

# Exact Simulation of Max-Infininitely Divisible Processes

Peng Zhong<sup>1</sup>, Raphaël Huser<sup>1</sup>, Thomas Opitz<sup>2</sup>

February 28, 2021

## Abstract

Max-infinitely divisible (max-id) processes play a central role in extreme-value theory and include the subclass of all max-stable processes. They allow for a constructive representation based on the componentwise maximum of random functions drawn from a Poisson point process defined on a suitable functions space. Simulating from a max-id process is often difficult due to its complex stochastic structure, while calculating its joint density in high dimensions is often numerically infeasible. Therefore, exact and efficient simulation techniques for max-id processes are useful tools for studying the characteristics of the process and for drawing statistical inferences. Inspired by the simulation algorithms for max-stable processes, we here develop theory and algorithms to generalize simulation approaches tailored for certain flexible (existing or new) classes of max-id processes. Efficient simulation for a large class of models can be achieved by implementing an adaptive rejection sampling scheme to sidestep a numerical integration step in the algorithm. We present the results of a simulation study highlighting that our simulation algorithm works as expected and is highly accurate and efficient, such that it clearly outperforms customary approximate sampling schemes. As a byproduct, we also develop here new max-id models, which can be represented as pointwise maxima of general location scale mixtures, and which possess flexible tail dependence structures capturing a wide range of asymptotic dependence scenarios.

**Keywords:** Adaptive rejection sampling; Exact simulation; Extremal function; Max-infinitely divisible process; Max-stable process.

---

<sup>1</sup>Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. E-mails: peng.zhong@kaust.edu.sa; raphael.huser@kaust.edu.sa

<sup>2</sup>BioSP, INRAE, Avignon, 84914, France, E-mail: thomas.opitz@inrae.fr

# 1 Introduction

Max-infinitely divisible processes (max-id) have gained much popularity in recent years for the flexible modeling of spatial extremes in phenomena such as precipitation, wind speeds and temperatures (Padoan, 2013; Huser et al., 2020; Bopp et al., 2020; Huser and Wadsworth, 2020; Zhong et al., 2020). Such models encompass max-stable processes (de Haan, 1984) as a subclass, which are currently the mainstream for modeling spatial extremes, as demonstrated by a large body of theoretical and applied literature (see, e.g., Davison et al., 2012; Dey and Yan, 2016; Davison et al., 2019). Max-stability is a property that arises in asymptotic models for multivariate block maxima, but that is however often violated in finite samples (Huser and Wadsworth, 2020). Precisely, the dependence strength in max-stable processes does not depend on event magnitude, while most environmental and climatic processes suggest that it should become weaker at higher thresholds, with the most extreme events being typically more localized. To circumvent this limitation, more flexible max-id processes, which drop the restrictive max-stability assumption but retain natural properties of multivariate block maxima, have been considered. In particular, the max-id models proposed by Huser et al. (2020), Bopp et al. (2020) and Zhong et al. (2020), extend certain popular classes of max-stable models (obtained as limiting cases on the boundary of the parameter space), in order to achieve a good compromise between the pragmatism of flexible max-id models and the strong theoretical foundations of the max-stable sub-class.

The theory behind max-id processes has been studied in depth (Giné et al., 1990; Resnick, 1987; Brown and Resnick, 1977). In particular, Giné et al. (1990) showed that max-id processes possess a functional Poisson point process (PPP) representation. More precisely, each max-id process  $Z = \{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  defined over the region  $\mathcal{S} \subset \mathbb{R}^d$ , can be defined as the pointwise maximum of a potentially infinite number of random functions  $\{\eta_i; i = 1, 2, \dots\}$  defined over  $\mathcal{S}$ , which are sampled according to a PPP with mean measure  $\nu$ , i.e.,

$$Z(\mathbf{s}) = \max_{i=1,2,\dots} \eta_i(\mathbf{s}), \quad \mathbf{s} \in \mathcal{S}. \quad (1)$$

However, analytical forms of the multivariate distribution functions of such processes exist only in special cases, and by analogy with max-stable processes their numerical evaluation in high dimensions becomes impossible due to the combinatorial explosion of the number of terms to be calculated (Padoan [et al.](#), 2010; Castruccio [et al.](#), 2016; Huser [et al.](#), 2020). Consequently, simulation of max-stable processes is often crucial to study their dependence characteristics, or for simulation-based inference (Erhardt and Smith, 2012; Hainy [et al.](#), 2016; Lee [et al.](#), 2018), and the same is true for general classes of max-id processes. However, exact simulation of  $Z(\mathbf{s})$  in (1) is not trivial, especially when the number of functions  $\eta_i$  is infinite and when these random functions are tricky to sample.

For the subclass of max-stable models, there are essentially two types of generic simulation algorithm that have been proposed based on variants of model representations stemming from (1) (Oesting and Strokorb, 2019). These algorithms exploit the specific structure of max-stable processes—the so-called spectral representation (de Haan, 1984), whereby on the unit Fréchet scale (i.e.,  $\Pr\{Z(\mathbf{s}) > z\} = 1/z$ ,  $z > 0$ ) the points of the Poisson process  $\{\eta_i; i = 1, 2, \dots\}$  in (1) can be decomposed into two multiplicative and stochastically independent components as  $\eta_i(\mathbf{s}) = R_i W_i(\mathbf{s})$ . Precisely,  $\{R_i; i = 1, 2, \dots\}$  here denotes a Poisson point process on  $[0, \infty)$  with measure  $r^{-2}dr$ , independent of the random functions  $\{W_i; i = 1, 2, \dots\}$ , which are independent copies of a random process  $W(\mathbf{s})$  defined on  $\mathcal{S}$  such that  $E[\max\{0, W(\mathbf{s})\}] = 1$ . This decomposition always exists for max-stable processes with unit Fréchet margins but is not unique. When the points  $\{R_i; i = 1, 2, \dots\}$  have heavy-tailed intensity  $r^{-(\alpha+1)}dr$  for some  $\alpha > 0$ , the resulting process  $Z$  remains max-stable, but with  $\alpha$ -Fréchet margins. Moreover, the points  $\{R_i^{-1}; i = 1, 2, \dots\}$  can be defined as the arrival times of a renewal process such that we can simulate them in decreasing order, i.e.,  $R_1 > R_2 > \dots$ . When this is the case, the componentwise maximum of  $\{R_j W_j; j = 1, \dots, i-1\}$  is less and less impacted by the  $i$ -th process  $\eta_i = R_i W_i$  as  $i > 1$  is iteratively incremented. If  $W(\mathbf{s})$  is an almost surely bounded random process on  $\mathcal{S}$ , the contribution of  $\eta_i = R_i W_i$  becomes even ultimately completely irrelevant. Based on this idea, Schlather (2002) introduced the first

exact simulation algorithm for max-stable processes by setting an appropriate stopping rule for the simulation of the Poisson points  $R_i W_i$ . If  $W(\mathbf{s})$  is unbounded, we can use certain alternative representations to make it a bounded random process under a suitable change of measure (Dieker and Mikosch, 2015; Oesting et al., 2018; Dombry et al., 2016), but simulation from the bounded process is intricate and numerically challenging even in moderate dimensions for many of the customary models (Kabluchko et al., 2009; Opitz, 2013). The second type of exact simulation algorithm for max-stable processes, introduced by Dombry et al. (2016), relies on the concept of extremal functions (Dombry and Eyi-Minko, 2013). With this method, it is possible to directly simulate the random functions  $\{\eta_i\}$  on  $\mathcal{S}$  that contribute to the maximum at a finite number of locations. In practice, efficient conditional simulation of the random functions  $\eta_i(\mathbf{s})$  given its value at a single location  $\mathbf{s}_0$  is a prerequisite for efficient simulation with this approach.

However, none of the approaches described above tackled exact simulation for general max-id processes, and we intend to fill this gap. In this work, we demonstrate that the idea of the exact simulation algorithm based on extremal functions, developed in Dombry et al. (2016) for max-stable processes carries over to the general max-id case after suitable adjustments, and we illustrate it with max-id models that have found interest in practical applications. The paper is organized as follows. In §2, we recall some theoretical background and formally define the notions of max-id processes and extremal functions. Then, in §3, we derive the specificities of the theory of max-id processes required for exact simulation based on extremal functions, and we present simulation algorithms focusing on two major classes of max-id processes. In particular, we start with some known max-id models of the form (1) where the random functions  $\eta_i(\mathbf{s})$  can be represented as Gaussian scale mixtures, and we then also explore new max-id models based on Gaussian location mixtures that possess appealing tail dependence characteristics. In §4, we conduct a simulation study to demonstrate the performance of our proposed simulation algorithm. We conclude with a discussion in §5.

## 2 Max-id Processes and Extremal Functions

A process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  is called max-infinitely divisible (max-id) if the joint cumulative distribution function (cdf)  $G$  of  $\mathbf{Z}(K_N) = \{Z(\mathbf{s}_1), \dots, Z(\mathbf{s}_N)\}^\top$  at any finite collection of sites  $K_N = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N\} \subset \mathcal{S}$  defines a valid cdf  $G^t$  (with the notation  $G^t(\mathbf{z}) = \{G(\mathbf{z})\}^t$ ,  $\mathbf{z} = (z_1, \dots, z_N)^\top \in \mathbb{R}^N$ ) for any  $t > 0$ . The distribution  $G^t$  does not necessarily stay within the same location-scale family; this property is only satisfied for max-stable distributions  $\tilde{G}$ , for which  $\tilde{G}^m(\mathbf{a}_m \mathbf{z} + \mathbf{b}_m) = \tilde{G}(\mathbf{z})$ , for any integer  $m = 1, 2, \dots$ , and appropriately defined normalizing vectors  $\mathbf{a}_m \in (0, \infty)^N$  and  $\mathbf{b}_m \in \mathbb{R}^N$ . As shown in Equation (1), any max-id process can be constructed by taking pointwise maxima over a Poisson point process defined on a suitable functions space. We assume that  $\{\eta_i(\mathbf{s}); i = 1, 2, \dots\}_{\mathbf{s} \in \mathcal{S}}$  in (1) is a Poisson point process with mean measure  $\Lambda$  on  $\mathcal{C}$ , where  $\mathcal{C}$  denotes the space of functions with compact support  $\mathcal{S}$ . Then, the extremal functions are defined as follows.

**Definition 1.** Let  $K \subset \mathcal{S}$  be a nonempty subset of the domain  $\mathcal{S}$ , and  $Z$  a max-id process. A function  $\phi \in \mathcal{C}$  is called  $K$ -extremal if there exists  $\mathbf{s} \in K$  such that  $Z(\mathbf{s}) = \phi(\mathbf{s})$ , otherwise  $\phi$  is called  $K$ -subextremal. We denote  $\mathcal{C}_K^+$  the set of  $K$ -extremal functions and  $\mathcal{C}_K^-$  the set of  $K$ -subextremal functions.

In other words, an extremal function is a Poisson point  $\eta_i$  in  $\mathcal{C}$  that contributes to the maximum process  $Z$  at one or more locations. We further make an assumption ensuring the continuity of univariate marginal distributions of  $Z$ . Let  $\ell$  denote the lower bound of the marginal support of  $Z$  such that  $\Pr\{Z(\mathbf{s}) > \ell\} = 1$  for all  $\mathbf{s} \in \mathcal{S}$ . We assume that for any fixed  $\mathbf{s}_0 \in \mathcal{S}$ ,

$$\Lambda_{\mathbf{s}_0}(z) = \Lambda(\{\eta \in \mathcal{C}; \eta(\mathbf{s}_0) \in [z, +\infty)\}) \text{ is continuous on } (\ell, +\infty), \lim_{z \downarrow \ell} \Lambda_{\mathbf{s}_0}(z) = +\infty. \quad (2)$$

Under this assumption, [Dombry and Eyi-Minko \(2013\)](#) showed that for singletons  $K = \{\mathbf{s}_0\}$  the set of extremal functions  $\mathcal{C}_{\{\mathbf{s}_0\}}^+$  contains almost surely a single point, denoted by  $\phi_{\{\mathbf{s}_0\}}^+$ , i.e., the resulting max-id process  $Z(\mathbf{s})$  is realized by one and only one Poisson point  $\eta_i$  at a

given location  $\mathbf{s}_0$ . This assumption implies that  $Z$  has continuous univariate margins. This restriction is not problematic since all strictly monotone marginal transformations of  $Z$  yield again a max-id process. In particular, any max-id process  $Z$  with continuous univariate distributions can be marginally transformed to satisfy assumption (2).

Let  $P_{\mathbf{s}_0}(x, \cdot)$  denote the conditional distribution of  $\eta$  given  $\eta(\mathbf{s}_0) = x$ . It turns out that the extremal function  $\phi_{\{\mathbf{s}_0\}}^+$  at location  $\mathbf{s}_0 \in \mathcal{S}$ , given  $Z(\mathbf{s}_0) = z$ , follows the same distribution as  $P_{\mathbf{s}_0}(z, \cdot)$  (see Proposition 4.1 in [Dombry and Eyi-Minko \(2013\)](#)). In other words, the distribution of a Poisson point  $\eta_i$  conditional on the value at a particular location  $\eta_i(\mathbf{s}_0) = z$  is the same if we pick an arbitrary point  $\eta_i$  or a point for which  $z$  is the maximum at  $\mathbf{s}_0$  over all points  $\{\eta_i\}$  of the Poisson process. This property, which is related to the independent occurrence of points in Poisson processes, is key for the exact simulation algorithm proposed by [Dombry et al. \(2016\)](#). Essentially, this algorithm requires simulating the extremal functions at the simulation sites  $K_N = \{\mathbf{s}_1, \dots, \mathbf{s}_N\}$  sequentially, which therefore involves sampling from  $P_{\mathbf{s}_i}(z, \cdot)$ ,  $i = 1, \dots, N$ . If conditional simulation for an arbitrary Poisson point  $\eta_i$  is straightforward, then this advantage will thus also automatically carry over to the simulation of the corresponding max-id process constructed as  $\max_{i=1,2,\dots} \eta_i(\mathbf{s})$ . The algorithm for general max-id processes is described in more detail in §3.

In particular, when the process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  is max-stable with unit Fréchet margins, the Poisson point process  $\{\eta_i; i = 1, 2, \dots\}$  can be decomposed as  $\eta_i(\mathbf{s}) = R_i W_i(\mathbf{s})$ ,  $i = 1, 2, \dots$ , where  $\{R_i; i = 1, 2, \dots\}$  is a Poisson point process on  $(0, \infty)$  with mean measure  $\kappa([r, \infty]) = r^{-1}$ ,  $r > 0$ , and  $\{W_i(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$ ,  $i = 1, 2, \dots$ , are independent copies of a random process  $\{W(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  with  $E[\max\{W(\mathbf{s}), 0\}] = 1$ , which are also independent of the points  $\{R_i; i = 1, 2, \dots\}$  ([de Haan, 1984](#); [Schlather, 2002](#)). A number of parametric max-stable models have been proposed for statistical applications, such as the Brown–Resnick model ([Brown and Resnick, 1977](#); [Kabluchko et al., 2009](#)), and the extremal- $t$  model ([Opitz, 2013](#)), where the intensity function associated to the mean measure  $\Lambda$  has a closed form expression. In these special cases, the multivariate conditional distributions required for exact simulation are

relatively easy to simulate from, and [Dombry et al. \(2016\)](#) provide the exact forms of the distributions  $P_{\mathbf{s}}(z, \cdot)$  required for simulating these max-stable models.

However, such simplifications are no longer available for general max-stable processes and for the classes of max-id (but not max-stable) process models proposed in the recent literature. Very often, the mean measure  $\Lambda(\cdot)$  does not have a closed-form intensity function, and the corresponding conditional distributions are complicated; rather, calculations of relevant measures of the form  $\Lambda(B)$  and intensities often involve computing a one-dimensional integral ([Huser et al., 2020](#); [Zhong et al., 2020](#)). Therefore, simulating from  $P_{\mathbf{s}}(z, \cdot)$  is more involved than in the max-stable case and requires further investigation into its form to propose efficient simulation techniques.

## 3 Exact Simulation

### 3.1 General simulation algorithm

We now extend the generic max-stable simulation algorithm based on extremal functions, as proposed by [Dombry et al. \(2016\)](#), to the general max-id case. We aim to simulate the max-id process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  at a finite number  $N$  of sites given as  $K_N = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N\} \subset \mathcal{S}$ . We define the subsets  $K_n = \{\mathbf{s}_1, \dots, \mathbf{s}_n\}$  of the first  $n$  sites for  $n = 1, \dots, N$ , and we denote by  $Z_n = \max_{\eta_i \in \mathcal{C}_{K_n}^+} \eta_i$  the process constructed from the extremal functions at the sites in  $K_n$ , such that  $\mathbf{Z}(K_N) = \mathbf{Z}_N(K_N)$ . To simulate from the joint distribution of the process  $\{Z(\mathbf{s})\}$  at sites in  $K_N$ , we can iterate through the set of locations from  $\mathbf{s}_1$  to  $\mathbf{s}_N$  in order to simulate the extremal functions at each location. Therefore, a user-defined order of sites must be fixed before simulating, and below we will comment on how this order may be chosen in practice. Recall that we denote by  $\{\phi_{\mathbf{s}_n}^+\}_{1 \leq n \leq N}$  the sequence of extremal functions associated to the singletons  $\{\mathbf{s}_n\}$ ,  $n = 1, \dots, N$ , where some of these functions may be identical to each other. The following theorem characterizes the conditional distributions of  $\phi_{\mathbf{s}_n}$  given  $\{\phi_{\mathbf{s}_k}^+\}_{1 \leq k \leq n-1}$  for extremal functions at position  $n = 2, \dots, N-1$ , which are required for iterative simulation.

**Theorem 2** (Conditional distributions of extremal functions). *Consider a max-id process  $\{Z(\mathbf{s})\}_{\mathbf{s} \in \mathcal{S}}$  as defined in Section 2. Given sites  $K_N = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_N\}$ , the distribution of its extremal function  $\phi_{\mathbf{s}_1}$  and of its conditional extremal functions  $\phi_{\mathbf{s}_n} \mid \{\phi_{\mathbf{s}_k}^+\}_{1 \leq k \leq n-1}$  for  $n > 1$  can be characterized as follows:*

- (1) *Initial extremal function: given a realization  $z$  of  $Z(\mathbf{s}_1)$ , the extremal function  $\phi_{\mathbf{s}_1}^+$  is distributed according to  $P_{\mathbf{s}_1}(z, \cdot)$ .*
- (2) *Conditional extremal functions for  $n \geq 1$ : given the extremal functions  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$  for sites in  $K_n$ , consider the Poisson measure of functions that do not exceed the already simulated max-id values of sites in  $K_n$  but that exceed the maximum of already simulated values at  $\mathbf{s}_{n+1}$ :*

$$\tilde{\Lambda}(\mathrm{d}f) = \mathbb{I}(f(\mathbf{s}_i) < Z_n(\mathbf{s}_i), 1 \leq i \leq n) \mathbb{I}(f(\mathbf{s}_{n+1}) > Z_n(\mathbf{s}_{n+1})) \Lambda(\mathrm{d}f), \quad (3)$$

where the indicator function  $\mathbb{I}(A)$  is 1 if the condition  $A$  is satisfied and 0 otherwise.

Then, the extremal function  $\phi_{\mathbf{s}_{n+1}}^+$  conditional on  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$  is given as follows:

$$\tilde{\phi}_{\mathbf{s}_{n+1}}^+ = \begin{cases} \arg \max_{\phi \in PPP(\tilde{\Lambda})} \phi(\mathbf{s}_{n+1}), & \tilde{\Lambda}(\mathcal{C}) \neq 0, \\ \arg \max_{\phi \in \{\phi_{\mathbf{s}_1}^+, \dots, \phi_{\mathbf{s}_n}^+\}} \phi(\mathbf{s}_{n+1}), & \tilde{\Lambda}(\mathcal{C}) = 0. \end{cases}$$

*Proof.* The proof follows closely the proof of Theorem 2 in [Dombry et al. \(2016\)](#), which concerns max-stable processes, while we here treat max-id processes.

To simulate extremal functions, we proceed sequentially. The distribution of the first extremal function  $\phi_{\mathbf{s}_1}^+$  given  $\phi_{\mathbf{s}_1}^+(\mathbf{s}_1) = z$  is exactly  $P_{\mathbf{s}_1}(z, \cdot)$ , as explained in §2 above, and  $\phi_{\mathbf{s}_1}^+(\mathbf{s}_1)$  thus follows the marginal distribution of  $Z(\mathbf{s}_1)$ , which has now a more general form than for max-stable processes. Notice also that given the continuity assumption in (2), there is almost surely exactly one extremal function at each location.

Then, to simulate the extremal function  $\phi_{\mathbf{s}_{n+1}}^+$  at site  $\mathbf{s}_{n+1}$ , for  $n \geq 1$ , we take into account that it can be generated independently conditional on the already simulated extremal functions  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$ , thanks to the Poisson process property. Therefore,  $\phi_{\mathbf{s}_{n+1}}^+$  either corresponds



to one of the functions in  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$ , or it is a new extremal function, stochastically independent of the first  $n$  extremal functions  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$ . The mean measure  $\tilde{\Lambda}(\mathcal{C})$  in (3) defines the restricted Poisson process of candidate functions with positive Poisson intensity, and a new extremal function  $\phi_{\mathbf{s}_{n+1}}^+$  can be generated according to this restricted point process with finite intensity measure. If the measure  $\tilde{\Lambda}(\mathcal{C})$  is zero, then there are almost surely no points in the Poisson process  $\text{PPP}(\tilde{\Lambda})$  of points that exceed the extremal functions  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$  at the location  $\mathbf{s}_{n+1}$ . Hence, in this case the value  $Z(\mathbf{s}_{n+1})$  of the max-id process corresponds to the maximum of the previous extremal functions  $\{\phi_{\mathbf{s}_i}^+\}_{1 \leq i \leq n}$ , i.e.,  $Z(\mathbf{s}_{n+1}) = Z_n(\mathbf{s}_{n+1})$ . If  $\tilde{\Lambda}(\mathcal{C}) > 0$ , then the realization of the restricted point process contains at least one point with positive probability. If there is at least one point, then the new extremal function at  $\mathbf{s}_{n+1}$  corresponds to the point realizing the maximum value at  $\mathbf{s}_{n+1}$ .  $\square$

An exact simulation algorithm based on Theorem 2 is given in Algorithm 1. Although the main structure of the algorithm resembles that proposed by Dombry et al. (2016) for max-stable processes, the key difficulty for general max-id processes resides in efficiently simulating from the distribution  $P_{\mathbf{s}}(z, \cdot)$  given  $Z(\mathbf{s}) = z$ . For suitable constructions of the Poisson point process  $\{\eta_i; i = 1, 2, \dots\}$ , we can investigate the specific structure of this conditional distribution. In this paper, we consider two different and very general structures for  $\eta_i$ , which are inspired from the construction of max-stable processes, yet they give rise to max-id models with much more flexibility in their joint tail decay rates. These two types of models, which encompass all of the commonly used max-stable models from the spatial extremes literature, are described in the following two subsections.

## 3.2 Representation of $\eta_i$ as a Gaussian scale mixture

### 3.2.1 Model description

One possible structure of the general max-id process is achieved by representing  $\eta_i$  as a Gaussian scale mixture, i.e.,  $\eta_i = R_i W_i$ , where  $\{R_i\}$  are the points of a Poisson point process on  $\mathbb{R}^+$  with  $\Lambda(0, \infty) = \infty$ , and  $\{W_i\}$  are independent Gaussian random fields on  $\mathcal{S}$ , which are

---

**Algorithm 1** Exact simulation of a max-id process  $Z$  at locations  $K_N = \{\mathbf{s}_1, \dots, \mathbf{s}_N\}$ .

---

**Input:** Stationary marginal distribution  $G_0$  of  $Z$ .

**Input:** Dependence characterization of  $Z$  allowing simulation of  $P_{\mathbf{s}_n}$ ,  $n = 1, \dots, N$ .

```

1: Simulate  $E_0 \sim \text{Exp}(1)$  and set  $z = G_0^{-1}(\exp(-E_0))$ .
2: Simulate  $Y \sim P_{\mathbf{s}_1}(z, \cdot)$ .
3: Set  $Z(K_N) = Y(K_N)$ .
4: for all  $n = 2, \dots, N$  do
5:   Simulate  $E_0 \sim \text{Exp}(1)$  and set  $z = G_0^{-1}(\exp(-E_0))$ .
6:   while  $\{z > Z(\mathbf{s}_n)\}$  do
7:     Simulate  $Y \sim P_{\mathbf{s}_n}(z, \cdot)$ .
8:     if  $Y(\mathbf{s}_i) < Z(\mathbf{s}_i)$  for all  $i = 1, \dots, n-1$  then
9:       update  $Z(K_N)$  by the componentwise  $\max(Z(K_N), Y(K_N))$ .
10:    end if
11:   Simulate  $E_1 \sim \text{Exp}(1)$  and update  $E_0$  and  $z$  by setting  $E_0 = E_0 + E_1$  and  $z = G_0^{-1}(\exp(-E_0))$ .
12:   end while
13: end for
Output: Return  $Z(K_N)$ 

```

---

not necessarily identically distributed and may depend on the overall random “magnitude”  $R_i$ . This construction mimics the max-stable spectral construction of extremal- $t$  processes (Opitz, 2013) arising for a specific choice of  $\Lambda$  with  $\{R_i\}$  and  $\{W_i\}$  being independent. To extend max-stable processes, Huser et al. (2020) proposed using a more flexible mean measure for  $\{R_i\}$  and Zhong et al. (2020) further relaxed the independence assumption between  $\{R_i\}$  and  $\{W_i\}$ . Specifically, they proposed a Weibull-tailed mean measure  $\kappa$  for  $\{R_i\}$ , where

$$\kappa(r) = \kappa([r, \infty)) = r^{-\beta} \exp\{-\alpha(r^\beta - 1)/\beta\}, \quad r > 0, \quad (\alpha, \beta)^\top \in (0, \infty)^2. \quad (4)$$

With this specification,  $\kappa(r) \rightarrow r^{-\alpha}$  as  $\beta \downarrow 0$ . Therefore, the model remains in the neighborhood of max-stable processes for which  $\kappa(r) = r^{-\alpha}$ . Given  $\{R_i\}$ , the processes  $\{W_i\}$  may be chosen as standard Gaussian random fields with correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i)$ , where  $R_i$  modulates the dependence range of  $W_i$  in such a way that large magnitudes  $R_i$  create processes  $W_i$  with weaker dependence. This gives extra flexibility for capturing the joint tail decay rate of the resulting max-id process  $Z$  in the asymptotic independence setting, i.e., when  $\lim_{u \rightarrow 1} \Pr[G_0\{Z(\mathbf{s}_1)\} > u \mid G_0\{Z(\mathbf{s}_2)\} > u] = 0$  with  $G_0$  being the marginal distribution of  $Z$ . For example, Zhong et al. (2020) chose a non-stationary correlation  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i)$ ,

which may be specified as

$$\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) = (\lambda_{\mathbf{s}_1} \lambda_{\mathbf{s}_2}) \left( \frac{\lambda_{\mathbf{s}_1}^2 + \lambda_{\mathbf{s}_2}^2}{2} \right)^{-1} \exp \left\{ - \left( \frac{\lambda_{\mathbf{s}_1}^2 + \lambda_{\mathbf{s}_2}^2}{2} \right)^{-1/2} (1 + R_i)^\nu \|\mathbf{s}_1 - \mathbf{s}_2\| \right\}, \quad (5)$$

where  $\lambda_{\mathbf{s}} > 0$  is a spatially-varying range parameter surface that can be associated with spatial covariates, and  $\nu \geq 0$  is a parameter that controls how the scaling variable  $R_i$  influences the dependence range of  $W_i$ . In the simulation examples below, we will illustrate some specific structures for  $\lambda_{\mathbf{s}}$ . When the value of  $R_i$  is large and  $\nu > 0$ , the dependence within the Gaussian random field  $W_i$  indeed decreases with increasing “magnitude”  $R_i$ . In this model, the marginal exponent measure for the locations in  $K \equiv K_N$  may be written as

$$\Lambda_K([\mathbf{z}_K, \infty)) = \int_0^\infty \{1 - \Phi(\mathbf{z}_K/r; \boldsymbol{\rho}(K; r))\} \kappa(dr)$$

where  $\Phi(\cdot; \boldsymbol{\rho}(K; r))$  is the multivariate normal distribution function with zero mean and correlation matrix  $\boldsymbol{\rho}(K; r)$  determined by the correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2; r)$  for  $\mathbf{s}_1, \mathbf{s}_2 \in K$ .

To illustrate the tail dependence properties of this max-id process, we use the level-dependent extremal coefficient (Padoan, 2013; Huser et al., 2020). For a collection of sites  $K_N \subset \mathcal{S}$  and  $|K_N| = N$ , the level-dependent extremal coefficient is defined as

$$\theta_N(z_0) = \frac{\log\{G(\mathbf{z}_0)\}}{\log\{G_0(z_0)\}} \quad (6)$$

where  $\mathbf{z}_0 = (z_0, \dots, z_0)^T \in \mathbb{R}^N$ ,  $G_0$  is the marginal distribution of  $Z$  and  $G$  is the joint distribution of the random vector  $\mathbf{Z}(K_N)$  with observations at the locations in the set  $K \subset \mathcal{S}$ . This coefficient can be expressed in terms of the measure  $\Lambda$ , and in the case where the margins are unit Fréchet it is specifically equal to  $\theta_N(z_0) = z_0 \Lambda_{K_N} \{(-\infty, \infty)^N \setminus (-\infty, z_0]^N\}$ . From (6) it is easy to see that  $\Pr\{\mathbf{Z}(K_N) \leq \mathbf{z}_0\} = G_0(z_0)^{\theta_N(z_0)}$ , such that this coefficient can be interpreted as the effective number of independent variables among the components of the vector  $\mathbf{Z}(K_N)$  at level  $z_0$ . Higher values of  $\theta_N(z_0)$  imply weaker dependence at quantile level  $z_0$ . This dependence measure is widely used in extreme-value studies (mainly with max-stable models), where modeling the tail dependence structure is the main interest. As an example, we consider the correlation function of  $W_i$  conditional to the event magnitude  $R_i$

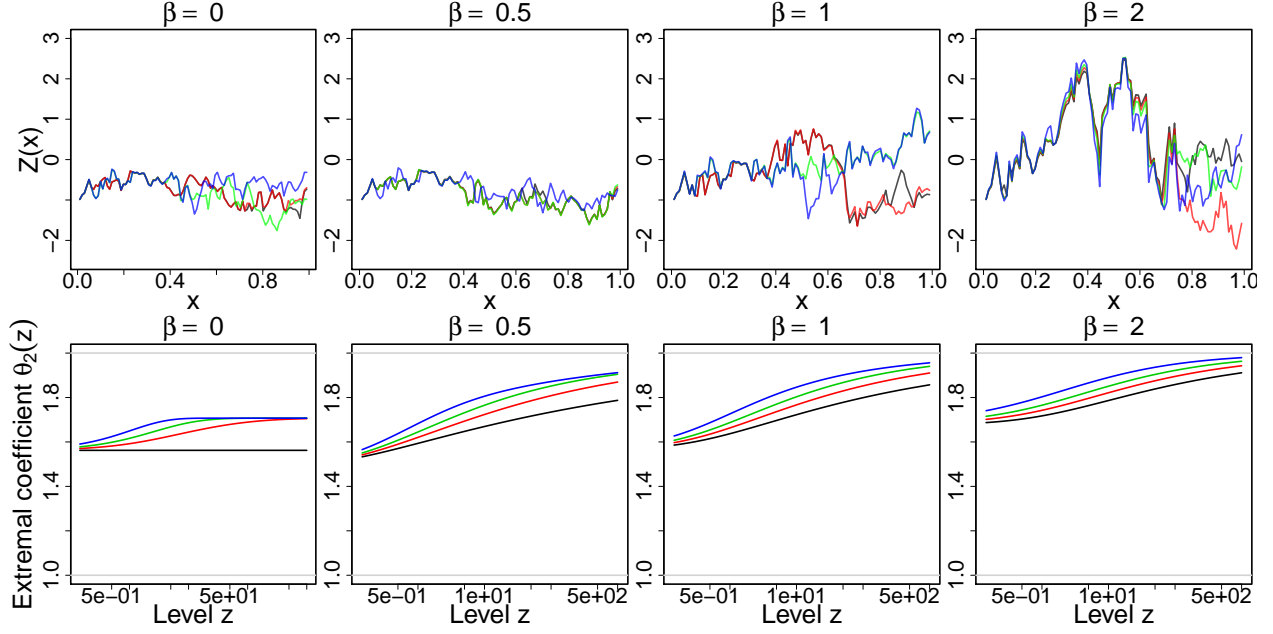


Figure 1: Simulations (top row) and bivariate level-dependent extremal coefficients (bottom row) for models based on the Gaussian scale mixture construction. The parameters are chosen as  $\nu = 0, 0.25, 0.5, 1$  (black, red, green, blue curves), where  $\nu$  controls the dependence between  $R_i$  and  $W_i$  by using correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) = \exp \{-2(1 + R_i)^\nu \|\mathbf{s}_1 - \mathbf{s}_2\|\}$  conditional to  $R_i$ . The level  $z$  is shown on a logarithmic scale.

chosen as  $\rho(\mathbf{s}_1, \mathbf{s}_2; R_i) = \exp \{-2(1 + R_i)^\nu \|\mathbf{s}_1 - \mathbf{s}_2\|\}$ , which corresponds to the stationary and isotropic counterpart of (5) with spatially-constant  $\lambda_s = 1/2$ . Figure 1 shows the realizations of  $\{Z(s)\}_{s \in (0,1)}$  on a standard Gumbel scale generated using our proposed exact simulation algorithm, as well as the bivariate level-dependent extremal coefficients between two sites at distance  $h = 0.5$  for the parameters  $\alpha = 1$ ,  $\beta = 0, 0.5, 1, 2$  and  $\nu = 0, 0.25, 0.5, 1$ . When  $\beta = 0$  (interpreted here as  $\beta \downarrow 0$ ), the extremal dependence persists with increasing level  $z$ . The curve of extremal dependence becomes a horizontal line when we  $\nu$  and  $\beta$  are both zero. In this case, the max-id process is the max-stable extremal- $t$  process. When  $\beta$  is different from zero, we get asymptotic independence (since  $\theta_2(z_0) \rightarrow 2$  when  $z \rightarrow \infty$ ), and the dependence level decreases as the values of  $\beta$  and  $\nu$  increase. The realizations were generated using the same random seed, and we can discern stronger fluctuations of the realizations as the value of  $\beta$  increases, corresponding to a decreasing dependence strength.

### 3.2.2 Simulation algorithm

The generic Algorithm 1 based on extremal functions can be applied, and we now specifically describe how to sample from  $P_{\mathbf{s}_0}(z, \cdot)$  for some fixed site  $\mathbf{s}_0$  (see lines 2 and 7 of the algorithm) for the max-id model constructed from Gaussian scale mixtures  $\eta_i = R_i W_i$  as described in §3.2.1. Simulation according to  $P_{\mathbf{s}_0}(z, \cdot)$  at locations in  $K$  implies that we have to simulate  $\eta_i = R_i W_i$  at  $K$  conditional on  $\eta_i(\mathbf{s}_0) = R_i W_i(\mathbf{s}_0) = z$ . Given the intensity function  $f_R(r) = d\Lambda(r)/dr$  and the univariate standard Gaussian density  $f_W$ , it is straightforward to use variable transformations to derive the joint intensity of  $R_i$  and  $R_i W_i(\mathbf{s}_0)$  as  $f_{(R, RW)}(r, z) = f_R(r) \times f_W(z/r)/r$ . Then, using the definition of conditional intensities, the conditional density  $f_{R|RW}(r | z)$  of  $R_i | R_i W_i(\mathbf{s}_0) = z$  is  $f_{(R, RW)}(r, z)/f_{RW}(z)$ , where the marginal intensity arising in the denominator is  $f_{RW}(z) = \int f_{(R, RW)}(r, z) dr$ . Since this involves an integral, which typically cannot be obtained in closed form and whose numerical calculation in practice may be relatively slow and potentially unstable, we treat it as an unknown constant. We here explore the use of two simulation techniques where knowledge of the value of this constant is not required. Precisely, we can proceed according to the following three general steps:

- (1) Simulate a value  $R$  from the distribution of  $R_i | R_i W_i(\mathbf{s}_0) = z$ , using either
  - the Metropolis–Hastings algorithm (e.g., [Robert and Casella, 2013](#))
  - the adaptive rejection sampling method proposed in [Gilks and Wild \(1992\)](#).
- (2) Given  $R_i = R$ , simulate a vector  $\mathbf{W}$  from the distribution of  $W_i | RW_i(\mathbf{s}_0) = z$  at the locations in  $K_N$ .
- (3) Return  $R\mathbf{W}$ .

Notice that the second step corresponds to simulating from a conditional multivariate normal distribution and can easily be performed exactly (e.g., [Varadhan, 2015](#)) in dimension  $N$  up to

several thousands. Therefore, the key difficulty lies in the first step, where we have to simulate from the probability density  $f_{R|RW}(r | z)$  to simulate the conditional event magnitudes.

### 3.2.3 Options for simulating conditional event magnitudes

With the Metropolis–Hastings (MH) method in Step 2 of the above algorithm in §3.2.2, we use the following symmetric random walk proposal:

$$\log(R_i^j) | R_i^{j-1} \sim \mathcal{N}(\log(R_i^{j-1}), \sigma^2), \quad j = 1, 2, \dots,$$

where  $\sigma^2$  is the proposal variance and has to be tuned, and the initial value  $R_i^0$  is either fixed by the user or drawn from some distribution that we can easily simulated from. MH is approximate (thus not exact) when the distribution of the initial value is different from the target distribution (which is always the case in practice), but its error becomes negligible when a sufficient number of iterations are performed. Here, the MH algorithm is performed on a single variable  $R_i$  at a time, so each Markov chain Monte Carlo (MCMC) chain mixes well and converges very quickly to its stationary distribution, while iterations are of very small computational cost. Our experience shows that very few iterations are in fact needed in practice to get accurate results. We therefore regard Algorithm 1 combined with MH for simulating  $R_i | R_i W_i(\mathbf{s}_0) = z$  as a general “quasi-exact” simulation method.

As an alternative to MH, we also consider adaptive rejection sampling (ARS), which is an exact simulation algorithm that is often used in combination with the Gibbs sampler. It can generate exact independent samples from the target distribution, here supposed to be the (unnormalized) density  $f(r | z) = f_{R|RW}(r | z) \propto f_{R,RW}(r, z)$  with fixed  $z$ . The only requirement for using this method is that  $h(r) = \log f(r | z)$  is concave in  $r$  with a connected support, and that it is continuous and differentiable everywhere within the support. In particular, for the previously presented parametric family of Gaussian scale mixture models based on (4), it is straightforward to show that  $h(r)$  is indeed a continuous, differentiable and concave function over its support  $\mathbb{R}$ . The ARS algorithm adaptively constructs a piecewise upper hull function  $u(r)$  and a piecewise lower bounding function  $l(r)$  of  $h(r)$ , respectively,

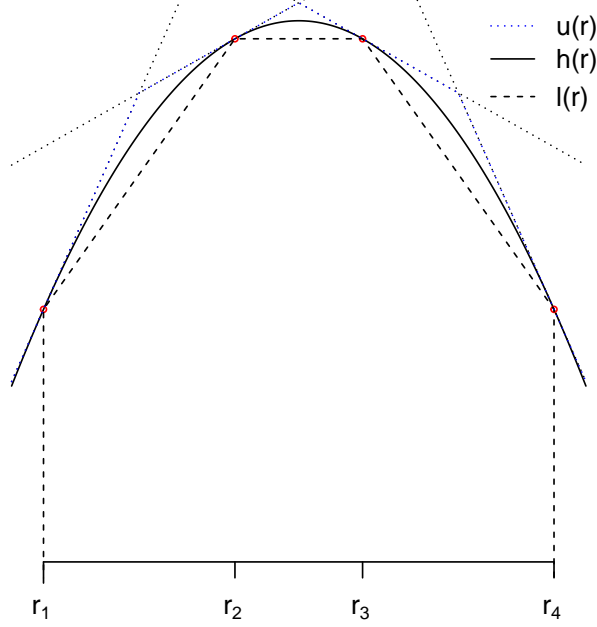


Figure 2: An illustration of the adaptive rejection sampling scheme, where the upper hull function  $u(r)$  is tangent to the logarithmic target density  $h(r)$  at locations  $\mathbf{r} = \{r_1, r_2, r_3, r_4\}$ .

and then uses the distribution  $s(r) = \exp\{u(r)\} / \int \exp\{u(r')\} dr'$  to generate proposals in the exact rejection sampling scheme. Figure 2 illustrates the adaptive rejection sampling scheme, where the upper hull function  $u(r)$  is tangent to the target  $h(r)$  at locations  $\mathbf{r} = \{r_1, r_2, r_3, r_4\}$ . Therefore,  $u(r)$  and  $l(r)$  are uniquely determined by the locations in  $\mathbf{r}$  and the target function  $f(r | z)$ . As we increase the number of points in  $\mathbf{r}$ , the functions  $u(r)$  and  $l(r)$  will form an increasingly narrow envelope around the function  $h(r)$ , and the acceptance rate of the algorithm will increase. The steps of the ARS algorithm are as follows:

- (0) Initialization: choose a set of anchor points  $\mathbf{r}$  and construct the upper and lower bounds  $u(r)$  and  $l(r)$ , respectively, as illustrated in Figure 2.
- (1) Sampling: sample  $R^* \sim s(r)$ , and sample  $U \sim \text{Unif}(0, 1)$  independently of  $R^*$ .
- (2) First acceptance test: if  $U \leq \exp\{l(R^*) - u(R^*)\}$  then accept  $R^*$  and go to 5.
- (3) Second acceptance test: If  $U \leq \exp\{h(R^*) - u(R^*)\}$  then accept  $R^*$  and go to 5.

- (4) If the first and second acceptance tests both fail, reject  $R^*$ ; add  $R^*$  to list of anchor points and update the bounds  $u(r)$  and  $l(r)$ . Go to 1.
- (5) Return  $R^*$ .

The distribution of the simulated value given acceptance is thus

$$\begin{aligned} \Pr(R^* \in A \mid \text{acceptance of } R^*) &= \frac{\int_A s(r) \left\{ \int_0^{\exp\{l(r)-u(r)\}} dw + \int_{\exp\{l(r)-u(r)\}}^{\exp\{h(r)-u(r)\}} dw \right\} dr}{\int_{\mathbb{R}} s(r) \left\{ \int_0^{\exp\{l(r)-u(r)\}} dw + \int_{\exp\{l(r)-u(r)\}}^{\exp\{h(r)-u(r)\}} dw \right\} dr} \\ &= \frac{\int_A \exp\{h(r)\} dr}{\int_{\mathbb{R}} \exp\{h(r)\} dr}, \end{aligned}$$

which corresponds exactly to the target distribution with density  $f(r \mid z)$ . By exploiting the concavity property of  $h$ , ARS usually only needs a relatively small number of iterations before acceptance, and it may therefore be faster and more accurate than MH, as we will illustrate through the simulation experiments in §4.

### 3.3 Representation of $\eta_i$ as a Gaussian location mixture

#### 3.3.1 Model description

Another flexible and general max-id model family is the class of Gaussian location mixtures, where  $\eta_i$  has the form  $\eta_i = R_i + W_i$ . Here, the location variables  $\{R_i\}$  are the points of a Poisson point process over  $\mathbb{R}$  with mean measure  $\bar{\kappa}$ , and  $\{W_i\}$  are independent Gaussian random fields. Again, the random variables  $\{R_i\}$  can be interpreted as overall event magnitudes, while the random processes  $\{W_i\}$  vary over  $\mathcal{S}$  and can be viewed as spatial “profiles”. For a general parametric form of the distribution of event magnitudes  $\{R_i\}$ , we propose to define  $\bar{\kappa}$  as

$$\bar{\kappa}(r) = \exp \left[ -\alpha \left\{ \mathbb{I}(r > 0) |r|^{\beta_1} + \mathbb{I}(r < 0) |r|^{\beta_2} \right\} \right], \quad r \in \mathbb{R}, \quad (\alpha, \beta_1, \beta_2)^\top \in (0, \infty)^2 \times (0, 2), \quad (7)$$

where  $\bar{\kappa}(r)$  is the short-hand notation for  $\bar{\kappa}([r, \infty))$ . The parameter restrictions for  $\alpha, \beta_1$  and  $\beta_2$  ensure that the mean measure  $\Lambda$  of the points  $\{\eta_i\}$  is indeed Radon (i.e., locally finite) as required to defined a valid Poisson point process. Max-id models based on Gaussian location



mixtures have never been explored in the literature, and the new model (7) is especially interesting in view of the flexible dependence structure that it yields and its close connection to max-stable Brown–Resnick processes (Kabluchko *et al.*, 2009). The parameter  $\beta_1$  controls the dependence strength in the upper tail, while  $\beta_2$  controls the dependence strength in the lower tail. As in §3.2.1, we may further allow for dependence between the point  $R_i$  and the Gaussian field  $W_i$  for each  $i$ . For a general and flexible correlation function of  $W_i$ , we may again use the correlation function in (5) to drive the range of the Gaussian fields  $W_i$  through the location variable  $R_i$ , though other options are of course possible. In this case, the mean measure of  $\{\eta_i; i = 1, 2, \dots\}$  at a set of locations  $K$  is given by

$$\Lambda_K([z_K, \infty)) = \int_{\mathbb{R}} \{1 - \Phi(z_K - r; \boldsymbol{\rho}(K; r))\} \bar{\kappa}(\mathrm{d}r).$$

When  $\alpha = \beta_1 = \beta_2 = 1$  and  $\{R_i\}$  and  $\{W_i\}$  are independent of each other, the proposed model reduces to the popular Brown–Resnick max-stable model (Kabluchko *et al.*, 2009; Wadsworth and Tawn, 2014), although this model is often defined through intrinsically stationary Gaussian processes  $\{W_i\}$  with some variogram function  $\gamma(h)$  rather than through a correlation function  $\rho(\mathbf{s}_1, \mathbf{s}_2)$ .

### 3.3.2 Simulation algorithm

By analogy with the approach developed for Gaussian scale mixtures, we can here again use Algorithm 1 for the max-id location mixture process. Moreover, to sample from  $P_{\mathbf{s}_0}(z, \cdot)$ , we similarly propose a simulation algorithm, where the first step consists of sampling from the conditional (unnormalized) density  $f(r \mid z)$  of  $R_i \mid R_i + W_i(\mathbf{s}_0) = z$ . Again, change-of-variable formulas allow writing  $f(r \mid z)$  as a density with a normalizing constant defined through a univariate integral. However, here we observe that the conditional density is not logarithmically concave, and therefore we cannot use the adaptive resampling scheme to generate extremal functions from  $P_{\mathbf{s}_0}(z, \cdot)$ , but we can still use the quasi-exact Metropolis–Hastings algorithm, which provides high accuracy at low computational cost. Thus, our proposed procedure for simulating max-id processes of Gaussian location mixture type at a

collection of sites  $K_N$  relies on Algorithm 1 combined with the following three-step approach to sample from  $P_{\mathbf{s}_0}(z, \cdot)$ :

- (1) Simulate a value  $R$  according to the distribution of  $R_i \mid R_i + W_i(\mathbf{s}_0) = z$  by using the Metropolis–Hastings method with random walk proposal distribution  $R_i^j \mid R_i^{j-1} \sim \mathcal{N}(R_i^{j-1}, \sigma^2)$ ,  $j = 1, 2, \dots$ , and initial point  $R_i^0$ .
- (2) Given  $R_i = R$ , simulate a vector  $\mathbf{W}$  according to the distribution of  $W_i \mid R + W_i(\mathbf{s}_0) = z$  at the locations in  $K_N$ .
- (3) Return  $R + \mathbf{W}$ .

As in the Gaussian scale mixture case, the second step consists in simulating from a conditional multivariate normal distribution, which is fast and straightforward. Figure 3 shows examples of realizations of  $\{Z(s)\}_{s \in (0,1)}$  with standard Gumbel marginal distributions (generated using the proposed quasi-exact algorithm), as well as the bivariate level-dependent extremal coefficients at distance  $h = 0.5$  for the parameters  $\beta_1 = 0.5, 1, 1.5$ ,  $\beta_2 = 0.5, 1, 1.5$  and  $\alpha = 1$ . The correlation function for  $W_i$  is here chosen to be exponential as  $\rho(h) = \exp(-2h)$ . We see that when both  $\beta_1$  and  $\beta_2$  are equal 1, the curve of the extremal coefficients is a horizontal line. In this case, the max-id process is the Brown–Resnick max-stable process. As the values for  $\beta_1$  and  $\beta_2$  increase, the dependence levels in the lower tail and the upper tail, which are controlled by  $\beta_2$  and  $\beta_1$ , respectively, decrease. When  $\beta_1 > 1$ , the extremal coefficient increases as the level  $z \rightarrow \infty$ . By contrast, when  $\beta_1 < 1$ , the extremal coefficient decreases as the level  $z \rightarrow \infty$ . Similar results are observed for the dependence level in the lower tail with  $\beta_2$ . Therefore, this model shows great flexibility through its ability to capture asymptotic dependence and asymptotic independence in the upper tail and different extremal dependence structures in the lower tail. In these realizations generated with the same random seed, we discern that the fluctuations are increasing as the dependence level decreases, as is expected.

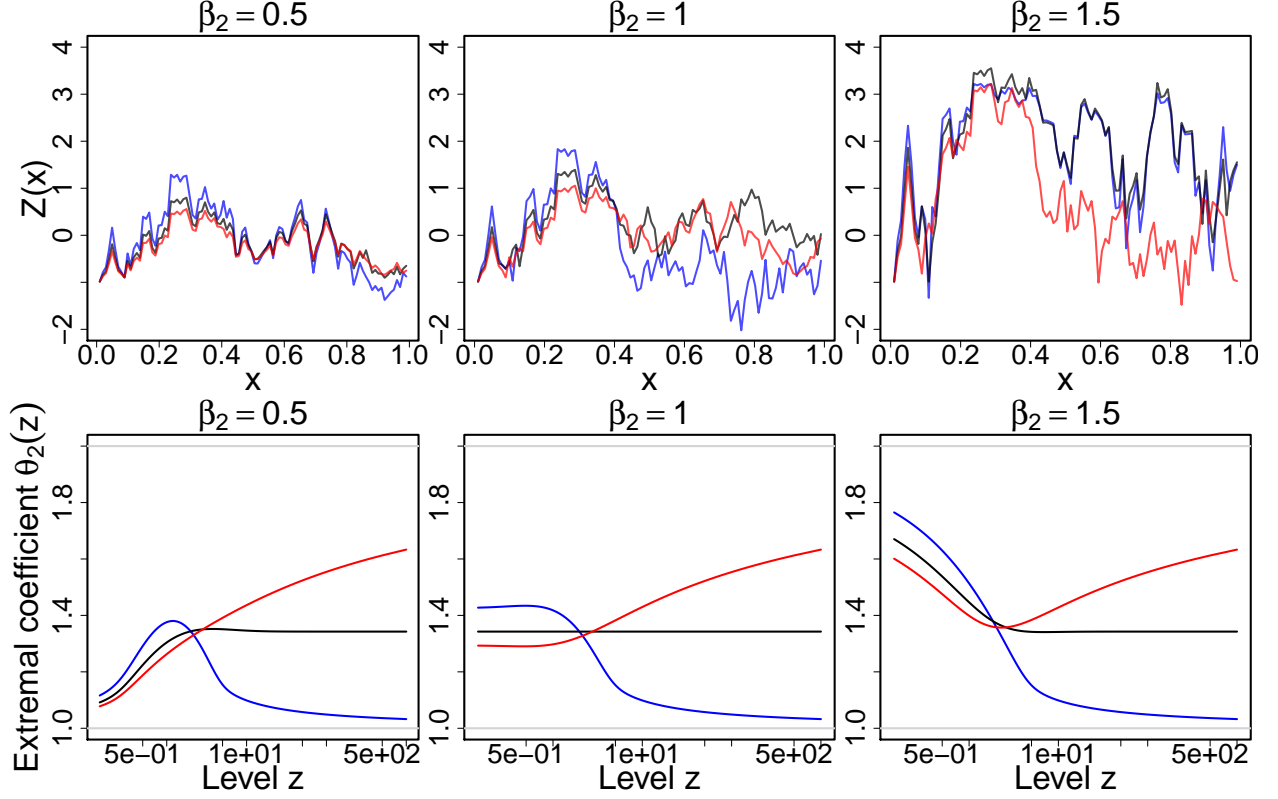


Figure 3: Simulations (top row) and bivariate level-dependent extremal coefficients (bottom row) for models based on the Gaussian location mixture construction. Parameters are chosen as  $\alpha = 1$  and  $\beta_1 = 0.5, 1, 1.5$  (blue, black, red). The correlation function for  $W_i$  is exponential and given by  $\rho(h) = \exp(-2h)$ . The level  $z$  is shown in logarithmic scale.

## 4 Simulation Study

In this section, we use the Gaussian scale mixture representation described in §3.2 as an example to illustrate Algorithm 1 and to assess the performance of the two algorithmic choices detailed in §3.2.2 for sampling from  $P_{s_0}(z, \cdot)$ , i.e., of the Metropolis–Hastings scheme and the adaptive rejection sampling scheme. We also provide a comparison with respect to a naive simulation approach inspired from Schlather (2002), which consists in simulating the magnitudes  $R_i$  in descending order and truncating the maximum in (1) as  $Z(\mathbf{s}) \approx \max_{i=1, \dots, n} \eta_i(\mathbf{s})$  with  $\eta_i = R_i W_i(\mathbf{s})$  for a sufficiently large integer  $n$ .

We perform the simulation on a  $7 \times 7$  regular grid in  $(0, 1)^2$ , corresponding to 49 locations in  $K$ . We fix  $\alpha = 5, \beta = 2$  in (4) and  $\nu = 3, \lambda_s = \exp[2 - 0.5\Phi\{(s_x - 0.5)/0.25\}]$  in (5),

where  $\mathbf{s} = (s_x, s_y)^\top$  and  $\Phi(\cdot)$  is the standard normal distribution function. We implemented three algorithms using the R software: the approximate naive sampling method (APPROX), Algorithm 1 combined with the quasi-exact random walk Metropolis–Hastings (MH), and Algorithm 1 combined with the exact adaptive rejection sampling (ARS). Our aim is to compare the simulation accuracy with respect to univariate distributions and to dependence properties. We tuned the settings of these three algorithms such that they require approximately the same amount of running time to generate  $N = 10^5$  independent samples of the max-id process on the grid. These settings are given as follows:

- (1) Naive method: Simulate  $\eta_i = R_i W_i$ ,  $i = 1, 2, \dots, n$  on the grid  $K$ , where  $n = 100$ .
- (2) Metropolis–Hastings scheme: In the MCMC, run 100 iterations to generate one sample from  $R_i \mid R_i W_i(\mathbf{s}_0) = z$ .
- (3) Adaptive rejection sampling scheme: Use 6 points to initialize the upper hull function and lower hull function.

Based on the generated samples, we compute the empirical Kullback–Leibler (KL) divergence  $D_{\mathbf{s}_0}(p\|q)$  (Kullback and Leibler, 1951) in order to compare the accuracy of simulated samples at each site  $\mathbf{s}_0 \in K \subset \mathcal{S}$ . Precisely, we compute

$$D_{\mathbf{s}_0}(p\|q) = \sum_{i=1}^N p\{Z_i(\mathbf{s}_0)\} \log \left[ \frac{q\{Z_i(\mathbf{s}_0)\}}{p\{Z_i(\mathbf{s}_0)\}} \right], \quad (8)$$

where  $q$  is the empirical density of the samples while  $p$  is the exact marginal density of  $Z$ . Therefore, we have 49 empirical KL divergences in total, one for each point on the grid. Figure 4 shows the boxplot of the 49 empirical KL divergences using the simulated datasets generated by the three algorithms. Both the adaptive rejection sampling scheme and the Metropolis–Hastings scheme outperform the naive method by a significant margin. Moreover, the median of the empirical KL divergence for the adaptive rejection sampling scheme is notably lower than with the Metropolis–Hastings scheme. Generating a single max-id replication on the grid took approximately 1.2 seconds using a single-core machine.

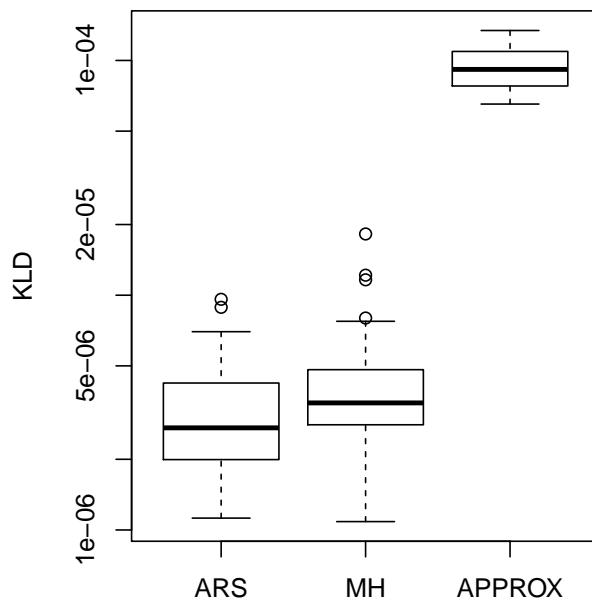


Figure 4: Boxplots of empirical KL divergences for the 49 grid locations based on the simulated data sets generated by the naive method (APPROX), the Metropolis–Hastings scheme (MH) and the adaptive rejection sampling scheme (ARS).

To further evaluate the performance of each sampling scheme with respect to the dependence structure, we calculated the bivariate (level-dependent) empirical extremal coefficients  $\hat{\theta}_2$ . For each pair of sites, we compared the empirical extremal coefficient computed from the  $N = 10^5$  simulated replicates with the theoretical counterpart computed by evaluating the mean measure  $\Lambda$  with a very accurate numerical unidimensional integral. Figure 4 shows the results for the three simulation techniques. The adaptive rejection sampling scheme and the Metropolis–Hastings scheme appear to perform much better than the naive scheme, which is significantly biased at low quantile levels  $z_0$ , while the adaptive rejection sampling scheme seems to be perhaps slightly better than the Metropolis–Hastings scheme. Our results imply that if we draw statistical inferences based on simulated samples from the naive method, the strength of the dependence at relatively low quantile levels of extremes would be significantly overestimated in the current setting. Better results might be obtained for the naive method by increasing the number  $n$  of simulated fields  $\{\eta_i\}$ , but this comes with a computational cost. By contrast, our proposed algorithm always provides exact max-id simulations at a

low computational cost, which may prove essential for simulation-based statistical inference.

As in [Dombry et al. \(2016\)](#), it can be shown that the expected number of random fields  $W_i$  to be simulated in the exact simulation algorithm equals to the number of simulation sites, i.e., the dimension  $|K|$ . To further illustrate the efficiency of our proposed algorithm, we simulated 10 replicas for different numbers of randomly sampled locations  $K$  on the unit square  $(0, 1)^2$  using the ARS scheme and calculated the time and number of random fields  $W_i$  that needed to be simulated. The results are shown in Figure 6. The parameters are set to be  $\alpha = 1, \beta = 1$  and the correlation function for  $W_i$  is chosen to be  $\rho(h) = \exp(-h)$ . Since we use the Cholesky decomposition for simulating the process  $W_i$  on  $K$ , whose complexity is  $\mathcal{O}(|K|^3)$ , the expected overall complexity for simulating a max-id replicate is  $\mathcal{O}(|K|^4)$ . The red curve on the left panel of Figure 6 illustrates this. Here, the computational time is about two minutes for one replicate in dimension  $|K| = 900$ , which is already quite fast. If further speeds-up are required, approximate algorithms for simulating the Gaussian processes  $W_i$  may be employed (e.g., based on Gaussian Markov random fields). As expected, the right panel of Figure 6 shows that the expected number of simulations of  $W_i$  that are needed is indeed equal to  $|K|$ . We also investigated how the ordering of locations  $\mathbf{s}_1, \dots, \mathbf{s}_n$  chosen in Algorithm 1 to iteratively sample from  $P_{\mathbf{s}_0}(z, \cdot)$  might affect the algorithm's efficiency. Our results (not shown) suggest that there was very little difference among the four orderings considered (coordinate-wise, random, middle-out, or maximum-minimum orderings), in terms of the mean and variance of the number of processes  $W_i$  to sample in each simulation. In our simulations, we used a coordinate-wise ordering for simplicity.

## 5 Discussion

In this paper, we have generalized the exact simulation algorithm for max-stable processes based on extremal functions to the more general class of max-id processes. Whenever possible, we recommend using the adaptive rejection sampling scheme for the conditional simulation of magnitude variables, in order to achieve fast and exact simulations in cases where

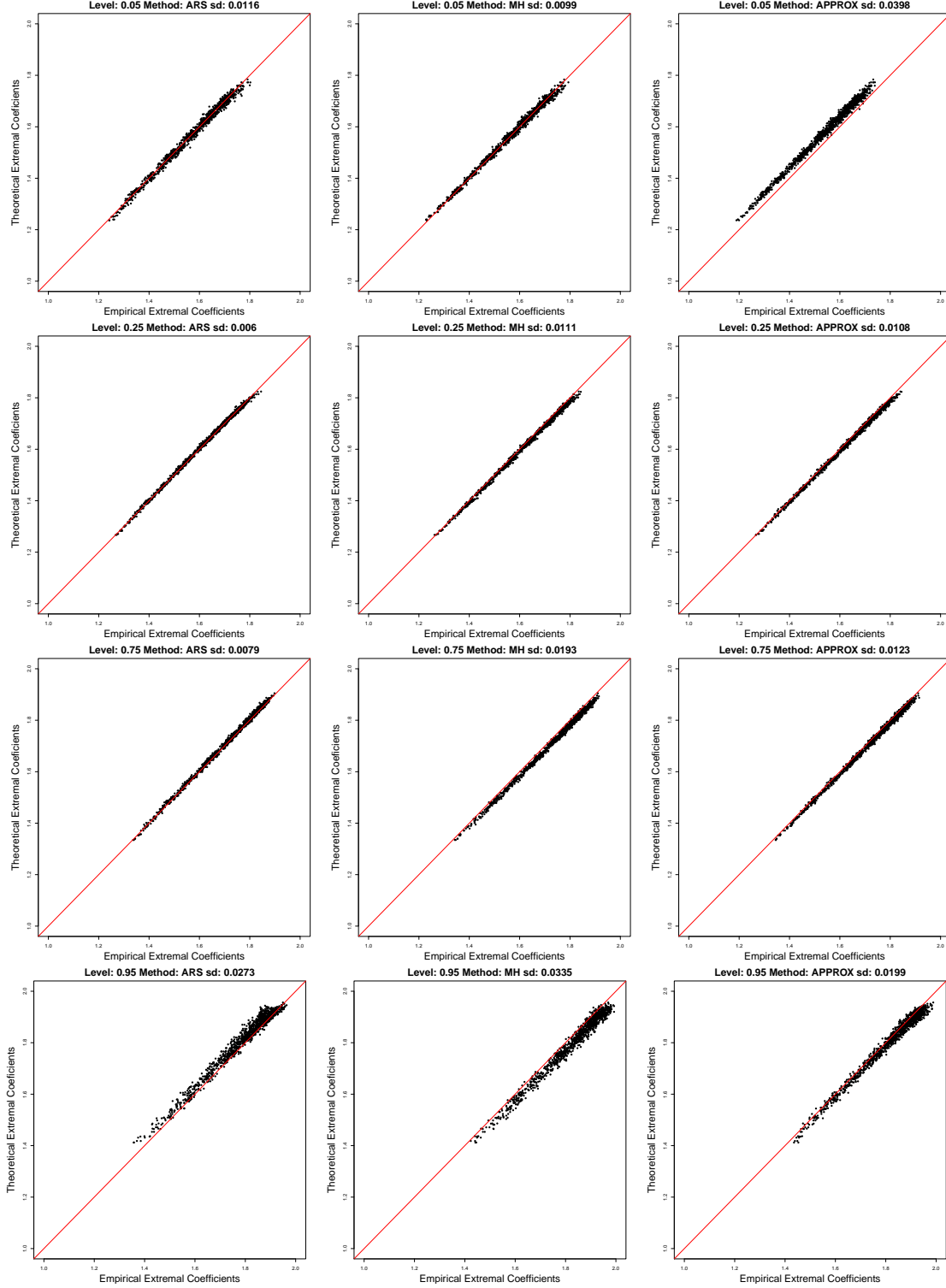


Figure 5: Bivariate scatterplots of theoretical against empirical bivariate level-dependent extremal coefficients at quantile levels 0.05, 0.25, 0.5, 0.75, 0.95 (from top to bottom) for all pairs of locations, based on the three sampling schemes (left: Algorithm 1 with adaptive rejection sampling (ARS), middle: Algorithm 1 with Metropolis–Hastings (MH), right: naive approximate method).

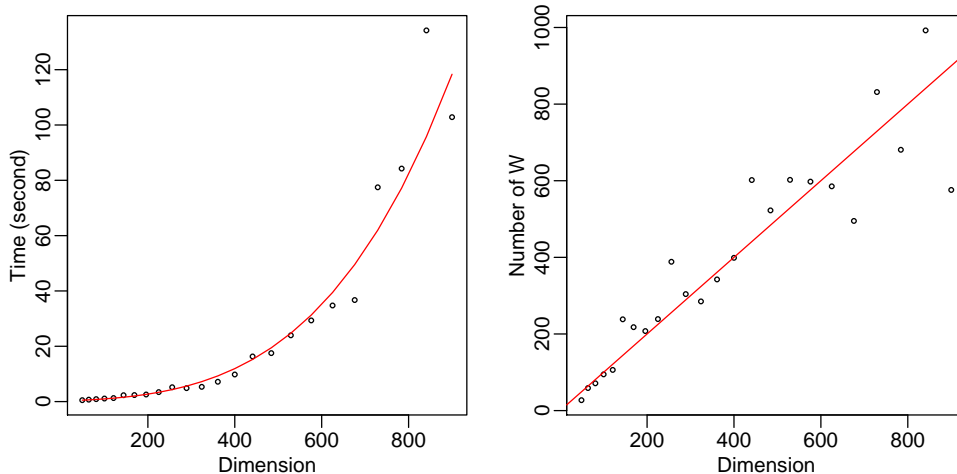


Figure 6: Left: Time in second needed to simulate one max-id replicate for different numbers of locations  $|K|$ . The red curve is the fit from a 4-th order polynomial. Right: Number of random fields  $W_i$  that were simulated for one max-id replicate for different numbers of locations  $|K|$ . The red line is the diagonal. The parameters are here set to be  $\alpha = 1, \beta = 1$  and the correlation function for  $W_i$  is chosen to be  $\rho(h) = \exp(-h)$ .

variables to be simulated do not correspond to some standard distribution from which samples are easily drawn. Although the specific implementation of the algorithm still depends on the chosen representation for  $\eta_i$ , our approach is general and can be applied to very broad classes of max-id processes while preserving computational efficiency. Moreover, the specific examples showcased in this paper already have very flexible dependence structures. In a simulation study, we used three different sampling algorithms to compare their speed and accuracy. Our novel exact simulation algorithm using the adaptive rejection sampling scheme clearly outperforms the two alternatives. In cases where the target density in the adaptive rejection sampling scheme is not logarithmically concave, it is possible to use an appropriately defined quasi-exact Metropolis–Hastings scheme to achieve high accuracy.

Besides standard uses of such fast and exact simulation techniques in Monte–Carlo-based statistical inferences involving quantities that cannot be calculated analytically, another interesting perspective concerns Approximate Bayesian Computation (ABC) for general max-id processes, which could be implemented by analogy with the approach developed for max-stable processes in [Erhardt and Smith \(2012\)](#). As highlighted above, the computation of



multivariate expression of densities and distributions functions is often numerically very expensive due to integrals that have to be evaluated numerically, such that likelihood-free inference techniques are promising alternatives. Fast simulation of large numbers of samples of the model is required in likelihood-free estimation methods such as ABC, and in other simulation-based inference methods such as maximum simulated likelihood (e.g., [Hajivassiliou, 2000](#)).

## References

- Bopp, G., Shaby, B. A. and Huser, R. (2020) A hierarchical max-infinitely divisible spatial model for extreme precipitation. Journal of American Statistical Association To appear.
- Brown, B. M. and Resnick, S. I. (1977) Extreme values of independent stochastic processes. Journal of Applied Probability **14**, 732–739.
- Castruccio, S., Huser, R. and Genton, M. G. (2016) High-order composite likelihood inference for max-stable distributions and processes. Journal of Computational and Graphical Statistics **25**, 1212–129.
- Davison, A. C., Huser, R. and Thibaud, E. (2019) Spatial extremes. In Handbook of Environmental and Ecological Statistics, eds A. E. Gelfand, M. Fuentes, J. A. Hoeting and R. L. Smith, pp. 711–744. CRC Press.
- Davison, A. C., Padoan, S. and Ribatet, M. (2012) Statistical modelling of spatial extremes (with Discussion). Statistical Science **27**, 161–186.
- Dey, D. K. and Yan, J. (2016) Extreme value modeling and risk analysis: methods and applications. CRC Press.
- Dieker, A. B. and Mikosch, T. (2015) Exact simulation of brown-resnick random fields at a finite number of locations. Extremes **18**, 301–314.
- Dombry, C., Engelke, S. and Oesting, M. (2016) Exact simulation of max-stable processes. Biometrika **103**, 303–317.
- Dombry, C. and Eyi-Minko, F. (2013) Regular conditional distributions of continuous max-infinitely divisible random fields. Electronic Journal of Probability **18**, 1–21.
- Erhardt, R. J. and Smith, R. L. (2012) Approximate Bayesian computing for spatial extremes. Computational Statistics & Data Analysis **56**, 1468–1481.
- Gilks, W. R. and Wild, P. (1992) Adaptive rejection sampling for gibbs sampling. Journal of the Royal Statistical Society. Series C (Applied Statistics) **41**, 337–348.

- Giné, E., Hahn, M. G. and Vatan, P. (1990) Max-infinitely divisible and max-stable sample continuous processes. Probability Theory and Related Fields **87**, 139–165.
- de Haan, L. (1984) A spectral representation for max-stable processes. The Annals of Probability **12**, 1194–1204.
- Hainy, M., Müller, W. G. and Wagner, H. (2016) Likelihood-free simulation-based optimal design with an application to spatial extremes. Stochastic Environmental Research and Risk Assessment **30**, 481–492.
- Hajivassiliou, V. (2000) Some practical issues in maximum simulated likelihood. Simulation-Based inference in econometrics: Methods and Applications pp. 71–99.
- Huser, R., Opitz, T. and Thibaud, E. (2020) Max-infinitely divisible models and inference for spatial extremes. Scandinavian Journal of Statistics pp. 1–28.
- Huser, R. and Wadsworth, J. L. (2020) Advances in statistical modeling of spatial extremes. WIREs Computational Statistics p. e1537.
- Kabluchko, Z., Schlather, M. and de Haan, L. (2009) Stationary max-stable fields associated to negative definite functions. Annals of Probability **37**, 2042–2065.
- Kullback, S. and Leibler, R. A. (1951) On information and sufficiency. Annals of Mathematical Statistics **22**, 79–86.
- Lee, X. J., Hainy, M., McKeone, J. P., Drovandi, C. C. and Pettitt, A. N. (2018) ABC model selection for spatial extremes models applied to South Australian maximum temperature data. Computational Statistics & Data Analysis **128**, 128–144.
- Oesting, M., Schlather, M. and Zhou, C. (2018) Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation. Bernoulli **24**, 1497–1530.
- Oesting, O. and Storkorb, K. (2019) A comparative tour through the simulation algorithms for max-stable processes. arXiv preprint 1809.09042.
- Opitz, T. (2013) Extremal  $t$  processes: Elliptical domain of attraction and a spectral representation. Journal of Multivariate Analysis **122**, 409–413.
- Padoan, S. A. (2013) Extreme dependence models based on event magnitude. Journal of Multivariate Analysis **122**, 1 – 19.
- Padoan, S. A., Ribatet, M. and Sisson, S. A. (2010) Likelihood-based inference for max-stable processes. Journal of the American Statistical Association **105**, 263–277.
- Resnick, S. I. (1987) Extreme Values, Regular Variation and Point Processes. Springer-Verlag New York.

- Robert, C. and Casella, G. (2013) Monte Carlo statistical methods. Springer Science & Business Media.
- Schlather, M. (2002) Models for stationary max-stable random fields. Extremes **5**, 33–44.
- Varadhan, R. (2015) condMVNorm: Conditional Multivariate Normal Distribution. R package version 2015.2-1.
- Wadsworth, J. L. and Tawn, J. A. (2014) Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. Biometrika **101**, 1–15.
- Zhong, P., Huser, R. and Opitz, T. (2020) Modeling non-stationary temperature maxima based on extremal dependence changing with event magnitude. arXiv preprint 2006.01569.