Master Thesis

Comparison of Simulation Methods of Brown-Resnick processes

Dennis Leber

31.10.2015

Supervisor: Prof. Dr. Claudia Klüppelberg, M.Sc. Sven Buhl

Chair of Mathematical Statistics

Center for Mathematical Sciences

Technische Universität München

Hiermit erkläre ich, dass ich die Masterarbeit selbstständig angefertigt und nur die angegebenen Quellen verwendet habe.

Garching, den 30.Oktober 2015

Abstract

This thesis deals with different simulation methods for a specific max-stable field, the so called **Brown-Resnick process**. It was first introduced by Brown and Resnick [6] in 1977. Brown-Resnick processes are strongly useful for modelling extremes in space and time and are used in several fields, like geostatistics, meteorology and physics.

After an introduction to extreme value theory and max-stable processes, we examine the different methods for speed, efficiency and max-stability. One of the main problem we face with the simulation of Brown-Resnick processes is the confrontation with a pointwise maximum over an infinite number of random functions as in the spectral representation, which makes simulation highly non-trivial. Consequently, most of the literature and methods till today, only present an approximation of the extreme processes.

The first method we look at - the **Gaussian approximation** method - constructs max-stable random fields based on a result by Hüsler and Reiss [30] as limits of normalized and rescaled pointwise maxima of Gaussian random fields, where the underlying correlation function has to satisfy a certain regularity condition.

The second approach - the **Naive simulation** - examines the problems, which can occur, if we try to simulate the Brown-Resnick process directly from its spectral representation.In many cases this leads to a bad approximation, but we still want to find out, if the result is satisfying regarding max-stability, since we expect an advantage in speed.

The third method - the method by **Dieker and Mikosch** [20] - is the first to devise an exact simulation for Brown-Resnick random fields. The main idea is based on a change of measure argument. We extend the method to space-time dimension and clarify if the proof of the change of measure argument still holds in the space-time setting and for certain class of dependence model.

The fourth and last method, which has been published by **Dombry, Engelke and Oesting** [23] during the writing of my thesis, focus only on so called extremal functions, which can be shortly described as functions, which actually contribute to the pointwise maximum.

Since most of the simulations of the corresponding Gaussian fields are done with fractional Brownian motions this thesis also deals with self-similar processes.

Acknowledgement

First and foremost, I want to thank Prof. Claudia Klüppelberg for the great opportunity to pursue this topic at her department and also for her guidance. Moreover I am grateful for the supervision by my advisor Sven Buhl who was both very helpful and supportive with new ideas. He always had time for intensive discussion as well as for time-consuming simulations.

I also want to thank Dr. Marius Hofert from the University of Waterloo and Dr. Marco Oesting from the Universität Siegen who gave me computational advice for my simulations.

Finally I want to express my deep gratitude for my parents and my brother for their unconditional support, not only during the writing of my thesis, but also about my complete studies, where I am also thankful to my good friend Julian Grauer for sharing these last five years with me.

Contents

Chapter 1

Extreme Value Theory: Basics

This chapter is dealing with some central topics of extreme value theory in the univariate case as well in the multivariate case. One of the main result is the Fisher-Tippett theorem which specifies the asymptotic behaviour of sample extremes and the so resulting families of extreme value distribution. There will be no focus on proofs, which can be gleaned in Embrechts et al. [24] and de Haan and Ferreira [17].

1.1 Univariate Extreme Value Theory

1.1.1 Setup

Let $X_1, X_2, X_3, ...$ be independent and identically distributed random variables with distribution function F. The Central Limit Theorem is concerned with limit behaviour of the normalized partial sums of Random Variables. In Extreme value theory the focus lies on the sample maximum $M_n = \bigvee_{i=1}^{n} X_i$ so we want to find the real normalizing and centering sequences $(a_n)_{n\in\mathbb{N}}, a_n > 0$ for all n, and $(b_n)_{n\in\mathbb{N}}$ such that

$$\frac{M_n - b_n}{a_n} \xrightarrow{\mathcal{D}} Y, \quad n \to \infty, \tag{1.1}$$

where $Y \sim G$ for a non-degenerate distribution function G and $\xrightarrow{\mathcal{D}}$ means convergence in distribution. Since the distribution of M_n is determined by

$$P\left(M_n \leq x\right) = P\left(X_1 \leq x, ..., X_n \leq x\right) = F^n(x)$$

we can rewrite (1.1) by

$$F^n\left(a_n x + b_n\right) \to G(x), \quad n \to \infty, \tag{1.2}$$

for all points x, where $F(x) \neq 0$. Next for each of those limit distributions, we shall find necessary and sufficient conditions on the initial distribtuion F such that (1.2) holds.

Definition 1.1. *The class of distributions of F satisfying (1.1) or (1.2) is called the* ***maximum domain of attraction*** *or simply domain of attraction of G and write $F \in MDA(G)$.*

Possible candidates for the distribution of G will be discussed right after another important property in Extreme Value Theory.

Definition 1.2 (cf. Embrechts et al. [24], Definition 3.2.1)**.** *The distribution Function F is called* ***max-stable*** *if for some sequences $(a_n)_{n\in\mathbb{N}}, a_n > 0$ for all n, and $(b_n)_{n\in\mathbb{N}}$,*

$$F^n(a_n x + b_n) = P\left(\frac{M_n - b_n}{a_n} \leq x\right) = P\left(X_1 \leq x\right) = F(x)$$

for all $x \in \mathbb{R}$ and $n \in \mathbb{N}$.

It turns out, that only a few distributions G can be considered as the asymptotic limit distribution of the standardized maximum M_n. They are referred to as the extreme value distriubtions and first introduced by Fisher and Tippett [1928]. These are the following three distribution functions:

Theorem 1.3 (Fisher-Tippett Theorem, see for example Embrechts et al. [24], Theorem 3.2.3)**.** *Let $X_1, X_2, X_3, ...$ be a sequence of independent identical distributed Random variables. If (1.1) is satisfied then the limit distribtuion G belongs to the type of the following three distribution functions:*

$$\textit{Fréchet}: \quad \Phi_\alpha(x) = \begin{cases} 0, & x = 0 \\ \exp\left\{-x^{-\alpha}\right\}, & x > 0 \end{cases} \quad \alpha > 0 \tag{1.3}$$

$$\textit{Weibull}: \quad \Psi_\alpha(x) = \begin{cases} \exp\left\{-(-x)^{\alpha}\right\}, & x = 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0 \tag{1.4}$$

$$\textit{Gumbel}: \Lambda(x) = \exp\left\{-e^{-x}\right\}, \quad x \in \mathbb{R} \tag{1.5}$$

Note that the class of max-stable distributions is the same as the class of extreme value distributions. An alternative approach with a finite dimensional parametrization of G is stated in de Haan and Ferreira [17] (Theorem 1.1.6):

Theorem 1.4. *For $\gamma \in \mathbb{R}$ the following statements are equivalent:*

1. *There exist real constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that*

$$\lim_{n\to\infty} F^n\left(a_n x + b_n\right) = G_y(x) = \begin{cases} \exp\left\{-(1+\gamma x)^{-\frac{1}{\gamma}}\right\}, & \text{if } \gamma \neq 0 \\ \exp\left\{-e^{-x}\right\}, & \text{if } \gamma = 0, \end{cases} \tag{1.6}$$

   for all x with $1 + \gamma x > 0$.

2. *There exists a positive function f such that*

$$\lim_{t\uparrow x^\star} \frac{1 - F(t + x f(t))}{1 - F(t)} = \begin{cases} (1+\gamma x)^{-\frac{1}{\gamma}}, & \text{if } \gamma \neq 0 \\ \exp(-x), & \text{if } \gamma = 0, \end{cases} = 1 - GPD_y(x), \tag{1.7}$$

   for all x for which $1 + \gamma x > 0$, where $x^\star = \sup\{x : F(x) < 1\}$, so the right endpoint of the distribution function F.

1.2 Multivariate Extreme Value Theory

Now we want to expand to the d-dimensional case, which is the one we will concentrate on the further chapters, especially when we are dealing with random fields in space and time. Most of the theory is similar to the univariate case, but now also marginal behaviour will play an important role. This part is based on Beirlant et al. [4], chapter 8.

1.2.1 Setup

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, ...$ be independent and identically distributed d-variate random variables with distribution function F. Like in the univariate case the focus in multivariate Extreme value theory lies on limit behaviour of the sample maximum $\boldsymbol{M}_n = \bigvee_{i=1}^n \boldsymbol{X}_i$ so we want to find the real normalizing and centering sequences $(\boldsymbol{a}_n)_{n\in\mathbb{N}}, \boldsymbol{a}_n > 0$ for all $n \in \mathbb{N}$, and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$ such that

$$\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n} \xrightarrow{\mathcal{D}} \boldsymbol{Y}, \quad n \to \infty, \tag{1.8}$$

where $\boldsymbol{Y} \sim G$ for a non-degenerate distribution function G. Since the distribution of $\boldsymbol{M}_n$ is determined by

$$P\left(\boldsymbol{M}_n \leq \boldsymbol{x}\right) = P\left(\boldsymbol{X}_1 \leq \boldsymbol{x}, ..., \boldsymbol{X}_n \leq \boldsymbol{x}\right) = F^n(\boldsymbol{x})$$

we can rewrite (1.8) by

$$F^n\left(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n\right) \to G(\boldsymbol{x}), \quad n \to \infty, \tag{1.9}$$

for all points $\boldsymbol{x}$, where $F(\boldsymbol{x}) \neq 0$. Next for each of those limit distributions, we shall find necessary and sufficient conditions on the initial distribution F such that (1.9) holds.

Definition 1.5. *The class of multivariate distributions of F satisfying (1.8) or (1.9) is called the* ***maximum domain of attraction*** *or simply domain of attraction of G and write $F \in MDA(G)$. G is called extreme value distribution.*

Since a sequence of random vectors can only converge in distribution if the corresponding marginal sequences do, we obtain for $j = 1, ..., d$

$$F_j^n\left(a_n, jx_j + b_n, j\right) \underset{\to}{\mathcal{D}} G_j(x_j), \quad n \to \infty$$

where F_j and G_j denote the j-th marginal distribution functions of F and G respectively. It follows that each G_j is by itself a univariate extreme value distribution from the family of Theorem 1.3 and F_j is in its domain of attraction.

The meaning for max-stability is the same as in the univariate case.

Definition 1.6. *The multivariate distribution Function F is called* ***max-stable*** *if for some sequences $(\boldsymbol{a}_n)_{n\in\mathbb{N}}, \boldsymbol{a}_n > 0$ for all $n \in \mathbb{N}$, and $(\boldsymbol{b}_n)_{n\in\mathbb{N}}$,*

$$F^n\left(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n\right) = P\left(\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n} \leq \boldsymbol{x}\right) = P\left(\boldsymbol{X}_1 \leq \boldsymbol{x}\right) = F(\boldsymbol{x}),$$

for all $\boldsymbol{x} \in \mathbb{R}^d$ and $n \in \mathbb{N}$. So again, a max- stable distribution is in its own domain of attraction.

1.2.2 Ways to display Multivariate Extreme Distributions

In the univariate case, we determined by the Fisher-Tippett Theorem (Theorem 1.3) a parametrization of the extreme value distributions. This is not possible in the d-dimensional case, because the class of the dependence structure is too large. In the following we describe the multivariate distribution function G by a special measure, the exponent measure. But first we standardize the margins of G.

Reduction to standard Fréchet margins

To study the dependence of a max-stable distribution, it is convenient to standardize the margins so that they are all the same. First we want to concentrate on standard Fréchet margins.

Let $\boldsymbol{X}$ be a d-dimensional random vector with distribution function G and let $G_\star$ be

$$G_\star(\boldsymbol{z}) = G\left[G_1^{\leftarrow}\left(e^{-1/z_1}\right), G_2^{\leftarrow}\left(e^{-1/z_2}\right), ..., G_d^{\leftarrow}\left(e^{-1/z_d}\right)\right], \quad \boldsymbol{z} \in (\boldsymbol{0}, \boldsymbol{\infty}). \tag{1.10}$$

Then the margins of $G_\star$ are standard Fréchet, so the distribution function of the j-th margin is $G_{\star,j}(z) = \exp(-1/z)$. Furthermore, G_j^{-1} is the generalized inverse of the j-th margin, which is defined by

$$G_i^{\leftarrow}(y) = \inf\{z : G_i(z) \geq y\}$$

Conversely,

$$G(\boldsymbol{x}) = G_\star\left[-1/\log G_1(x_1), ..., -1/\log G_d(x_d)\right], \quad x \in \mathbb{R}^d, \tag{1.11}$$

where, taking appropriate limits, $-1/\log(0) := 0$ and $-1/\log(1) := \infty$.

Both directions, (1.10) and (1.11), can be easily proofed with the help of the following Lemma.

Theorem 1.7 (Probability Integral Transform, see for example Czado [14], Lemma 1.3)**.** *If X is a continious random variable with distribution function F and inverse F^{-1}, then $U := F(X) \sim U[0,1]$. Therefore if $X := F^{-1}(U)$, then $X \sim F$.*

Now we rewrite (1.10) to

$$G_\star = P\left[-\log\left(G_1(X_1)\right)^{-1} \leq z_1, -\log\left(G_2(X_2)\right)^{-1} \leq z_2, ..., -\log\left(G_d(X_d)\right)^{-1} \leq z_d\right],$$

with $\boldsymbol{z} \in (\boldsymbol{0}, \boldsymbol{\infty})$. Since $t \to -\log(t)^{-1}$ for $t \in (0,1)$ is the inverse function of the standard Fréchet distribution function, we can use Theorem 1.7 for claiming that $-\log\left(G_j(X_j)\right)^{-1}$, for $j = 1, .., d$ is standrad Fréchet distributed, and so $G_\star$ has standard Fréchet margins. Not only does $G_\star$ has max-stable margins, it itself is max-stable and so also an extreme value distribution.

With this simplification to standard Fréchet margins a first representation of the Extrme value distribution is possible. A max-stable distribtuion function can be represented via an **exponent measure**. A proof for the bivariate case can be found for example in Balkema and Resnick [2].

Theorem 1.8. *A d-variate distribution function G with standard Fréchet margins is max-stable if and only if there exists a measure μ on $[\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ such that*

$$G_\star(\boldsymbol{z}) = \exp\left(-\mu_\star\left([\mathbf{0}, \infty) \setminus [\mathbf{0}, \boldsymbol{z}]\right)\right) \quad \boldsymbol{z} \in [\mathbf{0}, \infty]. \tag{1.12}$$

Definition 1.9. *The measure $\mu_\star$ in (1.12) is called exponent measure.*

Remark 1.10. *One important property of the exponent measure $\mu_\star$ is the homogeneity, that is,*

$$\mu_\star(s\cdot) = s^{-1}\mu_\star(\cdot), \quad 0 < s\infty,$$

which holds for all Borel subsets of $[\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$.

Before a transformation to Gumbel margins is introduced, we can determine with the exponent measure the so called stable tail dependence function (see for instance Beirlant et al. [4], page 257).

Definition 1.11. *The **stable tail dependence function** is defined by*

$$\begin{aligned} l(\boldsymbol{v}) &= \mu_\star\left([\mathbf{0}, \boldsymbol{\infty}] \setminus [\mathbf{0}, (1/v_1, ..., 1/v_d)]\right) \\ &= \mu_\star\left([\mathbf{0}, \boldsymbol{v}^{-1}]^c\right) \\ &= -\log G\left(\boldsymbol{v}^{-1}\right), \quad \boldsymbol{v} \in [\mathbf{0}, \boldsymbol{\infty}] \end{aligned}$$

So every max-stable distribution function can be built by its margins and its stable tail dependence function through

$$-\log G(\boldsymbol{x}) = l\left\{-\log G_1(x_1), ..., -\log G_d(x_d)\right\}, \quad \boldsymbol{x} \in \mathbb{R}^d.$$

Reduction to standard Gumbel margins

Annother way to transform the marginal distribution is into the Gumbel margins. Recall that the Gumbel distribution is defined as $\Lambda(x) = \exp(-e^{-x})$ for $x \in \mathbb{R}$, with inverse function $\Lambda^{-1}(y) = -\log(-\log(y))$. Although, the Fréchet distribtuion is the most popular form to transform the margins, we shortly introduce here the Gumbel case.

Theorem 1.12 (see for instance Beirlant et al. [4], page 272)**.** *If G is a multivariate extreme value distribution, then the distribution function of the random vector*

$$\left(-\log\left(-\log(G_1(X_1))\right), ..., -\log\left(-\log(G_d(X_d))\right)\right)$$

is given by

$$\Lambda_\star(\boldsymbol{x}) = \exp\left[-l(e^{-x_1}, ..., e^{-x_d})\right].$$

1.2.3 The Domain of Attraction

At the beginning of this section, we introduced the domain-of-attraction equation

$$\lim_{n\to\infty} F^n\left(\boldsymbol{a}_n\boldsymbol{x}+\boldsymbol{b}_n\right)=G(\boldsymbol{x}),\quad \boldsymbol{x}\in[-\boldsymbol{\infty},\boldsymbol{\infty}], \tag{1.13}$$

with $\boldsymbol{a}_n \in (\boldsymbol{0},\boldsymbol{\infty})$ and $\boldsymbol{b}_n \in \mathbb{R}^d$, for the multivariate case. Until now, we only concentrated on the right hand side of this equation. In the section before (1.2.2) we tried to find an alternative to the Fisher-Tippet Theorem in the univariate case to display the multivariate extreme distribution function $G(\boldsymbol{x})$.

To get more information about the distribution function F in (1.13), we have to manipulate it a little bit. Some of these manipulation-methods are the following (see therefore Beirlant et al. [4], chapter 8.3.1):

1. **Representation with stable tail dependence function**

   We „manipulate“ $F^n = \left[1-\frac{1}{n}\left(n(1-F)\right)\right]^n$ and using the fact from basic analysis that $\left(1-\frac{1}{n}x_n\right)^n \to e^{-x}$ if and only if $x_n \to x \in [0,\infty]$ as $n\to\infty$, so writing the exponential function as the limit of a sequence. Having this in mind, (1.13) only holds if and only if

   $$\lim_{n\to\infty} n\left[1-F\left(\boldsymbol{a}_n\boldsymbol{x}+\boldsymbol{b}_n\right)\right]=-\log G(\boldsymbol{x}),\quad \boldsymbol{x}\in[-\boldsymbol{\infty},\boldsymbol{\infty}], \tag{1.14}$$

   where we can rewrite by max-stability (1.14) as

   $$\begin{aligned} 1-F\left(\boldsymbol{a}_x\boldsymbol{x}+\boldsymbol{b}_n\right) &\sim -\log G\left(\boldsymbol{\alpha}_n\boldsymbol{x}+\boldsymbol{\beta}_n\right) \\ &\sim 1-G\left(\boldsymbol{\alpha}_n\boldsymbol{x}+\boldsymbol{\beta}_n\right),\quad n\to\infty \end{aligned} \tag{1.15}$$

   for $\boldsymbol{x}\in \mathit{supp}(G)$. This relation can be used as a starting point for statistical inference on $F(\boldsymbol{x})$ in $\boldsymbol{x}$-regions for which each marginal distribution function $F_j(x_j)$ is close to one. If we choose a vector $\boldsymbol{u}$ such that $F_j(u_j)$ is sufficiently close to one for every margin $j=1,...,d$, we can conclude from (1.15) the approximation

   $$F(\boldsymbol{x})\approx G\left(\boldsymbol{\alpha}_n\frac{1}{\boldsymbol{a}_n}\boldsymbol{x}-\boldsymbol{\alpha}_n\frac{1}{\boldsymbol{a}_n}\boldsymbol{b}_n+\boldsymbol{\beta}_n\right):=\tilde{G}(\boldsymbol{x}),\quad \boldsymbol{x}\geq\boldsymbol{u}. \tag{1.16}$$

   $\tilde{G}$ is also an extreme value distribution, since G and $\tilde{G}$ differ only in scale and location. They both have the same stable tail dependence function l introduced in definition 1.11 so we finally can display the distribution function F approximately in terms of l as

   $$F(\boldsymbol{x})\approx \exp\left(-l(\boldsymbol{v})\right),\quad \boldsymbol{x}\geq\boldsymbol{u}, \tag{1.17}$$

   where, for $j=1,...,d$ and $x_j \geq u_j$, we can write the one-dimensional entries v_j similar to definition 1.11 as

   $$v_j=-\log\left(\tilde{G}_j(x_j)\right) \tag{1.18}$$

   There are several models for the stable tail dependence function, for example in Gumbel [28] or Coles and Tawn [12]. Together with equations (1.17) and (1.18) this can form a fully parametric model for F in the region $[\boldsymbol{u},\boldsymbol{\infty})$.

2. **Representation with the exponent measure**

We focus again on the „manipulation“ $F^n = \left[1 - \frac{1}{n}\left(n(1-F)\right)\right]^n$ and recall that right hand side of (1.13) has an exponent measure, μ, similar to (1.12), which is given by $\mu\left([\mathbf{0}, \boldsymbol{\infty}) \setminus [\mathbf{0}, \boldsymbol{x}]\right) = -\log\left(G(\boldsymbol{x})\right)$ for $\boldsymbol{x} \geq 0$. We define the measures μ_n on $[\mathbf{0}, \boldsymbol{\infty}) \setminus \{\boldsymbol{q}\}$ as

$$\mu_n(\cdot) = nP\left[\frac{(\boldsymbol{X} - \boldsymbol{b}_n)}{\boldsymbol{a}_n} \vee \mathbf{0} \in \cdot\right], \tag{1.19}$$

where $\boldsymbol{X}$ has distribution function F. Moreover

$$\mu_n\left([\mathbf{0}, \boldsymbol{\infty}) \setminus [\mathbf{0}, \boldsymbol{x}]\right) = n\left[1 - F\left(\boldsymbol{a}_x \boldsymbol{x} + \boldsymbol{b}_n\right)\right], \quad \boldsymbol{x} \in [\mathbf{0}, \boldsymbol{\infty}].$$

We can express (1.14) now in terms of the measures μ_n and μ and vague convergence.

Definition 1.13 (see for instance Resnick [40])**.** *Let ν_n and ν be non-negative Radon measures on $[\mathbf{0}, \boldsymbol{\infty}] \setminus \{\mathbf{0}\}$. Then ν_n converges vaguely to ν if*

$$\int_{[\mathbf{0},\boldsymbol{\infty}]\setminus\{\mathbf{0}\}} f(\boldsymbol{x})\nu_n(d\boldsymbol{x}) \underset{n\to\infty}{\to} \int_{[\mathbf{0},\boldsymbol{\infty}]\setminus\{\mathbf{0}\}} f(\boldsymbol{x})\nu(d\boldsymbol{x})$$

for all continuous non-negative functions f with compact support. We notate vague convergence as $\nu_n \overset{v}{\to} \nu$.

For our case we get

$$\mu_n\left([\mathbf{0}, \boldsymbol{\infty}) \setminus [\mathbf{0}, \boldsymbol{x}]\right) = n\left[1 - F\left(\boldsymbol{a}_x \boldsymbol{x} + \boldsymbol{b}_n\right)\right] \overset{v}{\to} \mu\left([\mathbf{0}, \boldsymbol{\infty}) \setminus [\mathbf{0}, \boldsymbol{x}]\right), \quad \text{for } n \to \infty$$

Before we come to a third representation, we give a short introduction to Point Processes at this point. For a deeper study of the topic, see for example on Resnick [40] or Snyder and Miller [44]. We refer to Embrechts et al. [24]. The concept of point processes is not only important for the third representation method but also in later chapters to display Brown-Resnick fields.

Definition 1.14. *Let I be a Indexset, and $\{X_i : i \in I\}$ represent the location of points, occurring randomly in a state space S. A **point process** N counts the number of points in regions of S:*

$$N(A) = \sum_{i \in I} \mathbf{1}_{X_i \in A} \tag{1.20}$$

*for a sub-set $A \subseteq S$. Furthermore, the **intensity measure***

$$\Lambda(A) = E[N(A)] \tag{1.21}$$

gives the expected number of points in a set A. Λ has a density function $\lambda : S \to [0, \infty)$, that is

$$\Lambda(A) = \int_A \lambda(x)dx,$$

*then λ is called the **intensity function** of the process.*

The most common type of point processes are Poisson processes. These also will be processes, which are a part of Brown-Resnick processes.

Definition 1.15. *A point process N with intensity measure Λ is called a* ***Poisson process*** *if these two conditions are fulfilled:*

- *for each set A such that $\Lambda(A) < \infty$ is $N(A)$ a Poisson random variable with mean $\Lambda(A)$ so that:*
$$P\left(N(A) = k\right) = \exp\left(-\Lambda(A)\right) \frac{(\Lambda(A))^k}{k}.$$
Moreover we assume that Λ is a radon measure, which means it is finite for all compact subsets $A \subseteq E$.
- *for all positive integer k and all disjoint sets $A_1, ..., A_k$ are the random variables $N(A_1), ..., N(A_k)$ independent.*

Remark 1.16. *A Poisson Process on a Euclidiean space is called* ***homogeneous*** *if its intensity function λ is constant, $\lambda(x) \equiv \lambda$. When it is not constant we call the Poisson process* ***inhomogeneous****.*

These first definitions give us a recall for point processes and especially Poisson processes, which brings us back to the third representation problem.

3. **Representation via point processes**

   Consider a sequence of random vectors $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$ with a state space $E \subset \mathbb{R}^d$ which is equipped with the σ-algebra $\varepsilon := \mathcal{B}(E)$ and all random vectors are identical distributed with distribution function F. Then we stick to (1.19) and set $\boldsymbol{X}_{i,n} = \frac{(\boldsymbol{X}-\boldsymbol{b}_n)}{\boldsymbol{a}_n} \vee \boldsymbol{0}$. Furthermore we define by
   $$N_n(\cdot) := \sum_{i=1}^{\infty} \mathbf{1}_{\{\left(\frac{i}{n}, \boldsymbol{X}_{i,n}\right)\in\cdot\}}$$
   a point process on $[0,\infty) \times [\boldsymbol{0}, \boldsymbol{\infty})$ It follows that (1.13) is equivalent to
   $$N_n \xrightarrow{\mathcal{D}} N, \quad n \to \infty$$
   where N is a Poisson process with mean measure Λ. Λ is defined in terms of the exponent measure μ of the extreme value distribution G with standard Fréchet margins as
   $$\begin{aligned} E\left[N\left((s_1, s_2] \times [\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]\right)\right] &= \Lambda\left((s_1, s_2] \times [\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]\right) \\ &:= \int_{[\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]} \int_{(s_1, s_2]} du\ \mu(d\boldsymbol{x}) \\ &= (s_1 - s_2) \cdot \mu\left([\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]\right), \end{aligned}$$
   where $((s_1, s_2] \times [\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2]) \subset ([0,\infty) \times [\boldsymbol{0}, \boldsymbol{\infty}])$.

There are several other representation concering the distribution function F, see therefore Beirlant et al. [4], Chapter 8.3.1, but at this point and the following chapters it should be enough.

Chapter 2

Stochastic Processes

This section is mostly based on a mixture of a paper from Adler and Taylor [1] on random fields and the book statistics for spatial data by Cressie [13]. It should gives us an introduction to stochastic processes and their properties. To handle Brown-Resnick processes we have to know the essential basics about Gaussian processes. Since we are not only dealing in our simulations with the most common Gaussian process - the Brownian motion (in one ore more dimensions) - but also with fractional Brownian motions, we give a short summary about self-similar processes.

2.1 Introduction to Stochastic processes and Space-time Processes

A **random field** is simply a stochastic process, taking values in a Euclidean space, and defined over some set $M \subset \mathbb{R}^d$ of dimensionality at least one.

Definition 2.1. *Given a set M, a **stochastic Process $\boldsymbol{X}$** over M is a collection of random variables*

$$\{X(\boldsymbol{m}) : \boldsymbol{m} \in M\}.$$

In most parts of this thesis the focus will lie on spatial data at different time points. So the set M will split into two parts, a one dimensional time-line $T = [0, \infty)$ and a space component $S \subset \mathbb{R}^d$. Usually the dimension of S will be $d = 2$.

Definition 2.2. *A **space-time process $\boldsymbol{Z}$** will be denoted as*

$$\boldsymbol{Z} = \{Z(\boldsymbol{s}, t) : \boldsymbol{s} \in S, t \in T\}. \tag{2.1}$$

A first small example of a space-time process could be the height of an ocean surface above some nominal mean plane. It is obviously a function of both time and space, and so we acknowledge this by representing it as in (2.1). Another example would be rainfall at a finite number of places throughout a year measured on a daily basis.

2.1.1 Gaussian Processes and Max-Stable Processes

The most common stochastic process is the Gaussian process. We use the introduction of the Gaussian process to recall some basics about univariate and multivariate Gaussian random variables.

Definition 2.3. *A real-valued random variable X is said to be **Gaussian** if it has the density function*

$$\phi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

for some $m \in \mathbb{R}$ and $\sigma > 0$. We write this as $X \sim \mathcal{N}(m, \sigma)$.

*An $\mathbb{R}^d$- valued random variable $\boldsymbol{X} = (X_1, ..., X_d)$ is said to be **multivariate Gaussian** if, for every $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_d) \in \mathbb{R}^d$, the real-valued variable $\langle \boldsymbol{\alpha}, \boldsymbol{X} \rangle = \sum_{i=1}^d \alpha_i X_i$ is a Gaussian. In this case there exists a mean vector $\boldsymbol{m} \in \mathbb{R}^d$ with $m_j = E[X_j]$ and a non-negative definite $d \times d$ covariance matrix Σ, such that the probability density of $\boldsymbol{X}$ is given by*

$$\phi(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{m})\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{m})^T\right), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

where $|\Sigma| = det(\Sigma)$ is the determinant of Σ. We write this as $\boldsymbol{X} \sim \mathcal{N}_d(\boldsymbol{m}, \Sigma)$.

We can now introduce Gaussian processes or Gaussian random fields.

Definition 2.4. *A **Gaussian process** or a **Gaussian random field** is a stochastic process or random field $\boldsymbol{X} = \{X(\boldsymbol{m}) : \boldsymbol{m} \in M\}$ on a set M, for which the (finite dimensional) distributions of $(X(m_1), ..., X(m_k))$ are multivariate Gaussian for each $1 \leq k < \infty$ and each $(m_1, ..., m_k) \in M^k$.*

The most famous Gaussian process is the Brownian motion. We want to introduce it in one dimension. For statements to Brownian we refer to Klenke [34], Chapter 21 and Chapter 22.

Definition 2.5. *A stochastic Process $B = \{B(t), t \in \mathbb{R}_0^+\}$ is called a Brownian motion if the four facts hold:*

- $B(0) = 0$
- *$B(t)$ is almost surely continuous*
- *$B(t)$ has independent increments*
- *$B(t) - B(s) \sim \mathcal{N}(0, t-s)$ for $0 \leq s \leq t$.*

For later convergence statements we need an almost sure upper bound for the Brownian motion. With Blumentahl's 0-1 law, one can easily show that

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{t}} = \infty \quad a.s..$$

The aim of the next theorem is to find an alternative function to $\sqrt{t}$ to get a almost sure finite limes superior.

Theorem 2.6 (Law of the iterated logarithm for Brownian motion)**.** *Let $B(t)$ be a Brownian motion. Then*

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t\log(\log(t))}} = 1 \quad a.s.$$

Proof: *For detailed proof see Klenke [34], Theorem 22.1* □

Since multivariate Gaussian distributions are determined by means and covariances, it is immediate that Gaussian random fields are determined by their mean and covariance functions. We will define these functions more general for stochastic processes in Section 2.1.2.

In our later simulations Gaussian processes will be one of the important components to get a Brown-Resnick process. These processes are part of a family of processes, the so called max-stable processes. We already introduced max-stable distribution functions in Chapter 1. Before we extend this theory to stochastic processes, we first need another property of stochastic processes.

Definition 2.7. *A stochastic process $\boldsymbol{X} = \{X(\boldsymbol{m}) : \boldsymbol{m} \in M\}$ has stationary increments if the law of*

$$\{X(\boldsymbol{m}+\boldsymbol{t}_0) - X(\boldsymbol{t}_0), \boldsymbol{m} \in \mathbb{R}^d\}$$

does not depend on the choice of $\boldsymbol{t}_0 \in \mathbb{R}^d$.

In other words, a stochastic process has stationary increments when its joint probability distribution does not change when it gets shifted. As a consequence, parameters such as the mean and the variance don't change. We often just write stationary process or stationary field.

Definition 2.8. *Let $\boldsymbol{X}$ be a stationary stochastic process and $\boldsymbol{X}^i = \{X^{(i)}(\boldsymbol{m}) : \boldsymbol{m} \in M\}$ for $i = 1, ..., n$ be independent copies of $\boldsymbol{X}$. $\boldsymbol{X}$ is called max-stable, if the process $\boldsymbol{M}^n = \{M^{(n)}(\boldsymbol{m}) : \boldsymbol{m} \in M\}$, which is defined by*

$$M^{(n)}(\boldsymbol{m}) := \bigvee_{i=1}^{n} X^{(i)}(\boldsymbol{m}), \quad \boldsymbol{m} \in \mathbb{R}^d$$

satisfies the following condition: There exist sequences $a^{(n)}(\boldsymbol{m})$ and $b^{(n)}(\boldsymbol{m})$ for $n \in \mathbb{N}$ with $a^{(n)}(\boldsymbol{m}) > 0$ for all n such that for all n it holds that,

$$\frac{M^{(n)}(\boldsymbol{m}) - b^{(n)}(\boldsymbol{m})}{a^{(n)}(\boldsymbol{m})} \underset{=}{\mathcal{D}} X(\boldsymbol{m}), \quad \boldsymbol{m} \in \mathbb{R}^d,$$

where $\mathcal{D}$ means equal in distribution.

As in the case for multivariate max-stable random variables the finite-dimensional margins of a max-stable process are max-stable, such that the margins

$$P\left(X(\boldsymbol{m}) \le x\right) = \exp\left(-\frac{1}{x}\right), \quad x > 0,$$

are Fréchet distributed or

$$P\left(X(\boldsymbol{m}) \leq x\right) = \exp\left(-\exp\left(-x\right)\right), \quad x \in \mathbb{R},$$

Gumbel distributed. Another possibility would be, that the margins are Weibull distributed, but we won't consider this case. One way to prove that a stochastic process is max-stable, was presented in de Haan [16],

$$\begin{aligned}
&P\left[X(\boldsymbol{m}_1) \leq x_1, ..., X(\boldsymbol{m}_k) \leq x_k\right] \\
&= P\left[\frac{M^{(n)}(\boldsymbol{m}_1)}{n} \leq x_1, ..., \frac{M^{(n)}(\boldsymbol{m}_k)}{n} \leq x_k\right] \\
&= P\left[\frac{1}{n}X^{(1)}(\boldsymbol{m}_1) \leq x_1, ..., \frac{1}{n}X^{(n)}(\boldsymbol{m}_1) \leq x_1, ..., \frac{1}{n}X^{(1)}(\boldsymbol{m}_k) \leq x_k, ..., \frac{1}{n}X^{(n)}(\boldsymbol{m}_k) \leq x_k\right] \\
&= P\left[\frac{1}{n}X^{(1)}(\boldsymbol{m}_1) \leq x_1, ..., \frac{1}{n}X^{(1)}(\boldsymbol{m}_k) \leq x_k, ..., \frac{1}{n}X^{(n)}(\boldsymbol{m}_1) \leq x_1, ..., \frac{1}{n}X^{(n)}(\boldsymbol{m}_k) \leq x_k\right] \\
&= \prod_{i=1}^{n} P\left[X^{(i)}(\boldsymbol{m}_1) \leq nx_1, ..., X^{(i)}(\boldsymbol{m}_k) \leq nx_k\right] \\
&= P^n\left[X(\boldsymbol{m}_1) \leq nx_1, ..., X(\boldsymbol{m}_k) \leq nx_k\right]
\end{aligned}$$

2.1.2 Properties of Stochastic Processes

Definition 2.9. *Let $\boldsymbol{X} = \{X(\boldsymbol{m}) : \boldsymbol{m} \in M\}$ be a stochastic process or a random field with parameter set M, then its mean function is defined by*

$$m(\boldsymbol{t}) = E\left[X(\boldsymbol{t})\right], \quad \boldsymbol{t} \in M$$

and its covariance function by

$$C(\boldsymbol{x}, \boldsymbol{y}) = E\left[\left(X(\boldsymbol{x}) - m(\boldsymbol{x})\right)\left(X(\boldsymbol{y}) - m(\boldsymbol{y})\right)\right] = Cov\left(X(\boldsymbol{x}), X(\boldsymbol{y})\right), \quad \boldsymbol{x}, \boldsymbol{y} \in M.$$

Remark 2.10. *As for multivariate distribution the correlation matrix, there exist a correlation function $\tilde{C}$ for stochastic processes, which is defined by*

$$\tilde{C}(\boldsymbol{x}, \boldsymbol{y}) = \frac{C(\boldsymbol{x}, \boldsymbol{y})}{\sqrt{Var(X(\boldsymbol{x})) \cdot Var(X(\boldsymbol{y}))}}.$$

Since we concentrate most of the time on space-time data, and the correlation function $\tilde{C}$ plays an important role in the connection between Gaussian space-time processes and max-stable processes, we split the paramter set M again into a space part S and a time dimension T as in Definition 2.2. For this part we stick close to Davis et al. [15].

Remark 2.11. *Let $M = S \times T$. For the correlation function of a space-time process $\boldsymbol{Z} = \{Z(\boldsymbol{s}, t) : \boldsymbol{s} \in S, t \in T\}$ the arguments of the correlation function $\tilde{C}$ split into*

$$\tilde{C}(\boldsymbol{x}, \boldsymbol{y}) \Longleftrightarrow \tilde{C}\left(\boldsymbol{s}_1, t_1; \boldsymbol{s}_2, t_2\right), \quad \boldsymbol{x} = (\boldsymbol{s}_1, t_1), \boldsymbol{y} = (\boldsymbol{s}_2, t_2) \in M.$$

With this alteration of the correlation function for space-time processes the question arises, if there are two dependence structures, one between different spatial locations, and another one between different time points. In later chapters we will discuss different types of correlation functions, where the correlation function can depend not only on the distance between two spatial locations, but also in a more complex way. But to begin with, some properties of space-time processes will be described next.

Definition 2.12. *A space-time Process* $\boldsymbol{Z} = \{Z(\boldsymbol{s}, t) : \boldsymbol{s} \in S, t \in T\}$ *with correlation function* $\tilde{C}$ *is called*

- ***stationary*** *if, its correlation function* $\tilde{C}$ *depends on the spatial lag* $\boldsymbol{h} := \boldsymbol{s}_1 - \boldsymbol{s}_2$ *and the temporal lag* $u := t_1 - t_2$ *such that*

$$\tilde{C}\left(\boldsymbol{s}_1, t_1; \boldsymbol{s}_2, t_2\right) = \tilde{C}\left(\boldsymbol{s}_1 - \boldsymbol{s}_2, t_1 - t_2; 0, 0\right) =: \gamma(\boldsymbol{h}, u)$$

  for all $\boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathbb{R}^d$ *and* $t_1, t_2 \in [0, \infty)$.

- ***isotropic*** *if, its correlation function* $\tilde{C}$ *only depends on the absolute spatial lag* $\|h\|$ *and the absolute temporal lag* $|u|$ *in such a way that that there exists another correlation function* $\tilde{\gamma}$ *that satisfies*

$$\gamma\left(\boldsymbol{h}, u\right) = \tilde{\gamma}\left(\|\boldsymbol{h}\|, |u|\right), \quad \boldsymbol{h} \in \mathbb{R}^d, u \in \mathbb{R}.$$

- ***anisotropic*** *if nit isotropic.*

In spatial statistics, not only the covariance function plays an important role to describe spatial dependence of a random field or a stochastic process, but also the so called **variogram** or **semivariogram**.

Definition 2.13. *The* ***variogram*** 2δ *is defined as the variance of the difference of two values of the space-time process* $\boldsymbol{Z} = \{Z(\boldsymbol{s}, t) : \boldsymbol{s} \in S, t \in T\}$*:*

$$\begin{aligned} 2\delta(\boldsymbol{s}_1, t_1; \boldsymbol{s}_2, t_2) &= Var\left[Z(\boldsymbol{s}_1, t_1) - Z(\boldsymbol{s}_2, t_2)\right] \\ &= E\left[\left(Z(\boldsymbol{s}_1, t_1) - Z(\boldsymbol{s}_2, t_2)\right)^2\right], \end{aligned}$$

with $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S, t_1, t_2 \in T$ *. If we just consider*

$$\delta(\boldsymbol{s}_1, t_1; \boldsymbol{s}_2, t_2) = \frac{1}{2} Var\left[Z(\boldsymbol{s}_1, t_1) - Z(\boldsymbol{s}_2, t_2)\right],$$

we speak of the ***semi-variogram****.*

If we don't look at space-time processes, but on general dimensions and processes with stationary increments, we can and mostly use the following notation:

$$2\delta(\boldsymbol{h}) = Var\left[X(\boldsymbol{t}_1) - X(\boldsymbol{t}_2)\right], \quad \text{where} \quad \boldsymbol{h} = \boldsymbol{t}_1 - \boldsymbol{t}_2$$

with $\boldsymbol{t}_1, \boldsymbol{t}_2 \in \mathbb{R}^d$or just

$$2\delta(\boldsymbol{t}) = Var\left[X(\boldsymbol{t}) - X(0)\right], \ \boldsymbol{t} \in \mathbb{R}^d$$

With the introduction of the variogram, there are some examples of isotropic variogram model to look at. Various parametric variogram models are presented in Journel and Huijbregts [31]. In the following, we concentrate on some basic isotropic models, given in terms of the semivariogram.

Example 2.14 (see Cressie [13], page 61)**.**

- *Linear model (valid in $\mathbb{R}^d$):*

$$\delta\left(\boldsymbol{h}; c_o, b_l\right) = \begin{cases} 0, & \boldsymbol{h} = \boldsymbol{0}, \\ c_0 + b_l \left\|\boldsymbol{h}\right\|, & \boldsymbol{h} \neq \boldsymbol{0} \end{cases}$$

  with $c_0 \geq 0$ and $b_l \geq 0$.

- *Spherical model (valid in $\mathbb{R}^d$, $d \leq 3$):*

$$\delta\left(\boldsymbol{h}, c_0, c_s, a_s\right) = \begin{cases} 0, & \boldsymbol{h} = \boldsymbol{0}, \\ c_0 + c_s \left[\frac{3}{2}(\left\|\boldsymbol{h}\right\| / a_s) - \frac{1}{2}(\left\|\boldsymbol{h}\right\| / a_s)^3\right], & 0 < \left\|\boldsymbol{h}\right\| \leq a_s, \\ c_0 + c_s, & \left\|\boldsymbol{h}\right\| \geq a_s, \end{cases}$$

  with $c_0 \geq 0, c_s \geq 0, a_s \geq 0$.

- *Exponential model (valid in $\mathbb{R}^d$):*

$$\delta\left(\boldsymbol{h}; c_0, c_e, a_e\right) = \begin{cases} 0, & \boldsymbol{h} = \boldsymbol{0}, \\ c_0 + c_e \left[1 - \exp\left(-\left\|\boldsymbol{h}\right\| / a_e\right)\right], & \boldsymbol{h} \neq \boldsymbol{0}, \end{cases}$$

  with $c_0 \geq 0, c_e \geq 0, a_e \geq 0$.

Cressie [13] describes in Chapter 2.3.2 a relationship between the covariance function and the variogram of a stochastic process $\boldsymbol{X}$. We could introduce this relationship for stationary processes, but for this we only need a more weak form.

Definition 2.15 (see for example Cressie [13], page 53)**.** *A stochastic process $\boldsymbol{X} = \{X(\boldsymbol{s}, t) : \boldsymbol{s} \in M, t \in T\}$ has second-order stationarity or weak-sense stationarity if it has the follwoing restriction on its mean and its covariance function:*

- $E\left[X(\boldsymbol{s}, t)\right] = E\left[X(\boldsymbol{s} + \boldsymbol{s}_o, t + t_0)\right]$, *for all* $\boldsymbol{s}_0 \in \mathbb{R}^d, t_0 \in [0, \infty)$
- $Cov\left(X(\boldsymbol{s}_1, t_1), X(\boldsymbol{s}_2, t_2)\right) = Cov\left(X(\boldsymbol{s}_1 - \boldsymbol{s}_2, t_1 - t_2), 0\right)$

So only the first moment and the autocovariance function have not to vary, regarding a shift.

Remark 2.16. *If $\boldsymbol{X}(\cdot)$ is second-order stationary then we can express its variogram 2δ with its covariance function C as follow:*

$$2\delta(\boldsymbol{h}) = 2\left(C(\boldsymbol{0} - C(\boldsymbol{h})\right)$$

2.2 Self-similar Processes

This section about self similar processes and especially the fractional Brownian motion is based on the book by Samoradnitsky and Taqqu [42], chapter 7.

Self similar processes are considered in many applications such as in physics, hydrology or economics and play an important role, if we are talking about **long-range dependence** or **long memory**.

We are now talking about processes only depending on one dimension. Furthermore, we introduce a positive number denoted by H.

Definition 2.17 (see for instance Samoradnitsky and Taqqu [42], Definition 7.1.1)**.** *Consider a process $X = \{X(t), t \in \mathbb{R}\}$. X is called self-similar with index $H > 0$, if its finite dimensional have scaling properties, namely for any $a \in \mathbb{R}$, the process*

$$\{X(at), t \in \mathbb{R}\}$$

has the same finite dimensional distribution as

$$\{a^H X(t), t \in \mathbb{R}\}, \quad H > 0.$$

In other words, consider $t_1, ..., t_n \in \mathbb{R}$ and $x_1, ..., x_n \in \mathbb{R}$, then

$$P\left(X(at_1) \leq x_1, ..., X(at_n) \leq x_n\right) = P\left(a^H X(t_1) \leq x_1, ..., a^H X(t_n) \leq x_n\right)$$

The nonnegative number H has various names. It is called the exponent or the index of self-similarity but mostly the „Hurst exponent“.

Definition 2.18. *The process $X = \{X(t), t \in \mathbb{R}\}$ is H-sssi if it is self-similar with Hurst index H and has stationary increments*

Next we introduce the definition of long-range dependence

Definition 2.19. *A stationary discrete time process $(X(t))_{t\in\mathbb{Z}}$ is said to be a process with long-range dependence when its autocovariance function $\kappa(\cdot)$ decays so slowly that*

$$\sum_{s=0}^{\infty} \kappa(s) = \infty$$

where $\kappa(s) = Cov\left(X_n, X_{n+s}\right)$.

So the **autocovariance function** of a stochastic process, coincides with the introduced convariance function in Definition 3.9.

We now want to get to the H-sssi process, which we will mainly focus on. We stick close to Samoradnitsky and Taqqu [42], page 318. Recall that a random variable $X \sim \mathcal{N}(0, \sigma^2)$ has characteristic function

$$E\left[\exp(ibX)\right] = \exp\left(-\frac{1}{2}\sigma^2 b^2\right), \quad b \in \mathbb{R}.$$

The finite dimensional distributions of a Gaussian Process $\{X(t), t \in \mathbb{R}\}$ satisfy

$$E\left[exp\left(i\sum_{j=1}^{m} b_j X(t_j)\right)\right] = \exp\left[-\frac{1}{2}\sum_{j=1}^{m}\sum_{k=1}^{m} A(t_j,t_k)b_j b_j + \sum_{j=1}^{m}\mu(t_j)b_j\right],$$

with $b_1, ..., b_m \in \mathbb{R}$ and $\mu(t), t \in \mathbb{R}$ is a real valued function. Furthermore, $\{A(t_i,t_j),\ t_i,t_j \in \mathbb{R}\}$ is non-negative definite. So to each μ and A belongs a Gaussian process, where μ is its expectation function and A its autocovariance fucntion.

We choose now

$$\left\{A(t_i,t_j) = |t_i|^{2H} + |t_j|^{2H} - |t_i - t_j|^{2H},\ t_i,t_j \in \mathbb{R}, 0 < H \leq 1\right\}, \tag{2.2}$$

which is non-negative definite (for a proof se Lemma 2.10.8 in Samoradnitsky and Taqqu [42]).

Definition 2.20. *Since (2.2) is non-negative definite, there exists a Gaussian process $\{X(t), t \in \mathbb{R}\}$ with mean zero and autocovariance function*

$$Cov\left(X(t_1), X(t_2)\right) = \frac{1}{2}\left[|t_1|^{2H} + |t_2|^{2H} - |t_1 - t_2|^{2H}\right] Var\left(X(1)\right), \quad 0 < H \leq 1.$$

This process is called fractional Brownian motion and is denoted by $\{B_H(t), t \in \mathbb{R}\}$. It is called standard fractional Brownian motion if $Var(X(1)) = 1$.

Furthermore, this process is H-sssi.

Example 2.21. *Every Brownian motion $\{B(t), t \in \mathbb{R}\}$ is also a fractional Brownian motion $\{B_H(t), t \in \mathbb{R}\}$, with $H = \frac{1}{2}$. The autocovariance function of $\{B_{\frac{1}{2}}(t), t \in \mathbb{R}\}$ reduces to*

$$E\left[B_{\frac{1}{2}}(t_1)B_{\frac{1}{2}}(t_2)\right] = \begin{cases} Var(B_{\frac{1}{2}}(t))\min(t_1,t_2) & \text{if } t_1 \text{ and } t_2 \text{ have the same signs,} \\ 0 & \text{if } t_1 \text{ and } t_2 \text{ have opposite signs} \end{cases}$$

which is the autocovariance function of a Brownian motion.

For our applications in further chapters especially the case $H = 1$ will be important.

Lemma 2.22 (see therefore Samoradnitsky and Taqqu [42], Lemma 7.2.1)**.** *Consider a fractional Brownian motion $\{B_H(t), t \in \mathbb{R}\}$, where $B_H(0) = 0$. In the case $H = 1$,*

$$B_H(t) = tB_H(1) \quad a.s.$$

Proof: *If $H = 1$ and $B_1(0) = 0$ a.s., then we can claim for the autocovariacne function*

$$Cov\left(B_1(t_1), B_1(t_2)\right) = E\left[B_1(t_1)B_1(t_2)\right] = t_1 t_2 E\left[B_1^2(1)\right]$$

With this in mind we can conclude

$$\begin{aligned} E\left[B_1(t) - tB_1(1)\right]^2 &= E\left[B_1^2(t)\right] - 2tE\left[B_1(t)B_1(1)\right] + t^2 E\left[B_1^2(1)\right] \\ &= \left(t^2 - 2t\cdot t + t^2\right) E\left[B_1^2(1)\right] \\ &= 0, \end{aligned}$$

which is nothing else then $X(t) - tX(1) = 0$ a.s. for all t. $\square$

We not only can determine the fractional Brownian motion by its covariance function, but there also exist several stochastic integral representations.

Theorem 2.23 (see for instance Samoradnitsky and Taqqu [42], Proposition 7.2.6)**.** *The following process is a fractional Brownian motion with Hurst parameter $H \in (0,1)$*

$$\begin{aligned} B_H(t) &= \frac{1}{\Gamma(H+1/2)} \int_{\mathbb{R}} \left[(t-s)_+^{H-1/2} - (-s)_+^{H-1/2} \right] dW_s \\ &= \frac{1}{\Gamma(H+\frac{1}{2})} \left[\int_{-\infty}^{0} \left[(t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}} \right] dW_s + \int_0^t (t-s)^{H-\frac{1}{2}} dW_s \right] \end{aligned}$$

In further chapters we are interested in the distribution of the supremum of a fractional Brownian motion. In the following theorem we want to give a upper bound for the tail probability. See therefore Vardar [46] Theorem 2.2.

Theorem 2.24. *Consider a fractional Brownian motion $\{B_H(t), t \in \mathbb{R}\}$ with Hurst parameter $H > \frac{1}{2}$ up to a time point a, then*

$$P\left(\sup_{0 \le t \le a} B_H(t) \ge x \right) = \frac{\sqrt{2} a^H}{x\sqrt{\pi}}$$

As for Brownian motions we also want to give a law of the iterated logarithm for fractional Brownian motions. There are many approaches in the literature. Barka and Mountford [3] present a law of the iterated logarithm for short term behaviour, by considering the local time of a fractional Brownian motion. In a similar way Chen et al. [9] derives laws of the iterated logarithm as a consequence of their estimate calculations for large deviations of local times of fractional Brownian motions. We continue with the book Fractional Fields and Applications by Cohen and Istas [10], which gives an excellent summary for properties of the fractional Brownian motions.

Theorem 2.25 (see Theorem 3.2.4 by Cohen and Istas [10])**.** *Let $(B_H(t))_{t \ge 0}$ be a fractional Brownian motion, then for all $t \in \mathbb{R}$*

$$\limsup_{\epsilon \to 0^+} \frac{B_H(t+\epsilon) - B_H(t)}{\sqrt{2\epsilon^{2H} \log\left(\log(\frac{1}{\epsilon})\right)}} = 1 \quad a.s. \tag{2.3}$$

We want to give a similar limes superior result as in Theorem 2.6. So we first „assume“ another way to display (2.3).

Theorem 2.26. *Let $(B_H(t))_{t \ge 0}$ be a fractional Brownian motion, then for all $t \in \mathbb{R}$*

$$\limsup_{t \to \infty} \frac{B_H(t)}{\sqrt{2t^{2H} \log\left(\log(t)\right)}} = 1 \quad a.s.$$

Before we proof this theorem we need the following Lemma. We do a similar approach as Klenke [34] does in his proof of Corollary 22.3, but we do it the other way around.

Lemma 2.27. *Let $(B_H(t))_{t \ge 0}$ be a fractional Brownian motion. The stochastic process*

$$X(t) = \begin{cases} t^{2H} B_H(\frac{1}{t}), & if\ t > 0 \\ 0, & if\ t = 0 \end{cases}$$

is again a fractional Brownian motion.

Proof: *First we have to show that* $(X(t))_{t\geq 0}$ *is characterized by covariance function*

$$Cov\left(X(t),X(s)\right)=|t|^{2H}+|s|^{2H}-|t-s|^{2H}\quad t,s>0,$$

as a fractional Brownian motion. We do the calculation:

$$\begin{aligned}Cov\left(X(t),X(s)\right)&=Cov\left(t^{2H}B_H(\frac{1}{t}),s^{2H}X(\frac{1}{s})\right)\\&=t^{2H}s^{2H}Cov\left(B_H(\frac{1}{t}),B_H(\frac{1}{s})\right)\\&=t^{2H}s^{2H}\left[|\frac{1}{t}|^{2H}+|\frac{1}{s}|^{2H}-|\frac{1}{t}-\frac{1}{s}|^{2H}\right]\\&=|t|^{2H}+|s|^{2H}-|t-s|^{2H}\quad t,s>0.\end{aligned}$$

The proof for Hoelder continuity is similiar to Klenke [34], proof of theorem 21.14 □

We now need this fractional Brownian motion $(X(t))_{\geq 0}$ to proove Theorem 2.24.

Proof of Theorem 2.24: *The first step is to use Theorem 2.24 with* $(X(t))_{\geq 0}$*. With some reformulation we get from Theorem 2.24 to Theorem 2.23 in the follwing way:*

$$\begin{aligned}\limsup_{t\to\infty}\frac{t^{2H}B_H(\frac{1}{t})}{\sqrt{2t^{2H}\log\left(\log(t)\right)}}&=_{\epsilon:=1/t}\limsup_{\epsilon\to 0^+}\frac{\epsilon^{-2H}B_H(\epsilon)}{\sqrt{2\epsilon^{-2H}\log\left(\log(\frac{1}{\epsilon})\right)}}\\&=\limsup_{\epsilon\to 0^+}\frac{B_H(\epsilon)}{\epsilon^{2H}\sqrt{2\epsilon^{-2H}\log\left(\log(\frac{1}{\epsilon})\right)}}\\&=\limsup_{\epsilon\to 0^+}\frac{B_H(\epsilon)}{\sqrt{2\epsilon^{2H}\log\left(\log(\frac{1}{\epsilon})\right)}}\end{aligned}$$

The last term is exactly the one from Theorem 2.23 when we assume, w.l.o.g that $t=0$ *since it holds* $\forall\ t\geq 0$*.* □

In Figure 2.1 we can see 4 draws of fractional Brownian motion paths with different Hurst parameters. The bigger H gets, the smoother the paths become. For $H=1$ we see a linear path, as stated in Lemma 2.22. Little deviations from a linear path in Figure 2.1 are due to rounding.

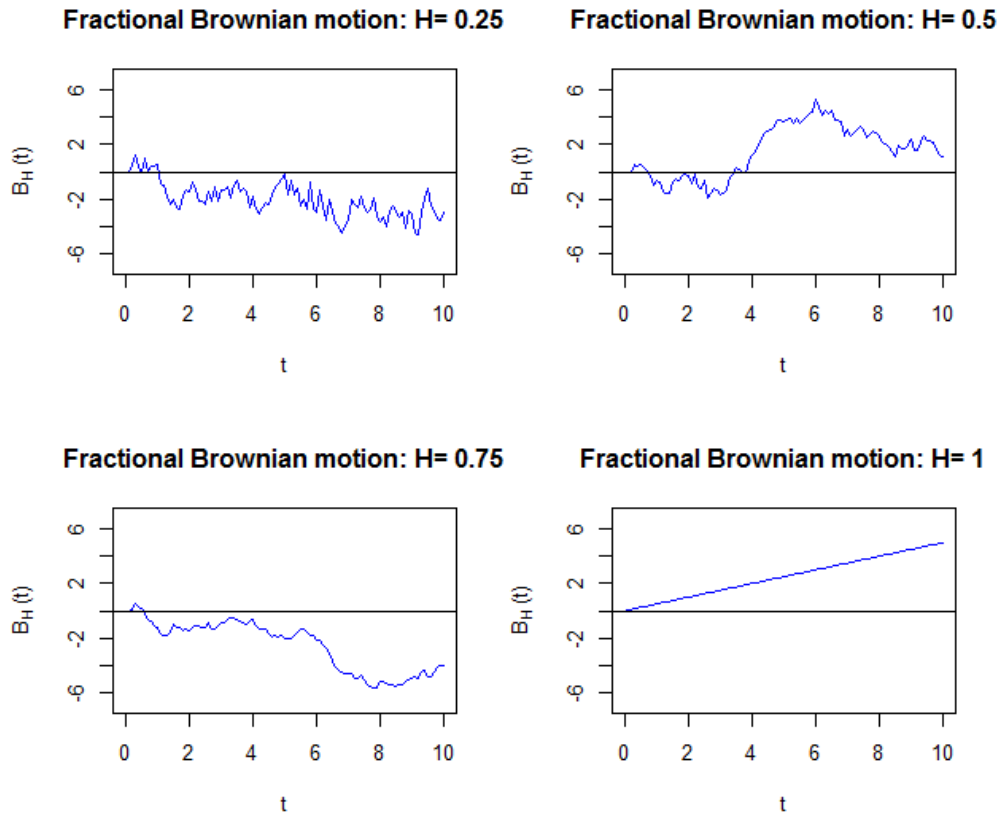

Figure 2.1: Fractional Brownian motions with Hurst parameters $H = 0.25$, $H = 0.5$, $H = 0.75$ and $H = 1$.

Self-similarity for Multidimensional Fields

Self-similar processes on d-dimensional spaces, with $d > 1$ are more difficult to study.

We already defined stationary increments in arbitrary dimension, so we start first with self-similarity in d dimensions.

Definition 2.28 (see for example Cohen and Istas [10], page 80)**.** *Consider a field* $\boldsymbol{X} = \{X(\boldsymbol{t}), t \in \mathbb{R}^d\}$. $\boldsymbol{X}$ *is called self-similar with Hurst parameter* $0 < H \leq 1$ *if*

$$\{X(a\boldsymbol{t}),\ \boldsymbol{t} \in \mathbb{R}^d\}$$

has the same distribution as

$$\{a^H X(\boldsymbol{t}),\ \boldsymbol{t} \in \mathbb{R}^d\}.$$

Since the definitions for stationary increments and self-similarity are clearly the same if we extend the dimensions, we also speak of H-ssi fields. For simulation of Brown-Resnick fields we need fractional Brownian fields, which are defined as follows.

Definition 2.29 (see for instance Cohen and Istas [10], page 81)**.** *The fractional Brownian field* $(\boldsymbol{B}_H(\boldsymbol{t}))_{\boldsymbol{t} \in \mathbb{R}^d_+}$ *is a Gaussian field with mean zero and autocovariance function*

$$Cov\left(B_H(\boldsymbol{t}), B_H(\boldsymbol{s})\right) = \frac{V}{2}\left[\|\boldsymbol{t}\|^{2H} + \|\boldsymbol{s}\|^{2H} - \|\boldsymbol{t} - \boldsymbol{s}\|^{2H}\right]$$

, *where* $V > 0$.

We see that we can transfer the basic results of fractional Brownian motions easily to multidimensional Fractional Brownian fields. For further statements to self-similar multidimensional fields, see Chapter 3.3 in Cohen and Istas [10].

As in the one-dimensional case, we plotted several realisations of fractional Brownian fields. In Figure 2.2, we can see four of them with different Hurst parameters. As for one dimension, the bigger H gets, the smoother the fields look. For the margin case $H = 1$ we even get a plane, which would be the counterpart of the linear path in Figure 2.1.

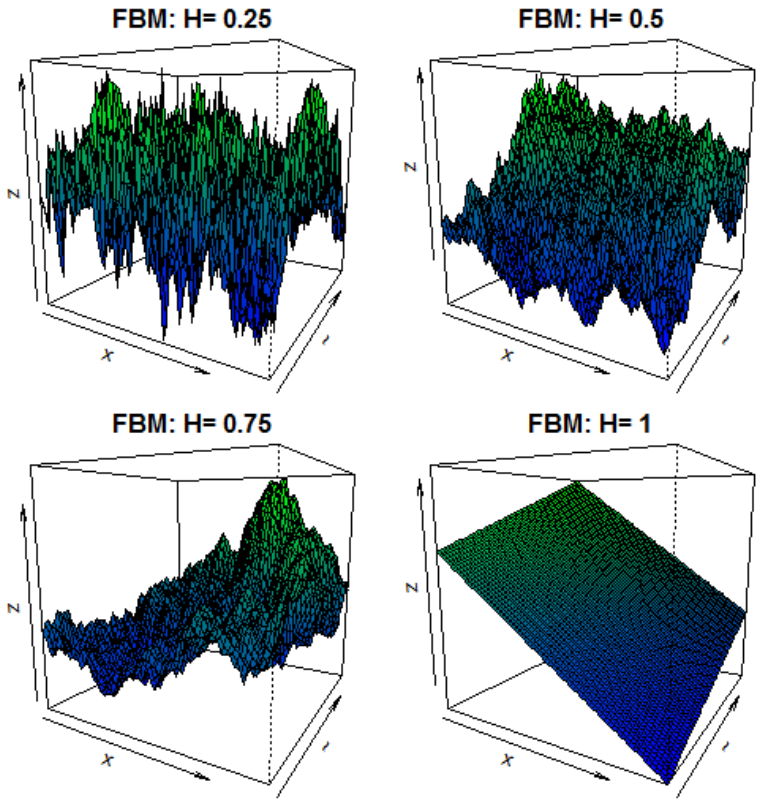

Figure 2.2: Fractional Brownian motions with Hurst parameters $H = 0.25$, $H = 0.5$, $H = 0.75$ and $H = 1$.

2.3 Brown-Resnick Processes

We now come to a special class of stationary max-stable processes. As we described at the beginning of Chapter 2, one can use Brown-Resnick processes for example to modell extrem rain-data. These processes also have been fitted to time series, spatial and latest also to space-time data. Brown-Resnick processes have been first introduced in 1977 by Brown and Resnick [6]. They only dealt in their paper with Brownian motions for the underlying Gaussian process. Kabluchko et al. [32] extend this concept to more generalized

Gaussian process. This generalization is often called „Brown-Resnick process associated to the variogram δ “. This paper is the more relevant for us.

A Brown-Resnick process is one of the most common max-stable process. Other popular examples are for instance the extremal-t process - see for instance Dombry et al. [23], chapter 5.3 - which also contains Gaussian processes or Smith's storm profile model, which was introduced by de Haan [16] in 1984 and extended by Smith [43] in 1990.

Since most of the time we deal with space-time simulations, we want to introduce a space-time version of a Brown-Resnick processes.

Definition 2.30. *Let $\{\xi_j, j \geq 1\}$ be points of a Poisson random measure on $\mathbb{R}_0^+$ with intensity measure $V^{-2}dV$ and let $Y_j(\boldsymbol{s},t), j = 1,2,...$ be independent copies of some nonnegative random field $\boldsymbol{Y} = \{Y(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\}$, i.e. $Y(\boldsymbol{s},t) > 0$ with finite mean function $E[Y(\boldsymbol{s},t)] < \infty$, which are independent of ξ_j. We call*

$$\eta(\boldsymbol{s},t) = \bigvee_{j=1}^{\infty} V_j Y_j(\boldsymbol{s},t), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+ \tag{2.4}$$

a Brown-Resnick process with Fréchet marginals.

Remark 2.31. *For the simulation in Chapter 3 we consider for the nonnegative random field $\boldsymbol{Y} = \{Y(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\}$ the following, such that (2.4) results in*

$$\eta(\boldsymbol{s},t) = \bigvee_{j=1}^{\infty} V_j \exp\left(W_j(\boldsymbol{s},t) - \delta(\boldsymbol{s},t)\right), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+, \tag{2.5}$$

where the W_i's are fractional Brownian fields introduced in Definition 2.29 and δ its corresponding semi-variogram.

At this point we also want to introduce a analytical representation of a bivariate distribution function of the Brown-Resnick process with Fréchet margins.

Remark 2.32. *The bivariate distribution function of the Brown-Resnick process, defined as in Remark 2.31, is given by*

$$\begin{aligned} F(y_1,y_2) &:= P\left(\eta(\boldsymbol{s}_1,t_1 \leq y_1, \eta(\boldsymbol{s}_2,t_2) \leq y_2\right) \\ &= \exp\left[-y_1^{-1}\Phi\left(\frac{\log(y_2/y_1)}{2\sqrt{\delta(\boldsymbol{h},u)}} + \sqrt{\delta(\boldsymbol{h},u)}\right) - y_2^{-1}\Phi\left(\frac{\log(y_1/y_2)}{2\sqrt{\delta(\boldsymbol{h},u)}} + \sqrt{\delta(\boldsymbol{h},u)}\right)\right] \end{aligned}$$

For $(\boldsymbol{s}_1,t_1)$ and $(\boldsymbol{s}_2,t_2) \in \mathbb{R}^d \times \mathbb{R}_0^+$,

It will be useful to check the exactness of our simulations. This can be done by comparing this theoretical bivariate distribution function to the empirical bivariate distribution function of the simulation.

Sometimes it is useful to consider Brown-Resnick processes with standard Gumbel margins. So we compute the distribution for this case.

$$\tilde{F}(y_1,y_2) := F\left(\exp(y_1),\exp(y_2)\right) = P\left(\eta(\boldsymbol{s}_1,t_1 \leq \exp(y_1), \eta(\boldsymbol{s}_2,t_2) \leq \exp(y_2)\right).$$

So we get

$$\tilde{F}(y_1, y_2) = \exp\left[-\exp(y)_1^{-1}\Phi\left(\frac{\log\left((\exp(y_2)/\exp(y_1))\right)}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right) - \exp(y)_2^{-1}\Phi\left(\frac{\log\left(\exp(y_1)/\exp(y_2)\right)}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right)\right]$$

$\Leftrightarrow$

$$\tilde{F}(y_1, y_2) = \exp\left[-\exp(y)_1^{-1}\Phi\left(\frac{y_2 - y_1}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right) - \exp(y)_2^{-1}\Phi\left(\frac{y_1 - y_2}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right)\right] \tag{2.6}$$

$\tilde{F}$ is the bivariate distribution function of a Brown-Resnick process with gumbel margins.

With a lot effort it is also possible to find an analytical form for a multivariate distribution function of a Brown-Resnick process ($d > 2$).

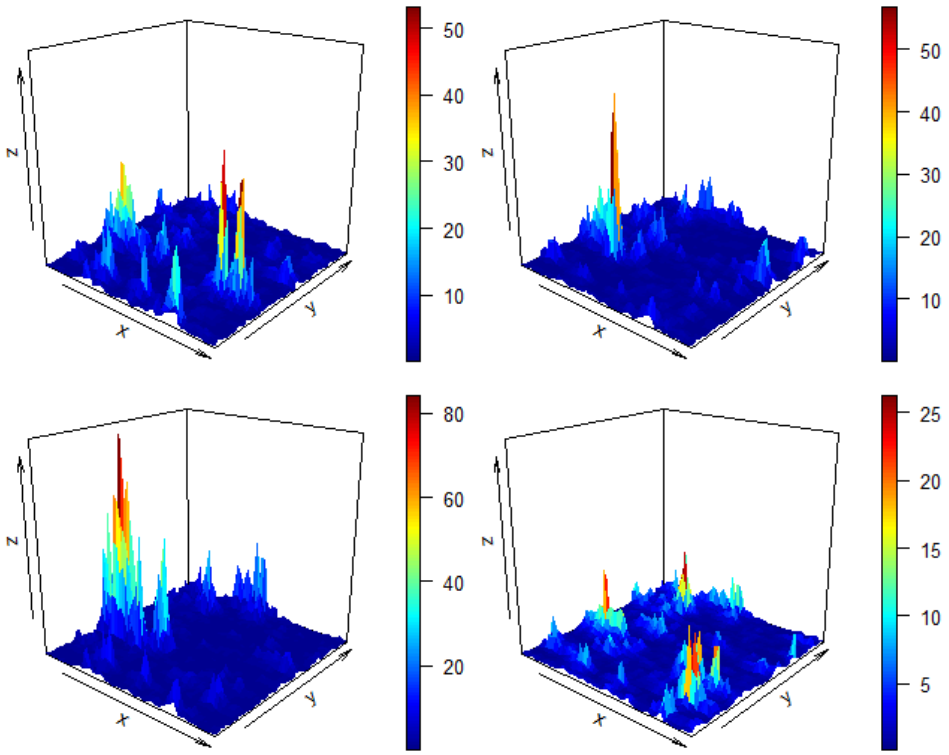

Figure 2.3: Anisotropic Brown-Resnick process with Fréchet margins at four consecutive time points

Chapter 3

Simulation Methods for Brown-Resnick-Processes

After all these introductions of the basics of extreme value theory and the essential background for max-stable processes, we now come to the more applied part of this thesis. We will introduce 4 different methods to simulate Brown-Resnick processes. Since they are all completely different in their execution, we always give a short summary of the theoretical part of these methods.

Despite the fact that these four ways cover a huge field of approaches to the simulation of Brown-Resnick processes, we recommend the reader to look at further computation methods. For instance, a conditional approach by Dombry et al. [22]. To sample from the regular conditional distribution of a Brown-Resnick process, they introduce a Markov Chain Monte-Carlo algorithm, to avoid a huge amount of computational effort.

Another example is Oesting et al. [37]. They give different methods, including random shifting and a mixed moving maxima representation which leads to approximations of Brown-Resnick processes.

We now present step by step our four different methods, where we immediately give the simulation results before we go to the next method.

3.1 Gaussian approximation method

Hüsler and Reis introduced in their paper [30] in 1989 a regularity condition of the correlation function of a Gaussian field, which created the foundation for a quite simpler way to simulate from our Brown-Resnick model. A simulation method was presented in Kabluchko et al. [32] in 2009, and extended by Davis et al. [15] in 2012 to space-time data.

Hüsler and Reis are also eponyms of the Hüsler-Reis distribution (see for example Engelke, Malinowski and Kabluchko [26] for a good summary), which coincides with the finite-dimensional distribution function of a Brown-Resnick process at given locations $(\boldsymbol{s}_1, t_1), ..., (\boldsymbol{s}_k, t_k)$, like we introduced in Remark 2.32 for the bivariate case.

To begin with we consider the case, where the Brown-Resnick process has Fréchet margins.

So we can take the Brown-Resnick process

$$\eta(\boldsymbol{s},t) = \bigvee_{j=1}^{\infty} V_j \exp\left(W_j(\boldsymbol{s},t) - \delta(\boldsymbol{s},t)\right), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+, \tag{3.1}$$

from Remark 2.31 with Fréchet margins.

3.1.1 General assumptions

Throughout this first presentation of a simulation method of a Brown-Resnick process, $\boldsymbol{Z} = \{Z(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\}$ be a stationary Gaussian process with

- mean function $m(\boldsymbol{s},t) = E\left[Z(\boldsymbol{s},t)\right] = 0$ for all $(\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+$
- variance $Var\left(Z(\boldsymbol{s},t)\right) = 1$ for all $(\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+$
- correlation function $\gamma = \gamma(\boldsymbol{s}_1,t_1;\boldsymbol{s}_2,t_2) = \gamma(\boldsymbol{h},u), \quad \boldsymbol{h} = \boldsymbol{s}_1 - \boldsymbol{s}_2, \quad u = t_1 - t_2$

Let V_j be the points of a Poisson random measure on $[0,\infty)$ with intensity $V^{-2}dV$. Since $Var\left(Z(\boldsymbol{s},t)\right) = 1$ in all points of our Gaussian process the correlation function is equal to the covariance function. The central assumption on the correlation function γ will be the following smoothness condition:

Assumption 3.1 (cf. Davis et al. [15], Ass 2.1)**.** *There exist nonnegative sequences* $(s_{n,m})_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$ *that satisfy* $s_{n,m} \to 0$ *and* $t_n \to 0$ *as* $n \to \infty$*, with* $m = 1,...,d$ *and a nonnegative function* δ *such that*

$$(\log(n))\left[1 - \gamma\left(s_{n,1}h_1,...,s_{n,d}h_d,t_nu\right)\right] \to \delta\left(h_1,h_2,...,h_d,u\right) > 0, \quad n \to \infty$$

where $\boldsymbol{h} = (h_1,...,h_d) \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}, u \in \mathbb{R}$.

With all these assumptions in mind we now look at a method to simulate the Brown-Resnick process presented as in (3.1).

3.1.2 Implementation

For the implementation we stick close to Davis et al. [15] who extended the spatial version by Kabluchko et al. [32] to space-time data.

Theorem 3.2. *Let* $Z_j(\boldsymbol{s},t), j = 1,2,...$ *be independent copies of a stationary Gaussian process for which all the assumptions introduced in section 3.1.1 hold. Additionally, assume the existence of a metric* D *on* $\mathbb{R}^d$ *such that the following condition holds:*

$$\delta\left(\boldsymbol{s}_1 - \boldsymbol{s}_2, t_1 - t_2\right) \leq D\left((\boldsymbol{s}_1,t_1),(\boldsymbol{s}_2 - t_2)\right)^2 \quad \textit{for all} \quad \boldsymbol{s}_1,\boldsymbol{s}_2 \in \mathbb{R}^d, t_1,t_2 \in [0,\infty)$$

Furthermore, set

$$\eta_n(\boldsymbol{s},t) = \frac{1}{n}\bigvee_{j=1}^{n} -\frac{1}{\log\left(\Phi(Z_j(s_{n,1}s_1,\ldots,s_{n,d}s_d,t_nt))\right)}, \quad \boldsymbol{s} \in \mathbb{R}^d, t \in [0,\infty) \tag{3.2}$$

and we get

$$\eta_n(\boldsymbol{s},t) \underset{\rightarrow}{\mathcal{L}} \eta(\boldsymbol{s},t) := \bigvee_{j=1}^{\infty} V_j Y_j(\boldsymbol{s},t), \quad n \to \infty \tag{3.3}$$

where the right side of the equation is the Brown-Resnick process introduced in (3.1), with $Y(\boldsymbol{s},t) := \exp\left[W(\boldsymbol{s},\boldsymbol{t}) - \delta(\boldsymbol{s},t)\right]$*, and* $\mathcal{L}$ *stands for weak convergence in* $C\left(\mathbb{R}^d \times \mathbb{R}_0^+\right)$*. The random field* $\boldsymbol{W} = \left\{W(\boldsymbol{s},t) : \boldsymbol{s} \in \mathbb{R}^d, t \in [0,\infty)\right\}$ *has mean function* $m(\boldsymbol{s},t) = 0$*, for all* $\boldsymbol{s} \in \mathbb{R}^d, t \in [0,\infty)$*. Furthermore, we can express its correlation function* γ *in terms of* δ *in the following way:*

$$\tilde{C}(\boldsymbol{s}_1,t_1;\boldsymbol{s}_2,t_2) = \delta(\boldsymbol{s}_1,t_1) + \delta(\boldsymbol{s}_2,t_2) - \delta(\boldsymbol{s}_1-\boldsymbol{s}_2,t_1-t_2)$$

So we can construct a max-stable random field as limits of normalized and rescaled pointwise maxima of Gaussian random fields, with having in mind that the underlying correlation function has to satisfy the regularity condition from Assumption 3.1.

After this presentation of the transformation from Theorem 3.2 we can implement this algorithm in **R**. We now want to state how we do this in **R**.

Remark 3.3. *This remark holds for all simulation methods in this thesis. For each of the four approaches, we have to deal with Gaussian fields. We use the package* ***RandomFields*** *by Schlather, Malinowski, Oesting et. al. (Version 3.0.62). We want to do the simulation with this specific semi-variogram for space-time processes:*

$$\delta(h_1,h_2,u) = C_1|h_1|^{\alpha_1} + C_2|h_2|^{\alpha_2} + C_3|u|^{\alpha_3}. \tag{3.4}$$

The package ***RandomFields*** *offers several models with a given covariance function, but only a few with a given variogram. The most suitable for us, is the model of the fractional Brownian motion,*

```
> RMfbm(alpha=alpha,Aniso=Aniso)
```

where the alpha argument matches with the Hurst parameter introduced in Definition 2.17, such that $\alpha = 2H$*. So* $\alpha \in (0,2]$*. With the argument Aniso, we hand over the parameters* C_1, C_2, C_3 *in a diagonal matrix, such that* $Aniso = diag(C_1, C_2, C_3)$*. To achieve the exact variogram, we have to choose* $\alpha = 2$*, which leads to the variogram*

$$\delta(h_1,h_2,u) = C_1|h_1|^2 + C_2|h_2|^2 + C_3|u|^2, \tag{3.5}$$

which doesn't give us much variability in the α_i*,* $i = 1,2,3$*, but will be enough for our simulations, since we are mainly interested in the exactness of the algorithms.*

Having this Remark 3.3 in mind we now come to the first method by Hüsler and Reis:

- We simulate 1000 independent Gaussian-space-time processes $Z_j(\boldsymbol{s},t), j = 1,...,1000$ with semi-variogram

  $$\delta(h_1,h_2,u) = C_1|h_1|^2 + C_2|h_2|^2 + C_3|u|^2,$$

  as described in Remark 3.3. We choose arbitrary for $C_1 = 0.02, C_2 = 0.06, C_3 = 0.04$. The space grid is quadratic and both time and space points are equidistant.

- We transform the simulated processes to standard Fréchet margings by setting

$$\tilde{Z}_j(\boldsymbol{s},t) := -\frac{1}{\log\left(\Phi\left(Z_j(\boldsymbol{s},t)\right)\right)}, \quad \boldsymbol{s} \in \{1,...,M\}^2,\ t \in \{1,...,T\}, j = 1,...,1000.$$

- Since we are still confronted with an infinite maximum over the Gaussian processes in the Gaussian approximation method, we only get an approximation of the Brown-Resnick paths. We take the pointwise maximum over the generations $\tilde{Z}_j(s_{n,1}s_1,\ldots,s_{n,d}s_d,t_nt)$ in space and time and divide by the number of simulated Gaussian space-time processes,

$$\eta\left(\boldsymbol{s},t\right) = \frac{1}{1000}\bigvee_{j=1}^{1000}\tilde{Z}_j(s_{n,1}s_1,\ldots,s_{n,d}s_d,t_nt),\ \boldsymbol{s} \in \{1,...,m\}^2, t \in \{1,...,T\}.$$

  For our procedure, we concentrate on a space-time grid with $10 \times 10 \times 5$ points and repeat this 100 times.

3.1.3 Results of the Gaussian approximation method

One advantage of the Gaussian approximation method is certainly the fact, that we can predict the simulation effort, because we choose the number of Gaussian fields n, which have to be generated at the beginning of our simulation. We do not want to go into details regarding speed, because this can differ a lot, depending which computer is used. More information on this topic is given in Chapter 4. But knowing how long the simulation of a single field takes time, we can approximate the total time amount of the Gaussian approximation method, because the time of the transformation into Fréchet margins can be neglected.

It is more interesting for us to check, if the algorithm output is max-stable. To begin with, one of the easiest way is to check if the univariate margins are max-stable. We pick out random locations, all at the same time point. We take the 100 simulations from these random locations and plot these empirical quantiles, after a log-transformation against the theoretical Gumbel quantiles. The log-transformation is necessary, since the simulated Brown-Resnick process with the Gaussian approximation method has Fréchet margins. With Gumbel margins we better display the result in a qq-plot, since the Gumbel quantiles are smaller than the Fréchet quantiles. The result for the Gaussian approximation method is found in Figure 3.1.

The red line in figure 3.1 stands for no deviation between the theoretical quantiles and the quantiles, which result from the simulation. The log-transformed univariate margins can be seen as standard Gumbel distributed. However, we note that the deviation increases with increasing quantiles.

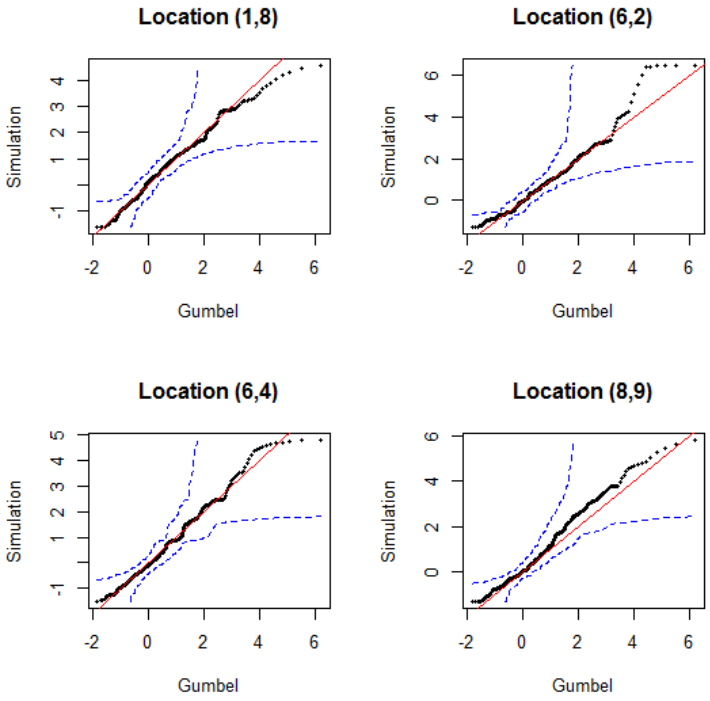

Figure 3.1: qq-plot of univariate margins of the log-transformed simulated process by Gaussian approximation method for a $10 \times 10 \times 5$-field versus the standard Gumbel distribution for four random space-time locations: (1,8,1), (6,2,1), (6,4,1), (8,9,1). The red line would be no deviation and the blue lines are the 95 % confidence bounds.

Now we want to introduce a more complex test for max-stability.

Max-stability-Test

Remark 3.4. *We use the method, which is also presented in Buhl and Klüppelberg [8] and applied there on real data, and goes back to Gabda et al. [27]. We denote the regular grid by*

$$S_M \times T_K = \{1, ..., M\}^d \times \{1, ..., K\}.$$

The hypothesis test is based in general on the simulation result $(\eta(\boldsymbol{s}, t))_{\boldsymbol{s},t \in \mathbb{R}^d \times \mathbb{R}_0^+}$ *with Gumbel margins (possibly transformed). We define the null-hypothesis* H_0 *by,*

$$H_0 \ : \ \{\eta(\boldsymbol{s}, t) : (\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}_0^+\} \text{ is max stable} \tag{3.6}$$

Under H_0 *all finite-dimensional margins are max-stable and for a concrete subset* $D \subseteq S_M \times T_K$ *we can determine the distribution function on* D *by*

$$F_D\left(y_1, ..., y_{|D|}\right) = \exp\{-\mu_D\left(e^{y_1}, ..., e^{y_{|D|}}\right)\}, \quad \left(y_1, ..., y_{|D|}\right) \in \mathbb{R}^{|D|},$$

where μ_D *is the exponent measure, introduced in (1.12). Let* $\psi_D := \log\left(\mu_D(1, ..., 1)\right)$. *Then* $0 \leq \psi_D \leq \log(|D|)$, *because* $1 \leq \mu_D(1, ..., 1) \leq |D|$. *Since* μ_D *is homogeneous of order* -1, *we get, that the random variable*

$$\eta_D := \max\{\eta(\boldsymbol{s}, t) \ : \ (\boldsymbol{s}, t) \in D\},$$

has a univariate Gumbel distribution, written as

$$P\left(\eta_D \leq y\right) = F_D(y, ..., y) = \exp\left(-e^{-y}\mu_D(1, ..., 1)\right) = \exp\left(-e^{-(y-\psi_D)}\right), \ y \in \mathbb{R}.$$

So ψ_D is the location parameter. We can use this to construct a graphical test for max-stability. We simulate for certain subsets D with fixed cardinality several independent realisations of η_D and test by means of a qq-plot, if they follow a Gumbel distribution. For more details see Buhl and Klüppelberg [8], Section 5.2.

Let $K = |D|$ be the cardinality of the subset D. We also apply non parametric block bootstrap methods to get 95 % confidence bounds, as a measure for variability. A consequence of the bootstrap method is that we preserve the dependence between different subsets D in the confidence intervals. To check if the null hypothesis is fulfilled, the bisecting line of the qq-plot should lie within these confidence bounds.

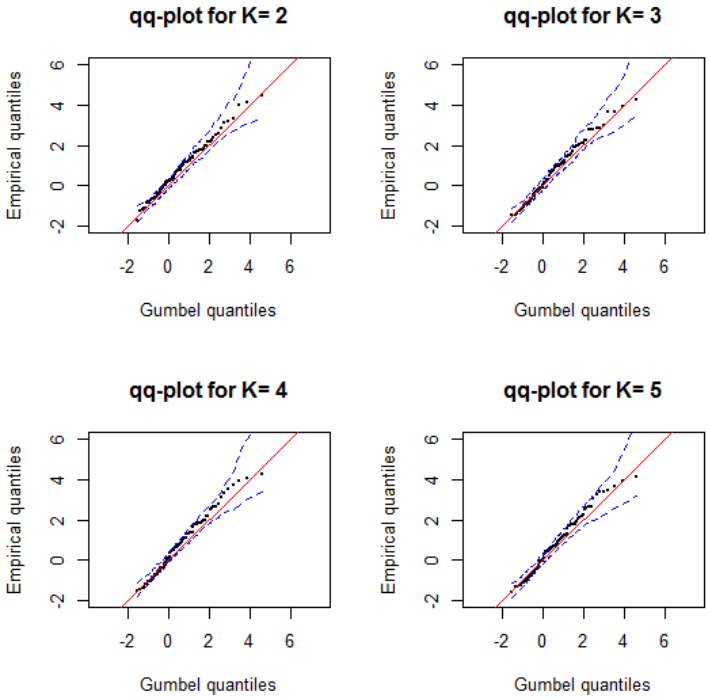

Figure 3.2: qq-plots of theoretical standard Gumbel quantiles versus maxima over subsets of the processes,simulated by the Gaussian approximation method. From top left to bottom right the cardinality K of the subsets changes from $K = 2$ to $K = 5$. Blue dots describe the 95 % pointwise confidence bounds, obtained by block bootstrap. Red line represents no deviation.

We can observe the result for the Gaussian approximation method of the max-stability test in Figure 3.1. It is no problem that our process has Fréchet margins. With an easy log-transformation, as for univariate qq-plots before, we get a Brown-Resnick process with Gumbel margins, which can be compared to Gumbel quantiles. We see four different qq-plots with $K = 2, 3, 4, 5$. We can not reject for any cardinality the null hypothesis (3.6), that the

simulated process $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^2\times\mathbb{R}_0^+}$ is max-stable. Hence there is no statistically significant evidence that the simulated space-time process is not max-stable.

We can already see from checking the univariate margins and the max-stability test, that the simulated process $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^2\times\mathbb{R}_0^+}$ can be seen as a good approximation but deviates from an exact result.

We now can compare the theoretical **bivariate distribution function** of a Brown-Resnick process with Gumbel margins $\tilde{F}^{TH}$ as in (2.4), with the empirical distribution $\tilde{F}^{SM}$ function which we get from the simulation, so

$$\tilde{F}^{TH}(y_1,y_2) = P\left(\eta^{TH}(\boldsymbol{s}_1,t_1)\leq y_1, \eta^{TH}(\boldsymbol{s}_2,t_2)\leq y_2\right) \tag{3.7}$$

compared to

$$\tilde{F}^{EM}(y_1,y_2) = P\left(\eta^{EM}(\boldsymbol{s}_1,t_1)\leq y_1, \eta^{EM}(\boldsymbol{s}_2,t_2)\leq y_2\right), \tag{3.8}$$

$(\boldsymbol{s}_i,t_i)\in\mathbb{R}^d\times\mathbb{R}_0^+$, for $i=1,2$. A comparison is possible with a **contour plot** at random locations. In Figure 3.2 the random locations are $(\boldsymbol{s}_1,t_1)=(4,3,1)$ and $(\boldsymbol{s}_2,t_2)=(8,9,1)$. The red contour lines should be similar to the green line of the theoretical distribution function. But as we can see, we have deviation for bigger quantiles. It is flashy, that the red contour line for all the % quantiles lies above the theoretical green line, so we can say the simulated values with the Gaussian approximation method tend to be too large. But since it is an approximation method, we can be satisfied with the result.

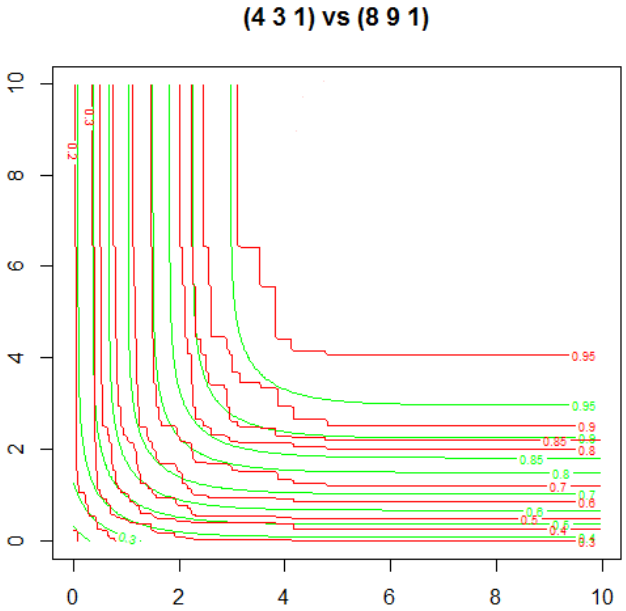

Figure 3.3: Contour Plot of the theoretical bivariate distribution of a Brown-Resnick process with Fréchet margins (green lines), compared to the empirical distribtuion of our Simulation with the Hüsler and Reis method (red lines.)

3.2 Naive approach with spectral representation

The next method isn't based on an actual paper. The idea to try the naive approach has come up during the writing of my thesis. This section can be seen more as a method, which can display problems, that can occur when simulating a Brown-Resnick process. In all the other methods we actually try to avoid these problems.

3.2.1 General assumptions

We recall the Brown-Resnick process with Fréchet margins as in (2.5).

$$\eta(\boldsymbol{s},t) = \bigvee_{j=1}^{\infty} V_j \exp\left(W_j(\boldsymbol{s},t) - \delta(\boldsymbol{s},t)\right), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+. \tag{3.9}$$

The random fields $\boldsymbol{W}_j = \left\{W_j(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\right\}$ are independent copies of the process $\boldsymbol{W} = \left\{W(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\right\}$. The random field $\boldsymbol{W}$ has mean function $m(\boldsymbol{s},t) = 0$, for all $\boldsymbol{s} \in \mathbb{R}^d, t \in [0,\infty)$ and variance $Var[W(\boldsymbol{s},t)] = \sigma^2(\boldsymbol{s},t)$, for all $\boldsymbol{s} \in \mathbb{R}^d, t \in [0,\infty)$. We can also rewrite (3.9) into

$$\eta(\boldsymbol{s},t) = \bigvee_{j=1}^{\infty} V_j \exp\left[W_j(\boldsymbol{s},t) - \sigma^2(\boldsymbol{s},t)/2\right], \quad (\boldsymbol{s},t) \in \mathbb{R} \times \mathbb{R}_0^+. \tag{3.10}$$

As the headline of this section already points out, we want to check how good a naive simulation of the Brown-Resnick process is. But first what do we understand by a naive approach? Recall that in the Gaussian approximation method, we avoid to create Poisson points V_i, with the help of the transformation of the Gaussian fields as in (3.2) to standard Fréchet margins. If we look ahead to the next chapter, where we focus on the method by Dieker and Mikosch [20], we can even avoid the pointwise maximum over an infinite number of random functions with a change of measure argument.

3.2.2 Implementation

For the naive approach now, we do not consider any transformations, but create the fields $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ directly from its spectral representation. In other words, we create both the Poisson points V_i and the Gaussian fields W_i (and its variogram δ) separately and combine them together to the random field

$$\varsigma_i(\boldsymbol{s},t) = V_i \exp\left(W_i(\boldsymbol{s},t) - \delta(\boldsymbol{s},t)\right), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+. \tag{3.11}$$

As in the section before, we now have to choose the number of simulations N of the field $(\varsigma_i(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ up to which we determine the maximum over. We finally assume, that the Brown-Resnick Process $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ can be approximated by

$$\eta(\boldsymbol{s},t) \approx \bigvee_{j=1}^{N} \varsigma_j(\boldsymbol{s},t), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+. \tag{3.12}$$

Before we test this simulation idea for max-stability as for the Gaussian approximation method, we first want to bring to attention a problem, which we face when we deal with this naive approach.

3.2.3 Convergence behaviour of $W - \delta$

In the naive approach we are confronted with the difference of a Gaussian field and its variogram, namely the term

$$W(\boldsymbol{s}, t) - \delta(s, t),\ (\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}_0^+.$$

For further investigations we focus only on one-dimensional random fields for care of notation, i.e. on the term $W(t) - \delta(t),\ t \in \mathbb{R}$.

The choice of the variogram δ depends on the model we choose for our simulation. It is a deterministic function, so it is simple to determine its convergence behaviour. This doesn't hold for the Gaussian field $\boldsymbol{W}$, which is stochastic.

We now want to examine the convergence behaviour for Gaussian processes. We consider two different convergence types. Since the Gaussian field is a stochastic process, we first want to look at the convergence behaviour of the expectation.

1.) Convergence of the expectation of $W - \delta$

Theorem 3.5. *Let $\boldsymbol{W} = \{W(t) : t \in \mathbb{R}_0^+\}$ be a Gaussian random field, $\delta(t)$ its variogram and we only consider variograms, for which $\delta(t) \to \infty$ for $t \to \infty$. Then we can claim that,*

$$\lim_{t\to\infty} E\left[\exp\left(W(t) - \delta(t)\right)\right] = 1. \tag{3.13}$$

Proof: *Equation (3.13) can be shown, with some simple descriptions.*

$$\begin{aligned}
\lim_{t\to\infty} E\left[\exp\left(W(t) - \delta(t)\right)\right] &= \lim_{t\to\infty} \exp\left(-\delta(t)\right) E\left[\exp\left(W(t)\right)\right] \\
&= \lim_{t\to\infty} \exp\left(-\delta(t)\right) \exp\left(\frac{1}{2}\sigma^2(t)\right) \\
&= \lim_{t\to\infty} \exp\left(0\right) = 1,
\end{aligned}$$

where we used that $W(t) \sim \mathcal{N}\left(0, \sigma^2(t)\right)$ and that $\sigma^2(t)/2 = \delta(t)$. $\square$

Remark 3.6. *Following Theorem 3.5, for large t, $\varsigma_i(t)$ from (3.11), will converge to the expectation of the Poisson points V_i. We see, if we take the expectation of $\varsigma_i(t)$ as follow:*

$$E\left[\varsigma_i(t)\right] = E\left[V_i \cdot \exp\left(W_i(t) - \delta(t)\right)\right] = E\left[V_i\right] \cdot E\left[\exp\left(W_i(t) - \delta(t)\right)\right] = E\left[V_i\right],$$

where we used, that the Poisson process V and the Gaussian field $\boldsymbol{W}$ are independent.

To rephrase Remark 3.6, for large time points, we see, that we only take the maximum of the spectral representation over the expectation of the Poisson points V_i.

2.) Almost sure convergence of $W - \delta$

Dieker and Mikosch [20] say that $W(t) - \delta(t)$ converges almost surely to $-\infty$, for $t \to \infty$, where $(W(t))_{t\geq 0}$ is a standard Brownian motion. Since most of our simulations are done with fractional Brownian motions, we want to examine almost sure convergence for these self-similar processes. Where for the convergence of the expectation it wasn't relevant to distinguish between these two types of Gaussian processes, it can be relevant for the almost sure convergence. Simultaneously we want to check the rate of convergence, if we can extend the assumption that $W(t) - \delta(t)$ converges almost surely to $-\infty$, for $t \to \infty$, to fractional Brownian motions.

Rate of convergence

We do this examination for three different cases. Simultaneously we display this graphically in Figure 3.4-3.11. The Gaussian field $\boldsymbol{W}$ is a fractional Brownian motion, introduced as in Chapter 2.2. Remember for simulation details Remark 3.3.

We use the function **RMfbm** of the package RandomFields to define an intrinsically stationary variogram model. The corresponding centered variogram only depends on the distance $r \geq 0$ between two points and is given by $\delta(r) = r^\alpha/2$. Additional parameters, like the *scale* and *var* parameter transform the variogram δ into:

$$\bar{\delta}(r) = var \cdot \delta(r/scale).$$

Before we begin, we have to define the term asymptotically equivalent.

Definition 3.7. *Two functions f, g with $f, g : \mathbb{R} \to \mathbb{R}$ are called **asymptotically equivalent**, if and only if,*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

We notate asymptotically equivalent as $f \sim g$.

In what follows, we now show how different *var* and *scale* parameter influence the behaviour of the fractional Brownian motion. But mostly we point out how or if different α (and so different Hurst parameters H) impact the convergence rate of the term $W(t) - \delta(t)$.

- In Figure 3.4 - 3.9 we do the simulation with $H = 1$, so $W_i(t) = B_1(t)$. Recall Lemma 2.22. If $H = 1$, it follows that

  $$W_i(t) = B_1(t) = tB(1) \quad a.s.,$$

  which means, that the Gaussian field is a linear function. It follows that

  $$W(t) - \delta(t) = tB(1) - t^2/2 \quad a.s.,$$

  where $\alpha = 2$ and w.l.o.g $scale = var = 1$. Since the Gaussian field $\boldsymbol{W}$ is a linear function, we can conclude for the convergence rate of the term $W_i(t) - \delta(t)$,

  $$\mathcal{O}(-\delta(t)) = -\mathcal{O}(t^2).$$

- In Figure 3.10, we are dealing with the standard Brownian motion. With the law of the iterated logarithm (Theorem 2.6),

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t\log(\log(t))}} = 1 \quad a.s.,$$

we can claim that $W(t)$ has an almost sure maximal fluctuation up to time t of order $\sqrt{2t\log(\log(t))}$ for every path $W(t,\omega)$, where $\omega \in \Omega$ and Ω the underlying probability space. Since the corresponding variogram for a standard Brownian motion is $\delta(t) = \sigma^2(t)/2 = t/2$, we can follow that $W(t) - \delta(t)$ asymptotic behaves like

$$[W(t,\omega) - \delta(t)] = [B(t,\omega) - t/2] \sim \left[\sqrt{2t\log(\log(t))} - t/2\right], \quad a.s.,$$

for all $\omega \in \Omega$, which means, that it holds almost surely for every sample path of the Brownian motion. To see the limit of the right hand side clearer we can give an upper bound

$$\lim_{t\to\infty}\left[\sqrt{2t\log(\log(t))} - t/2\right] \leq \lim_{t\to\infty}\left[\sqrt{2t\log(t)} - \frac{t}{2}\right] \leq \lim_{t\to\infty}\left[\sqrt{2t\sqrt{t}} - \frac{t}{2}\right]$$

Since this right hand side already tends to $-\infty$ for $t \to \infty$, we can conclude $\lim_{t\to\infty}[W(t) - t/2]$ goes to $-\infty$ $a.s.$ with linear convergence rate $-\mathcal{O}(t)$.

- In Figure 3.11 we look at a fractional Brownian motion with a very small $H = 0.005$. At first sight we can't see any convergence. It needs a little more effort to find a concrete convergence rate for the expression $W(t) - \delta(t)$. Since the fractional Brownian motion doesn't have a linear behaviour anymore and we don't consider the special case for $H = \frac{1}{2}$, it depends on H how fast the convergence is. Theorem 2.20 can give a first idea, how we can give an upper bound for the distribution of the supremum of the fractional Brownian motion. Recall

$$P\left(\sup_{0\leq t\leq a} B_H(t) \geq x\right) = \frac{\sqrt{2}a^H}{x\sqrt{\pi}}$$

from Theorem 2.24, which would hold for large H, close to one. This upper bound is based on the idea of the reflection principle for fractional Brownian motion.

For every $H \in (0,1)$ we can use alternatively as in the standard Brownian motion case a law of the iterated logarithm. This time we use Theorem 2.25 which gives a law of the iterated logarithm for fractional Brownian motions. Similar to the standard Brownian motion case

$$\limsup_{t\to\infty} \frac{B_H(t)}{\sqrt{2t^{2H}\log(\log(t))}} = 1 \quad a.s.,$$

we can say that the maximal fluctuation of a fractional Brownian motion is almost surely $\sqrt{2t^{2H}\log(\log(t))}$ for every path $W(t,\omega)$, where $\omega \in \Omega$. So we can interpret this as an almost sure asymptotic behaviour. It follows that

$$W(t,\omega) - \delta(t) = B_H(t,\omega) - t^{2H}/2 \sim \sqrt{2t^{2H}\log(\log(t))} - t^{2H}/2 \; a.s.,$$

for all $\omega \in \Omega$, which means, that it holds almost surely for every sample path of the fractional Brownian motion, as in the standard case. For the right hand side, we can calculate

$$\begin{aligned}
\lim_{t\to\infty}\left(\sqrt{2t^{2H}\log\left(\log(t)\right)} - t^{2H}/2\right) &= \lim_{t\to\infty}\left(t^{2H}\left[\frac{\sqrt{2\log\left(\log(t)\right)}}{t^H} - \frac{1}{2}\right]\right) \\
&= \lim_{t\to\infty}\left(\frac{\frac{\sqrt{2\log(\log(t))}}{t^H} - \frac{1}{2}}{t^{-2H}}\right) \\
&\Leftrightarrow \lim_{t\to\infty}\left(\frac{\overbrace{\left[\frac{2\log\left(\log(t)\right)}{t^{2H}}\right]}^{\to\ 0\ (\star)} - \frac{1}{2}}{\underbrace{t^{-2H}}_{\to\ 0}}\right) \\
&= -\infty \quad a.s.
\end{aligned}$$

It is left to show why $(\star)$ converges to 0 for $t \to \infty$. We show this with the rule of l'Hospital:

$$\lim_{t\to\infty}(\star) = \lim_{t\to\infty}\frac{2\log\left(\log(t)\right)}{t^{2H}} \underset{=}{\text{l'Hospital}} \lim_{t\to\infty}\frac{\frac{2}{\log(t)}\cdot t^{-1}}{2Ht^{2H-1}} = \lim_{t\to\infty}\frac{1}{Ht^{2H}\cdot\log(t)} = 0.$$

As in the cases before

$$\lim_{t\to\infty} W(t) - \delta(t) = \lim_{t\to\infty} B_H(t) - t^{2H} = -\infty \quad a.s.$$

The reason why we can't see any convergence in Figure 3.11 is the impact of H on the convergence rate. The convergence rate for the term $W(t) - \delta(t)$ for the fractional Brownian case is $-\mathcal{O}(t^{2H})$. The smaller H gets, the more slowly $W(t) - \delta(t)$ will converge almost surely to $-\infty$ for $t \to \infty$.

We can summarize these results in the following remark.

Remark 3.8. *We can conclude from the rate of convergence observations, that the term $W(t) - \delta(t)$ converges to $-\infty$ almost surely for $t \to \infty$ for every fractional Brownian motion, independent of the choice of the Hurst parameter H. This will lead to an almost sure convergence of the*

$$\varsigma_i(t) = V_i \exp\left(W_i(t) - \delta(t)\right), \quad t \in \mathbb{R}_0^+,$$

to 0. In other words, for large t, we take the maximum in the spectral representation over values close to zero.

But we also observe, that we can influence the rate of convergence with the Hurst parameter. Since we usually don't have an arbitrary large number of timepoints, we can choose the Hurst parameter in such a way, that the rate of convergence is slow enough. Slow enough means in this context, that the $\varsigma_i(t)$ are significantly larger than 0.

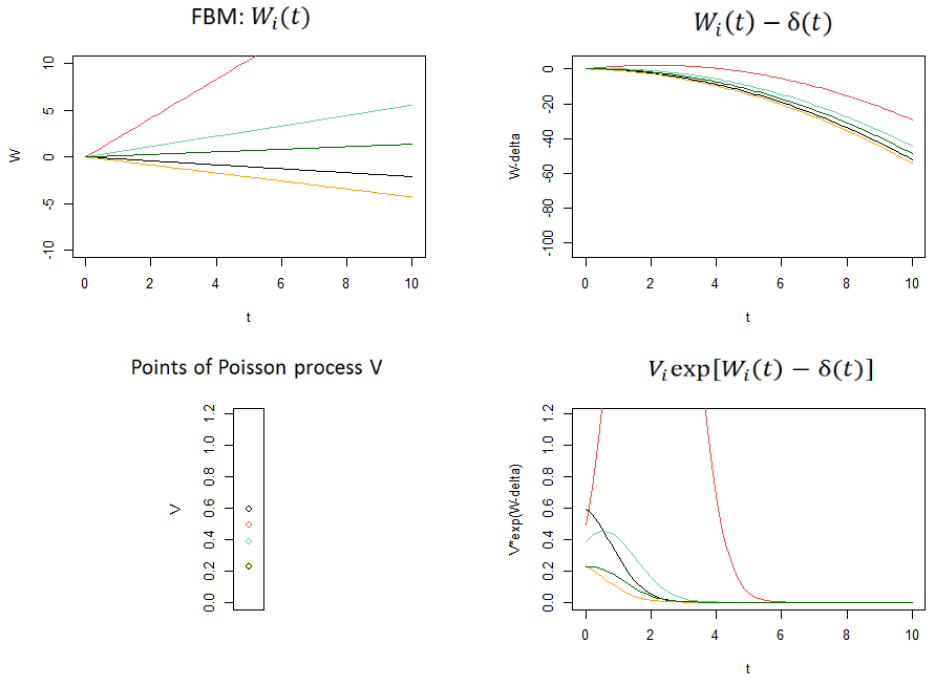

Figure 3.4: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 0.5, scale = 1$.

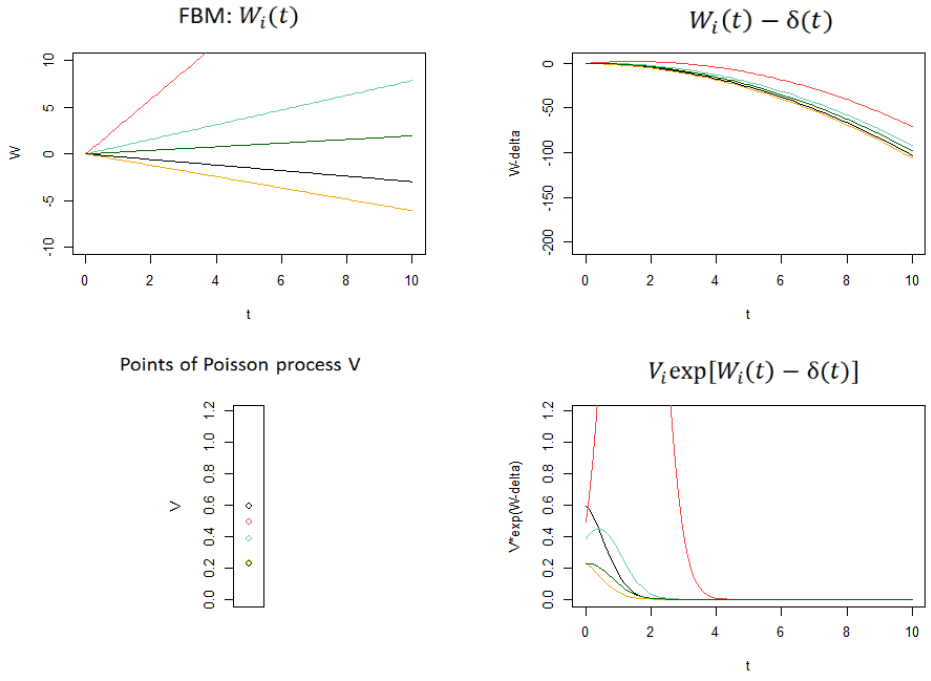

Figure 3.5: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 1, scale = 1$.

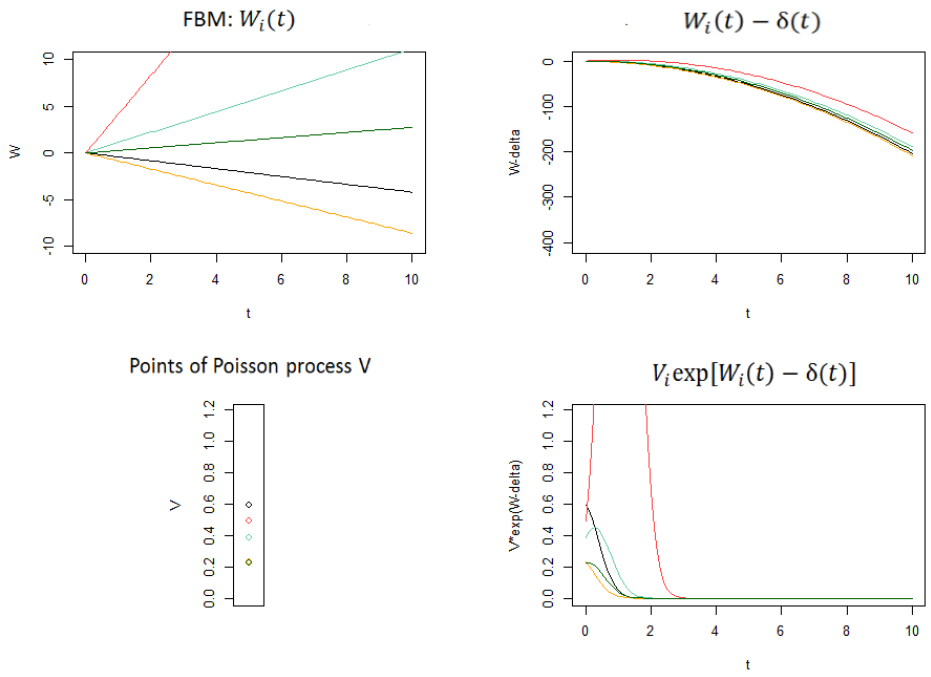

Figure 3.6: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 2, scale = 1$.

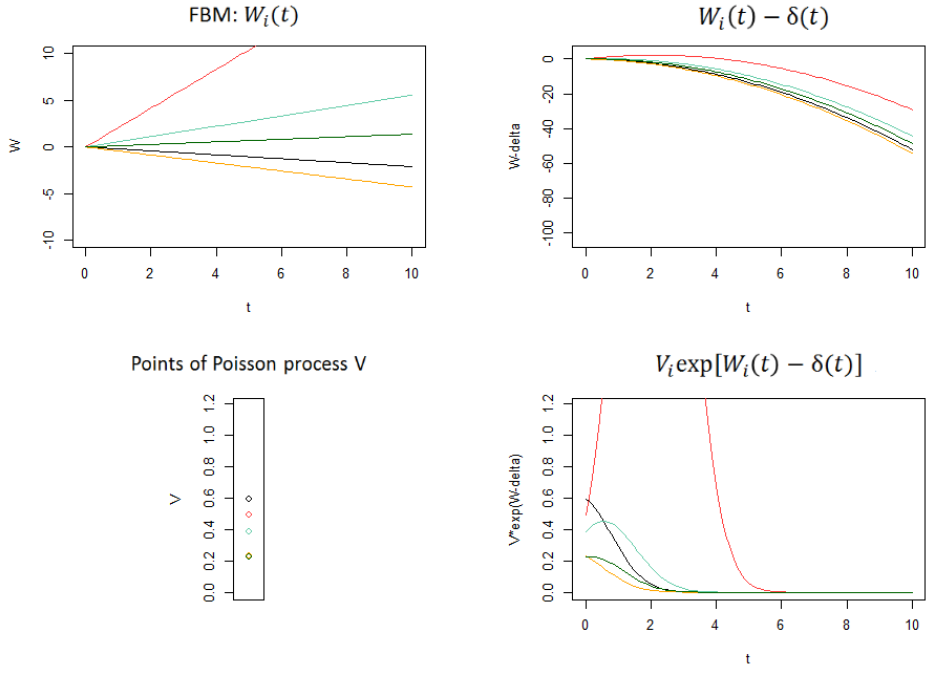

Figure 3.7: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 0.5, scale = 1$.

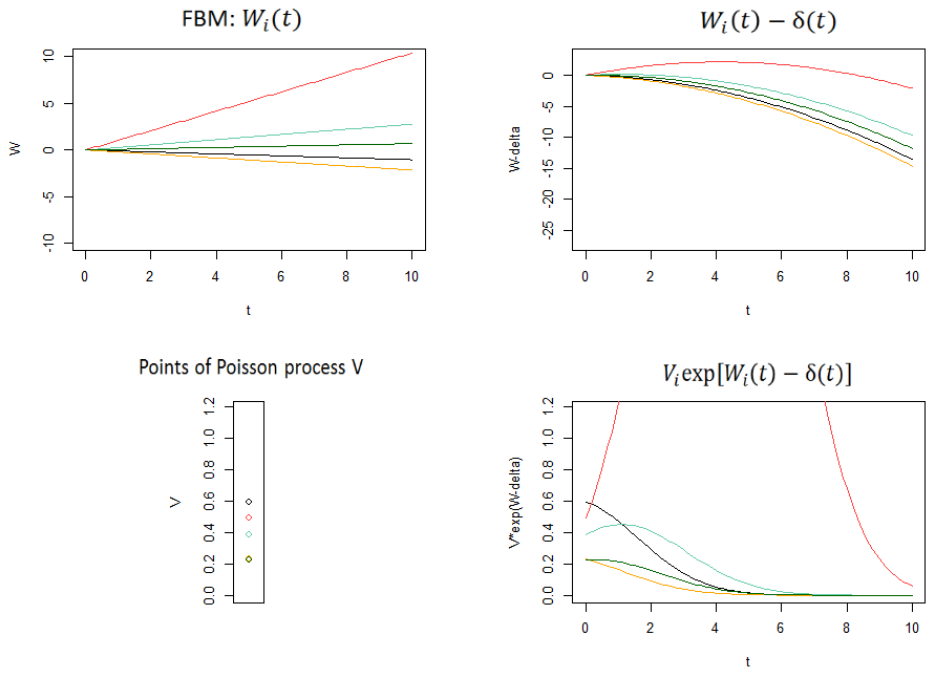

Figure 3.8: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 0.5, scale = 2$.

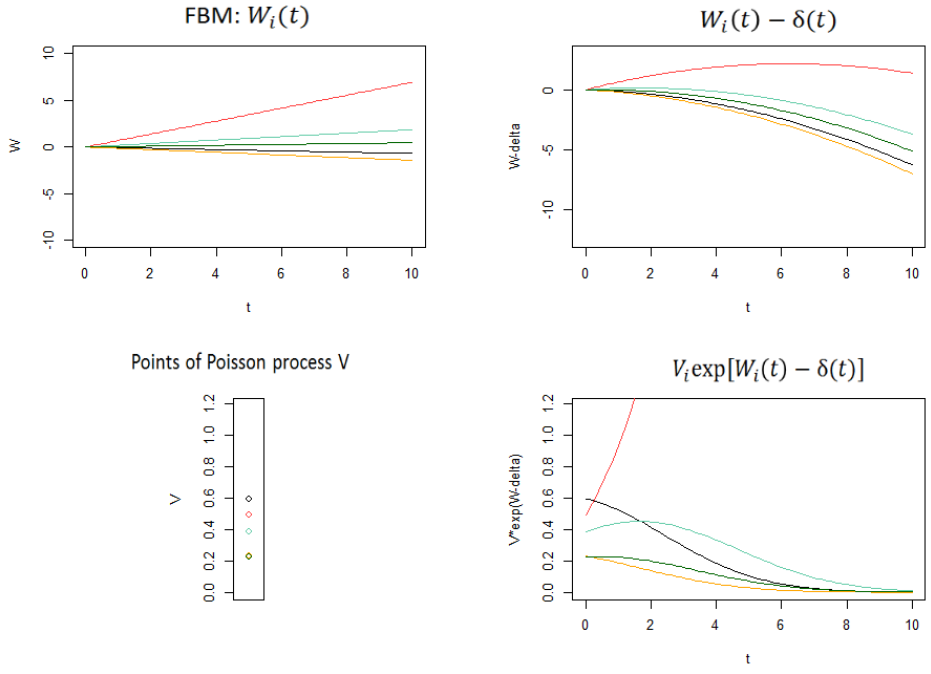

Figure 3.9: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 2, var = 0.5, scale = 3$.

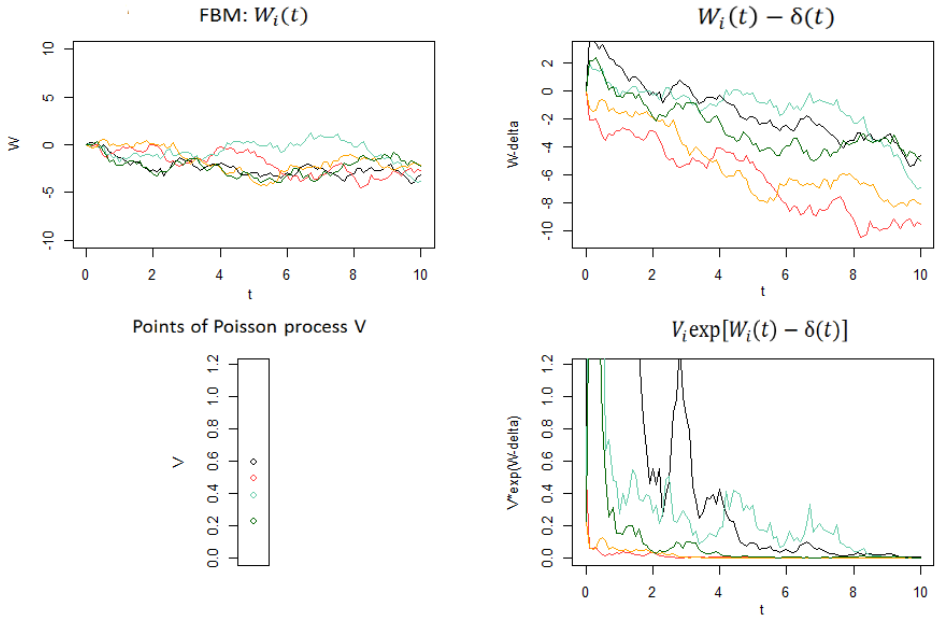

Figure 3.10: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 1, var = 1, scale = 2$.

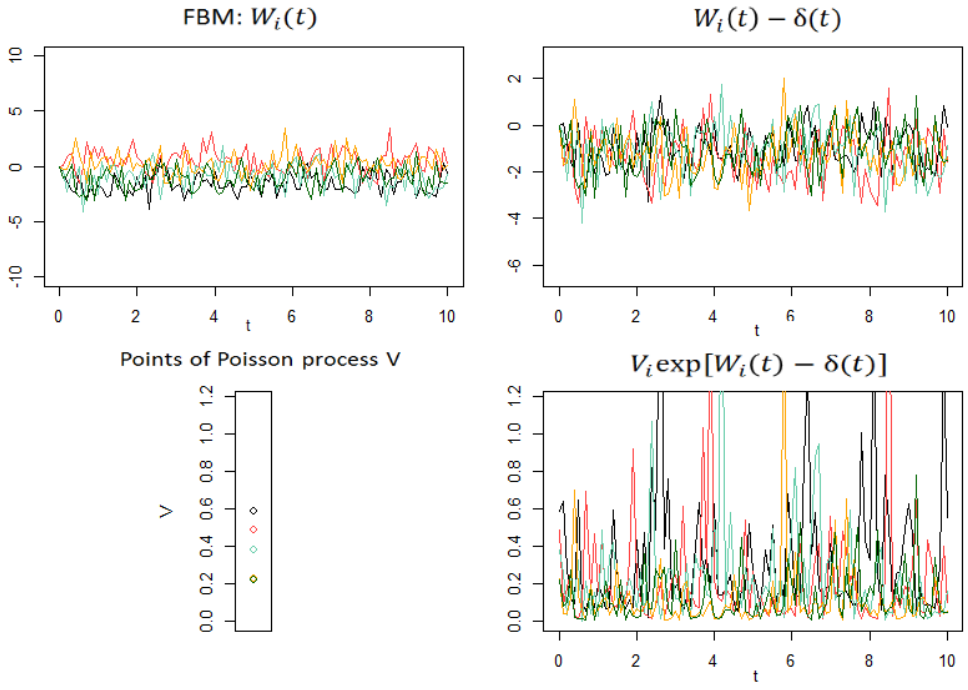

Figure 3.11: Random draws of the Gaussian field W_i, the corresponding difference $W_i - \delta$, the Poisson process V and $\varsigma_i = V_i \exp(W_i - \delta)$ with parameter for the Gaussian field given as follow: $\alpha = 0.01, var = 1, scale = 2$.

3.2.4 Results of the naive approach

We now come to the simulation results for the naive method. We do exactly the same as for the Gaussian approximation method. The Gaussian fields are getting simulated as in Remark 3.3. As for the Gaussian approximation method, we also have in the naive simulation the advantage to predict the simulation effort, since we are free to choose the number of simulated Gaussian fields N, like in (3.12).

We start with the univariate margins at different locations. For comparison, we choose the same locations. Since we only expect an approximation, the qq-plots in Figure 3.12 are quite satisfying. Only for larger quantiles, we have an obvious deviation.

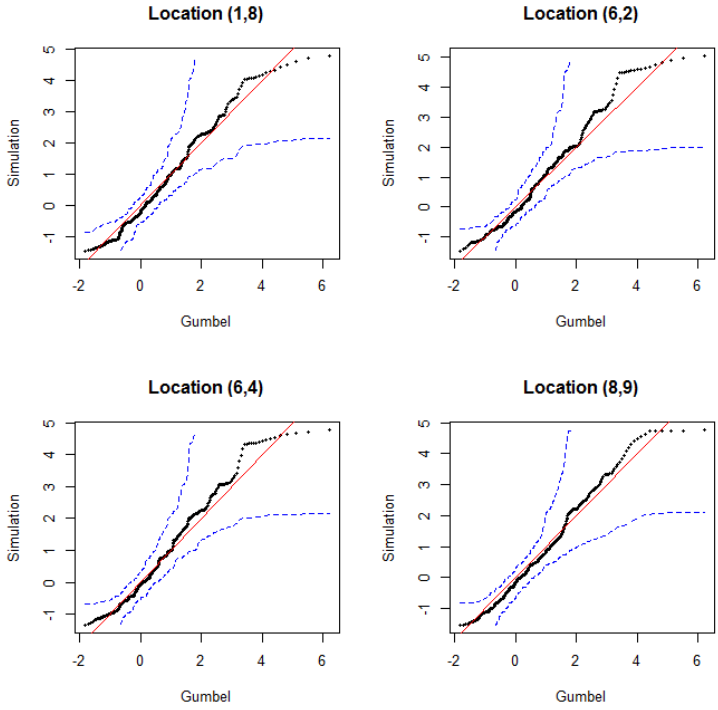

Figure 3.12: qq-plot of univariate margins of the log-transformed simulated process by the naive approach for a $10 \times 10 \times 5$-field versus the standard Gumbel distribution for four random space-time locations: (1,8,1), (6,2,1), (6,4,1), (8,9,1). The red line would be no deviation and the blue lines are the 95 % confidence bounds.

Again, we do the max-stability test, presented in Remark 3.4. If we compare the results from the max-stability test of the Gaussian approximation method in Figure 3.2 to the results from the naive approach in Figure 3.13, we can see an deterioration. The empirical quantiles tend to be bigger than the theoretical qunatiles of the Gumbel. Furthermore we can reject the null hypothesis that the simulated process from (3.12) is max-stable to the confidence level of 95 %, since the bisectings lines of the qq-plots don't lie within the bootstrap confidence bounds.

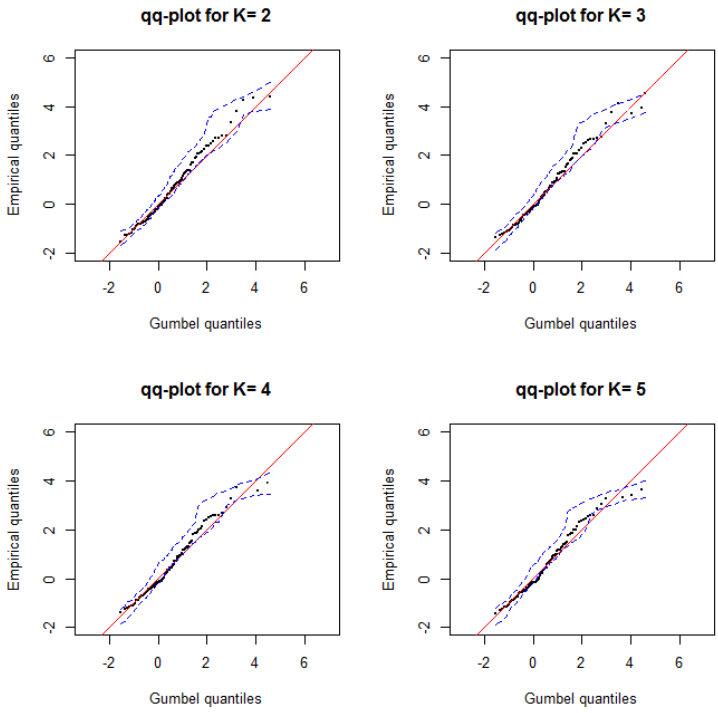

Figure 3.13: qq-plots of theoretical standard Gumbel quantiles versus maxima over subsets of the processes,simulated by the naive approach. From top left to bottom right the cardinality K of the subsets changes from $K = 2$ to $K = 5$. Blue dots describe the 95 % pointwise confidence bounds, obtained by block bootstrap. Red line represents no deviation.

The contour plot of the empirical and the theoretical distribution function is left to compare. If we recall the contour plot of the Gaussian approximation method (Figure 3.3, we remember that it was a satisfying approximation, but the empirical quantiles tend to be too large . In Figure 3.14 we can see that for the naive approach the probability mass is distributed over the complete representing subset of the $\mathbb{R}^2$. But the empirical quantiles compared to the theoretical distribution function of a Brown-Resnick process with Gumbel margins are too big especially for the second argument y_2 as in

$$\tilde{F}(y_1, y_2) = \exp\left[-\exp(y)_1^{-1}\Phi\left(\frac{y_2 - y_1}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right) - \exp(y)_2^{-1}\Phi\left(\frac{y_1 - y_2}{2\sqrt{\delta(\boldsymbol{h}, u)}} + \sqrt{\delta(\boldsymbol{h}, u)}\right)\right].$$

where $\tilde{F}$ is the bivariate distribution function of a Brown-Resnick process with Gumbel margins.

After presenting these results for the first two methods, we can give an **intermediate result**. Both methods are similar in computational effort, since we have to choose the number of Gaussian fields, which have to be simulated to approximate the infinite number of functions in the spectral representation of a Brown-Resnick process.

After all we can only expect an approximation.

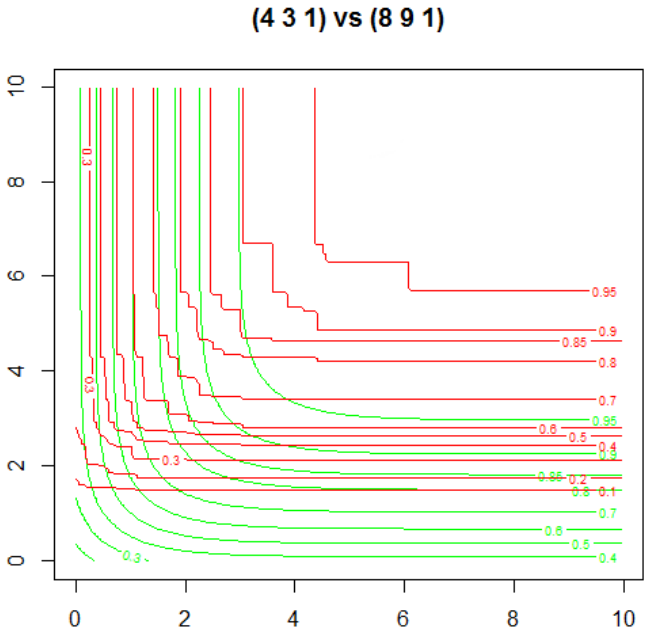

Figure 3.14: Contour Plot of the theoretical bivariate distribution of a Brown-Resnick process with Fréchet margins (green lines), compared to the empirical distribtuion of our Simulation with the naive method (red lines.)

3.3 Method by Dieker and Mikosch

We now come to the first exact algorithm. The methods before only give us an approximation of a Brown-Resnick process. The main idea is to apply a suitable change of measure to avoid the maximum over an infinite number of functions. We will extend the theory to a space-time version.

3.3.1 Theoretical background

In what follows, we focus completely on the paper from Dieker and Mikosch [20].

Dieker and Mikosch deel with a Brown-Resnick process $(\eta(\boldsymbol{t}))_{\boldsymbol{t}\in\mathbb{R}^d}$ defined as

$$\eta(\boldsymbol{t}) = \sup_{i\geq 1}\left(V_i + W_i(\boldsymbol{t}) - \sigma^2(\boldsymbol{t})/2\right), \quad \boldsymbol{t}\in\mathbb{R}^d, \tag{3.14}$$

where $\sigma^2(\boldsymbol{t}) = Var\left(W_1(\boldsymbol{t})\right)$, (W_i) is a sequence of iid. centered gaussian processes with stationary increments on $\mathbb{R}^d$ and (V_i) are the points of a poisson process V with intensity measure $e^{-x}dx$. So this time we focus on a Brown-Resnick process with Gumbel margins.

Since we implement space-time processes we also want to introduce $(\eta(\boldsymbol{t}))_{\boldsymbol{t}\in\mathbb{R}^d}$ in such a way and we speak of $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\ \in\mathbb{R}^d\times\mathbb{R}_0^+}$:

$$\eta(\boldsymbol{s},t) = \sup_{i\geq 1}\left(V_i + W_i(\boldsymbol{s},t) - \sigma^2(\boldsymbol{s},t)/2\right), \quad (\boldsymbol{s},t)\in\mathbb{R}^d\times\mathbb{R}_0^+, \tag{3.15}$$

where $\sigma^2(\boldsymbol{s},t) = Var\left(W_1(\boldsymbol{s},t)\right)$, (W_i) is a sequence of iid. centered gaussian processes with stationary increments on $\mathbb{R}^d\times\mathbb{R}_0^+$ and (V_i) is defined as in (3.11).

Throughout this section we consider this space-time version of $\boldsymbol{\eta}$. It is obvious that we are again faced with the problem, that we have to calculate an infinite supremum and so this representation is not suitable for exact sampling. We now come to the central theorem of this method.

For simplicity, we sometimes write $Z_i(\boldsymbol{s},t) = W_i(\boldsymbol{s},t) - \sigma^2(\boldsymbol{s},t)/2$.

Theorem 3.9 (see Dieker and Mikosch [20], Theorem 2.1)**.** *Suppose we are given an arbitrary probability product measure μ on $\mathbb{R}^d\times\mathbb{R}_0^+$, with $\mu = \mu_1\times\mu_2$. Consider*

$$\begin{aligned}\zeta(\boldsymbol{s},t) = \sup_{i\geq 1}\big(&V_i + Z_i(\boldsymbol{s}-\boldsymbol{S}_i, t-T_i)\\ &- \log\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z_i\left(\boldsymbol{u}-\boldsymbol{S}_i, v-T_i\right)\right)\mu_2(dv)\mu_1(d\boldsymbol{u})\right)\big),\end{aligned} \tag{3.16}$$

where $((\boldsymbol{P}_i,V_i))_{i\geq 1}$ are the points of a Poisson Process on $(\mathbb{R}^d\times\mathbb{R}_0^+)\times\mathbb{R}$ with intensity measure $\mu(d\boldsymbol{x})\times\exp(-\nu)d\nu$, where $\boldsymbol{P}_i = (\boldsymbol{S}_i,T_i)$. Then the random fields $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d}$ and $(\zeta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d}$ have the same distribution.

For the proof of this central theorem, we first need some Lemmas.

Lemma 3.10. *Let (X_i) be iid. copies of some random field $\boldsymbol{X}$ on $\mathbb{R}^d \times \mathbb{R}_0^+$ and (V_i) the points of a Poisson process on $\mathbb{R}$ with intensity measure $e^{-x}dx$. If we write*

$$\xi(\boldsymbol{s},t) = \sup_{i\geq 1}\left(V_i + X_i(\boldsymbol{s},t)\right), \quad (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+$$

then we have for $y_j \in \mathbb{R}, (\boldsymbol{s}_j, t_j) \in \mathbb{R}^d \times \mathbb{R}_0^+, i = 1, ..., n,$

$$P\left(\xi(\boldsymbol{s}_1,t_1) \leq y_1, ..., \xi(\boldsymbol{s}_n,t_n) \leq y_n\right) = \exp\left(-E\exp\left(\max_{j=1,...,n}\left(X(\boldsymbol{s}_j,t_j) - y_j\right)\right)\right).$$

Proof: *We start without the supremum and get*

$$\begin{aligned} P\left(V_i + X_i(\boldsymbol{s}_1,t_1) \leq y_1, ..., V_i + X_i(\boldsymbol{s}_n,t_n) \leq y_n\right) &= P\left(V_i \leq y_1 - X_i(\boldsymbol{s}_1,t_1), ..., V_i - X_i(\boldsymbol{s}_n,t_n) \leq y_n\right) \\ &= P\left(\sup_{i\leq 1}(V_i) \leq \min_{j=1,...,n}\left(y_j - X(\boldsymbol{s}_j,t_j)\right)\right) \end{aligned}$$

We further assume there exists a finite measure dP_X corresponding to the random field X on $\mathbb{R}^d$. Since V_i are the points of a poisson process with intensity measure $e^{-x}dx$ we can look at the following probability for the supremum of the points V_i:

$$\begin{aligned} P\left(\sup_{i\leq 1}(V_i) \leq \min_{j=1,...,n}\left(y_j - X(\boldsymbol{s}_j,t_j)\right)\right) &= P\left(\textit{no points above } x \textit{ for } V_i\right) \\ &= P\left(V_i(x,\infty) = 0\right) \\ &= \exp\left(-\int_\Omega \int_{x>\min_{j=1,...,n}((y_j - X(\boldsymbol{s}_j,t_j))} e^{-x}dx\ dP_X\right) \\ &= \exp\left(-\int_\Omega \exp\left(-\min_{j=1,...,n}\left(y_j - X(\boldsymbol{s}_j,t_j)\right)\right)\ dP_X\right) \\ &= \exp\left(-E\left[\exp\left(\max_{j=1,...,n}\left(X(\boldsymbol{s}_j,t_j) - y_j\right)\right)\right]\right) \end{aligned}$$

□

Lemma 3.11 (see Dieker and Mikosch [20], Lemma 5.2)**.** *Fix $(\boldsymbol{u}, v) \in \mathbb{R}^d \times \mathbb{R}_0^+$. For a measurable functional $F : f \to \mathbb{R}$, with $f : \mathbb{R}^d \times \mathbb{R}_0^+ \to \mathbb{R}$ that is translation invariant, we have for a random field $\boldsymbol{Z}$*

$$E\left[e^{Z(\boldsymbol{u},v)}F(Z)\right] = E\left[F(\theta_{\boldsymbol{u},v}Z)\right],$$

where the shift $\theta_{\boldsymbol{u},v}$ is defined through $(\theta_{\boldsymbol{u},v}Z)(\boldsymbol{s},t) = Z\left(\boldsymbol{s}-\boldsymbol{u}, t-v\right)$. We understand by translation invariant, that the value of F doesn't change if we add a constant function. In our case, we focus on $Z(\boldsymbol{s},t) = W(\boldsymbol{s},t) - \sigma^2(\boldsymbol{s},t)$,

$$\begin{aligned} E\left[e^{W(\boldsymbol{u},v)-\sigma^2(\boldsymbol{u},\boldsymbol{v})/2}F(W-\sigma^2/2)\right] &= E\left[F\left(\theta_{\boldsymbol{u},v}\left[W(\boldsymbol{s},t) - \sigma^2(\boldsymbol{s},\boldsymbol{t})/2\right]\right)\right] \\ &= E\left[F\left(W(\boldsymbol{s}-\boldsymbol{u},t-v) - \sigma^2(\boldsymbol{s}-\boldsymbol{u},\boldsymbol{t}-\boldsymbol{v})/2\right]\right) \end{aligned}$$

Proof: *We introduce the measure Q with*

$$\frac{d\mathcal{Q}}{dP} = \exp\left(W(\boldsymbol{u}, v) - \sigma^2(\boldsymbol{u}, v)/2\right), \quad (\boldsymbol{u}, v) \in \mathbb{R}^d \times \mathbb{R}_0^+$$

while P is the typical Wiener measure of the Gaussian process $\boldsymbol{W}$. By comparing generating functions, we want to show that $W(\boldsymbol{s}, t) - \sigma^2(\boldsymbol{s}, t)/2$ under Q has the same distribution as $W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) + \sigma^2(\boldsymbol{u}, v)/2$ under P. This also similar to the proof of Lemma 1 in Dieker and Yakir [21]. For better comparison we look at the logarithm of the moment generating function, where $\lambda \in \mathbb{R}$ and get for $W(\boldsymbol{s}, t) - \sigma^2(\boldsymbol{s}, t)/2$ with regard to $\mathcal{Q}$:

$$\begin{aligned}
&\log E^{Q} \exp\left[\lambda(W(\boldsymbol{s}, t) - \sigma^2(\boldsymbol{s}, t)/2)\right] \\
&= \log E\left[\exp\left(\lambda(W(\boldsymbol{s}, t) - \sigma^2(\boldsymbol{s}, t)/2)\right) \cdot \exp\left(W(\boldsymbol{u}, v) - \sigma^2(\boldsymbol{u}, v)/2\right)\right] \\
&= \log\left[E\left[e^{\lambda \cdot \sigma^2(\boldsymbol{s}, t)/2))}\right]\right] + \log\left[E\left[e^{\sigma^2(\boldsymbol{u}, v)/2))}\right]\right] + \log\left[E\left[e^{\lambda W(\boldsymbol{s}, t) + W(\boldsymbol{u}, v)}\right]\right] \\
&= -\frac{\lambda}{2}\sigma^2(\boldsymbol{s}, t) - \frac{1}{2}\sigma^2(\boldsymbol{u}, v) + \frac{1}{2} Var\left[W(\boldsymbol{u}, v) + \lambda W(\boldsymbol{s}, t)\right] \\
&= -\frac{\lambda}{2}\sigma^2(\boldsymbol{s}, t) + \lambda C\left(W(\boldsymbol{u}, v), W(\boldsymbol{s}, t)\right) + \frac{1}{2} Var\left[\lambda W(\boldsymbol{s}, t)\right] \\
&= \lambda\left[\frac{1}{2}\sigma^2(\boldsymbol{u}, v) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v)\right] + \frac{1}{2} Var\left[\lambda W(\boldsymbol{s}, t)\right]
\end{aligned}$$

where we used in the third line for the third summand the moment generating function of a Gaussian process. Now similar is the moment generating function of $W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) + \sigma^2(\boldsymbol{u}, v)/2$ with regard to P the following:

$$\begin{aligned}
&\log E \exp\left[\lambda(W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) + \sigma^2(\boldsymbol{u}, v)\right] \\
&= \log\left[E\left[e^{\lambda \cdot W(\boldsymbol{s}, t)/2))}\right]\right] + \log\left[E\left[e^{\lambda \cdot \delta(\boldsymbol{s} - \boldsymbol{u}, t - v))}\right]\right] + \log\left[E\left[e^{\lambda \cdot \sigma^2(\boldsymbol{u}, v)}\right]\right] \\
&= \frac{1}{2} Var\left[\lambda W(\boldsymbol{s}, t)\right] + \frac{1}{2}\lambda \cdot \sigma^2(\boldsymbol{u}, v) - \lambda \cdot \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) \\
&= \lambda\left[\frac{1}{2}\sigma^2(\boldsymbol{u}, v) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v)\right] + \frac{1}{2} Var\left[\lambda W(\boldsymbol{s}, t)\right].
\end{aligned}$$

So we can already claim that,

$$E^{\mathcal{Q}}\left[F(W(\boldsymbol{s}, t) - \sigma^2(\boldsymbol{s}, t)/2\right] = E\left[F(W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) + \sigma^2(\boldsymbol{u}, v)/2\right].$$

for all $(\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}_0^+$ and fixed $(\boldsymbol{u}, v) \in \mathbb{R}^d \times \mathbb{R}_0^+$. When we use that F is translation invariant,

$$F\left[W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) + \sigma^2(\boldsymbol{u}, v)\right] = F\left[W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) - W(\boldsymbol{u}, v)\right]$$

for all $(\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}_0^+$ and fixed $(\boldsymbol{u}, v) \in \mathbb{R}^d \times \mathbb{R}_0^+$. We can do this, because both $\sigma^2(\boldsymbol{u}, v)$ and $W(\boldsymbol{u}, v)$ are constant. We finally can claim, that

$$W(\boldsymbol{s}, t) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v) - W(\boldsymbol{u}, v)$$

has the same distribution as

$$Z(\boldsymbol{s} - \boldsymbol{u}, t - v) = W(\boldsymbol{s} - \boldsymbol{u}, t - v) - \delta(\boldsymbol{s} - \boldsymbol{u}, t - v),$$

since we assume that the Gaussian field $\boldsymbol{W}$ *has stationary increments.* □

With the help of Lemma 3.10 and Lemma 3.11 we are able to give a proof, which shows that the random fields $(\eta(\boldsymbol{s},t))_{(\boldsymbol{s},t)\in\mathbb{R}^d\times\mathbb{R}_0^+}$ and $(\zeta(\boldsymbol{s},t))_{(\boldsymbol{s},t)\in\mathbb{R}^d\times\mathbb{R}_0^+}$ have the same distribution.

Proof of Theroem 3.9: *Recall, we want to prove*

$$P\left(\eta(\boldsymbol{s}_1,t_1)\leq y_1,....,\eta(\boldsymbol{s}_n,t_n)\leq y_n\right)=P\left(\zeta(\boldsymbol{s}_1,t_1)\leq y_1,....,\zeta(\boldsymbol{s}_n,s_n)\leq y_n\right), \tag{3.17}$$

with $y_1,...y_n\in\mathbb{R}$. *First, we can now determine a closed expression of the n-dimensional distribution of the random field* $(\eta(\boldsymbol{s},t))_{(\boldsymbol{s},t)\in\mathbb{R}^d\times\mathbb{R}_0^+}$ *since Lemma 3.10 holds for every random field* $\boldsymbol{X}$ *on* $\mathbb{R}^{d+1}$. *Let* $(\boldsymbol{s}_i,t_i)\in\mathbb{R}^d\times\mathbb{R}_0^+, i=1,...,n$ *and* $y_i\in\mathbb{R}, i=1,...,n$ *arbitrary, then*

$$P\left(\eta(\boldsymbol{s}_1,t_1)\leq y_1,....,\eta(\boldsymbol{s}_n,t_n)\leq y_n\right)=\exp\left(-\underbrace{E\left[\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)\right]}_{=(\star)}\right).$$

What we have to do is, compare $(\star)$ *with the result, if we apply Lemma 3.10 on the fields*

$$X_i(\boldsymbol{s},t)=Z_i(\boldsymbol{s}-\boldsymbol{S}_i,t-T_i)-\log\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z_i\left(\boldsymbol{u}-\boldsymbol{S}_i,v-T_i\right)\right)\mu_2(dv)\mu_1(d\boldsymbol{u})\right). \tag{3.18}$$

This is where Lemma 3.11 presents a method to rewrite $(\star)$. *If we choose for the Functional the following*

$$F(x)=\frac{\max_{j=1,...,n}\exp\left(x(\boldsymbol{s}_j,t_j)-y_j\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(x(\boldsymbol{a},b)\right)\mu_2(db)\mu_1(d\boldsymbol{a})}=\frac{\exp\left(\max_{j=1,...,n}x(\boldsymbol{s}_j,t_j)-y_j\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(x(\boldsymbol{a},b)\right)\mu_2(db)\mu_1(d\boldsymbol{a})},$$

where we can switch the maximum and the exponential function, since the exponential func-

tion is a strictly monotone function, we got for the reformulation of ($\star$):

$$
\begin{aligned}
&E\left[\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)\right]\\
&=E\left[\frac{E\left[\exp(W(\boldsymbol{u},v)-\sigma^2(\boldsymbol{u},v)/2)\right]}{E\left[\exp(W(\boldsymbol{u},v)-\sigma^2(\boldsymbol{u},v)/2)\right]}\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)\right]\\
&=\int_\Omega\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp(W(\boldsymbol{u},v)-\sigma^2(\boldsymbol{u},v)/2)\cdot\\
&\qquad\frac{\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(W(\boldsymbol{a},b)-\sigma^2(\boldsymbol{a},b)/2\right)\mu_2(db)\mu_1(d\boldsymbol{a})}\mu_2(dv)\mu_1(d\boldsymbol{u})dP\\
&\overset{1)}{=}\int_{\mathbb{R}^d}\int_{\mathbb{R}}\int_\Omega\exp(W(\boldsymbol{u},v)-\sigma^2(\boldsymbol{u},v)/2)\cdot\\
&\qquad\frac{\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(W(\boldsymbol{a},b)-\sigma^2(\boldsymbol{a},b)/2\right)\mu_2(db)\mu_1(d\boldsymbol{a})}dP\mu_2(dv)\mu_1(d\boldsymbol{u})\\
&=\int_{\mathbb{R}^d}\int_{\mathbb{R}}E\left[\exp\left(W(\boldsymbol{u},v)\right.\right.\\
&\qquad\left.\left.-\sigma^2(\boldsymbol{u},v)/2\right)\frac{\exp\left(\max_{j=1,..,n}\left(W(\boldsymbol{s}_j,t_j)-\sigma^2(\boldsymbol{s}_j,t_j)/2-y_j\right)\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(W(\boldsymbol{a},b)-\sigma^2(\boldsymbol{a},b)/2\right)\mu_2(db)\mu_1(d\boldsymbol{a})}\right]\mu_2(db)\mu_1(d\boldsymbol{a})\\
&\overset{2)}{=}\int_{\mathbb{R}^d}\int_{\mathbb{R}}E\left[\frac{\exp\left(\max_{j=1,...,n}\left(Z(\boldsymbol{s}_j-\boldsymbol{u},t_j-v)-y_j\right)\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z(\boldsymbol{a}-\boldsymbol{u},b-v)\right)\mu_2(db)\mu_1(d\boldsymbol{a})}\right]\mu_2(dv)\mu_1(d\boldsymbol{u})\\
&=\underbrace{E\left[\frac{\exp\left(\max_{j=1,..,n}(Z(\boldsymbol{s}_j-\boldsymbol{S},t_j-T)-y_j)\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z(\boldsymbol{a}-\boldsymbol{S},b-T)\right)\mu_2(db)\mu_1(d\boldsymbol{a})}\right]}_{=(\star\star)},
\end{aligned}
$$

where we used in 1) Fubini since all of the expression in integral are positive and in 2) Lemma 3.11. We compare this last part to the Application of Lemma 3.10 on (3.16), which results in the following :

$$
\begin{aligned}
&P\left(\zeta(\boldsymbol{s}_1,t_1)\le y_1,....,\zeta(\boldsymbol{s}_n,t_n)\le y_n\right)\\
&=\exp\left(-E\left[\exp\left(\max_{j=1,...,n}Z(\boldsymbol{s}_j-\boldsymbol{S},t_j-T)\right.\right.\right.\\
&\qquad\left.\left.\left.-\log\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z(\boldsymbol{a}-\boldsymbol{S},b-T)\right)\mu_2(db)\mu_1(d\boldsymbol{a})\right)-y_j\right)\right]\right)\\
&=\exp\left(-E\left[\frac{\exp\left(\max_{j=1,...,n}Z(\boldsymbol{s}_j-\boldsymbol{S},t_j-T)-y_j\right)}{\exp\left(\log(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp(Z(\boldsymbol{a}-\boldsymbol{S},b-T))\mu_2(db)\mu_1(d\boldsymbol{a}))\right)}\right]\right)\\
&=\exp\left(-\underbrace{E\left[\frac{\exp\left(\max_{j=1,...,n}Z(\boldsymbol{s}_j-\boldsymbol{S},t_j-T)-y_j\right)}{\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp(Z(\boldsymbol{a}-\boldsymbol{S},b-T))\mu_2(db)\mu_1(d\boldsymbol{a})}\right]}_{=(\star\star\star)}\right)
\end{aligned}
$$

Looking simultaneously at ($\star\star$) and ($\star\star\star$) we infer that the distributions of the random fields $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ and $(\zeta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ are the same. $\square$

Remark 3.12. *There is a continuum of random fields with the same distribution as η, one for each measure μ.*

In Section 3.3.4 we will say something to the choice of the measure μ.

Remark 3.13. *Under the Assumptions of Theorem 3.8, it is known that the field* $\left(\zeta^{'}(\boldsymbol{s},t)\right)_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ *with*

$$\zeta^{'}(\boldsymbol{s},t) = \sup_{i\geq 1}\left(V_i + Z_i(\boldsymbol{s}-\boldsymbol{S}_i, t-T_i)\right) \ , \quad (\boldsymbol{s},t)\in\mathbb{R}^d\times\mathbb{R}_0^+ \tag{3.19}$$

also has the same distribution as $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$. *Although the fields* $\boldsymbol{\zeta}$ *and* $\boldsymbol{\zeta}^{'}$ *differ due to the additional log-term, the theorem states they have the same distribution. This surprising fact becomes perhaps more plausible after noting that, for every* i

$$\log E\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(Z_i(\boldsymbol{u}-\boldsymbol{S}_i, v-T_i)\right)\mu_2(dv)\mu_1(d\boldsymbol{u})\right) = 0$$

Remark 3.14 (see Dieker and Mikosch [20], Remark 2.5)**.** *If* $\sigma^2/2 = \delta$, *then* $(W_i(\boldsymbol{s}-\boldsymbol{S}_i, t-T_i))$ *has the same distribution as* $(W_i(\boldsymbol{s},t) - W_i(\boldsymbol{S}_i,T_i))$. *The term* $W_i(\boldsymbol{S}_i,T_i)$ *drops out of the expression for* $\boldsymbol{\zeta}$, *so in that case the random field*

$$\sup_{i\geq i}\Big(V_i + W_i(\boldsymbol{s},t) - \delta(\boldsymbol{s}-\boldsymbol{S}_i, t-T_i) - \log\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}}\exp\left(W_i(\boldsymbol{u},v) - \delta(\boldsymbol{u}-\boldsymbol{S}_i, v-T_i)\right)\mu_2(dv)\mu_1(d\boldsymbol{u})\right)\Big) \ , \quad t\in\mathbb{R}^d$$

also has the same distribution as $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$.

Whereas Remark 3.13 seems at first sight, as an improvement, because the logarithm-term drops out, only with Remark 3.14 we have a crucial computational advantage for the algorithm which we want to present next.

3.3.2 Implementation

The theoretical background 3.3.1 gives us a new representation for Brown-Resnick random fields, where a change of measure argument was centrical. As in the methods before we still are confronted with an supremum over infinite many functions, which is not computable.

In the following we outline an algorithm, which gives us the opportunity to avoid an supremum over infinite many functions and so to avoid the choice of a suitable number N of Gaussian fields, which would lead to an approximation as in in the Gaussian approximation method and the naive approach.

Again, the algorithm focus on a discrete set of points $(\boldsymbol{s}_1,t_1)...,(\boldsymbol{s}_n,t_n)\in\mathbb{R}^d\times\mathbb{R}_0^+$. For now let μ be uniform on $(\boldsymbol{s}_1,t_1)...,(\boldsymbol{s}_n,t_n)\in\mathbb{R}^d\times\mathbb{R}_0^+$. To achieve the mentioned advantage of Remark 3.14, we choose for the measure μ the uniform distribution, with support on the grid points.

Remark 3.15 (see Dieker and Mikosch [20], page 4)**.** *Thereom 3.9 and the corresponding Remark 3.14 shows that the vector* $(N(\boldsymbol{s}_1,t_1),...,N(\boldsymbol{s}_n,t_n))$ *with,*

$$N(\boldsymbol{s}_j,t_j) = \sup_{i\geq 1}[V_i + W_i((\boldsymbol{s}_j,t_j) - \delta((\boldsymbol{s}_j-\boldsymbol{S}_i,t_j-T_i) - \log\left(\frac{1}{n}\sum_{l=1}^{n}\exp\left(W_i((\boldsymbol{s}_l,t_l)-\delta((\boldsymbol{s}_l-\boldsymbol{S}_i,t_l-T_i))\right)\right)], \tag{3.20}$$

for $j=1,...,n$*, has the same distribution as* $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$*, where* $((V_i,\boldsymbol{P}_i))_{i\geq 1}$ *belong to a Poisson process on* $\mathbb{R}\times\{(\boldsymbol{s}_1,t_1),...,(\boldsymbol{s}_n,t_n)\}$ *with intensity measure* $e^{-x}dx\times(n^{-1}\sum_{i=1}^{n}\mathbb{1}_{\boldsymbol{s}_i,t_i}(dy))$*.*

For simulation of space-time processes we look on the case $d=2$ as in the chapters before. For simplicity and to get the essential idea of the algorithm we rewrite (3.20) into

$$N(\boldsymbol{s}_j,t_j) = \sup_{i\geq 1}[V_i + \log(n) + W_i((\boldsymbol{s}_j,t_j) - \delta((\boldsymbol{s}_j-\boldsymbol{S}_i,t_j-T_i) - \log\left(\sum_{l=1}^{n}\exp\left(W_i((\boldsymbol{s}_l,t_l)-\delta((\boldsymbol{s}_l-\boldsymbol{S}_i,t_l-T_i))\right)\right)]. \tag{3.21}$$

The motivation for the recalculation is the following: Concentrating in (3.21) just on the cluster of points $\{C_i(\boldsymbol{s}_j,t_j): j=1,...,n\}$ with

$$C_i(\boldsymbol{s}_j,t_j) = V_i + \log(n) + W_i((\boldsymbol{s}_j,t_j) - \delta((\boldsymbol{s}_j-\boldsymbol{S}_i,t_j-T_i) - \log\left(\sum_{l=1}^{n}\exp\left(W_i((\boldsymbol{s}_l,t_l)-\delta((\boldsymbol{s}_l-\boldsymbol{S}_i,t_l-T_i))\right)\right), \tag{3.22}$$

we only have to focus on the simulation of the Poisson points V_i and the Gaussian fields W_i. But the advantage of this representation is the connection between C_i and $V_i+\log n$.

Corollary 3.16. *For* C_i *defined as in (3.22) the following inequality holds*

$$C_i(\boldsymbol{s}_j,t_j) \leq V_i + \log n, \quad \forall\ j=1,...,n.$$

Proof: *Since*

$$C_i(\boldsymbol{s}_j,t_j) = (V_i+\log n) + \underbrace{W_i(\boldsymbol{s}_j,t_j) - \delta\left(\boldsymbol{s}_j-\boldsymbol{S}_i,t_j-T_i\right)}_{x_j} - \log\left(\sum_{l=1}^{n}\exp\left(W_i(\boldsymbol{s}_l,t_l)-\delta\left(\boldsymbol{s}_l-\boldsymbol{S}_i,t_l-T_i\right)\right)\right),$$

it suffices to show, that

$$x_j < \log\left(\sum_{l=1}^{n}\exp\left(x_l\right)\right).$$

This is equivalent to

$$\exp\left(x_j\right) < \sum_{l=1}^{n}\exp\left(x_l\right),$$

where we used that the exponential function is a strictly increasing monotone function. □

Using 3.15, we can claim, that only a finite number of points (V_i, C_i) has to be simulated. One reason why this holds, is that the points ($(V_i + \log(n), C_i)$ are generated in decreasing order of $(V_i + \log(n))$. So we can stop the algorithm when

$$\max_{i=1,..,N} C_i(\boldsymbol{s}_j, t_j) > V_{N+1} \log(n), \quad \forall\ j = 1, ..., n$$

because every C_i for $i > N + 1$ won't change the maximum, since for all $i \geq N + 1$

$$C_i(\boldsymbol{s}_j, t_j) \leq V_i + \log(n) \leq V_{N+1} + \log(n) < \max_{i=1,...,N} C_i(\boldsymbol{s}_j, t_j), \text{ for all } j = 1, ..., n,$$

so that we can conclude

$$N(\boldsymbol{s}_j, t_j) = \sup_{i \geq 1} C_i(\boldsymbol{s}_j, t_j) = \sup_{i \leq N} C_i(\boldsymbol{s}_j, t_j).$$

An important part of this algorithm is the decreasing order in the simulation of the points V_i of the Poisson process V on $\mathbb{R}$ with intensity measure $e^{-v}dv$. An approximation of these V_i is given in Oesting et al. [37] Chapter 5.

Theorem 3.17 (Simulation of the V_i)**.** *Let $Y_1, Y_2, ...$ be independent exponentially distributed random variables with parameter $\lambda > 0$. Consider $S_n = \sum_{i=1}^{n} Y_i$ for every $n \in \mathbb{N}$. Then, $\sum_{i \geq 1} \mathbb{1}_{S_i}$ is a Poisson process on $(0, \infty)$ with intesity λ. Using the mapping theorem for Poisson processes (see for example Kingman [33]), we can say that $V = \sum_{i \in \mathbb{N}} \mathbb{1}_{-\log(S_i)}$ is a Poisson process on $\mathbb{R}$ with intensity measure $\lambda \exp(-v)dv$ and the sequence $(V_i)_{i \in \mathbb{N}}$ with $V_i = -\log(S_i)$ is monotonically decreasing.*

Remark 3.18. *In the current setting is $\lambda = 1$. To accelerate the algorithm, the $Y_1, Y_2, ...$ are computed with the help of probability integral transform method (Theorem 1.7), to avoid to simulate random exponetially distributed variables. This results in:*

$$Y_1, Y_2, ... = -\log(U_1), -\log(U_2), ... \quad \textit{with } U_i \sim \mathcal{U}[0, 1], \quad \textit{for } i \in \mathbb{N}.$$

Following Theorem 3.16 we get for V_1 and the following $V_i \geq 2$

$$\begin{aligned} V_1 &= -\log\left(-\log(U_1)\right) \\ V_{i+1} &= -\log\left(\sum_{j=1}^{i+1} -\log(U_j)\right) = -\log\left(\sum_{j=1}^{i+1} Y_j\right) = -\log\left(S_i + Y_{i+1}\right), \quad \textit{for } i \geq 1. \end{aligned}$$

Especially $V_{i+1} = -\log(S_i + Y_{i+1})$ gives a good intuition, how V_{i+1} and V_i differ and why V_i is a monotone decreasing sequence of points.

To summarize this implementation Section 3.3.2, we give a short intuitive dummy code for the algorithm by Dieker and Mikosch.

Intuitive Dummy Code for The Method of Dieker and Mikosch:

- Simulate the first point V_1 of the Poisson process V and a first Gaussian random field W_1 with variogram δ.
- Calculate the first clusterpoint C_1 defined as in (3.22) depending on V_1, W_1 and its variogram δ .
- Simulate a second point V_2 of the Poisson process V depending on V_1 and calculated as in Remark 3.18. Check if $V_2 + \log(n) < C_1(\boldsymbol{s}_j, t_j), \forall\ j = 1, ..., n$. If so, stop the algorithm. If not, simulate a second Gaussian random field W_2 and calculate C_2 depending on V_2, W_2 and δ. Define $S_2^{(j)} = \max_{i\in\{1,2\}} C(\boldsymbol{s}_j, t_j),\ j = 1, .., n$ which takes the supremum of the two existing clusterpoint in every coordinate.
- Repeat this iteration but from now on check $V_{i+1} + \log(n) < S_i^{(j)}$, for all $j = 1, ..., n$ and $i \geq 2$.
- Let K be the first time it is fulfilled, so $V_K + \log(n) < S_{K-1}^{(j)}$, for all $j = 1, ..., n$ we can determine N as in (3.21) with $N(\boldsymbol{s}_j, t_j) = S_{K-1}^{(j)}$, for all $j = 1, ..., n$ which would be a simulation of the Brown-Resnick process $(\eta(\boldsymbol{s}, t))_{\boldsymbol{s},t\in\mathbb{R}^d\times\mathbb{R}_0^+}$ with Gumbel distributed margins.

A visualization of the method by Dieker and Mikosch is also given in Figure 3.15 (see also [20], Figure 1). Computational effort and max-stability of this algorithm will discussed in the next Section 3.3.3.

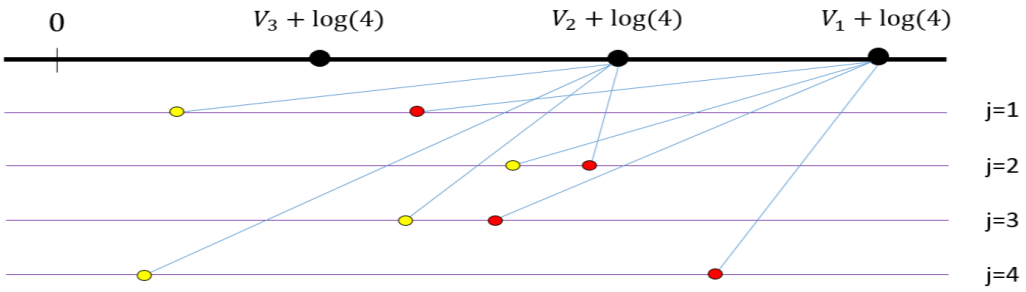

Figure 3.15: Visualization of the method by Dieker and Mikosch, for a space-time grid with $n = 4$ points. Each horizontal line below the axis (black line) stands for a value of j, and the corresponding dots on these lines for the cluster points $C_i, i = 1, 2$. Red dots stand for the current maximum on this space-time point.

3.3.3 Results of the method by Dieker and Mikosch

Dieker and Mikosch provide an algorithm to simulate exact Brown-Resnick fields in theory. We now want to show and test this, by running this algorithm. We concentrate again on a space-time field, where we assume two spatial dimensions ($d = 2$).

Our first survey of the algorithm is the number of cluster points C_i - defined as in (3.22) - we need to calculate, to get a Brown-Resnick field $(\eta(\boldsymbol{s}, t))_{\boldsymbol{s},t \in \mathbb{R}^d \times \mathbb{R}_0^+}$ with $n = n_1 \times n_2 \times n_3$ space-time points. For simplicity we sometimes call this number N_C. We start with 250 simulations of a space-time field with $n = 125 = 5 \times 5 \times 5$. To classify this result we do another 100 simulations with $n = 500 = 10 \times 10 \times 5$.. The corresponding Gaussian fields are simulated as in Remark 3.3.

Some statistics for the **number of cluster points** C_i are given in Table 3.1.

Table 3.1 gives a summary, with how much simulation effort we have to deal. We start with the smaller field with $n = 125$. The number N_C of possible cluster points lies between 9 and 935, which is almost $8 \cdot n$. The maximum of cluster points is even higher for the bigger field with $n = 500$, where up to 7223 C_i's can be observed, which is equivalent to almost $15 \cdot n$.

If we focus on the upper quantiles, we can observe that the high maximas can be considered as outliers, since for both fields 75 % of the observations are underneath $5 \cdot n$. This observation is also visualized in Figure 3.16. In both histograms lies some significantly probability mass in the right tail, after drastic falls around the upper quantiles.

Statistical key figures for N_C		
	$\mathbf{5 \times 5 \times 5}$	$\mathbf{10 \times 10 \times 5}$
$\boldsymbol{n}$	125	500
mean	269	1500
standard deviation	179	1065
min	9	51
max	935	7223
lower quantile	121	796
median	245	1282
upper quantile	374	2068

Table 3.1: Summary for the observed number N_C of cluster points C_i (3.22) for 250 simulation of a $5 \times 5 \times 5$-field and 100 simulations of a $10 \times 10 \times 5$-field.

These observations don't contradict the statement (see Dieker and Mikosch [20], page 5), concerning the convergence order of the cluster points. To get a impression how many points of V are needed to simulate, which is quivalent to the number of cluster points C_i consider the degenerate case, $(\boldsymbol{s}_1, t_1) = ... = (\boldsymbol{s}_n, t_n) = (\boldsymbol{s}, t)$. We get $C_i\,(\boldsymbol{s}_j, t_j)$ for $j = 1, ..., n$ so the algorithm stops after generating $\inf\{M : V_M + \log(n) < V_1\}$ points of V and for large n, this implies that the number of cluster points is of order n, as we observe. Of order n means in this context, that N_C can deviate from n, but only in a linear sense, so N_C is of $\mathcal{O}(n)$.

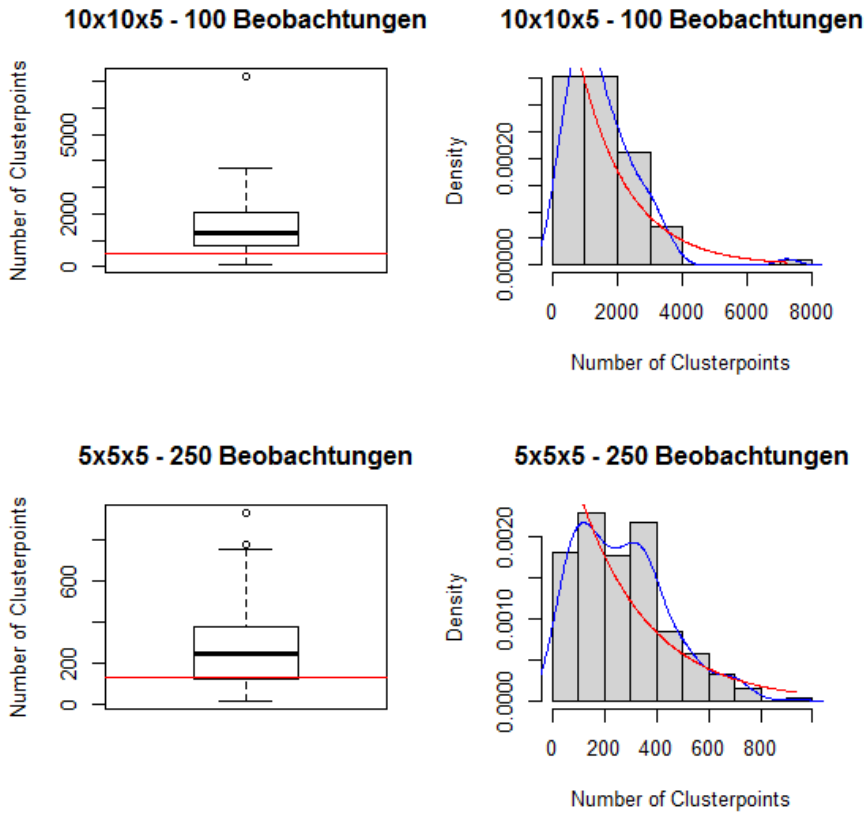

Figure 3.16: Box Plot and Histogram for the number N_C of cluster points C_i as in (3.22) for a $5 \times 5 \times 5$-field and a $10 \times 10 \times 5$-field. In the box plot, the red line stands for the number of space-time points n. Whereas in the histogram, the red line is the fitted exponential density and the blue line the actually observed density.

We can summarize, that we are not able to calculate any simulation effort, because we have to deal with a big variance in the number N_C of cluster points. So we loose this advantage, which we have in the two methods before. Based on Table 3.1 and on Figure 3.16, we can suggest at most graphically the upper quantiles as an upper bound for the numbers of generated cluster points in the algorithm for a Brown-Resnick field $(\eta(\boldsymbol{s},t))_{\boldsymbol{s},t\ \in\mathbb{R}^d\times\mathbb{R}_0^+}$. So instead of simulating arbitrary number of C_i's until

$$V_K + \log(n) \leq \max_{i=1,\ldots,K_1} C_i(\boldsymbol{s}_j, t_j),\ \forall\ j = 1, ..., n,$$

we could stop the algorithm at latest after $5 \cdot n$ and drop the simulation.

We now want to check, if the method provides an improvement **exactness**, as Dieker and Mikosch promise, compared to the Gaussian approximation method and the naive approach. As in the methods before, we start to asses the goodness of the marginal fits by qq-plot of the simulated quantiles versus the standard Gumbel quantiles for the same random location at the same time point as in Figure 3.1 and 3.12. We now look at Figure 3.17. In all four plots, the deviation from the Gumbel quantiles is actually very small, as it would be expected by an „exact“ algorithm.

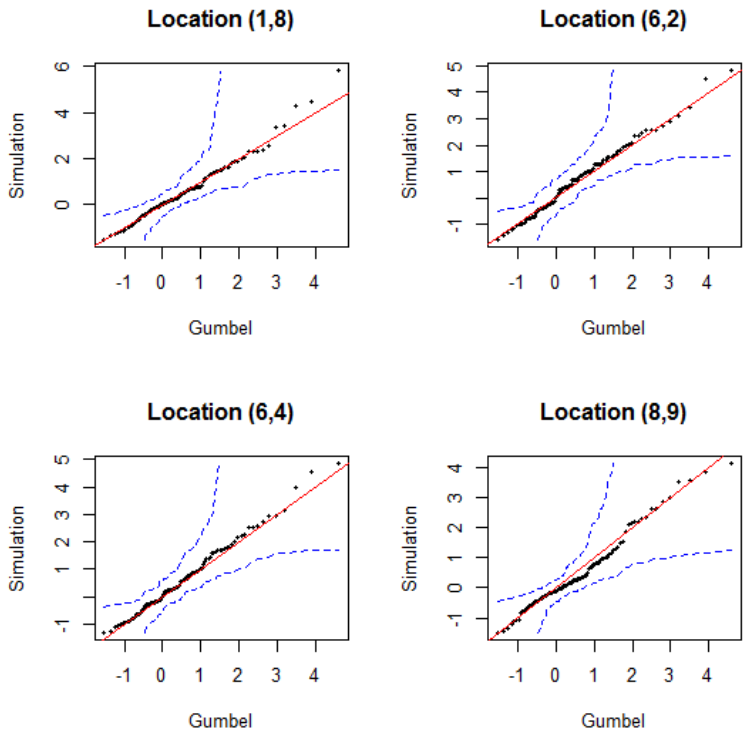

Figure 3.17: qq-plot of univariate margins of the log-transformed simulated process by the method of Dieker and Mikosch for a $10 \times 10 \times 5$-field versus the standard Gumbel distribution for four random space-time locations: (1,8,1), (6,2,1), (6,4,1), (8,9,1). The red line would be no deviation and the blue lines are the 95 % confidence bounds.

When we look at Figure 3.18, the exactness assumption also holds for the max-stability test explained in Remark 3.4. For cardinalities $K = 2, ..., 5$ the bisecting lines lies within the bootstrap confidence bounds. Again, as for the Gaussian approximation method, we can not reject the null hypothesis.

The improvement in exactness continues in the empirical bivariate distribution function. If we look at Figure 3.19 we can claim that

$$P\left(\eta^{TH}(\boldsymbol{s}_1, t_1) \leq y_1, \eta^{TH}(\boldsymbol{s}_2, t_2) \leq y_2\right) \text{ and } P\left(\eta^{EM}(\boldsymbol{s}_1, t_1) \leq y_1, \eta^{EM}(\boldsymbol{s}_2, t_2) \leq y_2\right),$$

roughly coincide.

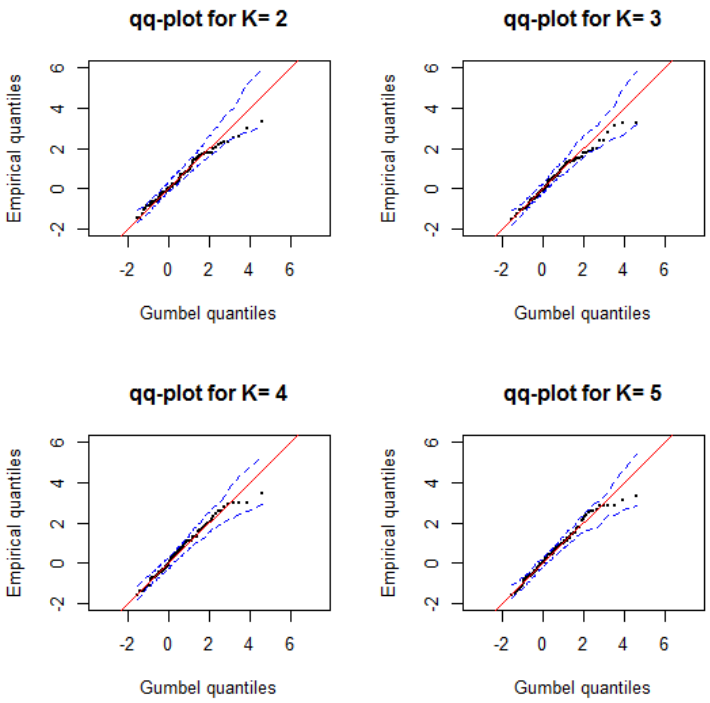

Figure 3.18: qq-plots of theoretical standard Gumbel quantiles versus maxima over subsets of the processes,simulated by the method of Dieker and Mikosch. From top left to bottom right the cardinality K of the subsets changes from $K = 2$ to $K = 5$. Blue dots describe the 95 % pointwise confidence bounds, obtained by block bootstrap. Red line represents no deviation.

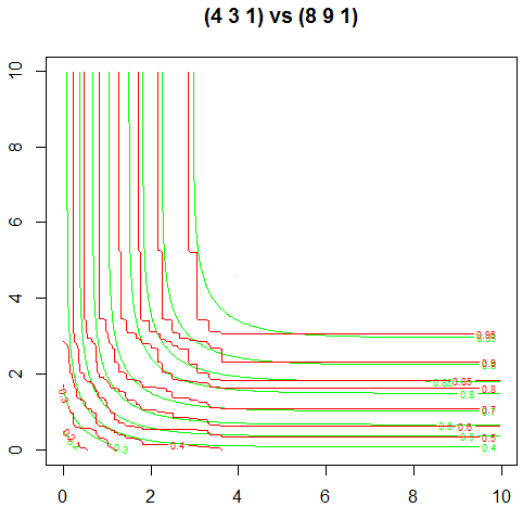

Figure 3.19: Contour Plot of the theoretical bivariate distribution of a Brown-Resnick process with Fréchet margins (green lines), compared to the empirical distribtuion of our Simulation with the method by Dieker and Mikosch (red lines.)

3.4 Method by Dombry, Engelke and Oesting

The following simulation method is based on the recently published paper from Dombry et al. [23]. At the time, this master thesis was written, the paper was still in progress. In the first part it actually takes the method by Dieker and Mikosch up and generalizes it to max-stable processes at a finite number of locations. The idea is based on simulating from the spectral measure. We leave this algorithm to the reader, because it is no significant improvement to the method by Dieker and Mikosch.

As to for the second algorithm in this paper, Dombry et. al. promise an improvement to the method by Dieker and Mikosch, especially regarding the number of cluster points have to be generated. We want to check if the second algortihm (which we from now call the method by Dombry, Engelke and Oesting) is really always more efficent than the method by Dieker and Mikosch.

The main idea basically is to only consider the **extremal functions**, which are those functions in the simulation of a Brown-Resnick process that effectively contribute to the pointwise maximum. We will give a brief introduction to the theoretical background and what Dombry et. al. understand exactly under extremal functions. Since the paper is written to more general max-stable fields, we will only focus on the parts, which we need for Brown-Resnick processes.

3.4.1 Theoretical background

We start with a general expression of a max-stable process $\{Z(\boldsymbol{s},t) : (\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}_0^+\}$. We consider a compact domain $\mathcal{X} \subset \mathbb{R}^d \times \mathbb{R}_0^+$. We can choose for the compact domain $\mathcal{X}$ for example as always the grid $\mathcal{X} = S_M \times T_K$, where $S_M = \{1, ..., M\}^d$ and $T_K = \{1, ..., K\}$. De Haan [1984] stated that every max stable process $\boldsymbol{Z}$ with unit Fréchet margins is characterized by a point process representation

$$Z(\boldsymbol{s},t) = \bigvee_{i \geq 1} V_i \psi_i(\boldsymbol{s},t), \quad (\boldsymbol{s},t) \in \mathcal{X}, \tag{3.23}$$

where $\{(V_i, \psi_i), i \geq 1\}$ is a Poisson process on $\mathbb{R}_0^+ \times C_+$ with intensity measure $V^{-2}dV \times \nu(d\psi)$ for some finite measure ν on the space C_+ which is the space of non-negative continious functions on $\mathcal{X}$ such that the following condtion is satisfied

$$\int_{C_+} \psi(\boldsymbol{s},t)\nu(d\psi) = 1, \quad (\boldsymbol{s},t) \in \mathcal{X}.$$

Again this spectral representation contains an infinite number of functions to determine the pointwise maximum. Dombry, Engelke and Oesting introduce now the concept of extremal functions to find a way to avoid this. For the sake of simplier notation, we introduce the following Lemma.

Lemma 3.19. *The C_+-valued point process $\Phi = \{\phi_i\}_{i \geq 1}$ with $\phi_i = V_i\psi_i$ is a Poisson point process with a combined intensity measure*

$$\mu(A) = \int_{C_+} \int_{\mathbb{R}_+^0} \mathbb{1}_{\{V\psi \in A\}} V^{-2} dV \nu(d\psi), \quad A \subset C_+.$$

It should be clear from the context, if we mean by Φ a Poisson process or the Gaussian distribution function.

Definition 3.20 (see Dombry et al. [23], Definition 1)**.** *Consider a compact subset $K \subset \mathcal{X}$, which is not the empty set. We call a function $\phi \in \Phi$ K-extremal if there is some $(\boldsymbol{s}, t) \in K$ such that $\phi(\boldsymbol{s}, t) = Z(\boldsymbol{s}, t)$. If there does not exist any $(\boldsymbol{s}, t)$ which fulfils the condition we call the function K-subextremal. We denote by Φ_K^+ the set of K-extremal functions and by Φ_K^- the set of K-subextremal functions.*

Remark 3.21. *It follows immediately from Definition 3.20 that we have the decomposition $\Phi = \Phi_K^+ \cup \Phi_K^-$.*

An example of the compact subset K can be for instance several points $\{\boldsymbol{x}_0, ..., \boldsymbol{x}_k\} \subset S_M \times T_K$, which would result in the sets $\Phi^+_{\{\boldsymbol{x}_0,...,\boldsymbol{x}_k\}}$ and $\Phi^-_{\{\boldsymbol{x}_0,...,\boldsymbol{x}_k\}}$. The algorithm otherwise is based on the idea to look just at one point $K = \{\boldsymbol{x}_i\}$ for $i \in 1, ..., k$.

When we reduce K to only a single point then $\Phi^+_{\{\boldsymbol{x}_i\}}$ is almost surely reduced to a single point that, which we denote by $\phi^+_{\boldsymbol{x}_i}$. Regarding the algorithm, we are mainly interested in the distribution of $\phi^+_{\boldsymbol{x}_i}$.

Theorem 3.22 (see Dombry et al. [23], Proposition 1)**.** *Recall that $Z(\boldsymbol{x}_i)$ has univariate Fréchet distribution, where $\boldsymbol{x}_i = (\boldsymbol{s}_i, t_i)$. Furthermore the distribution of $\phi^+_{\boldsymbol{x}_i}/Z(\boldsymbol{x}_i)$ is*

$$P_{\boldsymbol{x}_i}(A) = P\left(\phi^+_{\boldsymbol{x}_i}/Z(\boldsymbol{x}_i) \in A\right) = \int_{C_+} \mathbb{1}_{\{f/f(\boldsymbol{x}_i)\in A\}} f(\boldsymbol{x}_i)\nu(df), \quad A \subset C_+$$

The reason why we do not sample directly from $\phi^+_{\boldsymbol{x}_i}$ but using a normalization by $Z(\boldsymbol{x}_i)$, coincides with the problem described in Chapter 3.2.3 for the naive approach.

Algorithm idea for $d = 2$

This first introduction of extremal functions is already enough to explain the idea of the simulation method of Dombry, Engelke and Oesting.

1. First consider the maxstable process $\boldsymbol{Z}$ defined as in (3.14) on locations $\boldsymbol{x}_1, ..., \boldsymbol{x}_N$ being space-time points of the grid $\mathcal{X} = S_M \times T_K$, with $S_M = \{1, ..., M\}^2$ and $T_K = \{1, ..., K\}$, such that $n = M \cdot M \cdot K$.

2. For $k = 1, ..., n$ we recall the extremal point process for our here specific grid (grid correspond to the set K in Definition 3.21) which results in

$$\Phi_k^+ = \Phi^+_{\{\boldsymbol{x}_1,...,\boldsymbol{x}_k\}} \text{ or equivalent } \Phi_k^+ = \{\phi^+_{\boldsymbol{x}_i}\}_{1\le i\le k}.$$

3. To avoid repetitions on the right hand side of the equivalent representation, we introduce the n-th step maximum process

$$Z_k(\boldsymbol{x}) = \max_{\phi\in\Phi_k^+} \phi(\boldsymbol{x}) = \max_{1\le i\le n} \phi^+_{\boldsymbol{x}_i}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{X},$$

   where we take the maximum componentwise and here we consider an arbitrary grid point in $\mathcal{X}$. At a specific grid point we can point out that $Z(\boldsymbol{x}_i) = \phi^+_{\boldsymbol{x}_i}$, which just

follows from the definition of an extremal function with K as a single grid point, and thus it is obvious that for grid points up to point k:

$$Z(\boldsymbol{x}_i) = Z_k(\boldsymbol{x}_i), \quad i = 1, ..., k.$$

So we don't have to simulate $\boldsymbol{Z}$ at locations $\boldsymbol{x} = (\boldsymbol{s}, t)$, but with the introduction of extremal functions it is enough to simulate Φ_n^+ to get an exact simulation of our max stable process $\boldsymbol{Z}$.

But what do we understand by k-th step maximum process? The algorithm is in such a way designed that we have n steps, equal to the number of the space-time points. In the first step, we start at the first grid-point $\boldsymbol{x_1}$ and simulate $\phi^+_{\boldsymbol{x}_1}$. We already introduced in Theorem 3.20 the distribution of $\phi^+_{\boldsymbol{x}_0}$ for an arbitrary point $\boldsymbol{x}_0 \in \mathcal{X}$ which we can use as an initial distribution for $\phi^+_{\boldsymbol{x}_1}$. After the first step, we have to deal with a conditional-simulation, because $\phi^+_{\boldsymbol{x}_i}, i = 1, ..., k-1$ influences $\phi^+_{\boldsymbol{x}_k}$. We want to summarize this distribution assumptions in the following theorem.

Theorem 3.23 (see Dombry et al. [23], Theorem 2)**.** *The initial distribution $\phi^+_{\boldsymbol{x}_1}$ is given as follows:*

- *The extremal function $\phi^+_{\boldsymbol{x}_1}$ has the same distribtuion as $F_1 Y_1$ where F_1 is a unit Fréchet random variable and Y_1 an independent random process with distribution given in Theorem 3.22.*

For every further step we look at the conditional distribution of $\phi^+_{\boldsymbol{x}_{k+1}}$ with respect to the simulation of step $1, ..., k$.

- *For $1 \leq k \leq n-1$, the distribution of $\phi_{\boldsymbol{x}^+_{k+1}}$ conditional to $\left(\phi^+_{\boldsymbol{x}_i}\right)_{1 \leq i \leq k}$ is equal to the distribution of*

$$\tilde{\phi}^+_{\boldsymbol{x}_{k+1}} = \begin{cases} argmax_{\phi \in \tilde{\Phi}_{k+1}} \, \phi\left(\boldsymbol{x}_{k+1}\right) & \text{if } \tilde{\Phi}_{k+1} \neq \emptyset \\ argmax_{\phi \in \Phi^+_k} \, \phi\left(\boldsymbol{x}_{k+1}\right) & \text{if } \tilde{\Phi}_{k+1} = \emptyset \end{cases} \tag{3.24}$$

where $\tilde{\Phi}_{k+1}$ is a Poisson point process with intensity

$$\mathbb{1}_{\{f(\boldsymbol{x}_i) < Z_n(\boldsymbol{x}_i),\ 1 \leq i \leq k\}} \mathbb{1}_{\{f(\boldsymbol{x}_{k+1}) > Z_n(\boldsymbol{x}_{k+1})\}} \mu(df) \tag{3.25}$$

Proof: *For the proof of Theorem see Dombry et al. [23]* □

We want to clarify (3.27). We start with the question, how does all other $\phi^+_{\boldsymbol{x}_k}, k = 2, ..., n$ deviate from the initial distribtuion $\phi^+_{\boldsymbol{x}_1}$. The simulation is done exactly the same, but since $\phi^+_{\boldsymbol{x}_k}$ depends on $\phi^+_{\boldsymbol{x}_1}, ..., \phi^+_{\boldsymbol{x}_{k-1}}$, we only accept the realisation $\phi^+_{\boldsymbol{x}_k}$, if it fulfils certain conditions, which are summarized in the intensity (3.27), otherwise we withdraw the realisation. The conditions are

$$\phi^+_{\boldsymbol{x}_k}(\boldsymbol{x}_i) < Z_{n-1}(\boldsymbol{x}_i) \text{ for } i = 1, ..., k-1,$$

and

$$\phi^+_{\boldsymbol{x}_k}(\boldsymbol{x}_k) > Z_{k-1}(\boldsymbol{x}_k).$$

3.4.2 Implementation

We now want to transfer Theorem 3.23 from general max-stable processes to Brown-Resnick processes. Recall the spectral representation from (3.25),

$$Z(\boldsymbol{s},t) = \max_{i\geq 1} V_i \psi_i(\boldsymbol{s},t), \quad (\boldsymbol{s},t) \in \mathcal{X}.$$

where $\{(V_i,\psi_i), i \geq 1\}$ is a Poisson process on $\mathbb{R}_0^+ \times C_+$ with intensity measure $V^{-2}dV \times \nu(d\psi)$. To get the Brown-Resnick process model which we assume for the Gaussian approximation method and the naive approach, ν is a probability measure given by

$$\nu(A) = P\left[\exp\left(W(\cdot) - \delta(\cdot)/2\right) \in A\right], \quad A \subset C_+ \text{ Borel.} \tag{3.26}$$

What is left to transfer to a Brown-Resnick process is Theorem 3.22 and the corresponding random process Y from Theorem 3.23.

Theorem 3.24 (see Dombry et al. [23], Proposition 6)**.** *For all $x_i \in \mathcal{X}$, the distribution $P_{\boldsymbol{x}_i}$ from Theorem 3.23 is equal to the distribution of the log-normal process*

$$Y(\boldsymbol{s},t) = \exp\left[W(\boldsymbol{s},t) - W(\boldsymbol{x}_i) - \frac{1}{2}Var\left[W(\boldsymbol{s},t) - W\left(\boldsymbol{x}_i\right)\right]\right], \quad (\boldsymbol{s},t) \in \mathcal{X}. \tag{3.27}$$

or since we assume that $\boldsymbol{W}$ is a Gaussian process with stationary increments, we can rewrite (3.29) into

$$Y(\boldsymbol{s},t) = \exp\left[W(\boldsymbol{s},t) - W(\boldsymbol{x}_i) - \delta\left[(\boldsymbol{s},t) - \boldsymbol{x}_i\right]\right], \quad (\boldsymbol{s},t) \in \mathcal{X}. \tag{3.28}$$

Proof: *We recall the measure $\mathcal{Q}$ from the proof of Lemma 3.10. The following proof is similar to the proof of Lemma 3.10 but here we use the fact that the distribution of the fractional Brownian field $(W(\boldsymbol{s},t))_{\boldsymbol{s},t\in\mathcal{X}}$ under the transformed probability measure $\mathcal{Q} = \exp\left(W(\boldsymbol{x}_i) - \sigma^2(\boldsymbol{x}_i)/2\right) dP$ is equal to the distribution of the Gaussian Process*

$$W(\boldsymbol{s},t) + Cov\left(W(\boldsymbol{x}_i), W(\boldsymbol{s},t)\right), \ (\boldsymbol{s},t) \in \mathcal{X},$$

(see for proof details Dombry et al. [23], Lemma 1). So for all Borel sets $A \subset C_+$ it holds that,

$$\begin{aligned}
P_{\boldsymbol{x}_i} &= \int_{C_+} \mathbb{1}_{\{f/f(\boldsymbol{s},t)\in A\}} f(\boldsymbol{s},t)\nu(df) \\
&= E\left[\exp\left(W(\boldsymbol{x}_i) - \frac{1}{2}\sigma^2(\boldsymbol{x}_i)\right)\mathbb{1}_{\{e^{W(\cdot)-\frac{1}{2}\sigma^2(\cdot)}/e^{W(\boldsymbol{x}_i)-\frac{1}{2}\sigma^2(\boldsymbol{x}_i)}\in A\}}\right] \\
&= \mathcal{Q}\left[\exp\left(W(\cdot) - W(\boldsymbol{x}_i) - \frac{1}{2}(\sigma^2(\cdot) - \sigma^2(\boldsymbol{x}_i))\right) \in A\right] \\
&= P\left[\exp\left(W(\cdot) + Cov(W(\boldsymbol{x}_i), W(\cdot)) - W\left(\boldsymbol{x}_i\right)\right.\right. \\
&\qquad\qquad \left.\left. - \ Cov(W(\boldsymbol{x}_i), W(\boldsymbol{x}_i)) - \frac{1}{2}\left(\sigma^2(\cdot) - \sigma^2(\boldsymbol{x}_i)\right)\right) \in A\right] \\
&= P\left[\exp\left(W(\cdot) - W(\boldsymbol{x}_i) - \frac{1}{2}\left(\sigma^2(\cdot) + \sigma^2(\boldsymbol{x}_i) - 2Cov\left(W(\boldsymbol{x}_i), W(\cdot)\right)\right)\right) \in A\right].
\end{aligned}$$

When we use

$$\sigma^2(\boldsymbol{s},t)+\sigma^2(\boldsymbol{x}_i)-2Cov\left(W(\boldsymbol{x}_i),W(\boldsymbol{s},t)\right)=Var\left[W(\boldsymbol{s},t)-W(\boldsymbol{x}_i)\right]$$

for every $(\boldsymbol{s},t)\in\mathcal{X}$ *as we did in the proof of Lemma 3.10 we get that* $P_{\boldsymbol{x}_i}$ *is equal to the distribution of the log-normal process*

$$Y(\boldsymbol{s},t)=\exp\left[W(\boldsymbol{s},t)-W(\boldsymbol{x}_i)-\frac{1}{2}Var\left[W(\boldsymbol{s},t)-W\left(\boldsymbol{x}_i\right)\right]\right],\quad(\boldsymbol{s},t)\in\mathcal{X}.$$

□

We want to summarize this transformation from max-stable processes to a Brown-Resnick process and the algorithm idea from section 3.4.1 to a intuitive Dummy code:

Intuitive Dummy Code for the Method by Dombry, Engelke and Oesting

Consider a space-time grid $\mathcal{X}=S_M\times T_K$, with $S_M=\{1,...,M\}^2$ and $T_K=\{1,...,K\}$, such that $n=M\cdot M\cdot K$. The algorithm contains n steps, with a corresponding $k-$step maximum for $1\leq k\leq n$

- Starting with the first step maximum, for the first space-time point. Simulate the first Poisson point V_1 as a standard Fréchet random variable and Y defined as in (3.30). Set the current maximum to $Z(\boldsymbol{s},t)=V_1Y$.
- Since the further steps are to complex to describe it in a sentence, we give a short dummy code script:

```
For k = 2,...,n:
  Simulate V_k
  while(V_k > Z(x_k)){
   Simulate Y
   If(VY(x_i) < Z(x_i) for all i = 1,...,k-1 {
    Update Z(s,t) by the componentwise max(Z(s,t),VY(s,t))
   }
   Simulate E sim Exp(1) and update 1/V_n by 1/V_n + E
  }
Return Z.
```

  By E we determine a exponential random variable with parameter equal to one. All V_k are standard Fréchet random variables and Y is always simulated as in the first step. The while loop can be interpret as an termination criterion to enable, that the if-statement can be done several time, to translate the intensity from (3.27).

Similar to the method by Dieker and Mikosch it is not interesting to determine the exact speed of this algorithm but to examine how many cluster points have to be generated. Recall the cluster points $C_i(\boldsymbol{s},t)$ from (3.22). The pendant in the method by Dombry, Engelke and Oesting to the cluster points, would be the number of generated random processes Y, which we define by N_Y.

We already stated in the introduction of section 3.4 that Dombry et al. [23] generalized also the algorithm of Dieker and Mikosch [20]. Furthermore they examine the expectation of N_C and N_Y.

Theorem 3.25 (see Dombry et al. [23], Proposition 4)**.** *On a space-time grid $\mathcal{X} = S_M \times T_K$, where $S_M = \{1, ..., M\}^2$ and $T_K = \{1, ..., K\}$ such that $n = M \cdot M \cdot K$ the expected number N_C of cluster points C_i defined in (3.22) and the expected number N_Y of random processes $\boldsymbol{Y}$ defined as in (3.30) are:*

Dieker and Mikosch: $E(N_C) = nE\left[\max_{(\boldsymbol{s},t)\in\mathcal{X}} Z(\boldsymbol{s},t)^{-1}\right]$

Dombry, Engelke and Oesting: $E(N_Y) = n$

Proof: *The proof for the expectation of N_C is similar as the proof of Proposition 4.8 in Oesting et al. [38]. Since it is a lot of computation, we only want to give a proof sketch for the expectation of N_Y (see for more details Dombry et al. [23], proof of Proposition 4). First we check, how N_Y is compound. In the k-step, we simulate Poisson points V_k and stochastic processes $\boldsymbol{Y}$ until one of the following conditions holds*

1. *$V_k < Z_{k-1}(\boldsymbol{x}_k)$. If this condition holds no stochastic process $\boldsymbol{Y}$ has to be simulated.*
2. *$V_k > Z_{k-1}(\boldsymbol{x}_k)$ and $V_k Y(\boldsymbol{x}_i) \leq Z(\boldsymbol{x}_i)$ for all $1 \leq i < k-1$. In this case, $\boldsymbol{Z}$ is updated and $V_k Y$ is an extremal function as it contributes to $\boldsymbol{Z}$ at location $\boldsymbol{x}_k$.*

With this in mind the number N_Y of random processes $\boldsymbol{Y}$ satisfies

$$N_Y = |\Phi^+_{\{\boldsymbol{x}_1,...,\boldsymbol{x}_k\}}| + \sum_{k=2}^{n} \left| \left\{ i \geq 1 : V_k^{(i)} > Z(\boldsymbol{x}_k), V_k^{(i)} > \min_{j=1}^{k-1} \frac{Z(\boldsymbol{x}_j)}{Y_k^{(i)}(\boldsymbol{x}_j)} \right\} \right|. \tag{3.29}$$

To determine the expectation $E(N_Y)$, one can start with the first summand

$$E\left[|\Phi^+_{\{\boldsymbol{x}_1,...,\boldsymbol{x}_k\}}|\right] = E\left[\max_{j=1}^{k} Y_k(\boldsymbol{x}_j)/Z(\boldsymbol{x}_j)\right].$$

A proof therefore can be found in Oesting et al. [38], Lemma 4.7. The second summand needs a lot more effort to calculate.

$$\begin{aligned} &E\left(\left|\left\{(V_k^{(i)}, Y_k^{(i)}) : V_k^{(i)} > Z(\boldsymbol{x}_k), V_k^{(i)} > \min_{j=1}^{k-1} \frac{Z(\boldsymbol{x}_j)}{Y_k^{(i)}(\boldsymbol{x}_j)}\right\}\right|\right) \\ &= E\left(E\left(|\left\{(V,Y) \in \Phi_k^{(1)} : V > Z(\boldsymbol{x}_k)\right\}| \,|\Phi^+_{\{\boldsymbol{x}_1,...,\boldsymbol{x}_{k-1}\}}, \Phi_k^{(2)}\right)\right) \end{aligned} \tag{3.30}$$

where the two sets

$$\begin{aligned} \Phi_k^{(1)} &= \left\{(V_k^{(i)}, Y_k^{(i)} : V_k^{(i)} Y_k^{(i)}(\boldsymbol{x}_j) > Z(\boldsymbol{x}_j) \text{ for some } j = 1, ..., k-1\right\} \\ \Phi_k^{(2)} &= \left\{(V_k^{(i)}, Y_k^{(i)} : V_k^{(i)} Y_k^{(i)}(\boldsymbol{x}_j) \leq Z(\boldsymbol{x}_j) \text{ for all } j = 1, ..., k-1\right\} \end{aligned}$$

are independent Poisson point processes with intensities $V^{-2}\mathbb{1}_{\{V > \min_{j=1}^{k-1}(Z(\boldsymbol{x}_j)/Y(\boldsymbol{x}_j))\}} dV P_{\boldsymbol{x}_k}(dY)$ and $V^{-2}\mathbb{1}_{\{V < \min_{j=1}^{k-1}(Z(\boldsymbol{x}_j)/Y(\boldsymbol{x}_j))\}} dV P_{\boldsymbol{x}_k}(dY)$. Further calculation leads to the expectation $E[N_Y] = N$. □

3.4.3 Results of the method by Dombry, Engelke and Oesting

We now want to examine like for the methods in the sections before, if the algorithm is max-stable and if we can say something about computational effort. First we want to pick up Theorem 3.25.

As for the method by Dieker and Mikosch, we do 250 simulations for a $5 \times 5 \times 5$ Brown-Resnick field and 100 simulations for a $10 \times 10 \times 5$ Brown-Resnick field. The results are summarized in Table 3.2, where we can see, that the claim that $E[N_Y] = n$ from Theorem 3.25 holds. For both fields size, the mean of N_Y over all simulations is nearly exactly the same as n , the number of space-time grid points. Also the standard deviation is way smaller, as for the method by Dieker and Mikosch. So we do not expect too many large outliers, as we also can see in the boxplot of Figure 3.21 . Where as for the method by Dieker and Mikosch, where we approximated the histogram with a fitted exponential distribution ($\lambda = \frac{1}{\text{mean}}$), the histogram for the method by Dombry, Engelke and Oesting, tends to be more Gaussian distributed, with expectation $\mu =$ mean and $\sigma =$ std. deviation.

Statistical key figures for N_Y		
	$\mathbf{5 \times 5 \times 5}$	$\mathbf{10 \times 10 \times 5}$
$\boldsymbol{n}$	125	500
mean	125	501
standard deviation	57	63
min	11	368
max	298	690
lower quantile	86	452
median	120	496
upper quantile	167	542

Table 3.2: Summary for the observed number N_Y of processes Y defined as in Theorem 3.24, for 250 simulations of a $5 \times 5 \times 5$-field and 100 simulations of a $10 \times 10 \times 5$-field.

If we compare the simulation results by Dieker and Mikosch and Dombry, Engelke and Oesting, we see a clear improvement, regarding of prediction in computational effort. Furthermore, the theoretical claim in Theorem 3.25, about the expectation of the number N_C of cluster points and the number N_Y of stochastic processes, can also be confirmed for the simulations.

For checking max-stability of the simulation, we do the same procedure as for the methods before. We start again with the univariate margins in Figure 3.22, do the max-stability test with graphical result in Figure 3.23 and check the empirical bivariate distribution function against the theoretical bivariate distribution function of a Brown-Resnick process. In all three figures we can not see any significant difference to the simulation results by Dieker and Mikosch in Figure 3.17, 3.18 and 3.19. So regarding exactness we have the same satisfying result for method by Dombry, Engelke and Oesting.

Since both the Dombry, Engelke and Oesting method and the Dieker and Mikosch method do not show any difference in exactness, we prefer the lastly presented method. The decision results from the fact, that not only is the method by Dombry, Engleke and Oesting faster

than the method by Dieker and Mikosch, we are also able to approximate the computational effort.

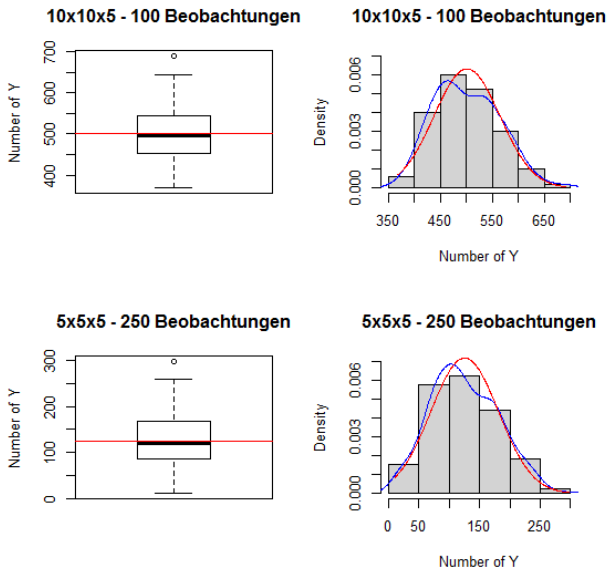

Figure 3.20: Box Plot and Histogram for the number N_Y of the stochastic process Y as in (3.27) for a $5 \times 5 \times 5$-field and a $10 \times 10 \times 5$-field. In the box plot, the red line stands for the number of space-time points n. Whereas in the histogram, the red line is the fitted Gaussian density and the blue line the actually observed density.

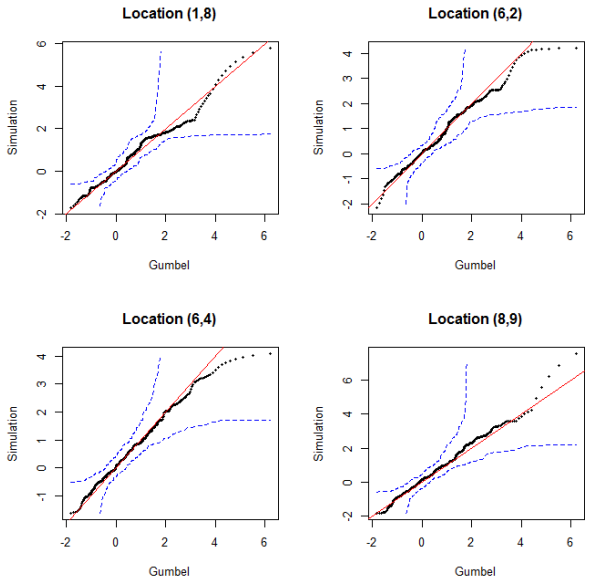

Figure 3.21: qq-plot of univariate margins of the log-transformed simulated process by the method of Dombry, Engelke and Oesting for a $10 \times 10 \times 5$-field versus the standard Gumbel distribution for four random space-time locations: (1,8,1), (6,2,1), (6,4,1), (8,9,1). The red line would be no deviation and the blue lines are the 95 % confidence bounds.

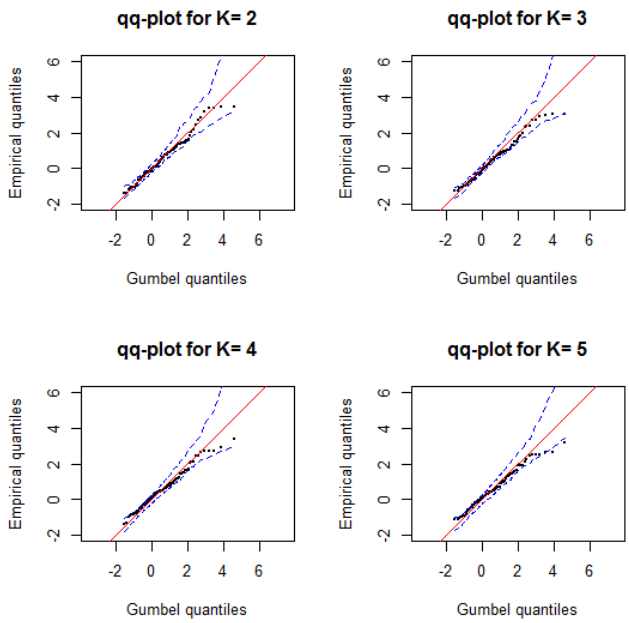

Figure 3.22: qq-plots of theoretical standard Gumbel quantiles versus maxima over subsets of the processes,simulated by the method of Dombry, Engelke and Oesting. From top left to bottom right the cardinality K of the subsets changes from $K = 2$ to $K = 5$. Blue dots describe the 95 % pointwise confidence bounds, obtained by block bootstrap. Red line represents no deviation.

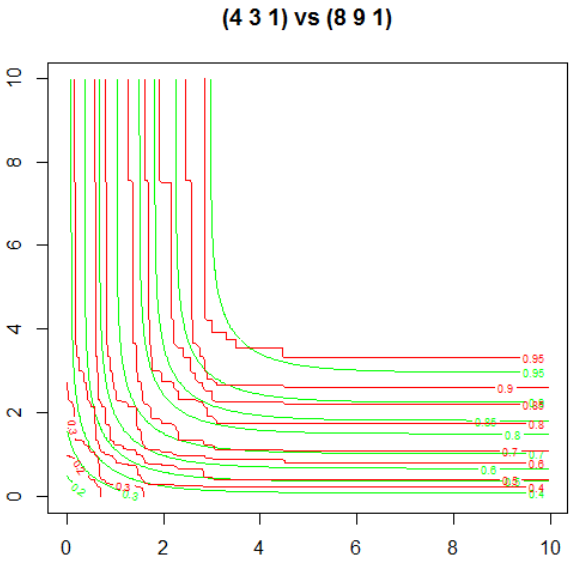

Figure 3.23: Contour Plot of the theoretical bivariate distribution of a Brown-Resnick process with Fréchet margins (green lines), compared to the empirical distribtuion of our Simulation with the method by Dombry, Engelke and Oesting (red lines.)

Chapter 4

Summary and Discussion

We saw a lot of different results during this thesis, so we want to give a short summary over the four different simulation methods. In our examination of the algorithms our main focus lies on the exactness, where we see a steady improvement from the first two methods - the Gaussian approximation method and the naive approach - where we have deviations from the theoretical model to the exact methods by Dieker and Mikosch [20] and the Dombry et al. [23].

There is one topic we have not yet discussed in detail, which would be the exact computation time. The reason why we did not specify the speed of the algorithm is simply that comparisons are not possible, because calculations were done with different computers. Since computational effort always depends on the choice of the computer, we see it more as an advantage, that we are able to predict the computational effort, no matter which computer is used, for the Gaussian approximation method, the naive approach and the method by Dombry, Engelke and Oesting. Only for the method by Dieker and Mikosch we are not able to determine a time horizon for calculations, because we observed huge variations (see again Table 3.1.) in computational effort.

In Remark 3.3 we mentioned that if we want to get a variogram as in (3.4),

$$\delta(h_1, h_2, u) = C_1|h_1|^{\alpha_1} + C_2|h_2|^{\alpha_2} + C_3|u|^{\alpha_3},$$

that we are restricted to only variability in the C_i's for $i \in \{1, 2, 3\}$, but the same fixed value for $\alpha_1 = \alpha_2 = \alpha_3 = 2$. At the end of the editing of this thesis and with the help of Dr. Marco Oesting from Universität Siegen we have been able to dissolve this restriction, and to do the simulation of the Brown-Resnick process with a variogram as in (3.4) for arbitrary $C_1, C_2, C_3 \in \mathbb{R}^+$ and $\alpha_1, \alpha_2, \alpha_3 \in (0, 2]$. The corresponding function can be seen in the appendix, where we give the space-time R-code for the algorithm by Dombry, Engelke and Oesting.

4.1 Goodness of fit Result

To finish this master thesis, we apply the most satisfying algorithm - the method by Dombry, Engelke and Oesting on a data analysis in Buhl and Klüppelberg [8]. They fitted a Brown-Resnick space-time process - which is the same, as we have for the Gaussian approximation method, the naive approach and the method for Dombry, Engelke and Oesting - to radar rainfall data, which were provided by the Southwest Florida Water Management District (SWFSWD). The data used for the analysis are rainfall measurements on a square of 120km $\times$ 120km in Florida over the years 1999-2004. The raw data consists of measurements in inches on a regular grid in space every two kilometres and every 15 minutes.

After some reduction and transformation of this data, e.g. taking block maxima in space and time (see for details Buhl and Klüppelberg [8], Section 5.1) Buhl and Klüppelberg apply the pairwise maximum likelihood estimation method to estimate the parameters of the function δ as in Remark 3.3, which are

$$C_1, C_2, C_3 \in \mathbb{R}^+ \quad \text{and} \quad \alpha_1, \alpha_2, \alpha_3 \in (0, 2].$$

Finally having fitted the Brown-Resnick space-time model to the precipitation data, they want to asses the quality of the fit. Similar as for the max-stability test of Remark 3.4, they consider subsets of the observations on a regular grid for J spatial locations and for time points $1, ..., B_1$,

$$D = \left\{ (s_1^{(l)}, s_2^{(l)}, 1), ..., (s_1^{(l)}, s_2^{(l)}, B_1) : l = 1, ..., J \right\}.$$

We extract R independent realisations out of the standard Gumbel transformed space-time observations $\{\eta_1(\boldsymbol{s}, t) : (\boldsymbol{s}, t) \in D\}$, where

$$\eta_1(\boldsymbol{s}, t) := \frac{\tilde{\eta}(\boldsymbol{s}, t) - \hat{\mu}}{\hat{\sigma}}, \quad t = 1, ..., 732,$$

where $\tilde{\eta}(\boldsymbol{s}, t)$ is the observed data, and $\hat{\mu}$ and $\hat{\sigma}$ are the corresponding empirical mean and empirical standard deviation. This yields in turn R independent realisations of the random variable $\eta_D = \max\{\eta_1(\boldsymbol{s}, t) : (\boldsymbol{s}, t) \in D\}$, which they summarise in the ordered vector $\eta_{data} = \left(\eta_D^{(1)}, ..., \eta_D^{(R)}\right)$. Now they simulate a corresponding vector, denoted by $\hat{\eta}_{sim} := \left(\hat{\eta}_D^{(1)}, ..., \hat{\eta}_D^{(R)}\right)$. To do this they need reliable Monte Carlo values as elements of $\hat{\eta}_{sim}$.

And here is the difference to Buhl and Klüppelberg. They do the simulations with the Gaussian approximation, where we now want to do this with the method by Dombry, Engelke and Oesting. We obatin the Monte-Carlo values by simulating empirical order statistics as follows. We simulate $m \cdot R$ independent copies of the Brown-Resnick space-time process on D with dependence structure δ as in Remark 3.3, with the pairwise likelihood estimates (see therefore Buhl and Klüppelberg [8], Table 5.2). We transform the univariate margins to standard Gumbel. This results in corresponding $m \cdot R$ independent simulations of η_D and we consider them as m blocks of size R. We order the R values in each block and define $\hat{\eta}_D^{(i)}$ as the mean of all simulated i-th order statistics for $i = 1, ..., R$, which gives $\hat{\eta}_{sim} := (\hat{\eta}_D^{(1)}, ..., \hat{\eta}_D^{(R)})$.

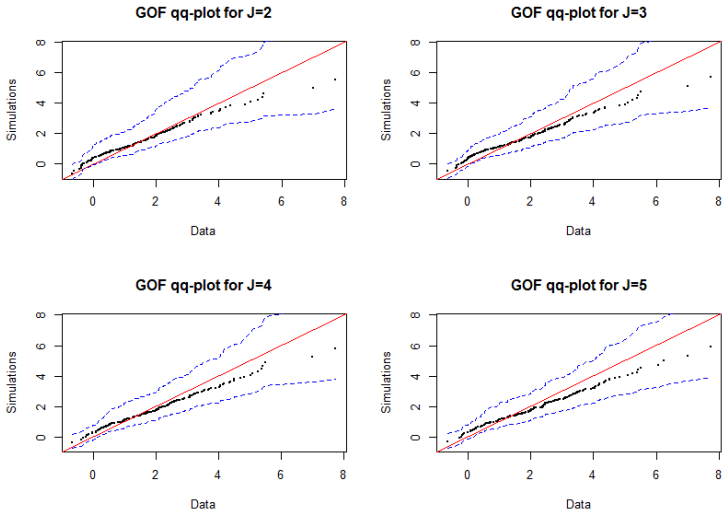

Figure 4.1: Goodness of fit qq-plots for different spatial locations and different J. Top left: $J = 2 : (2,5)$ and $(4,2)$. Top right: $J = 3 : (2,5), (4,2)$ and $(2,4)$. Bottom left: $J = 4 : (2,5), (4,2), (2,4)$ and $(5,2)$. Bottom right: $J = 5 : (2,5), (4,2), (2,4), (5,2)$ and $(4,4)$. PLME's underlying the simulations are based on maximum space-time lags 2. Blue lines mark 95 % pointwise confidence bounds. Red lines correspond to no deviation.

We now compare in Figure 4.1 the vectors η_{data} and $\hat{\eta}_{sim}$ by qq-plots. The fit is actually pretty good, since the black points in the plots lie approximately on the bisecting line. If we compare these plots to Figure 6 in Buhl and Klüppelberg [8], where $\hat{\eta}_{sim}$ is done with the Gaussian approximation method, we can see an improvement in Figure 4.1, regarding the small quantiles. Whereas for the qq-plots with the Gaussian approximation method the no deviation line does not always lie between the pointwise 95 %-confidence bounds, we can see in Figure 4.1 ,that this holds, if we use the method by Dombry, Engelke and Oesting.

Appendix

```
#Algorithm 2 from the paper "Exact simulation of max-stable processes"
#by C. Dombry, S. Engelke and M. Oesting 2015
#applied on the Brown-Resnick process, introduced in Chapter 5.2 for 3 dimensions
#with the variogram from Buhl and Cklu

require(RandomFields)
sim_BR <- function(alpha1, alpha2, alpha3, C1, C2, C3, x, y, z, n.BR)
{
  ## Setup
  RFoptions(spConform=FALSE)
  lx <- length(sx <- seq_along(x))
  ly <- length(sy <- seq_along(y))
  lz <- length(sz <- seq_along(z))

  ## Model-Variogram BuhlCklu
  modelBuhlCklu <- RMfbm(alpha=alpha1, var=C1, proj=1)
  + RMfbm(alpha=alpha2, var=C2, proj=2)
  + RMfbm(alpha=alpha3, var=C3, proj=3)


  ## Construct grid
  Nxy <- lx * ly
  N   <- Nxy * lz
  grid <- matrix(0, nrow=N, ncol=3) # (N,3)-matrix
  for (i in sx)
    for (j in seq_len(ly*lz))
      grid[i+(j-1)*ly, 1] <- i
  for (i in sy)
    for (j in sx)
      for(k in sz)
        grid[j+lx*(i-1)+(k-1)*Nxy, 2] <- i
  for (i in sz)
    for (j in seq_len(Nxy))
      grid[j+Nxy*(i-1), 3] <- i

  ## Construct shifted variogram
```

```
  Varm1 <- vapply(seq_len(N), function(n)
    RFvariogram(modelBuhlCklu,
                x=sx-grid[n,1],
                y=sy-grid[n,2],
                z=sz-grid[n,3]),
    array(NA_real_, dim=c(lx, ly, lz)))
  ## => (lx, ly, lz, N)-array

  ## Main
  Z <- array(, dim=c(lx, ly, lz, n.BR)) # 4d array
  E <- matrix(rexp(n.BR * N), nrow=n.BR, ncol=N)
  for (i in seq_len(n.BR))
  {
    ## n=1
    V <- 1/E[i,1]
    W <- RFsimulate(modelBuhlCklu, x, y, z, n=1)
    Y <- exp(W - W[1] - Varm1[,,,1])
    Z[,,,i] <- V * Y

    ## n in {2,..,N}
    for(n in 2:N)
    {
      Exp <- E[i,n]
      V <- 1/Exp
      while(V > Z[N*(i-1)+n])
      {
        W <- RFsimulate(modelBuhlCklu, x, y, z)
        Y <- exp(W - W[n] - Varm1[,,,n])
        if(all(V*Y[seq_len(n-1)] < Z[(N*(i-1)+1):(N*(i-1)+(n-1))]))
          Z[,,,i] <- pmax(V*Y, Z[,,,i])
        Exp <- Exp + rexp(1)
        V <- 1/Exp
      }
    }
  }

  ## Return
  Z
}
```

List of Figures

Bibliography

[1] R.J. Adler and J. Taylor. Random Fields and Geometry. 2005.

[2] A.A. Balkema and S. Resnick. Max-infinite divisibility. *Journal of Applied Probability*, 12(2):309–319, 1977.

[3] D. Barka and T. Mountford. A law of the iterated logarithm for fractional Brownian motions. *Séminaire de Probabilitiés*, XLI:161–179, 2008.

[4] J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels. *Statistics of Extremes: Theory and Applications*. John Wiley and Sons, Ltd, 2004.

[5] F. Biagini, Y. Hu, B. Oksenda, and T. Zhang. *Stochastic calculus for fractional Brownian motions and applications*. Springer, 2008.

[6] B.M. Brown and S.I. Resnick. Extreme values of independent stochastic processes. *Journal Applied Probabilities*, 14:732–739, 1977.

[7] S. Buhl. Modelling and estimation of extremes in space and time. *Master Thesis*, 2013.

[8] S. Buhl and C. Klüppelberg. Anisotropic Brown-Resnick space-time processes: estimation and model assessment. *Preprint*, 2015.

[9] X. Chen, W. V. Li, J. Rosinsiki, and Qi-Man Shao. Large deviations for local times and intersection local times of fractional Brownian motions and riemann-lioville processes. *The Annals of Probability*, 39:729–778, 2011.

[10] S. Cohen and J. Istas. *Fractional Fields and Applications*. Springer, 2013.

[11] S. Coles. *An introduction to statistical modeling of extreme values*. Springer Series in Statistics, 2001.

[12] S.G. Coles and J.A. Tawn. Statistics of Coastal Flood Prevention. *Philosophical Transactions: Physical Sciences and Engineering, Royal Society*, 332(1627):457–476, 1990.

[13] N.A.C. Cressie. *Statistics for spatial data*. John Wiley and Sons, Ltd, 1993.

[14] C. Czado. Stochastic Simulation and Algorithm with Applications. *Preprint*, 2007.

[15] R.A. Davis, C. Klüppelberg, and C. Steinkohl. Max-stable processes for modelling extremes observed in space and time. *J. Korean Statistical Society*, 42:399–414, 2013.

[16] L. de Haan. A spectral representation for max stable processes. *The Annals of Probability*, 12(4):1194–1204, 1984.

[17] L. de Haan and Ana Ferreira. *Extreme value theory - an introduction*. Springer, 2006.

[18] L. de Haan and S. I. Resnick. Limit theory for multivariate Sample Extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, by Springer Verlag*, 40:317–337, 1977.

[19] A. B. Dieker. Simulation of fractional Brownian motion. *Master Thesis*, 2002.

[20] A.B. Dieker and T. Mikosch. Exact simulation of Brown-Resnick random fields. *Extremes*, 18:301–314, 2015.

[21] A.B. Dieker and B. Yakir. On asymptotic constants in the theory of extremes for gaussian processes. *Bernoulli*, 20(3):1600–1619, 2014.

[22] C. Dombry, F. Evi-Minko, and M. Ribatet. Conditional simulations of Brown-Resnick processes. *Biometrika*, 100(1):111–124, 2013.

[23] C. Dombry, S. Engelke, and M. Oesting. Exact simulation of max-stable processes. *Avaiable from http://arxiv.org/abs/1506.04430*, 2015.

[24] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*. Springer, 1997.

[25] S. Engelke. Brown-Resnick processes: Analysis, Inference and Generalizations. *Dissertation*, 2012.

[26] S. Engelke, A. Malinowski, Z. Kabluchko, and M. Schlather. Estimation of Hüsler-Reiss distributions and Brown-Resnick processes. *Journal of the royal statistical scociety: Series B*, 77(1):239–265, 2012.

[27] D. Gabda, R. Towe, J.Wadsworth, and J.Tawn. Discussion of statistical modelling of spatial extremes, by A.C. Davison, S.A. Padoan and m. Ribatet. *Statistical Science*, 27 (2):189–192, 2012.

[28] E.J. Gumbel. Distributions des valeurs extrêmes en plusierus dimensions. *Publications de l'Institute de Statistique de L'Université de Paris*, 9:171–173, 1960.

[29] S. Haug. Programmieren mit R. *Lecture Notes*, 2015.

[30] J. Hüsler and R.D. Reiss. Maxima of normal random vectors: between independence and complete dependence. *Statistics and Probability Letters*, 7:283–286, 1989.

[31] A.G. Journel and Ch.J. Huijbregts. *Mining Geostatistics*. Academic Press, 1978.

[32] Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max stable fields associated to negative definite functions. *The Annals of Probability*, 37(5):2042–2065, 2009.

[33] J.F.C Kingman. *Poisson Processes: In Oxford Studies in Probability, vol .3*. The Clarendon Press Oxford University Press, 1999.

[34] A. Klenke. *Probability Theory - A comprehensive course*. Springer, 2006.

[35] A. Lodhia, S. Sheffield, X. Sun, and S.S. Watson. Fractional Gaussian Fields: a survey. *arxiv.org*, 2014.

[36] M. Oesting. Simulationsverfahren für Brown-Resnick Prozesse. *Diplomarbeit*, 2009.

[37] M. Oesting, Z. Kabluchko, and M. Schlather. Simulation of Brown-Resnick processes. *Extremes*, 15:89–107, 2012.

[38] M. Oesting, M.Schlather, and C.Zhou. On the normailzed spectral representation of max-stable process on a compact set. *Available from http://archiv.org/abs/1310.1813*, 2013.

[39] T. Opitz. Extremal t processes: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis*, 122:409–413, 2013.

[40] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, 1987.

[41] J. Ryvkina. Fractional Brownian motion with variable Hurst parameter: Definition and properties. *Journal of Theoretical Probability*, 28(3), 2013.

[42] G. Samoradnitsky and M.S. Taqqu. *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Chapman and Hall/CRC, 1994.

[43] R.L. Smith. *Max-stable processes and spatial extremes*. Unpublished manusscript, University of North Carolina, 1990.

[44] D.L. Snyder and M.I. Miller. *Random Point Processes in Time and Space*. Springer, 1991.

[45] C. Steinkohl. Statistical modelling of extremes in space and time using max stable processes. *Dissertation*, 2013.

[46] C. Vardar. Results on the supremum of fractional Brownian motion. *Hacettepe Journal of Mathematics and Statistics*, 40(2):255–264, 2011.

[47] Y. Xiao. Recent development on fractal properties of Gaussian random fields. *Further Developments in Fractals and Related Fields, Springer*, pages 255–288, 2013.