 0)^d$ *we have*

$$\left\| ST_{\lambda_1,\ldots,\lambda_{d-1}} (c, \psi_1, \ldots, \psi_{d-1}) \right\|_{\lambda_1,\ldots,\lambda_{d-1}} = c.$$

PROOF. The homogeneity of the norm implies

$$\left\| ST_{\lambda_1,\ldots,\lambda_{d-1}} (c, \psi_1, \ldots, \psi_{d-1}) \right\|_{\lambda_1,\ldots,\lambda_{d-1}} = c \left\| ST_{\lambda_1,\ldots,\lambda_{d-1}} (1, \psi_1, \ldots, \psi_{d-1}) \right\|_{\lambda_1,\ldots,\lambda_{d-1}},$$

so we only have to proof the assertion for $c = 1$.
We define $\mathbf{z}_k$ as in Remark 7.2 and in terms of $\mathbf{z}_k$ the above statement can be written as $\|\mathbf{z}_d\|_{\lambda_1,\ldots,\lambda_{d-1}} = 1$ and we will prove this by induction over $k$.
For $k = 2$ we have

$$\begin{aligned}
\|\mathbf{z}_2\|_{\lambda_1} &= \left( \left( \sin^{\frac{2}{\lambda_1}} \psi_1 \right)^{\lambda_1} + \left( \cos^{\frac{2}{\lambda_1}} \psi_1 \right)^{\lambda_1} \right)^{\frac{1}{\lambda_1}} \\
&= \left( \sin^2 \psi_1 + \cos^2 \psi_1 \right)^{\frac{1}{\lambda_1}} \\
&= 1.
\end{aligned}$$

Now assume that we have proven the statement for $k - 1 < d$, i.e.
$\|\mathbf{z}_{k-1}\|_{\lambda_1,\ldots,\lambda_{k-2}} = 1$. We have

$$\begin{aligned}
\|\mathbf{z}_k\|_{\lambda_1,\ldots,\lambda_{k-1}} &= \left\| \left( \left\| -\sin^{\frac{2}{\lambda_{k-1}}} \psi_k \mathbf{z}_{k-1} \right\|_{\lambda_1,\ldots,\lambda_{k-2}}, \cos^{\frac{2}{\lambda_{k-1}}} \psi_{k-1} \right) \right\|_{\lambda_{k-1}} \\
&= \left\| \left( \sin^{\frac{2}{\lambda_{k-1}}} \psi_{k-1} \underbrace{\|\mathbf{z}_{k-1}\|_{\lambda_1,\ldots,\lambda_{k-2}}}_{\substack{=1 \\ \text{by the induction hypothesis}}}, \cos^{\frac{2}{\lambda_{k-1}}} \psi_{k-1} \right) \right\|_{\lambda_{k-1}} \\
&= \left\| \left( \sin^{\frac{2}{\lambda_{k-1}}} \psi_{k-1}, \cos^{\frac{2}{\lambda_{k-1}}} \psi_{k-1} \right) \right\|_{\lambda_{k-1}} \\
&= 1
\end{aligned}$$

and thus the proof is finished.

□

LEMMA 7.4. *As in Lemma 3.1.7 of Michel (2006,* [**26**]*) we put* $\gamma_i := \cos\phi_i$, $\sigma_i := \sin\phi_i$ *and* $\alpha_i := \frac{2}{\lambda_i}$ *for* $i = 1, \ldots, d-1$*. Then the Jacobian matrix of the transformation* $ST_{\lambda_1,\ldots,\lambda_{d-1}}$ *has the following form*

$$A_{1,\ldots,d} := J_{ST_{\lambda_1,\ldots,\lambda_{d-1}}} (c, \psi_1, \ldots, \psi_{d-1}) =$$

$$
= -c \begin{pmatrix}
\frac{1}{c}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \alpha_1\gamma_1\sigma_1^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \dots\dots & \alpha_{d-1}\gamma_{d-1}\sigma_{d-1}^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} \\
\frac{1}{c}\gamma_1^{\alpha_1}\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & -\alpha_1\gamma_1^{\alpha_1-1}\sigma_1\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & \ddots & \alpha_{d-1}\gamma_1^{\alpha_1}\gamma_{d-1}\sigma_{d-1}^{-1}\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} \\
\vdots & 0 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \alpha_{d-1}\gamma_{d-2}^{\alpha_{d-2}}\gamma_{d-1}\sigma_{d-1}^{\alpha_{d-1}-1} \\
\frac{1}{c}\gamma_{d-1}^{\alpha_{d-1}} & 0 & \dots\ 0 & -\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}
\end{pmatrix}.
$$

Proof. We set $\gamma_0 := 1$, $\alpha_0 := 1$ and with $\Pi_i$ we denote the projection to the $i$-th element, i.e. $\Pi_i : \mathbb{R}^d \to \mathbb{R}$ with $(x_1, \dots, x_d) \mapsto x_i$. The first column of the Jacobian matrix is given by

$$
\begin{aligned}
\frac{\partial \Pi_i \left(ST_{\lambda_1,\dots,\lambda_{d-1}}(c, \psi_1, \dots, \psi_{d-1})\right)}{\partial c} &= \frac{\partial}{\partial c}\left(-c\gamma_{i-1}^{\alpha_{i-1}}\prod_{k=i}^{d-1}\sigma_k^{\alpha_k}\right) \\
&= -\gamma_{i-1}^{\alpha_{i-1}}\prod_{k=i}^{d-1}\sigma_k^{\alpha_k}
\end{aligned}
$$

for $i \in \{1, \dots, d\}$.
For the other elements of the Jacobian matrix we differ whether $i = j$, $i < j$ or $i > j$ where $i$ denotes the number of the row and $j$ the number of the column of the element.
In the case of $i > j$ we have

$$
\begin{aligned}
\frac{\partial \Pi_i \left(ST_{\lambda_1,\dots,\lambda_{d-1}}(c, \psi_1, \dots, \psi_{d-1})\right)}{\partial \psi_{j-1}} &= \frac{\partial}{\partial \psi_{j-1}}\left(-c\gamma_{i-1}^{\alpha_{i-1}}\prod_{k=i}^{d-1}\sigma_k^{\alpha_k}\right) \\
&= 0,
\end{aligned}
$$

if $i = j$ it is

$$
\begin{aligned}
\frac{\partial \Pi_i \left(ST_{\lambda_1,\dots,\lambda_{d-1}}(c, \psi_1, \dots, \psi_{d-1})\right)}{\partial \psi_{j-1}} &= \frac{\partial}{\partial \psi_{i-1}}\left(-c\gamma_{i-1}^{\alpha_{i-1}}\prod_{k=i}^{d-1}\sigma_k^{\alpha_k}\right) \\
&= (-c)\left(-\alpha_{i-1}\gamma_{i-1}^{\alpha_{i-1}-1}\sigma_{i-1}\prod_{k=i}^{d-1}\sigma_k^{\alpha_k}\right)
\end{aligned}
$$

and finally for $i < j$

$$\frac{\partial \Pi_i \left( ST_{\lambda_1,\ldots,\lambda_{d-1}} \left( c, \psi_1, \ldots, \psi_{d-1} \right) \right)}{\partial \psi_{j-1}} = \frac{\partial}{\partial \psi_{j-1}} \left( -c \gamma_{i-1}^{\alpha_{i-1}} \prod_{k=i}^{d-1} \sigma_k^{\alpha_k} \right)$$
$$= -c \alpha_{j-1} \gamma_{i-1}^{\alpha_{i-1}} \gamma_{j-1} \sigma_{j-1}^{-1} \prod_{k=i}^{d-1} \sigma_k^{\alpha_k}.$$

So we have calculated all entries of the Jacobian matrix as given in the lemma.

□

LEMMA 7.5. *With the abbreviation of Lemma 7.4 we have*

$$\det \left( A_{1,\ldots,d} \right) = (-1)\, c^{d-1} \prod_{i=1}^{d-1} \alpha_i \sigma_i^{i\alpha_i - 1} \gamma_i^{\alpha_i - 1}.$$

PROOF. We will prove this by induction. For $d = 2$ we have

$$A_{1,2} = -c \begin{pmatrix} \frac{1}{c} \sigma_1^{\alpha_1} & \alpha_1 \gamma_1 \sigma_1^{\alpha_1 - 1} \\ \frac{1}{c} \gamma_1^{\alpha_1} & -\alpha_1 \gamma_1^{\alpha_1 - 1} \sigma_1 \end{pmatrix}$$

and thus

$$\begin{aligned} \det \left( A_{1,2} \right) &= c^2 \left( -\frac{1}{c} \sigma_1^{\alpha_1} \alpha_1 \gamma_1^{\alpha_1 - 1} \sigma_1 - \frac{1}{c} \gamma_1^{\alpha_1} \alpha_1 \gamma_1 \sigma_1^{\alpha_1 - 1} \right) \\ &= -c \alpha_1 \sigma_1^{\alpha_1 - 1} \gamma_1^{\alpha_1 - 1} \left( \sigma_1^2 + \gamma_1^2 \right) \\ &= -c \alpha_1 \sigma_1^{\alpha_1 - 1} \gamma_1^{\alpha_1 - 1} \\ &= (-1)\, c^{2-1} \prod_{i=1}^{2-1} \alpha_i \sigma_i^{i\alpha_i - 1} \gamma_i^{\alpha_i - 1}. \end{aligned}$$

Now assume that the assertion holds for $d - 1$. We use the rules of determinant calculations (see for Example Section 4.2. in Fraleigh and Beauregard (1987, [**16**])). In the first step we expand by the last row and the next step we transform the matrices that the induction assumptions can be used. Thereby in the first determinant we interchange the last column with the preceding one and then the former last column which is now the second last with its preceding column and continue until we reached the first column. Furthermore we also use the scalar-multiplication property of a determinant on both determinants in several ways.

$$
\begin{aligned}
&\det\left(A_{1,\dots,d}\right) = \\
&=(-1)^d c^d \det\begin{pmatrix}
\frac{1}{c}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \alpha_1\gamma_1\sigma_1^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \dots & & \alpha_{d-1}\gamma_{d-1}\sigma_{d-1}^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} \\
\frac{1}{c}\gamma_1^{\alpha_1}\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & -\alpha_1\gamma_1^{\alpha_1-1}\sigma_1\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & \ddots & & \vdots \\
\vdots & 0 & \ddots & & \vdots \\
\vdots & \vdots & \ddots & & \alpha_{d-1}\gamma_{d-2}^{\alpha_{d-2}}\gamma_{d-1}\sigma_{d-1}^{\alpha_{d-1}-1} \\
\frac{1}{c}\gamma_{d-1}^{\alpha_{d-1}} & 0 & \dots & 0 & -\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}
\end{pmatrix} \\
&= (-1)^d c^d \Bigg( (-1)^{d+1}\frac{1}{c}\gamma_{d-1}^{\alpha_{d-1}} c^{d-1} \\
&\quad \det\begin{pmatrix}
\alpha_1\gamma_1\sigma_1^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \dots\dots & \dots & \alpha_{d-1}\gamma_{d-1}\sigma_{d-1}^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} \\
-\alpha_1\gamma_1^{\alpha_1-1}\sigma_1\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & \ddots & & \vdots \\
0 & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \dots\ 0 & -\alpha_{d-2}\gamma_{d-2}^{\alpha_{d-2}-1}\sigma_{d-2}\sigma_{d-1}^{\alpha_{d-1}} & \alpha_{d-1}\gamma_{d-2}^{\alpha_{d-2}}\gamma_{d-1}\sigma_{d-1}^{\alpha_{d-1}-1}
\end{pmatrix} \\
&\quad -\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}(-1)^{d-1}(-1)^{d-1}c^{d-1} \\
&\quad \det\begin{pmatrix}
\frac{1}{c}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \alpha_1\gamma_1\sigma_1^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} & \dots\dots & \alpha_{d-2}\gamma_{d-2}\sigma_{d-2}^{-1}\prod_{k=1}^{d-1}\sigma_k^{\alpha_k} \\
\frac{1}{c}\gamma_1^{\alpha_1}\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & -\alpha_1\gamma_1^{\alpha_1-1}\sigma_1\prod_{k=2}^{d-1}\sigma_k^{\alpha_k} & \ddots & \vdots \\
\vdots & 0 & \ddots & \vdots \\
\vdots & \vdots & \ddots & \alpha_{d-2}\gamma_{d-3}^{\alpha_{d-3}}\gamma_{d-2}\sigma_{d-2}^{-1}\prod_{k=d-2}^{d-1}\sigma_k^{\alpha_k} \\
\frac{1}{c}\gamma_1^{\alpha_1}\sigma_{d-1}^{\alpha_{d-1}} & 0 & \dots\ 0 & -\alpha_{d-2}\gamma_{d-2}^{\alpha_{d-2}-1}\sigma_{d-2}\sigma_{d-1}^{\alpha_{d-1}}
\end{pmatrix}\Bigg) \\
&= (-1)^d c\left(\gamma_{d-1}^{\alpha_{d-1}}\alpha_{d-1}(-1)^{d-2}\sigma_{d-1}^{(d-1)\alpha_{d-1}-1}\gamma_{d-1}\det\left(A_{1,\dots,d-1}\right)\right. \\
&\quad \left. -(-1)^{d-1}\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}\sigma_{d-1}^{(d-1)\alpha_{d-1}}\det\left(A_{1,\dots,d-1}\right)\right) \\
&= (-1)^d c\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}^{(d-1)\alpha_{d-1}-1}\det\left(A_{1,\dots,d-1}\right)\left((-1)^{d-2}\gamma_{d-1}^2+(-1)^d\sigma_{d-1}^2\right) \\
&= (-1)\, c\alpha_{d-1}\gamma_{d-1}^{\alpha_{d-1}-1}\sigma_{d-1}^{(d-1)\alpha_{d-1}-1}c^{d-2}\prod_{i=1}^{d-2}\alpha_i\sigma_i^{i\alpha_i-1}\gamma_i^{\alpha_i-1}
\end{aligned}
$$

$$= (-1)\, c^{d-1} \prod_{i=1}^{d-1} \alpha_i \sigma_i^{i\alpha_i - 1} \gamma_i^{\alpha_i - 1}$$

which completes the proof.

□

Let

$$B_r^{\lambda_1,\dots,\lambda_{d-1}} := \left\{ \mathbf{x} \in (-\infty, 0)^d : \|\mathbf{x}\|_{\lambda_1,\dots,\lambda_{d-1}} < r \right\}, \quad r > 0$$

be the ball in $(-\infty, 0)^d$ of radius $r$ with respect to the norm $\|\cdot\|_{\lambda_1,\dots,\lambda_{d-1}}$, centered at the origin.

The next results are up to now only proven for dimension 3.

LEMMA 7.6. *Let $(X_1, X_2, X_3) < 0$ be a random vector which is distributed according to a GPD $W_{\lambda_1,\lambda_2}$ of nested logistic type. Choose a number $c_0 > 0$ that $W_{\lambda_1,\lambda_2}$ has on $B_{c_0}^{\lambda_1,\lambda_2}$ the representation*

$$W_{\lambda_1,\lambda_2}(x_1, x_2, x_3) = 1 - \|x\|_{\lambda_1,\lambda_2}$$

*and denote by $w_{\lambda_1,\lambda_2}$ the density of $W_{\lambda_1,\lambda_2}$. Then the density of the Shi transformation $ST^{-1}_{\lambda_1,\lambda_2} : B_{c_0}^{\lambda_1,\lambda_2} \to (0, c_0) \times \left(0, \frac{\pi}{2}\right)^2$ of the random vector restricted to $B_{c_0}^{\lambda_1,\lambda_2}$ is independent of the radial component $c$ and factorizes with regard to the angular components $\psi_1, \psi_2$.*
*The function*

$$f(c, \psi_1, \psi_2) = \left(2 - \frac{2}{\lambda_2}\right)\left(2\frac{\lambda_1 - \lambda_2}{\lambda_1} + \sigma_2^2 \left(\frac{4\lambda_2 - 2}{\lambda_1}\right)\right) \sigma_1 \sigma_2 \gamma_1 \gamma_2$$

*is the density of $ST^{-1}_{\lambda_1,\lambda_2}(X_1, X_2, X_3)$ on $(0, c_0) \times \left(0, \frac{\pi}{2}\right)^2$ under the restriction $(X_1, X_2, X_3) \in B_{c_0}^{\lambda_1,\lambda_2}$.*

PROOF. Using the Density Transformation Theorem (see for example Section 9.5 in Fristedt and Gray (1997, [**18**])), Lemma 5.3.3 and Lemma 7.6 it is

$$\begin{aligned}
&f(c, \psi_1, \psi_2) \\
&= w_{\lambda_1,\lambda_2}(ST_{\lambda_1,\lambda_2}(c, \psi_1, \psi_2)) \left|\det\left(J_{ST_{\lambda_1,\lambda_2}}(c, \psi_1, \psi_2)\right)\right| \\
&= (\lambda_2 - 1)\left((c\sigma_2^{\alpha_2}\sigma_1^{\alpha_1})^{\lambda_1} + (c\sigma_2^{\alpha_2}\gamma_1^{\alpha_1})^{\lambda_1}\right)^{-2} (c\sigma_2^{\alpha_2}\sigma_1^{\alpha_1})^{\lambda_1 - 1} (c\sigma_2^{\alpha_2}\gamma_1^{\alpha_1})^{\lambda_1 - 1} (c\gamma_2^{\alpha_2})^{\lambda_2 - 1}
\end{aligned}$$

$$
\begin{aligned}
&\left(\left(\left(c\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(c\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}+\left(c\gamma_2^{\alpha_2}\right)^{\lambda_2}\right)^{\frac{1}{\lambda_2}-3}\left(\left(c\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(c\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\\
&\left(\left(c\gamma_2^{\alpha_2}\right)^{\lambda_2}\left(\lambda_1-\lambda_2\right)+\left(\left(c\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(c\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\left(\lambda_1+\lambda_2-1\right)\right)\\
&\left(c^2\alpha_1\alpha_2\sigma_1^{\alpha_1-1}\sigma_2^{2\alpha_2-1}\gamma_1^{\alpha_1-1}\gamma_2^{\alpha_2-1}\right)\\
=&\left(\lambda_2-1\right)\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{-2}c^{-2\lambda_1}c^{2\lambda_1-2}\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1-1}\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1-1}\\
&c^{\lambda_2-1}\left(\gamma_2^{\alpha_2}\right)^{\lambda_2-1}c^{1-3\lambda_2}\left(\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}+\left(\gamma_2^{\alpha_2}\right)^{\lambda_2}\right)^{\frac{1}{\lambda_2}-3}\\
&c^{\lambda_2}\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\\
&c^{\lambda_2}\left(\left(\gamma_2^{\alpha_2}\right)^{\lambda_2}\left(\lambda_1-\lambda_2\right)+\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\left(\lambda_1+\lambda_2-1\right)\right)\\
&\left(c^2\alpha_1\alpha_2\sigma_1^{\alpha_1-1}\sigma_2^{2\alpha_2-1}\gamma_1^{\alpha_1-1}\gamma_2^{\alpha_2-1}\right)\\
=&\left(\lambda_2-1\right)\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{-2}\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1-1}\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1-1}\left(\gamma_2^{\alpha_2}\right)^{\lambda_2-1}\\
&\left(\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}+\left(\gamma_2^{\alpha_2}\right)^{\lambda_2}\right)^{\frac{1}{\lambda_2}-3}\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\\
&\left(\left(\gamma_2^{\alpha_2}\right)^{\lambda_2}\left(\lambda_1-\lambda_2\right)+\left(\left(\sigma_2^{\alpha_2}\sigma_1^{\alpha_1}\right)^{\lambda_1}+\left(\sigma_2^{\alpha_2}\gamma_1^{\alpha_1}\right)^{\lambda_1}\right)^{\frac{\lambda_2}{\lambda_1}}\left(\lambda_1+\lambda_2-1\right)\right)\\
&\left(\alpha_1\alpha_2\sigma_1^{\alpha_1-1}\sigma_2^{2\alpha_2-1}\gamma_1^{\alpha_1-1}\gamma_2^{\alpha_2-1}\right)\\
=&\left(\lambda_2-1\right)\alpha_1\alpha_2\sigma_1\sigma_2\gamma_1\gamma_2\left(\lambda_1+\lambda_2\left(\sigma_2^2-\gamma_2^2\right)-\sigma_2^2\right)\\
=&\left(2-\frac{2}{\lambda_2}\right)\frac{2}{\lambda_1}\left(\lambda_1+\lambda_2\left(\sigma_2^2-\gamma_2^2\right)-\sigma_2^2\right)\sigma_1\sigma_2\gamma_1\gamma_2\\
=&\left(2-\frac{2}{\lambda_2}\right)\left(2\frac{\lambda_1-\lambda_2}{\lambda_1}+\sigma_2^2\left(\frac{4\lambda_2-2}{\lambda_1}\right)\right)\sigma_1\sigma_2\gamma_1\gamma_2
\end{aligned}
$$

as asserted.

□

THEOREM 7.7. *Let $(X_1, X_2, X_3)$ follow a nested logistic GPD with parameters $\lambda_1 \geq \lambda_2 > 1$ and which has the density*

$$f(\psi_1, \psi_2) = f(c, \psi_1, \psi_2)$$

*with regard to its Shi coordinates on $B_{c_0}^{\lambda}$ as in Lemma 7.6. Then $f$ has a positive mass on $\left(0, \frac{\pi}{2}\right)^2$:*

$$\nu := \int_{\left(0, \frac{\pi}{2}\right)^2} f(\psi_1, \psi_2) \,\mathrm{d}(\psi_1, \psi_2) = \frac{(\lambda_2 - 1)(2\lambda_1 - 1)}{2\lambda_1\lambda_2} > 0.$$

*Furthermore we have conditional on $C = \|X\|_{\lambda_1, \lambda_2} < c_0$:*

(i) *The Shi coordinates $C, \Psi_1, \Psi_2$ are independent.*
(ii) *The random variable $C$ is on $(0, c_0)$ uniformly distributed.*
(iii) *The angular component $\Phi_1$ has the density*

$$f_1(\psi_1) := 2\sigma_1\gamma_1$$

*and the angular component $\Phi_2$ has the density*

$$f_2(\psi_2) := 4\sigma_2\gamma_2 \left(\frac{\lambda_1 - \lambda_2}{2\lambda_1 - 1} + \frac{2\lambda_2 - 1}{2\lambda_1 - 1}\sigma_2^2\right)$$

*and therefore they have the distribution functions*

$$F_1(\psi_1) := \int_0^{\psi_1} f_1(t) \,\mathrm{d}\, t = \sin^2(\psi_1)$$

*and*

$$F_2(\psi_1) := \int_0^{\psi_2} f_2(t) \,\mathrm{d}\, t = 2\frac{\lambda_1 - \lambda_2}{2\lambda_1 - 1} \sin^2(\psi_2) + \frac{2\lambda_2 - 1}{2\lambda_1 - 1} \sin^4(\psi_2)$$

*with the corresponding quantile functions*

$$F_1^{-1}(\psi_1) = \arcsin\left(\sqrt{\psi_1}\right)$$

*and*

$$F_2^{-1}(\psi_2) = \arcsin\left(\sqrt{\frac{\lambda_2 - \lambda_1 + \sqrt{\psi_2 - 2\psi_2\lambda_1 + \lambda_1^2 - 2\psi_2\lambda_2 - 2\lambda_1\lambda_2 + 4\psi_2\lambda_1\lambda_2 + \lambda_2^2}}{2\lambda_2 - 1}}\right).$$

Proof. Note that for $\tau \in \mathbb{R} \setminus \{-1\}$ it is

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{\sin^{\tau+1} x}{\tau+1} = \sin^{\tau} x \cos x$$

and therefore using the First Fundamental Theorem of Calculus (see for example Chapter 14 in Spivak (2006, [**40**])) we obtain

$$\int_0^{\psi} \sin^{\tau} x \cos x \,\mathrm{d}\, x = \left[\frac{\sin^{\tau+1} x}{\tau+1}\right]_0^{\psi} = \frac{\sin^{\tau+1} \psi}{\tau+1}$$

and in the special case of $\psi = \frac{\pi}{2}$

$$\int_0^{\frac{\pi}{2}} \sin^{\tau} x \cos x \,\mathrm{d}\, x = \frac{1}{\tau+1}.$$

Using Fubini's Theorem (see for example Theorem 8 in Section II §6 of Shiryaev (1984, [**38**])) we get

$$\begin{aligned}
\nu &:= \int_{\left(0,\frac{\pi}{2}\right)^2} f\left(\psi_1, \psi_2\right) \mathrm{d}\, \psi \\
&= \int_{\left(0,\frac{\pi}{2}\right)^2} \left(2 - \frac{2}{\lambda_2}\right) \left(2\frac{\lambda_1 - \lambda_2}{\lambda_1} + \sigma_2^2 \left(\frac{4\lambda_2 - 2}{\lambda_1}\right)\right) \sigma_1 \sigma_2 \gamma_1 \gamma_2 \,\mathrm{d}\left(\psi_1, \psi_2\right) \\
&= \left(2 - \frac{2}{\lambda_2}\right) \left(\left(2\frac{\lambda_1 - \lambda_2}{\lambda_1}\right) \int_0^{\frac{\pi}{2}} \sigma_1 \gamma_1 \,\mathrm{d}\, \psi_1 \int_0^{\frac{\pi}{2}} \sigma_2 \gamma_2 \,\mathrm{d}\, \psi_2 \right. \\
&\qquad \left. + \left(\frac{4\lambda_2 - 2}{\lambda_1}\right) \int_0^{\frac{\pi}{2}} \sigma_1 \gamma_1 \,\mathrm{d}\, \psi_1 \int_0^{\frac{\pi}{2}} \sigma_2^3 \gamma_2 \,\mathrm{d}\, \psi_2 \right) \\
&= \left(2 - \frac{2}{\lambda_2}\right) \left(\left(2\frac{\lambda_1 - \lambda_2}{\lambda_1}\right) \frac{1}{2}\frac{1}{2} + \left(\frac{4\lambda_2 - 2}{\lambda_1}\right) \frac{1}{2}\frac{1}{4}\right) \\
&= \frac{\left(\lambda_2 - 1\right)\left(2\lambda_1 - 1\right)}{2\lambda_1 \lambda_2}.
\end{aligned}$$

Since $\lambda_1 \geq \lambda_2 \geq 1$ because of the definition of the nested logistic model (see Lemma 5.3.1) the condition $\nu > 0$ is equivalent to $\lambda_2 > 1$.
For $c < c_0$ one gets

$$\begin{aligned}
\mathrm{P}\left(C < c\right) &= \mathrm{P}\left(C < c, \left(\Psi_1, \Psi_2\right) \in \left(0, \frac{\pi}{2}\right)^2\right) \\
&= \int_0^c \int_{\left(0,\frac{\pi}{2}\right)^2} f\left(c, \psi_1, \psi_2\right) \mathrm{d}\left(\psi_1, \psi_2\right) \mathrm{d}\, c
\end{aligned}$$

$$= \int_0^c \int_{\left(0,\frac{\pi}{2}\right)^2} f\left(\psi_1,\psi_2\right) \mathrm{d}\left(\psi_1,\psi_2\right) \mathrm{d}\, c$$
$$= c\nu.$$

Therefore we have for $c \in (0, c_0)$

$$\begin{aligned}
\mathrm{P}\left(C < c | C < c_0\right) &= \frac{\mathrm{P}\left(C < c, C < c_0\right)}{\mathrm{P}\left(C < c_0\right)} \\
&= \frac{\mathrm{P}\left(C < c\right)}{\mathrm{P}\left(C < c_0\right)} \\
&= \frac{c\nu}{c_0\nu} \\
&= \frac{c}{c_0}.
\end{aligned}$$

Thus $C$ is uniformly distributed on $(0, c_0)$ and we have established (ii).
Let $B$ be a Borel set in $\left(0, \frac{\pi}{2}\right)$. Using again Fubini's Theorem we obtain

$$\begin{aligned}
&\mathrm{P}\left(\psi_1 \in B | C < c_0\right) \\
&\quad = \frac{\mathrm{P}\left(\psi_1 \in B, \psi_2 \in \left(0, \frac{\pi}{2}\right) C < c_0\right)}{\mathrm{P}\left(C < c_0\right)} \\
&\quad = \frac{1}{c_0\nu} \int_0^{c_0} \int_B \int_0^{\frac{\pi}{2}} f\left(c, \psi_1, \psi_2\right) \mathrm{d}\, \psi_2 d\psi_1 dc \\
&\quad = \frac{1}{c_0\nu} c_0 \left(2\frac{\lambda_2 - 1}{\lambda_2}\right) \left(\int_B 2\frac{\lambda_1 - \lambda_2}{\lambda_1} \sigma_1\gamma_1 \int_0^{\frac{\pi}{2}} \sigma_2\gamma_2 \,\mathrm{d}\, \psi_2 \,\mathrm{d}\, \psi_1\right. \\
&\qquad\qquad \left. + \int_B \left(\frac{4\lambda_2 - 2}{\lambda_1}\right) \sigma_1\gamma_1 \int_0^{\frac{\pi}{2}} \sigma_2^3\gamma_2 \,\mathrm{d}\, \psi_2 \,\mathrm{d}\, \psi_1\right) \\
&\quad = \frac{1}{\nu} \left(2\frac{\lambda_2 - 1}{\lambda_2}\right) \left(\int_B \frac{\lambda_1 - \lambda_2}{\lambda_1} \sigma_1\gamma_1 \,\mathrm{d}\, \psi_1 + \int_B \frac{1}{2} \left(\frac{2\lambda_2 - 1}{\lambda_1}\right) \sigma_1\gamma_1 \,\mathrm{d}\, \psi_1\right) \\
&\quad = \frac{1}{\nu} \frac{\left(\lambda_2 - 1\right)\left(2\lambda_1 - 1\right)}{\lambda_1\lambda_2} \int_B \sigma_1\gamma_1 \,\mathrm{d}\, \psi_1 \\
&\quad = \int_B \underbrace{2\sigma_1\gamma_1}_{=f_1(\psi_1)} \mathrm{d}\, \psi_1.
\end{aligned}$$

Using the above equation with $B = (0, \psi_1]$ we get

$$F_1\left(\psi_1\right) = \int_0^{\psi_1} f_1\left(t\right) \mathrm{d}\, t$$

$$
\begin{aligned}
&= \int_0^{\psi_1} 2\sigma_1\gamma_1 \,\mathrm{d}\, t \\
&= \sin^2(\psi_1)
\end{aligned}
$$

and therefore the quantile function is given by

$$F_1^{-1}(\psi_1) = \arcsin\left(\sqrt{\psi_1}\right).$$

So we have established the density, distribution and quantile function of the first angular component. Using the same approach as above we will continue with the second angular density though we will see that those functions are more complicated. We have

$$
\begin{aligned}
&\mathrm{P}\left(\psi_2 \in B | C < c_0\right) \\
&\quad = \frac{\mathrm{P}\left(\psi_1 \in \left(0, \frac{\pi}{2}\right), \psi_2 \in B, C < c_0\right)}{\mathrm{P}\left(C < c_0\right)} \\
&\quad = \frac{1}{c_0 \nu} \int_0^{c_0} \int_0^{\frac{\pi}{2}} \int_B f\left(c, \psi_1, \psi_2\right) \mathrm{d}\,\psi_2 \,\mathrm{d}\,\psi_1 \,\mathrm{d}\, c \\
&\quad = \frac{2\lambda_1\lambda_2}{\left(\lambda_2 - 1\right)\left(2\lambda_1 - 1\right)} \int_B \frac{2}{\lambda_2}\left(\lambda_2 - 1\right) \left( \underbrace{\int_0^{\frac{\pi}{2}} \sigma_1\gamma_1 \,\mathrm{d}\,\psi_1}_{=\frac{1}{2}} 2\frac{\lambda_1 - \lambda_2}{\lambda_1}\sigma_2\gamma_2 \right. \\
&\qquad\qquad \left. + \underbrace{\int_0^{\frac{\pi}{2}} \sigma_1\gamma_1 \,\mathrm{d}\,\psi_1}_{=\frac{1}{2}} 2\frac{2\lambda_2 - 1}{\lambda_1}\sigma_2^3\gamma_2 d \right) d\psi_2 \\
&\quad = \int_B \underbrace{4\sigma_2\gamma_2 \left(\frac{\lambda_1 - \lambda_2}{2\lambda_1 - 1} + \frac{2\lambda_2 - 1}{2\lambda_1 - 1}\sigma_2^2\right)}_{=f_2(\psi_2)} \mathrm{d}\,\psi_2.
\end{aligned}
$$

Setting again $B = (0, \psi_2]$ we obtain

$$
\begin{aligned}
F_2(\psi_2) &= \int_0^{\psi_2} f_2(t) \,\mathrm{d}\, t \\
&= 2\frac{\lambda_1 - \lambda_2}{2\lambda_1 - 1}\sin^2(\psi_2) + \frac{2\lambda_2 - 1}{2\lambda_1 - 1}\sin^4(\psi_2)
\end{aligned}
$$

and therefore

$$F_2^{-1}(\psi_2) = \arcsin\left(\sqrt{\frac{\lambda_2 - \lambda_1 + \sqrt{\psi_2 - 2\psi_2\lambda_1 + \lambda_1^2 - 2\psi_2\lambda_2 - 2\lambda_1\lambda_2 + 4\psi_2\lambda_1\lambda_2 + \lambda_2^2}}{2\lambda_2 - 1}}\right).$$

So we have proven (iii).
Let $A \subseteq \left(0, \frac{\pi}{2}\right)^2$ be a Borel set and $0 < c < c_0$. Once again using Fubini's Theorem we obtain

$$\begin{aligned}
\mathrm{P}\left(C < c, \Psi \in A | C < c_0\right) &= \frac{\mathrm{P}\left(C < c, \Psi \in A\right)}{\mathrm{P}\left(C < c_0\right)} \\
&= \frac{1}{c_0\nu}\int_A \int_{(0,c)} f(t,\psi)\,\mathrm{d}\,t\,\mathrm{d}\,\psi \\
&= \frac{c}{c_0\nu}\int_A f(\psi)\,\mathrm{d}\,\psi \\
&= \frac{1}{c_0\nu} c_0 \int_A f(\psi)\,\mathrm{d}\,\psi \frac{c}{c_0} \\
&= \frac{1}{c_0\nu}\int_{(0,c_0)} \int_A f(\psi)\,\mathrm{d}\,\psi\,\mathrm{d}\,t \frac{c}{c_0} \\
&= \frac{\mathrm{P}\left(C < c_0, \Psi \in A\right)}{\mathrm{P}\left(C < c_0\right)} \mathrm{P}\left(C < c | C < c_0\right) \\
&= \mathrm{P}\left(\Psi \in A | C < c_0\right) \mathrm{P}\left(C < c | C < c_0\right)
\end{aligned}$$

which shows the conditional independence of the Shi coordinates $\Psi$ and $C$.
Let $A_1$ and $A_2$ be Borel sets in $\left(0, \frac{\pi}{2}\right)$. With Fubini's Theorem we have

$$\begin{aligned}
&\mathrm{P}\left(\Psi_1 \in A_1, \Psi_2 \in A_2 | C < c_0\right) \\
&\quad = \frac{\mathrm{P}\left(\Psi_1 \in A_1, \Psi_2 \in A_2, C < c_0\right)}{\mathrm{P}\left(C < c_0\right)} \\
&\quad = \frac{1}{c_0\nu}\int_0^{c_0} \int_{A_1} \int_{A_2} f(c,\psi)\,\mathrm{d}\,\psi_2\,\mathrm{d}\,\psi_1\,\mathrm{d}\,c \\
&\quad = \frac{2\lambda_1\lambda_2}{(2\lambda_1 - 1)(\lambda_2 - 1)} \\
&\qquad\quad \int_{A_1}\int_{A_2} \left(\frac{4}{\lambda_1\lambda_2}(\lambda_2 - 1)\left((\lambda_1 - \lambda_2) + \sigma_2^2(2\lambda_2 - 1)\right)\sigma_1\sigma_2\gamma_1\gamma_2\right)\mathrm{d}\,\psi_2\,\mathrm{d}\,\psi_1
\end{aligned}$$

$$
\begin{aligned}
&= 2\frac{1}{2\lambda_1 - 1} \int_{A_1} \int_{A_2} \left(4\left((\lambda_1 - \lambda_2) + \sigma_2^2 (2\lambda_2 - 1)\right) \sigma_1 \sigma_2 \gamma_1 \gamma_2\right) \mathrm{d}\,\psi_2 \,\mathrm{d}\,\psi_1 \\
&= \int_{A_1} \underbrace{2\sigma_1\gamma_1}_{=f_1(\psi_1)} \mathrm{d}\,\psi_1 \int_{A_2} \underbrace{4\sigma_2\gamma_2 \left(\frac{\lambda_1 - \lambda_2}{2\lambda_1 - 1} + \frac{2\lambda_2 - 1}{2\lambda_1 - 1}\sigma_2^2\right)}_{=f_2(\psi_2)} \mathrm{d}\,\psi_2 \\
&= \mathrm{P}\left(\Psi_1 \in A_1 | C < c_0\right) \mathrm{P}\left(\Psi_2 \in A_2 | C < c_0\right).
\end{aligned}
$$

Thus the angular components are (conditionally) independent and the proof is finished.

□

Following the outline of the proof we see that things get more complicated in higher dimensions. In contrary to the logistic model as regarded in Michel (2006, [**26**]) the density of the nested logistic model can not expressed in a "neat" formula and is far more complex.
Using the theorem above we can give an algorithm to create random variables from a GPD of nested logistic type on $B_{c_0}^{\lambda_1,\lambda_2}$.

ALGORITHM 7.8.

(i) *Generate* $U_1$, $U_2$ *and* $U_3$ *uniformly distributed on* $(0,1)$, *all pairwise independent.*

(ii) *Compute*

$$
\Psi_1 := \arcsin\left(\sqrt{U_2}\right)
$$

*and*

$$
\Psi_2 := \arcsin\left(\sqrt{\frac{\lambda_2 - \lambda_1 + \sqrt{U_3 - 2U_3\lambda_1 + \lambda_1^2 - 2U_3\lambda_2 - 2\lambda_1\lambda_2 + 4U_3\lambda_1\lambda_2 + \lambda_2^2}}{2\lambda_2 - 1}}\right).
$$

(iii) *Return the vector* $(X_1, X_2, X_3) = ST_{\lambda_1,\lambda_2}\left(U_1, \Psi_1, \Psi_2\right)$.

CHAPTER 8

# Final remarks

The Main Theorem of this thesis completes the characterization of extreme value distributions with norms. This topic was started by Falk et al. (2004, [**13**]) where the $D$-Norm was introduced. But the question in which cases a norm is a $D$-Norm was still unresolved. Two years later Falk (2006, [**12**]) answered this question in the bivariate case which turns out to be much simpler then the case of dimension 3 or higher. In my diploma thesis (Hofmann, 2006, [**22**]) I showed that the condition given in Falk (2006, [**12**]) is not sufficient in arbitrary dimension. In this thesis I gave a necessary and sufficient condition of a norm to be a $D$-Norm and therefore extreme value distributions can be expressed in terms of norms and a new notation of extreme value distributions has been established.
As mentioned in Section 2.1 there are already different representations of a multivariate extreme value distribution. One big advantage of the representation via a norm is the fact that norms are well-known and this yields to a representation of an EVD every undergraduate student can understand. Also linking EVD to norms may give more insight in EVDs using results from other mathematical disciplines.

Furthermore as a byproduct the also open question which functions yield a Pickands dependence function is answered since the $D$-Norm and the Pickands dependence function are closely related to each other.

Besides, the Main Theorem can be used to answer an open question. It was unclear if the condition $\lambda_i \geq \lambda_{i+1}$, $1 \leq i \leq d-2$, for the nested logistic model is necessary. It is established in Lemma 5.3.1 that the condition of the Main Theorem is not fulfilled if the $\lambda_i$ are descending and therefore this condition for the $\lambda_i$ is not only sufficient as already known but it is also necessary.
Michel (2006, [**26**]) extended the counterexample for a GPF that is not a distribution function over its whole support from dimension 3 to higher dimension

and from the sum-norm to the $\lambda$-Norm where $\lambda$ is close to 1. He also raised the question whether there is a $\lambda$-Norm with $\lambda < \infty$ that really defines a distribution function over its whole support. This question is now also answered: there is no $\lambda$ for which this is the case.

The last two chapters are independent from the Main Theorem.

In Chapter 6 we specified the left neighborhood of $\mathbf{0}$ where the GPD defines a multivariate distribution function, namely the GPD is a distribution function on $\left[-\frac{1}{d}, 0\right]^d$.
Furthermore Chapter 6 deals with the GPD-Flow based on a method to create GPDs using random vectors coming from copulas. This approach goes back to Buishand et. al. (2008, [**7**]) and was extended by Aulbach, Bayer and Falk in an unpublished paper ([**1**]). Since a copula is linked to a GPD there is also a norm (the $D$-Norm) linked to the copula. We determined a formula for the $D$-Norm in dependency of the copula. Finally the iteration step named "GPD-Flow" was examined in more detail and first results are obtained. It is shown that if the GPD-Flow converges then it must be against the copula of complete dependence. However the convergence of the GPD-Flow is still an open question. Moreover it can be investigated for more different copulas which GPD they yield. Furthermore if we want to obtain a certain copula which multivariate distribution must be used to obtain this GPD is another open problem. There is a practical relevance of this question. The problem of creating random numbers that follow a GPD could then be translated to the problem of creating random numbers that follow this multivariate distribution function.

In Chapter 7 the algorithm of creating random vectors following a GPD of the logistic model using the Shi-Transformation as introduced by Michel (2006, [**26**]) was generalized in dimension 3 to the nested logistic model. Unfortunately the Pickands density and the angular density are not easy to express for the nested logistic model in a simple form in dimension 4 and higher.

APPENDIX A

# Definitions

Here we give several definitions used in this thesis. The source where they are taken from are quoted.

Definition A.1 (Semiring, see Halmos (1974 [**20**]) on page 22).
*A semiring is a nonempty class* $\mathbf{P}$ *of sets such that*

(i) *if* $E \in \mathbf{P}$ *and* $F \in \mathbf{P}$, *then* $E \cap F \in \mathbf{P}$, *and*
(ii) *if* $E \in \mathbf{P}$, $F \in \mathbf{P}$ *and* $E \subseteq F$, *then there is a finite class* $\{C_0, C_1, ..., C_n\}$ *of sets in* $\mathbf{P}$ *such that* $E = C_0 \subseteq C_1 \subseteq \cdots \subseteq C_n = F$ *and* $D_i = C_i \backslash C_{i-1} \in \mathbf{P}$ *for* $i = 1, \ldots, n$.

Definition A.2 (Measurable selection, see Molchanov (2005, [**27**]), page 26).
*A random element* $\xi$ *with values in* $\mathbb{E}$ *is called a (measurable) selection of* $X$ *if* $\xi(\omega) \in X(\omega)$ *for almost all* $\omega \in \Omega$. *The family of all selections of* $X$ *os denoted by* $\mathcal{S}(X)$.

Definition A.3 ($p$-integrable selections, see Molchanov (2005, [**27**]), page 146).
*If* $X$ *is a random closed set in* $\mathbb{E}$, *then* $\mathcal{S}^p(X)$, $1 \leq p \leq \infty$, *denotes the family of all selections of* $X$ *from* $\mathbf{L}^p$, *so that*

$$\mathcal{S}^p(X) = \mathcal{S}(X) \cap \mathbf{L}^p,$$

*where* $\mathcal{S}(X)$ *denotes the family of all (measurable) selections of* $X$ *and* $\mathbf{L}^p$ *denote the space of random elements with values in* $\mathbb{E}$ *such that the* $L^p$*-norm is finite. In particular* $\mathcal{S}^1(X)$ *is the family of integrable selections.*

Definition A.4 (Selection Expectation, see Molchanov (2005, [**27**]), page 151).
*Let* $X$ *be a random closed set in a separable Banach space* $\mathbb{E}$. *The selection expectation of* $X$ *is the closure of the set of all expectations of integrable selections, i.e.*

$$\mathrm{E}\, X = \overline{\{\mathrm{E}\,\xi : \xi \in \mathcal{S}^1(X)\}}.$$

*The selection expectation is also often called the Aumann expectation.*

# Bibliography


[1] Aulbach, S., Bayer, V., and Falk, M. A multivariate piecing-together approach with an application to operational risk data.

[2] Balkema, A., and Resnick, S. Max-infinite divisibility. J. Appl. Probab. 14 (1977), 309–319.

[3] Bauer, H. Probability theory and elements of measure theory. International Series in Decision Processes. New York etc.: Holt, Rinehart and Winston, 1972.

[4] Bauer, H. Maß- und Integrationstheorie., 2. ed. de Gruyter Lehrbuch, Berlin, 1992.

[5] Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., De Waal, D., and Ferro, C. Statistics of extremes. Theory and applications. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons. xiii , 2004.

[6] Berg, C., Christensen, J. P. R., and Ressel, P. Harmonic analysis on semigroups. Springer, Berlin, 1984.

[7] Buishand, T., de Haan, L., and Zhou, C. On spatial extremes: with application to a rainfall problem. Annals of Applied Statistics 2, 2 (2008), 624–642.

[8] Cairns, S. S. Introductory topology. New York: The Ronald Press Company. IX, 244 p., 1961.

[9] de Haan, L., and de Ronde, J. Sea and wind: multivariate extremes at work. Extremes 1, 1 (1998), 7–46.

[10] de Haan, L., and Ferreira, A. Extreme value theory. An introduction. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer., 2006.

[11] de Haan, L., and Resnick, S. I. Limit theory for multivariate sample extremes. Z. Wahrscheinlichkeitstheor. Verw. Geb. 40 (1977), 317–337.

[12] Falk, M. A representation of bivariate extreme value distributions via norms on $\mathbb{R}^2$. Extremes 9, 1 (2006), 63–68.

[13] Falk, M., Hüsler, J., and Reiss, R.-D. Laws of small numbers: extremes and rare events. 2nd edition. Basel: Birkhäuser, 2004.

[14] Falk, M., and Reiss, R.-D. On Pickands coordinates in arbitrary dimensions. J. Multivariate Anal. 92, 2 (2005), 426–453.

[15] Fisher, R. A., and Tippett, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings Cambridge 24 (1928), 180–190.

[16] Fraleigh, J. B., and Beauregard, R. A. Linear algebra. World Student Series. Reading, Mass., etc.: Addison-Wesley Publishing Company. XVII, 477, A-35, I-7 p. (TUB, Abt. Schiffstechnik: 8Qi 2959) , 1987.

[17] Fredricks, G. A., and Nelsen, R. B. Copulas constructed from diagonal sections. Beneš, Viktor (ed.) et al., Distributions with given marginals and moment problems. Proceedings of the 1996 conference, Prague, Czech Republic. Dordrecht: Kluwer Academic Publishers. 129-136 (1997)., 1997.

[18] Fristedt, B., and Gray, L. A modern approach to probability theory. Boston: Birkhäuser, 1997.

[19] Galambos, J. The asymptotic theory of extreme order statistics. Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons., 1978.

[20] Halmos, P. R. Measure Theory. Springer, Berlin, 1974.

[21] Harville, D. A. Matrix algebra from a statistician's perspective. New York, NY: Springer. xvii, 630 p., 1997.

[22] Hofmann, D. Über die darstellung multivariater extremwertverteilungen mittels normen. Master's thesis, University of Wuerzburg, 2006.

[23] Huang, X. Statistics of Bivariate Extreme Values. Thesis Publishers and Tinbergen Institute, 1992.

[24] Joe, H. Multivariate extreme-value distributions with applications to environmental data. Can. J. Stat. 22, 1 (1994), 47–64.

[25] Kaufmann, E., and Reiss, R.-D. Approximation rates for multivariate exceedances. J. Stat. Plann. Inference 45, 1-2 (1995), 235–245.

[26] Michel, R. Simulation and Estimation in Multivariate Generalized Pareto Models. Ph.D. dissertation, University of Wuerzburg, 2006.

[27] Molchanov, I. Theory of random sets. Probability and Its Applications. London: Springer, 2005.

[28] Molchanov, I. Convex and star-shaped sets associated with stable distributions. http://www.citebase.org/abstract?id=oai:arXiv.org:0707.0221, 2007.

[29] Molchanov, I. Convex geometry of max-stable distributions. Extremes 11, 3 (2008), 235–259.

[30] Molchanov, I., and Schmutz, M. Geometric extension of put-call symmetry in the multiasset setting. http://www.citebase.org/abstract?id=oai:arXiv.org:0806.4506, 2008.

[31] Nelsen, R. B. An introduction to copulas. 2nd edition. Springer Series in Statistics. New York, NY: Springer. , 2006.

[32] Reiss, R.-D. Approximate distributions of order statistics. (With applications to nonparametric statistics). Springer Series in Statistics, Springer, New York, 1989.

[33] Reiss, R.-D., and Thomas, M. Statistical analysis of extreme values. With applications to insurance, finance, hydrology and other fields. 3rd edition. Basel: Birkhäuser., 2007.

[34] Rivas, D., Caleyo, F., Valor, A., and Hallen, J. Extreme value analysis applied to pitting corrosion experiments in low carbon steel: Comparison of block maxima and peak over threshold aproaches. Corrosion Science 50, 11 (2008), 3193–3204.

[35] Rootzén, H., and Tajvidi, N. Multivariate generalized Pareto distributions. Bernoulli 12, 5 (2006), 917–930.

[36] Schneider, R. Convex Bodies. The Brunn-Minkowski Theory. Cambridge University Press, Cambridge, 1993.

[37] Shi, D. Multivariate extreme value distribution and its fisher information matrix. Acta Math. Appl. Sin., Engl. Ser. 11, 4 (1995), 421–428.

[38] Shiryaev, A. Probability. Transl. from the Russian by R. P. Boas. Graduate Texts in Mathematics, 95. New York etc.: Springer-Verlag. XI, 577 p., 54 figs. DM 148.00; $ 57.50 , 1984.

[39] Sibuya, M. Bivariate extreme statistics. I. Ann. Inst. Stat. Math. 11 (1960), 195–210.

[40] Spivak, M. Calculus. Corrected 3rd edition. Cambridge: Cambridge University Press. xii, 670 p., 2006.

[41] Tajvidi, N. Characterisation and Some Statistical Aspects of Univariate and Multivariate Generalised Pareto Distributions. Ph.D. dissertation, Chalmers Högskola Göteburg, 1996.

[42] Tawn, J. A. Modelling multivariate extreme value distributions. Biometrika 77, 2 (1990), 245–253.

[43] Tiago de Oliveira, J. Statistical decision for bivariate extremes. Port. Math. 24 (1966), 145–154.

## Acknowledgment

An dieser Stelle möchte ich allen Personen Dank sagen, die mich während der Entstehung dieser Arbeit unterstützt haben.
Vor allen möchte ich mich bei Herrn Prof. Dr. Falk für die sehr gute Betreuung, sein stets offenes Ohr und die vielen motivierenden Worte ganz herzlich bedanken. Über das mir entgegengebrachte Vertrauen in den letzten Jahren bin ich ihm sehr verbunden.
Mein besonderer Dank gilt auch Herrn Dr. Heins, Herrn Dr. Wensauer und Frau Distler, die mir die Zusammenarbeit mit AREVA ermöglicht haben und mir immer hilfreich zur Seite standen.
Des Weiteren möchte ich mich bei Dr. Julia Wissel und Dominik Streit für die moralische Unterstützung und die schöne Zeit in Erlangen bedanken.
Birgit Irblich, Christopher Klatt, Christian Appold und Sven Herzberg danke ich für die geistreichen abendlichen Gespräche bei meinen Besuchen in Würzburg.
Besonders erwähnen möchte ich meine Eltern Dagmar und Walter Hofmann, die mir mein Leben lang in jeglicher Hinsicht zur Seite gestanden haben und stets für mich da sind.
Letztendlich möchte ich mich noch bei allen bedanken, die mich die Jahre über unterstützt haben, in schwierigen Zeiten für mich da waren und mich nicht im Stich gelassen haben.