1

Conditional Simulation of Max-Stable Processes

Clément Dombry

Université de Franche-Comté

Marco Oesting

INRA, AgroParisTech

Mathieu Ribatet

Université Montpellier 2 and ISFA, Université Lyon 1

CONTENTS

Abstract

Max-stable processes are well established models for spatial extremes. In this chapter, we address the prediction problem: suppose a max-stable process is observed at some locations only, how can we use these observations to predict the behavior of

the process at other unobserved locations? Mathematically, the prediction problem is related to the conditional distribution of the process given the observations. Recently, Dombry and Eyi-Minko (2013) provided an explicit theoretical formula for the conditional distributions of max-stable processes. The result relies on the spectral representation of the max-stable process as the pointwise maxima over an infinite number of spectral functions belonging to a Poisson point process. The effect of conditioning on the Poisson point process is analyzed, resulting in the notions of hitting scenario and extremal or subextremal functions. Due to the complexity of the structure of the conditional distributions, conditional simulation appears at the same time challenging and important to assess characteristics that are analytically intractable such as the conditional median or quantiles. The issue of conditional simulation was considered by Dombry et al. (2012) who proposed a three step procedure for conditional sampling. As the conditional simulation of the hitting scenario becomes computationally very demanding even for a moderate number of conditioning points, a Gibbs sampler approach was proposed for this step. The results are illustrated on some simulation studies and we propose several diagnostics to check the performance.

1.1 Introduction: the prediction problem and conditional distribution

In classical geostatistics, Gaussian random fields play a central role in the statistical theory based on the Central Limit Theorem. In a similar manner, max-stable random fields turn out to be fundamental models for spatial extremes since they extend the well known univariate and multivariate extreme value theory to the infinite dimensional setting. Max-stable random fields arise naturally when considering the component-wise maxima of a large number of independent and identically random fields and seeking for a limit under a suitable affine normalization.

In this multivariate or functional setting, the notion of dependence is crucial: how does an extreme event occurring in some region affect the behavior of the random field at other locations? This is related to the *prediction problem* which is an important and long-standing challenge in extreme value theory. Suppose that a max-stable random field $Z = (Z(x))_{x\in\mathcal{X}}$ is observed at some stations $x_1, \ldots, x_k \in \mathcal{X}$ only, yielding

$$Z(x_1) = z_1, \ldots, Z(x_k) = z_k. \tag{1.1}$$

How can we take benefit from these observations and predict the random field Z at other places? We are naturally lead to consider the *conditional distribution* of $(Z(x))_{x\in\mathcal{X}}$ given the observations (1.1).

In the classical Gaussian framework, i.e., if Z is a centered Gaussian random field, it is well known that the corresponding conditional distribution remains Gaussian and simple formulas give the conditional mean and covariance structure. This theory is strongly linked with the theory of Hilbert spaces: for example, the conditional expectation of $Z(x)$ can be obtained as the L^2-projection of the random field η

onto the Gaussian subspace generated by the variables $\{Z(x_i),\ 1 \leq i \leq k\}$, resulting in the linear combination

$$\mathbb{E}\left[Z(x) \mid Z(x_1), \ldots, Z(x_k)\right] = \alpha_1(x)Z(x_1) + \cdots + \alpha_k(x)Z(x_k), \quad x \in \mathcal{X},$$

for some weight functions $\alpha_i(x)$, $i = 1, \ldots, n$, that are obtained via the kriging theory (cf. Chilès and Delfiner (1999), for example). A conditional simulation of Z given the observations (1.1) is then easily performed by setting

$$Z(x) = \tilde{Z}(x) + \alpha_1(x)(z_1 - \tilde{Z}(x_1)) + \cdots + \alpha_k(x)(z_k - \tilde{Z}(x_k)), \quad x \in \mathcal{X},$$

where $\tilde{Z}$ denotes the realization of an unconditional simulation of the random field Z.

In extreme value theory, the prediction problem turns out to be more difficult. A similar kriging theory for max-stable processes is very appealing and a first approach in that direction was done by Davis and Resnick (1989, 1993). They introduced a L^1-metric between max-stable variables and proposed a kind of projection onto max-stable spaces. To some extent, this work mimics the corresponding L^2-theory for Gaussian spaces. However, unlike the exceptional Gaussian case, there is no clear relationship between the predictor obtained by projection onto the max-stable space generated by the variables $\{Z(x_i),\ 1 \leq i \leq k\}$ and the conditional distributions of η with respect to these variables. Conditional distributions have been considered first by Weintraub (1991) in the case of a bivariate vector. A major contribution is the work by Wang and Stoev (2011) where the authors consider max-linear random fields, a special class of max-stable random fields with discrete spectral measure, and give an exact expression of the conditional distributions as well as efficient algorithms. The max-linear structure plays an essential role in their work and provides major simplifications since in this case Z admits the simple representation

$$Z(x) = \bigvee_{j=1}^{q} F_j f_j(x), \quad x \in \mathcal{X},$$

where the symbol $\bigvee$ denotes the maximum, $f_1, \ldots, f_q$ are deterministic functions and $F_1, \ldots, F_q$ are i.i.d. random variables with unit Fréchet distribution. The authors determine the conditional distributions of $(F_j)_{1 \leq j \leq q}$ given the observations (1.1) and deduce the conditional distribution of Z. Another approach by Oesting and Schlather (2014) deals with max-stable random fields with a mixed moving maxima representation.

We present and discuss here the recent result from Dombry and Eyi-Minko (2013) and Dombry et al. (2013) providing formulas for the conditional distributions of max-stable random fields as well as efficient algorithm for conditional sampling. Note that the results can be stated in the more general framework of max-infinitely divisible processes, but for the sake of simplicity, we stick here to the max-stable case. Section 2 reviews the theoretical results on conditional distribution: we introduce the notion of extremal functions and hitting scenarios and explicit the exact distribution of the max-stable process given the observations. Here, we focus on the

framework of regular models that yields tractable formulas. In Section 3, we put the emphasis on efficient simulation and discuss a 3-step procedure for conditional sampling. A difficult step is the conditional sampling for the hitting scenario for which a Gibbs sampler approach is proposed. Section 4 is devoted to simulation studies and we present several diagnostics to check the performance of the methods suggested.

1.2 Conditional distribution of max-stable processes

In the sequel, we consider Z a sample continuous max-stable random field on $\mathcal{X} \subset \mathbb{R}^d$. We can assume without loss of generality that Z has unit Fréchet margins. Thus, the process Z possesses a spectral representation (see for example de Haan (1984), Penrose (1992), Schlather (2002) or de Haan and Ferreira (2006))

$$Z(x) = \max_{i \geq 1} \zeta_i Y_i(x), \qquad x \in \mathcal{X}, \tag{1.2}$$

where $\{\zeta_i\}_{i\geq 1}$ are the points of a Poisson process on $(0, \infty)$ with intensity $\zeta^{-2}\mathrm{d}\,\zeta$, $(Y_i)_{i\geq 1}$ are independent replicates of a non-negative continuous sample path stochastic process Y such that $\mathbb{E}[Y(x)] = 1$ for all $x \in \mathcal{X}$, $(\zeta_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ are independent.

We now introduce a fundamental object in our analysis which is a function-valued Poisson point process associated with the representation (1.2). Let $\mathcal{C} = \mathcal{C}(\mathcal{X}, [0, +\infty))$ be the space of continuous non-negative functions on $\mathcal{X}$. We consider the $\mathcal{C}$-valued point process $\Phi = \{\phi_i\}_{i\geq 1}$ where $\phi_i(x) = \zeta_i Y_i(x)$ with ζ_i and Y_i as in (1.2). It is well known (de Haan and Ferreira, 2006) to verify that Φ is a Poisson point process with intensity measure Λ given by

$$\Lambda(A) = \int_0^\infty \mathbb{P}[\zeta Y \in A]\zeta^{-2}\mathrm{d}\zeta, \quad A \subset \mathcal{C} \text{ Borel}.$$

Our strategy is to work on the level of the Poisson point process Φ rather than on the level of the max-stable process Z and to derive the conditional distribution of Φ given the observations (1.1). The conditional distribution of Z is then deduced easily.

1.2.1 Extremal functions, subextremal functions and hitting scenarios

The observations (1.1) together with the spectral representation (1.2) yield the constraint

$$\max_{i\geq 1} \phi_i(x_1) = z_1, \ldots, \max_{i\geq 1} \phi_i(x_k) = z_k.$$

A first step in our strategy is the analysis of the way how these constraints are met, i.e. the way how the maxima are attained. An important preliminary remark is that for each $j = 1, \cdots, k$, the maximum $Z(x_j) = \max_{i\geq 1} \phi_i(x_j)$ is almost surely attained by a unique function ϕ_i, which can be easily seen by the fact that the point

process $\{\varphi_i(x_j), i \geq 1\}$ on $(0, \infty)$ has intensity $\zeta^{-2}d\zeta$. This leads to the following definition.

Definition 1.2.1. *The (almost surely) unique function $\phi \in \Phi$ such that $\phi(x_j) = Z(x_j)$ is called the extremal function at x_j and denoted by $\varphi^+_{x_j}$.*
Furthermore, we define the extremal point process $\Phi^+ = \{\varphi^+_{x_j}\}_{1\leq j\leq k}$ as the set of extremal functions with respect to the conditioning points $\{x_j\}_{1\leq j\leq k}$.

Note that the notation $\Phi^+ = \{\varphi^+_{x_j}\}_{1\leq j\leq k}$ may include the multiple occurrence of some functions for example if $\varphi^+_{x_1} = \varphi^+_{x_2}$. This is not taken into account in the extremal point process Φ^+ where each point has multiplicity one. The possible redundancies are captured by the so-called hitting scenario. We first introduce this notion with simple examples. Assume first that there are $k = 2$ conditioning points x_1, x_2 and hence two extremal functions $\varphi^+_{x_1}$ and $\varphi^+_{x_2}$. Two cases can occur:

- either $\varphi^+_{x_1} = \varphi^+_{x_2}$, i.e. the maxima at points x_1 and x_2 are reached by the same extremal function, in this case the hitting scenario is $\theta = \{x_1, x_2\}$;
- or $\varphi^+_{x_1} \neq \varphi^+_{x_2}$, i.e. the maxima at points x_1 and x_2 are reached by different extremal functions, in this case the hitting scenario is $\theta = (\{x_1\}, \{x_2\})$.

In the case of $k = 3$ conditioning points, there are 5 different possibilities for the hitting scenario:

$$(\{x_1, x_2, x_3\}),\ \ (\{x_1, x_2\}, \{x_3\}),\ \ (\{x_1, x_3\}, \{x_2\}),$$

$$(\{x_1\}, \{x_2, x_3\}) \text{ and } (\{x_1\}, \{x_2\}, \{x_3\}).$$

The interpretation is straightforward: for instance, the hitting scenario $\theta = (\{x_1, x_3\}, \{x_2\})$ corresponds to the case when the maxima at x_1 and x_3 are attained by the same extremal function but the maximum at x_2 corresponds to a different extremal event, i.e. $\varphi^+_{x_1} = \varphi^+_{x_3} \neq \varphi^+_{x_2}$. The general definition is as follows.

Definition 1.2.2. *The hitting scenario θ is a random partition $(\theta_1, \ldots, \theta_\ell)$ of the conditioning points $\{x_1, \cdots, x_k\}$ such that for any $j_1 \neq j_2$, x_{j_1} and x_{j_2} are in the same component of θ if and only if $\varphi^+_{j_1} = \varphi^+_{j_2}$.*

The hitting scenario θ takes into account the redundancies in $\Phi^+ = \{\varphi^+_{x_1}, \cdots, \varphi^+_{x_k}\}$ and it is straightforward that the number of blocks ℓ of θ is exactly the number of distinct extremal functions and hence the cardinality of Φ^+. Hence we can rewrite $\Phi^+ = \{\varphi^+_1, \cdots, \varphi^+_\ell\}$ where $\varphi^+_j = \varphi^+_x$ for all $x \in \theta_j$.

So far, we considered only those functions ϕ_i that hit the maximum Z at some conditioning points $\{x_j\}_{1\leq j\leq k}$. The remaining functions are called sub-extremal, as in the following definition.

Definition 1.2.3. *A function $\phi \in \Phi$ satisfying $\phi(x_j) < Z(x_j)$ for all $1 \leq j \leq k$ is called subextremal.*
The set of subextremal functions is called the subextremal point process Φ^-.

This yields a disjoint decomposition of the point process $\Phi = \Phi^+ \cup \Phi^-$ into its extremal and subextremal parts. This is illustrated in Figure 1.2.1 where the extremal functions (black) and the subextremal functions (gray) are depicted as well as the corresponding hitting scenarios.

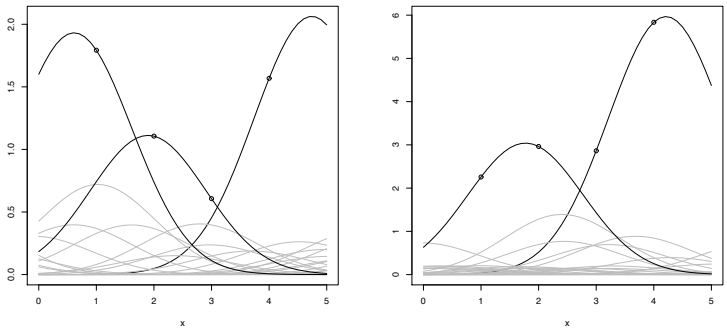

FIGURE 1.1
Two realizations of the Poisson point process Φ and of the corresponding hitting scenario θ with conditioning points $x_i = i$, $i = 1, \cdots, 4$ represented by the circles. Left: the hitting scenario is $\theta = (\{x_1\}, \{x_2, x_3\}, \{x_4\})$. Right: the hitting scenario is $\theta = (\{x_1, x_2\}, \{x_3, x_4\})$.

A key result in the study of the conditional distribution is the following theorem providing the joint distribution of (θ, Φ^+, Φ^-). We denote by $\mathcal{P}_k$ the set of partitions of $\{x_1, \ldots, x_k\}$ and by $\mathcal{M}_p(\mathcal{C})$ the set of $\mathcal{C}$-valued point measures. We introduce some vectorial notation. Let $\mathbf{x} = (x_1, \ldots, x_k)$ and for $\tau = (\tau_1, \ldots, \tau_\ell) \in \mathcal{P}_k$, we introduce $\mathbf{x}_{\tau_j} = (x)_{x \in \tau_j}$. For any vector $\mathbf{s} = (s_1, \ldots, s_m) \in \mathcal{X}^m$ and any function $f : \mathcal{X} \to \mathbb{R}$, we note $f(\mathbf{s}) = (f(s_1), \ldots, f(s_m))$. If $f_1, f_2 : \mathcal{X} \to \mathbb{R}^m$ are two functions, the notation $f_1(\mathbf{s}) < f_2(\mathbf{s})$ means that $f_1(s_j) < f_2(s_j)$, $j = 1, \ldots, m$.

Theorem 1.2.4. *For any partition $\tau = (\tau_1, \ldots, \tau_\ell)$ in $\mathcal{P}_k$ and any Borel sets $A \subset \mathcal{C}^\ell$ and $B \subset \mathcal{M}_p(\mathcal{C})$, we have*

$$\begin{aligned} &\mathbb{P}[\theta = \tau, (\varphi_1^+, \ldots, \varphi_\ell^+) \in A, \Phi^- \in B] \\ = &\int_{\mathcal{C}^\ell} 1_{\{\max_{j' \neq j} f_{j'}(\mathbf{x}_{\tau_j}) < f_j(\mathbf{x}_{\tau_j}),\ j=1,\ldots,\ell,\}} 1_{\{(f_1,\ldots,f_\ell) \in A\}} \cdots \\ &\qquad \cdots \mathbb{P}\left[\Phi \in B \text{ and } \forall \phi \in \Phi,\ \phi(\mathbf{x}) < \max_{1 \leq j \leq \ell} f_j(\mathbf{x})\right] \Lambda(df_1) \cdots \Lambda(df_\ell). \end{aligned}$$

The main technical tool for this proof is the Mecke-Slivniack formula from stochastic geometry, see e.g. Stoyan et al. (1987). We sketch here the main lines of the proof.

Proof. First note that the event $\{\theta = \tau, (\varphi_1^+, \ldots, \varphi_\ell^+) \in A,\ \Phi^- \in B\}$ is realized if and only if there exists a ℓ-tuple $(\phi_1, \cdots, \phi_\ell) \in \Phi^\ell$ satisfying the following conditions:

i) $\Phi^+ = \{\phi_1, \ldots, \phi_\ell\}$

ii) $\Phi^- = \Phi \setminus \{\phi_1, \ldots, \phi_\ell\}$;

iii) $\theta = \tau$;

iv) $(\phi_1, \ldots, \phi_\ell) \in A$;

v) $\Phi^- \in B$.

Clearly, if such a ℓ-tuple does exist, it is necessarily unique and equal to $(\varphi_1^+, \ldots, \varphi_\ell^+)$. We deduce that the probability of interest can be written in the form

$$\begin{aligned} & \mathbb{P}[\theta = \tau, (\varphi_1^+, \ldots, \varphi_\ell^+) \in A, \Phi^- \in B] \\ = \quad & \mathbb{E}\left[\sum_{(\phi_1, \cdots, \phi_\ell) \in \Phi^\ell} F(\phi_1, \ldots, \phi_\ell, \Phi \setminus \{\phi_1, \ldots, \phi_\ell\})\right] \end{aligned} \tag{1.3}$$

with $F : \mathcal{C}^\ell \times \mathcal{M}_p(\mathcal{C}) \to \{0, 1\}$ defined by

$$F(\phi_1, \ldots, \phi_\ell, \Phi \setminus \{\phi_1, \ldots, \phi_\ell\}) = \begin{cases} 1 & \text{if conditions i)-v) are satisfied} \\ 0 & \text{otherwise} \end{cases}.$$

By the Slyvniak-Mecke formula (Stoyan et al., 1987), we deduce that the expectation (1.3) is equal to

$$\int_{\mathcal{C}^\ell} \mathbb{E}[F(f_1, \ldots, f_\ell, \Phi)\Lambda(\mathrm{d}f_1) \cdots \Lambda(\mathrm{d}f_\ell).$$

The theorem follows after a more explicit description of the functional F. Condition i), ii) and iii) together are equivalent to

$$\max_{j' \neq j} \phi_{j'}(\mathbf{x}_{\tau_j}) < \phi_j(\mathbf{x}_{\tau_j}),\ j = 1, \ldots, \ell,$$

and

$$\forall \phi \in \Phi \setminus \{\phi_1, \ldots, \phi_\ell\},\ \phi(\mathbf{x}) < \max_{1 \leq j \leq \ell} \phi_j(\mathbf{x}).$$

Condition iv) is clear and condition v) is equivalent to $\Phi \setminus \{\phi_1, \ldots, \phi_\ell\} \in B$. □

1.2.2 The general structure of conditional distributions

In the previous section, we have seen that we are able to compute the joint distribution of the hitting scenario θ, the extremal functions Φ^+ and the subextremal functions Φ^-. We now consider the conditional joint distribution given the observations (1.1). It can be computed on the basis of Theorem 1.2.4 only, because the observations $(Z(x_1), \ldots, Z(x_k))$ can be expressed as a function of θ and Φ^+, i.e.

$$Z(x_i) = \varphi_j^+(x_i) \quad \text{with } j \text{ such that } x_i \in \theta_j.$$

Nevertheless, this computation is rather tedious and we give here the result without proof and we consider only the so-called regular case. We say that the intensity measure Λ of the point process Φ is regular at some point $\mathbf{s} = (s_1, \ldots, s_p) \in \mathcal{X}^p$ if the marginal spectral measure Λ is absolutely regular with respect to the Lebesgue measure on $(0, +\infty)^k$. More precisely, we define the marginal spectral measure

$$\Lambda_{\mathbf{s}}(\mathrm{d}z_1, \ldots, \mathrm{d}z_p) = \Lambda(f(s_1) \in \mathrm{d}z_1, \ldots, f(s_p) \in \mathrm{d}z_p)$$

and assume that on $(0, +\infty)^k$

$$\Lambda_{\mathbf{s}}(\mathrm{d}z_1, \ldots, \mathrm{d}z_p) = \lambda_{\mathbf{s}}(z_1, \ldots, z_p)\mathrm{d}z_1 \cdots \mathrm{d}z_p.$$

The proof of the following result and more details are to be found in Dombry and Eyi-Minko (2013). Note that, in the non-regular case, some further analysis of the different hitting scenarios is needed, as some hitting scenarios have to be excluded due to the conditioning data. For examples of non-regular cases and conditional sampling procedures in these cases, see Wang and Stoev (2011) and Oesting and Schlather (2014) who consider max-linear and mixed moving maxima processes, respectively.

Theorem 1.2.5. *Assume that Λ is regular at $\mathbf{x} = (x_1, \ldots, x_k)$. For $\tau = (\tau_1, \ldots, \tau_\ell) \in \mathcal{P}_k$ and $j = 1, \ldots, \ell$, define $I_j = \{i\colon x_i \in \tau_j\}$, $\mathbf{x}_{\tau_j} = (x_i)_{i \in I_j}$, $\mathbf{z}_{\tau_j} = (z_i)_{i \in I_j}$, $\mathbf{x}_{\tau_j^c} = (x_i)_{i \notin I_j}$ and $\mathbf{z}_{\tau_j^c} = (z_i)_{i \notin I_j}$. The conditional distribution of (θ, Φ^+, Φ^-) with respect to the observations* (1.1) *is obtained as follows:*

1. *For $\tau = (\tau_1, \ldots, \tau_\ell) \in \mathcal{P}_k$,*

$$\begin{aligned} & \mathbb{P}\left[\theta = \tau \mid Z(\mathbf{x}) = \mathbf{z}\right] \\ = \; & \frac{1}{C_{\mathbf{x},\mathbf{z}}} \prod_{j=1}^{\ell} \int_{\{\mathbf{u}_j < \mathbf{z}_{\tau_j^c}\}} \lambda_{(\mathbf{x}_{\tau_j}, \mathbf{x}_{\tau_j^c})}(\mathbf{z}_{\tau_j}, \mathbf{u}_j) d\mathbf{u}_j, \end{aligned} \tag{1.4}$$

   where the normalization constant $C_{\mathbf{x},\mathbf{z}}$ is such that

$$\sum_{\tau \in \mathcal{P}_k} \mathbb{P}\left[\theta = \tau \mid Z(\mathbf{x}) = \mathbf{z}\right] = 1.$$

2. *Conditionally on the observations* (1.1) *and on the hitting scenario $\tau = (\tau_1, \ldots, \tau_\ell) \in \mathcal{P}_k$, the extremal functions $\varphi_1^+, \ldots, \varphi_\ell^+$ are independent with distribution*

$$\begin{aligned} & \mathbb{P}[\varphi_j^+ \in df \mid Z(\mathbf{x}) = \mathbf{z}, \theta = \tau] \\ = \; & \Lambda[df \mid f(\mathbf{x}_{\tau_j}) = \mathbf{z}_{\tau_j}, f(\mathbf{x}_{\tau_j^c}) < \mathbf{z}_{\tau_j^c}], \quad j = 1, \ldots, \ell. \end{aligned} \tag{1.5}$$

3. *Conditionally on the observations* (1.1), *Φ^- is independent of $\Phi^+ = \{\varphi_1^+, \ldots, \varphi_\ell^+\}$ and has the same distribution as a Poisson Point process on $\mathcal{C}$ with intensity $1_{\{f(\mathbf{x}) < \mathbf{z}\}}\Lambda(df)$.*

This theorem allows to reconstruct the conditional distribution of the point process Φ given the observations $Z(\mathbf{x}) = \mathbf{z}$ via a three step procedure: construct first the conditional hitting scenario θ and the extremal functions $\varphi_1^+, \ldots, \varphi_\ell^+$ (step 1 and 2), this yields Φ^+; independently construct the conditional subextremal point process Φ^- (step 3); finally set $\Phi = \Phi^+ \cup \Phi^-$ to obtain the conditional point process Φ.

The conditional probability appearing in steps 2 and 3 are quite natural. Given the observations $Z(\mathbf{x}) = \mathbf{z}$ and the hitting scenario $\theta = \tau$, the extremal function φ_j^+ must satisfy the constraints $\varphi_j^+(\mathbf{x}_{\tau_j}) = \mathbf{z}_{\tau_j}$ and $\varphi_j^+(\mathbf{x}_{\tau_j^c}) < \mathbf{z}_{\tau_j^c}$ so that it seems natural to obtain the conditional intensity Λ given these constraints. Similarly, given $Z(\mathbf{x}) = \mathbf{z}$, the subextremal functions $\varphi \in \Phi^-$ must satisfy the constraint $\varphi(\mathbf{x}) < \mathbf{z}$ which naturally leads to the point process intensity Λ restricted to the set of functions $\{f \in \mathcal{C}, f(\mathbf{x}) < \mathbf{z}\}$.

1.2.3 Examples of regular models

Theorem 1.2.5 requires the assumption of regularity of the intensity measure Λ at point $\mathbf{x} = (x_1, \ldots, x_k)$. We now recall some popular models for max-stable process that are regular and give the corresponding intensity function $\lambda_{\mathbf{x}}$. These models are mainly based on Gaussian or related distributions and provide a nice framework for explicit computations.

Example 1: Brown-Resnick process
The Brown–Resnick process introduced by Kabluchko et al. (2009) is a stationary max-stable random field on $\mathcal{X} = \mathbb{R}^d$ corresponding to the case where $Y(x) = \exp\{W(x) - \gamma(x)\}$ in (1.2) with W a centered Gaussian process with stationary increments, semi-variogram γ and such that $W(o) = 0$ almost surely. One can show that the intensity measure Λ is regular at $\mathbf{x} \in \mathcal{X}^k$ as long as the covariance matrix $\Sigma_{\mathbf{x}}$ of the random vector $W(\mathbf{x})$ is positive definite. We denote by $g_{\mathbf{x}}$ the Gaussian density

$$g_{\mathbf{x}}(\mathbf{u}) = (2\pi)^{-k/2}\det(\Sigma_{\mathbf{x}})^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{u}^T\Sigma_{\mathbf{x}}^{-1}\mathbf{u}\right\}.$$

The marginal intensity measure is computed as follows: for all Borel set $A \subset (0, +\infty)^k$

$$\begin{aligned}
\Lambda_{\mathbf{x}}(A) &= \int_0^\infty \mathbb{P}[\zeta Y(\mathbf{x}) \in A]\zeta^{-2}\mathrm{d}\zeta \\
&= \int_0^\infty \mathbb{P}[\zeta \exp\{W(\mathbf{x}) - \gamma(\mathbf{x})\} \in A]\zeta^{-2}\mathrm{d}\zeta \\
&= \int_0^\infty \int_A g_{\mathbf{x}}(\log \mathbf{z} - \log\zeta + \gamma(\mathbf{x}))\prod_{i=1}^k z_i^{-1}\zeta^{-2}\mathrm{d}\zeta\mathrm{d}\mathbf{z} \\
&= \int_A \lambda_{\mathbf{x}}(\mathbf{z})\mathrm{d}\mathbf{z}
\end{aligned}$$

with

$$\lambda_{\mathbf{x}}(\mathbf{z}) = \int_0^\infty g_{\mathbf{x}}(\log \mathbf{z} - \log \zeta + \gamma(\mathbf{x})) \prod_{i=1}^k z_i^{-1} \zeta^{-2} \mathrm{d}\zeta.$$

Some standard but tedious computations for Gaussian integrals reveal that

$$\lambda_{\mathbf{x}}(\mathbf{z}) = C_{\mathbf{x}} \exp\left(-\frac{1}{2} \log \mathbf{z}^T Q_{\mathbf{x}} \log \mathbf{z} + L_{\mathbf{x}} \log \mathbf{z}\right) \prod_{i=1}^k z_i^{-1}, \ \ \mathbf{z} \in (0,\infty)^k,$$

with $1_k = (1)_{i=1,\ldots,k}$, $\gamma_{\mathbf{x}} = \{\gamma(x_i)\}_{i=1,\ldots,k}$,

$$\begin{aligned} Q_{\mathbf{x}} &= \Sigma_{\mathbf{x}}^{-1} - \frac{\Sigma_{\mathbf{x}}^{-1} 1_k 1_k^T \Sigma_{\mathbf{x}}^{-1}}{1_k^T \Sigma_{\mathbf{x}}^{-1} 1_k}, \qquad L_{\mathbf{x}} = \left(\frac{1_k^T \Sigma_{\mathbf{x}}^{-1} \gamma_{\mathbf{x}} - 1}{1_k^T \Sigma_{\mathbf{x}}^{-1} 1_k} 1_k - \gamma_{\mathbf{x}}\right)^T \Sigma_{\mathbf{x}}^{-1}, \\ C_{\mathbf{x}} &= (2\pi)^{(1-k)/2} \det(\Sigma_{\mathbf{x}})^{-1/2} (1_k^T \Sigma_{\mathbf{x}}^{-1} 1_k)^{-1/2} \cdots \\ &\quad \cdots \exp\left\{\frac{1}{2} \frac{(1_k^T \Sigma_{\mathbf{x}}^{-1} \gamma_{\mathbf{x}} - 1)^2}{1_k^T \Sigma_{\mathbf{x}}^{-1} 1_k} - \frac{1}{2} \gamma_{\mathbf{x}}^T \Sigma_{\mathbf{x}}^{-1} \gamma_{\mathbf{x}}\right\}. \end{aligned}$$

See Dombry et al. (2013) for more details.

Example 2: Schlather process
The Schlather process (Schlather, 2002), also called extremal Gaussian process, is a max-stable random field on $\mathcal{X} = \mathbb{R}^d$ corresponding to the case where $Y(x) = (2\pi)^{1/2} \max\{0, W(x)\}$ in (1.2) with W a standard Gaussian process with correlation function ρ. For $\mathbf{x} \in \mathcal{X}^k$ and provided the covariance matrix Σ_x of the random vector $W(\mathbf{x})$ is positive definite, the intensity function is

$$\lambda_{\mathbf{x}}(\mathbf{z}) = \pi^{-(k-1)/2} \det(\Sigma_{\mathbf{x}})^{-1/2} a_{\mathbf{x}}(\mathbf{z})^{-(k+1)/2} \Gamma\left(\frac{k+1}{2}\right), \qquad \mathbf{z} \in (0,+\infty)^k,$$

where $a_{\mathbf{x}}(\mathbf{z}) = \mathbf{z}^T \Sigma_{\mathbf{x}}^{-1} \mathbf{z}$. See Dombry et al. (2013) for more details on the computation.

Example 3: extremal t-process
The extremal t-process is a generalization of the Schlather process above obtained with an extra parameter $\nu > 0$ that yields more flexibility in the model. It is a max-stable random field on $\mathcal{X} = \mathbb{R}^d$ obtained with $Y(x) = c_\nu \max\{0, W(x)^\nu\}$ in (1.2) with W a standard Gaussian process with correlation function ρ and $c_\nu = \sqrt{\pi} 2^{-(\nu-2)/2} \Gamma((\nu+1)/2)^{-1}$. Provided the covariance matrix Σ_x of the random vector $W(\mathbf{x})$ is positive definite, the intensity function is

$$\begin{aligned} \lambda_{\mathbf{x}}(\mathbf{z}) \quad = \quad & c_\nu \nu^{-k+1} 2^{(\nu-2)/2} \pi^{-k/2} \det(\Sigma_{\mathbf{x}})^{-1/2} a_{\mathbf{x}}(\mathbf{z}, \nu)^{-(k+\nu)/2} \cdots \\ & \cdots \Gamma\left(\frac{k+\nu}{2}\right) \prod_{j=1}^k z_j^{(1-\nu)/\nu} \end{aligned}$$

for $\mathbf{z} \in (0,+\infty)^k$ and with $a_{\mathbf{x}}(\mathbf{z}, \nu) = (\mathbf{z}^{1/\nu})^T \Sigma_{\mathbf{x}}^{-1} \mathbf{z}^{1/\nu}$. See Ribatet (2013).

1.2.4 Distribution of the extremal functions

The distribution of the extremal function appearing in Equation (1.5) is still quite theoretical at this stage and it is unclear how one can sample from this distribution. There are also theoretical questions related to its definition: the intensity measure Λ has an infinite total mass and the conditioning event $\{f(\mathbf{x}_{\tau_j}) = \mathbf{z}_{\tau_j}, f(\mathbf{x}_{\tau_j^c}) < \mathbf{z}_{\tau_j^c}\}$ has zero measure.

A way to bypass these difficulties is to consider only the finite dimensional margins of the extremal functions φ_j^+ and not the full random process. In practice, this is enough for simulation purpose since one simulate the random field on a finite grid only and not on the whole state space $\mathcal{X}$. Let $\mathbf{s} = (s_1, \ldots, s_m) \in \mathcal{X}^m$ be the set of new locations for the conditional sampling, we focus on the conditional distribution of $Z(\mathbf{s}) \mid Z(\mathbf{x}) = \mathbf{z}$. Equation (1.5) can be simplified if we assume the regularity of the intensity measure Λ at $(\mathbf{x}, \mathbf{s}) \in \mathcal{X}^{k+m}$. We can then introduce the conditional intensity function

$$\lambda_{\mathbf{s}|\mathbf{x},\mathbf{z}}(\mathbf{u}) = \frac{\lambda_{(\mathbf{s},\mathbf{x})}(\mathbf{u}, \mathbf{z})}{\lambda_{\mathbf{x}}(\mathbf{z})}, \quad \mathbf{u} \in (0, \infty)^m.$$

Equation (1.5) can be rewritten as

$$\begin{aligned} & \mathbb{P}\left[\varphi_j^+(\mathbf{s}) \in \mathrm{d}\mathbf{v} \mid Z(\mathbf{x}) = \mathbf{z}, \theta = \tau\right] \\ = \quad & \frac{1}{C_j}\left(\int 1_{\{\mathbf{u_j} < \mathbf{z}_{\tau_j^c}\}} \lambda_{(\mathbf{s},\mathbf{x}_{\tau_j^c})|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{v}, \mathbf{u_j}) \mathrm{d}\mathbf{u_j}\right) \mathrm{d}\mathbf{v} \end{aligned} \tag{1.6}$$

with C_j the normalization constant

$$C_j = \int 1_{\{\mathbf{u_j} < \mathbf{z}_{\tau_j^c}\}} \lambda_{(\mathbf{s},\mathbf{x}_{\tau_j^c})|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{v}, \mathbf{u_j}) \mathrm{d}\mathbf{u_j} \mathrm{d}\mathbf{v}.$$

In words, the conditional law of $\varphi_j^+(\mathbf{s})$ is equal to the distribution of the random variable $\mathbf{V}$ obtained as the first component of $(\mathbf{V}, \mathbf{U_j})$ with density $\lambda_{(\mathbf{s},\mathbf{x}_{\tau_j^c})|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{v}, \mathbf{u})$ conditioned to the event $\mathbf{U_j} < \mathbf{z}_{\tau_j^c}$.

Example 1 continued: Brown-Resnick process
In the case of Brown-Resnick process, for all $(\mathbf{s}, \mathbf{x}) \in \mathcal{X}^{m+k}$, $(\mathbf{u}, \mathbf{z}) \in (0, \infty)^{m+k}$ and provided the covariance matrix $\Sigma_{(\mathbf{s},\mathbf{x})}$ is positive definite, the conditional intensity function corresponds to a multivariate log-normal probability density function

$$\begin{aligned} \lambda_{\mathbf{s}|\mathbf{x},\mathbf{z}}(\mathbf{u}) \quad = \quad & (2\pi)^{-m/2} \det(\Sigma_{\mathbf{s}|\mathbf{x}})^{-1/2} \cdots \\ & \cdots \exp\left\{-\frac{1}{2}(\log \mathbf{u} - \mu_{\mathbf{s}|\mathbf{x},\mathbf{z}})^T \Sigma_{\mathbf{s}|\mathbf{x}}^{-1} (\log \mathbf{u} - \mu_{\mathbf{s}|\mathbf{x},\mathbf{z}})\right\} \prod_{i=1}^m u_i^{-1}, \end{aligned}$$

where $\mu_{\mathbf{s}|\mathbf{x},\mathbf{z}} \in \mathbb{R}^m$ and $\Sigma_{\mathbf{s}|\mathbf{x}}$ are the mean and covariance matrix of the underlying normal distribution and are given by

$$\begin{aligned} \Sigma_{\mathbf{s}|\mathbf{x}}^{-1} \quad &= \quad J_{m,k}^T Q_{(\mathbf{s},\mathbf{x})} J_{m,k} \\ \mu_{\mathbf{s}|\mathbf{x},\mathbf{z}} \quad &= \quad \left\{L_{(\mathbf{s},\mathbf{x})} J_{m,k} - \log \mathbf{z}^T \tilde{J}_{m,k}^T Q_{(\mathbf{s},\mathbf{x})} J_{m,k}\right\} \Sigma_{\mathbf{s}|\mathbf{x}}, \end{aligned}$$

with

$$J_{m,k} = \begin{bmatrix} \mathrm{Id}_m \\ 0_{k,m} \end{bmatrix}, \qquad \tilde{J}_{m,k} = \begin{bmatrix} 0_{m,k} \\ \mathrm{Id}_k \end{bmatrix},$$

where Id_k denotes the $k \times k$ identity matrix and $0_{m,k}$ the $m \times k$ null matrix.

Example 2 continued: Schlather process
For $(\mathbf{s}, \mathbf{x}) \in \mathcal{X}^{m+k}$, $(\mathbf{u}, \mathbf{z}) \in \mathbb{R}^{m+k}$ and provided that the covariance matrix $\Sigma_{(\mathbf{s},\mathbf{x})}$ is positive definite, the conditional intensity function $\lambda_{\mathbf{s}|\mathbf{x},\mathbf{z}}(\mathbf{u})$ corresponds to a multivariate Student distribution

$$\begin{aligned} \lambda_{\mathbf{s}|\mathbf{x},\mathbf{z}}(\mathbf{u}) &= \pi^{-m/2}(k+1)^{-m/2}\det(\tilde{\Sigma})^{-1/2}\cdots \\ &\quad \cdots \left\{1 + \frac{(\mathbf{u}-\mu)^T\tilde{\Sigma}^{-1}(\mathbf{u}-\mu)}{k+1}\right\}^{-(m+k+1)/2} \frac{\Gamma\left(\frac{m+k+1}{2}\right)}{\Gamma\left(\frac{k+1}{2}\right)}, \end{aligned}$$

with $k+1$ degrees of freedom, location parameter $\mu = \Sigma_{\mathbf{s}:\mathbf{x}}\Sigma_{\mathbf{x}}^{-1}\mathbf{z}$ and scale matrix

$$\tilde{\Sigma} = \frac{a_{\mathbf{x}}(\mathbf{z})}{k+1}\left(\Sigma_{\mathbf{s}} - \Sigma_{\mathbf{s}:\mathbf{x}}\Sigma_{\mathbf{x}}^{-1}\Sigma_{\mathbf{x}:\mathbf{s}}\right)$$

where

$$\Sigma_{(\mathbf{s},\mathbf{x})} = \begin{bmatrix} \Sigma_{\mathbf{s}} & \Sigma_{\mathbf{s}:\mathbf{x}} \\ \Sigma_{\mathbf{x}:\mathbf{s}} & \Sigma_{\mathbf{x}} \end{bmatrix}.$$

Example 3 continued: extremal t-process
The results are similar and generalize the case of Schlather processes. For $(\mathbf{s}, \mathbf{x}) \in \mathcal{X}^{m+k}$, $(\mathbf{u}, \mathbf{z}) \in \mathbb{R}^{m+k}$ we set $\mu = \Sigma_{\mathbf{s}:\mathbf{x}}\Sigma_{\mathbf{x}}^{-1}\mathbf{z}^{1/\nu}$ and

$$\tilde{\Sigma} = (k+\nu)^{-1}a_{\mathbf{x}}(\mathbf{z},\nu)(\Sigma_{\mathbf{s}} - \Sigma_{\mathbf{s}:\mathbf{x}}\Sigma_{\mathbf{x}}^{-1}\Sigma_{\mathbf{x}:\mathbf{s}})$$

Then we have

$$\begin{aligned} \lambda_{\mathbf{s}|\mathbf{x},\mathbf{z}}(\mathbf{u}) &= \pi^{-m/2}(k+\nu)^{-m/2}\det(\tilde{\Sigma})^{-1/2}\cdots \\ &\quad \cdots \left\{1 + \frac{(\mathbf{u}^{1/\nu}-\mu)^T\tilde{\Sigma}^{-1}(\mathbf{u}^{1/\nu}-\mu)}{k+\nu}\right\}^{-(m+k+\nu)/2} \cdots \\ &\quad \cdots \frac{\Gamma\left(\frac{m+k+\nu}{2}\right)}{\Gamma\left(\frac{k+\nu}{2}\right)} \left\{\nu^{-m}\prod_{j=1}^{m} u_j^{-(\nu-1)/\nu}\right\}. \end{aligned}$$

The last term in bracket in the previous equation corresponds to the Jacobian of the mapping $\mathbf{u} \mapsto \mathbf{u}^{1/\nu}$. Hence we recognize that the conditional intensity function is the density of the random vector T^ν where T is a Student random vector with $k+\nu$ degrees of freedom, mean μ and scale matrix $\tilde{\Sigma}$.

1.3 Conditional sampling of max-stable processes

We now consider the conditional sampling of max-stable processes following Dombry et al. (2013). We present a 3-step sampling procedure that is quite straightforwardly derived from Theorem 1.2.5. We particularly focus on the first step and discuss a Gibbs sampling approach for the conditional hitting scenario.

1.3.1 A 3-step procedure for conditional sampling

Theorem 1.2.5 provides the conditional distribution of the point process Φ given the observation (1.1). In practice, we are rather interested in the conditional simulation of the max-stable process $Z(\mathbf{s})$ at some location $\mathbf{s} = (s_1, \ldots, s_m) \in \mathcal{X}^m$. The connection is rather straightforward and is made explicit in the next proposition.

Proposition 1.3.1. *Assume that Λ is regular at $(\mathbf{x}, \mathbf{s}) \in \mathcal{X}^{k+m}$. Consider the three–step procedure:*

1. *Draw a random partition $\theta \in \mathcal{P}_k$ with distribution* (1.4)*;*
2. *Given $\theta = (\tau_1, \ldots, \tau_\ell)$, draw ℓ independent random vectors $\varphi_j^+(\mathbf{s}), \ldots, \varphi_\ell^+(\mathbf{s})$ with distribution* (1.6) *and define the random vector*

$$Z^+(\mathbf{s}) = \max_{j=1,\ldots,\ell} \varphi_j^+(\mathbf{s}).$$

3. *Independently draw a Poisson point process $\{\zeta_i\}_{i\geq 1}$ on $(0, \infty)$ with intensity $\zeta^{-2}d\zeta$ and $\{Y_i\}_{i\geq 1}$ independent copies of Y, and define the random vector*

$$Z^-(\mathbf{s}) = \max_{i\geq 1} \zeta_i Y_i(\mathbf{s}) 1_{\{\zeta_i Y_i(\mathbf{x}) < \mathbf{z}\}}.$$

Then the random vector $\tilde{Z}(\mathbf{s}) = \max\{Z^+(\mathbf{s}), Z^-(\mathbf{s})\}$ follows the conditional distribution of $Z(s)$ given $Z(x) = z$.

This 3-step procedure — which, in general, can also be applied in the non-regular case — is illustrated on Figure 1.3.1 on the simple case of 1-D Smith storm process with Gaussian shape.

In step 3, it is worth noting that $Z^-(\mathbf{s})$ follows the conditional distribution of $Z(\mathbf{s})$ given $Z(\mathbf{x}) < \mathbf{z}$. It is not difficult to show that the conditional cumulative distribution function of $Z(\mathbf{s})$ given $Z(\mathbf{x}) = \mathbf{z}$ is given by

$$\mathbb{P}[Z(\mathbf{s}) \leq \mathbf{a} \mid Z(\mathbf{x}) = \mathbf{z}] = \mathbb{P}[Z(\mathbf{s}) \leq \mathbf{a} \mid Z(\mathbf{x}) < \mathbf{z}] \sum_{\tau \in \mathcal{P}_k} \pi_{\mathbf{x},\mathbf{z}}(\tau) \prod_{j=1}^{\ell} F_{\tau,j}(\mathbf{a}),$$

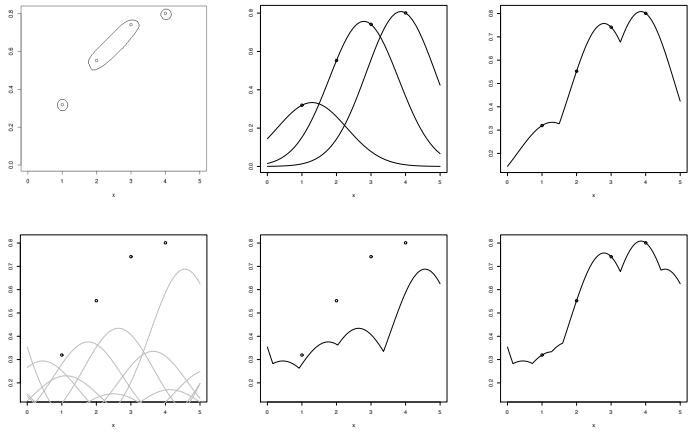

FIGURE 1.2
The 3-step procedure for simulation sampling. First line: construction of the conditional hitting scenario (step 1, left), of the extremal functions (step 2, middle) and of the associated process Z^+ (right). Second line: construction of the subextremal functions (step 3, left), of the associated process Z^- (middle) and of the conditioned max-stable process $\tilde{Z} = \max(Z^+, Z^-)$ (right).

where $\pi_{\mathbf{x},\mathbf{z}}(\tau) = \mathbb{P}[\theta = \tau \mid Z(\mathbf{x}) = \mathbf{z}]$ is given by Equation (1.4) and

$$\begin{aligned} F_{\tau,j}(\mathbf{a}) &= \mathbb{P}[\varphi_j^+(\mathbf{s}) \leq \mathbf{a} \mid Z(\mathbf{x}) = \mathbf{z},\, \theta = \tau] \\ &= \frac{\int_{\{\mathbf{u_j} < \mathbf{z}_{\tau_j^c}, \mathbf{v} < \mathbf{a}\}} \lambda_{(\mathbf{s},\mathbf{x}_{\tau_j^c})|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{v},\mathbf{u_j})\mathrm{d}\mathbf{u_j}\mathrm{d}\mathbf{v}}{\int_{\{\mathbf{u_j} < \mathbf{z}_{\tau_j^c}\}} \lambda_{\mathbf{x}_{\tau_j^c}|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{u_j})\mathrm{d}\mathbf{u_j}}. \end{aligned}$$

1.3.2 Gibbs sampling for the conditional hitting scenario

In the above 3-step procedure for conditional sampling, step 2 requires to sample the extremal functions. As we have seen in Examples 1, 2 and 3 where log-normal and Student distributions appear, these can be of a relatively simple structure although the imposed equality and inequality constraints may cause additional difficulties. Step 3 can be performed by an unconditional simulation of the max-stable process Z according to an appropriate spectral representation rejecting all those functions that do not respect the constraints given by the conditioning data. For details on unconditional simulation of max-stable processes we refer to corresponding chapter. However, the conditional sampling of the hitting scenario (step 1) remains the most challenging step. According to Equation (1.4), it is given by

$$\pi_{\mathbf{x},\mathbf{z}}(\tau) = \mathbb{P}[\theta = \tau \mid Z(\mathbf{x}) = \mathbf{z}] = \frac{1}{C_{\mathbf{x},\mathbf{z}}} \prod_{j=1}^{\ell} \omega_{\tau_j}$$

where

$$\omega_{\tau_j} = \lambda_{\mathbf{x}_{\tau_j}}(\mathbf{z}_{\tau_j}) \int_{\{\mathbf{u}_j < \mathbf{z}_{\tau_j^c}\}} \lambda_{\mathbf{x}_{\tau_j^c}|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}(\mathbf{u}_j)\mathrm{d}\mathbf{u}_j.$$

The last integral is the multivariate cumulative distribution of the conditional intensity function $\lambda_{\mathbf{x}^c_{\tau_j}|\mathbf{x}_{\tau_j},\mathbf{z}_{\tau_j}}$. Hence this is a multivariate log-normal cdf in the Brown-Resnick model and a multivariate Student cdf in the Schlather or extremal-t models. In the following, we assume that the weights ω_{τ_j} can be accurately computed numerically and we focus on how to sample from $\pi_{\mathbf{x},\mathbf{z}}$, a discrete probability measure on the set $\mathcal{P}_k$ of partition of $\{x_1, \cdots, x_k\}$. The main difficulty is that the norming constant $C_{\mathbf{x},\mathbf{z}}$ is unknown and a naive computation of this constant requires the computation of the weight $\prod_{j=1}^{\ell} \omega_{\tau_j}$ for all $\tau \in \mathcal{P}_k$. The number of terms, or equivalently the cardinality of $\mathcal{P}_k$ is given by the so-called Bell number B_k. The first 10 Bell numbers are

k	1	2	3	4	5	6	7	8	9	10
B_k	1	2	5	15	52	203	877	4140	21147	115975

Hence determining the discrete probability $\pi_{\mathbf{x},\mathbf{z}}$ requires the computation of 52 weights for $k = 5$ and 115975 for $k = 10$. Even worse, $B_{15} \approx 1.4\,10^9$ and $B_{20} \approx 5.2\,10^{13}$ so that the storage of all partitions is beyond the memory capacities of a standard computer. This is the so called phenomenon of combinatorial explosion and we need an alternative method to the naive discrete enumeration of all possibilities.

It is customary to use Monte Carlo Markov Chain sampling when the target distribution is known up to a multiplicative constant only. The aim is to construct a Markov chain with stationary distribution equal to the target distribution and with good mixing properties. Two main methods exist: the Metropolis-Hasting algorithm and the Gibbs sampler. We focus here on this second option.

For $\tau \in \mathcal{P}_k$ and $j \in \{1, \ldots, k\}$, let τ_{-j} be the restriction of τ to the set $\{x_1, \ldots, x_k\}$. As usual with Gibbs samplers, our goal is to simulate from

$$\mathbb{P}[\theta \in \cdot \mid \theta_{-j} = \tau_{-j}], \tag{1.7}$$

where $\theta \in \mathcal{P}_k$ is a random partition which follows the target distribution $\pi_{\mathbf{x},\mathbf{z}}(\cdot)$ and τ is typically the current state of the Markov chain. It is easy to see that the number of possible updates according to (1.7) is always less than k, so that the combinatorial explosion is avoided. Indeed, the point x_j can be reallocated to any of the components of τ_{-j} or to a new component with a single point. We deduce that the number of possible updates $\tau^* \in \mathcal{P}_k$ such that $\tau^*_{-j} = \tau_{-j}$ is

$$b^+ = \begin{cases} \ell & \text{if } \{x_j\} \text{ is a partitioning set of } \tau, \\ \ell + 1 & \text{if } \{x_j\} \text{ is not a partitioning set of } \tau, \end{cases}$$

For illustration, consider the set $\{x_1, x_2, x_3\}$ and let $\tau = (\{x_1, x_2\}, \{x_3\})$. Then the possible partitions τ^* such that $\tau^*_{-2} = \tau_{-2}$ are

$$(\{x_1, x_2\}, \{x_3\}), \qquad (\{x_1\}, \{x_2\}, \{x_3\}), \qquad (\{x_1\}, \{x_2, x_3\}), \tag{1.8}$$

while there exists only two partitions such that $\tau^*_{-3} = \tau_{-3}$, i.e.,

$$(\{x_1, x_2\}, \{x_3\}), \qquad (\{x_1, x_2, x_3\}).$$

For our work we use a random scan implementation of the Gibbs sampler (Liu et al., 1995), meaning that one iteration of the Gibbs sampler selects randomly an element of $j \in \{1, \ldots, k\}$ and then updates the current state τ according to the proposal distribution (1.7). For the sake of simplicity, we use the uniform random scan, i.e. j is selected according to the uniform distribution on $\{1, \ldots, k\}$. Figure 1.3.2 shows two successive iterations of this random scan Gibbs sampler.

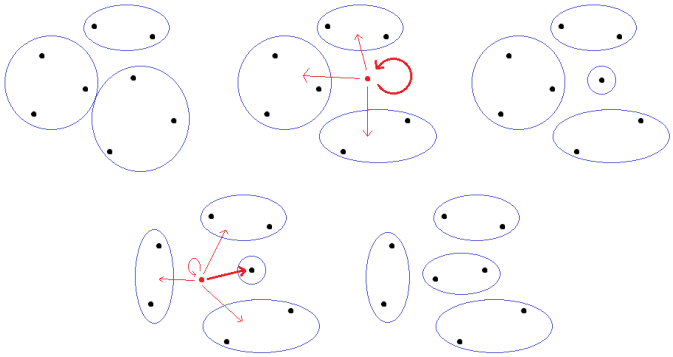

FIGURE 1.3
Two successive iterations of the Gibbs sampler for the conditional hitting scenario with $k = 8$ conditioning points: after choosing a random point, the arrows show the different possible reallocations, the bold arrow representing the chosen one.

The distribution (1.7) has nice properties. For all $\tau^* \in \mathscr{P}_k$ with $\tau^*_{-j} = \tau_{-j}$ we have

$$\mathbb{P}[\theta = \tau^* \mid \theta_{-j} = \tau_{-j}] = \frac{\pi_{\mathbf{x},\mathbf{z}}(\tau^*)}{\sum_{\tilde{\tau} \in \mathcal{P}_k} \pi_{\mathbf{x},\mathbf{z}}(\tilde{\tau}) 1_{\{\tilde{\tau}_{-j} = \tau_{-j}\}}} \propto \frac{\prod_{k=1}^{|\tau^*|} w_{\tau^*_k}}{\prod_{k=1}^{|\tau|} w_{\tau_k}}. \tag{1.9}$$

Since, τ and τ^* share many components, it can be seen that many factors in the right-hand side of (1.9) cancel out except at most four of them. In the previous example (1.8), the corresponding weights are

$$1, \quad \frac{w_{\{x_1,x_2\}}}{w_{\{x_1\}} w_{\{x_2\}}}, \quad \frac{w_{\{x_1,x_2\}} w_{\{x_3\}}}{w_{\{x_1\}} w_{\{x_2,x_3\}}} \qquad \text{respectively.}$$

This makes the Gibbs sampler especially convenient.

We finally provide some practical details on the computation of these conditional weights as well as a detailed algorithm (see Algorithm 1 in the supplementary material).

We first describe how each partition of $\{x_1, \ldots, x_k\}$ is stored. To illustrate consider the set $\{x_1, x_2, x_3\}$ and the partition $(\{x_1, x_2\}, \{x_3\})$. With our convention, this partition is defined as $(1, 1, 2)$ indicating that x_1 and x_2 belong to the same partitioning set labeled '1' and x_3 belongs to the partitioning set '2'. There exist several equivalent notations for this partition: for example one can use $(2, 2, 1)$ or $(1, 1, 3)$. However there is a one-one mapping between $\mathscr{P}_k$ and the set

$$\mathcal{P}_k^* = \left\{ (a_1, \ldots, a_k),\, \forall i \in \{2, \ldots, k\} \colon 1 = a_1 \leq a_i \leq \max_{1 \leq j < i} a_j + 1,\, a_i \in \mathbb{Z} \right\}.$$

Consequently we shall restrict our attention to the partitions that live in $\mathcal{P}_k^*$ and going back to our example we see that $(1, 1, 2)$ is valid but $(2, 2, 1)$ and $(1, 1, 3)$ are not.

For $\tau \in \mathcal{P}_k^*$ of size ℓ, let $r_1 = \sum_{i=1}^k \delta_{\tau_i = a_j}$ and $r_2 = \sum_{i=1}^k \delta_{\tau_i = b}$, i.e., the number of conditioning locations that belong to the partitioning sets 'a_j' and 'b' where $b \in \{1, \ldots, b^+\}$ and

$$b^+ = \begin{cases} \ell & (r_1 = 1), \\ \ell + 1 & (r_1 \neq 1). \end{cases}$$

Then the conditional probability distribution $\mathbb{P}[\tau_j = b \mid \tau_i = a_i, \forall i \neq j]$ given by Equation (1.9) is proportional to

$$\begin{cases} 1 & (b = a_j), & \text{(1.10a)} \\ w_{\tau_*, b} / (w_{\tau, b} w_{\tau, a_j}) & (r_1 = 1, r_2 \neq 0, b \neq a_j), & \text{(1.10b)} \\ w_{\tau_*, b} w_{\tau_*, a_j} / (w_{\tau, b} w_{\tau, a_j}) & (r_1 \neq 1, r_2 \neq 0, b \neq a_j), & \text{(1.10c)} \\ w_{\tau_*, b} w_{\tau_*, a_j} / w_{\tau, a_j} & (r_1 \neq 1, r_2 = 0, b \neq a_j), & \text{(1.10d)} \end{cases}$$

where $\tau_* = (a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_k)$ and $w_{\tau, a_i} = w_{\{x_k \colon \tau_k = a_i\}}$. It is worth stressing that although τ_* does not necessarily belongs to $\mathcal{P}_k^*$, it corresponds to a unique partition of $\mathcal{P}_k$ and we can use the bijection $\mathcal{P}_k \to \mathcal{P}_k^*$ to recode τ_* into an element of $\mathcal{P}_k^*$.

1.4 Simulation Study

1.4.1 Gibbs Sampler

In this section we check whether the Gibbs sampler is able to sample from $\pi_{\mathbf{x},\mathbf{z}}(\cdot)$. To illustrate a typical sample path, Figure 1 in the supplementary material shows the trace plot of a simulated chain of length 2000 with $k = 5$ conditioning locations and a thinning lag of 5 and compares the theoretical probabilities $\{\pi_{\mathbf{x},\mathbf{z}}(\tau), \tau \in \mathcal{P}_k\}$ to the empirical ones estimated from the Markov chain. As mentioned in the previous section it can be seen that only a few states have a significant probability to occur and these states most often differs only slightly. As expected for this particular simulation the empirical probabilities match the theoretical ones.

To better assess if our uniform random scan Gibbs sampler is able to sample from $\pi_{\mathbf{x},\mathbf{z}}(\cdot)$ we simulate 250 Markov chains of length 1000—after removing a burn-in period and thinning the chain, and similarly to Figure **??**, compare the theoretical probabilities $\{\pi_{\mathbf{x},\mathbf{z}}(\tau), \tau \in \mathcal{P}_k\}$ to the empirical ones using a χ^2 test for each of these chains. Since the computation of these theoretical probabilities is CPU intensive, the number of conditioning locations is at most 5 in our simulation study.

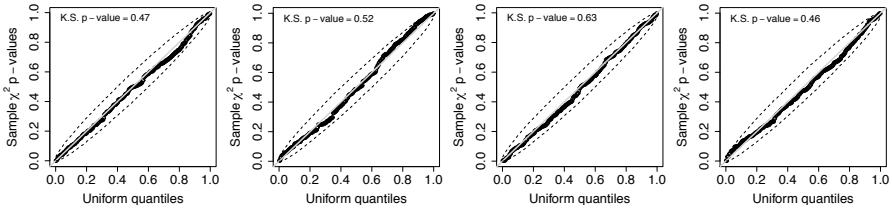

FIGURE 1.4
QQ-plots for the sample χ^2 test p-values against $U(0,1)$ quantiles with a varying number of conditioning locations k with an insert showing the p-values of a Kolmogorov Smirnov test for uniformity—from left to right, $k = 2, 3, 4$ and 5. The dashed lines show the 95% confidence envelopes.

Figure 1.4 plots the sample p-values of these χ^2 tests against the quantiles of a $U(0,1)$ distribution for a varying number of conditional locations. This figure corroborates what we saw in Figure **??** since the sample p-values seem to follow a $U(0,1)$ distribution indicating that the sampler is able to sample from $\pi_{\mathbf{x},\mathbf{z}}(\cdot)$.

1.4.2 Conditional simulations

In this section we check if our algorithm is able to produce realistic conditional simulations of Brown–Resnick processes with semi-variogram $\gamma(h) = (h/\lambda)^\nu$. In this case, the spectral process $Y(\cdot)$ in (1.2) is a fractional Brownian motion with Hurst index ν. To have a broad overview, we consider three different sample path properties as summarized below.

	Sample path properties		
	γ_1: Wiggly	γ_2: Smooth	γ_3: Very smooth
λ	25	54	69
ν	0·5	1·0	1·5

The variogram parameters are set to ensure that the extremal coefficient function satisfies $\theta(115) = 1.7$. Figure 1.5 shows one realization for each sample path configuration as well as the corresponding extremal coefficient function. These realizations will serve as the basis for our conditioning events.

In order to check if our sampling procedure is accurate, given a single conditional event $\{Z(\mathbf{x}) = \mathbf{z}\}$, we generated 1000 conditional realizations of a Brown–Resnick processes with standard Gumbel margins and semi-variograms γ_j $(j = 1, 2, 3)$.

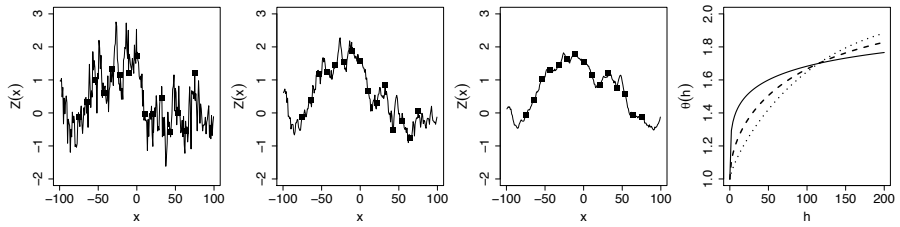

FIGURE 1.5
Three realizations of a Brown–Resnick process with standard Gumbel margins and semi-variograms γ_1, γ_2 and γ_3—from left to right. The squares correspond to the 15 conditioning values that will be used in the simulation study. The right panel shows the associated extremal coefficient functions where the solid, dashed and dotted lines correspond respectively to γ_1, γ_2 and γ_3.

Figure 1.6 shows the pointwise sample quantiles obtained from these 1000 simulated paths in comparison to unit Gumbel quantiles. As expected the conditional sample paths inherit the regularity driven by the Hurst index $\nu/2$ of the process $Y(\cdot)$ in (1.2) and there is less variability in regions close to some conditioning locations. On the opposite, although for this specific type of variogram $\theta(h) \to 2$ as $h \to \infty$, for regions far away from any conditioning location the sample quantiles converges to that of a standard Gumbel distribution indicating that the conditional event does not have any influence anymore. In addition the sample paths used to get the conditional events, see Figure 1.5, lay most of the time in the 95% pointwise confidence intervals corroborating that our sampling procedure seems to be accurate — the coverage ranges between 0·93 and 1·00 with a mean value of 0·96.

So far we have checked that the proposed sampling procedure yields the expected coverage as well as the right marginal properties as we move far away from any conditioning location. The last point to be fulfilled is to assess whether the simulation procedure honors the spatial dependence driven by the semi-variogram $\gamma(\cdot)$. To this aim we use the F-madogram (Cooley et al., 2006) to compare the pairwise extremal coefficient estimates to the theoretical extremal coefficient function. Since $Z(\cdot) \mid \{Z(\mathbf{x}) = \mathbf{z}\}$ is not max-stable the F-madogram cannot be used. However, by integrating out the conditional event we recover the original Brown-Resnick distribution and the max-stability property. So we generate independently 1000 conditional events $\{Z(\mathbf{x}) = \mathbf{z}\}$, $\mathbf{x}$ being fixed, and for each conditional event one conditional realization of a Brown–Resnick process. Figure 1.7 compares the pairwise extremal coefficient estimates based on these simulations to the theoretical extremal coefficient function. As expected, whatever the number of conditioning location and the semi-variograms are, the (binned) pairwise estimates match the theoretical curve indicating that the spatial dependence is honored.

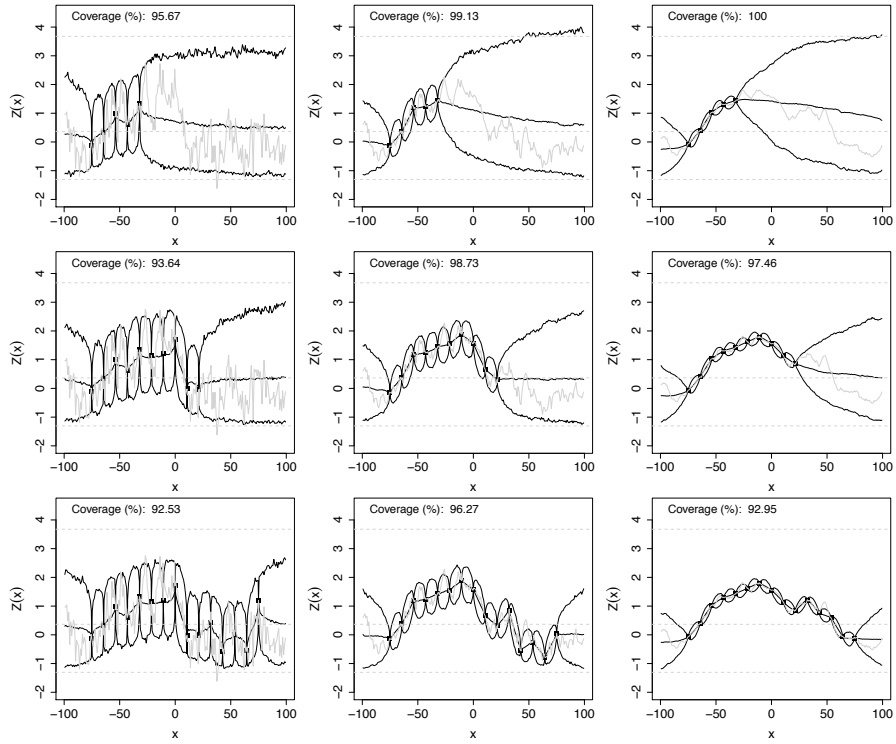

FIGURE 1.6

Pointwise sample quantiles estimated from 1000 conditional simulations of Brown–Resnick processes with standard Gumbel margins and semi-variograms γ_1, γ_2 and γ_3 (left to right) and with $k = 5, 10, 15$ conditioning locations—top to bottom. The solid black lines show the pointwise 0·025, 0·5, 0·975 sample quantiles and the dashed grey lines that of a standard Gumbel distribution. The squares show the conditional points $\{(x_i, z_i)\}_{i=1,\ldots,k}$. The solid grey lines correspond the simulated paths used to get the conditioning events. The inserts give the proportion of points lying in the 95% pointwise confidence intervals.

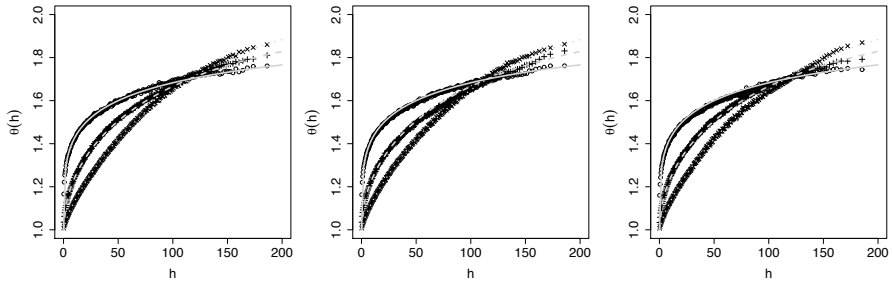

FIGURE 1.7

Comparison of the extremal coefficient estimates (using a binned F-madogram with 250 bins) and the theoretical extremal coefficient function for a varying number of conditioning locations and different semi-variograms with $k = 5$ (left) or $k = 15$ (right) conditioning points. The 'o', '+' and 'x' symbols correspond respectively to γ_1, γ_2 and γ_3. The solid, dashed and dotted grey lines correspond to the theoretical extremal coefficient functions for γ_1, γ_2 and γ_3.

Acknowledgements

M. Oesting and M. Ribatet acknowledge the ANR project McSim for financial support.

References

Chilès, J.-P. and Delfiner, P. (1999), *Geostatistics*, Wiley Series in Probability and Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, modeling spatial uncertainty, A Wiley-Interscience Publication.

Cooley, D., Naveau, P., and Poncet, P. (2006), "Variograms for spatial max-stable random fields," in *Dependence in probability and statistics*, Springer, New York, vol. 187 of *Lecture Notes in Statist.*, pp. 373–390.

Davis, R. A. and Resnick, S. I. (1989), "Basic properties and prediction of max-ARMA processes," *Adv. in Appl. Probab.*, 21, 781–803.

— (1993), "Prediction of stationary max-stable processes," *Ann. Appl. Probab.*, 3, 497–525.

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

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

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

Dombry, C., Éyi-Minko, F., and Ribatet, M. (2013), "Conditional simulation of max-stable processes," *Biometrika*, 100, 111–124.

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

Liu, J. S., Wong, W. H., and Kong, A. (1995), "Covariance structure and convergence rate of the Gibbs sampler with various scans," *J. Roy. Statist. Soc. Ser. B*, 57, 157–169.

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

Penrose, M. D. (1992), "Semi-min-stable processes," *Ann. Probab.*, 20, 1450–1463.

Ribatet, M. (2013), "Spatial extremes: max-stable processes at work," *J. SFdS*, 154, 156–177.

Schlather, M. (2002), "Models for stationary max-stable random fields," *Extremes*, 5, 33–44.

Stoyan, D., Kendall, W. S., and Mecke, J. (1987), *Stochastic geometry and its applications*, vol. 69 of *Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien [Mathematical Textbooks and Monographs, Part II: Mathematical Monographs]*, Akademie-Verlag, Berlin, with a foreword by David Kendall.

Wang, Y. and Stoev, S. A. (2011), "Conditional sampling for spectrally discrete max-stable random fields," *Adv. in Appl. Probab.*, 43, 461–483.

Weintraub, K. S. (1991), "Sample and ergodic properties of some min-stable processes," *Ann. Probab.*, 19, 706–723.