Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Lehrstuhl für Mathematik VIII
Statistik

Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften

Extreme Value Theory in Higher Dimensions

Max-Stable Processes and Multivariate Records

vorgelegt von

Maximilian Zott

am 03. Juni 2016 in

Würzburg

Acknowledgments

Over the past couple of years, many people have asked me what I do for a living. This should not be a difficult question to respond to, but I found it surprisingly hard to explain what my thesis was about. Eventually, I came up with a few lines that I repeated whenever I was asked about my research. I want to dedicate this thesis to all those close to me, who have had to endure listening to not always very authentic descriptions of my profession. This is what I did during the last couple of years.

First and foremost, I would like to thank my dear family, without whom this thesis would not have been possible, for their support on my way to this dissertation, as well as my girlfriend, who always encouraged me when I was doubtful.

Without doubt, the success of a thesis is highly dependent on the support of a supervisor. I am therefore very grateful to Michael Falk, who laid the foundation for this thesis by introducing me to the field of extreme value theory. I benefited a lot from his support, his helpful suggestions, and from several invitations to conferences, which had a large impact on this work. For instance, it was he who took a presentation given by my second advisor, Clément Dombry, at a conference in Pisa in December 2014 as an opportunity to initiate a fruitful collaboration. As a result, I spent 4 months in Besançon in 2015 to develop a joint research project, the outcomes of which have found an important place in this thesis. It was an amazing experience for me to work in Besançon and I enjoyed this cooperation a lot. For this, I would like to thank Clément Dombry very much.

Last, but not least, I express my gratitude to my colleagues at the University of Würzburg for valuable discussions, which certainly had a positive impact on this work. In particular, I would like to mention Martin Hofmann, whom I consider to be a mentor, and Stefan Aulbach.

Würzburg, May 2016 *Maximilian Zott*

Contents

List of Figures

1 Introduction

Being a sub-discipline of probability theory and statistics, extreme value theory is concerned with random observations. However, while fundamental theories of classical stochastics - such as the *laws of large numbers* or the *central limit theorem* - are used to investigate the behavior of a normalized sum of observations, extreme value theory focuses on the maximum or minimum of a set of observations. Thus the key task of extreme value theory is to model extremal (and hence rare) events. These events can be of any kind: Climatologists, for instance, are interested in extreme heatwaves, heavy rain events, or severe storms. In finance, on the other hand, huge losses on the stock market or complete financial fallouts, such as the 2008 Lehman Brothers bankruptcy, serve as examples of extremal events. Mathematically speaking, the following setup is the basis of classical (univariate) extreme value theory:

Take independent and identically distributed (iid) random variables $X, X^{(1)}, X^{(2)}, \ldots \in \mathbb{R}$ with common distribution function F. We will use an example from the field of climatology for way of illustration. Without doubt, it would not be realistic to model the daily rainfall at a weather station by an iid sequence $X^{(1)}, X^{(2)}, \ldots$ Neither will daily rainfall values be independent (the weather today has an influence on the weather tomorrow), nor will they be identically distributed (weather is highly dependent on the season in most parts of the world). However, we can think of $X^{(1)}, X^{(2)}, \ldots$ as the annual maxima of rainfall at a weather station. The iid assumption is now a lot more reasonable, despite the fact that there will be other disruptive factors, such as climate change. Now consider the random variables

$$M^{(n)} := \max\big(X^{(1)}, \ldots, X^{(n)}\big), \quad m^{(n)} := \min\big(X^{(1)}, \ldots, X^{(n)}\big).$$

It is sufficient to focus on the maximum since

$$\min\big(X^{(1)}, \ldots, X^{(n)}\big) = -\max\big(-X^{(1)}, \ldots, -X^{(n)}\big).$$

Given the iid property, calculating the distribution of these random variables is not difficult:

$$P\big(M^{(n)} \le x\big) = P\big(X^{(1)} \le x, \dots, X^{(n)} \le x\big) = F^n(x) \to_{n\to\infty} \begin{cases} 1, & \text{if } x \ge x_F, \\ 0, & \text{if } x < x_F, \end{cases}$$

where $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$ is the right endpoint of F. Clearly, this relation does not yield much information on the behavior of the extremal distribution. As in the central limit theorem, it is therefore necessary to *normalize* the maximum, i. e. we take $c_n > 0$, $d_n \in \mathbb{R}$, $n \in \mathbb{N}$, and consider the convergence

$$P\big((M^{(n)} - d_n)/c_n \le x\big) = F^n(c_n x + d_n) \to_{n\to\infty} G(x), \qquad x \in \mathbb{R},$$

where G is a non-degenerate distribution function, i. e. its mass is not concentrated on a single point. The fundamental theorem of extreme value theory by Fisher and Tippett (1928) and Gnedenko (1943) now states that the limit G is either a Fréchet, a Weibull or a Gumbel type distribution.

The next step is to go beyond the univariate world, and to observe several weather stations or a portfolio of shares simultaneously. This results in a sequence of iid random vectors $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots \in \mathbb{R}^d$, where the random variable $X_j^{(i)} \in \mathbb{R}$ represents the i-th observation at the j-th spot, $i \in \mathbb{N}$, $j = 1, \dots, d$. An extreme multivariate observation is meant to be extreme in each component. A heavy rainfall event, for instance, is observed at several weather stations simultaneously, or each share in a portfolio registers a huge loss at the same time etc. In accordance with that, we look at the *componentwise* maximum

$$\boldsymbol{M}^{(n)} = \big(M_1^{(n)}, \dots, M_d^{(n)}\big) := \left(\max_{i=1,\dots,n} X_1^{(i)}, \dots, \max_{i=1,\dots,n} X_d^{(n)}\right).$$

Unlike the sequence of observations itself, the components $X_1^{(i)}, \dots, X_d^{(i)}$ of each observation $\boldsymbol{X}^{(i)}$, $i \in \mathbb{N}$, are typically *not* iid. Modeling the dependence of the components is, in fact, one of the most crucial and challenging tasks in multivariate extreme value theory. For instance, if a portfolio contains shares of two banks - say bank A and bank B - it is likely that a significant loss made by bank A will have an influence also on bank B. Similarly, the weather observed at one weather station will have an impact on the weather at a second station, if they are not too far from each other.

As before, we examine the limit behavior of the componentwise maximum, i. e.

$$P\big((\boldsymbol{M}^{(n)} - \boldsymbol{d}_n)/\boldsymbol{c}_n \le \boldsymbol{x}\big) \to_{n\to\infty} G(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

where $\boldsymbol{c}_n \in [0,\infty)^d$, $\boldsymbol{d}_n \in \mathbb{R}^d$, $n \in \mathbb{N}$, are norming vectors and G is a d-variate distribution function with non-degenerate univariate margins. The characteristic property of G is its *max-stability*, i. e. there exist norming vectors $\boldsymbol{a}_n \in [0,\infty)^d$, $\boldsymbol{b}_n \in \mathbb{R}^d$, such that

$$G^n(\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n) = G(\boldsymbol{x}), \qquad n \in \mathbb{N},\ \boldsymbol{x} \in \mathbb{R}^d.$$

Note that all operations such as multiplication, addition, $\le$, and so on are defined componentwise. The theory of multivariate extremes can further be extended to *stochastic processes*. In that *functional* case, the observations are continuous real-valued functions on a compact metric space. Examples are the height of the sea level along a section of coast (modeled by $S = [0,1]$), or of the temperature over a certain part of a map (modeled by $S = [0,1]^2$). This leads to the theory of *max-stable processes*.

The representation of max-stable processes is much more complex than the univariate theory. Remember that the limit distribution function G is either a Fréchet, a Weibull, or a Gumbel type distribution in the one-dimensional case. In fact, the entire class of univariate max-stable distributions can be described by the three parameters of scale, location, and shape. In the multivariate or functional world, however, the complex dependence structure within the univariate margins leads to non-parametric classes of max-stable processes. Higher-dimensional max-stable distributions can be described by the so-called *angular measure*, which was described mainly by de Haan and Resnick (1977) in the multivariate, and by Giné et al. (1990) in the functional setup. Useful models of the dependence structure are also given by a certain class of norms, more precisely the *D-norms*, established by Falk et al. (2011). Max-stable processes are discussed in Chapter 2 extensively. Section 2.1 provides a recap of how max-stable distributions can be represented, and summarizes well-established facts, which are of crucial importance for the remainder of this thesis. In addition, different parametric models of multivariate max-stable distributions are presented. Section 2.2 introduces the new concept of dual D-norm functions, an important tool for higher-dimensional records, which are discussed later on. Finally, Section 2.3 places a focus on path properties of max-stable processes. More precisely, the concept of differentiability in distribution is introduced, and further results on max-stable processes - such as the distribution of the increments of a max-stable process - are collected.

The theory of max-stable processes with continuous sample paths has been developed a lot since the groundbreaking paper of de Haan (1984) was published. There is, however, a basic problem in practice, as it is not possible in reality to observe an entire process on an interval. Regardless of how many weather stations have been built in a country, it will only be possible to collect measurements from a finite number of locations. It must therefore be asked how stochastic processes with continuous sample paths on the interval $[0, 1]$ - or, more generally, on a compact metric space S - can be constructed, such that a given max-stable random vector is interpolated by this process while max-stability is being preserved. This leads to the *generalized max-linear models* that will be discussed in Chapter 3.

The last part of this dissertation, finally, takes a closer look on *multivariate records.* In daily life, it is impossible not to come across such records, given the diverse fields in which they appear. In 2009, Usain Bolt set the current 100 meters record; in 2015 the German stock market index DAX reached a new all time high; and global temperatures in 2015 were the warmest since modern record keeping began in 1880. From a mathematical point of view, it is quite straightforward to define a record from a sequence of random variables $X^{(1)}, X^{(2)}, \dots \in \mathbb{R}$. The n-th observation is called a record, if it is greater (or less) than all $n-1$ observations before. However, if $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ are random vectors or even stochastic processes, the lack of natural order in higher-dimensional spaces enables the definition of records in many different ways. For illustration, think of a decathlete. In our example, the decathlete sets a new record, if he or she sets a new decathlon record in at least one event, say javelin. Later on, we will name such an observation a *simple record.* A very strong decathlete might even be able to set a new record in all ten events at the same time. Naturally, such a *complete record* is much less likely. In Chapter 4, new results on simple and complete records, provided that the underlying sequence of observations is iid, are collected. Finally, the development of the componentwise maximum $\boldsymbol{M}^{(n)}$ over time will be investigated by means of so-called *hitting scenarios.*

2 Max-stable processes

2.1 Representations of max-stable processes

Let in what follows S be a compact metric space. Recall that, in particular, S is complete and also separable. To begin with, we define some function spaces that will be needed throughout the whole work. Let

$$\begin{aligned} C(S) &:= \{f : S \to \mathbb{R} : f \text{ continuous}\}, \\ C^+(S) := \{f \in C(S) : f > 0\} \quad &\text{and} \quad C^-(S) := \{f \in C(S) : f < 0\}, \\ \bar{C}^+(S) := \{f \in C(S) : f \geq 0\} \quad &\text{and} \quad \bar{C}^-(S) := \{f \in C(S) : f \leq 0\}, \\ E(S) &:= \{f : S \to \mathbb{R} : f \text{ bounded with only finitely many discontinuities}\}, \\ E^+(S) := \{f \in E(S) : f > 0\} \quad &\text{and} \quad E^-(S) := \{f \in E(S) : f < 0\}, \\ \bar{E}^-(S) := \{f \in E(S) : f \leq 0\} \quad &\text{and} \quad \bar{E}^+(S) := \{f \in E(S) : f \geq 0\}. \end{aligned}$$

All these function spaces can be equipped with the *supremum norm*

$$\|f\|_\infty := \sup_{s \in S} |f(s)|.$$

A typical choice of S would be the euclidean cube $S = [0,1]^k$ for some $k \in \mathbb{N}$, but we can also think of $S = \{1, \ldots, d\}$ equipped with the discrete metric. In that case, any function mapping from $\{1, \ldots, d\}$ to $\mathbb{R}$ is continuous, and hence $C(S) = E(S) = \mathbb{R}^d$, $C^+(S) = (0, \infty)^d$, and so on. To point out possible differences between the finite and the general case, we use the term *multivariate case* clarifying that the index set S is finite, whereas the *functional case* should be pictured as $S = [0,1]^k$.

Any operation on these function spaces is always meant componentwise, that is, for instance for $f_1, \ldots, f_n \in C(S)$,

$$\begin{aligned} f_1 \leq f_2 :\iff f_1(s) \leq f_2(s) \text{ for all } s \in S, \\ f_1 + f_2 := (f_1(s) + f_2(s))_{s \in S}, \end{aligned}$$

$$\max_{i=1,\dots,n} f_i := \left(\max_{i=1,\dots,n} f_i(s)\right)_{s\in S},$$

$$\vdots$$

We denote by $(\Omega, \mathcal{A}, P)$ a probability space, i. e. $\mathcal{A}$ is a σ-algebra on Ω and P a probability measure on $(\Omega, \mathcal{A})$. If $\xi : (\Omega, \mathcal{A}) \to (\Omega', \mathcal{A}')$ is a measurable mapping to another measure space $(\Omega', \mathcal{A}')$, then $P * \xi := P \circ \xi^{-1}$ marks the *distribution* of ξ.

Constant functions are often indicated by bold numbers, i. e. we write

$$\mathbf{1} := \mathbf{1}_S := (1)_{s\in S}, \quad \mathbf{0} := \mathbf{0}_S := (0)_{s\in S}, \dots$$

The *indicator function* corresponding to a subset A of an arbitrary set $\mathcal{X}$ is denoted by

$$\mathbb{1}_A(x) := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{else.} \end{cases}$$

If there is a σ-Algebra $\mathcal{B}$ available on $\mathcal{X}$, we can define the *Dirac measure* ϵ_x for a fixed $x \in \mathcal{X}$ via

$$\epsilon_x(A) := \mathbb{1}_A(x), \qquad A \in \mathcal{B}.$$

If $\mathcal{X}$ is a topological space, we denote by $\mathbb{B}(\mathcal{X})$ the *Borel σ-algebra*, i. e. the σ-algebra which is generated by the open subsets of $\mathcal{X}$. We will often discuss *weak convergence* of finite measures or random variables on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$, where $\mathcal{X}$ is a metric space. A sequence of finite measures μ_n on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$ *converges weakly* to a a finite measure μ on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$ (write $\mu_n \to_w \mu$), if

$$\int_{\mathcal{X}} f(x)\ \mu_n(\mathrm{d}x) \to_{n\to\infty} \int_{\mathcal{X}} f(x)\ \mu(\mathrm{d}x)$$

for every continuous and bounded function $f : \mathcal{X} \to \mathbb{R}$. Analogously, a sequence of random variables on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$ *converges weakly* or *in distribution* to a random variable X on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$, if $P * X_n \to_w P * X$, i. e.

$$E(f(X_n)) \to_{n\to\infty} E(f(X)),$$

where $E(\xi)$ denotes the expectation of an integrable real-valued random variable ξ. In that case, we also write

$$X_n \to_{\mathcal{D}} X.$$

A stochastic process $\boldsymbol{\vartheta} = (\vartheta_s)_{s\in S}$ with sample paths in $C(S)$ and non-degenerate univariate margins is called a *max-stable process* (MSP), if there are norming functions $a_n \in C^+(S)$, $b_n \in C(S)$, $n \in \mathbb{N}$, such that

$$\max_{i=1,\dots,n} \frac{\boldsymbol{\vartheta}^{(i)} - b_n}{a_n} =_{\mathcal{D}} \boldsymbol{\vartheta}, \qquad n \in \mathbb{N}, \tag{2.1}$$

where $\boldsymbol{\vartheta}^{(1)}, \boldsymbol{\vartheta}^{(2)}, \dots$ are independent and identically distributed (iid) copies of $\boldsymbol{\vartheta}$, and $=_{\mathcal{D}}$ denotes equality in distribution. The distribution of an MSP is called *max-stable* itself. Note again that in the case $S = \{1, \dots, d\}$, we deal in fact with random vectors in $\mathbb{R}^d$ rather than with actual stochastic processes.

Since max-stable distributions are the only possible non-degenerate limit distributions in a sequence of linearly standardized maxima of iid processes, MSP play an outstanding role in extreme value theory. Let $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid stochastic processes with sample paths in $C(S)$. Assume there exist norming functions $c_n \in C^+(S)$, $d_n \in C(S)$, $n \in \mathbb{N}$, and a stochastic process $\boldsymbol{Y}$ with sample paths in $C(S)$ and non-degenerate univariate margins such that

$$\max_{i=1,\dots,n} \frac{\boldsymbol{X}^{(i)} - d_n}{c_n} \to_{\mathcal{D}} \boldsymbol{Y} \tag{2.2}$$

in $C(S)$. The process $\boldsymbol{Y}$ in (2.2) is called *extreme value process*, and its distribution *extreme value distribution*. In fact, it is not difficult to see (cf. de Haan and Ferreira (2006, Section 9.2)) that the class of max-stable and extreme value processes coincide. Therefore, we will only refer to max-stable instead of extreme value processes from now on, even if we have the convergence (2.2) in mind. In Section 2.2, we will specifically bring relation (2.2) into focus.

Clearly, the univariate margins of an MSP are max-stable themselves, and hence belong to the class of either Fréchet, Weibull or Gumbel type distributions. In fact, the fundamental theorem in Extreme Value Theory of Fisher and Tippett (1928) and Gnedenko (1943) states that for every $s \in S$,

$$P(\vartheta_s \le a(s)x + b(s)) = G_{\gamma(s)}(x) = \exp\left(-\left(1 + \gamma(s)x\right)^{-1/\gamma(s)}\right), \qquad 1 + \gamma(s)x > 0, \tag{2.3}$$

where $a(s) > 0$, $b(s), \gamma(s) \in \mathbb{R}$ are parameters of scale, location and shape, and in the case $\gamma(s) = 0$, the right hand side is interpreted as $\exp(-e^{-x})$. The parameter $\gamma(s)$ is commonly known as the *extreme value index*. It was shown in Giné et al. (1990) that the functions $a(\cdot)$, $b(\cdot)$ and $\gamma(\cdot)$ are in fact continuous.

A stochastic process $\boldsymbol{\xi} = (\xi_s)_{s\in S}$ with sample paths in $C(S)$ is commonly called a *simple max-stable process* if it is max-stable with standard univariate Fréchet margins, i. e. $P(\xi_s \le x) = \exp(-x^{-1})$, $x > 0$, $s \in S$. In that case, we necessarily have $a(s) \equiv n$, $b(s) \equiv 0$, i. e.

$$\boldsymbol{\xi} =_{\mathcal{D}} \frac{1}{n} \max_{i=1,\dots,n} \boldsymbol{\xi}^{(i)},$$

where $\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(n)}$ are iid copies of $\boldsymbol{\xi}$. Different to that, we call a stochastic process $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ with sample paths in $C(S)$ a *standard max-stable process* (SMSP) if it is max-stable with standard negative exponential margins, i. e. $P(\eta_s \le x) = \exp(x)$, $x \le 0$, $s \in S$. In that case, we have

$$\boldsymbol{\eta} =_{\mathcal{D}} n \max_{i=1,\dots,n} \boldsymbol{\eta}^{(i)}, \tag{2.4}$$

where $\boldsymbol{\eta}^{(1)}, \dots, \boldsymbol{\eta}^{(n)}$ are iid copies of $\boldsymbol{\eta}$. Having a closer look to the univariate margins of simple and standard MSP, it seems to be obvious that $-1/\boldsymbol{\xi}$ is an SMSP if $\boldsymbol{\xi}$ is a simple MSP, and vice versa. However, to this end, it is necessary to prove that neither a simple MSP $\boldsymbol{\xi}$ nor an SMSP $\boldsymbol{\eta}$ attains the value zero with probability one. This is trivial in case the parameter space S is a finite set, since (2.3) implies that a max-stable random vector has a continuous distribution function (df). In addition, $P(\boldsymbol{\xi} > 0) = 1$ was shown in Giné et al. (1990, Corollary 3.4). For the sake of completeness, we state the equivalent result in the SMSP case, observing that the assertion has already been proven in Hofmann (2012, Lemma 2.2) in the case $S = [0, 1]$.

Lemma 2.1. *Let $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ be an SMSP. Then $P(\boldsymbol{\eta} < 0) = 1$.*

Proof. Denote by δ the metric pertaining to the compact metric space S. The max-stability of $\boldsymbol{\eta}$ (2.4) implies

$$P(\boldsymbol{\eta} < 0) = P\left(\max_{i=1,\dots,n} \boldsymbol{\eta}^{(i)} < 0\right) = P(\boldsymbol{\eta} < 0)^n,$$

hence

$$P(\boldsymbol{\eta} < 0) \in \{0, 1\}. \tag{2.5}$$

Now suppose $P(\boldsymbol{\eta} < 0) = 0 \iff P\left(\sup_{s\in S} \eta_s = 0\right) = 1$. Choose closed sets $\emptyset \ne A_1, B_1 \subsetneq S$ with $A_1 \cup B_1 = S$. Then (2.5) yields either $P\left(\sup_{s\in A_1} \eta_s = 0\right) = 1$ or $P\left(\sup_{s\in B_1} \eta_s = 0\right) = 1$, since the spaces (A_1, δ) and (B_1, δ) are compact metric spaces again, and it cannot occur that both probabilities are zero. Suppose without loss of generality that $P\left(\sup_{s\in A_1} \eta_s = 0\right) = 1$. Again, choose closed sets $\emptyset \ne A_2, B_2 \subsetneq A_1$ with $A_2 \cup B_2 = A_1$, repeat the preceding arguments and assume without loss of generality that $P\left(\sup_{s\in A_2}\eta_s=0\right)=1$. By iteration, it is possible to find a sequence of closed sets A_n, $n\in\mathbb{N}$, with $A_1\supset A_2\supset A_3\supset\cdots$ such that

$$\operatorname{diam}(A_n):=\sup\{\delta(x,y):\ x,y\in A_n\}\to_{n\to\infty}0$$

and $P\left(\sup_{s\in A_n}\eta_s=0\right)=1$, $n\in\mathbb{N}$. With S being a complete metric space, Cantor's intersection theorem implies that there is $s_0\in S$ with $\bigcap_{n\in\mathbb{N}}A_n=\{s_0\}$. Therefore, we obtain by the continuity from above of a probability measure and the fact that η_{s_0} is continuously distributed

$$0=P(\eta_{s_0}=0)=P\left(\bigcap_{n\in\mathbb{N}}\left\{\sup_{s\in A_n}\eta_t=0\right\}\right)=\lim_{n\to\infty}P\left(\sup_{s\in A_n}\eta_s=0\right)=1,$$

which is clearly a contradiction. □

Corollary 2.2. *A stochastic process $\boldsymbol{\eta}$ is an SMSP if and only if (iff) $-1/\boldsymbol{\eta}$ defines a simple MSP.*

Generators and D-norms

In the literature, there are various approaches that aim the characterization and representation of MSP. Depending on the nature of the index set S, different challenges have to be faced in order to understand the structure of max-stable distributions. For instance, considering either $S=\{1,\dots,d\}$ or $S=[0,1]$ will leave us with either regular random vectors or stochastic processes with continuous sample paths on $[0,1]$ - needless to say that not every result can be transferred from the finite-dimensional setup to the world of stochastic processes in a straightforward way, due to the different topological structure of $\mathbb{R}^d$ and $C([0,1])$.

In the beginnings, multivariate extreme value theory was mainly restricted to the finite-dimensional framework, where $S=\{1,\dots,d\}$. The first step on the way to the characterization of multivariate max-stability was in fact the investigation of a broader class of distributions, the *max-infinitely divisible* (max-id) distributions. The crucial observation is that every max-id distribution can be expressed by means of a so-called *exponent measure*. This was shown by Balkema and Resnick (1977) in the bivariate setup, and was later extended to higher (and even infinite) dimensions by Vatan (1985), see also Gerritse (1986). Since max-stable distributions are in particular max-id, it is clear that max-stable distributions have exponent measures as well. However, it was the groundbreaking paper of de Haan and Resnick (1977) that laid the foundation of multivariate extreme value theory by observing that the exponent measure satisfies some kind of homogeneity in the max-stable case. This property allows the factorization of the exponent measure into a radial and an angular part, where the latter is expressed by a finite measure on the unit circle, commonly known as the *angular measure*.

The strategy to deduce an angular measure which characterizes max-stable distributions has mainly been adapted to obtain analogous results in the case of sample path continuous stochastic processes on an arbitrary compact metric index space S. The papers of de Haan (1984) and Giné et al. (1990) are certainly two of the most seminal contributions in this matter.

Nice representations of MSP, in the finite-dimensional setup as well as in the functional context, can be obtained by using the angular measure to define a certain class of norms, the so-called D-norms, which traces back to Falk (2006) (see also Falk et al. (2011, Section 4.4) for details), and Aulbach et al. (2013) in the functional setup.

We will start with the representation of MSP established by de Haan and Resnick (1977) (in the finite-dimensional case) and Giné et al. (1990) (in the functional setup). In contrast to the above references, we will formulate the result for SMSP rather than simple MSP.

Theorem 2.3 (de Haan and Resnick (1977), Giné et al. (1990))**.** *Let $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ be an SMSP. There is a stochastic process $\boldsymbol{Z} = (Z_s)_{s\in S}$ with sample paths in $\bar{C}^+(S)$ which satisfies*

$$\|\boldsymbol{Z}\|_\infty = m \in [1,\infty) \text{ almost surely and } E(Z_s) = 1, \quad s \in S, \tag{2.6}$$

such that for compact subsets $K_1,\dots,K_d$ of S and $x_1,\dots,x_d \le 0$, $d \in \mathbb{N}$,

$$P\left(\eta_s \le x_j,\ s \in K_j,\ j = 1,\dots,d\right) = \exp\left(-E\left(\max_{1\le j\le d}\left(|x_j| \max_{s\in K_j} Z_s\right)\right)\right). \tag{2.7}$$

Conversely, every stochastic process $\boldsymbol{Z} \in \bar{C}^+(S)$ with (2.6) *gives rise to an SMSP via* (2.7)*, and $\boldsymbol{Z}$ is called a* generator.

The requirement $\|\boldsymbol{Z}\|_\infty = m$ almost surely arises because we have equipped the space $C(S)$ with the supremum norm to turn it into a complete and separable metric space. In the multivariate case, however, the choice of the norm does not matter since all norms are equivalent on $\mathbb{R}^d$. We can therefore replace the norm $\|\cdot\|_\infty$ by any norm $\|\cdot\|$, if S is a finite set.

Actually, it is not necessary to demand that the norm of a generator is almost surely constant, but only that it is integrable. Hence (2.6) can be replaced by the weaker assumption

$$E\left(\sup_{s\in S} Z_s\right) < \infty \text{ and } E(Z_s) = 1, \quad s \in S, \tag{2.6$'$}$$

cf. de Haan and Ferreira (2006, Corollary 9.4.5). Note that in case S is a finite set, $E(Z_s) = 1$ implies $E(\sup_{s\in S} Z_s) < \infty$. It depends on the application whether to use (2.6) or (2.6′) for the definition of a generator. Of course, it is more convenient to verify (2.6′) in order to provide a generator, whereas it can also be useful to know that there exists a generator whose supremum is constant almost surely. Every generator gives rise to the following definition of a D-norm, irrespective of whether condition (2.6) or (2.6′) serve as basis for the definition.

Definition 2.4. Let $\boldsymbol{Z} = (Z_s)_{s\in S}$ be a generator process with sample paths in $\bar{C}^+(S)$. The mapping

$$\|\cdot\|_D : E(S) \to [0,\infty), \quad f \mapsto \|f\|_D := E\left(\sup_{s\in S} |f(s)|\, Z_s\right),$$

is called *D-norm with generator $\boldsymbol{Z}$*. Indeed, $\|\cdot\|_D$ defines a norm on the linear space $E(S)$.

It is obvious from the definition that every D-norm $\|\cdot\|_D$ is *standardized*, i. e.

$$\left\|\mathbb{1}_{\{s\}}\right\|_D = 1, \qquad s \in S,$$

as well as *monotone*, that is, for arbitrary $f, g \in E(S)$,

$$|f(s)| \le |g(s)|, \quad s \in S \implies \|f\|_D \le \|g\|_D .$$

Clearly, the supremum norm $\|f\|_\infty = \sup_{s\in S} |f(s)|$ defines a D-norm, for instance induced by the trivial generator $\boldsymbol{Z} \equiv 1$. Beyond that, the supremum norm is generated by any constant stochastic process $Z_s = Z$, $s \in S$ that satisfies $E(Z) = 1$. Thus, the distribution of a generator is not unique. However, the number

$$\|\mathbf{1}\|_D = E\left(\sup_{s\in S} Z_s\right) < \infty$$

is obviously uniquely determined, which justifies the term *generator constant* for $\|\mathbf{1}\|_D$. It is an easy task to verify

$$\|f\|_\infty \le \|f\|_D \le \|\mathbf{1}\|_D \|f\|_\infty , \qquad f \in E(S), \tag{2.8}$$

see Hofmann (2012, Lemma 2.6). These inequalities are of particular interest if S is not a finite set. It is well-known that all norms on $\mathbb{R}^d$ are equivalent to each other which is not true on, say, $C([0,1])$. However, (2.8) actually shows that every D-norm is equivalent to the supremum norm. In particular, the common L^p-norms

$$\|f\|_p := \left(\int_0^1 |f(s)|^p \, \mathrm{d}s\right)^{1/p}, \qquad f \in C([0,1]),$$

do not define D-norms, in distinction from the case $S = \{1, \dots, d\}$, where we will see that the L^p-norms define D-norms indeed.

A slight modification of Theorem 2.3 now yields a representation of the distribution of SMSP via D-norms.

Corollary 2.5 (Aulbach et al. (2013, Lemma 2))**.** *Let $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ be an SMSP with generator $\boldsymbol{Z} = (Z_s)_{s\in S}$. Then for every $f \in \bar{E}^-(S)$,*

$$P(\boldsymbol{\eta} \le f) = \exp\left(-E\left(\sup_{s\in S} |f(s)|\, Z_s\right)\right) = \exp\left(-\|f\|_D\right). \tag{2.9}$$

Conversely, every generator $\boldsymbol{Z}$ gives rise to an SMSP via (2.9).

We prefer the space $E(S)$ to the space of all continuous functions on S because it allows the incorporation of the finite-dimensional distributions in the representation $P(\boldsymbol{\eta} \le f)$, even if S is e. g. a compact interval such as $[0,1]$. Take an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ with generator $\boldsymbol{Z} = (Z_s)_{s\in S}$ and D-norm $\|\cdot\|_D$. Choose pairwise different indices $s_1, \dots, s_d \in S$, $d \in \mathbb{N}$. Then $(\eta_{s_1}, \dots, \eta_{s_d})$ defines a standard max-stable random vector with pertaining D-norm

$$\|\boldsymbol{x}\|_{D_{s_1,\dots,s_d}} = E\left(\max_{1\le j\le d}\left(|x_j|\, Z_{s_j}\right)\right), \qquad \boldsymbol{x} = (x_1, \dots, x_d) \le \mathbf{0}.$$

Put $f(\cdot) = \sum_{j=1}^d x_j \mathbb{1}_{\{s_j\}}(\cdot) \in \bar{E}^-(S)$. Then

$$\begin{aligned} P(\boldsymbol{\eta} \le f) &= \exp\left(-\|f\|_D\right) \\ &= \exp\left(-E\left(\sup_{s\in S}\left(|f(s)|\, Z_s\right)\right)\right) \\ &= \exp\left(-E\left(\max_{1\le j\le d}\left(|x_j|\, Z_{s_j}\right)\right)\right) \\ &= \exp\left(-\|\boldsymbol{x}\|_{D_{s_1,\dots,s_d}}\right) \\ &= P(\eta_{s_1} \le x_1, \dots, \eta_{s_d} \le x_d). \end{aligned} \tag{2.10}$$

The function

$$G : \bar{E}^-(S) \to [0,1], \qquad f \mapsto G(f) = P(\boldsymbol{\eta} \le f) = \exp\left(-\|f\|_D\right),$$

is called the *distribution function* (df) of the SMSP $\boldsymbol{\eta}$. Clearly, if $S = \{1, \dots, d\}$, this definition coincides with the regular definition of a df on $\mathbb{R}^d$. In particular, in that case, it is obvious from the univariate margins that G is continuous. Interestingly, this result can be generalized to arbitrary index sets S. The function G is continuous with respect to the supremum norm, resulting in the fact that $G(f) = P(\boldsymbol{\eta} < f)$ for every $f \in \bar{E}^-(S)$, see Aulbach et al. (2013, Lemma 5). Just like in the multivariate case, the function G fully determines the distribution of $\boldsymbol{\eta}$ since we have seen previously that the finite-dimensional distributions are embedded by the choice of the function space $\bar{E}^-(S)$.

We end this section with the observation that the set of all D-norms forms a convex set.

Lemma 2.6. *Let $\|\cdot\|_{D_1}$ and $\|\cdot\|_{D_2}$ be two D-norms on $E(S)$. Then*

$$\|\cdot\|_{\lambda D_1 + (1-\lambda) D_2} := \lambda \|\cdot\|_{D_1} + (1-\lambda) \|\cdot\|_{D_2}$$

also defines a D-norm on $E(S)$ for each $\lambda \in [0,1]$.

Proof. Let X be a Bernoulli distributed random variable with success probability $P(X = 1) = \lambda$, independent of the generators $\boldsymbol{Z}^{(1)}$ and $\boldsymbol{Z}^{(2)}$ of $\|\cdot\|_{D_1}$ and $\|\cdot\|_{D_2}$. Then

$$\boldsymbol{Z}^* := X\boldsymbol{Z}^{(1)} + (1-X)\boldsymbol{Z}^{(2)}$$

is a generator of $\|\cdot\|_{\lambda D_1 + (1-\lambda) D_2}$, since (2.6′) is clearly satisfied, and

$$\begin{aligned} E\left(\sup_{s\in S} |f(s)|\, Z_s^*\right) &= E\left(\mathbb{1}_{\{X=1\}} \sup_{s\in S} |f(s)|\, Z_s^*\right) + E\left(\mathbb{1}_{\{X=0\}} \sup_{s\in S} |f(s)|\, Z_s^*\right) \\ &= \lambda E\left(\sup_{s\in S} |f(s)|\, Z_s^{(1)}\right) + (1-\lambda) E\left(\sup_{s\in S} |f(s)|\, Z_s^{(2)}\right). \end{aligned}$$

□

Examples of generators and D-norms

So far we have only seen one example of a D-norm, namely the supremum norm $\|\cdot\|_\infty$. Clearly, $\|\cdot\|_\infty$ is the D-norm of an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ iff the univariate margins of $\boldsymbol{\eta}$ are completely dependent, i. e. $\eta_s = \eta_t$ almost surely for $s, t \in S$. This makes $\|\cdot\|_\infty$ a rather trivial D-norm, as the complete dependence case basically reflects the univariate setup, where η is simply a standard negative exponentially distributed random variable. Convergence to the D-norm of complete dependence can be characterized very easily:

Lemma 2.7 (Aulbach et al. (2013, Lemma 3)). *Let $\|\cdot\|_{D_n}$, $n \in \mathbb{N}$, be a sequence of D-norms on $E(S)$. Then*

$$\|f\|_{D_n} \to_{n\to\infty} \|f\|_\infty \text{ for all } f \in E(S) \iff \|\mathbf{1}\|_{D_n} \to_{n\to\infty} \|\mathbf{1}\|_\infty = 1.$$

As an application of the preceding Lemma, we consider an SMSP with sample paths in $C([0,1])$, and show how it can be stretched to an SMSP with complete dependent univariate margins.

Example 2.8. Let $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ be a SMSP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$, D-norm $\|\cdot\|_D$, and generator constant $\|\mathbf{1}\|_D$. Choose a sequence of intervals $[a_{n+1}, b_{n+1}] \subset [a_n, b_n] \subset [0,1]$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} b_n - a_n = 0$. By Cantor's intersection theorem, there is a uniquely determined $c_0 \in \bigcap_{n\in\mathbb{N}}[a_n, b_n]$. Define for $n \in \mathbb{N}$ a stochastic process $\boldsymbol{\eta}^{(n)} = \left(\eta_s^{(n)}\right)_{s\in[0,1]} \in C[0,1]$ by

$$\eta_s^{(n)} = \eta_{a_n+s(b_n-a_n)}, \quad s \in [0,1].$$

More vividly, we select a piece $(\eta_s)_{s\in[a_n,b_n]}$ of the original SMSP $\boldsymbol{\eta}$, and stretch it to a stochastic process defined on the whole interval $[0,1]$. We then have for $f \in \bar{E}^-[0,1]$

$$\begin{aligned}
P\left(\boldsymbol{\eta}^{(n)} \le f\right) &= P\left(\eta_{a_n+s(b_n-a_n)} \le f(s), s \in [0,1]\right)\\
&= P\left(\eta_t \le f\left(\frac{t-a_n}{b_n-a_n}\right),\ t \in [a_n, b_n]\right)\\
&= \exp\left(-E\left(\sup_{t\in[a_n,b_n]}\left(\left|f\left(\frac{t-a_n}{b_n-a_n}\right)\right| Z_t\right)\right)\right)\\
&= \exp\left(-E\left(\sup_{s\in[0,1]}\left(|f(s)|\, Z_{a_n+s(b_n-a_n)}\right)\right)\right),
\end{aligned}$$

which immediately implies that $\boldsymbol{\eta}^{(n)}$, $n \in \mathbb{N}$, is an SMSP itself with generator $\boldsymbol{Z}^{(n)} = (Z_{a_n+s(b_n-a_n)})_{s\in[0,1]}$. Denote by $\|\cdot\|_{D_n}$ the D-norm corresponding to $\boldsymbol{\eta}^{(n)}$. We have for $n \in \mathbb{N}$

$$\|\mathbf{1}\|_{D_{n+1}} = E\left(\sup_{t\in[a_{n+1},b_{n+1}]} Z_t\right)$$

$$\leq E\left(\sup_{t\in[a_n,b_n]} Z_t\right)$$
$$= \|\mathbf{1}\|_{D_n}$$
$$\leq E\left(\sup_{t\in[0,1]} Z_t\right) = \|\mathbf{1}\|_D .$$

Hence, we obtain a decreasing sequence of generator constants

$$\|\mathbf{1}\|_D \geq \|\mathbf{1}\|_{D_1} \geq \|\mathbf{1}\|_{D_2} \geq \ldots \geq \|\mathbf{1}\|_{D_n} \geq \ldots \geq 1.$$

Therefore, $m_0 := \lim_{n\to\infty} \|\mathbf{1}\|_{D_n}$ exists, and $m_0 \geq 1$. Define $Y_n := \sup_{t\in[a_n,b_n]} Z_t$. Then $Y_n \downarrow Z_{c_0}$, and by the monotone convergence theorem,

$$\|\mathbf{1}\|_{D_n} = E(Y_n) \downarrow E(Z_{c_0}) = 1 = m_0.$$

We have shown $\|\mathbf{1}\|_{D_n} \to_{n\to\infty} 1$, which implies $\|f\|_{D_n} \to_{n\to\infty} \|f\|_\infty$ for every $f \in E[0,1]$ by Lemma 2.7.

The counterpart of *complete dependent* univariate margins is the case where all components η_s, $s \in S$, are *independent*. This in turn can only occur if S is finite. If, for instance, $S = [0,1]$, the assumption that all univariate margins are independent of each other contradicts the continuity of the sample paths.

This is a fundamental difference between max-stable random vectors and 'actual' max-stable processes. In fact, it is not easy to derive closed form representations for functional D-norms, whereas this is no problem in the multivariate setup. In the following, we will gather some examples of parametric families of D-norms, always assuming that $S = \{1,\ldots,d\}$. Note that we will still use the term SMSP, even though we actually consider random vectors.

Example 2.9 (Complete dependence and independence). Complete dependence of the univariate margins can be characterized via the supremum norm $\|\cdot\|_\infty$ which is, for instance, induced by the constant generator $\boldsymbol{Z} \equiv 1$. Also, it has been pointed out before that $\|\cdot\|_D \geq \|\cdot\|_\infty$. On the other hand, the SMSP $\boldsymbol{\eta} = (\eta_1,\ldots,\eta_d)$ has independent components, iff

$$P(\eta_1 \leq x_1,\ldots,\eta_d \leq x_d) = \prod_{i=1}^d P(\eta_i \leq x_i)$$

$$= \exp\left(-\sum_{i=1}^{d} |x_i|\right) = \exp\left(-\|\boldsymbol{x}\|_1\right), \qquad \boldsymbol{x} = (x_1, \dots, x_d) \le \mathbf{0}.$$

Thus, the independence case is characterized by the *sum norm* $\|\cdot\|_1$. It is further easy to see that $\|\cdot\|_D \le \|\cdot\|_1$ for any D-norm $\|\cdot\|_D$. To summarize, we have

$$\|\cdot\|_\infty \le \|\cdot\|_D \le \|\cdot\|_1,$$

where $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are D-norms themselves characterizing the two extremal cases of complete dependent and independent univariate margins. Hence the D-norm models the *dependence structure* within the components of the SMSP $\boldsymbol{\eta}$, which justifies the letter D. It is well-known from Takahashi's theorem that $\|\cdot\|_D = \|\cdot\|_\infty$ iff $\|\mathbf{1}\|_D = 1$ and $\|\cdot\|_D = \|\cdot\|_1$ iff $\|\mathbf{1}\|_D = d$, see Takahashi (1988) or Falk et al. (2011, Theorem 4.4.1). This makes the generator constant $\|\mathbf{1}\|_D \in [1, d]$ a popular dependence parameter, in the literature more commonly known as the *extremal coefficient*, cf. Smith (1990).

Now let us focus on the generator of $\|\cdot\|_1$. Different to the complete dependence case, where $\|\cdot\|_\infty$ has a generator that concentrates on the vector $\mathbf{1}$, the sum norm can be generated by a discrete random vector with all its mass on the axes. More precisely, let $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ be a random permutation of the vector $(d, 0, \dots, 0)$, such that $P(Z_i = d) = 1/d$, $i = 1, \dots, d$. It is not difficult to verify that

$$E\left(\max_{i=1,\dots,d} |x_i|\, Z_i\right) = \|\boldsymbol{x}\|_1, \qquad \boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

In fact, it can be shown quite easily that $\boldsymbol{\eta} = (\eta_1, \dots, \eta_d)$ is an SMSP with independent margins iff the (not necessarily uniquely determined) generator $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ satisfies

$$P(\min(Z_i, Z_j) = 0) = 1, \qquad 1 \le i \ne j \le d, \tag{2.11}$$

see Hofmann (2012).

Having established the two extremal cases, we will now give some examples of parametric families of multivariate D-norms covering the whole range from complete dependence to independence. We begin with the *Marshall-Olkin-norm*, which is a convex combination of the sum and the supremum norm.

Example 2.10 (Marshall-Olkin model). The *Marshall-Olkin* norm on $\mathbb{R}^d$ is defined by

$$\|\cdot\|_{M_\lambda} := \lambda \|\cdot\|_\infty + (1 - \lambda) \|\cdot\|_1, \qquad \lambda \in [0, 1].$$

According to Lemma 2.6, $\|\cdot\|_{M_\lambda}$ defines a D-norm and it is generated by

$$\boldsymbol{Z} := X\mathbf{1} + (1-X)\tilde{\boldsymbol{Z}},$$

where $\tilde{\boldsymbol{Z}}$ is a generator of $\|\cdot\|_1$ as in Example 2.9, and X is a Bernoulli random variable with success probability $P(X=1)=\lambda$, independent of $\tilde{\boldsymbol{Z}}$. Clearly, the Marshall-Olkin model covers the whole range from complete dependence ($\lambda = 1$) to independence ($\lambda = 0$).

Next we show that the L^p-norm is a D-norm, in the literature also known as the *logistic model.*

Example 2.11 (Logistic model). Consider for $p \in (1,\infty)$ the L^p-norm on $\mathbb{R}^d$, i.e.

$$\|\boldsymbol{x}\|_p := \left(\sum_{i=1}^d |x_i|^p\right)^{1/p}. \tag{2.12}$$

Let $\tilde{Z}_1, \dots, \tilde{Z}_d$ be independent and identically Fréchet distributed with parameter $p \in (1,\infty)$, that is,

$$P(X_1 \le t) = \exp(-t^{-p}), \qquad t > 0.$$

It is well known that $E(\tilde{Z}_1) = \Gamma(1-p^{-1})$, where $\Gamma(\cdot)$ denotes the gamma function. Define a random vector $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ by

$$Z_i := \frac{\tilde{Z}_i}{\Gamma(1-p^{-1})}, \qquad i = 1, \dots, d.$$

Then clearly $\boldsymbol{Z} \ge 0$ and $E(Z_i) = 1$, $i = 1, \dots, d$. Thus, $\boldsymbol{Z}$ defines a generator on $\mathbb{R}^d$, and it induces the L^p-norm indeed: Take $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$. We have by the independence of $\tilde{Z}_1, \dots, \tilde{Z}_d$

$$\begin{aligned}
E\left(\max_{i=1,\dots,d} |x_i|\, Z_i\right) &= \frac{1}{\Gamma(1-p^{-1})} E\left(\max_{i=1,\dots,d} |x_i|\, \tilde{Z}_i\right) \\
&= \frac{1}{\Gamma(1-p^{-1})} \int_0^\infty 1 - P\left(\max_{i=1,\dots,d} |x_i|\, \tilde{Z}_i \le t\right)\, \mathrm{d}t \\
&= \frac{1}{\Gamma(1-p^{-1})} \int_0^\infty 1 - P\left(\tilde{Z}_1 \le \frac{t}{|x_1|}, \dots, \tilde{Z}_d \le \frac{t}{|x_d|}\right)\, \mathrm{d}t \\
&= \frac{1}{\Gamma(1-p^{-1})} \int_0^\infty 1 - \exp\left(-t^{-p}\left(\sum_{i=1}^d |x_i|^p\right)\right)\, \mathrm{d}t \\
&= \frac{1}{\Gamma(1-p^{-1})} \|\boldsymbol{x}\|_p \int_0^\infty 1 - \exp\left(-u^{-p}\right)\, \mathrm{d}u
\end{aligned}$$

$$= \|\boldsymbol{x}\|_p \,.$$

The generator constant $\|\mathbf{1}\|_p$ pertaining to the D-norm $\|\cdot\|_p$ on $\mathbb{R}^d$ is given by $d^{1/p}$, obviously. This is also immediate by the well-known fact that the Fréchet distribution is max-stable, that is

$$\max(\tilde{Z}_1, \ldots, \tilde{Z}_d) =_{\mathcal{D}} d^{1/p} \tilde{Z}_1.$$

Furthermore, $\|\mathbf{1}\|_p$ is a strictly monotone decreasing function in p with the limits d and 1 as p tends to 1 and ∞, respectively. Hence, the whole range from complete dependence to independence is covered by the logistic model.

The next model is built on a Weibull distributed generator. Recall the useful formulas

$$\max_{i=1,\ldots,d} a_i = \sum_{\emptyset \neq T \subseteq \{1,\ldots,d\}} (-1)^{|T|-1} \min_{i \in T} a_i \quad \text{and} \quad \min_{i=1,\ldots,d} a_i = \sum_{\emptyset \neq T \subseteq \{1,\ldots,d\}} (-1)^{|T|-1} \max_{i \in T} a_i, \tag{2.13}$$

which hold for arbitrary vectors $\boldsymbol{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$. Here, $|T|$ denotes the number of elements of the set T. Furthermore, we introduce for a nonempty subset $T = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, d\}$ with $i_1 < \cdots < i_k$ the notation

$$\boldsymbol{a}_T := (a_{i_1}, \ldots, a_{i_k}) \in \mathbb{R}^k.$$

Example 2.12 (Weibull model). Let $\tilde{Z}_1, \ldots, \tilde{Z}_d$ be independent and identically Weibull distributed with parameter $\alpha \in (0, \infty)$, that is,

$$P(\tilde{Z}_1 > t) = \exp\left(-t^\alpha\right), \qquad t > 0.$$

It is well known that

$$E(\tilde{Z}_1) = \int_0^\infty \exp\left(-t^\alpha\right)\, dt = \Gamma(1 + \alpha^{-1}).$$

Therefore, the random vector $\boldsymbol{Z} = (Z_1, \ldots, Z_d)$ defined by

$$Z_i := \frac{\tilde{Z}_i}{\Gamma(1 + \alpha^{-1})}, \qquad i = 1, \ldots, d,$$

is a generator. Denote by $\|\cdot\|_{W_\alpha}$ the D-norm generated by $\boldsymbol{Z}$. Then for every $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$ with $x_i \neq 0$, $i = 1,\dots,d$,

$$\|\boldsymbol{x}\|_{W_\alpha} = \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \left(\|1/\boldsymbol{x}_T\|_\alpha\right)^{-1}, \qquad \alpha > 0.$$

Here we presume that $\|\cdot\|_\alpha$ is exactly defined as in (2.12), being aware of the fact that it is actually not a norm in the case $\alpha \in (0,1)$. Now (2.13) implies

$$\begin{aligned}
E\left(\max_{i=1,\dots,d} |x_i|\, Z_i\right) &= \frac{1}{E(\tilde{Z}_1)} \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} E\left(\min_{i\in T} |x_i|\, \tilde{Z}_i\right)\\
&= \frac{1}{E(\tilde{Z}_1)} \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \int_0^\infty P\left(\min_{i\in T} |x_i|\, \tilde{Z}_i > t\right) \,\mathrm{d}t\\
&= \frac{1}{E(\tilde{Z}_1)} \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \int_0^\infty P\left(\tilde{Z}_i > t/|x_i|, i \in T\right) \,\mathrm{d}t\\
&= \frac{1}{E(\tilde{Z}_1)} \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \int_0^\infty \exp\left(-t^\alpha \sum_{i\in T} |x_i|^{-\alpha}\right) \,\mathrm{d}t\\
&= \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \left(\|1/\boldsymbol{x}_T\|_\alpha\right)^{-1}.
\end{aligned}$$

The generator constant of $\|\cdot\|_{W_\alpha}$ is given by

$$\|\mathbf{1}\|_{W_\alpha} = \sum_{k=1}^d \binom{d}{k} (-1)^{k-1} k^{-1/\alpha}, \qquad \alpha > 0.$$

We have

$$\lim_{\alpha\to 0} \|\mathbf{1}\|_{W_\alpha} = d + \sum_{k=2}^d \binom{d}{k} (-1)^{k-1} \lim_{\alpha\to 0} k^{-1/\alpha} = d,$$

as well as

$$\lim_{\alpha\to\infty} \|\mathbf{1}\|_{W_\alpha} = \sum_{k=1}^d \binom{d}{k} (-1)^{k-1} = -\sum_{k=0}^d \binom{d}{k} (-1)^k 1^{d-k} + 1 = 1,$$

which means again that the Weibull model covers the whole range from independence to complete dependence.

In contrast to the last two examples, the next parametric family of D-norms is induced by a generator whose components are bounded.

Example 2.13. Choose $\gamma > 0$ and take iid random variables $\tilde{Z}_1, \dots, \tilde{Z}_d$ with df

$$P(\tilde{Z}_1 \le t) = t^\gamma, \qquad t \in [0,1].$$

Then $E(\tilde{Z}_1) = \gamma/(\gamma+1)$. In order to obtain a generator $\boldsymbol{Z} = (Z_1, \dots, Z_d)$, we normalize

$$Z_i := \frac{\gamma+1}{\gamma}\tilde{Z}_i, \qquad i = 1, \dots, d.$$

The D-norm generated by $\boldsymbol{Z}$ is given by

$$\|\boldsymbol{x}\|_{P_\gamma} = \frac{\gamma+1}{\gamma}\left(\|\boldsymbol{x}\|_\infty - \frac{\|\boldsymbol{x}\|_\infty^{d\gamma+1}}{(d\gamma+1)\,|x_1|^\gamma \cdot \ldots \cdot |x_d|^\gamma}\right)$$

for every $\boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ with $x_i \neq 0$, $i = 1, \dots, d$, since for such $\boldsymbol{x}$

$$\begin{aligned} E\left(\max_{i=1,\dots,d} |x_i|\, Z_i\right) &= \frac{\gamma+1}{\gamma}\int_0^{\|\boldsymbol{x}\|_\infty} 1 - P\left(\max_{i=1,\dots,d} |x_i|\, \tilde{Z}_i \le t\right)\, \mathrm{d}t \\ &= \frac{\gamma+1}{\gamma}\int_0^{\|\boldsymbol{x}\|_\infty} 1 - P\left(\tilde{Z}_i \le t/|x_i|\,, i = 1, \dots, d\right)\, \mathrm{d}t \\ &= \frac{\gamma+1}{\gamma}\int_0^{\|\boldsymbol{x}\|_\infty} 1 - \frac{t^{d\gamma}}{|x_1|^\gamma \cdot \ldots \cdot |x_d|^\gamma}\, \mathrm{d}t \\ &= \frac{\gamma+1}{\gamma}\left(\|\boldsymbol{x}\|_\infty - \frac{\|\boldsymbol{x}\|_\infty^{d\gamma+1}}{(d\gamma+1)\,|x_1|^\gamma \cdot \ldots \cdot |x_d|^\gamma}\right). \end{aligned}$$

The generator constant of $\|\cdot\|_{P_\gamma}$ on $\mathbb{R}^d$ is given by

$$\|\mathbf{1}\|_{P_\gamma} = \frac{d\gamma+d}{d\gamma+1}, \qquad \gamma > 0.$$

In particular, $\|\mathbf{1}\|_{P_\gamma} \to_{\gamma\to 0} d$ and $m\,\|\mathbf{1}\|_{P_\gamma} \to_{\gamma\to\infty} 1$, and $\|\mathbf{1}\|_{P_\gamma}$ is strictly monotonically decreasing in γ for $d > 1$.

Another example of a D-norm with a discrete generator is the *Bernoulli model.* Despite its very simple nature, it is still suitable to cover the whole dependence range.

Example 2.14 (Bernoulli model). Let $\tilde{Z}_1, \dots, \tilde{Z}_d$ be iid Bernoulli random variables with success probability $\beta \in (0,1]$. The standardization $Z_i := \tilde{Z}_i/\beta$, $i = 1, \dots, d$, yields a generator. The corresponding D-norm $\|\cdot\|_{B_\beta}$ is

$$\|\boldsymbol{x}\|_{B_\beta} = \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} \beta^{|T|-1}(1-\beta)^{d-|T|} \|\boldsymbol{x}_T\|_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

since we have

$$E\left(\max_{i=1,\dots,d}(|x_i|\, Z_i)\right) = \frac{1}{\beta} \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} \max_{i \in T} |x_i|\, P\left(\tilde{Z}_i = 1, i \in T, \tilde{Z}_j = 0, j \in T^{\complement}\right).$$

Note that $\|\boldsymbol{x}\|_{B_1} = \|\boldsymbol{x}\|_\infty$ and $\|\boldsymbol{x}\|_{B_0} := \lim_{q\to 0} \|\boldsymbol{x}\|_{D_q} = \|\boldsymbol{x}\|_1$. The generator constant $\|\mathbf{1}\|_{B_\beta}$ can easily be computed by

$$\|\mathbf{1}\|_{B_\beta} = \frac{1}{\beta} E\left(\max_{i=1,\dots,d} \tilde{Z}_i\right) = \frac{1}{\beta} P\left(\max_{i=1,\dots,d} \tilde{Z}_i = 1\right) = \frac{1}{\beta}\left(1 - (1-\beta)^d\right).$$

The bivariate case yields the Marshall-Olkin norm again, yet the distribution of the generator is clearly not the same as the one we have discussed in Example 2.10.

The calculation of concrete examples of functional D-norms is usually a very difficult task. Nevertheless, there exist numerous popular models of max-stable processes with a non-finite index set, two of which are introduced in the following. However, it is often only possible to give closed formulas of the bivariate marginal distributions and hence bivariate D-norms, which is still of value in many practical purposes.

In the next two examples, with a slight abuse of notation, we will discuss SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in\mathbb{R}}$ with the domain $\mathbb{R}$ instead of the compact metric space S. Note that the finite dimensional distributions of $\boldsymbol{\eta}$ still define standard max-stable random vectors.

We will need the following terminology. A stochastic process $\boldsymbol{X} = (X_s)_{s\in\mathbb{R}}$ is called *stationary*, if $\boldsymbol{X} =_{\mathcal{D}} (X_{s+h})_{s\in\mathbb{R}}$ for any $h \in \mathbb{R}$. Different to that, a stochastic process $\boldsymbol{Y} = (Y_s)_{s\in\mathbb{R}}$ has *stationary increments*, if the distribution of $(Y_{s+h} - Y_h)_{s\in\mathbb{R}}$ does not depend on the choice of $h \in \mathbb{R}$.

Example 2.15 (Brown-Resnick model). This model was originally created by Brown and Resnick (1977), and developed by Kabluchko et al. (2009) for MSP $\boldsymbol{\vartheta} = (\vartheta_s)_{s\in\mathbb{R}}$ with Gumbel margins, i.e. $P(\vartheta_s \leq x) = \exp(-e^{-x})$, $x \in \mathbb{R}$. Note that the transformation to SMSP is straightforward since $-\exp(-\boldsymbol{\vartheta})$ is an SMSP if $\boldsymbol{\vartheta}$ is an MSP with Gumbel margins. Let $\boldsymbol{W} = (W_s)_{s\in\mathbb{R}}$ be a centered (i.e. $E(W_s) = 0$, $s \in \mathbb{R}$) Gaussian process with continuous sample paths and stationary increments. Denote by σ_s^2 the variance of W_s, $s \in \mathbb{R}$. Then

$$Z_s = \exp\left(W_s - \sigma_s^2/2\right), \qquad s \in \mathbb{R},$$

defines a generator of a stationary SMSP $\boldsymbol{\eta}$. The distribution of $\boldsymbol{W}$, and hence of $\boldsymbol{Z} = (Z_s)_{s\in\mathbb{R}}$ is fully determined by the variances σ_s^2 and the *variogram*

$$\gamma(h) = E\left((W_{s_0+h} - W_{s_0})^2\right).$$

Note that the choice of $s_0 \in \mathbb{R}$ in the upper formula does not matter since $\boldsymbol{W}$ has stationary increments. The bivariate distributions of $\boldsymbol{\eta}$ are described by the bivariate Brown-Resnick D-norms

$$\begin{aligned}\|(x,y)\|_{BR_h} &= E\left(\max\left(|x|\, Z_s, |y|\, Z_{s+h}\right)\right)\\ &= |x|\,\Phi\left(\frac{\sqrt{\gamma(h)}}{2} + \frac{\log(x/y)}{\sqrt{\gamma(h)}}\right) + |y|\,\Phi\left(\frac{\sqrt{\gamma(h)}}{2} + \frac{\log(y/x)}{\sqrt{\gamma(h)}}\right),\end{aligned}$$

where Φ is the df of the standard normal distribution, see Kabluchko et al. (2009, Remark 24). The bivariate generator constants are given by

$$\|(1,1)\|_{BR_h} = 2\Phi\left(\frac{\sqrt{\gamma(h)}}{2}\right).$$

Hence, $\|(1,1)\|_{BR_h} \to 1$ as $h \to 0$ and $\|(1,1)\|_{BR_h} \to 2$ as $|h| \to \infty$ iff $\gamma(h) \to \infty$ as $|h| \to \infty$.

Example 2.16 (Schlather model). Schlather (2002) proposed to consider the generator

$$Z_s = \sqrt{2\pi}\max(0, W_s), \qquad s \in \mathbb{R},$$

where $\boldsymbol{W} = (W_s)_{s\in\mathbb{R}}$ is a stationary standard Gaussian process (i. e. $\boldsymbol{W}$ is stationary and W_s is standard normally distributed for each $s \in S$) with correlation function $\rho(h) = \mathrm{Cov}(W_s, W_{s+h})$ and continuous sample paths. The pertaining bivariate Schlather D-norms are given by

$$\begin{aligned}\|(x,y)\|_{SCH_h} &= E\left(\max\left(|x|\, Z_s, |y|\, Z_{s+h}\right)\right)\\ &= \frac{1}{2}\left(|x| + |y|\right)\left(1 + \sqrt{1 - \frac{2(1+\rho(h))}{|x|\,|y|\,(1/\,|x| + 1/\,|y|)^2}}\right),\end{aligned}$$

cf. Schlather (2002, Formula (7)), and the generator constant is

$$\|(1,1)\|_{SCH_h} = 1 + \sqrt{\frac{1-\rho(h)}{2}}.$$

Note that $\|(1,1)\|_{SCH_h} \leq 1 + \sqrt{1/2} \approx 1.7$ if $\rho(h) \geq 0$, such that the independence case cannot be approximated by the Schlather model with a positive correlation function.

Exponent measure and angular measure

As mentioned before, there is a key element in the representation of multivariate and functional max-stable distributions, which is the *exponent measure*, or finally the *angular measure*. In our setup, these measures are hidden in the distribution of a generator of a max-stable process. Since we will use the properties of an exponent measure at some point, it makes sense to introduce it carefully, that is, we will deduce it from the distribution of a generator. Historically, it worked the other way around: First the so-called *max-id* distributions where characterized by means of the exponent measure (Balkema and Resnick (1977), Giné et al. (1990)). In the smaller class of max-stable distributions, it turned out that the exponent measure can be factorized into a radial part and an angular measure (de Haan and Resnick (1977), de Haan (1984), Giné et al. (1990)), which in turn gave rise to a generator.

We now closely follow the discussion in de Haan and Lin (2001), see also de Haan and Ferreira (2006), Hult and Lindskog (2005) and Davis and Mikosch (2008) in order to establish a complete and separable metric space on which the exponent measure will be defined. Let $\bar{C}_1^+(S)$ be the unit sphere in $\bar{C}^+(S)$, i. e. $\bar{C}_1^+(S) = \{f \in \bar{C}^+(S) : \|f\|_\infty = 1\}$. Define

$$\mathbb{E} := (0, \infty] \times \bar{C}_1^+(S), \tag{2.14}$$

where $(0, \infty]$ is equipped with the metric $\varrho(x, y) = |1/x - 1/y|$, $x, y \in (0, \infty]$, to turn $(0, \infty]$ (and thus $\mathbb{E}$) into a complete and separable metric space. By the transformation to polar coordinates

$$T : \bar{C}^+(S) \setminus \{\mathbf{0}\} \to (0, \infty) \times \bar{C}_1^+(S), \quad f \mapsto \left(\|f\|_\infty, f/\|f\|_\infty\right),$$

the spaces $\bar{C}^+(S) \setminus \{\mathbf{0}\}$ (equipped with the relative topology of $C(S)$) and $(0, \infty) \times \bar{C}_1^+(S)$ (equipped with the relative topology of $\mathbb{E}$) are homeomorphic, and hence

$$\mathbb{B}\left(\mathbb{E} \cap [(0, \infty) \times \bar{C}_1^+(S)]\right) = \mathbb{B}\left(T\left(\bar{C}^+(S) \setminus \{\mathbf{0}\}\right)\right).$$

Note that the relative compact sets in $(0, \infty]$ are those that are bounded away from 0.

Let $\boldsymbol{Z} = (Z_s)_{s\in S}$ be a generator process and assume it satisfies $\|\boldsymbol{Z}\|_\infty = m$ almost surely. Define a finite measure ρ on $\bar{C}_1^+(S)$ via

$$\rho(A) := mP\left(\boldsymbol{Z}/m \in A\right), \qquad A \in \mathbb{B}\left(\bar{C}_1^+(S)\right). \tag{2.15}$$

The measure ρ is usually referred to as the *angular measure* or the *spectral measure* in the literature. Now we can define a measure ν on the product space $\mathbb{E} = (0, \infty] \times \bar{C}_1^+(S)$ such that

$$\mathrm{d}\nu = r^{-2}\,\mathrm{d}r \times \mathrm{d}\rho. \tag{2.16}$$

This is the *exponent measure* corresponding to the generator $\boldsymbol{Z}$, and hence corresponding to some SMSP. The characteristic property of the exponent measure is its *homogeneity of order* -1, i. e.

$$\nu(cA) = c^{-1}\nu(A), \qquad c > 0, \quad A \in \mathbb{B}(\mathbb{E}). \tag{2.17}$$

The connection between the exponent measure ν and the D-norm $\|\cdot\|_D$ generated by $\boldsymbol{Z}$ is as follows: Define for $h \in E(S)$ the set $A_h := \{(r, g) \in \mathbb{E} : rg(s) > h(s) \text{ for some } s \in S\}$. Then we have for every $f \in E^+(S)$ by (2.15) and (2.16)

$$\begin{aligned}
\nu(A_{1/f}) &= \nu\left(\left\{(r, g) \in \mathbb{E} : r > 1/\sup_{s\in S} f(s)g(s)\right\}\right) \\
&= \int_{\bar{C}_1^+(S)} \int_{\left(\sup_{s\in S} f(s)g(s)\right)^{-1}}^{\infty} r^{-2}\,\mathrm{d}r\ \rho(\mathrm{d}g) \\
&= E\left(\sup_{s\in S} f(s)Z_s\right) \\
&= \|f\|_D\,.
\end{aligned}$$

Poisson point process representations

A powerful tool in the analysis of max-stable processes is the representation via Poisson point processes tracing back to de Haan (1984). To establish that characterization, we follow the monographs of Resnick (2008) and de Haan and Ferreira (2006).

Let $\mathcal{X}$ be a complete and separable metric space. A *counting measure* m on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$ is a measure of the form

$$m(A) = \sum_{k\in\mathbb{N}} \epsilon_{x_k}(A), \qquad A \in \mathbb{B}(\mathcal{X}),$$

where $(x_k)_{k\in\mathbb{N}}$ is a collection of points in $\mathcal{X}$, and, as before, $\epsilon_x(\cdot)$ denotes the Dirac measure on $\mathbb{B}(\mathcal{X})$. If m is *boundedly finite*, i. e. $m(B) < \infty$ for all bounded sets $B \in \mathbb{B}(\mathcal{X})$, then we call m a *point measure*. Denote by $M_p(\mathcal{X})$ the set of all point measures on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$, and let $\mathcal{M}_p(\mathcal{X})$ be the smallest σ-Algebra such that all the mappings

$$f_A : M_p(\mathcal{X}) \to [0, \infty], \quad m \mapsto f_A(m) = m(A),$$

are measurable, i. e. it contains all sets of the form

$$\{m \in M_p(\mathcal{X}) : m(A) \in B\}, \qquad A \in \mathbb{B}(\mathcal{X}),\ B \in \mathbb{B}([0, \infty]).$$

A *point process* N with *state space* $\mathcal{X}$ is a measurable mapping

$$N : (\Omega, \mathcal{A}) \to (M_p(\mathcal{X}), \mathcal{M}_p(\mathcal{X}))$$

from a probability space $(\Omega, \mathcal{A}, P)$, yielding a point measure $N(\omega, \cdot)$ for each $\omega \in \Omega$, and a random variable $N(\cdot, A)$ with values in $\mathbb{N} \cup \{\infty\}$ for each $A \in \mathbb{B}(\mathcal{X})$. Hence, $N(\omega, A)$ is the number of points in A for the realization ω. The distribution $P * N$ on $(M_p(\mathcal{X}), \mathcal{M}_p(\mathcal{X}))$ is somewhat difficult to imagine, but luckily it is completely determined by its finite-dimensional distributions, i. e. two point processes N_1, N_2 have the same distribution iff

$$(N_1(A_1), \ldots, N_1(A_k)) =_{\mathcal{D}} (N_2(A_1), \ldots, N_2(A_k))$$

for all bounded $A_1, \ldots, A_k \in \mathbb{B}(\mathcal{X})$, see Daley and Vere-Jones (2008, Section 9.2).

Without loss of generality, we can assume that a point process N is of the form

$$N(\omega, A) = \sum_{k\in\mathbb{N}} \epsilon_{X_k(\omega)}(A),$$

where $(X_k)_{k\in\mathbb{N}}$ are random variables with values in $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$, often referred to as the *points of* N, see Daley and Vere-Jones (2008, Section 9.1). Point processes whose univariate margins are Poisson distributed random variables are particularly interesting in extreme value theory.

Definition 2.17. A point process N with state space $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$ is called *Poisson point process* with *intensity measure* μ (shortly PPP(μ)) on $\mathbb{B}(\mathcal{X})$, if

(i) For $A \in \mathbb{B}(\mathcal{X})$ and $k \in \mathbb{N} \cup \{0\}$

$$P(N(A) = k) = \begin{cases} \exp\left(-\mu(A)\right) \mu(A)^k / k!, & \text{if } \mu(A) < \infty, \\ 0, & \text{if } \mu(A) = \infty. \end{cases}$$

(ii) If $A_1, \ldots, A_n \in \mathbb{B}(\mathcal{X})$ are pairwise disjoint sets, then $N(A_1), \ldots, N(A_m)$ are independent random variables.

Clearly, $E(N(A)) = \mu(A)$ if $\mu(A) < \infty$, which is why μ is also often called *mean measure*.

To discuss the representation of max-stable processes via Poisson point processes, it is convenient to consider *simple* max-stable processes with univariate standard Fréchet margins. So let $\boldsymbol{\xi}$ be a simple MSP with a generator $\boldsymbol{Z} = (Z_s)_{s \in S}$, i. e. (2.9) is satisfied by the SMSP $\boldsymbol{\eta} = -1/\boldsymbol{\xi}$. Assume the generator fulfills $\|\boldsymbol{Z}\|_\infty = m$ almost surely. Denote by ν and ρ the exponent measure and the angular measure which have been derived in the previous section. The Poisson point process representation of a simple MSP is

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k \in \mathbb{N}} \zeta^{(k)} \boldsymbol{V}^{(k)}, \tag{2.18}$$

where $\left(\left(\zeta^{(k)}, \boldsymbol{V}^{(k)}\right)\right)_{k \in \mathbb{N}}$ are the points of PPP(ν) on $\mathbb{E} = (0, \infty] \times \bar{C}_1^+(S)$, and the intensity measure ν satisfies $\mathrm{d}\nu = r^{-2}\, \mathrm{d}r \times \mathrm{d}\rho$, see de Haan and Ferreira (2006, Corollary 9.4.2). Alternatively, we have

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k \in \mathbb{N}} \zeta^{(k)} \boldsymbol{W}^{(k)}, \tag{2.19}$$

where $\left(\left(\zeta^{(k)}, \boldsymbol{W}^{(k)}\right)\right)_{k \in \mathbb{N}}$ are the points of PPP$(\tilde{\nu})$ on $(0, \infty] \times m\bar{C}_1^+(S)$ with mean measure $\mathrm{d}\tilde{\nu} = r^{-2}\, \mathrm{d}r \times \mathrm{d}(P * \boldsymbol{Z})$. Even though technically, ν and $\tilde{\nu}$ are two measures defined on different spaces, we call $\tilde{\nu}$ *exponent measure* as well with abuse of notation. Lastly, it can be shown (de Haan and Ferreira (2006, Corollary 9.4.5)) that (2.19) is equivalent to the representation

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k \in \mathbb{N}} \zeta^{(k)} \boldsymbol{Z}^{(k)}, \tag{2.20}$$

where $\left(\zeta^{(k)}\right)_{k \in \mathbb{N}}$ are the points of PPP $\left(r^{-2}\, \mathrm{d}r\right)$ on $(0, \infty]$, and $\boldsymbol{Z}^{(1)}, \boldsymbol{Z}^{(2)}, \ldots$ are independent copies of the generator $\boldsymbol{Z}$, independent of $\left(\zeta^{(k)}\right)_{k \in \mathbb{N}}$. In the representation (2.20), it is no longer necessary to suppose $\|\boldsymbol{Z}\|_\infty = m$ almost surely. Again, this assumption can be replaced by $E\left(\|\boldsymbol{Z}\|_\infty\right) < \infty$.

Remark 2.18. In the multivariate case $S = \{1, \ldots, d\}$, the exponent measure is usually defined on the unfactorized set $\mathbb{E} = [0, \infty]^d \setminus \{\boldsymbol{0}\}$. While the representations (2.19) and

(2.20) remain essentially the same (choose an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^d$ and replace $\bar{C}_1^+(S)$ by $\mathbb{S}_{\mathbb{E}} := \{\boldsymbol{y} \in \mathbb{E} : \|\boldsymbol{y}\| = 1\}$), we usually prefer

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k\in\mathbb{N}} \boldsymbol{\vartheta}^{(k)}, \tag{2.18$'$}$$

where $\left(\boldsymbol{\vartheta}^{(k)}\right)_{k\in\mathbb{N}}$ are the points of PPP(ν) on $\mathbb{E}$. The connection between ν and a generator $\boldsymbol{Z}$ with $\|\boldsymbol{Z}\| = m$ almost surely is

$$\nu\left(\{\boldsymbol{y} \in \mathbb{E} : \|\boldsymbol{y}\| > mr, m\boldsymbol{y}/\|\boldsymbol{y}\| \in A\}\right) = r^{-1}P(\boldsymbol{Z} \in A).$$

2.2 Dual D-norm functions and domain of attraction

The dual D-norm function

While there was a strong connection between the df of a max-stable process and its D-norm, the *dual D-norm function* will occur in the context of exceedances over a threshold. This will be of particular interest in Chapter 4 when complete records are discussed. We start with a Lemma taken from Aulbach et al. (2013, Lemma 6), albeit the proof therein assumes the supremum of the generator to be constant, which we want to omit. For the sake of completeness, we show the result again.

Lemma 2.19. *Let $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ be an SMSP with generator $\boldsymbol{Z} = (Z_s)_{s\in S}$. Then, for $f \in \bar{E}^-(S)$:*

(i) $P(\boldsymbol{\eta} > f) \geq 1 - \exp\left(-E\left(\inf_{s\in S}\left(|f(s)|\, Z_s\right)\right)\right)$.

(ii) $\lim_{h\downarrow 0} h^{-1}P(\boldsymbol{\eta} > hf) = E\left(\inf_{s\in S}\left(|f(s)|\, Z_s\right)\right)$.

Proof. As we have seen in (2.20), $\boldsymbol{\xi} := -1/\boldsymbol{\eta} =_{\mathcal{D}} \sup_{k\in\mathbb{N}} \zeta^{(k)}\boldsymbol{Z}^{(k)}$, where $\left(\zeta^{(k)}\right)_{k\in\mathbb{N}}$ are the points of PPP $\left(r^{-2}\,\mathrm{d}r\right)$ on $(0,\infty]$, and $\left(\boldsymbol{Z}^{(k)}\right)_{k\in\mathbb{N}}$ are independent copies of $\boldsymbol{Z}$, independent of the point process. The point process N, defined by

$$N(A) := \sum_{k=1}^{\infty} \epsilon_{\left(\zeta^{(k)},\boldsymbol{Z}^{(k)}\right)}(A), \qquad A \in \mathbb{B}\left((0,\infty] \times \bar{C}^+(S)\right),$$

has the same distribution as a Poisson point process N^* on $(0,\infty]\times\bar{C}^+(S)$ whose intensity measure ν satisfies $\mathrm{d}\nu = r^{-2}\,\mathrm{d}r \times \mathrm{d}(P * \boldsymbol{Z})$, cf. de Haan and Ferreira (2006, Lemma

9.4.7). Therefore, for $f \in \bar{E}^-(S)$,

$$\begin{aligned}
P(\exists s \in S : \eta_s \le f(s)) &= P\left(\exists s \in S\ \forall k \in \mathbb{N} : \zeta^{(k)} Z_s^{(k)} \le \frac{1}{|f(s)|}\right) \\
&\le P\left(\forall k \in \mathbb{N}\ \exists s \in S : \zeta^{(k)} Z_s^{(k)} \le \frac{1}{|f(s)|}\right) \\
&= P\left(\nexists k \in \mathbb{N} : \zeta^{(k)} > \sup_{s\in S} \frac{1}{|f(s)|\, Z_s^{(k)}}\right) \\
&= P\left(N\left(\left\{(r,z) \in (0,\infty] \times \bar{C}^+(S) :\ r > \sup_{s\in S} \frac{1}{|f(s)|\, z(s)}\right)\right\} = 0\right) \\
&= P\left(N^*\left(\left\{(r,z) \in (0,\infty] \times \bar{C}^+(S) :\ r > \sup_{s\in S} \frac{1}{|f(s)|\, z(s)}\right)\right\} = 0\right) \\
&= \exp\left(-\int_{\bar{C}^+(S)} \int_{(\inf_{s\in S}|f(s)|z(s))^{-1}}^{\infty} r^{-2}\, \mathrm{d}r\ (P * \boldsymbol{Z})(\mathrm{d}z)\right) \\
&= \exp\left(-E\left(\inf_{s\in S} (|f(s)|\, Z_s)\right)\right),
\end{aligned}$$

which proofs (i). The proof in Aulbach et al. (2013, Lemma 6) can now be adapted to obtain (ii). □

The limit in Lemma 2.19 (ii) now serves as definition for the dual D-norm function.

Definition 2.20. Let $\|\cdot\|_D$ be a D-norm on $E(S)$, generated by $\boldsymbol{Z} = (Z_s)_{s\in S}$. The mapping

$$≀ \cdot ≀_D : E(S) \to [0,\infty), \quad f \mapsto ≀ f ≀_D := E\left(\inf_{s\in S} |f(s)|\, Z_s\right)$$

is called the *dual D-norm function* corresponding to $\|\cdot\|_D$.

Despite the fact that the generator of $\|\cdot\|_D$ is not uniquely determined, Lemma 2.19 (ii) guarantees that the dual D-norm function $≀ \cdot ≀_D$ does not depend on the choice of the generator of $\|\cdot\|_D$. Therefore, the mapping

$$\|\cdot\|_D \to ≀ \cdot ≀_D$$

is well-defined, yet not one-to-one, since different D-norms can lead to the same dual D-norm function. In fact, let $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ be an SMSP with some corresponding generator $\boldsymbol{Z} = (Z_s)_{s\in S}$. Takahashi's theorem (Falk et al. (2011, Theorem 4.4.1)) implies that two components η_s, η_t are independent iff $E(\max(Z_s, Z_t)) = 2 = E(Z_s + Z_t)$, resulting in $\max(Z_s, Z_t) = Z_s + Z_t$ almost surely, which in turn yields $\min(Z_s, Z_t) = 0$ almost surely.

Hence, the dual D-norm function is always the constant zero if there are at least two independent components of the SMSP $\boldsymbol{\eta}$.

In the multivariate case, where

$$≀\, \boldsymbol{x} \,≀_D = E\left(\min_{i=1,\dots,d} |x_i|\, Z_i\right), \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

a simple connection between the functions $\|\cdot\|_D$ and $≀\cdot≀_D$ is given by the equations (2.13). A multivariate dual D-norm function is also known as the *tail copula* in the literature, which was introduced by Schmidt and Stadtmüller (2006), see also de Haan et al. (2008). However, they define the tail copula to be the limit in (2.25) below without providing an explicit formula.

We continue with some examples of multivariate dual D-norm functions.

Example 2.21 (Complete dependence and independence). We have seen before that

$$≀ \cdot ≀_1 = 0$$

is the least dual D-norm function, corresponding to the case of independent univariate margins, where $\|\cdot\|_D = \|\cdot\|_1$. On the other hand, $|x_j|\, Z_j \geq \min_{i=1,\dots,d} |x_i|\, Z_i$ for each $j = 1, \dots, d$, and hence

$$≀\, \boldsymbol{x} \,≀_\infty = \min_{1\leq i\leq d} |x_i|, \qquad \boldsymbol{x} \in \mathbb{R}^d,$$

is the largest dual D-norm function, corresponding to the perfect dependence case, where $\|\cdot\|_D = \|\cdot\|_\infty$. Hence, we have for an arbitrary dual D-norm function the bounds

$$0 = ≀ \cdot ≀_1 \leq ≀ \cdot ≀_D \leq ≀ \cdot ≀_\infty.$$

Clearly, these bounds are also valid in the functional case, where $≀\, f \,≀_\infty = \inf_{s\in S} |f(s)|$ is the dual D-norm function corresponding to the complete dependence of the univariate margins.

Example 2.22 (Logistic model). We have seen in Example 2.11 that

$$\|\boldsymbol{x}\|_\lambda = \left(\sum_{i=1}^d |x_i|^\lambda\right)^{1/\lambda}, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \lambda \in (1, \infty),$$

defines a D-norm. Now by (2.13), the corresponding dual D-norm function is

$$≀\, \boldsymbol{x} \,≀_\lambda = \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} (-1)^{|T|-1} \left\|\boldsymbol{x}_T\right\|_\lambda, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \lambda \in (1,\infty).$$

EXAMPLE 2.23 (Weibull model). In Example 2.12, we have defined a generator $\boldsymbol{Z} = (Z_1,\dots,Z_d)$ by taking independent Weibull distributed random variables $\tilde{Z}_1,\dots,\tilde{Z}_d$, i. e. $P(\tilde{Z}_1 > t) = \exp(-t^\alpha)$, $t > 0$, $\alpha > 0$, and putting $Z_i := \tilde{Z}_i/\Gamma(1+1/\alpha)$. In fact, we have already shown that

$$≀\boldsymbol{x}\,≀_{W_\alpha} = E\left(\min_{i=1,\dots,d} |x_i|\, Z_i\right) = \left(\left\|1/\boldsymbol{x}\right\|_\alpha\right)^{-1} \tag{2.21}$$

for all $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$ with $x_i \neq 0$, $i = 1,\dots,d$.

EXAMPLE 2.24 (Bernoulli model). In Example (2.14), we have derived a D-norm based on Bernoulli distributed generator components

$$\left\|\boldsymbol{x}\right\|_{B_\beta} = \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} \beta^{|T|-1}(1-\beta)^{d-|T|} \left\|\boldsymbol{x}_T\right\|_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \beta \in (0,1].$$

Analogously, one can show that the attendant dual D-norm function is

$$≀\, \boldsymbol{x} \,≀_{B_\beta} = \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} \beta^{|T|-1}(1-\beta)^{d-|T|}\, ≀\boldsymbol{x}_T\, ≀_\infty, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \beta \in (0,1].$$

Domain of attraction

Max-stable processes are of outstanding interest in extreme value theory. However, the max-stability property itself, though being very handy and simple, does not explain the relevance of MSP. The importance is much rather explained by the fact that max-stable distributions are the only possible limit distributions in a sequence of linearly standardized maxima of iid processes. Hence, max-stable distributions play the same role in extreme value theory as the normal distribution in the central limit theorem. We will now focus on that property.

Let $\boldsymbol{Y} = (Y_s)_{s\in S}$ be an MSP. This is equivalent to the existence of a stochastic process $\boldsymbol{X}$ with sample paths in $C(S)$, and norming functions $c_n \in C^+(S)$, $d_n \in C(S)$, $n \in \mathbb{N}$, such that

$$\max_{i=1,\dots,n} \frac{\boldsymbol{X}^{(i)} - d_n}{c_n} \to_{\mathcal{D}} \boldsymbol{Y}, \tag{2.2}$$

where $\boldsymbol{X}^1, \boldsymbol{X}^{(2)}, \dots$ are iid copies of $\boldsymbol{X}$. In that case, we say that $\boldsymbol{X}$ is in the *domain of attraction* of the MSP $\boldsymbol{Y}$, and we write $\boldsymbol{X} \in \mathcal{D}(\boldsymbol{Y})$. If F and G are the df of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, we also say that F is in the domain of attraction of G, and write $F \in \mathcal{D}(G)$.

Another concept, yet closely related to the latter one, was introduced by Aulbach et al. (2013), and pursued in Aulbach et al. (2015). We say that $\boldsymbol{X} \in C(S)$ is in the *functional domain of attraction* of an MSP $\boldsymbol{Y}$, if there are norming functions $c_n \in C(S)$, $d_n \in C^+(S)$, $n \in \mathbb{N}$, such that

$$\lim_{n\to\infty} P\left(\max_{i=1,\dots,n} \frac{\boldsymbol{X}^{(i)} - d_n}{c_n} \le f\right) = P\left(\frac{\boldsymbol{X} - c_n}{d_n} \le f\right)^n = P\left(\boldsymbol{Y} \le f\right), \qquad f \in E(S), \tag{2.22}$$

where $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ are iid copies of $\boldsymbol{X}$. In that case, we write $\boldsymbol{X} \in \mathrm{FDA}(\boldsymbol{Y})$ or $F \in \mathrm{FDA}(G)$, if F and G are the df of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. It has been shown in Aulbach et al. (2013, Proposition 5) that $\boldsymbol{X} \in \mathcal{D}(\boldsymbol{Y})$ implies $\boldsymbol{X} \in \mathrm{FDA}(\boldsymbol{Y})$. It is also not difficult to see that $\boldsymbol{X} \in \mathrm{FDA}(\boldsymbol{Y})$ yields the convergence of the finite-dimensional distributions. As a matter of fact, in the multivariate case $S = \{1, \dots, d\}$, there is no difference between the concepts of domain of attraction and functional domain of attraction.

The concept of domain of attraction is closely related to the theory of regular variation which relies on *weak hash* and *vague* convergence of measures. For details on these types of convergence, see Daley and Vere-Jones (2003, Appendix 2) or Resnick (2008, Section 3.4). For nice reviews on regular variation in function spaces, see Hult and Lindskog (2005), Davis and Mikosch (2008) and Dombry and Ribatet (2015).

Definition 2.25 (Weak hash convergence)**.** Let $\mathcal{X}$ be a complete and separable metric space. Let μ_n, $n \in \mathbb{N} \cup \{0\}$, be boundedly finite measures on $\mathbb{B}(\mathcal{X})$, i. e. $\mu_n(B) < \infty$ for every bounded set $B \in \mathbb{B}(\mathcal{X})$, $n \in \mathbb{N} \cup \{0\}$. Then $\mu_n \to_{w\#} \mu_0$, iff

$$\int_{\mathcal{X}} f(x)\ \mu_n(\mathrm{d}x) \to_{n\to\infty} \int_{\mathcal{X}} f(x)\ \mu_0(\mathrm{d}x)$$

for every $f \in \bar{C}^+(\mathcal{X})$ vanishing outside a bounded set, i. e. there is a bounded set $B \in \mathbb{B}(\mathcal{X})$ with $f(x) = 0$ for all $x \in B^{\complement}$. Equivalently, $\mu_n \to_{w\#} \mu_0$ iff

$$\mu_n(B) \to_{n\to\infty} \mu_0(B)$$

for all bounded sets $B \in \mathbb{B}(\mathcal{X})$ with $\mu_0(\partial B) = 0$.

REMARK 2.26. In case the space $\mathcal{X}$ is locally compact, weak hash convergence coincides with *vague convergence*. Let μ_n, $n \in \mathbb{N} \cup \{0\}$, be *Radon measures* on $\mathbb{B}(\mathcal{X})$, i. e. $\mu_n(K) < \infty$ for every compact set $K \in \mathbb{B}(\mathcal{X})$, $n \in \mathbb{N} \cup \{0\}$. Then μ_n *converges vaguely* to μ_0 (write $\mu_n \to_v \mu_0$), iff

$$\int_{\mathcal{X}} f(x)\ \mu_n(\mathrm{d}x) \to_{n\to\infty} \int_{\mathcal{X}} f(x)\ \mu_0(\mathrm{d}x)$$

for every $f \in \bar{C}^+(\mathcal{X})$ vanishing outside a compact set. Equivalently, $\mu_n \to_v \mu_0$ iff

$$\mu_n(K) \to_{n\to\infty} \mu_0(K)$$

for all compact sets $K \in \mathbb{B}(\mathcal{X})$ with $\mu_0(\partial K) = 0$.

The following characterization of the domain of attraction of a simple MSP will be a very helpful tool in some proofs. It connects the concepts of domain of attraction, regular variation, and convergence of point processes. For a proof, see e. g. de Haan and Lin (2001).

Proposition 2.27. *Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots$ iid stochastic processes with sample paths in $C^+(S)$. Let $\mathbb{E} = (0,\infty] \times \bar{C}_1^+(S)$ be the complete and separable space from* (2.14)*. The following statements are equivalent:*

(i) Domain of attraction:

$$n^{-1} \max_{i=1,\ldots,n} \boldsymbol{X}^{(i)} \to_{\mathcal{D}} \boldsymbol{\xi},$$

that is, $\boldsymbol{X} \in \mathcal{D}(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is a simple MSP in $C^+(S)$.

(ii) Regular variation:

$$\nu_n(\cdot) := nP\left(n^{-1}\boldsymbol{\xi} \in \cdot\right) \to_{w\#} \nu(\cdot),$$

where ν_n, ν are defined on $(\mathbb{E}, \mathbb{B}(\mathbb{E}))$*. The measure ν is the exponent measure from* (2.16).

(iii) Convergence of point processes:

$$N_n := \sum_{i=1}^{n} \epsilon_{n^{-1}\boldsymbol{X}^{(i)}} \to_{\mathcal{D}} N,$$

in $(M_p(\mathbb{E}), \mathcal{M}_p(\mathbb{E}))$*, where N is* PPP(ν)*, and ν is the exponent measure from* (2.16).

REMARK 2.28. The preceding proposition explains why it is necessary to introduce the space $\mathbb{E} = (0,\infty] \times \bar{C}_1^+(S)$ from (2.14). Although $C^+(S)$ might be the natural choice considering probabilities like $P\left(n^{-1}\boldsymbol{\xi} \in \cdot\right)$, where $\boldsymbol{\xi}$ is a simple MSP, it has the major drawback that it is *not* a complete and separable metric space, if for instance $S = [0,1]$, such that assertions like (ii) and (iii) would not make any sense. By the transformation to polar coordinates and the extension to $\mathbb{E}$, we enlarged $C^+(S)$ to obtain a complete and separable metric space. It is also worth noticing that the metric $\varrho(x,y) = |1/x - 1/y|$ which is used on $(0,\infty]$ implies that the bounded sets in $\mathbb{E}$ are those that are bounded away from 0, i. e. the sets $B \in \mathbb{B}(\mathbb{E})$ that satisfy $\inf\{\|f\|_\infty : f \in B\} > 0$.

Another class of particularly interesting stochastic processes are *copula processes*, i. e. stochastic processes with continuous sample paths and uniformly on $(0,1)$ distributed univariate margins. Suppose a copula process $\boldsymbol{U} = (U_s)_{s\in S}$ is in the domain of attraction of an MSP $\boldsymbol{\eta}$, i. e.

$$n\left(\max_{i=1,\dots,n} \boldsymbol{U}^{(i)} - \mathbf{1}\right) \to_{\mathcal{D}} \boldsymbol{\eta}, \tag{2.23}$$

where $\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ are iid copies of $\boldsymbol{U}$. Then the limiting MSP is necessarily an SMSP, which explains why it seems natural to consider MSP with standard negative exponentially distributed margins. If $\boldsymbol{X} = (X_s)_{s\in S}$ is a stochastic process with continuous sample paths and continuous marginal df $F_s(x) = P(X_s \le x)$, $s \in S$, $x \in \mathbb{R}$, then $(F_s(X_s))_{s\in S}$ is the *copula process corresponding to* $\boldsymbol{X}$. In fact, it is not difficult to verify that the sample paths of $(F_s(X_s))_{s\in S}$ are continuous. It is shown in de Haan and Lin (2001) that $\boldsymbol{X}$ is in the domain of attraction of an MSP $\boldsymbol{Y}$ iff the corresponding copula process is in the domain of attraction of the SMSP $(\log(G_s(Y_s)))_{s\in S}$, where G_s is the df of Y_s, $s \in S$, and the univariate margins of $\boldsymbol{X}$ satisfy some uniformity condition. A similar statement is true in the case of functional domain of attraction, see Aulbach et al. (2015).

The following result will be crucial when it comes to multivariate records, being investigated in Chapter 4.

Proposition 2.29. *Let $\boldsymbol{U}$ be a copula process with $\boldsymbol{U} \in \mathcal{D}(\boldsymbol{\eta})$ (i. e. (2.23) holds), where $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ is an SMSP. Let $\|\cdot\|_D$ be the D-norm corresponding to $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ be a generator of $\|\cdot\|_D$. Then*

$$n\left(1 - P\left(n(\boldsymbol{U} - 1) \le f\right)\right) \to_{n\to\infty} E\left(\sup_{s\in S} |f(s)|\, Z_s\right) = \|f\|_D, \qquad f \in \bar{E}^-(S), \tag{2.24}$$

and

$$nP\left(n(\boldsymbol{U} - 1) > f\right) \to_{n\to\infty} E\left(\inf_{s\in S} |f(s)|\, Z_s\right) = \wr\, f \,\wr_D, \qquad f \in \bar{E}^-(S). \tag{2.25}$$

Proof. Condition (2.23) implies that $\boldsymbol{U}$ is in the functional domain of attraction of $\boldsymbol{\eta}$, i. e.

$$P\left(n(\boldsymbol{U}-1)\leq f\right)^n \to_{n\to\infty} P(\boldsymbol{\eta}\leq f)=\exp\left(-\|f\|_D\right), \qquad f\in\bar{E}^-(S).$$

Now (2.24) follows from Aulbach et al. (2013, Proposition 8). Next we verify (2.25). Having in mind that $\boldsymbol{U}<1$ a. s. (Hofmann (2012, Corollary 3.15)) and $\boldsymbol{\eta}<0$ a. s. (Lemma 2.1), it is easy to see that (2.23) is equivalent with

$$\frac{1}{n}\max_{i=1,\dots,n}\frac{1}{1-\boldsymbol{U}^{(i)}} \to_{\mathcal{D}} -\frac{1}{\boldsymbol{\eta}},$$

where $-1/\boldsymbol{\eta}$ is a *simple* MSP. Put $\mathbb{E}=(0,\infty]\times\bar{C}_1^+(S)$ as in (2.14). Denoting by ν the exponent measure from (2.16), we have by Proposition 2.27

$$\nu_n(A)=nP\left((n(1-\boldsymbol{U}))^{-1}\in A\right)\to_{n\to\infty}\nu(A) \tag{2.26}$$

for all Borel sets $A\in\mathbb{B}(\mathbb{E})$ with $\nu(\partial A)=0$ and $\inf\{\|f\|_\infty : f\in A\}>0$. Define for $h\in E(S)$ the set $A_h:=\{(r,g)\in\mathbb{E}: rg>h\}$. Let ρ be the angular measure from (2.15), and define a generator $\boldsymbol{Z}=(Z_s)_{s\in S}$ of $\boldsymbol{\eta}$ via $(P*\boldsymbol{Z})(A)=m^{-1}\rho\left(m^{-1}A\right)$, $A\in\mathbb{B}\left(m\bar{C}_1^+(S)\right)$. Now by (2.26), for all $f\in E^-(S)$,

$$\begin{aligned}
nP(n(\boldsymbol{U}-1)>f) &= \nu_n\left(A_{-1/f}\right)\\
&\to_{n\to\infty}\nu(A_{-1/f})\\
&=\nu\left(\{(r,g)\in\mathbb{E}: rg>1/|f|\}\right)\\
&=\int_{\bar{C}_1^+(S)}\int_{(\inf_{s\in S}|f(s)|g(s))^{-1}}^{\infty} r^{-2}\,\mathrm{d}r\,\rho(\mathrm{d}g)\\
&=E\left(\inf_{s\in S}|f(s)|\,Z_s\right).
\end{aligned}$$

It remains to show that $\nu(\partial A_h)=0$ for each $h\in E^+(S)$. The boundary of A_h can be expressed via

$$\partial A_h=\{(r,g)\in\mathbb{E}: rg(s)=h(s)\text{ for some } s\in S, rg(s)\geq h(s)\text{ for all } s\in S\}.$$

Hence,

$$\nu(\partial A_h)=\nu\left(\{(r,g)\in\mathbb{E}: rg\geq h\}\right)-\nu\left(\{(r,g)\in\mathbb{E}: rg>h\}\right)=0,$$

since the radial part of the exponent measure has the Lebesgue density $r^{-2}\,\mathrm{d}r$. □

Remark 2.30. We call a stochastic process $\boldsymbol{V}$ with sample paths in $\bar{C}^-(S)$ a *standard generalized Pareto process* (SGPP), if there is a D-norm $\|\cdot\|_D$ on $E(S)$ and some $c > 0$ such that

$$P(\boldsymbol{V} \leq f) = 1 - \|f\|_D$$

for all $f \in \bar{E}^-(S)$ with $\|f\|_\infty \leq c$. It is obvious that $\boldsymbol{V}$ is in the functional domain of attraction of an SMSP with D-norm $\|\cdot\|_D$. Note further that $\boldsymbol{V}$ is an SGPP iff there exists $M < 0$ such that V has in its upper tail the same distribution as $(\max(-U/Z_s, M)_{s \in S})$, more precisely there is $c > 0$ such that

$$P(\boldsymbol{V} \leq f) = P\left(\max(-U/Z_s, M) \leq f(s), s \in S\right)$$

for each $f \in \bar{E}^-(S)$ with $\|f\|_\infty \leq c$. It can easily be deduced that the survival function of $\boldsymbol{V}$ is given by

$$P(\boldsymbol{V} > tf) = t \wr\!\wr f \wr\!\wr_D$$

for $t > 0$ close enough to zero. Hence, condition (2.24) and (2.25) mean that the upper tail of the distribution of the copula process $\boldsymbol{U}$ is close to that of the shifted SGPP $\boldsymbol{V} + 1$. For details on generalized Pareto processes, see e. g. Buishand et al. (2008), Aulbach and Falk (2012*a*,*b*), Aulbach et al. (2013), Aulbach et al. (2015) and Ferreira and de Haan (2014).

2.3 Differentiability in distribution of max-stable processes

Even though max-stable processes have been studied quite extensively over the last decades, the focus is rarely put on path properties. However, while continuity of the sample paths is assumed throughout this work, many authors investigate max-stable processes with weaker sample path properties, starting with Norberg (1986) who studied max-id processes with upper semicontinuous sample paths. Others also examine max-stable processes with sample paths in $D([0,1])$, the space of right-continuous functions on $[0,1]$ with left limits, see e. g. de Haan and Lin (2001). A necessary and sufficient condition when a max-stable process has continuous sample paths is in turn provided by Resnick and Roy (1991). Moreover, in Hofmann (2013), the probability of an SMSP with continuous sample paths on $[0,1]$ hitting some constant function $x\mathbf{1}$ with $x < 0$ is investigated, which provides some insight in the path behavior of max-stable processes as well.

However, a topic that seems to be fairly new is the *differentiability* of max-stable processes with sample paths in $C([0,1])$. Whereas pathwise differentiability is likely to be too restrictive, we will establish *distributional differentiability* of an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$. This concept was originally presented in Aulbach et al. (2015, Section 4) and will be discussed here again.

In Proposition 2.34 we will prove the following result: Let $\boldsymbol{Z} = (Z_s)_{s\in[0,1]} \in C[0,1]$ be a generator process of $\boldsymbol{\eta}$. Suppose that $Z_s' = (\partial/\partial s)Z_s$ exists for $s = s_0$ almost surely. Then $(\eta_{s_0+h} - \eta_{s_0})/h$ converges in distribution to some random variable on the real line, as $h \to 0$, and we compute its df.

This is a first result on differentiability of max-stable processes. To the best of our knowledge, the question, under which conditions a max-stable process is differentiable at $s_0 \in [0,1]$ almost surely, is an open problem. However, if $\eta_s' = (\partial/\partial s)\eta_s$ actually exists at $s = s_0$ almost surely, the distribution of η_{s_0}' equals that of $-\eta_{s_0}\zeta_{s_0}$, where the random variable ζ_{s_0} is independent of η_{s_0} and has the df $\mathbb{F}_{s_0}(x) = E\left(\mathbb{1}_{\left\{Z_{s_0}' \le x Z_{s_0}\right\}} Z_{s_0}\right)$, $x \in \mathbb{R}$; see the discussion after Proposition 2.34.

As an auxiliary result, which is of interest on its own, we first compute the partial derivatives of a functional D-norm $\|\cdot\|_D$. For this purpose, we need the following definition. Let $\mathcal{X}$ be a normed function space, and $J : \mathcal{X} \to \mathbb{R}$ a functional. The *first variation* (or the *Gâteaux differential*) *of* J *at* $u \in \mathcal{X}$ *in the direction* $v \in \mathcal{X}$ is defined as

$$\nabla J(u)(v) := \lim_{\varepsilon \to 0} \frac{J(u+\varepsilon v) - J(u)}{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} J(u+\varepsilon v)\Big|_{\varepsilon=0}.$$

Moreover, the *right-hand* (*left-hand*) first variation of J at u in the direction v is defined as

$$\nabla^+ J(u)(v) := \lim_{\varepsilon \downarrow 0} \frac{J(u+\varepsilon v) - J(u)}{\varepsilon} \quad \text{and} \quad \nabla^- J(u)(v) := \lim_{\varepsilon \downarrow 0} \frac{J(u) - J(u-\varepsilon v)}{\varepsilon}.$$

Considering a D-norm $\|\cdot\|_D$ a functional on the space $E[0,1]$, we can calculate the first variation of $\|\cdot\|_D$. As discussed before, the choice of the space $E[0,1]$ allows us the incorporation of the fidis and therefore yields the partial derivatives of a *multivariate* D-norm. This finite-dimensional version of the following result has already been observed by Einmahl et al. (2012). Note that as a norm is a convex function, a multivariate D-norm $\|\boldsymbol{x}\|_D$ is for almost every $\boldsymbol{x} \in \mathbb{R}^d$ continuously differentiable.

Lemma 2.31 (Aulbach et al. (2015, Lemma 4.1))**.** *Let* $\|\cdot\|_D$ *be a* D*-norm on the function space* $E[0,1]$ *with generator* $\boldsymbol{Z} = (Z_s)_{s\in[0,1]} \in \bar{C}^+[0,1]$ *and choose* $s_0 \in [0,1]$*. Then for* *every* $f \in E[0,1]$,

$$\nabla^+ \|f\|_D \left(\mathbb{1}_{\{s_0\}}\right) = \lim_{\varepsilon \downarrow 0} \frac{\left\|f + \varepsilon \mathbb{1}_{\{s_0\}}\right\|_D - \|f\|_D}{\varepsilon}$$
$$= \begin{cases} -E\left(\mathbb{1}_{\{\sup_{s \neq s_0} |f(s)| Z_s < |f(s_0)| Z_{s_0}\}} Z_{s_0}\right), & f(s_0) < 0, \\ +E\left(\mathbb{1}_{\{\sup_{s \neq s_0} |f(s)| Z_s \leq |f(s_0)| Z_{s_0}\}} Z_{s_0}\right), & f(s_0) \geq 0, \end{cases}$$

and

$$\nabla^- \|f\|_D \left(\mathbb{1}_{\{s_0\}}\right) = \lim_{\varepsilon \downarrow 0} \frac{\|f\|_D - \left\|f - \varepsilon \mathbb{1}_{\{s_0\}}\right\|_D}{\varepsilon}$$
$$= \begin{cases} -E\left(\mathbb{1}_{\{\sup_{s \neq s_0} |f(s)| Z_s \leq |f(s_0)| Z_{s_0}\}} Z_{s_0}\right), & f(s_0) \leq 0, \\ +E\left(\mathbb{1}_{\{\sup_{s \neq s_0} |f(s)| Z_s < |f(s_0)| Z_{s_0}\}} Z_{s_0}\right), & f(s_0) > 0. \end{cases}$$

The first variation (or the partial derivatives, respectively) of a D-norm emerge in the easiest case of the so-called *prediction problem*, cf. Wang and Stoev (2011), Dombry et al. (2013) and Dombry and Éyi-Minko (2013). Suppose the distribution of an SMSP $\boldsymbol{\eta}$ is known, and the point $\{\eta_{s_0} = x\}$, $x < 0$, has already been observed. We are interested in the conditional distribution of $\boldsymbol{\eta}$, given $\{\eta_{s_0} = x\}$. The finite-dimensional version of the following Lemma is part of Proposition 4.2 in Dombry and Éyi-Minko (2013). Its proof is stated here again to point out the relevance of the first variation of a D-norm.

Lemma 2.32 (Aulbach et al. (2015, Lemma 4.2))**.** *Let* $\boldsymbol{\eta} = (\eta_s)_{s \in [0,1]}$ *be an SMSP with* D*-norm* $\|\cdot\|_D$ *generated by* $\boldsymbol{Z} = (Z_s)_{s \in [0,1]}$*. Choose an arbitrary* $s_0 \in [0,1]$*. Then for every* $f \in \bar{E}^-[0,1]$ *with* $f(s_0) = 0$ *and almost all* $y < 0$

$$P\left(\boldsymbol{\eta} \leq f \middle| \eta_{s_0} = y\right) = \exp\left(-\left(y + \left\|f + y\mathbb{1}_{\{s_0\}}\right\|_D\right)\right) \cdot E\left(\mathbb{1}_{\{\sup_{s \in [0,1]} |f(s)| Z_s \leq |y| Z_{s_0}\}} Z_{s_0}\right).$$

Proof. The random variable η_{s_0} has Lebesgue-density e^x, $x \leq 0$. Therefore, we have by basic rules of conditional distributions for almost all $y < 0$

$$P\left(\boldsymbol{\eta} \leq f \middle| \eta_{s_0} = y\right) = \lim_{\varepsilon \downarrow 0} \frac{\varepsilon^{-1} P(\boldsymbol{\eta} \leq f, \eta_{s_0} \in (y, y+\varepsilon])}{\varepsilon^{-1} P(\eta_{s_0} \in (y, y+\varepsilon])}$$
$$= \exp(-y) \lim_{\varepsilon \downarrow 0} \frac{P(\boldsymbol{\eta} \leq f, \eta_{s_0} \leq y + \varepsilon) - P(\boldsymbol{\eta} \leq f, \eta_{s_0} \leq y)}{\varepsilon}.$$

Now define the function $g := f + y\mathbb{1}_{\{s_0\}} \in E^-[0,1]$. Then we have by Lemma 2.31

$$\begin{aligned}
P\left(\boldsymbol{\eta} \le f \middle| \eta_{s_0} = y\right) &= \exp(-y) \lim_{\varepsilon \downarrow 0} \frac{\exp\left(-\left\|g + \varepsilon \mathbb{1}_{\{s_0\}}\right\|_D\right) - \exp\left(-\left\|g\right\|_D\right)}{\varepsilon} \\
&= -\exp(-y)\exp\left(-\left\|g\right\|_D\right) \cdot \nabla^+ \left\|g\right\|_D \left(\mathbb{1}_{\{s_0\}}\right) \\
&= \exp\left(-\left(y + \left\|g\right\|_D\right)\right) \cdot E\left(\mathbb{1}_{\{\sup_{s\in[0,1]}|g(s)|Z_s = |y|Z_{s_0}\}}\right) \\
&= \exp\left(-\left(y + \left\|f + y\mathbb{1}_{\{s_0\}}\right\|_D\right)\right) \cdot E\left(\mathbb{1}_{\{\sup_{s\in[0,1]}|f(s)|Z_s \le |y|Z_{s_0}\}}\right).
\end{aligned}$$

□

As a simple consequence of the preceding Lemma, we can calculate the distribution of the increments of an SMSP. This will clearly be needed to derive the distribution of the difference quotient $(\eta_{s_0+h} - \eta_{s_0})/h$. However, this result is also of its own interest. The proof is presented since it was dropped in the original paper.

Lemma 2.33 (Aulbach et al. (2015, Lemma 4.3))**.** *Consider an SMSP* $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ *with generator process* $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ *and choose arbitrary* $s, t \in [0,1]$, $s \neq t$. *Denote by* $\left\|\cdot\right\|_D$ *the D-norm pertaining to* (η_s, η_t). *Then for every* $x \in \mathbb{R}$

$$\begin{aligned}
&P\left(\eta_s - \eta_t \le x\right) \\
&= \begin{cases} \displaystyle\int_{-\infty}^{0} \exp\left(-\left\|(x+y, y)\right\|_D\right) \cdot E\left(\mathbb{1}_{\{yZ_t \le (x+y)Z_s\}} Z_t\right) \,\mathrm{d}y, & x < 0 \\ \displaystyle\int_{-\infty}^{-x} \exp\left(-\left\|(x+y, y)\right\|_D\right) \cdot E\left(\mathbb{1}_{\{yZ_t \le (x+y)Z_s\}} Z_t\right) \,\mathrm{d}y + 1 - \exp(-x), & x \ge 0. \end{cases}
\end{aligned}$$

Proof. Conditioning on $\eta_t = y$ yields

$$P(\eta_s - \eta_t \le x) = \int_{-\infty}^{0} P\left(\eta_s \le x + y \middle| \eta_t = y\right) e^y \,\mathrm{d}y.$$

In the case $x < 0$, we obtain by Lemma 2.32

$$P(\eta_s - \eta_t \le x) = \int_{-\infty}^{0} \exp\left(-\left\|(x+y, y)\right\|_D\right) \cdot E\left(\mathbb{1}_{\{yZ_t \le (x+y)Z_s\}} Z_t\right) \,\mathrm{d}y.$$

The case $x \ge 0$ works analogously, yet $P\left(\eta_s \le x + y \middle| \eta_t = y\right) = 1$ for $y \ge -x$. □

Note that the only possible point of discontinuity of the df $P(\eta_s - \eta_t \le x)$ is $x = 0$, where

$$P(\eta_s - \eta_t \le 0) = \left(\|(1,1)\|_D\right)^{-1} E\left(1\!\!1_{\{Z_t \ge Z_s\}} Z_t\right).$$

The preceding lemma allows us to introduce the following differentiability concept. Firstly, we call a stochastic process $(X_s)_{s\in[0,1]}$ *almost surely differentiable in* $s_0 \in [0,1]$, if the difference quotient $(X_{s_0+h} - X_{s_0})/h$ converges almost surely to some random variable X'_{s_0} on the real line for $h \to 0$. Different to that, we call a stochastic process $(Y_s)_{s\in[0,1]}$ *differentiable in distribution in* $s_0 \in [0,1]$, if the difference quotient $(Y_{s_0+h} - Y_{s_0})/h$ converges in distribution to some real-valued random variable for $h \to 0$. Lastly, we call a stochastic process $\boldsymbol{\xi} = (\xi_s)_{s\in[0,1]}$ *pathwise differentiable* on $[0,1]$ if every path $\boldsymbol{\xi}(\omega)$ is differentiable on $[0,1]$.

We proof the following Proposition once again to illustrate the connection to the previous results.

Proposition 2.34 (Aulbach et al. (2015, Proposition 4.3))**.** *Let* $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ *be an SMSP with generator process* $\boldsymbol{Z} = (Z_s)_{s\in[0,1]} \in \bar{C}^+[0,1]$*. Suppose that for some* $s_0 \in [0,1]$

$$\frac{Z_{s_0+h} - Z_{s_0}}{h} \to_{h\to 0} Z'_{s_0} \quad \textit{almost surely.} \tag{2.27}$$

Then for $x \ne 0$,

$$P\left(\frac{\eta_{s_0+h} - \eta_{s_0}}{h} \le x\right) \to_{h\to 0} H_{s_0}(x) := \int_{-\infty}^{0} \exp(y) E\left(1\!\!1_{\left\{Z'_{s_0} \le -\frac{x}{y} Z_{s_0}\right\}} Z_{s_0}\right) \,\mathrm{d}y.$$

Proof. We have for $x \ne 0$ and $h > 0$ by Lemma 2.33

$$\begin{aligned} &P\left(\eta_{s_0+h} - \eta_{s_0} \le hx\right) \\ &= \int_{-\infty}^{-h|x|} \exp\left(-\|(hx+y, y)\|_{D(h)}\right) \cdot E\left(1\!\!1_{\{yZ_{s_0} \le (hx+y)Z_{s_0+h}\}} Z_{s_0}\right) \,\mathrm{d}y + o(1) \end{aligned}$$

as $h \downarrow 0$, where $\|\cdot\|_{D(h)}$ is the D-norm generated by (Z_{s_0+h}, Z_{s_0}). Now we obtain for almost all $y < -h|x|$

$$\begin{aligned} &E\left(1\!\!1_{\left\{yZ_{s_0} \le (hx+y)Z_{s_0+h}\right\}} Z_{s_0}\right) \\ &= E\left(1\!\!1_{\left\{y\frac{Z_{s_0} - Z_{s_0+h}}{h} \le xZ_{s_0+h}\right\}} Z_{s_0}\right) \end{aligned}$$

$$= E\left(1\!\!1_{\left\{\frac{Z_{s_0+h}-Z_{s_0}}{h}\le -\frac{x}{y}Z_{s_0+h}\right\}} Z_{s_0}\right)$$

$$\to_{h\downarrow 0} E\left(1\!\!1_{\left\{Z'_{s_0}\le -\frac{x}{y}Z_{s_0}\right\}} Z_{s_0}\right)$$

by condition (2.27) which implies the assertion if $h \downarrow 0$. On the other hand, we have for $x \neq 0$ and $h < 0$ by Lemma 2.33, condition (2.27), and the fact that $E(Z_{s_0}) = 1$

$$P\left(\eta_{s_0+h} - \eta_{s_0} \ge hx\right) = 1 - P\left(\eta_{s_0+h} - \eta_{s_0} \le hx\right)$$

$$= 1 - \int_{-\infty}^{h|x|} \exp\left(-\left\|(hx+y, y)\right\|_{D(h)}\right) \cdot E\left(1\!\!1_{\{yZ_{s_0}\le (hx+y)Z_{s_0+h}\}} Z_{s_0}\right)\, dy + o(1)$$

$$\to_{h\uparrow 0} 1 - \int_{-\infty}^{0} \exp(y) E\left(1\!\!1_{\left\{Z'_{s_0}\ge -\frac{x}{y}Z_{s_0}\right\}} Z_{s_0}\right)\, dy$$

$$= 1 - \int_{-\infty}^{0} \exp(y)\, dy + \int_{-\infty}^{0} \exp(y) E\left(1\!\!1_{\left\{Z'_{s_0}\le -\frac{x}{y}Z_{s_0}\right\}} Z_{s_0}\right)\, dy$$

$$= \int_{-\infty}^{0} \exp(y) E\left(1\!\!1_{\left\{Z'_{s_0}\le -\frac{x}{y}Z_{s_0}\right\}} Z_{s_0}\right)\, dy.$$

□

Proposition 2.34 gives a sufficient condition on the differentiability in distribution of an SMSP. However, it does not imply differentiability of the path of $\boldsymbol{\eta}$ at s_0. But if $\boldsymbol{\eta}$ is differentiable at s_0 almost surely, then H_{s_0} is the df of the derivative $(\partial/\partial s)\eta_s$ of $\boldsymbol{\eta}$ at $s = s_0$. We, therefore, denote by η'_{s_0} a random variable which follows the df H_{s_0}.

Suppose that $\boldsymbol{Z}$ is almost surely differentiable in s_0. Then

$$\mathbb{F}_{s_0}(x) := E\left(1\!\!1_{\left\{Z'_{s_0}\le xZ_{s_0}\right\}} Z_{s_0}\right), \qquad x \in \mathbb{R},$$

defines a common df on $\mathbb{R}$. Denote by ζ_{s_0} a random variable which follows this df and which is independent of η_{s_0}. Then we obtain the equation

$$H_{s_0}(x) = P\left(-\eta_{s_0}\zeta_{s_0} \le x\right), \qquad x \in \mathbb{R},$$

i.e., we have

$$\eta'_{s_0} =_D -\eta_{s_0}\zeta_{s_0}.$$

Assuming its existence, the pathwise derivative of $\boldsymbol{\eta}$ at s_0, coincides therefore in distribution with $-\eta_{s_0}\zeta_{s_0}$.

Lemma 2.35 (Aulbach et al. (2015, Lemma 4.5))**.** *Suppose that $E\left(Z'_{s_0}\right)$ exists. Then the mean value of $\mathbb{F}_{s_0}$ exists as well and coincides with $E\left(Z'_{s_0}\right)$.*

As a consequence, we obtain in particular

$$E\left(\eta'_{s_0}\right) = -E\left(\eta_{s_0}\zeta_{s_0}\right) = -E\left(\eta_{s_0}\right)E\left(\zeta_{s_0}\right) = E\left(Z'_{s_0}\right).$$

We close this chapter by giving some examples how Proposition 2.34 can be applied.

Example 2.36 (Aulbach et al. (2015, Example 4.6)). Put for $\lambda \in \mathbb{R}$

$$Z_s := U\cos^2(\lambda s) + V\sin^2(\lambda s), \qquad s \in [0,1],$$

where $U \geq 0$, $V \geq 0$ are rv with $E(U) = E(V) = 1$. The process $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ is pathwise differentiable with

$$\frac{\partial}{\partial s} Z_s = \lambda \sin(2\lambda s)(V - U) =: Z'_s.$$

The distribution of the derivative in distribution η'_s is accessible under additional conditions on U and V, but it follows immediately from Lemma 2.35 that in general $E(\eta'_s) = E(Z'_s) = 0$.

Example 2.37 (Aulbach et al. (2015, Example 4.7)). The constant generator process $Z_s \equiv 1$, $s \in [0,1]$, gives rise to an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ with complete dependent univariate margins. The paths of this SMSP are constant almost surely, which means that $\eta'_s = 0$ with probability one. This fact is reflected in Proposition 2.34. If $Z_s \equiv 1$, $s \in [0,1]$, then $\mathbb{F}_s(x) = \mathbb{1}_{[0,\infty)}(x)$, which implies

$$H_s(x) = \int_{-\infty}^{0} \exp(y)\mathbb{F}_s(-x/y)\,\mathrm{d}y = \mathbb{1}_{[0,\infty)}(x).$$

Example 2.38 (Aulbach et al. (2015, Example 4.8)). Let (Z_0, Z_1) be the generator of a bivariate standard max-stable random vector (η_0, η_1) with independent margins, i. e.

$$P(Z_0 = 0, Z_1 = 2) = P(Z_0 = 2, Z_1 = 0) = 1/2,$$

cf. Example 2.9. Now define a generator process by

$$Z_s := Z_0 + s(Z_1 - Z_0) \, (= \max((1-s)Z_0, sZ_1)) \,, \qquad s \in [0,1],$$

and denote by $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ the corresponding SMSP. Then obviously $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ is pathwise differentiable with $Z_s' = Z_1 - Z_0$, $s \in [0,1]$. Lemma 2.35 instantly implies $E(\eta_s') = E(Z_s') = 0$. Furthermore, we have for $x \in \mathbb{R}$ and $s \in [0,1]$

$$\begin{aligned}
\mathbb{F}_s(x) &= E\left(\mathbb{1}_{\{Z_s' \le xZ_s\}} Z_s\right) \\
&= \mathbb{1}_{\{2 \le 2xs\}} \cdot 1/2 \cdot 2s + \mathbb{1}_{\{-2 \le x(2-2s)\}} \cdot 1/2 \cdot (2-2s) \\
&= s\mathbb{1}_{\{x \ge 1/s\}} + (1-s)\mathbb{1}_{\{x \ge 1/(s-1)\}} \\
&= \begin{cases} 1, & x \ge 1/s, \\ 1-s, & 1/(s-1) \le x < 1/s, \\ 0, & x < 1/(s-1). \end{cases}
\end{aligned}$$

Hence, the corresponding random variable ζ_s that follows the df $\mathbb{F}_s$ is discrete with $P(\zeta_s = 1/(s-1)) = 1-s$ and $P(\zeta_s = 1/s) = s$. Therefore, we obtain

$$\begin{aligned}
H_s(x) &= \int_{-\infty}^{0} \mathbb{F}_s(-x/y) \exp(y) \, \mathrm{d}y \\
&= \begin{cases} \displaystyle\int_{-\infty}^{x(1-s)} \exp(y)(1-s) \, \mathrm{d}y, & x < 0, \\ \displaystyle\int_{-\infty}^{-xs} \exp(y)(1-s) \, \mathrm{d}y + \int_{-xs}^{0} \exp(y) \, \mathrm{d}y, & x \ge 0 \end{cases} \\
&= \begin{cases} (1-s)\exp(x(1-s)), & x < 0, \\ 1 - s\exp(-xs), & x \ge 0. \end{cases}
\end{aligned}$$

The density of H_s is given by

$$h_s(x) = \begin{cases} (1-s)^2 \exp(x(1-s)), & x < 0, \\ s^2 \exp(-xs), & x \ge 0. \end{cases}$$

Example 2.39 (Aulbach et al. (2015, Example 4.9)). Let Z_0 be uniformly distributed on $(0,2)$ and Z_1 be a random variable with $Z_0 + Z_1 = 2$ almost surely. Clearly, Z_0 and Z_1 are identically distributed and (Z_0, Z_1) defines a (bivariate) generator. Define the generator process

$$Z_s := Z_0 + s(Z_1 - Z_0), \quad s \in [0,1].$$

Again, $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ is pathwise differentiable with $Z'_s = Z_1 - Z_0$, $s \in [0,1]$. In addition,

$$P(Z_1 - Z_0 \le x) = P(Z_0 \ge (2-x)/2) = x/4 + 1/2, \qquad x \in [-2,2],$$

which implies that Z'_s is uniformly distributed on $(-2,2)$. This, along with the fact that $Z_{1/2} = 1$ almost surely, yields

$$\mathbb{F}_{1/2}(x) = E\left(1_{\{Z'_{1/2}\le x\}}\right) = P\left(Z'_{1/2} \le x\right).$$

Hence, $\mathbb{F}_{1/2}$ is the df of the uniform distribution on $(-2,2)$. Now choose an arbitrary $y < 0$. In the case of $x < 0$ we have

$$-\frac{x}{y} \le -2 \iff y \ge \frac{x}{2},$$

whereas $x > 0$ yields

$$-\frac{x}{y} \ge 2 \iff y \ge -\frac{x}{2}.$$

Therefore, for $x < 0$,

$$\begin{aligned}
H_{1/2}(x) &= \int_{-\infty}^{0} \exp(y)\mathbb{F}_{1/2}\left(-\frac{x}{y}\right)\,\mathrm{d}y \\
&= \int_{-\infty}^{x/2} \exp(y)\left(-\frac{x}{4y} + \frac{1}{2}\right)\,\mathrm{d}y \\
&= -\frac{x}{4}\int_{-\infty}^{x/2} \frac{\exp(y)}{y}\,\mathrm{d}y + \frac{1}{2}\int_{-\infty}^{x/2} \exp(y)\,\mathrm{d}y \\
&= -\frac{x}{4}\,\mathrm{Ei}\left(\frac{x}{2}\right) + \frac{1}{2}\exp\left(\frac{x}{2}\right)
\end{aligned}$$

Here, $\mathrm{Ei}(x) = \int_{-\infty}^{x} \exp(t)/t\,\mathrm{d}t$ denotes the *exponential integral*, which is well-defined for negative values of x. Analogously, for $x > 0$,

$$\begin{aligned}
H_{1/2}(x) &= \int_{-\infty}^{-x/2} \exp(y)\left(-\frac{x}{4y} + \frac{1}{2}\right)\,dy + \int_{-x/2}^{0} \exp(y)\,dy \\
&= -\frac{x}{4}\,\mathrm{Ei}\left(-\frac{x}{2}\right) - \frac{1}{2}\exp\left(-\frac{x}{2}\right) + 1.
\end{aligned}$$

To summarize, we have

$$H_{1/2}(x) = \begin{cases} -\frac{x}{4}\operatorname{Ei}\left(\frac{x}{2}\right) + \frac{1}{2}\exp\left(\frac{x}{2}\right), & x < 0, \\ -\frac{x}{4}\operatorname{Ei}\left(-\frac{x}{2}\right) - \frac{1}{2}\exp\left(-\frac{x}{2}\right) + 1, & x > 0. \end{cases}$$

Furthermore, we have $H_{1/2}(0) = 1/2$. In particular, $H_{1/2}$ is continuous in 0 since the exponential integral satisfies $x\operatorname{Ei}(x) \to_{x\to 0} 0$. The density of $H_{1/2}$ is given by $h_{1/2}(x) = -\operatorname{Ei}\left(-|x|/2\right)/4$, $x \neq 0$.

3 Generalized max-linear models

There is a crucial problem in the theory on stochastic processes concerning its relevance in practice: as (continuous) processes on, say $[0,1]$, as a whole cannot be measured exactly, the question is how to construct these processes (with some characteristic stochastic behavior such as max-stability) from a finite set of observations. For instance, measuring the sea level along a coast is only possible at a finite number of stations. However, it is also important to predict the sea level between these stations. Hence, our aim is the 'prediction' of stochastic processes in space rather than in time.

In the case of max-stable processes, there are (partial) answers on the arising questions in Wang and Stoev (2011) and Dombry et al. (2013) based on conditional sampling. Different to that, our approach is conditionally deterministic.

In this chapter, we pick up the so-called *max-linear model* introduced in Wang and Stoev (2011). For the ease of notation, we consider SMSP with sample paths in $C\left([0,1]^k\right)$, $k \in \mathbb{N}$, being aware of the fact that we could replace the domain $[0,1]^k$ by any compact metric space S. For arbitrary nonnegative continuous functions $g_1, \dots, g_d$ satisfying condition (3.1) below, an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ is given by

$$\eta_s = \max_{i=1,\dots,d} \frac{X_i}{g_i(s)}, \qquad s \in [0,1]^k,$$

where $\boldsymbol{X} = (X_1, \dots, X_d)$ is a standard max-stable random vector with independent components. The obvious restriction of this model is the required independence of the margins of $\boldsymbol{X}$, which results in the fact that $\boldsymbol{X}$ always has a discrete angular measure (the angular measure is essentially the distribution of the generator $(Z_1, \dots, Z_d)$ of $\boldsymbol{X}$, cf. (2.15)).

In Section 3.1, we generalize this model by allowing arbitrary dependence structures of the margins of the max-stable random vector $\boldsymbol{X}$. This immediately leads to the main issue of this chapter, namely the reconstruction of a max-stable process with sample paths in $C\left([0,1]^k\right)$ which is observed only through a finite set of indices. Starting with the case $k = 1$, it is shown in Section 3.2 that if the random vector is some finite dimensional projection an SMSP, the processes resulting from a particular construction based on the model in Section 3.1 converge uniformly to the original process as the grid of indices gets finer. Moreover, the mean squared error between the predictive and the original process is computed at a fixed index, which is useful for practical purposes. Most of the results in Section 3.1 and Section 3.2 have already been published in Falk et al. (2015). Nevertheless, some proofs are repeated in order to be more elaborate than in the initial paper.

It is also possible to interpolate generalized Pareto processes (cf. Remark 2.30) with the same techniques as for MSP, which is the content of Section 3.3.

Having stated the model in its generality, we continue by considering SMSP with domain $[0,1]$. The extension of the results to more general domains (in particular higher dimensions of the domain) is not immediately obvious and the subject of Section 3.4.

3.1 The generalized max-linear model

Introduction of the model

Let in what follows $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]^k}$ be an SMSP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]^k}$ and D-norm $\|\cdot\|_D$. As shown in (2.10), the finite-dimensional projection $(\eta_{s_1},\dots,\eta_{s_d})$ defines a standard max-stable random vector with generator $(Z_{s_1},\dots,Z_{s_d})$ and D-norm $\|\cdot\|_{D_{s_1,\dots,s_d}}$ for pairwise different $s_1,\dots,s_d \in [0,1]^k$, that is,

$$P(\eta_{s_1} \le x_1,\dots,\eta_{s_d} \le x_d) = \exp\left(-E\left(\max_{i=1,\dots,d}\left(|x_i|\, Z_{s_i}\right)\right)\right) =: \exp\left(-\|\boldsymbol{x}\|_{D_{s_1,\dots,s_d}}\right),$$

where $\boldsymbol{x} = (x_1,\dots,x_d) \le \boldsymbol{0}$. Choose arbitrary deterministic functions $g_1,\dots,g_d \in \bar{C}^+[0,1]$ with the property

$$\|(g_1(s),\dots,g_d(s))\|_{D_{s_1,\dots,s_d}} = 1, \quad s \in [0,1]^k. \tag{3.1}$$

For instance, in case of independent margins of $(\eta_{s_1},\dots,\eta_{s_d})$, we have $\|\cdot\|_{D_{s_1,\dots,s_d}} = \|\cdot\|_1$, and condition (3.1) becomes

$$\sum_{i=1}^{d} g_i(s) = 1, \qquad s \in [0,1]^k,$$

i.e. $g_i(s)$, $i = 1,\dots,d$, defines a probability distribution on the set $\{1,\dots,d\}$ for each $s \in [0,1]^k$. This is the setup in the max-linear model introduced by Wang and Stoev

(2011). A onedimensional example for this case is given by the binomial distribution

$$g_i(s) := \binom{d-1}{i-1} s^{i-1}(1-s)^{d-i+1}, \qquad i = 1, \dots, d, \ s \in [0,1].$$

Put now

$$\hat{\eta}_s := \max_{i=1,\dots,d} \frac{\eta_{s_i}}{g_i(s)}, \qquad s \in [0,1]^k. \tag{3.2}$$

The model (3.2) is called *generalized max-linear model.* It defines an SMSP as the next lemma shows. The proof shall not be omitted in order to present it more extensively than in the original paper.

Lemma 3.1 (Falk et al. (2015, Lemma 1))**.** *The stochastic process* $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s\in[0,1]^k}$ *in* (3.2) *defines an SMSP with generator process* $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]^k}$ *given by*

$$\hat{Z}_s := \max_{i=1,\dots,d} \left(g_i(s) Z_{s_i}\right), \qquad s \in [0,1]^k. \tag{3.3}$$

Proof. At first we verify that the process $\hat{\boldsymbol{Z}}$ is a generator process indeed. It is obvious that the sample paths of $\hat{\boldsymbol{Z}}$ are in $\bar{C}^+\left([0,1]^k\right)$. Furthermore, we have by construction for each $s \in [0,1]^k$

$$E(\hat{Z}_s) = \|(g_1(s), \dots, g_d(s))\|_{D_{s_1,\dots,s_d}} = 1.$$

As $\|\cdot\|_\infty \le \|\cdot\|_D$ for an arbitrary D-norm, we have $\|(g_1(s), \dots, g_d(s))\|_\infty \le 1$, $s \in [0,1]^k$, and, thus, $\hat{Z}_s \le \max_{i=1,\dots,d} Z_{s_i}$, $s \in [0,1]^k$. In addition, we have for $f \in \bar{E}^-[0,1]$

$$\begin{aligned}
P(\hat{\boldsymbol{\eta}} \le f) &= P\left(\eta_{s_i} \le g_i(s) f(s), \ i = 1, \dots, d, \ s \in [0,1]^k\right) \\
&= P\Bigg(\eta_{s_i} \le -\sup_{s\in[0,1]^k} \left(g_i(s) \, |f(s)|\right), \ i = 1, \dots, d\Bigg) \\
&= \exp\Bigg(-\Bigg\|\Bigg(\sup_{s\in[0,1]^k} \left(g_1(s) \, |f(s)|\right), \dots, \sup_{s\in[0,1]^k} \left(g_d(s) \, |f(s)|\right)\Bigg)\Bigg\|_{D_{s_1,\dots,s_d}}\Bigg) \\
&= \exp\Bigg(-E\Bigg(\max_{i=1,\dots,d}\Bigg(\sup_{s\in[0,1]^k}\left(g_i(s) \, |f(s)|\right) Z_{s_i}\Bigg)\Bigg)\Bigg) \\
&= \exp\Bigg(-E\Bigg(\sup_{s\in[0,1]^k}\left(|f(s)| \max_{i=1,\dots,d} (g_i(s) Z_{s_i})\right)\Bigg)\Bigg) \\
&= \exp\Bigg(-E\Bigg(\sup_{s\in[0,1]^k}\left(|f(s)| \, \hat{Z}_s\right)\Bigg)\Bigg)
\end{aligned}$$

which completes the proof. □

Remark 3.2 (Falk et al. (2015, Remark 1)). Condition (3.1) ensures that the univariate margins $\hat{\eta}_s$, $s \in [0,1]^k$, of the process $\hat{\boldsymbol{\eta}}$ in model (3.2) follow the standard negative exponential distribution $P(\hat{\eta}_s \le x) = \exp(x)$, $x \le 0$. If we drop this condition, we still obtain a max-stable process: Take for $n \in \mathbb{N}$ iid copies $\hat{\boldsymbol{\eta}}^{(1)}, \dots, \hat{\boldsymbol{\eta}}^{(n)}$ of $\hat{\boldsymbol{\eta}}$. We have for $f \in \bar{E}^-[0,1]$

$$\begin{aligned} P\left(n \max_{1\le k\le n} \hat{\boldsymbol{\eta}}^{(k)} \le f\right) &= P\left(\hat{\eta}_{s_i} \le \inf_{s\in[0,1]^k} \left(\frac{g_i(s) f(s)}{n}\right),\ i = 1,\dots,d\right)^n \\ &= \exp\left(-\left\|\left(\sup_{s\in[0,1]^k} \left(g_1(s)\,|f(s)|\right), \dots, \sup_{s\in[0,1]^k} \left(g_d(s)\,|f(s)|\right)\right)\right\|_{D_{s_1,\dots,s_d}}\right) \\ &= P(\hat{\boldsymbol{\eta}} \le f). \end{aligned}$$

The univariate margins of $\hat{\boldsymbol{\eta}}$ are now given by

$$P(\hat{\eta}_s \le x) = \exp\left(x\, \|(g_1(s), \dots, g_d(s))\|_{D_{s_1,\dots,s_d}}\right), \quad x \le 0,\ s \in [0,1]^k. \tag{3.4}$$

Note that the above calculations also give an alternative proof of Lemma 3.1, except we do not obtain the generator process of $\hat{\boldsymbol{\eta}}$ with this approach.

In model (3.2), we have not made any further assumptions on the D-norm $\|\cdot\|_{D_{s_1,\dots,s_d}}$, that is, on the dependence structure of the random variables $\eta_{s_1}, \dots, \eta_{s_d}$. The special case $\|\cdot\|_{D_{s_1,\dots,s_d}} = \|\cdot\|_1$ characterizes the independence of $\eta_{s_1}, \dots, \eta_{s_d}$. This is the regular *max-linear model*, cf. Wang and Stoev (2011). On the contrary, $\|\cdot\|_{D_{s_1,\dots,s_d}} = \|\cdot\|_\infty$ provides the case of complete dependence $\eta_{s_1} = \dots = \eta_{s_d}$ a. s. with the constant generator $Z_{s_1} = \dots = Z_{s_d} = 1$. Thus, condition (3.1) becomes $\max_{i=1,\dots,d} g_i(s) = 1$, $s \in [0,1]^k$, and therefore

$$\hat{Z}_s = \max_{i=1,\dots,d} (g_i(s) Z_{s_i}) = \max_{i=1,\dots,d} g_i(s) = 1, \quad s \in [0,1]^k.$$

If we want $\hat{\boldsymbol{\eta}}$ to interpolate $(\eta_{s_1}, \dots, \eta_{s_d})$, then we only have to demand

$$g_i(s_j) = \delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \ne j, \end{cases} \quad 1 \le i, j \le d. \tag{3.5}$$

Recall that η_{s_i} is negative with probability one. We call $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{Z}}$ the *discretized version* of $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ with grid $\{s_1, \dots, s_d\}$ and weight functions $g_1, \dots, g_d$, when the weight functions satisfy both (3.1) and (3.5). We will see different examples of discretized versions in the next sections.

The mean squared error of the discretized version

We start this section with a result that applies to bivariate standard max-stable random vectors in general. The first part of the Lemma is Falk et al. (2015, Lemma 5), yet the proof was omitted therein.

Lemma 3.3. *Let (η_1, η_2) be standard max-stable with generator (Z_1, Z_2) and D-norm $\|\cdot\|_D$. Then*

$$E(\eta_1\eta_2) = \int_0^\infty \frac{1}{\|(1,t)\|_D^2}\, \mathrm{d}t. \tag{3.6}$$

In particular, the covariance and the correlation coefficient ϱ of η_1 and η_2 are given by

$$\mathrm{Cov}(\eta_1, \eta_2) = \int_0^\infty \frac{1}{\|(1,t)\|_D^2}\, \mathrm{d}t - 1 = \varrho(\eta_1, \eta_2). \tag{3.7}$$

Furthermore,

$$E(|Z_1 - Z_2|) = 2\left(\|(1,1)\|_D - 1\right). \tag{3.8}$$

Proof. Recall that the expected value of a negative exponentially distributed random variable η with parameter $\lambda > 0$ is given by

$$E(\eta) = \int_{-\infty}^0 x\lambda \exp(\lambda x)\; dx = -1/\lambda.$$

Hence, elementary calculations show

$$\begin{aligned}
E(\eta_1\eta_2) &= \int_{-\infty}^0 \int_{-\infty}^0 P(\eta_1 \le x, \eta_2 \le y)\; \mathrm{d}x\; \mathrm{d}y \\
&= \int_{-\infty}^0 \int_{-\infty}^0 \exp\left(-\|(x,y)\|_D\right)\; \mathrm{d}x\; \mathrm{d}y \\
&= \int_{-\infty}^0 \int_{-\infty}^0 \exp\left(x\,\|(1,y/x)\|_D\right)\; \mathrm{d}x\; \mathrm{d}y \\
&= -\int_0^\infty \int_{-\infty}^0 x \exp\left(x\,\|(1,t)\|_D\right)\; \mathrm{d}x\; \mathrm{d}t \\
&= -\int_0^\infty \frac{1}{\|(1,t)\|_D} \int_{-\infty}^0 x\,\|(1,t)\|_D \exp\left(x\,\|(1,t)\|_D\right)\; \mathrm{d}x\; \mathrm{d}t \\
&= \int_0^\infty \frac{1}{\|(1,t)\|_D^2}\; \mathrm{d}t,
\end{aligned}$$

which is (3.6). The assertions (3.7) follow from the fact that $E(\eta_1) = E(\eta_2) = -1$ and $\mathrm{Var}(\eta_1) = \mathrm{Var}(\eta_2) = 1$. Lastly, (3.8) follows from the general identity $\max(a,b) = \frac{1}{2}(a+b+|a-b|)$. □

EXAMPLE 3.4. In accordance to the characterization of the independence and complete dependence case in terms of D-norms, we obtain in the case $\|\cdot\|_D = \|\cdot\|_1$

$$\mathrm{Cov}(\eta_1,\eta_2) = \int_0^\infty \frac{1}{(u+1)^2}\,\mathrm{d}u - 1 = 0$$

and in the case $\|\cdot\|_D = \|\cdot\|_\infty$

$$\mathrm{Cov}(\eta_1,\eta_2) = \int_0^\infty \frac{1}{(\max(u,1))^2}\,\mathrm{d}u - 1 = 1.$$

In particular, we have $\mathrm{Cov}(\eta_1,\eta_2) = \varrho(\eta_1,\eta_2) \in [0,1]$ for every bivariate standard max-stable random vector (X,Y) since the maximum norm is the least D-norm and the sum norm is the largest D-norm. In addition to this, we obtain for $\|\cdot\|_D = \|\cdot\|_2$

$$\mathrm{Cov}(\eta_1,\eta_2) = \int_0^\infty \frac{1}{(u^2+1)}\,\mathrm{d}u - 1 = \Big[\arctan(u)\Big]_0^\infty - 1 = \pi/2 - 1.$$

For a general logistic D-norm with parameter $\lambda \in [1,\infty)$ we obtain by substituting $u \mapsto u^{1/\lambda}$

$$\begin{aligned}\mathrm{Cov}(\eta_1,\eta_2) &= \int_0^\infty \frac{1}{(u^\lambda+1)^{2/\lambda}}\,\mathrm{d}u - 1\\ &= \frac{1}{\lambda}\int_0^\infty \frac{u^{1/\lambda-1}}{(u+1)^{2/\lambda}}\,\mathrm{d}u - 1\\ &= \frac{1}{\lambda}B\left(\frac{1}{\lambda},\frac{1}{\lambda}\right) - 1,\end{aligned}$$

where $B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1}\,\mathrm{d}u = \int_0^\infty u^{x-1}(1+u)^{-x-y}\,du$ denotes the Euler beta function.

Let $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s\in[0,1]^k}$ be the discretized version of $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]^k}$ with grid $\{s_1,\dots,s_d\}$ and weight functions $g_1,\dots,g_d$. In order to calculate the mean squared error of $\hat{\eta}_s$, we need the following lemma.

Lemma 3.5. *Let $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]^k}$ be the generator of $\hat{\boldsymbol{\eta}}$ that is defined in* (3.3). *For each $s \in [0,1]^k$, the random vector $(\eta_s,\hat{\eta}_s)$ is standard max-stable with generator $(Z_s,\hat{Z}_s)$ and*

D-norm

$$\|(x,y)\|_{\hat{D}_s} := E\left(|x|\, Z_s, |y|\, \hat{Z}_s\right) = \|(x, g_1(s)y, \dots, g_d(s)y)\|_{D_{s,s_1,\dots,s_d}}, \tag{3.9}$$

where $\|\cdot\|_{D_{s,s_1,\dots,s_d}}$ *is the D-norm pertaining to* $(\eta_s, \eta_{s_1}, \dots, \eta_{s_d})$.

Proof. As $\boldsymbol{Z} = (Z_s)_{s\in[0,1]^k}$ is a generator of $\boldsymbol{\eta}$, we have for $x, y \le 0$

$$\begin{aligned} P(\eta_s \le x, \hat{\eta}_s \le y) &= P(\eta_s \le x, \eta_{s_1} \le g_1(s)y, \dots, \eta_{s_d} \le g_d(s)y) \\ &= \exp\left(-E\left(\max\left(|x|\, Z_s, |y| \max\left(g_1(s)Z_{s_1}, \dots, g_d(s)Z_{s_d}\right)\right)\right)\right) \\ &= \exp\left(-E\left(\max\left(|x|\, Z_s, |y|\, \hat{Z}_s\right)\right)\right). \end{aligned}$$

The assertion now follows from the fact that $\hat{Z}_s \ge 0$ and $E(\hat{Z}_s) = 1$. □

We can now use the preceding Lemmas to compute the mean squared error of the discretized version.

Proposition 3.6. *The mean squared error of* $\hat{\eta}_s$ *is given by*

$$\mathrm{MSE}\left(\hat{\eta}_s\right) := E\left((\eta_s - \hat{\eta}_s)^2\right) = 2\left(2 - \int_0^\infty \frac{1}{(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t\right), \qquad s \in [0,1]^k.$$

Proof. Due to Lemma 3.5, $(\eta_s, \hat{\eta}_s)$ is standard max-stable. Therefore, (3.6) and the fact that $E(\eta_s) = E(\hat{\eta}_s) = -1$ and $\mathrm{Var}(\eta_s) = \mathrm{Var}(\hat{\eta}_s) = 1$ yield

$$\mathrm{MSE}\left(\hat{\eta}_s\right) = E\left(\eta_s^2\right) - 2E\left(\eta_s\hat{\eta}_s\right) + E\left(\hat{\eta}_s^2\right) = 4 - 2\int_0^\infty \frac{1}{(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t.$$

□

Example 3.7. Suppose, $(\eta_s, \eta_{s_1}, \dots, \eta_{s_d})$ follows the logistic D-norm

$$\|\boldsymbol{x}\|_{D_{s,s_1,\dots,s_d}} = \|\boldsymbol{x}\|_\lambda = \left(\sum_{i=1}^{d+1} |x_i|^\lambda\right)^{1/\lambda}, \quad \boldsymbol{x} \in \mathbb{R}^{d+1},$$

for some $\lambda \in [1, \infty]$. Clearly, the D-norm $\|\cdot\|_{s_1,\dots,s_d}$ pertaining to $(\eta_{s_1}, \dots, \eta_{s_d})$ is the logistic norm with parameter λ on $\mathbb{R}^d$ and, thus,

$$\|(x,y)\|_{\hat{D}_s} = \|(x, g_1(s)y, \dots, g_d(s)y)\|_{D_{s,s_1,\dots,s_d}}$$

$$= \left(|x|^\lambda + |y|^\lambda \sum_{i=1}^d g_i(s)^\lambda \right)^{1/\lambda}$$
$$= \left(|x|^\lambda + |y|^\lambda \left\|(g_1(s), \dots, g_d(s))\right\|_{s_1,\dots,s_d}^\lambda \right)^{1/\lambda}$$
$$= \left(|x|^\lambda + |y|^\lambda \right)^{1/\lambda}.$$

Hence the norm $\|\cdot\|_{\hat{D}_s}$ is the bivariate logistic norm with parameter λ. In this particular case, we obtain by substituting $t \mapsto t^{1/\lambda}$

$$\mathrm{MSE}\left(\hat{\eta}_s\right) = 4 - 2\int_0^\infty \frac{1}{(t^\lambda + 1)^{2/\lambda}}\, \mathrm{d}t = 4 - \frac{2}{\lambda}\int_0^\infty \frac{t^{1/\lambda - 1}}{(t+1)^{2/\lambda}}\, \mathrm{d}t = 4 - \frac{2}{\lambda} B\left(\frac{1}{\lambda}, \frac{1}{\lambda}\right),$$

where $B(\cdot,\cdot)$ denotes the Euler beta function, cf. Example 3.4. It is in evidence that, under the above assumptions, the mean squared error does not depend on the weights $g_1(s), \dots, g_d(s)$.

Lemma 3.8. *The mean squared error of $\hat{\eta}_s$ satisfies*

$$\mathrm{MSE}\left(\hat{\eta}_s\right) \le 6E\left(\left|Z_s - \hat{Z}_s\right|\right), \qquad s \in [0,1]^k.$$

Proof. We have

$$\begin{aligned}
&2 - \int_0^\infty \frac{1}{(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t \\
&= \int_0^\infty \frac{1}{\|(1,t)\|_\infty^2}\, \mathrm{d}t - \int_0^\infty \frac{1}{(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t \\
&= \int_0^\infty \left(\|(1,t)\|_{\hat{D}_s} - \|(1,t)\|_\infty \right) \frac{\|(1,t)\|_{\hat{D}_s} + \|(1,t)\|_\infty}{(\|(1,t)\|_{\hat{D}_s})^2 \|(1,t)\|_\infty^2}\, \mathrm{d}t \\
&= \int_0^1 \left(\|(1,t)\|_{\hat{D}_s} - 1 \right) \frac{\|(1,t)\|_{\hat{D}_s} + 1}{(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t + \int_1^\infty \left(\|(1,t)\|_{\hat{D}_s} - t \right) \frac{\|(1,t)\|_{\hat{D}_s} + t}{t^2(\|(1,t)\|_{\hat{D}_s})^2}\, \mathrm{d}t \\
&\le 3\int_0^1 \left(\|(1,t)\|_{\hat{D}_s} - 1 \right)\, \mathrm{d}t + 2\int_1^\infty \frac{\|(1/t,1)\|_{\hat{D}_s} - 1}{t^2}\, \mathrm{d}t \\
&=: 3I_1 + 2I_2.
\end{aligned}$$

Since every D-norm is monotone, we have

$$\|(1,t)\|_{\hat{D}_s} \le \|(1,1)\|_{\hat{D}_s},\ t \in [0,1], \text{ and } \|(1/t,1)\|_{\hat{D}_s} \le \|(1,1)\|_{\hat{D}_s},\ t > 1,$$

and, thus, by (3.8),

$$I_1 + I_2 \le \|(1,1)\|_{\hat{D}_s} - 1 + \left(\|(1,1)\|_{\hat{D}_s} - 1\right) \int_0^\infty t^{-2}\,\mathrm{d}t = E\left(\left|Z_s - \hat{Z}_s\right|\right).$$

□

3.2 Reconstruction of SMSP in $C([0,1])$

The preceding approach offers a way to reconstruct an SMSP with sample paths in $C([0,1])$ in an appropriate way. Originally, the work on generalized max-linear models started with the following discussion, which is mainly taken from Falk et al. (2015). The generalization to higher dimensions of the domain was subject of further research. Let $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ be an SMSP with generator process $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ and D-norm $\|\cdot\|_D$. Choose a grid $0 =: s_1 < s_2 < \cdots < s_{d-1} < s_d := 1$ of points within $[0,1]$. Again, $(\eta_{s_1},\dots,\eta_{s_d})$ is a standard max-stable random vector in $\mathbb{R}^d$ with pertaining D-norm $\|\cdot\|_{D_{s_1,\dots,s_d}}$ generated by $(Z_{s_1},\dots,Z_{s_d})$.

The aim of this section is to define some discretized version $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s\in[0,1]}$ for which $\hat{\eta}_{s_i} = \eta_{s_i}$, $i = 1,\dots,d$, holds, i.e. $\hat{\boldsymbol{\eta}}$ *interpolates* the finite dimensional projections $(\eta_{s_1},\dots,\eta_{s_d})$ of the original SMSP $\boldsymbol{\eta}$ in an appropriate way. This will be done by means of a special case of the generalized max-linear model, i.e., by a particular choice of the functions g_i in equation (3.2). In a next step, we consider a sequence of discretized versions $\hat{\boldsymbol{\eta}}^{(d)}$, $d \in \mathbb{N}$, with a grid that gets finer and finer, and show that this way of predicting the original MSP $\boldsymbol{\eta}$ in space is reasonable, as the pointwise mean squared error

$$\mathrm{MSE}\left(\hat{\eta}_s^{(d)}\right) = E\left(\left(\eta_s - \hat{\eta}_s^{(d)}\right)^2\right)$$

diminishes for all $s \in [0,1]$ as d increases. Moreover, we establish uniform convergence of the 'predictive' processes and the corresponding generator processes to the original ones. In order to help the reader to keep track of the strategy, most of the proofs of the following results will be presented again, since some of them have been omitted in the original paper.

Uniform convergence of the star-discretized versions

As we have shown in Lemma 3.1, the stochastic process $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s\in[0,1]}$,

$$\hat{\eta}_s = \max_{i=1,\dots,d} \frac{\eta_{s_i}}{g_i(s)}, \qquad s \in [0,1],$$

defines an SMSP with generator process $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]}$, given by

$$\hat{Z}_s = \max_{i=1,\dots,d} \left(g_i(s) Z_{s_i}\right), \qquad s \in [0,1],$$

for arbitrary functions $g_1, \dots, g_d$ in $\bar{C}^+[0,1]$ that satisfy condition (3.1). We are going to specialize them now.

Denote by $\|\cdot\|_{D_{i,i+1}}$ the D-norm pertaining to the bivariate random vector $(\eta_{s_i}, \eta_{s_{i+1}})$, $i = 1, \dots, d-1$. Put

$$g_1^*(s) := \begin{cases} \dfrac{s_2 - s}{\|(s_2 - s, s)\|_{D_{s_1,s_2}}}, & s \in [0, s_2], \\ 0, & \text{else}, \end{cases}$$

$$g_i^*(s) := \begin{cases} \dfrac{s - s_{i-1}}{\|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}}}, & s \in [s_{i-1}, s_i], \\ \dfrac{s_{i+1} - s}{\|(s_{i+1} - s, s - s_i)\|_{D_{s_i,s_{i+1}}}}, & s \in [s_i, s_{i+1}], \quad i = 2, \dots, d-1, \\ 0, & \text{else}, \end{cases}$$

$$g_d^*(s) := \begin{cases} \dfrac{s - s_{d-1}}{\|(s_d - s, s - s_{d-1})\|_{D_{s_{d-1},s_d}}}, & s \in [s_{d-1}, 1], \\ 0, & \text{else}. \end{cases}$$

Clearly, $g_0^*, \dots, g_d^* \in \bar{C}^+[0,1]$ since the fact that a D-norm is standardized implies

$$\lim_{s\uparrow s_i} g_i^*(s) = \frac{s_i - s_{i-1}}{\|(0, s_i - s_{i-1})\|_{D_{s_{i-1},s_i}}} = 1 = \frac{s_{i+1} - s_i}{\|(s_{i+1} - s_i, 0)\|_{D_{s_{i-1},s_i}}} = \lim_{s\downarrow s_i} g_i^*(s).$$

Moreover, we have for $s \in [s_{i-1}, s_i]$, $i = 2, \dots, d$,

$$\|(g_0^*(s), \dots, g_d^*(s))\|_{D_{s_1,\dots,s_d}} = \left\|\left(g_{i-1}^*(s), g_i^*(s)\right)\right\|_{D_{s_{i-1},s_i}} = 1.$$

Hence, the functions $g_0^*, \dots, g_d^*$ are suitable for the generalized max-linear model (3.2). In addition, they have the following property:

Lemma 3.9 (Falk et al. (2015, Lemma 2)). *The functions $g_0^*, \dots, g_d^*$ defined above satisfy*

$$\|g_i^*\|_\infty = g_i^*(s_i) = 1, \quad i = 1, \dots, d.$$

In consideration of their properties described above, the functions g_i^* can be viewed as kernel functions quite similar to kernels in nonparametric kernel density estimators.

Each function $g_i^*(s)$ has maximum value 1 at $s = s_i$ and, with the distance between s and s_i increasing, the value $g_i^*(s)$ shrinks to zero. This view provides also the idea behind the extension of this approach to higher dimension as described in Section 3.4.

Proof of Lemma 3.9. From the fact that a D-norm is monotone and standardized, we obtain for $i = 2, \dots, d-1$ and $s \in [s_{i-1}, s_i)$

$$g_i^*(s) = \frac{s - s_{i-1}}{\|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}}} = \frac{1}{\left\|\left(\frac{s_i - s}{s - s_{i-1}}, 1\right)\right\|_{D_{s_{i-1},s_i}}} \le \frac{1}{\|(0,1)\|_{D_{s_{i-1},s_i}}} = 1,$$

and for $s \in [s_i, s_{i+1})$

$$g_i^*(s) = \frac{s_{i+1} - s}{\|(s_{i+1} - s, s - s_i)\|_{D_{s_i,s_{i+1}}}} = \frac{1}{\left\|\left(1, \frac{s - s_i}{s_{i+1} - s}\right)\right\|_{D_{s_i,s_{i+1}}}} \le \frac{1}{\|(1,0)\|_{D_{s_i,s_{i+1}}}} = 1.$$

Analogously, we have $g_0^* \le 1$ and $g_d^* \le 1$. The assertion now follows since $g_i^*(s_i) = 1$, $i = 0, \dots, d$. □

The SMSP $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s \in [0,1]}$ that is generated by the generalized max-linear model with these particular functions $g_1^*, \dots, g_d^*$ is given by

$$\begin{aligned} \hat{\eta}_s &= \max\left(\frac{\eta_{s_{i-1}}}{g_{i-1}^*(s)}, \frac{\eta_{s_i}}{g_i^*(s)}\right) \\ &= \|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}} \max\left(\frac{\eta_{s_{i-1}}}{s_i - s}, \frac{\eta_{s_i}}{s - s_{i-1}}\right), \quad s \in [s_{i-1}, s_i],\ i = 2, \dots, d. \end{aligned} \tag{3.10}$$

Note that $\eta_{s_i} < 0$ almost surely, $i = 1, \dots, d$. This implies that the maximum taken over d points in (3.2) breaks down to a maximum taken over only two points in (3.10) since all except two of the g_i vanish in $[s_{i-1}, s_i]$, $i = 1, \dots, d$. We have, moreover,

$$\hat{\eta}_{s_i} = \eta_{s_i}, \qquad i = 1, \dots, d,$$

so the above process interpolates the random vector $(\eta_{s_1}, \dots, \eta_{s_d})$.

To give an example, put $\|(x_1, x_2)\|_{D_{s_{i-1},s_i}} := \|(x_1, x_2)\|_\lambda := (|x_1|^\lambda + |x_2|^\lambda)^{1/\lambda}$, $1 \le \lambda \le \infty$ for every $1 \le i \le d$, i.e. each bivariate D-norm $\|\cdot\|_{D_{s_{i-1},s_i}}$ is a logistic one. In this case, we obtain the representation

$$\hat{\eta}_s = ((s_i - s)^\lambda) + (s - s_{i-1})^\lambda)^{1/\lambda} \max\left(\frac{\eta_{s_{i-1}}}{s_i - s}, \frac{\eta_{s_i}}{s - s_{i-1}}\right), \tag{3.11}$$

$$s \in [s_{i-1}, s_i],\ i = 2, \dots, d,$$

which we will illustrate later. To summarize, we state the following result.

Corollary 3.10 (Falk et al. (2015, Corollary 1)). *Let $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ be an SMSP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$, and let $0 := s_1 < s_2 <, \dots, < s_{d-1} < s_d := 1$ be a grid of points in the interval $[0,1]$. The process $\hat{\boldsymbol{\eta}} = (\hat{\eta}_s)_{s\in[0,1]}$ defined in (3.10) is an SMSP with generator process $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]}$, where*

$$\hat{Z}_s = \frac{\max\Big((s_i - s)Z_{s_{i-1}}, (s - s_{i-1})Z_{s_i}\Big)}{\|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}}}, \qquad s \in [s_{i-1}, s_i],\ i = 2, \dots, d. \tag{3.12}$$

The processes $\hat{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{Z}}$ interpolate the random vectors $(\eta_{s_1}, \dots, \eta_{s_d})$ and $(Z_{s_1}, \dots, Z_{s_d})$, respectively.

We call $\hat{\boldsymbol{\eta}}$ the *star-discretized version* of $\boldsymbol{\eta}$ and $\hat{\boldsymbol{Z}}$ the star-discretized version of $\boldsymbol{Z}$, both with grid $\{s_1, \dots, s_d\}$. Next we show that the preceding approach allows the approximation of an underlying SMSP based on multivariate observations; that is, the star-discretized version of the underlying SMSP converges to the original process in a strong sense. We need the following two lemmata which provide some technical insight in the structure of the chosen max-linear model.

Lemma 3.11 (Falk et al. (2015, Lemma 3)). *The SMSP defined in* (3.10) *fulfills for $i = 2, \dots, d$*

$$\sup_{s\in[s_{i-1},s_i]} \hat{\eta}_s = \max\left(\eta_{s_{i-1}}, \eta_{s_i}\right),$$

and

$$\inf_{s\in[s_{i-1},s_i]} \hat{\eta}_s = -\left\|(\eta_{s_{i-1}}, \eta_{s_i})\right\|_{D_{s_{i-1},s_i}}.$$

This minimum is attained for $s = (s_{i-1}\eta_{s_{i-1}} + s_i\eta_{s_i})/(\eta_{s_{i-1}} + \eta_{s_i})$.

Proof. We know from Lemma 3.9 that $g_i^*(s) \le 1$ for an arbitrary $i = 1, \dots, d$ and $s \in [0,1]$. Hence,

$$\hat{\eta}_s = \max\left(\frac{\eta_{s_{i-1}}}{g_{i-1}^*(s)}, \frac{\eta_{s_i}}{g_i^*(s)}\right) \le \max(\eta_{s_{i-1}}, \eta_{s_i})$$

for $i = 2, \dots, d$ and $s \in [s_{i-1}, s_i]$, which yields the first part of the assertion. Recall that $\eta_{s_i} < 0$ with probability one, $i = 1 \dots, d$.

Moreover, we have for $s \in (s_{i-1}, s_i)$

$$\frac{\eta_{s_{i-1}}}{s_i - s} \le \frac{\eta_{s_i}}{s - s_{i-1}} \iff \frac{s_i - s}{s - s_{i-1}} \le \frac{\eta_{s_{i-1}}}{\eta_{s_i}} \iff s \ge \frac{s_{i-1}\eta_{s_{i-1}} + s_i\eta_{s_i}}{\eta_{s_{i-1}} + \eta_{s_i}},$$

where equality in one of these expressions occurs iff it does in the other two. In this case of equality we have

$$\hat{\eta}_s = \|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}} \cdot \frac{\eta_{s_i}}{s - s_{i-1}} = -\left\|(\eta_{s_{i-1}}, \eta_{s_i})\right\|_{D_{s_{i-1},s_i}}.$$

On the other hand, the monotonicity of a D-norm implies for every $s \in (s_{i-1}, s_i)$ with $s \ge (s_{i-1}\eta_{s_{i-1}} + s_i\eta_{s_i})/(\eta_{s_{i-1}} + \eta_{s_i})$

$$\begin{aligned}
\hat{\eta}_s &\ge \|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}} \frac{\eta_{s_i}}{s - s_{i-1}} \\
&= \left\|\left(\frac{s_i - s}{s - s_{i-1}}, 1\right)\right\|_{D_{s_{i-1},s_i}} \eta_{s_i} \\
&\ge \left\|\left(\frac{\eta_{s_{i-1}}}{\eta_{s_i}}, 1\right)\right\|_{D_{s_{i-1},s_i}} \eta_{s_i} \\
&= -\left\|(\eta_{s_{i-1}}, \eta_{s_i})\right\|_{D_{s_{i-1},s_i}}.
\end{aligned}$$

Recall again that $\eta_{s_i} < 0$ almost surely. The case $s \le (s_{i-1}\eta_{s_{i-1}} + s_i\eta_{s_i})/(\eta_{s_{i-1}} + \eta_{s_i})$ works analogously. □

As an immediate consequence of the preceding result we obtain for $x \le 0$

$$\hat{\boldsymbol{\eta}} \le x \iff \max\left(\eta_{s_0}, \dots, \eta_{s_d}\right) \le x$$

and

$$\hat{\boldsymbol{\eta}} > x \iff \max_{1 \le i \le d} \left\|\left(\eta_{s_{i-1}}, \eta_{s_i}\right)\right\|_{D_{s_{i-1},s_i}} < -x.$$

In order to visualize the interpolation scheme of this particular generalized max-linear model, we plot some discretized versions with different grids and bivariate D-norms $\|\cdot\|_{D_{i-1,i}}$. For the sake of simplicity, the underlying path $\boldsymbol{\eta}(\omega)$ in this example shall not arise from a simulation of an actual SMSP, but rather is replaced by a smooth deterministic continuous function on $[0,1]$. More precisely, we choose in the following picture $\eta_s(\omega) := 7.5(0.16s - 0.5s^2 + s^3/3) - 0.125$, $s \in [0,1]$ which is represented by the dashed curve. The solid line in each plot is the discretized version $\hat{\boldsymbol{\eta}}(\omega)$ of this path. We use equidistant grids of dimension $d = 5$, $d = 10$ and $d = 20$. Each bivariate D-norm

$\|\cdot\|_{D_{i-1,i}}$ are logistic norms such that the discretized versions are given by formula (3.11) with $\lambda = 2$, $\lambda = 4$ and $\lambda = 8$.

The plots apparently show that the approximation of the original process through a discretized version improves as the dimension d gets higher and as the bivariate D-norms get closer to complete dependence case.

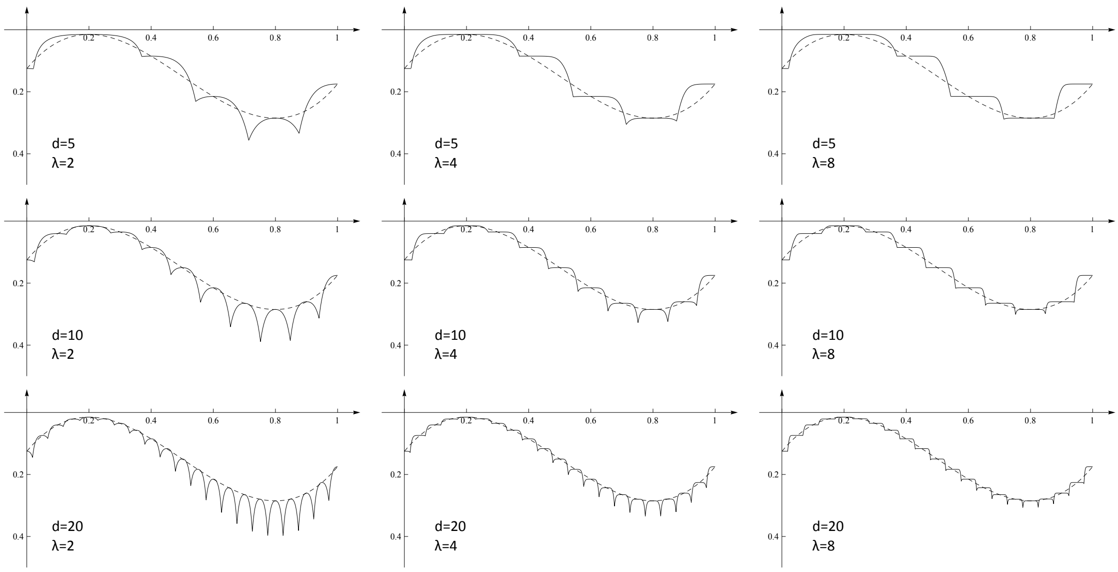

Fig. 3.1: *Plots of logistic type discretized versions (solid) of a deterministic function that stands for a path of an SMSP (dashed).*

The next lemma concerns the structure of the underlying generator processes.

Lemma 3.12 (Falk et al. (2015, Lemma 4))**.** *The generator process defined in* (3.12) *fulfills for* $i = 2, \dots, d$

$$\sup_{s\in[s_{i-1},s_i]} \hat{Z}_s = \max\left(Z_{s_{i-1}}, Z_{s_i}\right).$$

In particular, the extremal coefficient $E\left(||\hat{\boldsymbol{Z}}||_\infty\right)$ *of the SMSP* $\hat{\boldsymbol{\eta}}$ *coincides with the extremal coefficient* $E(\max_{i=1,\dots,d} Z_{s_i})$ *of the random vector* $(\eta_{s_1}, \dots, \eta_{s_d})$*. Moreover, for* $i = 2, \dots, d$*,*

$$\inf_{s\in[s_{i-1},s_i]} \hat{Z}_s = \begin{cases} \left(\left\|(1/Z_{s_{i-1}}, 1/Z_{s_i})\right\|_{D_{s_{i-1},s_i}}\right)^{-1} & \text{if} \quad Z_{s_{i-1}},\ Z_{s_i} > 0, \\ 0 & \text{else.} \end{cases}$$

In the first case, the minimum is attained for $s = (s_{i-1}Z_{s_i} + s_i Z_{s_{i-1}})/\left(Z_{s_{i-1}} + Z_{s_i}\right)$.

Proof. We know from Lemma 3.9 that $g_i^*(s) \le 1$ for an arbitrary $i = 1, \dots, d$ and $s \in [0,1]$. Hence,

$$\hat{Z}_s = \max\left(g_{i-1}^*(s)Z_{s_{i-1}}, g_i^*(s)Z_{s_i}\right) \le \max(Z_{s_{i-1}}, Z_{s_i}), \qquad i = 2, \dots, d,\ t \in [s_{i-1}, s_i],$$

which yields the first part of the assertion.

Moreover, we have for $s \in (s_{i-1}, s_i)$ in case of $Z_{s_{i-1}}, Z_{s_i} > 0$

$$(s_i - s)Z_{s_{i-1}} \le (s - s_{i-1})Z_{s_i} \iff \frac{s_i - s}{s - s_{i-1}} \le \frac{Z_{s_i}}{Z_{s_{i-1}}} \iff s \ge \frac{s_{i-1}Z_{s_i} + s_i Z_{s_{i-1}}}{Z_{s_{i-1}} + Z_{s_i}},$$

where equality in one of these expressions occurs iff it does in the other two. In this case of equality we have

$$\hat{Z}_s = \frac{(s - s_{i-1})Z_{s_i}}{\left\|(s_i - s, s - s_{i-1})\right\|_{D_{s_{i-1},s_i}}} = \frac{1}{\left\|(1/Z_{s_{i-1}}, 1/Z_{s_i})\right\|_{D_{s_{i-1},s_i}}}.$$

On the other hand, the monotonicity of a D-norm implies for every $s \in (s_{i-1}, s_i)$ with $s \ge (s_{i-1}Z_{s_i} + s_i Z_{s_{i-1}})/(Z_{s_{i-1}} + Z_{s_i})$

$$\begin{aligned}
\hat{Z}_s &\ge \frac{(s - s_{i-1})Z_{s_i}}{\left\|(s_i - s, s - s_{i-1})\right\|_{D_{s_{i-1},s_i}}} \\
&= \left(\left\|\left(\frac{s_i - s}{s - s_{i-1}}, 1\right)\right\|_{D_{s_{i-1},s_i}}\right)^{-1} Z_{s_i} \\
&\ge \left(\left\|\left(\frac{Z_{s_i}}{Z_{s_{i-1}}}, 1\right)\right\|_{D_{s_{i-1},s_i}}\right)^{-1} Z_{s_i} \\
&= \frac{1}{\left\|(1/Z_{s_{i-1}}, 1/Z_{s_i})\right\|_{D_{s_{i-1},s_i}}}.
\end{aligned}$$

The case $s \le (s_{i-1}Z_{s_i} + s_i Z_{s_{i-1}})/(Z_{s_{i-1}} + Z_{s_i})$ is shown analogously. □

So far we have only considered a *fixed* star-discretized version of an SMSP. The next step is to examine a *sequence* of star-discretized versions with certain grids whose diameter converges to zero. It turns out that such a sequence converges to the initial SMSP in the function space $C[0,1]$ equipped with the sup-norm. Thus, our method is suitable to reconstruct the initial process.

Let

$$\mathcal{G}_d := \left\{s_1^{(d)} \dots, s_d^{(d)}\right\}, \quad 0 =: s_1^{(d)} < s_2^{(d)} < \dots < s_{d-1}^{(d)} < s_d^{(d)} := 1, \qquad d \in \mathbb{N},$$

be a sequence of grids in $[0,1]$ with diameter

$$\kappa_d := \max_{i=2,\dots,d}\left(s_i^{(d)} - s_{i-1}^{(d)}\right) \to_{d\to\infty} 0.$$

Let $\hat{\boldsymbol{\eta}}^{(d)} = (\hat{\eta}_s^{(d)})_{s\in[0,1]}$ be the discretized version of an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]}$ with grid $\mathcal{G}_d$. Denote by $\hat{\boldsymbol{Z}}^{(d)} = (\hat{Z}_s^{(d)})_{s\in[0,1]}$ and $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$ the generator processes pertaining to $\hat{\boldsymbol{\eta}}^{(d)}$ and $\boldsymbol{\eta}$, respectively.

Theorem 3.13 (Falk et al. (2015, Theorem 1))**.** *The processes $\hat{\boldsymbol{\eta}}^{(d)}$ and $\hat{\boldsymbol{Z}}^{(d)}$, $d \in \mathbb{N}$, converge uniformly to $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ pathwise, i. e. $\left\|\hat{\boldsymbol{\eta}}^{(d)} - \boldsymbol{\eta}\right\|_\infty \to 0$ and $\left\|\hat{\boldsymbol{Z}}^{(d)} - \boldsymbol{Z}\right\|_\infty \to 0$ with probability one as $d \to \infty$.*

Proof. Denote by $[s]_d$ the left neighbor of $s \in [0,1]$ among $\mathcal{G}_d$, and by $\langle s\rangle_d$ the right neighbor of $s \in [0,1]$ among $\mathcal{G}_d$, $d \in \mathbb{N}$. Choose a sequence $s^{(d)} \in [0,1]$, $d \in \mathbb{N}$, with $s^{(d)} \to_{d\to\infty} s \in [0,1]$. Then obviously $[s^{(d)}]_d \to_{d\to\infty} s$ and $\langle s^{(d)}\rangle_d \to_{d\to\infty} s$. Hence we obtain by Lemma 3.11, and the continuity of the process $\boldsymbol{\eta}$

$$\hat{\eta}_{s^{(d)}}^{(d)} \le \max_{s\in\left[[s^{(d)}]_d, \langle s^{(d)}\rangle_d\right]} \hat{\eta}_s^{(d)} = \max\left(\eta_{[s^{(d)}]_d}, \eta_{\langle s^{(d)}\rangle_d}\right) \to_{d\to\infty} \eta_s,$$

as well as

$$\hat{\eta}_{s^{(d)}}^{(d)} \ge \min_{s\in\left[[s^{(d)}]_d, \langle s^{(d)}\rangle_d\right]} \hat{\eta}_s^{(d)} = -\left\|\left(\eta_{[s^{(d)}]_d}, \eta_{\langle s^{(d)}\rangle_d}\right)\right\|_{D_{[s^{(d)}]_d, \langle s^{(d)}\rangle_d}} \to_{d\to\infty} \eta_s,$$

where $\|\cdot\|_{D_{[s^{(d)}]_d, \langle s^{(d)}\rangle_d}}$ denotes the D-norm pertaining to $\left(\eta_{[s^{(d)}]_d}, \eta_{\langle s^{(d)}\rangle_d}\right)$. Hence the first part of the assertion is proven.

Now we show that $\hat{\boldsymbol{Z}}^{(d)} \to_{d\to\infty} \boldsymbol{Z}$ in $(C[0,1], \|\cdot\|_\infty)$. If $Z_s \ne 0$, the continuity of $\boldsymbol{Z}$ implies $Z_{[s^{(d)}]_d} \ne 0 \ne Z_{\langle s^{(d)}\rangle_d}$ for sufficiently large values of d. Repeating the above arguments, the assertion now follows by Lemma 3.12. If $Z_s = 0$, the continuity of $\boldsymbol{Z}$ implies

$$\hat{Z}_{s^{(d)}}^{(d)} \le 2\max\left(Z_{[s^{(d)}]_d}, Z_{\langle s^{(d)}\rangle_d}\right) \to_{d\to\infty} 2Z_s = 0,$$

which completes the proof. Check that $\left\|(\langle s^{(d)}\rangle_d - s, s - [s^{(d)}]_d)\right\|_D \ge 1/2$ since every D-norm is monotone and standardized. ☐

The preceding theorem is the main reason why we consider the discretized version $\hat{\boldsymbol{\eta}}$ of an SMSP $\boldsymbol{\eta}$ a reasonable predictor of this process, where the prediction is done in space, not in time. The predictions $\hat{\eta}_t$ of the points η_s, $s \in [0,1]$, only depend on the multivariate observations $(\eta_{s_1}, \dots, \eta_{s_d})$. More precisely, the only additional thing we need to know to make these predictions is the set of the adjacent bivariate marginal distributions of $(\eta_{s_1}, \dots, \eta_{s_d})$, that is, the bivariate D-norms $\|\cdot\|_{D_{i-1,i}}$, $i = 2, \dots, d$. This might, however, be a restrictive condition in praxis and suggests the problem to fit models of bivariate D-norms to data, which is, however, beyond the scope of the present discussion and requires future investigation.

The following results, however, are obvious. Let $\hat{\eta}_s$ be a point of the star-discretized version defined in (3.10) and define a *defective discretized version* via

$$\tilde{\eta}_s := \|(s_i - s, s - s_{i-1})\|_{\tilde{D}_i} \max\left(\frac{\eta_{s_{i-1}}}{s_i - s}, \frac{\eta_{s_i}}{s - s_{i-1}}\right), \qquad s \in [s_{i-1}, s_i],\ i = 2, \dots, d,$$

where $\|\cdot\|_{\tilde{D}_i}$ is an arbitrary norm on $\mathbb{R}^2$ which we call the *defective norm*. Then for every $s \in [s_{i-1}, s_i]$, $i = 1, \dots, d$,

$$|\hat{\eta}_s - \tilde{\eta}_s| = \\ \left| \|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_si}} - \|(s_i - s, s - s_{i-1})\|_{\tilde{D}_i} \right| \min\left(\frac{-\eta_{s_{i-1}}}{s_i - s}, \frac{-\eta_{s_i}}{s - s_{i-1}}\right).$$

In particular, we have $\tilde{\eta}_{s_i} = \hat{\eta}_{s_i} = \eta_{s_i}$, $i = 1, \dots, d$. This means that we obtain an interpolating process even if we replace the D-norm $\|\cdot\|_{D_{i-1,i}}$ by the defective norm $\|\cdot\|_{\tilde{D}_i}$. Furthermore, the defective discretized version still defines an MSP with sample paths in $\bar{C}^-[0,1]$. Its univariate marginal distributions are given by

$$P(\tilde{\eta}_s \le x) = \exp\left(\frac{\|(s_i - s, s - s_{i-1})\|_{D_{s_{i-1},s_i}}}{\|(s_i - s, s - s_{i-1})\|_{\tilde{D}_i}} x\right), \qquad x \le 0,\ s \in [s_{i-1}, s_i],\ i = 1, \dots, d.$$

In addition to this, the assertions in Lemma 3.11 also hold for the defective discretized version in case we know that each defective norm $\|\cdot\|_{\tilde{D}_i}$ is monotone and standardized. Repeating the arguments in the proof of Theorem 3.13 now shows that the uniform convergence towards the original process $\boldsymbol{\eta}$ is retained if we replace the norms $\|\cdot\|_{D_{i-1,i}}$ by arbitrary monotone and standardized norms (not necessarily D-norms) $\|\cdot\|_{\tilde{D}_i}$. Hence in that case, the only property of the discretized version that we have to drop is the standardization of the univariate margins.

The mean squared error of the star-discretized versions

In Proposition 3.6, the pointwise mean squared error of an arbitrary discretized version has been calculated. This can be used to derive convergence of the mean squared error of the star-discretized version to zero. As before, suppose $\boldsymbol{\eta}$ is an SMSP and choose a sequence of grids $\mathcal{G}_d$ of the interval $[0,1]$ with diameter $\kappa_d \to_{d\to\infty} 0$. Denote by $\hat{\boldsymbol{\eta}}^{(d)}$, $d \in \mathbb{N}$, the sequence of star-dicretized versions of $\boldsymbol{\eta}$ with grid $\mathcal{G}_d$. Denote further by $\|\cdot\|_{\hat{D}_s^{(d)}}$ the D-norm pertaining to $(\eta_s, \hat{\eta}_s^{(d)})$, $s \in [0,1]$, $d \in \mathbb{N}$.

Theorem 3.14 (Falk et al. (2015, Theorem 2))**.** *Let $\boldsymbol{\eta}$ and $\hat{\boldsymbol{\eta}}^{(d)}$, $d \in \mathbb{N}$, be as above. The mean squared error of $\hat{\eta}_s^{(d)}$ is given by*

$$\mathrm{MSE}\left(\hat{\eta}_s^{(d)}\right) = E\left(\left(\eta_s - \hat{\eta}_s^{(d)}\right)^2\right) = 2\left(2 - \int_0^\infty \frac{1}{(\|(1,u)\|_{\hat{D}_s^{(d)}})^2}\,\mathrm{d}u\right) \to_{d\to\infty} 0.$$

3.3 Reconstruction of SGPP

The preceding technique of discretizing and reconstructing a given SMSP can also be applied to the case of SGPP by simply replacing the standard max-stable random vector in the model (3.2) by a standard generalized Pareto distributed random vector. Again, the generalized max-linear model results in an SGPP. Once this statement is proven, most of the results of the previous sections carry over in a straightforward way.

Recall that a stochastic process $\boldsymbol{V}$ in $\bar{C}^-(S)$ is an SGPP, if there exists a D-norm $\|\cdot\|_D$ on $E(S)$ and some $c > 0$, such that $P(\boldsymbol{V} \le f) = 1 - \|f\|_D$ for all $f \in \bar{E}^-(S)$ with $\|f\|_\infty \le c$, cf. Remark 2.30. Note that this implies that each univariate marginal distribution of $\boldsymbol{V}$ coincides in the upper tail with the uniform distribution on $[-1,0]$. In the literature, there are different definitions of multivariate generalized Pareto distributions (GPD) available. For instance, Rootzén and Tajvidi (2006) define a d-dimensional generalized Pareto distribution function H to be

$$H(\boldsymbol{x}) = -\frac{1}{\log G(\mathbf{0})} \log \frac{G(\boldsymbol{x})}{G(\min(\boldsymbol{x}, \mathbf{0}))}, \qquad \boldsymbol{x} \in \mathbb{R}^d, \tag{3.13}$$

where G is a max-stable distribution function with $0 < G(\mathbf{0}) < 1$. Different to that, it is defined in Falk et al. (2011) that $\boldsymbol{Y} \in \mathbb{R}^d$ follows a standard GPD, if there exists a D-norm $\|\cdot\|_D$ on $\mathbb{R}^d$ and some $\boldsymbol{y}^{(0)} < \mathbf{0}$, such that

$$P(\boldsymbol{Y} \le \boldsymbol{y}) = 1 - \|\boldsymbol{y}\|_D, \qquad \boldsymbol{y}^{(0)} \le \boldsymbol{y} \le \mathbf{0}. \tag{3.14}$$

Even though both definitions are closely related, cf. Falk et al. (2011, Section 5.2), the approach of Rootzén and Tajvidi (2006) has the drawback that lower-dimensional margins of GPD are not necessarily GPD again, which obviously causes problems for the discretization and reconstruction of SGPP. For that reason, we prefer Definition (3.14) to (3.13), and work with that from now on.

Uniform Convergence of the Discretized Versions

Let $\boldsymbol{V} = (V_s)_{s\in[0,1]^k}$ be an SGPP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]^k}$ and D-norm $\|\cdot\|_D$. Choose arbitrary (and pairwise different) points $s_1, \dots, s_d \in [0,1]^k$. Then $(V_{s_1}, \dots, V_{s_d})$ is a standard GPD random vector in $\mathbb{R}^d$ with pertaining D-norm $\|\cdot\|_{D_{s_1,\dots,s_d}}$ generated by $(Z_{s_1}, \dots, Z_{s_d})$. Now choose deterministic functions $g_1, \dots, g_d \in \bar{C}^+\left([0,1]^k\right)$ with the property (3.1) and put

$$\hat{V}_s := \max_{i=1,\dots,d} \frac{V_{s_i}}{g_i(s)}, \qquad s \in [0,1]^k. \tag{3.15}$$

Lemma 3.15. *The stochastic process* $\hat{\boldsymbol{V}} = (V_s)_{s\in[0,1]^k}$ *in* (3.15) *defines an SGPP with generator process* $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]^k}$*,*

$$\hat{Z}_s = \max_{i=1,\dots,d} \left(g_i(s) Z_i\right), \quad s \in [0,1]^k. \tag{3.3}$$

Proof. We have already shown in Lemma 3.1 that $\hat{\boldsymbol{Z}}$ defines a generator process. Choose $c > 0$ such that $P(\boldsymbol{V} \le f) = 1 - \|f\|_D$ for $f \in \bar{E}^-\left([0,1]^k\right)$ with $\|f\|_\infty \le c$. Put $\hat{c} := c/\max_{i=1,\dots,d}\|g_i\|_\infty$. Then we have for all $f \in \bar{E}^-\left([0,1]^k\right)$

$$\|f\|_\infty \le \hat{c} \iff \inf_{s\in[0,1]^k} f(s) \ge -c/\max_{i=1,\dots,d}\|g_i\|_\infty .$$

Hence, for every f with $\|f\|_\infty \le \hat{c}$ and every $i = 1, \dots, d$,

$$\begin{aligned}
\inf_{s\in[0,1]^k}\left(g_i(s)f(s)\right) &\ge \inf_{s\in[0,1]^k} g_i(s) \cdot \left(-\frac{c}{\max_{j=1,\dots,d}\|g_j\|_\infty}\right) \\
&= \max_{j=1,\dots,d}\left(-c\frac{\inf_{s\in[0,1]^k} g_i(s)}{\sup_{s\in[0,1]^k} g_j(s)}\right) \\
&\ge -c\frac{\inf_{s\in[0,1]^k} g_i(s)}{\sup_{s\in[0,1]^k} g_i(s)} \\
&\ge -c.
\end{aligned}$$

To summarize, we have for all f close enough to zero (that is, $\|f\|_\infty \le \hat{c}$)

$$\begin{aligned}
P(\hat{\boldsymbol{V}} \le f) &= P\left(V_i \le g_i(s) f(s),\ i = 1, \dots, d,\ s \in [0,1]^k\right) \\
&= P\left(V_i \le \inf_{s\in[0,1]^k} \left(g_i(s) f(s)\right),\ i = 1, \dots, d\right) \\
&= 1 - \left\| \left(\sup_{s\in[0,1]^k} \left(g_1 \left|f(s)\right|\right), \dots, \sup_{s\in[0,1]^k} \left(g_d(s) \left|f(s)\right|\right) \right) \right\|_{D_{s_1,\dots,s_d}} \\
&= 1 - E\left(\max_{i=1,\dots,d} \left(\sup_{s\in[0,1]^k} \left(g_i(s) \left|f(s)\right|\right) Z_i \right) \right) \\
&= 1 - E\left(\sup_{s\in[0,1]^k} \left(\left|f(s)\right| \max_{i=1,\dots,d} \left(g_i(s) Z_i\right) \right) \right) \\
&= 1 - E\left(\sup_{s\in[0,1]^k} \left(\left|f(s)\right| \hat{Z}_s \right) \right),
\end{aligned}$$

which completes the proof. □

The generalized max-linear model is now applied to the case of SGPP. As before, the generalized max-linear model interpolates the underlying GPD random vector, the functions $g_1, \dots, g_d$ need to satisfy condition (3.5) in addition to (3.1). In particular, SGPP with sample paths in $C([0,1])$ can be reconstructed via the *star-discretized* versions considered in the previous section. By choosing a grid $0 := s_1 < s_2 < \dots < s_{d-1} < s_d := 1$ and the exact same functions $g_1^*, \dots, g_d^* \in \bar{C}^+([0,1])$ as in Section 3.2, the process $\hat{\boldsymbol{V}} = (\hat{V}_s)_{s\in[0,1]}$ from (3.15) turns into

$$\begin{aligned}
\hat{V}_s &= \max\left(\frac{V_{s_{i-1}}}{g_{i-1}^*(s)}, \frac{V_{s_i}}{g_i^*(s)} \right) \qquad (3.16) \\
&= \left\|(s_i - s, s - s_{i-1})\right\|_{D_{s_{i-1},s_i}} \max\left(\frac{V_{s_{i-1}}}{s_i - s}, \frac{V_{s_i}}{s - s_{i-1}} \right), \qquad s \in [s_{i-1}, s_i],\ i = 2, \dots, d.
\end{aligned}$$

Again, $\|\cdot\|_{D_{s_{i-1},s_i}}$ is the D-norm generated by $(Z_{s_{i-1}}, Z_{s_i})$, $i = 2, \dots, d$. In order to show that $\hat{\boldsymbol{V}}$ defines an SGPP as well, we only have to verify that the functions $g_0^*, \dots, g_d^*$ realize in $\bar{C}^+[0,1]$ and satisfy condition (3.1), which we have already done in Section 3.2. Thus, the following result is proven.

Corollary 3.16. *Let $\boldsymbol{V} = (V_s)_{s\in[0,1]}$ be an SGPP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]}$, and $0 := s_1 < s_2 <, \dots, < s_{d-1} < s_d := 1$ be a grid in the interval $[0,1]$. The process $\hat{\boldsymbol{V}} = (\hat{V}_s)_{s\in[0,1]}$ defined in* (3.16) *is an SGPP with the generator $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]}$ from* (3.12).

The processes $\hat{\boldsymbol{V}}$ *and* $\hat{\boldsymbol{Z}}$ *interpolate the random vectors* $(V_{s_1}, \ldots, V_{s_d})$ *and* $(Z_{s_1}, \ldots, Z_{s_d})$*, respectively.*

In complete accordance to the SMSP case we call $\hat{\boldsymbol{V}}$ the *star-discretized version* of $\boldsymbol{V}$ with grid $\{s_1, \ldots, s_d\}$.

REMARK 3.17. Let the original SGPP $\boldsymbol{V}$ satisfy $P(\boldsymbol{V} \le f) = 1 - \|f\|_D$ for all $f \in \bar{E}^-[0,1]$ with $\|f\|_\infty \le c$, with some D-norm $\|\cdot\|_D$ and some $c > 0$. Now denote by $\|\cdot\|_{\hat{D}}$ the D-norm pertaining to the discretized version $\hat{\boldsymbol{V}}$ of $\boldsymbol{V}$ with grid $\{s_1, \ldots, s_d\}$. Just like in the proof of Lemma 3.15, we obtain $P(\hat{\boldsymbol{V}} \le f) = 1 - \|f\|_{\hat{D}}$ for all $f \in \bar{E}^-([0,1])$ with $\|f\|_\infty \le \hat{c}$, where

$$\hat{c} = \frac{c}{\max_{i=1,\ldots,d} \|g_i^*\|_\infty} = c,$$

since $\|g_i^*\|_\infty = 1$ holds for all $i = 1, \ldots, d$ according to Lemma 3.9. Thus, the upper tail region where we know the distribution of the discretized version $\hat{\boldsymbol{V}}$ is just as large as that of the initial SGPP $\boldsymbol{V}$.

It is obvious that the pathwise structure of the star-discretized version of an SMSP we have established in Lemma 3.11 now carries over to the SGPP case since the assertion in this lemma follows solely from the structure of $g_1^*, \ldots, g_d^*$ and the fact that the initial process is nonpositive with probability one.

Lemma 3.18. *The SGPP defined in* (3.16) *fulfills for* $i = 2, \ldots, d$

$$\sup_{s \in [s_{i-1}, s_i]} \hat{V}_s = \max\left(V_{s_{i-1}}, V_{s_i}\right),$$

and

$$\inf_{s \in [s_{i-1}, s_i]} \hat{V}_s = -\left\|(V_{s_{i-1}}, V_{s_i})\right\|_{D_{s_{i-1}, s_i}}.$$

This minimum is attained for $s = (s_{i-1}V_{s_{i-1}} + s_i V_{s_i})/V_{s_{i-1}} + V_{s_i})$.

Now consider a sequence of star-discretized versions $\hat{\boldsymbol{V}}^{(d)}$ of an SGPP $\boldsymbol{V}$ with grid $\mathcal{G}_d$, where the diameter of $\mathcal{G}_d$ converges to zero. Repeating the arguments in the proof of Theorem 3.13 yields the following result.

Theorem 3.19 (Falk et al. (2015, Theorem 3))**.** *The sequence of processes* $\hat{\boldsymbol{V}}^{(d)}$*,* $d \in \mathbb{N}$*, converges uniformly to* $\boldsymbol{V}$ *pathwise, i. e.* $\left\|\hat{\boldsymbol{V}}^{(d)} - \boldsymbol{V}\right\|_\infty \to_{d \to \infty} 0$ *with probability one.*

The Mean Squared Error of the Discretized Version

The next aim is to calculate the mean squared error of the predictor $\hat{V}_s$ of V_s. We obtain again some kind of pointwise convergence in mean square of a sequence of star-discretized versions with decreasing diameter to the initial SGPP. Nevertheless, there is a difference to the considerations in the previous section. In contrast to the case of max-stable distributions, we typically only know the distribution of an SGPP in the upper tail. Note that the function $W(\boldsymbol{x}) := 1 - \|\boldsymbol{x}\|_D$, $\boldsymbol{x} \leq \boldsymbol{0}$, $\|\boldsymbol{x}\|_D \leq 1$, does not define a multivariate df in general, see cf. Falk et al. (2011). This fact forces us to consider *conditional expectations* in this section.

In the bivariate case, however, W defines a df, and we can assume that a GPD has this representation on the whole domain. The next Lemma is on some conditional moments of bivariate standard GPD random vectors in general.

Lemma 3.20. *Let (U, V) be a standard GPD random vector, i. e. there exists some D-norm $\|\cdot\|_D$ such that $P(U \leq u, V \leq v) = 1 - \|(u, v)\|_D$, $u, v \leq 0$, $\|(u, v)\|_D \leq 1$. Then we have for all such u, v*

(i)

$$P(U > u, V > v) = \wr\, (u, v) \,\wr_D = \|(u, v)\|_1 - \|(u, v)\|_D\,,$$

and, in case of $\|\cdot\|_D \neq \|\cdot\|_1$,

(ii)

$$\begin{aligned} &E(U^2 | U > u, V > v) \\ &\quad = -\frac{\frac{2}{3}u^3 + u^2\left(u + \|(u, v)\|_D\right) + v^3 \int_0^{u/v} \int_0^{u/v} \|(\max(t_1, t_2), 1)\|_D \, \mathrm{d}t_1 \, \mathrm{d}t_2}{\|(u, v)\|_1 - \|(u, v)\|_D}, \end{aligned}$$

(iii)

$$\begin{aligned} E(UV | U > u, V > v) = &- \frac{\int_v^0 \int_u^0 \|(t_1, t_2)\|_D \, \mathrm{d}t_1 \, \mathrm{d}t_2 + v^3 \int_0^{u/v} \|(t, 1)\|_D \, \mathrm{d}t}{\|(u, v)\|_1 - \|(u, v)\|_D} \\ &- \frac{u^3 \int_0^{v/u} \|(1, t)\|_D \, \mathrm{d}t + uv \|(u, v)\|_D}{\|(u, v)\|_1 - \|(u, v)\|_D}. \end{aligned}$$

Note that the case $\|\cdot\|_D = \|\cdot\|_1$ has to be treated with caution. It represents the case of uniform distribution on the line $\{(x,y) : x, y \le 0,\ x+y=-1\}$, which means that no observations fall in any rectangle $[u,0]\times[v,0]$, $u+v\ge -1$, cf. Falk et al. (2011).

Proof of Lemma 3.20. (i) The equation $P(U>u, V>v) = \wr\,(u,v)\,\wr_D$ is immediate from the inclusion-exclusion principle. Furthermore, we have

$$\begin{aligned}
P(U>u,V>v) &= 1 - P(U\le u) - P(V \le v) + P(U\le u, V\le v)\\
&= 1-(1+u)-(1+v)+1-\|(u,v)\|_D\\
&= \|(u,v)\|_1 - \|(u,v)\|_D\,.
\end{aligned}$$

(ii) We obtain by Fubini's theorem and elementary computations

$$\begin{aligned}
&E\left(\mathbb{1}_{\{U>u,V>v\}}U^2\right)\\
&= \int_{[u,0]\times[v,0]} x^2\ (P*(U,V))(\mathrm{d}(x,y))\\
&= \int_{[u,0]\times[v,0]} \left(\int_u^0\int_u^0 \mathbb{1}_{[x,0]}(t_1)\cdot\mathbb{1}_{[x,0]}(t_2)\ \mathrm{d}t_1\ \mathrm{d}t_2\right)\ (P*(U,V))(\mathrm{d}(x,y))\\
&= \int_u^0\int_u^0 \left(\int_{[u,0]\times[v,0]} \mathbb{1}_{[x,0]}(t_1)\cdot\mathbb{1}_{[x,0]}(t_2)\ (P*(U,V))(\mathrm{d}(x,y))\right)\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= \int_u^0\int_u^0 \left(\int_{[u,\min(t_1,t_2)]\times[v,0]} (P*(U,V))(\mathrm{d}(x,y))\right)\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= \int_u^0\int_u^0 P\left(U\in[u,\min(t_1,t_2)], V\in[v,0]\right)\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= \int_u^0\int_u^0 P\left(U\le \min(t_1,t_2)\right)\ \mathrm{d}t_1\ \mathrm{d}t_2 - \int_u^0\int_u^0 P\left(U\le u\right)\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&\quad + \int_u^0\int_u^0 P\left(U\le u, V\le v\right)\ \mathrm{d}t_1\ \mathrm{d}t_2 - \int_u^0\int_u^0 P\left(U\le \min(t_1,t_2), V\le v\right)\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= \int_u^0\int_u^0 1-\|(t_1,t_2)\|_\infty\ \mathrm{d}t_1\ \mathrm{d}t_2 - u^2\,(1+u)\\
&\quad + u^2\,(1-\|(u,v)\|_D) - \int_u^0\int_u^0 1-\|(\min(t_1,t_2),v)\|_D\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= -\frac{2}{3}u^3 - u^2\,(u+\|(u,v)\|_D) + \int_u^0\int_u^0 \|(\min(t_1,t_2),v)\|_D\ \mathrm{d}t_1\ \mathrm{d}t_2\\
&= -\frac{2}{3}u^3 - u^2\,(u+\|(u,v)\|_D) - v^3\int_0^{u/v}\int_0^{u/v} \|(\max(r_1,r_2),1)\|_D\ dr_1\ dr_2,
\end{aligned}$$

where we substitute $r_1 = t_1/v$ and $r_2 = t_2/v$ in the last equality. Together with assertion (i), the statement is proven.

(iii) Similar arguments as in the proof of (ii) yield

$$\begin{aligned}
E\left(\mathbb{1}_{\{U>u,V>v\}}UV\right) &= \int_v^0 \int_u^0 P\left(U \in [u,t_1], V \in [v,t_2]\right)\, \mathrm{d}t_1\, \mathrm{d}t_2 \\
&= \int_v^0 \int_u^0 P(U \le t_1, V \le t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 - \int_v^0 \int_u^0 P(U \le t_1, V \le v)\, \mathrm{d}t_1\, \mathrm{d}t_2 \\
&\quad - \int_v^0 \int_u^0 P(U \le u, V \le t_2)\, \mathrm{d}t_1\, \mathrm{d}t_2 + \int_v^0 \int_u^0 P(U \le u, V \le v)\, \mathrm{d}t_1\, \mathrm{d}t_2 \\
&= -\int_v^0 \int_u^0 \|(t_1,t_2)\|_D\, \mathrm{d}t_1\, \mathrm{d}t_2 + \int_u^0 \int_v^0 \|(t_1,v)\|_D\, \mathrm{d}t_2\, \mathrm{d}t_1 \\
&\qquad + \int_v^0 \int_u^0 \|(u,t_2)\|_D\, \mathrm{d}t_1\, \mathrm{d}t_2 - \int_v^0 \int_u^0 \|(u,v)\|_D\, \mathrm{d}t_1\, \mathrm{d}t_2 \\
&= -\int_v^0 \int_u^0 \|(t_1,t_2)\|_D\, \mathrm{d}t_1\, \mathrm{d}t_2 - v\int_u^0 \|(t_1,v)\|_D\, \mathrm{d}t_1 \\
&\qquad - u\int_v^0 \|(u,t_2)\|_D\, \mathrm{d}t_2 - uv\,\|(u,v)\|_D \\
&= -\int_v^0 \int_u^0 \|(t_1,t_2)\|_D\, \mathrm{d}t_1\, \mathrm{d}t_2 - v^3\int_0^{u/v} \|(r,1)\|_D\, \mathrm{d}r \\
&\qquad - u^3\int_0^{v/u} \|(1,r)\|_D\, \mathrm{d}r - uv\,\|(u,v)\|_D\,.
\end{aligned}$$

□

EXAMPLE 3.21. In case of total dependence of U and V (i.e. $\|\cdot\|_D = \|\cdot\|_\infty$) and $u = v =: c$, the formulas in Lemma 3.20 (ii) and (iii) become

$$E(U^2|U > u, V > v) = -\frac{\frac{2}{3}c^3 + c^2(c-c) + c^3}{-c} = \frac{5}{3}c^2$$

and

$$E(UV|U > c, V > c) = -\frac{-\int_c^0\int_c^0 \min(s,t)\; ds\; dt + c^3 + c^3 - c^3}{-c} = \frac{5}{3}c^2.$$

We now return to the discretized Version $\hat{\boldsymbol{V}}$ of an SGPP $\boldsymbol{V}$. Again, we have to show that for every $s \in [0,1]$ the random vector $(V_s, \hat{V}_s)$ follows a standard GPD. The exact same arguments as in Lemma 3.5 provide the bivariate df of this random vector, at least in the upper tail.

Lemma 3.22. *Let $\boldsymbol{V} = (V_s)_{s\in[0,1]^k}$ be an SGPP with generator $\boldsymbol{Z} = (Z_s)_{s\in[0,1]^k}$. Denote by $\hat{\boldsymbol{V}} = (\hat{V}_s)_{s\in[0,1]^k}$ its discretized version with grid $\{s_1, \dots, s_d\}$ and generator $\hat{\boldsymbol{Z}} = (\hat{Z}_s)_{s\in[0,1]}$. Then $(V_s, \hat{V}_s)$ defines a bivariate standard GPD random vector for every $s \in [0,1]$. Its df is given by*

$$P(V_s \le x, \hat{V}_s \le y) = 1 - \|(x,y)\|_{\hat{D}_s},$$

for x, y close enough to zero, where $\|\cdot\|_{\hat{D}_s}$ is the D-norm generated by $(Z_s, \hat{Z}_s)$, see (3.9).

Lemma 3.20 and Lemma 3.22 allow the calculation of the mean squared error of $\hat{V}_s$, under the condition that V_s and $\hat{V}_s$ attain values that are close enough to zero, such that there is a representation of the df of $(V_s, \hat{V}_s)$ available in this area. In particular, we can consider the star-discretized version and show that in this case, the mean squared error converges to zero.

Suppose $\boldsymbol{V}$ is an SGPP with sample paths in $C([0,1])$ with generator $\boldsymbol{Z}$ and choose, as in Section 3.2, a sequence of grids $\mathcal{G}_d$ of the interval $[0,1]$ with its fineness converging to zero as d increases. Denote by $\hat{\boldsymbol{V}}^{(d)}$, $d \in \mathbb{N}$, the sequence of star-dicretized versions of $\boldsymbol{V}$ with grid $\mathcal{G}_d$, and by $\hat{\boldsymbol{Z}}^{(d)}$, $d \in \mathbb{N}$, their generators. Denote further by $\|\cdot\|_{\hat{D}_s^{(d)}}$ the D-norm generated by $(Z_s, \hat{Z}_s^{(d)})$, $s \in [0,1]$, $d \in \mathbb{N}$.

Theorem 3.23. *Let $\boldsymbol{V}$ and $\hat{\boldsymbol{V}}^{(d)}$, $d \in \mathbb{N}$, be as above. Suppose $\|\cdot\|_{\hat{D}_s^{(d)}} \neq \|\cdot\|_1$ for d large enough. Then we have for c close enough to zero*

$$E\left(\left(V_s - \hat{V}_s^{(d)}\right)^2 \Big| V_s > c, \hat{V}_s^{(d)} > c\right) \to_{d\to\infty} 0.$$

Proof. According to Lemma 3.22, the random vector $\left(V_s, \hat{V}_s^{(d)}\right)$, $d \in \mathbb{N}$, is a standard GPD random vector with pertaining D-norm $\|\cdot\|_{\hat{D}_s^{(d)}}$. We have already shown in the proof of Theorem 3.14 that $\|\cdot\|_{\hat{D}_s^{(d)}} \to_{d\to\infty} \|\cdot\|_\infty$ pointwise. Substituting $\|\cdot\|_D$ by $\|\cdot\|_1$ in the numerators of Lemma 3.20 (ii) and (iii) leads to finite integrals exclusively. Hence, all these integrands are dominated by an integrable function, which is why we can apply the dominated convergence theorem in each of these integrals. Therefore, we obtain by the calculations in Example 3.21 for all $s \in [0,1]$

$$\begin{aligned}
&E\left(\left(V_s - \hat{V}_s^{(d)}\right)^2 \Big| V_s > c, \hat{V}_s^{(d)} > c\right) \\
&\quad = E\left(V_s^2 \big| V_s > c, \hat{V}_s^{(d)} > c\right) - 2E\left(V_s \hat{V}_s^{(d)} \big| V_s > c, \hat{V}_s^{(d)} > c\right) \\
&\qquad + E\left(\left(\hat{V}_s^{(d)}\right)^2 \Big| V_s > c, \hat{V}_s^{(d)} > c\right)
\end{aligned}$$

$$\to_{d\to\infty} \frac{10}{3}c^2 - \frac{10}{3}c^2 = 0.$$

□

3.4 Generalized max-linear models in arbitrary dimension

Although introduced quite generally with the domain $[0,1]^k$, $k \in \mathbb{N}$, the only concrete example of generalized max-linear models so far is based on the star-discretized versions, and restricted to the case $k = 1$. In this section, we will give possible extensions to higher-dimensional domains. As a start, we give a concrete example of an interpolating multivariate generalized max-linear model. As before, in the entire section $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]^k}$ is an SMSP with D-norm $\|\cdot\|_D$ generated by $\boldsymbol{Z} = (Z_s)_{s\in[0,1]^k}$. If $s_1, \dots, s_d$ are pairwise different points in $[0,1]^k$, the norm $\|\cdot\|_{D_{s_1,\dots,s_d}}$ denotes the D-norm generated by $(Z_{s_1}, \dots, Z_{s_d})$.

EXAMPLE 3.24. Choose pairwise different points $s_1, \dots, s_d \in [0,1]^k$ and an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^k$. Define

$$\tilde{h}_i(s) := \frac{\min_{j\neq i}(\|s - s_j\|)}{\min_{j\neq i}(\|s_i - s_j\|)}, \qquad s \in [0,1]^k,\ i = 1, \dots, d.$$

In order to normalize, put

$$\tilde{g}_i(s) := \frac{\tilde{h}_i(s)}{\left\|\left(\tilde{h}_1(s), \dots, \tilde{h}_d(s)\right)\right\|_{D_{s_1,\dots,s_d}}}, \qquad s \in [0,1]^k,\ i = 1, \dots, d.$$

These functions $\tilde{g}_i$ are well-defined since the denominator never vanishes: Suppose there is $s \in [0,1]^k$ with $\tilde{h}_1(s) = \dots = \tilde{h}_d(s) = 0$. Then $\min_{j\neq i}(\|s - s_j\|) = 0$ for all $i = 1, \dots, d$. Now fix $i \in \{1, \dots, d\}$. There is $j \neq i$ with $s = s_j$. But on the other hand, we have also $\min_{k\neq j}(\|s - s_k\|) = 0$ which implies that there is $k \neq j$ with $s = s_k = s_j$ which is a contradiction.

Clearly, g_i, $i = 1, \dots, d$, are functions in $\bar{C}^+\left([0,1]^k\right)$ that satisfy condition (3.1) and (3.5). Thus, we have found an interpolating generalized max-linear model on $C\left([0,1]^k\right)$.

A generalized max-linear model based on kernels

Unfortunately, the model from Example 3.24 is difficult to handle in terms of convergence to the original SMSP. Alternatively, one can consider the following model, which is at least 'close' to an interpolating generalized max-linear model. It was first introduced in Falk et al. (2015) and will be intensified here.

Let $K : [0, \infty) \to [0, 1]$ be a continuous and strictly monotonically decreasing function (kernel) with the two properties

$$K(0) = 1, \qquad \lim_{x\to\infty} \frac{K(ax)}{K(bx)} = 0, \qquad 0 \le b < a. \tag{3.17}$$

The *exponential kernel* $K_e(x) = \exp(-x)$, $x \ge 0$, is a typical example. Choose an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^k$ and a grid of pairwise different points $\{s_1, \dots, s_d\}$ in $[0, 1]^k$. Put for $i = 1, \dots, d$ and the *bandwidth* $h > 0$

$$g_{i,h}(s) := \frac{K(\|s - s_i\| /h)}{\|(K(\|s - s_1\| /h), \dots, K(\|s - s_d\| /h))\|_{D_{s_1,\dots,s_d}}}, \qquad s \in [0, 1]^k.$$

Define for $i = 1, \dots, d$

$$N(s_i) := \left\{ s \in [0, 1]^k : \|s - s_i\| \le \|s - s_j\|, \, j \ne i \right\}, \tag{3.18}$$

which is the set of those points $s \in [0, 1]^k$ that are closest to the grid point s_i.

Lemma 3.25. *For arbitrary $s \in [0, 1]^k$ and $i = 1, \dots, d$,*

$$g_{i,h}(s) \to_{h\downarrow 0} \begin{cases} 1, & \text{if } s = s_i \\ 0, & \text{if } s \notin N(s_i) \end{cases}$$

as well as $g_{i,h}(s) \le 1$.

Proof. The convergence $g_{i,h}(s_i) \to_{h\downarrow 0} 1$ follows from the fact that $K(0) = 1$ and that the D-norm is standardized. The fact that an arbitrary D-norm is bounded below by the sup-norm together with the monotonicity of K implies for $s \in [0, 1]^k$

$$g_{i,h}(s) \le \frac{K(\|s - s_i\| /h)}{\max_{1\le j\le d} K(\|s - s_j\| /h)} = \frac{K(\|s - s_i\| /h)}{K(\min_{1\le j\le d} \|s - s_j\| /h)} \le 1.$$

Note that $K(\|s - s_i\| /h) / K(\min_{1\le j\le d} \|s - s_j\| /h) \to_{h\downarrow 0} 0$ if $s \notin N(s_i)$ by the required growth condition on the kernel K in (3.17). □

The above Lemma shows in particular $g_{i,h}(s_j) \to_{h\downarrow 0} \delta_{ij}$ which is close to condition (3.5). Obviously, the functions $g_{i,h}$ are constructed in such a way that condition (3.1) holds, too. Therefore, we obtain the *kernel-based generalized max-linear model*

$$\hat{\eta}_{s,h} = \max_{i=1,\dots,d} \frac{\eta_{s_i}}{g_{i,h}(s)}, \qquad s \in [0,1]^k,$$

which does not interpolate $(\eta_{s_1}, \dots, \eta_{s_d})$ exactly, but $\hat{\eta}_{s_i,h}$ converges to η_{s_i} as $h \downarrow 0$. Note that the limit functions $\lim_{h\downarrow 0} g_{i,h}$ are not necessarily continuous: For instance, there may be $s_0 \in [0,1]^k$ with $\|s_0 - s_1\| = \dots = \|s_0 - s_d\|$. Then $s_0 \in \partial N(s_1)$ and $\lim_{h\downarrow 0} g_{1,h}(s_0) = 1/\|(1,\dots,1)\|_{D_{s_1,\dots,s_d}}$, but $\lim_{h\downarrow 0} g_{1,h}(s) = 0$ for all $s \notin N(s_1)$ due to Lemma 3.25.

Next we investigate a sequence of kernel-based generalized max-linear models, where the diameter of the grids decreases. We analyze under which conditions the integrated mean squared error of $(\hat{\eta}_{s,h})_{s\in[0,1]^k}$ converges to zero. We start with a general result on generator processes.

Lemma 3.26. *Let $(Z_s)_{s\in[0,1]^k}$ be a generator of an SMSP and ε_n, $n \in \mathbb{N}$, be a null sequence. Then*

$$E\left(\sup_{\|s-t\|\le\varepsilon_n} |Z_s - Z_t|\right) \to_{n\to\infty} 0,$$

where $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^k$.

Proof. The paths of $(Z_s)_{s\in[0,1]^k}$ are continuous, so they are also uniformly continuous. Therefore, $\sup_{\|s-t\|\le\varepsilon_n} |Z_s - Z_t| \to_{n\to\infty} 0$. Furthermore,

$$E\left(\sup_{\|s-t\|\le\varepsilon_n} |Z_s - Z_t|\right) \le 2E\left(\sup_{s\in[0,1]^k} Z_s\right) < \infty.$$

The assertion now follows from the dominated convergence theorem. □

Let $\mathcal{G}_n := \{s_{1,n}, \dots, s_{d(n),n}\}$, $n \in \mathbb{N}$, be a set of distinct points in $[0,1]^k$ with the property

$$\forall n \in \mathbb{N}\ \forall s \in [0,1]^k\ \exists\, s_{i,n} \in \mathcal{G}_n: \ \|s - s_{i,n}\| \le \varepsilon_n,$$

where $\varepsilon_n \to_{n\to\infty} 0$. Define, for instance, $\mathcal{G}_n$ in such a way that

$$\varepsilon_n := \max_{i=1,\dots,d} \sup_{s,t\in N(s_{i,n})} \|s - t\| \to_{n\to\infty} 0,$$

with $N(s_{i,n})$ as defined in (3.18). Clearly, $d := d(n) \to_{n\to\infty} \infty$. Denote by $\|\cdot\|_{D^{(n)}_{1,\dots,d}}$ the D-norm pertaining to $\eta_{s_{1,n}}, \dots, \eta_{s_{d,n}}$. Let further $\hat{\boldsymbol{\eta}}_n = (\hat{\eta}_{s,n})_{s\in[0,1]^k}$ be the kernel-based discretized version of $\boldsymbol{\eta}$ with grid $\mathcal{G}_n$, that is,

$$\hat{\eta}_{s,n} = \max_{i=1,\dots,d} \frac{\eta_{s_{i,n}}}{g_{i,n}(s)}, \qquad s \in [0,1]^k,$$

where for $i = 1, \dots, d$

$$g_{i,n}(s) = \frac{K(\|s - s_{i,n}\| / h_n)}{\|(K(\|s - s_{1,n}\| / h_n), \dots, K(\|s - s_{d,n}\| / h_n))\|_{D^{(n)}_{s_1,\dots,s_d}}}, \qquad s \in [0,1]^k,$$

where $K : [0, \infty) \to [0,1]$ is the continuous and strictly decreasing kernel function satisfying condition (3.17) and h_n, $n \in \mathbb{N}$, is some positive sequence. We have already seen in Lemma 3.25 that $g_{i,n}(s) \in [0,1]$, $s \in [0,1]^k$, $n \in \mathbb{N}$. Furthermore, we have the following result.

Lemma 3.27. *Choose $s \in [0,1]^k$. There is a sequence $i(n)$, $n \in \mathbb{N}$, such that $s \in \bigcap_{n\in\mathbb{N}} N(s_{i(n),n})$. Define $g_{i(n),n}$ and ε_n as above, $n \in \mathbb{N}$. Then*

$$\lim_{n\to\infty} g_{i(n),n}(s) = 1,$$

if $\varepsilon_n \to_{n\to\infty} 0$, $h_n \to_{n\to\infty} 0$, $\varepsilon_n / h_n \to_{n\to\infty} \infty$.

Proof. Let $s \in [0,1]^k$ and choose a sequence $i(n)$, $n \in \mathbb{N}$, as above. Put for simplicity $s_{i(n),n} =: s_{i,n}$ and $g_{i(n),n} =: g_{i,n}$. We have

$$\begin{aligned}
1 \ge g_{i,n}(s) &= \frac{K\left(\|s - s_{i,n}\| / h_n\right)}{E\left(\max_{j=1,\dots,d} K\left(\|s - s_{j,n}\| / h_n\right) Z_{s_{j,n}}\right)} \\
&\ge \left(\frac{E\left(\max_{j:\|s_{j,n} - s\| \ge 2\varepsilon_n} K\left(\|s - s_{j,n}\| / h_n\right) Z_{s_{j,n}}\right)}{K\left(\|s - s_{i,n}\| / h_n\right)} \right. \\
&\qquad \left. + \frac{E\left(\max_{j:\|s_{j,n} - s\| < 2\varepsilon_n} K\left(\|s - s_{j,n}\| / h_n\right) Z_{s_{j,n}}\right)}{K\left(\|s - s_{i,n}\| / h_n\right)} \right)^{-1} \\
&=: \left(A_{i,n}(s) + B_{i,n}(s)\right)^{-1}.
\end{aligned}$$

From $s \in N(s_{i,n})$ we conclude $\|s - s_{i,n}\| \le \varepsilon_n$. Hence, we have due to the properties of the kernel function K

$$0 \le A_{i,n}(s) \le \frac{K(2\varepsilon_n / h_n)}{K(\varepsilon_n / h_n)} E\left(\sup_{s\in[0,1]^k} Z_s \right) \to_{n\to\infty} 0,$$

since $\varepsilon_n/h_n \to_{n\to\infty} \infty$ by assumption. Furthermore, $s \in N(s_{i,n})$ and the fact that K is decreasing implies

$$\max_{j:\|s_{j,n}-s\|<2\varepsilon_n} K\left(\|s-s_{j,n}\|/h_n\right) = K\left(\|s-s_{i,n}\|/h_n\right).$$

Thus,

$$\begin{aligned}
1 \le B_{i,n}(s) &= \frac{1}{K\left(\|s-s_{i,n}\|/h_n\right)}\Bigg(E\bigg(\max_{j:\|s_{j,n}-s\|<2\varepsilon_n} K\left(\|s-s_{j,n}\|/h_n\right) Z_{s_{j,n}} \\
&\qquad\qquad - \max_{j:\|s_{j,n}-s\|<2\varepsilon_n} K\left(\|s-s_{j,n}\|/h_n\right) Z_{s_{i,n}}\bigg)\Bigg) + 1 \\
&\le \frac{E\left(\max_{j:\|s_{j,n}-s\|<2\varepsilon_n} K\left(\|s-s_{j,n}\|/h_n\right)\left|Z_{s_{j,n}} - Z_{s_{i,n}}\right|\right)}{K\left(\|s-s_{i,n}\|/h_n\right)} + 1 \\
&\le E\left(\max_{j:\|s_{j,n}-s\|<2\varepsilon_n} \left|Z_{s_{j,n}} - Z_{s_{i,n}}\right|\right) + 1 \\
&\le E\left(\sup_{\|r-t\|<3\varepsilon_n} |Z_r - Z_t|\right) + 1 \\
&\to_{n\to\infty} 1,
\end{aligned}$$

because of Lemma 3.26. Note that $\|s_{j,n}-s\| < 2\varepsilon_n$ and $s \in N(s_{i,n})$ imply $\|s_{j,n}-s_{i,n}\| < 3\varepsilon_n$. □

We have now gathered the tools to proof the convergence of the mean squared error to zero.

Theorem 3.28. *Define $\hat{\boldsymbol{\eta}}_n$ and ε_n as above, $n \in \mathbb{N}$. Then for every $s \in [0,1]^k$*

$$\mathrm{MSE}\left(\hat{\eta}_{s,n}\right) \to_{n\to\infty} 0,$$

and

$$\mathrm{IMSE}\left(\hat{\eta}_{s,n}\right) := \int_{[0,1]^k} \mathrm{MSE}\left(\hat{\eta}_{s,n}\right)\, \mathrm{d}s \to_{n\to\infty} 0,$$

if $\varepsilon_n \to_{n\to\infty} 0$, $h_n \to_{n\to\infty} 0$, $\varepsilon_n/h_n \to_{n\to\infty} \infty$.

Proof. Denote by

$$\hat{Z}_{s,n} = \max_{i=1,\dots,d}\left(g_{i,n}(s) Z_{s_{i,n}}\right), \qquad s \in [0,1]^k,$$

the generator of $\hat{\boldsymbol{\eta}}_n$. Choose $s \in [0,1]^k$ and a sequence $i := i(n)$, $n \in \mathbb{N}$, such that $s \in \bigcap_{n\in\mathbb{N}} N(s_{i,n})$. We have by Lemma 3.8, Lemma 3.27 and the continuity of $\boldsymbol{Z}$

$$\begin{aligned}
\operatorname{MSE}(\hat{\eta}_{s,n}) &\leq 6E\left(\left|Z_s - \hat{Z}_{s,n}\right|\right) \\
&\leq 6E\left(\left|Z_s - Z_{s_{i,n}}\right|\right) + 6E\left(\left|Z_{s_{i,n}} - g_{i,n}(s)Z_{s_{i,n}}\right|\right) \\
&\quad + 6E\left(\left|g_{i,n}(s)Z_{s_{i,n}} - \hat{Z}_{s,n}\right|\right) \\
&= 6E\left(\left|Z_s - Z_{s_{i,n}}\right|\right) + 12\left(1 - g_{i,n}(s)\right) \\
&\to_{n\to\infty} 0.
\end{aligned}$$

Next we prove the convergence of the integrated mean squared error. The sets $N(s_{i,n})$, as defined in (3.18), are typically not disjoint, but the intersections $N(s_{i,n}) \cap N(s_{j,n})$, $i \neq j$, have Lebesgue measure zero on $[0,1]^k$. Clearly, $\bigcup_{i=1}^d N(s_{i,n}) = [0,1]^k$. Therefore, applying Lemma 3.8 yields

$$\begin{aligned}
\operatorname{IMSE}(\hat{\eta}_{s,n}) &= \sum_{i=1}^d \int_{N(s_{i,n})} \operatorname{MSE}(\hat{\eta}_{s,n}) \, \mathrm{d}s \\
&\leq 6\sum_{i=1}^d \int_{N(s_{i,n})} E\left(\left|Z_s - \hat{Z}_{s,n}\right|\right) \, \mathrm{d}s \\
&\leq 6\Bigg(\sum_{i=1}^d \int_{N(s_{i,n})} E\left(\left|Z_s - Z_{s_{i,n}}\right|\right) \, \mathrm{d}s \\
&\qquad + \sum_{i=1}^d \int_{N(s_{i,n})} \left|1 - g_{i,n}(s)\right| E\left(Z_{s_{i,n}}\right) \, \mathrm{d}s \\
&\qquad + \sum_{i=1}^d \int_{N(s_{i,n})} E\left(\left|g_{i,n}(s)Z_{s_{i,n}} - \hat{Z}_{s,n}\right|\right) \, \mathrm{d}s\Bigg) \\
&=: 6\left(S_{1,n} + S_{2,n} + S_{3,n}\right)
\end{aligned}$$

due to Lemma 3.8. From Lemma 3.26 we conclude

$$\begin{aligned}
S_{1,n} &= \sum_{i=1}^d \int_{N(s_{i,n})} E\left(\left|Z_s - Z_{s_{i,n}}\right|\right) \, \mathrm{d}s \\
&\leq \sum_{i=1}^d \int_{N(s_{i,n})} E\left(\sup_{\|r-t\|\leq\varepsilon_n} |Z_r - Z_t|\right) \, \mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&= \int_{[0,1]^k} E\left(\sup_{\|r-t\|\le\varepsilon_n} |Z_r - Z_t|\right)\,\mathrm{d}s \\
&= E\left(\sup_{\|r-t\|\le\varepsilon_n} |Z_r - Z_t|\right) \\
&\to_{n\to\infty} 0.
\end{aligned}
$$

Define

$$
A_n := \frac{K(2\varepsilon_n/h_n)}{K(\varepsilon_n/h_n)} E\left(\sup_{s\in[0,1]^k} Z_s\right), \quad B_n := E\left(\sup_{\|r-t\|<3\varepsilon_n} |Z_r - Z_t|\right) + 1.
$$

As we have seen in the proof of Lemma 3.27, we have for $s \in N(s_{i,n})$

$$
1 \ge g_{i,n}(s) \ge (A_n + B_n)^{-1} \to 1,
$$

and therefore

$$
\begin{aligned}
S_{2,n} &= \sum_{i=1}^{d} \int_{N(s_{i,n})} (1 - g_{i,n}(s))\,\mathrm{d}s \\
&\le \sum_{i=1}^{d} \int_{N(s_{i,n})} 1 - (A_n + B_n)^{-1}\,\mathrm{d}s \\
&= \int_{[0,1]^k} 1 - (A_n + B_n)^{-1}\,\mathrm{d}s \\
&= 1 - (A_n + B_n)^{-1} \\
&\to_{n\to\infty} 0.
\end{aligned}
$$

Lastly, we have by the same argument as above

$$
S_{3,n} = \sum_{i=1}^{d} \int_{N(s_{i,n})} E\left(\hat{Z}_{s,n} - g_{i,n}(s) Z_{s_{i,n}}\right)\,\mathrm{d}s = S_{2,n} \to_{n\to\infty} 0,
$$

which completes the proof. □

Discretized versions of copula processes

The last task in that chapter will be the transformation of the generalized max-linear model to copula processes that are in a sense 'close' to max-stable processes. In Section 2.2 we have dealt with *copula processes* $\boldsymbol{U} = (U_s)_{s\in[0,1]^k}$, i. e. with stochastic processes with continuous sample paths, such that each random variable U_s is uniformly distributed on the interval $[0,1]$. We have also defined that $\boldsymbol{U}$ is in the *functional domain of attraction* of an SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in[0,1]^k}$, if

$$\lim_{n\to\infty} P\left(n\left(\boldsymbol{U}-1\right)\leq f\right)^n = P\left(\boldsymbol{\eta}\leq f\right) = \exp\left(-\left\|f\right\|_D\right), \qquad f\in \bar{E}^-\left([0,1]^k\right). \quad (3.19)$$

Define for any $s\in[0,1]^k$ and $n\in\mathbb{N}$

$$Y_s^{(n)} := n\left(\max_{i=1,\dots,n} U_s^{(i)} - 1\right),$$

with $\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ being independent copies of $\boldsymbol{U}$. In the sense of (3.19), the process $\boldsymbol{Y}^{(n)} = (Y_s^{(n)})_{s\in[0,1]^k}$ is close to the SMSP $\boldsymbol{\eta}$ for large values of n. Choose again pairwise different points $s_1,\dots,s_d\in[0,1]^k$ and functions $g_1,\dots,g_d\in\bar{C}^+\left([0,1]^k\right)$ with the properties (3.12) and (3.5). Condition (3.19) implies weak convergence of the finitedimensional distributions of $\boldsymbol{Y}^{(n)}$, i. e.

$$\left(Y_{s_1}^{(n)},\dots,Y_{s_d}^{(n)}\right) \to_{\mathcal{D}} \left(\eta_{s_1},\dots,\eta_{s_d}\right).$$

Just like before, we can define the *discretized version* $\hat{\boldsymbol{Y}}^{(n)} = (\hat{Y}_s^{(n)})_{s\in[0,1]^k}$ of $\boldsymbol{Y}^{(n)}$ with grid $\{s_1,\dots,s_d\}$ and weight functions $g_1,\dots,g_d$ to be

$$\hat{Y}_s^{(n)} := \max_{i=1,\dots,d} \frac{Y_{s_i}^{(n)}}{g_i(s)}, \qquad s\in[0,1]^k.$$

Elementary calculations show that (3.19) implies

$$\lim_{n\to\infty} P\left(\hat{\boldsymbol{Y}}^{(n)}\leq f\right) = P\left(\hat{\boldsymbol{\eta}}\leq f\right), \qquad f\in\bar{E}^-\left([0,1]^k\right),$$

where $\hat{\boldsymbol{\eta}}$ is the discretized version of $\boldsymbol{\eta}$ as defined in (3.2). Also, it is not difficult to see that for each $s\in[0,1]^k$,

$$\left(Y_s^{(n)},\hat{Y}_s^{(n)}\right) \to_{\mathcal{D}} \left(\eta_s,\hat{\eta}_s\right)$$

where $(\eta_s,\hat{\eta}_s)$ is the standard max-stable random vector from Lemma 3.5. Now applying the continuous mapping theorem, we obtain

$$\left(Y_s^{(n)}-\hat{Y}_s^{(n)}\right)^2 \to_{\mathcal{D}} \left(\eta_s-\hat{\eta}_s\right)^2.$$

It remains to prove uniform integrability of the sequence on the left hand side in order to obtain the next result.

Proposition 3.29. *Let* $s \in [0,1]^k$. *Then*

$$\mathrm{MSE}\left(\hat{Y}_s^{(n)}\right) = E\left(\left(Y_s^{(n)} - \hat{Y}_s^{(n)}\right)^2\right) \to_{n\to\infty} \mathrm{MSE}\left(\hat{\eta}_s\right).$$

Proof. Fix $s \in [0,1]^k$. It remains to show that the sequence $X_s^{(n)} := \left(Y_s^{(n)} - \hat{Y}_s^{(n)}\right)^2$ is uniformly integrable. A sufficient condition for uniform integrability is

$$\sup_{n\in\mathbb{N}} E\left(\left(X_s^{(n)}\right)^2\right) < \infty,$$

see Billingsley (1999, Section 3). Clearly, for every $n \in \mathbb{N}$,

$$E\left(\left(X_s^{(n)}\right)^2\right) \le E\left(\left(Y_s^{(n)}\right)^4\right) + E\left(\left(\hat{Y}_s^{(n)}\right)^4\right).$$

It is easy to verify that the random variable $Y_s^{(n)}$ has the density $(1+x/n)^{n-1}$ on $[-n,0]$. Therefore,

$$E\left(\left(Y_s^{(n)}\right)^4\right) = \int_{-n}^0 x^4 \left(1+\frac{x}{n}\right)^{n-1} \,\mathrm{d}x = \frac{24n^5(n-1)!}{(n+4)!} \le 24.$$

Moreover, putting $c := \min_{i=1,\dots,d} g_i(s) > 0$,

$$\left|\hat{Y}_s^{(n)}\right| = \min_{i=1,\dots,d} \frac{\left|Y_{s_i}^{(n)}\right|}{g_i(s)} \le \frac{\left|Y_{s_1}^{(n)}\right|}{c},$$

and hence

$$E\left(\left(\hat{Y}_s^{(n)}\right)^4\right) \le \frac{24}{c^4},$$

which completes the proof. □

4 Higherdimensional records

Records among a sequence of iid random variables $X^{(1)}, X^{(2)}, \dots$ on the real line have been investigated extensively over the past decades, see e. g. the monographs of Galambos (1987, Sections 6.2 and 6.3), Resnick (2008, Chapter 4) and Arnold et al. (1998). A record is defined as a random variable $X^{(n)}$ such that $X^{(n)} > \max(X^{(1)}, \dots, X^{(n-1)})$. Trying to generalize this concept to the case of random vectors, or even stochastic processes with continuous sample paths, the question arises how to define records in higher dimensions. While in the univariate case the definition of a record is rather inherently determined, the lack of a natural order in the case $d \geq 2$ gives rise to many possible definitions of records. In this chapter, we discuss different concepts of higherdimensional records and investigate their stochastic behavior. For instance, we are interested in the probability that a stochastic process $\boldsymbol{X}^{(n)}$ is a record as n tends to infinity.

The concept of records is closely related to classical extreme value theory. A crucial assumption for most of the results in this chapter will be that the observations are in the domain of attraction of an MSP. This is where extreme value theory will step in. Note that the domain-of-attraction condition is not very restrictive - it is actually not an easy task to find distributions that are *not* in the domain of attraction of some MSP. However, an example of such a distribution is given in Galambos (1987, Example 2.6.1).

Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be an iid sequence of stochastic processes in $C(S)$, where S denotes an arbitrary compact metric space as throughout Chapter 2. During the whole chapter, we assume that each univariate margin is continuously distributed in order to prevent ties within the sequence, i. e. we want to make sure that

$$P(X_s^{(i)} = X_s^{(j)}) = 0, \qquad i \neq j,\ s \in S.$$

We define

$$\boldsymbol{X}^{(n)} \textit{ simple record } :\iff \boldsymbol{X}^{(n)} \not\leq \max_{i=1,\dots,n-1} \boldsymbol{X}^{(i)}$$

and

$$\boldsymbol{X}^{(n)} \textit{ complete record } :\iff \boldsymbol{X}^{(n)} > \max_{i=1,\dots,n-1} \boldsymbol{X}^{(i)}.$$

Put further

$$\pi_n^s(\boldsymbol{X}) := P\left(\boldsymbol{X}^{(n)} \text{ is a simple record }\right),$$
$$\pi_n^c(\boldsymbol{X}) := P\left(\boldsymbol{X}^{(n)} \text{ is a complete record }\right).$$

By definition, the first observation $\boldsymbol{X}^{(1)}$ is always a record, so we have $\pi_1^s(\boldsymbol{X}) = \pi_1^c(\boldsymbol{X}) = 1$. In the univariate case, where $X, X^{(1)}, X^{(2)}, \dots$ are simply random variables on the real line, records are much easier to handle. It was Rényi (1962) who proved in that case

$$\pi_n^s(X) = \pi_n^c(X) = \frac{1}{n},$$

using a very simple combinatorial argument: There are $n!$ permutations of the observations $X^{(1)}, \dots, X^{(n)}$, including $(n-1)!$ permutations where $X^{(n)}$ is on the last place. Bearing in mind that $X^{(1)}, \dots, X^{(n)}$ are independent and continuously distributed yields $\pi_n^s(X) = \pi_n^c(X) = (n-1)!/n! = 1/n$.

As mentioned before, there are many detailed works on univariate records and record times. A lot of results can be found in the monographs of Resnick (2008, Chapter 4), Galambos (1987, Sections 6.2 and 6.3) and Arnold et al. (1998). Multivariate records have not been discussed that extensively, yet there are also many papers available on that subject. In the seminal paper of Goldie and Resnick (1989), for instance, it is shown that under a domain-of-attraction condition, there are either finitely many or infinitely many complete records almost surely among a sequence of iid bivariate random vectors, depending on whether or not the max-stable limit distribution has independent margins. Gnedin (1998) studies the asymptotic behavior of $\pi_n^c(\boldsymbol{X})$ under the assumption that the observations are iid normally distributed. In Hashorva and Hüsler (2005), the limit of $\pi_n^c(\boldsymbol{X})$ is investigated without assuming a certain dependence model for the margins. Hwang and Tsai (2010) provide a central limit theorem for the number of simple records within a sequence, albeit they make strong conditions on the dependence structure within the components by assuming the observations are uniformly distributed on the d-variate simplex. Similarly, Gnedin (2007) proves a central limit theorem for the number of so-called *chain records*, under the assumption that the observations follow a product distribution. In our work, we only touch on the asymptotic number of complete and simple records in Corollary 4.2 and 4.13 below, merely assuming the underlying copula process is in the domain of attraction of a max-stable process.

Unfortunately there is no consistent term for the different kinds of multivariate records, and often, they are just referred to as records. Gnedin (2007) proposes the terms *strong* and *weak records* instead of complete and simple records. However, this might cause confusion since there exists the same terminology in the univariate setup to distinguish whether or not records of discrete observations are strictly greater than the previous observations, cf. Vervaat (1973). Furthermore, complete and simple records are not the only possible definition of multivariate records. For other suggestions, see e. g. Kałuszka (1995) or Gnedin (2007). For more results on multivariate records, see e. g. Barndorff-Nielsen and Sobel (1966), Gnedin (1993, 1994a,b), Goldie and Resnick (1995), Deuschel and Zeitouni (1995), Arnold et al. (1998, Chapter 4) and Chen et al. (2012), as well as the references therein.

A concept that is closely related to the field of complete records is the so-called *concurrency of extremes*, which is due to Dombry et al. (2015). We say that $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ are *sample concurrent*, if

$$\max_{i=1,\dots,n} \boldsymbol{X}^{(i)} = \boldsymbol{X}^{(k)} \text{ for some } k \in \{1, \dots, n\}.$$

In that case, we call $\boldsymbol{X}^{(k)}$ the *champion* among $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$. We denote the sample concurrence probability by $p_n(\boldsymbol{X})$ and obtain due to the iid property

$$\begin{aligned} p_n(\boldsymbol{X}) &= P\left(\bigcup_{i=1}^{n} \left\{ \boldsymbol{X}^{(i)} > \max_{1 \le j \ne i \le n} \boldsymbol{X}^{(j)} \right\}\right) \qquad (4.1) \\ &= \sum_{i=1}^{n} P\left(\boldsymbol{X}^{(i)} > \max_{1 \le j \ne i \le n} \boldsymbol{X}^{(j)}\right) = nP\left(\boldsymbol{X}^{(n)} > \max_{j=1,\dots,n-1} \boldsymbol{X}^{(j)}\right) = n\pi_n^c(\boldsymbol{X}). \end{aligned}$$

Different to records, the concept of multivariate and functional champions is very recent. It has been established in the work of Dombry et al. (2015). In their paper, they derive the limit sample concurrence probability under iid random vectors $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ in $\mathbb{R}^d$. There are also some results on statistical inference in their work.

In Section 4.1, we generalize the limit sample concurrence probability which has been derived in Dombry et al. (2015, Theorem 2) to the case of stochastic processes with continuous sample paths on an arbitrary compact metric space S. Further, we compute the conditional distribution of a champion, given that there actually is one. Section 4.2 deals with simple record times and the distribution of simple records, where all considerations are restricted to the finite-dimensional case. Finally, we will specify simple records in Section 4.3 in order to get a better understanding of the evolution of the componentwise maxima $\max\left(\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}\right)$ as n increases.

4.1 The functional extremal concurrence probability

The aim of this section is to investigate the limit behavior of the sample concurrence probability. In Dombry et al. (2015), it is shown that the sample concurrence probability $p_n(\boldsymbol{X})$ of a random vector $\boldsymbol{X}$ converges, provided that $\boldsymbol{X}$ has continuously distributed margins and lies in the domain of attraction of a max-stable random vector. We generalize this assertion to stochastic process with continuous sample paths on a compact metric space S.

We start with a stochastic process $\boldsymbol{X} = (X_s)_{s\in S}$, such that $F_s(x) := P(X_s \le x)$, $x \in \mathbb{R}$, is continuous on $\mathbb{R}$ for each $s \in S$. Then the process $(F_s(X_s))_{s\in S}$ is a copula process, meaning that it has continuous sample paths and the univariate margins are uniformly distributed on $(0,1)$. Since continuity of the df F_s, $s \in S$, is assumed, we have

$$P(X_s > Y_s) = P\left(F_s(X_s) > F_s(Y_s)\right), \qquad s \in S.$$

if $\boldsymbol{Y} = (Y_s)_{s\in S}$ is an independent copy of $\boldsymbol{X}$. If F_s is strictly monotonically increasing in addition, we even have

$$P(\boldsymbol{X} > \boldsymbol{Y}) = P(F_s(X_s) > F_s(Y_s), s \in S). \tag{4.2}$$

Note that in the multivariate case $S = \{1, \dots, d\}$, (4.2) is still valid even if the strict monotonicity of the marginal df is omitted. Hence, dealing with records, it is reasonable to pay particular attention to the case where the observations follow a copula process. If a copula process $\boldsymbol{U}$ is in the domain of attraction of an MSP $\boldsymbol{\eta}$ ($\boldsymbol{U} \in \mathcal{D}(\boldsymbol{\eta})$), i. e.

$$n\left(\max_{i=1,\dots,n} \boldsymbol{U}^{(i)} - \boldsymbol{1}\right) \to_{\mathcal{D}} \boldsymbol{\eta}, \tag{2.23}$$

then it is clear from the univariate margins that $\boldsymbol{\eta}$ is necessarily an SMSP. Besides the continuity of the univariate margins, (2.23) will be the crucial assumption in the following. The SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ satisfies $P(\boldsymbol{\eta} \le f) = \exp\left(-\|f\|_D\right)$, $f \in \bar{E}^-(S)$, for some D-norm

$$\|f\|_D = E\left(\sup_{s\in S} |f(s)|\, Z_s\right), \qquad f \in E(S),$$

with generator $\boldsymbol{Z} = (Z_s)_{s\in S}$. In Section 2.2, we introduced the *dual D-norm function* $\wr \cdot \wr_D$ corresponding to $\|\cdot\|_D$ via

$$\wr f \wr_D = E\left(\inf_{s\in S} |f(s)|\, Z_s\right), \qquad f \in E(S).$$

In Proposition 2.29, we have seen that assuming $\boldsymbol{U} \in \mathcal{D}(\boldsymbol{\eta})$ for some SMSP with D-norm $\|\cdot\|_D$, the value $\|f\|_D$ arises as the limit of $nP(\boldsymbol{U} \not\leq 1 + f/n)$, as $n \to \infty$, whereas $\wr f \wr_D$ is the limit of $nP(\boldsymbol{U} > 1 + f/n)$, as $n \to \infty$, $f \in \bar{E}^-(S)$. This is a first hint on what role $\|\cdot\|_D$ and $\wr \cdot \wr_D$ will play in the world of records: It seems that $\|\cdot\|_D$ is related to *simple* records, whereas $\wr \cdot \wr_D$ corresponds to the case of *complete* records. The case of simple records is postponed to Section 4.2. In this section, we start with the limit sample concurrence probability, and hence by (4.1) with complete records. For the finitedimensional version of the following theorem, see Dombry et al. (2015, Theorems 1 and 2), or, in a different version, Hashorva and Hüsler (2005, Theorem 2.1).

Theorem 4.1. *Let $\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ be independent copies of a copula process $\boldsymbol{U}$, satisfying $\boldsymbol{U} \in \mathcal{D}(\boldsymbol{\eta})$, where $\boldsymbol{\eta}$ is an SMSP with corresponding D-norm $\|\cdot\|_D$. Then*

$$p_n(\boldsymbol{U}) = n\pi_n^c(\boldsymbol{U}) \to_{n\to\infty} E\left(\wr \boldsymbol{\eta} \wr_D\right),$$

where $\wr \cdot \wr_D$ is the dual D-norm function corresponding to $\|\cdot\|_D$.

We call $E\left(\wr \boldsymbol{\eta} \wr_D\right)$ the *extremal concurrence probability* corresponding to $\|\cdot\|_D$, in accordance with the terminology in Dombry et al. (2015). As it is shown therein, the extremal concurrence probability has the following interpretation. Remember the Poisson point process representation (2.18) of the simple MSP $\boldsymbol{\xi} := -1/\boldsymbol{\eta}$, that is

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k\in\mathbb{N}} \boldsymbol{\vartheta}^{(k)},$$

where $\boldsymbol{\vartheta}^{(k)} = \zeta^{(k)}\boldsymbol{V}^{(k)}$, $k \in \mathbb{N}$, and $\left(\left(\zeta^{(k)}, \boldsymbol{V}^{(k)}\right)\right)_{k\in\mathbb{N}}$ are the points of PPP(ν) on $(0,\infty] \times \bar{C}_1^+(S)$, with ν being the exponent measure, see (2.16). The extremal concurrence probability is now precisely the probability that only one function $\boldsymbol{\vartheta}^{(k)}$ contributes to the supremum in (2.18), see Dombry et al. (2015, Theorem 1) and also Remark 4.28 below. The concept of extremal concurrence is embedded in the more general framework of *extremal hitting scenarios*, which will be investigated in Section 4.3.

Note that one has to distinguish between $E\left(\wr \boldsymbol{\eta} \wr_D\right)$ and $E\left(\inf_{s\in S} |\eta_s|\, Z_s\right)$ in general. However, if $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ are independent, both terms coincide, cf. Lemma 4.3 below.

Proof of Theorem 4.1. Let $\boldsymbol{\eta}$ be an SMSP with D-norm $\|\cdot\|_D$ and put

$$\boldsymbol{M}^{(n)} := n \max_{i=1,\dots,n-1} \left(\boldsymbol{U}^{(i)} - \mathbf{1}\right) \to_{\mathcal{D}} \boldsymbol{\eta}$$

due to (2.23). Conditioning on $\boldsymbol{M}^{(n)} = f$ yields

$$\begin{aligned} n\pi_n^c(\boldsymbol{U}) &= \int_{C^-(S)} nP\left(n(\boldsymbol{U}-1) > f\right) \ \left(P * \boldsymbol{M}^{(n)}\right)(\mathrm{d}f) \\ &=: \int_{C^-(S)} G_n(f) \ \left(P * \boldsymbol{M}^{(n)}\right)(\mathrm{d}f), \end{aligned}$$

since $\boldsymbol{M}^{(n)}$ and $\boldsymbol{U}$ are independent. Setting $X_n := G_n \circ \boldsymbol{M}^{(n)}$, we need to show

$$n\pi_n^c(\boldsymbol{U}) = E(X_n) \to_{n\to\infty} = E\left(\wr\, \boldsymbol{\eta} \wr_D\right).$$

It is enough to verify (Billingsley (1968, p. 32)):

(i) $X_n \to_{\mathcal{D}} \wr\, \boldsymbol{\eta} \wr_D$.

(ii) There is $\varepsilon > 0$ with $\sup_{n\in\mathbb{N}} E\left(|X_n|^{1+\varepsilon}\right) < \infty$.

Note that (ii) implies the *uniform integrability* of the sequence $(X_n)_{n\in\mathbb{N}}$.

We first show (i). Let $f_n, f \in C^-(S)$ with $\|f_n - f\|_\infty \to_{n\to\infty} 0$. Choose $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $f - \varepsilon \le f_n \le f + \varepsilon$ for all $n \ge N$. Clearly, for such n,

$$G_n(f-\varepsilon) \ge G_n(f_n) \ge G_n(f+\varepsilon).$$

It has been shown in Proposition 2.29 that $G_n(f \pm \varepsilon) \to_{n\to\infty} \wr\, f \pm \varepsilon \wr_D$, which yields

$$\wr\, f - \varepsilon \wr_D \ge \lim_{n\to\infty} G_n(f_n) \ge \wr\, f + \varepsilon \wr_D.$$

Letting $\varepsilon \downarrow 0$, we obtain $\wr\, f \pm \varepsilon \wr_D \to \wr\, f \wr_D$ by the monotone convergence theorem, and hence $G_n(f_n) \to_{n\to\infty} \wr\, f \wr_D$. Now noticing that $\boldsymbol{M}^{(n)} \to_{\mathcal{D}} \boldsymbol{\eta}$, the assertion is immediate from the extended continuous mapping theorem, see cf. Billingsley (1968, Theorem 5.5).

Now we proof (ii). Elementary calculations show that for fixed $s_0 \in S$ and all $n \ge 2$

$$\begin{aligned} E\left(X_n^2\right) &= \int_{C^-(S)} n^2 P\left(n(\boldsymbol{U}-1) > f\right)^2 \ \left(P * \boldsymbol{M}^{(n)}\right)(\mathrm{d}f) \\ &\le \int_{C^-(S)} n^2 P\left(n\left(U_{s_0} - 1\right) > f(s_0)\right)^2 \ \left(P * \boldsymbol{M}^{(n)}\right)(\mathrm{d}f) \end{aligned}$$

$$= \int_{C^-(S)} f(s_0)^2 \left(P * \boldsymbol{M}^{(n)}\right)(\mathrm{d}f)$$
$$= E\left(\left(M_{s_0}^{(n)}\right)^2\right) = \frac{2n}{n+1} \leq 2.$$

To verify the second to last equality, check that $M_{s_0}^{(n)}$ is a random variable with Lebesgue density $h(x) = (n-1)/n \cdot (1 + x/n)^{n-2}$, $x \in [-n, 0]$. □

Corollary 4.2. *Denote by* $N_c(n) := \sum_{i=1}^n 1_{\left\{\boldsymbol{X}^{(i)} > \max_{1\leq j<i} \boldsymbol{X}^{(j)}\right\}}$ *the number of complete records among* $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)}$. *Then*

$$\frac{E(N_c(n))}{\log n} \to_{n\to\infty} E\left(\wr\, \boldsymbol{\eta} \,\wr_D\right).$$

Proof. The assertion follows from Theorem 4.1 and the fact that $\left(\sum_{i=1}^n \frac{a_i}{i}\right)/\log n$ $\to_{n\to\infty} a$, if $(a_n)_{n\in\mathbb{N}}$ is some real-valued sequence with $a_n \to_{n\to\infty} a$. □

The following lemma provides an alternative representation for the extremal concurrence probability.

Lemma 4.3. *Let* $\boldsymbol{\eta} = (\eta_s)_{s\in S}$ *be an SMSP with D-norm* $\|\cdot\|_D$ *generated by* $\boldsymbol{Z} = (Z_s)_{s\in S}$*, and* $f \in \bar{E}^-(S)$*. Then*

(i)

$$E\left(\wr\, \boldsymbol{\eta} \,\wr_D\right) = E\left(\|1/\boldsymbol{Z}\|_D^{-1}\, \mathbb{1}_{\{\boldsymbol{Z}>0\}}\right).$$

(ii)

$$E\left(\wr \max(\boldsymbol{\eta}, f) \wr_D\right) =$$
$$= E\left(\left(\|1/\boldsymbol{Z}\|_D\right)^{-1}\left(1 - \exp\left(\|1/\boldsymbol{Z}\|_D \sup_{s\in S}(f(s)Z_s)\right)\right)\mathbb{1}_{\{\boldsymbol{Z}>0\}}\right).$$

Proof. Without loss of generality, choose a generator $\boldsymbol{Z}$ of $\|\cdot\|_D$ which is independent of $\boldsymbol{\eta}$. Then

$$E\left(\inf_{s\in S} |\eta_s|\, Z_s\right) = \int_{\bar{C}^-(S)} \wr\, f \,\wr_D \, (P * \boldsymbol{\eta})(\mathrm{d}f) = E\left(\wr\, \boldsymbol{\eta} \,\wr_D\right).$$

Suppose $P(\boldsymbol{Z} > 0) = 1$ for ease of notation. Fubini's theorem and the fact that $\boldsymbol{\eta}$ and $\boldsymbol{Z}$ are independent, entail

$$E\left(\inf_{s\in S} (|\eta_s|\, Z_s)\right) = \int_0^\infty P\left(\inf_{s\in S} (|\eta_s|\, Z_s) > t\right) \,\mathrm{d}t$$

$$
\begin{aligned}
&= \int_0^\infty P\left(\eta_s < -t/Z_s,\ s \in S\right)\ \mathrm{d}t \\
&= E\left(\int_0^\infty \exp\left(-t\left\|1/\boldsymbol{Z}\right\|_D\right)\ \mathrm{d}t\right) \\
&= E\left(\left(\left\|1/\boldsymbol{Z}\right\|_D\right)^{-1}\int_0^\infty \exp\left(-t\right)\ \mathrm{d}t\right),
\end{aligned}
$$

which is (i). Assertion (ii) can be shown by similar arguments: Assuming $P(\boldsymbol{Z} > 0) = 1$, we have for $f \in \bar{E}^-(S)$

$$
\begin{aligned}
&E\left(\inf_{s\in S}\left(\left|\max(\eta_s, f(s))\right| Z_s\right)\right) \\
&= \int_0^\infty P\left(\inf_{s\in S}\left(\left|\max(\eta_s, f(s))\right| Z_s\right) > t\right)\ \mathrm{d}t \\
&= \int_0^\infty \int_{\bar{C}^-(S)} P\left(\left|\eta_s\right| z(s) > t, \left|f(s)\right| z(s) > t,\ s \in S\right)\ (P * \boldsymbol{Z})(\mathrm{d}z)\ \mathrm{d}t \\
&= \int_{\bar{C}^-(S)} \int_0^{\inf_{s\in S}|f(s)|z(s)} P\left(\eta_s < -\frac{t}{z(s)},\ s \in S\right)\ \mathrm{d}t\ (P * \boldsymbol{Z})(\mathrm{d}z) \\
&= \int_{\bar{C}^-(S)} \int_0^{\inf_{s\in S}|f(s)|z(s)} \exp\left(-t\left\|(1/z)\right\|_D\right)\ \mathrm{d}t\ (P * \boldsymbol{Z})(\mathrm{d}z) \\
&= \int_{\bar{C}^-(S)} \left(\frac{1-\exp\left(-\inf_{s\in S}\left(|f(s)|\, z(s)\right)\left\|(1/z)\right\|_D\right)}{\left\|(1/z)\right\|_D}\right)\ (P * \boldsymbol{Z})(\mathrm{d}z) \\
&= E\left(\left(\left\|1/\boldsymbol{Z}\right\|_D\right)^{-1}\right) - E\left(\left(\left\|1/\boldsymbol{Z}\right\|_D\right)^{-1}\exp\left(\left\|1/\boldsymbol{Z}\right\|_D \inf_{s\in S}\left(|f(s)|\, Z_s\right)\right)\right).
\end{aligned}
$$

□

EXAMPLE 4.4 (Independence and perfect dependence). A generator of the special D-norm $\|\cdot\|_D = \|\cdot\|_\infty$, which characterizes the complete dependence of the univariate margins of $\boldsymbol{\eta}$, is given by the constant $\boldsymbol{Z} \equiv 1$. In that case, Theorem 4.1 shows that the extremal concurrence probability is one, i. e. $p_n(\boldsymbol{U}) = n\pi_n^c(\boldsymbol{U}) \to_{n\to\infty} 1$. This is not at all surprising: in the univariate context, where $X^{(1)}, \dots, X^{(n)}$ are random variables on the real line, there clearly exists a champion with probability one - it is the maximum of $X^{(1)}, \dots, X^{(n)}$.

In contrast to that, we have

$$
E\left(\left(\left\|1/\boldsymbol{Z}\right\|_D\right)^{-1} \mathbb{1}_{\{\boldsymbol{Z}>0\}}\right) = 0 \iff \inf_{s\in S} Z_s = 0 \text{ a.s.} \tag{4.3}
$$

In particular, this is the case when at least two components η_s, η_t, $s \neq t$, are independent, see the discussion after Definition 2.20.

EXAMPLE 4.5 (Bernoulli model). Consider a standard max-stable random vector $\boldsymbol{\eta} \in \mathbb{R}^d$ with corresponding D-norm $\|\cdot\|_{B_\beta}$, $\beta \in (0,1]$, known from Example 2.14. The generator constant is given by

$$\|\mathbf{1}\|_{B_\beta} = \frac{1-(1-\beta)^d}{\beta}.$$

From the general equality

$$\begin{aligned} E\left(\wr\, \tilde{\boldsymbol{\eta}}\, \wr_\infty\right) &= \int_0^\infty P\left(\min_{i=1,\dots,d} |\tilde{\eta}_i| > t\right)\, \mathrm{d}t \\ &= \int_0^\infty P\left(\tilde{\eta}_i < -t, i = 1,\dots,d\right)\, \mathrm{d}t \\ &= \int_0^\infty \exp\left(-t\,\|\mathbf{1}\|_D\right)\, \mathrm{d}t \\ &= 1/\|\mathbf{1}\|_D, \end{aligned}$$

where $\tilde{\boldsymbol{\eta}}$ is some standard max-stable random vector with D-norm $\|\cdot\|_D$, we conclude

$$\begin{aligned} E\left(\wr\, \boldsymbol{\eta}\, \wr_{B_\beta}\right) &= \sum_{\emptyset \neq T \subseteq \{1,\dots,d\}} \beta^{|T|-1}(1-\beta)^{d-|T|} E\left(\wr\, \boldsymbol{\eta}_T\, \wr_\infty\right) \\ &= \sum_{k=1}^d \binom{d}{k} \beta^k \frac{(1-\beta)^{d-k}}{1-(1-\beta)^k}. \end{aligned}$$

For another example, namely the logistic model, we refer to Example 4.14 below.

REMARK 4.6. (i) Theorem 4.1 implies that the extremal concurrence probability, just like the dual D-norm function, does not depend on the choice of $\boldsymbol{Z}$, but only on $\|\cdot\|_D$.

(ii) In the preceding theorem, we can replace $\boldsymbol{U}, \boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ by a sequence of iid stochastic processes $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ whose univariate marginal df $F_s(x) = P(X_s \leq x)$, $s \in S$, are continuous and strictly monotonically increasing on their support. Condition (2.23) will then have to apply to the copula process $(F_s(X_s))_{s\in S}$. In that case,

$$nP\left(\boldsymbol{X} > \max_{i=1,\dots,n-1} \boldsymbol{X}^{(i)}\right) \to_{n\to\infty} E\left(\left(\|1/\boldsymbol{Z}\|_D\right)^{-1} \mathbb{1}_{\{\boldsymbol{Z}>0\}}\right),$$

where $\boldsymbol{Z}$ is a generator of the D-norm corresponding to the limit SMSP $\boldsymbol{\eta}$ in (2.23). Hence, the probability that there is a champion among $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ does not depend on the univariate margins, but only on the copula process of $\boldsymbol{X}$.

The above remark shows that we do not have to limit our considerations to copula processes. If, for instance, $\boldsymbol{X}$ is an MSP itself with univariate marginal distributions G_s, $s \in S$, then $\boldsymbol{\eta} := (\log(G_s(X_s)))_{s \in S}$ is an SMSP. Applying the max-stability of $\boldsymbol{\eta}$, we obtain

$$\begin{aligned}
\pi_n^c(\boldsymbol{X}) = \pi_n^c(\boldsymbol{\eta}) &= P\left(\boldsymbol{\eta} > \max_{i=1,\dots,n-1} \boldsymbol{\eta}^{(i)}\right) \\
&= P\left((n-1)\boldsymbol{\eta} > \boldsymbol{\eta}^{(1)}\right) \\
&= \int_{\bar{C}^-(S)} P\left((n-1)f > \boldsymbol{\eta}^{(1)}\right) \, (P * \boldsymbol{\eta})(\mathrm{d}f) \\
&= \int_{\bar{C}^-(S)} \exp\left(-(n-1)\left\|f\right\|_D\right) \, (P * \boldsymbol{\eta})(\mathrm{d}f) \\
&= E\left(\exp\left(-(n-1)\left\|\boldsymbol{\eta}\right\|_D\right)\right),
\end{aligned}$$

where $\left\|\cdot\right\|_D$ is the D-norm corresponding to $\boldsymbol{\eta}$, and $\boldsymbol{\eta}^{(1)}, \boldsymbol{\eta}^{(2)}, \dots$ are iid copies of $\boldsymbol{\eta}$.

Having established the functional extremal concurrence probability, we can now derive the limit survival function of a complete record. We will have to restrict to the case where $P(\boldsymbol{Z} > 0) > 0$, which is equivalent to the extremal concurrence probability being positive, cf. (4.3).

Just like before, we consider the copula process case first.

Proposition 4.7. *In addition to the assumptions of Theorem 4.1, suppose that the generator fulfills $P(\boldsymbol{Z} > 0) > 0$. Then, for $f \in \bar{E}^-(S)$,*

$$\begin{aligned}
&P\left(n\left(\boldsymbol{U}^{(n)} - \mathbf{1}\right) > f \,\middle|\, \boldsymbol{U}^{(n)} \text{ is a complete record}\right) \\
&\qquad =: \bar{H}_n(f) \to_{n\to\infty} \bar{H}_D(f) := \frac{E\left(\wr\!\wr \max(\boldsymbol{\eta}, f) \wr\!\wr_D\right)}{E\left(\wr\!\wr \boldsymbol{\eta} \wr\!\wr_D\right)},
\end{aligned}$$

where $\boldsymbol{\eta} = (\eta_s)_{s \in S}$ is an SMSP with corresponding D-norm $\left\|\cdot\right\|_D$.

Note that we avoid division by zero in the preceding formula since we assume $P(\boldsymbol{Z} > 0) > 0$.

Proof of Proposition 4.7. For the ease of notation, we write π_n^c instead of $\pi_n^c(\boldsymbol{U})$. We have

$$\bar{H}_n(f) = \frac{\Pi_n(f)}{\pi_n^c} := \frac{P\left(n(\boldsymbol{U}-1) > f, \boldsymbol{U} > \max_{i=1,\dots,n-1} \boldsymbol{U}^{(i)}\right)}{P\left(\boldsymbol{U} > \max_{i=1,\dots,n-1} \boldsymbol{U}^{(i)}\right)}.$$

By Theorem 4.1, it remains to show that for each $f \in \bar{E}^-(S)$

$$n\Pi_n(f) = nP\left(n(\boldsymbol{U}-1) > \max\left(f, \boldsymbol{M}^{(n)}\right)\right) \to_{n\to\infty} E\left(\wr\!\wr \max(\boldsymbol{\eta}, f) \wr\!\wr_D\right),$$

where $\boldsymbol{M}^{(n)} := n \max_{i=1,\dots,n-1}\left(\boldsymbol{U}^{(i)} - 1\right)$. This can be done by repeating the arguments of the proof of Theorem 4.1. □

Note that by Lemma 4.3, another representation of $\bar{H}_D(f)$ is given by

$$\bar{H}_D(f) = 1 - \frac{E\left(\left(\|1/\boldsymbol{Z}\|_D\right)^{-1} \exp\left(\|1/\boldsymbol{Z}\|_D \sup_{s\in S}(f(s)Z_s)\right) 1\!\!1_{\{\boldsymbol{Z}>0\}}\right)}{E\left(\left(\|1/\boldsymbol{Z}\|_D\right)^{-1} 1\!\!1_{\{\boldsymbol{Z}>0\}}\right)}, \tag{4.4}$$

where $\boldsymbol{Z}$ is a generator of $\|\cdot\|_D$.

Example 4.8. Remember the the Marshall-Olkin D-norm from Example 2.10, i. e.

$$\|\boldsymbol{x}\|_{M_\lambda} = \lambda \|\boldsymbol{x}\|_\infty + (1-\lambda)\|\boldsymbol{x}\|_1, \qquad \boldsymbol{x} \in \mathbb{R}^d,\ \lambda \in (0,1).$$

A generator of $\|\cdot\|_{M_\lambda}$ is given by $\boldsymbol{Z} = X\boldsymbol{1} + (1-X)\tilde{\boldsymbol{Z}}$, where $\tilde{\boldsymbol{Z}}$ is a generator of $\|\cdot\|_1$ (cf. Example 2.9), and X is a Bernoulli distributed random variable with $P(X=1) = \lambda$, independent of $\tilde{\boldsymbol{Z}}$. Obviously, $P(\boldsymbol{Z} > 0, X = 0) = 0$. On the other hand, $X = 1$ implies $\boldsymbol{Z} = 1$. Thus, we obtain by (4.4) for all $\boldsymbol{x} \le \boldsymbol{0}$

$$\begin{aligned}\bar{H}_{M_\lambda}(\boldsymbol{x}) &= 1 - \frac{E\left(\left(\|1/\boldsymbol{Z}\|_{M_\lambda}\right)^{-1} \exp\left(\|1/\boldsymbol{Z}\|_{M_\lambda} \max_{i=1,\dots,d}(x_i Z_i)\right) 1\!\!1_{\{\boldsymbol{Z}>0, X=1\}}\right)}{E\left(\left(\|1/\boldsymbol{Z}\|_{M_\lambda}\right)^{-1} 1\!\!1_{\{\boldsymbol{Z}>0, X=1\}}\right)} \\ &= 1 - \exp\left(\|\boldsymbol{1}\|_{M_\lambda} \max_{i=1,\dots,d} x_i\right).\end{aligned}$$

It is easy to see that $\bar{H}_{M_\lambda}$ is the survival function of the max-stable random vector $(\eta, \dots, \eta)/\|\boldsymbol{1}\|_{M_\lambda}$, where η is standard negative exponentially distributed and $\|\boldsymbol{1}\|_{M_\lambda} =$

$\lambda + d(1-\lambda)$. Clearly, this random vector has complete dependent and identically distributed univariate margins.

In order to generalize Proposition 4.7 to stochastic processes in $C(S)$ with arbitrary margins, the following lemma is needed.

Lemma 4.9. *Let f_n, $n \in \mathbb{N}$, be a sequence of functions in $\bar{E}^-(S)$ converging uniformly to $f \in \bar{E}^-(S)$. Then, under the conditions and notation of Proposition 4.7,*

$$\bar{H}_n(f_n) = \frac{\Pi_n(f_n)}{\pi_n^c} \to_{n\to\infty} \bar{H}_D(f).$$

Proof. Let $\varepsilon > 0$. Due to the uniform convergence of f_n, there exists $N \in \mathbb{N}$ such that $f - \varepsilon \le f_n \le f + \varepsilon$ for $n \ge N$. Assume without loss of generality $f + \varepsilon < 0$, otherwise consider $\min(f+\varepsilon, 0)$. Clearly, for such n,

$$\Pi_n(f+\varepsilon) \le \Pi_n(f_n) \le \Pi_n(f-\varepsilon).$$

Now with $n \to \infty$, Proposition 4.7 shows

$$E\left(\inf_{s\in S} \left|\max\left(\eta_s, f(s)-\varepsilon\right)\right| Z_s\right) \le \lim_{n\to\infty} \Pi_n(f_n) \le E\left(\inf_{s\in S} \left|\max\left(\eta_s, f(s)+\varepsilon\right)\right| Z_s\right).$$

Now check

$$\inf_{s\in S} \left|\max\left(\eta_s, f(s) \pm \varepsilon\right)\right| Z_s \le -\eta_{s_0} Z_{s_0}, \qquad s_0 \in S,$$

and let $\varepsilon \downarrow 0$. The assertion now follows from the dominated convergence theorem. □

We are now ready to generalize Proposition 4.7 to stochastic processes in $C(S)$ with arbitrary univariate margins. Let $\boldsymbol{X} = (X_s)_{s\in S}$ be a process in $C(S)$ whose univariate marginal df $F_s(x) = P(X_s \le x)$, $x \in \mathbb{R}$, $s \in S$, are continuous and strictly monotonically increasing on their support. Let $\boldsymbol{Y}$ be an MSP with univariate marginal df $G_s(x) = P(Y_s \le x)$, $x \in \mathbb{R}$, $s \in S$. We conclude from de Haan and Lin (2001, Theorem 2.8) that $\boldsymbol{X}$ is in the domain of attraction of $\boldsymbol{Y}$ (in the sense of (2.2)) iff the copula process corresponding to $\boldsymbol{X}$, i. e. $\boldsymbol{U} = (U_s)_{s\in S} = (F_s(X_s))_{s\in S}$, is in the domain of attraction of the SMSP $\boldsymbol{\eta} = (\eta_s)_{s\in S} =: (\log(G_s(Y_s)))_{s\in S}$ and the univariate margins fulfill

$$F_s(c_n(s)x + d_n(s))^n \to_{n\to\infty} G_s(x), \qquad x \in \mathbb{R}, \tag{4.5}$$

uniformly for $s \in S$ and locally uniformly for $x \in \mathbb{R}$, where $c_n \in C^+(S)$, $d_n \in C(S)$, $n \in \mathbb{N}$, are the norming functions from (2.2).

Corollary 4.10. *Let $\boldsymbol{Y}$ be an MSP with univariate marginal df G_s, $s \in S$, and $\boldsymbol{X}^{(1)}$, $\boldsymbol{X}^{(2)}, \dots$ be independent copies of a process $\boldsymbol{X} \in \mathcal{D}(\boldsymbol{Y})$ in $C(S)$. Let $c_n \in C^+(S)$, $d_n \in C(S)$, $n \in \mathbb{N}$, be the norming functions from* (2.2)*, and suppose the univariate margins of $\boldsymbol{X}$ satisfy* (4.5)*. Put*

$$\boldsymbol{U} = (U_s)_{s\in S} := (F_s(X_s))_{s\in S}, \quad \boldsymbol{\eta} = (\eta_s)_{s\in S} =: (\log(G_s(Y_s)))_{s\in S},$$

and let $\|\cdot\|_D$ be the D-norm of $\boldsymbol{\eta}$. Choose a generator $\boldsymbol{Z} = (Z_s)_{s\in S}$ of $\|\cdot\|_D$ and suppose that $P(\boldsymbol{Z} > 0) > 0$. Then, for $f \in E(S)$ with $\inf_{s\in S} G_s(f(s)) > 0$,

$$P\left(\frac{\boldsymbol{X}^{(n)} - d_n}{c_n} > f \,\middle|\, \boldsymbol{X}^{(n)} \text{ is a complete record}\right) \to_{n\to\infty} \bar{H}_D(\psi(f))$$

where $\psi(f)(s) = \log\left(G_s(f(s))\right)$, $s \in S$.

Proof. Denote by $\boldsymbol{U}^{(n)}$ the copula process corresponding to $\boldsymbol{X}^{(n)}$, $n \in \mathbb{N}$. Taking logarithms, (4.5) becomes

$$\sup_{s\in S}\left|n\left(F_s(c_n(s)x + d_n(s)) - 1\right) - \log\left(G_s(x)\right)\right| \to_{n\to\infty} 0. \tag{4.6}$$

It can be shown by elementary arguments that (4.6) is equivalent to

$$\begin{aligned}&\sup_{s\in S}\left|\psi_n(f(s)) - \psi(f(s))\right| := \\ &\quad \sup_{s\in S}\left|n\left(F_s(c_n(s)f(s) + d_n(s)) - 1\right) - \log\left(G_s(f(s))\right)\right| \to_{n\to\infty} 0\end{aligned}$$

for each $f \in E(S)$ with $\inf_{s\in S} G_s(f(s)) > 0$. Hence, Lemma 4.9 and the strict monotonicity of F_s entail

$$\begin{aligned}&nP\left(\frac{\boldsymbol{X} - d_n}{c_n} > f, \boldsymbol{X} > \max_{i=1,\dots,n-1} \boldsymbol{X}^{(i)}\right) \\ &= nP\left(n\left(U_s - 1\right) > \psi_n(f(s)),\ s \in S,\ \boldsymbol{U} > \max_{i=1,\dots,n-1} \boldsymbol{U}^{(i)}\right) \\ &\to_{n\to\infty} E\left(\wr \max(\boldsymbol{\eta}, \psi(f)) \wr_D\right).\end{aligned}$$

∎

4.2 Simple records for multivariate observations

Simple record probability

In the preceding section, we have investigated the (normalized) probability of a complete record and in particular, its limit, the extremal concurrence probability. Now we will repeat this procedure, this time for the simple record probability. Unlike in the previous section, where we were actually dealing with the probability of having a champion, normalizing the record probability with the factor n does not yield an interpretation in terms of a probability in the simple record case.

The following result is the equivalent of Theorem 4.1 and Proposition 4.7 in the context of multivariate simple records. Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid random vectors in $\mathbb{R}^d$ with common continuous df F. Recall that $\boldsymbol{X}^{(n)}$ is a simple record, if

$$\boldsymbol{X}^{(n)} \not\leq \max_{1 \leq i \leq n-1} \boldsymbol{X}^{(i)},$$

and $\pi_n^s(\boldsymbol{X})$ denotes the probability of $\boldsymbol{X}^{(n)}$ being a simple record within the iid sequence $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ As before, we initially focus on random vectors $\boldsymbol{U} = (U_1, \dots, U_d)$ that follow a *copula* $C(\boldsymbol{u}) = P(\boldsymbol{U} \leq \boldsymbol{u})$ on $\mathbb{R}^d$, i.e. each univariate margin U_i is uniformly distributed on $(0,1)$, $i = 1, \dots, d$.

Theorem 4.11. *Let $\boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \dots$ be independent copies of a random vector $\boldsymbol{U} \in \mathbb{R}^d$ following a copula C. Suppose that $C \in \mathcal{D}(G)$ with $G(\boldsymbol{x}) = \exp(-\|\boldsymbol{x}\|_D)$, $\boldsymbol{x} \leq \boldsymbol{0} \in \mathbb{R}^d$. Let $\boldsymbol{\eta}$ be a random vector with df G. Then*

$$n\pi_n^s(\boldsymbol{U}) \to_{n\to\infty} E\left(\|\boldsymbol{\eta}\|_D\right),$$

and

$$\begin{aligned} &P\left(n(\boldsymbol{U}^{(n)} - \boldsymbol{1}) \leq \boldsymbol{x} \middle| \boldsymbol{U}^{(n)} \text{ is a simple record}\right) \\ &\to_{n\to\infty} H_D(\boldsymbol{x}) := \frac{E(\|\min(\boldsymbol{x}, \boldsymbol{\eta})\|_D) - \|\boldsymbol{x}\|_D}{E\left(\|\boldsymbol{\eta}\|_D\right)}, \qquad \boldsymbol{x} \leq \boldsymbol{0} \in \mathbb{R}^d. \end{aligned}$$

In the one dimensional case $d = 1$ we obtain $H_D(x) = \exp(x)$, $x \leq 0$. Note, however, that H_D is not a probability df in general. Take, for instance, $\|\cdot\|_D = \|\cdot\|_1$, which is the largest D-norm. In this case the components $\eta_1, \dots, \eta_d$ of $\boldsymbol{\eta}$ are independent and we obtain for $\boldsymbol{x} = (x_1, \dots, x_d) \le \boldsymbol{0} \in \mathbb{R}^d$

$$H_1(\boldsymbol{x}) = \frac{\sum_{i=1}^d \Big(E(|\min(x_i, \eta_i)|) - |x_i|\Big)}{\sum_{i=1}^d E(|\eta_i|)} = \frac{\sum_{i=1}^d \exp(x_i)}{d}.$$

This is not a probability df on $(-\infty, 0]^d$ as, for example, $H_1(\boldsymbol{x})$ does not converge to zero if only one component x_i converges to $-\infty$. Even more, choose $\boldsymbol{a} \le \boldsymbol{b} \le \boldsymbol{0} \in \mathbb{R}^d$. If H_1 would define a probability measure Q on $(-\infty, 0]^d$, then the probability $Q([\boldsymbol{a}, \boldsymbol{b}])$ were given by

$$\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H_1 = \sum_{\boldsymbol{m} \in \{0,1\}^m} (-1)^{d - \sum_{1 \le j \le d} m_j} H_1 \left(b_1^{m_1} a_1^{1-m_1}, \dots, b_d^{m_d} a_d^{1-m_d}\right).$$

But elementary computations show that $\Delta_{\boldsymbol{a}}^{\boldsymbol{b}} H_1 = 0$, i.e., Q is the null measure on $(-\infty, 0]^d$.

Instead one can define Q on $[-\infty, 0]^d \backslash \{-\boldsymbol{\infty}\}$ by putting for $x_i \le 0$ and $i = 1, \dots, d$

$$Q\big(\{-\infty\} \times \dots \times \{-\infty\} \times (-\infty, x_i] \times \{-\infty\} \times \dots \times \{-\infty\}\big) := \frac{\exp(x_i)}{d}.$$

Then Q has its complete mass on the set $\left\{\bigcup_{i=1}^d \left(\{-\infty\}^{i-1} \times (-\infty, 0] \times \{-\infty\}^{d-i}\right)\right\}$ and

$$\begin{aligned}
&Q\left(\times_{i=1}^d [-\infty, x_i] \backslash \{-\boldsymbol{\infty}\}\right) \\
&= Q\left(\bigcup_{i=1}^d \left(\{-\infty\} \times \dots \times \{-\infty\} \times (-\infty, x_i] \times \{-\infty\} \times \dots \times \{-\infty\}\right)\right) \\
&= \sum_{i=1}^d Q\left(\{-\infty\} \times \dots \times \{-\infty\} \times (-\infty, x_i] \times \{-\infty\} \times \dots \times \{-\infty\}\right) \\
&= \frac{1}{d} \sum_{i=1}^d \exp(x_i).
\end{aligned}$$

This approach is closely related to the formulation of the exponent measure theorem as in Balkema and Resnick (1977) and Vatan (1985).

Take, on the other hand, $\|\cdot\|_D = \|\cdot\|_\infty$, which is the least D-norm. In this case, the components $\eta_1, \dots, \eta_d$ of $\boldsymbol{\eta}$ are completely dependent, i.e., $\eta_1 = \eta_2 = \dots = \eta_d$ a.s. and, thus,

$$\begin{aligned}
H_\infty(\boldsymbol{x}) &= E\left(\left\|(\min(x_i,\eta_1))_{i=1}^d\right\|_\infty\right) - \|\boldsymbol{x}\|_\infty\\
&= E\left(\max(\|\boldsymbol{x}\|_\infty, \eta_1)\right) - \|\boldsymbol{x}\|_\infty\\
&= \exp(-\|\boldsymbol{x}\|_\infty), \qquad \boldsymbol{x} = (x_1,\dots,x_d) \le \boldsymbol{0} \in \mathbb{R}^d,
\end{aligned}$$

which is a max-stable distribution.

Proof of Theorem 4.11. Let $\boldsymbol{Z}$ be a generator of $\|\cdot\|_D$, independent of $\boldsymbol{\eta}$. Theorem 4.1, the inclusion-exclusion principle and (2.13) yield

$$\begin{aligned}
n\pi_n^s(\boldsymbol{U}) &= nP\left(\boldsymbol{U} \not\le \max_{i=1,\dots,n-1} \boldsymbol{U}^{(i)}\right)\\
&= nP\left(\bigcup_{j=1}^d \left\{U_j > \max_{i=1,\dots,n-1} U_j^{(i)}\right\}\right)\\
&= \sum_{\emptyset\ne T\subseteq\{1,\dots,d\}} (-1)^{|T|-1} nP\left(U_j > \max_{i=1,\dots,n-1} U_j^{(i)},\ j\in T\right)\\
&\to_{n\to\infty} \sum_{\emptyset\ne T\subseteq\{1,\dots,d\}} (-1)^{|T|-1} E\left(\min_{j\in T} |\eta_j|\, Z_j\right)\\
&= E\left(\max_{j=1,\dots,d} |\eta_j|\, Z_j\right)\\
&= E\left(\|\boldsymbol{\eta}\|_D\right).
\end{aligned}$$

Similarly, one can use Proposition 4.7 in order to show for $\boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d$

$$nP\left(n(\boldsymbol{U}-1) \not\le \min\left(\boldsymbol{x}, \boldsymbol{M}^{(n)}\right)\right) \to_{n\to\infty} E\left(\|\min(\boldsymbol{x},\boldsymbol{\eta})\|_D\right),$$

where $\boldsymbol{M}^{(n)} := n\max_{i=1,\dots,n-1}\left(\boldsymbol{U}^{(n-1)}\right) \to_{\mathcal{D}} \boldsymbol{\eta}$. In summary, taking into account (2.24), we obtain

$$\begin{aligned}
&nP\left(\boldsymbol{U} \le 1 + \frac{\boldsymbol{x}}{n}, \boldsymbol{U} \not\le \max_{i=1,\dots,n-1} \boldsymbol{U}^{(i)}\right)\\
&= nP\left(n(\boldsymbol{U}-1) \not\le \min\left(\boldsymbol{x}, \boldsymbol{M}^{(n)}\right)\right) - nP\left(\boldsymbol{U} \not\le 1 + \frac{\boldsymbol{x}}{n}\right)\\
&\to_{n\to\infty} E\left(\|\min(\boldsymbol{x},\boldsymbol{\eta})\|_D\right) - \|\boldsymbol{x}\|_D.
\end{aligned}$$

□

The arguments in the preceding proof can easily be repeated to extend Theorem 4.11 to the case of a general random vectors $\boldsymbol{X} \in \mathbb{R}^d$, whose df is in the domain of attraction of a max-stable distribution. Denote by

$$C_F(\boldsymbol{u}) := F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right), \qquad \boldsymbol{u} = (u_1, \dots, u_d) \in [0,1]^d,$$

the *copula of a continuous df* F on $\mathbb{R}^d$, where F_i is the i-th univariate marginal df and F_i^{-1} its quantile function.

Corollary 4.12. *Let* $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ *be independent copies of a rv* $\boldsymbol{X} \in \mathbb{R}^d$, *whose df* F *is continuous and its copula* C_F *satisfies* $C_F \in \mathcal{D}(G)$, $G(\boldsymbol{x}) = \exp(-\|\boldsymbol{x}\|_D)$, $\boldsymbol{x} \leq \boldsymbol{0} \in \mathbb{R}^d$. *Assume in addition that each univariate margin* F_i *of* F *is in the domain of attraction of a univariate max-stable distribution* G_i, *i.e., there are constants* $a_{n,i} > 0$, $b_{n,i} \in \mathbb{R}$, $n \in \mathbb{N}$, *such that for* $i = 1, \dots, d$

$$n(1 - F(a_{n,i}x + b_{n,i})) \to_{n\to\infty} -\log(G_i(x)) =: -\psi_i(x), \qquad G_i(x) > 0.$$

Define for all $\boldsymbol{x} = (x_1, \dots, x_d)$ *with* $G_i(x_i) > 0$, $i = 1, \dots, d$,

$$\psi(\boldsymbol{x}) := (\psi_1(x_1), \dots, \psi_d(x_d)).$$

Then we obtain with $\boldsymbol{a}_n := (a_{n,1}, \dots, a_{n,d})$, $\boldsymbol{b}_n := (b_{n,1}, \dots, b_{n,d})$

$$P\left(\frac{\boldsymbol{X}^{(n)} - \boldsymbol{b}_n}{\boldsymbol{a}_n} \leq \boldsymbol{x} \middle| \boldsymbol{X}^{(n)} \text{ is a simple record}\right) \to_{n\to\infty} H_D(\psi(\boldsymbol{x})).$$

Note that in the case $d = 1$

$$H_D(\psi(x)) = \exp(\psi(x)) = G(x), \qquad G(x) > 0.$$

Note, moreover, that the assumptions on the df F in the preceding theorem are equivalent with the condition $F \in \mathcal{D}(G)$, where G is a d-dimensional max-stable df, together with the condition that F is continuous.

Proof of Corollary 4.12. Assume the representation

$$\boldsymbol{X} = \left(F_1^{-1}(U_1), \dots, F_d^{-1}(U_d)\right) =: \boldsymbol{F}^{-1}(\boldsymbol{U}),$$

where $\boldsymbol{U} = (U_1, \dots, U_d)$ follows the copula C of $\boldsymbol{X}$. Repeating the arguments in the proof of Theorem 4.11 now implies the assertion. □

In Corollary 4.2, we have investigated the expected number of complete records as the sample size goes to infinity, which can be done for simple records analogously.

Corollary 4.13. *Denote by* $N_s(n) := \sum_{i=1}^n 1_{\left\{\boldsymbol{X}^{(i)} \not\leq \max_{1\leq j<i} \boldsymbol{X}^{(j)}\right\}}$ *the number of simple records among* $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$*. Then*

$$\frac{E(N_s(n))}{\log n} \to_{n\to\infty} E\left(\|\boldsymbol{\eta}\|_D\right).$$

The next example shows some connection between the Fréchet and the Weibull model, and provides in particular closed formulas for $E\left(\wr\, \boldsymbol{\eta} \,\wr_\lambda\right)$ and $E\left(\|\boldsymbol{\eta}\|_\lambda\right)$.

Example 4.14 (Fréchet model). Choose $\lambda > 1$. Let $\boldsymbol{\eta}$ be a max-stable random vector in $\mathbb{R}^d$ with df $P(\boldsymbol{\eta} \leq \boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_\lambda\right)$, $\boldsymbol{x} \leq 0$. Let $\boldsymbol{Z}_F$ a Fréchet-based generator of $\|\cdot\|_\lambda$ (see Example 2.11 and Example 2.22), and $\boldsymbol{Z}_W$ a Weibull-based generator of $\|\cdot\|_{W_\lambda}$ (see Example 2.12 and Example 2.23). We know from Dombry et al. (2015, Example 1) that

$$E\left(\wr\, \boldsymbol{\eta} \,\wr_\lambda\right) = \frac{\Gamma(d-1/\lambda)}{(d-1)!\Gamma(1-1/\lambda)} = \prod_{i=1}^{d-1}\left(1 - \frac{1}{\lambda i}\right).$$

In the Fréchet model, a nice analogy between the formulas of the complete and simple record case occurs. We will prove in the following

$$E\left(\|\boldsymbol{\eta}\|_\lambda\right) = \frac{\Gamma(d+1/\lambda)}{(d-1)!\Gamma(1+1/\lambda)} = \prod_{i=1}^{d-1}\left(1 + \frac{1}{\lambda i}\right),$$

where the last equality can easily be shown by induction. To this end, we will show below

$$E\left(\|\boldsymbol{Z}_W\|_\lambda\right) = \frac{\Gamma(d+1/\lambda)}{(d-1)!\Gamma(1+1/\lambda)}. \tag{4.7}$$

Furthermore, (2.21) together with Lemma 4.3 yields

$$E\left(\wr\, \boldsymbol{Z}_F \,\wr_{W_\lambda}\right) = E\left(\|1/\boldsymbol{Z}_F\|_\lambda^{-1}\right) = E\left(\wr\, \boldsymbol{\eta} \,\wr_\lambda\right).$$

On the other hand, it is easy to see that $(\Gamma(1-1/\lambda)\Gamma(1+1/\lambda)\boldsymbol{Z}_W)^{-1}$ is also a generator of $\|\cdot\|_\lambda$, which yields in turn

$$E\left(\wr\, \boldsymbol{Z}_F \wr_{W_\lambda}\right) = E\left(\wr\, \boldsymbol{Z}_W \wr_\lambda\right).$$

Altogether, we obtain

$$E\left(\wr\, \boldsymbol{\eta} \wr_\lambda\right) = E\left(\wr\, \boldsymbol{Z}_W \wr_\lambda\right),$$

and hence by (2.13)

$$E\left(\|\boldsymbol{\eta}\|_\lambda\right) = E\left(\|\boldsymbol{Z}_W\|_\lambda\right) = \prod_{i=1}^{d-1}\left(1+\frac{1}{\lambda i}\right).$$

It remains to show (4.7). Let $\boldsymbol{Z}_W = (W_1,\dots,W_d)/\Gamma(1+1/\lambda)$, where $W_1,\dots,W_d$ are iid Weibull with $P(W_i > x) = \exp\left(-x^\lambda\right)$, $x > 0$. Then

$$E\left(\wr\, \boldsymbol{Z}_F \wr_{W_\lambda}\right) = \frac{1}{\Gamma(1+1/\lambda)} E\left(\left(\sum_{i=1}^{d} W_i^\lambda\right)^{1/\lambda}\right)$$

It is easy to see that W_i^λ is standard exponentially distributed, and by convolution, the sum $S := \sum_{i=1}^d W_i^\lambda$ follows the Gamma$(d,1)$-distribution, i. e. the density of S is given by

$$h(x;d,1) = \frac{x^{d-1}e^{-x}}{(d-1)!}, \qquad x > 0.$$

Hence,

$$E\left(S^{1/\lambda}\right) = \int_0^\infty \frac{x^{d-1+1/\lambda}e^{-x}}{(d-1)!}\,\mathrm{d}x = \frac{\Gamma(d+1/\lambda)}{(d-1)!},$$

which proofs (4.7).

Simple record times

So far we have investigated records within a sequence of iid random vectors, that is, the probability that the n-th observation is a record, and the conditional distribution of a record, given that there is one at hand. In this section, however, we will focus on the random indices at which a record occurs, the so-called *record times*. Generally speaking, records within an iid sequence are fairly uncommon, which makes the record times a random sequence of integers that grows very fast. This fact has already been touched on in the Corollaries 4.2 and 4.13. In fact, if $X, X^{(1)}, X^{(2)}, \dots$ are continuously distributed iid random variables on the real line, we have seen before that $\pi_n^s(X) = \pi_n^c(X) = 1/n$, resulting in the fact that we can only expect $\sum_{n=1}^{1000} 1/n \approx 7.49$ records within the first 1000 observations. Obviously, complete records are even more uncommon: If $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ are continuously distributed iid bivariate random vectors with independent margins, we have $\pi_n^c(\boldsymbol{X}) = 1/n^2$, and the expected number of complete records among the first 1000 observations is $\sum_{n=1}^{1000} 1/n^2 \approx 1.64$. Bearing in mind that the first observation is always a records by definition, this leaves us with only 0.64 non-trivial expected complete records. However, simple records are more frequent, depending on the dimension of the random vectors. Taking bivariate vectors $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ as above, it is easy to show that $\pi_n^s(\boldsymbol{X}) = 2/n - 1/n^2$, which yields an expectation of $\sum_{n=1}^{1000}(2/n - 1/n^2) \approx 13.33$ simple records within the first 1000 observations.

Denote by F the bivariate df of $\boldsymbol{X}$ and assume $F \in \mathcal{D}(G)$ for some max-stable df G. It has been shown in Goldie and Resnick (1989, Theorem 5.3), that the total number of *complete* records among $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ is either finite almost surely or infinite almost surely, depending on whether or not G has independent margins. In this section, we will prove a similar statement concerning the expectation of the record times. With complete records being so rare, we will focus on *simple* record times from now on, since most of the results we will derive will turn out to be trivial in the complete record case. It is well known that the record times have infinite expectation for a sequence of univariate iid random variables with a common continuous df, see e. g. the discussion in Arnold et al. (1998, Section 2.5) or Galambos (1987, Theorem 6.2.1). This is no longer true in the multivariate simple record case. Later, we give a precise characterization of when simple record times are integrable.

Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid random vectors in $\mathbb{R}^d$. Throughout the whole section, unless expressly stated otherwise, the df F of $\boldsymbol{X}$ is assumed to be continuous. We denote by $L(n)$, $n \geq 1$, the *simple record times*, i.e., those subsequent random indices at which a simple record occurs. Precisely, $L(1) = 1$, as $\boldsymbol{X}^{(1)}$ is by definition a simple record, and, for $n \geq 2$,

$$L(n) := \min\left\{ j : j > L(n-1), \boldsymbol{X}^{(j)} \not\leq \max_{1 \leq i \leq L(n-1)} \boldsymbol{X}^{(i)} \right\}.$$

The n-th simple record among $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ is obviously given by $\boldsymbol{X}^{(L(n))}$. Clearly, in contrast to the original observations, the sequence of simple records $\boldsymbol{X}^{(L(1))}, \boldsymbol{X}^{(L(2))}, \dots$ is *not* iid. (Note that in the univariate case, records actually do occur independently as mentioned in the beginning of Chapter 4.) However, the following result, even though being quite obvious, should be noticed. Exceptionally, it is not necessary to assume continuity of the df F.

Lemma 4.15. *Let $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid random vectors in $\mathbb{R}^d$ with df F.*

(i) $P\left(\boldsymbol{X}^{(L(n)+m)} \le \boldsymbol{x}\right) = F(\boldsymbol{x})$ *for every $n, m \in \mathbb{N}$ and $\boldsymbol{x} \in \mathbb{R}^d$.*

(ii) $\boldsymbol{X}^{(L(n))}$ *and* $\boldsymbol{X}^{(L(n)+m)}$ *are independent for $n, m \in \mathbb{N}$.*

Proof. Partitioning the sample space into disjoint events, we obtain

$$\begin{aligned}
&P\left(\boldsymbol{X}^{(L(n)+1)} \le \boldsymbol{x}\right)\\
&= \sum_{k_2=2}^{\infty}\sum_{k_3=1}^{\infty}\cdots\sum_{k_n=1}^{\infty} P\left(\boldsymbol{X}^{(\sum_{i=2}^n k_i+1)} \le \boldsymbol{x}, L(n) = \sum_{i=2}^n k_i, \dots, L(3) = k_2 + k_3, L(2) = k_2\right)\\
&= F(\boldsymbol{x})\sum_{k_2=2}^{\infty}\sum_{k_3=1}^{\infty}\cdots\sum_{k_n=1}^{\infty} P\left(L(n) = \sum_{i=2}^n k_i, \dots, L(3) = k_2 + k_3, L(2) = k_2\right)\\
&= F(\boldsymbol{x}),
\end{aligned}$$

since the observations are iid, and the events $\{L(n) = \sum_{i=2}^n k_i\}, \dots, \{L(2) = k_2\}$ only depend on $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(\sum_{i=2}^n k_i)}$. This shows (i). Similar arguments show

$$P\left(\boldsymbol{X}^{(L(n)+m)} \le \boldsymbol{x}, \boldsymbol{X}^{(L(n))} \le \boldsymbol{y}\right) = F(\boldsymbol{x})P\left(\boldsymbol{X}^{(L(n))} \le \boldsymbol{y}\right),$$

and hence (ii). □

Denote by $\boldsymbol{M}^{(n)} := \max_{i=1,\dots,n} \boldsymbol{X}^{(i)}$, $n \in \mathbb{N}$, the sequence of componentwise maxima within $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$. Then the independence of the observations implies that $\boldsymbol{M}^{(n)}$, $n \in \mathbb{N}$, is a homogenous Markov chain (cf. Resnick (2007, Section 5.6) and Resnick (2008, Section 4.1)), and it is easy to see that the transition is given by

$$P\left(\boldsymbol{M}^{(n+1)} \le \boldsymbol{x} \middle| \boldsymbol{M}^{(n)} = \boldsymbol{y}\right) = \begin{cases} F(\boldsymbol{x}), & \text{if } \boldsymbol{x} \ge \boldsymbol{y},\\ 0, & \text{else.}\end{cases} \tag{4.8}$$

The observations which generate the jumps in this Markov chain are precisely the simple records. Now put $\Delta_1 := 1$ and define by

$$\Delta_n := L(n) - L(n-1), \qquad n = 2, 3, \dots$$

the *interrecord waiting times.*

Lemma 4.16. *Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid random vectors in $\mathbb{R}^d$ following a continuous df F.*

(i) For every $n \in \mathbb{N}$,

$$P\left(\boldsymbol{X}^{(L(n+1))} \leq \boldsymbol{x} \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{y}\right) = P\left(\boldsymbol{X} \leq \boldsymbol{x} \middle| \boldsymbol{X} \not\leq \boldsymbol{y}\right).$$

(ii) For every $n \in \mathbb{N}$,

$$\Delta_{n+1} \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{x} \sim \mathrm{Geom}(1 - F(\boldsymbol{x})),$$

where $\mathrm{Geom}(p)$ *denotes the geometric distribution with support* $\{1, 2, \dots\}$ *and parameter* $p \in (0, 1]$.

REMARK 4.17. The above Lemma could just as easily be formulated for complete instead of simple records. Instead of $\{\boldsymbol{X} \not\leq \boldsymbol{y}\}$ in (i), we would have to condition on $\{\boldsymbol{X} > \boldsymbol{y}\}$ in the complete record case. Note that the n-th complete record $\boldsymbol{X}^{(L_c(n))}$ coincides with the maximum $\boldsymbol{M}^{(L_c(n))}$ at the n-th complete record time $L_c(n)$. Analogously, in (ii), the parameter of the geometric distribution would change from $1 - F(\boldsymbol{x})$ to $P(\boldsymbol{X} > \boldsymbol{x})$.

Proof of Lemma 4.16. We have due to Lemma 4.15

$$\begin{aligned}
&P\left(\boldsymbol{X}^{(L(n+1))} \leq \boldsymbol{x} \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{y}\right) \\
&= \sum_{k=1}^{\infty} P\left(\boldsymbol{X}^{(L(n+1))} \leq \boldsymbol{x}, \Delta_{n+1} = k \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{y}\right) \\
&= \sum_{k=1}^{\infty} P\Big(\boldsymbol{X}^{(L(n)+k)} \leq \boldsymbol{x}, \boldsymbol{X}^{(L(n)+1)} \leq \boldsymbol{y}, \dots \\
&\qquad\qquad \dots, \boldsymbol{X}^{(L(n)+k-1)} \leq \boldsymbol{y}, \boldsymbol{X}^{(L(n)+k)} \not\leq \boldsymbol{y} \Big| \boldsymbol{M}^{(L(n))} = \boldsymbol{y}\Big) \\
&= \sum_{k=1}^{\infty} P\left(\boldsymbol{X}_{(L(n)+k)} \leq \boldsymbol{x}, \boldsymbol{X}_{(L(n)+k)} \not\leq \boldsymbol{y}\right) F(\boldsymbol{y})^{k-1} \\
&= \frac{P(\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{X} \not\leq \boldsymbol{y})}{1 - F(\boldsymbol{y})} \\
&= P(\boldsymbol{X} \leq \boldsymbol{x} | \boldsymbol{X} \not\leq \boldsymbol{y}),
\end{aligned}$$

which shows (i). By similar arguments, we obtain

$$\begin{aligned}
&P\left(L(n+1) - L(n) = k \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{x}\right) \\
&= P\left(\boldsymbol{X}^{(L(n)+1)} \leq \boldsymbol{x}, \dots, \boldsymbol{X}^{(L(n)+k-1)} \leq \boldsymbol{x}, \boldsymbol{X}^{(L(n)+k)} \not\leq \boldsymbol{x} \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{x}\right) \\
&= F(\boldsymbol{x})^{k-1}(1 - F(\boldsymbol{x})),
\end{aligned}$$

which is (ii). □

Computing the distribution of the second record $\boldsymbol{X}^{(L(2))}$ is now an easy task: Conditioning on $\boldsymbol{X}^{(1)} = \boldsymbol{y}$ we obtain

$$\begin{aligned} P\left(\boldsymbol{X}^{(L(2))} \le \boldsymbol{x}\right) &= \int_{\mathbb{R}^d} P\left(\boldsymbol{X}^{(L(2))} \le \boldsymbol{x} \middle| \boldsymbol{X}^{(1)} = \boldsymbol{y}\right) \, F(\mathrm{d}\boldsymbol{y}) \\ &= \int_{\mathbb{R}^d} P(\boldsymbol{X} \le \boldsymbol{x} \big| \boldsymbol{X} \not\le \boldsymbol{y}) \, F(\mathrm{d}\boldsymbol{y}). \end{aligned}$$

The distribution of $\boldsymbol{X}^{(L(n))}$ for an arbitrary $n \ge 2$ is much more complex, but generally manageable. Computation of the limit distribution of $\boldsymbol{X}^{(N(n))}$, properly linearly standardized, as n tends to infinity, is an open problem. For the univariate case we refer to Galambos (1987, Section 6.4) or Resnick (2008, Section 4.2).

From now on, we will focus on simple record times and interrecord waiting times, whereas the sequence of simple records itself will take a back seat. As the df F is continuous, the distribution of $L(n)$ does not depend on F and, therefore, we assume in what follows without loss of generality that F is a copula C on $\mathbb{R}^d$, i. e. each component of $\boldsymbol{X}$ is uniformly distributed on $(0, 1)$.

Conditioning on $\boldsymbol{X}^{(1)} = \boldsymbol{u}$ yields for $j \ge 2$

$$\begin{aligned} P(L(2) = j) &= P\left(\boldsymbol{X}^{(2)} \le \boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(j-1)} \le \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(j)} \not\le \boldsymbol{X}^{(1)}\right) \\ &= \int_{[0,1]^d} C(\boldsymbol{u})^{j-2}(1 - C(\boldsymbol{u})) \, C(\mathrm{d}\boldsymbol{u}). \end{aligned}$$

Solving the geometric series, we get

$$E(L(2)) = \sum_{j=2}^{\infty} j P(L(2) = j) = \int_{[0,1]^d} \frac{1}{1 - C(\boldsymbol{u})} \, C(\mathrm{d}\boldsymbol{u}) + 1. \tag{4.9}$$

Now we generalize this formula. Choose $n \in \mathbb{N}$. Partitioning the sample space in disjoint events, we obtain for $k_n \ge 1$

$$\begin{aligned} &P\left(L(n+1) - L(n) = k_n\right) = \\ &= \sum_{k_1=2}^{\infty} \sum_{k_2=1}^{\infty} \cdots \sum_{k_{n-1}=1}^{\infty} P\left(L(n+1) = \sum_{j=1}^{n} k_j, L(n) = \sum_{j=1}^{n-1} k_j, \dots, L(2) = k_1\right), \end{aligned}$$

and further, similar to the calculation above,

$$P\left(L(n+1)=\sum_{j=1}^{n}k_j, L(n)=\sum_{j=1}^{n-1}k_j,\dots,L(2)=k_1\right)=$$
$$\int_{\{\boldsymbol{u}_n\not\leq\cdots\not\leq\boldsymbol{u}_1\}} C(\boldsymbol{u}_1)^{k_1-2}C(\boldsymbol{u}_2)^{k_2-1}\cdots C(\boldsymbol{u}_n)^{k_n-1}(1-C(\boldsymbol{u}_n))\ C(\mathrm{d}\boldsymbol{u}_1)\cdots C(\mathrm{d}\boldsymbol{u}_n).$$

Hence, solving all the occurring geometric series yields

$$E\left(L(n+1)-L(n)\right)=\int_{\{\boldsymbol{u}_n\not\leq\cdots\not\leq\boldsymbol{u}_1\}}\prod_{i=1}^{n}\frac{1}{1-C(\boldsymbol{u}_i)}\ C(\mathrm{d}\boldsymbol{u}_1)\cdots C(\mathrm{d}\boldsymbol{u}_n).$$

We summarize this result.

Lemma 4.18. *Let $\boldsymbol{X}^{(1)},\boldsymbol{X}^{(2)},\dots$ be iid random vectors following a copula C on $[0,1]^d$. For every $n\in\mathbb{N}$,*

$$E\left(\Delta_{n+1}\right)=\int_{\{\boldsymbol{u}_n\not\leq\cdots\not\leq\boldsymbol{u}_1\}}\prod_{i=1}^{n}\frac{1}{1-C(\boldsymbol{u}_i)}\ C(\mathrm{d}\boldsymbol{u}_1)\cdots C(\mathrm{d}\boldsymbol{u}_n). \tag{4.10}$$

Suppose now that $d=1$. Then we have $\boldsymbol{u}=u\in[0,1]$, $C(u)=u$ and

$$E(L(2))=\int_0^1\frac{1}{1-u}\,du+1=\infty,$$

which is well-known (Galambos (1987, Theorem 6.2.1)). Suppose next that $d\geq 2$ and that the margins of C are independent, i.e.,

$$C(\boldsymbol{u})=\prod_{i=1}^{d}u_i,\qquad \boldsymbol{u}=(u_1,\dots,u_d)\in[0,1]^d.$$

Then we obtain

$$\int_{[0,1]^d}\frac{1}{1-C(\boldsymbol{u})}\,C(\mathrm{d}\boldsymbol{u})=\int_0^1\dots\int_0^1\frac{1}{1-\prod_{i=1}^{d}u_i}\,\mathrm{d}u_1\dots\mathrm{d}u_d<\infty$$

by elementary arguments and, thus, $E(L(2))<\infty$. This observation gives rise to the problem of characterizing those copulas C on $[0,1]^d$ with $d\geq 2$, such that $E(L(2))$ is finite. Note that $E(L(2))=\infty$ if the components of C are completely dependent. The next lemma characterizes finiteness of $E(L(2))$. Note that in the complete record case, the second record time is never integrable, which is trivial since complete records are even less likely than univariate records.

Lemma 4.19. *Let $\boldsymbol{X} = (X_1, \dots, X_d)$ follow a copula C on $\mathbb{R}^d$. Then $E(L(2)) < \infty$ iff*

$$\int_1^\infty P\left(X_i > 1 - \frac{1}{t},\ 1 \le i \le d\right) \mathrm{d}t < \infty. \tag{4.11}$$

Condition (4.11) is trivially satisfied in case of independent components $X_1, \dots, X_d$ and $d \ge 2$. Below we will see that it is, roughly, in general satisfied, if there are at least two components that are asymptotically independent.

Proof of Lemma 4.19. Any copula C satisfies the Fréchet-Hoeffding bounds, that is, for $\boldsymbol{u} = (u_1, \dots, u_d) \in [0,1]^d$,

$$\max\left(1 - d + \sum_{i=1}^d u_i, 0\right) \le C(\boldsymbol{u}) \le \min\left(u_1, \dots, u_d\right). \tag{4.12}$$

Therefore, we obtain due to the upper bound in (4.12)

$$\begin{aligned} E(L(2)) - 1 &= \int_{[0,1]^d} \frac{1}{1 - C(\boldsymbol{u})}\, C(\mathrm{d}\boldsymbol{u}) \\ &= E\left(\frac{1}{1 - C(\boldsymbol{X})}\right) \\ &= \int_1^\infty P\left(C(\boldsymbol{X}) > 1 - \frac{1}{t}\right) \mathrm{d}t \\ &\le \int_1^\infty P\left(X_i > 1 - \frac{1}{t}, 1 \le i \le d\right) \mathrm{d}t. \end{aligned}$$

On the other hand, the lower bound in (4.12) yields

$$\begin{aligned} E(L(2)) - 1 &\ge \int_1^\infty P\left(\sum_{i=1}^d (1 - X_i) < \frac{1}{t}\right) \mathrm{d}t \\ &\ge \int_1^\infty P\left(1 - X_i < \frac{1}{dt}, 1 \le i \le d\right) \mathrm{d}t \\ &= d \int_{1/d}^\infty P\left(1 - X_i < \frac{1}{t}, 1 \le i \le d\right) \mathrm{d}t. \end{aligned}$$

□

Let C be a copula that is in the domain of attraction of a (standard) max-stable df G on $\mathbb{R}^d$, i. e.

$$C^n\left(1+\frac{\boldsymbol{x}}{n}\right) \to_{n\to\infty} G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right), \qquad \boldsymbol{x} \le \boldsymbol{0} \in \mathbb{R}^d,$$

which we abbreviate as before by $C \in \mathcal{D}(G)$. In order to prove the following result, we will benefit from Proposition 2.29, which dealt with convergence to the D-norm and the dual D-norm function, in case the copula is in the domain of attraction of a standard max-stable df.

Proposition 4.20. *Suppose that $C \in \mathcal{D}(G)$, where the D-norm corresponding to G satisfies $\wr \boldsymbol{1} \wr_D > 0$. Then $E(L(2)) = \infty$.*

Proof. Let $\boldsymbol{X}$ be a random vector with df C. It is well known from real analyis that

$$\int_1^\infty P\left(\boldsymbol{X} > 1 - \frac{1}{t}\right) \mathrm{d}t < \infty \iff \sum_{n=1}^\infty P\left(\boldsymbol{X} > 1 - \frac{1}{n}\right) < \infty.$$

From (2.25), we know $nP\left(\boldsymbol{X} > 1 - \frac{1}{n}\right) \to_{n\to\infty} \wr \boldsymbol{1} \wr_D > 0$. Now applying the limit comparison test for an infinite series, we deduce that $\sum_{n=1}^\infty P\left(\boldsymbol{X} > 1 - \frac{1}{n}\right)$ has the same limit behavior as the harmonic series $\sum_{i=1}^\infty \frac{1}{n}$, and hence, $E(L(2)) = \infty$ by Lemma 4.19. □

Suppose that $C \in \mathcal{D}(G)$. Due to Proposition 4.20, a finite expectation $E(L(2)) < \infty$ can only occur if $\wr \boldsymbol{1} \wr_D = 0$, which is true, for instance, if G has at least two independent margins.

Let $\boldsymbol{X}$ follow the df C. Next we show that $E(L(2))$ is typically finite if $\boldsymbol{X}$ has at least two components X_j, X_k which are *tail independent*, i. e.

$$\lim_{u\uparrow 1} P(X_k > u | X_j > u) = 0.$$

In case the limit exists, define the dependence measure

$$\bar{\chi} := \lim_{u\uparrow 1} \frac{2\log(1-u)}{\log(P(X_1 > u, X_2 > u))} - 1 \in [-1, 1],$$

where (X_1, X_2) follows some bivariate copula, cf. Coles et al. (1999) or Heffernan (2000). Note that we have $\bar{\chi} = 1$ if X_1, X_2 are *tail dependent* (meaning that they are *not* tail independent). In the class of (bivariate) copulas that are tail independent, however, $\bar{\chi}$ is a popular measure of tail comparison. For a bivariate normal copula with coefficient of correlation $\rho \in (-1,1)$ it is, for instance, well known that $\bar{\chi} = \rho$.

Proposition 4.21. *Let $\boldsymbol{X} = (X_1, \dots, X_d)$ follow a copula C in $\mathbb{R}^d$. Suppose that there exist indices $k \neq j$ such that*

$$\bar{\chi}_{k,j} = \lim_{u \uparrow 1} \frac{2\log(1-u)}{\log(P(X_k > u, X_j > u))} - 1 \in [-1, 1).$$

Then we have $E(L(2)) < \infty$.

Corollary 4.22. *We have $E(L(2)) < \infty$ for multivariate normal random vectors unless all components are completely dependent.*

Proof of Proposition 4.21. We have

$$\begin{aligned}
&\int_1^\infty P\left(X_i \geq 1 - \frac{1}{t},\, 1 \leq i \leq d\right) \mathrm{d}t \\
&\leq \int_1^\infty P\left(X_k \geq 1 - \frac{1}{t},\, X_j \geq 1 - \frac{1}{t}\right) \mathrm{d}t \\
&= \int_1^\infty \exp\left(\frac{\log\left(P\left(X_k \geq 1 - \frac{1}{t},\, X_j \geq 1 - \frac{1}{t}\right)\right)}{\log\left(\frac{1}{t^2}\right)} \log\left(\frac{1}{t^2}\right)\right) \mathrm{d}t,
\end{aligned}$$

where

$$\frac{\log\left(P\left(X_k \geq 1 - \frac{1}{t},\, X_j \geq 1 - \frac{1}{t}\right)\right)}{\log\left(\frac{1}{t^2}\right)} \to_{t \to \infty} \frac{1}{1 + \bar{\chi}} > \frac{1}{2}.$$

But this implies that the above integral is finite and, thus, the assertion is a consequence of Lemma 4.19. □

To complete this section, we investigate $E(L(n))$ for $n \geq 2$. Recall $\Delta_{n+1} = L(n+1) - L(n)$, $n \in \mathbb{N}$. Again, let $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be an iid sequence of random vectors on $\mathbb{R}^d$ following a copula C. Obviously, $E(L(2)) = \infty$ implies $E(L(n)) = \infty$ for $n \geq 2$, since $L(n) \geq L(2)$, $n \geq 2$. On the other hand, if $E(L(2)) < \infty$, we obtain due to (4.10) for all $n \geq 2$

$$\begin{aligned}
E\left(\Delta_{n+1}\right) &= E\left(\frac{1}{1 - C(\boldsymbol{X}^{(1)})} \cdots \frac{1}{1 - C(\boldsymbol{X}^{(n)})} \mathbb{1}_{\{\boldsymbol{X}^{(n)} \not\leq \dots \not\leq \boldsymbol{X}^{(1)}\}}\right) \\
&\leq E\left(\frac{1}{1 - C(\boldsymbol{X}^{(1)})} \cdots \frac{1}{1 - C(\boldsymbol{X}^{(n)})}\right) \\
&= \left[E\left(\frac{1}{1 - C(\boldsymbol{X}^{(1)})}\right)\right]^n
\end{aligned}$$

$$= [E(L(2)) - 1]^n,$$

which means that $E(\Delta_{n+1}) < \infty$ as well in that case. Furthermore, we will show below that $E(\Delta_{n+1}) = \infty$ for all $n \in \mathbb{N}$ if $E(L(2)) = \infty$. In conclusion, it is sufficient to decide whether $E(L(2))$ is finite or not if expectations of arbitrary simple record times are investigated. We summarize this discussion.

Proposition 4.23. *Let $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be iid random vectors following a copula C on $[0,1]^d$. Then the following implications hold:*

(i) If $E(L(2)) = \infty$, then $E(\Delta_n) = \infty$ for all $n = 2, 3, \dots$

(ii) If $E(L(2)) < \infty$, then $E(\Delta_n) < \infty$ for all $n \in \mathbb{N}$.

Proof. It remains to proof (i). We show that Δ_{n+1}, $n \in \mathbb{N}$, is *stochastically increasing*, i. e.

$$P(\Delta_{n+1} \leq t) \geq P(\Delta_{n+2} \leq t), \qquad t \in \mathbb{R},\ n \in \mathbb{N}. \tag{4.13}$$

Recall that the df of a random variable $X \sim \mathrm{Geom}(p)$ is given by $P(X \leq t) = 1-(1-p)^{\lfloor t \rfloor}$, where $\lfloor t \rfloor = \max\{m \in \mathbb{Z} : m \leq t\}$. Conditioning on $\boldsymbol{M}^{(L(n))} = \boldsymbol{x}$, we obtain by Lemma 4.16 (ii) for each $t \in \mathbb{R}$ and $n \in \mathbb{N}$

$$\begin{aligned}
&P(\Delta_{n+1} \leq t) \\
&= \int_{[0,1]^d} P\left(\Delta_{n+1} \leq t \middle| \boldsymbol{M}^{(L(n))} = \boldsymbol{x}\right) \left(P * \boldsymbol{M}^{(L(n))}\right)(\mathrm{d}\boldsymbol{x}) \\
&= \int_{[0,1]^d} 1 - C(\boldsymbol{x})^{\lfloor t \rfloor} \left(P * \boldsymbol{M}^{(L(n))}\right)(\mathrm{d}\boldsymbol{x}) \\
&= 1 - E\left(C\left(\boldsymbol{M}^{(L(n))}\right)^{\lfloor t \rfloor}\right),
\end{aligned}$$

which shows (4.13) since $\boldsymbol{M}^{(L(n))} \leq \boldsymbol{M}^{(L(n+1))}$. Hence,

$$E(\Delta_{n+1}) = \int_0^\infty P(\Delta_{n+1} > t)\ \mathrm{d}t \leq \int_0^\infty P(\Delta_{n+2} > t)\ \mathrm{d}t = E(\Delta_{n+2}).$$

□

4.3 Partial records and hitting scenarios

Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be a sequence of iid random vectors on $\mathbb{R}^d$, and denote by F the df of $\boldsymbol{X}$. As before, we assume throughout this section that F is continuous in order to avoid ties within the sequence. It has already been remarked that the sequence of componentwise maxima $\boldsymbol{M}^{(n)} = \max_{i=1,\dots,n} \boldsymbol{X}^{(i)}$, $n \in \mathbb{N}$, is a homogenous Markov chain (cf. (4.8)) which jumps at time $n \in \mathbb{N}$ iff $\boldsymbol{X}^{(n)}$ is a simple record. In this section, we will focus on the genesis and the development of $\boldsymbol{M}^{(n)}$ as n increases. In particular, we will specify the term of a simple record in order to state exactly on which subset T of $\{1, \dots, d\}$ a simple record exceeds the previous componentwise maximum. This will lead to the notion of *T-records*.

In the univariate setup, the structure of the Markov chain $M^{(n)}$, $n \in \mathbb{N}$, is rather simple. There is a jump at time n iff $X^{(n)}$ is a record, and in that case, the new maximum $M^{(n)}$ coincides with that record. In the multivariate world, however, both the composition of $\boldsymbol{M}^{(n)}$ and the evolution of $\boldsymbol{M}^{(n)}$ as n increases are much more complex, since there is possibly more than one observation that contributes to the componentwise maximum. More precisely, in $\mathbb{R}^d$, up to d different observations among $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ might contribute to $\boldsymbol{M}^{(n)}$. A precise description on the number of contributing random vectors and their impact can be given by the so-called *sample hitting scenario* (cf. Wang and Stoev (2011), Dombry et al. (2013), Dombry and Éyi-Minko (2013) and Dombry et al. (2015)), which we will investigate below.

T-records

It is obvious that the sequence $\boldsymbol{M}^{(n)}$, $n \in \mathbb{N}$, jumps at time n iff $\boldsymbol{X}^{(n)}$ is a simple record. However, the term of a simple record does not explain on which subset $T \subseteq \{1, \dots, d\}$ the observation $\boldsymbol{X}^{(n)}$ exceeds $\boldsymbol{M}^{(n-1)}$, which is why we will introduce a more precise definition. Remember the notation

$$\boldsymbol{a}_T = (a_{i_1}, \dots, a_{i_k}) \in \mathbb{R}^k$$

for a vector $\boldsymbol{a} = (a_1, \dots, a_d)$ and a nonempty subset $T = \{i_1, \dots, i_k\} \subseteq \{1, \dots, d\}$ with $i_1 < \dots < i_k$. We define for $\emptyset \neq T \subsetneq \{1, \dots, d\}$

$$\boldsymbol{X}^{(n)} \text{ is a } T\text{-}record \ :\Longleftrightarrow \ \ \boldsymbol{X}_T^{(n)} > \boldsymbol{M}_T^{(n-1)} \text{ and } \boldsymbol{X}_{T^\complement}^{(n)} \leq \boldsymbol{M}_{T^\complement}^{(n-1)}.$$

In case $T = \{1, \dots, d\}$, a T-record is meant to be a complete record. For the ease of notation, we introduce the convention of T-records with $T = \emptyset$: $\boldsymbol{X}^{(n)}$ is a $\emptyset$-record means that $\boldsymbol{X}^{(n)}$ is not a simple record. By $\pi_{n,T} = \pi_{n,T}(\boldsymbol{X})$ we denote the probability of $\boldsymbol{X}^{(n)}$ being a T-record. Our first aim is to investigate the limit behavior of $\pi_{n,T}$ as n tends to infinity.

Let us now assume that F is in the max-domain of attraction of a max-stable df G. It is clear that $\pi_{n,T}$ does not depend on the univariate margins of F, but only on its dependence structure, that is, on the copula of F. Thus, without loss of generality, G can be assumed to be standard max-stable, i. e. $G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right)$, $\boldsymbol{x} \leq \boldsymbol{0} \in \mathbb{R}^d$, for some D-norm

$$\|\boldsymbol{x}\|_D = E\left(\max_{i=1,\dots,d} |x_i|\, Z_i\right), \qquad \boldsymbol{x} \in \mathbb{R}^d.$$

Let $\boldsymbol{\eta}$ be a random vector following the standard max-stable df G. In Theorem 4.1 and Theorem 4.11 we have shown

$$n\pi_{n,\{1,\dots,d\}} \to_{n\to\infty} E\left(\wr\!\wr\, \boldsymbol{\eta} \,\wr\!\wr_D\right),$$

where $\wr\!\wr\, \boldsymbol{x} \,\wr\!\wr_D = E\left(\min_{i=1,\dots,d} |x_i|\, Z_i\right)$, $\boldsymbol{x} \in \mathbb{R}^d$, is the dual D-norm function, and

$$n\left(1 - \pi_{n,\emptyset}\right) \to_{n\to\infty} E\left(\|\boldsymbol{\eta}\|_D\right).$$

Now we derive an analogous result for T-records in general. To this end, the following technical lemma is helpful.

Lemma 4.24. *Choose $\emptyset \neq T \subsetneq \{1,\dots,d\}$ and let $(a_1,\dots,a_d)$ be an arbitrary vector in $\mathbb{R}^d$. Then*

$$\sum_{\emptyset \neq S \subseteq T^{\complement}} (-1)^{|S|-1} \min_{k\in T\cup S} a_k = \min\left(\min_{i\in T} a_i, \max_{j\in T^{\complement}} a_j\right).$$

Proof. Put $b_k := \min\left(\min_{i\in T} a_i, a_k\right)$, $k \in T^{\complement}$. Then clearly $\min_{k\in S} b_k = \min_{k\in T\cup S} a_k$ for $S \subseteq T^{\complement}$. Hence, we have by (2.13)

$$\begin{aligned}
\sum_{\emptyset \neq S \subseteq T^{\complement}} (-1)^{|S|-1} \min_{k\in T\cup S} a_k &= \sum_{\emptyset \neq S \subseteq T^{\complement}} (-1)^{|S|-1} \min_{k\in S} b_k \\
&= \max_{k\in T^{\complement}} b_k \\
&= \max_{k\in T^{\complement}} \min\left(\min_{i\in T} a_i, a_k\right) \\
&= \min\left(\min_{i\in T} a_i, \max_{k\in T^{\complement}} a_k\right).
\end{aligned}$$

□

In the following, we write with an abuse of notation $\wr\!\wr\, \boldsymbol{x}_T \,\wr\!\wr_D = E\left(\min_{i\in T} |x_i|\, Z_i\right)$ (and analogously $\|\boldsymbol{x}_T\|_D = E\left(\max_{i\in T} |x_i|\, Z_i\right)$), even though $\|\cdot\|_D$ and $\wr\!\wr \cdot \wr\!\wr_D$ are actually functions on $\mathbb{R}^d$ and not on $\mathbb{R}^{|T|}$.

Proposition 4.25. *Assume the underlying df F satisfies $F \in \mathcal{D}(G)$ for some standard max-stable df $G(\boldsymbol{x}) = \exp\left(-\|\boldsymbol{x}\|_D\right)$, $\boldsymbol{x} \le \mathbf{0} \in \mathbb{R}^d$. Choose $\emptyset \ne T \subseteq \{1,\dots,d\}$. Then*

$$n\pi_{n,T} \to_{n\to\infty} E\left(\max\left(\wr\, \boldsymbol{\eta}_T \,\wr_D - \left\|\boldsymbol{\eta}_{T^\complement}\right\|_D, 0\right)\right),$$

where $\boldsymbol{\eta}$ follows the df G and $\|\boldsymbol{\eta}_\emptyset\|_D$ is interpreted to be zero.

Proof. Define, for $\emptyset \ne T \subseteq \{1,\dots,d\}$,

$$\tilde{\pi}_{n,T} := P\left(\boldsymbol{X}_T^{(n)} > \max_{i=1,\dots,n-1} \boldsymbol{X}_T^{(i)}\right).$$

This corresponds to the probability that $\boldsymbol{X}_T^{(n)}$ is a complete record within $\boldsymbol{X}_T^{(1)},\dots,\boldsymbol{X}_T^{(n)}$. By Theorem 4.1, we obtain

$$n\tilde{\pi}_{n,T} \to_{n\to\infty} E\left(\wr\, \boldsymbol{\eta}_T \,\wr_D\right).$$

Thus, the inclusion exclusion principle yields

$$\begin{aligned}
n\pi_{n,T} &= n\left(\tilde{\pi}_{n,T} - P\left(\boldsymbol{X}_T^{(n)} > \max_{i=1,\dots,n-1} \boldsymbol{X}_T^{(i)}, \boldsymbol{X}_{T^\complement}^{(i)} \not\le \max_{i=1,\dots,n-1} \boldsymbol{X}_{T^\complement}^{(i)}\right)\right)\\
&= n\left(\tilde{\pi}_{n,T} - P\left(\bigcup_{j\in T^\complement}\left\{\boldsymbol{X}_{T\cup\{j\}} > \max_{i=1,\dots,n-1} \boldsymbol{X}_{T\cup\{j\}}^{(i)}\right\}\right)\right)\\
&= n\left(\tilde{\pi}_{n,T} - \sum_{\emptyset\ne S\subseteq T^\complement} (-1)^{|S|-1} P\left(\bigcap_{j\in T^\complement}\left\{\boldsymbol{X}_{T\cup\{j\}} > \max_{i=1,\dots,n-1} \boldsymbol{X}_{T\cup\{j\}}^{(i)}\right\}\right)\right)\\
&= n\left(\tilde{\pi}_{n,T} - \sum_{\emptyset\ne S\subseteq T^\complement} (-1)^{|S|-1} P\left(\boldsymbol{X}_{T\cup S} > \max_{i=1,\dots,n-1} \boldsymbol{X}_{T\cup S}^{(i)}\right)\right)\\
&\to_{n\to\infty} E\left(\wr\, \boldsymbol{\eta}_T \,\wr_D\right) - \sum_{\emptyset\ne S\subseteq T^\complement} (-1)^{|S|-1} E\left(\wr\, \boldsymbol{\eta}_{T\cup S} \,\wr_D\right)\\
&= E\left(\max\left(\wr\, \boldsymbol{\eta}_T \,\wr_D - \left\|\boldsymbol{\eta}_{T^\complement}\right\|_D, 0\right)\right),
\end{aligned}$$

where the last equality follows from Lemma 4.24. □

Hitting scenarios

A precise description of the composition of the componentwise maximum $\boldsymbol{M}^{(n)}$ can be given in terms of hitting scenarios. Take $\boldsymbol{x}^{(1)},\dots,\boldsymbol{x}^{(n)} \in \mathbb{R}^d$ and put $\boldsymbol{x} = \max_{i=1,\dots,n} \boldsymbol{x}^{(i)}$.

Define the (possibly empty) sets

$$C_i := \left\{ j : x_j = x_j^{(i)} \right\} \subseteq \{1, \ldots, d\}, \qquad i = 1, \ldots, n,$$

that consist of those indices at which $\boldsymbol{x}^{(i)}$ is contributing to $\boldsymbol{x}$. Assuming that the maximum $\boldsymbol{x}$ is attained uniquely, the set $\tau := \{C_i : C_i \neq \emptyset, i = 1, \ldots, \ell\}$ defines a partition of the index set $\{1, \ldots, d\}$ which we call the *hitting scenario* of $\max_{i=1,\ldots,n} \boldsymbol{x}^{(i)}$ or of $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}$, respectively.

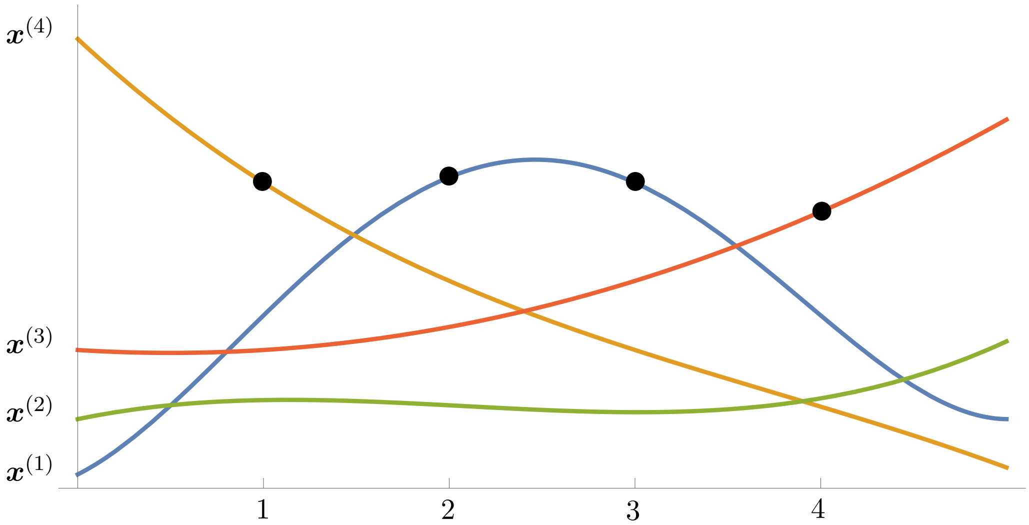

Fig. 4.1: *The functions $\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(4)}$ are observed at four sites, resulting in fourdimensional vectors $(x_1^{(i)}, \ldots, x_4^{(i)})$, $i = 1, \ldots, 4$. The hitting scenario τ is given by $C_1 = \{2, 3\}$, $C_2 = \emptyset$, $C_3 = \{4\}$, $C_4 = \{1\}$. Hence, $\tau = \{\{1\}, \{2, 3\}, \{4\}\}$.*

In Chapter 2, we have introduced point measures and point processes. We can also define hitting scenarios of point measures on $\mathbb{E} = [0, \infty]^d \setminus \{\mathbf{0}\}$, which will be an important space in the following, since max-stable random vectors with unit Fréchet margins can be expressed via Poisson point processes on $\mathbb{E}$, see Remark 2.18. Let $m = \sum_{k \in \mathbb{N}} \epsilon_{\boldsymbol{x}^{(k)}} \in M_p(\mathbb{E})$ be such a point measure. Choose $\varepsilon > 0$ such that $\sup_{k \in \mathbb{N}} x_j^{(k)} > \varepsilon$ for all $j = 1, \ldots, d$. Since bounded sets in $\mathbb{E}$ are those that are bounded away from $\mathbf{0}$, the set $K_\varepsilon := \mathbb{E} \setminus [0, \varepsilon)^d$ is bounded in $\mathbb{E}$, and hence, only finitely many points $\boldsymbol{x}^{(k_1)}, \ldots, \boldsymbol{x}^{(k_N)}$ fall in the set K_ε. By construction, $\sup_{k \in \mathbb{N}} \boldsymbol{x}^{(k)} = \max_{i=1,\ldots,N} \boldsymbol{x}^{(k_i)}$. Assuming this maximum is attained uniquely, its hitting scenario is defined to be the *hitting scenario of m* or of $\left(\boldsymbol{x}^{(k)}\right)_{k \in \mathbb{N}}$, respectively.

So far, hitting scenarios have only been introduced for deterministic sequences and point measures. Now let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be an iid sequence with continuous df F. Since there are no ties in that sequence almost surely, the random hitting scenario Π_n of the componentwise maximum $\boldsymbol{M}^{(n)}$ is well-defined with probability one. Following Dombry et al. (2015), this random partition is called the *sample hitting scenario*. Note that $\boldsymbol{X}^{(n)}$ being a T-record implies $T \in \Pi_n$. Again, the distribution of the hitting scenario does not depend on the univariate margins, but only on the copula of the underlying continuous df F. Thus, if we have two sequences $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ with continuous df F_1 and $\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}, \dots$ with continuous df F_2, such that F_1 and F_2 have the same copula, the sample hitting scenarios associated to $\boldsymbol{X}^{(1)}, \dots, \boldsymbol{X}^{(n)}$ and $\boldsymbol{Y}^{(1)}, \dots, \boldsymbol{Y}^{(n)}$ follow the same distribution for each $n \in \mathbb{N}$.

Different to $\boldsymbol{M}^{(n)}$, $n \in \mathbb{N}$, the sequence Π_n, $n \in \mathbb{N}$, does not define a Markov chain. However, the joint sequence $\left(\boldsymbol{M}^{(n)}, \Pi_n\right)_{n\in\mathbb{N}}$ does, since the observations $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ are independent. In the following, we will discuss the transition of this Markov chain.

Denote by $\mathbb{P}_d$ the power set of $\{1, \dots, d\}$ and by $\mathcal{P}_d$ the set of all partitions of $\{1, \dots, d\}$. By definition, there is exactly one $T \in \mathbb{P}_d$ (possibly $T = \emptyset$) such that $\boldsymbol{X}^{(n)}$ is a T-record. Knowing this set T and the n-th sample hitting scenario Π_n, one can easily deduce the $(n+1)$-th sample hitting scenario via

$$\Pi_n = \tau = \{S_1, \dots, S_\ell\} \in \mathcal{P}_d \Longrightarrow \Pi_{n+1} = \tau_T := \{S_1 \cap T^{\complement}, \dots, S_\ell \cap T^{\complement}, T\} \setminus \{\emptyset\} \in \mathcal{P}_d.$$

Now define for a fixed $\tau \in \mathcal{P}_d$ the mapping

$$\Theta_\tau : \mathbb{P}_d \to \mathcal{P}_d, \qquad \Theta_\tau(T) = \tau_T. \tag{4.14}$$

Clearly, this mapping is not one-to-one. For instance, if $\tau = \{\{1,2\}\}$, then $\tau_{\{1\}} = \tau_{\{2\}} = \{\{1\},\{2\}\}$. Hence, given that $\Pi_n = \{\{1,2\}\}$, either $\boldsymbol{X}^{(n)}$ being a $\{1\}$-record or a $\{2\}$-record results in the new sample hitting scenario $\Pi_{n+1} = \{\{1\},\{2\}\}$. More generally, not only a T-record, but every S-record, with S being an element of the preimage $\Theta_\tau^{-1}(\{\tau_T\})$ leads to the new sample hitting scenario τ_T. Hence, we have for each $\tau \in \mathcal{P}_d$ and $T \in \mathbb{P}_d$ due to the iid property of the observations

$$\begin{aligned} &P\left(\Pi_{n+1} = \tau_T \middle| \Pi_n = \tau, \boldsymbol{M}^{(n)} = \boldsymbol{y}\right) \\ &\quad = P\left(\bigcup_{S\in\Theta_\tau^{-1}(\{\tau_T\})} \left\{\boldsymbol{X}^{(n+1)} \text{ is an } S\text{-record}\right\} \middle| \Pi_n = \tau, \boldsymbol{M}^{(n)} = \boldsymbol{y}\right) \end{aligned}$$

$$= \sum_{S \in \Theta_\tau^{-1}(\{\tau_T\})} P\left(\boldsymbol{X}_S > \boldsymbol{y}_S, \boldsymbol{X}_{S^\complement} < \boldsymbol{y}_{S^\complement}\right)$$
$$= \sum_{S \in \Theta_\tau^{-1}(\{\tau_T\})} P\left(\boldsymbol{X}^{(n+1)} \text{ is an } S\text{-record} \middle| \boldsymbol{M}^{(n)} = \boldsymbol{y}\right).$$

We distinguish between three cases in order to visualize the transition from the n-th to the $(n+1)$-th sample hitting scenario. Let in the following $\Pi_n = \tau = \{S_1, \ldots, S_\ell\}$.

Case 1: Nothing changes. It might happen that the sample hitting scenario does not change at all, i. e. $\Pi_{n+1} = \tau = \{S_1, \ldots, S_\ell\}$. This is the case when $\boldsymbol{X}^{(n+1)}$ is either a $\emptyset$-record or a S_j-record for some $j = 1, \ldots, \ell$. It is easy to see that $\tau_\emptyset = \tau_{S_1} = \cdots = \tau_{S_\ell} = \tau$.

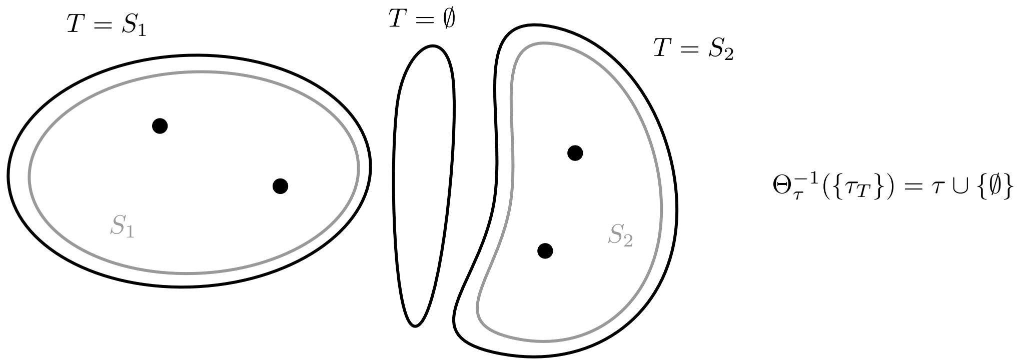

Fig. 4.2: *The sample hitting scenario does not change. The dots mark elements of the sets.*

Case 2: Cut. In that case, $\boldsymbol{X}^{(n+1)}$ is a T-record, where T 'cuts' some set $S_j \in \tau$ into two pieces, without intersecting a second set $S_k \in \tau$, $k \neq j$. For instance, if $\tau = \{\{1,2\}, \{3,4\}\}$, then either a $\{1\}$- or a $\{2\}$-record result in the new sample hitting scenario $\tau_{\{1\}} = \tau_{\{2\}} = \{\{1\}, \{2\}, \{3,4\}\}$.

Case 3: Cut & Paste. The third case works as follows: One can choose at least two different sets $S_{j_1}, \ldots, S_{j_k} \in \tau$, take as many elements (but at least one) from each of these sets, and paste them together in a new set T. In that case, $\tau_T = \tau_{T'}$ implies $T = T'$. For example, choose $\tau = \{\{1,2\}, \{3,4\}, \{5\}\}$. Then only a $\{2,3,5\}$-record will result in the new sample hitting scenario $\tau_{\{2,3,5\}} = \{\{1\}, \{2,3,5\}, \{4\}\}$.

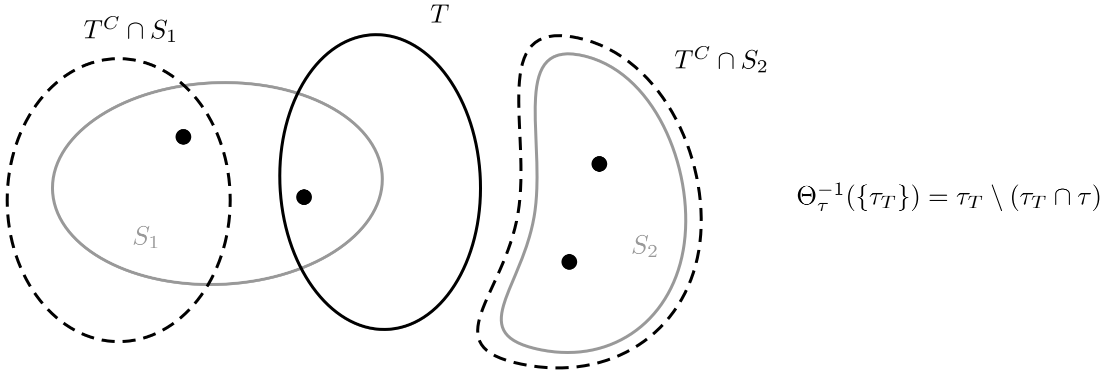

Fig. 4.3: *Every set in the new sample hitting scenario is a subset of an element of the preceding sample hitting scenario. The dots mark elements of the sets.*

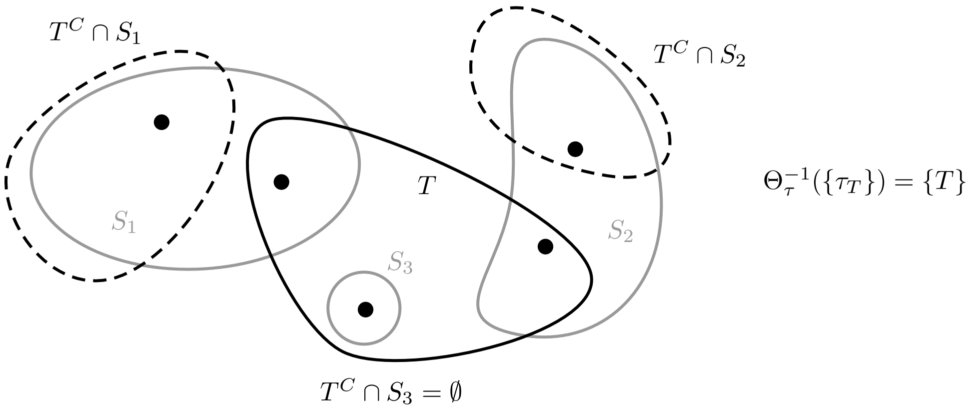

Fig. 4.4: *Figure of the Cut & Paste case. Only the new T-record can generate the new sample hitting scenario τ_T. The dots mark elements of the sets.*

Alternatively, let $\mathcal{T}_n \in \mathbb{P}_d$ the (possibly empty) random set such that $\boldsymbol{X}^{(n)}$ is a $\mathcal{T}_n$-record. Then

$$\begin{aligned} P\left(\Pi_{n+1} = \tau_T, \mathcal{T}_{n+1} = T \middle| \Pi_n = \tau, \boldsymbol{M}^{(n)} = \boldsymbol{y}\right) &= P\left(\boldsymbol{X}_T > \boldsymbol{y}_T, \boldsymbol{X}_{T^\complement} < \boldsymbol{y}_{T^\complement}\right) \\ &= P\left(\boldsymbol{X}_{n+1}\ T\text{-record} \middle| \boldsymbol{M}^{(n)} = \boldsymbol{y}\right). \end{aligned}$$

We summarize these results.

Proposition 4.26. *Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \dots$ be an iid sequence with continuous df.*

(i) *The sequence* $\left(\Pi_n, \boldsymbol{M}^{(n)}\right)_{n\in\mathbb{N}}$ *is a homogenous Markov chain on* $\mathcal{P}_d \times \mathbb{R}^d$ *with transition*

$$P\left(\Pi_{n+1} = \tau' \middle| \Pi_n = \tau, \boldsymbol{M}^{(n)} = \boldsymbol{y}\right)$$
$$= \begin{cases} \sum_{S\in\Theta_\tau^{-1}(\{\tau_T\})} P\left(\boldsymbol{X}_S > \boldsymbol{y}_S, \boldsymbol{X}_{S^\complement} < \boldsymbol{y}_{S^\complement}\right), & \text{if } \tau' = \tau_T \text{ for some } T \in \mathbb{P}_d, \\ 0, & \text{else.} \end{cases}$$

(ii) *The sequence* $\left(\Pi_n, \mathcal{T}_n, \boldsymbol{M}^{(n)}\right)_{n\in\mathbb{N}}$ *is a homogenous Markov chain on* $\mathcal{P}_d \times \mathbb{P}_d \times \mathbb{R}^d$ *with transition*

$$P\left(\Pi_{n+1} = \tau', \mathcal{T}_{n+1} = T \middle| \Pi_n = \tau, \boldsymbol{M}^{(n)} = \boldsymbol{y}\right)$$
$$= \begin{cases} P\left(\boldsymbol{X}_T > \boldsymbol{y}_T, \boldsymbol{X}_{T^\complement} < \boldsymbol{y}_{T^\complement}\right), & \text{if } \tau' = \tau_T \text{ for some } T \in \mathbb{P}_d, \\ 0, & \text{else.} \end{cases}$$

In general, the unconditional probability $P(\Pi_n = \tau)$ for some $\tau \in \mathcal{P}_d$ can not be calculated explicitly. However, in case the observations follow - or at least are in the domain of attraction - of a max-stable distribution themselves, a nice connection to the so-called extremal hitting scenario that has been developed by Wang and Stoev (2011) occurs (see also Dombry et al. (2013), Dombry and Éyi-Minko (2013) and Dombry et al. (2015)).

From now on, it will be more convenient to consider *simple* max-stable instead of *standard* max-stable random vectors. Note that this is just a matter of transforming the univariate margins, and does not affect the distribution of a hitting scenario. Remember the PPP representation

$$\boldsymbol{\xi} =_{\mathcal{D}} \sup_{k\in\mathbb{N}} \boldsymbol{\vartheta}^{(k)}, \tag{2.18'}$$

of a simple max-stable random vector $\boldsymbol{\xi}$ in $\mathbb{R}^d$, where $\left(\boldsymbol{\vartheta}^{(k)}\right)_{k\in\mathbb{N}}$ are the points of PPP(ν) on $\mathbb{E} = [0,\infty]^d \setminus \{\boldsymbol{0}\}$, and ν denotes the exponent measure, cf. Remark 2.18. The hitting scenario of $(\left(\boldsymbol{\vartheta}^{(k)}\right)_{k\in\mathbb{N}}$ is called the *extremal hitting scenario* of $\boldsymbol{\xi}$.

The following Lemma is Theorem 1 in Dombry et al. (2015), and establishes a connection between the sample and the extremal hitting scenario in the domain of attraction case.

Lemma 4.27. *Let* $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots$ *be an iid sequence following a continuous df* $F \in \mathcal{D}(H)$ *for some max-stable df* H. *Denote by* $H_1, \ldots, H_d$ *the univariate margins of* H. *Then the sample hitting scenario of* $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)}$ *converges weakly to the extremal hitting* *scenario of $\boldsymbol{\xi}$ which follows the simple max-stable df G defined by*

$$G(\boldsymbol{x}) := H\left(H_1^{-1}\left(e^{-x_1^{-1}}\right), \ldots, H_d^{-1}\left(e^{-x_d^{-1}}\right)\right), \qquad \boldsymbol{x} = (x_1, \ldots, x_d) > 0.$$

Remark 4.28. The preceding Lemma reflects again the convergence of the *champion probability* to something we called the *extremal concurrence probability* in Section 4.1 for a good reason. Let $\boldsymbol{X}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots \in \mathbb{R}^d$ be in the domain of attraction of a simple max-stable random vector $\boldsymbol{\xi}$. Denoting by Π_n the sample hitting scenario of $\max\left(\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)}\right)$ and by Π the extremal hitting scenario of $\boldsymbol{\xi}$, Lemma 4.27 implies in particular

$$P\left(|\Pi_n| = 1\right) \to_{n\to\infty} P\left(|\Pi| = 1\right).$$

But clearly, the left-hand side is exactly the probability $p_n(\boldsymbol{X})$ that there is a champion among $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(n)}$. Comparing this result with Theorem 4.1 yields that

$$E\left(\wr -1/\boldsymbol{\xi} \wr_D\right) = P\left(|\Pi| = 1\right),$$

i. e. the limit champion probability is indeed the probability that only one of the points of the PPP $\left(\boldsymbol{\vartheta}^{(k)}\right)_{k\in\mathbb{N}}$ actually contributes to the supremum in (2.18′), which explains the term 'extremal concurrence probability'.

Next we derive a connection between the sample and the extremal hitting scenario in case the observations are max-stable themselves. It is again sufficient to consider the simple max-stable case. The following Proposition extends Lemma 2 from Dombry et al. (2015). It was first derived by Stephenson and Tawn (2005) with a completely different and more heuristic proof based on likelihood analysis. A partition $\tau' \in \mathcal{P}_d$ is said to be finer than $\tau \in \mathcal{P}_d$ if every block of τ' is included in some block of τ. We write $\tau' \preceq \tau$ then.

Proposition 4.29. *Let Π_n be the sample hitting scenario of iid copies $\boldsymbol{\xi}^{(1)}, \ldots, \boldsymbol{\xi}^{(n)}$ of a simple max-stable random vector $\boldsymbol{\xi} \in \mathbb{R}^d$, and Π be the extremal hitting scenario of $\boldsymbol{\xi}$. For all $\tau \in \mathcal{P}_d$,*

$$P(\Pi_n = \tau) = \sum_{\tau' \preceq \tau} \ell! \binom{n}{\ell} n^{-\ell'} P(\Pi = \tau')$$

with ℓ and ℓ' being the sizes of τ and τ', respectively.

Proof. The proof in Dombry et al. (2015) can be adapted so as to obtain

$$P(\Pi_n = \tau) = \sum_{\tau' \preceq \tau} c_n(\tau', \tau) P(\Pi = \tau')$$

where $c_n(\tau', \tau)$ is the probability to recover τ from τ' by coloring independently the blocks of τ' with n colors and merging the blocks with the same color. This probability is computed easily. For $\tau = (S_1, \ldots, S_\ell)$, the blocks of $\tau' \preceq \tau$ can be labeled such that

$$\tau' = \{S_{1,1}, \ldots, S_{1,k_1}, S_{2,1}, \ldots, S_{2,k_2}, \ldots, S_{\ell,1}, \ldots, S_{\ell,k_\ell}\},$$

where $S_j = \bigcup_{k=1,\ldots,k_j} S_{j,k}$, $j = 1, \ldots, \ell$. There are $n(n-1)\ldots(n-\ell+1)$ choices of different colors for the blocks of τ and then, for each block S_j, the subblocks $S_{j,k}$, $k = 1, \ldots, k_j$ must be colored with the assigned color which happens with probability n^{-k_j}. Hence we have

$$c_n(\tau', \tau) = n(n-1)\ldots(n-\ell+1)\prod_{j=1}^{\ell} n^{-k_j} = \ell!\binom{n}{\ell} n^{-\ell'}.$$

□

Limit theory

In this section, we closely follow the discussion in Resnick (2008, Section 4.4) or Resnick (2007, Chapter 7). Consider an iid sequence $\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}, \ldots \in \mathbb{R}^d$ of continuously distributed random vectors. It is well known that these random vectors are in the max-domain of attraction of a simple max-stable random vector $\boldsymbol{\xi}$, i e.

$$n^{-1}\boldsymbol{M}^{(n)} = n^{-1} \max_{i=1,\ldots,n} \boldsymbol{X}^{(i)} \to_{\mathcal{D}} \boldsymbol{\xi} \tag{4.15}$$

iff

$$Y_n(\cdot) := n^{-1}\boldsymbol{M}^{([n\cdot])} \to_{\mathcal{D}} \sup_{k\in\mathbb{N}} \boldsymbol{j}^{(k)} \mathbb{1}_{\{t_k \le \cdot\}} =: Y(\cdot) \tag{4.16}$$

in the Skorokhod space $D\left([0,\infty), \mathbb{R}^d\right)$, where $[x]$ denotes the integer part of x, and the time-space point process

$$N := \sum_{k\in\mathbb{N}} \epsilon_{\left(t_k, \boldsymbol{j}^{(k)}\right)}$$

is PPP($\lambda \times \nu$) on $[0,\infty) \times \mathbb{E}$, with λ being the Lebesgue measure and ν being the exponent measure on $\mathbb{E} = [0,\infty]^d \setminus \{\mathbf{0}\}$, see e. g. Resnick (2007, Proposition 7.2). The limit process $\boldsymbol{Y}$ is called a multivariate *extremal process*. The equivalence of (4.15) and (4.16) can be shown by proving first that (4.15) is equivalent to

$$N_n := \sum_{k=1}^{n} \epsilon_{\left(k/n, \boldsymbol{X}^{(k)}/n\right)} \to_{\mathcal{D}} N, \tag{4.17}$$

where the convergence is meant in the space $M_p\left([0,\infty)\times\mathbb{E}\right)$ of point measures on $[0,\infty)\times\mathbb{E}$ equipped with the vague metric, see Resnick (2008, Section 3) for details. This is very similar to the equivalence of (i) and (iii) in Proposition 2.27. Applying the continuous mapping theorem then yields the assertion. We adapt this strategy to prove a similar result for hitting scenarios.

Note that it is also possible to define hitting scenarios of $\sup_{k\in\mathbb{N}} \boldsymbol{x}^{(k)} \mathbb{1}_{\{t_k\le t\}}$ for $t>0$, where $\left(t_k, \boldsymbol{x}^{(k)}\right)$, $k\in\mathbb{N}$, are the points of a point measure on $[0,\infty)\times\mathbb{E}$. Again, the supremum is actually a maximum due to the topological structure of $\mathbb{E}$, and we have to assume that it is reached uniquely.

Since we will work on the space $D\left((0,\infty),\mathcal{P}_d\right)$ below, a quick repetition on Skorokhod spaces will be useful. Excellent reviews on Skorokhod spaces can be found in Resnick (2008, Section 4.4.1) or Billingsley (1999, Sections 12 and 16). Let $(\mathcal{X},\delta)$ be a complete and separable metric space and choose $a,b\in\mathbb{R}$, $a<b$. We start with the space $D([a,b],\mathcal{X})$ of right-continuous functions on $[a,b)$ with finite left limits on $(a,b]$. Define the set of time deformations

$$\Lambda := \big\{\lambda : [a,b]\to[a,b] :\ \lambda(a)=a, \lambda(b)=b, \\ \lambda \text{ continuous and strictly increasing}\big\}.$$

Let $\mathrm{id}\in\Lambda$ be the identity transformation, i.e. $\mathrm{id}(t)=t$, $t\in[a,b]$. The *Skorokhod distance* of $f,g\in D([a,b],\mathcal{X})$ is

$$d_{a,b}(f,g) = \inf_{\lambda\in\Lambda} \max\left(\|\lambda-\mathrm{id}\|_\infty , \sup_{t\in[a,b]} \delta(f(t), g(\lambda(t))) \right).$$

Given a sequence $f_n, f\in D([a,b],\mathcal{X})$, $n\in\mathbb{N}$, we have $d_{a,b}(f_n,f)\to 0$ iff there exist $\lambda_n\in\Lambda$, $n\in\mathbb{N}$, such that

$$\|\lambda_n-\mathrm{id}\|_\infty \to 0, \qquad \sup_{t\in[a,b]} \delta(f_n(\lambda_n(t)), f(t)) \to 0.$$

It is possible to extend the definition of the Skorokhod metric to the space $D((0,\infty),\mathcal{X})$ of right-continuous functions $f:(0,\infty)\to\mathcal{X}$ with left limits. However, in order to show

$f_n \to f$ in $D((0,\infty),\mathcal{X})$, it is sufficient to show that for all $0 < a < b$ such that f is continuous in a, b, the restrictions of f_n to the interval $[a,b]$ converge to those of f in $D([a,b],\mathcal{X})$, see Resnick (2008, Lemma 4.16).

In the next Lemma, we will have to deal with a converging sequence of point measures on the locally compact space $[0,\infty) \times \mathbb{E}$. The suitable notion of convergence is *vague convergence*, introduced in Remark 2.26.

Lemma 4.30. *Let* $m = \sum_{k\in\mathbb{N}} \epsilon_{(t_k,\boldsymbol{x}_k)}$ *be a point measure on* $[0,\infty) \times \mathbb{E}$ *such that* $\sup_{k\in\mathbb{N}} \boldsymbol{x}_k \mathbb{1}_{\{t_k\le t\}}$ *is attained uniquely for all* $t > 0$. *Denote by* $\tau(t)$ *the hitting scenario of* $\sup_{k\in\mathbb{N}} \boldsymbol{x}_k \mathbb{1}_{\{t_k\le t\}}$. *The mapping*

$$\Psi : M_p\left([0,\infty) \times \mathbb{E}\right) \to D\left((0,\infty),\mathcal{P}_d\right), \quad \Psi(m)(t) = \tau(t), \quad t \in (0,\infty),$$

is continuous if m *satisfies* $m\left([0,a] \times \mathbb{E}\right) > 0$ *and* $m\left(\{b\} \times \mathbb{E}\right) = 0$ *for arbitrary* $a, b \in (0,\infty)$.

Proof. The proof is very similar to the arguments in Resnick (2008, p.214). Let $m, m_1, m_2, \ldots \in M_p\left([0,\infty) \times \mathbb{E}\right)$, $m = \sum_{k\in\mathbb{N}} \epsilon_{(t_k,\boldsymbol{x}_k)}$, $m_n = \sum_{k\in\mathbb{N}} \epsilon_{\left(t_k^{(n)},\boldsymbol{x}_k^{(n)}\right)}$, such that m_n converges vaguely to m satisfying the assumptions above. Denote by $\pi_n(t)$ and $\pi(t)$ the hitting scenarios of $\sup_{t_k^{(n)}\le t} \boldsymbol{x}_k^{(n)}$ and $\sup_{t_k\le t} \boldsymbol{x}_k$, respectively. We have to show that $\Psi(m_n) = \pi_n(\cdot) \to \pi(\cdot) = \Psi(m)$ in $D\left([a,b],\mathcal{P}_d\right)$. Choose $\varepsilon > 0$ such that

$$\sup_{t_k\le a} x_{k,j} > \varepsilon, \qquad j = 1,\ldots,d, \tag{4.18}$$

$$x_{k,j} \neq \varepsilon, \qquad j = 1,\ldots,d,\ k \in \mathbb{N}, \tag{4.19}$$

and put $K_\varepsilon := \mathbb{E} \setminus [0,\varepsilon)^d$. The set $[0,b] \times K_\varepsilon$ is compact in $[0,\infty) \times \mathbb{E}$. Furthermore, (4.19) and the assumption $m\left(\{b\} \times \mathbb{E}\right) = 0$ imply $m\left(\partial\left([0,b] \times K_\varepsilon\right)\right) = 0$. Therefore, we have for n large enough

$$m_n\left([0,b] \times K_\varepsilon\right) = m\left([0,b] \times K_\varepsilon\right) = N < \infty.$$

By (4.18), we have $N > 0$. Repeating this argument, we obtain for n large enough

$$m_n\left([0,a] \times K_\varepsilon\right) = m\left([0,a] \times K_\varepsilon\right) = M < N.$$

with $M > 0$. The points that fall in $[0,b] \times K_\varepsilon$ can be relabeled such that

$$0 < t_1^{(n)} < \cdots < t_M^{(n)} < a < t_{M+1}^{(n)} < \cdots < t_N^{(n)} < b,$$

$$0 < t_1 < \cdots < t_M < a < t_{M+1} < \cdots < t_N < b,$$

and

$$\left(t_i^{(n)}, \boldsymbol{x}_i^{(n)}\right) \to_{n\to\infty} (t_i, \boldsymbol{x}_i), \qquad i = 1, \ldots, N, \tag{4.20}$$

cf. Resnick (2008, Proposition 3.13). Let $t \in [a,b]$. By construction, $\sup_{t_k \le t} \boldsymbol{x}_k$ only depends on $\{(t_i, \boldsymbol{x}_i) : i = 1, \ldots, N\}$ and $\sup_{t_k^{(n)} \le t} \boldsymbol{x}_k^{(n)}$ only depends on $\left\{\left(t_i^{(n)}, \boldsymbol{x}_i^{(n)}\right) : i = 1, \ldots, N\right\}$. More precisely, denote by

$$t_L = \max\{t_i \in \{t_1, \ldots, t_N\} : t_i \le t\}$$

the left neighbor of t among $\{t_1, \ldots, t_N\}$, and analogously by $t_{L(n)}^{(n)}$ the left neighbor of t among $\left\{t_1^{(n)}, \ldots, t_N^{(n)}\right\}$. Then

$$\sup_{t_k \le t} \boldsymbol{x}_k = \max(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L),$$
$$\sup_{t_k^{(n)} \le t} \boldsymbol{x}_k^{(n)} = \max\left(\boldsymbol{x}_1^{(n)}, \ldots, \boldsymbol{x}_{L(n)}^{(n)}\right).$$

Note that in general, $L \neq L(n)$! For instance, if $t_L = t$, and $t_L^{(n)} = t + 1/n$, then $L(n) < L$, and the pointwise convergence $\pi_n(t_L) \to \pi(t_L)$ will generally fail in such a jump. This is where the Skorokhod time transformation steps in. Define a homeomorphism $\lambda_n : [a,b] \to [a,b]$ by

$$\lambda_n(a) = a, \quad \lambda_n(b) = b, \quad \lambda_n(t_i) = t_i^{(n)}, \quad i = M+1, \ldots, N,$$

and linearly interpolated elsewhere. Then $t \in [t_i, t_{i+1})$ iff $\lambda_n(t) \in \left[t_i^{(n)}, t_{i+1}^{(n)}\right)$, $i = M+1, \ldots, N-1$ (and analogously $t \in [a, t_{M+1}) \iff \lambda_n(t) \in \left[a, t_{M+1}^{(n)}\right)$ and $t \in [t_N, b] \iff \lambda_n(t) \in \left[t_N^{(n)}, b\right]$). This implies

$$\sup_{t_k^{(n)} \le \lambda_n(t)} \boldsymbol{x}_k^{(n)} = \max\left(\boldsymbol{x}_1^{(n)}, \ldots, \boldsymbol{x}_L^{(n)}\right).$$

Clearly, (4.20) now implies $\pi_n(\lambda_n(t)) = \pi(t)$ for $t \in [a,b]$ and n large enough. Furthermore, we have

$$\pi_n(\lambda_n(t)) = \pi_n\left(t_i^{(n)}\right), \quad \pi(t) = \pi(t_i), \qquad t \in [t_i, t_{i+1}),\ i = M, \ldots, N,$$

with a slight abuse of notation since λ_n is not defined on $[t_M, a) \cup (b, t_{N+1})$, where t_{N+1} has to be chosen such that $t_{N+1} > b$. Hence, denoting by δ the discrete metric on $\mathcal{P}_d$,

$$\begin{aligned}\sup_{t\in[a,b]} \delta\left(\pi_n(\lambda_n(t)), \pi(t)\right) &= \max_{i=M,\dots,N} \sup_{t\in[t_i,t_{i+1})} \delta\left(\pi_n(\lambda_n(t)), \pi(t)\right) \\ &= \max_{i=M,\dots,N} \delta\left(\pi_n\big(t_i^{(n)}\big), \pi(t_i)\right) \\ &= 0,\end{aligned}$$

for n large enough.

Lastly, it is easy to see that $\sup_{t\in[a,b]} |\lambda_n(t) - t| \to_{n\to\infty} 0$. Finally, we have proven $\pi_n(\cdot) \to \pi(\cdot)$ in the Skorokhod metric. □

The following theorem now extends Lemma 4.27 which states the weak convergence of the sample hitting scenario to the extremal hitting scenario. Now we are adding a time variable and proof convergence of entire hitting scenario functionals. This is a similar step as from (4.15) to (4.16).

Theorem 4.31. *Suppose* (4.15) *holds and denote by* $\Pi_{[nt]}$ *and* $\Pi(t)$ *the hitting scenarios of* $\boldsymbol{M}_{[nt]}$ *and* $Y(t)$ *from* (4.16). *Then*

$$\Pi_{[n\cdot]} \to_{\mathcal{D}} \Pi(\cdot)$$

in $D\left((0,\infty), \mathcal{P}_d\right)$.

Proof. It is sufficient to show $\Pi_{[n\cdot]} \to_{\mathcal{D}} \Pi(\cdot)$ in $D([a,b], \mathcal{P}_d)$ with $0 < a < b$ such that $\Pi(\cdot)$ is almost surely continuous in b, cf. Resnick (2008, Proposition 4.18). Condition (4.15) is equivalent with the convergence $N_n \to_{\mathcal{D}} N$ from (4.17). It is not difficult to see that N satisfies the assumptions of Lemma 4.30 with probability one, which means that the mapping Ψ is almost surely continuous at N. The continuous mapping theorem now yields

$$\Pi_{[n\cdot]} = \Psi(N_n) \to_{\mathcal{D}} \Psi(N) = \Pi(\cdot)$$

in $D([a,b], \mathcal{P}_d)$. □

Remark 4.32. Interestingly, the distribution of $\Pi(t)$ does not depend on t and is equal to the distribution of the extremal hitting scenario Π of $\boldsymbol{\xi}$. This can be seen as follows. Following the arguments in Resnick (2007, Section 5.5.2), it is easy to see that the extremal process $\boldsymbol{Y} = (Y_t)_{t\geq 0}$ has some *self-similarity* property, i. e. for any $c > 0$

$$Y(c\cdot) =_{\mathcal{D}} cY(\cdot). \tag{4.21}$$

This is mainly due to the fact that the exponent measure ν is homogenous of order -1, i. e for any $c > 0$

$$\nu(cA) = c^{-1}\nu(A), \qquad A \in \mathbb{B}(\mathbb{E}).$$

Therefore, $Y(ct) =_{\mathcal{D}} cY(t)$ for any $c, t > 0$. Hence the distribution of the hitting scenarios of both sides have to coincide as well, which implies $\Pi(ct) =_{\mathcal{D}} \Pi(t)$. In particular, it follows $\Pi(c) =_{\mathcal{D}} \Pi(1) =_{\mathcal{D}} \Pi$, where the last equality is due to the fact that

$$\boldsymbol{\eta} =_{\mathcal{D}} \sup_{k\in\mathbb{N}} \boldsymbol{\vartheta}^{(k)} \mathbb{1}_{\{t_k \leq 1\}},$$

with $\left((t_k, \boldsymbol{\vartheta}^{(k)})\right)_{k\in\mathbb{N}}$ being the points of $\mathrm{PPP}(\lambda \times \nu)$ on $[0, \infty) \times \mathbb{E}$, where λ denotes the Lebesgue measure.

Proposition 4.33. *Let $\Pi(t)$ be the hitting scenario of $Y(t)$ from* (4.16)*, $t > 0$. Then the time-changed processes $(e^{-s}Y(e^s))_{s\in\mathbb{R}}$ and $(\Pi(e^s))_{s\in\mathbb{R}}$ are stationary processes in $D\left(\mathbb{R}, \mathbb{R}^d\right)$ and $D\left(\mathbb{R}, \mathcal{P}_d\right)$, respectively.*

Proof. The stationarity of $(e^{-s}Y(e^s))_{s\in\mathbb{R}}$ directly follows from (4.21). Let

$$N = \sum_{k\in\mathbb{N}} \epsilon_{\left(t_k, \boldsymbol{j}^{(k)}\right)}$$

be $\mathrm{PPP}(\lambda\times\nu)$ on $(0, \infty)\times\mathbb{E}$, where λ is the Lebesgue measure and ν is the exponent measure. The homogenity of order -1 of the exponent measure implies that the transformed point process

$$N' = \sum_{k\in\mathbb{N}} \epsilon_{(t'_k, \boldsymbol{j}'^{(k)})} = \sum_{k\in\mathbb{N}} \epsilon_{(\log t_k, \boldsymbol{j}^{(k)}/t_k)}$$

is $\mathrm{PPP}(\lambda \times \nu)$ on $\mathbb{R} \times \mathbb{E}$, which can be seen as follows: Define the transformation

$$\Phi : (0, \infty) \times \mathbb{E} \to \mathbb{R} \times \mathbb{E}, \quad (t, j) \mapsto (\log t, \boldsymbol{j}/t).$$

Then the intensity measure of N' is given by

$$\begin{aligned}
\left((\lambda \times \mu) \circ \Phi^{-1}\right)(A) &= \int \mathbb{1}_A((\log t, \boldsymbol{j}/t)) \; \mathrm{d}t \; \nu(\mathrm{d}\boldsymbol{j}) \\
&= \int t^{-1} \mathbb{1}_A((\log t, \boldsymbol{j})) \; \mathrm{d}t \; \nu(\mathrm{d}\boldsymbol{j})
\end{aligned}$$

$$= \int e^{-t} \mathbb{1}_A((t, \boldsymbol{j})) e^t \, \mathrm{d}t \; \nu(\mathrm{d}\boldsymbol{j})$$
$$= (\lambda \times \nu)(A).$$

Next, define for $h \geq 0$ a mapping via

$$\tilde{\Phi}_h : \mathbb{R} \times \mathbb{E} \to D(\mathbb{R}, \mathbb{E} \cup \{\mathbf{0}\}), \quad (t', \boldsymbol{j}') \mapsto \left(\boldsymbol{j}' e^{t'-s-h} \mathbb{1}_{\{t' \leq s+h\}}\right)_{s \in \mathbb{R}}.$$

This yields another PPP

$$\tilde{N}_h = \sum_{k \in \mathbb{N}} \epsilon_{\left(\boldsymbol{j}'^{(k)} e^{t'_k - s - h} \mathbb{1}_{\{t'_k \leq s+h\}}\right)_{s \in \mathbb{R}}} = \sum_{k \in \mathbb{N}} \epsilon_{\left(\boldsymbol{j}^{(k)} e^{-s-h} \mathbb{1}_{\{t_k \leq e^{s+h}\}}\right)_{s \in \mathbb{R}}}$$

on $D(\mathbb{R}, \mathbb{E} \cup \{\mathbf{0}\})$. The intensity measure of $\tilde{N}_h$ is given by

$$\left((\lambda \times \nu) \circ \tilde{\Phi}_h^{-1}\right)(A) = \int \mathbb{1}_A \left(\left(\boldsymbol{j}' e^{t'-s-h} \mathbb{1}_{\{t' \leq s+h\}}\right)_{s \in \mathbb{R}} \right) \mathrm{d}t' \; \nu(\mathrm{d}\boldsymbol{j}')$$
$$= \int \mathbb{1}_A \left(\left(\boldsymbol{j}' e^{t'-s} \mathbb{1}_{\{t' \leq s\}}\right)_{s \in \mathbb{R}} \right) \mathrm{d}t' \; \nu(\mathrm{d}\boldsymbol{j}') = \left((\lambda \times \nu) \circ \tilde{\Phi}_0^{-1}\right)(A).$$

Hence, $\tilde{N}_h =_{\mathcal{D}} \tilde{N}_0$ for any $h \geq 0$. Since $(\Pi(e^{s+h}))_{s \in \mathbb{R}}$ is built on $\tilde{N}_h$, the stationarity of $(\Pi(e^s))_{s \in \mathbb{R}}$ follows. □

REMARK 4.34. The multivariate process $\tilde{Z}(s) = \sup_{k \in \mathbb{N}} \boldsymbol{j}'^{(k)} e^{t'_k - s} \mathbb{1}_{\{t'_k \leq s\}}$, $s \in \mathbb{R}$, that arises in the proof of Proposition 4.33 has an interesting interpretation in terms of so-called *moving maxima processes* which were firstly introduced by Deheuvels (1983), and later investigated by e. g. Schlather (2002) or Stoev and Taqqu (2005). Let f be a probability density on $\mathbb{R}^d$, and $\left(\zeta_k, \boldsymbol{T}^{(k)}\right)$ be the points of PPP$(r^{-2} \, \mathrm{d}r \times \lambda)$ on $(0, \infty) \times \mathbb{R}^d$, with λ denoting the Lebesgue measure. Then

$$Z(\boldsymbol{s}) = \sup_{k \in \mathbb{N}} \zeta_k f\left(\boldsymbol{s} - \boldsymbol{T}^{(k)}\right), \qquad \boldsymbol{s} \in \mathbb{R}^d,$$

defines a stationary max-stable process on $\mathbb{R}^d$ with unit Fréchet margins, which is called *moving maxima process*. It is obvious that the processes

$$\tilde{Z}_i(s) = \sup_{k \in \mathbb{N}} j_i'^{(k)} e^{t'_k - s} \mathbb{1}_{\{t'_k \leq s\}}, \qquad i = 1, \ldots, d,$$

with $\boldsymbol{j}'^{(k)} = \left(j_1'^{(k)}, \ldots, j_d'^{(k)}\right)$ define moving maxima processes on $\mathbb{R}$. The corresponding probability density is $f(s) = e^{-s} \mathbb{1}_{\{s \geq 0\}}$, which corresponds to the standard exponential distribution. In that sense, we can think of $\tilde{\boldsymbol{Z}} = (\tilde{Z}(s))_{s \in \mathbb{R}}$ as a multivariate version of a moving maxima processes on $\mathbb{R}$.

5 Summary and Outlook

In the beginning of this thesis, a review of representations of max-stable processes, mostly in terms of generators and D-norms, was given. Next, the concept of dual D-norm functions was introduced. It turned out that dual D-norm functions arise when discussing componentwise exceedances of max-stable processes over a given function. This was particularly important when we were later investigating complete records. Chapter 2 ended with an introduction of distributional differentiability of max-stable processes. Naturally, there are many questions open for future research. In Proposition 2.34, we started with a generator that is almost surely differentiable, and ended up with a max-stable process which is only differentiable in distribution. Despite there is not much hope that we can actually obtain a max-stable process which is differentiable in a stronger sense, it might be possible to reduce the assumptions on the generator.

In Chapter 3, it was proposed how to reconstruct a given max-stable process by discretizing and interpolating the resulting max-stable random vector again, such that max-stability is being preserved. In the onedimensional case, where the index set of the max-stable processes is $S = [0, 1]$, there has been developed a method, which is also applicable to the interpolation of generalized Pareto processes. This model however, can not be transferred to the case $S = [0, 1]^k$ since it relies explicitly on the natural order that is given in $\mathbb{R}$. Therefore, other models have been discussed to deal with the higherdimensional case. However, there is still a lot of work to do, and it would be helpful to come up with simulation studies.

Lastly, we presented many results on multivariate records in Chapter 4. For instance, we derived formulas for the weighted limit probability that the n-th observation is a record, as n tends to infinity. It would be very nice to develop methods how to estimate these probabilities based on the observations. Also, it is still an open problem to use the domain-of-attraction condition in order to determine the limit distribution of a suitably normed multivariate record. In the univariate case, this has already been achieved, see e. g. Resnick (2008, Section 4.2). Last but not least, there might be a lot of other possible definitions of multivariate records that are still not discussed yet!

Bibliography

ARNOLD, B. C., BALAKRISHNAN, N., and NAGARAJA, H. N. (1998). *Records.* Wiley Series in Probability and Statistics. Wiley, New York.

AULBACH, S., and FALK, M. (2012a). Local asymptotic normality in δ-neighborhoods of standard generalized Pareto processes. *J. Statist. Plann. Inference* **142**, 1339–1347. doi:10.1016/j.jspi.2011.12.011.

AULBACH, S., and FALK, M. (2012b). Testing for a generalized Pareto process. *Electron. J. Stat.* **6**, 1779–1802. doi:10.1214/12-EJS728.

AULBACH, S., FALK, M., and HOFMANN, M. (2013). On max-stable processes and the functional D-norm. *Extremes* **16**, 255–283. doi:10.1007/s10687-012-0160-3.

AULBACH, S., FALK, M., HOFMANN, M., and ZOTT, M. (2015). Max-stable processes and the functional D-norm revisited. *Extremes* **18**, 191–212. doi:10.1007/s10687-014-0210-0.

BALKEMA, A. A., and RESNICK, S. I. (1977). Max-infinite divisibility. *J. Appl. Probab.* **14**, 309–319. doi:10.2307/3213001.

BARNDORFF-NIELSEN, O. E., and SOBEL, M. (1966). On the distribution of the number of admissible points in a vector random sample. *Theory Probab. Appl.* **11**, 283–305. doi:10.1137/1111020.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley Series in Probability and Mathematical Statistics, 1st ed. Wiley.

BILLINGSLEY, P. (1999). *Convergence of Probability Measures.* Wiley Series in Probability and Statistics, 2nd ed. Wiley, New York. doi:10.1002/9780470316962.

BROWN, B. M., and RESNICK, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14**, 732–739. doi:10.2307/3213346.

Buishand, T. A., de Haan, L., and Zhou, C. (2008). On spatial extremes: with application to a rainfall problem. *Ann. Appl. Stat.* **2**, 624–642. doi:10.1214/08-AOAS159.

Chen, W.-M., Hwang, H.-K., and Tsai, T.-H. (2012). Maxima-finding algorithms for multidimensional samples: A two-phase approach. *Comp. Geom.* **45**, 33–53. doi:10.1016/j.comgeo.2011.08.001.

Coles, S. G., Heffernan, J. E., and Tawn, J. A. (1999). Dependence measure for extreme value analyses. *Extremes* **2**, 339–365.

Daley, D. J., and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods.* Probability and its Applications, 2nd ed. Springer, New York. doi:10.1007/b97277.

Daley, D. J., and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II: General Theory and Structure.* Probability and its Applications, 2nd ed. Springer, New York. doi:10.1007/978-0-387-49835-5.

Davis, R. A., and Mikosch, T. (2008). Extremal value theory for space-time processes with heavy-tailed distributions. *Stochastic Process. Appl.* **118**, 560–584. doi:10.1016/j.spa.2007.06.001.

Deheuvels, P. (1983). Point processes and multivariate extreme values. *J. Multivariate Anal.* **13**, 257–272. doi:10.1016/0047-259X(83)90025-8.

Deuschel, J.-D., and Zeitouni, O. (1995). Limiting curves for i.i.d. records. *Ann. Probab.* **23**, 852–878.

Dombry, C., and Éyi-Minko, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab.* **18**, 1–21. doi:10.1214/EJP.v18-1991.

Dombry, C., Éyi-Minko, F., and Ribatet, M. (2013). Conditional simulation of Brown-Resnick processes. *Biometrika* **100**, 111–124. doi:10.1093/biomet/ass067.

Dombry, C., and Ribatet, M. (2015). Functional regular variations, pareto processes and peaks over threshold. In *Special Issue on Extreme Theory and Application (Part II)* (Y. Wang and Z. Zhang, eds.), *Statistics and Its Interface*, vol. 8, 9–17. doi:10.4310/SII.2015.v8.n1.a2.

Dombry, C., Ribatet, M., and Stoev, S. (2015). Probabilities of concurrent extremes. Tech. Rep. arXiv:1503.05748 [math.ST].

Einmahl, J. H. J., Krajina, A., and Segers, J. (2012). An M-estimator for tail dependence in aribtrary dimensions. *Ann. Statist.* **40**, 1764–1793. doi:10.1214/12-AOS1023.

Falk, M. (2006). A representation of bivariate extreme value distributions via norms on R^2. *Extremes* **9**, 63–68. doi:10.1007/s10687-006-0010-2.

Falk, M., Hofmann, M., and Zott, M. (2015). On generalized max-linear models and their statistical interpolation. *J. Appl. Probab.* **52**, 736–751. doi:10.1239/jap/1445543843.

Falk, M., Hüsler, J., and Reiss, R.-D. (2011). *Laws of Small Numbers: Extremes and Rare Events.* 3rd ed. Springer, Basel. doi:10.1007/978-3-0348-0009-9.

Ferreira, A., and de Haan, L. (2014). The generalized Pareto process; with a view towards application and simulation. *Bernoulli* **20**, 1717–1737. doi:10.3150/13-BEJ538.

Fisher, R. A., and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Cambridge Philos. Soc.* **24**, 180–190. doi:10.1017/S0305004100015681.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics.* 2nd ed. Krieger, Malabar.

Gerritse, G. (1986). Supremum self-decomposable random vectors. *Prob. Th. Rel. Fields* **72**, 17–33. doi:10.1007/BF00343894.

Giné, E., Hahn, M., and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87**, 139–165. doi:10.1007/BF01198427.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math. (2)* **44**, 423–453. doi:10.2307/1968974.

Gnedin, A. V. (1993). On multivariate extremal processes. *J. Multivariate Anal.* **46**, 207–213. doi:10.1006/jmva.1993.1057.

Gnedin, A. V. (1994*a*). Conical extremes of a multivariate sample. *J. Res. Natl. Insl. Stand. Technol.* **99**, 511–519.

Gnedin, A. V. (1994*b*). On a best-choice problem with dependent criteria. *J. Appl. Probab.* **31**, 221–234. doi:10.2307/3215248.

GNEDIN, A. V. (1998). Records from a multivariate normal sample. *Statist. Probab. Lett.* **39**, 11–15. doi:10.1016/S0167-7152(98)00004-2.

GNEDIN, A. V. (2007). The chain records. *Electron. J. Probab.* **12**, 767–786. doi:10.1214/EJP.v12-410.

GOLDIE, C. M., and RESNICK, S. I. (1989). Records in a partially ordered set. *Ann. Probab.* **17**, 678–699. doi:10.1214/aop/1176991421.

GOLDIE, C. M., and RESNICK, S. I. (1995). Many multivariate records. *Stochastic Process. Appl.* **59**, 185–216. doi:10.1016/0304-4149(95)00047-B.

DE HAAN, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194–1204. doi:10.1214/aop/1176993148.

DE HAAN, L., and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction.* Springer Series in Operations Research and Financial Engineering. Springer, New York. doi:10.1007/0-387-34471-3. See http://people.few.eur.nl/ldehaan/EVTbook.correction.pdf and http://home.isa.utl.pt/~anafh/corrections.pdf for corrections and extensions.

DE HAAN, L., and LIN, T. (2001). On convergence toward an extreme value distribution in $C[0,1]$. *Ann. Probab.* **29**, 467–483. doi:10.1214/aop/1008956340.

DE HAAN, L., NEVES, C., and PENG, L. (2008). Parametric tail copula estimation and model testing. *J. Multivariate Anal.* **99**, 1260–1275. doi:10.1016/j.jmva.2007.08.003.

DE HAAN, L., and RESNICK, S. (1977). Limit theory for multivariate sample extremes. *Probab. Theory Related Fields* **40**, 317–337. doi:10.1007/BF00533086.

HASHORVA, E., and HÜSLER, J. (2005). Multiple maxima in multivariate samples. *Statist. Probab. Lett.* **75**, 11–17. doi:10.1016/j.spl.2005.05.014.

HEFFERNAN, J. E. (2000). A directory of coefficients of tail dependence. *Extremes* **3**, 279–290.

HOFMANN, M. (2012). *Contributions to Extreme Value Theory in the Space* $C[0,1]$. Ph.D. thesis, University of Würzburg, http://opus.bibliothek.uni-wuerzburg.de/volltexte/2012/7440/.

HOFMANN, M. (2013). On the hitting probability of max-stable processes. *Statist. Probab. Lett.* **83**, 2516–2521. doi:10.1016/j.spl.2013.07.012.

HULT, H., and LINDSKOG, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.* **115**, 249–274. doi:10.1016/j.spa.2004.09.003.

HWANG, H.-K., and TSAI, T.-H. (2010). Multivariate records based on dominance. *Electron. J. Probab.* **15**, 1863–1892. doi:10.1214/EJP.v15-825.

KABLUCHKO, Z., SCHLATHER, M., and DE HAAN, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37**, 2042–2065. doi:10.1214/09-AOP455.

KAŁUSZKA, M. (1995). Estimates of some probabilities in multidimensional convex records. *Applicationes Mathematicae* **23**, 1–11.

NORBERG, T. (1986). On the existence and convergence of probability measures on continuous semi-lattices. Tech. Rep., University of North Carolina, Center for Stochastic Processes, http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA179267.

RÉNYI, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. scient. Univ. de Clermont. Mathématiques* **8**, 7–13.

RESNICK, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

RESNICK, S. I. (2008). *Extreme Values, Regular Variation, and Point Processes.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

RESNICK, S. I., and ROY, R. (1991). Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Probab.* **1**, 267–292. doi:10.1214/aoap/1177005937.

ROOTZÉN, H., and TAJVIDI, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12**, 917–930. doi:10.3150/bj/1161614952.

SCHLATHER, M. (2002). Models for stationary max-stable random fields. *Extremes* **5**, 33–44. doi:10.1023/A:1020977924878.

SCHMIDT, R., and STADTMÜLLER, U. (2006). Non-parametric estimation of tail dependence. *Scand. J. Statist.* **33**, 307–335. doi:10.1111/j.1467-9469.2005.00483.x.

SMITH, R. L. (1990). Max-stable processes and spatial extremes. Preprint, Univ. North Carolina, `http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html`.

STEPHENSON, A., and TAWN, J. (2005). Exploiting occurrence times in likelihood inference for componentwise maxima. *Biometrika* **92**, 213–227. `doi:10.1093/biomet/92.1.213`.

STOEV, S. A., and TAQQU, M. S. (2005). Extremal stochastic integrals: a parallel between max-stable processes and α-stable processes. *Extremes* **8**, 237–266. `doi:10.1007/s10687-006-0004-0`.

TAKAHASHI, R. (1988). Characterizations of a multivariate extreme value distribution. *Adv. in Appl. Probab.* **20**, 235–236. `doi:10.2307/1427279`.

VATAN, P. (1985). Max-infinite divisibility and max-stability in infinite dimensions. In *Probability in Banach Spaces V: Proceedings of the International Conference held in Medford, USA, July 16–27, 1984* (A. Beck, R. Dudley, M. Hahn, J. Kuelbs, and M. Marcus, eds.), *Lecture Notes in Mathematics*, vol. 1153, 400–425. Springer, Berlin. `doi:10.1007/BFb0074963`.

VERVAAT, W. (1973). Limit theorems for records from discrete distributions. *Stochastic Process. Appl.* **1**, 317–334. `doi:10.1016/0304-4149(73)90015-X`.

WANG, Y., and STOEV, S. A. (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. in Appl. Probab.* **43**, 461–483, `http://projecteuclid.org/euclid.aap/1308662488`.