



Extreme-value limit of the convolution of exponential and multivariate normal distributions: Link to the Hüsler–Reiß distribution

Pavel Krupskii^{a,*}, Harry Joe^b, David Lee^b, Marc G. Genton^a

^a CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia

^b Department of Statistics, University of British Columbia, 2207 Main Mall, Vancouver, BC, Canada V6T 1Z4



ARTICLE INFO

Article history:

Received 14 March 2017

Available online 2 November 2017

AMS subject classifications:

62E20

62H99

Keywords:

Copula

Extreme-value limit

Parametric bootstrap

Parsimonious dependence

ABSTRACT

The multivariate Hüsler–Reiß copula is obtained as a direct extreme-value limit from the convolution of a multivariate normal random vector and an exponential random variable multiplied by a vector of constants. It is shown how the set of Hüsler–Reiß parameters can be mapped to the parameters of this convolution model. Assuming there are no singular components in the Hüsler–Reiß copula, the convolution model leads to exact and approximate simulation methods. An application of simulation is to check if the Hüsler–Reiß copula with different parsimonious dependence structures provides adequate fit to some data consisting of multivariate extremes.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

It is well known (see, e.g., [30]) that scaled vectors of component-wise maxima from the multivariate normal distributions are asymptotically independent, i.e., the limit in distribution is a product of univariate Gumbel or extreme-value distributions. Hüsler and Reiß [16] obtained a limiting non-trivial extreme-value distribution from scaled vectors of component-wise maxima from the multivariate normal by letting the correlation parameters increase to 1 at appropriate rates as the sample size $n \rightarrow \infty$. Nikoloulopoulos et al. [27] derived the multivariate extreme-value t copula (t -EV) from the extreme-value limit of the multivariate t distribution with degree of freedom parameter $\nu > 0$ and correlation matrix $\mathbf{R} = (\rho_{ij})$. By letting $\rho_{ij} \rightarrow 1$ and $\nu \rightarrow \infty$ at appropriate rates, the Hüsler–Reiß (HR) distribution was obtained in a simpler form.

Both of these limits are not direct extreme-value limits so that their constructions cannot be used to simulate from the Hüsler–Reiß distribution by approximation with the component-wise maxima of N independent and identically distributed (i.i.d.) random vectors for large N . It is however useful to be able to simulate, at least approximately, from the HR distribution. One example of use is to provide inferences of a fitted model by means of a parametric bootstrap.

In this paper, we show how the HR distribution can be obtained as a direct extreme-value limit from the convolution of a multivariate normal vector and an exponential random variable multiplied by a positive vector of constants. The convolution model is studied for spatial data with non-Gaussian dependence and tails in [21].

The derivation of the extreme-value limit of the convolution model is given in Section 2. In Sections 3 and 4, we study how the parameters of the HR distribution can be mapped to the parameters of the convolution model, starting with the trivariate case. Also, we derive inequalities for the parameters of the HR distribution based on partial correlations. In Section 5, we

* Corresponding author.

E-mail address: krupskii@yandex.ru (P. Krupskii).

give details on simulating random variates from the HR distribution and show an example where a parametric bootstrap is used for assessing the adequacy of fit of some fitted HR distributions with parsimonious dependence structures. Section 6 concludes with some discussion.

2. Convolution model and Hüsler–Rei extreme-value limit

Let $d \in \mathbb{N}$ be the multivariate dimension and let $\mathbf{Z} = (Z_1, \dots, Z_d)^\top$ be a multivariate normal random vector with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{R} = (\rho_{ij})$ with $i, j \in \{1, \dots, d\}$ which is assumed to be a correlation matrix. Let \mathcal{E} be a random variable with exponential distribution with mean 1. Let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^\top$ be a vector of positive constants. The convolution model is

$$\mathbf{W} = \mathbf{Z} + \boldsymbol{\alpha}\mathcal{E}. \tag{1}$$

In [21], it is shown that the multivariate distribution from this model has interesting upper tail dependence properties; the results were applied to the case when \mathbf{R} is a spatial autocorrelation matrix but the tail properties are valid for arbitrary \mathbf{R} .

In the derivations below, vectors are boldfaced and the product or ratio of two vectors implies component-wise product or division, respectively. If \mathbf{X} is a continuous random vector, the notation for its cumulative distribution function and copula are $F_{\mathbf{X}}$ and $C_{\mathbf{X}}$, respectively. Similarly, if Y is a random variable, its cumulative distribution function is denoted as F_Y unless otherwise specified.

If \mathbf{a} is a d -dimensional (random) vector, we use the notation $\mathbf{a}_{-j} = (a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_d)^\top$. The symbols ϕ and Φ are the standard normal density and cumulative distribution function, respectively; and $\phi_d(\cdot; \mathbf{R})$ and $\Phi_d(\cdot; \mathbf{R})$ are the multivariate counterparts when the mean vector is $\mathbf{0}$. For $d = 2$, we also use the notation $\Phi_2(\cdot; \rho)$, where ρ is a correlation. If \mathbf{R} is a $d \times d$ correlation matrix with $\text{diag}(\mathbf{R}) = \mathbf{1}_d$, then for each $j \in \{1, \dots, d\}$, we let $\mathbf{R}^{(-j)}$ denote the $(d - 1) \times (d - 1)$ matrix obtained from \mathbf{R} by deleting the j th row and column, and we let $\boldsymbol{\rho}^{(j)}$ denote the j th column of \mathbf{R} . Then $\boldsymbol{\rho}_{-j}^{(j)}$ is $\boldsymbol{\rho}^{(j)}$ without the j th element. If $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{R})$, then the conditional distribution of $(\mathbf{Z}_{-j} \mid Z_j = z_j)$ is $\mathcal{N}_{d-1}(\boldsymbol{\rho}_{-j}^{(j)}z_j, \mathbf{R}^{(-j)} - \boldsymbol{\rho}_{-j}^{(j)}(\boldsymbol{\rho}_{-j}^{(j)})^\top)$.

For the limiting extreme-value copula, we use the approach with upper tail dependence functions in [18,27]. We obtain the copula $C_{\mathbf{W}}$ of \mathbf{W} , and show that as $u \downarrow 0$, it has the asymptotic form given, for all $\mathbf{x} \in (0, \infty)^d$, by

$$C_{\mathbf{W}}(1 - ux_1, \dots, 1 - ux_d) \sim 1 - u\ell(\mathbf{x}), \tag{2}$$

where the asymptotic equivalence $a(x) \sim b(x)$ as $x \rightarrow m$ means that $\lim_{x \rightarrow m} a(x)/b(x) = 1$. Then, as $n \rightarrow \infty$,

$$C_{\mathbf{W}}^n(u_1^{1/n}, \dots, u_d^{1/n}) \rightarrow \exp\{-\ell(-\ln u_1, \dots, -\ln u_d)\}$$

and $C_{EV}(\mathbf{u}) = \exp\{-\ell(-\ln u_1, \dots, -\ln u_d)\}$ is an extreme-value limiting copula, with ℓ being the stable tail dependence function [6,10,14].

For Model (1), we will derive its copula and then ℓ , and show that ℓ matches the stable tail dependence function of the HR distribution in the form given in [27].

Theorem 1. Model (1), with correlation matrix $\mathbf{R} = (\rho_{ij})$ with $i, j \in \{1, \dots, d\}$, and positive scale vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)^\top$, multiplying \mathcal{E} , has upper tail dependence and its stable tail dependence function is

$$\ell(\mathbf{x}) = \sum_{i=1}^d x_i \Phi_{d-1}\{\eta_{ij}/2 + \eta_{ij}^{-1} \ln(x_i/x_j), j \neq i; \boldsymbol{\Gamma}_i\}, \tag{3}$$

where, for $i \neq j$,

$$\eta_{ij} = \frac{1}{\alpha_i \alpha_j} (\alpha_i^2 + \alpha_j^2 - 2\rho_{ij}\alpha_i\alpha_j)^{1/2}, \tag{4}$$

$\eta_{ii} = 0$ for $i \in \{1, \dots, d\}$, and the correlation matrix in the i th term of (3) is

$$\boldsymbol{\Gamma}_i = \left(\frac{\eta_{ij}^2 + \eta_{ik}^2 - \eta_{jk}^2}{2\eta_{ij}\eta_{ik}} \right), \quad j, k \in \{1, \dots, i - 1, i + 1, \dots, d\}. \tag{5}$$

Remark 1. With $\alpha_1 = \dots = \alpha_d = \alpha > 0$, $e^{\mathbf{W}} = (e^{W_1}, \dots, e^{W_d})^\top$ is the product of a log-normal random vector and an independent Pareto random variable, and therefore $e^{\mathbf{W}}$ belongs to the maximum domain of attraction of the Hüsler–Rei distribution. Genton et al. [12] used the Poisson process approach of constructing a max-stable process to prove the result of Theorem 1 for the Smith model and [15] extended this result to the Brown–Resnick process. In both cases, \mathbf{R} has a parsimonious form where the correlations depend on the distances between the corresponding spatial locations. At the same time, the result of Theorem 1 uses no additional assumptions on \mathbf{R} and is valid for an arbitrary positive definite correlation matrix \mathbf{R} and vector $\boldsymbol{\alpha}$ with arbitrary positive entries.

Proof of Theorem 1. For the multivariate distribution $F_{\mathbf{W}}$ of \mathbf{W} , with univariate margins F_1, \dots, F_d (so that $W_i \sim F_i$), the unique copula as derived from Sklar's theorem is given, for all $\mathbf{u} \in [0, 1]^d$, by

$$C_{\mathbf{W}}(\mathbf{u}) = F_{\mathbf{W}}\{F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\}, \quad (6)$$

where

$$F_{\mathbf{W}}(\mathbf{w}) = \int_0^\infty \Phi_d(\mathbf{w} - \alpha s; \mathbf{R}) e^{-s} ds. \quad (7)$$

For each $i \in \{1, \dots, d\}$, the i th univariate marginal distribution is

$$\begin{aligned} F_i(w) &= \int_{-\infty}^\infty \phi(z) F_{\mathcal{E}}\{(w-z)/\alpha_i\} dz = \int_{-\infty}^w \phi(z) \{1 - e^{-(w-z)/\alpha_i}\} dz \\ &= \Phi(w) - e^{-w/\alpha_i} e^{-1/(2\alpha_i^2)} \int_{-\infty}^w \phi(z - \alpha_i^{-1}) dz \\ &= \Phi(w) - \exp\{-w/\alpha_i + 1/(2\alpha_i^2)\} \Phi(w - \alpha_i^{-1}). \end{aligned} \quad (8)$$

To apply (2), we want to get $F_i^{-1}(1 - p_i)$ as $p_i \downarrow 0$. Let $w_i = F_i^{-1}(1 - p_i)$, then with $p_i \downarrow 0$ and $w_i \rightarrow \infty$,

$$\begin{aligned} 1 - p_i &= \Phi(w_i) - \exp\{-w_i/\alpha_i + 1/(2\alpha_i^2)\} \Phi(w_i - \alpha_i^{-1}) \\ &= 1 - \phi(w_i)/w_i - \exp\{-w_i/\alpha_i + 1/(2\alpha_i^2)\} \{1 - \phi(w_i - \alpha_i^{-1})/w_i - \alpha_i^{-1}\} + o\{\phi(w_i)/w_i\} \\ &= 1 - \exp\{-w_i/\alpha_i + 1/(2\alpha_i^2)\} + O\{w_i^{-1} \exp(-w_i^2/2)\}. \end{aligned}$$

Solving for w_i in the above leads to $w_i = 1/(2\alpha_i) - \alpha_i \ln p_i + o(1)$ as $p_i \downarrow 0$. With $p_i = ux_i$, let $w_i^* = 1/(2\alpha_i) - \alpha_i \ln x_i - \alpha_i \ln u$, so that as $u \downarrow 0$,

$$\begin{aligned} F_i(w_i^*) &= 1 - ux_i + O[(w_i^*)^{-1} \exp\{-(w_i^*)^2/2\}] \\ &= 1 - ux_i + O[(-\ln u)^{-1} \exp\{-(-\alpha_i \ln u)^2/2\}] = 1 - ux_i + o(u). \end{aligned}$$

Now, substitute $u_i = 1 - ux_i$ in (6)–(7) and take a tail expansion. Using component-wise operations on vectors,

$$\begin{aligned} C_{\mathbf{W}}(\mathbf{1}_d - \mathbf{u}\mathbf{x}) &= F_{\mathbf{W}}\{F_1^{-1}(1 - ux_1), \dots, F_d^{-1}(1 - ux_d)\} = F_{\mathbf{W}}(w_1^*, \dots, w_d^*) + o(u) \\ &= \int_0^\infty \Phi_d\{1/(2\alpha) - \alpha \ln \mathbf{x} - \alpha \ln u - \alpha s; \mathbf{R}\} e^{-s} ds + o(u) \\ &= u \int_{\ln u}^\infty \Phi_d\{1/(2\alpha) - \alpha \ln \mathbf{x} - \alpha t; \mathbf{R}\} e^{-t} dt + o(u). \end{aligned} \quad (9)$$

Next we make use of the fact that, for all $j \in \{1, \dots, d\}$,

$$\frac{\partial F_{\mathbf{Y}}(\mathbf{y})}{\partial y_j} = f_{Y_j}(y_j) F_{\mathbf{Y}_{-j}|Y_j}(\mathbf{y}_{-j} | y_j),$$

where $F_{\mathbf{Y}_{-j}|Y_j}(\cdot | y_j)$ is a conditional distribution, and apply integration by parts to (9) to get

$$\begin{aligned} C_{\mathbf{W}}(\mathbf{1}_d - \mathbf{u}\mathbf{x}) &= -u \left[\Phi_d\{1/(2\alpha) - \alpha \ln \mathbf{x} - \alpha t; \mathbf{R}\} e^{-t} \right]_{\ln u}^\infty \\ &\quad - u \sum_{i=1}^d \alpha_i \int_{\ln u}^\infty \Phi_{d-1}\{\mathbf{A}_i + \mathbf{B}_i t; \mathbf{R}^{(-i)} - \boldsymbol{\rho}_{-i}^{(i)}(\boldsymbol{\rho}_{-i}^{(i)})^\top\} \phi(D_i - \alpha_i t) e^{-t} dt + o(u) \\ &= 1 - u \sum_{i=1}^d \alpha_i \int_{\ln u}^\infty \Phi_{d-1}\{\mathbf{A}_i + \mathbf{B}_i t; \mathbf{R}^{(-i)} - \boldsymbol{\rho}_{-i}^{(i)}(\boldsymbol{\rho}_{-i}^{(i)})^\top\} \phi(D_i - \alpha_i t) e^{-t} dt + o(u), \end{aligned}$$

where

$$\begin{aligned} \mathbf{A}_i &= 1/(2\alpha_{-i}) - \alpha_{-i} \ln \mathbf{x}_{-i} - \boldsymbol{\rho}_{-i}^{(i)} \{1/(2\alpha_i) - \alpha_i \ln x_i\}, \\ \mathbf{B}_i &= -\alpha_{-i} + \boldsymbol{\rho}_{-i}^{(i)} \alpha_i, \\ D_i &= 1/(2\alpha_i) - \alpha_i \ln x_i. \end{aligned} \quad (10)$$

Let $\mathbf{R}_{-iji} = \mathbf{R}^{(-i)} - \boldsymbol{\rho}_{-i}^{(i)}[\boldsymbol{\rho}_{-i}^{(i)}]^\top$, where $\mathbf{R}^{(-i)}$ is \mathbf{R} without the i th row and column. Then with $\mathbf{Y}_i \sim \mathcal{N}_{d-1}(\mathbf{0}, \mathbf{R}_{-iji})$ for all $i \in \{1, \dots, d\}$ and $Z_0 \sim \mathcal{N}(0, 1)$ independent of each Y_i ,

$$\begin{aligned} \ell(\mathbf{x}) &\equiv \lim_{u \downarrow 0} \frac{1 - C_W(\mathbf{1}_d - u\mathbf{x})}{u} = \sum_{i=1}^d \alpha_i \int_{-\infty}^{\infty} \Phi_{d-1}(\mathbf{A}_i + \mathbf{B}_i t; \mathbf{R}_{-iji}) \phi(D_i - \alpha_i t) e^{-t} dt \\ &= \sum_{i=1}^d x_i \int_{-\infty}^{\infty} \Phi_{d-1} \left\{ \mathbf{A}_i + \mathbf{B}_i \left(\frac{s + D_i - 1/\alpha_i}{\alpha_i} \right); \mathbf{R}_{-iji} \right\} \phi(s) ds \\ &= \sum_{i=1}^d x_i \Pr \left(\mathbf{Y}_i + \frac{\mathbf{B}_i}{\alpha_i} Z_0 \leq \mathbf{A}_i + \mathbf{B}_i \frac{D_i - 1/\alpha_i}{\alpha_i} \right) \\ &= \sum_{i=1}^d x_i \Phi_{d-1} \left\{ \mathbf{A}_i + \mathbf{B}_i \left(\frac{D_i - 1/\alpha_i}{\alpha_i} \right); \mathbf{R}_{-iji}^* \right\}, \end{aligned} \tag{11}$$

where $\mathbf{R}_{-iji}^* = \mathbf{R}_{-iji} + \mathbf{B}_i \mathbf{B}_i^\top / \alpha_i^2$.

Note that if $\boldsymbol{\Omega} = (\omega_{jk})$ is a covariance matrix and $\text{diag}(\boldsymbol{\Omega}) = \boldsymbol{\omega}^2 = (\omega_1^2, \dots, \omega_m^2)^\top$, then

$$\Phi_m(\mathbf{y}; \boldsymbol{\Omega}) = \Phi_m(\mathbf{y}/\boldsymbol{\omega}; \boldsymbol{\Sigma}) \tag{12}$$

where $\boldsymbol{\Sigma} = (\omega_{jk} / \sqrt{\omega_{jj}\omega_{kk}})$.

We next rewrite (11) so that the last argument of Φ_{d-1} is a correlation matrix. We get

$$\begin{aligned} \mathbf{y}_i &\equiv \mathbf{A}_i + \mathbf{B}_i \left(\frac{D_i - 1/\alpha_i}{\alpha_i} \right) \\ &= \frac{1}{2\alpha_{-i}} - \frac{\boldsymbol{\rho}_{-i}^{(i)}}{2\alpha_i} - \alpha_{-i} \ln \mathbf{x}_{-i} + \alpha_i \boldsymbol{\rho}_{-i}^{(i)} \ln x_i + (-\alpha_{-i} + \boldsymbol{\rho}_{-i}^{(i)} \alpha_i) \left\{ \frac{-1/(2\alpha_i) - \alpha_i \ln x_i}{\alpha_i} \right\} \\ &= \frac{1}{2\alpha_{-i}} + \frac{\alpha_{-i}}{2\alpha_i^2} - \frac{\boldsymbol{\rho}_{-i}^{(i)}}{\alpha_i} - \alpha_{-i} \ln \mathbf{x}_{-i} + \alpha_{-i} \ln x_i \\ &= \frac{1}{2\alpha_i^2 \alpha_{-i}} (\alpha_i^2 + \alpha_{-i}^2 - 2\boldsymbol{\rho}_{-i}^{(i)} \alpha_i \alpha_{-i}) + \alpha_{-i} \ln(x_i / \mathbf{x}_{-i}) \\ &= \eta_i^2 \alpha_{-i} / 2 + \alpha_{-i} \ln(x_i / \mathbf{x}_{-i}), \end{aligned} \tag{13}$$

with

$$\eta_i \equiv \frac{(\alpha_i^2 \mathbf{1}_{d-1} + \alpha_{-i}^2 - 2\boldsymbol{\rho}_{-i}^{(i)} \alpha_i \alpha_{-i})^{1/2}}{\alpha_i \alpha_{-i}} \tag{14}$$

in vectorized mode. Let the elements of \mathbf{R}_{-iji}^* be denoted as $\omega_{i,jk}$ for $j \neq i, k \neq i$, and let $\boldsymbol{\omega}_i^2 = (\omega_{i,jj}, j \neq i)^\top$. With \mathbf{B}_i as defined in (10), the diagonal of \mathbf{R}_{-iji}^* is

$$\begin{aligned} \boldsymbol{\omega}_i^2 &\equiv \mathbf{1}_{d-1} - (\boldsymbol{\rho}_{-i}^{(i)})^2 + \frac{\mathbf{B}_i^2}{\alpha_i^2} = \mathbf{1}_{d-1} - (\boldsymbol{\rho}_{-i}^{(i)})^2 + \frac{\alpha_{-i}^2}{\alpha_i^2} + (\boldsymbol{\rho}_{-i}^{(i)})^2 - 2 \frac{\alpha_{-i} \boldsymbol{\rho}_{-i}^{(i)}}{\alpha_i} \\ &= \mathbf{1}_{d-1} + \frac{\alpha_{-i}^2}{\alpha_i^2} - 2 \frac{\alpha_{-i} \boldsymbol{\rho}_{-i}^{(i)}}{\alpha_i} = \frac{\alpha_i^2 \mathbf{1}_{d-1} + \alpha_{-i}^2 - 2 \boldsymbol{\rho}_{-i}^{(i)} \alpha_i \alpha_{-i}}{\alpha_i^2} \\ &= \eta_i^2 \boldsymbol{\alpha}_{-i}^2. \end{aligned} \tag{15}$$

Hence, using (12), (13) and (15),

$$\frac{\mathbf{y}_i}{\boldsymbol{\omega}_i} = \frac{1}{2} \boldsymbol{\eta}_i + \frac{1}{\boldsymbol{\eta}_i} \ln(x_i / \mathbf{x}_{-i}) = \left\{ \frac{1}{2} \boldsymbol{\eta}_{ij} + \boldsymbol{\eta}_{ij}^{-1} \ln(x_i / x_j) \right\}, \quad j \in \{1, \dots, i-1, i+1, \dots, d\},$$

where the $\boldsymbol{\eta}_{ij}$ in (14) match those in (4).

The final step is to derive the correlation matrix $\boldsymbol{\Gamma}_i = (\Gamma_{i,jk})$ from

$$\mathbf{R}_{-iji}^* = \mathbf{R}^{(-i)} - \boldsymbol{\rho}_{-i}^{(i)} (\boldsymbol{\rho}_{-i}^{(i)})^\top + \alpha_i^{-2} (-\alpha_{-i} + \boldsymbol{\rho}_{-i}^{(i)} \alpha_i) (-\alpha_{-i} + \boldsymbol{\rho}_{-i}^{(i)} \alpha_i)^\top.$$

For $j \neq k$ (and both not i), the (j, k) off-diagonal term of \mathbf{R}_{-iji}^* is

$$\omega_{i,jk} \equiv \frac{\rho_{jk} \alpha_i^2 + \alpha_j \alpha_k - \rho_{ij} \alpha_i \alpha_k - \rho_{ik} \alpha_i \alpha_j}{\alpha_i^2},$$

after some algebraic simplifications. Divide $\omega_{i,jk}$ by the square root of elements j and k in (15) to get

$$\Gamma_{i,jk} \equiv \frac{\rho_{jk}\alpha_i^2 + \alpha_j\alpha_k - \rho_{ij}\alpha_i\alpha_k - \rho_{ik}\alpha_i\alpha_j}{\alpha_i^2\eta_{ij}\eta_{ik}\alpha_j\alpha_k} = \frac{\eta_{ij}^2 + \eta_{ik}^2 - \eta_{jk}^2}{2\eta_{ij}\eta_{ik}},$$

where the last step consists of more algebraic simplification. This completes the proof of [Theorem 1](#). \square

An alternative parametrization of the HR distribution has the $d(d-1)/2$ dependence parameters labeled as $(\delta_{12}, \delta_{13}, \dots, \delta_{d-1,d})$, where $\delta_{ij} = 2\eta_{ij}^{-1}$ for all $i, j \in \{1, \dots, d\}$, and $\delta_{ji} = \delta_{ij}$ for $i < j$. This parametrization (see, e.g., Section 4.10 of [17]) is such that for the (i, j) bivariate margin with parameter δ_{ij} , the strength of dependence is increasing as δ_{ij} increases from 0 to infinity.

3. Mapping of parameters: trivariate case

In this section, we provide a detailed analysis for the trivariate case to show how the parameters of a HR distribution map to the parameters of the convolution model (1). This gives insight into the general multivariate case in the next section.

To reduce the number of subscripts, let the elements of Γ_i in (5) be denoted as $\gamma_{jk}^{(i)}$ for all $j, k \in \{1, \dots, i-1, i+1, \dots, d\}$. For $d = 3$, there are three parameters $\eta_{12}, \eta_{13}, \eta_{23}$.

Proposition 1. Consider the trivariate HR distribution with stable tail dependence function (3) so that $\Gamma_1, \Gamma_2, \Gamma_3$ in (5) are non-negative definite. Then for the three parameters, $\eta_{12}, \eta_{13}, \eta_{23}$, the sum of any two is larger than or equal to the third.

Proof. From the bivariate $\Gamma_1, \Gamma_2, \Gamma_3$ matrices in (5) we find $-1 \leq \gamma_{23}^{(1)} \leq 1$, $-1 \leq \gamma_{13}^{(2)} \leq 1$ and $-1 \leq \gamma_{12}^{(3)} \leq 1$. With i, j, k distinct, this becomes

$$-2\eta_{ij}\eta_{kj} \leq \eta_{ij}^2 + \eta_{kj}^2 - \eta_{ik}^2 \leq 2\eta_{ij}\eta_{kj},$$

$$\text{or } (\eta_{ij} - \eta_{kj})^2 \leq \eta_{ik}^2, \quad (\eta_{ij} + \eta_{kj})^2 \geq \eta_{ik}^2,$$

$$\text{or } |\eta_{ij} - \eta_{kj}| \leq \eta_{ik}, \quad \eta_{ij} + \eta_{kj} \geq \eta_{ik}.$$

This means that the sum of any two is larger than or equal to the third, and the absolute difference of any two is less than or equal to the third. If $0 < a \leq b \leq c$, then $a + b \geq c$ implies that $c - a \leq b$, and so the combination of the constraints is that the sum of the smaller two of $\eta_{12}, \eta_{13}, \eta_{23}$ is greater than or equal to the largest of the three. \square

We next consider the boundary case when the sum of two η s equals the third one, and then demonstrate the mapping of parameters from the convolution model with $d = 3$ to the HR parameters $\eta_{12}, \eta_{13}, \eta_{23}$.

3.1. Comment on the boundary case with singular component

At the boundary of the parameter space of the HR parameters, we show that the three correlations are ± 1 . For the boundary case, without loss of generality, suppose the indexing is such that $\eta_{23} = \eta_{12} + \eta_{13}$. Then $\gamma_{23}^{(1)} = -1$, $\gamma_{13}^{(2)} = 1$, $\gamma_{12}^{(3)} = 1$ since

$$\gamma_{23}^{(1)} = \frac{\eta_{12}^2 + \eta_{13}^2 - (\eta_{12} + \eta_{13})^2}{2\eta_{12}\eta_{13}} = \frac{-2\eta_{12}\eta_{13}}{2\eta_{12}\eta_{13}} = -1,$$

$$\gamma_{13}^{(2)} = \frac{\eta_{12}^2 + (\eta_{12} + \eta_{13})^2 - \eta_{13}^2}{2\eta_{12}\eta_{23}} = \frac{2\eta_{12}^2 + 2\eta_{12}\eta_{13}}{2\eta_{12}\eta_{23}} = \frac{\eta_{12} + \eta_{13}}{\eta_{23}} = 1,$$

and $\gamma_{12}^{(3)} = 1$ by switching subscripts 2,3 in the above.

For trivariate HR, (3) has the form

$$\begin{aligned} & x_1\Phi_2 \left\{ \frac{1}{2} \eta_{12} + \eta_{12}^{-1} \ln(x_1/x_2), \frac{1}{2} \eta_{13} + \eta_{13}^{-1} \ln(x_1/x_3); \gamma_{23}^{(1)} \right\} \\ & + x_2\Phi_2 \left\{ \frac{1}{2} \eta_{12} + \eta_{12}^{-1} \ln(x_2/x_1), \frac{1}{2} \eta_{23} + \eta_{23}^{-1} \ln(x_2/x_3); \gamma_{13}^{(2)} \right\} \\ & + x_3\Phi_2 \left\{ \frac{1}{2} \eta_{13} + \eta_{13}^{-1} \ln(x_3/x_1), \frac{1}{2} \eta_{23} + \eta_{23}^{-1} \ln(x_3/x_2); \gamma_{12}^{(3)} \right\}. \end{aligned}$$

For the boundary case with $\eta_{23} = \eta_{12} + \eta_{13}$, with the lower and upper Fréchet–Hoeffding bounds, this becomes

$$\begin{aligned} & x_1 \max \left\{ 0, \frac{1}{2} \eta_{12} + \eta_{12}^{-1} \ln(x_1/x_2) + \frac{1}{2} \eta_{13} + \eta_{13}^{-1} \ln(x_1/x_3) - 1 \right\} \\ & + x_2 \min \left\{ \frac{1}{2} \eta_{12} + \eta_{12}^{-1} \ln(x_2/x_1), \frac{1}{2} \eta_{23} + \eta_{23}^{-1} \ln(x_2/x_3) \right\} \\ & + x_3 \min \left\{ \frac{1}{2} \eta_{13} + \eta_{13}^{-1} \ln(x_3/x_1), \frac{1}{2} \eta_{23} + \eta_{23}^{-1} \ln(x_3/x_2) \right\}. \end{aligned}$$

Because there are “corner points” where left-sided and right-sided partial derivatives are not equal, the copula has non-singular components.

3.2. Mapping from the convolution model

If the parameters $\eta_{12}, \eta_{13}, \eta_{23}$ are such that the sum of the smaller two parameters is strictly larger than the largest of the three, then the singular component in the preceding subsection is avoided. In this case, we show that there is at least one mapping of parameter vector $(\rho_{12}, \rho_{13}, \rho_{23}, \alpha_1, \alpha_2, \alpha_3)$ in the normal–exponential convolution model (1) to the HR parameter vector $(\eta_{12}, \eta_{13}, \eta_{23})$. In fact, we show that this is possible with $\alpha_1 = \alpha_2 = \alpha_3 = \alpha > 0$.

Theorem 2. Assume that the sum of any two eta parameters, $\eta_{12}, \eta_{13}, \eta_{23}$, is strictly larger than the third one. In the convolution model (1), let $\mathbf{R} = (\rho_{ij})$, $i, j \in \{1, 2, 3\}$, with $\rho_{ij} = 1 - \beta^2 \eta_{ij}^2$ and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = \beta/\sqrt{2}$. Then \mathbf{R} is a positive definite matrix if $0 < \beta < \beta^* = \sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23})$, where

$$B_{32} = (\eta_{12} + \eta_{13} + \eta_{23})(\eta_{12} + \eta_{13} - \eta_{23})(\eta_{12} - \eta_{13} + \eta_{23})(-\eta_{12} + \eta_{13} + \eta_{23}),$$

and the limiting extreme-value copula in (1) is the absolutely continuous HR copula with parameters $\eta_{12}, \eta_{13}, \eta_{23}$.

Proof. From Theorem 1, the extreme-value limit in (1) with a valid correlation matrix \mathbf{R} and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha = \beta\sqrt{2}$ leads to the HR copula with parameters

$$\eta_{ij} = \frac{(\alpha_i^2 + \alpha_j^2 - 2\rho_{ij}\alpha_i\alpha_j)^{1/2}}{\alpha_i\alpha_j} = \frac{\sqrt{2}(1 - \rho_{ij})^{1/2}}{\alpha} = \frac{(1 - \rho_{ij})^{1/2}}{\beta}, \quad (i, j) \in \{(1, 2), (1, 3), (2, 3)\},$$

since we defined $\rho_{ij} = 1 - \beta^2 \eta_{ij}^2$. Now we need to check if $\mathbf{R} = (\rho_{ij})$, $i, j \in \{1, 2, 3\}$, is a valid correlation matrix. We have $-1 \leq \rho_{ij} \leq 1$ so that

$$\beta \leq \sqrt{2} / \max(\eta_{12}, \eta_{13}, \eta_{23}). \tag{16}$$

We next need to check when the determinant of \mathbf{R} is positive. We write $\rho_{ij} = \rho_{ij}(b) = 1 - \beta^2 \eta_{ij}^2 = 1 - bh_{ij}$, where $b = \beta^2$ and $h_{ij} = \eta_{ij}^2$; the matrix of the ρ_{ij} parameters is $\mathbf{R} = \mathbf{R}(b)$. We find

$$\begin{aligned} \det\{\mathbf{R}(b)\} &= 1 - \rho_{12}^2(b) - \rho_{13}^2(b) - \rho_{23}^2(b) + 2\rho_{12}(b)\rho_{13}(b)\rho_{23}(b) \\ &= 1 - (1 - bh_{12})^2 - (1 - bh_{13})^2 - (1 - bh_{23})^2 + 2(1 - bh_{12})(1 - bh_{13})(1 - bh_{23}) \\ &= b^2(2h_{12}h_{13} + 2h_{12}h_{23} + 2h_{13}h_{23} - h_{12}^2 - h_{13}^2 - h_{23}^2) - 2b^3h_{12}h_{13}h_{23}. \end{aligned} \tag{17}$$

Denote the coefficients of b^2 and b^3 in (17) as B_{32} and B_{33} , respectively. Then $B_{33} = -2h_{12}h_{13}h_{23}$, the determinant is $b^2(B_{32} + bB_{33})$ and

$$\begin{aligned} B_{32} &\equiv 2h_{12}h_{13} + 2h_{12}h_{23} + 2h_{13}h_{23} - h_{12}^2 - h_{13}^2 - h_{23}^2 \\ &= 2\eta_{12}^2\eta_{13}^2 + 2\eta_{12}^2\eta_{23}^2 + 2\eta_{13}^2\eta_{23}^2 - \eta_{12}^4 - \eta_{13}^4 - \eta_{23}^4 \\ &= (\eta_{12} + \eta_{13} + \eta_{23})(\eta_{12} + \eta_{13} - \eta_{23})(\eta_{12} - \eta_{13} + \eta_{23})(-\eta_{12} + \eta_{13} + \eta_{23}). \end{aligned} \tag{18}$$

Note that $B_{32} > 0$, with the constraints of $\eta_{12}, \eta_{13}, \eta_{23}$. The determinant of \mathbf{R} is positive if $B_{32} + B_{33}b > 0$ or $b < -B_{32}/B_{33}$. Combining with (16), we find that \mathbf{R} is positive definite when $b < b^*$ with

$$b^* = \min \left[2\{\max(h_{12}, h_{13}, h_{23})\}^{-1}, -B_{32}/B_{33} \right] = \min \left[2\{\max(h_{12}, h_{13}, h_{23})\}^{-1}, B_{32}/(2h_{12}h_{13}h_{23}) \right].$$

The upper bound on $\beta = \sqrt{b}$ is

$$\beta^* = \min \left[\sqrt{2} \{\max(\eta_{12}, \eta_{13}, \eta_{23})\}^{-1}, \sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23}) \right].$$

We next simplify β^* . Without loss of generality, we assume $\eta_{23} \geq \eta_{12}, \eta_{23} \geq \eta_{13}$, after permuting indices if needed, and $\eta_{12} + \eta_{13} > \eta_{23}$. Then

$$\beta^* = \min \left\{ \sqrt{2} \eta_{23}^{-1}, \sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23}) \right\}.$$

If $\beta = \sqrt{b} = \sqrt{2}\eta_{23}^{-1}$, then $\rho_{23} = 1 - 2\eta_{23}^2/\eta_{23}^2 = -1$ and $\det\{\mathbf{R}(b)\} = -(\rho_{12} + \rho_{13})^2 \leq 0$. This implies that $\rho_{12} = -\rho_{13}$ so that $\det\{\mathbf{R}(b)\} = 0$ and therefore from (17), $b = -B_{32}/B_{33}$, $\beta = \sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23})$. Hence it is not possible that $\sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23}) > \sqrt{2}\{\max(\eta_{12}, \eta_{13}, \eta_{23})\}^{-1}$, and $\beta^* = \sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23})$. This completes the proof of Theorem 2. \square

Remark 2. Let the right-hand side of (18) be denoted $B_{32}(\eta_{12}, \eta_{13}, \eta_{23})$. Note that if $\mathbf{R}(b)$ for $b < B_{32}(\eta_{12}, \eta_{13}, \eta_{23})/(2\eta_{12}^2\eta_{13}^2\eta_{23}^2)$ leads to the HR parameter vector $(\eta_{12}, \eta_{13}, \eta_{23})$, then for $a > 0$ and $\tilde{b} = b/a^2$, $\mathbf{R}(\tilde{b})$ for

$$\tilde{b} < B_{32}(a\eta_{12}, a\eta_{13}, a\eta_{23})/(2a^6\eta_{12}^2\eta_{13}^2\eta_{23}^2) = B_{32}(\eta_{12}, \eta_{13}, \eta_{23})/(2a^2\eta_{12}^2\eta_{13}^2\eta_{23}^2)$$

leads to the parameter vector $(a\eta_{12}, a\eta_{13}, a\eta_{23})$. The HR parameters imply there is weaker dependence for $(a\eta_{12}, a\eta_{13}, a\eta_{23})$ as a increases and the upper bound for $\tilde{\beta} = \sqrt{\tilde{b}}$ becomes $a^{-1}\sqrt{B_{32}/2}/(\eta_{12}\eta_{13}\eta_{23})$; i.e., as a increases, the multiplier on the additive factor ε becomes smaller. This makes sense as a larger $\beta = \alpha/\sqrt{2}$ value implies stronger (extreme-value) dependence.

4. Mapping of parameters: multivariate case

In this section, we deduce the constraints of the HR parameters for dimension 4 and higher, and show that for HR parameters $\{\eta_{ij} : i, j \in \{1, \dots, d\}, i < j\}$, that do not correspond to a distribution with a singular component, we can find a mapping to the convolution model (1) with constant α . The range of possible α parameters is narrower than those based on lower dimensional margins.

Remark 3. To see that there are extra constraints going from trivariate to general multivariate distributions, consider the exchangeable case. If $\eta_{ij} = \eta$ for all $i, j \in \{1, \dots, d\}$ with $i < j$, then the Γ_i matrices in (5) are all positive definite because the off-diagonal elements are all $1/2$. Hence, there exists a valid HR distribution with a constant parameter vector. With a constant α in (1), $\beta = \alpha/\sqrt{2}$, the extreme-value limit with a valid correlation matrix \mathbf{R} leads to the HR copula with parameters

$$\eta = \eta_{ij} = \frac{(\alpha_i^2 + \alpha_j^2 - 2\rho_{ij}\alpha_i\alpha_j)^{1/2}}{\alpha_i\alpha_j} = \frac{\sqrt{2}(1 - \rho_{ij})^{1/2}}{\alpha} = \frac{(1 - \rho_{ij})^{1/2}}{\beta}, \quad i, j \in \{1, \dots, d\}, i < j$$

so that $\rho_{ij} = 1 - \beta^2\eta_{ij}^2 = 1 - \beta^2\eta^2$ for all $i \neq j$. For dimension d , the resulting exchangeable correlation matrix is non-negative definite only if $1 - \beta^2\eta^2 \geq -1/(d - 1)$ or $\beta \leq \eta^{-1}\sqrt{d/(d - 1)}$. Hence the range of possible values of $\beta = \alpha/\sqrt{2}$ becomes narrower as d increases.

The next two theorems give some general results for the case with non-exchangeable dependence.

Theorem 3. Assume that Γ_d is a (strictly) positive definite matrix with $d \geq 3$. In the convolution model (1), let $\mathbf{R} = (\rho_{ij})$, $i, j \in \{1, \dots, d\}$, with $\rho_{ij} = 1 - \beta^2\eta_{ij}^2$ and $\alpha_1 = \alpha_2 = \dots = \alpha_d = \alpha = \beta\sqrt{2}$. Then \mathbf{R} is a positive definite matrix if $0 < \beta < \beta^*$ where $\beta^* = \sqrt{-B_{d,d-1}/B_{d,d}}$ and $B_{d,d-1}, B_{d,d}$ are given below in (20) and (21), respectively. In this case, the limiting extreme-value copula in (1) is the absolutely continuous HR copula with parameters (η_{ij}) , $i, j \in \{1, \dots, d\}$.

Proof. Similar to the proof of Theorem 2, the extreme-value limit in (1) with a valid correlation matrix \mathbf{R} and $\alpha_1 = \dots = \alpha_d = \alpha = \beta\sqrt{2}$ leads to the HR copula with parameters (η_{ij}) with $i, j \in \{1, \dots, d\}$. Let $b = \beta^2$ so that $\rho_{ij} = \rho_{ij}(b) = 1 - b\eta_{ij}^2$ for all $i \neq j$ (this is also well defined for $i = j$ with $\eta_{ii} = 0$ and $\rho_{ii} = 1$).

With $\eta_{ij} = \eta_{ij}^2$ as in the preceding section, write $\mathbf{R} = \mathbf{R}(b) = (1 - b\eta_{ij})$ with $i, j \in \{1, \dots, d\}$. We need to find b^* such that $\mathbf{R}(b)$ is a positive definite matrix for $0 < b < b^*$. By subtracting the last row of \mathbf{R} from rows $1, \dots, d - 1$, we get

$$\det[\mathbf{R}(b)] = \det \begin{pmatrix} b(h_{d1} - h_{11}) & b(h_{d2} - h_{12}) & \dots & b(h_{dd} - h_{1d}) \\ \vdots & \vdots & \ddots & \vdots \\ b(h_{d1} - h_{d-1,1}) & b(h_{d2} - h_{d-1,2}) & \dots & b(h_{dd} - h_{d-1,d}) \\ 1 - bh_{d1} & 1 - bh_{d2} & \dots & 1 - bh_{dd} \end{pmatrix} = b^{d-1}B_{d,d-1} + b^dB_{d,d}, \tag{19}$$

where

$$B_{d,d-1} = \det \begin{pmatrix} h_{d1} - h_{11} & h_{d2} - h_{12} & \dots & h_{dd} - h_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ h_{d1} - h_{d-1,1} & h_{d2} - h_{d-1,2} & \dots & h_{dd} - h_{d-1,d} \\ 1 & 1 & \dots & 1 \end{pmatrix}, \tag{20}$$

$$B_{d,d} = \det \begin{pmatrix} h_{d1} - h_{11} & h_{d2} - h_{12} & \dots & h_{dd} - h_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ h_{d1} - h_{d-1,1} & h_{d2} - h_{d-1,2} & \dots & h_{dd} - h_{d-1,d} \\ -h_{d1} & -h_{d2} & \dots & -h_{dd} \end{pmatrix}. \tag{21}$$

The determinant of the matrix in (20) does not change if we subtract its last column from columns 1, . . . , d − 1. Then the last row consists of d − 1 zeros and a one in the (d, d) position, so that B_{d,d−1} becomes the determinant of a (d − 1) × (d − 1) matrix, viz.

$$\begin{aligned}
 B_{d,d-1} &= \det \begin{pmatrix} \eta_{1d}^2 + \eta_{1d}^2 - \eta_{11}^2 & \eta_{1d}^2 + \eta_{2d}^2 - \eta_{12}^2 & \cdots & \eta_{1d}^2 + \eta_{d-1,d}^2 - \eta_{1,d-1}^2 \\ \eta_{2d}^2 + \eta_{1d}^2 - \eta_{12}^2 & \eta_{2d}^2 + \eta_{2d}^2 - \eta_{22}^2 & \cdots & \eta_{2d}^2 + \eta_{d-1,d}^2 - \eta_{2,d-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{d-1,d}^2 + \eta_{1d}^2 - \eta_{1,d-1}^2 & \eta_{d-1,d}^2 + \eta_{2d}^2 - \eta_{2,d-1}^2 & \cdots & \eta_{d-1,d}^2 + \eta_{d-1,d}^2 - \eta_{d-1,d-1}^2 \end{pmatrix} \\
 &= (\eta_{1d} \cdots \eta_{d-1,d})^2 \det \begin{pmatrix} \frac{\eta_{1d}^2 + \eta_{1d}^2 - \eta_{11}^2}{\eta_{1d}^2} & \frac{\eta_{1d}^2 + \eta_{2d}^2 - \eta_{12}^2}{\eta_{1d}\eta_{2d}} & \cdots & \frac{\eta_{1d}^2 + \eta_{d-1,d}^2 - \eta_{1,d-1}^2}{\eta_{1d}\eta_{d-1,d}} \\ \frac{\eta_{2d}^2 + \eta_{1d}^2 - \eta_{12}^2}{\eta_{2d}\eta_{1d}} & \frac{\eta_{2d}^2 + \eta_{2d}^2 - \eta_{22}^2}{\eta_{2d}^2} & \cdots & \frac{\eta_{2d}^2 + \eta_{d-1,d}^2 - \eta_{2,d-1}^2}{\eta_{2d}\eta_{d-1,d}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\eta_{d-1,d}^2 + \eta_{1d}^2 - \eta_{1,d-1}^2}{\eta_{d-1,d}\eta_{1d}} & \frac{\eta_{d-1,d}^2 + \eta_{2d}^2 - \eta_{2,d-1}^2}{\eta_{d-1,d}\eta_{2d}} & \cdots & \frac{\eta_{d-1,d}^2 + \eta_{d-1,d}^2 - \eta_{d-1,d-1}^2}{\eta_{d-1,d}^2} \end{pmatrix} \\
 &= (\eta_{1d} \cdots \eta_{d-1,d})^2 \times 2^{d-1} \times \det(\Gamma_d).
 \end{aligned}$$

Hence B_{d,d−1} > 0 if Γ_d is positive definite.

At the same time, after adding row d to rows 1, . . . , d − 1 in (21), one gets

$$B_{d,d} = \det \begin{pmatrix} 0 & -\eta_{12}^2 & \cdots & -\eta_{1d}^2 \\ -\eta_{21}^2 & 0 & \cdots & -\eta_{2d}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -\eta_{d-1,1}^2 & -\eta_{d-1,2}^2 & \cdots & -\eta_{d-1,d}^2 \\ -\eta_{d1}^2 & -\eta_{d2}^2 & \cdots & 0 \end{pmatrix}.$$

Now consider dimensions 2, . . . , d sequentially and define the corresponding values B_{j,j−1} and B_{j,j} based on the submatrix from the first j rows and columns of R(b) for j ∈ {2, . . . , d}. The determinants of the leading principal minors of R(b) are b^{j−1}B_{j,j−1} + b^jB_{j,j}. Let Γ_{d,j} be the submatrix of Γ_d from the first j rows and columns. Γ_d is positive definite so that Γ_{d,j} is positive definite. From the preceding results, B_{j,j−1} is equal in sign to det(Γ_{d,j}) so that it is positive for j ∈ {2, . . . , d}.

Because these B_{j,j−1} are positive, if we choose β² = b_ε > 0 small enough, then all leading principal minors of R(b_ε) have positive determinants, so that by Sylvester’s criterion (see [19]), R(b_ε) is positive definite. To find the upper bound b* of b for R(b) to retain positive definiteness, we increase b_ε to b₀ until the determinant of one (or several) leading principal minors is zero, while other leading principal minors have positive determinants. We know from Theorem 2 for j = 3 that b₀ is finite.

In this case, R(b₀) is positive semidefinite but it is not a positive definite matrix so that det{R(b₀)} = 0. This then implies that B_{d,d} < 0. A similar argument on the submatrices implies that B_{j,j} < 0 for j ∈ {3, . . . , d}. The definition of b₀ implies that b₀^{j−1}B_{j,j−1} + b₀^jB_{j,j} ≥ 0 for j ∈ {2, . . . , d − 1} and b₀^{d−1}B_{d,d−1} + b₀^dB_{d,d} = 0.

Hence, R is a positive definite correlation matrix if and only if 0 < b < b* = −B_{d,d−1}/B_{d,d}, and β* = √b*. The above also implies −B_{d,d−1}/B_{d,d} ≤ −B_{j,j−1}/B_{j,j} for j ∈ {2, . . . , d − 1}. By comparison on the submatrices, one also has that −B_{j,j−1}/B_{j,j} is decreasing as j increases. This completes the proof of Theorem 3. □

Theorem 4. Consider the multivariate HR copula with d ≥ 3 and stable tail dependence function given in (3) with parameters η_{ij} = η_{ji} for all i, j ∈ {1, . . . , d} with i < j, η_{ii} = 0 for i ∈ {1, . . . , d}, and with Γ_i defined in (5) for i ∈ {1, . . . , d}. The following properties hold:

- (a) If the parameters η_{ij} with i, j ∈ {1, . . . , d} are such that Γ_i, i ∈ {1, . . . , d}, are all positive definite, then the corresponding HR copula is an absolutely continuous distribution.
- (b) If one of the Γ_is is positive definite, then so are all the other Γ_is.
- (c) If one of the Γ_is is singular, then so are all the other Γ_is.

Proof. (a) A multivariate normal cumulative distribution function is absolutely continuous if and only if its correlation matrix parameter is positive definite. Furthermore, ℓ(x) in (3) is absolutely continuous only if each Γ_i for i ∈ {1, . . . , d} is a positive definite (non-singular) correlation matrix because these Γ_is are parameters of (d − 1)-dimensional Gaussian cumulative distribution functions.

(b) and (c): In the proof of [Theorem 3](#), it is shown that $B_{d,d-1} = (\eta_{1d} \cdots \eta_{d-1,d})^2 \times 2^{d-1} \times \det(\Gamma_d)$ and the positive definiteness of Γ_d implies $B_{d,d-1} > 0$. From [\(19\)](#), the left-hand side is invariant to the permutation of indices of the variables $1, \dots, d$, so that $B_{d,d}$ and $B_{d,d-1}$ are also invariant. Such permutation also allows us to obtain the following relationship:

$$B_{d,d-1} = 2^{d-1} \det(\Gamma_i) \cdot \prod_{j:1 \leq j \leq d, j \neq i} \eta_{ji}^2, \quad \text{for } i \in \{1, \dots, d\}.$$

Since $\eta_{ij} > 0$ for all $i \neq j$, $B_{d,d-1} > 0$ implies $\det(\Gamma_i) > 0$ for all $i \in \{1, \dots, d\}$. The leading submatrices based on the first $m \in \{2, \dots, d-1\}$ rows and columns of Γ_i appear in a marginal distribution, so that the leading minors all have positive determinants and Γ_i is positive definite. Hence if one of the Γ_i s is positive definite, then so are the other Γ_j for $j \neq i$. The result (c) is also immediate, as $\det(\Gamma_i) = 0$ for any i implies $B_{d,d-1} = 0$ and thus $\det(\Gamma_j) = 0$ for every $j \neq i$. This completes the proof of [Theorem 4](#). \square

Remark 4. The constraints on the η_{jk} parameters can be written in many different ways based on the conversion of a Γ_i matrix in [\(5\)](#) to correlations and partial correlations that are algebraically independent. Different ways of handling the constraints are useful if, for a simulation study, one wants to control the amount of dependence or conditional dependence. An illustration is given below for $d = 4$ to get a complete set of constraints for $(\eta_{ij} : i < j, i, j \in \{1, \dots, d\})$.

Consider Γ_1 matrices in [\(5\)](#), viz.

$$\begin{pmatrix} 1 & \gamma_{23}^{(1)} & \gamma_{24}^{(1)} \\ \gamma_{32}^{(1)} & 1 & \gamma_{34}^{(1)} \\ \gamma_{42}^{(1)} & \gamma_{43}^{(1)} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{\eta_{12}^2 + \eta_{13}^2 - \eta_{23}^2}{2\eta_{12}\eta_{13}} & \frac{\eta_{12}^2 + \eta_{14}^2 - \eta_{24}^2}{2\eta_{12}\eta_{14}} \\ \frac{\eta_{12}^2 + \eta_{13}^2 - \eta_{23}^2}{2\eta_{12}\eta_{13}} & 1 & \frac{\eta_{13}^2 + \eta_{14}^2 - \eta_{34}^2}{2\eta_{13}\eta_{14}} \\ \frac{\eta_{12}^2 + \eta_{14}^2 - \eta_{24}^2}{2\eta_{12}\eta_{14}} & \frac{\eta_{13}^2 + \eta_{14}^2 - \eta_{34}^2}{2\eta_{13}\eta_{14}} & 1 \end{pmatrix}.$$

Omitting the superscript in the γ_{ij} -s, the (sequential) constraints for non-negative definiteness can be written as $|\gamma_{23}| \leq 1$, $|\gamma_{24}| \leq 1$ and partial correlation

$$\gamma_{34;2} = [\gamma_{34} - \gamma_{23}\gamma_{24}] / \sqrt{(1 - \gamma_{23}^2)(1 - \gamma_{24}^2)}$$

satisfying $|\gamma_{34;2}| \leq 1$. From the trivariate case, the first two inequalities lead to

$$|\eta_{12} - \eta_{13}| \leq \eta_{23} \leq \eta_{12} + \eta_{13}, \quad |\eta_{12} - \eta_{14}| \leq \eta_{24} \leq \eta_{12} + \eta_{14},$$

with $\eta_{12}, \eta_{13}, \eta_{14}$ positive and algebraically independent. Next,

$$2\eta_{13}\eta_{14}[\gamma_{23}\gamma_{24} - \sqrt{(1 - \gamma_{23}^2)(1 - \gamma_{24}^2)}] \leq \eta_{13}^2 + \eta_{14}^2 - \eta_{34}^2 \leq 2\eta_{13}\eta_{14}[\gamma_{23}\gamma_{24} + \sqrt{(1 - \gamma_{23}^2)(1 - \gamma_{24}^2)}]$$

$$L \equiv A_1 - 2\eta_{13}\eta_{14}\sqrt{(1 - \gamma_{23}^2)(1 - \gamma_{24}^2)} \leq \eta_{34}^2 \leq A_1 + 2\eta_{13}\eta_{14}\sqrt{(1 - \gamma_{23}^2)(1 - \gamma_{24}^2)} \equiv U,$$

where $A_1 = \eta_{13}^2 + \eta_{14}^2 - 2\eta_{13}\eta_{14}\gamma_{23}\gamma_{24}$. The bounds of $L^{1/2}$ and $U^{1/2}$ on η_{34} are tighter than

$$\max\{|\eta_{14} - \eta_{13}|, |\eta_{24} - \eta_{23}|\} \leq \eta_{34} \leq \min\{\eta_{13} + \eta_{14}, \eta_{23} + \eta_{24}\},$$

which follows from the trivariate margins.

There are many ways to get a sequence of bounds for $\{\eta_{ij} : i, j \in \{1, \dots, d\}, i < j\}$ based on partial correlations; for any of the Γ_i matrices, a partial correlation vine [\[24\]](#) can be used for the sequence. For example, one sequence can be obtained if one of the Γ_i matrices is used with index order 4, 1, 2, 3 leading to bounds $L' \leq \eta_{23}^2 \leq U'$ in terms of the other five η parameters, similar to above. The symmetry in the Γ matrices implies that equivalent sets of inequalities arise from $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ in [\(5\)](#). If a set $\{\eta_{ij}\}$ leads to Γ_1 not being positive definite, then neither are Γ_2, Γ_3 and Γ_4 , based on [Theorem 4](#).

The above extends to higher dimensions, and the notation involves submatrices of Γ . If one of the constraints for $\{\eta_{ij} : i, j \in \{1, \dots, d\}, i < j\}$ becomes an equality, then the resulting HR distribution has a singular component; furthermore, for one of the principal minors of \mathbf{R} we have $B_{j,j-1} = 0$, where $B_{j,j-1}$ is defined in the proof of [Theorem 3](#). This implies that the determinant of this principal minor equals $B_{jj}b^j < 0$ so that there is no match from Model [\(1\)](#).

5. Application to simulation from the Hüsler–Reiß distribution

The spatial extreme-value models [\[5,31\]](#) lead to multivariate extreme-value limits that have multivariate HR distributions for any subsets of sites; see [\[12,15\]](#). The parameters of these HR distributions have a parsimonious form that depends on the distance of sites relative to a covariance matrix Ω .

The approximate simulation method based on the convolution model proceeds as follows: let $\mathbf{W}_1^{(i)}, \dots, \mathbf{W}_N^{(i)}$ be independent realizations of the random vectors from Model [\(1\)](#) for $i \in \{1, \dots, n\}$. Convert them to vectors of $\mathcal{U}(0, 1)$ random variables $\mathbf{U}_1^{(i)}, \dots, \mathbf{U}_N^{(i)}$ via the probability integral transform with [\(8\)](#). Let $M_{ij} = \max\{U_{ij}^{(1)}, \dots, U_{ij}^{(n)}\}$ for $i \in \{1, \dots, n\}$ and

$j \in \{1, \dots, d\}$. Let $\mathbf{M}_i = (M_{i1}, \dots, M_{id})^\top$ be the i th vector of dependent maxima; \mathbf{M}_i are independent for different i . Apply the Beta distribution $\mathcal{B}(N, 1)$ transform to the elements of \mathbf{M}_i resulting in \mathbf{V}_i that have $\mathcal{U}(0, 1)$ random components for $i \in \{1, \dots, n\}$. If N is large enough, $\mathbf{V}_1, \dots, \mathbf{V}_n$ is a random sample with a joint distribution that is approximately the HR copula.

Given parameters $\{\eta_{ij} : i, j \in \{1, \dots, d\}, i < j\}$ satisfying the relevant constraints (say Γ_1 is positive definite), we can choose parameters from (1) that have these particular parameters for the HR extreme-value limit. Hence this is a way to do a parametric bootstrap/simulation after fitting a HR copula with parsimonious dependence structure.

Remark 5. An alternative algorithm is to use the Poisson point process representation: let $\mathbf{Z}_1^{(i)}, \dots, \mathbf{Z}_N^{(i)}$ be independent realizations of the multivariate Gaussian distribution with correlation matrix \mathbf{R} . Let E_1, \dots, E_N be a random sample from $\mathcal{E}(1)$ and for each $j \in \{1, \dots, N\}$, let $P_j = E_1 + \dots + E_j$. With $\exp(\mathbf{z})$ of a vector \mathbf{z} denoting component-wise operation and maximum of a sequence of vectors denoting component-wise maxima, define

$$\mathbf{V}_i = \max \left[\exp\{\mathbf{Z}_1^{(i)}/\alpha - 1/(2\alpha^2)\}/P_1, \dots, \exp\{\mathbf{Z}_N^{(i)}/\alpha - 1/(2\alpha^2)\}/P_N \right]^\top.$$

If N is large enough, $\mathbf{V}_1, \dots, \mathbf{V}_n$ is a random sample with approximate unit Fréchet marginals and a joint distribution that has the HR copula as a limit; see [29].

5.1. Comparison with exact simulation methods

Recently, Dieker and Mikosch [8] proposed an exact simulation method for stationary Brown–Resnick random fields and Dombry et al. [9] extended their approach and used it for the simulation of max-stable processes using the spectral measure. The authors also proposed a new algorithm, based on the simulation of extremal functions, and showed that this algorithm is more efficient in terms of speed. For d spatial locations, this algorithm requires, on average, simulation of d d -variate normal vectors. This implies that the approximate method is theoretically slower if $d < N$ and can be more efficient if d is very large, for example when simulating a spatial process on a dense grid.

The exact method for simulating from the Hüsler–Reiß distribution (with the Brown–Resnick process as a special case) is summarized in Algorithm 1; this is a version of Algorithm 1 in [9] with \mathbf{R} as the input. For both the approximate and exact simulation methods, we check our implementations by applying Pearson’s chi-square goodness-of-fit test to discretized data from each bivariate margin: if $(U_i, U_j) \sim C_{ij}(\cdot; \eta_{ij})$ for the (i, j) margin, where C_{ij} is a marginal copula, then $(U_i, C_{j|i}(U_j|U_i))$ has a uniform distribution on the unit square where $C_{j|i}(y|x) = \partial C_{ij}(x, y)/\partial x$ is a conditional distribution. For many different parameter inputs and many replications of samples of size 1000, we find no significant departure of the distribution of the p -values based on the Pearson’s chi-square test from the uniform distribution on $(0, 1)$.

Algorithm 1 Algorithm to simulate from the Hüsler–Reiß distribution with Fréchet(1) margins and parameter matrix $\{\eta_{ij} : i, j \in \{1, \dots, d\}\}$, or based on the convolution model with correlation matrix $\mathbf{R} = (\rho_{ij})$ with $i, j \in \{1, \dots, d\}$ and constant $\alpha = 1/\sigma^2 > 0$ leading to $\{\eta_{ij} : i, j \in \{1, \dots, d\}\}$.

```

Simulate  $\zeta^{-1} \sim \mathcal{E}(1)$  and  $\mathbf{W} = (W_1, \dots, W_d)^\top \sim \mathcal{N}_d(\mathbf{0}, \mathbf{R})$ .
Set  $\mathbf{Y}_1 = (Y_{11}, \dots, Y_{1d})$  where  $Y_{1j} \leftarrow \exp\{\sigma^2(W_j - W_1 - 1 + \rho_{j1})\}$  for  $j \in \{1, \dots, d\}$ .
Set  $\mathbf{Z} = (Z_1, \dots, Z_d)^\top \leftarrow \zeta \mathbf{Y}_1$ .
for  $i \in \{2, \dots, d\}$ : do
  while  $\zeta > Z_i$  do
    Simulate  $\mathbf{W} = (W_1, \dots, W_d)^\top \sim \mathcal{N}_d(\mathbf{0}, \mathbf{R})$ .
    Set  $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{id})^\top$  where  $Y_{ij} \leftarrow \exp\{\sigma^2(W_j - W_i - 1 + \rho_{ji})\}$  for  $j \in \{1, \dots, d\}$ .
    If  $\zeta Y_{ij} < Z_j$  for all  $j \in \{1, \dots, i - 1\}$ , then update  $Z_k \leftarrow \max(Y_{ik}, Z_k)$  for  $k \in \{1, \dots, d\}$ .
    Simulate  $E \sim \mathcal{E}(1)$  and update  $\zeta \leftarrow 1/(\zeta^{-1} + E)$ .
  end while
end for
Return the  $d$ -vector  $\mathbf{Z}$  of dependent unit Fréchet random variables, or  $\mathbf{U} = 1 - \exp(-1/\mathbf{Z})$  with  $\mathcal{U}(0, 1)$  variables.

```

With $N = 1000$ we now simulate a random sample of size $n = 100$ using our approach. For comparison, we use the exact method employing the extremal functions. For simplicity, we simulate a Brown–Resnick process at d locations in \mathbb{R}^1 : $0, 1/(d - 1), 2/(d - 1), \dots, 1$. We assume that the Brown–Resnick process has a correlation matrix which is defined by the stationary covariance function $\rho(s_1, s_2) = \exp\{-\psi_1|s_1 - s_2|^{\psi_2}\}$, with $\psi_1 = 0.5, \psi_2 = 1.2$ and the variance $\sigma^2 = 1.25$. This corresponds to the extreme-value limit of (1) with the same correlation matrix, \mathbf{R} , and $\alpha = \sigma^{-2}\mathbf{1}_d$; see [12,20].

In Algorithm 1 and our approximate method, \mathbf{R} can be an arbitrary correlation matrix (not necessarily stationary or associated with a spatial process). We also tried other parameters ψ_1, ψ_2 and other (including non-stationary) covariance structures, with similar results. Fig. 1 shows the run time for the two methods for $d \in \{100, 200, \dots, 2500\}$.

Table 1
 RMSD_S, RMSD_L, RMSD_U and $\widehat{m}(\mathbf{w})$ for $\mathbf{w} = \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ (standard errors are shown in parentheses); $\alpha = 0.8$.

Method	$\widehat{m}(\mathbf{w}_1)$	$\widehat{m}(\mathbf{w}_2)$	$\widehat{m}(\mathbf{w}_3)$	$\widehat{m}(\mathbf{w}_4)$	$\widehat{m}(\mathbf{w}_5)$
$d = 1000$: RMSD _S = 0.01, RMSD _L = 0.01, RMSD _U = 0.02					
Exact	0.428(0.004)	0.270(0.003)	0.234(0.002)	0.321(0.003)	0.318(0.003)
Approximate	0.430(0.004)	0.269(0.003)	0.236(0.002)	0.323(0.003)	0.320(0.003)
$d = 2000$: RMSD _S = 0.00, RMSD _L = 0.01, RMSD _U = 0.02					
Exact	0.427(0.004)	0.268(0.003)	0.234(0.002)	0.316(0.003)	0.320(0.003)
Approximate	0.424(0.004)	0.265(0.003)	0.232(0.002)	0.315(0.003)	0.314(0.003)

With $N = 1000$ and $n = 50\,000$ replicates, we get $\text{RMSD}_{\rho_S} = 0.15$, $\text{RMSD}_L = 0.33$, $\text{RMSD}_U = 0.14$, so that the approximate method does not work well. With $\alpha = 0.30$, we obtain

$$\zeta \approx (0.96, 0.98, 0.96, 0.97, 0.94, 0.96, 0.95, 0.96, 0.96, 0.97, 0.96, 0.97, 0.96, 0.98, 0.98, 0.97, 0.98)^\top$$

and $\text{RMSD}_{\rho_S} = 0.04$, $\text{RMSD}_L = 0.09$, $\text{RMSD}_U = 0.03$, and the approximation works much better. If $\alpha \geq 0.5$, the approximation is very accurate because the correlation parameters are not close to ones. For example, if $\alpha = 1$, we obtain

$$\zeta \approx (0.59, 0.75, 0.50, 0.69, 0.34, 0.55, 0.45, 0.59, 0.55, 0.72, 0.59, 0.64, 0.55, 0.79, 0.75, 0.72, 0.82)^\top$$

and $\text{RMSD}_{\rho_S} = 0.00$, $\text{RMSD}_L = 0.01$, $\text{RMSD}_U = 0.01$.

We obtained similar results for other simulated data sets with different α_{\max} . We found $N = 1000$ and $\alpha > 0.5\alpha_{\max}$ is a good choice in most cases.

5.3. Accuracy of the approximation in high dimensions

We now assess the accuracy of the approximate method for high-dimensional data sets. With $N = 1000$, we simulate a random sample of size $n = 10,000$ using our approach and the exact method as described in Algorithm 1 in Section 5.1. As mentioned there, we simulate a Brown–Resnick process at d locations in $\mathbb{R} : 0, 1/(d - 1), 2/(d - 1), \dots, 1$, but with $d = 1000$ and $d = 2000$. Again, we assume that the Brown–Resnick process has the correlation matrix which is defined by the stationary covariance function $\rho(s_1, s_2) = \exp(-\psi_1|s_1 - s_2|^{\psi_2})$, with $\psi_1 = 0.5$, $\psi_2 = 1.2$ and the variance $\sigma^2 = 1.25$, so that $\alpha = 0.8$. This implies that $\alpha \approx 0.52\alpha_{\max}$ in this case.

We compute RMSD_S , RMSD_L and RMSD_U as before. In addition, to check if higher order dependencies are well approximated, we compute the weighted minima, $\mathbf{e}_i(\mathbf{w}) = \min_{j \in \{1, \dots, d\}}(w_j e_{ij})$, of the simulated vectors, $\mathbf{u}_i = (u_{i1}, \dots, u_{id})^\top$, $i \in \{1, \dots, n\}$, transformed to the unit-exponential variables $e_{ij} = -\ln(u_{ij})$, $i \in \{1, \dots, n\}$, $j \in \{1, \dots, d\}$. If \mathbf{u}_i follows the HR distribution, then $\mathbf{e}_1(\mathbf{w}), \dots, \mathbf{e}_n(\mathbf{w})$ are i.i.d. exponential random variables with mean $m(\mathbf{w}) = 1/\ell(1/\mathbf{w})$ where the stable tail dependence function, ℓ , is given in (3). We use the sample mean, $\widehat{m}(\mathbf{w}) = \sum_{i=1}^n \mathbf{e}_i(\mathbf{w})/n$, to estimate $m(\mathbf{w})$ using both the exact and the approximate methods, for different vectors $\mathbf{w} = \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5$ where

$$\begin{aligned} \mathbf{w}_1 &= (1, \dots, 1)^\top, \\ \mathbf{w}_2 &= (\underbrace{0.5, \dots, 0.5}_{d/4}, \underbrace{1, \dots, 1}_{d/2}, \underbrace{0.5, \dots, 0.5}_{d/4})^\top, \\ \mathbf{w}_3 &= (\underbrace{1, \dots, 1}_{d/4}, \underbrace{0.5, \dots, 0.5}_{d/2}, \underbrace{1, \dots, 1}_{d/4})^\top, \\ \mathbf{w}_4 &= (\underbrace{0.5, \dots, 0.5}_{d/4}, \underbrace{1, \dots, 1}_{d/4}, \underbrace{1.5, \dots, 1.5}_{d/4}, \underbrace{2, \dots, 2}_{d/4})^\top, \\ \mathbf{w}_5 &= (\underbrace{2, \dots, 2}_{d/4}, \underbrace{1.5, \dots, 1.5}_{d/4}, \underbrace{1, \dots, 1}_{d/4}, \underbrace{0.5, \dots, 0.5}_{d/4})^\top. \end{aligned}$$

Table 1 shows the results. We can see that the values of $m(\mathbf{w})$ are very close for the two methods and the RMSD is small for the approximate method. In addition, we use Q–Q plots to check if the two samples of the weighted minima, generated using the approximate and exact algorithms, come from the same (exponential) distribution. First, we compare the two samples and then we compare each of these samples to the respective theoretical quantiles, $-\widehat{m}(\mathbf{w}) \ln \{1 - (i - 0.5)/n\}$, $i \in \{1, \dots, n\}$. Fig. 2 shows the Q–Q plots for $d = 1000$ and $\mathbf{e}_i(\mathbf{w})$ with $\mathbf{w} = \mathbf{w}_1, \mathbf{w}_3$ and \mathbf{w}_5 ; we got similar Q–Q plots for $d = 2000$. We can see that the weighted minima data obtained using both samples fit the exponential distribution very well.

Similar results are obtained with different sets of parameters for \mathbf{R} and weights \mathbf{w} . One can therefore conclude that the approximate method also works well in high dimensions, and that higher order dependencies are well approximated when $N = 1000$ is used.

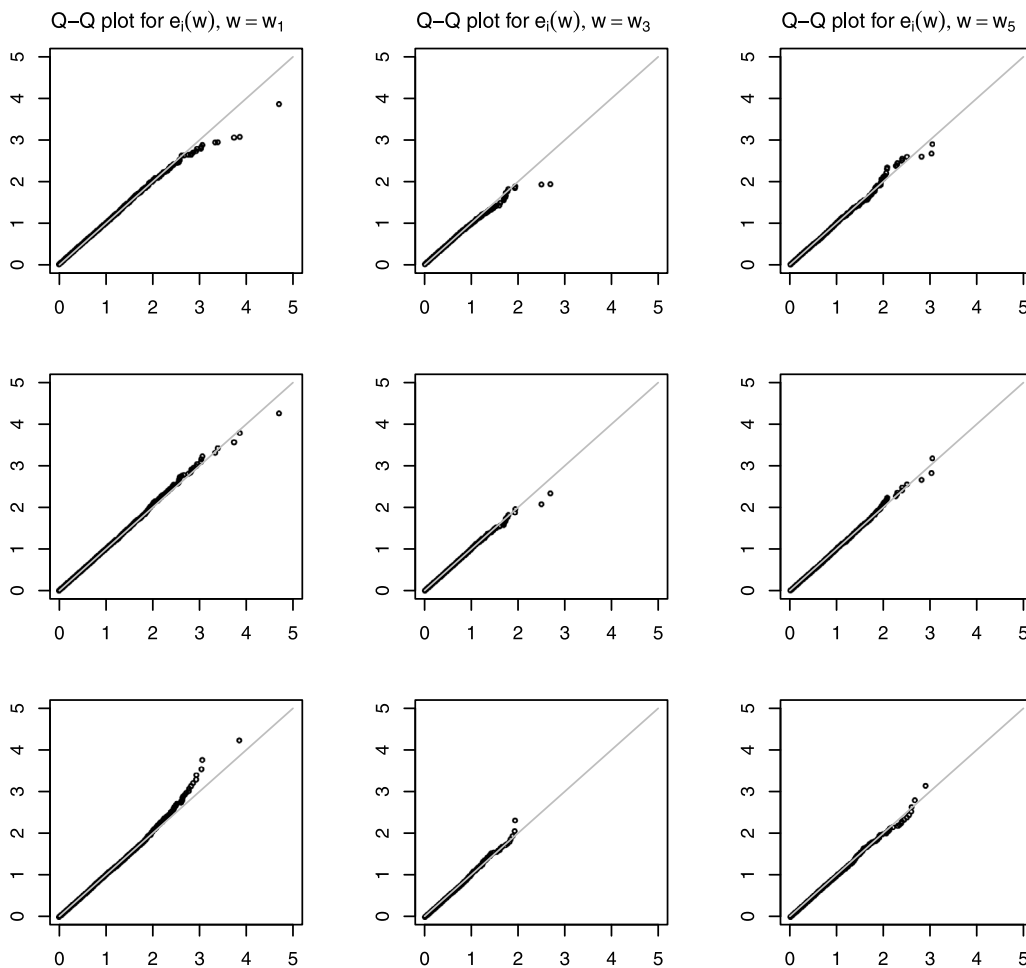


Fig. 2. Q–Q plots for $e_i(\mathbf{w})$ with $\mathbf{w} = \mathbf{w}_1$ (left), \mathbf{w}_3 (middle) and \mathbf{w}_5 (right). We compare two samples generated using the approximate and the exact methods (top row), first sample and theoretical quantiles (middle row) and second sample and theoretical quantiles (bottom row).

Remark 6. We obtained very similar results with $N = 1000$ for the alternative approximate simulation method based on the Poisson point process (see Remark 5). This is because the approximation method based on the Poisson point process is asymptotically equivalent to that using the convolution model (1). The choice of α is crucial for both approximation methods as small values of $\alpha \ll 0.5\alpha_{\max}$ can lead to a nearly singular correlation matrix \mathbf{R} . These approximate methods can be very inaccurate in this case, as we showed in Section 5.2 (we obtained similar results for the Schlather’s method).

Note that some other approximate methods are also available for the simulation of the Brown–Resnick process. These approximate methods assume stationarity of the underlying Gaussian process, implying a special structure for the correlation matrix \mathbf{R} . Therefore, they cannot be used for simulating data from the HR copula with arbitrary parameters; see [28] and the references therein.

5.4. Example of multivariate extremes with parsimonious dependence

In this subsection, we illustrate the use of simulation from the HR copula in evaluating model adequacy of a parsimonious dependence structure based on a goodness-of-fit statistic. The procedure used here extends to any multivariate model with parsimonious dependence and any goodness-of-fit statistic for which simulation, model fitting and obtaining features in the goodness-of-fit statistic are all computationally fast enough.

In the multivariate setting, we are often interested in statistical models with parsimonious dependence structures that offer intuitive interpretations. Two classes of simple structures are factor and Markov trees. With the factor structure, the dependence among variables is assumed to be driven through one or more latent factor(s); see [22]. With the Markov tree structure, the variables are joined together directly in the sense that neighboring variables have the strongest dependence, while the dependence between those further apart in the tree linkage is induced by that of successive neighboring variables

assuming conditional independencies. Markov trees are the special case of vine copulas [1] when the variables at higher levels of the vine are assumed to be conditionally independent; see [4]. These models are parsimonious as they require the use of only $O(d)$ variables, in contrast to $O(d^2)$ for a saturated model. One class of parsimonious dependence models for multivariate extremes can be obtained from HR distributions with the $d(d - 1)/2$ parameters being represented in terms of $O(d)$ parameters; see [25].

When fitting parsimonious dependence models, it is useful to inspect whether they are adequate in capturing the dependence characteristics expressed by the data. A natural statistic to consider is the root mean square differences (RMSD); here it is slightly different from the one introduced in Section 5.2 in the sense that we are interested in the discrepancy between the data (empirical measure) and the fitted model (model-based measure). Let $\rho_S^{i_1, i_2}$ (empirical) be the empirical Spearman rank correlation for the (i_1, i_2) margin, and $\rho_S^{i_1, i_2}$ (model) be the one calculated using the fitted model. The RMSD is defined as

$$\text{RMSD}_S = \left[\frac{2}{d(d-1)} \sum_{i_1 < i_2} \left\{ \rho_S^{i_1, i_2}(\text{empirical}) - \rho_S^{i_1, i_2}(\text{model}) \right\}^2 \right]^{1/2}.$$

This usage is analogous to a fit statistic used in the psychometrics literature; see [13]. A similar definition applies to Kendall's tau or other measures of dependence. To determine whether a model is adequate, one can compare the RMSD statistic to the typical behavior of such statistic assuming the fitted model is correctly specified. That is, we seek the sampling distribution of the RMSD statistic under the fitted model via a parametric bootstrap procedure, which requires a simulation from the model.

Extreme-value copulas with parsimonious structures allow us to describe the dependence structure among extremal observations. The HR copula with parameters (η_{ij}) and stable tail dependence function (3) is the limiting copula of a Gaussian random vector $(Y_1, \dots, Y_d)^T$ in which the elements of the correlation matrix $\mathbf{R}_Y(n) = (\rho_{ij}(n))$, a function of the sample size n , tend to 1 at the following rate:

$$4\{1 - \rho_{ij}(n)\} \ln n \rightarrow \eta_{ij}^2, \quad n \rightarrow \infty, \quad i < j, \quad i, j \in \{1, \dots, d\}.$$

Following [25], this relationship suggests a link between η_{ij} and the correlation of the underlying Gaussian vector: for n large, write $\eta_{ij} = \gamma(1 - \rho_{ij})^{1/2}$, where $\gamma > 0$ is a proportionality constant. A factor or Markov tree structure can be imposed on the Gaussian correlation matrix \mathbf{R}_Y ; the η_{ij} parameters are then represented as functions of $O(d)$ algebraically independent parameters characterizing \mathbf{R}_Y .

The data example we use contains the bimonthly minima of daily log returns of seven major stocks traded in the US stock markets between January 1997 and October 2016, for a total of 119 observations. The log return for day t is defined as $r_t = \ln(P_t/P_{t-1})$, where P_t is the closing stock price on day t . We sign-adjust these stock returns so that they exhibit upper tail dependence, i.e., extreme losses tend to occur together. These stocks are all chosen from the pharmaceutical sector: GlaxoSmithKline PLC; Johnson & Johnson; Eli Lilly and Company; Merck & Co., Inc.; Mylan NV; Novartis AG and Pfizer Inc. Since these stocks are from the same sector, the pairwise dependence of the normal scores of the returns is moderate to strong (between 0.4 and 0.6 for most pairs); see Fig. 3. The returns can be thought of driven through a sector-specific latent factor, and thus a factor structure is plausible. The Markov tree structure connects variables with the strongest dependence, and may be plausible if the performance of some stocks is an important indicator of the overall performance of the sector (and thus other stocks may be more strongly correlated with these “indicator” stocks).

We fit generalized extreme-value distributions to each of the seven univariate series (bimonthly minima are taken to reduce the effect of autocorrelation); diagnostic plots (not shown) indicate decent fit. The series are transformed to have $\mathcal{U}(0, 1)$ marginal distributions via the probability integral transform using the fitted parameters, and the resulting copula data are used to fit two HR copulas, one with a 1-factor structure and the other with a Markov tree structure. The tree is chosen using the minimum spanning tree algorithm assuming a Gaussian copula on the normal scores. Because of the intractability of the full likelihood of the HR copula, we adopt the method of composite likelihood [7,26]; in particular, we use the pairwise likelihood method as bivariate densities can be easily obtained.

Table 2 displays the fitting results in terms of the composite likelihood information criteria (CLIC and CLBIC), analogues of the AIC and BIC for the composite likelihood; see [11,32]. With these criteria, the HR copula with a Markov tree structure is slightly better than the 1-factor model; both are better than the saturated HR copula (with 21 parameters) which yields a CLIC and CLBIC of -549 and -452 , respectively. Based on the fitted models, we compute the RMSD using Kendall's tau and Spearman's rank correlation as the dependence measures. The critical values are defined as the 0.95 quantile of the sampling distribution of the RMSD statistic assuming that the fitted model is correctly specified, obtained using a parametric bootstrap (i.e., repeatedly simulating from the fitted model and fitting these data sets to the structured HR copulas). Simulation from the fitted HR copula is carried out using the exact and approximate methods described in the preceding subsections, with a value of α that is between $\alpha_{\max}/2$ and α_{\max} , where $\alpha_{\max} = \sqrt{-2B_{d,d-1}/B_{d,d}}$ with $B_{d,d-1}$ and $B_{d,d}$ defined as in Theorem 3 in Section 4.

We observe that both fitted models yield RMSD statistics that are above their respective critical values; this suggests that these structures may be too parsimonious and higher-order generalizations should be considered. More factors may be added to the factor structure, while more layers can be added for the tree structure by removing conditional independencies using the idea of truncated vines [2–4]. Alternatively, other multivariate extreme-value copulas with factor structure could be considered. The difference between the critical values based on the two simulation methods is very small, supporting the assertion that the approximation is sufficiently accurate when the Σ matrix is not close to singular.

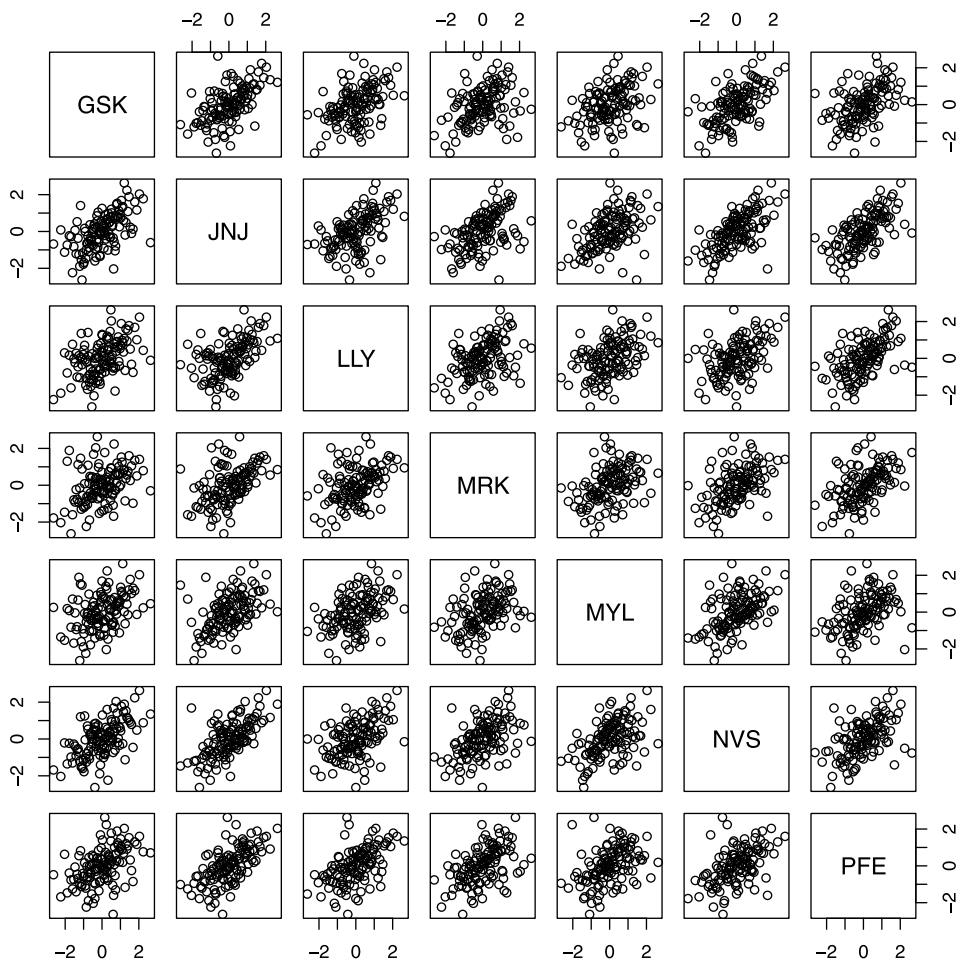


Fig. 3. Pairwise scatterplots of the normal scores of the US stock data, with observations being the sign-adjusted bimonthly minima of daily log returns.

Table 2

Fitting results of the HR 1-factor / Markov tree copula models. The columns are (left to right): model, number of parameters, composite likelihood information criterion, composite likelihood Bayesian information criterion, root mean square differences and critical values (CV) based on the Kendall's τ and Spearman's ρ_S (rank) correlation measures.

Model	# par	CLIC	CLBIC		RMSD (τ)	CV (τ)	RMSD (ρ_S)	CV (ρ_S)
1-factor	8	-562	-493	Exact	0.077	0.063	0.099	0.086
				Aprx.		0.061		0.086
Markov tree	7	-572	-513	Exact	0.077	0.064	0.099	0.089
				Aprx.		0.063		0.088

6. Discussion

In this paper we obtained an extreme-value limit of a convolution model that extends the factor copula model for replicated spatial data [21]. For the first time, the class of absolutely continuous HR copulas is derived as a direct extreme-value limit that covers the entire parameter space. We showed how the parameters of the limiting HR copula can be mapped to the parameters of this convolution model. Based on these results, we proposed an approximate simulation method for the HR copula with valid parameter vectors that do not yield singular components. This method is slower than some exact simulation methods proposed in the literature if the dimension $d < 1000$. However, in higher dimensions, the approximate method is more efficient in terms of computational speed. We applied the approximate and exact simulation procedures in a parametric bootstrap for diagnostic checks after fitting some parsimonious HR copulas to a data set of multivariate extremes.

The technique used to derive the HR copula from a convolution model can be extended to other convolution models to get other parametric models for multivariate extremes.

Acknowledgments

This research was supported by the King Abdullah University of Science and Technology (KAUST), Discovery Grant No. 8698 from the Natural Sciences and Engineering Research Council of Canada, a Collaborative Research Team grant for the project *Copula Dependence Modeling: Theory and Applications* of the Canadian Statistical Sciences Institute (CANSSI), and a University of British Columbia four-year doctoral fellowship. We are grateful to the two referees, the Associate Editor, and the Editor-in-Chief for their comments.

References

- [1] K. Aas, C. Czado, A. Frigessi, H. Bakken, Pair-copula constructions of multiple dependence, *Insurance Math. Econom.* 44 (2009) 182–198.
- [2] T. Bedford, R.M. Cooke, Probability density decomposition for conditionally dependent random variables modeled by vines, *Ann. Math. Artif. Intell.* 32 (2001) 245–268.
- [3] T. Bedford, R.M. Cooke, Vines — a new graphical model for dependent random variables, *Ann. Statist.* 30 (2002) 1031–1068.
- [4] E.C. Brechmann, C. Czado, K. Aas, Truncated regular vines in high dimensions with application to financial data, *Canad. J. Statist.* 40 (2012) 68–85.
- [5] B.M. Brown, S.I. Resnick, Extreme values of independent stochastic processes, *J. Appl. Probab.* 14 (1977) 732–739.
- [6] A. Charpentier, A.-L. Fougères, C. Genest, J.G. Nešlehová, Multivariate Archimax copulas, *J. Multivariate Anal.* 126 (2014) 118–136.
- [7] D.R. Cox, N.M. Reid, A note on pseudolikelihood constructed from marginal densities, *Biometrika* 91 (2004) 729–737.
- [8] A.B. Dieker, T. Mikosch, Exact simulation of Brown–Resnick random fields at a finite number of locations, *Extremes* 18 (2015) 301–314.
- [9] C. Dombry, S. Engelke, M. Oesting, Exact simulation of max-stable processes, *Biometrika* 103 (2016) 303–317.
- [10] H. Drees, X. Huang, Best attainable rates of convergence for estimators of the stable tail dependence function, *J. Multivariate Anal.* 64 (1998) 25–46.
- [11] X. Gao, P.X.-K. Song, Composite likelihood Bayesian information criteria for model selection in high-dimensional data, *J. Amer. Statist. Assoc.* 105 (2010) 1531–1540.
- [12] M.G. Genton, Y. Ma, H. Sang, On the likelihood function of Gaussian max-stable processes, *Biometrika* 98 (2011) 481–488.
- [13] L. Hu, P.M. Bentler, Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification, *Psychol. Methods* 3 (1998) 424–453.
- [14] X. Huang, Statistics of Bivariate Extremes, in: Tinbergen Institute Research Series, vol. 22, Erasmus Universiteit Rotterdam, 1922 (Ph.D. thesis).
- [15] R. Huser, A.C. Davison, Composite likelihood estimation for the Brown–Resnick process, *Biometrika* 100 (2013) 511–518.
- [16] J. Hüsler, R.-D. Reiß, Maxima of normal random vectors: between independence and complete dependence, *Statist. Probab. Lett.* 7 (1989) 283–286.
- [17] H. Joe, *Dependence Modeling with Copulas*, Chapman & Hall/CRC, Boca Raton, FL, 2014.
- [18] H. Joe, H. Li, A.K. Nikoloulopoulos, Tail dependence functions and vine copulas, *J. Multivariate Anal.* 10 (2010) 252–270.
- [19] C.R. Johnson, Positive definite matrices, *Amer. Math. Monthly* 7 (1970) 259–264.
- [20] Z. Kabluchko, M. Schlather, L. de Haan, Stationary max-stable fields associated to negative definite functions, *Ann. Probab.* 37 (2009) 2042–2065.
- [21] P. Krupskii, R. Huser, M.G. Genton, Factor copula models for replicated spatial data, *J. Amer. Statist. Assoc.* (in press).
- [22] P. Krupskii, H. Joe, Factor copula models for multivariate data, *J. Multivariate Anal.* 20 (2013) 85–101.
- [23] P. Krupskii, H. Joe, Tail-weighted measures of dependence, *J. Appl. Stat.* 42 (2015) 614–629.
- [24] D. Kurowicka, R.M. Cooke, *Uncertainty Analysis with High Dimensional Dependence Modelling*, Wiley, New York, 2006.
- [25] D. Lee, H. Joe, Multivariate extreme value copulas with factor and tree dependence structures, *Extremes* (in press).
- [26] B.G. Lindsay, Composite likelihood methods, *Contemp. Math.* 80 (1988) 221–239.
- [27] A.K. Nikoloulopoulos, H. Joe, H. Li, Extreme value properties of multivariate *t* copulas, *Extremes* 12 (2009) 129–148.
- [28] M. Oesting, M. Ribatet, C. Dombry, Simulation of max-stable processes, in: D.K. Dipak, J. Yan (Eds.), *Extreme Value Modeling and Risk Analysis: Methods and Applications*, Chapman and Hall/CRC, Boca Raton, FL, 2015.
- [29] M. Schlather, Models for stationary max-stable random fields, *Extremes* 5 (2002) 33–44.
- [30] M. Sibuya, Bivariate extreme statistics, *Ann. Inst. Statist. Math.* 11 (1960) 195–210.
- [31] R.L. Smith, Max-stable processes and spatial extremes, Technical Report, 1990.
- [32] C. Varin, P. Vidoni, A note on composite likelihood inference and model selection, *Biometrika* 92 (2005) 519–528.