

Spatial Interpolation and Prediction of Gaussian and Max-Stable Processes

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
an der Georg-August-Universität Göttingen

vorgelegt von

Marco Oesting

aus

Georgsmarienhütte

Göttingen 2012

Referent: Prof. Dr. Martin Schlather
Koreferent: Prof. Dr. Robert Schaback
Tag der mündlichen Prüfung: 03.05.2012

Preface

At this place I would like to thank a few of those people who supported me during my PhD studies. First of all, I owe my thanks to my supervisor Prof. Dr. Martin Schlather who was always available to answer my questions and give me support — even in busy times. In lots of fruitful discussions, he gave me new input and advice. I would also like to thank my co-supervisor Prof. Dr. Robert Schaback who introduced me to the field of approximation theory, answered my questions and drew my attention to numerous articles in numerical analysis related to this thesis. Both my supervisors provided me with lots of new ideas, but also gave me the freedom and support to develop and pursue my own ideas and interests.

I am very grateful to Dr. Michael Scheuerer. Besides my two supervisors, he showed me the connections between spatial statistics and numerical analysis and helped me to combine ideas from both fields. His preliminary work was the basis for Chapter 2. Furthermore, I would like to thank Jun.-Prof. Dr. Zakhar Kabluchko for the joint work, which led to Chapter 4, his annotations and help. I also thank Prof. Anthony C. Davison for his comments on simulations on a grid and Sebastian Engelke for helpful discussions on mixed moving maxima representations. Additionally, I am grateful to Thomas Rippl for drawing my attention to martingale arguments for computing conditional distributions in Chapters 5 and 6, and Kirstin Storkorb, Alexander Malinowski and my brother Henning for their careful proof-reading and helpful annotations.

I also gratefully acknowledge the financial support by the German Research Foundation DFG in form of a scholarship through the Research Training Group 1023 “Identification in Mathematical Models: Synergy of Stochastic and Numerical Methods”. Finally, I would like to thank my parents, colleagues and friends for their support.

Notation

$\Upsilon \subset \mathbb{R}^d$	domain of interest
Z	random field, defined on domain Υ
T	finite set of data locations; $T = \{t_1, \dots, t_n\}$
$Z(\mathbf{t})$	vector of values of the random field Z at data locations; $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$, page 6
\mathbb{P}	a probability measure
\mathbb{E}	expectation
Var	variance
Cov	covariance
i.i.d.	independent and identically distributed
a.s.	almost surely
\xrightarrow{p}	convergence in probability
\xrightarrow{d}	convergence in distribution (weak convergence)
$\stackrel{d}{=}$	equality in distribution
$\gamma(\cdot)$	a variogram, page 43
\mathbf{I}_n	$n \times n$ -identity matrix, page 20
$ A $	cardinality of a set A ; $ A = \#\{i : i \in A\}$
2^A	powerset of A
\mathbb{R}^Υ	set of all functions $f : \Upsilon \rightarrow \mathbb{R}$
$C(\Upsilon)$	set of all continuous functions $f : \Upsilon \rightarrow \mathbb{R}$
$\mathcal{C}(\Upsilon)$	σ -algebra on a space of functions $f : \Upsilon \rightarrow \mathbb{R}$; generated by cylinder sets, page 38
\mathcal{B}	Borel σ -algebra on \mathbb{R}
\mathcal{B}^d	Borel σ -algebra on \mathbb{R}^d
$\mathcal{B}(S)$	Borel σ -algebra on some space S

$L^p(\Upsilon)$	L^p -space w.r.t. domain Υ
$\mathcal{H}(K; \Upsilon)$	native space for the kernel $K : \Upsilon \times \Upsilon \rightarrow \mathbb{R}$, page 7
$H(K; \Upsilon)$	pre-Hilbert space spanned by the functions $K(\cdot - t)$, $t \in \Upsilon$, page 7
$W^{\tau,2}(\Upsilon)$	Sobolev space of order τ w.r.t. domain Υ , page 9
$\kappa_\nu, \tilde{\kappa}_\nu$	Whittle-Matérn covariance/correlation function with smoothness parameter ν , page 8
$\tilde{\kappa}_{\nu,c}$	Whittle-Matérn correlation function with smoothness parameter ν and scale parameter c , page 82
$s_{f,T,\nu}$	kernel interpolant w.r.t. data $\{(t, f(t)) : t \in T\}$ and kernel κ_ν , page 11
ϕ_ε	Gaussian kernel; scaled by $\varepsilon > 0$, page 13
\mathbf{C}_ε	kernel/covariance matrix w.r.t. kernel ϕ_ε and locations T , page 14
s_ε	kernel interpolant w.r.t. data $\{(t, Z(t)) : t \in T\}$ and kernel ϕ_ε , page 14
Φ	standard normal distribution function, page 82
φ	standard normal density, page 83

Contents

1	Introduction	1
2	Estimating the Whittle-Matérn Smoothness Parameter via the Native Space Norm	5
2.1	Spatial Interpolation for Second-Order Random Fields	5
2.2	Interpolation in Reproducing Kernel Hilbert Spaces	6
2.3	The Whittle-Matérn Kernel and its Native Space	8
2.4	On the Behaviour of the Native Space Norm	11
2.5	The Choice of the Tuning Parameter Function $\lambda(\nu)$	21
2.6	Comparison with MLE and Cross Validation Estimator	22
2.7	The Unknown Variance Case	24
2.8	An Alternative Approach Using Error Bounds	26
3	Max-Stable Processes Based on Flat Limits of Gaussian Random Fields	29
3.1	Max-Stable Processes	29
3.2	Max-Stable Processes Based on Universal Kriging with Two Data Locations	31
3.3	Max-Stable Processes Based on Simple Kriging with Two or Three Data Locations	35
4	Unconditional Simulation of Brown-Resnick Processes	43
4.1	Random Shifts	44
4.2	Mixed Moving Maxima Representation	46
4.3	Finite Approximations	51
4.4	Error Estimates	53
4.5	Simulation Study	55
5	Conditional Sampling of Mixed Moving Maxima Processes	59
5.1	Random Partition of Π and Measurability	60
5.2	Blurred Sets, Scenarios and Limit Considerations	62
5.3	Calculations in the Case of a Finite Number of Shape Functions	71
5.4	Comparison with Other Algorithms	82
5.5	Reduction of Computational Burden	85
5.6	Approximation in the Case of an Infinite Number of Shape Functions	89
5.7	Application to the Brown-Resnick Process	91
5.8	The Discretized Case	95
6	Conditional Sampling of Brown-Resnick and Extremal Gaussian Processes	97
6.1	Random Partition of Π	98
6.2	Blurred Sets	99
6.3	On the Distribution of Π_* within a Scenario	111

Bibliography

119

1 Introduction

This PhD thesis deals with different aspects of spatial interpolation and prediction of random fields. In general, we have the following situation: A random field $\{Z(t), t \in \Upsilon\}$, $\Upsilon \subset \mathbb{R}^d$, is measured at a finite set $T = \{t_1, \dots, t_n\}$ of locations. Based on these data $Z(t_1), \dots, Z(t_n)$, we aim to predict the random field at further locations taking into account spatial dependencies. This question of prediction is one of the main issues in spatial statistics and has various applications in sciences like mining, hydrology or meteorology. Methods of spatial statistics have also been adapted in a non-spatial framework, e.g. in animal breeding. In these applications, measurements are often very expensive. Therefore, we often have to cope with small and sparse sets T of data points. This makes an accurate prediction quite challenging.

In this context of prediction, basically two questions occur: First, one may ask for a pointwise predictor which is optimal w.r.t. some criterion, e.g. the conditional mean or the conditional median. In most applications, this question is answered by various kinds of kriging providing the best linear unbiased predictor which minimizes the predictive variance. Kriging (named after D. G. Krige) is a suitable tool to predict the variable of interest at some fixed location, e.g. the ore concentration in a deposit or the total genetic value of a single animal (cf. [Ober et al., 2011](#)). However, one should be aware of the fact that kriging yields a spatial interpolant being much smoother than a typical sample path of the random field. As kriging only accounts for the pointwise conditional variability of a random field, important quantities — like the probability that some critical amount of precipitation is exceeded within a whole region — cannot be determined by the means of kriging.

Thus, secondly, one may ask how to sample the random field conditionally on the data. By conditional sampling, we can get the whole conditional distribution of the random field and therefore entirely describe its stochastic properties.

In this thesis, we will deal with two types of random fields: (stationary) Gaussian and max-stable random fields.

The case of a stationary Gaussian random field is known best in spatial statistics as it covers the assumptions made in most applications. In this framework, various methods of kriging are used for prediction, providing the conditional mean of the random field. Conditional sampling is quite easy as the conditional random field can be described as a sum of the kriging result and a zero mean (non-stationary) Gaussian random field with both summands being stochastically independent. However, all these methods rely on the knowledge of the correct covariance structure of the random field. Therefore, tools are needed to identify the underlying covariance from the data. In this work, we will point out ways to incorporate results from kernel interpolation in numerical analysis into parameter estimation procedures used in spatial statistics. To this end, we make use of connections to kernel interpolation in numerical analysis which have recently been analysed by [Scheuerer \(2009\)](#).

Besides the most prominent class of Gaussian random fields, we also consider so-called max-stable random fields. Over the last decades, max-stable processes have entered numerous applications as models for extreme events like heavy storms or heat waves. In this case, the application of kriging methods is troublesome. In general, kriging methods require a second-order random field, i.e. the existence of second moments — a condition which is often not met in the max-stable setting. Bypassing this problem by transformations often leads to covariances which are numerically intractable. Naturally, conditional sampling is even more challenging. Promising approaches to tackle this problem have been rare for a long time. Recently, first results in this framework were obtained by [Wang and Stoev \(2011\)](#), [Dombry and Eyi-Minko \(2011\)](#), [Dombry et al. \(2011\)](#) and [Dombry and Ribatet \(2012\)](#). We present procedures for conditional simulation for three classes of max-stable processes which are used in spatial applications quite frequently: mixed moving maxima processes (e.g. [Smith, 1990](#)) and — independently of [Dombry et al. \(2011\)](#) and [Dombry and Ribatet \(2012\)](#) — extremal Gaussian processes ([Schlather, 2002](#)) and Brown-Resnick processes ([Brown and Resnick, 1977](#); [Kablichko et al., 2009](#)).

In more detail, this PhD thesis is organized as follows: Chapter 2 deals with an example of parameter estimation in case of a Gaussian random field. Here, we consider the very flexible class of Whittle-Matérn kernels and estimate the corresponding smoothness parameter. [Scheuerer \(2009\)](#) showed that this parameter is closely related to the smoothness of the sample paths of the random field expressed in terms of weak derivatives. The corresponding Sobolev spaces also occur as so-called native spaces of the Whittle-Matérn kernels in numerical analysis. In this framework, kernels and native spaces are well-studied as there are numerous applications, e.g. in machine learning and for solving partial differential equations numerically. We combine the stochastic properties of the random field with results on the behaviour of the interpolant from numerical analysis to create new estimators. We compare these estimators to classical ones like maximum likelihood and cross validation estimators. While deriving new estimators, we obtain results on the behaviour of the interpolant, i.e. the kriging result, as the smoothness parameter tends to infinity. As a tool for the analysis of these smooth limits we use results on flat limits of kernel interpolants.

Flat limits also occur in Chapter 3 which connects Gaussian random fields to max-stable random fields. Here, we construct max-stable processes based on flat limits of Gaussian processes. Furthermore, we study the stationarity of these processes. In some examples, the construction yields processes which belong to the class of Brown-Resnick processes. Brown-Resnick processes are the natural link between Gaussian and max-stable processes as they occur as limits of maxima of Gaussian processes.

Before tackling the problem of conditional simulation, we notice that even unconditional simulation of Brown-Resnick processes is quite involved. Although Brown-Resnick processes are stationary, finite approximations based on the definition indicate nonstationarity. Therefore, in Chapter 4 — which is also published in a slightly modified form as ([Oesting et al., 2012](#)) resulting from a diploma thesis ([Oesting, 2009](#)) — we present alternative representations of these processes. Based on these representations, we propose finite approximations and give error estimates. Finally, we compare different simulation techniques. Some techniques, in particular simulation by a mixed moving maxima representation, provide notable improvements.

The final Chapters 5 and 6 deal with conditional sampling for max-stable processes. In both chapters, we choose similar approaches making use of the underlying Poisson point process structure. In Chapter 5, which is also published as (Oesting and Schlather, 2012) in a slightly modified form, we consider the class of processes which allow for a mixed moving maxima representation. Besides general formulae and results in a broad setting, explicit and exact calculations are presented for processes on \mathbb{R} with a finite number of smooth shape functions. We compare our results to other algorithms for the Gaussian extreme value process (Smith, 1990) and the original Brown-Resnick process (Brown and Resnick, 1977) based on the mixed moving maxima representation by Engelke et al. (2011).

As the algorithm for mixed moving maxima cannot be applied exactly to Brown-Resnick processes, we present an exact procedure for conditional sampling for these processes – using a technically different approach than Dombry et al. (2011). The same techniques can be applied to the class of extremal Gaussian processes (Chapter 6; see also Dombry and Ribatet 2012).

In this thesis, we also have to deal with the fact that the literature on max-stable processes is quite heterogeneous w.r.t. the marginal distributions considered. Working with processes based on maxima of Gaussian processes, it is very natural to have Gumbel margins as the normal distribution is in the Gumbel max-domain of attraction. Therefore, we study processes with Gumbel margins in Chapters 3 and 4. Contrarily, Fréchet margins are assumed in most recent publications as these allow for convenient handling of exponent and spectral measures. For this reason, we work with Fréchet marginal distributions for conditional sampling of max-stable processes (Chapters 5 and 6). Thus, within this thesis, we deal with Brown-Resnick processes with Gumbel margins (Chapter 4) and Fréchet margins (Chapter 6). However, marginal transformation can be done very easily by the exponential and logarithmic function, respectively.

2 Estimating the Whittle-Matérn Smoothness Parameter via the Native Space Norm

In this chapter, we deal with spatial interpolation in the case of Gaussian random fields. As already mentioned in the introduction, in this case, procedures for optimal prediction and conditional sampling are well-studied provided that the underlying covariance structure is known. We aim to estimate the covariance from the broad class of Whittle-Matérn covariance functions making use of strong connections between methods from spatial statistics and numerical analysis, which were already treated in [Scheuerer \(2009\)](#). First, we give a short introduction to prediction for second-order random fields and conditional sampling for Gaussian random fields (Section 2.1) and present the basics of kernel interpolation in numerical analysis (Section 2.2). In Section 2.3, we point out that there is a strong connection in the case of Whittle-Matérn kernels. Employing this connection, we analyse the behaviour of the spatial interpolant as the smoothness parameter of the kernel used for interpolation tends to infinity (Section 2.4). In Sections 2.5–2.7, we develop new estimators based on these results and compare them with other estimators. Finally, we consider an estimator making use of error bounds known from numerical analysis (Section 2.8).

2.1 Spatial Interpolation for Second-Order Random Fields

Let $\{Z(t), t \in \Upsilon\}$, $\Upsilon \subset \mathbb{R}^d$, be a second-order random field, i.e. a random field with finite second moments, of the form

$$Z(t) = m(t) + \zeta(t), \quad t \in \Upsilon,$$

where $m : \Upsilon \rightarrow \mathbb{R}$ is the *mean function* and $\{\zeta(t), t \in \Upsilon\}$ is a zero mean second-order random field. By this decomposition, we have $\mathbb{E}(Z(t)) = m(t)$ and $\text{Cov}(Z(s), Z(t)) = \text{Cov}(\zeta(s), \zeta(t))$ for all $s, t \in \Upsilon$.

We assume that Z is known on some set $T = \{t_1, \dots, t_n\} \subset \Upsilon$. In this framework, a natural way to predict $Z(t_0)$, $t_0 \in \Upsilon$, based on $Z(t_1), \dots, Z(t_n)$ is to choose the *best linear unbiased predictor* (BLUP) which is a predictor $Z^*(t_0)$ of the form

$$Z^*(t_0) = \sum_{i=1}^n \alpha_i(t_0) Z(t_i), \quad t_0 \in \Upsilon, \quad \alpha_i(t_0) \in \mathbb{R},$$

such that $\mathbb{E}(Z^*(t_0)) = \sum_{i=1}^n \alpha_i(t_0) m(t_i) = m(t_0)$ (i.e. $Z^*(t_0)$ is unbiased) and

$$\text{Var}(Z^*(t_0) - Z(t_0)) \leq \text{Var} \left(\sum_{i=0}^n \beta_i(t_0) Z(t_i) - Z(t_0) \right)$$

for all $\beta_1(t_0), \dots, \beta_n(t_0) \in \mathbb{R}$ such that $\sum_{i=1}^n \beta_i(t_0)m(t_i) = m(t_0)$.

For second-order random fields, various kinds of best linear unbiased prediction are summarized by the term “kriging”. Based on the work of [Krige \(1951\)](#) and [Matheron \(1963\)](#), BLUPs based on different assumptions on the random field Z have been developed, e.g. simple kriging (known mean), ordinary kriging (unknown, but constant mean) or universal kriging (unknown mean function which is known to be in some given finite-dimensional linear space).

Here, we will focus on simple kriging. For details on other kinds of kriging, see [Chilès and Delfiner \(1999\)](#). In the framework of simple kriging, the mean function $m(t)$ is assumed to be known, w.l.o.g. $m(t) \equiv 0$. Then, the simple kriging predictor (BLUP) is given by $Z^*(t_0) = \sum_{i=0}^n \alpha_i(t_0)Z(t_i)$ where the vector of $\alpha_1(t_0), \dots, \alpha_n(t_0)$ is a solution of the linear system

$$\begin{pmatrix} K(t_1, t_1) & \cdots & K(t_1, t_n) \\ \vdots & \ddots & \vdots \\ K(t_n, t_1) & \cdots & K(t_n, t_n) \end{pmatrix} \cdot \begin{pmatrix} \alpha_1(t_0) \\ \vdots \\ \alpha_n(t_0) \end{pmatrix} = \begin{pmatrix} K(t_0, t_1) \\ \vdots \\ K(t_0, t_n) \end{pmatrix}, \quad (2.1)$$

where $K(s, t) = \text{Cov}(Z(s), Z(t))$ for all $s, t \in \Upsilon$.

If $K(\cdot, \cdot)$ is strictly positive definite, Equation (2.1) is uniquely solvable and we get

$$Z^*(t_0) = (K(t_0, t_1), \dots, K(t_0, t_n)) \cdot \begin{pmatrix} K(t_1, t_1) & \cdots & K(t_1, t_n) \\ \vdots & \ddots & \vdots \\ K(t_n, t_1) & \cdots & K(t_n, t_n) \end{pmatrix}^{-1} \cdot \begin{pmatrix} Z(t_1) \\ \vdots \\ Z(t_n) \end{pmatrix}. \quad (2.2)$$

Then, the kriging variance is given by

$$\text{Var}(Z^*(t_0) - Z(t_0)) = K(0, 0) - \sum_{i=1}^n \sum_{j=1}^n K(t_0, t_i) \left((K(t_k, t_l))_{k,l}^{-1} \right)_{ij} K(t_0, t_j). \quad (2.3)$$

Note that the simple kriging predictor relies on the knowledge of the true covariance function $K(\cdot, \cdot)$ of Z .

If Z is a Gaussian random field, the simple kriging predictor is not only the best linear unbiased estimator but even yields the conditional expectation $Z^*(t_0) = \mathbb{E}(Z(t_0) \mid Z(\mathbf{t}))$ where $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$. Furthermore, $Z(\cdot) \mid Z(\mathbf{t})$ is a Gaussian random field, by itself, and the random kriging error $Z^*(\cdot) - Z(\cdot)$ is independent of $Z^*(\cdot)$. Thus, we get

$$Z(\cdot) \mid Z(\mathbf{t}) \stackrel{d}{=} Z^*(\cdot) + \tilde{Z}(\cdot) - \tilde{Z}^*(\cdot),$$

where \tilde{Z} is a Gaussian random field with the same distribution but independent of Z ([Lantuéjoul, 2002](#)). Thus, the conditional distribution of Z can be simulated quite easily.

2.2 Interpolation in Reproducing Kernel Hilbert Spaces

In this section, we give a short introduction to kernel interpolation in numerical analysis which is a projection in a reproducing kernel Hilbert space. To this end, we need the notion of a positive definite kernel on some domain $\Upsilon \subset \mathbb{R}^d$.

Definition 2.1 (cf. [Wendland 2005](#), for example). A continuous and symmetric function $K : \Upsilon \times \Upsilon \rightarrow \mathbb{R}$ is called *positive semi-definite*, if for all finite sets of pairwise distinct locations $T = \{t_1, \dots, t_n\} \subset \Upsilon$, $n \in \mathbb{N}$, and coefficients $a_1, \dots, a_n \in \mathbb{R}$, we have

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j K(t_i, t_j) \geq 0. \quad (2.4)$$

The kernel K is called *strictly positive definite* if equality in (2.4) holds if and only if $a_1 = \dots = a_n = 0$.

Note that any (continuous) covariance function of a random field is positive semi-definite. Let $K : \Upsilon \times \Upsilon \rightarrow \mathbb{R}$ be a strictly positive-definite kernel. Then, following [Wendland \(2005\)](#), we can define the linear space

$$H(K; \Upsilon) = \text{span}\{K(\cdot, t) : t \in \Upsilon\}.$$

Thus, $H(K; \Upsilon)$ is a pre-Hilbert space w.r.t. the inner product $(\cdot, \cdot)_K$ defined by

$$\left(\sum_{i=1}^m a_i K(\cdot, x_i), \sum_{j=1}^n b_j K(\cdot, y_j) \right)_K = \sum_{i=1}^m \sum_{j=1}^n a_i b_j K(x_i, y_j).$$

Furthermore, for all $f \in H(K; \Upsilon)$ and $t \in \Upsilon$, the reproducing property

$$f(t) = (f, K(\cdot, t))_K \quad (2.5)$$

holds. Completing $H(K; \Upsilon)$ w.r.t. the norm induced by the inner product $(\cdot, \cdot)_K$, we get a Hilbert space $\mathcal{H}(K; \Upsilon)$. By continuity arguments, the reproducing property can be extended and the elements $f \in \mathcal{H}(K; \Upsilon)$ can be interpreted as functions defined on Υ via (2.5). The space $\mathcal{H}(K; \Upsilon)$ is called *reproducing kernel Hilbert space* (RKHS) associated to the kernel $K(\cdot, \cdot)$.

Now, let $f \in \mathcal{H}(K; \Upsilon)$ be given on the set $T = \{t_1, \dots, t_n\}$. A suitable approach to reconstruct f based on $f(t_1), \dots, f(t_n)$ is to find the best approximation in $H(K; T)$ w.r.t. the norm $\|\cdot\|_K$ defined by $\|g\|_K^2 = (g, g)_K$ for all $g \in \mathcal{H}(K; \Upsilon)$. Thus, we have to find $s_{f,T}(\cdot) = \sum_{i=1}^n a_i K(\cdot, t_i)$ such that

$$\|s_{f,T}(\cdot) - f(\cdot)\|_K \leq \left\| \sum_{i=1}^n b_i K(\cdot, t_i) - f(\cdot) \right\|_K \quad \text{for all } b_1, \dots, b_n \in \mathbb{R}.$$

By [Wendland \(2005\)](#), Thm. 13.1, the solution $s_{f,T}(\cdot)$ turns out to be the unique interpolant to f in $H(K; T)$ w.r.t. the data locations $(t_i, f(t_i))$, $i = 1, \dots, n$. Thus, $s_{f,T}(t_0) = \sum_{i=1}^n a_i K(t_0, t_i)$ with

$$\begin{pmatrix} K(t_1, t_1) & \cdots & K(t_1, t_n) \\ \vdots & \ddots & \vdots \\ K(t_n, t_1) & \cdots & K(t_n, t_n) \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_n) \end{pmatrix}.$$

2. Estimating the Whittle-Matérn Smoothness Parameter via the Native Space Norm

As K is strictly positive definite, this linear system has a unique solution, and we can rewrite

$$s_{f,T}(t_0) = (f(t_1), \dots, f(t_n)) \cdot \begin{pmatrix} K(t_1, t_1) & \cdots & K(t_1, t_n) \\ \vdots & \ddots & \vdots \\ K(t_n, t_1) & \cdots & K(t_n, t_n) \end{pmatrix}^{-1} \cdot \begin{pmatrix} K(t_0, t_1) \\ \vdots \\ K(t_0, t_n) \end{pmatrix}$$

which is exactly the same formula as (2.2). The pointwise interpolation error can be bounded uniformly by

$$|f(t) - s_{f,T}(t)| \leq P_{K,T}(t) \cdot \|f\|_K \quad \text{for all } f \in \mathcal{H}(K; \Upsilon) \quad (2.6)$$

where $P_{K,T}$, defined by

$$P_{K,T}^2(t) = K(0, 0) - \sum_{i=1}^n \sum_{j=1}^n K(t, t_i) \left((K(t_k, t_l))_{k,l}^{-1} \right)_{ij} K(t, t_j), \quad (2.7)$$

is called *power function*. The inequality (2.6) is sharp and equality holds for $f(\cdot) = K(\cdot, t) - s_{K(\cdot, t), T}(\cdot)$, for example (cf. Wendland, 2005). Note that Equations (2.3) and (2.7) look exactly the same, that is, the kriging variance equals the squared power function. In the following, we will make use of these connections between interpolation in second-order random fields and reproducing kernel Hilbert spaces which yield the same formulae although basic assumptions are different.

2.3 The Whittle-Matérn Kernel and its Native Space

Let Z be a stationary zero mean random field on an open and bounded domain $\Upsilon \subset \mathbb{R}^d$ with a Whittle-Matérn covariance function $\tilde{\kappa}_{\nu_0}$, $\nu_0 > 0$. The Whittle-Matérn covariance function is given by

$$\kappa_{\nu}(h) := \frac{\|h\|^{\nu}}{2^{\nu+d/2-1}\Gamma(\nu+d/2)} \mathcal{K}_{\nu}(\|h\|), \quad h \in \mathbb{R}^d, \quad \nu > 0,$$

where $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind. The analogous correlation function is defined by

$$\tilde{\kappa}_{\nu}(h) = \frac{\|h\|^{\nu}}{2^{\nu-1}\Gamma(\nu)} \mathcal{K}_{\nu}(\|h\|) = \frac{2^{d/2}\Gamma(\nu+d/2)}{\Gamma(\nu)} \kappa_{\nu}(h).$$

The class of Whittle-Matérn covariance functions is very flexible as it allows for modelling the regularity of the sample paths of Z by the smoothness parameter ν . It can be shown that the sample paths of the Gaussian field Z are k times differentiable (after some modification if necessary) if and only if $\nu_0 > k$ (cf. Gelfand et al., 2010). Because of this flexibility, Stein (1999) recommends to use this class for modelling spatial data in various applications. Furthermore, the Whittle-Matérn family contains other popular covariance functions like the exponential ($\nu = 1/2$) and the Gaussian covariance kernel which occurs as a scaled limit $\tilde{\kappa}_{\nu}(2\sqrt{\nu}r) \rightarrow \exp(-r^2)$ for all $r \geq 0$ as $\nu \rightarrow \infty$ (cf. Gelfand

et al., 2010). For further details on the Whittle-Matérn family see the review article by Guttorp and Gneiting (2006).

Now, we study the native space (reproducing kernel Hilbert space) belonging to $\tilde{\kappa}_\nu$. As the Fourier transform $\hat{\kappa}_\nu$ of κ_ν is given by

$$\hat{\kappa}_\nu(\omega) = \frac{1}{\sqrt{2\pi}^d} \int_{\mathbb{R}^d} \kappa_\nu(x) e^{-ix^T \omega} dx = (1 + \|\omega\|_2^2)^{-(\nu+d/2)}, \quad \omega \in \mathbb{R}^d, \quad (2.8)$$

the native space $\mathcal{H}(\kappa_\nu; \mathbb{R}^d) = \mathcal{H}(\tilde{\kappa}_\nu; \mathbb{R}^d)$ can be written as

$$\mathcal{H}(\kappa_\nu; \mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) \cap C(\mathbb{R}^d) : \int_{\mathbb{R}^d} |\hat{f}(\omega)|^2 (1 + \|\omega\|_2^2)^{\nu+d/2} d\omega < \infty \right\} \quad (2.9)$$

by Wendland (2005), Thm. 10.12.

On the other hand, we have the so-called *Sobolev space*

$$W^{\tau,2}(\Upsilon) = \left\{ f \in L^2(\Upsilon) : \sum_{|\alpha| \leq \lfloor \tau \rfloor} \|D^\alpha f\|_{L^2(\Upsilon)}^2 + \sum_{|\alpha| = \lfloor \tau \rfloor} \int_{\Upsilon} \int_{\Upsilon} \frac{|D^\alpha f(s) - D^\alpha f(t)|^2}{\|s - t\|^{d+2(\tau - \lfloor \tau \rfloor)}} ds dt < \infty \right\}, \quad \tau > 0,$$

where $D^\alpha f$ denotes the weak derivative of f of order $\alpha \in \mathbb{N}_0^d$ and $\lfloor \tau \rfloor$ denotes the largest integer less or equal to τ . If τ is an integer, the second term is dropped.

The Sobolev space is a Hilbert space with respect to the inner product $(\cdot, \cdot)_{W^{\tau,2}(\Upsilon)}$, which is given by

$$(f, g)_{W^{\tau,2}(\Upsilon)} = \sum_{|\alpha| \leq \lfloor \tau \rfloor} (D^\alpha f, D^\alpha g)_{L^2(\Upsilon)} + \sum_{|\alpha| = \lfloor \tau \rfloor} \int_{\Upsilon} \int_{\Upsilon} \frac{(D^\alpha f(x) - D^\alpha f(y))(D^\alpha g(x) - D^\alpha g(y))}{\|x - y\|^{d+2(\tau - \lfloor \tau \rfloor)}} dx dy$$

for $f, g \in W^{\tau,2}(\Upsilon)$ (cf. Kufner et al., 1977, Section 6.8).

If we replace Υ by \mathbb{R}^d , the Sobolev space $W^{\tau,2}(\mathbb{R}^d)$ and the corresponding inner product can be written via Fourier transforms yielding the space

$$H^\tau(\mathbb{R}^d) = \left\{ f \in L^2(\mathbb{R}^d) : \int_{\mathbb{R}^d} |\hat{f}(\omega)|^2 (1 + \|\omega\|_2^2)^\tau d\omega < \infty \right\}$$

and the inner product

$$(f, g)_{H^\tau(\mathbb{R}^d)} = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{g}(\omega)} (1 + \|\omega\|_2^2)^\tau d\omega, \quad f, g \in H^\tau(\mathbb{R}^d).$$

By Adams (1975), Theorem 7.63, the Hilbert spaces $W^{\tau,2}(\mathbb{R}^d)$ and $H^\tau(\mathbb{R}^d)$ coincide algebraically and the norms $\|\cdot\|_{H^\tau(\mathbb{R}^d)}$ and $\|\cdot\|_{W^{\tau,2}(\mathbb{R}^d)}$, induced by the inner products $(\cdot, \cdot)_{H^\tau(\mathbb{R}^d)}$ and $(\cdot, \cdot)_{W^{\tau,2}(\mathbb{R}^d)}$, respectively, are equivalent. Thus, we have

$$\mathcal{H}(\kappa_\nu; \mathbb{R}^d) = W^{\nu+d/2,2}(\mathbb{R}^d), \quad \nu > 0, \quad (2.10)$$

by Equation (2.9) (see also Wendland, 2005, Cor. 10.13).

We also aim to link native spaces of Whittle-Matérn kernels with Sobolev spaces for an appropriate class of domains $\Upsilon \subsetneq \mathbb{R}^d$. Following Grisvard (1985), we define domains with Lipschitz boundaries.

Definition 2.2 (cf. Grisvard, 1985, Def. 1.2.1.1). Let $\Upsilon \subset \mathbb{R}^d$ be an open set with boundary Γ . We call Υ *Lipschitz domain* or *domain with Lipschitz boundary* if for every $t \in \Upsilon$ there exists a neighbourhood $U_t \subset \mathbb{R}^d$ and a system $\{y_1, \dots, y_d\} = \{y_1(t), \dots, y_d(t)\}$ of orthogonal coordinates such that

1. $U_t = \{(y_1, \dots, y_d) : -a_j < y_j < a_j, j = 1, \dots, d\}$ for some $a_j = a_j(t) > 0, j = 1, \dots, d$
2. there is a Lipschitz function

$$\psi : \tilde{U}_t = \{(y_1, \dots, y_{d-1}) : a_j < y_j < a_j, j = 1, \dots, d-1\} \rightarrow \left[-\frac{a_d}{2}, \frac{a_d}{2}\right]$$

such that $\Upsilon \cap U_t = \{(\tilde{y}, y_d) \in U_t : y_d < \psi(\tilde{y})\}, \Gamma \cap U_t = \{(\tilde{y}, y_d) \in U_t : y_d = \psi(\tilde{y})\}$.

Note that the assumption that an open and bounded subset $\Upsilon \subset \mathbb{R}^d$ has Lipschitz boundary is relatively mild. For example, every bounded, open and convex subset of \mathbb{R}^d has a Lipschitz boundary (Grisvard, 1985, Cor. 1.2.2.3). Thus, the following statements hold if we consider the convex hull of our domain of interest.

Lemma 2.3 (cf. Grisvard, 1985, Thm. 1.4.3.1). *Let Υ be an open and bounded domain with Lipschitz boundary. Then, any function $f \in W^{\tau,2}(\Upsilon)$ can be extended to some function $\tilde{f} \in W^{\tau,2}(\mathbb{R}^d)$ for any $\tau > 0$.*

The next theorem and its proof are similar to Corollary 10.48 in Wendland (2005).

Theorem 2.4. *Let Υ be an open and bounded set with Lipschitz boundary. Then, we have*

$$\mathcal{H}(\kappa_\nu; \Upsilon) = W^{\nu+d/2,2}(\Upsilon)$$

for all $\nu > 0$.

Proof. By Equation (2.10), we have that these spaces coincide if we replace Υ by \mathbb{R}^d . Thus, we get

$$\mathcal{H}(\kappa_\nu; \Upsilon) \subset \{f|_\Upsilon : f \in \mathcal{H}(\kappa_\nu; \mathbb{R}^d)\} = \{f|_\Upsilon : f \in W^{\nu+d/2,2}(\mathbb{R}^d)\} \subset W^{\nu+d/2,2}(\Upsilon),$$

where we used that any $f \in \mathcal{H}(\kappa_\nu; \Upsilon)$ can be extended to a function $\tilde{f} \in \mathcal{H}(\kappa_\nu; \mathbb{R}^d)$ by Thm. 10.46 in Wendland (2005).

Now, let $f \in W^{\nu+d/2,2}(\Upsilon)$. Then, by Lemma 2.3, f can be extended to a function $\tilde{f} \in W^{\nu+d/2,2}(\mathbb{R}^d) = \mathcal{H}(\kappa_\nu; \mathbb{R}^d)$. This implies, as stated in Wendland (2005), Thm. 10.47, that the restriction $f = \tilde{f}|_\Upsilon$ is in the native space $\mathcal{H}(\kappa_\nu; \Upsilon)$. \square

In case of \mathbb{R}^d , the connection between native and Sobolev spaces is even stronger.

Theorem 2.5 (cf. Wendland, 2005, Thm. 10.12). *For each $f \in \mathcal{H}(\kappa_\nu; \mathbb{R}^d)$, the native space norm $\|f\|_{\kappa_\nu}$ and the Sobolev norm $\|f\|_{H^{\nu+d/2}(\mathbb{R}^d)}$ are the same, which means that both Hilbert spaces are identical.*

Proof. It suffices to show the proposition for $f \in \text{span}\{\kappa_\nu(\cdot - t), t \in \mathbb{R}^d\}$. Let $f(\cdot) = \sum_{i=1}^n a_i \kappa_\nu(\cdot - t_i)$ for some $a_i \in \mathbb{R}$, $t_i \in \mathbb{R}^d$, $n \in \mathbb{N}$. Then, with $\tau = \nu + d/2$, we have

$$\begin{aligned}
 \|f\|_{H^\tau(\mathbb{R}^d)}^2 &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left(\sum_{j=1}^n \sum_{k=1}^n a_j a_k \widehat{\kappa_\nu(\cdot - t_j)}(\omega) \overline{\widehat{\kappa_\nu(\cdot - t_k)}(\omega)} \right) (1 + \|\omega\|^2)^\tau \, d\omega \\
 &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \left(\sum_{j=1}^n \sum_{k=1}^n a_j a_k e^{-i\omega^T t_j} \widehat{\kappa_\nu}(\omega) e^{i\omega^T t_k} \widehat{\kappa_\nu}(\omega) \right) (1 + \|\omega\|^2)^\tau \, d\omega \\
 &\stackrel{(2.8)}{=} \sum_{j=1}^n \sum_{k=1}^n a_j a_k \left((2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i\omega^T (t_k - t_j)} \widehat{\kappa_\nu}(\omega) \, d\omega \right) \\
 &= \sum_{j=1}^n \sum_{k=1}^n a_j a_k \kappa_\nu(t_j - t_k) = \|f\|_{\kappa_\nu}^2
 \end{aligned}$$

□

In [Scheuerer \(2010, Corollary 1 and Proposition 1\)](#), for an open and bounded domain $\Upsilon \subset \mathbb{R}^d$ and a Gaussian random field Z with correlation function $\tilde{\kappa}_{\nu_0}$ the a.s. equivalence

$$\nu_0 > k \iff Z_\omega(\cdot) \in W^{k,2}(\Upsilon)$$

is shown for $k \in \mathbb{N}_0$ (although the kernel $\tilde{\kappa}_{\nu_0}$ generates the smaller RKHS $W^{\nu_0+d/2,2}(\Upsilon)$), i.e.

$$\begin{aligned}
 \nu_0 > k &\iff Z_\omega(\cdot) \in W^{k,2}(\Upsilon) \quad a.s. \\
 \nu_0 \leq k &\iff Z_\omega(\cdot) \notin W^{k,2}(\Upsilon) \quad a.s.
 \end{aligned}$$

For a fractional order τ of the Sobolev space the condition $\nu_0 > \tau$ is at least sufficient (see [Scheuerer \(2010\)](#), Remark 1).

2.4 On the Behaviour of the Native Space Norm

In the following, we will always assume Υ to be an open and bounded set with Lipschitz boundary. So, we can assess the smoothness parameter ν_0 by

$$\hat{\nu} := \sup \left\{ \tau > \frac{d}{2} : Z_\omega(\cdot) \in \mathcal{H}(\kappa_{\tau-\frac{d}{2}}; \Upsilon) = W^{\tau,2}(\Upsilon) \right\} \quad (2.11)$$

if it is greater than $d/2$.

In order to determine this value, let $T = \{t_1, \dots, t_n\} \subset \Upsilon$. Furthermore, we denote by $s_{Z,T,\nu}$ the unique interpolant to $Z_\omega(\cdot)$ on T in $H(\kappa_\nu; T)$

Theorem 2.6 ([Schaback and Wendland 2002](#), Thm. 5.1). *A continuous function $f : \Upsilon \rightarrow \mathbb{R}$ belongs to the native space $\mathcal{H}(\kappa_\tau; \Upsilon)$ if and only if there exists a constant c_f such that $\|s_{f,T,\tau}\|_{\kappa_\tau} \leq c_f$ for all finite subsets $T \subset \Upsilon$.*

2. Estimating the Whittle-Matérn Smoothness Parameter via the Native Space Norm

Actually, $\|s_{f,T,\tau}\|_{\kappa_\tau} = \|s_{f,T,\tau}\|_{H^{\tau+d/2}(\mathbb{R}^d)}$ turns out to be very large for $\tau + \frac{d}{2} \gg \nu_0$ and a sufficiently dense $T \subset \Upsilon$. However, $\widehat{\nu}$ is quite difficult to determine since, in general, data on a large set T are needed.

Therefore, we try to assess an appropriate smoothness parameter for Z in another way. For kriging and interpolation purposes, it is not really necessary to estimate the “real” parameter ν , but to assure a small interpolation error. The latter one is given by the power function, which equals the kriging variance (see Equations (2.3) and (2.7)),

$$P_{\tilde{\kappa}_\nu, T}^2(t) = \tilde{\kappa}_\nu(0) - \tilde{\mathbf{k}}_\nu(t)^T \tilde{\mathbf{K}}_\nu^{-1} \tilde{\mathbf{k}}_\nu(t) \quad (2.12)$$

where $\tilde{\mathbf{k}}_\nu(t) = (\tilde{\kappa}_\nu(t-t_i))_{i=1,\dots,n}^T$ and $\tilde{\mathbf{K}}_\nu = (\tilde{\kappa}_\nu(t_i-t_j))_{i,j=1,\dots,n}$ are the correlation matrix and vector for the set of locations $T = \{t_1, \dots, t_n\}$. Note, that $\tilde{\mathbf{K}}_\nu$ is invertible because of $\tilde{\kappa}_\nu$ being strictly positive definite.

Proposition 2.7. *Let $\{Z_\nu(t), t \in \Upsilon\}$ be a zero mean Gaussian random field with covariance function $\tilde{\kappa}_\nu$ and $T = \{t_1, \dots, t_n\} \subset \Upsilon$. Then, we have $\lim_{\nu \rightarrow \infty} P_{\tilde{\kappa}_\nu, T}^2(t) = \lim_{\nu \rightarrow \infty} \text{Var}(Z_\nu(t) - s_{Z_\nu, T, \nu}(t)) = 0$ for all $t \in \Upsilon$*

Proof. Since $s_{Z_\nu, T, \nu}(t)$ is the best linear unbiased predictor of $Z_\nu(t)$ and $\tilde{\kappa}_\nu(h) \xrightarrow{\nu \rightarrow \infty} 1$ for all $h \in \mathbb{R}^d$, we have

$$P_{\tilde{\kappa}_\nu, T}^2(t) = \text{Var}(Z_\nu(t) - s_{Z_\nu, T, \nu}(t)) \leq \text{Var}(Z_\nu(t) - Z_\nu(t_1)) = 2 - 2\tilde{\kappa}_\nu(t-t_1) \xrightarrow{\nu \rightarrow \infty} 0$$

for all $t \in \Upsilon$. □

One approach to estimate the true parameter ν_0 is to minimize some (continuous) error function \mathcal{E} , depending on the kernel κ_ν as a function of ν , like the power function or a modification of it. We combine this with (2.11) and minimize a function of the type

$$G_{T,\lambda}(\tau) = \mathcal{E}(\kappa_\tau) + \lambda \cdot \|s_{Z,T,\tau-d/2}\|_{\kappa_{\tau-d/2}}^2, \quad \tau > d/2.$$

Because of Theorem 2.6 and the following lemma, the penalty term $\|s_{Z,T,\tau-d/2}\|_{\kappa_{\tau-d/2}}^2$ is expected to be large for $Z_\omega(\cdot) \notin \mathcal{H}(\kappa_{\tau-d/2}; \Upsilon)$ — i.e. $\tau \geq \nu_0$ — and sufficiently large $T \subset \Upsilon$.

Lemma 2.8. *Let $\{T_m\}_{m \in \mathbb{N}}$ be an increasing sequence of subsets of Υ , i.e. $T_m \subset T_{m+1}$ for all $m \in \mathbb{N}$. Then, the sequence $\{\|s_{Z, T_m, \nu}\|_{\kappa_\nu}\}_{m \in \mathbb{N}}$ is monotonically increasing.*

Proof. We have $s_{Z, T_m, \nu} = s_{Z, T_{m+1}, \nu}$ on T_m . By Wendland (2005, Thm. 13.2), it holds

$$\|s_{Z, T_m, \nu}\|_{\kappa_\nu} = \min_{\substack{g \in \mathcal{H}(\kappa_\nu; \Upsilon) \\ g = s_{Z, T_m, \nu} \text{ on } T_m}} \|g\|_{\kappa_\nu}$$

and therefore we get $\|s_{Z, T_m, \nu}\|_{\kappa_\nu} \leq \|s_{Z, T_{m+1}, \nu}\|_{\kappa_\nu}$. □

The behaviour of $\|s_{Z, T, \nu}\|_{\kappa_\nu}^2$ as a function of ν is described by the following statement.

Proposition 2.9. *Let $T = \{t_1, \dots, t_n\} \subset \Upsilon$. The function $\|s_{Z, T, \nu}\|_{\kappa_\nu}^2$ is monotonically increasing in ν on $(0, \infty)$.*

Proof. First, we use the fact that each interpolating function

$$s_{Z,T,\nu}(\cdot) = \sum_{i=1}^n \lambda_i \kappa_\nu(\cdot - t_i)$$

can be extended naturally to \mathbb{R}^d and that

$$\|s_{Z,T,\nu}\|_{\kappa_\nu}^2 = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \kappa_\nu(t_i - t_j)$$

does not depend on Υ , as long as $T \subset \Upsilon$. Therefore, w.l.o.g. we may replace Υ by \mathbb{R}^d . Let $0 < \nu_1 < \nu_2$. By

$$\mathcal{H}(\kappa_{\nu_1}; \mathbb{R}^d) = W^{\nu_1+d/2,2}(\mathbb{R}^d) \supset W^{\nu_2+d/2,2}(\mathbb{R}^d) = \mathcal{H}(\kappa_{\nu_2}; \mathbb{R}^d)$$

the norm $\|s_{Z,T,\nu_2}\|_{\kappa_{\nu_1}}$ is well-defined and finite. By [Wendland \(2005, Thm. 13.2\)](#), we get

$$\|s_{Z,T,\nu_1}\|_{\kappa_{\nu_1}} = \min_{\substack{g \in \mathcal{H}(\kappa_{\nu_1}; \mathbb{R}^d) \\ g = s_{Z,T,\nu_2} \text{ on } T}} \|g\|_{\kappa_{\nu_1}} \leq \|s_{Z,T,\nu_2}\|_{\kappa_{\nu_1}}. \quad (2.13)$$

Furthermore, by [Theorem 2.5](#), we have

$$\begin{aligned} \|f\|_{\kappa_{\nu_1}}^2 &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} |\hat{f}(\omega)|^2 (1 + \|\omega\|_2^2)^{\nu_1+d/2} d\omega \\ &\leq (2\pi)^{-d/2} \int_{\mathbb{R}^d} |\hat{f}(\omega)|^2 (1 + \|\omega\|_2^2)^{\nu_2+d/2} d\omega = \|f\|_{\kappa_{\nu_2}}^2 \end{aligned}$$

for all $f \in \mathcal{H}(\kappa_{\nu_2}; \mathbb{R}^d)$. Applying this to $f = s_{Z,T,\nu_2}$ and using [Equation 2.13](#), we get the proposition. \square

In order to draw further conclusions on the behaviour of $\|s_{Z,T,\nu}\|_{\kappa_\nu}$ for large ν , we study the geometry of the set $T = \{t_1, \dots, t_n\} \subset \Upsilon \subset \mathbb{R}^d$ of data locations and the behaviour of *flat limits*. To this end, we consider the kernel interpolant

$$s_\varepsilon(\cdot) = \sum_{i=1}^n a_{i,\varepsilon} C(\varepsilon(\cdot - t_j)), \quad \varepsilon > 0,$$

to the data points $(t_1, z_1), \dots, (t_n, z_n) \in \mathbb{R}^d \times \mathbb{R}$ where C is a positive definite kernel, i.e. a covariance function. The flat limit s_0 is defined as the limit

$$s_0(t) = \lim_{\varepsilon \searrow 0} s_\varepsilon(t), \quad t \in \mathbb{R}^d.$$

Many results on the existence and behaviour of this limit can be found in the literature. For example, [Driscoll and Fornberg \(2002\)](#) gave a first condition on C which guarantees that the flat limit in \mathbb{R} exists and coincides with the Lagrange interpolating polynomial. If we consider the scaled Gaussian kernel

$$\phi_\varepsilon(h) = \exp\left(-\frac{(\varepsilon\|h\|)^2}{2}\right), \quad \varepsilon > 0, \quad h \in \mathbb{R}^d,$$

s_ε always converges to an interpolating polynomial (cf. [Schaback, 2005](#)).

In [Schaback \(2008\)](#), the existence of the flat limit was shown under some assumptions on C and a condition on the geometry of $T = \{t_1, \dots, t_n\}$. Therefore, following [Schaback \(2008\)](#), we define

$$k_0 = k_0(T) := \max\{k : p \in P_k^d, p(T) = \{0\} \Rightarrow p = 0\} \quad (2.14)$$

$$k_1 = k_1(T) := \min \left\{ k : n \leq \binom{k+d}{d} \right\} \quad (2.15)$$

$$k_2 = k_2(T) := \min\{k : \text{rank}((t_j^\alpha)_{1 \leq j \leq n, \alpha \in \mathbb{Z}_0^d, |\alpha| \leq k}) = n\} \quad (2.16)$$

where P_k^d denotes the linear space of multivariate polynomials on \mathbb{R}^d with degree at most k . These quantities can be interpreted as follows:

- k_0 is the largest degree which guarantees uniqueness of polynomial interpolation
- k_1 is the expected degree of an interpolating polynomial on a set of magnitude $|T| = n$ neglecting the geometry of T
- k_2 is the minimal degree which guarantees existence of an interpolating polynomial

We have $k_0 \leq k_1 \leq k_2$. If $d = 1$, we have $k_0 = k_1 = k_2 = n - 1$. For the points of T being on a line in \mathbb{R}^d , $d > 1$, we have $0 = k_0$ and $k_2 = n - 1$ (cf. [Schaback, 2008](#)).

By Theorem 2 in [Schaback \(2008\)](#), the flat limit s_0 exists for an analytic radial basis function C with positive Fourier transform on a set of positive measure if $k_2 \leq k_0 + 2$.

Here, we will use flat limits for the Gaussian kernel ϕ_ε to get results on the “smooth limit” behaviour of $s_{Z,T,\nu}$ as $\nu \rightarrow \infty$. Let $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$, $\mathbf{C}_\varepsilon = (\phi_\varepsilon(t_i - t_j))_{1 \leq i, j \leq n}$ and $v_\varepsilon(t) = (\phi_\varepsilon(t - t_j))_{j=1, \dots, n}$ for $\varepsilon > 0$, $t \in \mathbb{R}^d$. As the Gaussian kernel is strictly positive definite, the matrix \mathbf{C}_ε is non-singular and the interpolant $s_\varepsilon(\cdot)$ on T with respect to the kernel ϕ_ε can be written as

$$s_\varepsilon(t) = Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1} v_\varepsilon(t).$$

By Theorem 2 in [Schaback \(2005\)](#), we have that s_ε converges pointwise to a polynomial p interpolating the data points $(t_i, Z(t_i))$, $i = 1, \dots, n$. In the following, we will analyse the behaviour of s_ε for random data. We consider multivariate Gaussian data $Z(\mathbf{t}) \in \mathbb{R}^n$ with a non-degenerated covariance structure, which means $\text{supp}(d_{Z(\mathbf{t})}) = \mathbb{R}^n$ where $d_{Z(\mathbf{t})}$ is the probability density of $Z(\mathbf{t})$. Then, $p \in P_{k_1-1}^d$ with

$$p(t) = \lim_{\varepsilon \rightarrow 0} (Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1} v_\varepsilon(t)), \quad t \in \mathbb{R}^d,$$

implies that $Z(\mathbf{t}) \in \{(q(t_1), \dots, q(t_n)) : q \in P_{k_1-1}^d\}$. On the other hand, we have

$$\begin{aligned} \dim \left\{ (q(t_1), \dots, q(t_n)) : q \in P_{k_1-1}^d \right\} &\leq \dim(P_{k_1-1}^d) = \binom{k_1 - 1 + d}{d} \\ &< n = \dim(\text{supp}(d_{Z(\mathbf{t})})). \end{aligned}$$

This implies $\mathbb{P}(Z(\mathbf{t}) \in \{(q(t_1), \dots, q(t_n)) : q \in P_{k_1-1}^d\}) = 0$ and therefore

$$p \notin P_{k_1-1}^d \quad a.s. \quad (2.17)$$

Lemma 2.10. *Let $Z(\mathbf{t})$ be a random vector as above and $k_0 = k_0(T)$, $k_1 = k_1(T)$ and $k_2 = k_2(T)$ be defined as in Equations (2.14)–(2.16). Then, for any two norms $\|\cdot\|_V$ on \mathbb{R}^n and $\|\cdot\|_M$ on $\mathbb{R}^{n \times n}$, it holds*

1. $(\mathbf{C}_\varepsilon^{-1})_{ij} \in \mathcal{O}(\varepsilon^{-2k_2})$ for all $i, j \in \{1, \dots, n\}$.
2. $\|Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1}\|_V \notin o(\varepsilon^{-k_1})$ a.s.
3. With probability one there exist $0 < C_1(Z(\mathbf{t})) < \infty$ and $0 < C_2(Z(\mathbf{t})) < \infty$ such that

$$\lim_{\varepsilon \searrow 0} \frac{\|Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1}\|_V}{\|\mathbf{C}_\varepsilon^{-1}\|_M} = C_1(Z(\mathbf{t})) \quad \text{and} \quad \lim_{\varepsilon \searrow 0} \frac{Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1} Z(\mathbf{t})}{\|Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1}\|_V} = C_2(Z(\mathbf{t}))$$

for all $\varepsilon > 0$.

Proof. 1. Follows directly from Theorem 1 in [Schaback \(2008\)](#).

2. Assume that $\|Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1}\| \in o(\varepsilon^{-k_1})$. As each component of $Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1}$ can be written as the quotient of two power series with respect to ε^2 by the explicit series expansion

$$\phi_\varepsilon(h) = \sum_{k=0}^{\infty} (-1)^k \frac{\varepsilon^{2k}}{k!} \frac{\|h\|^{2k}}{2^k}, \quad h \in \mathbb{R}^d,$$

it has a power series expansion itself which implies $Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1} \in \mathcal{O}(\varepsilon^{-(k_1-1)})$. Using this series expansion $(Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1})_j = \sum_{n=-(k_1-1)}^{\infty} c_{nj} \varepsilon^n$ we get

$$s_\varepsilon(\cdot) = Z(\mathbf{t})^T \mathbf{C}_\varepsilon^{-1} v_\varepsilon(\cdot) = \sum_{j=1}^n \left(\sum_{l=-(k_1-1)}^{\infty} c_{ij} \varepsilon^l \right) \cdot \left(\sum_{k=0}^{\infty} (-1)^k \frac{\varepsilon^{2k}}{k!} \frac{\|\cdot - t_j\|^{2k}}{2^k} \right).$$

By [Schaback \(2005\)](#), we get that this product of series converges to a polynomial p , so it has the form

$$\begin{aligned} s_\varepsilon(\cdot) &\sim \sum_{k=0}^{\lfloor \frac{k_1-1}{2} \rfloor} \sum_{j=1}^n \tilde{c}_{jk} \|\cdot - t_j\|^{2k} + o(1) \\ &\xrightarrow{\varepsilon \rightarrow 0} \sum_{k=0}^{\lfloor \frac{k_1-1}{2} \rfloor} \sum_{j=1}^n \tilde{c}_{jk} \|\cdot - t_j\|^{2k} = p(\cdot) \in P_{k_1-1}^d \end{aligned}$$

for some appropriate $\tilde{c}_{jk} \in \mathbb{R}$. However, by (2.17) above, $p(\cdot) \notin P_{k_1-1}^d$ a.s.

3. We write $\mathbf{C}_\varepsilon^{-1} = \sum_{k=-2k_2}^{\infty} \mathbf{A}_k \varepsilon^k$ for matrices \mathbf{A}_k . By k^* we denote the smallest $k \in \mathbb{Z}$ such that $\mathbf{A}_k \neq 0$. Then, we immediately get that the limits $C_1(Z(\mathbf{t}))$ and $C_2(Z(\mathbf{t}))$ considered in the lemma exist and have the form

$$C_1(Z(\mathbf{t})) = \frac{\|Z(\mathbf{t})^T \mathbf{A}_{k^*}\|_V}{\|\mathbf{A}_{k^*}\|_M} \quad \text{and} \quad C_2(Z(\mathbf{t})) = \frac{Z(\mathbf{t})^T \mathbf{A}_{k^*} Z(\mathbf{t})}{\|Z(\mathbf{t})^T \mathbf{A}_{k^*}\|_V}.$$

We note that

$$\mathbf{A}_{k^*} = \lim_{\varepsilon \rightarrow 0} \mathbf{C}_\varepsilon^{-1} \varepsilon^{-k^*}$$

is a positive semi-definite and symmetric matrix and so there exists a lower triangular matrix \mathbf{L} such that $\mathbf{A}_{k^*} = \mathbf{L}^T \mathbf{L}$.

Assume $Z(\mathbf{t})^T \mathbf{A}_{k^*} Z(\mathbf{t}) = 0$. Then, we have $(\mathbf{L}Z(\mathbf{t}))^T (\mathbf{L}Z(\mathbf{t})) = 0$, so $\mathbf{L}Z(\mathbf{t}) = 0$ and $\mathbf{A}_{k^*} Z(\mathbf{t}) = \mathbf{L}^T \mathbf{L}Z(\mathbf{t}) = 0$. This means $Z(\mathbf{t}) \in \ker \mathbf{A}_{k^*}$. As $\mathbf{A}_{k^*} \neq 0$, we have $\dim(\ker \mathbf{A}_{k^*}) < n$, and therefore $Z(\mathbf{t}) \in \ker \mathbf{A}_{k^*}$ with probability zero. This means $Z(\mathbf{t})^T \mathbf{A}_{k^*} Z(\mathbf{t}) \neq 0$ and $Z(\mathbf{t})^T \mathbf{A}_{k^*} \neq \mathbf{0}$ with probability one. Therefore, $0 < C_1(Z(\mathbf{t})) < \infty$ and $0 < C_2(Z(\mathbf{t})) < \infty$ almost surely. \square

Lemma 2.11. *Let $T = \{t_1, \dots, t_n\} \subset \Upsilon$ and $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$ be normally distributed with non-degenerated density. Then,*

$$\lim_{\tau \rightarrow \infty} \|s_{Z,T,\tau}\|_{\kappa_\tau}^2 = \infty \quad a.s.$$

For $|T| > 1$, we also have

$$\lim_{\tau \rightarrow \infty} \|s_{Z,T,\tau}\|_{\tilde{\kappa}_\tau}^2 = \infty \quad a.s.$$

Proof. The idea of this proof is to consider some scaled Gaussian covariance function such that the native space norm of the corresponding interpolant is asymptotically smaller than the one corresponding to $\tilde{\kappa}_\tau$. The first one is increasing by a rate which can be assessed by Lemma 2.10.

In a first step we consider the Gaussian covariance function

$$\phi_{1/\sqrt{\tau}}(h) = \exp\left(-\frac{\|h\|^2}{2\tau}\right), \quad h \in \mathbb{R}^d, \quad \tau > \frac{d}{2}.$$

The corresponding Fourier transform is given by

$$\hat{\phi}_{1/\sqrt{\tau}}(\omega) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \phi_{1/\sqrt{\tau}} e^{-ix^T \omega} dx = \tau^{d/2} \exp\left(-\frac{\|\omega\|^2 \tau}{2}\right), \quad \omega \in \mathbb{R}^d.$$

As a second step we define a covariance function C_τ via its Fourier transform

$$\hat{C}_\tau(\omega) = (2\tau)^{d/2} \exp\left(-\frac{\|\omega\|^2 \tau}{2}\right) \mathbf{1}_{\{\|\omega\|^2 \leq 2\}} + 2^{d/2} \frac{\Gamma(\tau)}{\Gamma(\tau - d/2)} (1 + \|\omega\|^2)^{-\tau} \mathbf{1}_{\{\|\omega\|^2 > 2\}}.$$

Then, using inverse Fourier transforms, we get for $h \in \mathbb{R}^d$ that

$$\begin{aligned} & |2^{d/2} \phi_{1/\sqrt{\tau}}(h) - C_\tau(h)| \\ &= \left| \int_{\mathbb{R}^d} \left(\left(\frac{\tau}{\pi}\right)^{d/2} \exp\left(-\frac{\|\omega\|^2 \tau}{2}\right) - \frac{\Gamma(\tau) \pi^{-d/2}}{\Gamma(\tau - d/2)} (1 + \|\omega\|^2)^{-\tau} \right) \mathbf{1}_{\{\|\omega\|^2 > 2\}} \cos(h^T \omega) d\omega \right| \\ &\leq \int_{\mathbb{R}^d} \left(\frac{\tau}{\pi}\right)^{d/2} \exp\left(-\frac{\|\omega\|^2 \tau}{2}\right) \mathbf{1}_{\{\|\omega\|^2 > 2\}} d\omega \\ &\quad + \int_{\mathbb{R}^d} \left(\frac{1}{\pi}\right)^{d/2} \frac{\Gamma(\tau)}{\Gamma(\tau - d/2)} (1 + \|\omega\|^2)^{-\tau} \mathbf{1}_{\{\|\omega\|^2 > 2\}} d\omega \end{aligned} \quad (2.18)$$

The first summand of (2.18) can be rewritten as $\int_{\mathbb{R}^d} \frac{1}{\sqrt{\pi^d}} \exp\left(-\frac{\|\omega\|^2}{2}\right) \mathbf{1}_{\{\|\omega\|^2 > 2\tau\}} d\omega$ which is — up to the factor $2^{d/2}$ — the probability of a χ^2 -distribution with d degrees of freedom to be larger than 2τ (cf. Abramowitz and Stegun, 1965, Section 26.4). This can be bounded by $\frac{c_d}{\sqrt{\tau}} \exp\left(-\frac{\tau}{d}\right)$ for some $c_d > 0$, for instance.

Since $\frac{\Gamma(\tau)}{\Gamma(\tau-d/2)} \sim (\tau-d/2)^{d/2}$ for τ sufficiently large (cf. Gradshteyn and Ryzhik, 2007, Formula 8.328.2), the second summand of (2.18) equals asymptotically

$$\int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} \int_2^\infty \left(\frac{\tau-d/2}{\pi}\right)^{d/2} (1+r^2)^{-\tau} r^{d-1} dr d\varphi \prod_{j=1}^{d-2} ((\sin \theta_j)^j d\theta_j),$$

which can be seen via transformation to polar coordinates (cf. Amann and Escher, 2008, Section X.8). Thus, the absolute value of the inner integral can be bounded by $\int_2^\infty r^{-2\tau+d-1} (\tau-d/2)^{d/2} dr = \frac{(\tau-d/2)^{d/2}}{2\tau-d} 2^{-2\tau+d}$.

All in all, we have that $|C_\tau(h) - 2^{d/2} \phi_{\tau-1/2}(h)|$ decreases exponentially as $\tau \rightarrow \infty$ for all $h \in \mathbb{R}^d$. The same holds true for $\|(C_\tau(t_i - t_j))_{i,j} - (2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}\|$.

Using $\|2^{d/2} (\phi_{\tau-1/2}(x_i - x_j))_{i,j}^{-1}\| \sim \tilde{K} \tau^{k^*/2}$ for some $k^* \leq 2k_2$ and $\tilde{K} > 0$ (cf. the first part of Lemma 2.10) we get

$$\begin{aligned} & \frac{\|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1} Z(\mathbf{t}) - (C_\tau(x_i - x_j))_{i,j}^{-1} Z(\mathbf{t})\|}{\|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1} Z(\mathbf{t})\|} \\ & \leq \frac{\|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j} - (C_\tau(t_i - t_j))_{i,j}\| \cdot \|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1}\|}{1 - \|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1}\| \cdot \|(2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j} - (C_\tau(t_i - t_j))_{i,j}\|} \in o(1) \end{aligned}$$

for τ large enough by Lange (1999), Section 6.5. Thus, almost surely, there is $K = K(Z(\mathbf{t})) > 0$ such that

$$\begin{aligned} & Z(\mathbf{t})^T (C_\tau(t_i - t_j))_{i,j}^{-1} Z(\mathbf{t}) \\ & = Z(\mathbf{t})^T (2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1} Z(\mathbf{t}) + o(\|Z(\mathbf{t})^T (2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1}\|) \\ & = (C_1(Z(\mathbf{t})) + o(1)) \cdot \|Z(\mathbf{t})^T (2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1}\| \\ & \sim C_1(Z(\mathbf{t})) \cdot \|Z(\mathbf{t})^T (2^{d/2} \phi_{\tau-1/2}(t_i - t_j))_{i,j}^{-1}\| \geq K \cdot \tau^{k_1/2} \end{aligned} \tag{2.19}$$

for τ large enough by the second and third part of Lemma 2.10.

As a last step we consider the difference $C_\tau(\cdot) - \tilde{\kappa}_{\tau-d/2}(\cdot)$. Using again $2^{d/2} \frac{\Gamma(\tau)}{\Gamma(\tau-d/2)} \sim (2\tau-d)^{d/2}$ for τ sufficiently large, we can see that the Fourier transform

$$\widehat{C}_\tau(\omega) - \widehat{\tilde{\kappa}}_{\tau-d/2}(\omega) \sim \left((2\tau)^{d/2} \exp(-\|\omega\|^2 \tau/2) - (2\tau-d)^{d/2} (1+\|\omega\|^2)^{-\tau} \right) \mathbf{1}_{\{\|\omega\|^2 \leq 2\}}$$

is nonnegative everywhere and strictly positive on a Lebesgue non-null set and we have that the matrix

$$\underbrace{(C_\tau(t_i - t_j))_{i,j}}_{=: A_1(\tau)} - \underbrace{(\tilde{\kappa}_{\tau-d/2}(t_i - t_j))_{i,j}}_{=: A_2(\tau)}$$

is strictly positive definite (cf. Wendland, 2005, Thm. 6.8).

2. Estimating the Whittle-Matérn Smoothness Parameter via the Native Space Norm

It is well-known that this implies that $A_2^{-1}(\tau) - A_1^{-1}(\tau)$ is positive semidefinite (see [Horn and Johnson, 1996](#), Corollary 7.7.4). Hence, for τ large enough,

$$\|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2 = Z(\mathbf{t})^T A_2^{-1}(\tau) Z(\mathbf{t}) \geq Z(\mathbf{t})^T A_1^{-1}(\tau) Z(\mathbf{t}) \geq K \tau^{\frac{k_1}{2}} \xrightarrow{\tau \rightarrow \infty} \infty$$

with probability one if $k_1 > 0$. This condition holds if and only if $n > 1$ since $\binom{0+d}{d} = 1$.

Furthermore, we have

$$\|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2 \sim (2\tau - d)^{d/2} \|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2 \gtrsim (2\tau - d)^{d/2} \tau^{\frac{k_1}{2}} \xrightarrow{\tau \rightarrow \infty} \infty$$

for any $n \in \mathbb{N}$. □

The proof of Lemma 2.11 was based on results on flat limits for Gaussian kernels. Recently, [Song et al. \(2012\)](#) proved the existence of flat limits for the Whittle-Matérn kernel κ_ν , $\nu \in \mathbb{N}$, provided that the set T is unisolvent w.r.t. the set $P_{2\nu}^d$ of polynomials of degree up to 2ν . Here, we consider “smooth limits” for Whittle-Matérn kernels based on some fixed set T . Note that this set is not unisolvent w.r.t. $P_{2\nu}^d$ as $\nu \rightarrow \infty$. Therefore, we cannot resort to the results by [Song et al. \(2012\)](#), but again use results on Gaussian kernels instead.

Proposition 2.12. *Let $T = \{t_1, \dots, t_n\} \subset \Upsilon$ such that $k_2 \leq 2$ and $f : \Upsilon \rightarrow \mathbb{R}$. Then, the sequence of interpolants $(s_{f,T,m})_{m \in \mathbb{N}}$ w.r.t. the Whittle-Matérn kernel converges (point-wise) to a polynomial.*

Proof. For the proof we use the series expansion of the modified Bessel function given in [Abramowitz and Stegun \(1965\)](#):

$$\begin{aligned} \mathcal{K}_m(r) &= \frac{1}{2} \left(\frac{1}{2}r\right)^{-m} \sum_{k=0}^{m-1} \frac{(m-k-1)!}{k!} \left(-\frac{1}{4}r^2\right)^k + (-1)^{m+1} \log\left(\frac{1}{2}r\right) I_m(r) \\ &\quad + (-1)^m \frac{1}{2} \left(\frac{1}{2}r\right)^m \sum_{k=0}^{\infty} (\Psi(k+1) + \Psi(m+k+1)) \frac{(\frac{1}{4}r^2)^k}{k!(m+k)!}, \quad m \in \mathbb{N}, r \geq 0, \end{aligned}$$

with

$$I_m(r) = \left(\frac{1}{2}r\right)^m \sum_{k=0}^{\infty} \frac{(\frac{1}{4}r^2)^k}{k! \Gamma(m+k+1)}$$

and

$$\Psi(1) = -\gamma, \quad \Psi(m) = -\gamma + \sum_{k=1}^{m-1} k^{-1}$$

where $\gamma \approx 0.5772$ is the Euler-Mascheroni constant.

So we get

$$\tilde{\kappa}_m(h) = 2 \left(\frac{1}{2}\|h\|\right)^m \frac{\mathcal{K}_m(\|h\|)}{(m-1)!}$$

$$\begin{aligned}
 &= \sum_{k=0}^{m-1} \frac{(m-k-1)!}{(m-1)!} \frac{(-\frac{1}{4}\|h\|^2)^k}{k!} + \frac{2 \cdot (-1)^{m+1}}{(m-1)!} \log\left(\frac{1}{2}\|h\|\right) \sum_{k=0}^{\infty} \frac{1}{k!} \frac{(\frac{1}{4}\|h\|^2)^{m+k}}{(m+k)!} \\
 &\quad + \frac{(-1)^n}{(m-1)!} \sum_{k=0}^{\infty} \frac{\Psi(k+1) + \Psi(m+k+1)}{k!} \frac{(\frac{1}{4}\|h\|^2)^{m+k}}{(m+k)!},
 \end{aligned}$$

which means

$$\begin{aligned}
 \tilde{\kappa}_m(h) &= 1 - \frac{1}{m-1} \left(\frac{1}{4}\|h\|^2 \right) + \frac{1}{(m-1)(m-2)} \frac{(\frac{1}{4}\|h\|^2)^2}{2!} \\
 &\quad + \sum_{k=3}^{\infty} (-1)^k \frac{1}{(m-1) \cdots (m-k)} \frac{(\frac{1}{4}\|h\|^2)^k}{k!} + o\left(\frac{1}{(m-1)!}\right), \quad m \rightarrow \infty.
 \end{aligned}$$

Thus, we obtain

$$\begin{aligned}
 \left| \tilde{\kappa}_m(h) - \phi_{\frac{1}{\sqrt{2(m-1)}}}(h) \right| &= \left| \sum_{k=2}^{m-1} \left(\frac{1}{(m-1) \cdot (m-2) \cdots (m-k)} \right) \frac{(-\frac{1}{4}\|h\|^2)^k}{k!} \right. \\
 &\quad \left. - \sum_{k=m}^{\infty} \frac{(-1)^k}{(m-1)^k} \frac{(-\frac{1}{4}\|h\|^2)^k}{k!} + o\left(\frac{1}{(m-1)!}\right) \right| \in \mathcal{O}\left(\frac{1}{m^3}\right).
 \end{aligned}$$

On the other hand, by Lemma 2.10, we have

$$\left\| \left(\phi_{1/\sqrt{2(m-1)}}(t_i - t_j) \right)_{i,j=1,\dots,n}^{-1} \right\| \in \mathcal{O}(m^{k_2}).$$

Thus, by Schaback and Wendland (2004), Thm. 3.18, for $\tau_m = \frac{1}{\sqrt{2(m-1)}}$ and m large enough, we get the inequality

$$\begin{aligned}
 &\| \phi_{\tau_m}(t_i - t_j)_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i - (\tilde{\kappa}_m(t_i - t_j))_{i,j}^{-1} (\tilde{\kappa}_m(t - t_i))_i \| \\
 &\leq \frac{\| (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} \| \cdot \| (\phi_{\tau_m}(t_i - t_j))_{i,j} \|}{1 - \| (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} \| \cdot \| (\phi_{\tau_m}(t_i - t_j))_{i,j} - (\tilde{\kappa}_m(t_i - t_j))_{i,j} \|} \\
 &\quad \cdot \left(\frac{\| (\phi_{\tau_m}(t_i - t_j))_{i,j} - (\tilde{\kappa}_m(t_i - t_j))_{i,j} \|}{\| (\phi_{\tau_m}(t_i - t_j))_{i,j} \|} + \frac{\| (\phi_{\tau_m}(t - t_i))_i - (\tilde{\kappa}_m(t - t_i))_i \|}{\| (\phi_{\tau_m}(t - t_j))_i \|} \right) \\
 &\quad \cdot \| (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i \| \\
 &= o(\| (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i \|)
 \end{aligned}$$

for any $t \in \Upsilon$ since $k_2 < 3$.

As $\mathbf{F}^T (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i$ converges to a polynomial for any $\mathbf{F} \in \mathbb{R}^n$ by Theorem 1 in Schaback (2008), we get that $\| (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i \|$ is bounded. Therefore,

$$\| \phi_{\tau_m}(t_i - t_j)_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i - (\tilde{\kappa}_m(t_i - t_j))_{i,j}^{-1} (\tilde{\kappa}_m(t - t_i))_i \| \in o(1)$$

for any $t \in \Upsilon$.

Thus, for all $\mathbf{F} = (f(t_1), \dots, f(t_n)) \in \mathbb{R}^n$ we have

$$s_{f,T,m}(t) = \mathbf{F}^T (\tilde{\kappa}_m(t_i - t_j))_{i,j}^{-1} (\tilde{\kappa}_m(t - t_i))_i \sim \mathbf{F}^T (\phi_{\tau_m}(t_i - t_j))_{i,j}^{-1} (\phi_{\tau_m}(t - t_i))_i$$

and this term converges to a polynomial. \square

Lemma 2.13. *For any $f : \Upsilon \rightarrow \mathbb{R}$ with $f|_T \not\equiv 0$ it holds*

1. $\lim_{\nu \rightarrow 0} \|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2 \in (0, \infty)$
2. $\lim_{\nu \rightarrow 0} \|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2 = 0$

Proof. 1. Let $\tilde{\mathbf{K}}_\nu = (\tilde{\kappa}_\nu(t_i - t_j))_{i,j=1,\dots,n}$. Then, we have $\tilde{\mathbf{K}}_\nu \xrightarrow{\nu \rightarrow 0} \mathbf{I}_n$ where $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is the $n \times n$ -identity matrix. Therefore, as all matrices $\tilde{\mathbf{K}}_\nu$ are strictly positive definite and matrix multiplication and inversion are continuous,

$$\|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2 = \mathbf{F}^T \tilde{\mathbf{K}}_\nu^{-1} \mathbf{F} \xrightarrow{\nu \rightarrow 0} \mathbf{F}^T \mathbf{F} > 0 \quad \text{for } \mathbf{F} \neq \mathbf{0}$$

where $\mathbf{F} = (f(t_1), \dots, f(t_n))^T$.

2. By Formulae 6.2.1 and 6.2.2 in [Abramowitz and Stegun \(1965\)](#), we have $\frac{\Gamma(\nu)}{\Gamma(\nu+d/2)} = \Gamma(d/2) \cdot B(\nu, d/2)$, where B denotes the beta function, and therefore

$$\frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \sim \frac{1}{\int_0^1 t^{\nu-1} (1-t)^{d/2-1} dt} \leq \frac{1}{\int_0^{1/2} t^{\nu-1} (1/2)^{d/2-1} dt} \xrightarrow{\nu \rightarrow 0} 0.$$

Using $\|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2 \sim \frac{\Gamma(\nu+d/2)}{\Gamma(\nu)} \|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2$ and the first part of this lemma, we get $\lim_{\nu \rightarrow 0} \|s_{f,T,\nu}\|_{\tilde{\kappa}_\nu}^2 = 0$. □

By these considerations we can proof the existence of a minimum of

$$G_{T,\lambda}(\tau) := \mathcal{E}(\tau) + \lambda(\tau) \cdot \|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2,$$

where \mathcal{E} is some error function, under certain conditions.

Theorem 2.14. *Let $0 < \varepsilon < \frac{d}{2} < \nu_0$, and $Z(\cdot)$ be a stationary zero mean Gaussian random field with covariance function $\tilde{\kappa}_{\nu_0}$. Furthermore, let $\mathcal{E} : (0, \infty) \rightarrow [0, \infty)$ be continuous and $\lambda : (\frac{d}{2}, \infty) \rightarrow [0, \infty)$ be continuous and eventually larger than some $C > 0$. Let $T \subset \Upsilon$ be finite. Then, $\min_{\tau \in [d/2+\varepsilon, \infty)} G_{T,\lambda}(\tau)$ exists with probability one.*

Proof. Note that $\tau \mapsto \mathcal{K}_\tau(h)$ and $\tau \mapsto \kappa_\tau(h)$ are continuous on $(0, \infty)$ for all $h \in \mathbb{R}^d$. Employing

$$\|s_{Z,T,\tau}\|_{\tilde{\kappa}_\tau}^2 = Z(\mathbf{t})^T (\kappa_\tau)_{i,j}^{-1} Z(\mathbf{t})$$

where $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$, we get that the mapping $G_{T,\lambda} : (\frac{d}{2}, \infty) \rightarrow [0, \infty)$ is continuous since matrix multiplication and inversion are continuous operations. For $\mu - \frac{d}{2} \in (\varepsilon, \nu_0)$ we have $Z_\omega(\cdot) \in W^{\mu-d/2,2}(T)$ with probability one by [Scheuerer \(2010\)](#). Theorem 2.6 yields

$$\min_{\tau \in [\frac{d}{2}+\varepsilon, \mu]} G_{T,\lambda}(\tau) \leq G_{T,\lambda}(\mu) \leq \mathcal{E}(\mu) + \lambda(\mu) \cdot c_Z^2 =: \tilde{C} \quad \forall T \subset \Upsilon.$$

By the assumptions on $\lambda(\cdot)$ we have $\lambda(\tau) > C$ for all τ larger than some μ_1 . Furthermore, by Lemma 2.11, there exists some μ_2 such that $C \cdot \|s_{Z,T,\mu_2-d/2}\|_{\tilde{\kappa}_{\mu_2-d/2}}^2 > \tilde{C}$ and by

Proposition 2.9 the same inequality holds for all $\tau \geq \mu_2$. Since $\mathcal{E}(\tau)$ is nonnegative, we have $G_{T,\lambda}(\tau) > \tilde{C}$ for all $\tau > \mu^* = \max\{\mu_1, \mu_2\}$. Therefore,

$$\inf_{\tau \in [\frac{d}{2} + \varepsilon, \infty)} G_{T,\lambda}(\tau) = \inf_{\tau \in [\frac{d}{2} + \varepsilon, \mu^*]} G_{T,\lambda}(\tau) = \min_{\tau \in [\frac{d}{2} + \varepsilon, \mu^*]} G_{T,\lambda}(\tau).$$

The last equation holds because of the continuity of $G_{T,\lambda}(\cdot)$ and the compactness of $[\frac{d}{2} + \varepsilon, \mu^*]$. \square

Remark 2.15. 1. In fact, the condition $\frac{d}{2} < \nu_0$ is not necessary for the existence of the minimum stated above. Nevertheless, it guarantees that $Z_\omega(\cdot)$ is element of the Hilbert spaces $\mathcal{H}(\kappa_\mu; \Upsilon)$, $\mu < \nu_0 - \frac{d}{2}$. Only under this condition we can expect to re-estimate ν_0 correctly by minimizing $G_{T,\lambda}(\cdot)$.

2. In the case $|T| > 1$, Theorem 2.14 also holds true if we replace $G_{T,\lambda}(\nu)$ by

$$\tilde{G}_{T,\lambda}(\tau) := \mathcal{E}(\tau) + \lambda(\tau) \cdot \|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2.$$

Note that $\tau \mapsto \|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2$ is not monotonically increasing, but tends to infinity as $\tau \rightarrow \infty$ by Lemma 2.11 and therefore $C \cdot \|s_{Z,T,\tau-d/2}\|_{\tilde{\kappa}_{\tau-d/2}}^2$ is eventually larger than any \tilde{C} . In contrast to $G_{T,\lambda}$, the function $\tilde{G}_{T,\lambda}$ also allows for penalizing for small values of τ as $\lim_{\nu \searrow 0} \|s_{Z,T,\nu}\|_{\tilde{\kappa}_\nu} > 0 = \lim_{\nu \searrow 0} \|s_{Z,T,\nu}\|_{\kappa_\nu}$ (Lemma 2.13).

As the minimum exists under the conditions of Theorem 2.14, we can define the following estimators, which we call *native space penalty estimators*,

$$\hat{\nu}_{NP1} = \operatorname{argmin}_{\nu > d/2} G_{\lambda,T}(\nu) \quad \text{and} \quad \hat{\nu}_{NP2} = \operatorname{argmin}_{\nu > d/2} \tilde{G}_{\lambda,T}(\nu), \quad (2.20)$$

respectively. Considering these estimators, we have to deal with the choice of the error function \mathcal{E} and the tuning parameter function λ .

2.5 The Choice of the Tuning Parameter Function $\lambda(\nu)$

Remark 2.16. Let $(\tilde{\kappa}_\nu(t_i - t_j))_{i,j} = \mathbf{L}_\nu \mathbf{L}_\nu^T$ be the Cholesky decomposition of $(\tilde{\kappa}_\nu(t_i - t_j))_{i,j}$. Then, for a zero mean stationary Gaussian random field Z with covariance function κ_{ν_0} and a data vector $Z(\mathbf{t}) = (Z(t_1), \dots, Z(t_n))^T$ we have $(\mathbf{L}_{\nu_0}^{-1})Z(\mathbf{t}) \sim \mathcal{N}(0, \mathbf{I}_n)$ and therefore $\|s_{Z,T,\nu_0}\|_{\tilde{\kappa}_{\nu_0}}^2 \sim \chi_n^2$, i.e. $\|s_{Z,T,\nu_0}\|_{\tilde{\kappa}_{\nu_0}}^2$ is χ^2 -distributed with n degrees of freedom (cf. Abramowitz and Stegun, 1965, Section 26.4).

It seems reasonable to take a tuning parameter which represents the probability of the observed value $\|s_{Z,T,\nu}\|_{\tilde{\kappa}_\nu}^2$ under the assumption that $\tilde{\kappa}_\nu$ is the true covariance function, which means to choose $\lambda(\cdot)$ as a data-dependent function of $p(\nu) = \mathbb{P}(X_n \leq \|s_{Z,T,\nu}\|_{\tilde{\kappa}_\nu}^2 \mid Z(\mathbf{t}))$ where $X_n \sim \chi_n^2$ is independent of Z . Here, we want λ to be small for ‘‘common’’ values of $\|s_{Z,T,\nu}\|_{\tilde{\kappa}_\nu}^2$ and to be large for rare values, i.e. $p \approx 0$ or $p \approx 1$. One canonical choice of λ would be an indicator function of the type $\lambda(\nu) = \mathbf{1}_{|p(\nu) - \frac{1}{2}| > b}$ for some $b \in (0, \frac{1}{2})$. However, for Theorem 2.14 to hold, λ should be continuous. Therefore, we choose an continuous approximation to an indicator function:

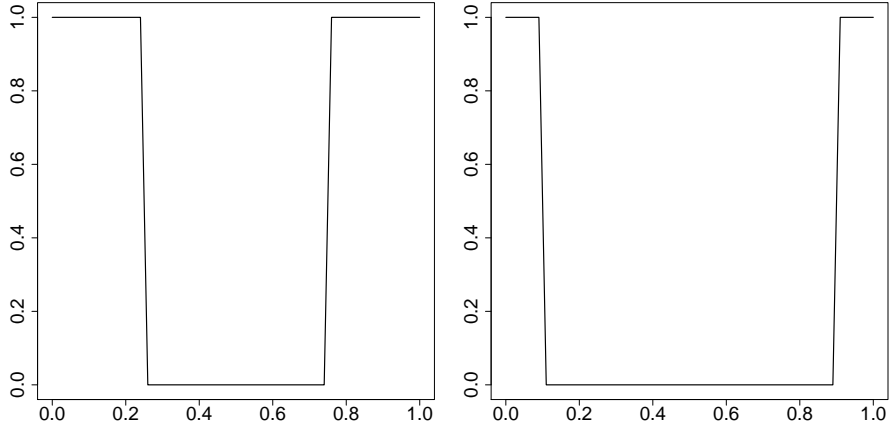


Figure 2.1: $\lambda(\nu)$ as a function of $p(\nu)$ for different parameters. Left: $b_1 = 0.24, b_2 = 0.26$. Right: $b_1 = 0.39, b_2 = 0.41$.

$$\lambda(\nu) = \begin{cases} 0, & |p(\nu) - \frac{1}{2}| \leq b_1 \\ -\frac{p(\nu) - (\frac{1}{2} - b_1)}{b_2 - b_1}, & \frac{1}{2} - b_2 < p(\nu) < \frac{1}{2} - b_1 \\ \frac{p(\nu) - (\frac{1}{2} + b_1)}{b_2 - b_1}, & \frac{1}{2} + b_1 < p(\nu) < \frac{1}{2} + b_2 \\ 1, & |p(\nu) - \frac{1}{2}| \geq b_2 \end{cases}$$

for appropriate parameters $0 \leq b_1 < b_2 \leq \frac{1}{2}$ (see Figure 2.1 for λ with two different parameter settings). Thus, λ is a nonnegative continuous function. By Lemma 2.11, we have $p(\nu) \xrightarrow{\nu \rightarrow \infty} 1$ a.s. if $|T| > 1$ and therefore λ is finally bounded away from zero.

2.6 Comparison with MLE and Cross Validation Estimator

Now, we compare the results of the native space penalty estimators $\hat{\nu}_{NP1}$ and $\hat{\nu}_{NP2}$ from Equation (2.20) to the results of other estimators. As an alternative estimator for ν_0 we consider the well-known *maximum likelihood* estimator (see Stein 1999, p. 169, for example)

$$\hat{\nu}_{MLE} = \operatorname{argmax}_{\nu > 0} \mathcal{L}(\nu).$$

Here, $\mathcal{L}(\cdot)$ denotes the Gaussian likelihood function

$$\mathcal{L}(\nu) = \frac{1}{\sqrt{2\pi}^n (\det(\tilde{\mathbf{K}}_\nu))^{1/2}} \exp\left(-\frac{1}{2} \mathbf{z}^T \tilde{\mathbf{K}}_\nu^{-1} \mathbf{z}\right),$$

where $\tilde{\mathbf{K}}_\nu = (\tilde{\kappa}_\nu(|t_i - t_j|))_{i,j=1,\dots,n}$ is the covariance matrix corresponding to the parameter $\nu > 0$ and $\mathbf{z} \in \mathbb{R}^n$ denotes the vector consisting of observations at locations t_1, \dots, t_n . Note that — for numerical reasons — often the log-likelihood function $\ell(\nu) = \log \mathcal{L}(\nu)$ is maximized instead of the likelihood function $\mathcal{L}(\nu)$ itself.

Furthermore, we consider the *leave-one-out cross validation* (LOOCV) estimator which was proposed by Rippa (1999) for scale parameter estimation. In contrast to $\hat{\nu}_{MLE}$, the

LOOCV estimator does not assume Z to be Gaussian. Here, one considers the error vector $\mathbf{E}(\nu) = (E_1(\nu), \dots, E_n(\nu))^T$ with $E_k(\nu) = Z(t_k) - s_{Z, T \setminus \{t_k\}, \kappa_\nu}(t_k)$ and defines

$$\hat{\nu}_{LCV} = \operatorname{argmin}_{\nu > 0} \|\mathbf{E}(\nu)\|_1.$$

Rightaway, the calculation of $\mathbf{E}(\nu)$ seems to be quite time-consuming because it involves solving n linear systems of dimension $(n-1) \times (n-1)$. However, [Rippa \(1999\)](#) showed that this can be done much faster: Let $\mathbf{a}(\nu)$ and $\mathbf{x}^{(k)}(\nu)$, $k = 1, \dots, n$, be the solution of the linear systems $\tilde{\mathbf{K}}_\nu \mathbf{a}(\nu) = \mathbf{z}$ and $\tilde{\mathbf{K}}_\nu \mathbf{x}^{(k)}(\nu) = \mathbf{e}_k$, respectively, where \mathbf{e}_k denotes the k -th standard basis vector in \mathbb{R}^n , $k = 1, \dots, n$. Then, one can rewrite $\mathbf{E}(\nu)$ in terms of $\mathbf{a}(\nu)$ and $\mathbf{x}^{(k)}(\nu)$ and one obtains $E_k(\nu) = \frac{\mathbf{a}_k(\nu)}{\mathbf{x}^{(k)}_k(\nu)}$. This simplification reduces the costs of calculating $\mathbf{E}(\nu)$ from $\mathcal{O}(n^4)$ to $\mathcal{O}(n^3)$.

However, this estimator may not take into account the geometry of $T = \{t_1, \dots, t_n\}$ in a suitable way. Therefore, [Scheuerer \(2011\)](#) proposes a *weighted cross validation* (WCV) estimator. Let

$$\epsilon_i(\nu) = \frac{|s_{f, \{t_1, \dots, t_{i-1}\}, \nu}(t_i) - f(t_i)|}{P_{\{t_1, \dots, t_{i-1}\}, \tilde{\kappa}_\nu}(t_i)}, \quad i = 1, \dots, N,$$

where $s_{f, \emptyset, \nu} \equiv 0$ and $P_{\emptyset, \tilde{\kappa}_\nu} \equiv 1$. By the weighting of the errors by the reciprocal power function we account for the expected accuracy of interpolation based on the geometry of T . Then, the weighted cross validation estimator is given by

$$\hat{\nu}_{WCV} = \operatorname{argmin}_{\nu > 0} \left(\sqrt[n]{\prod_{i=1}^n P_{\{t_1, \dots, t_{i-1}\}, \tilde{\kappa}_\nu}(t_i) \cdot \|\epsilon(\nu)\|_2^2} \right).$$

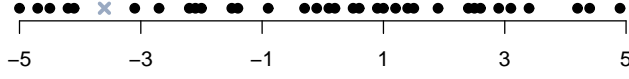
The computation can be simplified by using $\|\epsilon(\nu)\|_2^2 = \|s_{f, T, \nu}\|_{\tilde{\kappa}_\nu}^2$ (cf. [Scheuerer, 2011](#)).

Before we can apply the native space penalty estimators from Equation (2.20), we have to care for the choice of the error function $\mathcal{E}(\cdot)$ involved. The power function itself does not seem to be suitable as it determines the error only for the true smoothness parameter ν_0 . As the weighted cross validation estimator turns out to provide much better results than the leave-one-out cross validation estimator, we choose

$$\mathcal{E}(\nu) = \sqrt[n]{\prod_{i=1}^n P_{\{t_1, \dots, t_{i-1}\}, \tilde{\kappa}_\nu}(t_i) \cdot \|\epsilon(\nu)\|_2^2}.$$

The tuning parameter function λ is chosen as in Section 2.5 with parameters $b_1 = 0.24$ and $b_2 = 0.26$.

We do a performance study in order to compare the estimators $\hat{\nu}_{MLE}$, $\hat{\nu}_{LOOCV}$, $\hat{\nu}_{WCV}$, $\hat{\nu}_{NP1}$ and $\hat{\nu}_{NP2}$ as well as their impact on kriging errors. To this end, we simulate data Z_1, \dots, Z_K of a stationary Gaussian random field with covariance function $\tilde{\kappa}_{\nu_0}$ with $\nu_0 \in \{0.75, 1.25, 1.75, 2.25\}$ on the set $Y = \{-5, -4.9, \dots, 4.9, 5\} \subset \mathbb{R}$ using the software environment **R** ([Ihaka and Gentleman, 1996](#); [R Development Core Team, 2011](#)), in particular the **R** package **RandomFields** ([Schlather, 2012](#)). Each simulation is repeated independently $K = 500$ times. Estimation is performed based on the data on a set $T \subset Y$ with $|T| = 34$ (see Figure 2.2) yielding estimates $\hat{\nu}^{(1)}, \dots, \hat{\nu}^{(K)}$. In a first step we determine the mean squared error for each estimator. Furthermore, we use the estimated


 Figure 2.2: The set T (black dots) and the point t_0 (grey cross).

ν_0	$\hat{\nu}_{MLE}$ MSE	$\hat{\nu}_{NP1}$ MSE	$\hat{\nu}_{NP2}$ MSE	$\hat{\nu}_{WCV}$ MSE	$\hat{\nu}_{LCV}$ MSE	$\hat{\nu}_{MLE}$ MSKE	$\hat{\nu}_{NP1}$ MSKE	$\hat{\nu}_{NP2}$ MSKE	$\hat{\nu}_{WCV}$ MSKE	$\hat{\nu}_{LCV}$ MSKE
0.75	0.0059	0.0091	0.0061	0.0169	0.139	0.2555	0.2562	0.2563	0.2582	0.2998
1.25	0.0057	0.0064	0.0064	0.0214	0.282	0.0571	0.0572	0.0572	0.0573	0.0640
1.75	0.0058	0.0069	0.0069	0.0232	0.332	0.0123	0.0123	0.0123	0.0123	0.0142
2.25	0.0066	0.0074	0.0074	0.0252	0.369	0.0021	0.0021	0.0021	0.0021	0.0024

 Table 2.1: Estimation of ν_0 in the case of a known variance ($\sigma^2 = 1$) and a Gaussian random field. Results of a simulation study with 500 simulated data sets for each ν_0 . The mean squared error (MSE) for the estimation of ν and the mean squared kriging error (MSKE) at location t_0 are shown.

parameters for kriging based on the data on T and determine the *mean squared kriging error* at location t_0 (see Figure 2.2) which is

$$MSKE(t_0) = \frac{1}{K} \sum_{i=1}^K (s_{Z_i, T, \hat{\nu}^{(i)}}(t_0) - Z_i(t_0))^2.$$

The results are presented in Table 2.1. As a first result, the WCV estimator performs much better than the LOOCV estimator. Furthermore, the estimators $\hat{\nu}_{MLE}$, $\hat{\nu}_{NP1}$ and $\hat{\nu}_{NP2}$ which assume Gaussianity, provide better results than the cross validation ones which are not based on this assumption. The native space penalty estimators perform slightly worse than the maximum likelihood estimator, in particular with respect to the MSE. Among the native space penalty estimators, differences only occur if ν_0 is small as penalization of small τ is crucial in this case. Here, the results of $\hat{\nu}_{NP2}$ are a little better than the ones of $\hat{\nu}_{NP1}$.

Furthermore, we analyse the case where the assumption of Gaussianity is not perfectly satisfied. At the same time, the random field we consider should allow for controlling the sample path smoothness. Here, we choose Gaussian random fields and truncate with thresholds -1.5 and 1.5 . As truncation functions are once weakly differentiable, the smoothness of the sample paths should not be affected by the truncation for $\nu \leq 1$. We redo the same simulation study as before with truncated Gaussian random fields for $\nu_0 = 0.7, 0.9$. The results are shown in Table 2.2. In this case of slight deviations from Gaussianity, $\hat{\nu}_{NP2}$ performs as good as $\hat{\nu}_{MLE}$.

2.7 The Unknown Variance Case

In the previous sections, we considered Gaussian random fields with covariance function $\tilde{\kappa}_{\nu_0}$, that is, with variance $\sigma^2 = 1$. Now, we want to modify $\lambda(\nu)$ for the case of unknown

ν_0	$\widehat{\nu}_{MLE}$ MSE	$\widehat{\nu}_{NP1}$ MSE	$\widehat{\nu}_{NP2}$ MSE	$\widehat{\nu}_{WCV}$ MSE	$\widehat{\nu}_{LCV}$ MSE	$\widehat{\nu}_{MLE}$ MSKE	$\widehat{\nu}_{NP1}$ MSKE	$\widehat{\nu}_{NP2}$ MSKE	$\widehat{\nu}_{WCV}$ MSKE	$\widehat{\nu}_{LCV}$ MSKE
0.7	0.0116	0.0131	0.0107	0.0171	0.119	0.2431	0.2441	0.2434	0.2444	0.2741
0.9	0.0107	0.0142	0.0106	0.0200	0.192	0.1391	0.1395	0.1394	0.1405	0.1573

Table 2.2: Estimation of ν_0 in the case of a known variance ($\sigma^2 = 1$) and a truncated Gaussian random field. Results of a simulation study with 500 simulated data sets for each $\nu_0 = 0.7, 0.9$. The mean squared error (MSE) for the estimation of ν_0 and the mean squared kriging error (MSKE) at location t_0 are shown.

variance σ^2 , i.e. for the case of a stationary Gaussian random field with covariance function $\sigma^2 \cdot \tilde{\kappa}_{\nu_0}$. Therefore, we should choose $\lambda(\cdot)$ in a way that the behaviour of $G_{T,\lambda}$ (and $\tilde{G}_{T,\lambda}$, respectively) is invariant under multiplying Z with some norming constant (which is modifying the variance of the data). As $\mathcal{E}(\nu)$, $\|s_{Z,T,\nu-d/2}\|_{\tilde{\kappa}_\nu}^2$ and $\|s_{Z,T,\nu-d/2}\|_{\tilde{\kappa}_\nu}^2$ all depend on the data in a quadratic way, $\widehat{\nu}_{NP1}$ and $\widehat{\nu}_{NP2}$ become invariant w.r.t. the variance of the data if $\lambda(\cdot)$ does not depend on the variance.

Let \mathbf{L}_ν be as in Remark 2.16. Then, we have for $Z(\mathbf{t}) \sim \mathcal{N}(0, \sigma^2(\tilde{\kappa}_{\nu_0}(x_i - x_j))_{i,j})$ that

$$(g_1(\nu), \dots, g_n(\nu))^T = (\mathbf{L}_{\nu_0}^{-1})^T Z(\mathbf{t}) \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n).$$

So we get that

$$\frac{\sum_{i=1}^{\lfloor n/2 \rfloor} g_i^2(\nu_0)}{\sum_{i=\lfloor n/2 \rfloor + 1}^{2\lfloor n/2 \rfloor} g_i^2(\nu_0)} \sim F_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$$

is F-distributed and the distribution does not depend on the variance (cf. [Abramowitz and Stegun, 1965](#), Section 26.6). Thus, we can define

$$p(\nu) = \mathbb{P} \left(Y_n \leq \frac{\sum_{i=1}^{\lfloor n/2 \rfloor} g_i^2(\nu)}{\sum_{i=\lfloor n/2 \rfloor + 1}^{2\lfloor n/2 \rfloor} g_i^2(\nu)} \mid Z(\mathbf{t}) \right),$$

where $Y_n \sim F_{\lfloor n/2 \rfloor, \lfloor n/2 \rfloor}$ is independent of Z , and choose $\lambda(\nu)$ as a function of $p(\nu)$ as in Section 2.5 with parameters $b_1 = 0.39$ and $b_2 = 0.41$.

Note that the maximum likelihood estimator has to be modified for the unknown variance case, as well. Here, the profile log-likelihood is given by

$$\ell(\nu) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} (1 - \log n) - \frac{1}{2} \log(\det(\tilde{\mathbf{K}}_\nu)) - \frac{n}{2} \log(\mathbf{z}^T \tilde{\mathbf{K}}_\nu^{-1} \mathbf{z})$$

([Scheuerer, 2009](#)). The resulting estimator $\widehat{\nu}_{MLE} = \operatorname{argmin}_{\nu > 0} \ell(\nu)$ is equivalent to the weighted cross validation estimator (cf. [Scheuerer, 2011](#)) which motivates the accuracy of maximum likelihood estimation even in a non-Gaussian framework. The cross validation estimators are invariant under changes of the variance.

We redo the above simulation study with Gaussian random fields and the true but unknown variance $\sigma^2 = 10$. The results are shown in Table 2.3. Here, the estimators $\widehat{\nu}_{NP1}$ and $\widehat{\nu}_{NP2}$ perform similar to the ML/WCV estimator, and much better than the LOOCV estimator. Note that the parameters chosen for λ lead to penalization only if the observed value of $\|s_{Z,T,\nu-d/2}\|_{\tilde{\kappa}_\nu}^2$ is quite unlikely provided that it is χ_n^2 -distributed. Therefore, in most cases, the estimators $\widehat{\nu}_{MLE}$ ($= \widehat{\nu}_{WCV}$), $\widehat{\nu}_{NP1}$ and $\widehat{\nu}_{NP2}$ coincide.

ν_0	$\widehat{\nu}_{MLE}$ MSE	$\widehat{\nu}_{NP1}$ MSE	$\widehat{\nu}_{NP2}$ MSE	$\widehat{\nu}_{WCV}$ MSE	$\widehat{\nu}_{LCV}$ MSE	$\widehat{\nu}_{MLE}$ MSKE	$\widehat{\nu}_{NP1}$ MSKE	$\widehat{\nu}_{NP2}$ MSKE	$\widehat{\nu}_{WCV}$ MSKE	$\widehat{\nu}_{LCV}$ MSKE
0.75	0.0180	0.0198	0.0189	0.0180	0.162	2.602	2.601	2.596	2.602	2.933
1.25	0.0201	0.0207	0.0207	0.0201	0.257	0.5799	0.5814	0.5814	0.5799	0.6690
1.75	0.0217	0.0218	0.0218	0.0217	0.301	0.1136	0.1137	0.1137	0.1136	0.1251
2.25	0.0256	0.0263	0.0263	0.0256	0.331	0.0169	0.0169	0.0169	0.0169	0.0192

Table 2.3: Estimation of ν_0 in case of an unknown variance and a Gaussian random field. Results of a simulation study with 500 simulated data sets with variance $\sigma^2 = 10$ for each ν_0 . The mean squared error (MSE) for the estimation of ν_0 and the mean squared kriging error (MSKE) for the location t_0 are shown.

2.8 An Alternative Approach Using Error Bounds

Let $\Upsilon \subset \mathbb{R}^d$ be an open and bounded Lipschitz domain satisfying the *interior cone condition* (cf. [Wendland, 2005](#)): there exists an angle $\theta \in (0, 2\pi)$ and a radius $r > 0$ such that for every $t \in \Upsilon$ a unit vector $\psi(t)$ exists with

$$\{t + \lambda u : u \in \mathbb{R}^d, \|u\|_2 = 1, u^T \psi(t) \geq \cos \theta, \lambda \in [0, r]\} \subset \Upsilon.$$

Furthermore, for a finite subset $T = \{t_1, \dots, t_n\} \subset \Upsilon$, we define the *fill distance*

$$h_{T,\Upsilon} = \sup_{t \in \Upsilon} \min_{i=1,\dots,n} \|t_i - t\|$$

and the *mesh ratio*

$$\rho_{T,\Upsilon} = \frac{h_{T,\Upsilon}}{q_T}$$

where $q_T = \frac{1}{2} \min_{i,j=1,\dots,n, i \neq j} \|t_i - t_j\|$ is the so-called separation radius. That is, the fill distance denotes the radius of the largest ball in Υ which does not contain any element of T whereas the separation radius is the radius of the smallest ball which contains two elements of T . Thus, if T is a grid in \mathbb{R}^d , the mesh ratio $\rho_{T,\Upsilon}$ equals \sqrt{d} up to some boundary effects. We can assess the convergence rate of $s_{f,T,\tau}$ to f in terms of $h_{T,\Upsilon}$ and $\rho_{T,\Upsilon}$ via the following proposition.

Proposition 2.17. *Let $k > \frac{d}{2}$ be some positive integer, $0 < s \leq 1$ and $f \in W^{k+s,2}(\Upsilon)$. Furthermore, let $\tau > k + s$. Then, there exist $C, h_0 > 0$ such that*

1. $\|f - s_{f,T,\tau-\frac{d}{2}}\|_\infty \leq C \cdot h_{T,\Upsilon}^{k+s-d/2} \cdot \rho_{T,\Upsilon}^{\tau-k-s} \cdot \|f\|_{W^{k+s,2}(\Upsilon)}$
2. $\|f - s_{f,T,\tau-\frac{d}{2}}\|_2 \leq C \cdot h_{T,\Upsilon}^{k+s} \cdot \rho_{T,\Upsilon}^{\tau-k-s} \cdot \|f\|_{W^{k+s,2}(\Upsilon)}$

for all $T \subset \Upsilon$ with $h_{T,\Upsilon} \leq h_0$. Here, $\|\cdot\|_q$ denotes the $L^q(\Upsilon)$ -norm for $q \in \{2, \infty\}$.

Proof. As $\tau > k + s$, we have $u := f - s_{f,T,\tau-d/2} \in W^{k+s,2}(\Upsilon)$ and $u|_T = 0$. Applying Theorem 11.32 from [Wendland \(2005\)](#), we get

$$\|f - s_{f,T,\tau-d/2}\|_q \leq C_1 \cdot h_{T,\Upsilon}^{k+s-d(1/2-1/q)} \cdot \|f - s_{f,T,\tau-d/2}\|_{W^{k+s,2}(\Upsilon)}$$

ν_0	mean		MSE	
	$\hat{\nu}_{L_2}$	$\hat{\nu}_{MLE}$	$\hat{\nu}_{L_2}$	$\hat{\nu}_{MLE}$
1.1	1.17	1.10	0.0134	0.0006
1.3	1.36	1.30	0.0138	0.0007
1.5	1.55	1.50	0.0146	0.0006
1.7	1.76	1.70	0.0167	0.0006
1.9	1.96	1.90	0.0187	0.0007

Table 2.4: Mean and mean squared errors (MSE) for $\hat{\nu}_2$ and $\hat{\nu}_{MLE}$ based on 500 realisations of a Gaussian random field with covariance function $\tilde{\kappa}_{\nu_0}$.

for $q \in \{2, \infty\}$, for suitable $C_1 > 0$ and sufficiently small $h_{T,\Upsilon}$. Furthermore, Theorem 4.2 of [Narcowich et al. \(2006\)](#) yields

$$\|f - s_{f,T,\tau-d/2}\|_{W^{k+s,2}(\Upsilon)} \leq C_2 \cdot \rho_{T,\Upsilon}^{\tau-k-s} \cdot \|f\|_{W^{k+s,2}(\Upsilon)}$$

for some $C_2 > 0$. □

Thus, for sets T consisting of equispaced points — which means that $\rho_{T,\Upsilon}$ does not depend on $h_{T,\Upsilon}$ — we might expect that $\|f - s_{f,T,\tau-d/2}\|$ decays like $h_{T,\Upsilon}^{k+s}$ if $\|\cdot\| = \|\cdot\|_2$ or like $h_{T,\Upsilon}^{k+s-d/2}$ if $\|\cdot\| = \|\cdot\|_\infty$ for $\tau > k+s$, provided that the assessments in Proposition 2.17 are accurate. This motivates to estimate the true smoothness parameter ν_0 of a stationary Gaussian random field by choosing $\tau \gg \nu_0$ and determining the slope of a log-log-regression of $\|f - s_{f,T,\tau-d/2}\|$ on $h_{T,\Upsilon}$. As the error estimates in Proposition 2.17 only hold for small $h_{T,\Upsilon}$, one should only use small values of $h_{T,\Upsilon}$ for regression. However, we note that the $\|\cdot\|_\infty$ -norm of the kriging error can hardly be determined without knowing f exactly at least on a dense set near the boundary of Υ as the largest error is expected to occur in this region. Contrarily, as these boundary effects occur only in a small region, the $\|\cdot\|_2$ -norm of the kriging error might be approximated quite well by summing up and normalising the squared errors on a grid. Thus, in the following, we will restrict ourselves to the $L^2(\Upsilon)$ -norm.

We want to assess the performance of such an estimator by a simulation study based on Gaussian random fields with covariance function $\tilde{\kappa}_{\nu_0}$. Note that the $L^2(\Upsilon)$ -error based estimator can be applied only if $\nu_0 > \lceil d/2 \rceil$. For $\nu_0 = 1.1, 1.3, \dots, 1.9$ we simulate $k = 500$ realisations of a zero mean Gaussian random field with covariance function κ_ν on $Y = \{-0.5, -0.495, \dots, 0.495, 0.5\}$ using the R package `RandomFields` ([Schlather, 2012](#)). We apply simple kriging with $\tau = 2.6, 2.8, 3$ for the sets $T = Y \cap 0.01\mathbb{Z}, Y \cap 0.015\mathbb{Z}, Y \cap 0.02\mathbb{Z}, \dots$ and perform a log-log-regression for the $L^2(\Upsilon)$ -errors on the smallest $2, \dots, 10$ values of $h_{T,\Upsilon}$. Using the mean of these estimates respectively and averaging over the different values for τ we get an estimator denoted by $\hat{\nu}_{L_2}$.

As a reference method we use maximum likelihood estimation based on the set of locations $T = \{-0.5, -0.48, \dots, 0.5\}$ denoted by $\hat{\nu}_{MLE}$. The results — in terms of means and mean squared errors — are shown in Table 2.4. It can be seen that ν_0 seems to be overestimated by $\hat{\nu}_{L_2}$. Furthermore, $\hat{\nu}_{MLE}$ performs much better than the $L^2(\Upsilon)$ -error based estimator although fewer data are used for estimation.

3 Max-Stable Processes Based on Flat Limits of Gaussian Random Fields

In this chapter, we construct max-stable processes based on interpolated Gaussian random fields. We consider flat limits of interpolated fields which we also used to proof Lemma 2.10. More precisely, we deal with flat limits in the case of $n = 2$ or $n = 3$ points. The flat limits are calculated explicitly and employed in a construction of Brown-Resnick type (cf. [Kabluchko et al., 2009](#)).

This chapter is organized as follows: First, we give a short introduction to max-stable processes and the construction principle we consider in this chapter (Section 3.1). In Sections 3.2 and 3.3, we analyse max-stable processes based on flat limits for universal and simple kriging, respectively. Both sections are divided into three subsections, devoted to the flat limits involved, the max-stable processes occurring and considerations on the stationarity of these processes.

3.1 Max-Stable Processes

Having dealt with Gaussian processes in Chapter 2, we now advance to max-stable processes. These are well studied in extreme value theory and have found their way in numerous applications, see [de Haan and Ferreira \(2006\)](#), and [Ribatet \(2011\)](#), for instance. We start by giving the definition of max-stable processes.

Definition 3.1 ([De Haan, 1984](#)). A stochastic process $\{\eta(t), t \in \mathbb{R}^d\}$ is called *max-stable* if there exist functions $a_n : \mathbb{R}^d \rightarrow (0, \infty)$, $b_n : \mathbb{R}^d \rightarrow \mathbb{R}$ such that

$$\left\{ \max_{i=1, \dots, n} \left(\frac{\eta_i(t) - b_n(t)}{a_n(t)} \right), t \in \mathbb{R}^d \right\} \stackrel{d}{=} \{\eta(t), t \in \mathbb{R}^d\},$$

where $\{\eta_i(t), t \in \mathbb{R}^d\}$, $i \in \mathbb{N}$, are independent copies of η .

In [de Haan \(1984\)](#), stochastically continuous max-stable processes have been characterized entirely by a spectral representation. Based on this approach involving Poisson point processes, many models for stationary max-stable processes have been developed. Let us just mention some of these models which will be further analysed within this thesis. [Smith \(1990\)](#) introduced “rainfall-storm” models like the Gaussian and t extreme value processes. For this kind of models — which allow for a mixed moving maxima representation — conditional sampling will be done in Chapter 5. Another model, called extremal Gaussian process, was proposed by [Schlather \(2002\)](#) (cf. Chapter 6).

In this chapter, we consider processes which are constructed similarly to Brown-Resnick processes ([Brown and Resnick, 1977](#); [Kabluchko et al., 2009](#)). As we will see, Brown-Resnick processes arise naturally as stationary max-stable processes based on Gaussian processes. They will be further investigated in Chapter 4 and Chapter 6.

Let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$ and, independently, let $\{\xi_i(t), t \in \mathbb{R}^d\}$, $i \in \mathbb{N}$, be i.i.d. stochastic processes. Then, the process

$$\eta(t) = \max_{i \in \mathbb{N}} (U_i + \xi_i(t)), \quad t \in \mathbb{R}^d,$$

is max-stable and has finite-dimensional distributions

$$\mathbb{P}(\eta(t_1) \leq z_1, \dots, \eta(t_n) \leq z_n) = \exp \left(-\mathbb{E} \left(\max_{i=1, \dots, n} \exp(\xi_1(t_i) - z_i) \right) \right),$$

$t_1, \dots, t_n \in \mathbb{R}^d$, $z_1, \dots, z_n \in \mathbb{R}$, if

$$\mathbb{E}(\exp(\xi_1(t))) < \infty \tag{3.1}$$

for all $t \in \mathbb{R}^d$ (Kabluchko et al., 2009; de Haan, 1984). In particular, $\{\eta(t), t \in \mathbb{R}^d\}$ has Gumbel margins. The following lemma turns out to be useful to proof max-stability in the following sections.

Lemma 3.2. *Let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$. Furthermore, let $(Z_1^{(i)}, \dots, Z_m^{(i)})$, $i \in \mathbb{N}$, be independent copies of a Gaussian random vector (Z_1, \dots, Z_m) and $f_1, \dots, f_m : \mathbb{R}^d \rightarrow \mathbb{R}$ be measurable functions. Then,*

$$\begin{aligned} \zeta(t) &= \max_{i \in \mathbb{N}} \left(U_i + \sum_{j=1}^m f_j(t) Z_j^{(i)} \right), \quad t \in \mathbb{R}^d, \\ \zeta_+(t) &= \max_{i \in \mathbb{N}} \left(U_i + \sum_{j=1}^m |f_j(t)| \cdot |Z_j^{(i)}| \right), \quad t \in \mathbb{R}^d, \\ \zeta_-(t) &= \max_{i \in \mathbb{N}} \left(U_i - \sum_{j=1}^m |f_j(t)| \cdot |Z_j^{(i)}| \right), \quad t \in \mathbb{R}^d, \end{aligned}$$

define max-stable processes with Gumbel margins.

Proof. By the considerations above, it suffices to verify (3.1), that is, to show that

$$\begin{aligned} \mathbb{E} \left(\exp \left(- \sum_{j=1}^m |f_j(t)| \cdot |Z_j| \right) \right) &\leq \mathbb{E} \left(\exp \left(\sum_{j=1}^m f_j(t) Z_j \right) \right) \\ &\leq \mathbb{E} \left(\exp \left(\sum_{j=1}^m |f_j(t)| \cdot |Z_j| \right) \right) < \infty \end{aligned}$$

for all $t \in \mathbb{R}^d$. As the vector $(Z_j)_{j=1}^m$ is Gaussian, there exists a vector $\mu \in \mathbb{R}^m$ and a matrix $B = (b_{ij})_{ij} \in \mathbb{R}^{m \times m}$ such that

$$(Z_j)_{j=1}^m \stackrel{d}{=} \mu + B \cdot (X_j)_{j=1}^m$$

where $X_j \sim_{i.i.d.} \mathcal{N}(0, 1)$, $j = 1, \dots, m$. With $g_j(t) = \sum_{i=1}^m |b_{ij}| \cdot |f_i(t)|$ we can assess

$$\begin{aligned} \mathbb{E} \left(\exp \left(\sum_{j=1}^m |f_j(t)| \cdot |Z_j| \right) \right) &\leq \mathbb{E} \left(\exp \left(\sum_{j=1}^m |f_j(t)| \cdot |\mu_j| + \sum_{j=1}^m |g_j(t)| \cdot |X_j| \right) \right) \\ &= \exp \left(\sum_{j=1}^m |f_j(t)| \cdot |\mu_j| \right) \prod_{j=1}^m \mathbb{E} (\exp (|g_j(t)| \cdot |X_j|)) \\ &= \exp \left(\sum_{j=1}^m |f_j(t)| \cdot |\mu_j| \right) \prod_{j=1}^m \exp \left(\frac{1}{2} |g_j(t)|^2 \right) < \infty \end{aligned}$$

for all $t \in \mathbb{R}^d$. □

3.2 Max-Stable Processes Based on Universal Kriging with Two Data Locations

3.2.1 Flat Limits

Here, we consider a Gaussian random field of the form

$$Z(t) = c \cdot m(t) + \zeta(t), \quad t \in \mathbb{R},$$

where $c \in \mathbb{R}$ is an unknown constant, $m : \mathbb{R} \rightarrow \mathbb{R}$ is a known trend basis function and $\{\zeta(t), t \in \mathbb{R}\}$ is a stationary Gaussian random field with mean zero and variance one.

We aim to interpolate this random field w.r.t. its values at two locations t_1 and t_2 . W.l.o.g. we assume $t_1 = 1$, and $t_2 = -1$. Then, we have

$$(Z(1), Z(-1)) \sim \mathcal{N} \left(c \begin{pmatrix} m(1) \\ m(-1) \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right) \quad (3.2)$$

where $\rho = \text{Cov}(Z(-1), Z(1)) \in [-1, 1]$.

As the mean function $\mathbb{E}(Z(\cdot)) \in \text{span}\{m(\cdot)\}$ is only known up to some constant, interpolation is done by universal kriging which provides a best linear estimator for each $Z(t)$, $t \in \mathbb{R}$, if the true covariance of Z is used ([Chilès and Delfiner, 1999](#)). However, as we are interested in flat limits, we will perform universal kriging based on a scaled covariance function

$$C_\varepsilon(h) := C(\varepsilon h), \quad \varepsilon > 0, h \in \mathbb{R},$$

where $C(\cdot)$ is of the form

$$C(h) = 1 + a|h|^\alpha + o(|h|^\alpha) \quad (3.3)$$

for some $\alpha > 0$, $a \neq 0$.

By

$$\{Z_\varepsilon(t), t \in \mathbb{R}^d\}$$

we denote the random field we get by universal kriging based on the data $Z(1), Z(-1)$, the covariance function $C_\varepsilon(\cdot)$ and the trend function m , i.e.

$$Z_\varepsilon(t) = \lambda_{1,\varepsilon}(t)Z(1) + \lambda_{2,\varepsilon}(t)Z(-1)$$

where $(\lambda_{1,\varepsilon}(t), \lambda_{2,\varepsilon}(t), \mu_\varepsilon(t))$ is a solution of the linear system

$$\begin{pmatrix} C_\varepsilon(0) & C_\varepsilon(2) & m(1) \\ C_\varepsilon(-2) & C_\varepsilon(0) & m(-1) \\ m(1) & m(-1) & 0 \end{pmatrix} \begin{pmatrix} \lambda_{1,\varepsilon}(t) \\ \lambda_{2,\varepsilon}(t) \\ \mu_\varepsilon(t) \end{pmatrix} = \begin{pmatrix} C_\varepsilon(t-1) \\ C_\varepsilon(t+1) \\ m(t) \end{pmatrix} \quad (3.4)$$

(see [Chilès and Delfiner, 1999](#)).

Proposition 3.3. *Under the assumptions given above and for $m(1) = m(-1) = 1$, the random field $\{Z_0(t), t \in \mathbb{R}\}$ defined by*

$$Z_0(t) = \lim_{\varepsilon \searrow 0} Z_\varepsilon(t), \quad t \in \mathbb{R},$$

exists and coincides with the random function G given by

$$G(t) = \frac{1}{2}(Z(1) + Z(-1))m(t) + \frac{1}{2}(Z(1) - Z(-1))\frac{|t+1|^\alpha - |t-1|^\alpha}{2^\alpha}.$$

Proof. Employing (3.3) one can solve the universal kriging system (3.4) explicitly:

$$\underbrace{\begin{pmatrix} 1 & 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) & 1 \\ 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{=: \mathbf{A}_\varepsilon} \begin{pmatrix} \lambda_{1,\varepsilon}(t) \\ \lambda_{2,\varepsilon}(t) \\ \mu_\varepsilon(t) \end{pmatrix} = \begin{pmatrix} 1 + a|t-1|^\alpha \varepsilon^\alpha \\ 1 + a|t+1|^\alpha \varepsilon^\alpha \\ m(t) \end{pmatrix}.$$

The condition $a \neq 0$ ensures that \mathbf{A}_ε is invertible for small ε . For calculating $\lambda_{1,\varepsilon}(t)$ and $\lambda_{2,\varepsilon}(t)$ we need to determine the first and second row of $\mathbf{A}_\varepsilon^{-1}$ which is

$$\mathbf{A}_\varepsilon^{-1} = \frac{1}{\det(\mathbf{A}_\varepsilon)} \begin{pmatrix} -1 & 1 & a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) \\ 1 & -1 & a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) \\ \dots & \dots & \dots \end{pmatrix}$$

where $\det(\mathbf{A}_\varepsilon) = a(2\varepsilon)^\alpha + o(\varepsilon^\alpha)$. Thus, we get

$$\begin{aligned} Z_\varepsilon(t) &= \lambda_{1,\varepsilon}(t)Z(1) + \lambda_{2,\varepsilon}(t)Z(-1) \\ &= \left(-\frac{|t-1|^\alpha}{2^{\alpha+1}} + \frac{|t+1|^\alpha}{2^{\alpha+1}} + \frac{m(t)}{2} + o(1) \right) Z(1) \\ &\quad + \left(\frac{|t-1|^\alpha}{2^{\alpha+1}} - \frac{|t+1|^\alpha}{2^{\alpha+1}} + \frac{m(t)}{2} + o(1) \right) Z(-1). \end{aligned}$$

This yields $\lim_{\varepsilon \searrow 0} Z_\varepsilon(t) = G(t)$. □

3.2.2 Max-Stable Processes

Based on the flat limit G given in Proposition 3.3, we construct max-stable processes as described in Section 3.1.

Theorem 3.4. Let $(Z^{(i)}(1), Z^{(i)}(-1))$, $i \in \mathbb{N}$, be independent copies of

$$(Z(1), Z(-1)) \sim \mathcal{N} \left(c \begin{pmatrix} m(1) \\ m(-1) \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

and $Z_\varepsilon^{(i)}$ the corresponding random field based on universal kriging. Furthermore, independently of $(Z^{(i)}(1), Z^{(i)}(-1))$, $i \in \mathbb{N}$, let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$. Then,

$$Y(t) = \lim_{\varepsilon \searrow 0} \max_{i \in \mathbb{N}} \left(U_i + Z_\varepsilon^{(i)}(t) \right), \quad t \in \mathbb{R},$$

defines a max-stable random field with Gumbel margins on \mathbb{R} .

Proof. First, we rewrite

$$Z_0(t) = \lambda_{1,0}(t)Z(1) + \lambda_{2,0}(t)Z(-1), \quad t \in \mathbb{R}.$$

By the calculations in the proof of Proposition 3.3 there are functions $f_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$, $\varepsilon > 0$, such that

$$|\lambda_{j,\varepsilon}(t) - \lambda_{j,0}(t)| < f_\varepsilon(t) \cdot |\lambda_{j,0}(t)|, \quad j = 1, 2, \quad t \in \mathbb{R}^d,$$

and $f_\varepsilon(t) \rightarrow 0$ as $\varepsilon \searrow 0$ for all $t \in \mathbb{R}$. Thus, for any $t \in \mathbb{R}$, we can assess

$$\begin{aligned} & (1 + f_\varepsilon(t)) \cdot \max_{i \in \mathbb{N}} \left(U_i + Z_0^{(i)}(t) \right) \\ & - f_\varepsilon(t) \cdot \max_{i \in \mathbb{N}} \left(U_i + 2 \cdot |\lambda_{1,0}(t)| \cdot |Z^{(i)}(1)| + 2 \cdot |\lambda_{2,0}(t)| \cdot |Z^{(i)}(-1)| \right) \\ & \leq \max_{i \in \mathbb{N}} \left(U_i + Z_\varepsilon^{(i)}(t) \right) \\ & \leq (1 - f_\varepsilon(t)) \cdot \max_{i \in \mathbb{N}} \left(U_i + Z_0^{(i)}(t) \right) \\ & \quad + f_\varepsilon(t) \cdot \max_{i \in \mathbb{N}} \left(U_i + 2 \cdot |\lambda_{1,0}(t)| \cdot |Z^{(i)}(1)| + 2 \cdot |\lambda_{2,0}(t)| \cdot |Z^{(i)}(-1)| \right). \end{aligned}$$

By Lemma 3.2 all the maxima define max-stable processes. As $\varepsilon \searrow 0$, we get pointwise a.s. convergence to

$$Y(t) = \max_{i \in \mathbb{N}} \left(U_i + Z_0^{(i)}(t) \right)$$

which is

$$Y(t) = \max_{i \in \mathbb{N}} \left(U_i + \frac{Z^{(i)}(1) + Z^{(i)}(-1)}{2} \cdot m(t) + \frac{Z^{(i)}(1) - Z^{(i)}(-1)}{2} \cdot \frac{|t+1|^\alpha - |t-1|^\alpha}{2^\alpha} \right)$$

by Proposition 3.3. Lemma 3.2 yields that $\{Y(t), t \in \mathbb{R}\}$ is a max-stable process with Gumbel margins. \square

Remark 3.5. One can also calculate the scale parameters of the marginal Gumbel distributions and get

$$\mathbb{E}e^{Z_0(t)} = \exp \left(\frac{1}{4}(1 + \rho)m(t)^2 + c \cdot m(t) + \frac{1}{4}(1 - \rho) \left(\frac{|t+1|^\alpha - |t-1|^\alpha}{2^\alpha} \right)^2 \right). \quad (3.5)$$

3.2.3 Stationarity

There are some situations where we end up with a process $Y(\cdot)$ which is even stationary. For example, if $\rho = -1$, $\alpha = 2$, $m(t) = t^2$ and $c = -\frac{1}{2}$ we have

$$Y(t) \stackrel{d}{=} \max_{i \in \mathbb{N}} \left(U_i + Z^{(i)}(1) \cdot t \right) \stackrel{d}{=} \max_{i \in \mathbb{N}} \left(U_i + X_i \cdot t - \frac{1}{2} \cdot t^2 \right)$$

for $X_i \sim_{i.i.d.} \mathcal{N}(0, 1)$. As we will see in Chapter 4, this means that $Y(\cdot)$ is a Brown-Resnick process associated to the variogram $\gamma(h) = h^2$, which is indeed stationary as shown by [Kablichko et al. \(2009\)](#).

However, in general, $Y(t)$ does not even need to have identically distributed marginals which can be seen in Equation (3.5). One can overcome this drawback by marginal transformations

$$Y'(t) = Y(t) - \log \left(\mathbb{E} e^{Z_0(t)} \right).$$

Then, $Y'(\cdot)$ has standard Gumbel marginals.

In some examples this procedure leads to a stationary process. In the case $m(t) = |t|$, $\alpha = 2$, $c = 0$ and an arbitrary $-1 \leq \rho \leq 1$, we have

$$Y(t) = \max_{i \in \mathbb{N}} \left(U_i + |t| \cdot Z^{(i)}(\text{sgn}(t)) \right), \quad t \in \mathbb{R},$$

which is not stationary since we have $\mathbb{E} e^{Z_0(t)} = \exp(\frac{1}{2}t^2)$. But it is “stationarizable” at least on each half axis since $Y'(\cdot)$ coincides with the Brown-Resnick process associated to the variogram $\gamma(h) = h^2$ restricted to this domain.

In order to check whether all these processes are stationarizable, it might be useful to have a look at bivariate characteristics. A very weak form of stationarity — which is at least stronger than identically distributed margins — is the invariance of the *extremal coefficient function* (cf. [Schlather and Tawn, 2003](#))

$$\begin{aligned} \theta(s, t) &= \lim_{x \rightarrow \infty} \frac{\log \mathbb{P}(Y'(s) \leq x, Y'(t) \leq x)}{\log \mathbb{P}(Y'(t) \leq x)} \\ &= \mathbb{E} \max \left\{ \exp(Z_0^{(1)}(s)) \cdot \mathbb{E} \exp(-Z_0^{(1)}(s)), \exp(Z_0^{(1)}(t)) \cdot \mathbb{E} \exp(-Z_0^{(1)}(t)) \right\}, \end{aligned}$$

under translations of $s, t \in \mathbb{R}$.

But even this mild condition turns out not to be always fulfilled. For example, if $m(t) = t^2$, $\alpha = 2$, $c = 0$ and $\rho = 1$, which leads to

$$Y'(t) = \max_{i \in \mathbb{N}} \left(U_i + t^2 \cdot Z^{(i)}(1) - \frac{1}{2}t^4 \right), \quad t \in \mathbb{R},$$

we can compute

$$\begin{aligned} \theta(s, t) &= \int_{-\infty}^{s^2+t^2} e^{(s^2 \wedge t^2)w - \frac{1}{2}(s^2 \wedge t^2)^2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) dw \\ &\quad + \int_{s^2+t^2}^{\infty} e^{(s^2 \vee t^2)w - \frac{1}{2}(s^2 \vee t^2)^2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) dw = 1 + \Phi(s^2 \vee t^2) - \Phi(s^2 \wedge t^2), \end{aligned}$$

which is obviously not invariant under translations.

3.3 Max-Stable Processes Based on Simple Kriging with Two or Three Data Locations

3.3.1 Flat Limits

Here, we assume $\{Z(t), t \in \mathbb{R}\}$ to be a stationary random field with zero mean. Let $T = \{t_1, \dots, t_n\} \subset \mathbb{R}$ be a set of locations with corresponding data $Z(t_i)$, $i = 1, \dots, n$. As the mean is known, interpolation w.r.t. these data is done by the best linear estimator which is the result of simple kriging using the true covariance structure (cf. [Chilès and Delfiner 1999](#); see also Section 3.1). However, following the lines of flat limit theory, we use a scaled covariance function $C_\varepsilon(\cdot) = C(\varepsilon \cdot)$. We denote the result by $Z_\varepsilon(\cdot)$, i.e.

$$Z_\varepsilon(t) = \lambda_{1,\varepsilon}(t)Z(t_1) + \dots + \lambda_{n,\varepsilon}(t)Z(t_n), \quad t \in \mathbb{R},$$

where $\lambda_\varepsilon(\mathbf{t}) = (\lambda_{1,\varepsilon}(t), \dots, \lambda_{n,\varepsilon}(t))^T$ is the solution of the *simple kriging system*

$$\mathbf{A}_\varepsilon \lambda_\varepsilon(\mathbf{t}) = \mathbf{k}_\varepsilon(\mathbf{t}) \quad (3.6)$$

with $\mathbf{A}_\varepsilon = (C_\varepsilon(t_i - t_j))_{i,j=1,\dots,n}$ and $\mathbf{k}_\varepsilon(\mathbf{t}) = (C_\varepsilon(t_1 - t), \dots, C_\varepsilon(t_n - t))^T$.

Proposition 3.6. *Let Z_ε denote the simple kriging result w.r.t. the scaled kernel C_ε .*

If $T = \{-1, 1\}$ and $C(h) = 1 + a|h|^\alpha + o(|h|^\alpha)$ for some $\alpha > 0$, $a \neq 0$, then

$$Z_0(t) = \lim_{\varepsilon \searrow 0} Z_\varepsilon(t), \quad t \in \mathbb{R},$$

exists and coincides with the random function

$$H_1(t) = \frac{1}{2} \left(1 + \frac{|t-1|^\alpha}{2^\alpha} - \frac{|t+1|^\alpha}{2^\alpha} \right) Z(-1) + \frac{1}{2} \left(1 - \frac{|t-1|^\alpha}{2^\alpha} + \frac{|t+1|^\alpha}{2^\alpha} \right) Z(1).$$

If $T = \{-1, 0, 1\}$ and $C(h) = 1 + a|h|^\alpha + \mathcal{O}(|h|^{2\alpha})$ for some $0 < \alpha < 2$, $a \neq 0$, then

$$Z_0(t) = \lim_{\varepsilon \searrow 0} Z_\varepsilon(t), \quad t \in \mathbb{R},$$

exists and coincides with the random function

$$\begin{aligned} H_2(t) = & \left(\frac{2 \cdot (|t-1|^\alpha - |t+1|^\alpha)}{2^\alpha(4-2^\alpha)} + \frac{1 + |t|^\alpha - |t-1|^\alpha}{4-2^\alpha} \right) \cdot Z(-1) \\ & + \frac{2 - 2^\alpha + |t+1|^\alpha - 2|t|^\alpha + |t-1|^\alpha}{4-2^\alpha} \cdot Z(0) \\ & + \left(\frac{2 \cdot (|t+1|^\alpha - |t-1|^\alpha)}{2^\alpha(4-2^\alpha)} + \frac{1 + |t|^\alpha - |t+1|^\alpha}{4-2^\alpha} \right) \cdot Z(1) \end{aligned}$$

Proof. Employing C_ε one can solve the system (3.6) explicitly in both cases.

In the case of two locations $\{-1, 1\}$, we can write (3.6) as

$$\mathbf{A}_\varepsilon \cdot \begin{pmatrix} \lambda_{1,\varepsilon}(t) \\ \lambda_{2,\varepsilon}(t) \end{pmatrix} = \begin{pmatrix} 1 + a(|t+1|\varepsilon)^\alpha + o(\varepsilon^\alpha) \\ 1 + a(|t-1|\varepsilon)^\alpha + o(\varepsilon^\alpha) \end{pmatrix}$$

where

$$\mathbf{A}_\varepsilon = \begin{pmatrix} 1 & 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) \\ 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) & 1 \end{pmatrix}.$$

As $a \neq 0$, \mathbf{A}_ε is invertible for small ε yielding

$$\mathbf{A}_\varepsilon^{-1} = \frac{1}{2a(2\varepsilon)^\alpha + o(\varepsilon^\alpha)} \begin{pmatrix} -1 & 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) \\ 1 + a(2\varepsilon)^\alpha + o(\varepsilon^\alpha) & -1 \end{pmatrix}.$$

Thus, we get

$$\begin{aligned} Z_\varepsilon(t) = \lambda_{1,\varepsilon}(t)Z(-1) + \lambda_{2,\varepsilon}(t)Z(1) &= \left(\frac{1}{2} - \frac{|t+1|^\alpha}{2 \cdot 2^\alpha} + \frac{|t-1|^\alpha}{2 \cdot 2^\alpha} + o(1) \right) Z(-1) \\ &+ \left(\frac{1}{2} + \frac{|t+1|^\alpha}{2 \cdot 2^\alpha} - \frac{|t-1|^\alpha}{2 \cdot 2^\alpha} + o(1) \right) Z(1), \end{aligned}$$

which implies $\lim_{\varepsilon \searrow 0} Z_\varepsilon(t) = H_1(t)$ for all $t \in \mathbb{R}$.

In the case of three locations $\{-1, 0, 1\}$, we rewrite $C(h) = 1 + |h|^\alpha + b|h|^{2\alpha} + o(|h|^{2\alpha})$, $h \in \mathbb{R}$, with some $b \in \mathbb{R}$. This yields the simple kriging system

$$\mathbf{A}_\varepsilon \cdot \begin{pmatrix} \lambda_{1,\varepsilon} \\ \lambda_{2,\varepsilon} \\ \lambda_{3,\varepsilon} \end{pmatrix} = \begin{pmatrix} 1 + a(|t+1|\varepsilon)^\alpha + b(|t+1|\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}) \\ 1 + a(|t|\varepsilon)^\alpha + b(|t|\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}) \\ 1 + a(|t-1|\varepsilon)^\alpha + b(|t-1|\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}) \end{pmatrix}$$

with

$$\mathbf{A}_\varepsilon = \begin{pmatrix} 1 & C(\varepsilon) & C(2\varepsilon) \\ C(\varepsilon) & 1 & C(\varepsilon) \\ C(2\varepsilon) & C(\varepsilon) & 1 \end{pmatrix}.$$

This implies $\det(\mathbf{A}_\varepsilon) = 2^\alpha a^2 (4 - 2^\alpha) \varepsilon^{2\alpha} + o(\varepsilon^{2\alpha})$. As $a \neq 0$, $0 < \alpha < 2$, we get $\det(\mathbf{A}_\varepsilon) \neq 0$ for ε small enough. Inverting \mathbf{A}_ε yields $\mathbf{A}_\varepsilon^{-1} = \frac{1}{\det(\mathbf{A}_\varepsilon)} (\tilde{a}_{ij})_{i,j=1,2,3}$ where

$$\begin{aligned} \tilde{a}_{11} = \tilde{a}_{33} &= -2a\varepsilon^\alpha - (2b + a^2)\varepsilon^{2\alpha} + o(\varepsilon^{2\alpha}), \\ \tilde{a}_{12} = \tilde{a}_{21} &= a(2\varepsilon)^\alpha + a^2 2^\alpha \varepsilon^{2\alpha} + b(2\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}), \\ \tilde{a}_{13} = \tilde{a}_{31} &= 2a\varepsilon^\alpha + (a^2 + 2b)\varepsilon^{2\alpha} - a(2\varepsilon)^\alpha - b(2\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}), \\ \tilde{a}_{22} &= -2a(2\varepsilon)^\alpha - (a^2 + 2b)(2\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}), \\ \text{and } \tilde{a}_{23} = \tilde{a}_{32} &= a(2\varepsilon)^\alpha + a^2 2^\alpha \varepsilon^{2\alpha} + b(2\varepsilon)^{2\alpha} + o(\varepsilon^{2\alpha}). \end{aligned}$$

Thus, we compute

$$\begin{aligned} \lambda_{1,\varepsilon}(t) &= \frac{a^2 \varepsilon^{2\alpha}}{\det(\mathbf{A}_\varepsilon)} (2(|t-1|^\alpha - |t+1|^\alpha) + 2^\alpha(1 + |t|^\alpha - |t-1|^\alpha) + o(1)), \\ \lambda_{2,\varepsilon}(t) &= \frac{a^2 2^\alpha \varepsilon^{2\alpha}}{\det(\mathbf{A}_\varepsilon)} (2 - 2^\alpha + |t+1|^\alpha - 2|t|^\alpha + |t-1|^\alpha + o(1)), \\ \lambda_{3,\varepsilon}(t) &= \frac{a^2 \varepsilon^{2\alpha}}{\det(\mathbf{A}_\varepsilon)} (2(|t+1|^\alpha - |t-1|^\alpha) + 2^\alpha(1 + |t|^\alpha - |t+1|^\alpha) + o(1)). \end{aligned}$$

Hence, $\lim_{\varepsilon \searrow 0} Z_\varepsilon(t) = \lim_{\varepsilon \searrow 0} (\lambda_{1,\varepsilon}(t)Z(-1) + \lambda_{2,\varepsilon}(t)Z(0) + \lambda_{3,\varepsilon}(t)Z(1)) = H_2(t)$ for all $t \in \mathbb{R}$. \square

3.3.2 Max-Stable Processes

One way to construct max-stable processes based on flat limits of random fields interpolated by simple kriging is to imitate the construction in the previous section, which yields similar results. However, here, we follow the lines of neighbourhood kriging (see [Chilès and Delfiner, 1999](#)) which is based on the idea of predicting Z by data located within a moving neighbourhood.

Let $\{Z(n), n \in \mathbb{Z}\}$ be a stationary family of standard Gaussian random variables. For $t \in \mathbb{R}$, let T be the set of the two or three nearest neighbours of t in \mathbb{Z} , respectively, and $Z_\varepsilon(t)$ the corresponding spatially interpolated field based on simple kriging with covariance function C_ε with $C(h) = 1 + a|h|^\alpha + o(|h|^\alpha)$ or $C(h) = 1 + a|h|^\alpha + \mathcal{O}(|h|^{2\alpha})$, respectively.

Then, by Proposition 3.6, the flat limit

$$Z_0(t) = \lim_{\varepsilon \searrow 0} Z_\varepsilon(t), \quad t \in \mathbb{R},$$

exists and consists of pieces of functions of type H_1 or H_2 , respectively. Each of these pieces is defined on the set of points with the same nearest neighbours.

In the case of two nearest neighbours, these sets are the intervals $(z, z + 1)$, $z \in \mathbb{Z}$. As $Z_0|_{(z, z+1)}$ can be extended continuously to $[z, z + 1]$ by the values $Z(z)$ and $Z(z + 1)$, Z_0 is well-defined and continuous.

In the case of three neighbours, we have intervals of type $(z, z + \frac{1}{2})$, $z \in \mathbb{Z}$, and the pieces of Z_0 need not be glued together continuously at half-integer locations. Therefore, we define Z_0 as a càdlàg function, i.e.

$$Z_0\left(z + \frac{1}{2}\right) = \lim_{\delta \searrow 0} Z_0\left(z + \frac{1}{2} + \delta\right), \quad z \in \mathbb{Z}.$$

Furthermore, we note that the distribution of $Z_0(\cdot)$ is invariant under translations by integers by construction, but $Z_0(\cdot)$ is not necessarily stationary.

Theorem 3.7. *Let $\{Z^{(i)}(n), n \in \mathbb{N}\}$, $i \in \mathbb{N}$, be independent copies of $\{Z(n), n \in \mathbb{N}\}$ and $Z_\varepsilon^{(i)}$ the corresponding random field based on simple kriging as described above. Furthermore, independently of $\{Z^{(i)}(n), n \in \mathbb{Z}\}$, $i \in \mathbb{N}$, let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$. Then,*

$$Y(t) = \lim_{\varepsilon \searrow 0} \max_{i \in \mathbb{N}} \left(U_i + Z_\varepsilon^{(i)}(t) \right), \quad t \in \mathbb{R},$$

defines a max-stable random field with Gumbel margins on \mathbb{R} .

Proof. By similar arguments as in the proof of Theorem 3.4, we get

$$Y(t) = \max_{i \in \mathbb{N}} \left(U_i + Z_0^{(i)}(t) \right), \quad t \in \mathbb{R}.$$

By Lemma 3.2, this defines a max-stable process. □

3.3.3 Stationarity

As in the section beforehand, Y is not stationary in general. In the case of simple kriging with a neighbourhood of two points, we get

$$Z_0(t) = \frac{1}{2}(1 + |t - 1|^\alpha - |t|^\alpha)Z(0) + \frac{1}{2}(1 - |t - 1|^\alpha + |t|^\alpha)Z(1), \quad 0 \leq t \leq 1,$$

and

$$\mathbb{E}e^{Z_0(t)} = \exp\left(\frac{1}{4}(1 + \rho) + \frac{1}{4}(1 - \rho)(|t|^\alpha - |t - 1|^\alpha)^2\right).$$

In a general setting, this term is not constant with respect to $t \in \mathbb{R}$, but there is a possibility to stationarize Y .

Therefore, we revisit the notion of *Brown-Resnick stationarity* introduced by [Kabluchko et al. \(2009\)](#). Let $\xi_i(\cdot)$ be independent copies of a stochastic process $\xi(\cdot)$ on \mathbb{R}^d with $\mathbb{E}e^{\xi(\cdot)} < \infty$. Independently, let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$.

Then, ξ is called Brown-Resnick stationary if the max-stable process $\eta(\cdot)$ defined by

$$\eta(t) = \max_{i \in \mathbb{N}} (U_i + \xi_i(t)), \quad t \in \mathbb{R}^d,$$

is stationary.

$\xi(\cdot)$ is Brown-Resnick stationary if and only if the Poisson point process $\sum_{i \in \mathbb{N}} \delta_{U_i + \xi_i(\cdot)}$ on $\mathbb{R}^{\mathbb{R}^d}$ is translation-invariant. Working with this point process we have to endow $\mathbb{R}^{\mathbb{R}^d}$ with a σ -algebra. To this end, we define the cylinder sets

$$C_{t_1, \dots, t_n}(B) = \{f \in \mathbb{R}^{\mathbb{R}^d} : (f(t_1), \dots, f(t_n)) \in B\}$$

for $t_1, \dots, t_n \in \mathbb{R}$, $n \in \mathbb{N}$, and Borel sets $B \subset \mathbb{R}^n$. These sets generate the appropriate σ -algebra $\mathcal{C}(\mathbb{R}^d)$ for stochastic processes. It can be shown that $\mathcal{C}(\mathbb{R}^d)$ — restricted to subsets of $C(\mathbb{R}^d)$ — is the Borel σ -algebra on $C(\mathbb{R}^d)$ w.r.t. the topology of uniform convergence on compact sets. We will work with this σ -algebra in various situations within this thesis. Note that $\mathcal{C}(\mathbb{R}^d)$ is used as σ -algebra on $\mathbb{R}^{\mathbb{R}^d}$ and $C(\mathbb{R}^d)$.

Theorem 3.8. *Let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ and $\{Z_0(t), t \in \mathbb{R}\}$ be as in Theorem 3.7 and, independently, $S_i \sim_{i.i.d.} \text{Unif}([0, 1])$. Then,*

$$Y'(t) = \max_{i \in \mathbb{N}} \left(U_i + Z_0^{(i)}(t - S_i) \right), \quad t \in \mathbb{R},$$

is a stationary and max-stable process.

Proof. Max-stability can be shown analogously to the proof of Theorem 3.7. We show stationarity by showing that the Poisson point process $\Pi = \sum_{i \in \mathbb{N}} \delta_{U_i + Z_0^{(i)}(\cdot - S_i)}$ is translation invariant. The intensity measure Λ of Π is given by

$$\Lambda(C_{t_1, \dots, t_n}(B)) = \int_0^1 \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-s, \dots, t_n-s}(dw_1, \dots, dw_n) du ds,$$

where $\mathbb{P}_{t_1, \dots, t_n}$ denotes the probability measure belonging to $(Z_0(t_1), \dots, Z_0(t_n))$.

Let $t \in \mathbb{R}$. We have to show that $\Lambda(C_{t_1-t, \dots, t_n-t}(B)) = \Lambda(C_{t_1, \dots, t_n}(B))$. First, we note that the mapping $\psi_t : [0, 1] \rightarrow \mathbb{Z}$, $s \mapsto \lfloor t + s \rfloor$ (where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x) has a range consisting of exactly two values, $x_1, x_2 \in \mathbb{Z}$. Let $A_1 = \psi_t^{-1}(x_1)$, $A_2 = \psi_t^{-1}(x_2)$, $B_1 = t + A_1 - x_1$ and $B_2 = t + A_2 - x_2$. By construction, the interior of B_1 and the interior of B_2 are disjoint and we have $\overline{B_1 \cup B_2} = [0, 1]$.

These considerations yield

$$\begin{aligned}
 & \Lambda(C_{t_1-t, \dots, t_n-t}(B)) \\
 &= \int_0^1 \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-(t+s), \dots, t_n-(t+s)}(dw_1, \dots, dw_n) du ds \\
 &= \int_{A_1} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-x_1-(t+s-x_1), \dots, t_n-x_1-(t+s-x_1)}(dw_1, \dots, dw_n) du ds \\
 &\quad + \int_{A_2} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-x_2-(t+s-x_2), \dots, t_n-x_2-(t+s-x_2)}(dw_1, \dots, dw_n) du ds \\
 &= \int_{B_1} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-x_1-v, \dots, t_n-x_1-v}(dw_1, \dots, dw_n) du dv \\
 &\quad + \int_{B_2} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-x_2-v, \dots, t_n-x_2-v}(dw_1, \dots, dw_n) du dv \\
 &= \int_{B_1} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-v, \dots, t_n-v}(dw_1, \dots, dw_n) du dv \\
 &\quad + \int_{B_2} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-v, \dots, t_n-v}(dw_1, \dots, dw_n) du dv \\
 &= \int_0^1 \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{t_1-v, \dots, t_n-v}(dw_1, \dots, dw_n) du dv \\
 &= \Lambda(C_{t_1, \dots, t_n}(B)),
 \end{aligned}$$

where we use the fact that $\mathbb{P}_{t_1, \dots, t_n}$ is invariant under translations by integers. This means that Π is translation-invariant, so $Y'(\cdot)$ is stationary. \square

This way to stationarize processes leads to the following definition.

Definition 3.9. Let $\{\xi_i(t), t \in \mathbb{R}^d\}$ be independent copies of a stochastic process $\{\xi(t), t \in \mathbb{R}^d\}$ with $\mathbb{E} \exp(\xi(t)) < \infty$ for all $t \in \mathbb{R}^d$. Independently, let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $e^{-u} du$. Then, we call $\xi(\cdot)$ *Brown-Resnick stationarizable* if there exists a family $\{S_i, i \in \mathbb{N}\}$ of i.i.d. random variables with distribution F and independent of $\sum_{i \in \mathbb{N}} \delta_{U_i}$ such that the max-stable process

$$v(t) = \max_{i \in \mathbb{N}} (U_i + \xi_i(t - S_i)), \quad t \in \mathbb{R}^d,$$

is stationary. We call F a *stationarizing distribution* of ξ .

Remark 3.10. Obviously, any Brown-Resnick stationary process is Brown-Resnick stationarizable with stationarizing distribution δ_0 , for example.

Using the same techniques as in the proof of Theorem 3.8, we can generalize the result in two steps. First, we consider processes on \mathbb{R} allowing for invariance under translations on a more general grid. Then, we advance to processes on \mathbb{R}^d . Furthermore, we give some results on the class of stationarizing distribution.

Proposition 3.11. *Let $\{\xi(t), t \in \mathbb{R}\}$ satisfy $\mathbb{E} \exp(\xi(t)) < \infty$ for all $t \in \mathbb{R}$. If there exists a $\lambda > 0$ such that the distribution of $\xi(\cdot)$ is invariant under translations by each $\tilde{t} \in \lambda\mathbb{Z}$, then $\xi(\cdot)$ is Brown-Resnick stationarizable.*

Proof. Let $S_i \sim_{i.i.d.} \text{Unif}([0, \lambda])$, $i \in \mathbb{N}$. Then, one can proof that $\sum_{i \in \mathbb{N}} \delta_{U_i + \xi_i(\cdot - S_i)}$ is translation invariant similarly to the proof of Theorem 3.8. We only have to redefine the mapping $\psi_t : [0, \lambda] \rightarrow \lambda\mathbb{Z}$, $s \mapsto \lambda \lfloor \frac{s+t}{\lambda} \rfloor$. \square

Proposition 3.12. *Let $\{\xi(t), t \in \mathbb{R}^d\}$ satisfy $\mathbb{E} \exp(\xi(t)) < \infty$ for all $t \in \mathbb{R}^d$. If there exist $\lambda_1, \lambda_2, \dots, \lambda_d > 0$ such that the distribution of $\xi(\cdot)$ is invariant under translations by each $\tilde{t} \in \lambda_1\mathbb{Z} \times \dots \times \lambda_d\mathbb{Z}$, then $\xi(\cdot)$ is Brown-Resnick stationarizable.*

Proof. Let $S_i = (S_{i,1}, \dots, S_{i,d})$ where $S_{i,j} \sim \text{Unif}([0, \lambda_j])$, $i \in \mathbb{N}$, $j = 1, \dots, d$, are independent. Considering mappings as in the proof of Proposition 3.11 for each component separately, Brown-Resnick stationarizability can be shown in the same way as before. \square

Lemma 3.13. *Let $\{\xi(t), t \in \mathbb{R}^d\}$ be a stochastic process with $\mathbb{E} \exp(\xi(t)) < \infty$ for all $t \in \mathbb{R}^d$. If $\xi(\cdot - S)$ is stationary for some random variable S which is independent of $\xi(\cdot)$, then $\xi(\cdot)$ is Brown-Resnick stationarizable.*

Proof. Let S_i , $i \in \mathbb{N}$, be i.i.d. random variables with the same distribution as S . Then, $\xi_i(\cdot - S_i)$, $i \in \mathbb{N}$, are independent copies of the stationary field $\xi(\cdot - S)$ and therefore $\max_{i \in \mathbb{N}} (U_i + \xi_i(\cdot - S_i))$ is stationary, as well. \square

Proposition 3.14. *Let $\{\xi(t), t \in \mathbb{R}^d\}$ be a stochastic process with $\mathbb{E} \exp(\xi(t)) < \infty$ for all $t \in \mathbb{R}^d$ and $N \in \mathcal{B}^d$ a Lebesgue null set such that $\xi(\cdot)|_{\mathbb{R}^d \setminus N}$ is stationary. Then, $\xi(\cdot)$ is Brown-Resnick stationarizable.*

Proof. Let S be an arbitrary continuous random variable on \mathbb{R}^d and let $t, t_1, \dots, t_m \in \mathbb{R}^d$, $B \subset \mathcal{B}^m$ and $m \in \mathbb{N}$. Then, we have $\mathbb{P}(t_i - t - S \notin N, t_i - S \notin N, i = 1, \dots, m) = 1$ and therefore

$$\begin{aligned} & \mathbb{P}((\xi(t_1 - t - S), \dots, \xi(t_m - t - S)) \in B) \\ &= \mathbb{P}((\xi(t_1 - t - S), \dots, \xi(t_m - t - S)) \in B, t_i - t - S \notin N, t_i - S \notin N, i = 1, \dots, m) \\ &= \mathbb{P}((\xi(t_1 - S), \dots, \xi(t_m - S)) \in B, t_i - t - S \notin N, t_i - S \notin N, i = 1, \dots, m) \\ &= \mathbb{P}((\xi(t_1 - S), \dots, \xi(t_m - S)) \in B) \end{aligned}$$

where we use the stationarity of ξ on $\mathbb{R}^d \setminus N$. Hence, $\xi(\cdot - S)$ is stationary and Lemma 3.13 yields the proposition. \square

Lemma 3.15. *Let $\{\xi(t), t \in \mathbb{R}^d\}$ be a Brown-Resnick stationary process. Then, any probability distribution F on \mathbb{R}^d is a stationarizing distribution for ξ .*

Proof. Let F be an arbitrary probability distribution on \mathbb{R}^d and $S_1, S_2, \dots \sim_{i.i.d.} F$. Furthermore, let Λ be the intensity measure of the Poisson point process $\sum_{i \in \mathbb{N}} \delta_{U_i + \xi_i(\cdot - S_i)}$. Then, for any t_1, t_2, \dots, t_n , $t \in \mathbb{R}^d$, $B \in \mathcal{B}^n$, we have

$$\begin{aligned} \Lambda(C_{t_1-t, \dots, t_n-t}(B)) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{\xi(t_1-(t+s)), \dots, \xi(t_n-(t+s))}(dw_1, \dots, dw_n) du F(ds) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{\xi(t_1), \dots, \xi(t_n)}(dw_1, \dots, dw_n) du F(ds) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}} \int_{B-u\mathbf{1}} e^{-u} \mathbb{P}_{\xi(t_1-s), \dots, \xi(t_n-s)}(dw_1, \dots, dw_n) du F(ds) \\ &= \Lambda(C_{t_1, \dots, t_n}(B)). \end{aligned}$$

Here, we used that $\xi(\cdot)$ is Brown-Resnick stationary. □

Corollary 3.16. *The stationarizing distribution F of a Brown-Resnick stationarizable process $\xi(\cdot)$ is not unique. In particular, the convolution $F \star G$ is a stationarizing distribution, as well, for any probability distribution G .*

Now, let us summarize what we know about the class of stationarizing distributions:

- For “trivial cases” where $\xi(\cdot)$ is (Brown-Resnick) stationary, any distribution is stationarizing. If $\xi(\cdot)$ is stationary at least almost everywhere, any absolutely continuous distribution is stationarizing (cf. Proposition 3.14).
- There exist some non-trivial cases which show that uniform distributions can occur as stationarizing distributions (cf. Theorem 3.8, Propositions 3.11 and 3.12). Corollary 3.16 yields that the class of stationarizing distributions in non-trivial cases contains at least all convolutions of a uniform distribution and any other probability distribution, as well.

4 Unconditional Simulation of Brown-Resnick Processes

In this chapter, which is based on the article of (Oesting et al., 2012) resulting from a diploma thesis (Oesting, 2009), we will focus on a class of max-stable processes that has been introduced by Brown and Resnick (1977) and generalized by Kabluchko et al. (2009). This class is notable, as Brown-Resnick processes also occur as the limit of maxima of independent copies of stationary and appropriately scaled Gaussian random fields (Kabluchko et al., 2009). Thus, Brown-Resnick processes are the natural link of max-stable processes to Gaussian processes as we have already seen in Chapter 3.

Let $\{W(t), t \in \mathbb{R}^d\}$ be a Gaussian process with stationary increments, that is, the law of $\{W(t+h) - W(h), t \in \mathbb{R}^d\}$ does not depend on the choice of $h \in \mathbb{R}^d$. For any second-order process $W(\cdot)$ with stationary increments the *variogram* $\gamma(\cdot)$ (see Chilès and Delfiner, 1999) is defined by

$$\gamma(h) = \mathbb{E}(W(h) - W(0))^2, \quad h \in \mathbb{R}^d.$$

If $W(0) = 0$, then $\sigma^2(t) = \gamma(t)$, where $\sigma^2(t)$ is the variance $\text{Var}(W(t))$. With the following theorem, we review the construction of stationary max-stable processes, which has originally been introduced by Brown and Resnick (1977) for the special case of $W(\cdot)$ being a Brownian motion, i.e. for $\gamma(h) = |h|$.

Theorem 4.1 (Kabluchko et al. 2009). *Let $\{W_i(t), t \in \mathbb{R}^d\}$, $i \in \mathbb{N}$, be independent copies of a Gaussian random field $\{W(t), t \in \mathbb{R}^d\}$ with zero mean and variance $\sigma^2(\cdot)$. Independently, let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on \mathbb{R} with intensity measure $\exp(-u) du$.*

1. *The process $\{Z(t), t \in \mathbb{R}^d\}$, defined by*

$$Z(t) = \max_{i \in \mathbb{N}} \left(U_i + W_i(t) - \frac{\sigma^2(t)}{2} \right),$$

is a max-stable process with standard Gumbel margins.

2. *If, additionally, $W(\cdot)$ has stationary increments, then the process $Z(\cdot)$ is stationary and its law only depends on the variogram $\gamma(\cdot)$ of $W(\cdot)$. The process $Z(\cdot)$ is called Brown-Resnick process associated to the variogram $\gamma(\cdot)$.*
3. *Moreover, under the assumptions of the second part of this theorem,*

$$\sum_{i \in \mathbb{N}} \delta_{U_i + W_i(\cdot) - \sigma^2(\cdot)/2}$$

is a translation invariant Poisson point process on $\mathbb{R}^{\mathbb{R}^d}$.

Brown-Resnick processes have been further analysed over the last years. For example, [Kabluchko \(2009a\)](#) showed that they also occur, in a modified form, as the limit of empirical distribution functions. The decomposition of the processes into conservative and dissipative components is discussed in [Kabluchko \(2009b\)](#) and [Wang and Stoev \(2010\)](#); ergodicity and mixing properties are investigated by [Kabluchko and Schlather \(2010\)](#). Recently, [Dombry et al. \(2011\)](#) presented an algorithm to perform conditional sampling. We will also deal with this issue in Chapter 6. Furthermore, Brown-Resnick processes have found their way to applications. [Buishand et al. \(2008\)](#) use them for modelling spatial rainfall.

Here, we aim at developing algorithms which allow for efficient simulation of Brown-Resnick processes as it turns out that finite approximations based on the definition indicate the presence of a drift although Brown-Resnick processes are stationary. Therefore, we present alternative representations, e.g. by random shifts, which we introduce in Section 4.1. Section 4.2 deals with Brown-Resnick processes which are generated by a dissipative flow and presents further representations for them. All these different representations lead to different kinds of finite approximations introduced in Section 4.3. Error estimates for these approximations are given in Section 4.4 for the original process of [Brown and Resnick \(1977\)](#). In Section 4.5, we analyse the results of a simulation study and compare the quality of different simulation techniques resulting from the finite approximations.

We restrict ourselves to max-stable processes with Gumbel margins as these naturally occur when considering maxima of Gaussian processes. Fréchet and Weibull margins can be obtained by marginal transformation.

4.1 Random Shifts

Figure 4.1 shows that a finite approximation of the Brown-Resnick process based on the definition (cf. Subsection 4.3) turns out to appear non-stationary on large intervals if the equation

$$\lim_{\|t\| \rightarrow \infty} \left(W(t) - \frac{\sigma^2(t)}{2} \right) = -\infty \quad \mathbb{P} - a.s. \quad (4.1)$$

holds. Therefore, we seek equivalent representations of Brown-Resnick processes that avoid this drawback. A first possibility is to include some “random shifting” into the construction.

Theorem 4.2. *Let $W_i(\cdot)$, $i \in \mathbb{N}$, be independent copies of a Gaussian random field $\{W(t), t \in \mathbb{R}^d\}$ with zero mean, stationary increments and variance $\sigma^2(\cdot)$ and let Q be a probability measure on \mathbb{R}^d . Independently of $W_i(\cdot)$, let $\sum_{i \in \mathbb{N}} \delta_{(U_i, S_i)}$ be a Poisson point process on $\mathbb{R} \times \mathbb{R}^d$ with intensity measure $\exp(-u) du \times Q(ds)$. Then,*

$$\tilde{Z}(t) = \max_{i \in \mathbb{N}} \left(U_i + W_i(t - S_i) - \frac{\sigma^2(t - S_i)}{2} \right), \quad t \in \mathbb{R}^d,$$

is a Brown-Resnick process associated to the variogram $\gamma(\cdot)$, i.e. $\tilde{Z} \stackrel{d}{=} Z$ with Z as in Theorem 4.1.

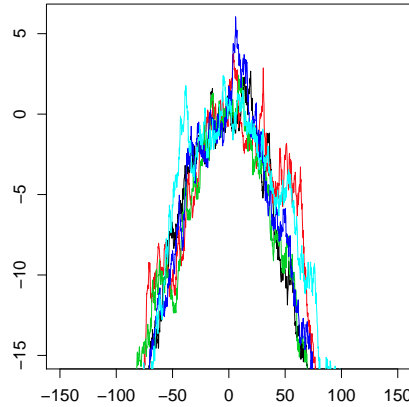


Figure 4.1: Five finite approximations of the original Brown-Resnick process, each based on the largest 1000 values of the underlying Poisson point process.

Proof. Let $t_1, \dots, t_m \in \mathbb{R}^d$, $y_1, \dots, y_m \in \mathbb{R}$ and $m \in \mathbb{N}$, be arbitrary and $\mathbb{P}_{t_1, \dots, t_m}$ be the law of the random vector $(W(t_1), \dots, W(t_m))$.

Then, $\Pi = \sum_{i \in \mathbb{N}} \delta_{(U_i, S_i, W_i)}$ is a Poisson point process on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{\mathbb{R}^d}$ with intensity measure $\exp(-u) du \times Q(ds) \times \mathbb{P}_W(dw)$ and

$$\begin{aligned} & \tilde{Z}(t_1) \leq y_1, \dots, \tilde{Z}(t_m) \leq y_m \\ \Leftrightarrow & \Pi \left(\left\{ (u, s, w) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{\mathbb{R}^d} : u > \min_{k=1, \dots, m} \left(y_k - w(t_k - s) + \frac{\sigma^2(t_k - s)}{2} \right) \right\} \right) = 0. \end{aligned}$$

Thus, we obtain

$$\begin{aligned} & -\log(\mathbb{P}(\tilde{Z}(t_1) \leq y_1, \dots, \tilde{Z}(t_m) \leq y_m)) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^m} \int_{\min_{k=1, \dots, m} \left(y_k - w_k + \frac{\sigma^2(t_k - s)}{2} \right)}^{\infty} \exp(-u) du \mathbb{P}_{t_1 - s, \dots, t_m - s}(dw_1, \dots, dw_m) Q(ds) \\ &= \int_{\mathbb{R}^d} -\log(\mathbb{P}(Z(t_1 - s) \leq y_1, \dots, Z(t_m - s) \leq y_m)) Q(ds). \end{aligned}$$

Due to the stationarity of $Z(\cdot)$ the right hand side equals

$$\begin{aligned} & \int_{\mathbb{R}^d} -\log(\mathbb{P}(Z(t_1 - s) \leq y_1, \dots, Z(t_m - s) \leq y_m)) Q(ds) \\ &= \int_{\mathbb{R}^d} -\log(\mathbb{P}(Z(t_1) \leq y_1, \dots, Z(t_m) \leq y_m)) Q(ds) \\ &= -\log(\mathbb{P}(Z(t_1) \leq y_1, \dots, Z(t_m) \leq y_m)) \end{aligned}$$

since Q is a probability measure. □

This theorem can be used for representing Brown-Resnick processes in many different ways. Here, we give two corollaries as applications.

Corollary 4.3. Let $W(\cdot)$ be as in Theorem 4.1, $W_i^{(j)} \sim_{i.i.d.} W$, $i \in \mathbb{N}$, $j = 1, \dots, n$. Independently of $W_i^{(j)}$, $i \in \mathbb{N}$, let $\Pi^{(j)} = \sum \delta_{U_i^{(j)}}$, $j = 1, \dots, n$, be independent Poisson point processes on \mathbb{R} with intensity measure $n^{-1} \exp(-u) du$ and $s_1, \dots, s_n \in \mathbb{R}^d$. Then,

$$Z_1(t) = \max_{j=1, \dots, n} \max_{i \in \mathbb{N}} \left(U_i^{(j)} + W_i^{(j)}(t - s_j) - \frac{\sigma^2(t - s_j)}{2} \right), \quad t \in \mathbb{R}^d,$$

is a Brown-Resnick process associated to the variogram $\gamma(\cdot)$, i.e. $Z_1 \stackrel{d}{=} Z$.

Proof. Note that the superposition $\sum_{j=1}^n \sum_{i \in \mathbb{N}} \delta_{(U_i^{(j)}, s_j)}$ is a Poisson point process on $\mathbb{R} \times \mathbb{R}^d$ with intensity measure $\exp(-u) du \times (\frac{1}{n} \sum_{j=1}^n \delta_{s_j})$, and apply Theorem 4.2. \square

Corollary 4.4. Let $W_i(\cdot)$ be as in Theorem 4.1, and $I \subset \mathbb{R}^d$ a finite cuboid. Independently of W_i let $\Pi = \sum \delta_{(U_i, S_i)}$ be a Poisson point process on $\mathbb{R} \times I$ with intensity measure $\exp(-u) du \times |I|^{-1} ds$. Then, $Z \stackrel{d}{=} Z_2$, where

$$Z_2(t) = \max_{i \in \mathbb{N}} \left(U_i + W_i(t - S_i) - \frac{\sigma^2(t - S_i)}{2} \right), \quad t \in \mathbb{R}^d.$$

Proof. With $Q(ds) = |I|^{-1} \mathbf{1}_{s \in I} ds$ the assertion follows from Theorem 4.2. \square

4.2 Mixed Moving Maxima Representation

The notion of max-stable processes generated by non-singular flows has been introduced by [de Haan and Pickands \(1986\)](#); further results on the representations of max-stable processes have been obtained in [Kabluchko \(2009b\)](#) and [Wang and Stoev \(2010\)](#) by transferring some work of [Rosiński \(1995\)](#) on $S\alpha S$ -processes.

[Kabluchko et al. \(2009, Theorem 14\)](#) showed that a Brown-Resnick process is generated by a dissipative flow if (4.1) holds. In the case $d = 1$, condition (4.1) is satisfied if $\liminf_{t \rightarrow \infty} \gamma(t) / \log t > 8$.

Using the stationarity criterion from the third part of Theorem 4.1, we provide equivalent representations of Brown-Resnick processes given on the following theorems.

Theorem 4.5. Let $W_i^{(j)}$, $i \in \mathbb{N}$, $j \in \mathbb{Z}^d$, be independent copies of a Gaussian random field $W(\cdot)$ with continuous sample paths, stationary increments, zero mean, variance $\sigma^2(\cdot)$ and variogram $\gamma(\cdot)$ on \mathbb{R}^d . Furthermore, let

$$T_i^{(j)} = \inf \left(\operatorname{argsup}_{t \in \mathbb{R}^d} \left(W_i^{(j)}(t) - \frac{\sigma^2(t)}{2} \right) \right)$$

where the “inf” is understood in the lexicographic sense if $d > 1$. We assume (4.1), so that $T_i^{(j)}$ is well-defined a.s.

Independently of $W_i^{(j)}$, let $\Pi^{(j)} = \sum_{i \in \mathbb{N}} \delta_{U_i^{(j)}}$, $j \in \mathbb{Z}^d$, be independent Poisson point processes on \mathbb{R} with intensity measure $m^{-d} \exp(-u) du$ for some $m \in \mathbb{N}$. Furthermore, let $p > 0$. Then,

$$Z_3(t) = \max_{j \in \mathbb{Z}^d} \max_{\substack{i \in \mathbb{N} \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d}} \left(U_i^{(j)} + W_i^{(j)}(t - pj) - \frac{\sigma^2(t - pj)}{2} \right), \quad t \in \mathbb{R}^d,$$

is a Brown-Resnick process associated to the variogram $\gamma(\cdot)$, i.e. $Z_3 \stackrel{d}{=} Z$.

Proof. We endow $C(\mathbb{R}^d)$ with the Borel- σ -algebra w.r.t. uniform convergence on compact sets $\mathcal{C}(\mathbb{R}^d)$ (see Section 3.3). Furthermore, we define $\xi_i^{(j)}(t) = W_i^{(j)}(t) - \sigma^2(t)/2$. Because of condition (4.1), each $T_i^{(j)}$ is finite \mathbb{P} -a.s. and $M_i^{(j)} = \sup_{t \in \mathbb{R}^d} (U_i^{(j)} + \xi_i^{(j)}(t))$ is well-defined. The mapping

$$\Theta : C(\mathbb{R}^d) \rightarrow \mathbb{R}^d \times C(\mathbb{R}^d), \quad U_i^{(j)} + \xi_i^{(j)}(\cdot) \mapsto (T_i^{(j)}, U_i^{(j)} + \xi_i^{(j)}(\cdot)),$$

is measurable since $\sup_{t \in \mathbb{R}^d} \xi_i^{(j)}(t) = \sup_{t \in \mathbb{Q}^d} \xi_i^{(j)}(t)$ and $T_i^{(j)}$ is the first root of $\xi_i^{(j)} - \sup(\xi_i^{(j)})$. Therefore, the mapping theorem for Poisson point processes (Kingman, 1993) yields that $\sum_{i \in \mathbb{N}} \delta_{(T_i^{(j)}, U_i^{(j)} + \xi_i^{(j)}(\cdot))}$ is a Poisson point process with intensity measure

$$\Psi(A) = \int_{\mathbb{R}} \frac{1}{m^d} \exp(-u) \mathbb{P}_W(u + \xi \in \Theta^{-1}(A)) du, \quad A \in \mathcal{B}^d \times \mathcal{C},$$

where \mathbb{P}_W is the law of the process $W(\cdot)$.

Now we define $\mathcal{U}_t : C(\mathbb{R}^d) \rightarrow C(\mathbb{R}^d)$, $f(\cdot) \mapsto f(\cdot - t)$ and $\mathcal{V}_t : \mathbb{R}^d \times C(\mathbb{R}^d) \rightarrow \mathbb{R}^d \times C(\mathbb{R}^d)$, $(s, f(\cdot)) \mapsto (s + t, f(\cdot - t))$ as translations by $t \in \mathbb{R}^d$. Then we obtain

$$(\Theta \circ \mathcal{U}_t)(U_i^{(j)} + \xi_i^{(j)}(\cdot)) = (T_i^{(j)} + t, U_i^{(j)} + \xi_i^{(j)}(\cdot - t)) = (\mathcal{V}_t \circ \Theta)(U_i^{(j)} + \xi_i^{(j)}(\cdot)).$$

The intensity measure of the Poisson point process $\sum \delta_{U_i^{(j)} + \xi_i^{(j)}(\cdot)}$ is translation invariant (with respect to \mathcal{U}_t) by Theorem 4.1. Because of the fact that Θ commutes with the translation operators, Ψ is translation invariant (with respect to \mathcal{V}_t), as well.

Thus, for any $j \in \mathbb{Z}^d$, we obtain

$$\max_{\substack{i \in \mathbb{N} \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d}} \left(U_i^{(j)} + \xi_i^{(j)}(\cdot - pj) \right) \stackrel{d}{=} \max_{T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj} \left(U_i^{(j)} + \xi_i^{(j)}(\cdot) \right). \quad (4.2)$$

Now we consider each side of (4.2) separately. For different $j \in \mathbb{Z}^d$ we get stochastically independent processes. This yields

$$\begin{aligned} Z_3(\cdot) &= \max_{j \in \mathbb{Z}^d} \max_{\substack{i \in \mathbb{N} \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d}} \left(U_i^{(j)} + \xi_i^{(j)}(\cdot - pj) \right) \\ &\stackrel{d}{=} \max_{j \in \mathbb{Z}^d} \max_{\substack{i \in \mathbb{N} \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj}} \left(U_i^{(j)} + \xi_i^{(j)}(\cdot) \right). \end{aligned}$$

Furthermore, by replacing $T_i^{(j)}$, $\xi_i^{(j)}$, and $U_i^{(j)}$ by $T_i^{(j \bmod m)}$, $\xi_i^{(j \bmod m)}$, and $U_i^{(j \bmod m)}$, respectively, we obtain

$$\max_{\substack{i \in \mathbb{N} \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj}} \left(U_i^{(j)} + \xi_i^{(j)}(\cdot) \right) \stackrel{d}{=} \max_{\substack{i \in \mathbb{N} \\ T_i^{(j \bmod m)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj}} \left(U_i^{(j \bmod m)} + \xi_i^{(j \bmod m)}(\cdot) \right)$$

where ‘‘mod’’ is understood as a componentwise operation.

For $j_1 \equiv j_2 \pmod{m}$, $j_1 \neq j_2$ we have

$$\left((-mp/2, mp/2]^d + pj_1\right) \cap \left((-mp/2, mp/2]^d + pj_2\right) = \emptyset,$$

which guarantees that the processes $\xi_i^{(j_1 \bmod m)}$ with $T_i^{(j_1 \bmod m)} \in (-mp/2, mp/2]^d + pj_1$ and $\xi_i^{(j_2 \bmod m)}$ with $T_i^{(j_2 \bmod m)} \in (-mp/2, mp/2]^d + pj_2$ are independent. By these considerations we get

$$\begin{aligned} Z_3(\cdot) &\stackrel{d}{=} \max_{j \in \mathbb{Z}^d} \max_{\substack{i \in \mathbb{N} \\ T_i^{(j \bmod m)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj}} \left(U_i^{(j \bmod m)} + \xi_i^{(j \bmod m)}(\cdot) \right) \\ &\stackrel{d}{=} \max_{k \in \{0, \dots, m-1\}^d} \max_{j \in \mathbb{Z}^d} \max_{\substack{i \in \mathbb{N} \\ j \bmod m \equiv k \\ T_i^{(k)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d + pj}} \left(U_i^{(k)} + \xi_i^{(k)}(\cdot) \right) \\ &= \max_{k \in \{0, \dots, m-1\}^d} \max_{i \in \mathbb{N}} \left(U_i^{(k)} + \xi_i^{(k)}(\cdot) \right) \stackrel{d}{=} Z(\cdot). \end{aligned}$$

The last step is based on the fact that $\sum_{k \in \{0, \dots, m-1\}^d} \sum_{i \in \mathbb{N}} \delta_{U_i^{(k)}}$ is a Poisson point process with intensity measure $\sum_{k \in \{0, \dots, m-1\}^d} m^{-d} \exp(-u) du = \exp(-u) du$. \square

By [Kabluchko \(2009b\)](#), condition (4.1) holds only if $Z(\cdot)$ has a mixed moving maxima representation, i.e. $Z(\cdot)$ is of the form

$$Z(\cdot) = \max_{i \in \mathbb{N}} \left(\tilde{U}_i + \tilde{F}_i(\cdot - S_i) \right), \quad t \in \mathbb{R}^d.$$

where (S_i, \tilde{U}_i) , $i \in \mathbb{N}$, are the points of a Poisson point process on $\mathbb{R}^d \times \mathbb{R}$ with intensity measure $ds \times e^{-u} du$ and \tilde{F}_i , $i \in \mathbb{N}$, are i.i.d. random measurable functions such that $\mathbb{E}(\int_{\mathbb{R}^d} e^{\tilde{F}_1(t)} dt) < \infty$. In order to construct such a representation, we repeat results from the proof of Theorem 14 in [Kabluchko et al. \(2009\)](#).

Theorem 4.6. *Let $\{W_i(t), t \in \mathbb{R}^d\}$, $i \in \mathbb{N}$, be independent copies of a Gaussian random field $\{W(t), t \in \mathbb{R}^d\}$ with continuous sample paths, stationary increments, zero mean, variance $\sigma^2(\cdot)$ and variogram $\gamma(\cdot)$ on \mathbb{R}^d . We assume that condition (4.1) is satisfied.*

Furthermore, let $T_i = \inf \left(\operatorname{argsup}_{t \in \mathbb{R}^d} \left(W_i(t) - \frac{\sigma^2(t)}{2} \right) \right)$, $M_i = \sup_{t \in \mathbb{R}^d} \left(W_i(t) - \frac{\sigma^2(t)}{2} \right)$ and $F_i(\cdot) = W_i(\cdot + T_i) - \frac{\sigma^2(\cdot + T_i)}{2} - M_i$.

Independently of W_i , let $\sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process with intensity measure $\exp(-u) du$. Then, the random measure $\sum_{i \in \mathbb{N}} \delta_{(T_i, U_i + M_i, F_i)}$ defines a Poisson point process on $\mathbb{R}^d \times \mathbb{R} \times C(\mathbb{R}^d)$ with intensity measure $\lambda^ dt \times e^{-y} dy \times \tilde{Q}(dF)$ for some $\lambda^* > 0$ and a probability measure \tilde{Q} on $C(\mathbb{R}^d)$.*

Furthermore, let $\sum_{i \in \mathbb{N}} \delta_{(S_i, \tilde{U}_i)}$ be a Poisson point process on $\mathbb{R}^d \times \mathbb{R}$ with intensity measure $\lambda^ ds \times e^{-\tilde{u}} d\tilde{u}$ and $\tilde{F}_i \sim_{i.i.d.} \tilde{Q}$. Then, we have $Z_4 \stackrel{d}{=} Z$ for*

$$Z_4(t) = \max_{i \in \mathbb{N}} \left(\tilde{U}_i + \tilde{F}_i(t - S_i) \right), \quad t \in \mathbb{R}^d.$$

Proof. The first part is shown in the proof of Theorem 14 in [Kabluchko et al. \(2009\)](#).

For the second part note that $\sum_{i \in \mathbb{N}} \delta_{(T_i, U_i + M_i, F_i)}$ and $\sum_{i \in \mathbb{N}} \delta_{(S_i, \tilde{U}_i, \tilde{F}_i)}$ are Poisson point processes on $\mathbb{R}^d \times \mathbb{R} \times C(\mathbb{R}^d)$ with the same intensity measure. Furthermore, we have $Z(\cdot) = \max_{i \in \mathbb{N}} \Gamma(T_i, U_i + M_i, F_i)$ and $Z_4(\cdot) = \max_{i \in \mathbb{N}} \Gamma(S_i, U_i, \tilde{F}_i)$ with the transformation $\Gamma : \mathbb{R}^d \times \mathbb{R} \times C(\mathbb{R}^d) \rightarrow C(\mathbb{R}^d)$, $(s, y, f) \mapsto y + f(\cdot - s)$. \square

Remark 4.7. A similar result holds if we consider all the processes from Theorem 4.6 restricted to $p\mathbb{Z}^d$, $p > 0$, instead of \mathbb{R}^d . Then, for

$$T_i^{(p)} = \inf \left(\operatorname{argsup}_{t \in p\mathbb{Z}^d} \left(W_i(t) - \frac{\sigma^2(t)}{2} \right) \right), \quad M_i^{(p)} = \sup_{t \in p\mathbb{Z}^d} \left(W_i(t) - \frac{\sigma^2(t)}{2} \right),$$

$$\text{and } F_i^{(p)}(\cdot) = W_i(\cdot + T_i^{(p)}) - \frac{\sigma^2(\cdot + T_i^{(p)})}{2} - M_i^{(p)}, \quad t \in p\mathbb{Z}^d,$$

the random measure $\sum_{i \in \mathbb{N}} \delta_{(T_i^{(p)}, U_i + M_i^{(p)}, F_i^{(p)})}$ defines a Poisson point process on $p\mathbb{Z}^d \times \mathbb{R} \times \mathbb{R}^{p\mathbb{Z}^d}$ with intensity measure $\lambda^{(p)} p^d \delta_t \times e^{-y} dy \times \tilde{Q}^{(p)}(dF)$ for some $\lambda^{(p)} > 0$ and some probability measure $\tilde{Q}^{(p)}$ on $\mathbb{R}^{p\mathbb{Z}^d}$. An equivalent representation $Z_4^{(p)}$ of $Z|_{p\mathbb{Z}^d}$ can be given analogously to Theorem 4.6. Even more easily, all the other results from this chapter up to here can be transferred to processes on a lattice.

For approximating Z via the representation Z_4 , the law \tilde{Q} is needed explicitly. Note that, in general, \tilde{Q} is not the law of $W(\cdot + T) - \sigma^2(\cdot + T) - M$ (and $\tilde{Q}^{(p)}$ is not the law of $W(\cdot + T^{(p)}) - \sigma^2(\cdot + T^{(p)}) - M^{(p)}$). If we assume $W(0) = 0$ — which can be done w.l.o.g. by replacing $W(\cdot)$ by $W(\cdot) - W(0)$ — and restrict ourselves to processes on a lattice $p\mathbb{Z}^d$, we get the following result.

Theorem 4.8. *Let $p > 0$ and $\{W(t), t \in p\mathbb{Z}^d\}$ be as in Theorem 4.6 and*

$$T^{(p)} = \inf \left(\operatorname{argsup}_{t \in p\mathbb{Z}^d} \left(W(t) - \frac{\sigma^2(t)}{2} \right) \right).$$

Furthermore, assume $W(0) = 0$. Then, $Q^{(p)}$ is the law of $W(\cdot) - \frac{\sigma^2(\cdot)}{2} \mid T^{(p)} = 0$.

Proof. Let $A \in \mathcal{B}(\mathbb{R}^{p\mathbb{Z}^d})$ and $V \in \mathcal{B}$ such that $0 < \int_V e^{-u} du < \infty$. Furthermore, let $\Pi = \sum_{i \in \mathbb{N}} \delta_{(T_i^{(p)}, U_i + M_i^{(p)}, F_i^{(p)})}$ be the Poisson point process on $p\mathbb{Z}^d \times \mathbb{R} \times \mathbb{R}^{p\mathbb{Z}^d}$ with the notation from Remark 4.7. As the intensity measure of Π is a product measure, we have

$$\tilde{Q}^{(p)}(A) = \mathbb{P}(\Pi(\{0\} \times V \times A) = 1 \mid \Pi(\{0\} \times V \times \mathbb{R}^{p\mathbb{Z}^d}) = 1), \quad (4.3)$$

Furthermore, we may assume that the points $(T_i^{(p)}, U_i + M_i^{(p)}, F_i^{(p)})$ are numbered such that the sequence $(U_i)_{i \in \mathbb{N}}$ is decreasing (cf. Section 4.3). Then, we get

$$\begin{aligned} \mathbb{P}(\Pi(\{0\} \times V \times A) = 1 \mid \Pi(\{0\} \times V \times \mathbb{R}^{p\mathbb{Z}^d}) = 1) & \quad (4.4) \\ = \sum_{i \in \mathbb{N}} \mathbb{P}(T_i^{(p)} = 0, U_i + M_i^{(p)} \in V \mid \#\{i : (T_i^{(p)}, U_i + M_i^{(p)}) \in \{0\} \times V\} = 1) \\ \cdot \mathbb{P}(F_i^{(p)} \in A \mid T_i^{(p)} = 0, U_i + M_i^{(p)} \in V, \#\{i : (T_i^{(p)}, U_i + M_i^{(p)}) \in \{0\} \times V\} = 1) \end{aligned}$$

with

$$\begin{aligned}
 & \mathbb{P}(F_i^{(p)} \in A \mid T_i^{(p)} = 0, U_i + M_i^{(p)} \in V, \#\{i : (T_i^{(p)}, U_i + M_i^{(p)}) \in \{0\} \times V\} = 1) \\
 &= \mathbb{P}(F_i^{(p)} \in A \mid T_i^{(p)} = 0, U_i + M_i^{(p)} \in V, (T_j^{(p)}, U_j + M_j^{(p)}) \notin \{0\} \times V \forall j \neq i) \\
 &= \mathbb{P}(F_i^{(p)} \in A \mid T_i^{(p)} = 0, U_i \in V, (T_j^{(p)}, U_j + M_j^{(p)}) \notin \{0\} \times V \forall j \neq i) \\
 &= \mathbb{P}(F_i^{(p)} \in A \mid T_i^{(p)} = 0),
 \end{aligned}$$

where we use the fact that W_i is independent of U_i, U_j and W_j for all $j \neq i$. Employing (4.3), (4.4), and

$$\sum_{i \in \mathbb{N}} \mathbb{P}\left(T_i^{(p)} = 0, U_i + M_i^{(p)} \in V \mid \Pi(\{0\} \times V \times \mathbb{R}^{p\mathbb{Z}^d}) = 1\right) = 1,$$

we get

$$\tilde{Q}^{(p)}(A) = \mathbb{P}(F_i^{(p)} \in A \mid T_i = 0) = \mathbb{P}\left(W(\cdot) - \frac{\sigma^2(\cdot)}{2} \in A \mid T^{(p)} = 0\right)$$

for all $A \in \mathcal{B}(\mathbb{R}^{p\mathbb{Z}^d})$. □

Remark 4.9. Let Π be defined as in the proof of Theorem 4.8. Considering the intensity $\lambda^{(p)} p^d$ of the restriction of Π on the set $\{0\} \times [0, \infty) \times C(\mathbb{R}^d)$ we get the equality $\lambda^{(p)} p^d = \mathbb{P}(T^{(p)} = 0)$.

Using only the assumptions of Theorem 4.6, \tilde{Q} can be described as the law of F_i conditional on $U_i + M_i$ and T_i . Let $\Pi = \sum_{i \in \mathbb{N}} \delta_{(T_i, U_i + M_i, F_i)}$ and $E \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$ such that $\int_E e^{-u} (dt \times du) \in (0, \infty)$. Furthermore, let $N = \Pi(E \times \mathcal{C}(\mathbb{R}^d))$ and $i_1 < \dots < i_N$ such that $(T_{i_k}, U_{i_k} + M_{i_k}) \in E$ for $k = 1, \dots, N$. By G_1, \dots, G_N we denote a random permutation of F_{i_1}, \dots, F_{i_N} .

Theorem 4.10. *Conditional on $N = n$, the processes G_1, \dots, G_n are i.i.d. with law \tilde{Q} .*

Proof. We have to proof that all finite dimensional margins of G_1, \dots, G_n are products of one dimensional margins with law \tilde{Q} . By decomposing the sets of $\mathcal{C}(\mathbb{R}^d)$ and changing numbering of indices, it suffices to proof that $\mathbb{P}(G_1 \in A_1, \dots, G_{n_1} \in A_1, G_{n_1+1} \in A_2, \dots, G_{n_1+n_2+\dots+n_l} \in A_l \mid N = n)$ equals $\prod_{i=1}^l \tilde{Q}(A_i)^{n_i}$ for pairwise disjoint sets $A_1, \dots, A_l \in \mathcal{C}$, $n_1, \dots, n_l \in \mathbb{N}$ with $n_1 + \dots + n_l \leq n$. Let $m = n_1 + \dots + n_l$ and $A = \bigcup_{i=1}^l A_i$. Then, we have

$$\begin{aligned}
 & \mathbb{P}(G_1 \in A_1, \dots, G_{n_1} \in A_1, G_{n_1+1} \in A_2, \dots, G_m \in A_l \mid N = n) \\
 &= \sum_{\substack{k_1 \geq n_1, \dots, k_l \geq n_l \\ k_1 + \dots + k_l \leq n}} \mathbb{P}\left(G_1 \in A_1, \dots, G_m \in A_l \mid \bigcap_{j=1}^l \Pi(E \times A_j) = k_j, N = n\right) \\
 & \quad \cdot \mathbb{P}(\Pi(E \times A_j) = k_j, j = 1, \dots, l \mid N = n) \\
 &= \sum_{\substack{k_1 \geq n_1, \dots, k_l \geq n_l \\ k_1 + \dots + k_l \leq n}} \frac{k_1 \dots k_l}{n} \frac{k_1 - n_1 + 1}{n - n_1 + 1} \frac{k_2}{n - n_1} \dots \frac{k_l - n_l + 1}{n - m + 1}
 \end{aligned}$$

$$\begin{aligned}
 & \cdot \binom{n}{k_1, \dots, k_l, n - k_1 - \dots - k_l} \cdot \tilde{Q}(A_1)^{k_1} \dots \tilde{Q}(A_l)^{k_l} \cdot (1 - \tilde{Q}(A))^{n - k_1 - \dots - k_l} \\
 = & \left(\prod_{i=1}^l \tilde{Q}(A_i)^{n_i} \right) \cdot \left(\sum_{\substack{k_1 \geq n_1, \dots, k_l \geq n_l \\ k_1 + \dots + k_l \leq n}} \binom{n - m}{k_1 - n_1, \dots, k_l - n_l, n - k_1 - \dots - k_l} \right) \\
 & \cdot \tilde{Q}(A_1)^{k_1 - n_1} \dots \tilde{Q}(A_l)^{k_l - n_l} \cdot (1 - \tilde{Q}(A))^{n - k_1 - \dots - k_l} \\
 = & \left(\prod_{i=1}^l \tilde{Q}(A_i)^{n_i} \right) \cdot \left(\sum_{\substack{k_1 \geq 0, \dots, k_l \geq 0 \\ k_1 + \dots + k_l \leq n - m}} \binom{n - m}{k_1, \dots, k_l, n - m - k_1 - \dots - k_l} \right) \\
 & \cdot \tilde{Q}(A_1)^{k_1} \dots \tilde{Q}(A_l)^{k_l} \cdot (1 - \tilde{Q}(A))^{n - m - k_1 - \dots - k_l} = \prod_{i=1}^l \tilde{Q}(A_i)^{n_i}.
 \end{aligned}$$

which is the assertion of the theorem. \square

Remark 4.11. In case of the original Brown-Resnick process, \tilde{Q} can be described even more explicitly than in Theorem 4.10, see [Engelke et al. \(2011\)](#) for details.

4.3 Finite Approximations

Let Y_1, Y_2, \dots be independent exponentially distributed random variables with parameter $\lambda > 0$ and define $R_n = \sum_{i=1}^n Y_i$ for $n \in \mathbb{N}$. Then, $\sum_{i \in \mathbb{N}} \delta_{R_i}$ is a Poisson point process on $(0, \infty)$ with intensity λ . Applying the mapping theorem ([Kingman, 1993](#)) we get that $\sum_{i \in \mathbb{N}} \delta_{-\log R_i}$ is a Poisson point process on \mathbb{R} with intensity measure $\lambda \exp(-u) du$ and the sequence $(U_i)_{i \in \mathbb{N}}$ with $U_i = -\log R_i$ is monotonically decreasing.

For simplicity we will only consider approximations of Z on a symmetric cuboid $[-b, b]$, $b \in (0, \infty)^d$, based on the definition of Z and the representations Z_1, Z_2, Z_3 , and Z_4 , respectively.

1. Approximation method based on the definition of $Z(\cdot)$

For the Poisson point process $\Pi = \sum_{i \in \mathbb{N}} \delta_{U_i}$ with intensity measure $\exp(-u) du$ and independent copies $\{W_i(t), t \in [-b, b]\}$ of a Gaussian process with stationary increments, let

$$Z^{(k)}(t) = \max_{i=1, \dots, k} \left(U_i + W_i(t) - \frac{\sigma^2(t)}{2} \right), \quad t \in [-b, b], \quad k \in \mathbb{N}.$$

2. Approximation method based on the representation $Z(\cdot) = Z_1(\cdot)$

Let $\{U_i^{(j)}\}_{i \in \mathbb{N}}, j = 1, \dots, n$, be decreasing sequences of points of Poisson point processes $\sum_{i \in \mathbb{N}} \delta_{U_i^{(j)}}$ with intensity measure $n^{-1} \exp(-u) du$ and $\{W_i^{(j)}(t), t \in [-b - s_j, b - s_j]\}$ be independent copies of Gaussian processes. Then, for $k \in \mathbb{N}$, let

$$Z_1^{(k)}(t) = \max_{j=1, \dots, n} \max_{i=1, \dots, k} \left(U_i^{(j)} + W_i^{(j)}(t - s_j) - \frac{\sigma^2(t - s_j)}{2} \right), \quad t \in [-b, b].$$

3. Approximation method based on the representation $Z(\cdot) = Z_2(\cdot)$

Let $\sum_{i \in \mathbb{N}} \delta_{(S_i, U_i)}$ be a Poisson point process on $I \times \mathbb{R}$ with intensity measure $|I|^{-1} ds \times \exp(-u) du$. For each $i \in \mathbb{N}$, let $\{W_i(t), t \in [-b - I_{\max}, b - I_{\min}]\}$ be independent copies of a Gaussian process where I_{\min} and I_{\max} are the lower and upper end point of I , and let

$$Z_2^{(k)}(t) = \max_{i=1, \dots, k} \left(U_i + W_i(t - S_i) + \frac{\sigma^2(t - S_i)}{2} \right), \quad t \in [-b, b], \quad k \in \mathbb{N}.$$

 4. Approximation method based on the representation $Z(\cdot) = Z_3(\cdot)$

For $j_{\min} \leq j \leq j_{\max} \in \mathbb{Z}^d$, let $\{U_i^{(j)}\}_{i \in \mathbb{N}}$ be descending sequences of points of the Poisson point processes $\sum_{i \in \mathbb{N}} \delta_{U_i^{(j)}}$ with intensity measure $m^{-d} \exp(-u) du$. We assume $pj_{\min} < a < b < pj_{\max}$. Furthermore, we have independent copies $\{W_i^{(j)}, t \in [-b - pj, b - pj]\}$ of Gaussian process and define $T_i^{(j)} = \inf(\operatorname{argsup}(W_i^{(j)}(t) - \frac{\sigma^2(t)}{2}))$. For $k \in \mathbb{N}$, $t \in [-b, b]$, let

$$Z_3^{(k)}(t) = \max_{j=j_{\min}, \dots, j_{\max}} \max_{\substack{i=1, \dots, k \\ T_i^{(j)} \in (-\frac{m}{2}p, \frac{m}{2}p]^d}} \left(U_i^{(j)} + W_i^{(j)}(t - pj) - \frac{\sigma^2(t - pj)}{2} \right).$$

 5. Approximation method based on the representation $Z(\cdot) = Z_4(\cdot)$

Let I be a finite cuboid in \mathbb{R}^d with $[-b, b] \subset I$. Let (\tilde{U}_i, S_i) be descending in \tilde{U}_i such that $\sum_{i \in \mathbb{N}} \delta_{(\tilde{U}_i, S_i)}$ is a Poisson point process on $\mathbb{R} \times I$ with intensity measure $\lambda^* \exp(-\tilde{u}) d\tilde{u} \times ds$, i.e. $\tilde{U}_1 \geq \tilde{U}_2 \geq \tilde{U}_3 \geq \dots$. For each $i \in \mathbb{N}$, let \tilde{F}_i be an independent sample path with law \tilde{Q} . Then, for $k \in \mathbb{N}$, $t \in [-b, b]$, let

$$Z_4^{(k)}(t) = \max_{i=1, \dots, k} \left(\tilde{U}_i + \tilde{F}_i(t - S_i) \right).$$

This construction is illustrated by Figure 4.2.

For simulating $\tilde{F}_i \sim_{i.i.d.} \tilde{Q}$ or the discretized version $\tilde{F}_i^{(p)} \sim_{i.i.d.} \tilde{Q}^{(p)}$ we can either use Theorem 4.8 or 4.10. In the first case we simulate independent copies $W_j(\cdot)$ of $W(\cdot)$ and reject all those processes with $T_j^{(p)} \neq 0$. In the second case we simulate a Brown-Resnick process (e.g. using the standard approximation method) and use all those processes F_j with $(U_j + M_j, T_j) \in E$ in a random order. We have to choose E carefully such that we have to simulate as few processes $W_j(\cdot)$ as possible to get a realization of \tilde{F}_i .

The constants λ^* or $\lambda^{(p)} p^d$ can be estimated by counting $\#\{i \in \mathbb{N} : U_i + M_i > 0, T_i \in [0, 1]^d\}$ or $\#\{i \in \mathbb{N} : U_i > 0, T_i^{(p)} = 0\}$ (i.e. estimating $\mathbb{P}(T^{(p)} = 0)$), respectively.

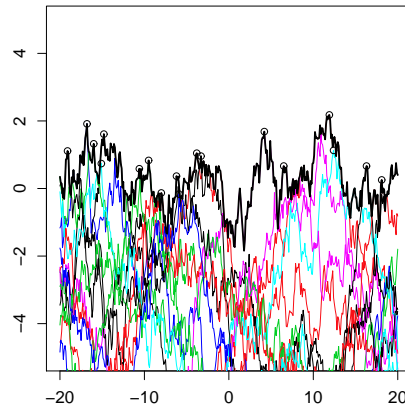


Figure 4.2: Construction of one sample path of the original Brown-Resnick process based on the representation $Z(\cdot) = Z_4(\cdot)$. The circles mark the points (S_i, \tilde{U}_i) . The resulting process is displayed by the bold line.

4.4 Error Estimates

In this section we assume that $\{W(t), t \in \mathbb{R}\}$ is a one-dimensional standard Brownian motion. Furthermore, we consider a symmetric interval $[-b, b]$ on which the simulation is performed. Under these assumptions, we give some basic error estimates for the approximations from Section 4.3. For details see Oesting (2009), Version 2.

In order to estimate the error for the approximation of $Z(\cdot)$ by $Z^{(k)}(\cdot)$, we define the random variable $C_k = \inf_{t \in [-b, b]} (Z^{(k)}(t) + \sigma^2(t)/2)$. Then, we get the following result.

Proposition 4.12. (Oesting, 2009, Version 2, Proposition 3.8) *Let $u < c < 0$. Then,*

$$\begin{aligned} & \mathbb{P}(Z^{(k)}(t) \neq Z(t) \text{ for some } t \in [-b, b] \mid U_k \leq u, C_k > c) \\ & \leq 4e^{-c} \frac{b}{c-u} \exp\left(\frac{b}{2}\right) \left(1 - \Phi\left(\frac{c-u-b}{\sqrt{b}}\right)\right). \end{aligned} \quad (4.5)$$

Furthermore, we have

$$\mathbb{P}(C_k \leq c) \leq 1 - \left(1 - \exp\left(-\exp\left(-\frac{c}{2}\right)\right)\right) \cdot \left(2\Phi\left(-\frac{c}{2\sqrt{b}}\right) - 1\right)^2 \quad (4.6)$$

and

$$\mathbb{P}(U_k > u) \leq \frac{\exp(-(k-1)u)}{(k-1)!} (1 - \exp(-\exp(-u))). \quad (4.7)$$

The unconditional error probability $\mathbb{P}(Z^{(k)}(t) \neq Z(t) \text{ for some } t \in [-b, b])$ can be bounded by the sum of the rhs of (4.5), (4.6) and (4.7).

If the right-hand sides of (4.5), (4.6) and (4.7) tend to 0 for $k \rightarrow \infty$, e.g. for $c = c(k) = -\log \log(\log(k)/2)$ and $x = x(k) = -\log(k)/2$, we get

$$\lim_{k \rightarrow \infty} \mathbb{P}(Z^{(k)}(t) \neq Z(t) \text{ for some } t \in [-b, b]) = 0.$$

Similar results hold for $Z_1^{(k)}(\cdot)$ and $Z_2^{(k)}(\cdot)$, cf. [Oesting \(2009\)](#), Version 2, Propositions 3.9 and 3.10. In particular, these bounds have the same asymptotic behaviour.

To give bounds for $Z_4(\cdot)$, we consider the standard Brownian motion restricted to the lattice $p\mathbb{Z}$ for some $p > 0$. Here, we also have to adjust the Poisson point process $\sum_{i \in \mathbb{N}} \delta_{(\tilde{U}_i, S_i)}$ to the lattice and to restrict its second component to some finite interval $[-v, v]$. Let $\sum_{i \in \mathbb{N}} \delta_{(Y_i^{(p)}, R_i^{(p)})}$ be a Poisson point process on $\mathbb{R} \times [-v, v]$ with intensity measure $\lambda^{(p)} p \exp(-y) dy \times \sum_{k \in \mathbb{Z} \cap [-v/p, v/p]} \delta_{pk}(ds)$. Then, we get an approximation of Z_4 by

$$Z_4^{(k,v,p)}(t) = \max_{i=1, \dots, k} \left(Y_i^{(p)} + \tilde{F}_i^{(p)}(t - R_i^{(p)}) \right)$$

for $t \in [-b, b] \cap p\mathbb{Z}$, where $\tilde{F}_i^{(p)} \sim_{i.i.d.} \tilde{Q}^{(p)}$. For the following result we use that $\tilde{Q}^{(p)}$ is the law of $W(t) - \frac{|t|}{2} \mid T^{(p)} = 0$, where W is a standard Brownian motion and $T^{(p)} = \inf(\operatorname{argsup}_{t \in p\mathbb{Z}}(W(t) - |t|/2))$.

Proposition 4.13. ([Oesting, 2009](#), Version 2, Proposition 3.14) *Let $b, v \in p\mathbb{N}$ with $v - b \geq 16$ and $b \geq 1$. Furthermore, let $C_k = \min_{t \in [-b, b] \cap p\mathbb{Z}} Z_4^{(k,v,p)}(t)$ and $c < 0$. Then, we have*

$$\begin{aligned} & \mathbb{P}(Z_4^{(k,v,p)}(t) \neq Z_4(t) \text{ for some } t \in [-b, b] \cap p\mathbb{Z} \mid C_k > c, Y_k^{(p)} \leq c) \\ & \leq \frac{64\lambda^{(p)}e^{-c}}{(1 - \exp(-p/2))^2} \left(\exp\left(-\frac{\sqrt{v-b}}{2}\right) + 7\sqrt{b^3} \exp\left(-\frac{v-b}{48b}\right) \right). \end{aligned} \quad (4.8)$$

Furthermore, it holds that

$$\begin{aligned} \mathbb{P}(C_k \leq c) & \leq 1 - \left(1 - \exp\left(-\lambda^p(2b+p)e^{-c/2}\right) \right) \\ & \quad \cdot \left(1 - (2(v-b)\lambda^{(p)})^k \frac{\exp\left(-\frac{(k-1)c}{2}\right)}{(k-1)!} (1 - \exp(-e^{-(v-b)\lambda^{(p)}c})) \right) \\ & \quad \cdot \left(1 - \frac{4}{1 - e^{-p/2}} \left(1 - \Phi\left(-\frac{c}{\sqrt{8b}} - \sqrt{\frac{b}{2}}\right) \right. \right. \\ & \quad \quad \left. \left. - \exp\left(-\frac{c}{2}\right) \left(1 - \Phi\left(-\frac{c}{\sqrt{8b}} + \sqrt{\frac{b}{2}}\right) \right) \right) \right)^2 \end{aligned} \quad (4.9)$$

and

$$\mathbb{P}(Y_k^{(p)} > c) \leq ((2v+p)\lambda^{(p)})^k \frac{\exp(-(k-1)c)}{(k-1)!} (1 - \exp(-e^{-(2v+p)\lambda^{(p)}c})). \quad (4.10)$$

The unconditional error probability $\mathbb{P}(Z_4^{(k,v,p)}(t) \neq Z_4(t) \text{ for some } t \in [-b, b])$ can be bounded by the sum of the rhs of (4.8), (4.9) and (4.10).

If we choose $v = v(k) = k^{1/4}$ and $c = c(k) = -\log(k)/4$ for example, again all the error bounds given in Proposition 4.13 tend to 0 as $k \rightarrow \infty$.

Here, the unconditional bound is given for fixed $k \in \mathbb{N}$. However, we can also use a stopping rule similar to [Schlather \(2002\)](#) and consider $Z_4^{(k)}$ with $k \in \mathbb{N}$ such that

$C_k > Y_k^{(p)}$. Then, the error can be assessed by the bound for the probability (4.8) which is independent of k .

Note that the upper bounds given in Proposition 4.13 can be made explicit by employing $\lambda^{(p)} = \mathbb{P}(T^{(p)} = 0)/p$ (cf. Remark 4.9) and using further bounds like $(1 - \exp(-p/2))^2/4 \leq \mathbb{P}(T^{(p)} = 0) \leq 1$. Similar bounds can be given for $Z_3^{(k)}$, cf. Oesting (2009), Propositions 3.11 and 3.12.

4.5 Simulation Study

In order to compare the different simulation techniques described in Section 4.3 we perform a simulation study on \mathbb{R}^1 using the software environment R (Ihaka and Gentleman, 1996; R Development Core Team, 2011). Functions to simulate Brown-Resnick processes based on different approximations will be available in a future version of the R package `RandomFields` (Schlather, 2012). We consider symmetric intervals $[-b, b]$ with $b \in \{1, 2, 5, 10, 20, 30, 50\}$ and the variogram $\gamma(h) = 2|h|^\alpha$ for $\alpha \in \{0.2, 0.4, 0.6, \dots, 1.8\}$. This means that for all these Brown-Resnick processes condition (4.1) holds. We always consider the process on a grid of length 0.1.

In order to get a fair criterion for the comparison of the different methods, we have fixed the number q of simulated sample paths of $W(\cdot)$ on $[-b, b]$ per realization of the Brown-Resnick process, $q = 100, 500$ and 2500 . Simulation techniques for W are given in Lantuéjoul (2002) and Schlather (2012). In order to approximate Z_3 and Z_4 the paths $W(\cdot)$ have to be computed on an interval larger than $[-b, b]$. Here, we assume that the computing time depends linearly on the length of this interval and modify the number of simulated sample paths to have an approximately equal computing time for all the approximations. We repeat every simulation of a Brown-Resnick process $N = 5000$ times and call this a run.

For ease of notation, we call the approximation of Z_i “method i ” ($i = 1, 2, 3, 4$) and the approximation of Z “method 0”.

Applying method 1, we have to choose s_1, \dots, s_n depending on b . It seems to be reasonable to distribute s_1, \dots, s_n equally on $[-b, b]$. Furthermore, the distance $\Delta s = s_2 - s_1$ should neither be too large — because we want to cover the interval with good approximations — nor too small — in order to get a method distinct from method 2. Here, depending on b , we choose some $n \in \{5, 10, 20, 50\}$ such that $0.5 \leq \Delta s \leq 2$.

In order to approximate $Z_2(\cdot)$, let Q be the uniform distribution on $[-b, b]$. When approximating $Z_3(\cdot)$, we always set p as the mesh width, i.e. $p = 0.1$, and set $j_{\max} = 200$ for $b = 1, 2$, $j_{\max} = 250$ for $b = 5$ and $j_{\max} = 10b + 300$ otherwise. Note that for the choice of j_{\max} (and also for the choice of v in method 4) not its absolute value, but the difference $pj_{\max} - b$ is important. In practice, a difference of 30 (or larger) provides very good results. However, one should be aware of the fact, that increasing j_{\max} is quite expensive in terms of computing times if k is fixed. We also varied the intensity parameter m and got best results for $m = 31$.

In case of method 4 we set $v = 20$ for $b = 1, 2$, $v = 25$ for $b = 5$ and $v = b + 30$ otherwise. Here, the choice of the set E is crucial. A large domain of E requires the simulation of low values of U_i , involving high computational costs. A very small domain, however, leads to a high rejection rate. We choose E in the following way. Let $\tilde{\Pi} = \sum_{i \in \mathbb{N}} \delta_{(U_i, T_i, U_i + M_i)}$

be a Poisson point process on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ with intensity measure $\tilde{\Lambda}$. From Theorem 4.6 it is known that $\tilde{\Lambda}(\mathbb{R} \times E) = \lambda^* \int_E e^{-y} dt \times dy$. We compare this to $\tilde{\Lambda}([u_0, \infty) \times E)$ for some fixed $u_0 \in \mathbb{R}$. The latter one can be easily estimated by simulation. Then, the probability of drawing \tilde{F}_i incorrectly when restricting our simulation to processes with $U_j > x_0$ given that $\tilde{\Pi}([u_0, \infty) \times E) > 0$ can be bounded by

$$\mathbb{P}(\tilde{\Pi}((-\infty, u_0) \times E) > 0) = 1 - \exp(-\tilde{\Lambda}((-\infty, u_0) \times E)).$$

For our simulation study we choose $u_0 = -2$ and approximate the area of highest intensity with cubes.

As already mentioned before, for methods 3 and 4, the (location of the) maximum of the Gaussian process is needed. To this end, we simulate the Gaussian process on a larger interval, which is at least of length $pj_{\max} + b$ or $v + b$, respectively, and take the maximum of the process restricted to this area which implies additional errors. Note that we do not get any additional error by discretization as the equivalent representations $Z_3(\cdot)$ and $Z_4(\cdot)$ also hold for Brown Resnick processes restricted to a lattice (cf. Remark 4.7).

As a measure of approximation quality we take the largest distance between the empirical cumulative distribution function of the approximated process at the interval bounds and the standard Gumbel distribution function. This is motivated by the fact that we expect the largest deviations from the original process at the interval bounds. Both bounds are taken into account in order to get a lower volatility of the results: For independent realizations $Z_{i,1}^{(k)}(\cdot), \dots, Z_{i,N}^{(k)}(\cdot)$, let $Z_{i,(1)}^{(k)}(t), \dots, Z_{i,(N)}^{(k)}(t)$ be the order statistic at location $t \in p\mathbb{Z}$. Then, we define the deviation of approximation as

$$\begin{aligned} \epsilon = & \frac{1}{2} \max_{j=1, \dots, N} \left| \exp(-\exp(-Z_{i,(j)}^{(k)}(-b))) - \frac{j}{N} \right| \\ & + \frac{1}{2} \max_{j=1, \dots, N} \left| \exp(-\exp(-Z_{i,(j)}^{(k)}(b))) - \frac{j}{N} \right|. \end{aligned}$$

We perform all the simulations up to 50 times. After each run of all methods we calculate the p -values for pairwise t-tests between the different methods assuming that ϵ is normally distributed. We stop simulating a method whenever $p < 0.005$. Figure 4.3 depicts the methods for each pair (b, α) that have not been rejected after 50 repetitions.

In general, methods 0, 1 and 2 perform best, if α or b is small. If both are large, method 4 is the best one. The area where method 4 performs best increases for q growing. For large q methods 0 and 2 have the same behaviour; if q is small, there is a sharper distinction between these methods. Method 0 provides better results for small b , method 2 for small α . Method 3 only works well if q gets large. Then, we get best results for b large.

The typical behaviour for large b , α and q is shown in Figures 4.1 and 4.4 for the standard Brownian motion. The development of the deviation ϵ for growing q and different b and α is shown in Table 4.1. More generally, there are the following recommendations concerning the choice of methods in practice: If the variogram value evaluated at the diameter of the simulated area is low, then use the original definition; simulation by random shifting is also appropriate if an imprecise simulation is sufficient; if the variogram tends to infinity and the value evaluated at the diameter of the simulated window is high, then the mixed moving maxima representation is best.

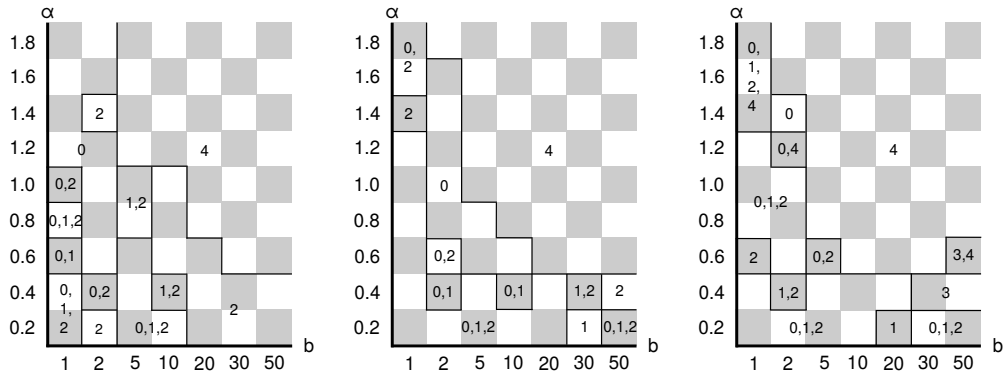


Figure 4.3: Methods providing best results depending on the interval bound b and variogram parameter α using $q = 100$ (left), 500 (middle), and 2500 (right) simulated sample paths of $W(\cdot)$ on $[-b, b]$ per realization of the Brown-Resnick process.

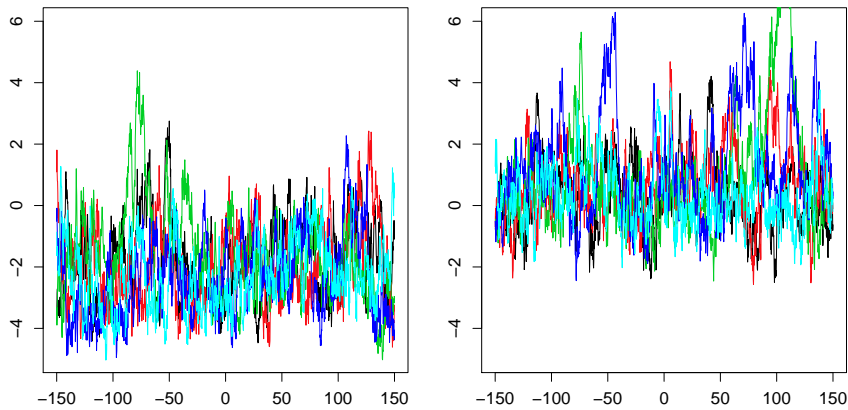


Figure 4.4: Finite approximations of the original Brown-Resnick process; five realizations of $Z_2^{(1000)}(\cdot)$ (left) and $Z_4^{(1000)}(\cdot)$ (right).

	$\alpha = 0.4, b = 30$			$\alpha = 1, b = 10$			$\alpha = 1.6, b = 2$		
q	100	500	2500	100	500	2500	100	500	2500
ϵ_0	0.148	0.066	0.027	0.662	0.514	0.367	0.085	0.030	0.014
ϵ_1	0.153	0.064	0.026	0.429	0.345	0.281	0.137	0.101	0.065
ϵ_2	0.135	0.063	0.030	0.423	0.339	0.280	0.129	0.084	0.046
ϵ_3	0.816	0.472	0.023	0.833	0.493	0.026	0.848	0.519	0.025
ϵ_4	0.379	0.099	0.066	0.514	0.076	0.014	0.357	0.045	0.012

Table 4.1: The mean deviation of approximation is shown for different (α, b, q) . By ϵ_i we denote the deviation of approximation by method i , $i = 0, \dots, 4$.

5 Conditional Sampling of Mixed Moving Maxima Processes

This chapter, which is based on the article [Oesting and Schlather \(2012\)](#), and Chapter [6](#) are devoted to conditional simulation of some classes of max-stable processes. In view of the variety of models, which have been developed over the last decades, and the wide range of potential applications of max-stable processes for modelling extreme events, the question of prediction and conditional sampling arises. [Davis and Resnick \(1989, 1993\)](#) proposed prediction procedures for time series which basically aim to minimize a suitable distance between observation and prediction. Further approaches for time series or random fields have been rare for a long time, apart from a few exceptions. [Cooley et al. \(2007\)](#) introduced an approximation of the conditional density. Recently, [Wang and Stoev \(2011\)](#) proposed an exact and efficient algorithm for conditional sampling for max-linear models

$$Z_i = \max_{j=1, \dots, p} a_{ij} Y_j, \quad i = 1, \dots, n,$$

where Y_j are independent Fréchet random variables. Algorithms for conditional simulation of Brown-Resnick processes ([Dombry et al., 2011](#)) and for extremal Gaussian processes ([Dombry and Ribatet, 2012](#)) were developed based on more general results on conditional distributions of max-stable processes given in [Dombry and Eyi-Minko \(2011\)](#).

In this chapter, we consider stationary max-stable processes with standard Fréchet margins that allow for a mixed moving maxima representation (see, for instance, [Schlather 2002](#), [Stoev and Taqqu 2005](#)). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $F : (\Omega, \mathcal{F}) \rightarrow (G, \mathcal{G})$ be a random function such that $\mathbb{E}(\int_{\mathbb{R}^d} F(x) dx) = 1$. We assume that G is a countable set of measurable functions $f : \mathbb{R}^d \rightarrow [0, \infty)$ and $\mathcal{G} = 2^G$.

Then, we consider the stationary max-stable process

$$Z(t) = \max_{i \in \mathbb{N}} (U_i \cdot F_i(t - S_i)), \quad t \in \mathbb{R}^d, \quad (5.1)$$

where (S_i, U_i, F_i) , $i \in \mathbb{N}$, are the elements of a Poisson point process $\Pi = \sum_{i \in \mathbb{N}} \delta_{(S_i, U_i, F_i)}$ on $S = \mathbb{R}^d \times (0, \infty) \times G$ with intensity measure

$$\Lambda(A \times B \times C) = |A| \cdot \mathbb{P}_F(C) \cdot \int_B \frac{du}{u^2}, \quad A \in \mathcal{B}^d, B \in \mathcal{B} \cap (0, \infty), C \in \mathcal{G}, \quad (5.2)$$

and \mathbb{P}_F is the probability measure belonging to F .

We aim to sample from the conditional distribution of Z given $Z(t_1), \dots, Z(t_n)$ for fixed $t_1, \dots, t_n \in \mathbb{R}^d$. As Z is entirely determined by the Poisson point process Π , we analyse the distribution of Π given some values of Z . The idea to use a Poisson point process structure for calculating conditional distributions has already been implemented in the

case of a bivariate min-stable random vector (Weintraub, 1991). A very general Poisson point process approach was recently used by Dombry and Eyi-Minko (2011) yielding formulae for conditional distributions in terms of the exponent measure. Some of these results are independently found here, as well.

This chapter is organized as follows. In Section 5.1, we introduce a random partition of Π into three measurable point processes. This partition allows to focus on the critical points of Π which determine $Z(t_1), \dots, Z(t_n)$. Similarly to Wang and Stoev (2011) and Dombry and Eyi-Minko (2011), we will call realizations of these point configuration scenarios and figure out the conditional distribution of these scenarios coping with the problem that we work on events of probability zero (Section 5.2). In Section 5.3, the conditional distribution of Π is calculated explicitly for the case $d = 1$ and under some regularity assumptions on a finite number of random shape functions. In Section 5.4, the results are applied to the Gaussian extreme value process (Smith, 1990) and compared to other algorithms. Section 5.5 deals with sampling techniques which are needed in the case that the number of shape functions is large. In Section 5.6, an approximation procedure is introduced for the case of a countable and uncountable number of random shape functions. A prominent example, the Brown-Resnick process (Brown and Resnick, 1977), is the matter of a comparison study for different algorithms in Section 5.7. In the last section, we give a brief overview of the results for a discrete mixed moving maxima process restricted to $p\mathbb{Z}^d$.

5.1 Random Partition of Π and Measurability

In this section, we will consider random sets of points within Π which essentially determine the process Z . Separating these critical points of Π from the other ones, we get a random partition of Π . We will show that this partition is measurable, which allows for further investigation of this partition.

For some fixed $(t, z) \in \mathbb{R}^d \times (0, \infty)$, define the set

$$K_{t,z} = \left\{ (x, y, f) \in S : f(t-x) > 0, y = \frac{z}{f(t-x)} \right\},$$

which we call the *set of points generating* (t, z) due to the fact that

$$Z(t) = z \iff \Pi(K_{t,z}) \geq 1 \wedge \Pi(\overline{K_{t,z}}) = 0.$$

Here, $\overline{K} = \bigcup_{(x,y,f) \in K} \{x\} \times (y, \infty) \times \{f\}$ for a set $K \subset S$.

In a next step, we consider n fixed points $(t_1, z_1), \dots, (t_n, z_n) \in \mathbb{R}^d \times (0, \infty)$. For any vector $\mathbf{t} = (t_1, \dots, t_n) \in \mathbb{R}^{d \cdot n}$ and any mapping g with domain $\text{dom}(g) \subset \mathbb{R}^d$, we write $g(\mathbf{t})$ instead of $(g(t_1), \dots, g(t_n))$, for short. Similarly, $\mathbf{t} > 0$ is understood as $t_i > 0$, $i = 1, \dots, n$.

Now, we define the *set of points generating* (\mathbf{t}, \mathbf{z}) as

$$K_{\mathbf{t}, \mathbf{z}} = \left\{ (x, y, f) \in S : \max_{i=1, \dots, n} f(t_i - x) > 0, y = \min_{i=1, \dots, n} \frac{z_i}{f(t_i - x)} \right\}$$

$$= \left\{ (x, y, f) \in S : yf(\mathbf{t} - x) \leq \mathbf{z}, yf(t_j - x) = z_j \text{ for some } j \in \{1, \dots, n\} \right\}.$$

This implies

$$\overline{K_{\mathbf{t}, \mathbf{z}}} = \left\{ (x, y, f) \in S : yf(t_j - x) > z_j \text{ for some } j \in \{1, \dots, n\} \right\}$$

and

$$K_{\mathbf{t}, \mathbf{z}} \cap K_{t_i, z_i} = \left\{ (x, y, f) \in S : yf(t_i - x) = z_i, yf(\mathbf{t} - x) \leq \mathbf{z} \right\}.$$

Therefore, we have

$$Z(\mathbf{t}) \leq \mathbf{z} \iff \Pi(\overline{K_{\mathbf{t}, \mathbf{z}}}) = 0$$

and

$$Z(\mathbf{t}) = \mathbf{z} \iff \Pi(K_{t_i, z_i} \cap K_{\mathbf{t}, \mathbf{z}}) \geq 1, i = 1, \dots, n \wedge \Pi(\overline{K_{\mathbf{t}, \mathbf{z}}}) = 0. \quad (5.3)$$

Now we define a random partition of Π by

$$\begin{aligned} \Pi_1(\cdot) &:= \Pi(\cdot \cap \overline{K_{\mathbf{t}, Z(\mathbf{t})}}) \\ \Pi_2(\cdot) &:= \Pi(\cdot \cap K_{\mathbf{t}, Z(\mathbf{t})}), \\ \text{and } \Pi_3(\cdot) &:= \Pi(\cdot) - \Pi_1(\cdot) - \Pi_2(\cdot). \end{aligned}$$

Relation (5.3) implies that $\Pi_1(\cdot) \equiv 0$ and $\Pi_2(K_{t_i, Z(t_i)}) \geq 1$ a.s. for $i \in \{1, \dots, n\}$. Note that the partition of Π into Π_1 , Π_2 and Π_3 corresponds to the classification of Φ_K^+ and Φ_K^- in [Dombry and Eyi-Minko \(2011\)](#).

Before proceeding any further, we need to proof that Π_1 , Π_2 and Π_3 are well-defined. We will do this by showing the measurability of a further partition of Π_2 , namely the restriction of Π to intersection sets. For any $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, these are defined as

$$\begin{aligned} I_A(\mathbf{z}) &= K_{\mathbf{t}, \mathbf{z}} \cap \left(\bigcap_{i \in A} K_{t_i, z_i} \setminus \bigcup_{j \in A^c} K_{t_j, z_j} \right) \\ &= \left\{ (x, y, f) \in S : yf(t_i - x) = z_i, i \in A, yf(t_j - x) < z_j, j \notin A \right\} \end{aligned}$$

By construction, $K_{\mathbf{t}, \mathbf{z}}$ is a disjoint union of $I_A(\mathbf{z})$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$.

Proposition 5.1. *Let $t_1, \dots, t_n \in \mathbb{R}^d$ be fixed.*

1. *The mapping*

$$\Psi : S \rightarrow \mathbb{R}^{\mathbb{R}^d}, (x, y, f) \mapsto yf(\cdot - x)$$

is $(\mathcal{B}^d \times (\mathcal{B} \cap (0, \infty)) \times 2^G, \mathcal{C}(\mathbb{R}^d))$ -measurable.

2. *Let $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$ and $B \subset S$ a bounded Borel set. Then, $\Pi(I_A(Z(\mathbf{t})) \cap B)$ is a random variable.*

3. *Π_1, Π_2 and Π_3 are point processes (cf. [Dombry and Eyi-Minko, 2011](#)).*

Proof. 1. It suffices to verify that $\Psi^{-1}(C_{s_1, \dots, s_m}(\chi_{i=1}^m(a_i, b_i)))$ is measurable for any $s_j \in \mathbb{R}^d$, $a_j < b_j \in \mathbb{R}$, $j = 1, \dots, m$, $m \in \mathbb{N}$. We have

$$\begin{aligned} & \Psi^{-1}(C_{s_1, \dots, s_m}(\chi_{i=1}^m(a_i, b_i))) \\ &= \bigcup_{f \in G} \bigcup_{t \in \mathbb{R}^d} \{t\} \times \left(\max_{i=1, \dots, m} \frac{a_i}{f(t_i - t)}, \min_{i=1, \dots, m} \frac{b_i}{f(t_i - t)} \right) \times \{f\} \\ &= \bigcup_{f \in G} \bigcup_{\substack{q_1, q_2 \in \mathbb{Q}_+ \\ q_1 < q_2}} \left\{ t : (q_1, q_2) \subset \left(\max_{i=1}^m \frac{a_i}{f(t_i - t)}, \min_{i=1}^m \frac{b_i}{f(t_i - t)} \right) \right\} \times (q_1, q_2) \times \{f\}. \end{aligned}$$

As each $f \in G$ is measurable, sets of the type $\{t \in \mathbb{R}^d : f(t_i - t) \in B\}$ are measurable for any $B \in \mathcal{B}$. Therefore, $\Psi^{-1}(C_{s_1, \dots, s_m}(\chi_{i=1}^m(a_i, b_i))) \in \mathcal{B}^d \times (\mathcal{B} \cap (0, \infty)) \times 2^G$.

2. We consider

$$\begin{aligned} & \{\omega : \Pi(I_A(Z(\mathbf{t})) \cap B) = k\} \\ &= \bigcup_{n_0 \in \mathbb{N}} \bigcap_{m=n_0}^{\infty} \left(\bigcup_{\mathbf{y} \in \mathbb{Q}^n} \left\{ \omega \in \Omega : Z(\mathbf{t}) \in \chi_{i=1}^n \left(y_i - \frac{1}{m}, y_i + \frac{1}{m} \right), \right. \right. \\ & \quad \left. \left. \Pi \left(\Psi^{-1} \left(\left\{ f \in \mathbb{R}^{\mathbb{R}^d} : f(t_i) \in \left(y_i - \frac{1}{m}, y_i + \frac{1}{m} \right), i \in A, \right. \right. \right. \right. \right. \\ & \quad \left. \left. \left. \left. f(t_j) \leq y_j - \frac{1}{m}, j \notin A \right\} \right) \cap B \right) = k \right\} \right) \end{aligned}$$

By the first part of this proposition, Ψ is a measurable mapping and we get that $\{\omega : \Pi(I_A(Z(\mathbf{t})) \cap B) = k\}$ is measurable.

3. For any bounded Borel set $B \subset S$ the second part of this proposition yields that $\Pi_1(B) = 0$, $\Pi_2(B) = \sum_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \Pi(I_A(Z(\mathbf{t})) \cap B)$ and $\Pi_3(B) = \Pi(B) - \Pi_2(B)$ are measurable. Thus, Π_1 , Π_2 and Π_3 are point processes (Daley and Vere-Jones, 1988, Cor. 6.1.IV). □

5.2 Blurred Sets, Scenarios and Limit Considerations

This section mainly deals with the analysis of the distribution of the set of critical points, Π_2 . As this set has the intensity measure zero conditional on an event $Z(\mathbf{t}) = \mathbf{z}$ of probability zero, this distribution cannot be calculated straightforward. We need to borrow arguments from martingale theory, taking limits of probabilities conditional on the observations being in small intervals containing \mathbf{z} . By this condition, the set of critical points gets blurred. We distinguish between different scenarios denoting which points influence the different observations. Using general bounds for the rate of convergence of the intensity of these sets, we can prove that each observation is generated by exactly one point of Π (Proposition 5.5), which restricts the number of scenarios that occur with positive probability. According to the blurred sets, the scenarios get blurred, as

well. The blurred scenarios are not exactly the same as the scenarios conditional on blurred observations, but much more tractable. However, it can be shown that both events asymptotically yield the same conditional probability (Theorem 5.6). Based on these considerations, the independence of Π_1 , Π_2 and Π_3 conditional on $Z(\mathbf{t})$ is shown (Theorem 5.7). This allows us to simulate Π_2 and Π_3 independently. Furthermore, Π_3 turns out to be easily simulated (Theorem 5.7).

Let

$$\mathcal{F}_m = \sigma \left(\left\{ Z(t_i) \in \left(\frac{k}{2^m}, \frac{k+1}{2^m} \right], i = 1, \dots, n, k \in \mathbb{N}_0 \right\} \right).$$

Then, $\{\mathcal{F}_m\}_{m \in \mathbb{N}}$ is a filtration and $\mathcal{F}_\infty := \bigcap_{m \in \mathbb{N}} \mathcal{F}_m = \sigma(Z(\mathbf{t}))$. For $z > 0$, let $j_m(z) \in \mathbb{N}$ be such that $z \in A_m(z)$ with

$$A_m(z) = \left(\frac{j_m(z)}{2^m}, \frac{j_m(z) + 1}{2^m} \right].$$

Then, we have $A_m(z) \xrightarrow{m \rightarrow \infty} \{z\}$ monotonically and $\{\omega \in \Omega : Z(\mathbf{t}) \in A_m(\mathbf{z})\} \in \mathcal{F}_m$. Furthermore, let

$$K_{t,z}^{(m)} = \bigcup_{\bar{z} \in A_m(z)} K_{t,\bar{z}} = \{(x, y, f) \in S : yf(t-x) \in A_m(z)\},$$

$$K_{\mathbf{t},\mathbf{z}}^{(m)} = \bigcup_{\bar{\mathbf{z}} \in A_m(\mathbf{z})} K_{\mathbf{t},\bar{\mathbf{z}}} = \left\{ (x, y, f) \in S : yf(\mathbf{t}-x) \leq \frac{j_m(\mathbf{z}) + 1}{2^m}, \right.$$

$$\left. yf(t_i - x) \in A_m(z_i) \text{ for some } i \in \{1, \dots, n\} \right\},$$

$$\text{and } \overline{K_{\mathbf{t},\mathbf{z}}^{(m)}} = \bigcap_{\bar{\mathbf{z}} \in A_m(\mathbf{z})} \overline{K_{\mathbf{t},\bar{\mathbf{z}}}} = \left\{ (x, y, f) \in S : yf(t_i - x) > \frac{j_m(z_i) + 1}{2^m} \right.$$

$$\left. \text{for some } i \in \{1, \dots, n\} \right\} = \overline{K_{\mathbf{t}, \frac{j_m(\mathbf{z})+1}{2^m}}}.$$

These definitions imply that

$$K_{\mathbf{t},\mathbf{z}}^{(m)} \cap \overline{K_{\mathbf{t},\mathbf{z}}^{(m)}} = \emptyset \quad \text{and} \quad K_{\mathbf{t},\mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t},\mathbf{z}}^{(m)}} = \overline{K_{\mathbf{t}, \frac{j_m(\mathbf{z})+1}{2^m}}}.$$

We call these sets the blurred sets belonging to $Z(\mathbf{t})$ conditional on $Z(\mathbf{t}) \in A_m(\mathbf{z})$. This notation is due to the fact that we have

$$Z(\mathbf{t}) \leq \frac{j_m(\mathbf{z}) + 1}{2^m} \iff \Pi(\overline{K_{\mathbf{t},\mathbf{z}}^{(m)}}) = 0.$$

Furthermore, as

$$K_{\mathbf{t},\mathbf{z}}^{(m)} \cap K_{t_i,z_i}^{(m)} = \left\{ (x, y, f) \in S : yf(t_i - x) \in A_m(z_i), yf(\mathbf{t}-x) \leq \frac{j_m(\mathbf{z}) + 1}{2^m} \right\},$$

we get that

$$\Pi(K_{t_i,z_i}^{(m)} \cap K_{\mathbf{t},\mathbf{z}}^{(m)}) \geq 1, \quad i = 1, \dots, n \implies Z(\mathbf{t}) > \frac{j_m(\mathbf{z})}{2^m}.$$

Thus, we obtain

$$Z(\mathbf{t}) \in A_m(\mathbf{z}) \iff \Pi(K_{t_i, z_i}^{(m)} \cap K_{\mathbf{t}, \mathbf{z}}^{(m)}) \geq 1, \quad i = 1, \dots, n \quad \wedge \quad \Pi(\overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}}) = 0. \quad (5.4)$$

In particular, for fixed $\mathbf{z} \in (0, \infty)^n$, the point process $\Pi(\cdot \cap S \setminus (K_{\mathbf{t}, \mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}}))$ is independent of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$.

Based on these blurred sets, we define the *blurred intersection sets*

$$I_A^{(m)}(\mathbf{z}) = K_{\mathbf{t}, \mathbf{z}}^{(m)} \cap \bigcap_{i \in A} K_{t_i, z_i}^{(m)} \setminus \bigcup_{j \in A^c} K_{t_j, z_j}^{(m)}, \quad A \in 2^{\{1, \dots, n\}} \setminus \emptyset.$$

We note that $K_{\mathbf{t}, \mathbf{z}}^{(m)}$ can be written as a disjoint union of $I_A^{(m)}(\mathbf{z})$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$.

Lemma 5.2. *For any $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$ and $\mathbf{z} > \mathbf{0}$, we have $\Lambda(I_A^{(m)}(\mathbf{z})) \in \mathcal{O}(2^{-m})$, where $\Lambda(\cdot)$ is given by (5.2).*

Proof. It suffices to show that

$$\Lambda(K_{t_i, z_i}^{(m)}) \in \mathcal{O}(2^{-m}), \quad i = 1, \dots, n. \quad (5.5)$$

By a straightforward computation we get

$$\begin{aligned} \Lambda(K_{t_i, z_i}^{(m)}) &= \mathbb{E}_F \left(\int_{\mathbb{R}^d} \int_{\frac{j_m(z_i)}{2^m F(t_i - t)}}^{\frac{j_m(z_i) + 1}{2^m F(t_i - t)}} \frac{1}{u^2} \, du \, dx \right) \\ &= \mathbb{E}_F \left(\int_{\mathbb{R}^d} F(t_i - x) \, dx \right) \cdot \left(\frac{1}{\frac{j_m(z_i)}{2^m}} - \frac{1}{\frac{j_m(z_i) + 1}{2^m}} \right) = \frac{\frac{1}{2^m}}{\frac{j_m(z_i)}{2^m} \frac{j_m(z_i) + 1}{2^m}}. \end{aligned}$$

Using the fact that $\lim_{m \rightarrow \infty} \frac{j_m(z_i)}{2^m} = z_i$, the assertion of the Lemma follows. \square

In Section 5.3, a more precise notion about the speed of convergence of $2^{-m} j_m(z_i) \rightarrow z_i$ will be useful.

Lemma 5.3. *For any $\varepsilon > 0$, with probability one we have*

$$\begin{aligned} \liminf_{m \rightarrow \infty} 2^{m(1+\varepsilon)} \min_{i=1, \dots, n} \left(\frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \right) &= \infty \\ \text{and } \limsup_{m \rightarrow \infty} 2^{m(1+\varepsilon)} \max_{i=1, \dots, n} \left(\frac{j_m(Z(t_i))}{2^m} - Z(t_i) \right) &= -\infty. \end{aligned}$$

Proof. We present the proof of the first assertion. The second one can be shown analogously. For the first assertion, it suffices to show that

$$\liminf_{m \rightarrow \infty} 2^{m(1+\varepsilon)} \left(\frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \right) = \infty, \quad i = 1, \dots, n.$$

Let $a \in (0, \frac{1}{2})$ and $m \in \mathbb{N}$ large enough such that

$$\sum_{k=1}^{\infty} 2^{-m} \frac{1}{(k2^{-m})^2} \exp\left(-\frac{1}{k2^{-m}}\right) \leq 2 \int_0^{\infty} \frac{1}{x^2} \exp\left(-\frac{1}{x}\right) \, dx = 2. \quad (5.6)$$

Then, we have

$$\begin{aligned}
 & \mathbb{P} \left(\frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \leq a2^{-m} \right) \\
 &= \sum_{k=1}^{\infty} \mathbb{P} \left((k-a)2^{-m} \leq Z(t_i) \leq k2^{-m} \right) \\
 &= \sum_{k=1}^{\infty} \exp \left(-\frac{1}{k2^{-m}} \right) \cdot \left(1 - \exp \left(-\frac{1}{(k-a)2^{-m}} + \frac{1}{k2^{-m}} \right) \right) \\
 &= \sum_{k=1}^{\infty} \exp \left(-\frac{1}{k2^{-m}} \right) \cdot \left(1 - \exp \left(-\frac{1}{k2^{-m}} \cdot \frac{a/k}{1-a/k} \right) \right) \\
 &\leq \sum_{k=1}^{\infty} \exp \left(-\frac{1}{k2^{-m}} \right) \cdot \left(\frac{1}{k2^{-m}} \cdot 2\frac{a}{k} \right),
 \end{aligned}$$

where we used the fact that $1 - \exp(-x) \leq x$ for all $x > 0$. Employing (5.6), we get for $a = C2^{-m\varepsilon}$ with $C > 0$ and m large enough that

$$\begin{aligned}
 & \mathbb{P} \left(\frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \leq C \cdot 2^{-m(1+\varepsilon)} \right) \\
 &\leq 2C \cdot 2^{-m\varepsilon} \sum_{k=1}^{\infty} 2^{-m} \exp \left(-\frac{1}{k2^{-m}} \right) \cdot \left(\frac{1}{(k2^{-m})^2} \right) \leq 4C \cdot (2^\varepsilon)^{-m}.
 \end{aligned}$$

Therefore, the probabilities above are summable with respect to m and the Borel-Cantelli lemma yields

$$\begin{aligned}
 & \mathbb{P} \left(\liminf_{m \rightarrow \infty} 2^{m(1+\varepsilon)} \left(\frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \right) < C \right) \\
 &= \mathbb{P} \left(\limsup_{m \rightarrow \infty} \left\{ \omega \in \Omega : \frac{j_m(Z(t_i)) + 1}{2^m} - Z(t_i) \leq C \cdot 2^{-m(1+\varepsilon)} \right\} \right) = 0
 \end{aligned}$$

for any $C > 0$. This completes the proof. \square

Now, we relate the points in $K_{\mathbf{t}, \mathbf{z}}^{(m)}$ to the corresponding “blurred” intersection sets by introducing disjoint “blurred” scenarios

$$\begin{aligned}
 E_{\{n_A\}}^{(m)}(\mathbf{z}) &= E_{\{n_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}}^{(m)}(\mathbf{z}) \\
 &= \{ \omega \in \Omega : \Pi(I_A^{(m)}(\mathbf{z})) = n_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, \Pi(\overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}}) = 0 \}
 \end{aligned}$$

with $\{n_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\} \in N_1$ where

$$N_1 = \left\{ \{n_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}, \sum_{A: i \in A} n_A \geq 1, i = 1, \dots, n \right\}.$$

Thus, we have

$$\{ \omega : Z(\mathbf{t}) \in A_m(\mathbf{z}) \} = \bigcup_{\{n_A\} \in N_1} E_{\{n_A\}}^{(m)}(\mathbf{z})$$

and the union is obviously disjoint. In the same way, dropping the (m) in the definition, we specify scenarios $E_{\{n_A\}}(\mathbf{z})$.

Now, we show that

$$\mathbb{P}(\exists i \in \{1, \dots, n\} : \Pi(K_{t_i, Z(t_i)}) \geq 2) = 0.$$

To this end, we first verify that, with probability one, Π_2 has no points which can be removed without any effect on $Z(\mathbf{t})$. We consider scenarios $E_{\{n_A\}}(Z(\mathbf{t}))$ with $\{n_A\} \in N_2$. Here, N_2 is the ensemble of all sets $\{n_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}$ such that there exists some $A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset$ with

$$n_A^* = \begin{cases} n_{A^*} - 1, & A = A^*, \\ n_A, & \text{else.} \end{cases}$$

and $\{n_A^*\} \in N_1$. Thus, $\{n_A\} \in N_2$ if and only if $E_{\{n_A\}}(Z(\mathbf{t}))$ remains an allowable scenario after removing one point.

Lemma 5.4. *With probability one, we have*

$$\begin{aligned} & \mathbb{P} \left(\bigcup_{\{n_A\} \in N_2} E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z} \right) \\ &= \lim_{m \rightarrow \infty} \mathbb{P} \left(\bigcup_{\{n_A\} \in N_2} E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) = 0. \end{aligned}$$

In particular, $\lim_{m \rightarrow \infty} \mathbb{P} \left(\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0 \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) = 0$.

Proof. By Lévy's ‘‘Upward’’ Theorem (see [Rogers and Williams, 2000](#), Thm. 50.3), we a.s. have

$$\sum_{\{n_A\} \in N_2} \mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \sum_{\{n_A\} \in N_2} \mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})).$$

Noting that

$$\bigcup_{\{n_A\} \in N_2} E_{\{n_A\}}(Z(\mathbf{t})) \subset \bigcup_{\{n_A\} \in N_2} E_{\{n_A\}}^{(m)}(\mathbf{z}),$$

it suffices to show the equation

$$\lim_{m \rightarrow \infty} \sum_{\{n_A\} \in N_2} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) = \lim_{m \rightarrow \infty} \frac{\sum_{\{n_A\} \in N_2} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))}{\sum_{\{n_A\} \in N_1} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))} = 0. \quad (5.7)$$

for every $\mathbf{z} > \mathbf{0}$. Corresponding to each $E_{\{n_A\}}^{(m)}(\mathbf{z})$ with $\{n_A\} \in N_2$ there exists some $E_{\{n_A^*\}}^{(m)}(\mathbf{z})$ with $\{n_A^*\} \in N_1$ as in the definition of N_2 . For $A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset$, let N_{2, A^*} be the set of all $\{n_A\} \in N_2$ with $n_{A^*}^* = n_{A^*} - 1$. Then, for fixed $m \in \mathbb{N}$, the left-hand side of (5.7) is less than or equal to

$$\sum_{A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\sum_{\{n_A\} \in N_{2, A^*}} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))}{\sum_{\{n_A\} \in N_1} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))}$$

$$\begin{aligned}
 &= \sum_{A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\sum_{\{n_A\} \in N_{2, A^*}} \left(\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} e^{-\Lambda(I_A^{(m)})} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A}}{n_A!} \right) \cdot \exp(-\Lambda(\overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}}))}{\sum_{\{n_A\} \in N_1} \left(\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} e^{-\Lambda(I_A^{(m)})} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A}}{n_A!} \right) \cdot \exp(-\Lambda(\overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}}))} \\
 &= \sum_{A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\sum_{\{n_A\} \in N_{2, A^*}} \frac{\Lambda(I_{A^*}^{(m)}(\mathbf{z}))}{n_{A^*}} \cdot \left(\prod_{\substack{A \in 2^{\{1, \dots, n\}} \setminus \emptyset \\ A \neq A^*}} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A}}{n_A!} \right) \cdot \frac{\Lambda(I_{A^*}^{(m)}(\mathbf{z}))^{n_{A^*}-1}}{(n_{A^*}-1)!}}{\sum_{\{n_A\} \in N_1} \prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A}}{n_A!}} \\
 &\leq \sum_{A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\Lambda(I_{A^*}^{(m)}(\mathbf{z})) \sum_{\{n_A^*\} \in N_1} \frac{1}{n_{A^*}^*+1} \prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A^*}}{n_A^*!}}{\sum_{\{n_A\} \in N_1} \prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{n_A}}{n_A!}} \\
 &\leq \sum_{A^* \in 2^{\{1, \dots, n\}} \setminus \emptyset} \Lambda(I_{A^*}^{(m)}(\mathbf{z})),
 \end{aligned}$$

where we embedded N_{2, A^*} into N_1 by identifying $\{n_A\}$ with the corresponding $\{n_A^*\}$. By Lemma 5.2, we get equality (5.7), i.e. the first assertion of the lemma. Furthermore, $\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0$ and $Z(\mathbf{t}) \in A_m(\mathbf{z})$ imply that $\Pi(\cdot \cap K_{\mathbf{t}, \mathbf{z}}^{(m)})$ contains points which can be removed without affecting $Z(\mathbf{t}) \in A_m(\mathbf{z})$. This is,

$$\begin{aligned}
 &\lim_{m \rightarrow \infty} \mathbb{P} \left(\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0 \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) \\
 &\leq \lim_{m \rightarrow \infty} \sum_{\{n_A\} \in N_2} \mathbb{P} \left(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) = 0.
 \end{aligned}$$

Thus, the second assertion of this lemma is verified. \square

The following proposition is also stated in a more general setting in [Dombry and Eyi-Minko \(2011\)](#), Prop. 2.2.

Proposition 5.5. *For any fixed $t_1, \dots, t_n \in \mathbb{R}^d$ we have*

$$\mathbb{P}(\Pi(K_{t_i, Z(t_i)}) \geq 2 \text{ for some } i \in \{1, \dots, n\}) = 0.$$

Proof. It suffices to show $\mathbb{P}(\Pi(K_{t_i, Z(t_i)}) \geq 2) = 0$ for all $i \in \{1, \dots, n\}$. Now, let $i \in \{1, \dots, n\}$ be fixed. Then, conditioning on $Z(t_i)$ only, we get

$$\mathbb{P}(\Pi(K_{t_i, Z(t_i)}) \geq 2 \mid Z(t_i) = z) = \mathbb{P} \left(\bigcup_{\{n_A\} \in N_2} E_{\{n_A\}}(Z(t_i)) \mid Z(t_i) = z \right) = 0$$

for almost every $z > 0$ by Lemma 5.4. This yields the desired result. \square

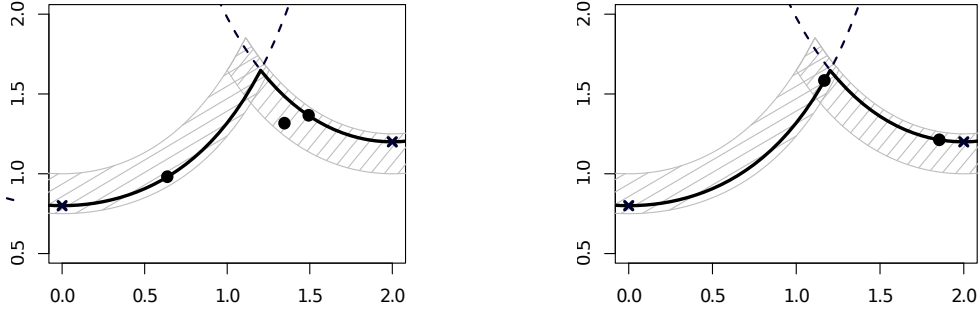


Figure 5.1: The event Δ_m in the proof of Thm. 5.6. Black crosses: data $(t_i, Z(t_i))$, black line: $K_{\mathbf{t}, Z(\mathbf{t})}$, grey area: $K_{\mathbf{t}, Z(\mathbf{t})}^{(m)}$, black dots: $\Pi(\cdot \cap K_{\mathbf{t}, Z(\mathbf{t})}^{(m)})$. Left: $\Pi(K_{\mathbf{t}, Z(\mathbf{t})}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0$. Right: $\Pi(I_{\{1\}}(Z(\mathbf{t}))) = 1 > 0 = \Pi(I_{\{1\}}^{(m)}(Z(\mathbf{t})))$

By Proposition 5.5, almost surely one of the scenarios $E_{\{n_A\}}(\mathbf{z})$ with $\sum_{A: i \in A} n_A = 1$ for all $i \in \{1, \dots, n\}$ occurs.

Theorem 5.6. *With probability 1 we have*

1. $\mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$
for any $\{n_A\} \in N_1$,
2. $\mathbb{P}(\Pi_2(B_j) = r_j, j = 1, \dots, k, \Pi_3(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z})$
 $= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \cap B_j) = r_j, j = 1, \dots, k,$
 $\Pi(B_j \setminus (K_{\mathbf{t}, \mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}})) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$
for any $B_j \subset S$, $r_j \in \mathbb{N}$, $j \in \{1, \dots, l\}$.

Proof. 1. Lévy's ‘‘Upward’’ Theorem (Rogers and Williams, 2000, Thm. 50.3) yields

$$\mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$$

with probability 1. It remains to verify

$$\lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}^{(m)}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) = \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \quad (5.8)$$

where $\mathbb{P}(E_{\{n_A\}}^{(m)}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$ equals $\mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$ by definition.

To this end, we consider the symmetric difference Δ_m of the events $E_{\{n_A\}}^{(m)}(Z(\mathbf{t}))$ and $E_{\{n_A\}}(Z(\mathbf{t}))$. Note that any element of Δ_m satisfies $\Pi(K_{\mathbf{t}, Z(\mathbf{t})}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0$ (cf. Figure 5.1, left) or $\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t})))$ for some $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$ (cf. Figure 5.1, right). The second kind of event occurs if there is a point of Π in $I_A(Z(\mathbf{t})) \cap (\bigcup_{j \notin A} K_{t_j, Z(t_j)}^{(m)})$. As this set vanishes for any $Z(\mathbf{t}) > \mathbf{0}$ as $m \rightarrow \infty$, we get that

$$\{\omega \in \Omega : \Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t})))\} \searrow \emptyset, \quad m \rightarrow \infty,$$

for any $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, and therefore

$$\lim_{m \rightarrow \infty} \mathbb{P}(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t})))) = 0.$$

This yields

$$\begin{aligned} & \int_0^\infty \lim_{m \rightarrow \infty} \mathbb{P}\left(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(\mathbf{z})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \mathbb{P}(Z(\mathbf{t}) \in d\mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t})))) = 0 \end{aligned}$$

using dominated convergence and the fact that

$$\begin{aligned} & \int_0^\infty \mathbb{P}\left(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(\mathbf{z})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \mathbb{P}(Z(\mathbf{t}) \in d\mathbf{z}) \\ &= \sum_{\mathbf{z} \in (2^{-m}\mathbb{N})^n} \mathbb{P}\left(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t}))) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \mathbb{P}(Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \mathbb{P}(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(Z(\mathbf{t})))) . \end{aligned}$$

Therefore, we have

$$\lim_{m \rightarrow \infty} \mathbb{P}\left(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(\mathbf{z})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) = 0 \quad (5.9)$$

for any $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$ and almost all $\mathbf{z} > \mathbf{0}$. All in all, we end up with

$$\begin{aligned} & \lim_{m \rightarrow \infty} \left| \mathbb{P}\left(E_{\{n_A\}}^{(m)}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) - \mathbb{P}\left(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \right| \\ & \leq \lim_{m \rightarrow \infty} \sum_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \mathbb{P}\left(\Pi(I_A(Z(\mathbf{t}))) > \Pi(I_A^{(m)}(\mathbf{z})) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \\ & \quad + \lim_{m \rightarrow \infty} \mathbb{P}\left(\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0 \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) = 0 \quad a.s. \end{aligned}$$

by (5.9) and by the second part of Lemma 5.4. Thus, we get (5.8).

2. Let $B_1, \dots, B_k, B_{k+1}, \dots, B_l \in \mathcal{B}^d \times (\mathcal{B} \cap (0, \infty)) \times 2^G$. Then, each of the events

$$\begin{aligned} & \left\{ \Pi_2(B_j) \neq \Pi(K_{\mathbf{t}, Z(\mathbf{t})}^{(m)} \cap B_j) \text{ for any } j = 1, \dots, k \right\} \\ \text{and} & \left\{ \Pi_3(B_j) \neq \Pi(B_j \setminus (K_{\mathbf{t}, Z(\mathbf{t})}^{(m)} \cup \overline{K_{\mathbf{t}, Z(\mathbf{t})}^{(m)}})) \text{ for any } j = k+1, \dots, l \right\} \end{aligned}$$

implies that $\Pi(K_{\mathbf{t}, Z(\mathbf{t})}^{(m)} \setminus K_{\mathbf{t}, Z(\mathbf{t})}) > 0$. Therefore, for any $r_1, \dots, r_l \in \mathbb{N}$, we have

$$\begin{aligned} & \lim_{m \rightarrow \infty} \left| \mathbb{P}\left(\Pi_2(B_j) = r_j, j = 1, \dots, k, \right. \right. \\ & \quad \left. \Pi_3(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \\ & \quad - \mathbb{P}\left(\Pi(K_{\mathbf{t}, \mathbf{z}}^{(m)} \cap B_j) = r_j, j = 1, \dots, k, \right. \\ & \quad \left. \Pi(B_j \setminus (K_{\mathbf{t}, \mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}})) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right) \left| \right. \end{aligned}$$

$$\leq \lim_{m \rightarrow \infty} \mathbb{P} \left(\Pi(S \cap K_{(\mathbf{t}, \mathbf{z})}^{(m)}) - \Pi_2(S) > 0 \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) = 0.$$

by the second part of Lemma 5.4. This verifies the assertion. \square

These results enable us to show the independence of Π_1 , Π_2 and Π_3 conditional on $Z(\mathbf{t})$ and to calculate the conditional distribution of Π_3 . The results can also be found in [Dombry and Eyi-Minko \(2011\)](#), Thm. 3.1, obtained by a different approach.

Theorem 5.7. 1. *With probability 1, the point processes Π_1 , Π_2 and Π_3 conditional on $Z(\mathbf{t})$ are stochastically independent.*

2. *The process $\Pi_3 \mid Z(\mathbf{t}) = \mathbf{z}$ has the same distribution as $\Pi \setminus (K_{\mathbf{t}, \mathbf{z}} \cup \overline{K_{\mathbf{t}, \mathbf{z}}})$ with probability 1.*

Proof. 1. Since $\Pi_1 \equiv 0$ a.s., we only have to show the independence of Π_2 and Π_3 conditional on $Z(\mathbf{t})$. Let $B_1, \dots, B_k, B_{k+1}, \dots, B_l \in \mathcal{B}^d \times ((0, \infty) \cap \mathcal{B}) \times \mathcal{G}$ and $r_1, \dots, r_l \in \mathbb{N}_0$. The second part of Theorem 5.6 yields

$$\begin{aligned} & \mathbb{P}(\Pi_2(B_j) = r_j, j = 1, \dots, k, \Pi_3(B_j) = r_j, j = k + 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cap B_j) = r_j, j = 1, \dots, k, \\ & \quad \Pi(B_j \setminus (K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}})) = r_j, j = k + 1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})). \end{aligned}$$

By (5.4) the process $\Pi(\cdot \cap S \setminus (K_{\mathbf{t}, \mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t}, \mathbf{z}}^{(m)}}))$ is independent of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$. Hence,

$$\begin{aligned} & \mathbb{P}(\Pi_2(B_j) = r_j, j = 1, \dots, k, \Pi_3(B_j) = r_j, j = k + 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cap B_j) = r_j, j = 1, \dots, k \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ & \quad \cdot \mathbb{P}(\Pi(B_j \setminus (K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}})) = r_j, j = k + 1, \dots, l) \\ &= \mathbb{P}(\Pi_2(B_j) = r_j, j = 1, \dots, k \mid Z(\mathbf{t}) = \mathbf{z}) \\ & \quad \cdot \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(B_j \setminus (K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}})) = r_j, j = k + 1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \mathbb{P}(\Pi_2(B_j) = r_j, j = 1, \dots, k \mid Z(\mathbf{t}) = \mathbf{z}) \mathbb{P}(\Pi_3(B_j) = r_j, j = k + 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}), \end{aligned}$$

where we use the same arguments as before.

2. For any sets $B_1, \dots, B_l \in \mathcal{B}^d \times (\mathcal{B} \cap (0, \infty)) \times \mathcal{G}$ and $r_1, \dots, r_l \in \mathbb{N}_0$, the second part of Theorem 5.6 implies

$$\begin{aligned} & \mathbb{P}(\Pi_3(B_j) \leq r_j, j = 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(B_j \setminus (K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}})) \leq r_j, j = 1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi(B_j \setminus (K_{(\mathbf{t}, \mathbf{z})}^{(m)} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}})) \leq r_j, j = 1, \dots, l) \\ &= \mathbb{P}(\Pi(B_j \setminus (K_{\mathbf{t}, \mathbf{z}} \cup \overline{K_{\mathbf{t}, \mathbf{z}}})) \leq r_j, j = 1, \dots, l). \end{aligned}$$

Here, we use that the process $\Pi(\cdot \cap S \setminus (K_{\mathbf{t},\mathbf{z}}^{(m)} \cup \overline{K_{\mathbf{t},\mathbf{z}}^{(m)}}))$ is independent of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$ and the fact that we have a monotone limit. \square

By the second part of Theorem 5.7 the process $\Pi_3 \mid Z(\mathbf{t}) = \mathbf{z}$ can be easily simulated by unconditionally simulating Π and restricting it to $\mathbb{R}^d \times (0, \infty) \times G \setminus (K_{\mathbf{t},\mathbf{z}} \cup \overline{K_{\mathbf{t},\mathbf{z}}})$.

We close this section by noting that there exists a more general version of Theorem 5.6 which we will need for simulation. Let $B_1, \dots, B_k, \in \mathcal{B}^d \times ((0, \infty) \cap \mathcal{B}) \times \mathcal{G}$ be pairwise disjoint with $\bigcup_{j=1}^k B_j = S$. We introduce *generalized “blurred” scenarios*

$$\begin{aligned} E_{\{n_A^{(j)}\}}^{(m)}(\mathbf{z}) &= E_{\{n_A^{(j)}\}_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset, j=1, \dots, k}}^{(m)}(\mathbf{z}) \\ &= \{\Pi(I_A^{(m)}(\mathbf{z}) \cap B_j) = n_A^{(j)}, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, j = 1, \dots, k, \Pi(\overline{K_{(\mathbf{t}, \mathbf{z})}^{(m)}}) = 0\} \end{aligned}$$

with $n_A^{(j)} \in \mathbb{N}_0$ such that $\sum_{j=1}^k \sum_{A: i \in A} n_A^{(j)} \geq 1$ for $i \in \{1, \dots, n\}$.

Analogously, *generalized scenarios* $E_{\{n_A^{(j)}\}}(\mathbf{z})$ are defined. Then, in exactly the same way as Theorem 5.6, the following theorem can be shown.

Theorem 5.8. *With probability 1 we have*

$$\mathbb{P}(E_{\{n_A^{(j)}\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A^{(j)}\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$$

for any scenario $E_{\{n_A^{(j)}\}}(\mathbf{z})$ with $\sum_{j=1}^k \sum_{A: i \in A} n_A^{(j)} \geq 1$ for all $i \in \{1, \dots, n\}$.

The remainder of the chapter will address the problem of simulating $\Pi_2 \mid Z(\mathbf{t}) = z$. We propose a procedure consisting of two steps. First, we draw a scenario $E_{\{n_A\}}(Z(\mathbf{t}))$ conditional on $Z(\mathbf{t}) = \mathbf{z}$. Then, the points of Π_2 corresponding to this scenario are simulated.

5.3 Calculations in the Case of a Finite Number of Shape Functions

As shown in Section 5.2, all we need for calculating $\mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z})$ is the exact asymptotic behaviour of $\Lambda(I_A^{(m)}(\mathbf{z}))$. In particular, we have to analyse the behaviour of the intersection of two curves $K_{t_i, z_i + \delta_i} \cap K_{t_j, z_j + \delta_j}$ for $|\delta_i|, |\delta_j|$ small. Explicit calculations turn out to be quite involved. Therefore, we restrict ourselves to the case $d = 1$. Furthermore, the intersection depends on the derivative of the shape function. We calculate the asymptotics of $\Lambda(I_A^{(m)}(\mathbf{z}))$ for $|A| = 1$ (Proposition 5.13), $|A| = 2$ (Proposition 5.9) and $|A| \geq 3$ (Proposition 5.11), see Figure 5.2. In the latter case, the rate of convergence of $\Lambda(I_A^{(m)})$ cannot be determined exactly. Nevertheless, the conditional probability of any scenario can be calculated (Theorem 5.16).

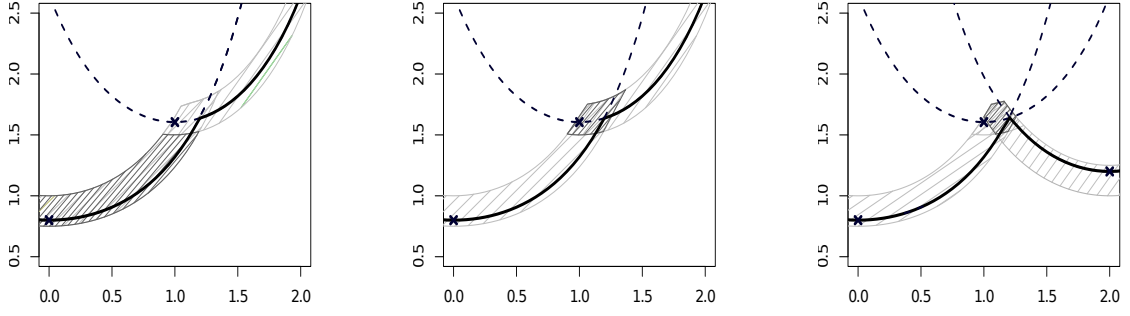


Figure 5.2: Blurred intersection sets for $|A| = 1$, $|A| = 2$ and $|A| = 3$. Black crosses: data $(t_i, Z(t_i))$, dashed black lines: $K_{t_i, Z(t_i)}$, black line: $K_{t, Z(t)}$, grey area: $K_{t, Z(t)}^{(m)}$, black area: $I_A^{(m)}(Z(t))$ with $|A| = 1$ (left), $|A| = 2$ (middle) and $|A| = 3$ (right).

First we assume that G is a finite space of functions $f : \mathbb{R} \rightarrow (0, \infty)$ such that the intersections

$$M_{c, t_0} = \{t \in \mathbb{R} : f(t) = cf(t_0 + t)\} \quad (5.10)$$

are finite for all $c > 0$, $t_0 \in \mathbb{R}$, $f \in G$. This implies that each set $I_A(\mathbf{z})$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, $|A| \geq 2$, $\mathbf{z} > \mathbf{0}$, is finite. W.l.o.g. we assume that $\mathbb{P}_F(\{f\}) > 0$ for all $f \in G$.

Proposition 5.9. *Let $t_1, t_2 \in \mathbb{R}$, $z_1, z_2 > 0$ such that*

$$I_{\{1, 2\}}(\mathbf{z}) = \{(t_0, y_0, f)\}.$$

Furthermore, let f be continuously differentiable in a neighbourhood of $t_i - t_0$ and $t_2 - t_0$ with

$$z_1 f'(t_2 - t_0) \neq z_2 f'(t_1 - t_0). \quad (5.11)$$

Then, we have

$$\Lambda(I_{\{1, 2\}}^{(m)}(\mathbf{z})) = \frac{2^{-2m}}{y_0^2 |z_1 f'(t_2 - t_0) - z_2 f'(t_1 - t_0)|} \mathbb{P}_F(\{f\}) + o(2^{-2m}).$$

Proof. We note that (t_0, y_0, f) satisfies the equation

$$\frac{f(t_1 - t_0)}{z_1} = \frac{f(t_2 - t_0)}{z_2} = y_0^{-1}.$$

Let

$$H : (-\mathbf{z}, \infty) \times \mathbb{R} \rightarrow \mathbb{R}, \quad (\delta, t) \mapsto \frac{f(t_1 - t)}{z_1 + \delta_1} - \frac{f(t_2 - t)}{z_2 + \delta_2}.$$

Then, $H(\mathbf{0}, t_0) = 0$ and

$$\frac{\partial H}{\partial t}(\mathbf{0}, t_0) = -\frac{f'(t_1 - t_0)}{z_1} + \frac{f'(t_2 - t_0)}{z_2} \neq 0$$

due to (5.11). The implicit function theorem yields the existence of a neighbourhood V of $\mathbf{0}$ and a continuously differentiable function $h : V \rightarrow \mathbb{R}$ such that $H(\delta, h(\delta)) = 0$. Using the notation $(t_\delta, y_\delta, f) = I_{\{1,2\}}(\mathbf{z} + \delta)$ we get $h(\delta) = t_\delta$ and the equality

$$\frac{f(t_1 - t_\delta)}{z_1 + \delta_1} = \frac{f(t_2 - t_\delta)}{z_2 + \delta_2} = y_\delta^{-1}.$$

As h is continuously differentiable, we obtain $t_0 - t_\delta \in \mathcal{O}(\|\delta\|)$, and a Taylor expansion of f yields

$$f(t_i - t_\delta) = f(t_i - t_0) - f'(t_i - t_0) \cdot (t_\delta - t_0) + o(\|\delta\|). \quad (5.12)$$

Let $g(t) = f(t - t_0)$. Thus, using (5.12), t_δ is given implicitly by

$$\frac{g(t_1) - g'(t_1) \cdot (t_\delta - t_0)}{z_1 + \delta_1} = \frac{g(t_2) - g'(t_2) \cdot (t_\delta - t_0)}{z_2 + \delta_2} + o(\|\delta\|),$$

which implies the explicit representation

$$t_\delta = t_0 + \frac{\delta_1 g(t_2) - \delta_2 g(t_1)}{z_1 g'(t_2) - z_2 g'(t_1)} + o(\|\delta\|). \quad (5.13)$$

Plugging in (5.13) into (5.12) yields

$$y_\delta^{-1} = \frac{f(t_1 - t_\delta)}{z_1 + \delta_1} = \frac{g(t_1)}{z_1 + \delta_1} - \frac{g'(t_1)}{z_1} \cdot \frac{\delta_1 g(t_2) - \delta_2 g(t_1)}{z_1 g'(t_2) - z_2 g'(t_1)} + o(\|\delta\|).$$

As f and $\delta \mapsto t_\delta = h(\delta)$ are C^1 -functions, all the terms $o(\|\delta\|)$ are continuously differentiable for small $\|\delta\|$. Therefore, the mapping

$$\Phi : V \rightarrow \mathbb{R} \times (0, \infty), \quad \delta \mapsto (t_\delta, y_\delta^{-1})$$

is continuously differentiable near the origin. Calculating the partial derivatives explicitly, we obtain

$$\det(D\Phi(\delta)) = -\frac{g^2(t_1)}{z_1^2 \cdot (z_1 g'(t_2) - z_2 g'(t_1))} + o(1). \quad (5.14)$$

As $\det(D\Phi(\mathbf{0})) \neq 0$, the inverse function theorem allows to regard Φ as a diffeomorphism restricted to an appropriate neighbourhood of $\mathbf{0}$. Thus, considering the transformed Poisson point process $\tilde{\Pi} = \sum_{i \in \mathbb{N}} \delta_{(S_i, U_i^{-1})}$ on $\mathbb{R} \times (0, \infty)$ whose intensity measure is the Lebesgue measure, we get

$$\begin{aligned} & \Lambda(\{I_{\{1,2\}}(\mathbf{z} + \delta), z_i + \delta_i \in A_m(z_i), i = 1, 2\}) \\ &= \int_{\Phi((A_m(z_1) - z_1) \times (A_m(z_2) - z_2))} \mathbb{P}_F(\{f\}) \, d(t, y) \\ &= \int_{(A_m(z_1) - z_1) \times (A_m(z_2) - z_2)} |\det(D\Phi(\delta))| \cdot \mathbb{P}_F(\{f\}) \, d\delta \\ &= \int_{(A_m(z_1) - z_1) \times (A_m(z_2) - z_2)} \frac{1/y_0^2 + o(1)}{|z_1 g'(t_2) - z_2 g'(t_1)|} \mathbb{P}_F(\{f\}) \, d\delta. \end{aligned}$$

We note that the term $o(1)$ is continuous w.r.t. δ and therefore the integrands can be locally bounded by

$$\left(\frac{1/y_0^2}{|z_1 g'(t_2) - z_2 g'(t_1)|} - \varepsilon_m \right) \mathbb{P}_F(\{f\})$$

from below, and by

$$\left(\frac{1/y_0^2}{|z_1 g'(t_2) - z_2 g'(t_1)|} + \varepsilon_m \right) \mathbb{P}_F(\{f\})$$

from above for all $(\delta_1, \delta_2) \in (A_m(z_1) - z_1) \times (A_m(z_2) - z_2)$ with m large enough and an appropriate sequence $(\varepsilon_m)_{m \in \mathbb{N}}$ with $\varepsilon_m \searrow 0$. This implies that the integral has the form

$$2^{-2m} \cdot \frac{1/y_0^2}{|z_1 g'(t_2) - z_2 g'(t_1)|} \cdot \mathbb{P}_F(\{f\}) + o(2^{-2m})$$

which is the desired result. \square

Remark 5.10. 1. Using $\frac{z_1}{f(t_1-t_0)} = \frac{z_2}{f(t_2-t_0)} = y_0$ we get that the equality

$$z_1 f'(t_2 - t_0) = z_2 f'(t_1 - t_0)$$

holds if and only if

$$\left. \frac{\partial}{\partial t} \frac{z_1}{f(t_1 - t)} \right|_{t=t_0} = \left. \frac{\partial}{\partial t} \frac{z_2}{f(t_2 - t)} \right|_{t=t_0},$$

i.e. if and only if the two sets of admissible points, K_{t_1, z_1} and K_{t_2, z_2} , are tangents to each other in (t_0, y_0, f) which is an event of probability zero by Assumption (5.10). Therefore, (5.11) is satisfied a.s.

2. If $I_{\{1,2\}}(\mathbf{z})$ consists of a finite number of points,

$$I_{\{1,2\}}(z_1, z_2) = \{(t_0^{(1)}, y_0^{(1)}, f_1), \dots, (t_0^{(k)}, y_0^{(k)}, f_k)\},$$

we get

$$\Lambda(I_{\{1,2\}}(\mathbf{z})) = 2^{-2m} \cdot \sum_{j=1}^k \frac{\mathbb{P}(\{f_j\})}{y_0^2 \cdot |z_1 f_j'(t_2 - t_0^{(j)}) - z_2 f_j'(t_1 - t_0^{(j)})|} + o(2^{-2m}).$$

Proposition 5.11. Let $t_1, \dots, t_l \in \mathbb{R}$, $z_1, \dots, z_l > 0$, $l \geq 3$ such that

$$I_{\{1, \dots, l\}}(\mathbf{z}) = \{(t_0, y_0, f)\}.$$

Let $f : \mathbb{R} \rightarrow (0, \infty)$ be continuously differentiable in a neighbourhood of $t_1 - t_0, \dots, t_l - t_0$ with (5.11). Then, we have

$$\Lambda(I_{\{1, \dots, l\}}^{(m)}(\mathbf{z})) \leq \frac{2^{-2m}}{y_0^2 |z_1 f'(t_2 - t_0) - z_2 f'(t_1 - t_0)|} \mathbb{P}_F(\{f\}) + o(2^{-2m}).$$

For any $C > 0$, $\varepsilon > 0$, there exists $m_{C, \varepsilon} \in \mathbb{N}$ such that

$$\Lambda(I_{\{1, \dots, l\}}^{(m)}(\mathbf{z})) \geq C 2^{-2m(1+\varepsilon)}$$

for all $m \geq m_{C, \varepsilon}$.

Proof. The first assertion follows immediately from Proposition 5.9 by the fact that

$$\bigcap_{i=1}^l K_{t_i, z_i}^{(m)} \subset K_{t_1, z_1}^{(m)} \cap K_{t_2, z_2}^{(m)}.$$

In order to verify the second assertion, we recall results from the proof of Proposition 5.9: we showed the existence of a C^1 -function $(\delta_1, \delta_2) \mapsto t_{\delta_1 \delta_2}$ defined in an appropriate neighbourhood V of $(0, 0)$ such that

$$\frac{f(t_1 - t_{\delta_1 \delta_2})}{z_1 + \delta_1} = \frac{f(t_2 - t_{\delta_1 \delta_2})}{z_2 + \delta_2}.$$

Now, we consider the C^1 -functions

$$H_i : V \times (-z_i, \infty) \rightarrow \mathbb{R}, \quad (\delta_1, \delta_2, \delta_i) \mapsto \frac{f(t_1 - t_{\delta_1 \delta_2})}{z_1 + \delta_1} - \frac{f(t_i - t_{\delta_1 \delta_2})}{z_i + \delta_i}, \quad i \in \{3, \dots, l\}.$$

As $H_i(0, 0, 0) = 0$ and $\frac{\partial H_i}{\partial \delta_i}(0, 0, 0) = \frac{f(t_i - t_0)}{z_i^2} \neq 0$, we get the existence of a continuously differentiable function h_i defined on a neighbourhood of $(0, 0)$ such that

$$\frac{f(t_1 - t_{\delta_1 \delta_2})}{z_1 + \delta_1} = \frac{f(t_i - t_{\delta_1 \delta_2})}{z_i + h_i(\delta_1, \delta_2)}. \quad (5.15)$$

Using Taylor expansions of $g(\cdot) = f(\cdot - t_0)$ of first order, employing Equation (5.13), and solving Equation (5.15) yields

$$h_i(\delta_1, \delta_2) = \frac{g(t_i)}{g(t_1)} \delta_1 + \frac{z_i g'(t_1) - z_1 g'(t_i)}{g(t_1)} \frac{g(t_2) \delta_1 - g(t_1) \delta_2}{z_1 g'(t_2) - z_2 g'(t_1)} + o(|\delta_1|) + o(|\delta_2|). \quad (5.16)$$

So, there are constants $c_{1,i}, c_{2,i}$ such that $h_i(\delta_1, \delta_2) = c_{1,i} \delta_1 + c_{2,i} \delta_2 + o(|\delta_1|) + o(|\delta_2|)$. Let $A_m^{(i)} = A_m(z_i) - z_i$ for $i \in \{1, \dots, n\}$. We are interested in those pairs $(\delta_1, \delta_2) \in A_m^{(1)} \times A_m^{(2)}$ with $h_i(\delta_1, \delta_2) \in A_m^{(i)}$. By Lemma 5.3, for any $C' > 0$, $\varepsilon > 0$, we have that $(-C' 2^{-m(1+\varepsilon)}, C' 2^{-m(1+\varepsilon)}) \in A_m^{(i)}$, $i = 1, \dots, n$, for m large enough. Therefore, $h_i(\delta_1, \delta_2) \in A_m^{(i)}$ is guaranteed for $|\delta_1| < \frac{C' 2^{-m(1+\varepsilon)}}{3|c_{1,i}|}$ and $|\delta_2| < \frac{C' 2^{-m(1+\varepsilon)}}{3|c_{2,i}|}$ if m is sufficiently large.

By the same argumentation for all $i \in \{3, \dots, l\}$, we get that the existence of all $h_i(\delta_1, \delta_2)$ is ensured for

$$|\delta_1| < \frac{C' 2^{-m(1+\varepsilon)}}{3 \max_{i=3, \dots, l} |c_{1,i}|}, \quad |\delta_2| < \frac{C' 2^{-m(1+\varepsilon)}}{3 \max_{i=3, \dots, l} |c_{2,i}|} \quad (5.17)$$

for m large enough. Furthermore, to ensure $\delta_1 \in A_m^{(1)}$, $\delta_2 \in A_m^{(2)}$, we have to add the conditions $|\delta_1|, |\delta_2| < C' 2^{-m(1+\varepsilon)}$. With $C_j = \max\{1, 3 \max_{i=3, \dots, l} |c_{j,i}|\}$ for $j = 1, 2$, this yields

$$\begin{aligned} & \Lambda(\{I_{\{1, \dots, l\}}(\mathbf{z} + \delta), \delta_i \in A_m^{(i)}, i = 1, \dots, l\}) \\ & \geq \Lambda\left(\left\{\left(t_{\delta_1 \delta_2}, \frac{z_1 + \delta_1}{f(t_1 - t_{\delta_1 \delta_2})}, f\right), |\delta_j| < \frac{C'}{C_j} 2^{-m(1+\varepsilon)}, j = 1, 2\right\}\right) \\ & = \frac{(C')^2 2^{-2m(1+\varepsilon)} \cdot \mathbb{P}_F(\{f\})}{y_0^2 C_1 C_2 |z_1 f'(t_2 - t_0) - z_2 f'(t_1 - t_0)|} + o(2^{-2m(1+\varepsilon)}) \end{aligned}$$

where we use the same argumentation as in the proof of Proposition 5.9. This completes the proof. \square

Remark 5.12. Note that Proposition 5.11 does not provide exact asymptotics which we need to determine $\mathbb{P}(\Pi(\{(t_0, y_0, f)\}) = 1 \mid \Pi(I_{\{1, \dots, l\}}(\mathbf{z})) = 1, Z(\mathbf{t}) = \mathbf{z})$ for $(t_0, y_0, f) \in I_{\{1, \dots, l\}}(\mathbf{z})$ and $|I_{\{1, \dots, l\}}(\mathbf{z})| > 1$. However, we can get exact results by conditioning on $Z(t_i)$ being in intervals of different size for each $i \in \{1, \dots, l\}$ instead of $Z(t_i) \in A_m(z_i)$ for all $i = 1, \dots, l$. We will choose these intervals such that some restrictions on the intersection sets vanish asymptotically and we can resort to the results on the intersection of two curves.

The calculations in the proof of Proposition 5.11 yield

$$|h_i(\delta_1, \delta_2)| \leq (|c_{1,i}| + o(1)) \cdot |\delta_1| + (|c_{2,i}| + o(1)) \cdot |\delta_2| \leq 2^{-m}(|c_{1,i}| + |c_{2,i}| + o(1))$$

for $(\delta_1, \delta_2) \in A_m^{(1)} \times A_m^{(2)}$. Thus, for any $\varepsilon > 0$, we can replace m by $\lfloor m(1-\varepsilon) \rfloor$ in Equation (5.17) and get that $h_i(\delta_1, \delta_2) \in A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}$ holds for $|\delta_1| < \frac{C'2^{-\lfloor m(1-\varepsilon) \rfloor(1+\varepsilon)}}{3|c_{1,i}|} \sim 2^{\varepsilon^2 m} 2^{-m}$ and $|\delta_2| < \frac{C'2^{-\lfloor m(1-\varepsilon) \rfloor(1+\varepsilon)}}{3|c_{2,i}|} \sim 2^{\varepsilon^2 m} 2^{-m}$ for m large enough. Therefore,

$$h_i(\delta_1, \delta_2) \subset A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}, \quad i = 3, \dots, l$$

for all $\delta_1 \in A_m^{(1)} \subset (-2^{-m}, 2^{-m}]$, $\delta_2 \in A_m^{(2)} \subset (-2^{-m}, 2^{-m}]$ if m is sufficiently large. This implies

$$\begin{aligned} & \left\{ I_{\{1, \dots, l\}}(\mathbf{z} + \delta), \delta_1 \in A_m^{(1)}, \delta_2 \in A_m^{(2)}, \delta_i \in A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}, i = 3, \dots, l \right\} \\ &= \left\{ I_{\{1, \dots, l\}}(\mathbf{z} + \delta), (\delta_1, \delta_2) \in A_m^{(1)} \times A_m^{(2)}, h_i(\delta_1, \delta_2) \in A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}, i = 3, \dots, l \right\} \\ &= \left\{ \left(t_{\delta_1 \delta_2}, \frac{z_1 + \delta_1}{f(t_1 - t_{\delta_1 \delta_2})}, f \right), \delta_1 \in A_m^{(1)}, \delta_2 \in A_m^{(2)} \right\} \end{aligned}$$

and, therefore

$$\begin{aligned} & \Lambda \left(\left\{ I_{\{1, \dots, l\}}^{(m)}(\mathbf{z} + \delta), \delta_1 \in A_m^{(1)}, \delta_2 \in A_m^{(2)}, \delta_i \in A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}, i = 3, \dots, l \right\} \right) \\ &= \frac{2^{-2m} \mathbb{P}_F(\{f\})}{y_0^2 |z_1 f'(t_2 - t_0) - z_2 f'(t_1 - t_0)|} + o(2^{-2m}). \end{aligned}$$

By conditioning on $Z(\mathbf{t}) \in A_m(z_1) \times A_m(z_2) \times \times_{i=3, \dots, l} A_{\lfloor m(1-\varepsilon) \rfloor}(z_i)$, for $I_{\{1, \dots, l\}}(\mathbf{z}) = \{(t_0^{(1)}, y_0^{(1)}, f_1), \dots, (t_0^{(k)}, y_0^{(k)}, f_k)\}$ for some $l \geq 3$, we can apply Lévy's "Upward" Theorem (Rogers and Williams, 2000, Thm. 50.3) and end up with

$$\begin{aligned} & \mathbb{P}(\Pi(\{(t_0, y_0, f)\}) = 1 \mid \Pi(I_{\{1, \dots, l\}}(Z(\mathbf{t}))) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P} \left(\Pi \left(\left\{ (t_{\delta_1 \delta_2}, y_\delta, f) : \delta_1 \in A_m^{(1)}, \delta_2 \in A_m^{(2)}, \delta_i \in A_{\lfloor m(1-\varepsilon) \rfloor}^{(i)}, i \geq 3 \right\} \right) = 1 \mid \right. \\ & \quad \left. \Pi(I_{\{1, \dots, l\}}(Z(\mathbf{t}))) = 1, Z(\mathbf{t}) \in A_m(z_1) \times A_m(z_2) \times \times_{i=3, \dots, l} A_{\lfloor m(1-\varepsilon) \rfloor}(z_i) \right) \\ &= \frac{\mathbb{P}_F(\{f\})}{y_0^2 |z_1 f'(t_2 - t_0) - z_2 f'(t_1 - t_0)|} \cdot \left(\sum_{j=1}^k \frac{\mathbb{P}_F(\{f_j\})}{(y_0^{(j)})^2 |z_1 f_j'(t_2 - t_0^{(j)}) - z_2 f_j'(t_1 - t_0^{(j)})|} \right)^{-1}. \end{aligned} \tag{5.18}$$

Note that Lévy's Upward Theorem implies that, with probability one, the right hand side of (5.18) does not depend on the choice of the labelling.

Thus, despite the lack of exact convergence rates of $\Lambda(I_{\{1,\dots,l\}}^{(m)}(\mathbf{z}))$, the distribution of $\Pi(\cdot \cap I_{\{1,\dots,l\}}(Z(\mathbf{t}))) \mid \Pi(I_{\{1,\dots,l\}}(Z(\mathbf{t}))) = 1$ can be determined exactly.

Proposition 5.13. *Let $f \in G$, $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{z} > \mathbf{0}$ such that f is continuously differentiable in a neighbourhood of $t_i - t_0$ for all $(t_0, y_0, f) \in K_{\mathbf{t}, \mathbf{z}} \cap \bigcup_{\substack{l=1 \\ l \neq i}}^n (K_{t_i, z_i} \cap K_{t_l, z_l})$, i.e. all $(t_0, y_0, f) \in K_{t_i, z_i}$ involved in an intersection, for all $i \in \{1, \dots, n\}$. We denote the projection of the set $I_{\{i\}}(\mathbf{z}) \cap (\mathbb{R} \times (0, \infty) \times \{f\})$ onto its first component in \mathbb{R} by*

$$D_i^{(f)} = \{t \in \mathbb{R} : (t, y, f) \in I_{\{i\}}(\mathbf{z}) \text{ for some } y > 0\}.$$

Then, we have

$$\Lambda\left(I_{\{i\}}^{(m)}(\mathbf{z}) \cap (\mathbb{R} \times (0, \infty) \times \{f\})\right) = 2^{-m} \cdot \mathbb{P}_F(\{f\}) \cdot \int_{D_i^{(f)}} \frac{f(t_i - t)}{z_i^2} dt + o(2^{-m}). \quad (5.19)$$

Proof. First, we note that by renumbering, it suffices to show the result for $i = 1$. The idea of this proof is to assess the set $D_1^{(f)}$ by the sets $D_{1,\min}^{(m)}$ from below and $D_{1,\max}^{(m)}$ from above. Here, $D_{1,\min}^{(m)}$ consists of all first components of $I_{\{1\}}^{(m)}(\mathbf{z})$ which are not part of any intersections $I_A^{(m)}(\mathbf{z})$, $A \supseteq \{1\}$ and $D_{1,\max}^{(m)}$ is the set of the first components of $\bigcup_{A \supseteq \{1\}} I_A^{(m)}(\mathbf{z})$. Analogously, $\Lambda\left(I_{\{i\}}^{(m)}(\mathbf{z}) \cap (\mathbb{R} \times (0, \infty) \times \{f\})\right)$ can be bounded from below and above by replacing $D_1^{(f)}$ in (5.19) by the sets $D_{1,\min}^{(m)}$ and $D_{1,\max}^{(m)}$, respectively. We will show that the difference, which consists of blurred intersections $I_A^{(m)}(\mathbf{z})$, $A \supseteq \{1\}$, vanishes asymptotically.

Let $A_m^{(i)} = A_m(z_i) - z_i$. Then, for any $\delta \in \times_{i=1}^n (-z_i, \infty)$ we define

$$\begin{aligned} D_{1,\delta}^{(f)} &= \left\{ t \in \mathbb{R} : \left(t, \frac{z_1 + \delta_1}{f(t_1 - t)}, f \right) \in I_{\{1\}}(\mathbf{z} + \delta) \right\} \\ &= \left\{ t \in \mathbb{R} : \frac{z_1 + \delta_1}{f(t_1 - t)} < \min_{i=2,\dots,n} \frac{z_i + \delta_i}{f(t_i - t)} \right\}. \end{aligned} \quad (5.20)$$

Thus, defining $D_{1,\min}^{(m)} = \bigcap_{\delta \in \times_{i=1}^n A_m^{(i)}} D_{1,\delta}^{(f)}$ and $D_{1,\max}^{(m)} = \bigcup_{\delta \in \times_{i=1}^n A_m^{(i)}} D_{1,\delta}^{(f)}$, we get

$$D_{1,\min}^{(m)} \subset D_1^{(f)} \subset D_{1,\max}^{(m)}.$$

On the other hand, we have

$$\begin{aligned} &\left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in D_{1,\min}^{(f)}, yf(t_1 - t) \in A_m(z_1) \right\} \\ &\subset I_{\{1\}}^{(m)}(\mathbf{z}) \subset \left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in D_{1,\max}^{(f)}, yf(t_1 - t) \in A_m(z_1) \right\}. \end{aligned} \quad (5.21)$$

Now, let $t \in D_{1,\max}^{(m)} \setminus D_{1,\min}^{(m)}$. Then, by definition of $D_{1,\min}^{(f)}$ and $D_{1,\max}^{(f)}$, there exist $\delta^{(1)}, \delta^{(2)} \in \times_{i=1}^n A_m^{(i)}$ such that $t \in D_{1,\delta^{(1)}}^{(f)}$, but $t \notin D_{1,\delta^{(2)}}^{(f)}$. That is, by Equation (5.20),

$$\frac{z_1 + \delta_1^{(1)}}{f(t_1 - t)} < \min_{i=2,\dots,n} \frac{z_i + \delta_i^{(1)}}{f(t_i - t)} \quad \text{and} \quad \frac{z_1 + \delta_1^{(2)}}{f(t_1 - t)} \geq \min_{i=2,\dots,n} \frac{z_i + \delta_i^{(2)}}{f(t_i - t)}.$$

By continuity arguments, a $\delta \in \times_{i=1}^n A_m^{(i)}$ exists such that

$$\frac{z_1 + \delta_1}{f(t_1 - t)} = \min_{i=2, \dots, n} \frac{z_i + \delta_i}{f(t_i - t)},$$

i.e. $t \in T_1^{(m)} = \{t \in \mathbb{R} : (t, y, f) \in \bigcup_{A: \{1\} \subsetneq A} I_A^{(m)}(\mathbf{z}) \text{ for some } y > 0\}$. Thus,

$$D_{1, \max}^{(m)} \setminus D_{1, \min}^{(m)} \subset T_1^{(m)}. \quad (5.22)$$

By definition, $T_1^{(m)}$ denotes the set of first components involved in any blurred intersection and we have $T_1^{(m)} \searrow T_1 = \{t \in \mathbb{R} : (t, y, f) \in \bigcup_{A: \{1\} \subsetneq A} I_A(\mathbf{z}) \text{ for some } y > 0\}$ as $m \rightarrow \infty$ and T_1 is finite by Assumption (5.10). Therefore, dominated convergence yields

$$\int_{T_1^{(m)}} f(t_1 - t) dt \searrow 0, \quad m \rightarrow \infty. \quad (5.23)$$

Thus, by Equations (5.21) and (5.22) we get

$$\begin{aligned} & \Lambda \left(I_{\{1\}}^{(m)}(\mathbf{z}) \Delta \left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in D_1^{(f)}, yf(t_1 - t) \in A_m(z_1) \right\} \right) \\ & \leq \Lambda \left(\left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in D_{1, \max}^{(f)}, yf(t_1 - t) \in A_m(z_1) \right\} \setminus \right. \\ & \quad \left. \left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in D_{1, \min}^{(f)}, yf(t_1 - t) \in A_m(z_1) \right\} \right) \\ & \leq \Lambda \left(\left\{ (t, y, f) \in \mathbb{R} \times (0, \infty) \times \{f\} : t \in T_1^{(m)}, yf(t_1 - t) \in A_m(z_1) \right\} \right) \\ & = \mathbb{P}_F(\{f\}) \int_{A_m^{(1)}} \int_{T_1^{(m)}} \frac{f(t_1 - t)}{(z_1 + \delta_1)^2} dt d\delta_1 = \mathbb{P}_F(\{f\}) \int_{A_m^{(1)}} \frac{o(1)}{(z_1 + \delta_1)^2} d\delta_1 \in o(2^{-m}). \end{aligned}$$

The last equality follows from Equation (5.23). Hence, we have

$$\begin{aligned} \Lambda(I_{\{1\}}^{(m)}(\mathbf{z})) &= \mathbb{P}_F(\{f\}) \int_{D_1^{(f)}} \int_{A_m^{(1)}} \frac{f(t_1 - t)}{(z_1 + \delta_1)^2} d\delta_1 dt + o(2^{-m}) \\ &= 2^{-m} \cdot \mathbb{P}_F(\{f\}) \cdot \left(\int_{D_1^{(f)}} \frac{f(t_1 - t)}{z_1^2} dt + o(1) \right) \end{aligned}$$

which completes the proof. \square

Now, we can use Theorem 5.6 in order to compute the conditional probabilities

$$\begin{aligned} \mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) &= \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \lim_{m \rightarrow \infty} \frac{\mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))}{\sum_{\{\tilde{n}_A\} \in N_0} \mathbb{P}(E_{\{\tilde{n}_A\}}^{(m)}(\mathbf{z}))} \end{aligned} \quad (5.24)$$

where $N_0 = \{\{n_A : A \in 2^{\{1, \dots, n\}} \setminus \emptyset\} : \sum_{A: i \in A} n_A = 1, i = 1, \dots, n\}$.

As the sets $I_A^{(m)}(\mathbf{z})$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, are pairwise disjoint and the sequences $\Lambda(I_A^{(m)}(\mathbf{z}))$ tend to zero for $m \rightarrow \infty$ by Lemma 5.2, we get

$$\mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z})) \sim \exp(-\Lambda(\overline{K}_{(t, \mathbf{z})})) \prod_{A: n_A=1} \Lambda(I_A(\mathbf{z})). \quad (5.25)$$

Considering (5.24), we can restrict ourselves to those scenarios with the slowest rate of convergence to zero. Propositions 5.9, 5.11 and 5.13 yield that scenarios involving intersections of at least three sets are always of a dominating order. Therefore, the unknown terms from Proposition 5.11 are cancelled out.

Example 5.14. Let $F(x) = f(x) =: \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$ a.s. Furthermore, let $Z(t_1) = Z(t_2) = a$ where $t_1 = -b$ and $t_2 = b$ for constants $a, b > 0$. Then, we get

$$\begin{aligned} I_{\{1\}}(a, a) &= \left\{ \left(t, a\sqrt{2\pi} \exp\left(\frac{(-b-t)^2}{2}\right), f \right) : t < 0 \right\} \\ I_{\{2\}}(a, a) &= \left\{ \left(t, a\sqrt{2\pi} \exp\left(\frac{(b-t)^2}{2}\right), f \right) : t > 0 \right\} \\ I_{\{1,2\}}(a, a) &= \left\{ \left(0, a\sqrt{2\pi} \exp\left(\frac{b^2}{2}\right), f \right) \right\}. \end{aligned}$$

By the formulae from Propositions 5.9 and 5.13, we get

$$\Lambda(I_{\{1\}}^{(m)}(a, a)) = \frac{2^{-m}}{a^2} \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(-b-t)^2}{2}\right) dt + o(2^{-m}) \sim \frac{2^{-m}}{a^2} \Phi(b) \quad (5.26)$$

$$\Lambda(I_{\{2\}}^{(m)}(a, a)) = \frac{2^{-m}}{a^2} \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(b-t)^2}{2}\right) dt + o(2^{-m}) \sim \frac{2^{-m}}{a^2} \Phi(b) \quad (5.27)$$

$$\Lambda(I_{\{1,2\}}^{(m)}(a, a)) = \frac{2^{-2m} \left(\frac{1}{a} \frac{1}{\sqrt{2\pi}} \exp(-\frac{b^2}{2})\right)^2 + o(2^{-2m})}{\left| a \frac{b}{\sqrt{2\pi}} \exp(-\frac{b^2}{2}) - a \frac{-b}{\sqrt{2\pi}} \exp(-\frac{b^2}{2}) \right|} \sim \frac{2^{-2m}}{2a^3 b \sqrt{2\pi} \exp(\frac{b^2}{2})}. \quad (5.28)$$

Here, there are two scenarios satisfying $|\Pi \cap K_{t_1, z_1}| = |\Pi \cap K_{t_2, z_2}| = 1$ and Equations (5.26)–(5.28) yield

$$\begin{aligned} \mathbb{P}(\Pi(I_{\{1,2\}}(a, a)) = 1 \mid Z(-b) = Z(b) = a) &= \frac{a(2b\sqrt{2\pi} \exp(\frac{b^2}{2}))^{-1}}{\Phi(b)^2 + a(2b\sqrt{2\pi} \exp(\frac{b^2}{2}))^{-1}}, \\ \mathbb{P}(\Pi(I_{\{1\}}(a, a)) = \Pi(I_{\{2\}}(a, a)) = 1 \mid \\ Z(-b) = Z(b) = a) &= \frac{\Phi(b)^2}{\Phi(b)^2 + a(2b\sqrt{2\pi} \exp(\frac{b^2}{2}))^{-1}}. \end{aligned}$$

Note that the probability that the observations are generated by two different points of Π increases as a gets smaller. This is due to the fact that Π gets more intense as the second component decreases.

Using the formulae above, the limits of the conditional probabilities can always be calculated explicitly except for those cases where two scenarios exist, both involving different

terms which cannot be determined exactly (cf. Proposition 5.11). This may happen only if two sets $A_1 = \{i_1, \dots, i_r\}$ and $A_2 = \{j_1, \dots, j_s\}$, $r, s \geq 3$, $A_1 \cap A_2 \neq \emptyset$, exist such that

$$J_{A_1}(Z(\mathbf{t})) \neq \emptyset, \quad J_{A_2}(Z(\mathbf{t})) \neq \emptyset \quad \text{and} \quad J_{A_1 \cup A_2}(Z(\mathbf{t})) = \emptyset, \quad (5.29)$$

where

$$J_A(Z(\mathbf{t})) = \bigcup_{B \supset A} I_B(Z(\mathbf{t})) = K_{\mathbf{t}, Z(\mathbf{t})} \cap \bigcap_{i \in A} K_{t_i, Z(t_i)}, \quad A \in 2^{\{1, \dots, n\}} \setminus \emptyset.$$

In all other cases, the terms as in Proposition 5.11 are cancelled out. Note that we work with sets of the type $J_A(Z(\mathbf{t}))$ in order to avoid case-by-case analysis for all the sets $I_B(Z(\mathbf{t}))$ with $B \supset A$.

Lemma 5.15. *Let G consist of functions which are continuously differentiable a.e. Then, for any fixed set $\{t_1, \dots, t_n\} \subset \mathbb{R}$ we have*

$$\mathbb{P}(Z(\mathbf{t}) \text{ satisfies (5.29)}) = 0.$$

Proof. We proof that condition (5.29) has probability 0 for all fixed index sets $A_1, A_2 \subset \{1, \dots, n\}$. By renumbering, we may assume $A_1 = \{1, \dots, r\}$ and $A_2 = \{q, \dots, q+s-1\}$ with $q \leq r$. Assume that $\mathbb{P}(Z(\mathbf{t}) \text{ satisfies (5.29)}) > 0$. In a first step we only consider those realizations of $Z(t_1), \dots, Z(t_r)$ with $J_{A_1}(Z(t_1), \dots, Z(t_r)) \neq \emptyset$. Then, by the calculations in Propositions 5.9, 5.11 and 5.13, we get that

$$\mathbb{P}(\Pi(I_{A_1}^{(m)}(Z(t_1), \dots, Z(t_r))) = 1) \notin \mathcal{O}(2^{-2m(1+\varepsilon)})$$

for any $\varepsilon > 0$ and

$$\mathbb{P}(\Pi(I_{B_1}^{(m)}(Z(t_1), \dots, Z(t_r))) = 1, \Pi(I_{B_2}^{(m)}(Z(t_1), \dots, Z(t_r))) = 1) \in \mathcal{O}(2^{-3m})$$

for any $B_1, B_2 \subset A_1$, $B_1 \cap B_2 = \emptyset$. This yields $\Pi(J_{A_1}(Z(t_1), \dots, Z(t_r))) = 1$ a.s. Similarly, for almost every Z_q, \dots, Z_{q+s-1} such that $J_{A_2}(Z(t_q), \dots, Z(t_{q+s-1})) \neq \emptyset$, we have $\Pi(J_{A_2}(Z(t_q), \dots, Z(t_{q+s-1}))) = 1$. As

$$\begin{aligned} & \{\omega : Z(\mathbf{t}) \text{ satisfies (5.29)}\} \\ & \subset \{\omega : J_{A_1}(Z(t_1), \dots, Z(t_r)) \neq \emptyset\} \cap \{\omega : J_{A_2}(Z(t_q), \dots, Z(t_{q+s-1})) \neq \emptyset\}, \end{aligned}$$

we have $\Pi(J_{A_1}(Z(t_1), \dots, Z(t_r))) = 1$ and $\Pi(J_{A_2}(Z(t_q), \dots, Z(t_{q+s-1}))) = 1$ for $Z(\mathbf{t})$ satisfying (5.29) almost surely. Therefore, we get $\mathbb{P}(\Pi(K_{t_i, Z(t_i)}) \geq 2) > 0$ for every $i \in A_1 \cap A_2$ since $J_{A_1 \cup A_2}(Z(\mathbf{t})) = \emptyset$. This is a contradiction to Proposition 5.5. \square

From the considerations above and Lemma 5.15 we immediately derive the following result.

Theorem 5.16. *Let G consist of functions which are continuously differentiable a.e. Then, with probability one,*

$$\mathbb{P}(E_{\{n_A\}}(Z(\mathbf{t})) \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z}))$$

can be calculated explicitly via the formulae given in Propositions 5.9, 5.11, 5.13 and Remark 5.12.

Remark 5.17. We may also consider the case that G is countable. However, to transfer the results of the finite case, we have to ensure uniform convergence of the blurred intersection sets which is needed to compute $\sum_{n_A: \sum_{A: i \in A} n_A = 1} \mathbb{P}(E_{\{n_A\}}^{(m)}(\mathbf{z}))$ in the denominator of Equation (5.24). To this end, we have to impose some additional conditions. For example, we could assume that, for almost every $\mathbf{z} > \mathbf{0}$, there is only a finite number of shape functions involved in the intersection sets $I_A(\mathbf{z})$, $|A| \geq 2$.

We are still left with simulating $\Pi_2 \mid Z(\mathbf{t}) = \mathbf{z}$ given the occurrence of a scenario $E_{\{n_A\}}(\mathbf{z})$ with $\sum_{A: i \in A} n_A = 1$ for all $i \in \{1, \dots, n\}$, that is, we are interested in

$$\mathbb{P} \left(\bigcap_{A: n_A=1} \{\Pi_2(C_A \times (0, \infty) \times \{f\}) = 1\} \mid E_{\{n_A\}}(\mathbf{z}) \right)$$

for $C_A \subset \mathbb{R}$, $f \in G$ with $(C_A \times (0, \infty) \times \{f\}) \cap I_A(\mathbf{z}) \neq \emptyset$. Using Theorem 5.8 with sets $B_A = C_A \times (0, \infty) \times \{f\}$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, we get that

$$\mathbb{P} \left(\bigcap_{A: n_A=1} \{\Pi_2(I_A(\mathbf{z}) \cap B_A) = 1\} \mid E_{\{n_A\}}(\mathbf{z}) \right) = \lim_{m \rightarrow \infty} \prod_{A: n_A=1} \frac{\Lambda(I_A^{(m)}(\mathbf{z}) \cap B_A)}{\Lambda(I_A^{(m)}(\mathbf{z}))}.$$

Thus, each random vector $(T_A, F_A) \in \mathbb{R} \times G$, $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, which is defined by $\Pi_2(\{T_A\} \times (0, \infty) \times \{F_A\}) \cap I_A(Z(\mathbf{t})) = 1$ if $\Pi_2(I_A(Z(\mathbf{t}))) = 1$, can be simulated independently.

The distribution of (T_A, F_A) depends on the cardinal number of A . If $A = \{i\}$ for some $i \in \{1, \dots, n\}$, we have

$$\mathbb{P}(T_A \in B, F_A = f) = \frac{\mathbb{P}_F(\{f\}) \int_{D_i^{(f)} \cap B} f(t_i - t) dt}{\sum_{g \in G} \mathbb{P}_F(\{g\}) \int_{D_i^{(g)}} g(t_i - t) dt}, \quad B \in \mathcal{B}.$$

For $|A| \geq 2$, let $I_A(\mathbf{z}) = \{(t_0^{(1)}, y_0^{(1)}, f_1), \dots, (t_0^{(k)}, y_0^{(k)}, f_k)\}$. Then, we get

$$\mathbb{P}(T_A = t_0^{(j)}, F_A = f_j) = \frac{\frac{\mathbb{P}_F(\{f_j\})}{(y_0^{(j)})^2 |z_1 f_j'(t_2 - t_0^{(j)}) - z_2 f_j'(t_1 - t_0^{(j)})|}}{\sum_{j'=1}^k \frac{\mathbb{P}_F(\{f_{j'}\})}{(y_0^{(j')})^2 |z_1 f_{j'}'(t_2 - t_0^{(j')}) - z_2 f_{j'}'(t_1 - t_0^{(j')})|}}, \quad j = 1, \dots, k.$$

Thus, we end up with the following procedure for calculating the conditional distribution of $Z(t_0)$ given $Z(\mathbf{t}) = \mathbf{z}$ with $t_0, t_1, \dots, t_n \in \mathbb{R}$, $\mathbf{z} > \mathbf{0}$.

1. Compute the conditional probabilities (5.24) for all the scenarios $E_{\{n_A\}}(\mathbf{z})$ and generate a random scenario following this distribution.
2. For a given scenario $E_{\{n_A\}}(\mathbf{z})$, set $\Pi_2 = \sum_{\{A: n_A=1\}} \delta_{(T_A, U_A, F_A)}$, where the law of (T_A, F_A) is given above and $U_A = \min_{i=1}^n \frac{z_i}{F_A(t_i - T_A)}$.
3. Independently, sample points $(S_i, U_i, F_i)_{i \in \mathbb{N}}$ from $\Pi_3(\cdot) = \Pi(\cdot \setminus (K_{(\mathbf{t}, \mathbf{z})} \cup \overline{K_{(\mathbf{t}, \mathbf{z})}}))$. Then, $Z(t_0) = \max_{\{A: n_A=1\}} (U_A \cdot F_A(t_0 - T_A)) \vee \max_{i \in \mathbb{N}} (U_i \cdot F_i(t_0 - S_i))$.

In the next section, we will point out the capability of this exact approach by comparing it to other algorithms in the simple case of a deterministic shape function which is continuously differentiable. In Section 5.5, we address the computational burden of this algorithm.

5.4 Comparison with Other Algorithms

Recently, Wang and Stoev (2011) proposed an algorithm for exact and efficient conditional sampling for max-linear models

$$Z(t_i) = \max_{j=1, \dots, p} a_{ij} Y_j, \quad i = 0, \dots, n,$$

where Y_j , $j = 1, \dots, p$, are independent standard Fréchet random variables.

Rewriting Z from (5.1) as an extremal integral (see Stoev and Taqqu, 2005)

$$Z(t) = \sum_{f \in G} \mathbb{P}_F(\{f\}) \int_{\mathbb{R}}^e f(t-u) M_1(du), \quad t \in \mathbb{R},$$

where M_1 is a random sup-measure on \mathbb{R} w.r.t. the Lebesgue measure, we can see that Z can be approximated arbitrarily well by a max-linear model, e.g. by

$$Z_{M,h}(t) = h \max_{\substack{l=-M, \dots, M-1 \\ j=1, \dots, k}} \mathbb{P}_F(\{f_j\}) \cdot f_j \left(t - \left(l + \frac{1}{2} \right) h \right) \cdot Y_{j,l}, \quad M \in \mathbb{N}, h > 0,$$

where $G = \{f_1, \dots, f_k\}$. Then, we have $Z_{M,h}(t) \xrightarrow{p} Z(t)$, $t \in \mathbb{R}$, as $M \rightarrow \infty$, $h \rightarrow 0$.

We also consider another approach based on the assumption of a multi-gaussian model (cf. Chilès and Delfiner, 1999, p. 381). The data are transformed such that the marginal distribution is Gaussian. As the marginals of Z are known to be standard Fréchet, the corresponding transformation is given by

$$\Psi : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \Phi^{-1}(\Phi_1(x)),$$

where Φ is the standard normal distribution function and $\Phi_1 = \exp(-1/x)$ is the standard Fréchet distribution function. The transformed random field $Y = \Psi(Z)$ is stationary and second-order. As the covariance function C of Y can hardly be computed for general shape functions f_1, \dots, f_k , we estimate it using maximum likelihood techniques, for instance, from a large parametric class like the Whittle-Matérn class

$$\tilde{\kappa}_{\nu,c}(h) = \frac{(c||h||)^\nu}{2^{\nu-1}\Gamma(\nu)} \mathcal{K}_\nu(c||h||), \quad \nu, c > 0,$$

assuming that Y is a Gaussian random field. Under this assumption, the conditional distribution can be sampled easily as described in Section 2.1. Afterwards, the sample has to be retransformed via

$$\Psi^{-1} : \mathbb{R} \rightarrow (0, \infty), \quad y \mapsto \Phi_1^{-1}(\Phi(y)).$$

To compare different methods, we need a measure for the goodness-of-fit of a distribution. Here, we use the *continuous ranked probability score* (CRPS) which is defined as

$$CRPS(F_1, x) = - \int_{-\infty}^{\infty} (F_1(y) - \mathbf{1}_{\{y \geq x\}})^2 dy,$$

where F_1 is a cumulative distribution function and $x \in \mathbb{R}$ (see [Gneiting and Raftery, 2007](#)). Note that $CRPS(F_1, F_2) := \int CRPS(F_1, x) F_2(dx)$ is a strictly proper scoring rule, i.e.

$$CRPS(F_2, F_2) \geq CRPS(F_1, F_2)$$

for all cumulative distribution functions F_1, F_2 . If F_1 and F_2 both belong to measures with finite first moment, equality holds if and only if $F_1 = F_2$. Assuming that F_1 has a finite first moment, by Lemma 2.1 in [Baringhaus and Franz \(2004\)](#), the CRPS can be calculated via

$$CRPS(F_1, x) = \frac{1}{2} \mathbb{E}_{F_1} |X - X'| - \mathbb{E}_{F_1} |X - x|, \quad (5.30)$$

which shows $CRPS(F_1, F_1) = -\frac{1}{2} \mathbb{E}_{F_1} |X - X'| \leq 0$. Here, $X, X' \sim F_1$ are independent random variables.

In order to compare different algorithms for getting a realization from the conditional distribution $Z(t_0) \mid Z(\mathbf{t})$, we consider K samples Z_1, \dots, Z_K of the random field Z . For each method m , we get an empirical distribution function $F_i^{(m)}$ by drawing k times from the (approximated) conditional distribution of $\log(Z_i(t_0)) \mid Z_i(\mathbf{t})$, $i = 1, \dots, K$, and calculate $CRPS(F_i^{(m)}, \log(Z_i(t_0)))$ via (5.30). Here, we do the log-transformation to Gumbel marginals to ensure that the conditional distribution has finite expectation.

Then, a measure for the goodness-of-fit is the *mean score* ([Gneiting and Raftery, 2007](#))

$$CRPS_{K,m} = \frac{1}{K} \sum_{i=1}^K CRPS(F_i^{(m)}, \log(Z_i(t_0))).$$

Furthermore, we consider the mean absolute error of the logarithmic conditional median

$$MAE_{K,m} = \frac{1}{K} \sum_{i=1}^K \left| \log \left((F_i^{(m)})^{-1}(0.5) \right) - \log(Z_i(t_0)) \right|.$$

Here, for computational reasons, we choose the Gaussian extreme value process ([Smith, 1990](#)) which has the deterministic shape function

$$f(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} x^2 \right).$$

Furthermore, let $n = 4$, $\mathbf{t} = (-2, -1, 1, 2)$ and $t_0 = 0$. Figure 5.3 shows two realizations of $Z(\cdot)$, the first one is sampled unconditionally and the second one is based on conditional sampling of the first one.

The conditional distribution $F_i^{(m)}$ is calculated based on a sample of size $k = 100$ simulated in R ([Ihaka and Gentleman, 1996](#); [R Development Core Team, 2011](#)). The performance is measured via $CRPS_{K,1} / MAE_{K,1}$ (conditional sampling via the Poisson point

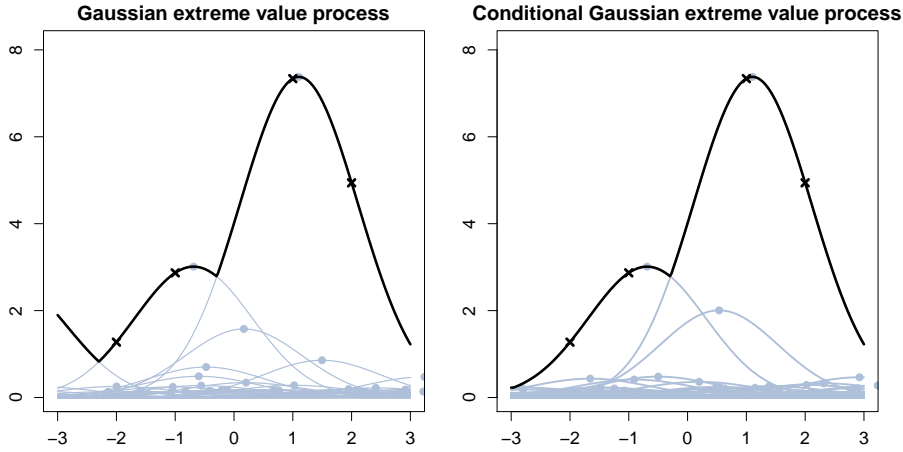


Figure 5.3: Left: Construction of Z . The grey dots represent the points $(S_i, U_i \cdot F_i(0))$ with $\Pi = \sum_{i \in \mathbb{N}} \delta_{(S_i, U_i, F_i)}$, the black line is one realization of Z . The black dots mark $Z(\mathbf{t})$. Right: Construction of Z conditional on $Z(\mathbf{t})$.

	$m = 1$	$m = 2$	$m = 3$
$CRPS_{K,m}$	-0.135	-0.359	-0.251
$MAE_{K,m}$	0.197	0.506	0.338

Table 5.1: Results of the simulation study for $f(x) = \varphi(x)$ and $K = 1000$.

process), $CRPS_{K,2} / MAE_{K,2}$ (conditional sampling for a max-linear model with $M = 5$ and $h = 0.1$ using the R package `maxLinear` (Wang, 2010)) and $CRPS_{K,3} / MAE_{K,3}$ (conditional sampling via transformation to Gaussian marginals) with $K = 1000$ samples. As already mentioned, the latter approach requires the knowledge of the covariance structure of the transformed random field. This is assessed by first simulating data from this model on a dense grid repeatedly and then estimating the scale and the smoothness parameter of a Whittle-Matérn covariance model based on maximum likelihood techniques implemented in the R package `RandomFields` (Schlather, 2012). The parameters for the first and second method are chosen such that these methods have a similar running time, while the last method runs much faster.

The results of the simulation study are shown in Table 5.1. Here, $CRPS_{K,1}$ and $MAE_{K,1}$ can be interpreted as reference values as the first method is exact. We note that conditional sampling for max-linear models performs worse than conditional sampling via transformation to Gaussian marginals.

For further analysis and comparison of these methods we do not restrict ourselves to pointwise prediction, but have a look at the sample paths. Additionally, pointwise quantile estimation of the conditional distribution can be done including the special case of the conditional median which can be seen as an analogon to kriging. In case of conditional sampling via the Poisson point process and conditional sampling of a max-linear model, the quantiles have to be estimated from the empirical conditional distribution. For sampling via Gaussian transformation, the quantiles can be calculated directly by the means of Kriging result and variance.

Figure 5.4 shows five sample paths and the median of the Gaussian extreme value process conditional on

- a. observations at four locations $-2, -1, 1, 2$,
- b. observations at eleven locations $-2.5, -2, \dots, 2, 2.5$.

In general, conditional simulation via the Poisson point process yields sample paths which capture the main features of the process quite well. Even in the case of four observations, parts of the sample path are reconstructed exactly with a positive probability. For eleven observations most of the sample path is restored with high probability.

The results of conditional sampling of the max-linear model are similar to the first method in case of four observations. For eleven observations, however, the method fails because of model misspecification. As the data do not match the discretized model, some observations cannot be reconstructed. For some realizations of the Gaussian extreme value process, this problem even occurs in case of four observations. This is the main reason for the bad results of this method in the simulation study above.

Gaussian transformation yields conditional sample paths which are structurally very different from the true ones. However, for eleven observations the deviations from the original sample path are quite small.

5.5 Reduction of Computational Burden

Now, we deal with the computational costs of the above algorithm. The computing time increases linearly with $|G|$ and exponentially with n due to the fact that, for any function $f \in G$, the intersections of the curves

$$\left\{ \left(t, \frac{Z(t_i)}{f(t_i - t)} \right) : t \in \mathbb{R}, f(t_i - t) > 0 \right\}, \quad i = 1, \dots, n,$$

and the corresponding intensities have to be calculated. Furthermore, these intersection sets have to be combined to scenarios. The following example shows that the number of scenarios with positive probability may grow exponentially:

Example 5.18. *Assume that $t_1 < t_2 < \dots < t_n \in \mathbb{R}$, \mathbb{P}_F is degenerated and $Z(\mathbf{t}) = \mathbf{z}$ such that there are no intersections of three or more curves and two curves K_{t_i, z_i} and K_{t_j, z_j} intersect if and only if $|i - j| = 1$. Furthermore, we assume that $I_{\{i\}}(\mathbf{z})$ is non-empty for $i = 1, \dots, n$.*

Let $R(n)$ denote the number of scenarios with a positive probability conditional on n observations. Then, there are $R(n - 1)$ scenarios (with positive probability) satisfying $\Pi(I_{\{n\}}(\mathbf{z})) = 1$ and $R(n - 2)$ scenarios with $\Pi(I_{\{n-1, n\}}(\mathbf{z})) = 1$. Thus, we get the recurrence formula $R(n) = R(n - 1) + R(n - 2)$ with $R(1) = 1$ and $R(2) = 2$. Hence, $R(n + 1)$ equals the n -th Fibonacci number and grows exponentially.

We note that the calculation of all the intersection sets of dominating order, i.e. all the intersections of at least three curves, is unavoidable as these will be included in Π_2 with probability one — if non-empty. Depending on the algorithm used for determining

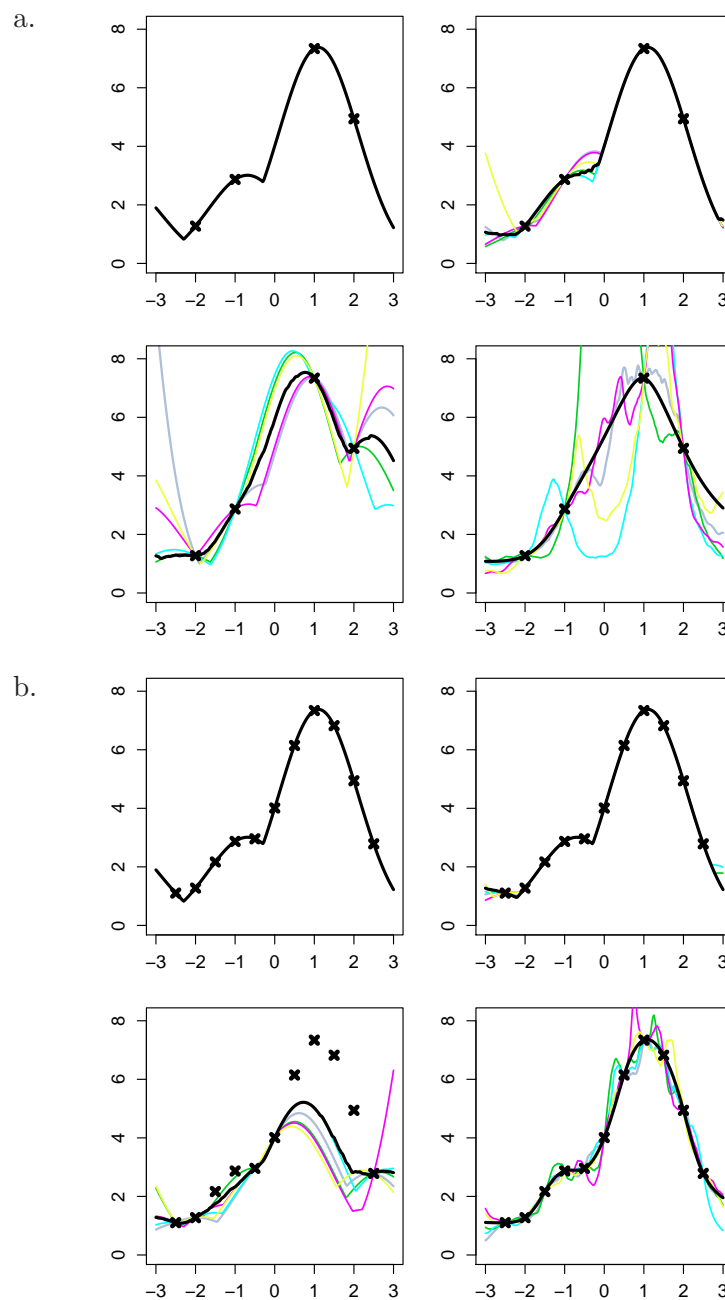


Figure 5.4: Comparison of the Gaussian extreme value process with different types of conditional simulations: a. simulations conditional on four observations at $-2, -1, 1, 2$, b. simulations conditional on eleven observations at $-2.5, -2, \dots, 2, 2.5$. In both cases the original Gaussian extreme value process (top left), conditional samples via the Poisson point process (top right) and conditional results for a max-linear approximation (bottom left) and an approximation by Gaussian transformation (bottom right) are shown. Black crosses: observations, coloured lines: conditional sample paths, black line: conditional mean.

these intersections, the intersections of two curves may be found with low additional computational costs, e.g. by evaluating the function

$$t \mapsto \left(\mathbf{1}_{\left\{ \frac{z_1}{f(t_1-t)} = \min_{i=1, \dots, n} \frac{z_i}{f(t_i-t)} \right\}}, \dots, \mathbf{1}_{\left\{ \frac{z_n}{f(t_n-t)} = \min_{i=1, \dots, n} \frac{z_i}{f(t_i-t)} \right\}} \right).$$

However, storing all these intersections which are needed for conditional sampling might be very demanding. Therefore, we aim to reduce at least the required memory.

To this end, we consider the shape functions involved in the point process Π_2 . As $\Pi(K_{t_1, Z(t_1)}) = \dots = \Pi(K_{t_n, Z(t_n)}) = 1$ a.s. by Proposition 5.5, there are well-defined random variables $\Theta_1, \dots, \Theta_n$ such that $\Pi(K_{t_i, Z(t_i)} \cap \mathbb{R} \times (0, \infty) \times \{\Theta_i\}) = 1$, i.e. Θ_i denotes the shape function generating the i th observation, $i = 1, \dots, n$. By the results of Section 5.2 and 5.3, the computation would become easy if $\Theta = (\Theta_1, \dots, \Theta_n)$ was known. Therefore, it might be promising to do conditional sampling of $Z \mid Z(\mathbf{t}) = \mathbf{z}$ in the following way:

1. Choose $\theta = (\theta_1, \dots, \theta_n)$ according to the distribution $\pi(\cdot)$ of $\Theta \mid Z(\mathbf{t}) = \mathbf{z}$
2. Sample $\Pi_2 \mid Z(\mathbf{t}) = \mathbf{z}, \Theta = \theta$.
3. Simulate Π_3 .

So, we are stuck to the problem to compute the distribution π of Θ given $Z(\mathbf{t})$. By summing up all the generalized scenarios including $\Theta = \theta$, we get

$$\mathbb{P}(\Theta = \theta \mid Z(\mathbf{t}) = \mathbf{z}) = \lim_{m \rightarrow \infty} \frac{\mathbb{P}(\Theta = \theta, Z(\mathbf{t}) \in A_m(\mathbf{z}))}{\sum_{\theta \in G^n} \mathbb{P}(\Theta = \theta, Z(\mathbf{t}) \in A_m(\mathbf{z}))}.$$

as an immediate consequence of Theorem 5.8. The explicit calculations can be done by means of the results in Section 5.3.

As already mentioned above, we have to consider the intersections of at least three curves K_{t_i, z_i} , $i = 1, \dots, n$, separately. There are sets $B_1(\mathbf{z}), \dots, B_l(\mathbf{z}) \subset \{1, \dots, n\}$ such that $|B_j(\mathbf{z})| > 2$, $I_{B_j(\mathbf{z})}(\mathbf{z}) \neq \emptyset$ and $I_{\bar{B}_j}(\mathbf{z}) = \emptyset$ for all $B_j \supseteq B_j(\mathbf{z})$, $j = 1, \dots, l$. We assume that $l = l(\mathbf{z})$ is maximal, which means that the set of all indices involved in intersections of at least three curves is given by $B_{\geq 3}(\mathbf{z}) = \bigcup_{j=1}^l B_j(\mathbf{z})$. By Lemma 5.15, the sets $B_1(\mathbf{z}), \dots, B_l(\mathbf{z})$ are pairwise disjoint and we get $\Pi(I_{B_j(Z(\mathbf{t}))}(Z(\mathbf{t}))) = 1$ a.s., in particular $|\{\Theta_i : i \in B_j(Z(\mathbf{t}))\}| = 1$ a.s. for any $j \in \{1, \dots, l\}$. As $\Pi(K_{t_i, Z(t_i)}) = 1$ a.s., we have that $\mathbb{P}(\Pi(J_{\{i, j\}}(Z(\mathbf{t}))) > 0 \mid Z(\mathbf{t}) = \mathbf{z}) = 0$ for any $i \in B_{\geq 3}(\mathbf{z})$, $j \notin B_{\geq 3}(\mathbf{z})$ which means that only intersection sets $I_B(\mathbf{z})$ satisfying either $B \subset B_{\geq 3}(\mathbf{z})$ or $B \subset B_{\leq 2}(\mathbf{z}) = B_{\leq 2}^c(\mathbf{z})$ have to be considered. Therefore, the random vectors $(\Theta_i)_{i \in B_{\geq 3}(\mathbf{z})}$ and $(\Theta_j)_{j \in B_{\leq 2}(\mathbf{z})}$ are independent conditional on $Z(\mathbf{t}) = \mathbf{z}$. In practice, the set of shape functions involved in intersections $I_B(\mathbf{z})$, $|B| \geq 3$, is very small. Thus, the simulation of $\{\Theta_i \mid Z(\mathbf{t}), i \in B_{\geq 3}(\mathbf{z})\}$ is fast.

By the considerations above we get

$$\begin{aligned} & \mathbb{P}(\Theta = \theta, Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= c_m(\{\theta_j, j \in B_{\leq 2}(\mathbf{z})\}) \cdot \prod_{j=1}^l \left(\sum_{f_j \in G} \Lambda(I_{B_j}^{(m)}(\mathbf{z}) \cap (\mathbb{R} \cap (0, \infty) \cap \{f_j\})) \prod_{i \in B_j} \mathbf{1}_{\theta_i = f_j} \right). \end{aligned}$$

where $c_m(\{\theta_j, j \in B_{\leq 2}(\mathbf{z})\}) \geq 0$ is a constant depending only on $\{\theta_j, j \in B_{\leq 2}(\mathbf{z})\}$. As no three elements of $\{K_{t_j, z_j}(\mathbf{z}), j \in B_{\leq 2}(\mathbf{z})\}$ intersect, Propositions 5.9 and 5.13 yield that $c_m(\{\theta_j, j \in B_{\leq 2}(\mathbf{z})\}) \in \mathcal{O}(2^{-|B_{\leq 2}(\mathbf{z})| \cdot m})$ and we get

$$\begin{aligned} & c_m(\{\theta_j, j \in B_{\leq 2}(\mathbf{z})\}) \notin o(2^{-|B_{\leq 2}(\mathbf{z})| \cdot m}) \\ \iff & K_{t_j, z_j} \cap K_{\mathbf{t}, \mathbf{z}} \cap (\mathbb{R} \times (0, \infty) \cap \{\theta_j\}) \neq \emptyset \text{ for all } j \in B_{\leq 2}(\mathbf{z}). \end{aligned} \quad (5.31)$$

For simulating $\pi(\cdot \mid \{\theta_i, i \in B_{\geq 3}(\mathbf{z})\})$ we notice that it might still be very challenging to calculate the probabilities for large G , in particular if $B_{\leq 2}(\mathbf{z})$ is large. The computational burden can be eased by MCMC techniques like Gibbs sampling (cf. Lantuéjoul, 2002; Gaetan and Guyon, 2010) restricted to components in $B_{\leq 2}(\mathbf{z})$ using that $Q = \{(\theta_i)_{i \in B_{\leq 2}(\mathbf{z})} : \pi(\mathbf{x} \mid \{\theta_j, j \in B_{\geq 3}(\mathbf{z})\}) > 0\} = \times_{i \in B_{\leq 2}(\mathbf{z})} Q_i$ with $Q_i = \{\theta_i \in G : K_{t_i, z_i} \cap K_{\mathbf{t}, \mathbf{z}} \cap (\mathbb{R} \times (0, \infty) \cap \{\theta_i\}) \neq \emptyset\}$ by Equation (5.31).

The convergence of Gibbs sampling is guaranteed by Theorem 4.4 in Gaetan and Guyon (2010) – assuming that all the conditional distributions $\pi(\cdot \mid \theta^{(i)})$ can be simulated exactly, where $\theta^{(i)} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n)$.

Again, as G can be very large or even infinite, the calculation of the probabilities involved in $\pi(\cdot \mid \theta^{(i)})$ might cause some problems. Therefore, sampling from $\pi(\cdot \mid \theta^{(i)})$ is done by a standard Metropolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) with proposal transition probability $q(f, g) = \mathbb{P}_F(\{g\})$, $f \in Q_i$, $g \in G \supset Q_i$, and acceptance probability $\min \left\{ 1, \frac{\pi(g|\theta^{(i)})\mathbb{P}_F(\{f\})}{\pi(f|\theta^{(i)})\mathbb{P}_F(\{g\})} \right\}$ (see Gaetan and Guyon, 2010).

Thus, we get the probability of transition from f to g by

$$P(f, g) = \begin{cases} \mathbb{P}_F(\{g\}) \min \left\{ 1, \frac{\mathbb{P}_F(\{f\}) \pi(g|\theta^{(i)})}{\mathbb{P}_F(\{g\}) \pi(f|\theta^{(i)})} \right\}, & f \neq g, \\ 1 - \sum_{h \in Q_i \setminus \{f\}} P(f, h), & f = g, \end{cases} \quad (5.32)$$

where $\frac{\pi(g|\theta^{(i)})}{\pi(f|\theta^{(i)})} = \frac{\pi(g|\theta^{(i)}, \Theta_i \in \{f, g\})}{\pi(f|\theta^{(i)}, \Theta_i \in \{f, g\})}$ can be calculated by the formulae from Section 5.3.

Proposition 5.19. *For $\pi(\cdot \mid \theta^{(i)})$ -a.e. initial value $f \in Q_i$, the Markov chain with transition kernel P converges to $\pi(\cdot \mid \theta^{(i)})$ in total variation norm, i.e.*

$$\sup_{A \in \mathcal{G}} \left| P^k(f, A) - \pi(A \mid \theta^{(i)}) \right| \xrightarrow{k \rightarrow \infty} 0.$$

Proof. By Theorem 4.1 in Gaetan and Guyon (2010) (see also Tierney, 1994) we have to verify that P is $\pi(\cdot \mid \theta^{(i)})$ -irreducible, $\pi(\cdot \mid \theta^{(i)})$ -invariant and aperiodic. To show irreducibility, we just note that $P(f, g) = 0$ implies $\pi(g \mid \theta^{(i)}) = 0$ as $\mathbb{P}_F(\{f\}), \mathbb{P}_F(\{g\}) > 0$. Invariancy w.r.t. $\pi(\cdot \mid \theta^{(i)})$ holds because of $\pi(\cdot \mid \theta^{(i)})$ -reversibility

$$\pi(f \mid \theta^{(i)})P(f, g) = \min\{\mathbb{P}_F(\{g\})\pi(f \mid \theta^{(i)}), \mathbb{P}_F(\{f\})\pi(g \mid \theta^{(i)})\} = \pi(g \mid \theta^{(i)})P(g, f).$$

Furthermore, we get aperiodicity because of $P(f, f) \geq \mathbb{P}_F(\{f\}) > 0$ for all $f \in G$. \square

We end up with the following MCMC algorithm for $\pi(\cdot)$:

1. Determine $B_{\geq 3}(\mathbf{z})$ and simulate $\Theta_i, i \in B_{\geq 3}(\mathbf{z})$.

2. Start with initial values $\theta \in Q$.
3. Choose $i \in B_{\leq 2}(\mathbf{z})$ according to a uniform distribution on G and simulate g according to $\pi(\cdot | \theta^{(i)})$ by the above Metropolis algorithm.
4. Update $\theta_i := g$ and go to 3.

The main advantage of drawing $\Theta | Z(\mathbf{t})$ first, instead of drawing scenarios directly, is that intersection sets are much smaller and therefore less memory is needed. However, still all the intersections of at least three curves have to be calculated first.

5.6 Approximation in the Case of an Infinite Number of Shape Functions

Here, we drop the assumption that G is finite. Note that the measurability of Π_2 and Π_3 is still an open question if G is uncountable. We present an approximation of the distribution of $Z(t_0)$ given $Z(\mathbf{t})$ based on a finite number of shape functions.

Let $\Gamma_1, \Gamma_2, \dots$ be independent copies of F where F is defined as at the beginning of this chapter. Then, given $\Gamma_1, \dots, \Gamma_N$, we define

$$Z_N(t) = \max_{i \in \mathbb{N}} U_i^{(N)} F_i^{(N)}(t - S_i^{(N)}), \quad t \in \mathbb{R}^d, \quad (5.33)$$

where $(S_i^{(N)}, U_i^{(N)}, F_i^{(N)})$, $i \in \mathbb{N}$, are the points of a Poisson point process $\Pi^{(N)} = \sum_{i \in \mathbb{N}} \delta_{(S_i^{(N)}, U_i^{(N)}, F_i^{(N)})}$ on $\mathbb{R}^d \times (0, \infty) \times \{\Gamma_1, \dots, \Gamma_N\}$ with intensity measure

$$\Lambda(A \times B \times \{\Gamma_k\}) = \frac{1}{N} \int_A \int_B \frac{1}{u^2} du ds, \quad A \in \mathcal{B}^d, B \in \mathcal{B} \cap (0, \infty), k \in \{1, \dots, n\}.$$

Theorem 5.20. *For any $\mathbf{z} > \mathbf{0}$, it holds*

$$\Pi^{(N)} | Z_N(\mathbf{t}) \leq \mathbf{z} \xrightarrow{d} \Pi | Z(\mathbf{t}) \leq \mathbf{z}$$

as $N \rightarrow \infty$. In particular, $Z_N(t_0) | Z_N(\mathbf{t}) \leq \mathbf{z} \xrightarrow{d} Z(t_0) | Z(\mathbf{t}) \leq \mathbf{z}$.

Proof. We note that it suffices to show

$$\begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{P}(\Pi^{(N)}(M_j) = n_j, j = 1, \dots, l | Z_N(\mathbf{t}) \leq \mathbf{z}) \\ &= \mathbb{P}(\Pi(M_j) = n_j, j = 1, \dots, l | Z(\mathbf{t}) \leq \mathbf{z}) \end{aligned} \quad (5.34)$$

for $\mathbf{z} > \mathbf{0}$ and sets $M_j = (a_j, b_j) \times (c_j, \infty) \times G_j \subset \mathbb{R}^d \times (0, \infty) \times G$ which are pairwise disjoint. First, we consider the distributions of Z and Z_N , respectively. It holds that $\mathbb{P}(Z(\mathbf{t}) \leq \mathbf{z}) = \exp(-\mathbb{E}H)$, where

$$H = \int_{\mathbb{R}^d} \max_{i=1, \dots, n} \frac{F(t_i - t)}{z_i} dt.$$

We define H_1, \dots, H_N replacing F by $\Gamma_1, \dots, \Gamma_N$, respectively. Then, we have

$$\mathbb{P}(Z_N(\mathbf{t}) \leq \mathbf{z}) = \mathbb{E}_{\Gamma_1, \dots, \Gamma_N}(\mathbb{P}(Z_N(\mathbf{t}) \leq \mathbf{z} \mid \Gamma_1, \dots, \Gamma_N)) = E \left(\exp \left(-\frac{1}{N} \sum_{k=1}^N H_k \right) \right).$$

As the sets M_1, \dots, M_l are pairwise disjoint, the joint distribution of Π and $Z(\mathbf{t})$ can be rewritten as

$$\begin{aligned} & \mathbb{P}(\Pi(M_j) = n_j, j = 1, \dots, l, Z(\mathbf{t}) \leq \mathbf{z}) \\ &= \mathbb{P}(Z(\mathbf{t}) \leq \mathbf{z}) \cdot \prod_{j=1}^l \mathbb{P} \left(\Pi \left(M_j \cap \left\{ (x, y, f) \in S : y \leq \min_{i=1, \dots, n} \frac{z_i}{f(t_i - x)} \right\} \right) = n_j \right) \\ &= \mathbb{P}(Z(\mathbf{t}) \leq \mathbf{z}) \\ & \quad \cdot \prod_{j=1}^l \mathbb{P} \left(\Pi \left(\left\{ (x, y, f) \in (a_j, b_j) \times (c_j, \infty) \times G_j : y \leq \min_{i=1, \dots, n} \frac{z_i}{f(t_i - x)} \right\} \right) = n_j \right). \end{aligned}$$

Thus, we get

$$\mathbb{P}(\Pi(M_j) = n_j, j = 1, \dots, l, Z(\mathbf{t}) \leq \mathbf{z}) = \exp(-\mathbb{E}H) \prod_{j=1}^l \frac{(\mathbb{E}H^{(j)})^{n_j}}{n_j!} \exp(-\mathbb{E}H^{(j)}),$$

where

$$H^{(j)} = \int_{a_j}^{b_j} \left(\left(\frac{1}{c_j} - \max_{i=1, \dots, n} \frac{F(t_i - t)}{z_i} \right) \vee 0 \right) \mathbf{1}_{\{F \in G_j\}} dt, \quad j = 1, \dots, l.$$

Analogously, we can define $H_k^{(j)}$ replacing F by Γ_k in the definition of $H^{(j)}$ for $j = 1, \dots, l, k = 1, \dots, N$. This yields

$$\begin{aligned} & \mathbb{P}(\Pi^{(N)}(M_j) = n_j, j = 1, \dots, l, Z_N(\mathbf{t}) \leq \mathbf{z}) \\ &= \mathbb{E} \left(\exp \left(-\frac{1}{N} \sum_{k=1}^N H_k \right) \prod_{j=1}^l \frac{1}{n_j!} \left(\frac{1}{N} \sum_{k=1}^N H_k^{(j)} \right)^{n_j} \exp \left(-\frac{1}{N} \sum_{k=1}^N H_k^{(j)} \right) \right). \end{aligned}$$

As $H_1, \dots, H_N \sim_{i.i.d.} H$ and $H_1^{(j)}, \dots, H_N^{(j)} \sim_{i.i.d.} H^{(j)}$ for $j = 1, \dots, l$ with $E|H| \leq \sum_{i=1}^n \frac{1}{z_i} < \infty$ and $H^{(j)} < \frac{\prod_{i=1}^d (b_{ji} - a_{ji})}{c_j}$, the Strong Law of Large Numbers yields that $\frac{1}{N} \sum_{k=1}^N H_k \rightarrow \mathbb{E}H$ and $\frac{1}{N} \sum_{k=1}^N H_k^{(j)} \rightarrow \mathbb{E}H^{(j)}$ a.s. as $N \rightarrow \infty$. Thus, dominated convergence yields (5.34) and $\mathbb{P}(Z_N(\mathbf{t}) \leq \mathbf{z}) \rightarrow \mathbb{P}(Z(\mathbf{t}) \leq \mathbf{z})$, which verifies the first assertion.

The second assertion follows immediately by rewriting

$$\{\omega \in \Omega : Z(t_0) \leq z\} = \{\omega \in \Omega : \Pi(\{(x, y, f) \in S : yf(t_0 - x) > z\}) = 0\}.$$

□

If G is countable, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \Pi_3^{(N)}(\cdot) \mid (Z_N(\mathbf{t}) = \mathbf{z}) &\stackrel{d}{=} \lim_{N \rightarrow \infty} \Pi^{(N)}(\cdot \setminus (K_{\mathbf{t}, \mathbf{z}} \cup \overline{K_{\mathbf{t}, \mathbf{z}}})) \\ &\stackrel{d}{=} \Pi(\cdot \setminus (K_{\mathbf{t}, \mathbf{z}} \cup \overline{K_{\mathbf{t}, \mathbf{z}}})) \stackrel{d}{=} \Pi_3(\cdot) \mid (Z(\mathbf{t}) = \mathbf{z}), \end{aligned}$$

where we used the first part of Theorem 5.20 and applied the second part of Theorem 5.7 to the processes Z and Z_N . This motivates to improve the approximation $\Pi \approx \Pi^{(N)}$ by $\Pi(\cdot) \approx \Pi_2^{(N)}(\cdot) + \Pi_3(\cdot)$, i.e. by the following procedure:

1. Simulate points (T_A, U_A, F_A) of $\Pi_2^{(N)} \mid Z(\mathbf{t}) = \mathbf{z}$ as described in Sections 5.3 and 5.5.
2. Independently of $\Pi_2^{(N)}$, sample points (S_i, U_i, F_i) from $\Pi_3 \mid Z(\mathbf{t}) = \mathbf{z}$ which is defined by $\Pi_3(\cdot) = \Pi(\cdot \cap (\mathbb{R} \times (0, \infty) \times G) \setminus (K_{\mathbf{t}, \mathbf{z}} \cup \overline{K_{\mathbf{t}, \mathbf{z}}}))$ analogously to the second part of Theorem 5.7.

Then, $Z(t_0) \approx \max_{\{A: n_A=1\}} (U_A \cdot F_A(t_0 - T_A)) \vee \max_{i \in \mathbb{N}} (U_i \cdot F_i(t_0 - S_i))$.

5.7 Application to the Brown-Resnick Process

We will apply the method of conditional sampling via the Poisson point process to the original Brown-Resnick process (Brown and Resnick, 1977).

Let $\{W_i(t), t \in \mathbb{R}\}$, $i \in \mathbb{N}$, be independent copies of a standard Brownian motion and — independently of the W_i 's — let $\sum_{i \in \mathbb{N}} \delta_{\tilde{U}_i}$ be a Poisson point process on $(0, \infty)$ with intensity measure $u^{-2} du$. Then,

$$Z(t) = \max_{i \in \mathbb{N}} \left(\tilde{U}_i \exp \left(W_i(t) - \frac{|t|}{2} \right) \right), \quad t \in \mathbb{R},$$

defines a stationary max-stable process with standard Fréchet margins.

Recently, this process was generalized (Kablichko et al. 2009; for further details see Chapter 4) and its mixed moving maxima representation (5.1) was given explicitly (Engelke et al., 2011).

This is,

$$Z(t) \stackrel{d}{=} \max_{i \in \mathbb{N}} \left(U_i \cdot \frac{1}{2} \exp(-R_i(t - S_i)) \right), \quad t \in \mathbb{R}, \quad (5.35)$$

where $\sum_{i \in \mathbb{N}} \delta_{(S_i, U_i, R_i)}$ is a Poisson point process on $\mathbb{R} \times (0, \infty) \times C(\mathbb{R})$ with intensity measure $ds u^{-2} du \mathbb{P}_R(df)$ and \mathbb{P}_R is the law of the process

$$R(t) = \mathbf{1}_{t < 0} R^{(-)}(-t) + \mathbf{1}_{t \geq 0} R^{(+)}(t).$$

Here, $\{R^{(-)}(t), t \geq 0\}$, $\{R^{(+)}(t), t > 0\}$ are independent Bessel processes of a three-dimensional Brownian motion with drift $\frac{1}{2}$ in its first component (cf. Rogers and Pitman, 1981), i.e.

$$R^{(-)}(t) \stackrel{d}{=} R^{(+)}(t) \stackrel{d}{=} \sqrt{(W_1(t) + |t|/2)^2 + W_2(t)^2 + W_3(t)^2},$$

where W_1, W_2 and W_3 are independent standard Brownian motions.

We will use the results obtained in the previous sections to sample from the conditional distribution of the Brown-Resnick process. However, the sample paths of $\exp(-R(\cdot))$ do not satisfy the assumptions of Propositions 5.9, 5.11 and 5.13. In particular, the sample paths are not continuously differentiable almost everywhere. To overcome this drawback, we do not use the exact sample paths $F(\cdot) = \exp(-R(\cdot))$, but the sample paths evaluated on a grid and interpolated linearly in between.

We show that this procedure is correct in the limit. Let $T = \{t_z, z \in \mathbb{Z}\} \subset \mathbb{R}$ with $\dots < t_{-2} < t_{-1} < t_0 < t_1 < t_2 < \dots$ such that $\lim_{z \rightarrow -\infty} t_z = -\infty$, $\lim_{z \rightarrow \infty} t_z = \infty$ and $\|T\| = \sup_{z \in \mathbb{Z}} (t_z - t_{z-1})$. Let $\{(t, F_T(t)), t \in \mathbb{R}\}$ be the polygonal line through the points $\{(t, F(t)), t \in T\}$. Furthermore, define $Z_T(t)$ as in (5.35), replacing F by F_T .

Proposition 5.21. *For $\|T\| \rightarrow 0$ we have $Z_T(t) \rightarrow Z(t)$ in probability for all $t \in \mathbb{R}$. In particular,*

$$Z_T(t_0) \mid (Z_T(\mathbf{t}) \in B) \xrightarrow{d} Z(t_0) \mid (Z(\mathbf{t}) \in B)$$

for all Borel sets $B \subset \mathbb{R}^n$ with $\mathbb{P}(Z(\mathbf{t}) \in B) > 0$ and $\mathbb{P}(Z(\mathbf{t}) \in \partial B) = 0$.

Proof. W.l.o.g. we assume $\|T\| \leq \frac{1}{2}$. Then, we have

$$\begin{aligned} \int_{\mathbb{R}} \frac{1}{2} F_T(t) dt &\leq \sum_{z \in \mathbb{Z}} \sup_{t \in [z-1/2, z+1/2]} \frac{1}{2} F_T(t) \\ &\leq \sum_{z \in \mathbb{Z}} \left(\sup_{t \in [z-1, z]} \frac{1}{2} F(t) + \sup_{t \in [z, z+1]} \frac{1}{2} F(t) \right) \\ &= \sum_{z \in \mathbb{Z}} \sup_{t \in [0, 1]} F(z+t) = \sum_{z \in \mathbb{Z}} \exp \left(- \inf_{t \in [0, 1]} R(z+t) \right). \end{aligned}$$

We show that the rhs is L^1 -integrable. To this end, we assess

$$\begin{aligned} &\mathbb{E} \left(\sum_{z \in \mathbb{Z}} \exp \left(- \inf_{t \in [0, 1]} R(z+t) \right) \right) \\ &\leq 2 \limsup_{n \rightarrow \infty} \sum_{k=0}^n \mathbb{E} \left(\exp \left(- \inf_{t \in [0, 1]} R(k+t) \right) \right) \\ &\leq 2 \limsup_{n \rightarrow \infty} \sum_{k=0}^n \mathbb{E} \left(\exp \left(- \inf_{t \in [0, 1]} \left| W(k+t) + \frac{k+t}{2} \right| \right) \right) \\ &\leq 2 \limsup_{n \rightarrow \infty} \sum_{k=0}^n \mathbb{E} \left(\exp \left(- \left| W(k) + \frac{k}{2} \right| + \sup_{t \in [0, 1]} \left| W(k+t) - W(k) + \frac{t}{2} \right| \right) \right) \\ &\leq 2 \mathbb{E} \left(\exp \left(\sup_{t \in [0, 1]} \left| W(t) + \frac{t}{2} \right| \right) \right) \cdot \limsup_{n \rightarrow \infty} \left\{ 1 + \sum_{k=1}^n \mathbb{E} \left(\exp \left(- \left| W(k) + \frac{k}{2} \right| \right) \right) \right\}. \end{aligned}$$

where W is a standard Brownian motion.

By Landau and Shepp (1970) there exists some $\varepsilon > 0$ such that

$$\mathbb{E} \left\{ \exp \left(\varepsilon \left(\sup_{t \in [0, 1]} \left| W(t) + \frac{t}{2} \right| \right)^2 \right) \right\} < \infty.$$

Since $|x| \leq \frac{1}{4\varepsilon} + \varepsilon x^2$ for any $x \in \mathbb{R}$, this implies that $\mathbb{E}(\exp(\sup_{t \in [0,1]} |W(t) + \frac{t}{2}|))$ is finite. Furthermore, using that $\int_{x_0}^{\infty} \exp(-\frac{1}{2}x^2) dx < \frac{1}{x_0} \exp(-\frac{x_0^2}{2})$ for all $x_0 > 0$, we have

$$\begin{aligned} & \mathbb{E} \left(\exp \left(- \left| W(k) + \frac{k}{2} \right| \right) \right) \\ &= \int_{-k/2}^{\infty} e^{-x-k/2} \frac{1}{\sqrt{2\pi k}} \exp \left(-\frac{x^2}{2k} \right) dx + \int_{-\infty}^{-k/2} e^{x+k/2} \frac{1}{\sqrt{2\pi k}} \exp \left(-\frac{x^2}{2k} \right) dx \\ &= \frac{1}{\sqrt{2\pi k}} \left(\int_{-k/2}^{\infty} \exp \left(-\frac{(x+k)^2}{2k} \right) dx + \exp(k) \int_{-\infty}^{-k/2} \exp \left(-\frac{(x-k)^2}{2k} \right) dx \right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_{\frac{\sqrt{k}}{2}}^{\infty} \exp \left(-\frac{x^2}{2} \right) dx + \exp(k) \int_{\frac{3\sqrt{k}}{2}}^{\infty} \exp \left(-\frac{x^2}{2} \right) dx \right) \\ &\leq \frac{2}{\sqrt{2\pi k}} \exp \left(-\frac{k}{8} \right) + \exp(k) \frac{2}{3\sqrt{2\pi k}} \exp \left(-\frac{9k}{8} \right) = \frac{8}{3\sqrt{2\pi k}} \exp \left(-\frac{k}{8} \right). \end{aligned}$$

All in all, we get

$$\mathbb{E} \left(\sum_{z \in \mathbb{Z}} \exp \left(- \inf_{t \in [0,1]} R(z+t) \right) \right) \leq 2C \limsup_{n \rightarrow \infty} \left(1 + \sum_{k=1}^n \exp \left(-\frac{k}{8} \right) \right) < \infty.$$

As we have an integrable majorizing random variable, dominated convergence yields

$$\begin{aligned} & \lim_{\|T\| \rightarrow 0} \mathbb{E} \left(\int_{\mathbb{R}} \left| \frac{1}{2} F_T(s-t) - \frac{1}{2} F(s-t) \right| dt \right) \\ &= \mathbb{E} \left(\lim_{\|T\| \rightarrow 0} \int_{\mathbb{R}} \left| \frac{1}{2} F_T(s-t) - \frac{1}{2} F(s-t) \right| dt \right) = 0, \quad s \in \mathbb{R}, \end{aligned} \quad (5.36)$$

as F is continuous and integrable.

By Theorem 2.1 in [Stoev and Taqqu \(2005\)](#), Equation (5.36) is equivalent to $Z_T(s) \xrightarrow{p} Z(s)$ for all $s \in \mathbb{R}$. As

$$\mathbb{P}(\|(Z_t(s_1), \dots, Z_T(s_m)) - (Z(s_1), \dots, Z(s_m))\| > \varepsilon) \leq \sum_{i=1}^m \mathbb{P} \left(|Z_T(s_i) - Z(s_i)| > \frac{\varepsilon}{m} \right),$$

we get that all finite-dimensional marginal distributions of Z_T converge in probability and therefore also weakly. Thus, for Borel sets $A \subset \mathbb{R}$, $B \subset \mathbb{R}^n$ with $\mathbb{P}(Z(\mathbf{t}) \in B) > 0$ and $\mathbb{P}(Z(t_0) \in \partial A) = \mathbb{P}(Z(\mathbf{t}) \in \partial B) = 0$, we get

$$\begin{aligned} & \lim_{\|T\| \searrow 0} \mathbb{P}(Z_T(t_0) \in A \mid Z_T(\mathbf{t}) \in B) = \lim_{\|T\| \searrow 0} \frac{\mathbb{P}(Z_T(t_0) \in A, Z_T(\mathbf{t}) \in B)}{\mathbb{P}(Z_T(\mathbf{t}) \in B)} \\ &= \frac{\mathbb{P}(Z(t_0) \in A, Z(\mathbf{t}) \in B)}{\mathbb{P}(Z(\mathbf{t}) \in B)} = \mathbb{P}(Z(t_0) \in A \mid Z(\mathbf{t}) \in B) \end{aligned}$$

which completes the proof. \square

Still, for any fixed T , the range of $\frac{1}{2}F_T$ is uncountable. Therefore, we have to use the approximation introduced in Section 5.6.

We compare this approximation to conditional sampling based on the approach of Wang and Stoev (2011) and the Gaussian marginals approach, see Section 5.4.

To compare these procedures, we simulate $K = 100$ independent samples of Z on the set $\{t_0, t_1, t_2, t_3, t_4\}$ with $t_0 = 0$, $t_1 = -2$, $t_2 = -1$, $t_3 = 1$ and $t_4 = 2$. We calculate the CRPS by sampling $k = 50$ times from the (approximate) conditional distribution of $\log(Z(t_0))$ given $Z(\mathbf{t})$. Let $CRPS_{K,1a} / MAE_{K,1a}$ and $CRPS_{K,1b} / MAE_{K,1b}$ denote the CRPS / MAE based on two variants of conditional sampling of the Poisson point process for K samples, $CRPS_{K,2} / MAE_{K,2}$ the CRPS / MAE of the approach by Wang and Stoev (2011) and $CRPS_{K,3} / MAE_{K,3}$ the CRPS / MAE based on transformation to Gaussian marginals. For the Poisson point process approach, we chose $N = 1000$ for the number of shape functions on the grid $T = \{-7, -6.95, -6.9, \dots, 6.9, 6.95, 7\}$. However, if we restrict ourselves to a finite number of shape functions, the intersection set $I_B(\mathbf{z})$ with $|B| \geq 3$ is most likely empty, even though $Z(t_i)$, $i \in B$, may be determined by the same $(s, u, f) \in \Pi$. Therefore, we do not only consider “exact” intersections, but also intersections which occur if the function values differ up to a given tolerance, i.e. we assume $(t, y, f) \in I_B(\mathbf{z})$ with $y = \min_{i=1, \dots, n} \frac{z_i}{f(t_i - t)}$ if

$$\frac{z_i}{f(t_i - t)} < \min \{y + \varepsilon, y(1 + \varepsilon)\} \iff i \in B$$

for some given tolerance $\varepsilon > 0$. The simulation study is performed for $\varepsilon = 10^{-6}$ ($CRPS_{K,1a} / MAE_{K,1a}$) and for $\varepsilon = 10^{-3}$ ($CRPS_{K,1b} / MAE_{K,1b}$). Thus, by construction, $CRPS_{K,1a} / MAE_{K,1a}$ does not contain any intersection of more than two curves, but $CRPS_{K,1b} / MAE_{K,1b}$ does. For the Wang-Stoev-approach, we choose $M = 5$, $h = 0.05$ and the same $N = 1000$ shape functions. The parameters are chosen such that the first and the second method have similar running times. For the transformation to Gaussian marginals, we can calculate the bivariate marginal distributions of $Y(\cdot) = \Phi^{-1}(\Phi_1(Z(\cdot)))$ and get

$$\mathbb{P}(Y(t_1) \leq y_1, Y(t_2) \leq y_2) = \prod_{i=1}^2 \Phi(y_i) \Phi\left(\frac{1}{2}\sqrt{|t_1 - t_2|} - (-1)^i \log(\log \Phi(y_1) / \log \Phi(y_2)) / \sqrt{|t_1 - t_2|}\right).$$

Deriving the bivariate density, we obtain

$$\begin{aligned} & \frac{\partial^2}{\partial y_1 \partial y_2} \mathbb{P}(Y(t_1) \leq y_1, Y(t_2) \leq y_2) \\ = & \mathbb{P}(Y(t_1) \leq y_1, Y(t_2) \leq y_2) \frac{\varphi(y_1)\varphi(y_2)}{\Phi(y_1)\Phi(y_2)} \\ & \cdot \left[\prod_{i=1,2} \Phi\left(\sqrt{|t_1 - t_2|}/2 - (-1)^i \log(\log \Phi(y_1) / \log \Phi(y_2)) / \sqrt{|t_1 - t_2|}\right) \right. \\ & \left. + \frac{\exp\left(-\frac{1}{2}\left(\frac{|t_1 - t_2|}{4} + \log(\log \Phi(y_1) \log \Phi(y_2)) + \log^2\left(\frac{\log \Phi(y_1)}{\log \Phi(y_2)}\right) / |t_1 - t_2|\right)\right)}{\sqrt{2\pi|t_1 - t_2|}} \right]. \end{aligned}$$

This enables us to compute the covariances $\mathbb{E}Y(t_1)Y(t_2)$ by numerical integration. However, as this is very time-consuming and inaccurate, we just fit a covariance function $\tilde{\kappa}_{\nu,c}$ of Whittle-Matérn type by maximum likelihood estimation in `RandomFields`.

	$m = 1a$	$m = 1b$	$m = 2$	$m = 3$
$CRPS_{K,m}$	-0.347	-0.413	-0.332	-0.341
$MAE_{K,m}$	0.487	0.526	0.466	0.469

Table 5.2: Results of the simulation study for the Brown-Resnick process with $N = 1000$ and $K = 100$.

The results are presented in Table 5.2. Here, all the methods have almost the same accuracy, at least if $\varepsilon = 10^{-6}$ ($m = 1a$). However, the last method runs much faster than the others. Furthermore, we notice the difference between $CRPS_{K,1a}$ and $CRPS_{K,1b}$ indicating that considering approximate intersections of at least three curves yields worse results. This is because these intersections involve incorrect shape functions. Furthermore, they lead to degenerated conditional distributions which are not supposed to occur in the case of the Brown-Resnick process. Thus, in this case, the approximation of the mixed moving maxima process seems to be appropriate only if we do not consider intersections of three or more curves.

5.8 The Discretized Case

By now, we have considered the general model (5.1). The procedure we proposed is exact in the case of a finite number of shape functions which are sufficiently smooth. However, as the example of the Brown-Resnick process in Section 5.7 illustrates, we may run into problems if these assumptions are violated.

Now, we modify our general model (5.1) and use a discretized version

$$Z(t) = \max_{i \in \mathbb{N}} \left(U_i^{(p)} \cdot F_i^{(p)}(t - S_i^{(p)}) \right), \quad t \in p\mathbb{Z}^d, \quad (5.37)$$

where $(S_i^{(p)}, U_i^{(p)}, F_i^{(p)})_{i \in \mathbb{N}}$ are the points of a simple Poisson point process $\Pi^{(p)}$ on $p\mathbb{Z}^d \times (0, \infty) \times G$ with $p > 0$ and a countable set $G \subset (0, \infty)^{p\mathbb{Z}^d}$. The intensity measure of $\Pi^{(p)}$ is given by

$$\Lambda(\{s\} \times B \times \{g\}) = \sum_{z \in \mathbb{Z}^d} \delta_{pz}(ds) \times \int_B u^{-2} du \times \mathbb{P}_F(\{g\})$$

where \mathbb{P}_F is a probability measure belonging to a G -valued random variable F with $\mathbb{E}(\sum_{z \in \mathbb{Z}^d} F(pz)) = 1$.

Using the same notation as before, we get the same results as in Section 5.2. However, all the calculations can be done explicitly without any further assumptions on $f \in G$. We get the following results:

Proposition 5.22. *Let $B = \{i\} \subset \{1, \dots, n\}$, $\mathbf{z} \in \mathbb{R}^n$ and*

$$D_i(\mathbf{z}) = \{(x, f) \in p\mathbb{Z}^d \times G : (x, y, f) \in I_{\{i\}}(\mathbf{z}) \text{ for some } y \in \mathbb{R}\}.$$

Then, we have

$$\Lambda(I_{\{i\}}^{(m)}(\mathbf{z})) = \frac{1}{2^m} \sum_{(x, f) \in D_i(\mathbf{z})} \frac{f(t_i - x)}{z_i} \mathbb{P}_F(\{f\}) + o(2^{-m}).$$

Proposition 5.23. *Let $B \in 2^{\{1, \dots, n\}} \setminus \emptyset$, $|B| > 1$ and $\mathbf{z} \in \mathbb{R}^n$ such that*

$$I_A(\mathbf{z}) = \{(x_j, y_j, f_j), j = 1, \dots, l(\mathbf{z})\}$$

with $l(\mathbf{z}) > 0$. Then, for m large enough, we have

$$\Lambda(I_B^{(m)}(\mathbf{z})) = \sum_{j=1}^{l(\mathbf{z})} \frac{1}{y_j} \mathbb{P}_F(\{f_j\}) \cdot \left(\min_{i \in A} \frac{2^m z_i}{j_m(z_i)} - \max_{i \in A} \frac{2^m z_i}{j_m(z_i) + 1} \right)$$

In particular, $\Lambda(I_B^{(m)}(\mathbf{z})) \in \mathcal{O}(2^{-m})$, but $\Lambda(I_B^{(m)}(\mathbf{z})) \notin \mathcal{O}(2^{-m(1+\varepsilon)})$ for any $\varepsilon > 0$.

By these formulae, all the scenario probabilities can be calculated. As the intensity of each intersection set has the same rate of convergence, only scenarios with minimal $\Pi(K_{\mathbf{t}, \mathbf{z}})$ occur.

We note that our model is very close to the model investigated by [Wang and Stoev \(2011\)](#). To see this, we calculate that

$$\mathbb{P}(Z(t_1) \leq z_1, \dots, Z(t_n) \leq z_n) = \exp \left(- \sum_{f \in G} \sum_{z \in \mathbb{Z}^d} \max_{i=1, \dots, n} \frac{f(t_i - pz) \mathbb{P}_F(\{f\})}{z_i} \right).$$

Therefore, we get that

$$Z \stackrel{d}{=} \max_{z \in \mathbb{Z}^d} \max_{f \in F} \left(f(\cdot - pz) \mathbb{P}_F(\{f\}) Z_f^{(z)} \right),$$

where the random variables $Z_f^{(z)}$, $z \in \mathbb{Z}^d$, $f \in G$, are independently standard Fréchet distributed.

This means, the model (5.37) is a max-linear model and the exact conditional distribution can be calculated via the algorithm of [Wang and Stoev \(2011\)](#) if G is finite and the support of each $f \in G$ is finite. However, the algorithm of conditional sampling of the Poisson point process and the algorithm of [Wang and Stoev \(2011\)](#) do not work in exactly the same way. According to the latter algorithm one samples from each random variable $Z_f^{(z)}$. This procedure corresponds to simulating the largest point of Π in $(0, \infty) \times \{pz\} \times \{f\}$ for each $z \in \mathbb{Z}$, $f \in G$. The first algorithm includes the simulation of points in Π until a terminating condition given in Theorem 4 of [Schlather \(2002\)](#) is met. Computational experiments show that both algorithms yield identical results. Also the technical results are related.

For example, the occurrence of a scenario $J \subset \mathbb{Z}^d \times G$ (in the notation of Wang/Stoev) corresponds to the event that Π_2 consists of $|J|$ points $(pz, \min_{i=1, \dots, n} \frac{Z(t_i)}{f(t_i - pz)}, f)$ with $(z, f) \in J$. By this correspondence, the statements

- $\Pi(K_{\mathbf{t}, \mathbf{z}})$ is minimal a.s.
- an occurring hitting scenario J satisfies $|J| = r(\mathcal{J}(A, \mathbf{x}))$ a.s. ([Wang and Stoev, 2011](#))

are equivalent, both claiming that the number of points generating the observation (\mathbf{t}, \mathbf{z}) is minimal. Hence, although the approaches look quite different, there are similar observations and results in [Wang and Stoev \(2011\)](#) and in this section.

6 Conditional Sampling of Brown-Resnick and Extremal Gaussian Processes

Besides mixed moving maxima processes, there are basically two further models for max-stable processes which are frequently used in applications: so-called extremal Gaussian processes (Schlather 2002; for applications see Blanchet and Davison 2011, Davison and Gholamrezaee 2011, for example) and Brown-Resnick processes (Kablichko et al. 2009; for applications see Buishand et al. 2008 and Eckert et al. 2011), which we already considered in Chapter 4.

In this chapter, we will analyse the problem of conditional sampling for both types of processes. As both models are also based on Poisson point processes, we will adapt arguments and methods from Chapter 5 to this framework. However, as the finite-dimensional marginal distributions of extremal Gaussian and Brown-Resnick processes are absolutely continuous w.r.t. the Lebesgue measure, formulae for conditional distributions can be derived more directly using derivatives of the exponent measure. Recently, this way has been taken successfully by Dombry et al. (2011) and Dombry and Ribatet (2012) based on the work of Dombry and Eyi-Minko (2011). But, as we have seen in Chapter 5, the results of Dombry and Eyi-Minko (2011) cannot be applied directly if the finite-dimensional marginals of the max-stable random field are not absolutely continuous w.r.t. the Lebesgue measure. Therefore, we will independently derive the same results as Dombry et al. (2011) and Dombry and Ribatet (2012), using the arguments from Chapter 5 which can be used in a more general context.

After a short introduction to a general model pooling both types of processes, we consider a partition of the underlying Poisson point process (Section 6.1). Following the approach of Chapter 5, in Section 6.2, we analyse “blurred” intersection sets to calculate probabilities for different scenarios. Finally, we deal with the distribution of those points of the underlying Poisson point process which generate the observations within a scenario (Section 6.3).

Let $\tilde{\Pi} = \sum_{i \in \mathbb{N}} \delta_{U_i}$ be a Poisson point process on $(0, \infty)$ with intensity $u^{-2} du$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Independently, let $\{W_i(t), t \in \mathbb{R}^d\}_{i \in \mathbb{N}}$ be independent copies of a Gaussian random field $\{W(t), t \in \mathbb{R}^d\}$, which will be specified separately for the two types of processes.

Then, the extremal Gaussian process (Schlather, 2002) is defined by

$$Z_S(t) = \max_{i \in \mathbb{N}} \left(U_i \cdot \max \left\{ \sqrt{2\pi} W_i(t), 0 \right\} \right), \quad t \in \mathbb{R}^d, \quad (6.1)$$

where W is a stationary zero mean Gaussian random field with standard variance.

The Brown-Resnick process is given by

$$Z_{BR}(t) = \max_{i \in \mathbb{N}} \left(U_i \cdot \exp \left(W_i(t) - \frac{\sigma^2(t)}{2} \right) \right), \quad t \in \mathbb{R}^d, \quad (6.2)$$

where W is a zero mean Gaussian random field with stationary increments and variance $\sigma^2(\cdot)$. Recall that the distribution does not depend on the variance $\sigma^2(\cdot)$, but only on the variogram γ of the underlying Gaussian process W , i.e.

$$Z_{BR}(t) \stackrel{d}{=} \max_{i \in \mathbb{N}} \left(U_i \cdot \exp \left(W_i(t) - W_i(0) - \frac{\gamma(t)}{2} \right) \right), \quad t \in \mathbb{R}^d,$$

(cf. Thm. 4.1). Henceforth, we assume $W(0) = 0$ a.s. for the Brown-Resnick process. Both the extremal Gaussian process and the Brown-Resnick process define stationary max-stable processes with standard Fréchet margins. Pooling both models, we write

$$Z(t) = \max_{i \in \mathbb{N}} (U_i \cdot f(t, W_i(t))), \quad t \in \mathbb{R}^d, \quad (6.3)$$

where the link function $f : \mathbb{R}^d \times \mathbb{R} \rightarrow [0, \infty)$ either has the form $f(t, y) = \max\{\sqrt{2\pi}y, 0\}$ (for the extremal Gaussian process) or $f(t, y) = \exp(y - \gamma(t)/2)$ (for the Brown-Resnick process).

This chapter aims to sample from $Z(\cdot)$ conditional on its values at some fixed locations $t_1, \dots, t_n \in \mathbb{R}^d$. In the following, we will assume that the joint distribution of $(W(t_1), \dots, W(t_n))$ is non-degenerated and that $W(\cdot)$ has continuous sample paths. Under these assumptions, we will determine the distribution of the Poisson point process

$$\Pi = \sum_{i \in \mathbb{N}} \delta_{(U_i, W_i)}$$

on $(0, \infty) \times C(\mathbb{R}^d)$ conditional on $Z(t_1), \dots, Z(t_n)$. The intensity measure Λ of Π is given by

$$\Lambda(A \times B) = \int_A u^{-2} du \cdot \mathbb{P}_W(B), \quad A \subset (0, \infty), \quad B \in \mathcal{C}(\mathbb{R}^d), \quad (6.4)$$

where \mathbb{P}_W is the probability measure belonging to $W(\cdot)$ and $\mathcal{C}(\mathbb{R}^d)$ denotes the Borel- σ -algebra on $C(\mathbb{R}^d)$ w.r.t. uniform convergence on compact sets (see Section 3.3).

6.1 Random Partition of Π

For $A \in 2^{\{1, \dots, n\}}$, we consider the random measure Π_A defined by

$$\Pi_A(\cdot) = \Pi(\cdot \cap \{(u, w) \in (0, \infty) \times C(\mathbb{R}^d) : uf(t_i, w(t_i)) = Z(t_i), i \in A, \\ uf(t_j, w(t_j)) < Z(t_j), j \notin A\}).$$

We will show that the random sets Π_A are point processes.

Proposition 6.1. *For any $A \in 2^{\{1, \dots, n\}}$, the random set Π_A is a point process.*

Proof. First, we note that the mapping

$$\Psi : (0, \infty) \times C(\mathbb{R}^d) \rightarrow C(\mathbb{R}^d), \quad (u, w(\cdot)) \mapsto uf(\cdot, w(\cdot))$$

is well-defined as $W(\cdot)$ has continuous sample paths. This also implies the continuity of $\gamma(\cdot)$. Furthermore, Ψ is $((\mathcal{B} \cap (0, \infty)) \times C(\mathbb{R}^d), \mathcal{C}(\mathbb{R}^d))$ -measurable since the mapping

$w(\cdot) \mapsto f(\cdot, w(\cdot))$ is continuous, i.e. uniform convergence of a sequence $\{w_n(\cdot)\}_{n \in \mathbb{N}} \subset C(\mathbb{R}^d)$ on compact sets implies uniform convergence of $\{f(\cdot, w_n(\cdot))\}_{n \in \mathbb{N}} \subset C(\mathbb{R}^d)$ on compact sets.

Now, let $M \in (\mathcal{B} \cap (0, \infty)) \times \mathcal{C}(\mathbb{R}^d)$ be bounded and $k \in \mathbb{N}$. Then,

$$\begin{aligned} & \{\omega \in \Omega : \Pi_A(M) = k\} \\ &= \bigcup_{n_0=1}^{\infty} \bigcap_{m=n_0}^{\infty} \bigcup_{\mathbf{y} \in \mathbb{Q}^n} \left\{ \omega \in \Omega : Z(\mathbf{t}) \in \times_{i=1}^n \left(y_i - \frac{1}{m}, y_i + \frac{1}{m} \right), \right. \\ & \quad \left. \Pi \left(\left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : uf(t_i, w(t_i)) \in \left(y_i - \frac{1}{m}, y_i + \frac{1}{m} \right), i \in A, \right. \right. \right. \\ & \quad \left. \left. \left. uf(t_j, w(t_j)) < y_j - \frac{1}{m}, j \notin A \right\} \cap M \right) = k \right\} \end{aligned}$$

is measurable, i.e. Π_A is a point process. \square

Proposition 6.1 allows for considering the point process

$$\Pi_* = \bigcup_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \Pi_A$$

which consists of all the points within Π which contribute to the maximum (6.3) at locations t_1, \dots, t_n . In particular, we have $\sum_{A: i \in A} \Pi_A((0, \infty) \times C(\mathbb{R}^d)) \geq 1$ for all $i \in \{1, \dots, n\}$. By the definition of Z , there is no point of Π in the set

$$\mathcal{X}(Z(\mathbf{t})) = \{(u, w) \in (0, \infty) \times C(\mathbb{R}^d) : uf(t_i, w(t_i)) > Z(t_i) \text{ for some } i \in \{1, \dots, n\}\}.$$

Therefore, we get $\Pi = \Pi_* \cup \Pi_\emptyset$, which corresponds to the classification of Φ_K^+ and Φ_K^- in Dombry and Eyi-Minko (2011).

6.2 Blurred Sets

We will calculate the explicit distribution of Π_* and Π_\emptyset conditional on $Z(\mathbf{t}) = \mathbf{z}$. However, $Z(\mathbf{t}) = \mathbf{z}$ is an event of probability zero. To end up with a regular version of the conditional probability, we will use martingale arguments similar to Chapter 5. Most of the notation is the same as in Chapter 5, however, as we deal with different spaces, we recap and modify the definitions we need. Again, we consider filtrations

$$\mathcal{F}_m = \sigma(\{Z(t_i) \in (k_i \cdot 2^{-m}, (k_i + 1) \cdot 2^{-m}), i = 1, \dots, n\}, (k_1, \dots, k_n) \in \mathbb{N}^n)$$

for $m \in \mathbb{N}$. Furthermore, for $z \in (0, \infty)$, let $j_m(z) \in \mathbb{N}$ such that $z \in A_m(z) = (j_m(z) \cdot 2^{-m}, (j_m(z) + 1) \cdot 2^{-m}]$. Then, by Lévy's "Upward" Theorem (cf. Rogers and Williams, 2000, Thm. 50.3) we get that

$$\begin{aligned} & \mathbb{P}(\Pi_A(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_A(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \end{aligned} \quad (6.5)$$

for $B_A \in (\mathcal{B} \cap (0, \infty)) \times C(\mathbb{R}^d)$, $k_A \in \mathbb{N}_0$, $A \in 2^{\{1, \dots, n\}}$.

Conditioning on $Z(\mathbf{t}) \in A_m(\mathbf{z})$, the *blurred point processes* $\Pi_{A, \mathbf{z}}^{(m)}(\cdot) = \Pi(\cdot \cap I_A^{(m)}(\mathbf{z}))$ with

$$I_A^{(m)}(\mathbf{z}) = \left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : uf(t_i, w(t_i)) \in A_m(z_i), i \in A, \right. \\ \left. uf(t_j, w(t_j)) \leq \frac{j_m(z_j)}{2^m}, j \notin A \right\}$$

for $A \in 2^{\{1, \dots, n\}}$ (cf. Chapter 5), turn out to be useful tools to analyse the distribution of Π .

Similar to the decomposition above, conditional on $Z(\mathbf{t}) \in A_m(\mathbf{z})$, we have that $\Pi = \bigcup_{A \in 2^{\{1, \dots, n\}}} \Pi_A^{(m)}(\mathbf{z})$ since $Z(\mathbf{t}) \in A_m(\mathbf{z})$ implies that the set

$$\mathcal{X}_m(\mathbf{z}) = \left\{ (u, w) \in (0, \infty) : uf(t_i, w(t_i)) > \frac{j_m(z_i) + 1}{2^m} \text{ for some } i \in \{1, \dots, n\} \right\}$$

does not contain any points of Π .

Remark 6.2. For any $\mathbf{z} \in (0, \infty)^n$, the equivalence

$$Z(\mathbf{t}) \in A_m(\mathbf{z}) \iff \sum_{A: i \in A} \Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) \geq 1, i = 1, \dots, n \wedge \Pi(\mathcal{X}_m(\mathbf{z})) = 0$$

holds. In particular, $\Pi_{\emptyset, \mathbf{z}}^{(m)}$ is independent of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$ for any fixed $\mathbf{z} \in (0, \infty)^n$.

By Remark 6.2, the point processes $\{\Pi_{A, \mathbf{z}}^{(m)}, A \in 2^{\{1, \dots, n\}}\}$ are directly linked to $Z(\mathbf{t})$. To describe the joint distribution of $\Pi_{A, \mathbf{z}}^{(m)}$, $A \in 2^{\{1, \dots, n\}}$, we need to calculate the intensity of sets of type $\{(u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot f(t_i, w(t_i)) \in [a_i, b_i], i = 1, \dots, n\}$ for some $0 \leq a_i < b_i < \infty$.

Let $(W(t_1), \dots, W(t_n)) \sim \mathcal{N}(\mathbf{0}, \Sigma)$, $\Sigma \in \mathbb{R}^{n \times n}$ with $|\Sigma| = \det(\Sigma) > 0$. Then, in the case of extremal Gaussian processes, we get

$$\Lambda \left(\left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot \max\{\sqrt{2\pi}w(t_i), 0\} \in [a_i, b_i], i = 1, \dots, n \right\} \right) \\ = \int_0^\infty \frac{1}{u^2} \int_{\mathcal{X}_{i=1}^n(\tilde{a}_i, b_i)} \frac{1}{(2\pi u)^n |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y^T \Sigma^{-1} y}{2\pi u^2}\right) dy du,$$

where $\tilde{a}_i = a_i$ for $a_i > 0$ or $\tilde{a}_i = -\infty$ for $a_i = 0$.

Change of variables $v = u^{-2}$ and employing the definition of the Gamma function (Abramowitz and Stegun, 1965, Formula 6.1.1) lead to

$$\Lambda \left(\left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot \max\{\sqrt{2\pi}w(t_i), 0\} \in [a_i, b_i], i = 1, \dots, n \right\} \right) \\ = \int_{\mathcal{X}_{i=1}^n(\tilde{a}_i, b_i)} \int_0^\infty \frac{v^{\frac{n-1}{2}}}{2(2\pi)^n |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y^T \Sigma^{-1} y}{2\pi} v\right) dv dy$$

$$= \int_{\times_{i=1}^n (\tilde{a}_i, b_i]} \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n-1}{2}} (y^T \Sigma^{-1} y)^{\frac{n+1}{2}} |\Sigma|^{\frac{1}{2}}} dy = \int_{\times_{i=1}^n (\tilde{a}_i, b_i]} \frac{C(n)}{|\Sigma|^{\frac{1}{2}} (y^T \Sigma^{-1} y)^{\frac{n+1}{2}}} dy \quad (6.6)$$

with $C(n) = \Gamma(\frac{n+1}{2}) \sqrt{\pi}^{1-n}$ (cf. [Dombry and Ribatet, 2012](#)).

In the case of Brown-Resnick processes, we transform to Gumbel margins and get

$$\begin{aligned} & \Lambda \left(\left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot \exp \left(w(t_i) - \frac{\gamma(t_i)}{2} \right) \in [a_i, b_i], i = 1, \dots, n \right\} \right) \\ &= \int_0^\infty \frac{1}{u^2} \mathbb{P} \left(u \cdot \exp \left(W(t_i) - \frac{\gamma(t_i)}{2} \right) \in [a_i, b_i], i = 1, \dots, n \right) du \\ &= \int_{-\infty}^\infty e^{-u} \mathbb{P} \left(u + W(t_i) - \frac{\gamma(t_i)}{2} \in [\log a_i, \log b_i], i = 1, \dots, n \right) du \\ &= \int_{-\infty}^\infty e^{-u} \int_{\times_{i=1}^n [\tilde{a}_i, \tilde{b}_i]} \frac{1}{\sqrt{2\pi}^n |\Sigma|^{1/2}} \exp \left(-\frac{1}{2} (y - \mathbf{1}_n u)^T \Sigma^{-1} (y - \mathbf{1}_n u) \right) dy du \end{aligned}$$

where $\tilde{a}_i = \log(a_i) + \frac{\gamma(t_i)}{2}$, $\tilde{b}_i = \log(b_i) + \frac{\gamma(t_i)}{2}$ for $i = 1, \dots, n$. This yields

$$\begin{aligned} & \Lambda \left(\left\{ (u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot \exp \left(w(t_i) - \frac{\gamma(t_i)}{2} \right) \in [a_i, b_i], i = 1, \dots, n \right\} \right) \\ &= \int_{\times_{i=1}^n [\tilde{a}_i, \tilde{b}_i]} \int_{-\infty}^\infty \exp \left(-\frac{1}{2} (\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n) \left(\frac{y^T \Sigma^{-1} \mathbf{1}_n - 1}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} - u \right)^2 \right) du \\ & \quad \cdot \frac{1}{\sqrt{2\pi}^n |\Sigma|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} y^T \Sigma^{-1} y \right) \exp \left(\frac{1}{2} \frac{(y^T \Sigma^{-1} \mathbf{1}_n - 1)^2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) dy \\ &= \int_{\times_{i=1}^n [\tilde{a}_i, \tilde{b}_i]} \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} y^T \Sigma^{-1} y \right) \exp \left(\frac{1}{2} \frac{(y^T \Sigma^{-1} \mathbf{1}_n - 1)^2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) dy \\ &= \int_{\times_{i=1}^n [\tilde{a}_i, \tilde{b}_i]} \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} y^T \left(\Sigma^{-1} - \frac{\Sigma^{-1} \mathbf{1}_n \mathbf{1}_n^T \Sigma^{-1}}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) y \right) \\ & \quad \cdot \exp \left(-\frac{\mathbf{1}_n^T \Sigma^{-1}}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} y \right) \exp \left(\frac{1}{2} \frac{1}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) dy. \quad (6.7) \end{aligned}$$

To compare our result to [Dombry et al. \(2011\)](#), we apply a logarithmic transformation and get the density

$$\lambda(\mathbf{z}) = C \exp \left(-\frac{1}{2} (\log \mathbf{z})^T Q \log \mathbf{z} + L \log z \right) \prod_{i=1}^n \frac{1}{z_i}, \quad \mathbf{z} > \mathbf{0},$$

where

$$\begin{aligned} Q &= \Sigma^{-1} - \frac{\Sigma^{-1} \mathbf{1}_n \mathbf{1}_n^T \Sigma^{-1}}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n}, \\ L &= \frac{1}{2} \left(\frac{\mathbf{1}_n^T \Sigma^{-1} \gamma(\mathbf{t}) - 2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \mathbf{1}_n^T - \gamma(\mathbf{t})^T \right) \Sigma^{-1}, \\ C &= \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(-\frac{1}{8} \gamma(\mathbf{t})^T Q \gamma(\mathbf{t}) + \frac{1}{2} \frac{\mathbf{1}_n^T \Sigma^{-1} \gamma(\mathbf{t}) + 1}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) \end{aligned}$$

$$= \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(\frac{1}{2} \frac{(\mathbf{1}_n^T \Sigma^{-1} \frac{\gamma(\mathbf{t})}{2} - 1)^2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} - \frac{1}{2} \frac{\gamma(\mathbf{t})^T}{2} \Sigma^{-1} \frac{\gamma(\mathbf{t})}{2} \right).$$

Note that the normalizing constant C differs from the one calculated by [Dombry and Eyi-Minko \(2011\)](#) and [Dombry et al. \(2011\)](#), which is

$$\tilde{C} = \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(\frac{1}{2} \frac{(\mathbf{1}_n^T \Sigma^{-1} \gamma(\mathbf{t}) - 1)^2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} - \frac{1}{2} \gamma(\mathbf{t})^T \Sigma^{-1} \gamma(\mathbf{t}) \right).$$

This difference also affects the scenario probabilities calculated later and yields a different conditional distribution.

We are interested in the asymptotic behaviour of these intensity measures for $|b_i - a_i| \rightarrow 0$ for $i \in A \subset \{1, \dots, n\}$. The assessments, which need to be done for specifying the asymptotics, rely on the following lemma.

Lemma 6.3. *Let $t_i \in \mathbb{R}^d$, $0 < a_i < b_i$ for $i = 1, \dots, n$ and $(W(t_i))_{i=1}^n \sim \mathcal{N}(\mathbf{0}, \Sigma)$ with $|\Sigma| > 0$. Furthermore, let $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$, $y_A = (y_i)_{i \in A}$ for any vector $y = (y_1, \dots, y_n)^T \in \mathbb{R}^n$. Then, the functions*

$$\begin{aligned} \Psi_S &: X_{i \in A} [a_i, b_i] \rightarrow (0, \infty), \quad y_A \mapsto \int_{X_{i \notin A} (0, b_i]} \frac{C(n)}{|\Sigma|^{\frac{1}{2}} (y^T \Sigma^{-1} y)^{\frac{n+1}{2}}} dy_{A^c}, \\ \Psi_{BR} &: X_{i \in A} [\tilde{a}_i, \tilde{b}_i] \rightarrow (0, \infty), \\ & y_A \mapsto \int_{X_{i \notin A} (-\infty, \tilde{b}_i]} \frac{(\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n)^{-\frac{1}{2}}}{\sqrt{2\pi}^{n-1} |\Sigma|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} y^T \Sigma^{-1} y + \frac{1}{2} \frac{(y^T \Sigma^{-1} \mathbf{1}_n - 1)^2}{\mathbf{1}_n^T \Sigma^{-1} \mathbf{1}_n} \right) dy_{A^c}, \end{aligned}$$

are continuous. Here $\tilde{a}_i = \log a_i + \frac{\gamma(t_i)}{2}$ and $\tilde{b}_i = \log b_i + \frac{\gamma(t_i)}{2}$ for $i = 1, \dots, n$. For $A = \{1, \dots, n\}$, Ψ_S and Ψ_{BR} are understood as the mapping of y_A to the respective integrand.

Proof. First, we note that Ψ_S and Ψ_{BR} are strictly positive as they are integrals of strictly positive functions over a set of positive measure. The finiteness of the integrals will be shown as a byproduct of the proof of the continuity of Ψ_S and Ψ_{BR} . For showing the continuity, we first note that both functions have the form $y_A \mapsto \int \psi(y_A, y_{A^c}) dy_{A^c}$, and thus fit in the framework of parameter-dependent integrals. Therefore, we can recourse to well-known continuity results in this area, e.g. [Schilling \(2005\)](#), Thm. 11.4. As the continuity of $y_{A^c} \mapsto \psi(y_A, y_{A^c})$ is obvious for any $y_A \in \text{dom}(\Psi_S)$ and $y_A \in \text{dom}(\Psi_{BR})$, respectively, and as the function ψ is measurable in both cases, it suffices to show that there is an integrable majorant $\eta(y_{A^c})$ with $|\psi(y_A, y_{A^c})| < \eta(y_{A^c})$ for all $y_A \in \text{dom}(\Psi_S)$ with $y_{A^c} \in X_{i \notin A} (0, b_i]$ and $y_A \in \text{dom}(\Psi_{BR})$ with $y_{A^c} \in X_{i \notin A} (-\infty, \tilde{b}_i]$, respectively.

In both cases, the construction will be done in the following way: First, we use Equations (6.6) and (6.7) to write ψ as an integral w.r.t. some variable u ranging from 0 to ∞ . The integrand contains some multivariate Gaussian density $f_{(Y_1, \dots, Y_n)}$ which we decompose to $f_{Y_{A^c}} \cdot f_{Y_{i^*} | Y_{A^c}} \cdot f_{Y_{A \setminus \{i^*\}} | Y_{\{i^*\} \cup A^c}}$. The last factor is bounded by a constant, the second one by a function not depending on y_{i^*} . Finally, the integral is decomposed to show finiteness.

First, we construct a majorant for ψ_S : Let $i^* \in A$. We start with (6.6) and decompose the joint density to get

$$\begin{aligned}
 \psi_S(y_A, y_{A^c}) &\leq \int_0^\infty \frac{1}{u^2} \frac{1}{(2\pi u)^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y_{A^c}^T \Sigma_{A^c}^{-1} y_{A^c}}{2\pi u^2}\right) \\
 &\quad \cdot \frac{1}{(2\pi u)^{|A|} |\Sigma_{\{i^*\}|A^c}|^{\frac{1}{2}} |\Sigma_{A \setminus \{i^*\}|A^c \cup \{i^*\}}|^{\frac{1}{2}}} \\
 &\quad \cdot \exp\left(-\frac{1}{2} \frac{(y_{i^*} - \text{Cov}(W(t_{i^*}), (W(t_i))_{i \in A^c}) \Sigma_{A^c}^{-1} y_{A^c})^2}{2\pi u^2 \Sigma_{\{i^*\}|A^c}}}\right) du \\
 &\leq \int_0^\infty \frac{1}{u^2} \frac{1}{(2\pi u)^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y_{A^c}^T \Sigma_{A^c}^{-1} y_{A^c}}{2\pi u^2}\right) \\
 &\quad \cdot \frac{C}{u^{|A|-1}} \frac{1}{2\pi \tau u} \exp\left(-\frac{1}{2} \frac{(y_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right) du \quad (6.8)
 \end{aligned}$$

for an appropriate constant $C > 0$, the variance $\tau^2 := \Sigma_{\{i^*\}|A^c} > 0$ and the vector $\mathbf{c} := \text{Cov}(W(t_{i^*}), (W(t_i))_{i \in A^c}) \Sigma_{A^c}^{-1} \in \mathbb{R}^{n-|A|}$. Here, by an expression of the form Σ_{A_1} , we denote the covariance matrix of $(W(t_i))_{i \in A_1}$ and, by $\Sigma_{A_1|A_2}$, the covariance matrix of $(W(t_i))_{i \in A_1}$ conditional on $(W(t_i))_{i \in A_2}$ for sets $A_1, A_2 \subset \{1, \dots, n\}$.

Note that $\exp\left(-\frac{1}{2} \frac{(y_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right)$ is the only term in (6.8) that depends on y_A , and that it can be bounded by

$$\tilde{\eta}_S(y_A^c) = \begin{cases} \exp\left(-\frac{1}{2} \frac{(a_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right) + \exp\left(-\frac{1}{2} \frac{(b_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right), & \mathbf{c}^T y_{A^c} \notin [a_{i^*}, b_{i^*}] \\ 1, & \mathbf{c}^T y_{A^c} \in [a_{i^*}, b_{i^*}]. \end{cases}$$

Thus, we have

$$\psi_S(y_A, y_{A^c}) \leq \int_0^\infty \frac{1}{u^2} \frac{C \tilde{\eta}_S(y_{A^c})}{(2\pi)^{n-|A|+1} u^n \tau |\Sigma_{A^c}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y_{A^c}^T \Sigma_{A^c}^{-1} y_{A^c}}{2\pi u^2}\right) du := \eta_S(y_{A^c}).$$

Now, we show that η_S is integrable.

$$\begin{aligned}
 &\int_{\times_{i \in A} (0, b_i]} \eta_S(y_{A^c}) dy_{A^c} \\
 &\leq \int_0^\infty \frac{C}{u^{2+|A|-1}} \int_{\mathbb{R}^{n-|A|}} \frac{1}{(2\pi u)^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{y_{A^c}^T \Sigma_{A^c}^{-1} y_{A^c}}{2\pi u^2}\right) \\
 &\quad \cdot \frac{1}{2\pi \tau u} \left(\exp\left(-\frac{1}{2} \frac{(a_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right) + \exp\left(-\frac{1}{2} \frac{(b_{i^*} - \mathbf{c}^T y_{A^c})^2}{2\pi \tau^2 u^2}\right) \right) dy_{A^c} du \\
 &\quad + \int_0^\infty \frac{C}{2\pi \tau u^{2+|A|}} \int_{\{y_{A^c} \in \mathbb{R}^{n-|A|}: \mathbf{c}^T y_{A^c} \in [a_{i^*}, b_{i^*}]\}} \frac{1}{(2\pi u)^{n-|A|} |\Sigma|^{\frac{1}{2}}} \\
 &\quad \cdot \exp\left(-\frac{1}{2} \frac{y_{A^c}^T \Sigma_{A^c}^{-1} y_{A^c}}{2\pi u^2}\right) dy_{A^c} du
 \end{aligned}$$

$$\begin{aligned}
 &= \int_0^\infty \frac{C}{u^{2+|A|-1}} \int_{\mathbb{R}^{n-|A|}} \frac{1}{(2\pi u)^{n-|A|+1} |\Sigma_{A^c \cup \{i^*\}}|^\frac{1}{2}} \\
 &\quad \cdot \left(\exp\left(-\frac{1}{2} \frac{(y_{A^c}^T, a_{i^*}) \Sigma_{A^c \cup \{i^*\}}^{-1} (y_{A^c}^T, a_{i^*})^T}{2\pi u^2}\right) \right. \\
 &\quad \left. + \exp\left(-\frac{1}{2} \frac{(y_{A^c}^T, b_{i^*}) \Sigma_{A^c \cup \{i^*\}}^{-1} (y_{A^c}^T, b_{i^*})^T}{2\pi u^2}\right) \right) dy_{A^c} du \\
 &+ \int_0^\infty \frac{C}{2\pi\tau u^{2+|A|}} \mathbb{P}(\mathbf{u}\mathbf{c}^T \sqrt{2\pi}(W(t_i))_{i \in A^c} \in [a_{i^*}, b_{i^*}]) du \\
 &\leq \int_0^\infty \frac{C}{2\pi u^{2+|A|}} \left(\exp\left(-\frac{1}{2} \frac{a_{i^*}^2}{2\pi u^2}\right) + \exp\left(-\frac{1}{2} \frac{b_{i^*}^2}{2\pi u^2}\right) \right) du \\
 &+ \sum_{i \in A^c} \int_0^\infty \frac{C}{2\pi\tau u^{2+|A|}} \mathbb{P}\left(|W(t_i)| > \frac{a_{i^*}}{(n-|A|)\|\mathbf{c}\|_\infty \sqrt{2\pi}u}\right) du \\
 &\leq \int_0^\infty \frac{C}{2\pi u^{2+|A|}} \left(\exp\left(-\frac{1}{2} \frac{a_{i^*}^2}{2\pi u^2}\right) + \exp\left(-\frac{1}{2} \frac{b_{i^*}^2}{2\pi u^2}\right) \right) du + \int_\varepsilon^\infty \frac{C(n-|A|)}{2\pi\tau u^{2+|A|}} du \\
 &+ \int_0^\varepsilon \frac{2C\sqrt{2\pi}(n-|A|)\|\mathbf{c}\|_\infty}{2\pi\tau u^{2+|A|-1} a_{i^*}} \exp\left(-\frac{1}{2} \frac{a_{i^*}^2}{2\pi\|\mathbf{c}\|_\infty^2 (n-|A|)^2 u^2}\right) du \tag{6.9}
 \end{aligned}$$

for some sufficiently small $\varepsilon > 0$ as $1 - \Phi(x) \leq x^{-1}\varphi(x)$ for any $x > 0$. Here, Φ and φ are the standard normal distribution and density function, respectively. The integrals on the right hand side of (6.9) are finite. Therefore, Ψ_S is continuous.

Next, we assess ψ_{BR} : Again, with (6.7) and $i^* \in A$, the common density can be decomposed yielding that $\psi_{BR}(y_A, y_{A^c})$ can be bounded by

$$\begin{aligned}
 &\int_{-\infty}^\infty \frac{\exp(-u)}{\sqrt{2\pi}^{n-|A|} |\Sigma_{A^c}|^\frac{1}{2}} \exp\left(-\frac{1}{2} (y_{A^c} - \mathbf{1}_{n-|A|}u)^T \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}_{n-|A|}u)\right) \\
 &\quad \cdot \frac{1}{\sqrt{2\pi}^{|A|} |\Sigma_{\{i^*\}|_{A^c}}|^\frac{1}{2} |\Sigma_{A \setminus \{i^*\}|_{A^c \cup \{i^*\}}|^\frac{1}{2}} \\
 &\quad \cdot \exp\left(-\frac{(y_{i^*} - (u + \text{Cov}(W(t_{i^*}), (W(t_i))_{i \in A^c}) \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}_{n-|A|}u)))^2}{2\Sigma_{\{i^*\}|_{A^c}}}\right) du \\
 &\leq \int_{-\infty}^\infty \frac{\exp(-u)}{\sqrt{2\pi}^{n-|A|} |\Sigma_{A^c}|^\frac{1}{2}} \exp\left(-\frac{1}{2} (y_{A^c} - \mathbf{1}_{n-|A|}u)^T \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}_{n-|A|}u)\right) \\
 &\quad \cdot \frac{C}{\sqrt{2\pi}\tau} \exp\left(-\frac{(y_{i^*} - (1 - \mathbf{c}^T \mathbf{1}_{n-|A|})u - \mathbf{c}^T y_A^c)^2}{2\tau^2}\right) du \tag{6.10}
 \end{aligned}$$

for an appropriate constant $C > 0$, the variance $\tau^2 := \Sigma_{\{i^*\}|_{A^c}} > 0$ and the vector $\mathbf{c} := \text{Cov}(W(t_{i^*}), (W(t_i))_{i \in A^c}) \Sigma_{A^c}^{-1} \in \mathbb{R}^{n-|A|}$. To dispose of the dependency on y_A , we bound the term $\exp\left(-\frac{(y_{i^*} - (1 - \mathbf{c}^T \mathbf{1}_{n-|A|})u - \mathbf{c}^T y_A^c)^2}{2\tau^2}\right)$ in (6.10) uniformly for all $y_A \in \chi_{i \in A}[\tilde{a}_i, \tilde{b}_i]$ by

$$\tilde{\eta}_{BR}(y_A^c, u) = \begin{cases} \exp\left(-\frac{(\tilde{a}_{i^*} - (1 - \mathbf{c}^T \mathbf{1})u - \mathbf{c}^T y_{A^c}^c)^2}{2\tau^2}\right) + \exp\left(-\frac{(\tilde{b}_{i^*} - (1 - \mathbf{c}^T \mathbf{1})u - \mathbf{c}^T y_{A^c}^c)^2}{2\tau^2}\right), & (1 - \mathbf{c}^T \mathbf{1})u + \mathbf{c}^T y_{A^c} \notin [\tilde{a}_{i^*}, \tilde{b}_{i^*}] \\ 1, & (1 - \mathbf{c}^T \mathbf{1})u + \mathbf{c}^T y_{A^c} \in [\tilde{a}_{i^*}, \tilde{b}_{i^*}] \end{cases} \cdot$$

Thus, we get

$$\begin{aligned} \psi_{BR}(y_A, y_{A^c}) &\leq \int_{-\infty}^{\infty} \frac{C\tilde{\eta}_{BR}(y_{A^c}, u)}{\sqrt{2\pi\tau}} \frac{\exp(-u)}{\sqrt{2\pi}^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \\ &\quad \cdot \exp\left(-\frac{1}{2}(y_{A^c} - \mathbf{1}u)^T \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}u)\right) du =: \eta_{BR}(y_{A^c}). \end{aligned}$$

We assess the integral $\int_{\chi_{i \notin A}[\tilde{a}_i, \tilde{b}_i]} \eta_{BR}(y_{A^c}) dy_{A^c}$ by

$$\begin{aligned} &\int_{-\infty}^{\infty} C e^{-u} \int_{\mathbb{R}^{n-|A|}} \frac{1}{\sqrt{2\pi}^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(y_{A^c} - \mathbf{1}u)^T \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}u)\right) \\ &\quad \cdot \frac{1}{\sqrt{2\pi\tau}} \left(\exp\left(-\frac{(\tilde{a}_{i^*} - (1 - \mathbf{c}^T \mathbf{1})u - \mathbf{c}^T y_{A^c})^2}{2\tau^2}\right) \right. \\ &\quad \left. + \exp\left(-\frac{(\tilde{b}_{i^*} - (1 - \mathbf{c}^T \mathbf{1})u - \mathbf{c}^T y_{A^c})^2}{2\tau^2}\right) \right) dy_{A^c} du \\ &+ \int_{-\infty}^{\infty} \frac{C e^{-u}}{\sqrt{2\pi\tau}} \int_{\{y_{A^c}: (1 - \mathbf{c}^T \mathbf{1})u + \mathbf{c}^T y_{A^c} \in [\tilde{a}_{i^*}, \tilde{b}_{i^*}]\}} \frac{1}{\sqrt{2\pi}^{n-|A|} |\Sigma_{A^c}|^{\frac{1}{2}}} \\ &\quad \cdot \exp\left(-\frac{1}{2}(y_{A^c} - \mathbf{1}_{n-|A|}u)^T \Sigma_{A^c}^{-1} (y_{A^c} - \mathbf{1}_{n-|A|}u)\right) dy_{A^c} du \\ &= \int_{-\infty}^{\infty} C e^{-u} \int_{\mathbb{R}^{n-|A|}} \frac{1}{\sqrt{2\pi}^{n-|A|+1} |\Sigma_{A^c \cup \{i^*\}}|^{\frac{1}{2}}} \\ &\quad \left(\exp\left(-\frac{1}{2} \begin{pmatrix} y_{A^c} - \mathbf{1}u \\ \tilde{a}_{i^*} - u \end{pmatrix}^T \Sigma_{A^c \cup \{i^*\}}^{-1} \begin{pmatrix} y_{A^c} - \mathbf{1}u \\ \tilde{a}_{i^*} - u \end{pmatrix}\right) \right. \\ &\quad \left. + \exp\left(-\frac{1}{2} \begin{pmatrix} y_{A^c} - \mathbf{1}u \\ \tilde{b}_{i^*} - u \end{pmatrix}^T \Sigma_{A^c \cup \{i^*\}}^{-1} \begin{pmatrix} y_{A^c} - \mathbf{1}u \\ \tilde{b}_{i^*} - u \end{pmatrix}\right) \right) dy_{A^c} du \\ &+ \int_{-\infty}^{\infty} \frac{C e^{-u}}{\sqrt{2\pi\tau}} \mathbb{P}(u + \mathbf{c}^T(W(t_i))_{i \in A^c} \in [\tilde{a}_{i^*}, \tilde{b}_{i^*}]) du \\ &\leq \int_{-\infty}^{\infty} C e^{-u} \frac{1}{\sqrt{2\pi\gamma(t_{i^*})}} \left(\exp\left(-\frac{(\tilde{a}_{i^*} - u)^2}{2\gamma(t_{i^*})}\right) + \exp\left(-\frac{(\tilde{b}_{i^*} - u)^2}{2\gamma(t_{i^*})}\right) \right) du \\ &\quad + \int_{-\infty}^{\infty} \frac{C e^{-u}}{\sqrt{2\pi\tau}} \mathbb{P}(\mathbf{c}^T(W(t_i))_{i \in A^c} > \tilde{a}_{i^*} - u) du \\ &= C \exp\left(\frac{\gamma(t_{i^*})}{2}\right) \left(\exp(-\tilde{a}_{i^*}) + \exp(-\tilde{b}_{i^*}) \right) + \int_{u_0}^{\infty} \frac{C e^{-u}}{\sqrt{2\pi\tau}} du \\ &\quad + \int_{-\infty}^{u_0} \frac{C e^{-u}}{\sqrt{2\pi\tau}} \sum_{i \in A^c} \mathbb{P}\left(|W(t_i)| > \frac{\tilde{a}_{i^*} - u}{(n - |A|)\|\mathbf{c}\|_{\infty}}\right) du \\ &\leq C \exp\left(\frac{\gamma(t_{i^*})}{2}\right) \left(\exp(-\tilde{a}_{i^*}) + \exp(-\tilde{b}_{i^*}) \right) + \frac{C e^{-u_0}}{\sqrt{2\pi\tau}} \\ &\quad + \frac{2C}{\sqrt{2\pi\tau}} \sum_{i \in A^c} \int_{-\infty}^{u_0} e^{-u} \frac{(n - |A|)\|\mathbf{c}\|_{\infty} \sqrt{\gamma(t_i)}}{\sqrt{2\pi}(\tilde{a}_{i^*} - u)} \cdot \exp\left(-\frac{1}{2\gamma(t_i)} \left(\frac{\tilde{a}_{i^*} - u}{(n - |A|)\|\mathbf{c}\|_{\infty}}\right)^2\right) du \end{aligned}$$

for $u_0 < 0$ and $|u_0|$ large enough, where we used that $1 - \Phi(x) \leq x^{-1}\varphi(x)$ for all $x > 0$. As these integrals are finite, Ψ_{BR} is continuous. \square

Lemma 6.4 describes the joint distribution of $\Pi_A^{(m)}$, $A \in 2^{\{1, \dots, n\}}$.

Lemma 6.4. *For any $\mathbf{t} \in \mathbb{R}^{d \cdot n}$, $\mathbf{z} \in \mathbb{R}^n$, $\{k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\} \subset \mathbb{N}$ and Λ defined as in (6.4), we have*

$$\begin{aligned} & \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \exp(-\Lambda(\mathcal{X}(\mathbf{z}))) \cdot \left(\prod_{|A| \geq 1} 2^{|A|k_A} \right)^{-m} \cdot \left(\prod_{|A| \geq 1} \frac{c_A(\mathbf{z})^{k_A}}{k_A!} + o(1) \right), \end{aligned} \quad (6.11)$$

where

$$c_A(\mathbf{z}) = c_A^S(\mathbf{z}) = \Psi_S(\mathbf{z}_A)$$

in case of extremal Gaussian processes and

$$c_A(\mathbf{z}) = c_A^S(\mathbf{z}) = \frac{\Psi_S(\log(\mathbf{z})_A + (\gamma(t_i)/2)_{i \in A})}{\prod_{i \in A} z_i}$$

in case of Brown-Resnick processes.

Proof. By Equations (6.6) and (6.7) we get

$$\Lambda(I_A^{(m)}) = \int_{\chi_{i \in A} 2^{-m \cdot (j_m(z_i), j_m(z_i)+1)}} \Psi_S(\mathbf{z}) \, dy_A$$

and

$$\Lambda(I_A^{(m)}) = \int_{\chi_{i \in A} 2^{-m \cdot (\log(2^{-m} j_m(z_i)), \log(2^{-m} (j_m(z_i)+1)))] + \gamma(t_i)/2} \Psi_{BR}(\mathbf{z}) \, dy_A,$$

respectively. As Ψ_S and Ψ_{BR} are continuous on any compact set by Lemma 6.3, the integrals equal $2^{-m|A|} c_A(\mathbf{z}) + o(2^{-m|A|})$. The assertion follows from the facts that

$$\begin{aligned} & \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \exp(-\Lambda(\mathcal{X}_m(\mathbf{z}))) \cdot \prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{\Lambda(I_A^{(m)}(\mathbf{z}))^{k_A}}{k_A!} \exp(-\Lambda(I_A^{(m)}(\mathbf{z}))), \end{aligned}$$

that $\Lambda(\mathcal{X}_m(\mathbf{z})) \rightarrow \Lambda(\mathcal{X}(\mathbf{z}))$ and that $\Lambda(I_A^{(m)}(\mathbf{z})) \rightarrow 0$ as $m \rightarrow \infty$. \square

Now, we consider the distribution of each Π_A . As Lemma 6.4 already suggests, Π_* does not contain any points which can be removed without affecting $Z(\mathbf{t})$, i.e. for each $i \in \{1, \dots, n\}$, there is exactly one $j \in \mathbb{N}$ such that (U_j, W_j) satisfies

$$U_j \cdot f(t_i, W_j(t_i)) = Z(t_i).$$

The following result is also stated in [Dombry and Eyi-Minko \(2011\)](#), Prop. 2.2, in a slightly modified form and more general framework.

Proposition 6.5. *Let $\mathbf{t} \in \mathbb{R}^{d \cdot n}$ and*

$$N = \left\{ \{k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\} : \sum_{A: i \in A} k_A \geq 1 \text{ for all } i \in \{1, \dots, n\}, \right. \\ \left. \sum_{A: i \in A} k_A > 1 \text{ for some } i \in \{1, \dots, n\} \right\}.$$

Then, with probability one,

$$\sum_{\{k_A\} \in N} \mathbb{P} \left(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) = \mathbf{z} \right) \\ = \lim_{m \rightarrow \infty} \sum_{\{k_A\} \in N} \mathbb{P} \left(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) = 0.$$

Proof. By Equation (6.5), we have

$$\sum_{\{k_A\} \in N} \mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset \mid Z(\mathbf{t}) = \mathbf{z}) \\ = \lim_{m \rightarrow \infty} \sum_{\{k_A\} \in N} \mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset \mid Z(\mathbf{t}) \in A_m(\mathbf{z})).$$

Furthermore, as $\sum_{A: i \in A} \Pi_A(\cdot) \leq \sum_{A: i \in A} \Pi_A^{(m)}(\cdot)$, we get that

$$\bigcup_{\{k_A\} \in N} \{\omega \in \Omega : \Pi_{A, \mathbf{z}}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\} \\ \subset \bigcup_{\{k_A\} \in N} \{\omega \in \Omega : \Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\}$$

and both sets are subsets of $E^{(m)}(\mathbf{z}) = \bigcup_{\{k_A\} \in N \cap \{0, 1, 2\}^{2^n - 1}} E_{\{k_A\}}^{(m)}(\mathbf{z})$, where

$$E_{\{k_A\}}^{(m)}(\mathbf{z}) = \{\omega \in \Omega : \Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) \geq k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\}.$$

Therefore, it suffices to show

$$\lim_{m \rightarrow \infty} \mathbb{P}(E^{(m)}(\mathbf{z})) = 0. \quad (6.12)$$

By Lemma 6.4, we can assess

$$\mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) \geq k_A) \\ \leq \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A) + \frac{\Lambda(I_A^{(m)})^{k_A}}{k_A!} \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) > 0) \\ = 2^{-|A|m} \left(\frac{C_A^{k_A}}{k_A!} + o(1) \right)$$

and therefore, for any $\{k_A\} \in N \cap \{0, 1, 2\}^{2^n-1}$,

$$\begin{aligned} \mathbb{P}(E_{\{k_A\}}^{(m)}(\mathbf{z})) &= 2^{-m \sum_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} (|A| \cdot k_A)} \cdot \left(\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \frac{c_A^{k_A}}{k_A!} + o(1) \right) \cdot \exp(-\Lambda(\mathcal{X}(\mathbf{z}))) \\ &\leq 2^{-m(n+1)} \cdot \left(\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \max(c_A^2, 1) + o(1) \right) \cdot \exp(-\Lambda(\mathcal{X}(\mathbf{z}))). \end{aligned}$$

Hence, we get

$$\mathbb{P}(E^{(m)}(\mathbf{z})) \leq 3^{2^n-1} \cdot 2^{-m(n+1)} \cdot \left(\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \max(c_A^2, 1) + o(1) \right) \cdot \exp(-\Lambda(\mathcal{X}(\mathbf{z}))).$$

Furthermore,

$$\begin{aligned} &\mathbb{P}(Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &\geq \mathbb{P}(\Pi_{\{i\}}((0, \infty) \times C(\mathbb{R}^d)) = 1, i = 1, \dots, n, |\Pi_A| = 0, |A| > 1, Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= 2^{-mn} \left(\prod_{i=1}^n c_{\{i\}} + o(1) \right) \cdot \exp(-\Lambda(\mathcal{X}(\mathbf{z}))). \end{aligned}$$

Therefore, we get

$$\mathbb{P}(E^{(m)}(\mathbf{z}) \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \leq 2^{-m} \frac{\prod_{A \in 2^{\{1, \dots, n\}} \setminus \emptyset} \max(c_A^2, 1)}{3 \cdot 3^{-2^n} \cdot \prod_{i=1}^n c_{\{i\}}} + o(2^{-m}),$$

which implies Equation (6.12) and hence closes the proof. \square

Thus, we may restrict our attention to events $\{\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}$ with $\sum_{A: i \in A} k_A = 1$ for $i = 1, \dots, n$. Replacing Π_A by $\Pi_A^{(m)}(\mathbf{z})$, Lemma 6.4 yields that events of the type $\{\Pi_{A, \mathbf{z}}^{(m)}((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\}$ with k_A satisfying these constraints, all have the same rate of convergence 2^{-nm} . We aim to describe events involving Π_A in terms of $\Pi_A^{(m)}(Z(\mathbf{t}))$. This will be done in Theorem 6.7. For the proof of this theorem we will need the following lemma.

Lemma 6.6. *For almost every $\mathbf{z} \in \mathbb{R}^n$, the term*

$$\mathbb{P}\left(\Pi_{A, \mathbf{z}}^{(m)} \neq \Pi_A \text{ for some } A \in 2^{\{1, \dots, n\}} \setminus \emptyset \mid Z(\mathbf{t}) \in A_m(\mathbf{z})\right)$$

vanishes as $m \rightarrow \infty$.

Proof. We consider the set

$$\Delta = \bigcup_{A: |A| \geq 1} \{\omega \in \Omega : \Pi_{A, \mathbf{z}}^{(m)} \neq \Pi_A, Z(\mathbf{t}) \in A_m(\mathbf{z})\}.$$

Furthermore, we define $S = (0, \infty) \times C(\mathbb{R}^d)$ and

$$\Delta_1 = \left\{ \omega \in \Omega : \sum_{A: |A| \geq 1} \Pi_{A, \mathbf{z}}^{(m)}(S) > \sum_{A: |A| > 1} \Pi_A(S), Z(\mathbf{t}) \in A_m(\mathbf{z}) \right\}.$$

First, let $\omega \in \Delta_1$. Then, as $\sum_{A: i \in A} \Pi_{A, \mathbf{z}}^{(m)}(S) \geq \sum_{A: i \in A} \Pi_A(S) \geq 1$ for all $i \in \{1, \dots, n\}$ by construction, there exists an index $i_0 \in \{1, \dots, n\}$ with $\sum_{A: i_0 \in A} \Pi_{A, \mathbf{z}}^{(m)}(S) > 1$. Therefore,

$$\omega \in \bigcup_{\{k_A\} \in N} \{\omega \in \Omega : \Pi_{A, \mathbf{z}}^{(m)}(S) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\}.$$

For any $\omega \in \Delta \setminus \Delta_1$, we have $\sum_{A: |A| \geq 1} \Pi_A(S) = \sum_{A: |A| \geq 1} \Pi_A^{(m)}(S)$ (and this means $\sum_{A: |A| \geq 1} \Pi_A(\cdot) \equiv \sum_{A: |A| \geq 1} \Pi_A^{(m)}(\cdot)$ since we have $\sum_{A: |A| \geq 1} \Pi_A(\cdot) < \sum_{A: |A| \geq 1} \Pi_A^{(m)}(\cdot)$ by definition), but that there are some $(u, w) \in S$ and some A such that $\Pi_A(\{(u, w)\}) = 1$, but $\Pi_{A, \mathbf{z}}^{(m)}(\{(u, w)\}) = 0$. Then, $\Pi_A(\cdot) \leq \sum_{A' \supseteq A} \Pi_{A', \mathbf{z}}^{(m)}(\cdot)$ implies that there is a set $A' \supseteq A$ such that $\Pi_{A', \mathbf{z}}^{(m)}(\{(u, w)\}) = 1$. Let $j \in A' \setminus A$. Then, as $\Pi_A(\{(u, w)\}) = 1$ and $j \notin A$,

$$\sum_{\tilde{A}: j \in \tilde{A}} \Pi_{\tilde{A}}^{(m)}(S) \geq \sum_{\tilde{A}: j \in \tilde{A}} \Pi_{\tilde{A}}(S) + |\{(u, w)\}| > \sum_{\tilde{A}: j \in \tilde{A}} \Pi_{\tilde{A}}(S) \geq 1.$$

All in all, we end up with

$$\Delta \subset \bigcup_{\{k_A\} \in N} \{\omega \in \Omega : \Pi_A^{(m)}(S) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})\}$$

and the assertion follows by Proposition 6.5. \square

Theorem 6.7. For $A \in 2^{\{1, \dots, n\}}$, let $k_A \in \mathbb{N}$ and $B_A \in \mathcal{B} \times \mathcal{C}(\mathbb{R}^d)$. Then, with probability one,

$$\begin{aligned} & \mathbb{P}(\Pi_A(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) \in A_m(\mathbf{z})). \end{aligned}$$

Proof. By Equation (6.5), it suffices to show

$$\begin{aligned} & \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_A(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) \in A_m(\mathbf{z})) \\ &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}(B_A) = k_A, A \in 2^{\{1, \dots, n\}} \mid Z(\mathbf{t}) \in A_m(\mathbf{z})). \end{aligned} \quad (6.13)$$

Noting that $\Pi_{\emptyset, \mathbf{z}}^{(m)}(\cdot) \leq \Pi_{\emptyset}(\cdot) \leq \Pi_{\emptyset, \mathbf{z}}^{(m)}(\cdot) + \sum_{A: |A| \geq 1} \Pi_{A, \mathbf{z}}^{(m)}(\cdot)$ for $Z(\mathbf{t}) = \mathbf{z}$, we have

$$\begin{aligned} & \left| \mathbb{P}(\Pi_A(B_A) = k_A, A \in 2^{\{1, \dots, n\}}, Z(\mathbf{t}) \in A_m(\mathbf{z})) \right. \\ & \quad \left. - \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)}(B_A) = k_A, A \in 2^{\{1, \dots, n\}}, Z(\mathbf{t}) \in A_m(\mathbf{z})) \right| \\ & \leq \mathbb{P}(\Pi_{A, \mathbf{z}}^{(m)} \neq \Pi_A \text{ for some } A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) \in A_m(\mathbf{z})). \end{aligned}$$

Thus, Lemma 6.6 yields (6.13), i.e. the assertion. \square

Plugging in the results of Lemma 6.4 into Theorem 6.7 for all $\{k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}$ with $\sum_{A: i \in A} k_A = 1$ for all $i \in \{1, \dots, n\}$, we get that

$$\begin{aligned} & \mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d))) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \frac{\prod_{A: k_A=1} c_A}{\sum \left\{ \begin{array}{l} \tilde{k}_A: \sum_{A: i \in A} \tilde{k}_A=1 \\ \text{for all } i \in \{1, \dots, n\} \end{array} \right\} \prod_{A: \tilde{k}_A=1} c_A}. \end{aligned} \quad (6.14)$$

Equation (6.14) enables us to sample from the distribution of the (finitely many) events $\{\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset, Z(\mathbf{t}) = \mathbf{z}\}$ conditional on $Z(\mathbf{t}) = \mathbf{z}$. Given such an event, the formulae given in Lemma 6.4 and Theorem 6.7 allow for sampling the points of $\{\Pi_A, k_A = 1\}$ independently. To achieve conditional sampling from $\Pi(\cdot) = \Pi_*(\cdot) + \Pi_\emptyset(\cdot)$, we have to analyse the distribution of Π_\emptyset given $Z(\mathbf{t}) = \mathbf{z}$. The following result is also stated in Dombry and Eyi-Minko (2011), Thm. 3.1.

Theorem 6.8. *Let $\mathbf{t} \in \mathbb{R}^{d \cdot n}$ be fixed.*

1. *Conditional on $Z(\mathbf{t}) = \mathbf{z}$, the point processes Π_* and Π_\emptyset are stochastically independent.*
2. *With probability one, $\Pi_\emptyset(\cdot) \mid Z(\mathbf{t}) = \mathbf{z}$ has the same distribution as $\Pi(\cdot \setminus \mathcal{X}(\mathbf{z}))$.*

Proof. 1. Let $B_1, \dots, B_k, B_{k+1}, \dots, B_l \in (\mathcal{B} \cap (0, \infty)) \times \mathcal{C}(\mathbb{R}^d)$ and $r_1, \dots, r_l \in \mathbb{N}$. Then, by Theorem 6.7, we get

$$\begin{aligned} & \mathbb{P}(\Pi_*(B_j) = r_j, j = 1, \dots, k, \Pi_\emptyset(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P} \left(\begin{array}{l} \sum_{A: |A| \geq 1} \Pi_{A, \mathbf{z}}^{(m)}(B_j) = r_j, j = 1, \dots, k, \\ \Pi_{\emptyset, \mathbf{z}}^{(m)}(B_j) = r_j, j = k+1, \dots, l \end{array} \middle| Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) \end{aligned}$$

By Remark 6.2, $\Pi_{\emptyset, \mathbf{z}}^{(m)}$ is independent of $\sum_{A: |A| \geq 1} \Pi_{A, \mathbf{z}}^{(m)}$ and of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$. Hence,

$$\begin{aligned} & \mathbb{P}(\Pi_*(B_j) = r_j, j = 1, \dots, k, \Pi_\emptyset(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \\ &= \lim_{m \rightarrow \infty} \mathbb{P} \left(\begin{array}{l} \sum_{A: |A| \geq 1} \Pi_{A, \mathbf{z}}^{(m)}(B_j) = r_j, j = 1, \dots, k \\ \cdot \mathbb{P} \left(\Pi_{\emptyset, \mathbf{z}}^{(m)}(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z}) \right) \end{array} \right) \\ &= \mathbb{P}(\Pi_*(B_j) = r_j, j = 1, \dots, k \mid Z(\mathbf{t}) = \mathbf{z}) \\ & \quad \cdot \mathbb{P}(\Pi_\emptyset(B_j) = r_j, j = k+1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) \end{aligned}$$

which shows the independence of $\Pi_* \mid Z(\mathbf{t}) = \mathbf{z}$ and $\Pi_\emptyset \mid Z(\mathbf{t}) = \mathbf{z}$.

2. Let $B_1, \dots, B_k \in \mathcal{B} \times \mathcal{C}(\mathbb{R}^d)$ and $r_1, \dots, r_l \in \mathbb{N}$. Then, Theorem 6.7 yields

$$\mathbb{P}(\Pi_\emptyset(B_j) = r_j, j = 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z})$$

$$= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_{\emptyset, \mathbf{z}}^{(m)}(B_j) = r_j, j = 1, \dots, l \mid Z(\mathbf{t}) \in A_m(\mathbf{z})).$$

From Remark 6.2 we get that $\Pi_{\emptyset, \mathbf{z}}^{(m)}$ is independent of the event $Z(\mathbf{t}) \in A_m(\mathbf{z})$, which implies

$$\begin{aligned} \mathbb{P}(\Pi_{\emptyset}(B_j) = r_j, j = 1, \dots, l \mid Z(\mathbf{t}) = \mathbf{z}) &= \lim_{m \rightarrow \infty} \mathbb{P}(\Pi_{\emptyset, \mathbf{z}}^{(m)}(B_j) = r_j, j = 1, \dots, l) \\ &= \mathbb{P}(\Pi(B_j \setminus \mathcal{X}(\mathbf{z})) = r_j, j = 1, \dots, l). \end{aligned}$$

Here, we use that $\Pi_{\emptyset, \mathbf{z}}^{(m)}(\cdot) = \Pi(\cdot \setminus \mathcal{X}^{(m)}(\mathbf{z}))$ and that $\mathcal{X}^{(m)}(\mathbf{z}) \xrightarrow{m \rightarrow \infty} \mathcal{X}(\mathbf{z})$. \square

By the second part of Theorem 6.8 we have

$$\max_{i \in \mathbb{N}} \left(\tilde{U}_i \cdot f(\cdot, \tilde{W}_i(\cdot)) \right) \mid (Z(\mathbf{t}) = \mathbf{z}) \stackrel{d}{=} Z(\cdot) \mid Z(\mathbf{t}) < \mathbf{z},$$

where $(\tilde{U}_i, \tilde{W}_i)_{i \in \mathbb{N}}$ are the points of Π_{\emptyset} . The process on the rhs can be easily simulated using any Poisson point process representation of the extremal Gaussian process or the Brown-Resnick process as described in Chapter 4, respectively, and rejecting all the points not satisfying $Z(\mathbf{t}) < \mathbf{z}$.

Composing all these considerations, we end up with the following procedure to sample from $Z(\cdot) \mid Z(\mathbf{t}) = \mathbf{z}$ (Dombry et al., 2011; Dombry and Ribatet, 2012):

1. Calculate the probabilities of all the scenarios $\{\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = k_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset\}$ conditional on $Z(\mathbf{t}) = \mathbf{z}$ with $\sum_{A: i \in A} k_A = 1$ for $i \in \{1, \dots, n\}$ (cf. Equation (6.14)). Draw a scenario from this distribution.
2. Independently, sample points (U_A, W_A) from each $\Pi_A, A \in 2^{\{1, \dots, n\}} \setminus \emptyset$ with $k_A = 1$.
3. Generate an independent simulation of $Z(\cdot) \mid Z(\mathbf{t}) < \mathbf{z}$. Take the maximum of all the samples.

An efficient implementation of the first step is not straightforward as one has to deal with an exploding number of scenarios with a positive probability as n increases. Dombry et al. (2011) propose an MCMC algorithm to cope with this problem. While the third step can be done by the considerations above, the implementation of the second step needs some further work. This will be the issue of the next section.

6.3 On the Distribution of Π_* within a Scenario

In this section, we will deal with the distribution of $\Pi_A \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1$ for $A \in 2^{\{1, \dots, n\}} \setminus \emptyset$. More precisely, we consider the distribution of the random vector $(U_A, W_A) \in (0, \infty) \times C(\mathbb{R}^d)$, which is well-defined by $\Pi_A(\{(U_A, W_A)\}) = 1$ conditional on $|\Pi_A| = 1$. In Dombry et al. (2011) and Dombry and Ribatet (2012), all finite-dimensional marginal distributions of the random field $\{U_A \cdot f(t, W_A(t)), t \in \mathbb{R}^d\}$ are given. These are multivariate Student distributions in the case of extremal Gaussian processes and multivariate log-normal distributions in the case of Brown-Resnick processes. However,

there are only few details how to sample from these distributions. Therefore, we have another look at this problem reducing it to the well-known problem of conditional sampling for Gaussian processes. We distinguish between extremal Gaussian processes and Brown-Resnick processes.

6.3.1 Extremal Gaussian Processes

We simulate from (U_A, W_A) in a two-step procedure, sampling from both components separately. Considering U_A first, for an arbitrary set $B \in \mathcal{B} \cap (0, \infty)$ we get

$$\begin{aligned} & \mathbb{P}(U_A \in B \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \mathbb{P}(\Pi_A(B \times C(\mathbb{R}^d)) = 1 \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \frac{\mathbb{P}(\Pi_A(B \times C(\mathbb{R}^d)) = 1 \mid Z(\mathbf{t}) = \mathbf{z})}{\mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1 \mid Z(\mathbf{t}) = \mathbf{z})} \\ &= \frac{\int_B u^{-2} \int_{\times_{i \notin A}(-\infty, z_i]} \frac{1}{(2\pi u)^n |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{\mathbf{y}^T \Sigma^{-1} \mathbf{y}}{2\pi u^2}\right) dy_{A^c} du}{\int_0^\infty u^{-2} \int_{\times_{i \notin A}(-\infty, z_i]} \frac{1}{(2\pi u)^n |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \frac{\mathbf{y}^T \Sigma^{-1} \mathbf{y}}{2\pi u^2}\right) dy_{A^c} du} \end{aligned}$$

where we used Equation (6.8), Lemma 6.4 and Theorem 6.7. Thus, the density of U_A has the form

$$\begin{aligned} & \mathbb{P}(U_A \in du \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \kappa(A, \mathbf{z}) \cdot u^{-2} \cdot \mathbb{P}(\sqrt{2\pi}uW(t_i) \in dz_i, i \in A, \sqrt{2\pi}uW(t_j) < z_j, j \notin A) du, \end{aligned} \tag{6.15}$$

for some $\kappa(A, \mathbf{z}) > 0$ and any $u > 0$. Sampling from this density is quite involved, but can be done by MCMC methods, for example.

In a second step, we consider the distribution of W_A conditional on U_A . For a set $C \in \mathcal{C}(\mathbb{R}^d)$, we have

$$\begin{aligned} & \mathbb{P}(U_A \in B, W_A \in C \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \mathbb{P}(\Pi_A(B \times C) = 1 \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z}) \\ &= \frac{\mathbb{P}(\Pi_A(B \times C) = 1 \mid Z(\mathbf{t}) = \mathbf{z})}{\mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1 \mid Z(\mathbf{t}) = \mathbf{z})} \\ &= \kappa(A, \mathbf{z}) \int_B u^{-2} \cdot \mathbb{P}(W(\cdot) \in C, \sqrt{2\pi}uW(t_i) \in dz_i, i \in A, \sqrt{2\pi}uW(t_j) < z_j, j \notin A) du. \end{aligned}$$

By Equation (6.15), we get the conditional distribution

$$\begin{aligned} & \mathbb{P}(W_A \in C \mid U_A = u, Z(\mathbf{t}) = \mathbf{z}) \\ &= \mathbb{P}\left(W \in C \mid W(t_i) = \frac{z_i}{\sqrt{2\pi}u}, i \in A, W(t_j) < \frac{z_j}{\sqrt{2\pi}u}, j \notin A\right). \end{aligned}$$

Thus, conditional on U_A , the distribution of W_A equals a conditional distribution of the stationary Gaussian random field $W(\cdot)$. One way to sample from this distribution is conditioning on $(W(t_i))_{i \in A}$ by kriging (cf. Section 2.1) and rejecting all the realizations with $W(t_j) \geq \frac{z_j}{\sqrt{2\pi}u}$ for some $j \in A^c$.

6.3.2 Brown-Resnick Processes

One way to tackle the problem in the case of Brown-Resnick processes is to use an analogous procedure to the case of extremal Gaussian processes. Then, the first step of this procedure could be skipped if we had $t_i = 0$ for some $i \in A$ as $W(0) = 0$ a.s., which means that $U_A = Z(0)$.

In general, we have $0 \notin \{t_1, \dots, t_n\}$. Nevertheless, we can make use of the degenerated distribution at the origin if we consider the Poisson point process

$$\widehat{\Pi} = \sum_{i \in \mathbb{N}} \delta_{U_i \exp(W_i(\cdot) - \gamma(\cdot)/2)}$$

instead of $\Pi = \sum_{i \in \mathbb{N}} \delta_{(U_i, W_i)}$ and use the fact that this process is stationary by Proposition 5 in [Kabluchko et al. \(2009\)](#). Let $i^* \in A$ and $C \in \mathcal{C}(\mathbb{R}^d)$. Then, we have

$$\begin{aligned} & \mathbb{P} \left(U_A \cdot \exp \left(W_A(\cdot) - \frac{\gamma(\cdot)}{2} \right) \in C \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t}) = \mathbf{z} \right) \\ &= \mathbb{P} \left(\widehat{\Pi} \left(\{g \in C(\mathbb{R}^d) : g \in C, g(t_i) = z_i, i \in A, g(t_j) < z_j, j \notin A\} \mid 1 \right) \right. \\ & \quad \widehat{\Pi} \left(\{g \in C(\mathbb{R}^d) : g(t_i) = z_i, i \in A, g(t_j) < z_j, j \notin A\} \right) = 1, \\ & \quad \widehat{\Pi}(\{g \in C(\mathbb{R}^d) : g(t_i) = z_i\}) = 1 \text{ for all } i \in \{1, \dots, n\}, \\ & \quad \widehat{\Pi}(\{g \in C(\mathbb{R}^d) : g(t_j) > z_j \text{ for some } j \in \{1, \dots, n\}\}) = 0 \left. \right) \\ &= \mathbb{P} \left(\widehat{\Pi} \left(\{g \in C(\mathbb{R}^d) : g(\cdot - t_{i^*}) \in C, g(t_i - t_{i^*}) = z_i, i \in A, \right. \right. \\ & \quad \left. \left. g(t_j - t_{i^*}) < z_j, j \notin A\} \mid 1 \right) \right. \\ & \quad \widehat{\Pi}(\{g \in C(\mathbb{R}^d) : g(t_i - t_{i^*}) = z_i, i \in A, g(t_j - t_{i^*}) < z_j, j \notin A\}) = 1, \\ & \quad \widehat{\Pi}(\{g \in C(\mathbb{R}^d) : g(t_i - t_{i^*}) = z_i\}) = 1 \text{ for all } i \in \{1, \dots, n\}, \\ & \quad \widehat{\Pi}(\{g \in C(\mathbb{R}^d) : g(t_j - t_{i^*}) > z_j \text{ for some } j \in \{1, \dots, n\}\}) = 0 \left. \right) \\ &= \mathbb{P} \left(U_A \cdot e^{W_A(\cdot - t_{i^*}) - \gamma(\cdot - t_{i^*})/2} \in C \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t} - t_{i^*}) = \mathbf{z} \right) \end{aligned}$$

Thus, we have to consider the distribution of Π conditional on $Z(0)$. Up to now, we excluded this case. However, we can redo all the computations above, reducing the dimension of the domain of integration w.r.t. y by one and restricting the domain of integration w.r.t. u . We end up with

$$\begin{aligned} & \mathbb{P} \left(U_A \cdot e^{W_A(\cdot - t_{i^*}) - \gamma(\cdot - t_{i^*})/2} \in C \mid \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t} - t_{i^*}) = \mathbf{z} \right) \\ &= \mathbb{P} \left(\Pi_A(\{(u, w) \in (0, \infty) \times C(\mathbb{R}^d) : u \cdot \exp(w(\cdot - t_{i^*}) - \gamma(\cdot - t_{i^*})/2) \in C\}) = 1 \mid \right. \\ & \quad \left. \Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1, Z(\mathbf{t} - t_{i^*}) = \mathbf{z} \right) \\ &= \lim_{m \rightarrow \infty} \frac{\mathbb{P}(\Pi_A(\{(u, w) : u e^{w(\cdot - t_{i^*}) - \frac{\gamma(\cdot - t_{i^*})}{2}} \in C\}) = 1 \mid Z(\mathbf{t} - t_{i^*}) \in A_m(\mathbf{z}))}{\mathbb{P}(\Pi_A((0, \infty) \times C(\mathbb{R}^d)) = 1 \mid Z(\mathbf{t} - t_{i^*}) \in A_m(\mathbf{z}))} \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{P} \left(e^{W(\cdot - t_{i^*}) - \frac{\gamma(\cdot - t_{i^*})}{2}} \in \frac{C}{z_{i^*}}, W(t_i - t_{i^*}) - \frac{\gamma(t_i - t_{i^*})}{2} \in d \log \left(\frac{z_i}{z_{i^*}} \right), i \in A, \right. \\
 &\quad \left. W(t_j - t_{i^*}) - \frac{\gamma(t_j - t_{i^*})}{2} < \log \left(\frac{z_j}{z_{i^*}} \right), j \notin A \right) \\
 &\quad \cdot \mathbb{P} \left(W(t_i - t_{i^*}) - \frac{\gamma(t_i - t_{i^*})}{2} \in d \log \left(\frac{z_i}{z_{i^*}} \right), i \in A, \right. \\
 &\quad \left. W(t_j - t_{i^*}) - \frac{\gamma(t_j - t_{i^*})}{2} < \log \left(\frac{z_j}{z_{i^*}} \right), j \notin A \right)^{-1} \\
 &= \mathbb{P} \left(e^{W(\cdot - t_{i^*}) - \frac{\gamma(\cdot - t_{i^*})}{2}} \in \frac{C}{z_{i^*}} \mid W(t_i - t_{i^*}) - \frac{\gamma(t_i - t_{i^*})}{2} = \log \left(\frac{z_i}{z_{i^*}} \right), i \in A, \right. \\
 &\quad \left. W(t_j - t_{i^*}) - \frac{\gamma(t_j - t_{i^*})}{2} < \log \left(\frac{z_j}{z_{i^*}} \right), j \notin A \right)
 \end{aligned}$$

which is the conditional distribution of a transformed Gaussian random field. We can sample from this distribution using standard techniques for conditional sampling of Gaussian intrinsic random functions (e.g. conditioning by kriging, cf. [Delfiner, 1976](#); [Chilès and Delfiner, 1999](#)). Recall that this is not the distribution of Π_A itself, but all we need to sample from $Z(\cdot) \mid Z(\mathbf{t})$.

Thus, we are able to simulate from extremal Gaussian processes and Brown-Resnick processes conditional on data $(t_i, Z(t_i))_{i \in \mathbb{N}}$. For results on simulation studies and applications to real data see [Dombry et al. \(2011\)](#) and [Dombry and Ribatet \(2012\)](#).

Bibliography

- Abramowitz, M. and Stegun, I. A., editors (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. New York: Dover Publications.
- Adams, R. A. (1975). *Sobolev Spaces*. Academic Press, New York-London. Pure and Applied Mathematics, Vol. 65.
- Amann, H. and Escher, J. (2008). *Analysis. III. Grundstudium Mathematik*. [Basic Study of Mathematics]. Basel: Birkhäuser Verlag, second edition.
- Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *J. Multivariate Anal.*, 88(1), 190–206.
- Blanchet, J. and Davison, A. (2011). Spatial Modelling of Extreme Snow Depth. *Ann. Appl. Stat.*, 5(3), 1699–1725.
- Brown, B. M. and Resnick, S. I. (1977). Extreme Values of Independent Stochastic Processes. *J. Appl. Probab.*, 14(4), 732–739.
- Buishand, T., de Haan, L., and Zhou, C. (2008). On Spatial Extremes: with Applications to a Rainfall Problem. *Ann. Appl. Statist.*, 2(2), 624–642.
- Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. New York: John Wiley & Sons Inc. Modeling spatial uncertainty, A Wiley-Interscience Publication.
- Cooley, D., Davis, R. A., and Naveau, P. (2007). Prediction for Max-Stable Processes via an Approximated Conditional Density. Technical Report 2007-3, Colorado State University.
- Daley, D. J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. Springer Series in Statistics. New York: Springer-Verlag.
- Davis, R. A. and Resnick, S. I. (1989). Basic properties and prediction of max-ARMA processes. *Adv. in Appl. Probab.*, 21(4), 781–803.
- (1993). Prediction of stationary max-stable processes. *Ann. Appl. Probab.*, 3(2), 497–525.
- Davison, A. C. and Gholamrezaee, M. M. (2011). Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 468(2138), 581–608.
- de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.*, 12(4), 1194–1204.

- de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. New York: Springer.
- de Haan, L. and Pickands, J., III (1986). Stationary min-stable stochastic processes. *Probab. Theory Relat. Fields*, 72(4), 477–492.
- Delfiner, P. (1976). Linear estimation of nonstationary spatial phenomena. In Guarascio, M., David, M., and Huijbregts, C., editors, *Advanced Geostatistics in the Mining Industry*, pp. 49–68. Dordrecht: Reidel.
- Dombry, C. and Eyi-Minko, F. (2011). Regular conditional distributions of max infinitely divisible processes. *ArXiv e-prints*, [1109.6492](#).
- Dombry, C., Eyi-Minko, F., and Ribatet, M. (2011). Conditional simulations of Brown-Resnick processes. *ArXiv e-prints*, [1112.3891](#).
- Dombry, C. and Ribatet, M. (2012). Conditional simulation of extremal Gaussian processes. *ArXiv e-prints*, [1202.4737](#).
- Driscoll, T. A. and Fornberg, B. (2002). Interpolation in the limit of increasingly flat radial basis functions. *Comput. Math. Appl.*, 43(3-5), 413–422. Radial basis functions and partial differential equations.
- Eckert, N., Gaume, J., and Castebrunet, H. (2011). Using spatial and spatial-extreme statistics to characterize snow avalanche cycles. *Procedia Environmental Sciences*, 7(0), 224 – 229.
- Engelke, S., Kabluchko, Z., and Schlather, M. (2011). An equivalent representation of the Brown-Resnick process. *Statist. Probab. Lett.*, 81(8), 1150–1154.
- Gaetan, C. and Guyon, X. (2010). *Spatial Statistics and Modeling*. Springer Series in Statistics. New York: Springer.
- Gelfand, A. E., Diggle, P. J., Fuentes, M., and Guttorp, P., editors (2010). *Handbook of spatial statistics*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton, FL: CRC Press.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.*, 102(477), 359–378.
- Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, seventh edition.
- Grisvard, P. (1985). *Elliptic Problems in Nonsmooth Domains*, volume 24 of *Monographs and Studies in Mathematics*. Boston, MA: Pitman (Advanced Publishing Program).
- Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika*, 93(4), 989–995.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.

-
- Horn, R. A. and Johnson, C. R. (1996). *Matrix Analysis*. Cambridge: Cambridge University Press.
- Ihaka, R. and Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *J. Comput. Graph. Statist.*, 5(3), 299–314.
- Kabluchko, Z. (2009a). Extremes of Space-Time Gaussian Processes. *Stoch. Proc. Appl.*, 119(11), 3962–3980.
- (2009b). Spectral Representations of Sum- and Max-Stable Processes. *Extremes*, 12(4), 401–424.
- Kabluchko, Z. and Schlather, M. (2010). Ergodic properties of max-infinitely divisible processes. *Stoch. Proc. Appl.*, 120(3), 281–295.
- Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.*, 37(5), 2042–2065.
- Kingman, J. F. C. (1993). *Poisson Processes*, volume 3 of *Oxford Studies in Probability*. New York: The Clarendon Press Oxford University Press. Oxford Science Publications.
- Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwaterstrand. *J. of the Chem., Metal. and Mining Soc. of South Africa*, 52(6), 119–139.
- Kufner, A., John, O., and Fučík, S. (1977). *Function spaces*. Leyden: Noordhoff International Publishing. Monographs and Textbooks on Mechanics of Solids and Fluids; Mechanics: Analysis.
- Landau, H. J. and Shepp, L. A. (1970). On the Supremum of a Gaussian Process. *Sankhyā Ser. A*, 32, 369–378.
- Lange, K. (1999). *Numerical Analysis for Statisticians*. Statistics and Computing. New York: Springer-Verlag.
- Lantuéjoul, C. (2002). *Geostatistical Simulation: Models and Algorithms*. Statistics and Computing. New York: Springer.
- Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58(8), 1246–1266.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equations of State Calculations by fast Computing Machines. *Journal Of Chemical Physics*, 21(6), 1087–1092.
- Narcowich, F. J., Ward, J. D., and Wendland, H. (2006). Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions. *Constr. Approx.*, 24(2), 175–186.
- Ober, U., Erbe, M., Long, N., Porcu, E., Schlather, M., and Simianer, H. (2011). Predicting Genetic Values: a Kernel-Based Best Linear Unbiased Prediction with Genomic Data. *Genetics*, 188(3), 695–708.

- Oesting, M. (2009). *Simulationsverfahren für Brown-Resnick-Prozesse*. Master's thesis, Universität Göttingen. *ArXiv e-prints*, [0911.4389](https://arxiv.org/abs/0911.4389). Version 1 (2009): diploma thesis; Version 2 (2010): revised technical report.
- Oesting, M., Kabluchko, Z., and Schlather, M. (2012). Simulation of Brown–Resnick processes. *Extremes*, 15(1), 89–107.
- Oesting, M. and Schlather, M. (2012). Conditional Sampling for Max-Stable Processes with a Mixed Moving Maxima Representation. *ArXiv e-prints*, [1202.5023](https://arxiv.org/abs/1202.5023). URL <http://arxiv.org/abs/1202.5023>.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.r-project.org>.
- Ribatet, M. (2011). *SpatialExtremes: Modelling Spatial Extremes*. R Package Version 1.8-1, URL <http://cran.r-project.org/package=SpatialExtremes>.
- Rippa, S. (1999). An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Adv. Comput. Math.*, 11(2-3), 193–210.
- Rogers, L. C. G. and Pitman, J. W. (1981). Markov functions. *Ann. Probab.*, 9(4), 573–582.
- Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 1*. Cambridge Mathematical Library. Cambridge: Cambridge University Press. Foundations, Reprint of the second (1994) edition.
- Rosiński, J. (1995). On the Structure of Stationary Stable Processes. *Ann. Probab.*, 23(3), 1163–1187.
- Schaback, R. (2005). Multivariate interpolation by polynomials and radial basis functions. *Constr. Approx.*, 21(3), 293–317.
- (2008). Limit problems for interpolation by analytic radial basis functions. *J. Comput. Appl. Math.*, 212(2), 127–149.
- Schaback, R. and Wendland, H. (2002). Inverse and saturation theorems for radial basis function interpolation. *Math. Comp.*, 71(238), 669–681 (electronic).
- (2004). *Numerische Mathematik*. Springer-Lehrbuch. Springer.
- Scheuerer, M. (2009). *A Comparison of Models and Methods for Spatial Interpolation in Statistics and Numerical Analysis*. Ph.D. thesis, Universität Göttingen. URL <http://webdoc.sub.gwdg.de/diss/2010/scheuerer/>.
- (2010). Regularity of the sample paths of a general second order random field. *Stoch. Proc. Appl.*, 120(10), 1879–1897.
- (2011). An alternative procedure for selecting a good value for the parameter c in RBF-interpolation. *Adv. Comput. Math.*, 34(1), 105–126.

-
- Schilling, R. L. (2005). *Measures, Integrals and Martingales*. New York: Cambridge University Press.
- Schlather, M. (2002). Models for Stationary Max-Stable Random Fields. *Extremes*, 5(1), 33–44.
- (2012). *RandomFields: Simulation and Analysis of RandomFields*. R Package Version 2.0.54, URL <http://cran.r-project.org/package=RandomFields>.
- Schlather, M. and Tawn, J. A. (2003). A dependence measure for multivariate and spatial extreme values: properties and inference. *Biometrika*, 90(1), 139–156.
- Smith, R. L. (1990). Max-Stable Processes and Spatial Extremes. Unpublished manuscript.
- Song, G., Riddle, J., Fasshauer, G., and Hickernell, F. (2012). Multivariate interpolation with increasingly flat radial basis functions of finite smoothness. *Adv. Comput. Math.*, 36(3), 485–501.
- Stein, M. L. (1999). *Interpolation of Spatial Data*. Springer Series in Statistics. New York: Springer-Verlag. Some Theory for Kriging.
- Stoev, S. A. and Taquq, M. S. (2005). Extremal stochastic integrals: a parallel between max-stable processes and α -stable processes. *Extremes*, 8(4), 237–266.
- Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.*, 22(4), 1701–1762. With discussion and a rejoinder by the author.
- Wang, Y. (2010). *maxLinear: Conditional Sampling for Max-Linear Models*. R Package Version 1.0, URL <http://cran.r-project.org/package=maxLinear>.
- Wang, Y. and Stoev, S. A. (2010). On the structure and representations of max-stable processes. *Adv. in Appl. Probab.*, 42(3), 855–877.
- (2011). Conditional sampling for spectrally discrete max-stable random fields. *Adv. in Appl. Probab.*, 43(2), 461–483.
- Weintraub, K. S. (1991). Sample and ergodic properties of some min-stable processes. *Ann. Probab.*, 19(2), 706–723.
- Wendland, H. (2005). *Scattered Data Approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge: Cambridge University Press.