

# Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation

MARCO OESTING<sup>1</sup>, MARTIN SCHLATHER<sup>2</sup> and CHEN ZHOU<sup>3,4</sup>

<sup>1</sup>Department Mathematik, Universität Siegen, Walter-Flex-Str. 3, 57072 Siegen, Germany.  
E-mail: [oesting@mathematik.uni-siegen.de](mailto:oesting@mathematik.uni-siegen.de)

<sup>2</sup>Institut für Mathematik, Universität Mannheim, A 5, 6, 68161 Mannheim, Germany.  
E-mail: [schlather@math.uni-mannheim.de](mailto:schlather@math.uni-mannheim.de)

<sup>3</sup>Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000DR Rotterdam, The Netherlands. E-mail: [zhou@ese.eur.nl](mailto:zhou@ese.eur.nl)

<sup>4</sup>Economics and Research Division, De Nederlandsche Bank, P.O. Box 98, 1000AB Amsterdam, The Netherlands. E-mail: [c.zhou@dnb.nl](mailto:c.zhou@dnb.nl)

The efficiency of simulation algorithms for max-stable processes relies on the choice of the spectral representation: different choices result in different sequences of finite approximations to the process. We propose a constructive approach yielding a normalized spectral representation that solves an optimization problem related to the efficiency of simulating max-stable processes. The simulation algorithm based on the normalized spectral representation can be regarded as max-importance sampling. Compared to other simulation algorithms hitherto, our approach has at least two advantages. First, it allows the exact simulation of a comprising class of max-stable processes. Second, the algorithm has a stopping time with finite expectation. In practice, our approach has the potential of considerably reducing the simulation time of max-stable processes.

*Keywords:* importance sampling; mixed moving maxima; optimal simulation

## 1. Introduction

Max-stable processes have become a popular tool for modeling spatial extremes, particularly in environmental sciences, see, for example, Coles [3], Coles and Tawn [4] and Padoan, Ribatet and Sisson [22]. A stochastic process  $\{Z(y) : y \in K\}$  with standard Fréchet margins defined on an index set  $K$ , that is,  $\mathbb{P}(Z(y) \leq z) = \exp(-z^{-1})$ ,  $z > 0$ , for all  $y \in K$ , is called max-stable if

$$\frac{1}{n} \max_{i=1}^n Z_i =_d Z$$

for any  $n \in \mathbb{N}$  and independent copies  $Z_i$ ,  $i = 1, \dots, n$ , of  $Z$ , where the maximum is taken pointwise. Max-stable processes occur naturally as limits of suitably normalized pointwise maxima of stochastic processes which motivates their usage in the context of spatial extremes.

Simulating max-stable processes is an important step in application for the following three reasons. First, while bivariate marginal distributions can be calculated frequently, higher dimensional marginal distributions do not have, in nearly all the cases, explicit formulae. Consequently,

they can be addressed only by simulation. Second, most applications require the estimation of characteristics of max-stable processes that cannot be explicitly calculated. That leaves simulation as the only option, see, for example, Buishand, de Haan and Zhou [2] and Blanchet and Davison [1]. Finally, unconditional simulation appears as part of the conditional simulation of max-stable processes (Dombry, Éyi-Minko and Ribatet [13], Oesting and Schlather [19]).

Schlather [26] suggested an algorithm to simulate max-stable processes. However, the simulation is exact only under substantial restrictions, namely when the shape function is bounded and has compact support. In this case, the algorithm ends in finite time almost surely. Compact support can be enforced by cutting the shape function, which introduces an approximation error. In some cases, such an error is not negligible. For example, consider moving maximum processes with monotone shape functions that have a pole at the origin. Strokorb, Ballani and Schlather [31] provided a derivation of such processes whose realizations have poles on a dense subset of the space. In particular, they are discontinuous everywhere with probability one. For such processes, any modification of the shape function towards a bounded shape function will alter dramatically the properties of the process. Consequently, the use of Schlather’s [26] algorithm becomes doubtful. By contrast, this paper deals with the question of drawing random samples of max-stable processes in an exact and efficient way, including shape functions that are unbounded and do not have compact support.

Simulation of max-stable processes is based on their spectral representation (see de Haan [6], Giné, Hahn and Vatan [14], Kabluchko [16], Wang and Stoev [32] for instance): For any max-stable process  $Z$  with standard Fréchet margins defined on an index set  $K$ , there exists a spectral measure  $H$  defined on an appropriate set  $\mathbb{H}$  of non-negative functions such that

$$Z(y) = \max_{(t,f) \in \Pi} tf(y), \quad y \in K, \tag{1.1}$$

where  $\Pi$  is a Poisson point process on  $(0, \infty) \times \mathbb{H}$  with intensity  $t^{-2} dt H(df)$  and

$$\int_{\mathbb{H}} f(y)H(df) = 1 \tag{1.2}$$

for all  $y \in K$ . The functions  $f$  in  $\mathbb{H}$  are the *spectral functions* of the max-stable process  $Z$ . As any max-stable process can be obtained from a max-stable process with standard Fréchet marginals via marginal transformations, we will henceforth assume processes with standard Fréchet marginals, that is, processes with representation (1.1) and (1.2).

According to the spectral representation (1.1), the construction of a max-stable process involves infinitely many points  $(t, f) \in \Pi$ . Nevertheless, since only the maximum over all functions  $tf$  counts, the number of points  $(t, f)$  that contribute to  $Z$ , that is,  $Z(y) = tf(y)$  for at least one point  $y \in K$ , is finite under mild conditions, see de Haan and Ferreira [7], Cor. 9.4.4. However, their statement is a theoretical one that does not help for simulation purposes because one cannot determine *ex ante* which functions  $f$  will contribute.

Schlather’s [26] algorithm for simulating max-stable processes requires that  $K$  is compact, the shape functions  $f : \mathbb{R}^d \rightarrow [0, \infty)$  are bounded by some  $C \in (0, \infty)$  and the support is within a ball of radius  $r$  centered at the origin. For example, consider a stationary moving maxima processes with the measure  $H$  given by

$$H(\{f(\cdot - x) : f \in B, x \in A\}) = \Lambda(A)H_f(B) \tag{1.3}$$

for some probability measure  $H_f$  and the Lebesgue measure  $\Lambda$ . Let  $K_r = \{x \in \mathbb{R}^d : \|x - y\| \leq r \text{ for some } y \in K\}$ . Under the aforementioned assumptions, Schlather [26] showed that the right-hand side of (1.3) can be restricted to  $\Lambda(A \cap K_r)H_f(B)$ . Hence, the intensity for  $(t, f(\cdot - x))$  equals

$$t^{-2} \Lambda(K_r) dt \cdot \frac{\Lambda(dx \cap K_r)}{\Lambda(K_r)} H_f(df),$$

where  $f(\cdot - x)$  can be reinterpreted as realization of a random function  $F$  with law  $\Lambda(dx \cap K_r)/\Lambda(K_r) \cdot H_f(df)$  that is independent of the Poisson point process with intensity  $t^{-2} \Lambda(K_r) dt$ .

Schlather [26] suggested to start with those points  $(t, f)$  that will contribute most likely to  $Z$ , i.e., with those points  $(t, f)$  that have the highest values of  $t$ . By ranking the points  $t$  in a descending order  $t_1 > t_2, \dots$ , we have that  $t_i = {}^d 1/(\sum_{j=1}^i E_j)$ , where  $E_j$  are independent and identically distributed random variables with standard exponential distribution. Let  $F_i \sim_{i.i.d.} F$  be independent of the  $E_j$  and

$$Z^{(m)}(y) = \max_{1 \leq i \leq m} \frac{1}{\sum_{j=1}^i E_j} F_i(y), \quad y \in K,$$

be a finite approximation for  $Z$ . Then,  $Z = {}^d Z^{(\infty)}$ . Assume that for some  $m$  the inequality

$$Z^{(m)}(y) \geq \frac{1}{\sum_{j=1}^{m+1} E_j} \sup_{f \in \mathbb{H}} f(y) \quad \text{for all } y \in K, \tag{1.4}$$

holds. Then, obviously,  $Z^{(n)}(y) = Z^{(\infty)}(y)$  for all  $y \in K$  and all  $n \geq m$ . In other words, any spectral function  $f_i$  with  $i > m$  cannot contribute to  $Z$ . This results in a *stopping rule* for a “ $m$ -step representation” of  $Z$  which can be used to construct an exact simulation procedure. Here, the minimal number of steps  $M = \min\{m \in \mathbb{N} : m \text{ satisfies (1.4)}\}$  is a random integer. The properties of  $M$  depend on the choices of the spectral functions.

Schlather [26] proposed to replace inequality (1.4) by a stronger stopping rule  $\inf_{y \in K} Z^{(m)}(y) \geq C/\sum_{j=1}^{m+1} E_j$ . As a generalization of Schlather [26], we consider stopping rule (1.4) with  $H$  being any locally finite measure, not necessarily of the form (1.3).

The random number of steps in the simulation of a max-stable process depends on the choice of the corresponding ensemble of spectral functions which is not unique (cf. de Haan and Ferreira [7], Remark 9.6.2). Some specific choices may bear severe disadvantages for the accuracy and speed of the simulation. For instance, finite approximations based on the original definition of the Brown–Resnick process are usually far from the actual process (Kablichko, Schlather and de Haan [17], Oesting, Kablichko and Schlather [18]). The optimality of the choice of spectral functions with respect to the number of steps in simulating max-stable processes has not been discussed in literature yet. Here, we propose a choice of spectral functions, the normalized spectral representation and show that it is the solution to an optimization problem related to the number of steps in the simulation.

An illustrative example may clarify why the choice of spectral functions can have a significant impact on the distribution of the stochastic number  $M$ . Consider the simplest case where  $Z$

is univariate. Specializing (1.1) to  $K = \{y_0\}$  and  $f \equiv 1$ , the random variable  $Z(y_0)$  follows a univariate Fréchet distribution. It has a representation given by

$$Z(y_0) =^d \max_{t \in \Pi} t, \quad (1.5)$$

where  $\Pi$  is the Poisson point process on  $(0, \infty)$  with intensity  $t^{-2} dt$ . By ranking the points  $t$ , we get that

$$Z^{(m)}(y_0) = \max_{1 \leq i \leq m} \frac{1}{\sum_{j=1}^i E_j} = \frac{1}{E_1} = Z^{(1)}(y_0).$$

In other words,  $M \equiv 1$  based on the constant spectral function  $f \equiv 1$ . Differently,  $Z(y_0)$  can also be constructed by (1.1) with  $f(y_0)$  being a non-degenerate random variable satisfying  $\mathbb{E}f(y_0) = 1$ . In that case, the stochastic number  $M$  is greater than 1 with positive probability. Therefore, for simulating  $Z(y_0)$ , the spectral representation in (1.5) would be considered as optimal. This example illustrates the optimality we intend to achieve by the choice of spectral functions for an arbitrary max-stable process.

The very general optimality problem for general index sets  $K$  and arbitrary random functions  $f$  seems to be rather complicated. Therefore, we shall suggest a modified optimization problem and shall demonstrate that its solution is explicit and unique for each given max-stable process and index set  $K$ . The optimality is achieved by transforming any ensemble of spectral functions to a new ensemble of spectral functions satisfying  $\sup_{y \in K} f(y) = c$ , for all  $f \in \mathbb{H}$ . We call such a representation with all spectral functions sharing the same supremum the *normalized spectral representation*. This representation was initially used in constructing the spectral representation for max-stable processes with a continuous sample path on  $K = [0, 1]$ , see, for example, de Haan and Lin [8] and de Haan and Ferreira [7], Cor. 9.4.5. In this paper, we provide a theoretical justification on the use of this normalized spectral representation in simulation algorithms.

This paper is organized as follows. In Section 2, we revisit de Haan's [6] spectral representation of max-stable processes and give the formula of spectral representation transformation, i.e. transforming one ensemble of spectral functions under a given spectral measure to another ensemble under a different spectral measure. Based on the transformed representation, we formulate a stopping rule that allows an exact simulation of max-stable processes. In Section 3, we pose an optimization problem for selecting the spectral representation that yields the most efficient simulation procedure. This problem is closely related to the problem of importance sampling. In addition, we give the explicit solution of a modified optimization problem. This results in the normalized spectral representation. Differences between the modified problem and the original optimization problem are evaluated in Section 4. Section 5 deals with the normalized spectral representation for moving maxima processes. For moving maxima processes and Brown–Resnick processes, the performance of the simulation procedure based on the normalized spectral representation is compared to other algorithms in Section 6. In Section 7, we summarize and discuss our results.

## 2. Transformation of spectral representations

Throughout the paper, we consider a max-stable process  $Z$  on some index set  $K$  that is assumed to be a compact Polish space. Further, we will assume that the spectral functions  $f$  from representation (1.1) lie in some Polish space  $\mathbb{H} \subset [0, \infty)^K$  equipped with a  $\sigma$ -algebra  $\mathcal{H}$  such that the mapping  $f \mapsto \sup_{y \in K} f(y)$  is  $(\mathcal{H}, \mathcal{B} \cap [0, \infty))$ -measurable where  $\mathcal{B}$  denotes the standard Borel  $\sigma$ -algebra on  $\mathbb{R}$ .

The subsequent proposition presents a general procedure to transform one spectral representation to another yielding the same max-stable process. It can be proved by standard arguments from extreme value theory, see Oesting, Schlather and Zhou [20] for details.

**Proposition 2.1.** *Let  $Z$  be a max-stable process with standard Fréchet margins defined as in (1.1) and (1.2). Suppose that  $H$  is a locally finite measure on  $\mathbb{H}$ . Let  $g$  be some probability density on  $\mathbb{H}$  w.r.t.  $H$ , that is,  $g \geq 0$  and  $\int_{\mathbb{H}} g(f)H(df) = 1$ , such that*

$$H\left(\left\{f : g(f) = 0, \sup_{y \in K} f(y) > 0\right\}\right) = 0. \tag{2.1}$$

Then, by excluding all the functions  $f \in \mathbb{H}$  with  $g(f) = 0$ , we get

$$Z(y) =^d \max_{(t,f) \in \tilde{\Pi}} t \frac{f(y)}{g(f)}, \quad y \in K, \tag{2.2}$$

where  $\tilde{\Pi}$  is a Poisson point process with intensity  $t^{-2} dt g(f)H(df)$ .

Applying Proposition 2.1, a given ensemble of spectral functions  $\{f\}_{(t,f) \in \Pi}$  can be transformed to a new ensemble  $\{f/g(f)\}_{(t,f) \in \tilde{\Pi}}$ , where  $f$  follows the transformed probability measure  $gH$  defined by

$$gH(A) = \int_A g(f)H(df)$$

for all measurable sets  $A \subset \mathbb{H}$ . For this transformed spectral representation (2.2), the stopping rule (1.4) can be formulated as follows. Denote

$$Z^{(m)}(y) = \max_{1 \leq i \leq m} \frac{1}{\sum_{j=1}^i E_j} \cdot \frac{F_i(y)}{g(F_i)}, \quad y \in K, \tag{2.3}$$

for standard exponentially distributed random variables  $E_j$  and  $F_j \sim gH$ , which are all independent. Let

$$Z^{(\infty)} = \lim_{m \rightarrow \infty} Z^{(m)}. \tag{2.4}$$

Then,  $Z^{(\infty)} =^d Z$  and, for fixed  $\omega \in \Omega$ , we have  $Z^{(m)} \equiv Z^{(\infty)}$  on  $K$  if

$$\text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(m)}(y)} \leq \sum_{j=1}^{m+1} E_j, \tag{2.5}$$

where the essential supremum is taken w.r.t. the probability measure  $gH$ . Note that, by (2.1), up to a  $H$  null set, the set  $\{f \in \mathbb{H} : g(f) = 0\}$  consists of functions  $f \in \mathbb{H}$  with  $f|_K \equiv 0$ . Thus, for fixed  $g$ , we may exclude all the functions  $f \in \mathbb{H}$  with  $g(f) = 0$ .

If the number  $M_g := \min\{m \in \mathbb{N} : m \text{ satisfies (2.5)}\}$  is finite, the max-stable process  $Z$  can be simulated exactly in finite time via the following algorithm.

**Algorithm 2.2.**

Set  $m = 0$  and  $Z^{(0)}(y) = 0$  for all  $y \in K$ .

Simulate a standard exponentially distributed random variable  $E_1$ .

while (2.5) is not satisfied {

Update  $m$  by  $m + 1$ .

Sample  $F_m$  from  $gH$ .

Set  $Z^{(m)}(y) = \max\{Z^{(m-1)}(y), \frac{1}{\sum_{j=1}^m E_j} \frac{F_m(y)}{g(F_m)}\}$  for all  $y \in K$ .

Simulate a standard exponentially distributed random variable  $E_{m+1}$ .

}

Return  $Z^{(m)}$ .

Algorithm 2.2 requires the evaluation of the stopping rule (2.5) which involves suprema with respect to  $y \in K$  and  $f \in \mathbb{H}$ . While, in practice, the set  $K$  often is finite, that is, the supremum w.r.t.  $y$  is a maximum, the evaluation of the essential supremum w.r.t.  $f$  is more difficult. However, with the normalized spectral representation (cf. Algorithm 3.11), it is possible to provide a bound for the essential supremum facilitating this evaluation; see Remark 3.12.

The exact simulation based on the stopping rule (2.5) requires the simulation of  $M_g$  processes from law  $gH$  and  $M_g + 1$  exponentially distributed random variables. Therefore, the time costs of Algorithm 2.2 depend on the stopping rule, which essentially depends on the choice of  $g$ . Consequently, the following question arises: which density  $g$  minimizes the random number of steps  $M_g$ ? This optimization problem will be formulated in a more rigorous way in the next section.

### 3. The optimization problem

Henceforth, we will always assume that we are in the framework of Proposition 2.1. Further, the process  $Z$  is assumed to be almost surely strictly positive on  $K$ , that is,

$$\mathbb{P}\left(\inf_{y \in K} Z(y) > 0\right) = 1. \tag{3.1}$$

Note that this assumption is valid for processes with continuous sample paths.<sup>1</sup>

<sup>1</sup>Assuming sample continuity, we even have

$$\mathbb{E}\left(\left(\inf_{y \in K} Z(y)\right)^{-1}\right) < \infty$$

(Dombry and Eyi-Minko [11], Thm. 2.2). It can be shown that this also holds true under weaker assumptions than sample continuity (see Oesting, Schlather and Zhou [20]).

We are interested in minimizing the number of steps in Algorithm 2.2,  $M_g$ , that is, the minimal number  $m$  such that (2.5) holds. Since  $M_g$  is a random variable, we aim at minimizing some of its mathematical characteristics. As many applications require a large number of simulations, a natural choice to start with is to minimize the mean  $Q_g$  of  $M_g$  by choosing a proper density  $g$ , that is,

$$Q_g = \mathbb{E}M_g, \quad M_g = \min \left\{ m \in \mathbb{N} : \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(m)}(y)} \leq \sum_{j=1}^{m+1} E_j \right\}. \quad (3.2)$$

In addition, we aim to determine at least one member of

$$\mathcal{G} = \arg \min_g Q_g.$$

We remark that the finiteness of  $Q_g$  is not ensured. As a first step, the following proposition provides a sufficient condition for  $M_g$  being finite almost surely. It follows directly from the definition of  $M_g$ .

**Proposition 3.1.** *Assume that (3.1) holds. Then  $M_g$  is finite a.s. if*

$$\text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} < \infty. \quad (3.3)$$

Proposition 4.6 below assures the finiteness of  $Q_g$  provided that condition (3.3) is fulfilled and  $Z$  is sample-continuous.

There is a close relation between our approach and the importance sampling (Hastings [15]). We interpret (2.2) as “max-importance sampling” as follows. Importance sampling is targeted on calculating an integral  $\int f(x)\lambda(dx)$  in an efficient way for a function  $f : \mathcal{X} \rightarrow \mathbb{R}$  and a measure  $\lambda$  on some space  $\mathcal{X}$ . It uses the fact that

$$\int f(x)\lambda(dx) = \int \frac{f(x)}{g(x)}(g\lambda)(dx), \quad (3.4)$$

where  $g : \mathcal{X} \rightarrow (0, \infty)$  is a probability density w.r.t.  $\lambda$  and  $(g\lambda)(A) = \int_A g(x)\lambda(dx)$ . Then the integral can be approximated by  $\hat{I}^{(n)} = \frac{1}{n} \sum_{i=1}^n \frac{f(X_i)}{g(X_i)}$ , where  $(X_i)_{i=1}^n$  is an i.i.d. sequence of random variables on  $\mathcal{X}$  with law  $g\lambda$ . Importance sampling considers a practically advantageous choice of the density  $g$  such that  $\text{Var}(f(X_1)/g(X_1))$  and hence  $\text{Var}(\hat{I}^{(n)})$  are small. Exchanging the underlying space  $\mathcal{X}$  with the function space  $\mathbb{H}$ , and replacing the integral with the max-integral (de Haan [6]), we receive a similar description of our approach:

$$Z(y) = \int^{\vee} y(f)H^*(df) = \int^{\vee} \frac{y(f)}{g(f)}(gH^*)(df), \quad (3.5)$$

where  $y : \mathbb{H} \rightarrow [0, \infty)$ ,  $f \mapsto y(f)$ , is the punctual evaluation of a function  $f$  at  $y$ ,  $H^*$  is a random discrete measure on  $\mathbb{H}$  defined as  $H^*({f}) = \sup_{(t,f) \in \Pi} t$  and the integral  $\int^{\vee}$  is understood as taking the maximum of the integrand weighted by the measure  $H^*$ .

Despite the formal correspondence between (3.4) and (3.5), there are some notable differences. First, in our approach,  $y$ , which is formally an element of the dual space of  $\mathbb{H}$ , is not fixed. Second, importance sampling is always an approximation to the integral value of interest. Under mild conditions, see Section 4, a finite approach to the right-hand side of equality (3.5) renders the exact value of the max-integral. Hence, we intend to choose  $g$  so that the required number of steps is minimized in order to eliminate the error of the finite approximation. In contrast, importance sampling searches for a  $g$  that leads to a high speed of convergence, that is, a small variance of  $\hat{I}^{(n)}$ .

Whilst the theoretical optimum is well known in importance sampling, and the difficulty there is to find a numerically advantageous function  $g$ , the optimization problem (3.2) itself is difficult to solve, since both the numerator and the denominator of  $f(y)/(g(f)Z^{(m)}(y))$  depend on  $y$  and the denominator is stochastic. To circumvent this difficulty, we modify the optimization problem in Section 3.1 and solve the modified problem in Section 3.2. The solution of the modified problem leads to the normalized spectral representation discussed in Section 3.3.

### 3.1. A modified optimization problem

We first motivate the modification of the optimization problem. Recall the stopping rule (2.5) as

$$\text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(m)}(y)} \leq \sum_{j=1}^{m+1} E_j.$$

A stronger inequality that implies this stopping rule is

$$\text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f) \inf_{\tilde{y} \in K} Z^{(m)}(\tilde{y})} \leq \sum_{j=1}^{m+1} E_j, \tag{3.6}$$

while a weaker inequality that is implied by the stopping rule is

$$\text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f) \sup_{\tilde{y} \in K} Z^{(m)}(\tilde{y})} \leq \sum_{j=1}^{m+1} E_j. \tag{3.7}$$

The actual stopping time is larger than that under the weaker rule and smaller than that under the stronger rule. In other words, if we consider (3.6) and (3.7) as stopping rules, we might simulate too many or too few spectral functions, respectively. Next, we define an ensemble of alternative stopping rules that also lie in between the rules (3.6) and (3.7). Suppose  $T : [0, \infty)^K \rightarrow [0, \infty)$  is a functional that satisfies  $T(\mathbf{1}) = 1$  and that is max-linear, that is,

$$T(\max\{a_1 h_1, a_2 h_2\}) = \max\{a_1 T(h_1), a_2 T(h_2)\}$$

for all  $a_1, a_2 \geq 0$  and  $h_1, h_2 : K \rightarrow [0, \infty)$ . Then, we have that  $T(h) \leq T(g)$  for all  $h \leq g$ , which leads to

$$\inf_{y \in K} h(y) \leq T(h) \leq \sup_{y \in K} h(y) \tag{3.8}$$



for all  $h : K \rightarrow [0, \infty)$ . Therefore, any max-linear functional  $T$  delivers a stopping rule

$$\text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)T(Z^{(m)})} \leq \sum_{j=1}^{m+1} E_j \tag{3.9}$$

that lies also in between the aforementioned stronger and weaker stopping rules.

We regard these new conditions (3.9) as surrogates for the actual stopping rule (2.5). The corresponding modified optimization problem is then

$$\begin{aligned} \mathcal{G}^* &= \arg \min_g Q_g^*, \\ Q_g^* &= \mathbb{E} M_g^* = \mathbb{E} \min \left\{ m \in \mathbb{N} : \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)T(Z^{(m)})} \leq \sum_{j=1}^{m+1} E_j \right\}. \end{aligned} \tag{3.10}$$

Examples of  $T$  are  $T(h) = \sup_{y \in K} h(y)$  and  $T(h) = h(y_0)$  for some  $y_0 \in K$ . The corresponding modified problems based on these two specific  $T$  minimize the quantities

$$Q_g^{(1)} = \mathbb{E} \min \left\{ m \in \mathbb{N} : \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f) \sup_{\tilde{y} \in K} Z^{(m)}(\tilde{y})} \leq \sum_{j=1}^{m+1} E_j \right\}, \tag{3.11}$$

and 
$$Q_g^{(2)}(y_0) = \mathbb{E} \min \left\{ m \in \mathbb{N} : \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)Z^{(m)}(y_0)} \leq \sum_{j=1}^{m+1} E_j \right\}. \tag{3.12}$$

Note that the modified condition (3.9) does not correspond to the stopping rule under consideration (Algorithm 2.2).

### 3.2. The solution of the modified optimization problem

The following proposition provides a first but also a key step towards the solution of (3.10). It shows that the solution of (3.10) is independent of  $T$ .

**Proposition 3.2.** *The solution of the modified optimization problem in (3.10) satisfies*

$$\mathcal{G}^* = \arg \min_g \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)}.$$

**Proof.** If there exists some  $g$  such that  $Q_g^*$  is finite, then necessarily

$$\text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)} < \infty.$$

Thus, we can restrict ourselves to

$$g \in D = \left\{ g : \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)} < \infty \right\}$$

and assume w.l.o.g. that  $D \neq \emptyset$ . For  $c = \int_{\mathbb{H}} \sup_{y \in K} f(y) H(df)$  and any  $g \in D$ , we have

$$c \leq \int_{\mathbb{H}} \text{esssup}_{h \in \mathbb{H}} \frac{\sup_{y \in K} h(y)}{g(h)} g(f) H(df) = \text{esssup}_{h \in \mathbb{H}} \frac{\sup_{y \in K} h(y)}{g(h)} < \infty. \quad (3.13)$$

Thus, by (3.8), for  $c_T = \int_{\mathbb{H}} T(f) H(df)$ , we obtain  $c_T \leq c < \infty$ .

Next, we prove  $c_T > 0$  by contradiction. Assume that  $c_T = 0$ . This yields  $T(f) = 0$  for  $H$ -a.e.  $f \in \mathbb{H}$  which – by the max-linearity of  $Z$  – implies  $T(Z) = 0$  a.s. in contradiction to  $\inf_{y \in K} Z(y) > 0$  a.s. and (3.8). Thus, we conclude that  $c_T \in (0, \infty)$ .

Now, let  $g \in D$ . Using the max-linearity of  $T$  and the fact that  $Q_g^* = \mathbb{E}M_g^* = \sum_{m=0}^{\infty} \mathbb{P}(M_g^* > m)$ , we have

$$\begin{aligned} Q_g^* - 1 &= \sum_{m=1}^{\infty} \mathbb{P} \left( \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} / \sum_{j=1}^{m+1} E_j > \max_{1 \leq k \leq m} \frac{1}{\sum_{j=1}^k E_j} \frac{T(f_k)}{g(f_k)} \right) \\ &= \sum_{m=1}^{\infty} \mathbb{P} \left( \frac{\sum_{j=1}^k E_j}{\sum_{j=1}^{m+1} E_j} > \frac{T(f_k)}{g(f_k)} / \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)}, 1 \leq k \leq m \right). \end{aligned} \quad (3.14)$$

Note that

$$\frac{T(f_k)}{g(f_k)} / \text{esssup}_{f \in H} \sup_{y \in K} \frac{f(y)}{g(f)} \in [0, 1].$$

As the joint distribution of  $(\sum_{j=1}^k E_j / \sum_{j=1}^{m+1} E_j)_{k=1, \dots, m}$  equals the joint distribution of the order statistics of  $m$  independent random variables  $U_1, \dots, U_m$  which are uniformly distributed on  $[0, 1]$ , and as, by exchangeability,

$$\mathbb{P}(U_{(1)} > X_1, \dots, U_{(m)} > X_m) = \mathbb{P}(U_1 > X_1, \dots, U_m > X_m)$$

holds for i.i.d.  $[0, 1]$ -valued random variables  $X_1, \dots, X_m$ , we obtain

$$\begin{aligned} Q_g^* &= 1 + \sum_{m=1}^{\infty} \left[ 1 - \mathbb{E} \left( \frac{T(f_1)}{g(f_1)} / \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} \right)^m \right] \\ &= \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} / \mathbb{E} \left( \frac{T(f_1)}{g(f_1)} \right) \\ &= \text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)} / \int_{\mathbb{H}} T(f_1) H(df_1). \end{aligned} \quad (3.15)$$

This finishes the proof since  $c_T \in (0, \infty)$ . □

**Remark 3.3.** In the proof of Proposition 3.2, the distribution of the random variable  $M_g^*$  defined in (3.10) is calculated. We see that  $M_g^*$  follows a geometric distribution with parameter  $\int_{\mathbb{H}} T(h)H(dh) / \text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} (f(y)/g(f))$ . Therefore, a density  $g$  that minimizes  $\text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} (f(y)/g(f))$  does not only minimize the expectation of  $M_g^*$ , but also other characteristics. For example, the probability  $\mathbb{P}(M_g^* > m_0)$  for a given  $m_0 \in \mathbb{N}$ , or the quantile of  $M_g^*$  at a given probability level. However, this property does not necessarily hold for the stochastic number  $M_g$  in the actual stopping rule (2.5).

We carry on to find the density  $g$  that minimizes  $\text{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)}$ . Instead of considering the supremum in the numerator, we deal with a broader class of functionals in the following proposition.

**Proposition 3.4.** *Let  $L : \mathbb{H} \rightarrow (0, \infty)$  be measurable and  $c_L := \int_{\mathbb{H}} L(f)H(df) < \infty$ . Then,*

$$g_L(f) = c_L^{-1}L(f) \tag{3.16}$$

is an element of

$$\mathcal{G}^{(L)} = \arg \min_g \text{esssup}_{f \in \mathbb{H}} \frac{L(f)}{g(f)}.$$

Furthermore, for every  $g \in \mathcal{G}^{(L)}$ , Equation (3.16) holds for  $H$ -a.e.  $f \in \mathbb{H}$ .

**Proof.** First, by contradiction, we show that the inequality

$$\text{esssup}_{f \in \mathbb{H}} \frac{L(f)}{g(f)} \geq c_L \tag{3.17}$$

holds for all  $g$ . So, assume that (3.17) does not hold for some  $g$  considered in Proposition 2.1. Then some  $\varepsilon > 0$  and some density  $g$  with  $\int g(f)H(df) = 1$  exist such that, for almost all  $f \in \mathbb{H}$ , we have  $L(f)/g(f) \leq c_L - \varepsilon$ . Hence,

$$c_L = \int_{\mathbb{H}} L(f)H(df) \leq (c_L - \varepsilon) \int_{\mathbb{H}} g(f)H(df) < c_L$$

which is a contradiction. Hence, (3.17) is proved. Note that the choice  $g(f) = c_L^{-1}L(f)$  implies equality in (3.17). The first assertion follows.

For the proof of the second assertion, assume that there is some  $g \in \mathcal{G}^{(L)}$  such that (3.16) does not hold for  $H$ -a.e.  $f \in \mathbb{H}$ . Then, as

$$\int_{\mathbb{H}} g(f)H(df) = 1 = \int_{\mathbb{H}} c_L^{-1}L(f)H(df),$$

there is some set  $A \subset \mathbb{H}$  with  $H(A) > 0$  such that, for all  $f \in A$ ,  $g(f) < c_L^{-1}L(f)$ , but  $g(f) > 0$  by (2.1). This yields  $gH(A) > 0$  and, hence,  $\text{esssup}_{f \in \mathbb{H}} L(f)/g(f) > c_L$ , which is a contradiction to  $g \in \mathcal{G}^{(L)}$ . □

**Remark 3.5.** Let  $\mathcal{L}_p(H)$  be the space of  $p$ -integrable functionals with respect to  $H$ . Then  $c_L = \|L\|_{\mathcal{L}_1(H)}$  and  $\text{esssup}_{f \in \mathbb{H}} \frac{L(f)}{g(f)} = \|L/g\|_{\mathcal{L}_\infty(H)}$ . Proposition 3.4 states a special case of Hölder’s inequality for all density functions  $g$ ,  $c_L \leq \|L/g\|_{\mathcal{L}_\infty(H)} \|g\|_{\mathcal{L}_1(H)} = \|L/g\|_{\mathcal{L}_\infty(H)}$  and equality holds if and only if  $g$  is proportional to  $L$  ( $H$ -a.e.).

Similarly, in importance sampling, the second moment  $\mathbb{E}(\tilde{f}(X_1)/g(X_1))^2 = \|\tilde{f}/\sqrt{g}\|_2^2$  is intended to be small. Now,  $c_g := \|\tilde{f}\|_1 \leq \|\tilde{f}/\sqrt{g}\|_2 \|\sqrt{g}\|_2 = \mathbb{E}(\tilde{f}(X_1)/g(X_1))^2$  and equality holds if and only if  $g$  is proportional to  $\tilde{f}$ .

Both  $c_L$  and  $c_g$  can simply be regarded as a factor that normalize  $L$  and  $\tilde{f}$ , respectively.

The results stated above enable us to describe the solution of the optimization problem (3.10). To be rigorous, we first give a necessary and sufficient condition for the solvability of the problem. Here an optimization problem

$$\arg \min_{x \in A} h(x), \quad h : A \rightarrow \mathbb{R} \cup \{\infty\},$$

is *solvable* if  $\inf_{x \in A} h(x) \in (-\infty, \infty)$  and there exists some  $x_0 \in A$  such that  $h(x_0) = \inf_{x \in A} h(x)$ . Our key theorem is given as follows.

**Theorem 3.6.** *The optimization problem (3.10) is solvable if and only if*

$$c = \int_{\mathbb{H}} \sup_{y \in K} f(y) H(df) < \infty.$$

Assuming  $c < \infty$ , the solution is given as

$$g^*(f) := c^{-1} \sup_{y \in K} f(y), \quad f \in \mathbb{H}. \tag{3.18}$$

The solution is unique  $H$ -a.s.

**Proof.** If  $c = \infty$ , Equation (3.13) and Proposition 3.2 yield that (3.10) is not solvable. For  $c < \infty$ , the solution and its uniqueness follow directly from Propositions 3.2 and 3.4 with  $L(f) = \sup_{y \in K} f(y)$ .  $\square$

**Remark 3.7.** It is obvious that  $g^*$  is also the  $H$ -a.s. solution for the two examples of the modified optimization problem in (3.11) and (3.12) and these problems are solvable if and only if  $c < \infty$ . Consequently, the original optimization problem (3.2) is not solvable if  $c = \infty$ , because  $Q_g^{(1)} \leq Q_g$  for any  $g$ .

We close this subsection with analyzing under which conditions  $c$  is finite, that is, the modified optimization problem is solvable. For instance, if  $Z$  is sample-continuous or if  $K$  is finite,  $c < \infty$  follows from a result by Resnick and Roy [24] (see also de Haan and Ferreira [7], Thm. 9.6.1) who showed the equivalence of the first and third assertion in the following proposition. Our result is more general as it shows the equivalence to further conditions and  $\sup_{y \in K} f(y)$  is replaced by a general max-linear functional  $L$ .

**Proposition 3.8.** *Assume the framework of Proposition 2.1. Furthermore, assume that the functional  $L : \mathbb{H} \rightarrow (0, \infty)$  is measurable and max-linear. Then the following conditions are equivalent:*

1.  $c_L := \int_{\mathbb{H}} L(f)H(df) < \infty$ .
2.  $\mathbb{P}(L(Z) \leq a) > 0$  for some  $a > 0$ .
3.  $\mathbb{P}(L(Z) < \infty) = 1$  (or, equivalently,  $\mathbb{P}(L(Z) < \infty) > 0$ ).

If there is some stochastic process  $W$  such that

$$Z =^d \max_{t \in \Pi_0} t W_t, \tag{3.19}$$

where  $\Pi_0$  is a Poisson point process on  $(0, \infty)$  with intensity  $t^{-2} dt$  and  $W_t, t \in \Pi_0$ , are independent copies of  $W$ , we get another equivalent condition:

4.  $\mathbb{E}L(W) < \infty$ .

**Proof.** The assertion follows from the following continued equality:

$$\begin{aligned} \exp\left(-\frac{c_L}{a}\right) &= \exp\left(-\int_{\mathbb{H}} \int_{a/L(f)}^{\infty} t^{-2} dt H(df)\right) = \mathbb{P}(L(Z) \leq a) \\ &= \exp\left(-\mathbb{E}_W\left(\int_{a/L(W)}^{\infty} u^{-2} du\right)\right) = \exp(-a^{-1}\mathbb{E}L(W)) \end{aligned}$$

for any  $a > 0$ . The equivalence to the third assertion follows from the relation  $\mathbb{P}(L(Z) < \infty) = \lim_{a \rightarrow \infty} \mathbb{P}(L(Z) \leq a)$ . □

### 3.3. The normalized spectral representation

Plugging in the solution  $g^*$  of (3.10) given by (3.18) into (2.2), we obtain  $Z =_d \tilde{Z}$  with

$$\tilde{Z}(y) = \max_{t \in \Pi_0} t \frac{c F_t(y)}{\sup_{\tilde{y} \in K} F_t(\tilde{y})}, \quad y \in K, \tag{3.20}$$

where  $\Pi_0$  is a Poisson point process on  $(0, \infty)$  with intensity  $t^{-2} dt$  and  $F_t, t \in \Pi_0$ , are independent random processes with density  $g^*(f)H(df)$ . It can be verified that the transformed spectral functions  $\{c F_t / \sup_{y \in K} F_t(y)\}$  are independent copies of a stochastic process  $F^*$  with

$$\sup_{y \in K} F^*(y) \equiv c. \tag{3.21}$$

We define such a representation as the normalized spectral representation as follows.

**Definition 3.9.** Let  $Z$  be a max-stable process on  $K$  satisfying

$$Z =^d \max_{t \in \Pi_0} t F_t^*. \tag{3.22}$$

Here,  $\Pi_0$  is a Poisson point process on  $(0, \infty)$  with intensity  $t^{-2}dt$  and  $F_t^*$ ,  $t \in \Pi_0$ , are independent copies of a stochastic process  $F^*$  satisfying (3.21) for some  $c \in (0, \infty)$ . Then, the right-hand side of (3.22) is called normalized spectral representation of  $Z$ .

Theorem 3.6 implies that the normalized spectral representation exists if and only if  $c < \infty$ . The following proposition implies that the constant  $c$  and the finite-dimensional distributions of the processes  $F_t^*$  in the normalized spectral representation are uniquely determined by the process  $Z$ .

**Proposition 3.10.** *Let  $Z$  be a max-stable process with a normalized spectral representation. Furthermore, let  $Z^K := \sup_{y \in K} Z(y)$ . Then, we have*

1.  $c = -\log \mathbb{P}(Z^K \leq 1)$ .
2. For any  $y_1, \dots, y_n \in K$ ,  $v_1, \dots, v_n > 0$ , it holds

$$\mathbb{P}(F^*(y_i) \leq v_i, 1 \leq i \leq n) = \lim_{z \rightarrow \infty} \mathbb{P}\left(\frac{Z(y_i)}{Z^K} \leq \frac{v_i}{c}, 1 \leq i \leq n \mid Z^K > z\right). \tag{3.23}$$

**Proof.** The first part is a consequence of the proof of Proposition 3.8. In order to prove the second part, we provide an lower and an upper bound for the probability of the event

$$A = \left\{ Z(y_i) \leq v_i Z^K / c, i = 1, \dots, n, Z^K > z \right\}.$$

To this end, let  $\tilde{\Pi} = \{(t, F_t^*) : t \in \Pi_0\}$  and note that  $Z^K > z$  if and only if there is some  $(u, w) \in \tilde{\Pi}$  such that  $Z^K = uc = u \sup_{y \in K} w(y) > z$ . Now, suppose that further  $Z(y_i) \leq v_i Z^K / c$  holds for all  $i = 1, \dots, n$ . Then, this point  $(u, w)$  necessarily satisfies  $w(y_i) \leq v_i$  for all  $i = 1, \dots, n$ . This yields an upper bound for  $\mathbb{P}(A)$ . If, on the other hand,  $Z(y_i) > v_i Z^K / c$  for some  $i = 1, \dots, n$ , then one point  $(u, w) \in \tilde{\Pi}$  with  $uc > z$  satisfies  $w(y_i) > v_i$  for some  $i = 1, \dots, n$  or there is some point  $(u, w) \in \tilde{\Pi}$  with  $uc \leq z$ , but  $uw(y_i)/z > v_i/c$  for some  $i = 1, \dots, n$ . Considering the complementary probabilities, we obtain a lower bound. Summing up, we have

$$\begin{aligned} & \mathbb{P}\left(\left| \tilde{\Pi} \cap \left\{ (u, w) : u > \frac{z}{c} \right\} \right| > 0, \left| \tilde{\Pi} \cap \left\{ (u, w) : u > \frac{z}{c}, 1 > \min_{1 \leq i \leq n} \frac{v_i}{w(y_i)} \right\} \right| = 0, \right. \\ & \left. \left| \tilde{\Pi} \cap \left\{ (u, w) : u \leq \frac{z}{c}, \frac{u}{z} > \min_{1 \leq i \leq n} \frac{v_i}{cw(y_i)} \right\} \right| = 0\right) \\ & \leq \mathbb{P}(A) \\ & \leq \mathbb{P}\left(\left| \tilde{\Pi} \cap \left\{ (u, w) : u > \frac{z}{c}, \max_{1 \leq i \leq n} \frac{w(y_i)}{v_i} \leq 1 \right\} \right| > 0\right). \end{aligned} \tag{3.24}$$

The lower bound in (3.24) equals

$$\begin{aligned} & \left( 1 - \exp\left(-\frac{c}{z} \mathbb{P}(F^*(y_i) \leq v_i, 1 \leq i \leq n)\right) \right) \\ & \cdot \exp\left(-\frac{c}{z} \mathbb{P}\left(\max_{1 \leq i \leq n} \frac{F^*(y_i)}{v_i} > 1\right)\right) \exp\left(-\mathbb{E} \int_{\frac{z}{c} \wedge \min_{1 \leq i \leq n} \frac{zv_i}{cF^*(y_i)}}^{\frac{z}{c}} u^{-2} du\right) \end{aligned}$$

$$= \left( 1 - \exp\left(-\frac{c}{z}\mathbb{P}(F^*(y_i) \leq v_i, 1 \leq i \leq n)\right) \right) \cdot \exp\left(-\frac{c}{z}\mathbb{P}\left(\max_{1 \leq i \leq n} \frac{F^*(y_i)}{v_i} > 1\right)\right) \exp\left(-\frac{c}{z}\mathbb{E}\left(\max_{1 \leq i \leq n} \frac{F^*(y_i)}{v_i} - 1\right)_+\right),$$

while the upper bound equals  $1 - \exp(-\frac{c}{z}\mathbb{P}(F^*(y_i) \leq v_i, 1 \leq i \leq n))$ . By using that  $\mathbb{P}(Z^K > z) = 1 - e^{-c/z}$  and taking the limit  $z \rightarrow \infty$ , inequality (3.24) yields (3.23).  $\square$

Proposition 3.10 implies that the solution of the optimization problem (3.10) is unique in two different aspects. First, as stated in Theorem 3.6, any solution  $g \in \mathcal{G}^*$  satisfies  $g = g^*$   $H$ -a.s. Second, the finite-dimensional distributions of the normalized spectral functions  $\{f/g^*(f)\}$  do not depend on the initial choice of the spectral functions. In particular, the normalized spectral representation is unique if  $\mathbb{H}$  is the space of continuous functions on  $K$  equipped with the Borel  $\sigma$ -algebra, that is, the product  $\sigma$ -algebra on  $[0, \infty)^K$ .

Summarizing the results in this section, we suggest to make use of the normalized spectral representation for exact and efficient simulation as it is the unique solution to the modified optimization problem provided that  $c < \infty$ . Analogously to Algorithm 2.2, the algorithm for simulation via the normalized spectral representation is given as follows.

**Algorithm 3.11.**

Set  $m = 0$  and  $Z^{(0)}(y) = 0$  for all  $y \in K$ .

Simulate a standard exponentially distributed random variable  $E_1$ .

while  $\text{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{cf(y)}{\sup_{\tilde{y} \in K} f(\tilde{y})Z^{(m)}(y)} > \sum_{j=1}^{m+1} E_j$  {

Update  $m$  by  $m + 1$ .

Sample  $F_m$  with density  $c^{-1} \sup_{y \in K} f(y)H(df)$ .

Set  $Z^{(m)}(y) = \max\{Z^{(m-1)}(y), \frac{c}{\sum_{j=1}^m E_j} \frac{F_m(y)}{\sup_{\tilde{y} \in K} F_m(\tilde{y})}\}$  for all  $y \in K$ .

Simulate a standard exponentially distributed random variable  $E_{m+1}$ .

}

Return  $Z^{(m)}$ .

**Remark 3.12.** In practice, with the normalized spectral representation, the criterion in the while-loop in Algorithm 3.11 might be replaced by the weaker criterion  $\frac{c}{\inf_{y \in K} Z^{(m)}(y)} > \sum_{j=1}^{m+1} E_j$ . Although the use of the weaker stopping rule based on this criterion may increase the number of iterations, it simplifies the evaluation of the essential supremum.

The constant  $c$  plays the role analogous to the bound of the quotient of the genuine density function and the proposed density function in rejection sampling. In particular,  $c$  controls the speed of the algorithm, while (2.5) guarantees the exactness.

Finally, we investigate the sampling of  $F_m$  in Algorithm 3.11. The simulation of the stochastic processes  $F_m = \{F_m(y), y \in K\}$ ,  $m \in \mathbb{N}$ , according to the transformed measure  $g^*H$  may not be always straightforward. For some processes, such as the moving maxima processes (cf. Section 5), the distribution of  $F_m$  can be calculated explicitly, which allows for efficient sampling. For many other processes, such as Brown–Resnick or extremal  $t$  processes, there is no

direct way to simulate the normalized spectral functions. One solution is to use the following Metropolis–Hastings algorithm.

**Algorithm 3.13.**

Simulate  $f^{(1)}$  according to the law  $H$ .

for  $k = 1, \dots, n_{MCMC} - 1$  {

Sample  $f^{\text{prop}}$  from  $H$  and set

$$f^{(k+1)} = \begin{cases} f^{\text{prop}} & \text{with probability } \min \left\{ \frac{\sup_{y \in K} f^{\text{prop}}(y)}{\sup_{y \in K} f^{(k)}(y)}, 1 \right\}, \\ f^{(k)} & \text{with probability } \max \left\{ 1 - \frac{\sup_{y \in K} f^{\text{prop}}(y)}{\sup_{y \in K} f^{(k)}(y)}, 0 \right\}. \end{cases}$$

}

Return  $f^{(n_{MCMC})}$ .

If  $H$  is a probability measure, Algorithm 3.13 generates a Markov chain of length  $n_{MCMC}$  whose stationary distribution is  $g^*H$  based on simulations from  $H$ .

### 4. Evaluating the modified optimization problem

In this section, we discuss how close the modified optimization problem and the original problem are. Considering the two examples (3.11) and (3.12) of the modified problem, we have that  $Q_g^{(1)} \leq Q_g^{(2)}(y_0) \leq Q_g$  for all  $g$  under some mild condition on  $\mathbb{H}$  (see the proof of Proposition 4.1). Therefore, the modified optimization problem is in fact minimizing a lower bound of  $g \mapsto Q_g$ . This section proceeds in two steps. First, we will improve the lower bound and show that the normalized spectral representation also minimizes the improved lower bound function. Second, we give a formula for calculating  $Q_g$ . In particular, this formula allows the calculation of  $Q_{g^*}$ , that is, the expected number of steps when simulating  $Z$  using the normalized spectral representation and the *original* stopping rule (2.5).

We start giving bounds for the expected number of iterations in the simulation algorithm. The proof is given in the appendix.

**Proposition 4.1.** *Assume the existence of a countable subset  $K_0 \subset K$  such that  $\sup_{y \in K} f(y) = \sup_{y \in K_0} f(y)$  for  $H$ -a.e.  $f \in \mathbb{H}$ . Denoting by  $g_0 \in \mathcal{G}$  a solution of the original optimization problem, we get for all  $y_0 \in K$ ,*

$$1 = Q_{g^*}^{(1)} \leq Q_{g^*}^{(2)}(y_0) \leq Q_{g_0} \leq Q_{g^*}.$$

As Proposition 4.1 shows, approximating the optimal number of steps in the original problem (3.2) by the solution for the two example problems (3.11) and (3.12) might be quite vague. In particular, the achieved minimum of  $Q_g^{(1)}$  always equals to 1. In other words, some spectral functions that in fact contribute to  $Z$  are not taken into account in the calculation of  $Q_g^{(1)}$  or  $Q_g^{(2)}$ .



To overcome the aforementioned pitfall, we proceed with a theoretical investigation to improve the lower bound of  $Q_g$  for a given density  $g$ . The idea is to evaluate  $Q_g$  by separating two types of spectral functions. First, we consider those functions that contribute to the max-stable process under the transformed spectral representation. Second, we deal with those functions that do not contribute, but are taken into consideration because of the stopping rule. To this end, we replace the processes  $Z^{(m)}$  occurring in the construction of  $Z$  by the final process  $Z^{(\infty)}$  given by (2.4). Theoretically, this will not affect the construction of the process because once we stop after  $m$  steps according to the stopping rule (2.5), we have  $Z^{(m)} = Z^{(\infty)}$ . However, the modified inequality is not a stopping rule. This is why we call the analysis below a “theoretical investigation”.

We further assume that  $\mathbb{H}$  corresponds to the space  $C_+(K)$  of nonnegative continuous functions on  $K$  endowed with the Borel  $\sigma$ -algebra  $\mathcal{H}$ . This ensures the finiteness of  $c$ , and thus the solvability of all the optimization problems. Note that all the results of this section hold true for more general spaces  $\mathbb{H}$ , see Oesting, Schlather and Zhou [20].

In order to separately consider these two types of functions, we adopt the concepts of  $K$ -extremal and  $K$ -subextremal points introduced by Dombry and Eyi-Minko [12] and Dombry and Eyi-Minko [11] as follows.

**Definition 4.2.** Let  $\Phi$  be some Poisson point process on  $(0, \infty) \times C_+(K)$  with intensity measure  $u^{-2} du \times \nu(dh)$  where  $\nu$  is a locally finite measure on  $C_+(K)$ . We call  $(t^*, h^*) \in \Phi$  a  $K$ -extremal point and write  $(t^*, h^*) \in \Phi_K^+$  if and only if

$$t^*h^*(y) = \max_{(t,h) \in \Phi} th(y) \quad \text{for some } y \in K.$$

Otherwise, i.e. if  $t^*h^*(y) < \max_{(t,h) \in \Phi} th(y)$  for all  $y \in K$ ,  $(t^*, h^*) \in \Phi$  is called a  $K$ -subextremal point and we write  $(t^*, h^*) \in \Phi_K^-$ .

In contrast to Dombry and Eyi-Minko [12], we are interested in tuples  $(t, h)$  instead of the product  $th$ . Therefore, we generalize a result given in Dombry and Eyi-Minko [12] and show that the random sets  $\Phi_K^+$  and  $\Phi_K^-$  are point processes on  $(0, \infty) \times C_+(K)$ , i.e.  $\Phi_K^+(S)$  and  $\Phi_K^-(S)$  are random variables for any bounded set  $S \in \mathcal{B} \times C_+(K)$ ; see the following proposition. The proof runs analogously using the fact that the mapping  $\phi : (0, \infty) \times C_+(K) \rightarrow C_+(K)$ ,  $(t, h) \mapsto th(\cdot)$  is measurable and is therefore omitted.

**Proposition 4.3.**  $\Phi_K^+$  and  $\Phi_K^-$  are point processes on  $(0, \infty) \times C_+(K)$ .

To apply the theory of extremal and subextremal points in the construction of the process  $Z^{(\infty)}$ , we define the Poisson point process

$$\Phi = \{(t, f/g(f)) : (t, f) \in \tilde{\Pi}\}.$$

Similar to the proof of Lemma 3.2 in Dombry and Eyi-Minko [11], the following lemma characterizes the points of the point process  $\Phi_K^-$ .

**Lemma 4.4.** *Conditional on  $Z^{(\infty)}$ , the point process  $\Phi_K^-$  is a Poisson point process on  $(0, \infty) \times C_+(K)$  with intensity measure*

$$\frac{d\tilde{\Lambda}^-}{dt \times dH}(t, h) = t^{-2}g(h) \cdot \mathbf{1}_{th(\cdot)/g(h) < Z^{(\infty)}(\cdot)}.$$

In addition, we calculate the expected number of points in the point process  $\Phi_K^+$  in the following lemma.

**Lemma 4.5.** *We have*

$$\mathbb{E}|\Phi_K^+| = \mathbb{E}_Z \left( \int_{C_+(K)} \sup_{y \in K} \frac{f(y)}{Z(y)} H(df) \right)$$

which does not depend on the choice of  $g$ .

**Proof.** Let  $B = [t_0, \infty) \times C_+(K)$  with  $t_0 > 0$ . Then, we have  $\mathbb{E}|\Phi_K^+ \cap B| = \mathbb{E}|\Phi \cap B| - \mathbb{E}|\Phi_K^- \cap B|$ . Conditioning on  $Z^{(\infty)}$ , Lemma 4.4 yields

$$\begin{aligned} \mathbb{E}|\Phi_K^+ \cap B| &= \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t \geq t_0} dt g(f) H(df) \\ &\quad - \mathbb{E}_{Z^{(\infty)}} \left( \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t \geq t_0} \mathbf{1}_{\frac{1}{t} > \sup_{y \in K} \frac{f(y)}{g(f)Z^{(\infty)}(y)}} dt g(f) H(df) \right) \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t \geq t_0} \mathbf{1}_{\frac{1}{t} \leq \sup_{y \in K} \frac{f(y)}{g(f)Z(y)}} dt g(f) H(df) \right). \end{aligned}$$

Considering a monotone sequence  $t_{0,n} \searrow 0$  as  $n \rightarrow \infty$ , the monotone convergence theorem yields

$$\begin{aligned} \mathbb{E}|\Phi_K^+| &= \mathbb{E}_Z \left( \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t > 1/\sup_{y \in K} \frac{f(y)}{g(f)Z(y)}} dt g(f) H(df) \right) \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \sup_{y \in K} \frac{f(y)}{Z(y)} H(df) \right), \end{aligned}$$

which completes the proof. □

The first type of spectral functions that contribute to  $M_g$  correspond to the points  $(t^*, h^*) \in \Phi_K^+$ . Thus, we can rewrite (3.2) as

$$Q_g = \mathbb{E}|\Phi_K^+| + \mathbb{E} \left( \left| \left\{ (t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(\infty)}(y)} > \frac{1}{t} \right\} \right| \right). \tag{4.1}$$

The second term in the right-hand side of (4.1) corresponds to the number of the second type of spectral functions: they do not contribute to the max-stable process but are counted due to the

stopping criterion (2.5). Notice that the component  $\mathbb{E}|\Phi_K^+|$  is independent of the choice of  $g$ . We thus modify the optimization problems by maintaining this component, while refining the second component in an analogous way as the modification in (3.10). This results in a refined version of the modified optimization problem (3.10) as

$$\tilde{Q}_g^* = \mathbb{E}|\Phi_K^+| + \mathbb{E}\left(\left|\left\{(t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f)T(Z^{(\infty)})} > \frac{1}{t}\right\}\right|\right). \tag{4.2}$$

The following proposition relates the minimizer of (4.2) to the solution of our previously modified optimization problem,  $g^* \in \mathcal{G}^*$ . In addition, it provides a formula for  $Q_g$  for any given  $g$  based on the results in Lemma 4.4 and 4.5.

**Proposition 4.6.**

1. For any  $g$ , we have

$$Q_g = \mathbb{E}_Z\left(\text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)Z(y)}\right). \tag{4.3}$$

2. For any max-linear function  $T$ , it holds

$$\arg \min_g \tilde{Q}_g^* \supset \arg \min_g \text{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f)} = \mathcal{G}^*,$$

where  $\tilde{Q}_g^*$  is as in (4.2) and  $\mathcal{G}^* = \arg \min_g Q_g^*$ .

**Proof.** Let  $Z^{(\infty)}$  be given by (2.4). By Lemma 4.5, we obtain

$$Q_g = \mathbb{E}_Z \int_{C_+(K)} \sup_{y \in K} \frac{f(y)}{Z(y)} H(df) + \mathbb{E}\left(\left|\left\{(t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(\infty)}(y)} > t^{-1}\right\}\right|\right).$$

Conditioning on  $Z^{(\infty)}$ , Lemma 4.4 yields

$$\begin{aligned} & \mathbb{E}\left(\left|\left\{(t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(\infty)}(y)} > t^{-1}\right\}\right|\right) \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t > 1/\text{esssup}_{h \in C_+(K)} \sup_{y \in K} \frac{h(y)}{g(h)Z(y)}} \mathbf{1}_{t < 1/\sup_{y \in K} \frac{f(y)}{g(f)Z(y)}} dt g(f) H(df) \right) \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \left\{ \text{esssup}_{h \in C_+(K)} \sup_{y \in K} \frac{h(y)g(f)}{g(h)Z(y)} - \sup_{y \in K} \frac{f(y)}{Z(y)} \right\}_+ H(df) \right) \\ &= \mathbb{E}_Z \left[ \text{esssup}_{h \in C_+(K)} \sup_{y \in K} \frac{h(y)}{g(h)Z(y)} \right] - \mathbb{E}_Z \int_{C_+(K)} \sup_{y \in K} \frac{f(y)}{Z(y)} H(df). \end{aligned}$$

In the last step we used the fact that

$$\text{esssup}_{h \in C_+(K)} \sup_{y \in K} \frac{h(y)g(f)}{g(h)Z(y)} - \sup_{y \in K} \frac{f(y)}{Z(y)} \geq 0$$

for  $H$ -a.e.  $f \in C_+(K)$ . The first assertion follows.

Analogously to the first part, we get that

$$\begin{aligned} \tilde{Q}_g^* &= \mathbb{E}_Z \int_{C_+(K)} \sup_{y \in K} \frac{f(y)}{Z(y)} H(df) \\ &\quad + \mathbb{E} \left( \left| \left\{ (t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f)T(Z(\infty))} > t^{-1} \right\} \right| \right) \end{aligned}$$

and

$$\begin{aligned} &\mathbb{E} \left( \left| \left\{ (t, h) \in \Phi_K^- : \text{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f)T(Z(\infty))} > t^{-1} \right\} \right| \right) \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \int_0^\infty t^{-2} \mathbf{1}_{t > 1/\text{esssup}_{h \in C_+(K)} \sup_{y \in K} \frac{h(y)}{g(h)T(Z)}} \mathbf{1}_{t < 1/\sup_{y \in K} \frac{f(y)}{g(f)Z(y)}} dt g(f) H(df) \right) \tag{4.4} \\ &= \mathbb{E}_Z \left( \int_{C_+(K)} \left\{ \text{esssup}_{h \in C_+(K)} \frac{\sup_{y \in K} h(y)}{g(h)T(Z)} g(f) - \sup_{y \in K} \frac{f(y)}{Z(y)} \right\}_+ H(df) \right) \\ &\geq \mathbb{E}_Z \left( \int_{C_+(K)} \left\{ \frac{\sup_{y \in K} f(y)}{T(Z)} - \sup_{y \in K} \frac{f(y)}{Z(y)} \right\}_+ H(df) \right). \end{aligned}$$

Now, let  $g \in \mathcal{G}^* = \arg \min_g \text{esssup}_{f \in C_+(K)} (\sup_{y \in K} f(y)/g(f))$ . Then, by Theorem 3.4, we have that  $g(f) = c^{-1} \sup_{y \in K} f(y)$  for all  $f \in C_+(K)$ ,  $H$ -a.e. Thus, we get equality in Equation (4.4) and hence  $\mathcal{G}^* \subset \arg \min_g \tilde{Q}_g^*$ . □

Proposition 4.6 leads to two implications in applications. First, it facilitates the numerical calculation of  $Q_g = \mathbb{E}(M_g)$  by simulation. While Equation (4.3) is difficult to be evaluated exactly in many cases, it may be used to obtain bounds for  $\mathbb{E}M_g$  such as

$$\text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)} \leq Q_g = \mathbb{E}M_g \leq \text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)} \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right]. \tag{4.5}$$

This confirms the finding of Proposition 3.1 that  $\text{esssup}_{f \in C_+(K)} \sup_{y \in K} \frac{f(y)}{g(f)} < \infty$ , is a necessary condition for  $\mathbb{E}M_g$  to be finite.

In particular, the result can be applied to analyze  $Q_{g^*}$ . With  $\frac{f(y)}{g^*(f)} \leq c$ , we obtain

$$Q_{g^*} \leq c \cdot \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right]. \tag{4.6}$$

This yields that  $Q_{g^*} < \infty$  if  $Z$  is sample-continuous, as in this case we have  $c < \infty$  and  $\mathbb{E}((\inf_{y \in K} Z(y))^{-1}) < \infty$  holds. In other words, the expectation of the stochastic number  $M_{g^*}$  based on the normalized representation is finite. Note that the upper bound in (4.6) is reached under some mild conditions, see Oesting, Schlather and Zhou [20]. We will see that all examples considered in the present paper meet these conditions.

Second, Proposition 4.6 implies that the minimal value for  $\tilde{Q}_g^*$  can be achieved by any  $g^* \in \mathcal{G}^*$ . As further  $Q_g^* \leq \tilde{Q}_g^*$  by construction, we obtain the following corollary.

**Corollary 4.7.** *The optimization problem given in (4.2) is solvable if and only if the optimization problem (3.10) is solvable (cf. Theorem 3.6). With solvability, the normalized spectral representation is an optimal solution to (4.2) and (3.10).*

This corollary further confirms that using the normalized spectral representation may lead to an efficient and exact simulation, because the refined optimization problem (4.2) is closer to the original optimization problem than the modified problem (3.10).

Lastly, we improve the lower bounds for  $Q_{\tilde{g}}$ . To this end, refined versions of the examples in (3.11) and (3.12) are considered:

$$\tilde{Q}_g^{(1)} = \mathbb{E}|\Phi_K^+| + \mathbb{E}\left(\left|\left\{(t, h) \in \Phi_K^- : \operatorname{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f) \sup_{\tilde{y} \in K} Z^{(\infty)}(\tilde{y})} > \frac{1}{t}\right\}\right|\right), \quad (4.7)$$

$$\tilde{Q}_g^{(2)}(y_0) = \mathbb{E}|\Phi_K^+| + \mathbb{E}\left(\left|\left\{(t, h) \in \Phi_K^- : \operatorname{esssup}_{f \in C_+(K)} \frac{\sup_{y \in K} f(y)}{g(f)Z^{(\infty)}(y_0)} > \frac{1}{t}\right\}\right|\right). \quad (4.8)$$

**Proposition 4.8.** *For any  $g_0 \in \mathcal{G}$ , we have*

$$1 \leq \tilde{Q}_{g^*}^{(1)} \leq \inf_{y_0 \in K} \tilde{Q}_{g^*}^{(2)}(y_0) \leq Q_{g_0} \leq Q_{g^*},$$

where  $g^* \in \mathcal{G}^*$  is given by (3.18).

By definition,  $\tilde{Q}_g^{(1)} \geq Q_g^{(1)}$  and  $\tilde{Q}_g^{(2)} \geq Q_g^{(2)}$  for all  $g$ . Hence, the results in Proposition 4.8 give improved lower bounds of  $Q_{g_0}$ . The proof is analogous to Proposition 4.1 and is thus omitted.

## 5. Example: The normalized spectral representation of moving maxima processes

In this section, we discuss the normalized spectral representation for the class of moving maxima processes which can be simulated via an algorithm of Schlather [26]. In this case, the normalized spectral functions can be calculated explicitly and are convenient to handle, which allows a rather general implementation of the simulation procedure for processes on a grid, see the R package `RandomFields` (R Core Team [23], Schlather *et al.* [27]). The procedure is further compared to Schlather’s [26] algorithm both from a theoretical point of view and in a simulation study in Section 6.

Here, we focus on moving maxima processes on a compact set  $K \subset \mathbb{R}^d$ , i.e. processes of the form

$$Z(y) =_d \max_{(t,x) \in \Pi_{M2}} th(y-x), \quad y \in K, \tag{5.1}$$

where  $\Pi_{M2}$  is a Poisson point process on  $(0, \infty) \times \mathbb{R}^d$  with intensity  $t^{-2} dt \Lambda(dx)$ ,  $\Lambda$  denotes the Lebesgue measure on  $\mathbb{R}^d$  and  $h : \mathbb{R}^d \rightarrow [0, \infty)$  is a so-called shape function satisfying  $\int_{\mathbb{R}^d} h(x) dx = 1$ . Thus,  $Z$  has a spectral representation of form (1.1) with  $H(A) = \lambda(\{x \in \mathbb{R}^d : h(\cdot - x) \in A\})$ ,  $A \in \mathbb{H}$ .

In the following, we will explicitly calculate the distribution of the normalized spectral functions for moving maxima processes. First, we note that, due to the specific structure of a moving maxima process  $Z$  defined in (5.1), its normalized spectral representation can be written as

$$Z(y) =_d \max_{t \in \Pi_0} ct \frac{h(y - X_t)}{\tilde{h}(X_t)}, \quad y \in K,$$

where  $\Pi_0$  is a Poisson point process on  $(0, \infty)$  with intensity measure  $t^{-2} dt$  and  $X_t, t \in \Pi_0$ , are independent random vectors with Lebesgue density  $c^{-1} \tilde{h}(x) dx$ ,  $\tilde{h}(x) = \sup_{y \in K} h(y-x)$  and  $c = \int_{\mathbb{R}^d} \tilde{h}(x) dx$ . Thus, both the function  $\tilde{h}$  and the constant  $c$  are crucial for the simulation of the normalized spectral functions. Further,  $c$  also occurs in the stopping rule and thus influences the number of spectral functions considered in Algorithm 3.11. If  $h$  is continuous or  $K$  is discrete, the upper bound (4.6) is reached, i.e.

$$Q_{g^*} = \mathbb{E}M_{g^*} = c \cdot \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right].$$

In general, the term  $\mathbb{E}[(\inf_{y \in K} Z(y))^{-1}]$  cannot be calculated analytically, but needs to be estimated via simulations. For the implementation of Algorithm 3.11, however, only the constant  $c$  and the function  $\tilde{h}$  are needed, both of them depending on the shape function  $h$  and the geometry of the set  $K$ .

In the following, we will calculate  $c$  and  $\tilde{h}$  under different assumptions on the domain  $K$ . We restrict ourselves to the case where the shape function  $h$  is radial symmetric and non-increasing, that is,  $h(x) = f_0(\|x\|)$  for a non-increasing function  $f_0 : [0, \infty) \rightarrow [0, \infty)$ . Then, in general,

$$\tilde{h}(x) := \sup_{y \in K} f_0(\|y-x\|) = f_0\left(\min_{y \in K} \|y-x\|\right).$$

First, consider the case that  $K$  is a  $d$ -dimensional ball  $b(0, R)$  centered at the origin with radius  $R$ , that is,  $K = \{x \in \mathbb{R}^d : \|x\| \leq R\}$ . Then,  $\tilde{h}(x) = f_0(0) \mathbf{1}_{\|x\| \leq R} + \mathbf{1}_{\|x\| > R} f_0(\|x\| - R)$  and  $c = f_0(0)R^d + d \int_0^\infty (\tilde{r} + R)^{d-1} f_0(\tilde{r}) d\tilde{r} < \infty$ .

Second, consider the case that  $K$  is a  $d$ -dimensional cube, that is, the case that  $K = [-R, R]^d$  for some  $R > 0$ . Then, we get

$$\tilde{h}((x_1, \dots, x_d)) = f_0(\|((|x_1| - R) \vee 0, \dots, (|x_d| - R) \vee 0)\|). \tag{5.2}$$

We consider the subcases  $d = 1$  and  $d = 2$  to derive explicit formulae. If  $d = 1$ , then  $K = [-R, R] = b(0, R)$ , and, according to the formulae above, we get that  $\tilde{h}(x) = \mathbf{1}_{|x| \leq R} f_0(0) + \mathbf{1}_{|x| > R} f_0(|x| - R)$  and thus,

$$c = \int_{\mathbb{R}} \tilde{h}(x) dx = 2Rf_0(0) + \int_{|x|>0} f_0(|x|) dx = 2Rf_0(0) + 1.$$

If  $d = 2$ , we obtain

$$\begin{aligned} \tilde{h}(x) &= \mathbf{1}_{|x_1| \vee |x_2| \leq R} f_0(0) + 2 \cdot \mathbf{1}_{|x_1| \wedge |x_2| \leq R, |x_1| \vee |x_2| > R} f_0(|x_1| \wedge |x_2| - R) \\ &\quad + \mathbf{1}_{|x_1| \wedge |x_2| > R} f_0(\|(|x_1| - R, |x_2| - R)\|). \end{aligned}$$

Thus,

$$c = (2R)^2 \cdot f_0(0) + 4R \cdot \int_{\mathbb{R}} f_0(|x|) dx + \int_{\mathbb{R}^2} f_0(\|x\|) dx = 4R^2 f_0(0) + 4R \int_{\mathbb{R}} f_0(|x|) dx + 1.$$

Next, we further specify explicit examples on the function  $f_0$ , under which the constant  $c$  can be further calculated.

**Example 5.1.**

1. *Indicator function.* We consider the case that the shape function is the indicator function of a ball  $b(0, r)$  with radius  $r > 0$  centered at the origin, i.e.  $f_0(\|x\|) = \mathbf{1}_{\|x\| \leq r}$ . In this case, we have  $\tilde{h}(x) = \mathbf{1}_{K \oplus b(0,r)}(x)$  and  $c = \text{vol}(K \oplus b(0, r))$  where  $\oplus$  denotes morphological dilation and  $\text{vol}$  the  $d$ -dimensional volume. Here, all the finite approximations derived from the normalized spectral representation coincide with the corresponding approximations resulting from the algorithm proposed by Schlather [26]. See Section 6 for details on this algorithm.
2. *Smith model.* As a second example, we consider the Gaussian extreme value process (Smith [29]) where  $f_0$  is a Gaussian density function. Here, for simplicity, we assume the shape function is the density of a multivariate normal random vector  $Y \sim \mathcal{N}(\mathbf{0}, \sigma^2 \text{Id})$  with  $\sigma > 0$ . Thus, it is a radial symmetric monotone function. Let  $K = [-R, R]^d$  for some  $R > 0$ . Then, by the considerations above, we get that  $\tilde{h}$  is of type (5.2) and for  $d = 1, 2$ , we obtain

$$c = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{R}{\sigma} + 1, & d = 1, \\ \frac{2}{\pi} \left(\frac{R}{\sigma}\right)^2 + 2\sqrt{\frac{2}{\pi}} \frac{R}{\sigma} + 1, & d = 2. \end{cases}$$

**Remark 5.2.** Note that the results can easily be generalized to the case of mixed moving maxima processes (Schlather [26], Stoev and Taquq [30]), that is, the case where the deterministic shape function  $h$  is replaced by independent copies of a random function. See Oesting, Schlather and Zhou [20] for details.

## 6. Simulation: Comparison to other algorithms

In this section, we investigate the number of spectral functions needed in simulating a max-stable process, considering the suggested normalized spectral representation as well as other algorithms. In Section 6.1, we compare it with the analogous number in Schlather’s [26] algorithm for moving maxima processes. Then, we compare simulations via the normalized spectral representation to the recent algorithms devised by Dieker and Mikosch [9] and Dombry, Engelke and Oesting [10] focusing on Brown–Resnick processes in Section 6.2. Besides the number of spectral functions, we further consider the actual computational costs accounting for the fact that simulation of the normalized spectral functions is not straightforward in this case.

### 6.1. Comparison to the algorithm proposed in Schlather [26]

Let  $Z$  be a moving maxima process on  $K \subset \mathbb{R}^d$  as defined in (5.1). Assume that the shape function  $h$  is bounded and has compact support, that is,  $h(x) < C$  for all  $x \in \mathbb{R}^d$  for some  $C > 0$  and  $\text{supp}(h) \subset b(0, r)$  for some  $r > 0$ . Schlather’s [26] algorithm considers the following equivalent representation

$$Z(y) =_d |K \oplus b(0, r)| \cdot \max_{1 \leq k \leq M} \frac{h(y - U_k)}{\sum_{i=1}^k E_i}, \quad y \in K,$$

where  $E_i$  are standard exponentially distributed random variables,  $U_i$  are uniformly distributed on  $K \oplus b(0, r)$  and all these random variables are independent. Following Schlather’s [26] algorithm, the number of simulated spectral functions is then a random number defined as

$$M = \min \left\{ m \in \mathbb{N} : \frac{C}{\sum_{i=1}^{m+1} E_i} \leq \inf_{x \in K} \max_{1 \leq k \leq m} \frac{h(x - U_k)}{\sum_{i=1}^k E_i} \right\}.$$

Here, analogously to Proposition 4.6, the following result can be shown.

**Proposition 6.1.** *The expectation of  $M$ , defined as above, equals*

$$\mathbb{E}M = \mathbb{E} \left( \frac{|K \oplus b(0, r)| \cdot C}{\inf_{y \in K} Z(y)} \right).$$

Thus, the ratio between the expected numbers of spectral functions considered by the normalized spectral representation and by Schlather’s [26] algorithm is

$$\frac{Q_{g^*}^*}{\mathbb{E}M} = \frac{\mathbb{E}M_{g^*}}{\mathbb{E}M} = \frac{c}{|K \oplus b(0, r)| \cdot C}.$$

If the shape function  $h$  is bounded, but not compactly supported, the max-stable process  $Z$  can be approximated using a shape function which is cut off outside a compact set  $J$ , i.e.  $h_{\text{cut}}(x) =$



$h(x) \cdot \mathbf{1}_{x \in J}$ . Let  $\tilde{U}_k \sim_{i.i.d.} \text{Unif}(K \oplus \check{J})$  where  $\check{J} = \{-x : x \in J\}$ , and  $Z_J(\cdot)$  be given by

$$Z_J(y) = |K \oplus \check{J}| \cdot \max_{n \in \mathbb{N}} \frac{h_{\text{cut}}(y - \tilde{U}_n)}{\sum_{k=1}^n E_k}, \quad y \in K,$$

Then, the number  $M_J$  of shape functions that need to be considered is finite a.s., and, by Proposition 6.1, its expectation equals  $\mathbb{E}M_J = \mathbb{E}((\inf_{y \in K} Z_J(y))^{-1} \cdot |K \oplus \check{J}| \cdot C)$ . Hence, in the approximative case, the ratio of expected numbers of spectral functions considered by the two algorithms, can be written as the product

$$\frac{Q_{g^*}}{\mathbb{E}M_J} = \frac{\mathbb{E}M_{g^*}}{\mathbb{E}M_J} = A_{K,J} \cdot P_{K,J}, \tag{6.1}$$

where  $A_{K,J} = \frac{c}{|K \oplus \check{J}| \cdot C}$  and  $P_{K,J} = \frac{\mathbb{E}(\sup_{y \in K} Z(y)^{-1})}{\mathbb{E}(\sup_{y \in K} Z_J(y)^{-1})}$ .

The first factor  $A_{K,J}$  refers to the domain of the Poisson point process, and the second factor  $P_{K,J}$  refers to the precision of the approximation by Schlather’s [26] algorithm. As  $h_{\text{cut}}(\cdot) \leq h(\cdot)$ , we have that  $P_{K,J} \leq 1$  with  $\lim_{J \nearrow \mathbb{R}^d} P_{K,J} = 1$ . Thus, we obtain the upper bound  $Q_{g^*}/\mathbb{E}M_J \leq A_{K,J}$  which can be calculated via the formulae for  $c$  obtained in Section 5. The factor  $P_{K,J}$ , however, cannot be calculated explicitly in general, but needs to be accessed via simulation.

In view of these theoretical observations, we perform a simulation study for Smith’s [29] model described in Example 5.1 on a rectangle  $[-R, R]^d$  for  $d = 1, 2$ . For the simulation algorithm of Schlather [26], we need an approximation as described above. Here, a natural choice for cutting off the shape function is  $J = [-k\sigma, k\sigma]^d$  for some  $k \in \mathbb{N}$ . Then, by Example 5.1, the first factor  $A_{R,k} := A_{[-R,R]^d, [-k\sigma, k\sigma]^d}$  in (6.1) equals  $(R/\sigma + \sqrt{\pi/2})^d / (R/\sigma + k)^d$ , i.e.  $A_{R,k} < 1$  if and only if  $k > \sqrt{\frac{\pi}{2}}$ . In order to access  $P_{R,k} := P_{[-R,R]^d, [-k\sigma, k\sigma]^d}$  and the exact values of  $Q_{g^*}$  and  $\mathbb{E}M_{[-k\sigma, k\sigma]^d}$ , we choose  $\sigma = 1$  and simulate  $Z$  and  $Z_{[-k, k]^d}$  for  $k = 2, 3$  on a grid  $K = \{-R, -R + h, \dots, R - h, R\}^d$ ,  $d = 1, 2$ . For simplicity, the normalized spectral representation is chosen as if  $K$  was the rectangle  $[-R, R]^d$ .

In the case  $d = 1$ , for  $h = 0.1$  and  $R \in \{1, 2, 5, 10, 50, 100\}$  we simulate each process  $N = 5000$  times. The values of  $Q_{g^*}$  and  $\mathbb{E}M_{[-k, k]^d}$  are estimated via the corresponding empirical means denoted by  $\hat{Q}_{g^*}$  and  $\widehat{\mathbb{E}M}_k$  (the corresponding empirical standard deviations are denoted by  $s(\widehat{M}_{g_j})$  and  $s(\widehat{M}_k)$ , respectively). We use a plug-in estimator  $\hat{P}_{R,k}$  for  $P_{R,k}$  that is based on the empirical means of  $\sup_{y \in K} Z(y)^{-1}$  and  $\sup_{y \in K} Z_{[-k, k]^d}(y)^{-1}$ . The results of the simulation study are shown in Table 1.

First, we note that – in accordance to Equation (6.1) –  $Q_{g^*}$  is always smaller than  $\mathbb{E}M_{[-k, k]^d}$ . For instance, for  $R = 1$ , the number of considered shape functions is decreased by 29% ( $k = 2$ ) and 43% ( $k = 3$ ), respectively. Furthermore, we observe that  $P_{R,k}$  seems to be almost constant in  $R$ , namely  $P_{R,2} \approx 0.95$  and  $P_{R,3} \approx 1$  which shows that the approximation of  $Z$  by  $Z_{[-3, 3]}$  is sufficiently good for  $h = 0.1$ . Thus, the behavior of  $Q_{g^*}/\mathbb{E}M_{[-k, k]}$  is basically driven by  $A_{R,k}$  which tends to 1 as  $R \rightarrow \infty$ . For large  $R$ ,  $Q_{g^*}/\mathbb{E}M_{[-k, k]} \approx P_{R,k}$ . This explains the surprising fact that  $\mathbb{E}M_{[-2, 2]} > \mathbb{E}M_{[-3, 3]}$  even though the approximation of  $Z$  by  $Z_{[-2, 2]}$  is less accurate than that by  $Z_{[-3, 3]}$ .

**Table 1.** Results for simulations of  $Z$  (via the normalized spectral representation) compared to  $Z_{[-k,k]}$  with  $k = 2$  (top) and  $k = 3$  (bottom), on  $\{-R, -R + 0.1, \dots, R - 0.1, R\}$  for different  $R$ . For each case,  $A_{R,k}$  and the estimates for  $Q_{g^*}$ ,  $\mathbb{E}M_{[-k,k]}$  and  $P_{R,k}$  as well as the corresponding sample standard deviations are displayed, based on  $N = 5000$  simulations of each process

$R$	$\widehat{\mathbb{E}M_2}$	$\hat{Q}_{g^*}$	$A_{R,2}$	$\hat{P}_{R,2}$	$\widehat{s(M_2)}$	$\widehat{s(M_{g^*})}$
1	4.38	3.12 (-29%)	0.75	0.94	2.84	1.85
2	7.57	5.73 (-24%)	0.81	0.94	4.10	3.06
5	18.83	15.82 (-16%)	0.89	0.95	7.93	6.68
10	40.57	35.63 (-12%)	0.94	0.94	14.03	12.55
50	257.61	239.75 (-7%)	0.99	0.94	62.63	60.96
100	579.11	540.44 (-7%)	0.99	0.94	124.66	117.76

  

$R$	$\widehat{\mathbb{E}M_3}$	$\hat{Q}_{g^*}$	$A_{R,3}$	$\hat{P}_{R,3}$	$\widehat{s(M_3)}$	$\widehat{s(M_{g^*})}$
1	5.46	3.12 (-43%)	0.56	1.00	3.78	1.85
2	8.93	5.73 (-36%)	0.65	0.98	5.13	3.06
5	19.98	15.82 (-21%)	0.78	1.02	8.71	6.68
10	41.16	35.63 (-13%)	0.87	1.00	14.82	12.55
50	247.35	239.75 (-3%)	0.97	1.00	60.14	60.96
100	550.70	540.44 (-2%)	0.98	1.00	114.36	117.76

Next, we perform the simulation for  $d = 2$ ,  $R \in \{1, 2, 5, 10\}$  and  $h = 0.25$ . Each process is simulated  $N = 2500$  times. The results are shown in Table 2. In general, the results are similar to our observations for  $d = 1$ . However, for  $d = 2$  the improvements compared to Schlather’s [26] algorithm are even more distinct. In the case  $R = 1$ , the number of considered spectral functions is decreased by 45% ( $k = 2$ ) and 69% ( $k = 3$ ), respectively.

### 6.2. Comparison to the algorithms proposed in Dieker and Mikosch [9] and Dombry, Engelke and Oesting [10]

Recently, Dieker and Mikosch [9] proposed an exact algorithm for the simulation of Brown–Resnick processes on a finite set. In Dombry, Engelke and Oesting [10], a generalization of this algorithm and a novel exact simulation procedure based on extremal functions have been presented. Denote the number of stochastic processes to be simulated for obtaining an exact simulation on  $K$  via the (generalized version of the) Dieker–Mikosch algorithm and via the extremal functions approach by  $M^{(DM)}$  and  $M^{(EF)}$ , respectively. Then, it can be shown that

$$\mathbb{E}M^{(DM)} = |K| \cdot \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right] \quad \text{and} \quad \mathbb{E}M^{(EF)} = |K| \tag{6.2}$$

(cf. Dombry, Engelke and Oesting [10]). As  $\mathbb{E}[(\inf_{y \in K} Z(y))^{-1}] \geq 1$  for any max-stable process with standard Fréchet margins, we have  $\mathbb{E}M^{(DM)} \geq \mathbb{E}M^{(EF)}$ . Further, the underlying stochastic

**Table 2.** Results for simulations of  $Z$  (via the normalized spectral representation) compared to  $Z_{[-k,k]^2}$  with  $k = 2$  (top) and  $k = 3$  (bottom), on  $\{-R, -R + 0.25, \dots, R - 0.25, R\}^2$  for different  $R$ . For each case,  $A_{R,k}$  and the estimates for  $Q_{g^*}$ ,  $\mathbb{E}M_{[-k,k]^2}$  and  $P_{R,k}$  as well as the corresponding sample standard deviations are displayed, based on  $N = 2500$  simulations of each process

$R$	$\widehat{\mathbb{E}M_2}$	$\hat{Q}_{g^*}$	$A_{R,2}$	$\hat{P}_{R,2}$	$\widehat{s(M_2)}$	$\widehat{s(M_{g^*})}$
1	14.86	8.14 (−45%)	0.56	0.96	9.26	4.65
2	40.17	26.32 (−34%)	0.66	1.00	18.07	10.96
5	189.83	150.89 (−21%)	0.80	0.99	49.32	40.96
10	727.33	636.03 (−13%)	0.88	0.99	146.88	127.36

---

$R$	$\widehat{\mathbb{E}M_3}$	$\hat{Q}_{g^*}$	$A_{R,3}$	$\hat{P}_{R,3}$	$\widehat{s(M_3)}$	$\widehat{s(M_{g^*})}$
1	26.37	8.14 (−69%)	0.32	0.96	16.63	4.65
2	61.07	26.32 (−57%)	0.42	1.03	26.99	10.96
5	247.10	150.89 (−39%)	0.61	1.00	65.75	40.96
10	839.44	636.03 (−24%)	0.75	1.01	168.31	127.36

processes follow mixtures of the same laws. Consequently, between these two approaches, the one via extremal functions is always preferred in terms of the average computational costs of simulation.

In this subsection, we will compare the computational costs of these algorithms to that of simulation via the normalized spectral representation for simulating Brown–Resnick processes. This is what the original Dieker–Mikosch algorithm was designed for. Let  $Z$  be a Brown–Resnick process on  $K \subset \mathbb{R}^d$  associated to a variogram  $\gamma$  (Kablichko, Schlather and de Haan [17]), that is, a max-stable process with representation (1.1) with  $H$  being the probability measure of the stochastic process

$$W(y) = \exp(G(y) - \text{Var}(G(y))/2), \quad y \in \mathbb{R}^d.$$

Here  $G$  is a centered Gaussian process with stationary increments and variogram  $\gamma(h) = \mathbb{E}(W(y+h) - W(y))^2$ . In the following, we consider a Brown–Resnick process associated to the variogram  $\gamma(h) = \|h\|$  on the rectangle  $[0, 1]^2$ . We simulate  $N = 500$  realizations of the process on the grids  $\{0, 0.05, \dots, 0.95\}^2$  (400 points) and  $\{0, 0.01, \dots, 0.99\}^2$  (10 000 points), respectively, via each of the three algorithms on a 2.90 GHz processor.

We start with comparing the average numbers of spectral processes to be simulated, i.e.  $Q_{g^*} = \mathbb{E}M_{g^*}$ ,  $\mathbb{E}M^{(DM)}$  and  $\mathbb{E}M^{(EF)}$ . Since the terms  $c$  in (4.6) and  $\mathbb{E}[(\inf_{y \in K} Z(y))^{-1}]$  in (4.6) and (6.2) often cannot be calculated explicitly, the above expectations need to be estimated via simulations. Furthermore, in the Dieker–Mikosch algorithm and the extremal functions approach, we can directly simulate the underlying spectral functions as a single log-Gaussian process, whereas we have to use the Algorithm 3.13 to simulate the normalized spectral functions. For each process  $F^{(m)}$ , we simulate  $n_{MCMC}$  log-Gaussian processes. Consequently, a fair comparison across the three algorithms should be based on the average computational costs accounting for the number

of simulated log-Gaussian processes. The costs for the simulation via the normalized spectral representation, the Dieker–Mikosch algorithm and the extremal functions approach are thus

$$C^{(NSR)} = n_{MCMC} \cdot \mathbb{E}M_{g^*} = n_{MCMC} \cdot c \cdot \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right],$$

$$C^{(DM)} = \mathbb{E}M^{(DM)} = |K| \cdot \mathbb{E} \left[ \left( \inf_{y \in K} Z(y) \right)^{-1} \right]$$

and  $C^{(EF)} = \mathbb{E}M^{(EF)} = |K|.$

We use  $n_{MCMC} = 100, 500$  and  $1000$  in our simulation. The normalization constant  $c$  is finally estimated by the average of the maxima of the  $n_{MCMC}$  log-Gaussian processes simulated in each Markov chain.

Finally, we consider the exactness of the simulations. Notice that the Dieker–Mikosch algorithm and the extremal functions approach yield exact realizations of the Brown–Resnick process. This is true for the normalized spectral representation approach only if the Markov chains have converged. Therefore, we need to evaluate the quality of the simulations. As a measure of exactness, we first calculate the Kolmogorov–Smirnov distance between the standard Fréchet distribution  $\Phi_1$  and the empirical cumulative distribution function  $F_{(500)}^y$  obtained from the  $N = 500$  realizations  $Z_1(y), \dots, Z_N(y)$  at each location  $y$ . Then we calculate the average of these distances across all  $y \in K$  i.e.

$$\overline{d_{KS}} = \frac{1}{|K|} \sum_{y \in K} \|F_{(500)}^y - \Phi_1\|_\infty.$$

Further, we consider the extremal coefficient (Smith [29], Schlather and Tawn [28]),  $\theta(y_1, y_2)$  defined by

$$\mathbb{P}(Z(y_1) \leq z, Z(y_2) \leq z) = \mathbb{P}(Z(y_1) \leq z)^{\theta(y_1, y_2)}, \quad z > 0,$$

as a measure of extremal dependence between  $Z(y_1)$  and  $Z(y_2)$ . By definition,  $\theta(y_1, y_2) \in [1, 2]$  with  $\theta(y_1, y_2) = 1$  if  $Z(y_1) = Z(y_2)$  a.s. and  $\theta(y_1, y_2) = 2$  if  $Z(y_1)$  and  $Z(y_2)$  are independent. For a Brown-Resnick process  $Z$ , we have  $\theta(y_1, y_2) = 2\Phi(\sqrt{\gamma(h)}/2)$  where  $\Phi$  denotes the standard normal distribution function. We estimate the extremal coefficients making use of the relation

$$\theta(y_1, y_2) = \frac{1 + 2v^F(y_1, y_2)}{1 - 2v^F(y_1, y_2)}, \quad y_1, y_2 \in K, \tag{6.3}$$

between  $\theta(y_1, y_2)$  and the  $F$ -madogram (Cooley, Naveau and Poncet [5])

$$v^F(y_1, y_2) = \frac{1}{2} \mathbb{E} |F^{y_1}(Z(y_1)) - F^{y_1}(Z(y_1))|, \quad y_1, y_2 \in K,$$

with  $F^{(y)}$  being the cumulative distribution function of  $Z(y)$ ,  $y \in K$ . The  $F$ -madogram can be estimated non-parametrically by

$$\widehat{v^F}(y_1, y_2) = \frac{1}{2N(N+1)} \sum_{i=1}^N |R_i(y_1) - R_i(y_2)|, \quad y_1, y_2 \in K,$$

where  $R_i(y)$  denotes the rank of  $Z_i(y)$  (Ribatet [25]). Plugging the estimator  $\widehat{v^F}(y_1, y_2)$  into relation (6.3), we obtain an estimator  $\hat{\theta}(y_1, y_2)$  and  $\theta(y_1, y_2)$ . As a measure of quality of the simulations, we calculate the root-mean-square error between the estimated and the theoretical extremal coefficients, that is,

$$\overline{d_{EC}} = \left[ \frac{1}{|\mathcal{Y}|} \sum_{(y_1, y_2) \in \mathcal{Y}} (\hat{\theta}(y_1, y_2) - 2\Phi(\sqrt{\|y_1 - y_2\|/2}))^2 \right]^{1/2}$$

for some finite set  $\mathcal{Y} \subset K \times K \setminus \{(y, y) : y \in K\}$ .

Table 3 reports the average number of simulated spectral functions  $\overline{M}$ , the average computational costs both in terms of the number of log-Gaussian processes simulated  $\overline{C}$  and in terms of the CPU time for a single simulation (in seconds), the average Kolmogorov–Smirnov distance  $\overline{d_{KS}}$  and the root-mean-square error  $\overline{d_{EC}}$  based on 79 800 pairs. We note that the average CPU time  $\overline{t}$  is effectively proportional to the average number  $\overline{C}$  of simulated log-Gaussian processes.

**Table 3.** Results for simulations of Brown–Resnick processes on the grids  $\{0, 0.05, \dots, 0.95\}^2$  (top) and  $\{0, 0.01, \dots, 0.99\}^2$  (bottom). For simulation via the normalized spectral representation with  $n_{MCMC} = 100, 500$  and  $1000$  ( $NSR_{100}, NSR_{500}$  and  $NSR_{1000}$ ), the Dieker–Mikosch algorithm (DM) and simulation via extremal functions (EF), the average number  $\overline{M}$  of simulated spectral functions, the average computational costs  $\overline{C}$  (with sample standard deviation  $s(C)$ ), the average CPU time  $\overline{t}$  (with sample standard deviation  $s(t)$ ), the average Kolmogorov–Smirnov distance  $\overline{d_{KS}}$  and the root-mean-square error  $\overline{d_{EC}}$  are displayed

	$\overline{M}$	$\overline{C}$	$\overline{t}$	$\overline{d_{KS}}$	$\overline{d_{EC}}$	$s(C)$	$s(t)$
$NSR_{100}$	7	704	0.13	0.030	0.011	492	0.07
$NSR_{500}$	7	3677	0.56	0.050	0.014	2462	0.36
$NSR_{1000}$	7	7244	1.07	0.047	0.020	4971	0.71
$DM$	932	932	0.17	0.036	0.020	713	0.11
$EF$	390	390	0.09	0.036	0.013	355	0.05

  

	$\overline{M}$	$\overline{C}$	$\overline{t}$	$\overline{d_{KS}}$	$\overline{d_{EC}}$	$s(C)$	$s(t)$
$NSR_{100}$	9	896	108	0.041	0.012	666	81
$NSR_{500}$	9	4357	522	0.028	0.013	2888	346
$NSR_{1000}$	9	9194	1095	0.040	0.011	6131	737
$DM$	26 270	26 270	3126	0.027	0.012	21 199	2542
$EF$	9361	9361	1116	0.054	0.015	8448	1014

For the grid  $\{0, 0.05, \dots, 0.95\}^2$ , we observe  $\bar{t} \approx 2 \cdot 10^{-4} \bar{C}$  while we have  $\bar{t} \approx 0.12 \bar{C}$  for the grid  $\{0, 0.01, \dots, 0.99\}^2$ . Smaller deviations from this proportion in case of the first grid are mainly due to some preparatory computations which have a larger effect in case of small  $\bar{C}$ . Thus, we can conclude that  $\bar{C}$  is an appropriate measure of the computational costs of the simulations.

Comparing the results for the different algorithms in more detail, we first focus on the exactness of the simulations via the normalized spectral representation. As both the Kolmogorov–Smirnov distances and the root-mean-square errors for the extremal coefficients indicate, the Markov chains converge quite fast. In all the cases, both measures of the quality of the simulations via the normalized spectral representation are comparable to those for the exact simulations via the Dieker–Mikosch [9] algorithm or via extremal functions. Therefore, the algorithms perform with equivalent exactness even though the algorithm based on the normalized spectral representation provides approximations only.

Secondly, the number of simulated normalized spectral functions in our algorithm is much lower than that in the other two algorithms. This difference becomes even more pronounced in the case of a dense grid: While the numbers of simulated processes in the Dieker–Mikosch [9] algorithm and the algorithm via extremal functions grow (at least) linearly in  $|K|$ , that in our algorithm remains bounded and practically stable (increased from 7 to 9), as long as  $K \subset [0, 1]^2$ .

Lastly, even though the computational costs of a single normalized spectral function in our algorithm is  $n_{MCMC}$  times higher than that for the other two algorithms, on a dense grid, the total computational costs is much lower than that via the other two algorithms.

The aforementioned three features make the simulation via the normalized spectral representation very attractive, particularly when the process should be simulated at a large number of locations on a dense grid. Simulation via extremal functions (Dombry, Engelke and Oesting [10]) is preferred only if the process should be simulated at a small or moderate number of locations.

## 7. Summary and discussion

Whilst in the definition of a max-stable process an infinite number of spectral functions is involved, the minimal expected number of spectral functions that are actually needed for a simulation is an open problem. We consider two substitution problems, problems (3.10) and (4.2), and show that the unique normalized spectral representation is a solution in both cases. Consequently, we propose a simulation algorithm based on the normalized spectral representation.

Our simulation result reveals two advantages of the proposed algorithm. First, it improves the algorithm of Schlather [26] for (mixed) moving maxima processes, because in this case the normalized spectral functions can be simulated easily.

Second, it is competitive to other algorithms even if simulation of the underlying normalized spectral functions from the transformed measure  $g^*H$  is not straightforward, such as in the simulation of Brown–Resnick processes, even though simulations are no longer exact in this case. Similar results are also expected for other popular max-stable models, such as the extremal Gaussian (Schlather [26]) and extremal  $t$  processes (Opitz [21]).

Although the problem (4.2) is rather close to the original problem (3.2), it remains unclear whether the normalized spectral representation is also the solution to the original one. It is even not known whether different initial choices of the spectral representation in (1.1) may lead to the same solution via renormalizations  $g$  in (2.2) and whether the solution is unique.

Other representations of max-stable processes may also allow for exact simulations. For example, Wang and Stoev [32] considered a representation based on a finite number of Fréchet variables when the spectral measure  $H$  is discrete. Our approach may be advantageous if the number of Fréchet variables gets large while the domain  $K$  is bounded. The assessment of the relative performance is left to future research.

## Appendix: Proof of Proposition 4.1

**Proof of Proposition 4.1.** First, the last part of the inequality,  $Q_{g_0} \leq Q_{g^*}$ , holds automatically due to the optimality of  $g_0$ . Second, note that

$$\operatorname{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f) \sup_{\tilde{y} \in K} Z^{(m)}(\tilde{y})} \leq \operatorname{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f) Z^{(m)}(y_0)}$$

and thus  $Q_g^{(1)} \leq Q_g^{(2)}(y_0)$  for any  $g$  and any  $y_0 \in K$ . In particular, these inequalities hold for  $g = g^*$ . Therefore, we get that  $Q_{g^*}^{(1)} \leq Q_{g^*}^{(2)}(y_0)$  for any  $y_0 \in K$ .

Third, we show that  $Q_{g^*}^{(1)} = 1$ . On the one hand, we have that

$$\sup_{y \in K} \frac{f(y)}{g^*(f)} = c \quad \text{for } g^*H\text{-a.e. } f \in \mathbb{H}. \tag{A.1}$$

On the other hand, with (2.3), we obtain

$$\sup_{y \in K} Z^{(m)}(y) = \sup_{y \in K} \max_{i \in \mathbb{N}} \frac{1}{\sum_{j=1}^i E_j} \frac{f_i(y)}{g^*(f_i)} = c E_1^{-1} \quad \text{for all } m \in \mathbb{N}. \tag{A.2}$$

Combining equations (A.1) and (A.2) yields

$$\begin{aligned} Q_{g^*}^* &= \mathbb{E} \min \left\{ m \in \mathbb{N} : \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g^*(f)} \leq \sup_{y \in K} Z^{(m)}(y) \sum_{j=1}^{m+1} E_j \right\} \\ &= \mathbb{E} \min \left\{ m \in \mathbb{N} : c \leq c \cdot \sum_{j=1}^{m+1} E_j / E_1 \right\} = 1. \end{aligned}$$

To complete the proof of the proposition, we show that  $Q_g^{(2)}(y_0) \leq Q_g$  for all  $g$ . We first consider the case that  $\operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} = \infty$ . Since  $c < \infty$ , by Proposition 3.8, we get that  $\sup_{y \in K} Z^{(m)}(y) \leq \sup_{y \in K} Z(y) < \infty$  with probability one. Thus, by the definition of  $Q_g^{(1)}$  in (3.11), we get that  $Q_g^{(1)} = \infty$ . Consequently,  $Q_g \geq Q_g^{(1)} = \infty$ . It is thus proved that  $Q_g^{(2)} \leq Q_g$ .

Next, consider the case  $\operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} < \infty$ . Since there exists some countable set  $K_0 \subset K$  such that  $\sup_{y \in K} f(y) = \sup_{y \in K_0} f(y)$  for  $H$ -a.e.  $f \in \mathbb{H}$ , we have that

$$\operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} = \sup_{y \in K_0} \operatorname{esssup}_{f \in \mathbb{H}} \frac{f(y)}{g(f)}.$$

Therefore, for every  $\varepsilon > 0$ , there exists some  $y(\varepsilon) \in K_0$  such that

$$\begin{aligned} \frac{1}{1+\varepsilon} \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(m)}(y(\varepsilon))} &\leq \operatorname{esssup}_{f \in \mathbb{H}} \frac{f(y(\varepsilon))}{g(f)Z^{(m)}(y(\varepsilon))} \\ &\leq \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)Z^{(m)}(y)}. \end{aligned} \tag{A.3}$$

Analogously to the proof of Proposition 3.2, we have that

$$\begin{aligned} &\mathbb{E} \min \left\{ m \in \mathbb{N} : \frac{1}{1+\varepsilon} \operatorname{esssup}_{f \in \mathbb{H}} \frac{\sup_{y \in K} f(y)}{g(f)Z^{(m)}(y(\varepsilon))} \leq \sum_{j=1}^{m+1} E_j \right\} \\ &= 1 + \sum_{m=1}^{\infty} \left[ 1 - \mathbb{E} \left( 1 \wedge \left( \frac{(1+\varepsilon) \cdot f_1(y(\varepsilon))}{g(f_1)} / \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} \right) \right) \right]^m \\ &= \left[ \mathbb{E} \left( 1 \wedge \left( \frac{(1+\varepsilon) \cdot f_1(y(\varepsilon))}{g(f_1)} / \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)} \right) \right) \right]^{-1} \geq \frac{\operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)}}{1+\varepsilon}, \end{aligned}$$

where the last step follows from  $\int_{\mathbb{H}} f(y(\varepsilon))H(df) = 1$ . From (3.15), it is straightforward to verify that, for any  $y_0 \in K$ ,  $Q_g^{(2)}(y_0) = \operatorname{esssup}_{f \in \mathbb{H}} \sup_{y \in K} \frac{f(y)}{g(f)}$ , which is independent from the choice of  $y_0 \in K$ . Hence, by using (A.3) and taking  $\varepsilon \rightarrow 0$ , we obtain  $Q_g^{(2)}(y_0) \leq Q_g$  for any  $g$ , which implies that  $Q_{g^*}^{(2)}(y_0) \leq Q_{g_0}$ , due to the optimality of  $g^*$ .  $\square$

## Acknowledgements

The authors are grateful to Prof. C. Dombry, an associate editor and two referees for valuable suggestions and comments. M. Oesting and M. Schlather have been financially supported by Volkswagen Stiftung within the project ‘Mesoscale Weather Extremes – Theory, Spatial Modeling and Prediction (WEX-MOP)’. The work of M. Oesting has also partly been funded by the ANR project ‘McSim’. Views expressed do not necessarily reflect official positions of De Nederlandsche Bank.

## References

- [1] Blanchet, J. and Davison, A.C. (2011). Spatial modeling of extreme snow depth. *Ann. Appl. Stat.* **5** 1699–1725. [MR2884920](#)
- [2] Buishand, T.A., de Haan, L. and Zhou, C. (2008). On spatial extremes: With application to a rainfall problem. *Ann. Appl. Stat.* **2** 624–642. [MR2524349](#)
- [3] Coles, S.G. (1993). Regional modelling of extreme storms via max-stable processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **55** 797–816. [MR1229882](#)
- [4] Coles, S.G. and Tawn, J.A. (1996). Modelling extremes of the areal rainfall process. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **58** 329–347. [MR1377836](#)



- [5] Cooley, D., Naveau, P. and Poncet, P. (2006). Variograms for spatial max-stable random fields. In *Dependence in Probability and Statistics* 373–390. New York: Springer.
- [6] de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. [MR0757776](#)
- [7] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. New York: Springer. [MR2234156](#)
- [8] de Haan, L. and Lin, T. (2001). On convergence toward an extreme value distribution in  $C[0, 1]$ . *Ann. Probab.* **29** 467–483.
- [9] Dieker, A.B. and Mikosch, T. (2015). Exact simulation of Brown–Resnick random fields at a finite number of locations. *Extremes* **18** 301–314.
- [10] Dombry, C., Engelke, S. and Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317.
- [11] Dombry, C. and Eyi-Minko, F. (2012). Strong mixing properties of max-infinitely divisible random fields. *Stochastic Process. Appl.* **122** 3790–3811. [MR2965926](#)
- [12] Dombry, C. and Eyi-Minko, F. (2013). Regular conditional distributions of continuous max-infinitely divisible random fields. *Electron. J. Probab.* **18** no. 7. [MR3024101](#)
- [13] Dombry, C., Éyi-Minko, F. and Ribatet, M. (2013). Conditional simulation of max-stable processes. *Biometrika* **100** 111–124. [MR3034327](#)
- [14] Giné, E., Hahn, M.G. and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165.
- [15] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109.
- [16] Kabluchko, Z. (2009). Spectral representations of sum- and max-stable processes. *Extremes* **12** 401–424. [MR2562988](#)
- [17] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. [MR2561440](#)
- [18] Oesting, M., Kabluchko, Z. and Schlather, M. (2012). Simulation of Brown–Resnick processes. *Extremes* **15** 89–107. [MR2891311](#)
- [19] Oesting, M. and Schlather, M. (2014). Conditional sampling for max-stable processes with a mixed moving maxima representation.
- [20] Oesting, M., Schlather, M. and Zhou, C. (2013). On the normalized spectral representation of max-stable processes on a compact set. Available at [arXiv:1310.1813v1](#).
- [21] Opitz, T. (2013). Extremal  $t$  processes: Elliptical domain of attraction and a spectral representation. *J. Multivariate Anal.* **122** 409–413. [MR3189331](#)
- [22] Padoan, S.A., Ribatet, M. and Sisson, S.A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105** 263–277.
- [23] R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- [24] Resnick, S.I. and Roy, R. (1991). Random usc functions, max-stable processes and continuous choice. *Ann. Appl. Probab.* 267–292.
- [25] Ribatet, M. (2013). Spatial extremes: Max-stable processes at work. *J. Soc. Fr. Stat.* **154** 156–177.
- [26] Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5** 33–44.
- [27] Schlather, M., Malinowski, A., Oesting, M., Boecker, D., Strokorb, K., Engelke, S., Martini, J., Ballani, F., Moreva, O., Menck, P.J., Gross, S., Ober, U., Berreth, C., Burmeister, K., Manitz, J., Ribeiro, P., Singleton, R., Pfaff, B. and R Core Team (2015). RandomFields: Simulation and analysis of random fields. R package version 3.1.3.
- [28] Schlather, M. and Tawn, J.A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90** 139–156.

- [29] Smith, R.L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript.
- [30] Stoev, S.A. and Taqqu, M.S. (2005). Extremal stochastic integrals: A parallel between max-stable processes and  $\alpha$ -stable processes. *Extremes* **8** 237–266. [MR2324891](#)
- [31] Strokorb, K., Ballani, F. and Schlather, M. (2015). Tail correlation functions of max-stable processes. *Extremes* **18** 241–271.
- [32] Wang, Y. and Stoev, S.A. (2010). On the structure and representations of max-stable processes. *Adv. in Appl. Probab.* **42** 855–877.

*Received March 2014 and revised September 2016*