Spatial point process models with applications to max-stable random fields

Inauguraldissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften der Universität Mannheim

vorgelegt von

Martin Dirrler
aus Karlsruhe

Mannheim 2017

Dekan: Dr. Bernd Lübcke, Universität Mannheim
Referent: Prof. Dr. Martin Schlather, Universität Mannheim
Korreferentin: Dr. Kirstin Strokorb, Cardiff University

Tag der mündlichen Prüfung: 25. Oktober 2017

Abstract

In Part I of this thesis, we briefly summarize some theory of point processes which is crucial for the subsequent parts.
We introduce a class of spatial stochastic processes in the max-domain of attraction of familiar max-stable processes in Part II. The new class is based on Cox processes instead of Poisson processes. We show that statistical inference is possible within the given framework, at least under some reasonable restrictions.
The Matérn hard-core processes are classical examples for point process models obtained from (marked) Poisson point processes. Points of the original Poisson process are deleted according to a dependent thinning rule, resulting in a process whose points have a prescribed hard-core distance. In Part III, we present a new model which generalizes the underlying point process, the thinning rule and the marks attached to the original process. The new model further reveals several connections to mixed moving maxima processes, e.g. a process of visible storm centres.

Zusammenfassung

Im ersten Teil dieser Dissertation fassen wir einige grundlegende Resultate zu Punktprozessen zusammen, diese sind für alle nachfolgenden Teile essentiell.
In Teil II stellen wir eine Klasse räumlicher stochastischer Prozesse vor, die sich im Max-Anziehungsbereich bekannter max-stabiler Prozesse befindet. Diese neue Klasse basiert auf Cox Prozessen anstatt von Poisson Punktprozessen. Wir zeigen, dass Inferenz zumindest unter einigen sinnvollen Beschränkungen möglich ist.
Die Matérn hard-core Prozesse sind ein klassisches Beispiel für Punktprozesse, die von markierten Poisson Punktprozessen abgeleitet sind. Punkte des ursprünglichen Poisson Prozesses werden gemäß eines Ausdünnungsalgorithmus entfernt, was zur Folge hat, dass die verbliebenen Punkte einen vorgegebenen Mindestabstand haben. Im dritten Teil präsentieren wir ein neues Modell, das den zu Grunde liegenden Punktprozess, den Ausdünnungsalgorithmus und die Marken der Punkte verallgemeinert. Dieses Modell ermöglicht eine Verbindung zu max-stabilen Prozessen.

Acknowledgements

Once I was a little boy in class one of primary school, when asked about my future career aspirations I used to say 'I want to invent math'. My younger self might have been a little too ambitious since I figured out pretty fast that math had already been 'invented'. The years went by and interests changed several times, but finally the boy – not so little anymore – decided to become a mathematician. This endeavour would not have been successfully completed without the support of many people.

First of all, I am indebted to my supervisor Prof. Dr. Martin Schlather for giving me the opportunity to do this PhD. His constant support, patience and confidence entailed an excellent working atmosphere. He was always available to help and advice, despite of being busy almost all the time.
I am much obliged to my second supervisor Dr. Kirstin Strokorb. She made huge efforts and spent much of her rare time to improve my skills and sharpen my mind. Especially in the first year of this project, she encouraged me a lot and helped me attaining a much higher level of mathematical knowledge. Without her support, this PhD project would have failed most likely.
Special thanks also go to Anja Gilliar for taking care of all administrative work and the constantly excellent collaboration concerning the organization of numerous exercise sessions.
I thank all my colleagues for many cheerful talks and discussions (especially Dr. Martin Kroll for excessive conversations about Prussian noblemen and their moustached descendants).
Finally, I am deeply indebted to my parents for their unconditional support throughout my whole life. My wife Vera deserves my greatest gratitude for bearing with me even in these strenuous times.

Financial support by the Deutsche Forschungsgemeinschaft through 'RTG 1953 - Statistical Modeling of Complex Systems and Processes' and by Volkswagen Stiftung within the project 'Mesoscale Weather Extremes - Theory, Spatial Modeling and Prediction' is gratefully acknowledged.

Contents

1. Introduction

Da steh ich nun, ich armer Tor!
Und bin so klug als wie zuvor
Heiße Magister, heiße Doktor gar
Und ziehe schon an die zehen Jahr
Herauf, herab und quer und krumm
Meine Schüler an der Nase herum
Und sehe, dass wir nichts wissen können!

(aus Faust I, Johann Wolfgang von Goethe)

Spatial point patterns occur in many applications from different areas, such as environmental sciences (Stoyan and Penttinen, 2000), finance (Chavez-Demoulin et al., 2005), physics (Babu and Feigelson, 1996; Scargle and Babu, 2003) and information technology (Ibrahim et al., 2013). Such point patterns are commonly modelled by so-called *point processes*. Point processes are the fundamental building blocks of this thesis. In the sequel, we give a short summary of the topics covered by this work. We try to give the summary without introducing too much mathematical theory and postpone the rigorous mathematics to the subsequent parts.

Part I: Point processes

Roughly speaking, a point process is a random variable whose realization is not a real number but a point pattern. A simple but rather artificial example of a point process is to roll a dice and interpret the resulting point pattern as realization – see Figure 1.1.

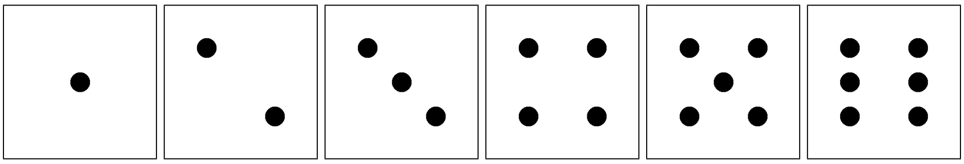

Figure 1.1.: All realizations of the point process 'roll an ordinary dice' that occur with positive probability.

In general, a point process has infinitely many possible outcomes, that is we may think of rolling a dice with infinitely many sides and with different point patterns on each side (Figure 1.2).

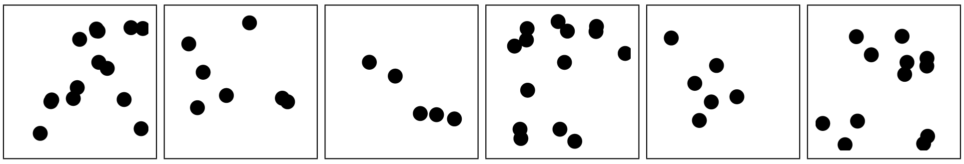

Figure 1.2.: Six arbitrary realizations of a point process on a bounded set.

These point patterns are not limited to be subsets of $\mathbb{R}^d$ as in the toy examples above. Indeed the points might also be functions or even point processes themselves. This flexibility is one reason for point processes being used in many different areas. Though it comes with the price that the theoretical treatment of point processes is a rather difficult task. We therefore give a brief introduction to point processes in Part I of this dissertation and refer to Karr (1986); Stoyan and Stoyan (1992); Daley and Vere-Jones (2003); Møller and Waagepetersen (2004); Daley and Vere-Jones (2008); Chiu et al. (2013) for detailed descriptions of that topic.

Part II: Conditionally Max-stable Random Fields

Point processes are the building blocks for *max-stable* processes. A random field Z is called max-stable if there exist an i.i.d. sequence of random fields $Y_1, Y_2, Y_3, \dots \sim Y$ and sequences of norming functions $a_n(\cdot) > 0$, $b_n(\cdot) \in \mathbb{R}$ such that

$$\left\{ \frac{\bigvee_{i=1}^{n} Y_i(t) - b_n(t)}{a_n(t)} \right\}_{t \in \mathbb{R}^d} \xrightarrow{\mathcal{D}} \{Z(t)\}_{t \in \mathbb{R}^d}. \tag{1.1}$$

Furthermore, we say Y lies in the *max-domain of attraction* (MDA) of the max-stable random field Z.
Max-stable processes are commonly used to model spatial extremal events, for instance maximum precipitation or maximum wind speed, on an annual scale. During the last decades, many different models for max-stable processes have been proposed (Smith, 1990; Schlather, 2002; Stoev and Taqqu, 2005; Kabluchko et al., 2009). A well-known and commonly used max-stable process is the mixed moving maxima process

$$Z(t) = \bigvee_{(s,u,X) \in \Phi} uX(t-s). \tag{1.2}$$

Here, the final shape of the process is determined by the maxima of scaled and shifted functions X where (s, u, X) are points of some point process Φ. In most cases, only some of the points (s, u, X) of Φ contribute to the final process Z – we call these points the *contributing points* or *extremal functions* (Dombry and Eyi-Minko, 2013). A special case of the mixed moving maxima process is the *Smith model* (Smith, 1990), where X equals the density function of a (multivariate) standard normal distribution - see also right plot in Figure 1.3.

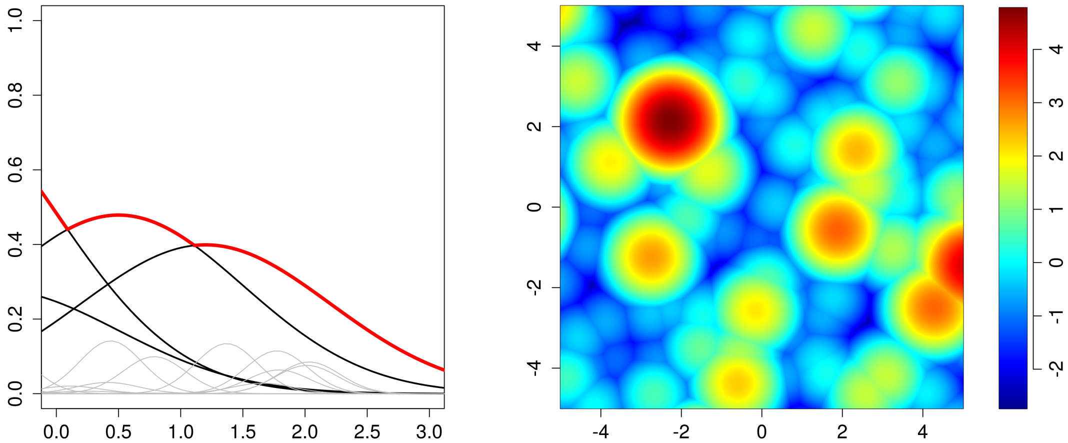

Figure 1.3.: The red curve in the left plot describes a realization of a mixed moving maxima process on $[0, 3]$. Each curves corresponds to one point of the underlying point process. Note that the grey curves do not contribute to the final process. The right plot depicts a two-dimensional Smith model.

However, the task of describing processes in the MDA has been far less examined. Obviously, each max-stable process lies in its own MDA – that is the trivial case. Our idea is that, since max-stable processes are used to model extremes on an annual scale, a process in the MDA might have the potential to model extremes on much smaller time scale. The α-stable processes (Samorodnitsky and Taqqu, 1994; Stoev and Taqqu, 2005) are known to lie in the MDA, but they are not suitable for modelling real data.
In Part II of this dissertation, we derive the new class of *conditionally max-stable random fields* which are in the MDA of max-stable random fields, but which are not max-stable themselves. The main contribution of this part is the proof that our new process is indeed in the MDA of a familiar mixed moving maxima process. Besides, we show that inference is still feasible under some usual and reasonable conditions. This part is based on a joint work (Dirrler et al., 2016) with Martin Schlather and Kirstin Strokorb.

Part III: On a generalization of the Matérn hard-core process

The work on conditionally max-stable random fields inspired me to this last and most theoretical part of the thesis. While working on estimation procedures for our newly presented model, it turned out that it would simplify that task a lot if an explicit characterization (e.g. in terms of an intensity function) of the contributing points of the mixed moving maxima process (1.2) was known. This is a rather difficult problem on its own, since the contributing points are highly correlated with each other. However, the way the contributing points dominate the non-contributing points, is related to an early approach of Matérn (1960). The *Matérn hard-core processes* encompass different models for point processes where the original point pattern is thinned by a dependent thinning algorithm, i.e. some of the points are deleted and the probability that an individual point

is removed, depends on the other points of the sample.
However, both the thinning algorithms in the classical Matérn model and those in more recent generalizations (Månsson and Rudemo, 2002; Kuronen and Leskelä, 2013; Teichmann et al., 2013; Andersen and Hahn, 2016) are quite restrictive and not suitable for our demand. Therefore, the original aim fades a bit from the spotlight – we first generalize these models but also establish some connections to mixed moving maxima processes. Our model comprises the recent generalizations of the Matérn model mentioned above. Still we are able to keep most of our proofs quite short due to the usage of *Palm calculus*. First and second order statistics can be explicitly derived under rather mild conditions and the results of this part can be used to improve the estimation procedures of Part II. Most of the results in Part III have already been published in Dirrler and Schlather (2017).

Part I.

Point processes

2. Preliminaries on point processes

In this chapter, we briefly summarize the mathematical fundamentals which will be necessary in the subsequent parts. We hereby follow closely Møller and Waagepetersen (2004), Daley and Vere-Jones (2008) and Chiu et al. (2013) in the first two sections. The third section is loosely based on Møller and Waagepetersen (2004) but extended by ideas of my own.

2.1. Basic properties and notation

Let S be a metric space and $\mathscr{B} = \mathscr{B}(S)$ its Borel σ-field. We define the subset of bounded Borel sets by

$$\mathscr{B}_0 = \{B \in \mathscr{B} : \text{B is bounded}\}.$$

For a subset $\varphi \subset S$, we denote by $n(x)$ the number of points in φ. We call $\varphi \subset S$ *locally finite* if

$$n(\varphi \cap B) < \infty, \quad \forall B \in \mathscr{B}_0.$$

We define the space of locally finite subsets of S by

$$N_{\text{lf}} = \{\varphi \subset S : n(\varphi \cap B) < \infty, \text{ for all bounded } B \subset S\}$$

and the corresponding σ-algebra

$$\mathcal{N}_{\text{lf}} = \sigma(\{\varphi \in N_{\text{lf}} : n(\varphi \cap B) = m\} : \ B \subset S \text{ bounded and } m \in \mathbb{N}).$$

A point process Φ is a measurable mapping from a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ to $(N_{\text{lf}}, \mathcal{N}_{\text{lf}})$. That is, we regard point processes as random countable subsets of S. The distribution P of Φ is determined by

$$\mathrm{P}(F) = \mathbb{P}(\Phi \in F), \quad F \in \mathcal{N}_{\text{lf}}.$$

We further define the count function as

$$N(B) = n(\Phi \cap B). \tag{2.1}$$

Definition 1. *We define the nth-order moment measure $\mu^{(n)}$ of Φ as*

$$\mu^{(n)}(B) = \mathbb{E}\left(\sum_{(\xi_1,\dots,\xi_n)\in\Phi} \mathbb{1}_B(\xi_1,\dots,\xi_n)\right), \quad B \subset S^n \tag{2.2}$$

and the nth-order factorial moment measure as

$$\alpha^{(n)}(B) = \mathbb{E}\left(\sum_{(\xi_1,\ldots,\xi_n)\in\Phi}^{\neq} \mathbb{1}_B(\xi_1,\ldots,\xi_n) \right), \quad B \subset S^n. \tag{2.3}$$

The first order moment measure $\mu(B) = \mu^{(1)}(B)$ is also called *intensity measure* and can be interpreted as the mean number of points of Φ hitting the set B

$$\mu(B) = \mathbb{E}\sum_{\xi\in\Phi} \mathbb{1}_B(\xi) = \mathbb{E}N(B).$$

If the nth-order factorial moment measure can be written as

$$\alpha^{(n)}(B) = \int\cdots\int \mathbb{1}_B(\xi_1,\ldots,\xi_n)\rho^{(n)}(\xi_1,\ldots,\xi_n)\mathrm{d}\xi_1\cdots\mathrm{d}\xi_n, \quad B \subset S^n, \tag{2.4}$$

with some non-negative function $\rho^{(n)}$, then $\rho^{(n)}$ is called *nth-order intensity function.*

Definition 2. *The pair correlation function of a point process Φ, with existing first and second order density functions, is defined as*

$$g(\xi_1,\xi_2) = \frac{\rho^{(2)}(\xi_1,\xi_2)}{\rho(\xi_1)\rho(\xi_2)}. \tag{2.5}$$

Definition 3. *Consider a (possibly random) function $p : S \to [0,1]$. The point process $p\Phi$ obtained from Φ by independently deleting every point $\xi \in \Phi$ with probability $1-p(\xi)$ is called p-thinning of Φ.*

The p-thinning is an important tool to transform point processes. Since the points are independently deleted, p-thinning is sometimes also called *independent thinning.*

Definition 4. *Let Φ be a point process. A marked point process Φ_M is defined by randomly attaching marks m_ξ from some Polish space $\mathbb{M}$ to each point $\xi \in \Phi$. That is*

$$\Phi_M = \{(\xi, m_\xi),\ \xi \in \Phi\}$$

is a mapping into $(M_{\mathrm{lf}}, \mathcal{M}_{\mathrm{lf}})$, with the set of point configurations

$$M_{\mathrm{lf}} = \left\{\varphi \subset S \times \mathbb{M} : \{\xi \in S, (\xi, m_\xi) \in \varphi\} \in N_{\mathrm{lf}} \text{ and } (\xi, m_\xi), (\xi, m'_\xi) \in \varphi \Rightarrow m_\xi = m'_\xi\right\}$$

and its σ-algebra $\mathcal{M}_{\mathrm{lf}}$ which is defined analogous to $\mathcal{N}_{\mathrm{lf}}$.

2.2. Campbell measure and Palm distribution

Definition 5. *Let h be a non-negative and measurable function on $S \times N_{\mathrm{lf}}$. The reduced Campbell measure $C^!$ is a measure on $(S \times N_{\mathrm{lf}}, S \times \mathcal{N}_{\mathrm{lf}})$ defined by*

$$\int \sum_{\xi\in\varphi} h(\xi, \varphi \setminus \xi)\ \mathrm{P}(\mathrm{d}\varphi) = \int h(\xi,\varphi) C^!(\mathrm{d}(\xi,\varphi)).$$

By choosing $h(\xi, \varphi) = \mathbb{1}_{(\xi,\varphi)\in D}$ it is an immediate consequence of this definition, that

$$C^!(D) = \mathbb{E}\sum_{\xi\in\Phi} \mathbb{1}_{(\xi,\Phi\setminus\xi)\in D}, \quad D \subset S \times N_{\text{lf}}.$$

We henceforth assume that the intensity measure μ is σ-finite. Then the Campbell measure is also σ-finite and, in its first component, absolutely continuous with respect to μ. Its Radon-Nikodym density $P^!_\xi$ is called *reduced Palm distribution*. Therefore, the Campbell measure can be decomposed to

$$C^!(B \times F) = \int_B P^!_\xi(F)\mathrm{d}\mu(\xi), \quad B \subset S, F \in \mathcal{N}_{\text{lf}}$$

and we obtain that for non-negative functions $h : S \times N_{\text{lf}} \to [0, \infty)$

$$\mathbb{E}\sum_{\xi\in\Phi} h(\xi, \Phi \setminus \{\xi\}) = \int\int h(\xi, \eta)\mathrm{d}P^!_\xi(\eta)\mathrm{d}\mu(\eta).$$

Hence $P^!_\xi$ can be interpreted as the conditional distribution of $\Phi \setminus \{\xi\}$ given $\xi \in \Phi$.

2.3. Special point processes

Definition 6. *Let f be a density function on $B \in \mathscr{B}(S)$. Consider a point process Φ consisting of $n \in \mathbb{N}$ i.i.d. points $\xi_1, \ldots, \xi_n$ distributed according to f in B. Then Φ is called binomial point process and we write $\Phi \sim BP(f, n)$.*

The binomial process is quite restrictive and rarely used, but it is the starting point to derive more complex point process models. A canonical extension is to allow n to be random, which leads to the following definition.

Definition 7. *Let $\lambda(B) = \int_B \psi(s)\ \mathrm{d}s$ for a non-negative function ψ. A point process Φ is called Poisson point process with intensity (function) ψ if*

(i) $N(B) \sim poi(\lambda(B))$ for all $B \subset S$ with $\lambda(B) < \infty$,

(ii) for all $n \in \mathbb{N}$ and $B \subset S$ with $\mu(B) \in (0, \infty)$ it holds true that

$$\Phi \cap B|_{N(B)=n} \sim\ BP(\psi/\lambda(B), n).$$

We write $\Phi \sim PP(\psi)$ for short.

We call $PP(\psi)$ *homogeneous* if ψ is constant – otherwise *inhomogeneous*. Note that a single realization of a Poisson point process cannot be distinguished from a realization of a binomial point process. This is a consequence of the second condition in the definition above. The Poisson point process plays a fundamental role within the scope of point processes – comparable with the importance of the normal distribution for probability distributions.

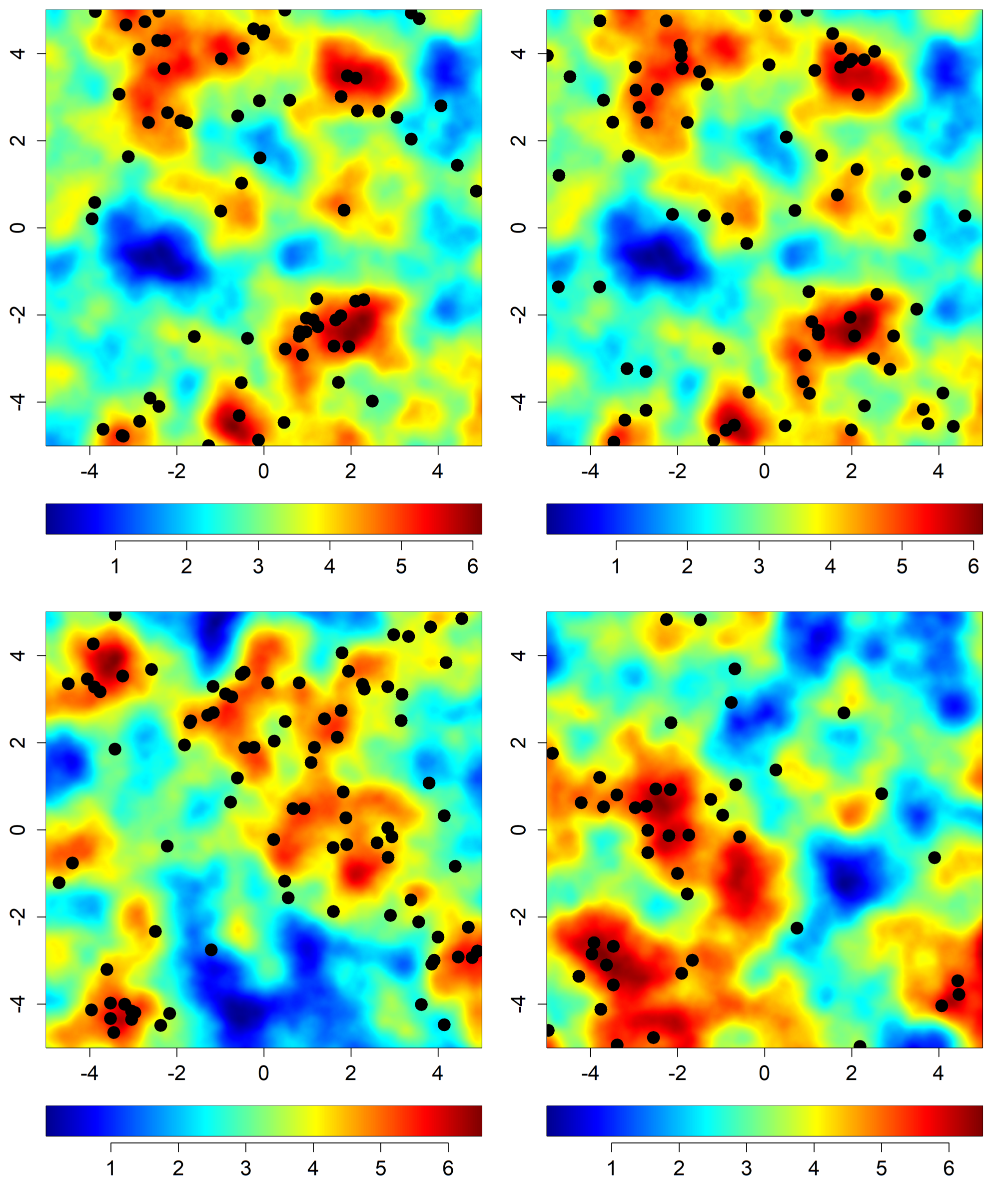

Figure 2.1.: Two realizations of an inhomogeneous Poisson point process and the underlying intensity function (upper left and upper right). The lower plots depict two realizations of a stationary Cox process and the underlying realizations of the intensity function.

The Poisson point process is probably the most commonly used point process in practice, but the assumption of a deterministic intensity function might be still too artificial for certain applications. Therefore it is a quite natural extension to allow the intensity function to be random itself.

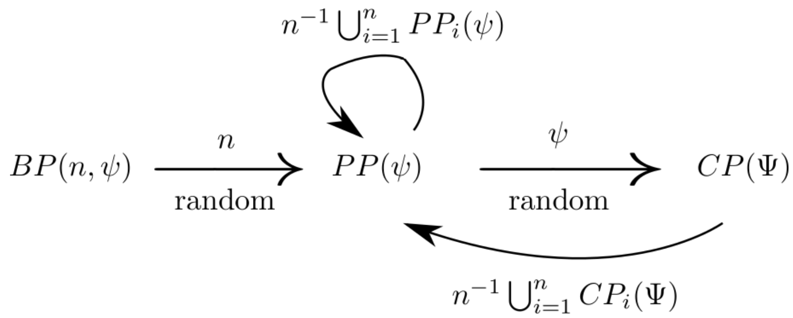

Figure 2.2.: Connections between binomial, Poisson and Cox process. A Poisson process is derived from a binomial process by imposing a Poisson distribution on n. A Cox process may be regarded as Poisson process with the additional property that its intensity function is allowed to be random. On the contrary, the n^{-1}-thinning of n i.i.d. Cox processes converges to a Poisson process if $\mathbb{E}\Psi(\cdot) = \psi(\cdot)$. The n^{-1}-thinning of Poisson processes remains a Poisson process.

Definition 8. *Consider an almost surely locally integrable random field Ψ. The point process Φ is called Cox process with (random) intensity function Ψ ($\Phi \sim CP(\Psi)$), if $\Phi|_{\Psi=\psi}$ is a Poisson process with intensity function ψ.*

For a Poisson process $\Phi \sim PP(\psi)$ the intensity measure $\mu(B)$ equals $\lambda(B)$. Note that the measure $\Lambda(B) = \int_B \Psi(s)\ \mathrm{d}s$ is random if $\Phi \sim CP(\Psi)$. We call Λ the *directing measure* of the Cox process Φ. The intensity measure of a Cox process is the mean of its directing measure, $\mu(B) = \mathbb{E}(\Lambda(B))$.

By definition, a Poisson process is a Cox process with deterministic directing measure. Still, a single realization of a Cox process cannot be distinguished from a single realization of a Poisson point process (or even a binomial point process). The following lemma underlines the importance of the Poisson point process and may be regarded as a central limit theorem for point processes. The lemma is an immediate consequence of Theorem 11.3.III in Daley and Vere-Jones (2008) and crucial for our work in Part II of this thesis.

Lemma 9. *Let $\Phi_i \overset{i.i.d.}{\sim} CP(\Psi)$, $i = 1, \dots, n$ be an i.i.d. sequence of Cox processes with directing measure $\Lambda(A) = \int_A \Psi(s)\ \mathrm{d}s$. Then, for $n \to \infty$*

$$n^{-1} \bigcup_{i=1}^{n} \Phi_i \to PP(\lambda), \quad \textit{where} \quad \lambda(A) = \mathbb{E}\Lambda(A), \quad \forall A \in \mathscr{B}(S).$$

Proof. The point process on the left-hand side is the $1/n$-thinning of $\bigcup_{i=1}^{n} \Phi_i$. Hence, by Theorem 11.3.III in Daley and Vere-Jones (2008) the desired convergence holds true if and only if $n^{-1} \sum_{i=1}^{n} \Psi_i$ converges to λ for i.i.d. copies Ψ_i of Ψ. This follows from the multivariate law of large numbers and Theorem 11.1.VII in Daley and Vere-Jones (2008). □

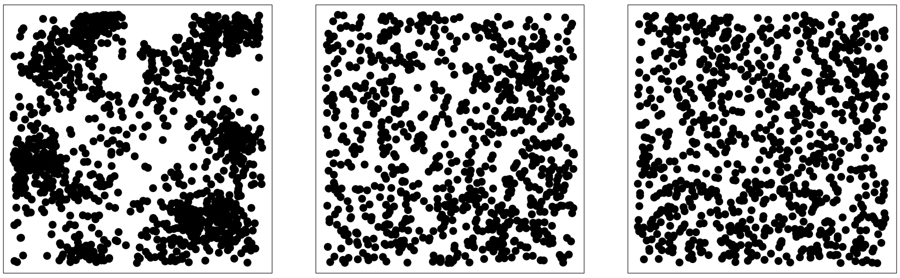

Figure 2.3.: Superposition of 1 (left), 30 (centre) and 300 (right) thinned stationary Cox processes.

2.4. Hard-core point processes

In this section, the independent p-thinning introduced in Definition 3 is generalized. Matérn introduced point process models which are obtained from a homogeneous Poisson process by a dependent thinning method (Matérn, 1960). Let Φ be a Poisson process on $S = \mathbb{R}^d$ with intensity λ. In the Matérn I model, all points $\xi \in \Phi$ that have neighbours within a deterministic hard-core distance R are deleted. The remaining points can be described by

$$\Phi_{\text{MatI}} = \{\xi \in \Phi : \Phi \cap B_R(\xi) \setminus \{\xi\} = \emptyset\}.$$

The Matérn II model considers a marked point process Φ_M where each point $\xi \in \Phi$ is independently endowed with a random mark $m_\xi \sim \mathcal{U}[0,1]$. A point $(\xi, m_\xi) \in \Phi_M$ is retained in the thinned process if the sphere $B_R(\xi)$ contains no points $\xi' \in \Phi \setminus \{\xi\}$ with $m_{\xi'} < m_\xi$. That is, the remaining points are

$$\Phi_{\text{MatII}} = \{(\xi, m_\xi) \in \Phi_M : m_\xi < m_{\xi'},\ \forall \xi' \in \Phi \cap B_R(\xi) \setminus \{\xi\}\}.$$

We revisit the Matérn hard-core processes in Part III of this thesis.

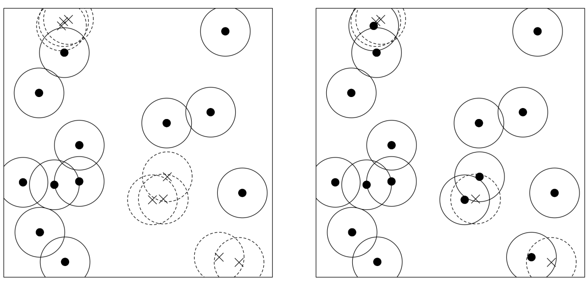

Figure 2.4.: Matérn hard-core model I (left) and II (right) with hard-core distance $R = 1$ based on the same Poisson point process with intensity $\lambda = 0.25$.

Part II.

Conditionally Max-stable Random Fields

3. Cox extremal process

Probabilistic modelling of spatial extremal events is often based on the assumption that daily observations lie in the max-domain of attraction of a max-stable random field which justifies statistical inference by means of block maxima procedures. This methodology is applied in various branches of environmental sciences, for instance, heavy precipitation (Cooley, 2005), extreme wind speads (Engelke et al., 2015; Genton et al., 2015; Oesting et al., 2015) and forest fire danger (Stephenson et al., 2015).
At the same time modelling extreme observations on a smaller time scale is a much more intricate issue and to date only few non-trivial processes are known to lie in the max-domain of attraction (MDA) of familiar max-stable processes. Among them α-stable processes (Samorodnitsky and Taqqu, 1994) form a natural class which may be rich enough to cover a wide range of environmental sample path behaviour (Stoev and Taqqu, 2005) and scale mixtures of Gaussian processes with regularly varying scale, are known to lie in the domain of attraction of extremal t-processes (Opitz, 2013). However, stable processes are themselves complicated objects whose statistical inference is a challenging research topic (Nolan, 2016) and scale mixtures of Gaussian processes have an unnatural degree of long-range dependence.
Our objective in this part of the thesis is to introduce another class of spatial processes in the MDA of familiar max-stable models, which encompasses processes with short-range dependence and to explore whether statistical inference on them is feasible, at least under some reasonable restrictions.
It is well-understood that max-stable processes can be built from Poisson point processes (de Haan, 1984; Giné et al., 1990; Stoev and Taqqu, 2006). In order to define our new class of models, we modify the underlying Poisson point process such that its intensity function is no longer fixed, but may depend on some spatial random effects. We pursue this idea by introducing conditionally max-stable processes based on Cox processes (Cox, 1955) which naturally generalize the class of mixed moving maxima processes (Smith, 1990; Schlather, 2002; Zhang and Smith, 2004; Stoev, 2008). Section 3.1 contains the definition of our proposed model. A functional convergence theorem shows that these processes lie in the MDA of familiar mixed moving maxima processes (Section 3.2). From a practical point of view, we choose to model the spatial random effects that influence the intensity function by a log Gaussian random field which makes the theory and application of log Gaussian Cox processes conveniently available for our setting, cf. Møller et al. (1998); Møller and Waagepetersen (2004); Møller and Schoenberg (2010); Diggle et al. (2013). We discuss in Section 3.3 how exact simulation of our proposed model can be traced back to exact simulation of max-stable random fields as in Schlather (2002). Inference on our new model is postponed to the subsequent chapter.
This chapter is based on the first part of Dirrler et al. (2016), where I am responsible

for the main contribution – but particularly Lemma 10 is strongly influenced by my co-authors.
Please note that we switch the notation of point processes within this part of the thesis – here we regard point processes as special cases of random measures, see for instance Daley and Vere-Jones (2008).

3.1. Model specification

Let X be a (possibly deterministic) non-negative stochastic process on $\mathbb{R}^d$ that we call *storm process* or *shape*. We assume X to have continuous sample paths. Based on its law $\mathbb{P}_X$ (on the complete separable metric space $\mathbb{X} = C(\mathbb{R}^d)$ with the usual Fréchet metric) and another sample-continuous positive stochastic process Ψ on $\mathbb{R}^d$, to be called *spatial intensity process*, and a positive *scaling constant* μ_Y, we define a random field Y on $\mathbb{R}^d$ by

$$Y(t) = \bigvee_{i=1}^{\infty} u_i X_i(t - s_i), \qquad t \in \mathbb{R}^d, \tag{3.1}$$

where $N = \sum_{i=1}^{\infty} \delta_{(s_i,u_i,X_i)}$ is a Cox-process on $S = \mathbb{R}^d \times (0,\infty] \times \mathbb{X}$, directed by the random measure

$$\mathrm{d}\Lambda(s,u,X) = \mu_Y^{-1}\, \Psi(s)\mathrm{d}s\, u^{-2}\mathrm{d}u\, \mathrm{d}\mathbb{P}_X. \tag{3.2}$$

The randomness of the measure Λ is due to the randomness of the spatial intensity process Ψ. Similarly to the situation with mixed moving maxima processes (Smith, 1990; Zhang and Smith, 2004; Stoev, 2008) or, more generally, extremal shot noise (Serra, 1984; Jourlin et al., 1988; Heinrich and Molchanov, 1994; Dombry, 2012), we will think of the processes X_i as being random storms centred around s_i that will affect its surroundings with severity u_i. In case, the intensity process is almost surely identically one ($\Psi \equiv 1$), the construction of Y is indeed the usual mixed moving maxima process

$$Z(t) = \bigvee_{i=1}^{\infty} u_i X_i(t - s_i), \qquad t \in \mathbb{R}^d, \tag{3.3}$$

where $\sum_{i=1}^{\infty} \delta_{(s_i,u_i,X^{(i)})}$ is the Poisson process on S with directing measure

$$\mathrm{d}\lambda(s,u,X) = {\mu_Z}^{-1}\, \mathrm{d}s\, u^{-2}\mathrm{d}u\, \mathrm{d}\mathbb{P}_X.$$

Note that, conditional on its intensity process Ψ, the extremal process Y is a (non-stationary) max-stable mixed moving maxima process. In the sequel, we call Y a *conditionally max-stable random field* or *Cox extremal process*.

3.2. Properties of Cox extremal processes

Continuity, Stationarity and Max-Domain of Attraction. Even though the Cox extremal process Y in (3.1) itself is not max-stable, we show in this section that it lies in the max-domain of attraction of an associated mixed moving maxima random field Z under rather general conditions. To show this, we first clarify some technical requirements that guarantee the finiteness and the continuity of sample paths of Y and Z.

Lemma 10 (Finiteness and Sample-Continuity). *Let K be a compact subset of $\mathbb{R}^d$.*

1. *If the integrability condition*

$$\mathbb{E}_\Psi\left(\mathbb{E}_X\left(\int_{\mathbb{R}^d}\sup_{t\in K} X(t-s)\Psi(s)\ \mathrm{d}s\right)\right) < \infty \tag{3.4}$$

holds, then $\sup_{t\in K} Y(t)$ is almost surely finite.

2. *If, additionally, the support of X contains some r-ball around the origin $o \in \mathbb{R}^d$ with positive probability, that is*

$$\exists\, r > 0 \text{ such that } \mathbb{P}_X\big(B_r(o) \subset \operatorname{supp}(X)\big) > 0 \tag{3.5}$$

with $\operatorname{supp}(X) = \{s \in \mathbb{R}^d : X(s) > 0\}$, then the sample paths of the process Y are almost surely continuous on K.

3. *If both (3.4) and (3.5) are satisfied for any compact $K \subset \mathbb{R}^d$, then Y is almost surely finite on compact sets and sample-continuous on $\mathbb{R}^d$. In this case only finitely many points of N contribute to Y on K.*

Proof. We follow closely the arguments of (Kabluchko et al., 2009, Proposition 13). For $K \subset \mathbb{R}^d$ and $c > 0$, set

$$I_c(K) = \left\{ i \in \mathbb{N} : \sup_{t\in K} u_i X_i(t-s_i) > c \right\}.$$

1. Conditional on the the process Ψ, the number of points in $I_c(K)$ is Poisson distributed with parameter

$$\Lambda\left(\left\{(s,u,X) : \sup_{t\in K} uX(t-s) > c\right\}\right) = c^{-1}\,\mathbb{E}_X \int_{\mathbb{R}^d}\sup_{t\in K} X(t-s)\Psi(s)\mathrm{d}s,$$

which is $\mathbb{P}_\Psi$-almost surely finite by the integrability condition (3.4). Hence, the number of points in $I_c(K)$ is almost surely finite, which entails that

$$\sup_{t\in K} Y(t) \le \bigvee_{i\in I_c(K)} \sup_{t\in K} u_i X_i(t-s_i) \vee c$$

is almost surely finite.

2. Due to its compactness, K can be split into finitely many (possibly overlapping) compact pieces $K_1, \dots, K_p$ of diameter less than r. Since the intensity process Ψ is positive and continuous almost surely, we also know $\inf_{s \in K} \Psi(s) > 0$ almost surely. Hence, there are almost surely infinitely many elements in $I_j := \{i \in \mathbb{N} : s_i \in K_j\}$ of the Cox process N (that underlies the construction of Y) in each of these pieces K_j, $j = 1, \dots, p$. Since there exists an $r > 0$ such that $\mathbb{P}_X(B_r(o) \subset \operatorname{supp}(X)) > 0$, there exists almost surely an element (in fact, infinitely many elements) $i_j \in I_j$ within the Cox process, such that $B_r(o) \subset \operatorname{supp}(X_{i_j})$. Summarizing, K is almost surely covered by

$$K \subset \bigcup_{j=1}^{p} K_j \subset \bigcup_{j=1}^{p} B_r(s_{i_j}) \subset \bigcup_{j=1}^{p} \operatorname{supp}(X_{i_j}(\cdot - s_{i_j})).$$

Setting $n := \max_{j=1}^{p} i_j$ and $c_j := \inf_{s \in B_r(o)} X_{i_j}(s) > 0$, we deduce that

$$\inf_{t \in K} \bigvee_{i=1}^{n} u_i X_i(t - s_i) \geq \inf_{t \in K} \bigvee_{j=1}^{p} u_{i_j} X_{i_j}(t - s_{i_j}) \geq \inf_{t \in K} \bigvee_{j=1}^{p} u_{i_j} c_j \mathbf{1}_{B_r(s_{i_j})}(t) \geq \bigvee_{j=1}^{p} u_{i_j} c_j > 0$$

is strictly greater than zero. Hence, there exists $n \in \mathbb{N}$, such that

$$Y(t) = \bigvee_{i \in I_c(K) \cup \{1,\dots,n\}} u_i X_i(t - s_i) \quad \forall t \in K$$

almost surely. That is, the process Y can be represented on K as the maximum of a finite number of continuous functions, which ensures the continuity of Y on K.

□

Remark 11. *The Cox extremal process Y is in general not uniquely determined by the choice of its shape X and intensity process Ψ. For instance, let $\tilde{X}$ be a process which satisfies the same assumptions as X, and independently of $\tilde{X}$, let ξ be a random variable, such that X can be decomposed into*

$$X(t) = \tilde{X}(t)\xi, \quad t \in \mathbb{R}^d.$$

Then choosing $\tilde{X}$ as shape and $\Psi \cdot \xi$ as intensity process does not alter the finite dimensional marginal distributions of the Cox extremal process Y, since

$$\begin{aligned} &\mathbb{P}(Y(t_1) \leq y_1, \dots, Y(t_n) \leq y_n) \\ &\quad = \mathbb{E}_\Psi \exp\left(-\mu_Y^{-1} \mathbb{E}_X \int \max_{i=1,\dots,n} \frac{X(t_i - s)}{y_i} \Psi(s) \, \mathrm{d}s \right). \end{aligned} \tag{3.6}$$

In the sequel, we will always assume that the intensity process Ψ is strictly stationary with

$$c_\Psi = \mathbb{E}_\Psi \Psi(o) < \infty. \tag{3.7}$$

This assumption simplifies some requirements of the preceding lemma. For instance, by Tonelli's theorem, condition (3.4) will be equivalent to

$$\mathbb{E}_X \left(\int_{\mathbb{R}^d} \sup_{t \in K} X(t-s) \mathrm{d}s \right) < \infty. \tag{3.8}$$

For $K = \{t\}$, we obtain that

$$\mathbb{E}_\Psi \left(\mathbb{E}_X \int_{\mathbb{R}^d} X(t-s)\Psi(s) \mathrm{d}s \right) = c_\Psi \cdot \mathbb{E}_X \left(\int_{\mathbb{R}^d} X(s) \ \mathrm{d}s \right) < \infty \tag{3.9}$$

entails the finiteness of $Y(t)$ as well as $Z(t)$ for $t \in \mathbb{R}^d$. In fact, the mixed moving maxima field Z in (3.3) has standard Fréchet margins if its scaling constant μ_Z equals (3.9) with $\Psi \equiv 1$, that is $c_\Psi = 1$. Condition (3.4) will be trivially satisfied for compact subsets K of $\mathbb{R}^d$ if Ψ is stationary, $c_\Psi \in (0, \infty)$ and

$$X \leq C \mathbf{1}_{B_R(o)}, \quad \mathbb{P}_X\text{-almost surely} \tag{3.10}$$

for some positive constants $C, R > 0$, where $B_R(o) \subset \mathbb{R}^d$ denotes the closed ball of radius R centred at $o \in \mathbb{R}^d$. Finally, stationarity of Ψ ensures that also the Cox extremal process Y built on the intensity process Ψ is stationary.

Lemma 12 (Stationarity)**.** *If the intensity process Ψ is stationary, then the Cox extremal process Y is stationary.*

Proof. The stationarity of Ψ and the invariance of the Lebesgue measure with respect to translations, gives that

$$\begin{aligned}
&\mathbb{P}\Big(Y(t_1 + h) \leq y_1, \ldots, Y(t_k + h) \leq y_k \Big) \\
&= \ \mathbb{E}_\Psi \Big[\exp\big(-\mu_Y^{-1} \mathbb{E}_X \int_{\mathbb{R}^d} \bigvee_{j=1}^k \frac{X(t_j + h - s)}{y_j} \Psi(s) \ \mathrm{d}s \big) \Big] \\
&= \ \mathbb{E}_\Psi \Big[\exp\big(-\mu_Y^{-1} \mathbb{E}_X \int_{\mathbb{R}^d} \bigvee_{j=1}^k \frac{X(t_j - s)}{y_j} \Psi(s + h) \ \mathrm{d}s \big) \Big]
\end{aligned}$$

equals (3.6). $\square$

Remark 13. *In the definition of the Cox extremal process Y it is also possible to work with storm processes X that may attain negative values, such as Gaussian processes. If at least* (3.5) *is satisfied, the resulting random field Y will be almost surely strictly positive.*

The following theorem is the main result of this section.

Theorem 14 (Max-Domain of Attraction). *Let the (sample-continuous) intensity process Ψ be stationary and almost surely strictly positive satisfying* (3.7) *and let the (sample-continuous) storm process X satisfy conditions* (3.8) *and* (3.5)*. Then the random fields Y and Z are finite on compact sets and sample-continuous, and the random field Y lies in the max-domain of attraction of Z. More precisely, if the scaling constant μ_Y equals the integral* (3.9) *and $\mu_Z = \mu_Y / c_\Psi$, then the following convergence holds weakly in $C(\mathbb{R}^d)$*

$$n^{-1} \bigvee_{i=1}^{n} Y_i \to Z,$$

where Y_i are i.i.d. copies of Y.

We will now prepare to prove Theorem 14. To this end we set the left-hand-side $Y^{(n)} := n^{-1} \left(\bigvee_{i=1}^{n} Y_i\right)$ which can be more conveniently represented as

$$Y^{(n)}(t) \stackrel{d}{=} \bigvee_{i=1}^{\infty} u_i X_i(t - s_i), \qquad t \in \mathbb{R}^d,$$

where $N_n = \sum_{i=1}^{\infty} \delta_{(s_i, u_i, X_i)}$ is a Cox-process on $\mathbb{R}^d \times (0, \infty] \times \mathbb{X}$, directed by the random measure

$$\mathrm{d}\Lambda_n(s, u, X) = \mu_Y^{-1}\, n^{-1} \sum_{i=1}^{n} \Psi_i(s) \mathrm{d}s\, u^{-2} \mathrm{d}u\, \mathrm{d}\mathbb{P}_X,$$

and Ψ_i, $i = 1, \dots, n$ represent i.i.d. copies of Ψ. The random measure Λ_n is the directing measure of the union of the underlying independent Cox-processes of the random fields Y_i, $i = 1, \dots, n$, scaled by n^{-1}. By Lemma 9, the point process N_n converges weakly to the Poisson-process with directing measure

$$\mathrm{d}\lambda(s, u, X) = {\mu_Z}^{-1}\, \mathrm{d}s\, u^{-2} \mathrm{d}u\, \mathrm{d}\mathbb{P}_X$$

that underlies the mixed moving maxima random field Z. The latter convergence indicates already the result of Theorem 14. In order to prove Theorem 14, we show first the convergence of the finite dimensional distributions and then the tightness of the sequence $Y^{(n)}$, $n = 1, 2, \dots$.

Lemma 15 (Convergence of finite-dimensional distributions). *Let the random fields Y and Z be specified as in Theorem 14, then the finite dimensional distributions of $Y^{(n)}$ converge to those of Z as $n \to \infty$.*

Proof. We fix $t_1, \dots, t_k \in \mathbb{R}^d$ and show that the random vector $(Y(t_1), \dots, Y(t_k))$ lies in the max-domain of attraction of the random vector $(Z(t_1), \dots, Z(t_k))$. It then automatically follows that the finite dimensional distributions $Y^{(n)}$ converge to those of Z, since the scaling constants for each individual $t \in \mathbb{R}^d$ are chosen appropriately.

For $y = (y_1, \dots, y_k) \in (0,\infty)^d$, it follows from (3.9) that the non-negative random variable

$$H_\Psi(y) := \mu_Y^{-1} \mathbb{E}_X \int \max_{1\le j\le k} \frac{X(t_j - s)}{y_j} \Psi(s)\, \mathrm{d}s$$

satisfies that its first moment $\mathbb{E}_\Psi(H_\Psi(y)) < \infty$ is finite and can be gained from its Laplace transform via

$$\mathbb{E}_\Psi(H_\Psi(y)) = -\lim_{t\downarrow 0} \frac{d}{\mathrm{d}t} \mathbb{E}_\Psi\left(e^{-t\, H_\Psi(y)}\right).$$

Hence, by l'Hôpital's rule

$$\lim_{\lambda\to\infty} \frac{1 - \mathbb{P}(Y(t_1) \le \lambda y_1, \dots, Y(t_k) \le \lambda y_k)}{1 - \mathbb{P}(Y(t_1) \le \lambda, \dots, Y(t_k) \le \lambda)} = \lim_{t\to 0} \frac{1 - \mathbb{E}_\Psi(e^{-t\, H_\Psi(y)})}{1 - \mathbb{E}_\Psi(e^{-t\, H_\Psi(1)})}$$
$$= \frac{\mathbb{E}_\Psi(H_\Psi(y))}{\mathbb{E}_\Psi(H_\Psi(1))} =: V(y)$$

with $V(cy) = c^{-1}y$. Moreover, $V(y)$ is a multiple of the exponent function of the max-stable random vector $(Z(t_1), \dots, Z(t_k))$

$$-\log \mathbb{P}(Z(t_1) \le y_1, \dots, Z(t_k) \le y_k) = \mu_Z^{-1} \mathbb{E}_X \int \max_{1\le j\le k} \frac{X(t_j - s)}{y_j}\, \mathrm{d}s = \mathbb{E}_\Psi(H_\Psi(y)).$$

Hence, by (Resnick, 2008, Corollary 5.18 (a)), the random vector $(Y(t_1), \dots, Y(t_k))$ lies in its domain of attraction. ☐

The following lemma will be useful to prove the tightness of the sequence $Y^{(n)}$.

Lemma 16. *Let a_n and b_n be bounded sequences of non-negative real numbers, then*

$$\left| \bigvee_{n=1}^\infty a_n - \bigvee_{n=1}^\infty b_n \right| \le \bigvee_{n=1}^\infty |a_n - b_n|.$$

Proof. The statement is the triangle inequality $|\|a\|_\infty - \|b\|_\infty| \le \|a - b\|_\infty$ with $\|\cdot\|_\infty$ the ℓ^∞ norm in the space of bounded sequences. ☐

Lemma 17 (Tightness)**.** *Let the random field Y be specified as in Theorem 14, then the sequence of random fields $Y^{(n)}$ is tight.*

Proof. Since the finiteness of Y does also ensure the finiteness of each $Y^{(n)}$, it suffices to show that, for a compact set $K \subset \mathbb{R}^d$, the modulus of continuity

$$\omega_K\left(Y^{(n)}, \delta\right) := \sup_{t_1, t_2 \in K \,:\, \|t_1 - t_2\| \le \delta} \left| Y^{(n)}(t_1) - Y^{(n)}(t_2) \right|$$

satisfies the convergence

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\omega_K\left(Y^{(n)},\delta\right)>\varepsilon\right)=0. \tag{3.11}$$

To simplify the notation, we introduce

$$K_\delta := \left\{(t_1,t_2)\in\mathbb{R}^d\times\mathbb{R}^d \,:\, \|t_1-t_2\|\le\delta, t_1,t_2\in K\right\}.$$

By the definition of $Y^{(n)}$ and the preceding Lemma 16, we have

$$\begin{aligned}\mathbb{P}\left(\omega_K\left(Y^{(n)},\delta\right)\le\varepsilon\right) &= \mathbb{P}\left(\sup_{(t_1,t_2)\in K_\delta}\left|\bigvee_{i=1}^{\infty}u_iX_i(t_1-s_i)-\bigvee_{i=1}^{\infty}u_iX_i(t_2-s_i)\right|\le\varepsilon\right)\\ &\ge \mathbb{P}\left(\sup_{(t_1,t_2)\in K_\delta}\bigvee_{i=1}^{\infty}u_i\left|X_i(t_1-s_i)-X_i(t_2-s_i)\right|\le\varepsilon\right).\end{aligned}$$

As the tuples (s_i,u_i,X_i), $i\in\mathbb{N}$, are the points of the Cox process N_n, we can compute the latter probability as expected void-probability. To this end, let us denote the joint probability law of the i.i.d. intensity processes Ψ_i, $i=1,2,\dots$ and its expectation by $\mathbb{P}_\Psi$ and $\mathbb{E}_\Psi$, respectively. Setting $\Psi^{(n)}(s):=n^{-1}\sum_{i=1}^n\Psi_i(s)$, we obtain

$$\begin{aligned}&\liminf_{n\to\infty}\mathbb{P}\left(\omega_K\left(Y^{(n)},\delta\right)\le\varepsilon\right)\\ &\ge \liminf_{n\to\infty}\mathbb{E}_\Psi\left[\exp\left(-\varepsilon^{-1}\mu_Y^{-1}\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{(t_1,t_2)\in K_\delta}\left|X(t_1-s)-X(t_2-s)\right|\Psi^{(n)}(s)\mathrm{d}s\right)\right]\\ &\ge \mathbb{E}_\Psi\left[\liminf_{n\to\infty}\exp\left(-\varepsilon^{-1}\mu_Y^{-1}\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{(t_1,t_2)\in K_\delta}\left|X(t_1-s)-X(t_2-s)\right|\Psi^{(n)}(s)\mathrm{d}s\right)\right],\end{aligned}$$

where the last inequality follows from Fatou's Lemma. Moreover, the strong law of large numbers and condition (3.4) (which ensures the existence and finiteness of the following right-hand side) yield that $\mathbb{P}_\Psi$-almost surely

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{(t_1,t_2)\in K_\delta}\left|X(t_1-s)-X(t_2-s)\right|\Psi^{(n)}(s)\mathrm{d}s\\ &= c_\Psi\,\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{(t_1,t_2)\in K_\delta}\left|X(t_1-s)-X(t_2-s)\right|\mathrm{d}s\\ &\le c_\Psi\,\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{t\in B_\delta(o)}\left|X(s-t)-X(s)\right|\mathrm{d}s,\end{aligned}$$

which entails

$$\liminf_{n\to\infty}\mathbb{P}\left(\omega_K\left(Y^{(n)},\delta\right)\le\varepsilon\right)\ge\exp\left(-\varepsilon^{-1}\mu_Y^{-1}c_\Psi\,\mathbb{E}_X\int_{\mathbb{R}^d}\sup_{t\in B_\delta(o)}\left|X(s-t)-X(s)\right|\mathrm{d}s\right).$$

Finally, in order to establish (3.11), it remains to be shown that

$$\lim_{\delta \to 0} \mathbb{E}_X \int_{\mathbb{R}^d} \sup_{t \in B_\delta(o)} |X(s-t) - X(s)| \, \mathrm{d}s = 0.$$

This, however, follows from the dominated convergence theorem, since for any fixed $X \in C(\mathbb{R}^d)$ and any fixed $s \in \mathbb{R}^d$ the convergence of the integrand to 0 holds true and by

$$\sup_{t \in B_\delta(o)} |X(s-t) - X(s)| \leq \sup_{t \in B_\delta(o)} X(s-t) + X(s)$$

and condition (3.8), there exists an integrable upper bound. □

We are now in position to prove the main result of this section.

Proof of Theorem 14. The finiteness and sample-continuity of the random fields Y and Z are an immediate consequence of Lemma 10. While Lemma 15 shows that the finite-dimensional distributions of the random fields $Y^{(n)}$ converge to those of the process Z, Lemma 17 establishes the tightness of the sequence $Y^{(n)}$. Collectively, this proves the assertions. □

Choices for the intensity process. For inference reasons we shall further assume henceforth that the intensity process Ψ is a stationary log Gaussian random field, that is,

$$\Psi(s) = \exp(W(s)), \qquad s \in \mathbb{R}^d,$$

where W is stationary and Gaussian. Thereby, all requirements for Ψ from the preceding Theorem 14 are guaranteed as long as W has continuous sample paths. Moreover, the latter also ensures that the distribution of the random measure Λ, cf. (3.2), is uniquely determined by the distribution of W. By Møller et al. (1998) (see also Adler (1981), page 60), a Gaussian process W is indeed sample-continuous if its correlation function C satisfies $1 - C(h) < M\|h\|^\alpha$, $h \in \mathbb{R}^d$, for some $M > 0$ and $\alpha > 0$. This condition holds for most common correlation functions, for instance, the stable model $C(h) = \exp(-\|h\|^\alpha)$, $\alpha \in (0, 2], h \in \mathbb{R}^d$, and the Whittle-Matérn model

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} (\sqrt{2\nu} h)^\nu K_\nu(\sqrt{2\nu} h), \quad \nu > 0, h \in \mathbb{R}^d, \tag{3.12}$$

see Guttorp and Gneiting (2006).

Choices for the storm profiles. For statistical inference, we rely on identifying at least some of the centres of the storms X_i from observations of Y. As a starting point, it is reasonable to assume that the paths of X satisfy a monotonicity condition, for instance that for each path X_ω there exist some monotonously decreasing functions f_ω and g_ω such that

$$g_\omega(\|t\|) \leq X_\omega(t) \leq f_\omega(\|t\|) \tag{3.13}$$

and $g_\omega(0) = X_\omega(0) = f_\omega(0)$. For the purpose of illustration, we will use in most of our examples a deterministic shape $X = \varphi$, with φ being the density of the d-dimensional standard normal distribution as in Smith (1990). See also Section 5 for a discussion of this choice and the recovery of storm centres.

3.3. Simulation

In many cases, functionals of max-stable processes cannot be explicitly calculated, e.g., for most models only the bivariate marginal distributions are known while the higher dimensional distributions do not have a closed-form expression. Therefore and in order to test estimation procedures, efficient and sufficiently exact simulation algorithms are desirable. However, exact simulation of (conditionally) max-stable random fields can be challenging, since a priori, its series representation (3.1) involves taking maxima over infinitely many storm processes. A first approach in order to simulate mixed moving maxima processes and some other max-stable processes was presented in Schlather (2002). Meanwhile, several improvements with respect to exactness and efficiency have been proposed in Engelke et al. (2011); Oesting et al. (2012, 2013); Dieker and Mikosch (2015); Dombry et al. (2016) and Liu et al. (2016). Since our focus in this work is not on the simulation algorithm, it will be sufficient for us to extend the straightforward approach of Schlather (2002) in this article.
Under the (mild) conditions of Lemma 10 only finitely many of the storms in (3.1) contribute to the maximum if we restrict the random field to a compact domain $D \subset \mathbb{R}^d$, see also de Haan and Ferreira (2006). Still, the centres of these contributing storms could be located on the whole $\mathbb{R}^d$. In order to define a feasible and exact algorithm we consider bounded storm profiles X which satisfy condition (3.10). In such a situation only storms with centres within the enlarged region

$$D_R = D \oplus B_R(o) = \bigcup_{s \in D} B_R(s),$$

can contribute to the maximum (3.1).

Proposition 18 (Simple Simulation Algorithm)**.** *Let $D \subset \mathbb{R}^d$ be a compact subset and assume that the conditions of Lemma 10 hold true and additionally the storm profile X satisfies almost surely* (3.10)*. Then the following construction leads to an exact simulation algorithm on D for the associated Cox extremal process Y.*

- *Let ψ be a realization of the intensity process Ψ and $\nu_\psi(\cdot) = \int_\cdot \psi(s)\,\mathrm{d}s$ the associated measure on $\mathbb{R}^d$.*
- *Let $S_i \overset{i.i.d.}{\sim} \psi/\nu_\psi(D_R)$, $i = 1, 2, \ldots$ be an i.i.d. sequence of random variables from the probability measure $\psi/\nu_\psi(D_R)$ on D_R.*
- *Let $X_i \overset{i.i.d.}{\sim} X$, $i = 1, 2, \ldots$ be an i.i.d. sequence of storm profiles.*

- *Let ξ_i, $i = 1, 2, \dots$ be an i.i.d. sequence of standard exponentially distributed random variables and set $\Gamma_n = \sum_{i=1}^n \xi_i$ for $n = 1, 2, \dots$.*

Based on the stopping time

$$T = \inf\left\{n \geq 1 : \Gamma_{n+1}^{-1} C \leq \inf_{t \in D} \bigvee_{i=1}^{n} \Gamma_i^{-1} X_i(t - S_i)\right\},$$

we define the random field $\widetilde{Y}$ on D via

$$\widetilde{Y}(t) = \frac{\nu_\psi(D_R)}{\mu_Y} \bigvee_{i=1}^{T} \Gamma_i^{-1} X_i(t - S_i), \quad t \in D.$$

In this situation the following holds true.

1. *The stopping time T is almost surely finite.*
2. *The law of the process $\widetilde{Y}$ coincides with the law of the Cox extremal process Y restricted to D.*

Proof. First note that $\sum_{i=1}^\infty \delta_{\Gamma_i}$ is a Poisson process on $\mathbb{R}_+$ with intensity 1. Hence $\sum_{i=1}^\infty \delta_{\Gamma_i^{-1}}$ is a Poisson process on $(0, \infty]$ with intensity $u^{-2}du$. Attaching the independent markings X_i and, for fixed $\Psi = \psi$, the markings $S_i \sim \psi(s)/\nu(D_R)$ yields that, for fixed $\Psi = \psi$, the point process $\sum_{i=1}^\infty \delta_{(S_i, \nu_\psi(D_R)\mu_Y^{-1}\Gamma_i^{-1}, X_i)}$ is a Poisson process directed by the measure $\mu_Y^{-1}\psi(s)\ \mathrm{d}s\, u^{-2}\ \mathrm{d}u\ \mathrm{d}\mathbb{P}_X$ on $D_R \times [0, \infty) \times \mathbb{X}$.
Since in the construction of Y only storms with center in D_R can contribute to the process Y on D, the law of Y on D and the law of

$$\frac{\nu_\psi(D_R)}{\mu_Y} \bigvee_{i=1}^{\infty} \Gamma_i^{-1} X_i(t - S_i), \quad t \in D$$

coincide. By definition of the stopping time T and since X is uniformly bounded by C, the latter has the same law as the process $\widetilde{Y}$ on D. So, it remains to be shown that T is almost surely finite. Similar to the proof of Lemma 10, it can be shown that

$$\exists\, n \in \mathbb{N} \,:\, \inf_{t \in D} \bigvee_{i=1}^{n} \Gamma_i^{-1} X_i(t - S_i) > 0 \quad \text{almost surely.} \tag{3.14}$$

Together with the decrease of the sequence Γ_{n+1}^{-1} this implies the a.s.-finiteness of T. □

Beyond this extension, we would like to point out that in fact all previous procedures for simulation of non-stationary mixed moving maxima processes can be adapted for the simulation of Cox extremal processes in a similar way. For instance, by using a transformed representation of the original process Y, the efficiency improvement of Oesting et al. (2013) can be transferred as well.

Remark 19. *When condition* (3.10) *is not satisfied, we choose R and C such that*

$$\mathbb{P}\left(\sup_{t\in\mathbb{R}^d\setminus B_R(o)} X(t) > \varepsilon\right) \leq \alpha \quad \textit{and} \quad \mathbb{P}\left(\sup_{t\in B_R(o)} X(t) > C\right) \leq \alpha \tag{3.15}$$

hold true for some prescribed small $\varepsilon > 0$ and $\alpha > 0$ and approximate X by its truncation $\min(X\mathbb{1}_{B_R}, C)$ in the preceding algorithm, whence simulation will be only approximately exact. For example, let us consider a generalization of the Smith model in $\mathbb{R}^d$, i.e. $X = \varphi$ with φ the d-variate standard normal density. Then for arbitrary $\varepsilon > 0$, (3.15) *is satisfied with $\alpha = 0, R = \sqrt{-d\log(2\pi) - 2\log(\varepsilon)}$ and $C = (2\pi)^{-d/2}$. Figure 3.1 depicts two plots of a Cox extremal process Y and its underlying log Gaussian random field Ψ.*

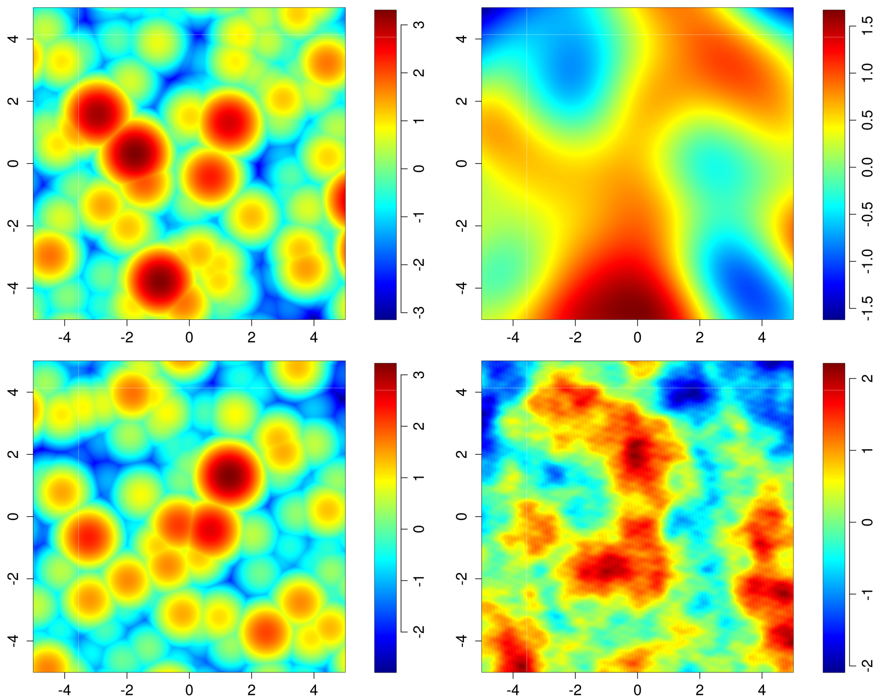

Figure 3.1.: Cox extremal processes Y (left) and underlying log Gaussian random fields Ψ (right). The covariance of $\log\Psi$ is of Whittle-Matérn type with var $= 1$, scale $= 2$ and $\nu = \infty$ (upper plots) and $\nu = 1$ (lower plots) respectively. The plots have been transformed to a logarithmic scale and the storm profiles have deterministic shape $X = \varphi$.

4. Inference on the underlying Cox process

This chapter is based on the second part of Dirrler et al. (2016). We further examine the Cox process N that underlies the Cox extremal process (3.1). More specifically, we want to perform inference on the intensity process Ψ that influences the intensity of N. The process Ψ is modelled by a log Gaussian random field – we take a closer look at two important aspects in the recovery of the underlying Gaussian process.
In Section 4.1, we deal with non-parametric estimation of realizations of the random intensity of the Cox process from observations of the conditionally max-stable processes and their storm centres. Based on the outcome of this procedure, we consider parametric estimation of the covariance function of the Gaussian process in Section 4.2. The performance of these procedures is examined in Section 4.3 in a simulation study.

4.1. Non-parametric inference on the realization ψ of the intensity process Ψ

For practical purposes it is critical to understand how one can recover (i) the storm profile X and (ii) the intensity process Ψ from i.i.d. replicates of the Cox extremal process Y that they induce via (3.1).
For inference on X note that, by Theorem 14 the process Y as in (3.1) lies in the MDA of the ordinary mixed moving maxima process Z as in (3.3) . Then $m^{-1} \bigvee_{j=1}^{m} Y_j$ equals approximately Z for sufficiently large m and the distribution of X can be estimated using methods for estimating the shape of Z itself (note that Z does not depend on Ψ). Among these are for instance *madograms* (see Matheron (1987) and Cooley (2005)), *censored likelihood* (Nadarajah et al., 1998; Schlather and Tawn, 2003) or *composite likelihood* (Castruccio et al., 2015). The recent article of Huser et al. (2016) gives an overview over likelihood methods. We henceforth assume that the distribution of X is known and focus on the second question (ii), the inference on the intensity process Ψ.
To understand the stochastic mechanism behind Ψ, we first need to understand how we can recover a single realization ψ of Ψ from a corresponding single realization y of Y. To this end, we assume the following general strategy.

1. **Determination of the visible storm centres.** If the storm profiles X_i assume their global maximum at the origin and decay monotonously, then the local maxima of y are the visible storm centres. They constitute approximately a sample of a point process n_K^y whose intensity has a close link to the intensity ψ.

2. **Estimation of the realization ψ.** We use a kernel estimator to get a first estimate for ψ. Such an estimator is necessarily biased.

3. **Correcting the bias.** We will make an artificial assumption on the observations to obtain a reasonable approximation of the spatially varying bias factor $b^y_K(s)$. The original kernel estimator is divided by the bias factor to obtain the final estimator for ψ.

We will show in a simulation study (Section 4.3) that our approach works reasonably well. In this section we underpin it theoretically.
To this end, we consider the point process N^Y_K of locations whose corresponding shape functions contribute to Y

$$N^Y_K = \sum_{i=1}^{\infty} \delta_{s_i} \mathbb{1}_{\{\sup_{t\in K} X_i(t-s_i)Y(t)^{-1} \geq u_i^{-1}\}}, \tag{4.1}$$

on a compact set $K \subset \mathbb{R}^d$, see also Figure 4.1. That is, N^Y_K equals the location component of the process of extremal functions introduced by Dombry and Eyi-Minko (2013) and Oesting and Schlather (2014). We call N^Y_K the contributing storm centres of Y on K and denote a realization of N^Y_K by n^y_K. Note that ψ, y and n^y_K are directly related, that is, ψ is the realization of the intensity process Ψ which leads to the realization y of the Cox extremal process Y whose contributing storm centres are n^y_K.
What complicates statistical inference is that the process N^Y_K is not a Cox process anymore and hence there is no straightforward way to derive, for instance, its intensity. We circumvent this problem by considering the following modified process

$$N^{Y^*}_K = \sum_{i=1}^{\infty} \delta_{s_i} \mathbb{1}_{\{\sup_{t\in K} X_i(t-s_i)Y^*(t)^{-1} \geq u_i^{-1}\}} \tag{4.2}$$

where Y^* is an almost surely positive random field. The following proposition states conditions which enable us to derive some useful properties of $N^{Y^*}_K$.

Proposition 20. *Assume that the conditions* (3.10) *and* (3.5) *are satisfied. Let* $Y^*|_{\Psi=\psi}$ *be an independent copy of* $Y|_{\Psi=\psi}$ *which is independent of* $N|_{\Psi=\psi}$. *Let* $K \subset \mathbb{R}^d$ *be a compact set. Then* $N^{Y^*}_K$ *is a Cox process on* $\mathbb{R}^d$. *More specifically*

$$N^{Y^*}_K|_{\Psi=\psi, Y^*=y} \sim PP\left(b^y_K(s)\psi(s)\right)$$

is a Poisson point process, whose intensity function equals $\psi(s)$ *up to the correcting factor*

$$b^y_K(s) = \mu_Y^{-1}\mathbb{E}_X\left[\sup_{t\in K}\frac{X(t-s)}{y(t)}\right], \quad s \in \mathbb{R}^d. \tag{4.3}$$

Proof. Since $Y^*|_{\Psi=\psi}$ is independent of $N|_{\Psi=\psi}$, the process $N^{Y^*}_K|_{\Psi=\psi,Y^*=y}$ is an independent thinning of the Poisson process $N|_{\Psi=\psi}$. The number of points in the set

$$\left\{s \in K_R : (s,u,X) \in N|_{\Psi=\psi},\ u^{-1} \leq \sup_{t\in K}\frac{X(t-s)}{y(t)}\right\}$$

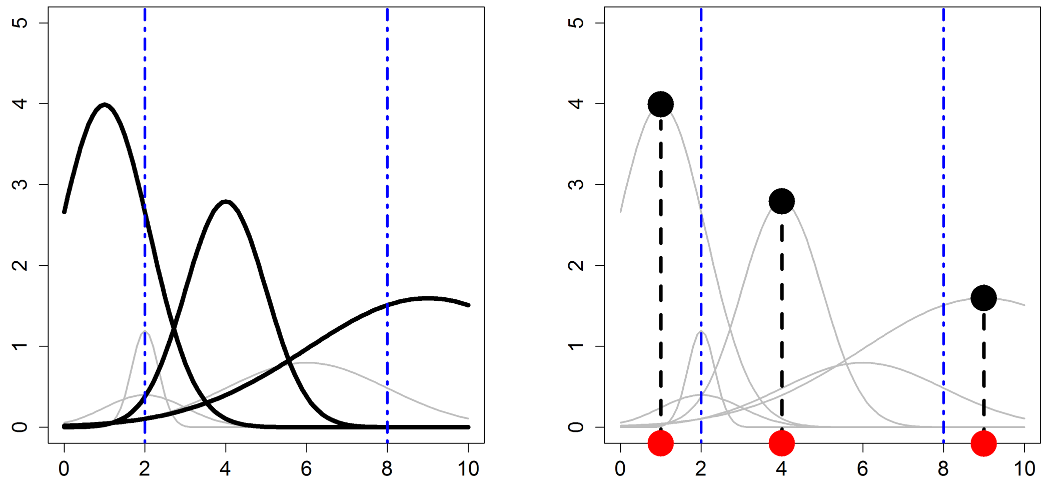

Figure 4.1.: Left plot: Contributing storms (black), the grey ones do not contribute to the final process on $K = [2, 8]$. Right plot: Storm centres (black dots) and location of the storms (red). The red dots correspond to the point process N_K^Y. Some points of N_K^Y are outside of K.

is Poisson distributed with parameter

$$\mu_Y{}^{-1} \int_{K_R} \mathbb{E}_X \sup_{t \in K} \frac{X(t-s)}{y(t)} \psi(s) \, \mathrm{d}s.$$

This finishes the proof. □

Remark 21. *This result can be stated in a more general setting. Let f and $\tilde{f}$ be arbitrary functions which satisfy $\sigma(f(Y^*), \tilde{f}(Y^*)) = \sigma(Y^*)$. Suppose that $\sum_{i=1}^{\infty} \delta_{Y_i^*}$ is a Cox process with intensity $\int \mathrm{d}\mathbb{P}_{Y^*|\tilde{f}(Y^*)}$. Then*

$$N_K^{Y^*} \sim CP\left(\mathbb{E}_{Y^*}\left[\mathbb{E}_X\left(\sup_{t \in K} X(t-s)(Y^*(t))^{-1}\right) \middle| \tilde{f}(Y^*)\right] \Psi(s)\right),$$

if additionally

$$Y^*|f(Y^*), \Psi \perp\!\!\!\perp N|f(Y^*), \Psi.$$

This implies the statement of Proposition 20 by choosing $\tilde{f} = \mathrm{id}$.

The correcting factor $b_K^y(s)$ is in principle known if the distribution of the shape process X is known and can be evaluated numerically. We henceforth use n_K^y as an estimate of a realization of $N_K^{Y^*}|_{\Psi=\psi, Y^*=y}$, assuming that $N_K^{\psi,y} := N_K^Y|_{\Psi=\psi, Y=y}$ approximates $N_K^{\psi,y^*} := N_K^{Y^*}|_{\Psi=\psi, Y^*=y}$ sufficiently well in practice, even if the independence assumption of Proposition 20 is violated. See Section 5 for a discussion of this assumption. In particular, simulation results are promising that the error made is not too big compared to other effects.

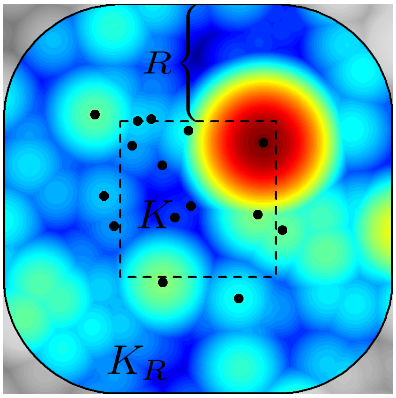

Figure 4.2.: Domain of observation $\mathcal{D}$ (big square) and domain of estimation K_R. The process N_K^Y takes value on K_R and its points are marked by black circles.

As a consequence of our approach, we have that $N_K^{\psi,y} \approx PP(b_K^y(s)\psi(s))$.
Note that a single observation of N_K^Y cannot be distinguished from a single observation of $N_K^{\psi,y}$ or $N_K^y := N_K^Y|_{Y=y}$. That is, n_K^y can be regarded as a realization of each of these processes.
We assume now that a single realization of y is observed on a set $\mathcal{D}$ which fulfils the equation $K = \mathcal{D} \ominus B_R(o)$ for a compact set K. We further assume that n_K^y can be recovered from y. The conditions (3.10) and (3.5) imply that N_K^Y is almost surely a finite point process. Furthermore, with $K_R = K \oplus B_R(o)$ we have $\operatorname{supp}\left(N_K^Y\right) \subset K_R$ almost surely, that is, the support of the correction factor b_K^y lies in K_R.
We derive a non-parametric estimator of ψ on a set $D \subset K_R = K \oplus B_R(o)$, see Figure 4.2 for illustration. The distribution of the shape function X is assumed to be known. Then $\psi_K^y(s) = b_K^y(s)\psi(s)$ can be estimated non-parametrically by the kernel estimator (Diggle, 1985)

$$\widehat{\psi}_D^y(s) = h^{-d} \sum_{t \in N_K^{\psi,y^*} \cap D} c_D(t)^{-1} k\left(\frac{s-t}{h}\right), \quad s \in D, \quad D \subset K_R, \tag{4.4}$$

with bandwidth h and the Epanechnikov kernel

$$k(s) = \frac{d+2}{2|B_1(o)|}(1 - \|s\|^2)\mathbb{1}_{B_1(o)}(s).$$

To compensate edge effects, weights $c_D(t) = h^{-d}\int_D k\left(\frac{s-t}{h}\right) \mathrm{d}s$ are included as proposed in Ripley (1977). The impact of the bandwidth h is strong and several approaches of figuring out a reasonable bandwidth can be found in Diggle (1985) and Stoyan and Stoyan (1992).

Lemma 22. *The estimator $\int_D \widehat{\psi}_D^y(s)\,\mathrm{d}s$ is unbiased for $\int_D \psi^y(s)\,\mathrm{d}s$, that is*

$$\mathbb{E}\int_D \widehat{\psi}_D^y(s)\,\mathrm{d}s = \int_D \psi^y(s)\,\mathrm{d}s \quad \forall h \in \mathbb{R}_+, \quad \forall D \subset K_R.$$

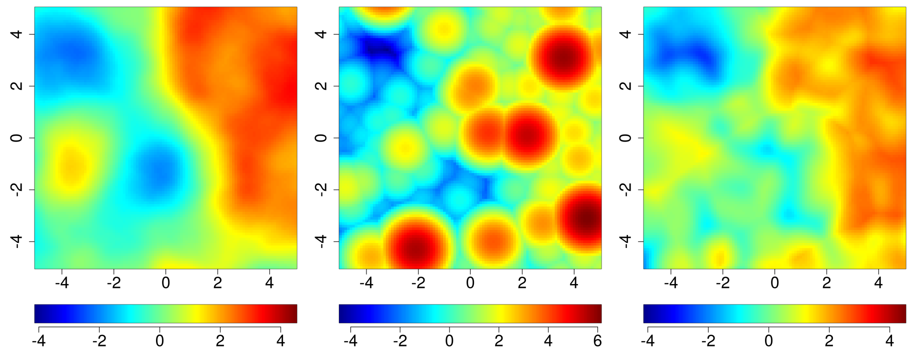

Figure 4.3.: True realization ψ of Ψ (left), associated Cox extremal process Y (centre) and estimated intensity $\hat{\psi}$ (right). The intensity process Ψ is log Gaussian with Matérn covariance function with parameters $\nu = 2, \text{scale} = 3, \text{var} = 2$.

Proof. The assertion follows from the straight forward computation

$$\mathbb{E} \int_D \widehat{\psi_h^y}(s) \, \mathrm{d}s = \mathbb{E} \int_D h^{-d} \sum_{t \in N_K^{\psi,y^*} \cap D} c_D(t)^{-1} k\left(\frac{s-t}{h}\right) \, \mathrm{d}s$$

$$= \mathbb{E} \sum_{t \in N_K^{\psi,y^*} \cap D} 1 = \mathbb{E} N_K^{\psi,y^*}(D).$$

Since the number of points in $N_K^{\psi,y^*}(D)$ is Poisson distributed with parameter

$$\mu_Y^{-1} \int_D \mathbb{E}_X \sup_{t \in K} X(t-s) y(t)^{-1} \psi(s) \, \mathrm{d}s,$$

we conclude

$$\mathbb{E} \int_D \widehat{\psi_h^y}(s) \, \mathrm{d}s = \mu_Y^{-1} \int_D \mathbb{E}_X \left(\sup_{t \in K} X(t-s) y(t)^{-1} \right) \psi(s) \mathrm{d}s = \int_D \psi_K^y(s) \, \mathrm{d}s.$$

$\square$

Finally, we divide $\widehat{\psi}_D^y$ by the correcting factor b_K^y and use

$$\widehat{\psi}_D(s) = b_K^y(s)^{-1} \widehat{\psi}_D^y(s) \tag{4.5}$$

to estimate ψ - see Figure 4.3 for an illustration.

Remark 23. *The integral of the estimator $\widehat{\psi}_{K_R}^y$ is unbiased for the integral of ψ_K^y. That said, ψ_K^y and $\widehat{\psi}_{K_R}^y$ are rather small near the boundary ∂K_R of K_R. Condition (3.13) implies that $b_K^y(s)$ is also small for s close to ∂K_R. Since $\widehat{\psi}_{K_R}$ is defined as $\widehat{\psi}_{K_R} = \widehat{\psi}_{K_R}^y / b_K^y$, the estimates are highly unstable in these regions. The severeness of this effect depends mainly on the shape function X and can a priori be avoided by restricting $\widehat{\psi}_D$ to $D = K$ or using a smaller radius $\tilde{R} < R$ instead of the exact R, such that $\mathbb{E} \inf_{s \in B_{\tilde{R}}(o)} X(s) > \alpha$ for some sufficiently large $\alpha > 0$.*

4.2. Parametric estimation of the covariance function of the intensity process Ψ

As described in Section 3.1, we model the intensity process Ψ that underlies our Cox extremal process Y by a log Gaussian process $\Psi = \exp(W)$. Let $\sigma^2 C_\beta$ be the covariance function of the Gaussian random field W with correlation function C_β and unknown parameters $\sigma^2 > 0$ and $\beta \in \mathbb{R}^p$. To estimate σ^2 and β from a realization y of Y that is observed on $\mathcal{D}$, the following steps are carried out.

1. **Estimate a sample of $CP(\Psi)$.**
    a) Obtain the visible storm centres from y, see Section 4.1.
    b) Modify the sample of storm centres such that its theoretical intensity equals ψ. That is, points are deleted at areas where the original intensity is too high and additional points are simulated at areas where the original intensity is too low.
2. **Estimate σ^2 and β by applying the minimum contrast method (Møller et al., 1998) to the (estimated) sample of $CP(\Psi)$.**
    a) Estimate the pair correlation function of the modified sample of storm centres by kernel methods.
    b) Minimize the distance between the theoretical pair correlation function and its estimate to obtain estimates $\widehat{\sigma^2}$ and $\widehat{\beta}$ for σ^2 and β.

In case of n observations $y_1, \dots, y_n$ we define $\widehat{\sigma_i^2}$ and $\widehat{\beta}_i$ for each $i = 1, \dots, n$ as described above. Then, the final estimates of σ^2 and β are $\widehat{\sigma^2} = n^{-1} \sum_{i=1}^n \widehat{\sigma_i^2}$ and $\widehat{\beta} = n^{-1} \sum_{i=1}^n \widehat{\beta}_i$, respectively. In the sequel we provide detailed descriptions of step 1 and 2 from above. As in Section 4.1, let $K = \mathcal{D} \ominus B_R(o)$ where $\mathcal{D}$ is such that K is compact. We assume again that the visible storm centres can be recovered from a realization y of Y and are regarded as sample of the process N_K^y.

Estimate a sample of $CP(\Psi)$ by modifying N_K^y. As a consequence of Proposition 20, the point process N_K^y is a Cox process with intensity function $b_K^y \Psi$. Compared to the original point process $N_0 \sim CP(\Psi)$ on which the Cox extremal process Y is based, it is very likely that N_K^y possesses more points in the region $\{b_K^y \geq 1\}$ and fewer points in the region $\{b_K^y < 1\}$.
To adjust for this discrepancy we delete some points of N_K^y when $b_K^y > 1$ and add points to N_K^y when $b_K^y < 1$. The first adjustment on $\{b_K^y \geq 1\}$ is done by independent thinning. If p is a measurable function on $\mathbb{R}^d$ with $p(s) \in [0, 1]$, then $p \cdot N_K^y$ denotes the point process where every point of N_K^y is independently deleted with probability $1 - p(\cdot)$ (see (Daley and Vere-Jones, 2008) Chapter 11.3 for details). Figure 4.4 depicts a plot of a sample of the original N_K^y, the thinning probabilities and the thinned sample N_K^y. We

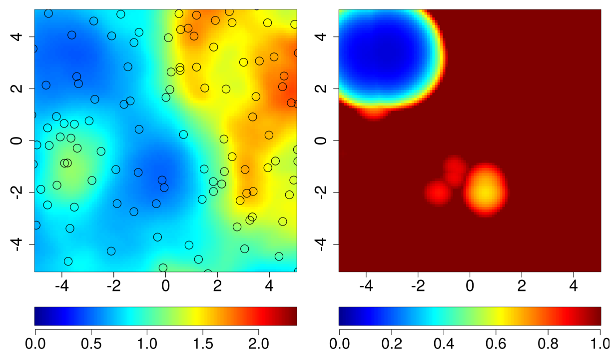

Figure 4.4.: Realization ψ of Ψ and the original sample of N_K^y (left). The retaining probabilities p are plotted in the middle. Thinned sample of N_K^y (circles), the deleted points are marked with crosses (right).

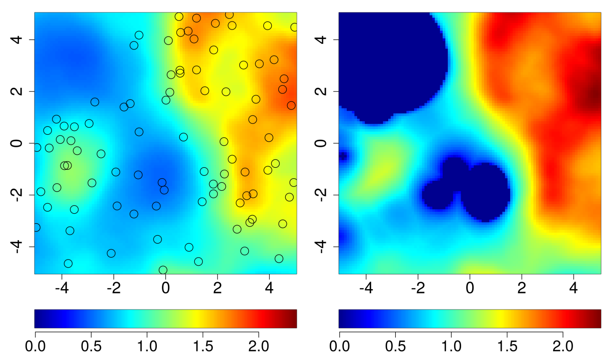

Figure 4.5.: Realization ψ of Ψ and the thinned sample $p \cdot N_K^y$ (left). Additional points are simulated with intensity function $(1 - b_K^y)_+ \Psi$ (middle). Superposition of $p \cdot N_K^y$ with the additional points (filled circles) is plotted in the right.

choose $p = 1/b_K^y$ on $\{b_K^y \geq 1\}$ such that the random intensity function of the thinned process equals Ψ on $\{b_K^y \geq 1\}$. The second adjustment, adding points where $b_K^y < 1$, is achieved by simulating additional points in such way that the sum of the intensity functions equals Ψ on $\{b_K^Y < 1\}$, see also Figure 4.5. The following lemma summarizes and justifies this procedure.

Proposition 24. *Let $CP(f\Psi)$ be a finite Cox process on $\mathbb{R}^d$ and $p = f^{-1} \cdot \mathbb{1}_{\{f\geq 1\}} + \mathbb{1}_{\{f<1\}}$. Then, p is a measurable function on $\mathbb{R}^d$ with $p(s) \in [0,1]$ for all $s \in \mathbb{R}^d$ and*

$$\begin{aligned} &p \cdot CP(f\Psi) + CP((1-f)_+\Psi) \\ &= \underbrace{p \cdot CP(f\Psi)|_{\{f\geq 1\}}}_{\substack{\textit{p-thinning of original } CP(f\Psi) \\ \textit{on } \{f\geq 1\}}} + \underbrace{CP(f\Psi)|_{\{f<1\}}}_{\substack{\textit{original } CP(f\Psi) \\ \textit{on } \{f<1\}}} + \underbrace{CP((1-f)\Psi)|_{\{f<1\}}}_{\substack{\textit{additional points} \\ \textit{on } \{f<1\}}} \sim CP(\Psi). \end{aligned} \tag{4.6}$$

That is, the left-hand side is distributed like a Cox process with intensity process Ψ.

Proof. A simple calculation shows that $p \cdot CP(f\Psi) = p \cdot CP(f\Psi)|_{\{f\geq 1\}} + CP(f\Psi)|_{\{f<1\}}$ and
$(1-f)_+\Psi = (1-f)\Psi|_{\{f<1\}}$ which implies

$$\begin{aligned} &p \cdot CP(f\Psi) + CP((1-f)_+\Psi) \\ &= p \cdot CP(f\Psi)|_{\{f\geq 1\}} + CP(f\Psi)|_{\{f<1\}} + CP((1-f)\Psi)|_{\{f<1\}}. \end{aligned}$$

Since $p = f^{-1} \cdot \mathbb{1}_{\{f\geq 1\}} + \mathbb{1}_{\{f<1\}}$ we obtain

$$p \cdot CP(f\Psi)|_{\{f\geq 1\}} = f^{-1} \cdot CP(f\Psi)|_{\{f\geq 1\}} = CP(\Psi)\mathbb{1}_{\{f\geq 1\}}$$

for the first part of the sum on the set $\{f \geq 1\}$. Furthermore, the remaining parts satisfy $CP(f\Psi)|_{\{f<1\}} + CP((1-f)\Psi)|_{\{f<1\}} = CP(\Psi)|_{\{f<1\}}$ on the set $\{f < 1\}$ which entails the assertion (4.6). □

In our situation we apply Proposition 24 to N_K^y by choosing $f = b_K^y$ and restricting the resulting process to K. That is,

$$(p \cdot N_K^y + CP((1-b_K^y)_+\Psi))|_K \sim CP(\Psi)|_K =: \Phi_K.$$

The first two components considered in (4.6) form the thinned point process $p \cdot N_K^y$. To add the additional points on $\{b_K^y < 1\}$ we rely on our estimate of ψ from Section 4.1.

Minimum contrast method. The so-called pair correlation function (Stoyan and Stoyan, 1992) of a Cox process on $K \subset \mathbb{R}^d$ with random intensity function $\Psi = \exp(W)$ is given by

$$g(s_1, s_2) = \frac{\mathbb{E}[\Psi(s_1)\Psi(s_2)]}{\mathbb{E}\Psi(s_1)\mathbb{E}\Psi(s_2)}, \quad s_1, s_2 \in K.$$

A remarkable property of a log Gaussian Cox process is that its distribution is fully characterized by its first and second order product density. We refer to Theorem 1 in Møller et al. (1998), which also covers the following lemma.

Lemma 25 (Stationarity and second order properties). *A log Gaussian Cox process is stationary if and only if the corresponding Gaussian random field is stationary. Then, its pair correlation function equals*

$$g(s_1 - s_2) = \exp(\sigma^2 C(s_1 - s_2)),$$

where $\sigma^2 C(\cdot)$ is the covariance function of the associated Gaussian random field.

Hence, a log Gaussian Cox process enables a one-to-one mapping between its pair correlation function and the covariance function of the associated Gaussian random field. Therefore, the spatial random effects influencing the random intensity function of the Cox process can be studied by properties of the Cox process itself. The minimum contrast method (Diggle and Gratton, 1984; Møller et al., 1998) exploits this fact.

Proposition 26 (Minimum contrast method, (Møller et al., 1998)). *Suppose that $T_{\sigma^2,\beta}(h) = \sigma^2 C_\beta(h)$ is the covariance function of a Gaussian random field W. Let g be the pair correlation function of the log Gaussian Cox process associated to W. If $\widehat{g}$ is an estimator for g and $\widehat{T}(h) = \log \widehat{g}(h)$, then the distance*

$$d(T_{\sigma^2,\beta}, \hat{T}) = \int_\varepsilon^{r_0} \left(T_{\sigma^2,\beta}(r)^\alpha - \hat{T}(r)^\alpha \right)^2 \mathrm{d}r, \tag{4.7}$$

with tuning parameters $0 \le \varepsilon < r_0$ and $\alpha > 0$, is minimized by the minimum contrast estimators

$$\hat{\beta} = \arg\max_\beta \frac{A(\beta)^2}{B(\beta)}, \quad \hat{\sigma}^2 = \left(\frac{A(\hat{\beta})}{B(\hat{\beta})} \right)^{1/\alpha}, \tag{4.8}$$

with

$$A(\beta) = \int_\varepsilon^{r_0} \left[\log\left(\hat{g}(r)\right) C_\beta(r) \right]^\alpha \mathrm{d}r, \quad B(\beta) = \int_\varepsilon^{r_0} C_\beta(r)^{2\alpha} \mathrm{d}r.$$

The minimum contrast method minimizes the distance of the pair correlation function g and its estimator $\widehat{g}$. Thus, the task of estimating the covariance parameters of W is transformed to a non-parametric estimation of g.

Combined procedure for estimation of β and σ^2. Proposition 24 justifies to approximate a realization of Φ_K by a realization of

$$\widehat{\Phi}_K = p \cdot N_K^y + PP((1 - b_K^y)_+ \widehat{\psi}_{K_R})|_K.$$

We interpret the observed n_K^y as realization of N_K^y and simulate additional points from the point process $PP((1 - b_K^y)_+ \widehat{\psi}_{K_R})|_K$ where $\widehat{\psi}_{K_R}$ is the estimator described in Section 4.1.

Next, we estimate the pair correlation function g of Φ_K by a non-parametric kernel estimator based on the realization $\widehat{\phi}_K$ of $\widehat{\Phi}_K$. Finally, the minimum contrast method can be applied to $\widehat{g}$ to obtain estimates of the parameters σ^2 and β of the log Gaussian Cox process.

Remark 27. *Besides using $\widehat{\phi}_K$ to estimate g, it is also possible to simulate $\widetilde{\phi}_K \sim PP(\widehat{\psi}_{K_R}(s)\,\mathrm{d}s)$ on the whole set K and build estimators for g from samples of $\widetilde{\phi}_K$. However, this leads to a higher bias since the intensity of $\widehat{\phi}_K$ is exactly equal to ψ on $\{b_K^y \geq 1\}$ if N_K^y is known. Additionally, computing the thinning of N_K^y on $\{b_K^y \geq 1\}$ is much faster than simulating $PP(\widehat{\psi}_{K_R}(s)\,\mathrm{d}s)$ on $\{b_K^y \geq 1\}$.*

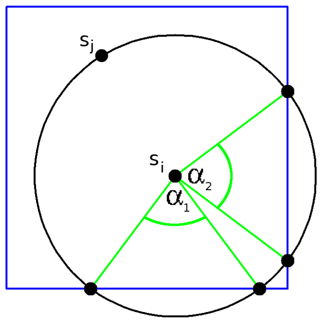

Figure 4.6.: The edge correction b_{ij} is the ratio of the whole circumference 2π and the circle arcs $\gamma_{ij} = 2\pi - \alpha_1 - \alpha_2$ within the square K.

Practical aspects of implementation. We propose to use a non-parametric kernel estimator as discussed by Stoyan and Stoyan (1992) (Part III, Chapter 5.4.2). Consider $\widehat{\phi}_K = \sum_{i=1}^n \delta_{s_i}$, then we estimate the pair correlation function g by

$$\widehat{g}(r) = \frac{|K|}{2\pi n^2 r} \sum_{\substack{i,j=1 \\ i \neq j}}^{n} k_h(r - \|s_i - s_j\|) b_{ij},$$

with the Epanechnikov kernel $k_h(r) = 0.75h^{-1}(1 - r^2/h^2)\mathbb{1}_{|r|<h}$, and kernel weights $b_{ij} \geq 0$ for edge correction (see Ripley (1977)). These are defined as $b_{ij} = 2\pi/\gamma_{ij}$ which is the ratio of the whole circumference of $B_{\|s_i - s_j\|}(s)$ to the circumference within K, i.e γ_{ij} is the sum of all angles, for which the associated non-overlapping circle arcs are within K, see Figure 4.6.

Remark 28. *The estimates of σ^2 and β obtained from $\widehat{g}$ by the minimum contrast method, have a very high variance. Therefore, this procedure is only recommended if we observe several i.i.d. realizations $y_1, \ldots, y_n$ of Y and the associated $N_K^{y_1}, \ldots, N_K^{y_n}$. The final estimates of σ^2 and β may then be defined as the mean or median of the estimates obtained from $\widehat{g}_1, \ldots, \widehat{g}_n$.*

Plots of the estimated pair correlation functions via a sample of $N_K^Y|_{Y=y}$ and by points of $\widehat{\Phi}_K$ are compared in Figure 4.7. They are also compared to the true pair correlation function and the natural benchmark which is obtained from a direct sample of Φ_K instead of $\widehat{\Phi}_K$. Numerical experiments such as reported in Figure 4.7 and Section 4.3 support that our proposed modification works surprisingly well.

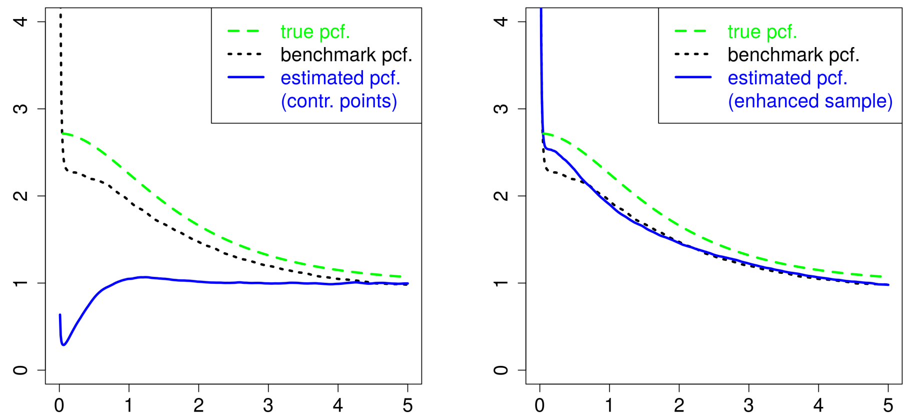

Figure 4.7.: Estimated pair correlation functions via a sample of $N_K^Y|_{Y=y}$ (left) and by points of the modified process $\widehat{\Phi}_K$ (right). They are pointwise averages of $n = 50$ experiments.

4.3. Simulation study

We survey the performance of our proposed non-parametric estimator $\widehat{\psi}_D$ (4.5) of the realization ψ of Ψ and that of the estimators $\widehat{\beta}$ and $\widehat{\sigma^2}$ (4.8) of the parameters of the covariance function of $\Psi = \exp(W)$ in a simulation study.

Setting. In our numerical experiments we choose the covariance of the underlying Gaussian random field W to be the Whittle-Matérn model (3.12) with known smoothness parameter $\nu \in \{0.5, 1, 2, \infty\}$ and unknown variance σ^2 and scale β. The scale β will control the size of clusters in our point processes and the variance σ^2 directs the variability of the number of points within the local clusters. The performances of the associated estimators are compared for different choices $\sigma^2 \in \{1, 2\}$ and $\beta \in \{1, 2\}$. As shape mechanism we consider the fixed storm process $X = \varphi$ where φ is the density of the standard normal distribution. We simulate $n = 1000$ realizations $y_1, \ldots, y_n$ of the corresponding Cox extremal process Y on an equidistant grid with 101^2 grid points in $[-5, 5]^2$.
Henceforth, we simplify the notation from the previous section by writing $\widehat{\psi}$ instead of $\widehat{\psi}_D$ for the estimated intensity function. A natural benchmark of our estimation procedures from Sections 4.1 and 4.2 are such estimators which are derived from direct samples of a Cox process $N_0 \sim CP(\Psi)$ with spatial intensity process Ψ. We denote the benchmark kernel estimator by

$$\widehat{\psi_0}(s) = h^{-d} \sum_{t \in N_0 \cap D} c_{D,h}(t)^{-1} k\left(\frac{s-t}{h}\right), \quad s \in D. \tag{4.9}$$

Accordingly, let $\widehat{\sigma_0^2}$ and $\widehat{\beta}_0$ be the minimum contrast estimators obtained from direct samples of N_0.

Measures for evaluation. To assess the performance of our non-parametric estimates we use the following *mean relative variance*

$$\widehat{\mathrm{MRV}}(\widehat{\psi}, \psi) := n^{-1}|D|^{-1} \sum_{i=1}^{n} \sum_{s \in D} (c_i \widehat{\psi}_i(s)/\psi_i(s) - 1)^2, \tag{4.10}$$

with $c_i^{-1} = |D|^{-1} \sum_{s \in D} \frac{\widehat{\psi}_i(s)}{\psi_i(s)}$ for $i = 1, \ldots, n$. Up to a scaling constant, $\widehat{\mathrm{MRV}}(\widehat{\psi}, \psi)$ is an empirical version of $\mathrm{MISE}(c \cdot \widehat{\psi}/\psi, 1)$ with $\mathrm{MISE}(\widehat{\phi}, \phi) := \mathbb{E} \int (\widehat{\phi}(s) - \phi(s))^2 \, \mathrm{d}s$. We compare the MRV of our estimated intensity $\hat{\psi}$ with that of the benchmark $\widehat{\psi}_0$. The corresponding relative MRV of $\widehat{\psi}_0$ and $\widehat{\psi}$ is defined as the ratio $\widehat{\mathrm{MRV}}(\widehat{\psi})/\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$.
The goodness of fit of the parametric estimates is measured in terms of the empirical mean squared error $\widehat{\mathrm{MSE}}(\widehat{\theta}) := n^{-1} \sum_{i=1}^{n} (\widehat{\theta}_i - \theta)^2$. Again the MSE of $\widehat{\sigma^2}$ and $\widehat{\beta}$ are compared with those of the benchmark estimators $\widehat{\sigma_0^2}$ and $\widehat{\beta}_0$, respectively.

Results. The results of our simulation study are reported in the tables of Figures 4.8 and 4.9. The best performance we can hope for is to be as good as the benchmark estimators that are applied to samples of the original point process $N_0 \sim CP(\Psi)$. Hence, in the case of our non-parametric estimation of the realisations of the intensity processes we can expect the ratios $\widehat{\mathrm{MRV}}(\widehat{\psi})/\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$ to be always greater or equal to 1 and at best even close to 1. Indeed, this is confirmed by the simulation study as can be seen from Figure 4.8. All ratios (except one) lie slightly above 1. The exceptional case occurs when the standard error of $\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$ is relatively high, where we even outperform the benchmark. This is quite remarkable given that we infer the intensity under an independence assumption that is not necessarily satisfied (cf. Section 5 for a discussion) and secondly, we correct it by a data driven quotient as in (4.3).
Likewise we observe that the standard errors for the MRV are close to the benchmark when $\beta = 1$ and much smaller – sometimes even half the size – in the case $\beta = 2$. This indicates that our estimation procedure for ψ is relatively stable compared to the benchmark. In general, both estimators perform better for the larger value of the scale parameter β, that is for larger cluster sizes in the point processes, whereas a higher variance σ^2 naturally leads to a worse performance. The influence of the smoothness parameter is not entirely clear. Looking at the values $\nu \in \{0.5, 1, 2\}$ one might conclude that the estimation improves for a smoother intensity. But in case of the smoothest field ($\nu = \infty$) the MRVs get larger again. However, what is more important is that both procedures, the one that we proposed for inference on ψ for Cox extremal processes and the benchmark $\widehat{\psi}_0$, behave coherently as the parameters vary across different smoothness classes, cluster sizes and variability of number of points within local clusters.
The non-parametric estimates are further used to obtain the parametric estimates of σ^2 and β. Here, estimation of the pair correlation function is very sensitive to the choice of the scale. Our maximal scale $\beta = 2$ is large in relation to the size of the observation window $[-5, 5]^2$ which causes a bias in the estimation of all pair correlation functions. Therefore, all parametric estimates – the benchmarks $\widehat{\sigma^2}_0, \widehat{\beta}_0$ as well as our estimates $\widehat{\sigma^2}, \widehat{\beta}$ – are also biased when $\beta = 2$. The estimation of σ^2 is volatile if $\sigma^2 = 2$, this applies

in particular to our $\widehat{\sigma^2}$ which fails when both $\beta = 1$ and $\sigma^2 = 2$. Still, in all other cases the MSE of our multi-stage estimators is close to that of the benchmark. There are even some cases when we outperform the benchmark, which is not surprising as the standard errors are very high in general.

	$\beta = 1$		$\beta = 2$	
$\nu = 0.5$	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 1$	$\sigma^2 = 2$
$\widehat{\mathrm{MRV}}(\widehat{\psi})$	0.7531 (0.1071)	1.1069 (0.1407)	0.5838 (0.0964)	0.7968 (0.1285)
$\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$	0.6040 (0.0893)	0.9197 (0.1099)	0.5212 (0.148)	0.713 (0.1967)
ratio	1.2469	1.2036	1.1201	1.1175

	$\beta = 1$		$\beta = 2$	
$\nu = 1$	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 1$	$\sigma^2 = 2$
$\widehat{\mathrm{MRV}}(\widehat{\psi})$	0.6533 (0.1211)	0.9246 (0.1628)	0.4852 (0.0948)	0.6057 (0.1291)
$\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$	0.5227 (0.1026)	0.7629 (0.1379)	0.4538 (0.1629)	0.5518 (0.2277)
ratio	1.2499	1.2120	1.0692	1.0977

	$\beta = 1$		$\beta = 2$	
$\nu = 2$	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 1$	$\sigma^2 = 2$
$\widehat{\mathrm{MRV}}(\widehat{\psi})$	0.5986 (0.1345)	0.8254 (0.1771)	0.4452 (0.1038)	0.5062 (0.1241)
$\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$	0.4649 (0.101)	0.6431 (0.1532)	0.4399 (0.1809)	0.5145 (0.2614)
ratio	1.2876	1.2835	1.0125	0.9839

	$\beta = 1$		$\beta = 2$	
$\nu = \infty$	$\sigma^2 = 1$	$\sigma^2 = 2$	$\sigma^2 = 1$	$\sigma^2 = 2$
$\widehat{\mathrm{MRV}}(\widehat{\psi})$	0.6679 (0.1538)	0.9343 (0.2145)	0.4575 (0.1106)	0.5452 (0.1484)
$\widehat{\mathrm{MRV}}(\widehat{\psi}_0)$	0.4925 (0.0968)	0.7402 (0.1645)	0.4127 (0.142)	0.4906 (0.2174)
ratio	1.3562	1.2623	1.1085	1.1113

Figure 4.8.: Results of the simulation study for the non-parametric estimators. The estimator $\widehat{\psi}$ is compared with its benchmark estimator $\widehat{\psi}_0$. The standard errors are reported in brackets.

$\nu = 0.5$	$\beta = 1$, $\sigma^2 = 1$	$\beta = 1$, $\sigma^2 = 2$	$\beta = 2$, $\sigma^2 = 1$	$\beta = 2$, $\sigma^2 = 2$
$\widehat{\text{MSE}}(\widehat{\beta})$	0.0881 (0.1251)	0.2123 (0.0399)	0.8324 (0.5805)	0.6392 (0.2628)
$\widehat{\text{MSE}}(\widehat{\beta}_0)$	0.1074 (0.1357)	0.1300 (0.0321)	1.077 (0.8921)	0.9117 (0.5578)
$\widehat{\text{MSE}}(\widehat{\sigma^2})$	0.0942 (0.1250)	0.8352 (0.1796)	0.0970 (0.0227)	0.7621 (0.2471)
$\widehat{\text{MSE}}(\widehat{\sigma_0^2})$	0.0980 (0.1487)	0.2701 (0.062)	0.1545 (0.1316)	0.4714 (0.1526)

$\nu = 1$	$\beta = 1$, $\sigma^2 = 1$	$\beta = 1$, $\sigma^2 = 2$	$\beta = 2$, $\sigma^2 = 1$	$\beta = 2$, $\sigma^2 = 2$
$\widehat{\text{MSE}}(\widehat{\beta})$	0.1248 (0.016)	0.1106 (0.0229)	0.7407 (0.8757)	0.5305 (0.392)
$\widehat{\text{MSE}}(\widehat{\beta}_0)$	0.1439 (0.0298)	0.0926 (0.0148)	0.8045 (0.6625)	0.6221 (0.2936)
$\widehat{\text{MSE}}(\widehat{\sigma^2})$	0.0954 (0.0151)	0.8925 (0.2141)	0.1095 (0.0212)	0.8513 (0.2881)
$\widehat{\text{MSE}}(\widehat{\sigma_0^2})$	0.1076 (0.0296)	0.3446 (0.0899)	0.1399 (0.0418)	0.6780 (0.2548)

$\nu = 2$	$\beta = 1$, $\sigma^2 = 1$	$\beta = 1$, $\sigma^2 = 2$	$\beta = 2$, $\sigma^2 = 1$	$\beta = 2$, $\sigma^2 = 2$
$\widehat{\text{MSE}}(\widehat{\beta})$	0.0900 (0.0126)	0.0866 (0.0218)	0.6989 (0.9267)	0.5648 (0.3897)
$\widehat{\text{MSE}}(\widehat{\beta}_0)$	0.0993 (0.0165)	0.0681 (0.0057)	0.6214 (0.4832)	0.5346 (0.2335)
$\widehat{\text{MSE}}(\widehat{\sigma^2})$	0.097 (0.0125)	0.9633 (0.2071)	0.1123 (0.0269)	0.8731 (0.293)
$\widehat{\text{MSE}}(\widehat{\sigma_0^2})$	0.0943 (0.0221)	0.4013 (0.1165)	0.1543 (0.0356)	0.8601 (0.376)

$\nu = \infty$	$\beta = 1$, $\sigma^2 = 1$	$\beta = 1$, $\sigma^2 = 2$	$\beta = 2$, $\sigma^2 = 1$	$\beta = 2$, $\sigma^2 = 2$
$\widehat{\text{MSE}}(\widehat{\beta})$	0.1195 (0.0127)	0.1242 (0.0125)	0.4111 (1.4327)	0.2941 (0.5625)
$\widehat{\text{MSE}}(\widehat{\beta}_0)$	0.0663 (0.0132)	0.0412 (0.0064)	0.3355 (0.1851)	0.2676 (0.0767)
$\widehat{\text{MSE}}(\widehat{\sigma^2})$	0.1153 (0.0126)	1.1322 (0.2121)	0.105 (0.016)	0.8436 (0.2605)
$\widehat{\text{MSE}}(\widehat{\sigma_0^2})$	0.0825 (0.0149)	0.3606 (0.0901)	0.1373 (0.0224)	0.6939 (0.2627)

Figure 4.9.: Results of the simulation study for the parametric estimators. The estimators $\widehat{\beta}$ and $\widehat{\sigma^2}$ are compared with their benchmark estimators $\widehat{\beta}_0$ and $\widehat{\sigma_0^2}$. The standard errors are reported in brackets.

5. Discussion

In Part II of this thesis, we present the new class of conditionally max-stable random fields based on Cox processes, which we therefore also call Cox extremal processes. We prove in Theorem 14 that these processes are in the MDA of familiar max-stable models. Hence, they have the potential to model spatial extremes on a smaller time scale.
An objective of practical importance is to identify the random effects influencing the underlying Cox process from the centres of the contributing storms. In order to make inference feasible, we impose an additional independence assumption on our observed data (see Proposition 20) that allows to derive a non-parametric kernel estimator (4.5) for the realization ψ of the intensity process Ψ. Imposing such an independence assumption can be seen in a similar manner to the composite likelihood method that ignores dependence among higher order tupels. We believe that our condition is sufficiently well satisfied in most situations, since only a small number of large storms from the Cox process N approximate the Cox extremal process Y already quite well.
For parametric estimation the non-parametric estimator (4.5) can be used to correct the observed point process $N_K^{y,\psi}$ of contributing storm centres in order to obtain a sample of $CP(\Psi)$ (Proposition 24). If Ψ is log Gaussian, the minimum contrast method can be applied subsequently to obtain estimates for the parameters of the covariance function of $\log \Psi$. The performance of our proposed estimation procedures is addressed in a simulation study (Section 4.3). Here, the best we can hope for is that our estimators can compete with the benchmark estimators applied to the original point process $CP(\Psi)$. Indeed, our non-parametric procedure is usually relatively close to the benchmark which is quite remarkable in view of the necessary adjustments we have to make. Also, looking at different kinds of smoothness, cluster sizes and variances we find evidence for the stability of our proposed estimation procedure when compared to the benchmark. Both (our procedure and the benchmark) behave coherently across different choices of these properties. Similar behaviour can be observed for the parametric estimates, even though they are more volatile and the estimation of the pair correlation function is generally very sensitive to the choice of scale.
Within our simulation study and all other illustrations, we consider deterministic storm processes $X = \varphi$ where φ is the density of the standard normal distribution. This restriction is only done to reduce the computing time. Indeed, all estimators presented in Chapter 4 are valid for much more general X and simulations showed that the specification of X only slightly influences the inference on Ψ as long as enough centres of contributing storms can be identified. For instance, if we impose the monotonicity assumption (3.13) on the storm process X, the majority can be recovered as local maxima of the realization y of Y. Computational methods for identification of such points are left for further research.

Part III.

Generalization of the Matérn hard-core process

6. Generalized Matérn model

Point process models obtained by dependent thinning of homogeneous Poisson point processes have been extensively examined during the last decades. The Matérn hard-core processes (Matérn, 1960) are classical examples for such processes, where the thinning probability of an individual point depends on the other points of the original point pattern. The Matérn models and slight modifications of them are applied to real data in various branches, for instance ecological science (Stoyan, 1987; Picard, 2005), geographical analysis (Stoyan, 1988) and computer science (Ibrahim et al., 2013).
There already exist several extensions of Matérns models (Kuronen and Leskelä, 2013), concerning the hard-core distance (Stoyan and Stoyan, 1985; Månsson and Rudemo, 2002), the thinning rule (Teichmann et al., 2013) or the generalization the underlying Poisson process (Andersen and Hahn, 2016). We present a new model which encompasses all these approaches and further generalizes the underlying point process, the thinning rule and the marks attached to the original process.
This chapter is based on the first part of Dirrler and Schlather (2017). In Section 6.1, we shortly review the Matérn hard-core processes from a different point of view and state more details on Palm calculus which will be used throughout this part of the thesis. Our general model is defined in Section 6.2. We restrict the underlying ground process to a log Gaussian Cox process in Section 6.3 and calculate first and second order properties for this model. In Chapter 7, we establish a connection between our model and mixed moving maxima (M3) processes.

6.1. Matérn hard-core processes and Palm calculus

In Section 2.4 we gave a brief summary of the Matérn hard-core processes I and II. We now present a different kind of representation of these processes.
Let Φ be a homogeneous Poisson process on $S = \mathbb{R}^d$ with intensity λ. Consider the function $f_{\mathrm{MatI}}(\Phi;\xi) = \prod_{\xi' \in \Phi \setminus \{\xi\}} (1 - \mathbb{1}_{\xi' \in B_R(\xi)})$, then

$$\Phi_{\mathrm{MatI}} = \{\xi \in \Phi : f_{\mathrm{MatI}}(\Phi;\xi) = 1\}.$$

We now regard the marked point process Φ_M where each point $\xi \in \Phi$ is independently endowed with a random mark $m_\xi \sim \mathcal{U}[0,1]$. Let

$$f_{\mathrm{MatII}}(\Phi_M;\xi,m_\xi) = \prod_{(\xi',m_{\xi'}) \in \Phi_M \setminus \{(\xi,m_\xi)\}} (1 - \mathbb{1}_{\xi' \in B_R(\xi)} \mathbb{1}_{m_{\xi'} < m_\xi}),$$

then

$$\Phi_{\mathrm{MatII}} = \{(\xi,m_\xi) \in \Phi_M : f_{\mathrm{MatII}}(\Phi_M;\xi,m_\xi) = 1\}.$$

It is often of particular interest to compute the probability that a given point $(\xi, m_\xi) \in \Phi_M$ is retained in the thinned process Φ_{MatII}. This probability can be calculated using Palm calculus (Mecke, 1967; Møller and Waagepetersen, 2004; Daley and Vere-Jones, 2008; Chiu et al., 2013), we summarize the basic results in Section 2.2.
The Palm distribution $P^!_\xi$ can be interpreted as the conditional distribution of $\Phi \setminus \{\xi\}$ given $\xi \in \Phi$. Thereby the retaining probability of a point $(\xi, m_\xi) \in \Phi_M$ equals

$$r(\xi, m_\xi) = \int_{M_{\mathrm{lf}}} f_{\mathrm{MatII}}(\varphi; \xi, m_\xi)\ P^!_{\xi, m_\xi}(\mathrm{d}\varphi)$$

where M_{lf} is the suitably defined space of point configurations of the marked process Φ_M, for details see Definition 4 in Section 2.1. The *generating functional* of a point process Φ is defined as

$$G_\Phi(u) = \mathbb{E} \prod_{\xi \in \Phi} u(\xi) \tag{6.1}$$

for functions $u : S \to [0, 1]$, see (Westcott, 1972). The Palm distribution $P^!_{\xi, m_\xi}$ equals the distribution of Φ_M since Φ_M is a Poisson process - see Example 4.3 in Chiu et al. (2013). As a consequence of this, $r(\xi, m_\xi)$ is the generating functional of Φ_M evaluated at f_{MatII}. Therefore

$$r(\xi, m_\xi) = \exp(-\lambda |B_R(o)| \cdot m_\xi)$$

and the intensity of the thinned process equals

$$\lambda_{\mathrm{MatII}} = \lambda \int_0^1 r(m_\xi)\ \mathrm{d}m_\xi = |B_R(o)|^{-1}(1 - \exp(-\lambda |B_R(o)|)).$$

The Palm distribution of a general point process is more difficult to handle, however it can be explicitly calculated for many Cox process models (Møller, 2003; Coeurjolly et al., 2015). Besides, Mecke (1967) indicates how to calculate the Palm distribution of an infinitely divisible point process. A point process Φ is called infinitely divisible if, for all $n \in \mathbb{N}$ there exist iid. processes $\Phi_1, \ldots, \Phi_n$ such that $\Phi \stackrel{d}{=} \Phi_1 + \cdots + \Phi_n$.

6.2. Generalizing the Matérn hard-core processes

We present a new point process model, obtained by dependent thinning of a ground process Φ, which generalizes the Matérn model in several ways. We therefore call the new model *generalized Matérn model*.
Suppose that Φ is a locally finite point process on S. Each point ξ of Φ is independently attached with a random mark m_ξ. We allow these marks to be continuous functions from S to $\mathbb{R}$, i.e. an element of the space of continuous functions $\mathbb{M} = C(S, \mathbb{R})$ with law ν. Then,

$$\Phi_M = \{(\xi, m_\xi) : \xi \in \Phi\}$$

	p	m_ξ	ζ
Matérn I	1	-	$\mathbb{1}_{\xi' \in B_R(\xi)}$
Generalized Matérn I	$(1 - \|\xi - \xi'\|/R)_+$	-	$\mathbb{1}_{\xi' \in B_R(\xi)}$
Matérn II	1	$\mathcal{U}[0,1]$	$\mathbb{1}_{\xi' \in B_R(\xi)} \mathbb{1}_{m'_\xi < m_\xi}$
Generalized Matérn II	$(1 - \|\xi - \xi'\|/R)_+$	$\mathcal{U}[0,1]$	$\mathbb{1}_{\xi' \in B_R(\xi)} \mathbb{1}_{m'_\xi < m_\xi}$

Table 6.1.: Let $S = \mathbb{R}^d$. The classical Matérn models are obtained as special case of our general model. We call the models resulting from Matérn I or II by including an additional stochastic thinning, *generalized Matérn I* and *II* model respectively.

is a marked point process, i.e. a mapping into $(M_{\text{lf}}, \mathcal{M}_{\text{lf}})$. The Bernoulli random variable $\tau_{\Phi_M;\xi,m_\xi}$ shall indicate whether a point of Φ_M is retained in the thinned process. We define the thinned marked process

$$\Phi_{\text{th}} = \{(\xi, m_\xi) \in \Phi_M : (\xi, m_\xi) \in \Phi_M, \tau_{\Phi_M;\xi,m_\xi} = 1\}, \tag{6.2}$$

which we call *generalized Matérn process*, and the thinned ground process

$$\Phi^0_{\text{th}} = \{\xi : (\xi, m_\xi) \in \Phi_{\text{th}}\}. \tag{6.3}$$

The success probability of $\tau_{\Phi_M;\xi,m_\xi}$ equals the *thinning function*

$$f_{\text{th}}(\Phi_M; \xi, m_\xi) = \prod_{(\xi', m'_\xi) \in \Phi_M} (1 - \zeta(\xi, m_\xi, \xi', m_{\xi'}) p(\xi, m_\xi, \xi', m_{\xi'})).$$

Here, $\zeta : (S, \mathbb{M})^2 \to \{0, 1\}$ is a measurable function which we call *competition function* and which specifies the inferior points which are endangered to be deleted. We call a point ξ *inferior* if $\zeta(\xi, m_\xi, \xi', m'_\xi) = 1$ for some $(\xi', m'_\xi) \in \Phi_M \setminus \{\xi, m_\xi\}$. Likewise, $p : (S, \mathbb{M})^2 \to [0, 1]$ is a measurable function determining the probability that an inferior point is deleted. We henceforth fix

$$\zeta(\xi, m_\xi, \xi, m_\xi) = 1, \quad p(\xi, m_\xi, \xi, m_\xi) = 1 - p_0 \in [0, 1]$$

and thereby include independent p_0-thinning in our model. In order to simplify notation, we will henceforth use abbreviations like $\boldsymbol{\xi} = (\xi, m_\xi)$ and $\zeta(\boldsymbol{\xi}, \boldsymbol{\xi}') = \zeta(\xi, m_\xi, \xi', m_{\xi'})$.

Example 29. *The Matérn hard-core models I and II can be easily derived from our model, see Table 6.1. There, we also give generalizations where $p \neq 1$.*

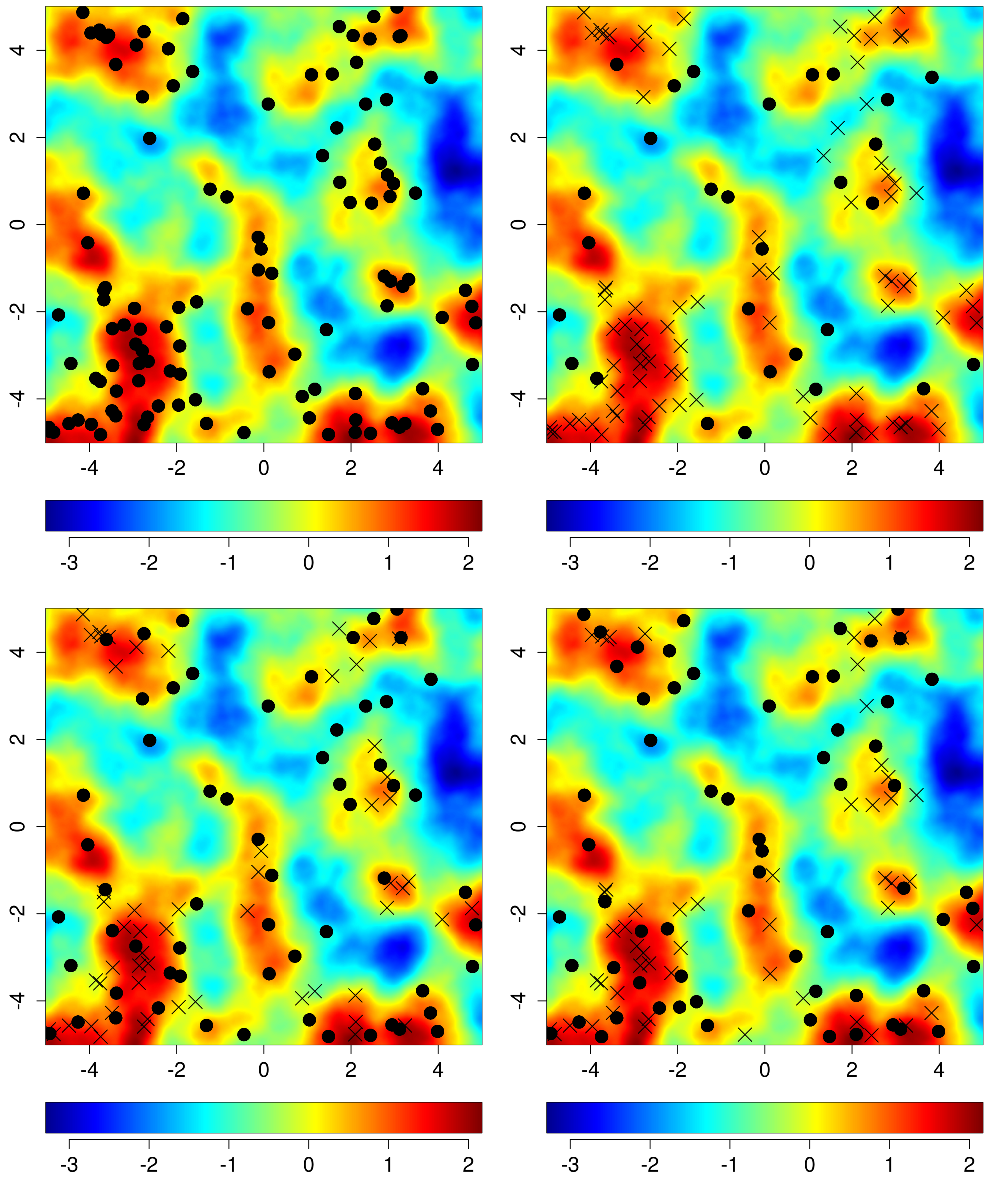

Figure 6.1.: Plot of the original Poisson process Φ and the underlying intensity function (upper left). Thinned points Φ_{th}^0 of a generalized Matérn I model (upper right) and a generalized Matérn II model (lower left) with $R = 1$ and $p(\boldsymbol{\xi}, \boldsymbol{\xi}') = \max(0, 1 - \|\xi - \xi'\|)$. The last plot shows the thinned points of a generalized hard-core process with competition function $\zeta(\boldsymbol{\xi}, \boldsymbol{\xi}') = \mathbb{1}_{m_{\xi'}(\xi - \xi') > m_{\xi}(o)}$, random mark functions $m_\xi(\cdot) = u \cdot \varphi(\cdot)$, $u \sim \mathcal{U}[0, 1]$ with the two-dimensional standard-normal density φ and thinning probability $p(\boldsymbol{\xi}, \boldsymbol{\xi}') = \max(0, 1 - \|\xi - \xi'\|)$.

Example 30. *A further generalization of the Matérn I model in $S = \mathbb{R}^d$ was presented in Teichmann et al. (2013). According to their thinning rule, a point ξ of the ground process Φ is retained with probability*

$$p_0 \prod_{\xi' \in \Phi_M \setminus \{\xi\}} (1 - f(\|\xi - \xi'\|)),$$

with $p_0 \in (0, 1]$ and some deterministic function $f : [0, \infty) \to [0, 1]$. This equals our model with the choice $\zeta \equiv 1$ and $p(\xi, m_\xi, \xi', m'_\xi) = f(\|\xi - \xi'\|)$.

Example 31. *Consider now $S = \mathbb{R}^d$, marks m_ξ in $\mathbb{M} = \mathbb{R}^{\{0,1\}}$ with $m_\xi(0) \sim \mu$ and $m_\xi(1) \sim \nu$ for probability measures μ and ν. Let $\zeta(\xi, m_\xi, \xi', m_{\xi'}) = \mathbb{1}_{m_\xi(0) \geq m_{\xi'}(0)}$ and $p(\xi, m_\xi, \xi', m_{\xi'}) = f(\|\xi - \xi'\|, m_\xi(1), m_{\xi'}(1))$. Then*

$$f_{\text{th}}(\Phi_M; \xi, m_\xi) = p_0 \prod_{\boldsymbol{\xi}' \in \Phi_M \setminus \{\boldsymbol{\xi}\}} \left[1 - \mathbb{1}_{m_\xi(0) \geq m_{\xi'}(0)} f\big(\|\xi - \xi'\|, m_\xi(1), m_{\xi'}(1)\big)\right].$$

This model was presented by Teichmann et al. (2013) as an extension of the Matérn II model.

Example 32. *Let Φ be an inhomogeneous Poisson process in $\mathbb{R}^d$, attached with random mark functions $m_\xi(\cdot) = u \cdot \varphi(\cdot)$, $u \sim \mathcal{U}[0, 1]$ with the d-dimensional standard-normal density φ. Consider the competition function $\zeta(\boldsymbol{\xi}, \boldsymbol{\xi}') = \mathbb{1}_{m_{\xi'}(\xi - \xi') > m_\xi(o)}$ and $p(\boldsymbol{\xi}, \boldsymbol{\xi}') = \max(0, 1 - \|\xi - \xi'\|)$. This leads to a soft-core model where inferior points are the more likely to be thinned the closer they are to superior points. See Figure 6.1 for a plot of this model in $d = 2$ and Figure 6.2 for a plot of arbitrary points (ξ, m_ξ) and (η, m_η) with $d = 1$.*

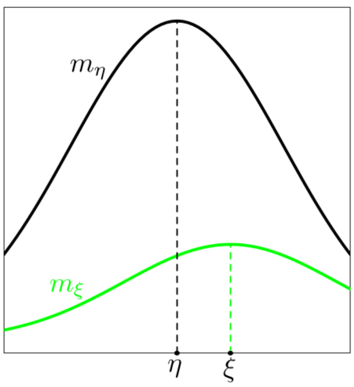

Figure 6.2.: Arbitrary points (ξ, m_ξ) and (η, m_η) of Φ_M from Example 32 with $d = 1$. Since $m_\eta(\xi - \eta) > m_\xi(0)$, the point (ξ, m_ξ) is inferior to (η, m_η) and hence endangered to be thinned with probability $p(\xi, m_\xi, \eta, m_\eta) = \max(0, 1 - |\xi - \eta|)$.

6.3. Generalized Matérn model based on log Gaussian Cox processes

Let $P^!_{\boldsymbol{\xi}}(\cdot)$ be the reduced Palm distribution of Φ_M, that is $P^!_{\boldsymbol{\xi}}(\cdot)$ is a probability measure on $(M_{\text{lf}}, \mathcal{M}_{\text{lf}})$ for each $\boldsymbol{\xi} \in S \times \mathbb{M}$. The retaining probability of a point $\xi \in \Phi$ with mark function m_ξ can then be calculated by

$$r(\xi, m_\xi) = \int_{M_{\text{lf}}} f_{\text{th}}(\varphi \cup (\xi, m_\xi); \xi, m_\xi)\ P^!_{\boldsymbol{\xi}}(\mathrm{d}\varphi).$$

Since our new model is defined in a rather general setting, reasonable restrictions are needed in order to calculate the reduced Palm distribution of Φ_M and thereby first and seconder order properties of Φ_{th}. We henceforth assume that Φ is a log Gaussian Cox process, though all results may be derived in a similar way for other Cox processes or infinitely divisible point processes (Mecke, 1967), when $P^!_{\boldsymbol{\xi}}$ is known.
Let $\Psi = \exp(W)$ be the random intensity function of Φ where W is a Gaussian random field with mean function μ and covariance function C. We write $\Phi \sim \text{LGCP}(\mu, C)$ for short.

Proposition 33. *Let* $\Phi \sim \text{LGCP}(\mu, C)$ *and* $h(\boldsymbol{\xi}, \boldsymbol{\xi}') = 1 - \zeta(\boldsymbol{\xi}, \boldsymbol{\xi}')p(\boldsymbol{\xi}, \boldsymbol{\xi}')$*, then the retaining probability is*

$$r(\xi, m_\xi) = \mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M} h(\boldsymbol{\xi}, \boldsymbol{\xi}'),$$

with $\widetilde{\Phi} \sim \text{LGCP}(\widetilde{\mu}, C)$ *and* $\widetilde{\mu}(\cdot) = \mu(\cdot) + C(\cdot, \xi)$*. Furthermore, if* $\widetilde{\Psi}$ *is the random intensity function of* $\widetilde{\Phi}$*, the first order intensity of* Φ^0_{th} *is given by*

$$\rho_{th}(\xi) = p_0 \rho_\Phi(\xi) \int_{\mathbb{M}} \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_S \int_{\mathbb{M}} \zeta(\boldsymbol{\xi}, \boldsymbol{\xi}')p(\boldsymbol{\xi}, \boldsymbol{\xi}')\widetilde{\Psi}(\xi')\nu(dm_{\xi'})\mathrm{d}\xi' \right) \nu(\mathrm{d}m_\xi), \quad (6.4)$$

where ρ_Φ *is the intensity of* Φ*.*

Proof. The reduced Palm distribution $P^!_\xi$ of Φ equals the distribution of $\widetilde{\Phi} \sim \text{LGCP}(\widetilde{\mu}, C)$ since $\Phi \sim \text{LGCP}(\mu, C)$ - see Proposition 1 in Coeurjolly et al. (2015). Therefore

$$\begin{aligned}
r(\xi, m_\xi) &= \int_{M_{\text{lf}}} f_{\text{th}}(\varphi \cup \{\boldsymbol{\xi}\}; \boldsymbol{\xi})\ P^!_{\boldsymbol{\xi}}(\mathrm{d}\varphi) \\
&= \mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M \cup \{\boldsymbol{\xi}\}} [1 - \zeta(\boldsymbol{\xi}, \boldsymbol{\xi}')p(\boldsymbol{\xi}, \boldsymbol{\xi}')] \\
&= (1 - \zeta(\boldsymbol{\xi}, \boldsymbol{\xi})p(\boldsymbol{\xi}, \boldsymbol{\xi}))\mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M} [1 - \zeta(\boldsymbol{\xi}, \boldsymbol{\xi}')p(\boldsymbol{\xi}, \boldsymbol{\xi}')] \\
&= p_0\ \mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M} h(\boldsymbol{\xi}, \boldsymbol{\xi}') \\
&= p_0\ \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{S \times \mathbb{M}} \left(1 - h(\boldsymbol{\xi}, \boldsymbol{\xi}')\right) \widetilde{\Psi}(\xi')\mathrm{d}\xi' \nu(\mathrm{d}m_{\xi'}) \right),
\end{aligned}$$

where the last equality follows from calculating the generating functional of $\widetilde{\Phi}_M$.

$\square$

Proposition 34. *Consider* $\widetilde{\Phi} \sim \mathrm{LGCP}(\tilde{\mu}, C)$ *with* $\widetilde{\mu}(\cdot) = \mu(\cdot) + C(\cdot, \xi) + C(\cdot, \eta)$ *and let* $\rho_\Phi^{(2)}$ *be the second order intensity of* Φ*. Then, the second order intensity of the thinned process* Φ_{th}^0 *equals*

$$\begin{aligned}\rho_{\mathrm{th}}^{(2)}(\xi,\eta) =& \rho_\Phi^{(2)}(\xi,\eta) p_0^2 \int_{\mathbb{M}} \int_{\mathbb{M}} \Bigg[h(\boldsymbol{\xi},\boldsymbol{\eta}) h(\boldsymbol{\eta},\boldsymbol{\xi}) \\ & \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{S\times\mathbb{M}} \left(1 - h(\boldsymbol{\xi},\boldsymbol{\xi}') h(\boldsymbol{\eta},\boldsymbol{\xi}')\right) \widetilde{\Psi}(\xi') \mathrm{d}\xi' \nu(\mathrm{d}m'_\xi) \right) \Bigg] \nu(\mathrm{d}m_\xi) \nu(\mathrm{d}m_\eta).\end{aligned}$$

Proof. The probability that none of the two arbitrary points (ξ, m_ξ) and (η, m_η) is deleted by any point of the point configuration $\varphi \in M_{\mathrm{lf}}$ is

$$f_{\mathrm{th}}^{(2)}(\varphi; \boldsymbol{\xi}, \boldsymbol{\eta}) = \prod_{\boldsymbol{\xi}' \in \varphi} (1 - \zeta(\boldsymbol{\xi},\boldsymbol{\xi}') p(\boldsymbol{\xi},\boldsymbol{\xi}'))(1 - \zeta(\boldsymbol{\eta},\boldsymbol{\xi}') p(\boldsymbol{\eta},\boldsymbol{\xi}')).$$

Thus, the probability that $(\xi, m_\xi), (\eta, m_\eta) \in \Phi_M$ are retained in Φ_{th} equals

$$r(\boldsymbol{\xi},\boldsymbol{\eta}) = \int_{M_{\mathrm{lf}}} f_{\mathrm{th}}^{(2)}(\varphi \cup \{\boldsymbol{\xi},\boldsymbol{\eta}\}; \boldsymbol{\xi},\boldsymbol{\eta}) P^!_{\boldsymbol{\xi},\boldsymbol{\eta}}(\mathrm{d}\varphi),$$

where $P^!_{\boldsymbol{\xi},\boldsymbol{\eta}}$ is the two-point reduced Palm distribution of Φ_M, which is also the distribution of a log Gaussian Cox process $\widetilde{\Phi} \sim \mathrm{LGCP}(\tilde{\mu}, C)$ with $\widetilde{\mu}(\cdot) = \mu(\cdot) + C(\cdot, \xi) + C(\cdot, \eta)$ - see again Proposition 1 in Coeurjolly et al. (2015). Therefore

$$\begin{aligned}&\int_{M_{\mathrm{lf}}} f_{\mathrm{th}}^{(2)}(\varphi \cup \{\boldsymbol{\xi},\boldsymbol{\eta}\}; \boldsymbol{\xi},\boldsymbol{\eta}) P^!_{\boldsymbol{\xi},\boldsymbol{\eta}}(\mathrm{d}\varphi) \\ &= \mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M \cup \{\boldsymbol{\xi},\boldsymbol{\eta}\}} \Big(1 - \zeta(\boldsymbol{\xi},\boldsymbol{\xi}') p(\boldsymbol{\xi},\boldsymbol{\xi}')\Big)\Big(1 - \zeta(\boldsymbol{\eta},\boldsymbol{\xi}') p(\boldsymbol{\eta},\boldsymbol{\xi}')\Big) \\ &= p_0^2 h(\boldsymbol{\xi},\boldsymbol{\eta}) h(\boldsymbol{\eta},\boldsymbol{\xi}) \; \mathbb{E} \prod_{\boldsymbol{\xi}' \in \widetilde{\Phi}_M} h(\boldsymbol{\xi},\boldsymbol{\xi}') h(\boldsymbol{\eta},\boldsymbol{\xi}') \\ &= p_0^2 h(\boldsymbol{\xi},\boldsymbol{\eta}) h(\boldsymbol{\eta},\boldsymbol{\xi}) \; \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{S\times\mathbb{M}} \Big(1 - h(\boldsymbol{\xi},\boldsymbol{\xi}') h(\boldsymbol{\eta},\boldsymbol{\xi}')\Big) \widetilde{\Psi}(\xi') \mathrm{d}\xi' \nu(\mathrm{d}m_\xi) \right).\end{aligned}$$

$\square$

7. Application to mixed moving maxima processes

This chapter is based on the second part of Dirrler and Schlather (2017). We establish a connection between the generalized Matérn model and mixed moving maxima processes in this section. In the first section, we choose a specific thinning function and prove that a process based on the corresponding generalized Matérn process (6.2) converges to known (conditional) mixed moving maxima processes. We slightly modify this thinning function in the second section to obtain a process whose first and second order properties can be derived and whose points can be recovered from observations of the mixed moving maxima process itself under rather mild assumptions.

General framework. Let $S = \mathbb{R}^d$, $K \subset S$ be compact and let X be a stochastic process whose paths are almost surely in $\mathbb{X} = C(S, \mathbb{R})$ and which fulfils the condition

$$\mathbb{E}_X \int_S \sup_{t \in K} X(t - \xi) \, \mathrm{d}\xi < \infty. \tag{7.1}$$

Then a mixed moving maxima process (Smith, 1990) is defined by

$$Z(t) = \bigvee_{(s,u,X) \in \Theta} uX(t - s), \qquad t \in S, \tag{7.2}$$

where Θ is a Poisson process on $S \times (0, \infty] \times \mathbb{X}$ with directing measure

$$\mathrm{d}\lambda(s, u, X) = \mathrm{d}s \, u^{-2} \mathrm{d}u \, \mathrm{d}\mathbb{P}_X.$$

Further, we assume that Ψ is a non-negative process with

$$\mathbb{E}_\Psi \mathbb{E}_X \int_S \sup_{t \in K} X(t - \xi) \Psi(\xi) \mathrm{d}\xi < \infty. \tag{7.3}$$

Then the Cox extremal process (3.1) is defined in an analogous way

$$Y(t) = \bigvee_{(s,u,X) \in \widetilde{\Theta}} uX(t - s), \tag{7.4}$$

where $\widetilde{\Theta}$ is a Cox process directed by the random measure

$$\mathrm{d}\Lambda(s, u, X) = \Psi(s) \mathrm{d}s u^{-2} \mathrm{d}u \mathrm{d}\mathbb{P}_X.$$

We henceforth consider $\Phi \sim \mathrm{LGCP}(\mu - \log(\tau), C)$ with random intensity function $\tau^{-1}\Psi$ and deterministic $\tau > 0$. Each point ξ of Φ is independently attached with a random mark function $m_\xi(\cdot) = U_\xi X_\xi(\cdot)$, where $U \sim \tau u^{-2}\mathbb{1}_{(\tau,\infty)}\mathrm{d}u$ and $X \sim \mathrm{d}\mathbb{P}_X$. We assume that for each path $X(\omega, \cdot)$ of $X(\cdot)$ there exist monotonously decreasing functions f_ω and g_ω such that

$$g_\omega(\|t\|) \leq X(\omega, t) \leq f_\omega(\|t\|), \quad \forall t \in S, \tag{7.5}$$

and $g_\omega(0) = X(\omega, 0) = f_\omega(0)$.

7.1. Matérn extremal process

We choose the competition function ζ such that a point $\xi \in \Phi$ is deleted if its corresponding mark function m_ξ is - at each point - strictly smaller than the mark function $m_{\xi'}$ of some other point $\xi' \in \Phi$. That is, the thinning function can be written as

$$f_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = \prod_{\xi' \in \Phi_M} \left[1 - \mathbb{1}_{u_{\xi'} > \sup_{t \in S} u_\xi X_\xi(t-\xi) X_{\xi'}(t-\xi')^{-1}}\right] \tag{7.6}$$

and the process resulting from dependent thinning is

$$\Phi_{\mathrm{th}} = \{(\xi, m_\xi) \in \Phi_M : f_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = 1\}.$$

We introduce the *Matérn extremal process* defined by

$$\Pi(t) = \bigvee_{(\xi, m_\xi) \in \Phi_{\mathrm{th}}} m_\xi(t - \xi). \tag{7.7}$$

Let $\mathrm{ex}(\Phi_M)$ be the set of extremal functions of Φ_M as introduced by Dombry and Eyi-Minko (2013), that is

$$\mathrm{ex}(\Phi_M) = \{(\xi, m_\xi) \in \Phi_M : \exists t \in S, m_\xi(t - \xi) \geq \Pi(t)\}. \tag{7.8}$$

The (ξ, m_ξ) in Φ_{th} are closely related to the extremal functions of Φ_M, though the set Φ_{th} is usually much larger than the set of extremal functions. The intensity of Φ_{th} is finite, this is a fundamental difference compared to the Cox extremal process Y whose underlying point process $\widetilde{\Theta}$ is infinite. Still, the following lemma shows that the two processes coincide in the limiting case $\tau \to 0$.

Lemma 35. *Let the conditions* (7.1), (7.3) *and* (7.5) *hold true and assume that* $\Phi \sim \mathrm{LGCP}(\mu - \log\tau, C)$. *If* $\tau \to 0$, *the convergence*

$$\Pi \to Y$$

holds weakly in $C(S, \mathbb{R})$.

Proof. The sample paths of the extremal hard-core process Π are continuous, since the marks m_ξ are continuous and Φ is locally finite. The finite-dimensional distributions of Π are given by

$$\begin{aligned}
&\mathbb{P}(\Pi(t_1) \le y_1, \dots, \Pi(t_n) \le y_n) \\
&= \mathbb{P}(U_\xi X_\xi(t_1 - \xi) \le y_1, \dots, U_\xi X_\xi(t_n - \xi) \le y_n, \forall (\xi, m_\xi) \in \Phi_{\text{th}}) \\
&= \mathbb{P}\left(U_\xi \le \min_{1 \le i \le n} (y_i X_\xi(t_i - \xi)^{-1}), \forall (\xi, m_\xi) \in \Phi_{\text{th}} \right) \\
&= \mathbb{P}\left(U_\xi \le \min_{1 \le i \le n} (y_i X_\xi(t_i - \xi)^{-1}), \forall (\xi, m_\xi) \in \Phi_M \right) \\
&= \mathbb{E}_\Psi \exp\left[-\int_{\mathbb{X}} \int_S \int_{\min_{1 \le i \le n}(y_i X_\xi(t_i - \xi)^{-1})} u^{-2} \mathbb{1}_{(\tau,\infty)}(u) \, \mathrm{d}u \Psi(\xi) \mathrm{d}\xi \mathrm{d}\mathbb{P}_X \right] \\
&= \mathbb{E}_\Psi \exp\left[-\int_{\mathbb{X}} \int_S \max\left(\tau, \min_{1 \le i \le n} (y_i X_\xi(t_i - \xi)^{-1}) \right)^{-1} \Psi(\xi) \, \mathrm{d}\xi \, \mathrm{d}\mathbb{P}_X \right] \\
&= \mathbb{E}_\Psi \exp\left[-\int_{\mathbb{X}} \int_S \min\left(1/\tau, \max_{1 \le i \le n} ({y_i}^{-1} X_\xi(t_i - \xi)) \right) \Psi(\xi) \, \mathrm{d}\xi \, \mathrm{d}\mathbb{P}_X \right]
\end{aligned}$$

Hence, with condition (7.3)

$$\begin{aligned}
&\lim_{\tau \to 0} \ \mathbb{P}(\Pi(t_1) \le y_1, \dots, \Pi(t_n) \le y_n) \\
&= \mathbb{E}_\Psi \exp\left[-\int_{\mathbb{X}} \int_S \max_{1 \le i \le n} \left({y_i}^{-1} X_\xi(t_i - \xi) \right) \Psi(\xi) \, \mathrm{d}\xi \, \mathrm{d}\mathbb{P}_X \right]
\end{aligned}$$

which equals the finite-dimensional distribution of Y, see Remark 3 in Dirrler et al. (2016). It remains to prove the tightness of Π, that is

$$\lim_{\delta \to 0} \limsup_{\tau \to 0} \mathbb{P}(\omega_K(\Pi, \delta) > \varepsilon) = 0$$

with an arbitrary compact set $K \subset S$ and

$$\omega_K(\Pi, \delta) = \sup_{t_1, t_2 \in K : \|t_1 - t_2\| \le \delta} |\Pi(t_1) - \Pi(t_2)| \, .$$

This can be proven in the same way as in Theorem 7 of Dirrler et al. (2016).

□

We have just proven the convergence of Π to the Cox extremal process Y in the limiting case $\tau \to 0$. Since Y is in the MDA of Z, it seems natural to check whether Π is also in the MDA of Z for fixed $\tau > 0$.

Theorem 36. *Let the assumptions of Lemma 35 hold true and let Ψ be stationary with $\mathbb{E}\Psi(o) = 1$. The Matérn extremal process Π is in the max-domain of attraction of the mixed moving maxima process Z given by Equation (7.2). That is, if Π_i are iid. copies of Π, the convergence*

$$n^{-1} \bigvee_{i=1}^{n} \Pi_i \to Z,$$

holds weakly in $C(S, \mathbb{R})$.

Proof. Consider the sequence $\Pi^{(n)} = n^{-1} \bigvee_{i=1}^{n} \Pi_i$. We have to prove that $\Pi^{(n)}$ is tight, and that its marginal distributions converge to that of Z which are given by

$$\mathbb{P}(Z(t_1) \leq z_1, \ldots, Z(t_m) \leq z_m) = \exp\left[- \int_{\mathbb{X}} \int_S \max_{1 \leq i \leq n} \left({y_i}^{-1} X_\xi(t_i - \xi) \right) \, \mathrm{d}\xi \, \mathrm{d}\mathbb{P}_X \right].$$

The tightness can be derived by similar arguments as in the proof of Theorem 7 in Dirrler et al. (2016). The finite-dimensional distributions of $\Pi^{(n)}$ are given by

$$\begin{aligned}
&\mathbb{P}(\Pi^{(n)}(t_1) \leq z_1, \ldots, \Pi^{(n)}(t_m) \leq z_m) \\
&= \prod_{i=1}^{n} \mathbb{E}_\Psi \exp\left(- \int_{\mathbb{X}} \int_S \min\left(\frac{1}{\tau}, \max_{1 \leq j \leq m} \frac{X_\xi(t_j - \xi)}{n z_j} \right) \Psi_i(\xi) \, \mathrm{d}\xi \mathrm{d}\mathbb{P}_{X_\xi} \right) \\
&= \mathbb{E}_\Psi \exp\left(- \int_{\mathbb{X}} \int_S \min\left(\frac{n}{\tau}, \max_{1 \leq j \leq m} \frac{X_\xi(t_j - \xi)}{z_j} \right) n^{-1} \sum_{i=1}^{n} \Psi_i(\xi) \, \mathrm{d}\xi \mathrm{d}\mathbb{P}_{X_\xi} \right).
\end{aligned}$$

Since

$$\lim_{n \to \infty} \min\left(\frac{n}{\tau}, \max_{1 \leq j \leq m} \frac{X_\xi(t_j - \xi)}{z_j} \right) n^{-1} \sum_{i=1}^{n} \Psi_i(\xi) = \max_{1 \leq j \leq m} \frac{X_\xi(t_j - \xi)}{z_j},$$

we obtain

$$\begin{aligned}
&\lim_{n \to \infty} \mathbb{P}(\Pi^{(n)}(t_1) \leq z_1, \ldots, \Pi^{(n)}(t_m) \leq z_m) \\
&= \exp\left[- \int_{\mathbb{X}} \int_S \max_{1 \leq i \leq n} \left({z_j}^{-1} X_\xi(t_j - \xi) \right) \, \mathrm{d}\xi \, \mathrm{d}\mathbb{P}_X \right].
\end{aligned}$$

□

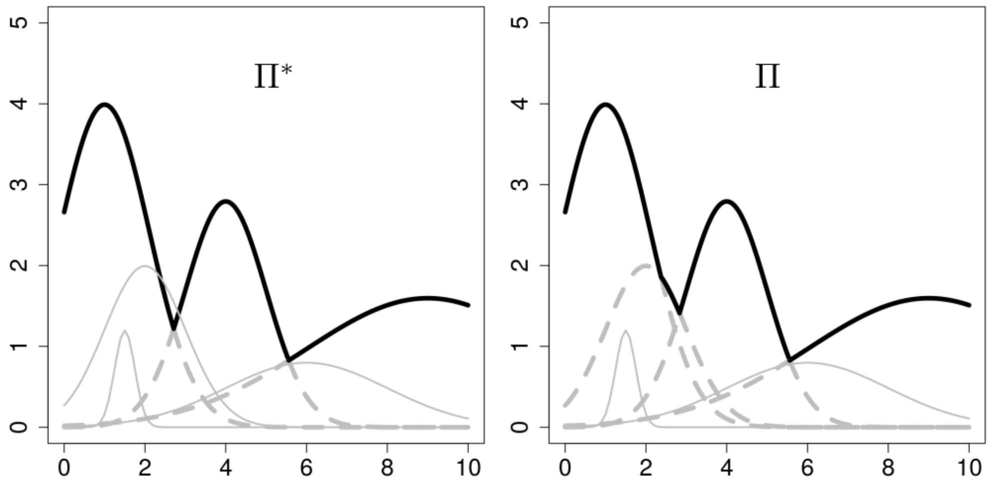

Figure 7.1.: The black solid lines form the final processes Π^* (left) and Π (right). The shape with centre $\xi = 2$ does not contribute to Π^*, since its centre is covered by an other shape.

7.2. Process of visible storm centres

We now consider the thinning function

$$f^*_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = \prod_{\boldsymbol{\xi}' \in \Phi_M} \left[1 - \mathbb{1}_{u_{\xi'} > u_\xi X_\xi(0) X_{\xi'}(\xi - \xi')^{-1}}\right].$$

The process resulting from dependent thinning equals

$$\begin{aligned} \Phi^*_{\mathrm{th}} &= \{(\xi, m_\xi) \in \Phi_M : f^*_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = 1\} \\ &= \{(\xi, m_\xi) \in \Phi_{\mathrm{th}} : \Pi(\xi) = m_\xi(o)\}. \end{aligned}$$

That is, a point (ξ, m_ξ) is retained if $m_\xi(o) \geq m_{\xi'}(\xi - \xi')$ for all other (ξ', m'_ξ) in Φ_M. This condition is sharper than (7.6) in the preceding section where points (ξ, m_ξ) are retained if there is an arbitrary t such that $m_\xi(t - \xi) \geq m_{\xi'}(t - \xi')$ for all other $(\xi', m_{\xi'})$ and also sharper than the condition for extremal functions (7.8) - therefore $\Phi^*_{\mathrm{th}} \subset \mathrm{ex}(\Phi_M) \subset \Phi_{\mathrm{th}}$.

Remark 37. *The set Φ^*_{th} is a subset of the set of extremal functions of Φ_M introduced by Dombry and Eyi-Minko (2013). If m_ξ is an extremal function which is not included in Φ^*_{th}, then $\Pi^*(\xi) > m_\xi(0)$, i.e. the centre of m_ξ is covered by other storms, see Figure 7.1.*

We introduce the *process of visible storm centres* defined by

$$\Pi^*(t) = \bigvee_{(\xi, m_\xi) \in \Phi^*_{\mathrm{th}}} m_\xi(t - \xi). \tag{7.9}$$

This process is closely related to the extremal hard-core process Π defined in (7.7), see also Figure 7.1.

We apply Proposition 33 and 34 from Section 6.3 to calculate first and second order properties of the thinned process $\Phi_0^* = \{\xi : (\xi, m_\xi) \in \Phi_{\text{th}}^*\}$.

Lemma 38. *Let $\Phi \sim \mathrm{LGCP}(\mu - \log\tau, C)$ with random intensity function $\tau^{-1}\Psi$. The intensity of Φ_0^* is given by*

$$\rho_{\Phi_0^*}(\xi) = p_0 \mathbb{E}\Psi(\xi) \int_{\mathbb{X}} X_\xi(o) \mathbb{E}_{\widetilde{\Psi}} \left[\frac{1 - \exp(-\tau^{-1} X_\xi(0)^{-1} \cdot c_{\widetilde{\Psi}})}{c_{\widetilde{\Psi}}} \right] \mathrm{d}\mathbb{P}_{X_\xi},$$

where

$$c_{\widetilde{\Psi}} = \int_S \mathbb{E}_X X_{\xi'}(\xi - \xi') \widetilde{\Psi}(\xi') \mathrm{d}\xi'$$

and $\tau^{-1}\widetilde{\Psi}$ is the random intensity function of $\widetilde{\Phi} \sim \mathrm{LGCP}(\widetilde{\mu} - \log\tau, C)$ and $\widetilde{\mu}(\cdot) = \mu(\cdot) + C(\cdot, \xi)$.

Proof. The retaining probability can be calculated by

$$\begin{aligned}
r(\xi, m_\xi) &= \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_S \int_{\mathbb{M}} \zeta(\boldsymbol{\xi}, \boldsymbol{\xi}') p(\boldsymbol{\xi}, \boldsymbol{\xi}') \tau^{-1} \widetilde{\Psi}(\xi') \mu(dm_{\xi'}) \mathrm{d}\xi' \right) \\
&= \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_S \int_{\mathbb{M}} \mathbb{1}_{u_{\xi'} > u_\xi X_\xi(0) X_{\xi'}(\xi - \xi')^{-1}} \tau^{-1} \widetilde{\Psi}(\xi') \mu(dm_{\xi'}) \mathrm{d}\xi' \right) \\
&= \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_S \int_{\mathbb{X}} \int_\tau^\infty \mathbb{1}_{u_{\xi'} > u_\xi X_\xi(0) X_{\xi'}(\xi - \xi')^{-1}} \widetilde{\Psi}(\xi') u_{\xi'}^{-2} \mathrm{d}u_{\xi'} \mathrm{d}\mathbb{P}_{X_{\xi'}} \mathrm{d}\xi' \right).
\end{aligned}$$

Due to the condition (7.5), $u_\xi X_\xi(0) X_{\xi'}(\xi - \xi')^{-1} \geq \tau$ and therefore

$$\begin{aligned}
r(\xi, m_\xi) &= \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{\mathbb{X}} \int_S \int_{u_\xi X_\xi(0) X_{\xi'}(\xi - \xi')^{-1}}^\infty u_{\xi'}^{-2} \widetilde{\Psi}(\xi') \mathrm{d}u_{\xi'} \mathrm{d}\mathbb{P}_{X_{\xi'}} \mathrm{d}\xi' \right) \\
&= \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{\mathbb{X}} \int_S \frac{X_{\xi'}(\xi - \xi')}{u_\xi X_\xi(0)} \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \right).
\end{aligned}$$

The intensity then equals

$$\begin{aligned}
\rho_{\Phi_{\text{th}}}(\xi) &= p_0 \rho_\Phi(\xi) \int_{\mathbb{M}} \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{\mathbb{X}} \int_S \frac{X_{\xi'}(\xi - \xi')}{u_\xi X_\xi(0)} \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \right) \mu(\mathrm{d}m_\xi) \\
&= p_0 \mathbb{E}\Psi(\xi) \int_{\mathbb{X}} \int_\tau^\infty \mathbb{E}_{\widetilde{\Psi}} \exp\left(-u^{-1} X_\xi(0)^{-1} c_{\widetilde{\Psi}, X_\xi} \right) u^{-2} \mathrm{d}u \mathrm{d}\mathbb{P}_{X_\xi},
\end{aligned}$$

with $c_{\widetilde{\Psi}} = \int_{\mathbb{X}} \int_S X_{\xi'}(\xi - \xi') \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}}$. By calculating the integral with respect to u we finally obtain

$$\rho_{\Phi_0^*}(\xi) = p_0 \mathbb{E}\Psi(\xi) \int_{\mathbb{X}} X_\xi(0) \mathbb{E}_{\widetilde{\Psi}} \left[\frac{1 - \exp(-\tau^{-1} X_\xi(0)^{-1} \cdot c_{\widetilde{\Psi}})}{c_{\widetilde{\Psi}}} \right] \mathrm{d}\mathbb{P}_{X_\xi}.$$

□

Lemma 39. *Let $\Phi \sim \mathrm{LGCP}(\mu - \log \tau^{-1}, C)$ with random intensity function Ψ. The second order intensity of Φ_0^* equals*

$$\begin{aligned}\rho^{(2)}_{\Phi_0^*}(\xi,\eta) = p_0^2 \mathbb{E}[\Psi(\xi)\Psi(\eta)] \int_{\mathbb{X}} \int_{\mathbb{X}} \Bigg[&\int_\tau^\infty \int_\tau^{\frac{u_\eta X_\eta(0)}{X_\xi(\eta-\xi)}} r(\boldsymbol{\xi}) r(\boldsymbol{\eta}) r(\boldsymbol{\xi},\boldsymbol{\eta}) u_\xi^{-2} u_\eta^{-2} \ \mathrm{d}u_\xi \ \mathrm{d}u_\eta \\ &- \int_\tau^\infty \int_{\frac{u_\xi X_\xi(0)}{X_\eta(\xi-\eta)}}^\infty r(\boldsymbol{\xi}) r(\boldsymbol{\eta}) r(\boldsymbol{\xi},\boldsymbol{\eta}) u_\xi^{-2} u_\eta^{-2} \ \mathrm{d}u_\eta \ \mathrm{d}u_\xi \Bigg] \mathrm{d}\mathbb{P}_{X_\xi} \mathrm{d}\mathbb{P}_{X_\eta}\end{aligned}$$

with

$$r(\boldsymbol{\xi}) = \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{\mathbb{X}} \int_S \frac{X_{\xi'}(\xi - \xi')}{u_\xi X_\xi(0)} \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \right)$$

and

$$r(\boldsymbol{\xi},\boldsymbol{\eta}) = \mathbb{E}_{\widetilde{\Psi}} \exp\left(\int_X \int_S \min\left(\frac{X_{\xi'}(\xi - \xi')}{u_\xi X_\xi(0)}, \frac{X_{\xi'}(\eta - \xi')}{u_\eta X_\eta(0)} \right) \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \right).$$

Here $\tau^{-1}\widetilde{\Psi}$ is the random intensity function of $\widetilde{\Phi} \sim \mathrm{LGCP}(\widetilde{\mu} - \log \tau^{-1}, C)$ and $\widetilde{\mu}(\cdot) = \mu(\cdot) + C(\cdot,\xi) + C(\cdot,\eta)$.

Proof. Due to Proposition 34

$$\begin{aligned}\rho^{(2)}_{\Phi_0^*}(\xi,\eta) = \rho^{(2)}_{\Phi}(\xi,\eta) p_0^2 \int_{\mathbb{M}} \int_{\mathbb{M}} \Bigg[& h(\boldsymbol{\xi},\boldsymbol{\eta}) h(\boldsymbol{\eta},\boldsymbol{\xi}) \\ & \mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{S\times\mathbb{M}} \big(1 - h(\boldsymbol{\xi},\boldsymbol{\xi}') h(\boldsymbol{\eta},\boldsymbol{\xi}')\big) \widetilde{\Psi}(\xi') \mathrm{d}\xi' \nu(\mathrm{d}m'_\xi) \right) \Bigg] \nu(\mathrm{d}m_\xi) \nu(\mathrm{d}m_\eta).\end{aligned}$$

Then

$$\begin{aligned}&\mathbb{E}_{\widetilde{\Psi}} \exp\left(- \int_{S\times\mathbb{M}} \big(1 - h(\boldsymbol{\xi},\boldsymbol{\xi}') h(\boldsymbol{\eta},\boldsymbol{\xi}')\big) \tau^{-1} \widetilde{\Psi}(\xi') \mathrm{d}\xi' \nu(\mathrm{d}m'_\xi) \right) \\ &= \mathbb{E}_{\widetilde{\Psi}} \exp\Bigg(- \int_X \int_S \int_\tau^\infty [\mathbb{1}_{u > u_\xi X_\xi(0)(X_{\xi'}(\xi-\xi'))^{-1}} + \mathbb{1}_{u > u_\eta X_\eta(0)(X_{\xi'}(\eta-\xi'))^{-1}} \\ &\quad - \mathbb{1}_{u > u_\xi X_\xi(0)(X_{\xi'}(\xi-\xi'))^{-1}} \mathbb{1}_{u > u_\eta X_\eta(0)(X_{\xi'}(\eta-\xi'))^{-1}}] \widetilde{\Psi}(\xi') u^{-2} \ \mathrm{d}u \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \Bigg) \\ &= r(\boldsymbol{\xi}) r(\boldsymbol{\eta}) \mathbb{E}_{\widetilde{\Psi}} \exp\Bigg(- \int_X \int_S \int_\tau^\infty \Big[\mathbb{1}_{u > \max[u_\xi X_\xi(0)(X_{\xi'}(\xi-\xi'))^{-1}, u_\eta X_\eta(0)(X_{\xi'}(\eta-\xi'))^{-1}]} \\ &\qquad \widetilde{\Psi}(\xi') u^{-2} \Big] \ \mathrm{d}u \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \Bigg) \\ &= r(\boldsymbol{\xi}) r(\boldsymbol{\eta}) \mathbb{E}_{\widetilde{\Psi}} \exp\left(\int_X \int_S \min\left(\frac{X_{\xi'}(\xi - \xi')}{u_\xi X_\xi(0)}, \frac{X_{\xi'}(\eta - \xi')}{u_\eta X_\eta(0)} \right) \widetilde{\Psi}(\xi') \mathrm{d}\xi' \mathrm{d}\mathbb{P}_{X_{\xi'}} \right)\end{aligned}$$

with $r(\boldsymbol{\xi}) = \mathbb{E}_{\widetilde{\Psi}} \exp\left(-\int_{\mathbb{X}}\int_S \frac{X_{\xi'}(\xi-\xi')}{u_\xi X_\xi(0)}\widetilde{\Psi}(\xi')\mathrm{d}\xi'\mathrm{d}\mathbb{P}_{X_{\xi'}}\right)$. Furthermore

$$\begin{aligned} h(\boldsymbol{\xi},\boldsymbol{\eta})h(\boldsymbol{\eta},\boldsymbol{\xi}) &= 1 - \mathbb{1}_{u_\eta > u_\xi X_\xi(0) X_\eta(\xi-\eta)^{-1}} - \mathbb{1}_{u_\xi > u_\eta X_\eta(0) X_\xi(\eta-\xi)^{-1}} \\ &\quad + \mathbb{1}_{u_\eta > u_\xi X_\xi(0) X_\eta(\xi-\eta)^{-1}} \mathbb{1}_{u_\xi > u_\eta X_\eta(0) X_\xi(\eta-\xi)^{-1}} \\ &= 1 - \mathbb{1}_{u_\eta > u_\xi X_\xi(0) X_\eta(\xi-\eta)^{-1}} - \mathbb{1}_{u_\xi > u_\eta X_\eta(0) X_\xi(\eta-\xi)^{-1}} \end{aligned}$$

and

$$\begin{aligned} \rho^{(2)}_{\Phi_0^*}(\xi,\eta) &= p_0^2 \mathbb{E}[\Psi(\xi)\Psi(\eta)]\Bigg[\int_{\mathbb{X}}\int_\tau^\infty\int_{\mathbb{X}}\int_\tau^\infty r(\boldsymbol{\xi})r(\boldsymbol{\eta})r(\boldsymbol{\xi},\boldsymbol{\eta})u_\xi^{-2}u_\eta^{-2}\ \mathrm{d}u_\xi\ \mathrm{d}u_\eta \mathrm{d}\mathbb{P}_{X_\xi}\mathrm{d}\mathbb{P}_{X_\eta} \\ &\qquad - \int_{\mathbb{X}}\int_\tau^\infty\int_{\mathbb{X}}\int_{\frac{u_\eta X_\eta(0)}{X_\xi(\eta-\xi)}}^\infty r(\boldsymbol{\xi})r(\boldsymbol{\eta})r(\boldsymbol{\xi},\boldsymbol{\eta})u_\xi^{-2}u_\eta^{-2}\ \mathrm{d}u_\xi \mathrm{d}\mathbb{P}_{X_\xi}\ \mathrm{d}u_\eta \mathrm{d}\mathbb{P}_{X_\eta} \\ &\qquad - \int_{\mathbb{X}}\int_\tau^\infty\int_{\mathbb{X}}\int_{\frac{u_\xi X_\xi(0)}{X_\eta(\xi-\eta)}}^\infty r(\boldsymbol{\xi})r(\boldsymbol{\eta})r(\boldsymbol{\xi},\boldsymbol{\eta})u_\xi^{-2}u_\eta^{-2}\ \mathrm{d}u_\eta \mathrm{d}\mathbb{P}_{X_\eta}\ \mathrm{d}u_\xi \mathrm{d}\mathbb{P}_{X_\xi}\Bigg] \\ &= p_0^2 \mathbb{E}[\Psi(\xi)\Psi(\eta)]\int_{\mathbb{X}}\int_{\mathbb{X}}\Bigg[\int_\tau^\infty\int_\tau^{\frac{u_\eta X_\eta(0)}{X_\xi(\eta-\xi)}} r(\boldsymbol{\xi})r(\boldsymbol{\eta})r(\boldsymbol{\xi},\boldsymbol{\eta})u_\xi^{-2}u_\eta^{-2}\ \mathrm{d}u_\xi\ \mathrm{d}u_\eta \\ &\qquad - \int_\tau^\infty\int_{\frac{u_\xi X_\xi(0)}{X_\eta(\xi-\eta)}}^\infty r(\boldsymbol{\xi})r(\boldsymbol{\eta})r(\boldsymbol{\xi},\boldsymbol{\eta})u_\xi^{-2}u_\eta^{-2}\ \mathrm{d}u_\eta\ \mathrm{d}u_\xi\Bigg]\mathrm{d}\mathbb{P}_{X_\xi}\mathrm{d}\mathbb{P}_{X_\eta}. \end{aligned}$$

□

Remark 40. *Let Z be the classical Smith model (Smith, 1990) in $\mathbb{R}^2$, that is X is the density of the two-dimensional standard-normal distribution. Then, the intensity of the process of visible storm centres of Z can be calculated as*

$$\lim_{\tau\to 0}\rho_{\Phi_0^*}(\xi) = X_\xi(o)\left[\int_S X_{\xi'}(\xi-\xi')\mathrm{d}\xi'\right]^{-1} = (2\pi)^{-1}.$$

In general, $\rho_{\Phi_0^*}$ and $\rho^{(2)}_{\Phi_0^*}$ cannot be explicitly calculated if Ψ is random. However, numerical calculation of $\rho_{\Phi_0^*}$ is feasible in most cases. For certain choices of X, e.g. $X(t) = (1-t^2)_+$ in $\mathbb{R}$, even $\rho^{(2)}_{\Phi_0^*}$ is numerically tractable.

8. Discussion

The initial motivation of Part III was to improve the estimation procedures from Chapter 4, especially the non-parametric estimation presented in Section 4.1. The generalized Matérn model occurred as side effect of these reflections and since - to our best knowledge - nothing similar was known, we decided to spend some extra time in order to examine the new model. In the first part of this chapter, we propose an alternative approach for the non-parametric estimation in Section 4.1 by using the results of the previous chapter. In the second part, we discuss the problem of modelling the contributing storms.

Alternative approach to the non-parametric estimation of ψ. In Section 4.1 we estimate the realization ψ of the intensity process Ψ. To this end we first extract the visible storm centres from our observed data and pretend that those points are the centres of all contributing storms. We approximate the intensity function ψ_K^y of the true centres of contributing points by $\psi_K^y \approx \psi(s) b_K^y(s)$ with

$$b_K^y(s) = \mathbb{E}_X \left[\sup_{t \in K} \frac{X(t-s)}{y(t)} \right].$$

Finally, we compute kernel estimators for ψ_K^y and use $\hat{\psi}(s) = b_K^y(s)^{-1} \hat{\psi}_K^y(s)$ as estimator for ψ. This procedure has two big weaknesses. For one thing, our observed data is misspecified, for another thing we also approximate the true intensity of the contributing storm centres. These problems can be substantially reduced by using an alternative procedure which exploits the results of the previous section and which we sketch in the following.

Let Y be the Cox extremal process (7.4) and let $\widetilde{N}$ be the point process that encompasses all visible storm centres of Y. We define the intensity function of $\widetilde{N}$ as $\Psi_{\widetilde{N}}$. Then $\widetilde{N}$ can be described in terms of the process Φ_0^* from Section 7.2, more specifically

$$\lim_{\tau \to 0} \Phi_0^* = \widetilde{N},$$

in terms of weak convergence. Therefore, the intensity function $\Psi_{\widetilde{N}}$ of $\widetilde{N}$ can be calculated by

$$\begin{aligned} \Psi_{\widetilde{N}}(\xi) = & \lim_{\tau \to 0} \Psi(\xi) \int_{\mathbb{X}} X_\xi(o) \mathbb{E}_{\widetilde{\Psi}} \left[\frac{1 - \exp(-\tau^{-1} X_\xi(0)^{-1} \cdot c_{\widetilde{\Psi}})}{c_{\widetilde{\Psi}}} \right] \mathrm{d}\mathbb{P}_{X_\xi} \\ = & \Psi(\xi) \int_{\mathbb{X}} X_\xi(o) \mathbb{E}_{\widetilde{\Psi}}(c_{\widetilde{\Psi}}{}^{-1}) \mathrm{d}\mathbb{P}_{X_\xi} = \Psi(\xi) \frac{\mathbb{E}_X X(o)}{\mathbb{E}_{\widetilde{\Psi}} c_{\widetilde{\Psi}}}. \end{aligned}$$

That is, a realization ψ of Ψ fulfils the equality

$$\psi(\xi) = \psi_{\widetilde{N}}(\xi)\frac{\mathbb{E}_{\widetilde{\Psi}} c_{\widetilde{\Psi}}}{\mathbb{E}_X X(o)},$$

where $\psi_{\widetilde{N}}$ is a realization of the intensity function of the process of visible storm centres of the Cox extremal process Y. Hence we can estimate ψ by computing a kernel estimator for $\psi_{\widetilde{N}}$ and calculating $\mathbb{E}X(o)$ (which is assumed to be known) and $\mathbb{E}c_{\widetilde{\Psi}}$. So the original problem is reduced to the calculation of $\mathbb{E}c_{\widetilde{\Psi}} = \mathbb{E}_{\widetilde{\Psi}} \int_S \mathbb{E}_X X(\xi - \xi')\widetilde{\Psi}(\xi')\mathrm{d}\xi'$. Note that $c_{\widetilde{\Psi}}$ does also depend on ξ – see Lemma 38 – therefore its handling is difficult as well. Still, this procedure should be superior to the one proposed in Section 4.1.

Exact representation of the process of contributing storm centres. Our generalization of the Matérn model which we present in Chapter 6, is based on thinning functions of type

$$f_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = \prod_{(\xi', m'_\xi)\in\Phi_M} (1 - \zeta(\xi, m_\xi, \xi', m_{\xi'})p(\xi, m_\xi, \xi', m_{\xi'})).$$

This representation as product of functions over the points $(\xi', m_{\xi'}) \in \Phi_M$ is essential, since it enables the use of the generating functional (6.1) and thereby the calculation of first and second order properties of the thinned process. Unfortunately, the product representation also entails that our thinning algorithm is only suitable to model pairwise interaction between the points.

In order to model the contributing storms, we can modify our approach by choosing the new thinning function

$$\widetilde{f}_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = 1 - \widetilde{\zeta}(\xi, m_\xi, \Phi_M)$$

with $\widetilde{\zeta} : (S, \mathbb{M}) \times M_{\mathrm{lf}} \to \{0, 1\}$ and

$$\widetilde{\zeta}(\xi, m_\xi, \Phi_M) = \begin{cases} 1, & \text{if } \forall t \in S : m_\xi(t-\xi) < \bigvee_{(\xi', m_{\xi'})\in\Phi_M} m'_\xi(t-\xi') \\ 0, & \text{else.} \end{cases}$$

Then the thinned process

$$\widetilde{\Phi_{\mathrm{th}}} = \{(\xi, m_\xi) \in \Phi_M : (\xi, m_\xi) \in \Phi_M, \widetilde{f}_{\mathrm{th}}(\Phi_M; \xi, m_\xi) = 1\},$$

fulfils $\lim_{\tau\to 0} \widetilde{\Phi_{\mathrm{th}}} \overset{\mathcal{D}}{=} \mathrm{ex}(\Phi_M)$, as desired. Though, the retaining probability of an individual point $(\xi, m_\xi) \in \Phi_M$ is then given by

$$r(\xi, m_\xi) = \int_{M_{\mathrm{lf}}} \widetilde{f}_{\mathrm{th}}(\varphi; \xi, m_\xi)P^!_{\boldsymbol{\xi}}(\mathrm{d}\varphi) = 1 - \mathbb{E}\widetilde{\zeta}(\boldsymbol{\xi}, \widetilde{\Phi}_M)$$

with $\widetilde{\Phi}_M$ as defined in Proposition 33. The expected value of $\widetilde{\zeta}(\boldsymbol{\xi}, \widetilde{\Phi}_M)$ cannot be further simplified, that is we do not get any knowledge gain from $\widetilde{\Phi_{\mathrm{th}}}$ compared to $\mathrm{ex}(\Phi_M)$. This is the reason we introduced the process of visible storm centres in Section 7.2 instead.

Bibliography

R. J. Adler. *The geometry of Random fields.* John Wiley & SonsLtd, 1981.

I. T. Andersen and U. Hahn. Matérn thinned Cox processes. *Spat. Stat.*, 15:1–21, 2016.

Gutti Jogesh Babu and Eric D. Feigelson. Spatial point processes in astronomy. *J. Statist. Plann. Inference*, 50(3):311–326, 1996.

Stefano Castruccio, Raphael Huser, and Marc G. Genton. High-order composite likelihood inference for max-stable distributions and processes. *J. Comput. Graph. Statist.*, 2015.

V. Chavez-Demoulin, A. C. Davison, and A. J. McNeil. Estimating value-at-risk: a point process approach. *Quant. Finance*, 5(2):227–234, 2005. ISSN 1469-7688.

S. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, third edition, 2013.

J. Coeurjolly, J. Møller, and R. P. Waagepetersen. Palm distributions for log Gaussian Cox processes. *arXiv:1506.04576*, 2015.

D. Cooley. *Statistical Analysis of Extremes Motivated by Weather and Climate Studies: Applied and Theoretical Advances.* PhD thesis, University of Colorado, 2005.

D. R. Cox. Some statistical models connected with series of events. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 17:129–164, 1955.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods.* Springer, 2003.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure.* Springer, 2008.

L. de Haan. A spectral representation for max-stable processes. *Ann. Probab.*, 12(4): 1194–1204, 1984.

L. de Haan and A. Ferreira. *Extreme value theory.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. An introduction.

A. B. Dieker and T. Mikosch. Exact simulation of Brown-Resnick random fields at a finite number of locations. *Extremes*, 18(2):301–314, 2015.

P. J. Diggle. A kernel method for smoothing point process data. *J. Roy. Statist. Soc. Ser. C*, 34:138–147, 1985.

P. J. Diggle and R. J. Gratton. Monte carlo methods of inference for implicit statistical models. *J. R. Stat. Soc. Ser. B Stat. Methodol*, 46(2):193–227, 1984.

P. J. Diggle, P. Moraga, B. Rowlingson, and B. M. Taylor. Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Statist. Sci.*, 28 (4):542–563, 2013.

M. Dirrler and M. Schlather. On a generalization of Matérn hard-core processes with applications to max-stable processes. *arXiv:1709.05868*, 2017.

M. Dirrler, M. Schlather, and K. Strokorb. Conditionally max-stable random fields based on log gaussian cox processes. *arXiv:1612.04576*, 2016.

C. Dombry. Extremal shot noises, heavy tails and max-stable random fields. *Extremes*, 15(2):129–158, 2012.

C. Dombry and F. Eyi-Minko. Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab*, 18:no. 7, 21, 2013.

C. Dombry, S. Engelke, and M. Oesting. Exact simulation of max-stable processes. *Biometrika*, 2016.

S. Engelke, Z. Kabluchko, and M. Schlather. An equivalent representation of the Brown-Resnick process. *Statist. Probab. Lett.*, 81(8):1150–1154, 2011.

S. Engelke, A. Malinowski, Z. Kabluchko, and M. Schlather. Estimation of Hüsler-Reiss distributions and Brown-Resnick processes. *J. R. Stat. Soc. Ser. B Stat. Methodol*, 77 (1):239–265, 2015.

M. G. Genton, S. A. Padoan, and H. Sang. Multivariate max-stable spatial processes. *Biometrika*, 102(1):215–230, 2015.

E. Giné, M. G. Hahn, and P. Vatan. Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields*, 87(2):139–165, 1990.

P. Guttorp and T. Gneiting. Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika*, 93(4):989–995, 2006.

L. Heinrich and I. S. Molchanov. Some limit theorems for extremal and union shot-noise processes. *Math. Nachr.*, 168:139–159, 1994.

Raphaël Huser, Anthony C. Davison, and Marc G. Genton. Likelihood estimators for multivariate extremes. *Extremes*, 19(1):79–103, 2016.

A.M. Ibrahim, T. ElBatt, and A. El-Keyi. Coverage probability analysis for wireless networks using repulsive point processes. *Proc. IEEE PIMRC*, pages 1002–1007, 2013.

M. Jourlin, B. Laget, G. Matheron, F. Meyer, F. Preteux, M. Schmitt, and J. Serra. *Image Analysis and Mathematical Morphology. Volume 2: Theoretical Advances.* Academic Press, Inc., London, 1988.

Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max-stable fields associated to negative definite functions. *Ann. Probab.*, 37(5):2042–2065, 2009.

A. F. Karr. *Point processes and their statistical inference.* Probability, pure and applied. Marcel Dekker, 1986.

M. Kuronen and L. Leskelä. Hard-core thinnings of germ-grain models with power-law grain sizes. *Adv. in Appl. Probab.*, 45(3):595–625, 2013.

Z Liu, J Blanchet, A.B. Dieker, and T. Mikosch. Optimal exact simulation of max-stable and related random fields. *arXiv:1609.06001*, 2016.

M. Månsson and M. Rudemo. Random patterns of nonoverlapping convex grains. *Adv. in Appl. Probab.*, 34(4):718–738, 2002.

B. Matérn. *Spatial variation: Stochastic models and their application to some problems in forest surveys and other sampling investigations.* Meddelanden Fran Statens Skogsforskningsinstitut, Band 49, Nr. 5, Stockholm, 1960.

G. Matheron. Suffit-il, pour une covariance, d'etre de type positif? *Sciences de la terre, serie informatique geologique*, 26, 1987.

J. Mecke. Stationäre zufällige Maße auf lokalkompakten Abelschen Gruppen. *Z. Wahrscheinlichkeitsth.*, 9:36–58, 1967.

J. Møller. Shot noise Cox processes. *Adv. in Appl. Probab.*, 35(3):614–640, 2003. ISSN 0001-8678. doi: 10.1239/aap/1059486821. URL `http://dx.doi.org/10.1239/aap/1059486821`.

J. Møller and F. P. Schoenberg. Thinning spatial point processes into poisson process. *Adv. in Appl. Probab.*, 42(2):347–358, 2010.

J. Møller and R. P. Waagepetersen. *Statistical inference and simulation for spatial point processes*, volume 100 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, Boca Raton, FL, 2004.

J. Møller, A. R. Syversveen, and R. P. Waagepetersen. Log Gaussian Cox Processes. *Scand. J. Statist.*, 25:451–482, 1998.

S. Nadarajah, C.W. Anderson, and J. A. Tawn. Ordered multivariate extremes. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 60:473–496, 1998.

J. Nolan. *Stable Distributions : Models for Heavy-Tailed Data.* Birkhauser, 2016.

M. Oesting and M. Schlather. Conditional sampling for max-stable processes with a mixed moving maxima representation. *Extremes*, 17(1):157–192, 2014.

M. Oesting, Z. Kabluchko, and M. Schlather. Simulation of Brown-Resnick processes. *Extremes*, 15(1):89–107, 2012.

M. Oesting, M. Schlather, and C. Zhou. On the normalized spectral representation of max-stable processes on a compact set. *arXiv:1310.1813v1*, 2013.

M. Oesting, M. Schlather, and P. Friederichs. Statistical post-processing of forecasts for extremes using bivariate brown-resnick processes with an application to wind gusts. *arXiv:1312.4584v2*, 2015.

T. Opitz. Extremal t processes: elliptical domain of attraction and a spectral representation. *J. Multivariate Anal.*, 122:409–413, 2013.

N. Picard. Tree density estimations using a distance method in Mali savanna. *Forest Sci.*, 51:7–18, 2005.

S. I. Resnick. *Extreme values, regular variation and point processes.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2008. Reprint of the 1987 original.

B. D. Ripley. Modelling spatial patterns. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 39(2): 172–212, 1977. With discussion.

G. Samorodnitsky and M. S. Taqqu. *Stable non-Gaussian random processes.* Stochastic Modeling. Chapman & Hall, New York, 1994. Stochastic models with infinite variance.

Jeffrey D. Scargle and Gutti Jogesh Babu. Point processes in astronomy: exciting events in the universe. In *Stochastic processes: modelling and simulation*, volume 21 of *Handbook of Statist.*, pages 795–825. North-Holland, Amsterdam, 2003.

M. Schlather. Models for stationary max-stable random fields. *Extremes*, pages 33–44, 2002.

M. Schlather and J. A. Tawn. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90(1):139–156, 2003.

J. Serra. *Image analysis and mathematical morphology.* Academic Press, Inc., London, 1984. English version revised by Noel Cressie.

R. L. Smith. Max-stable processes and spatial extremes. Unpublished Manuscript, 1990.

A. G. Stephenson, B. A. Shaby, B. J. Reich, and A. L. Sullivan. Estimating spatially varying severity thresholds of a forest fire danger rating system using max-stable extreme-event modeling. *J. Appl. Meteor.*, 54(2), 2015.

S. A. Stoev. On the ergodicity and mixing of max-stable processes. *Stochastic Process. Appl.*, 118(9):1679–1705, 2008.

S. A. Stoev and M. S. Taqqu. Extremal stochastic integrals: a parallel between max-stable processes and α-stable processes. *Extremes*, 8(4):237–266 (2006), 2005.

S. A. Stoev and M. S. Taqqu. How rich is the class of multifractional Brownian motions? *Stochastic Process. Appl.*, 116(2):200–221, 2006.

D. Stoyan. Statistical analysis of spatial point processes: A soft-core model and cross-correlations of marks. *Biom. J.*, 29:971–980, 1987. ISSN 0323-3847; 1521-4036/e.

D. Stoyan. Thinnings of point processes and their use in the statistical analysis of a settlement pattern with deserted tillages. *Statistics*, 19:45–56, 1988.

D. Stoyan and H. Stoyan. On one of Matérn's hard-core point process models. *Math. Nachr.*, 122:205–214, 1985.

D. Stoyan and H. Stoyan. *Fraktale - Formen - Punktbilder*. 1992.

Dietrich Stoyan and Antti Penttinen. Recent applications of point process methods in forestry statistics. *Statist. Sci.*, 15(1):61–78, 2000. ISSN 0883-4237.

J. Teichmann, F. Ballani, and K.G. van den Boogaart. Generalizations of Matérn's hard-core point processes. *Spatial Statistics*, 3:33 – 53, 2013.

M. Westcott. The probability generating functional. *J. Austral. Math. Soc.*, 14:448–466, 1972.

Z. Zhang and R. L. Smith. The behavior of multivariate maxima of moving maxima processes. *J. Appl. Probab.*, 41(4):1113–1123, 2004.

Hiermit erkläre ich, dass ich die vorliegende Arbeit mit dem Titel “Spatial point process models with applications to max-stable random fields “ selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Mannheim, den 26.09.2017

. . . .
Martin Dirrler