Exact Simulation of Max-Infinitely Divisible Processes

October 21, 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 function 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.

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., 2021; Huser and Wadsworth, 2020; Zhong et al., 2021). 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. (2021) and Zhong et al. (2021), 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 $Z = \{Z(\boldsymbol{s})\}_{\boldsymbol{s} \in \mathcal{S}}$ 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 $\{\eta_i;\ i = 1, 2, \ldots\}$ defined over $\mathcal{S}$, which are sampled according to a PPP with mean measure Λ, i.e.,

$$Z(\boldsymbol{s}) = \max_{i=1,2,\ldots} \eta_i(\boldsymbol{s}), \quad \boldsymbol{s} \in \mathcal{S}, \tag{1}$$

where, by convention, $Z(\boldsymbol{s})$ takes value at the lower boundary of the support of $\eta_i(\boldsymbol{s})$ when the PPP $\{\eta_i\}$ contains no point (which can happen when Λ is finite). To avoid intricate theoretical issues dealing with the lower boundary, we here assume that Λ is an infinite measure, such that the number of Poisson points η_i in (1) is infinite. However, analytical forms of the multivariate density 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 $Z(\boldsymbol{s})$ in (1) is not trivial, especially when the number of functions η_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, 2021). 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., $\Pr\{Z(\boldsymbol{s}) \leq z\} = \exp(-1/z)$, $z > 0$) 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 $\eta_i(\boldsymbol{s}) = R_i W_i(\boldsymbol{s})$. Precisely, $\{R_i; \ i = 1, 2, \dots\}$ here denotes a Poisson point process on $[0, \infty)$ with mean measure $r^{-2}\mathrm{d}r$, independent of the random functions $\{W_i; \ i = 1, 2, \dots\}$, which are independent copies of a random process $W(\boldsymbol{s})$ defined on $\mathcal{S}$ such that $\mathrm{E}[\max\{0, W(\boldsymbol{s})\}] = 1$. 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 $\alpha r^{-(\alpha+1)}\mathrm{d}r$ for some $\alpha > 0$, the resulting process Z remains max-stable, but with α-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., $R_1 > R_2 > \ldots$. When this is the case, the componentwise maximum of $\{R_j W_j; \; j = 1, \ldots, i-1\}$ is less and less impacted by the i-th process $\eta_i = R_i W_i$ as $i > 1$ is iteratively incremented. If $W(\boldsymbol{s})$ is an almost surely bounded random process on $\mathcal{S}$, the contribution of $\eta_i = R_i W_i$ becomes even ultimately completely irrelevant, as i grows. 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 $W(\boldsymbol{s})$ 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 $\eta_i(\boldsymbol{s})$ given its value at a single location $\boldsymbol{s}_0$ 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 $\eta_i(\boldsymbol{s})$ 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 $\{Z(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ is called max-infinitely divisible (max-id) if the joint cumulative distribution function (cdf) G of $\boldsymbol{Z}(K_N) = \{Z(\boldsymbol{s}_1),\ldots,Z(\boldsymbol{s}_N)\}^\top$ at any finite collection of sites $K_N = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N\} \subset \mathcal{S}$ defines a valid cdf G^t (with the notation $G^t(\boldsymbol{z}) = \{G(\boldsymbol{z})\}^t$, $\boldsymbol{z} = (z_1,\ldots,z_N)^\top \in \mathbb{R}^N$) 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 $\tilde{G}^m(\boldsymbol{a}_m\boldsymbol{z} + \boldsymbol{b}_m) = \tilde{G}(\boldsymbol{z})$, for any integer $m = 1, 2, \ldots$, and appropriately defined normalizing vectors $\boldsymbol{a}_m \in (0,\infty)^N$ and $\boldsymbol{b}_m \in \mathbb{R}^N$. 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 function space. We assume that $\{Z(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ is a sample-continuous max-id process in a function space $\mathcal{C}$, where $\mathcal{C}$ denotes the space of continuous functions with compact support $\mathcal{S} \subset \mathbb{R}^d$ endowed with the uniform norm, i.e., $\|f\| = \sup_{\boldsymbol{s}\in\mathcal{S}} f(\boldsymbol{s}), f \in \mathcal{C}$, and $\{\eta_i(\boldsymbol{s});\ i = 1, 2, \ldots\}_{\boldsymbol{s}\in\mathcal{S}}$ in (1) is a Poisson point process with mean measure Λ on $\mathcal{C}$. 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 $\boldsymbol{s} \in K$ such that $Z(\boldsymbol{s}) = \phi(\boldsymbol{s})$, otherwise ϕ is called K-subextremal. We denote $\mathcal{C}_K^+(Z)$ the set of K-extremal functions and $\mathcal{C}_K^-(Z)$ the set of K-subextremal functions.*

In other words, an extremal function is a Poisson point η_i in $\mathcal{C}$ that contributes to the maximum process Z at one or more locations. Notice that both $\mathcal{C}_K^+(Z)$ and $\mathcal{C}_K^-(Z)$ depend on the max-id process Z, even though the realization of the max-id process Z on K is fully

determined by $\mathcal{C}_K^+(Z)$ alone. Let ℓ be the vertex function (i.e., the lower boundary) of the max-id process Z, defined by

$$\ell(\boldsymbol{s}) = \sup\{z \in \mathbb{R} : \Pr(Z(\boldsymbol{s}) \geq z) = 1\} \in [-\infty, \infty), \quad \boldsymbol{s} \in \mathcal{S}. \tag{2}$$

Giné et al. (1990) showed that for sample continuous max-id processes in $\mathcal{C}$, if the vertex function ℓ is continuous, then it can be subtracted from Z. Therefore, we here assume $\ell(\boldsymbol{s}) = 0$ without loss of generality, and denote $\mathcal{C}_0 = \{f \in \mathcal{C}; f \neq 0, f \geq 0\}$. We further assume that for any fixed $\boldsymbol{s}_0 \in \mathcal{S}$,

$$\Lambda_{\boldsymbol{s}_0}(z) = \Lambda(\{\eta \in \mathcal{C}_0; \eta(\boldsymbol{s}_0) \in [z, \infty)\}) \text{ is continuous on } (0, \infty), \lim_{z \downarrow 0} \Lambda_{\boldsymbol{s}_0}(z) = \infty, \tag{3}$$

$$\Lambda_{\boldsymbol{s}_0}(\{\eta \in \mathcal{C}_0; \eta(\boldsymbol{s}_0) > \varepsilon\}) < \infty, \forall \; \varepsilon > 0. \tag{4}$$

Under these assumptions, Dombry and Eyi-Minko (2013) showed that for singletons $K = \{\boldsymbol{s}_0\}$ the set of extremal functions $\mathcal{C}_{\{\boldsymbol{s}_0\}}^+$ contains almost surely a single point (a function), denoted by $\phi_{\boldsymbol{s}_0}^+$, i.e., the resulting max-id "process" $Z(\boldsymbol{s}_0)$ is realized by one and only one Poisson point η_i at a given location $\boldsymbol{s}_0$. This assumption implies that Z has continuous margins and the marginal distribution has no mass at the lower boundary $\ell(\boldsymbol{s}) = 0$.

Under the standing assumptions, Dombry and Eyi-Minko (2013, Lemma A.2) show that the mean measure Λ possesses a regular conditional probability measure $P_{\boldsymbol{s}_0}(z, \cdot)$ with respect to the location $\boldsymbol{s}_0$; i.e., given that $\eta_i(\boldsymbol{s}_0) = z > 0$ for a point η_i of the Poisson point process, the random function $\eta_i(\boldsymbol{s})$ for $\boldsymbol{s} \neq \boldsymbol{s}_0$ has (conditional) probability distribution given by $P_{\boldsymbol{s}_0}(z, \cdot)$. Dombry and Eyi-Minko (2013) further show in their Theorem 3.2 that $P_{\boldsymbol{s}_0}(z, \cdot)$ is also the conditional distribution for the extremal function $\phi_{\boldsymbol{s}_0}^+$ at $\boldsymbol{s}_0$ given $Z(\boldsymbol{s}_0) = z > 0$; moreover, conditionally on the values $Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)$ of the max-id process at $K = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$, the subextremal functions in $\mathcal{C}_K^-(Z)$ define a Poisson point process that is independent of the extremal functions in $\mathcal{C}_K^+(Z)$. This restricted Poisson point process has again mean measure Λ but now with all the mass "removed" when one of the conditioning values $Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)$ is exceeded, i.e., it has no mass on $\{\eta \in \mathcal{C}_0 : \max_{i=1}^n \eta(\boldsymbol{s}_i)/Z(\boldsymbol{s}_i) > 1\}$. This remarkable result is the key for the proposed simulation algorithm that consists in

6

iteratively simulating extremal functions at a set of locations (see §3.1), and we summarize it in the following lemma.

Lemma 2. *(Dombry and Eyi-Minko, 2013, Theorem 3.2) Given the extremal functions $\mathcal{C}_K^+(Z)$ on K, the subextremal functions $\mathcal{C}_K^-(Z)$ form a Poisson point process on $\mathcal{C}_0$ with intensity $\mathbb{I}(f(K) < Z(K))\Lambda(\mathrm{d}f)$, where $\mathbb{I}(\cdot)$ is the indicator function.*

In particular, when the process $\{Z(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ is max-stable with unit Fréchet margins, the Poisson point process $\{\eta_i;\ i=1,2,\dots\}$ can be decomposed as $\eta_i(\boldsymbol{s}) = R_i W_i(\boldsymbol{s})$, $i = 1,2,\dots$, 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 $\{W_i(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$, $i = 1,2,\dots$, are independent copies of a random process $\{W(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ with $\mathrm{E}[\max\{W(\boldsymbol{s}),0\}] = 1$, 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 Λ 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 $P_{\boldsymbol{s}}(z,\cdot)$ 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., 2021). Therefore, simulating from $P_{\boldsymbol{s}}(z,\cdot)$ 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. While our proposed algorithm in the max-id case is almost identical in its overall idea and general structure to Dombry et al.'s algorithm in the max-stable case (modulo certain technicalities and specificities in how the different steps are executed), it is still important to note that generalizing a specific simulation algorithm that works for max-stable processes to the much wider class of max-id models, while remaining an exact and feasible simulation procedure, is not necessarily trivial and requires careful investigation. This has never been explored before.

Our aim is, thus, to simulate the max-id process $\{Z(\boldsymbol{s})\}_{s \in \mathcal{S}}$ at a the finite number N of sites given as $K_N = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N\} \subset \mathcal{S}$. We define the subsets $K_n = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ of the first n sites for $n = 1, \ldots, N$, 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 $\boldsymbol{Z}(K_N) = \boldsymbol{Z}_N(K_N) := \{Z_N(\boldsymbol{s}_1), \ldots, Z_N(\boldsymbol{s}_N)\}^\top$. To simulate from the joint distribution of the process $\{Z(\boldsymbol{s})\}$ at sites in K_N, we can iterate through the set of locations from $\boldsymbol{s}_1$ to $\boldsymbol{s}_N$ in order to simulate the extremal functions at each location. More precisely, suppose we have already simulated the extremal functions on K_n. Recall that we denote by $\{\phi_{\boldsymbol{s}_n}^+\}_{1 \leq n \leq N}$ the sequence of extremal functions associated to the singletons $\{\boldsymbol{s}_n\}$, $n = 1, \ldots, N$, where some of these functions may be identical to each other. The extremal function $\phi_{\boldsymbol{s}_{n+1}}^+$ at a new location $\boldsymbol{s}_{n+1} \notin K_n$ either coincides with one of the extremal functions on K_n already simulated, i.e., $\phi_{\boldsymbol{s}_{n+1}}^+ \in \mathcal{C}_{K_n}^+(Z)$, or it corresponds to a subextremal function on K_n, i.e., $\phi_{\boldsymbol{s}_{n+1}}^+ \in \mathcal{C}_{K_n}^-(Z)$, in which case it is a point of the restricted Poisson point process defined in Lemma 2. Moreover, at $\boldsymbol{s}_{n+1}$ the maximum of the values of the extremal functions on K_n, i.e., $Z_n(\boldsymbol{s}_{n+1})$, already provides a lower bound for the values of the new extremal function that could arise at $\boldsymbol{s}_{n+1}$. Therefore, we proceed as follows to find the extremal function $\phi_{\boldsymbol{s}_{n+1}}^+$, which directly provides $Z(\boldsymbol{s}_{n+1})$. We simulate from the unrestricted point process by assuming that its points are ordered in

descending order of the values at $\boldsymbol{s}_{n+1}$, i.e., $\eta_1(\boldsymbol{s}_{n+1}) > \eta_2(\boldsymbol{s}_{n+1}) > \ldots$ Simulation starts from the point η_1 with the largest value at $\boldsymbol{s}_{n+1}$ and then continues in descending order. To keep only the points from the restricted point process of the subextremal functions of K_n, we reject the points violating the upper bound conditions at $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ set by the restrictions given in Lemma 2. We simulate a point η_i by first simulating its value $\eta_i(\boldsymbol{s}_{n+1})$ at $\boldsymbol{s}_{n+1}$ according to the marginal (unrestricted) Poisson point process with mean measure $\Lambda_{\boldsymbol{s}_{n+1}}$, and we then simulate the values at all the other sites in K_N using the regular conditional probability distribution $P_{\boldsymbol{s}_{n+1}}(\eta_i(\boldsymbol{s}_{n+1}), \cdot)$. Moreover, we can stop simulating as soon as a value $\eta_i(\boldsymbol{s}_{n+1})$ falls below the lower bound $Z_n(\boldsymbol{s}_{n+1})$ and we reject it, since this value (as well as all the following smaller ones) cannot correspond to the extremal function at $\boldsymbol{s}_{n+1}$. In any case, we stop as soon as we obtain a point η_i that is not rejected, and this point then corresponds to $\phi_{\boldsymbol{s}_{n+1}}^+$. If we have rejected all points until we stop, there is no new extremal function at $\boldsymbol{s}_{n+1}$, such that $\phi_{\boldsymbol{s}_{n+1}}^+ \in \mathcal{C}_{K_n}^+(Z)$.

The benefit of this approach is that it requires to simulate only a finite number of Poisson points to either obtain a new extremal function at $\boldsymbol{s}_{n+1}$, or to confirm that the extremal function at $\boldsymbol{s}_{n+1}$ coincides with one of the already simulated extremal functions on K_n. Therefore, the approach exploits the fact that the already simulated extremal functions provide upper bounds at $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ and a lower bound at $\boldsymbol{s}_{n+1}$. This idea then allows us to iteratively simulate all of the extremal functions at a finite set of locations $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N$, where we condition on the extremal functions at $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ when identifying the extremal function at $\boldsymbol{s}_{n+1}$, $1 \leq n < N$. Since the extremal functions define the values of the max-id process at the locations $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N$, this is all we need. Dombry et al. (2016) have proposed this approach for max-stable processes, but their algorithm generalizes straightforwardly to max-id processes thanks to the general results of Dombry and Eyi-Minko (2013). The following theorem formalizes the simulation procedure detailed above by characterizing the conditional distributions of $\phi_{\boldsymbol{s}_n}$ given $\{\phi_{\boldsymbol{s}_k}^+\}_{1 \leq k \leq n-1}$ for extremal functions at position $n = 2, \ldots, N - 1$, which are required for iterative simulation.

Theorem 3 (Conditional distributions of extremal functions). *Consider a max-id process* $\{Z(\boldsymbol{s})\}_{s \in \mathcal{S}}$ *as defined in* §2. *Given sites* $K_N = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_N\}$, *the distribution of its extremal function* $\phi^+_{\boldsymbol{s}_1}$ *and of its conditional extremal functions* $\phi^+_{\boldsymbol{s}_{n+1}} \mid \{\phi^+_{\boldsymbol{s}_k}\}_{1 \le k \le n}$ *for* $n \ge 1$ *can be characterized as follows:*

(1) *Initial extremal function: given a realization* z *of* $Z(\boldsymbol{s}_1)$, *the extremal function* $\phi^+_{\boldsymbol{s}_1}$ *is distributed according to* $P_{\boldsymbol{s}_1}(z, \cdot)$.

(2) *Conditional extremal functions for* $n \ge 1$: *given the extremal functions* $\{\phi^+_{\boldsymbol{s}_i}\}_{1 \le i \le n}$ *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* $\boldsymbol{s}_{n+1}$:

$$\tilde{\Lambda}(\mathrm{d}f) = \mathbb{I}\left(f(\boldsymbol{s}_i) < Z_n(\boldsymbol{s}_i), 1 \le i \le n\right) \mathbb{I}\left(f(\boldsymbol{s}_{n+1}) > Z_n(\boldsymbol{s}_{n+1})\right) \Lambda(\mathrm{d}f). \tag{5}$$

Then, we denote the Poisson point process with mean measure $\tilde{\Lambda}$ *by* $\mathrm{PPP}(\tilde{\Lambda})$, *and the extremal function* $\phi^+_{\boldsymbol{s}_{n+1}}$ *conditional on* $\{\phi^+_{\boldsymbol{s}_i}\}_{1 \le i \le n}$ *is given as follows:*

$$\tilde{\phi}^+_{\boldsymbol{s}_{n+1}} = \begin{cases} \arg\max_{\phi \in \mathrm{PPP}(\tilde{\Lambda})} \phi(\boldsymbol{s}_{n+1}), & \mathrm{PPP}(\tilde{\Lambda}) \ne \emptyset, \\ \arg\max_{\phi \in \{\phi^+_{\boldsymbol{s}_1}, \ldots, \phi^+_{\boldsymbol{s}_n}\}} \phi(\boldsymbol{s}_{n+1}), & \mathrm{PPP}(\tilde{\Lambda}) = \emptyset, \end{cases}$$

Proof. The proof essentially follows the intuitive description of the algorithm above. It is almost identical (modulo some technicalities) to the proof of Theorem 2 in Dombry et al. (2016), which concerns max-stable processes, while we here treat max-id processes.

As mentioned, we proceed sequentially to simulate extremal functions. The distribution of the first extremal function $\phi^+_{\boldsymbol{s}_1}$ given $\phi^+_{\boldsymbol{s}_1}(\boldsymbol{s}_1) = z$ is exactly $P_{\boldsymbol{s}_1}(z, \cdot)$, as explained in §2 above, and $\phi^+_{\boldsymbol{s}_1}(\boldsymbol{s}_1)$ thus follows the marginal distribution of $Z(\boldsymbol{s}_1)$, which has now a more general form than for max-stable processes. Notice also that given the continuity assumption in (3), there is almost surely exactly one extremal function at each location.

Then, to simulate the extremal function $\phi^+_{\boldsymbol{s}_{n+1}}$ at site $\boldsymbol{s}_{n+1}$, for $n \ge 1$, we recall that $\phi^+_{\boldsymbol{s}_{n+1}}$ is either a subextremal function in $\mathcal{C}^-_{K_n}$ or an extremal function in $\mathcal{C}^+_{K_n}$. According to

Lemma 2, given $\mathcal{C}_{K_n}^+$, the conditional distribution of $\mathcal{C}_{K_n}^-$ follows the distribution of a Poisson point process with intensity

$$\mathbb{I}\left(f(\boldsymbol{s}_i) < Z_n(\boldsymbol{s}_i), 1 \leq i \leq n\right) \Lambda(\mathrm{d}f). \tag{6}$$

If $\phi_{\boldsymbol{s}_{n+1}}^+ \in \mathcal{C}_{K_n}^-$, we need $\phi_{\boldsymbol{s}_{n+1}}^+(\boldsymbol{s}_{n+1}) > Z_n(\boldsymbol{s}_{n+1})$, such that $\phi_{\boldsymbol{s}_{n+1}}^+$ is the extremal function at $\boldsymbol{s}_{n+1}$. Therefore, we can further restrict the Poisson process with intensity in (6) and consider the Poisson process $\mathrm{PPP}(\tilde{\Lambda}) = \mathcal{C}_{K_n}^- \cap \{f \in \mathcal{C}_0 : f(\boldsymbol{s}_{n+1}) > Z_n(\boldsymbol{s}_{n+1})\}$, which has the mean measure $\tilde{\Lambda}$ given in Equation (5). If there is no point in $\mathrm{PPP}(\tilde{\Lambda})$, i.e. $\phi_{\boldsymbol{s}_{n+1}}^+ \notin \mathcal{C}_{K_n}^- \cap \{f \in \mathcal{C}_0 : f(\boldsymbol{s}_{n+1}) > Z_n(\boldsymbol{s}_{n+1})\}$, it must be true that $\phi_{\boldsymbol{s}_{n+1}}^+ \in \mathcal{C}_{K_n}^+$ and $\phi_{\boldsymbol{s}_{n+1}}^+ = \arg\max_{\phi \in \{\phi_{\boldsymbol{s}_1}^+, \dots, \phi_{\boldsymbol{s}_n}^+\}} \phi(\boldsymbol{s}_{n+1})$. Otherwise, $\phi_{\boldsymbol{s}_{n+1}}^+(\boldsymbol{s}_{n+1})$ is a point of $\mathrm{PPP}(\tilde{\Lambda}) \neq \emptyset$, and it can be identified according to the procedure outlined above: we simulate from the unrestricted Poisson process and reject the points violating the restriction in Equation (5) until either the restriction is satisfied or the simulated value falls below the lower bound $Z_n(\boldsymbol{s}_{n+1})$. $\quad\square$

The pseudo code of an exact simulation algorithm based on Theorem 3 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 $P_{\boldsymbol{s}}(z, \cdot)$ given $Z(\boldsymbol{s}) = z$. For suitable constructions of the PPP $\{\eta_i; i = 1, 2, \dots\}$, we can investigate the specific structure of this conditional distribution. In this paper, we consider two different and very general structures for η_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 η_i as a Gaussian scale mixture

3.2.1 Model description

One possible structure of the general max-id process is achieved by representing η_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

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

Input: Stationary marginal distribution G_0 of Z.
Input: Dependence characterization of Z allowing simulation of $P_{\boldsymbol{s}_n}$, $n = 1, \ldots, N$.
 1: Simulate $E_0 \sim \text{Exp}(1)$ and set $z = G_0^{-1}(\exp(-E_0))$.
 2: Simulate $Y \sim P_{\boldsymbol{s}_1}(z, \cdot)$ over K_N, which yields $\boldsymbol{Y}(K_N) = \{Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_N)\}^\top$.
 3: Set $\boldsymbol{Z}(K_N) = \boldsymbol{Y}(K_N)$.
 4: **for all** $n = 2, \ldots, N$ **do**
 5:     Simulate $E_0 \sim \text{Exp}(1)$ and set $z = G_0^{-1}(\exp(-E_0))$.
 6:     **while** $\{z > Z(\boldsymbol{s}_n)\}$ **do**
 7:         Simulate $Y \sim P_{\boldsymbol{s}_n}(z, \cdot)$ over K_N, which yields $\boldsymbol{Y}(K_N) = \{Y(\boldsymbol{s}_1), \ldots, Y(\boldsymbol{s}_N)\}^\top$.
 8:         **if** $Y(\boldsymbol{s}_i) < Z(\boldsymbol{s}_i)$ for all $i = 1, \ldots, n-1$ **then**
 9:             update simulated values as $\boldsymbol{Z}(K_N) = \max(\boldsymbol{Z}(K_N), \boldsymbol{Y}(K_N))$ (componentwise).
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 $\boldsymbol{Z}(K_N)$

on $\mathbb{R}^+$ with mean measure κ such that $\kappa([0, \infty)) = \infty$, and $\{W_i\}$ are independent standard Gaussian random fields with mean zero and unit variance on $\mathcal{S}$, which are 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 Λ in (1) 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. (2021) further relaxed the independence assumption between $\{R_i\}$ and $\{W_i\}$. Specifically, they proposed a Weibull-tailed mean measure κ for $\{R_i\}$, where

$$\kappa(r) := \kappa\left([r, \infty)\right) = r^{-\beta} \exp\{-\alpha(r^\beta - 1)/\beta\}, \quad r > 0, \quad (\alpha, \beta)^\top \in (0, \infty)^2. \tag{7}$$

With this specification, $\kappa(r) \to r^{-\alpha}$ as $\beta \downarrow 0$. Therefore, the model remains in the "neighborhood" of max-stable processes for which $\kappa(r) = r^{-\alpha}$. Recall that the extremal-t max-stable process with α-Fréchet margins can indeed be obtained by taking $\eta_i = R_i W_i$ in (1) with $\{R_i\}$ a Poisson point process on $(0, \infty)$ with mean measure $\kappa(r) = r^{-\alpha}$, and $\{W_i\}$, independent zero mean Gaussian processes, that are rescaled such that $\text{E}[\max\{0, W_i(\boldsymbol{s})\}] = 1$, and that are also independent of $\{R_i\}$. In other words, if $\{R_i\}$ and $\{W_i\}$ are independent of each other and

W_i is Gaussian, then the max-id model constructed from (7) can be arbitrarily close to the max-stable extremal-t dependence structure by varying the parameter β. Given $\{R_i\}$, a natural choice is, thus, to specify the processes $\{W_i\}$ to be standard Gaussian random fields with some correlation function $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; R_i)$, 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 $\lim_{u \to 1} \Pr[G_0\{Z(\boldsymbol{s}_1)\} > u \mid G_0\{Z(\boldsymbol{s}_2)\} > u] = 0$ with G_0 being the marginal distribution of Z. For example, Zhong et al. (2021) chose a non-stationary correlation $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; R_i)$, which may be specified as

$$\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; R_i) = (\lambda_{\boldsymbol{s}_1} \lambda_{\boldsymbol{s}_2}) \left(\frac{\lambda_{\boldsymbol{s}_1}^2 + \lambda_{\boldsymbol{s}_2}^2}{2} \right)^{-1} \exp \left\{ - \left(\frac{\lambda_{\boldsymbol{s}_1}^2 + \lambda_{\boldsymbol{s}_2}^2}{2} \right)^{-1/2} (1 + R_i)^{\nu} \|\boldsymbol{s}_1 - \boldsymbol{s}_2\| \right\}, \quad (8)$$

where $\lambda_{\boldsymbol{s}} > 0$ 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 particular, when $\nu = 0$, R_i and W_i are independent. In the simulation examples below, we will illustrate some specific structures for $\lambda_{\boldsymbol{s}}$. 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 mean measure of $\{\eta_i; i = 1, 2, \dots\}$ at the finite set of locations in $K \equiv K_N$ may be written as

$$\Lambda_K([\boldsymbol{0}, \boldsymbol{z}_K]^C) = \int_0^\infty \{1 - \Phi(\boldsymbol{z}_K/r; \boldsymbol{\rho}(K, K; r))\} \kappa(\mathrm{d}r), \quad \boldsymbol{z}_K > \boldsymbol{0}, \quad (9)$$

where $\Phi(\cdot; \boldsymbol{\rho}(K, K; r))$ is the multivariate normal distribution function with zero mean and correlation matrix $\boldsymbol{\rho}(K, K; r)$ determined by the correlation function $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; r)$ for $\boldsymbol{s}_1, \boldsymbol{s}_2 \in K$, and $[\boldsymbol{0}, \boldsymbol{z}_K]^C \equiv ([0, z_1] \times \cdots \times [0, z_N])^C$. Huser et al. (2020) showed that the marginal mean measure $\Lambda_{\boldsymbol{s}_0}$ is an infinite continuous measure and locally finite, thus it satisfies the assumptions in (3) stated in §2. The intensity function $g_{\Lambda_K}(\boldsymbol{z}_K)$ of Λ_K can be written as follows, where we use notation φ for the multivariate Gaussian density corresponding to Φ in (9):

$$g_{\Lambda_K}(\boldsymbol{z}_K) = \int_0^\infty r^{-N} \varphi(\boldsymbol{z}_K/r; \boldsymbol{\rho}(K, K; r)) \kappa(\mathrm{d}r), \quad \boldsymbol{z}_K > \boldsymbol{0}. \quad (10)$$

For marginal intensities arising for singletons $K = \{\boldsymbol{s}\}$, we write $g_{\Lambda_{\boldsymbol{s}}}$. The density function of the probability measure $P_{\boldsymbol{s}_1}(z, \cdot)$ is then given by the conditional intensity

$$\boldsymbol{z}_{K \setminus \{\boldsymbol{s}_1\}} \mapsto g_{\Lambda_K}(z, \boldsymbol{z}_{K \setminus \{\boldsymbol{s}_1\}})/g_{\boldsymbol{s}_1}(z) \propto g_{\Lambda_K}(z, \boldsymbol{z}_{K \setminus \{\boldsymbol{s}_1\}}). \tag{11}$$

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

$$\theta_N(z_0) = \frac{\log\{G(\boldsymbol{z}_0)\}}{\log\{G_0(z_0)\}} \tag{12}$$

where $\boldsymbol{z}_0 = (z_0, \ldots, z_0)^T \in \mathbb{R}^N$, G_0 is the marginal distribution of Z and G is the joint distribution of the random vector $\boldsymbol{Z}(K_N)$ with observations at the locations in the set $K \subset \mathcal{S}$. This coefficient can be expressed in terms of the measure Λ, 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 (12) it is easy to see that $\Pr\{\boldsymbol{Z}(K_N) \leq \boldsymbol{z}_0\} = G_0(z_0)^{\theta_N(z_0)}$, such that this coefficient can be interpreted as the effective number of independent variables among the components of the vector $\boldsymbol{Z}(K_N)$ 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 on the event magnitude R_i chosen as $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; R_i) = \exp\{-2(1 + R_i)^\nu \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|\}$, which corresponds to the stationary and isotropic counterpart of (8) with spatially-constant $\lambda_{\boldsymbol{s}} = 1/2$. 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 ν and β are both zero. In this case, the max-id process is the max-stable extremal-t process. When β is different from zero, we get asymptotic independence (since $\theta_2(z_0) \to 2$ when $z \to \infty$), and the
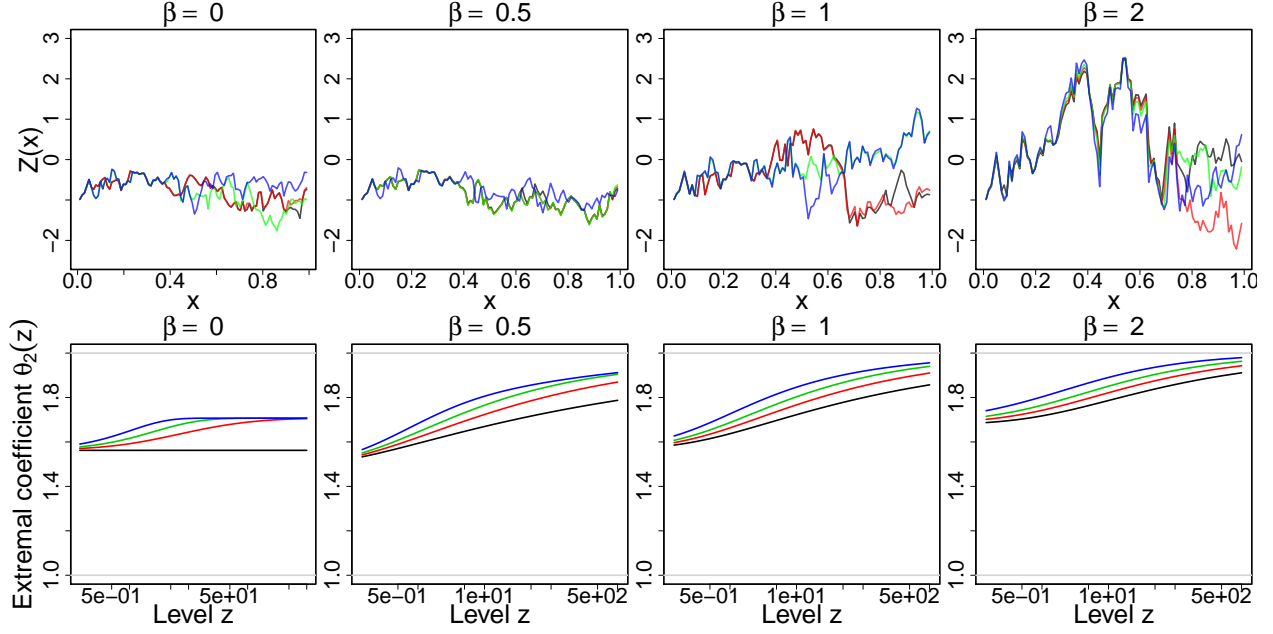
14

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 ν controls the dependence between R_i and W_i by using correlation function $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2; R_i) = \exp\left\{-2(1 + R_i)^\nu \|\boldsymbol{s}_1 - \boldsymbol{s}_2\|\right\}$ conditional on R_i. The level z is shown on a logarithmic scale.

dependence level decreases as the values of β and ν increase. The realizations were generated using the same random seed, and we can discern stronger fluctuations of the realizations as the value of β 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 $P_{\boldsymbol{s}_1}(z, \cdot)$ at a finite set of locations K (need in lines 2 and 7 of Algorithm 1) for the max-id model constructed from Gaussian scale mixtures $\eta_i = R_i W_i$ as described in §3.2.1. Intuitively speaking, we first sample a random variable $\tilde{R}_z$ whose probability distribution corresponds to the intensity of R_i conditional on $\eta_i(\boldsymbol{s}_1) = R_i W_i(\boldsymbol{s}_1) = z$, and in the second step we sample the Gaussian field W_i at locations in K conditional on $\eta_i(\boldsymbol{s}_1) = z$ and $\tilde{R}_z$, i.e., we sample a Gaussian random vector conditional on its value $z/\tilde{R}_z$ at $\boldsymbol{s}_1$.

To obtain the distribution of $\tilde{R}_z$ in the first step, we consider the joint intensity function of points $\{R_i\}$ and $\{W_i(\boldsymbol{s}_1)\}$ given by $(r, w) \mapsto \kappa(\mathrm{d}r)\varphi(w)$. Next, we perform a change of variables from (r, w) to $(r, z) = (r, rw)$, which yields the intensity $g_{R,\eta}(r, z) = \kappa(\mathrm{d}r)\varphi(z/r)/r$. Finally, we obtain the desired conditional density as $g_{\tilde{R}_z}(r; z) = g_{R,\eta}(r, z)/g_{\Lambda_{\boldsymbol{s}_1}}(z)$. Note that this defines indeed a valid probability density function since it corresponds to a conditional distribution of $P_{\boldsymbol{s}_1}(z, \cdot)$, which has been shown to be a proper probability measure.

In the second step, conditional on $\tilde{R}_z$, we write $\tilde{\eta}(K)$ for the conditional Gaussian random vector with mean vector $\boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)z$, common variance $\tilde{R}_z^2$, and correlation matrix $\boldsymbol{\rho}(K, K; \tilde{R}_z) - \boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)\boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)^\top$, where $\boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)$ denotes the column vector of correlation values $\rho(\boldsymbol{s}_1, \boldsymbol{s}_j; \tilde{R}_z)$, $j = 1, \ldots, N$, and $\boldsymbol{\rho}(K, K; \tilde{R}_z)$ is a matrix with entries $\rho(\boldsymbol{s}_i, \boldsymbol{s}_j; \tilde{R}_z)$, $i = 1, \ldots, N$, $j = 1, \ldots, N$. We denote this joint Gaussian density function as $g_{\tilde{\eta}(K)|\tilde{R}_z}$. Note that, after integrating out $\tilde{R}_z$, it is easy to formally show that the (unconditional) distribution of $\tilde{\eta}(K)$ is

$$\tilde{\eta}(K) \sim P_{\boldsymbol{s}_1}(z, \cdot).$$

Simulation from $P_{\boldsymbol{s}_1}(z, \cdot)$ at locations in K can thus be performed using a two-step procedure, where we need to simulate $\tilde{R}_z$ first, and conditional on $\tilde{R}_z$, we then simulate $\tilde{\eta}(K)$ from a multivariate Gaussian distribution. We here explore the use of two simulation techniques where $g_{\tilde{R}_z}(r; z)$ only needs to be known up to a constant that depends on the value of z. Precisely, simulation from $P_{\boldsymbol{s}_1}(z, \cdot)$ proceeds according to the following successive steps:

(1) Simulate a value $\tilde{R}_z$ from the density function $g_{\tilde{R}_z}(r; z)$, $r > 0$, 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 $\tilde{R}_z$, simulate a random vector $\tilde{\eta}(K)$ from the multivariate Gaussian density function $g_{\tilde{\eta}(K)|\tilde{R}_z}$.

(3) Return $\tilde{\eta}(K)$.

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 for an exact simulation procedure lies in the first step, where we have to simulate from the probability density $g_{\tilde{R}_z}$, known up to a constant only.

3.2.3 Options for simulating $\tilde{R}_z$

With the Metropolis–Hastings (MH) method in Step 1 of the above algorithm in §3.2.2, we use the following symmetric random walk proposal on a logarithmic scale:

$$\log(\tilde{R}_z^j) \mid \tilde{R}_z^{j-1} \sim \mathcal{N}(\log(\tilde{R}_z^{j-1}), \sigma^2), \quad j = 1, 2, \dots,$$

where σ^2 is the proposal variance and has to be tuned to achieve good mixing, and the initial value $\tilde{R}_z^0$ is either fixed by the user or drawn from some distribution that we can easily simulate from. MH is usually 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 $\tilde{R}_z$ at a time, so each Markov chain Monte Carlo (MCMC) chain involved in Algorithm 1 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 $\tilde{R}_z$ as a general "quasi-exact" simulation method. Propp and Wilson (1996) further showed how to perform exact MCMC-based sampling by coupling Markov chains, but for simplicity, we shall here only consider the standard MCMC procedure outlined above, since it is easy to implement and already able to deliver very accurate simulation samples.

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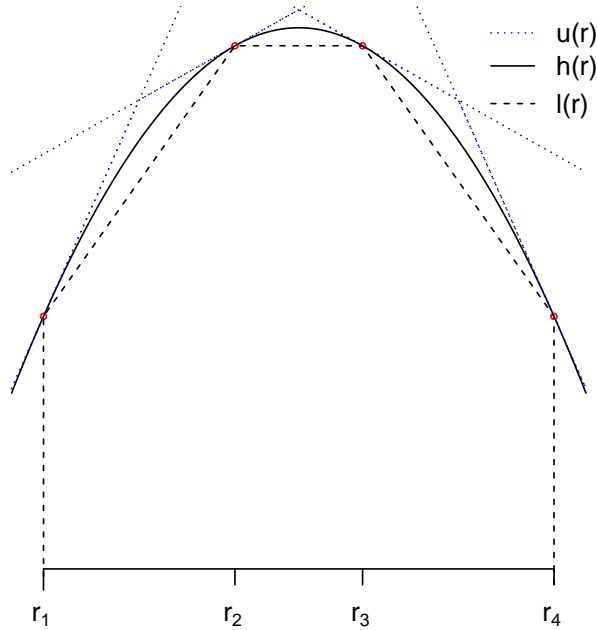
Figure 2: An illustration of the adaptive rejection sampling scheme, where the upper hull function $u(r)$ is tangent to the logarithmic target density $h(r)$ at locations $\boldsymbol{r} = \{r_1, r_2, r_3, r_4\}$.

$g_{\tilde{R}_z}(r; z)$, $r > 0$ with fixed z. The only requirement for using this method is that $h(r) = \log g_{\tilde{R}_z}(r; z)$, $r > 0$ 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 (7), 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, and then uses the distribution $s(r) = \exp\{u(r)\}/\int \exp\{u(r')\}\mathrm{d}r'$ 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 the points $\boldsymbol{r} = \{r_1, r_2, r_3, r_4\}$. Therefore, $u(r)$ and $l(r)$ are uniquely determined by the points in $\boldsymbol{r}$ and the target function $g_{\tilde{R}_z}(r; z)$. As we increase the number of points in $\boldsymbol{r}$, 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 $\boldsymbol{r}$ and construct the upper and lower bounds $u(r)$ and $l(r)$, respectively, as illustrated in Figure 2.

(1) Sampling: sample $\tilde{R}_z^* \sim s(r)$, and sample $U \sim \text{Unif}(0,1)$ independently of $\tilde{R}_z^*$.

(2) First acceptance test: if $U \leq \exp\{l(\tilde{R}_z^*) - u(\tilde{R}_z^*)\}$ then accept $\tilde{R}_z^*$ and go to 5.

(3) Second acceptance test: If $U \leq \exp\{h(\tilde{R}_z^*) - u(\tilde{R}_z^*)\}$ then accept $\tilde{R}_z^*$ and go to 5.

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

(5) Return $\tilde{R}_z^*$.

The distribution of the simulated value given acceptance is thus

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

which corresponds exactly to the target distribution with density $g_{\tilde{R}_z}(r; z), r > 0$. 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 η_i as a Gaussian location mixture

3.3.1 Model description

Another flexible and general max-id model family is the class of <u>Gaussian location mixtures</u>, where η_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 $\overline{\kappa}$, and $\{W_i\}$ are independent standard Gaussian random fields with mean zero and unit variance. 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 and flexible parametric form of the distribution of event magnitudes $\{R_i\}$, we propose to define $\overline{\kappa}$ as

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

(13)

where $\overline{\kappa}(r)$ is the short-hand notation for $\overline{\kappa}([r, \infty))$. The parameter restrictions for α, β_1 and β_2 ensure that the mean measure Λ 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 (13) is especially interesting in view of the flexible dependence structure that it yields and its close connection to the max-stable Brown–Resnick process (Kabluchko et al., 2009). The parameter β_1 controls the dependence strength in the upper tail, while β_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 (8) 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 the finite set of locations $K \equiv K_N$ is given by

$$\Lambda_K([0, \boldsymbol{z}_K]^C) = \int_{\mathbb{R}} \{1 - \Phi(\boldsymbol{z}_K - r; \boldsymbol{\rho}(K, K; r))\}\overline{\kappa}(\mathrm{d}r),$$

where $\Phi(\cdot; \boldsymbol{\rho}(K, K; r)))$ and $\boldsymbol{\rho}(K, K; r)$ are defined as before. The expressions for unconditional and conditional intensity functions can be derived by analogy with the scale mixture case. 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 $\rho(\boldsymbol{s}_1, \boldsymbol{s}_2)$.

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 $P_{\boldsymbol{s}_1}(z, \cdot)$ at a finite set of locations, we propose a similar simulation algorithm, where the first step consists in sampling a random variable $\tilde{R}_z$ from the probability density $g_{\tilde{R}_z}(r; z) \propto \varphi(z - r)\mathrm{d}\overline{\kappa}(\mathrm{d}r)$, $r > 0$, and the second step is to sample a Gaussian random vector $\tilde{\eta}(K)$, conditional on $\tilde{R}_z$, with mean $\tilde{R}_z + \boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)(z - \tilde{R}_z)$, common unit variance, and correlation matrix $\boldsymbol{\rho}(K, K; \tilde{R}_z) - \boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)\boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)^{\top}$, where $\boldsymbol{\rho}(\boldsymbol{s}_1, K; \tilde{R}_z)$ denotes the column vector of correlation values $\rho(\boldsymbol{s}_1, \boldsymbol{s}_j; \tilde{R}_z)$, $j = 1, \dots, N$, and $\boldsymbol{\rho}(K, K; \tilde{R}_z)$ is the corresponding correlation matrix. As above, we denote the density of $\tilde{\eta}(K)$ conditional on $\tilde{R}_z$ as $g_{\tilde{\eta}(K)|\tilde{R}_z}$. Again, the resulting (unconditional) distribution of $\tilde{\eta}(K)$ can be shown to be $P_{\boldsymbol{s}_1}(z, \cdot)$ at locations in K. However, here we observe that the density $g_{\tilde{R}_z}(r; z)$ is not logarithmically concave, and therefore we cannot use the adaptive resampling scheme to generate extremal functions at a finite set of locations in K from $P_{\boldsymbol{s}_1}(z, \cdot)$, 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 relies on Algorithm 1 combined with the following successive steps to sample from $P_{\boldsymbol{s}_1}(z, \cdot)$:

(1) Simulate a value $\tilde{R}_z$ according to the density $g_{\tilde{R}_z}(r; z)$ by using the Metropolis–Hastings method with random walk proposal distribution $\tilde{R}_z^j \mid \tilde{R}_z^{j-1} \sim \mathcal{N}(\tilde{R}_z^{j-1}, \sigma^2)$, $j = 1, 2, \dots$, and initial point $\tilde{R}_z^0$.

(2) Given $\tilde{R}_z$, simulate a vector $\tilde{\eta}(K)$ according to the multivariate Gaussian density function $g_{\tilde{\eta}(K)|R_z}$.

(3) Return $\tilde{\eta}(K)$.

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 β_1 and β_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 β_1 and β_2 increase, the dependence levels in the lower tail and the upper tail, which are controlled by β_2 and β_1, respectively, decrease. When $\beta_1 > 1$, the extremal coefficient increases as the level $z \to \infty$. By contrast, when $\beta_1 < 1$, the extremal coefficient decreases as the level $z \to \infty$. Similar results are observed for the dependence level in the lower tail with β_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.

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 $Z(\boldsymbol{s}) \approx \max_{i=1,\ldots,n} \eta_i(\boldsymbol{s})$ with $\eta_i = R_i W_i(\boldsymbol{s})$ for a sufficiently large integer n.

We perform the simulation on a 7×7 regular grid in $(0,1)^2$, corresponding to 49 locations in K. We fix $\alpha = 5, \beta = 2$ in (7) and $\nu = 3, \lambda_{\boldsymbol{s}} = \exp[2 - 0.5\Phi\{(s_x - 0.5)/0.25\}]$ in (8), where $\boldsymbol{s} = (s_x, s_y)^\top$ and $\Phi(\cdot)$ is the standard normal distribution function. We implemented
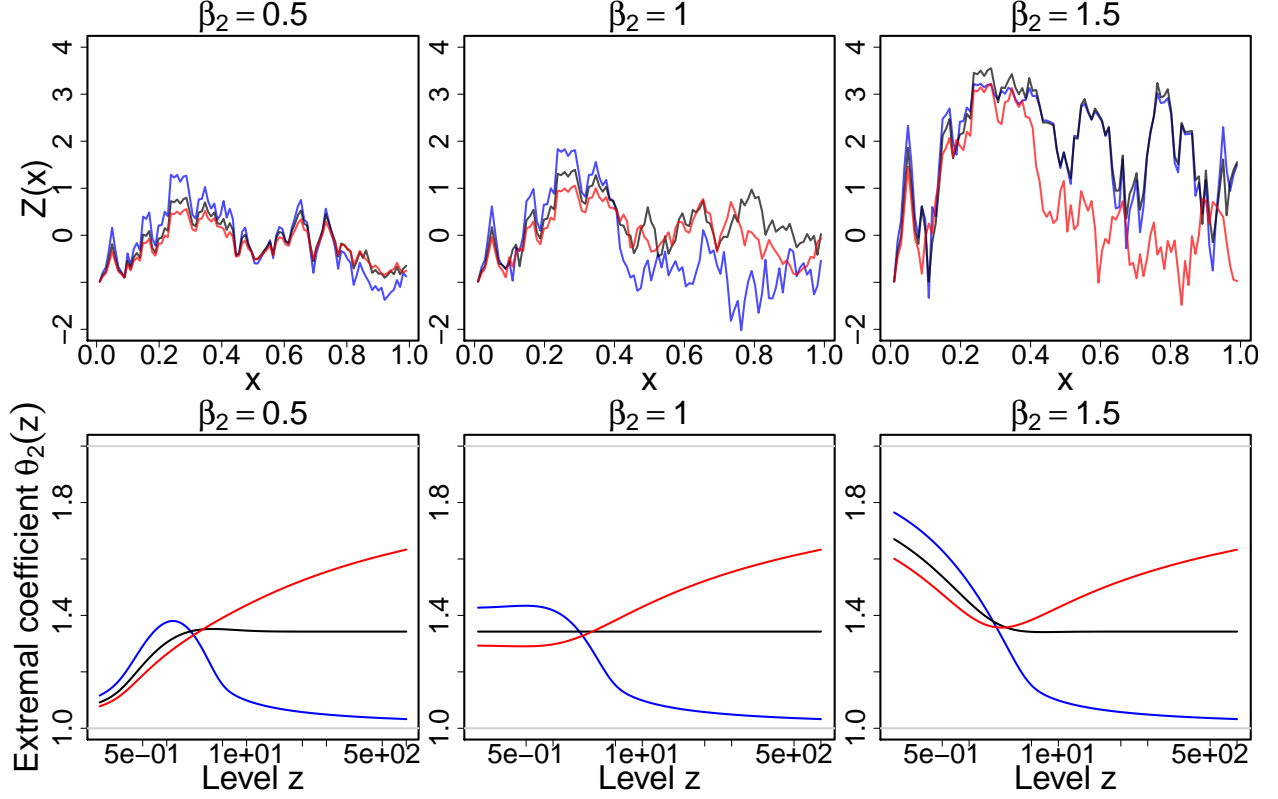
Figure 3: Simulations (top row) and bivariate level-dependent extremal coefficients (bottom row) for models based on the Gaussian location mixture construction. Parameters are chosen as $\alpha = 1$ and $\beta_1 = 0.5, 1, 1.5$ (blue, black, red). The correlation function for W_i is exponential and given by $\rho(h) = \exp(-2h)$. The level z is shown in logarithmic scale.

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 $\eta_i = R_i W_i$, $i = 1, 2, \ldots, n$ on the grid K, where $n = 100$.

(2) Metropolis–Hastings scheme: In the MCMC, run 100 iterations to generate one sample $\tilde{R}_z$ from the density function $g_{\tilde{R}_z}(r; z)$.

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

$$D_{\boldsymbol{s}_0}(p\|q) = \sum_{i=1}^{N} \widehat{p}\{Z_i(\boldsymbol{s}_0)\} \log \left[\frac{\widehat{p}\{Z_i(\boldsymbol{s}_0)\}}{q\{Z_i(\boldsymbol{s}_0)\}} \right], \quad \widehat{p}\{Z_i(\boldsymbol{s}_0)\} = \frac{p\{Z_i(\boldsymbol{s}_0)\}}{\sum_{i=1}^{N} p\{(Z_i(\boldsymbol{s}_0))\}}, \tag{14}$$

where q places mass $1/N$ on each simulated sample as the empirical cdf, while p is the exact marginal density of Z. In practice, the samples are always treated as discrete random variables, and $\widehat{p}$ and q need to be probability mass functions instead of density functions. 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.

To further evaluate the performance of each sampling scheme with respect to the dependence structure, we calculated the bivariate (level-dependent) empirical extremal coefficients $\widehat{\theta}_2$. 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 Λ with a very accurate numerical unidimensional integral. Figure 5 shows the results for the three simulation techniques with the computed standard error, sd, of the approximation error shown in the title of each panel. The adaptive rejection sampling scheme appears to perform generally better than the naive scheme at all levels, especially at low quantile levels when the naive scheme is significantly biased. At high quantile levels $(q = 0.95)$, all methods have a comparable performance. The pseudo-exact Metropolis–
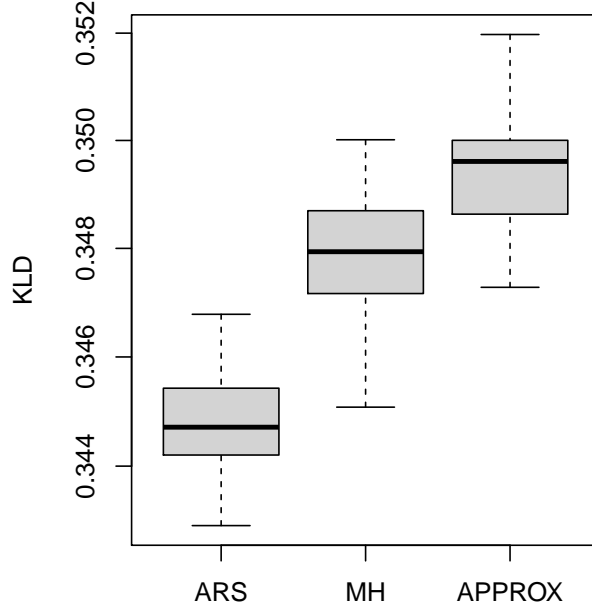
Figure 4: Boxplots of empirical KL divergences for the 49 grid locations based on the simulated data sets generated by the naive method (APPROX), the Metropolis–Hastings scheme (MH) and the adaptive rejection sampling scheme (ARS).

Hastings scheme appears to be as accurate as the adaptive rejection sampling method. 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
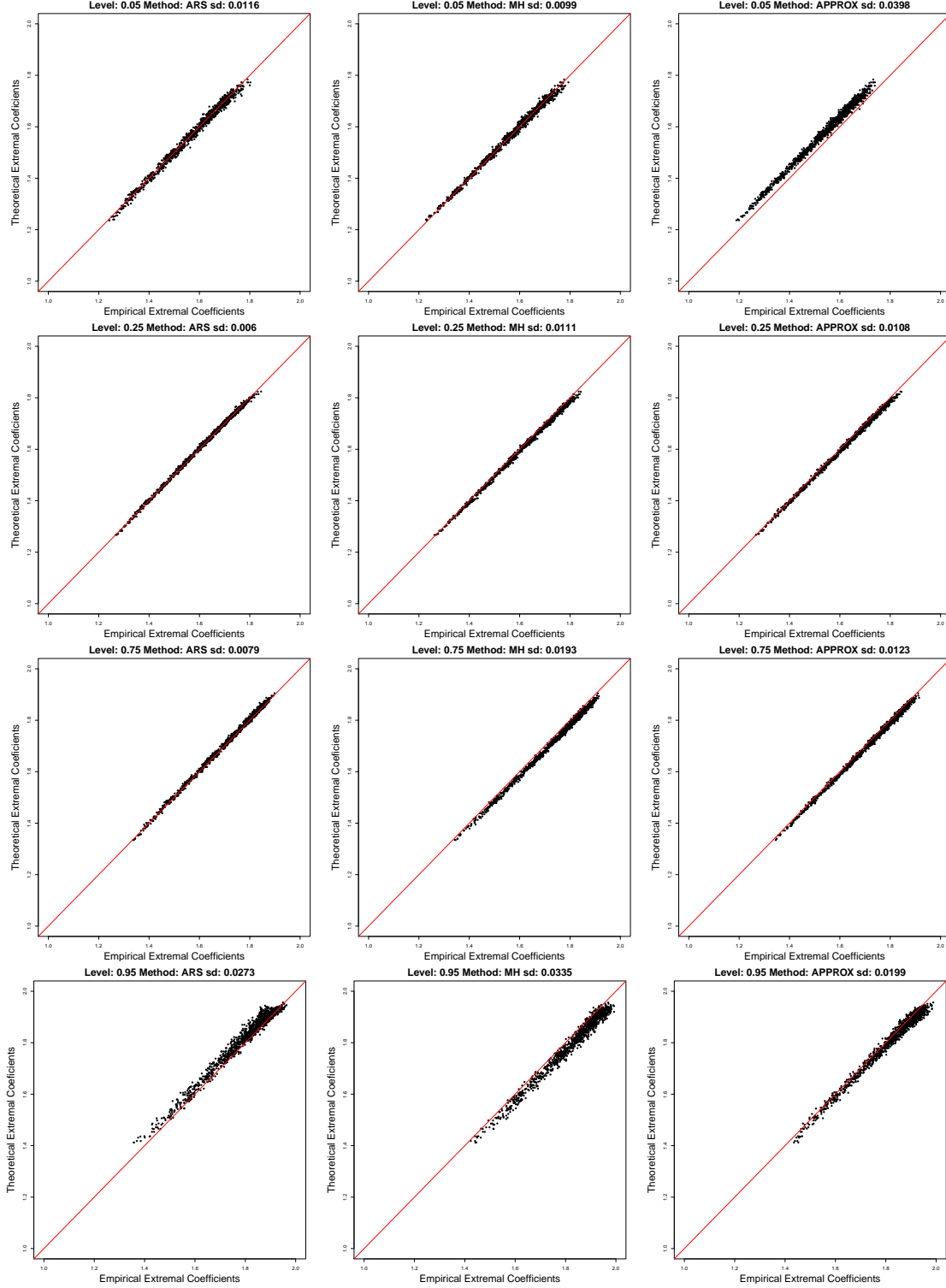
Figure 5: Bivariate scatterplots of theoretical against empirical bivariate level-dependent extremal coefficients at quantile levels $0.05, 0.25, 0.5, 0.75, 0.95$ (from top to bottom) for all pairs of locations, based on the three sampling schemes (left: Algorithm 1 with adaptive rejection sampling (ARS), middle: Algorithm 1 with Metropolis–Hastings (MH), right: naive approximate method).
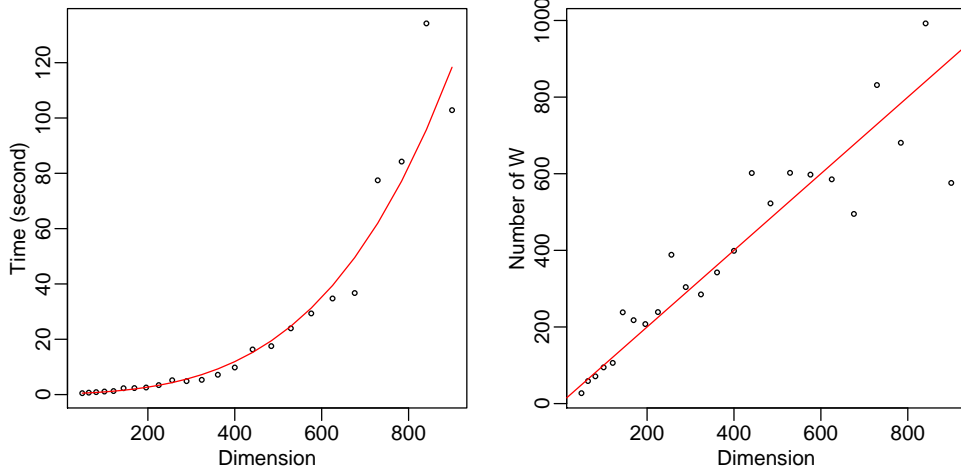
Figure 6: Left: Time in second needed to simulate one max-id replicate for different numbers of locations $|K|$. The red curve is the fit from a 4-th order polynomial. Right: Number of random fields W_i that were simulated for one max-id replicate for different numbers of locations $|K|$. The red line 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 $\rho(h) = \exp(-h)$.

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 $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ chosen in Algorithm 1 to iteratively sample from $P_{\boldsymbol{s}_0}(z, \cdot)$ 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 all simulations presented here, 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 $\tilde{R}_z$, in order to achieve fast and exact simulations in cases where 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 η_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. Exploring exact MCMC procedures based on coupling of Markov chains (Propp and Wilson, 1996) would also be an interesting future research direction.

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 described above, the computation of multivariate density and distribution 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).

Acknowledgments

References

Bopp, G., Shaby, B. A. and Huser, R. (2021) A hierarchical max-infinitely divisible spatial model for extreme precipitation. Journal of American Statistical Association **116**, 93–106.

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. In Simulation-Based inference in Econometrics: Methods and Applications, eds R. Mariano, M. Weeks and T. Schuermann, pp. 71–99. Cambridge University Press.

Huser, R., Opitz, T. and Thibaud, E. (2020) Max-infinitely divisible models and inference for spatial extremes. Scandinavian Journal of Statistics **48**, 321–348.

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, M. and Strokorb, K. (2021) A comparative tour through the simulation algorithms for max-stable processes. Statistical Science To appear.

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.

Propp, J. G. and Wilson, D. B. (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics. Random Structures & Algorithms **1–2**, 223–252.

Resnick, S. I. (1987) Extreme Values, Regular Variation and Point Processes. Spinger-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. (2021) Modeling non-stationary temperature maxima based on extremal dependence changing with event magnitude. Annals of Applied Statistics To appear.