

# Exact unconditional sampling from max-stable random vectors

Léo Belzile

EPFL-SB-MATH-STAT  
Station 8, Bâtiment MA  
CH-1015, Switzerland,  
leo.belzile@epfl.ch

Originally, the ‘mev’ package was introduced to implement the exact unconditional sampling algorithms in Dombry et al. (2016) to simulate simple max-stable random vectors. The implementation will work efficiently for moderate dimensions and does not currently take improvements for large spatial domains.

## 1 Functions and use

There are two main functions, `rmev` and `rmevspec`. `rmev` samples from simple max-stable processes, meaning it will return an  $n \times d$  matrix of samples, where each of the column has a sample from a unit Frechet distribution. In contrast, `rmevspec` returns sample on the unit simplex from the spectral (or angular) measure. One could use this to test estimation based on spectral densities, or to construct samples from Pareto processes.

The syntax is

```
library(mev)
#Sample of size 1000 from a 5-dimensional logistic model
x <- rmev(n=1000, d=5, param=0.5, model="log")
#Marginal parameters are all standard Frechet, meaning GEV(1,1,1)
apply(x, 2, function(col){ismev::gev.fit(col, show=FALSE)$mle})

##      [,1] [,2] [,3] [,4] [,5]
## [1,] 1.003 0.9724 0.9780 0.9995 1.0216
## [2,] 0.988 1.0059 0.9586 1.0107 1.0145
## [3,] 1.010 1.0538 1.0143 1.0171 0.9739

#Sample from the corresponding spectral density
w <- rmevspec(n=1000, d=5, param=0.5, model="log")
#All rows sum to 1 by construction
head(rowSums(w))

## [1] 1 1 1 1 1 1
```

```
#The marginal mean is 1/d
round(colMeans(w),2)

## [1] 0.20 0.20 0.21 0.20 0.20
```

## 2 Description of the models implemented

The different models implemented are described in Dombry et al. (2016), but some other models can be found and are described here.

1. **logistic distribution** (log) The logistic model of Gumbel (1960) (the Gumbel Archimedean copula)

$$P(\mathbf{X} \leq \mathbf{x}) = \exp \left[ - \left( \sum_{i=1}^n \left( \frac{1}{x_i} \right)^\alpha \right)^{\frac{1}{\alpha}} \right]$$

for  $\alpha > 1$ . By default, rmev will transform an argument in  $(0, 1)$  without warning, to conform with the implementation. The spectral measure density is

$$h_W(\mathbf{w}) = \frac{1}{d} \frac{\Gamma(d - \alpha)}{\Gamma(1 - \alpha)} \alpha^{d-1} \left( \prod_{j=1}^d w_j \right)^{-(\alpha+1)} \left( \sum_{j=1}^d w_j^{-\alpha} \right)^{1/\alpha-d}, \quad \mathbf{w} \in \mathbb{S}_d$$

2. **asymmetric logistic distribution** (alog) This model proposed by Tawn (1990) uses the parametrization of the evd package, merely replacing the algorithm for the generation of logistic variates. The distribution function is

$$P(\mathbf{X} \leq \mathbf{x}) = \exp \left[ - \sum_{b \in \mathbb{B}} \left( \sum_{i \in b} \left( \frac{\theta_{i,b}}{x_i} \right)^{\alpha_b} \right)^{\frac{1}{\alpha_b}} \right]$$

while the spectral density corresponds to

$$h_W(\mathbf{w}) = \frac{1}{d} \frac{\prod_{j=1}^{d-1} (j\alpha_b - 1)}{\prod_{i \in b} w_i} \left( \prod_{i \in b} \frac{\theta_{i,b}}{w_i} \right)^{\alpha_b} \left( \sum_{i \in b} \left( \frac{\theta_{i,b}}{w_i} \right)^{\alpha_b} \right)^{1/\alpha_b-d}$$

The parameters  $\theta_{i,b}$  must be provided in a list and represents the asymmetry parameter. In the  $d$ -variate case, where  $\mathbb{B}$  is the power set of the collection  $\{1, \dots, d\}$ , without the empty set. This corresponds to generating samples  $\mathbf{Z}_b$  from a logistic distribution of dimension  $|b|$  (or Fréchet variates if  $|b| = 1$ ) with parameter  $\alpha_b$  (possibly recycled). Then, each marginal value corresponds to the maximum of the weighted corresponding entry. That is,  $X_i = \max_{b \in \mathbb{B}} \theta_{i,b} Z_{i,b}$  for all  $i = 1, \dots, d$ . The max-mixture is valid provided that  $\sum_{b \in \mathbb{B}} \theta_{i,b} = 1$  for  $i = 1, \dots, d$ . As such, empirical estimates of the spectral measure will almost surely place mass on the inside of the simplex rather than on subfaces.

3. **negative logistic distribution** (`neglog`) The distribution function of the min-stable distribution due to Galambos (1975) is

$$P(\mathbf{X} \leq \mathbf{x}) = \exp \left[ - \sum_{b \in \mathbb{B}} (-1)^{|b|} \left( \sum_{i \in b} x_i^{-\alpha} \right)^{\frac{1}{\alpha}} \right]$$

for  $\alpha_b \leq 0$  and corresponds to the model in Dombry et al. (2016).

$$h_{\mathbf{W}}(\mathbf{w}) = \frac{1}{d} \sum_{b \in \mathbb{B}: c \subset b} (-1)^{|b|} \left( \prod_{j=1}^{d-1} (1 - j\alpha) \right) \left( \prod_{i \in c} w_j \right)^{-(\alpha+1)} \left( \sum_{i \in c} w_i^{-\alpha} \right)^{1/\alpha-d}$$

4. **asymmetric negative logistic distribution** (`aneglog`) The construction of this distribution is exactly the same as for the asymmetric logistic distribution, with distribution function

$$P(\mathbf{X} \leq \mathbf{x}) = \exp \left[ - \sum_{b \in \mathbb{B}} \sum_{c \in b} (-1)^{|c|} \left( \sum_{i \in c} \left( \frac{x_i}{\theta_{i,b}} \right)^{\alpha_b} \right)^{-\frac{1}{\alpha_b}} \right].$$

In particular, it does not correspond to the “negative logistic distribution” given in Joe (1990), which is not a valid distribution function according to Stephenson.

5. **multilogistic distribution** (`bilog`) This multivariate extension of the logistic, proposed by Boldi (2009), places mass on the interior of the simplex. Let  $\mathbf{W} \in \mathbb{S}_d$  be the solution of

$$\frac{W_j}{W_d} = \frac{C_j U_j^{-\alpha_j}}{C_d U_d^{-\alpha_d}}, \quad j = 1, \dots, d$$

where  $C_j = \Gamma(d - \alpha_j) / \Gamma(1 - \alpha_j)$  for  $j = 1, \dots, d$  and  $\mathbf{U} \in \mathbb{S}_d$  follows a  $d$ -mixture of Dirichlet with the  $j$ th component being  $\mathcal{D}(\mathbf{1} - \delta_j \alpha_j)$ , so that the mixture has density function

$$h_{\mathbf{U}}(\mathbf{u}) = \frac{1}{d} \sum_{j=1}^d \frac{\Gamma(d - \alpha_j)}{\Gamma(1 - \alpha_j)} u_j^{-\alpha_j}$$

for  $0 < \alpha_j < 1, j = 1, \dots, d$ . The spectral density is given by

$$h_{\mathbf{W}}(\mathbf{w}) = \frac{1}{d} \left( \sum_{j=1}^d \alpha_j u_j \right)^{-1} \left( \prod_{j=1}^d \alpha_j u_d \right) \left( \sum_{j=1}^d \frac{\Gamma(d - \alpha_j)}{\Gamma(1 - \alpha_j)} u_j^{-\alpha_j} \right) \prod_{j=1}^d w_j^{-1}$$

for  $\alpha_j \in (0, 1)$

6. **Coles and Tawn Dirichlet distribution** (`ct`) The Dirichlet model of Coles & Tawn (1991)

$$h_{\mathbf{W}}(\mathbf{w}) = \frac{1}{d} \frac{\Gamma(1 + \sum_{j=1}^d \alpha_j)}{\prod_{j=1}^d \alpha_j w_j} \left( \sum_{j=1}^d \alpha_j w_j \right)^{-(d+1)} \prod_{j=1}^d \alpha_j \prod_{j=1}^d \left( \frac{\alpha_j w_j}{\sum_{k=1}^d \alpha_k w_k} \right)^{\alpha_j - 1}$$

for  $\alpha_j > 0$ .

7. **scaled extremal Dirichlet** (`dir`) The angular density of the scaled extremal Dirichlet model with parameters  $\rho > 0$  and  $\boldsymbol{\alpha} \in \mathbb{R}_+^d$  is given, for all  $\boldsymbol{w} \in \mathbb{S}_d$ , by

$$h_W(\boldsymbol{w}) = \frac{\Gamma(\bar{\alpha} + \rho)}{d\rho^{d-1} \prod_{i=1}^d \Gamma(\alpha_i)} \langle \{c(\boldsymbol{\alpha}, \rho)\}^{1/\rho}, \boldsymbol{w}^{1/\rho} \rangle^{-\rho - \bar{\alpha}} \prod_{i=1}^d \{c(\alpha_i, \rho)\}^{\alpha_i/\rho} w_i^{\alpha_i/\rho - 1}.$$

where  $c(\boldsymbol{\alpha}, \rho)$  is the  $d$ -vector with entries  $\Gamma(\alpha_i + \rho)/\Gamma(\alpha_i)$  for  $i = 1, \dots, d$ .

8. **scaled negative extremal Dirichlet** (`negdir`) The angular density of the scaled negative extremal Dirichlet model with parameters  $0 < \rho < \min(\alpha_1, \dots, \alpha_d)$  and  $\boldsymbol{\alpha}$  is given, for all  $\boldsymbol{w} \in \mathbb{R}_+^d$ , by

$$h_W(\boldsymbol{w}) = \frac{\Gamma(\bar{\alpha} - \rho)}{d\rho^{d-1} \prod_{i=1}^d \Gamma(\alpha_i)} \langle \{c(\boldsymbol{\alpha}, -\rho)\}^{-1/\rho}, \boldsymbol{w}^{-1/\rho} \rangle^{\rho - \bar{\alpha}} \prod_{i=1}^d \{c(\alpha_i, -\rho)\}^{-\alpha_i/\rho} w_i^{-\alpha_i/\rho - 1}.$$

9. **Hüsler–Reiss** (`hr`), due to Hüsler & Reiss (1989). It is a special case of the Brown–Resnick process. While Engelke et al. (2015) state that Hüsler–Reiss variates can be sampled following the same scheme, the spatial analog is conditioned on a particular site ( $s_0$ ), which complicates the comparisons with the other methods.

Let  $I_{-j} = \{1, \dots, d\} \setminus \{j\}$  and  $\lambda_{ij}^2 \geq 0$  be entries of a strictly conditionally negative definite matrix  $\boldsymbol{\Lambda}$ , for which  $\lambda_{ij}^2 = \lambda_{ji}^2$ . Then, following Nikoloulopoulos et al. (2009) (Remark 2.5) and Huser & Davison (2013), we can write the distribution function as

$$\mathbb{P}(\boldsymbol{X} \leq \boldsymbol{x}) = \exp \left[ - \sum_{j=1}^d \frac{1}{x_j} \Phi_{d-1, \boldsymbol{\Sigma}_{-j}} \left( \lambda_{ij} - \frac{1}{2\lambda_{ij}} \log \left( \frac{x_j}{x_i} \right), i \in I_{-j} \right) \right].$$

where the partial correlation matrix  $\boldsymbol{\Sigma}_{-j}$  has elements

$$\varrho_{i,k;j} = \frac{\lambda_{ij}^2 + \lambda_{kj}^2 - \lambda_{ik}^2}{2\lambda_{ij}\lambda_{kj}}$$

and  $\lambda_{ii} = 0$  for all  $i \in I_{-j}$  so that the diagonal entries  $\varrho_{i,ij} = 1$ .<sup>1</sup>

The `evd` package implementation has a bivariate implementation of the Hüsler–Reiss distribution with dependence parameter  $r$ , with  $r_{ik} = 1/\lambda_{ik}$  or  $2/r = \sqrt{2\gamma(\boldsymbol{h})}$  for  $\boldsymbol{h} = \|\boldsymbol{s}_i - \boldsymbol{s}_j\|$  for the Brown–Resnick model. In this setting, it is particularly easy since the only requirement is non-negativity of the parameter. For inference in dimension  $d > 2$ , one needs to impose the constraint  $\boldsymbol{\Lambda} = \{\lambda_{ij}^2\}_{i,j=1}^d \in \mathcal{D}$  (cf. Engelke et al. (2015), p.3), where

$$\mathcal{D} = \left\{ \boldsymbol{A} \in [0, \infty)^{d \times d} : \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} < 0, \forall \boldsymbol{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\} \right. \\ \left. \text{with } \sum_{i=1}^d x_i = 0, a_{ij} = a_{ji}, a_{ii} = 0 \forall i, j \in \{1, \dots, d\} \right\}.$$

<sup>1</sup>Engelke et al. (2015) uses the covariance matrix with entries are  $\varsigma = 2(\lambda_{ij}^2 + \lambda_{kj}^2 - \lambda_{ik}^2)$ , so the resulting expression is evaluated at  $2\lambda_{ij}^2 - \log \left( \frac{x_j}{x_i} \right)$  instead. We recover the same expression by standardizing, since this amounts to division by the standard deviations  $2\lambda_{ij}$

An avenue to automatically satisfy these requirements is to optimize over a symmetric positive definite matrix parameter  $\Sigma = \mathbf{L}^\top \mathbf{L}$ , where  $\mathbf{L}$  is an upper triangular matrix whose diagonal elements are on the log-scale to ensure uniqueness of the Cholesky factorization; see ?. By taking

$$\Lambda(\Sigma) = \begin{pmatrix} 0 & \text{diag}(\Sigma)^\top \\ \text{diag}(\Sigma) & \mathbf{1} \text{diag}(\Sigma)^\top + \text{diag}(\Sigma) \mathbf{1}^\top - 2\Sigma \end{pmatrix}$$

one can perform unconstrained optimization for the non-zero elements of  $\mathbf{L}$  which are in one-to-one correspondence with those of  $\Lambda$ .

It easily follows that generating  $\mathbf{Z}$  from a  $d - 1$  dimensional log-Gaussian distribution with covariance  $\text{Cov}(Z_i, Z_k) = 2(\lambda_{ij}^2 + \lambda_{kj}^2 - \lambda_{ik}^2)$  for  $i, k \in I_{-j}$  with mean vector  $-2\lambda_{\bullet j}^2$  gives the finite dimensional analog of the Brown–Resnick process in the mixture representation of Dombry et al. (2016).

The `rmev` function checks conditional negative definiteness of the matrix. The easiest way to do so negative definiteness of  $\Lambda$  with real entries is to form  $\tilde{\Lambda} = \mathbf{P}\Lambda\mathbf{P}^\top$ , where  $\mathbf{P}$  is an  $d \times d$  matrix with ones on the diagonal,  $-1$  on the  $(i, i + 1)$  entries for  $i = 1, \dots, d - 1$  and zeros elsewhere. If the matrix  $\Lambda \in \mathcal{D}$ , then the eigenvalues of the leading  $(d - 1) \times (d - 1)$  submatrix of  $\tilde{\Lambda}$  will all be negative.

10. **Brown–Resnick** (`br`) The Brown–Resnick process is the extension of the Hüsler–Reiss distribution, and is a max-stable process associated with the log-Gaussian distribution.

It is often in the spatial setting conditioned on a location (typically the origin). Users can provide a variogram function or the resulting covariance function. See Engelke et al. (2015) or Dombry et al. (2016) for more information.

11. **Extremal Student** (`extstud`) of Nikoloulopoulos et al. (2009), eq. 2.8, with unit Fréchet margins is

$$\mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \exp \left[ - \sum_{j=1}^d \frac{1}{x_j} T_{d-1, \nu+1, \mathbf{R}_{-j}} \left( \sqrt{\frac{\nu+1}{1-\rho_{ij}^2}} \left[ \left( \frac{x_i}{x_j} \right)^{1/\nu} - \rho_{ij} \right], i \in I_{-j} \right) \right],$$

where  $T_{d-1}$  is the distribution function of the  $d - 1$  dimensional Student- $t$  distribution and the partial correlation matrix  $\mathbf{R}_{-j}$  has diagonal entry

$$r_{i,ij} = 1, \quad r_{i,kj} = \frac{\rho_{ik} - \rho_{ij}\rho_{kj}}{\sqrt{1-\rho_{ij}^2}\sqrt{1-\rho_{kj}^2}}$$

for  $i \neq k, i, k \in I_{-j}$ .

The user must provide a valid correlation matrix (the function checks for diagonal elements), which can be obtained from a variogram.

12. **Dirichlet mixture** (`dirmix`) proposed by Boldi & Davison (2007), see Dombry et al. (2016) for details on the mixture. The spectral density of the model is

$$h_{\mathbf{W}}(\mathbf{w}) = \sum_{k=1}^m \pi_k \frac{\Gamma(\alpha_{1k} + \dots + \alpha_{dk})}{\prod_{i=1}^d \Gamma(\alpha_{ik})} \left( 1 - \sum_{i=1}^{d-1} w_i \right)^{\alpha_{dk}-1} \prod_{i=1}^{d-1} w_i^{\alpha_{ik}-1}$$

The argument param is thus a  $d \times m$  matrix of coefficients, while the argument for the  $m$ -vector weights gives the relative contribution of each Dirichlet mixture component.

13. **Smith model** (`smith`), from the unpublished report of Smith (1990). It corresponds to a moving maximum process on a domain  $\mathbb{X}$ . The de Haan representation of the process is

$$Z(x) = \max_{i \in \mathbb{N}} \zeta_i h(x - \eta_i), \quad \eta_i \in \mathbb{X}$$

where  $\{\zeta_i, \eta_i\}_{i \in \mathbb{N}}$  is a Poisson point process on  $\mathbb{R}_+ \times \mathbb{X}$  with intensity measure  $\zeta^{-2} d\zeta d\eta$  and  $h$  is the density of the multivariate Gaussian distribution. Other  $h$  could be used in principle, but are not implemented.

## References

- Boldi, M.-O. (2009). A note on the representation of parametric models for multivariate extremes. *Extremes*, 12(3), 211–218.
- Boldi, M.-O. & Davison, A. C. (2007). A mixture model for multivariate extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2), 217–229.
- Coles, S. G. & Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2), 377–392.
- Dombry, C., Engelke, S., & Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika*, 103(2), 303–317.
- Engelke, S., Malinowski, A., Kabluchko, Z., & Schlather, M. (2015). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(1), 239–265.
- Galambos, J. (1975). Order statistics of samples from multivariate distributions. *J. Amer. Statist. Assoc.*, 70(351, part 1), 674–680.
- Gumbel, É. J. (1960). Distributions des valeurs extrêmes en plusieurs dimensions. *Publ. Inst. Statist. Univ. Paris*, 9, 171–173.
- Huser, R. & Davison, A. C. (2013). Composite likelihood estimation for the brown–resnick process. *Biometrika*, 100(2), 511–518.
- Hüsler, J. & Reiss, R.-D. (1989). Maxima of normal random vectors: between independence and complete dependence. *Statist. Probab. Lett.*, 7(4), 283–286.
- Joe, H. (1990). Multivariate concordance. *Journal of Multivariate Analysis*, 35(1), 12–30.
- Nikoloulopoulos, A. K., Joe, H., & Li, H. (2009). Extreme value properties of multivariate t copulas. *Extremes*, 12(2), 129–148.
- Smith, R. L. (1990). Max-stable processes and spatial extremes.
- Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika*, 77(2), 245–253.