Bernoulli **21**(1), 2015, 38–61
DOI: 10.3150/13-BEJ560

Maxima of independent, non-identically distributed Gaussian vectors

SEBASTIAN ENGELKE[1,2], ZAKHAR KABLUCHKO[3]
and MARTIN SCHLATHER[4]

[1]*Institut für Mathematische Stochastik, Georg-August-Universität Göttingen, Goldschmidtstr. 7, D-37077 Göttingen, Germany. E-mail: sengelk@uni-goettingen.de*
[2]*Faculté des Hautes Etudes Commerciales, Université de Lausanne, Extranef, UNIL-Dorigny, CH-1015 Lausanne, Switzerland*
[3]*Institut für Stochastik, Universität Ulm, Helmholtzstr. 18, D-89069 Ulm, Germany.
E-mail: zakhar.kabluchko@uni-ulm.de*
[4]*Institut für Mathematik, Universität Mannheim, A5, 6, D-68131 Mannheim, Germany.
E-mail: schlather@math.uni-mannheim.de*

Let $X_{i,n}, n \in \mathbb{N}, 1 \le i \le n$, be a triangular array of independent $\mathbb{R}^d$-valued Gaussian random vectors with correlation matrices $\Sigma_{i,n}$. We give necessary conditions under which the row-wise maxima converge to some max-stable distribution which generalizes the class of Hüsler–Reiss distributions. In the bivariate case, the conditions will also be sufficient. Using these results, new models for bivariate extremes are derived explicitly. Moreover, we define a new class of stationary, max-stable processes as max-mixtures of Brown–Resnick processes. As an application, we show that these processes realize a large set of extremal correlation functions, a natural dependence measure for max-stable processes. This set includes all functions $\psi(\sqrt{\gamma(h)}), h \in \mathbb{R}^d$, where ψ is a completely monotone function and γ is an arbitrary variogram.

Keywords: extremal correlation function; Gaussian random vectors; Hüsler–Reiss distributions; max-limit theorems; max-stable distributions; triangular arrays

1. Introduction

It is well known that the standard normal distribution Φ is in the max-domain of attraction of the Gumbel distribution, that is,

$$\lim_{n\to\infty} \Phi(b_n + x/b_n)^n = \exp\bigl(-\exp(-x)\bigr), \qquad \text{for all } x \in \mathbb{R},$$

where $b_n, n \in \mathbb{N}$, is a sequence of normalizing constants defined by $b_n = n\phi(b_n)$, where ϕ is the standard normal density. By Theorem 1.5.3 in Leadbetter *et al.* [21], it is given as

$$b_n := \sqrt{2\log n} - \frac{(1/2)\log\log n + \log(2\sqrt{\pi})}{\sqrt{2\log n}} + \mathrm{o}\bigl((\log n)^{-1/2}\bigr). \tag{1}$$

Sibuya [25] showed that the maxima of i.i.d. bivariate normal random vectors with correlation $\rho < 1$ asymptotically always become independent. However, for triangular arrays with i.i.d. entries within each row where the correlation in the nth row approaches 1, as $n \to \infty$, with an

1350-7265

appropriate speed, Hüsler and Reiss [17] proved that the row-wise maxima converge to a new class of max-stable bivariate distributions, namely

$$F_\lambda(x, y) = \exp\left[-\Phi\left(\lambda + \frac{x-y}{2\lambda}\right)e^{-y} - \Phi\left(\lambda + \frac{y-x}{2\lambda}\right)e^{-x}\right], \qquad x, y \in \mathbb{R}. \tag{2}$$

Here, $\lambda \in [0, \infty]$ parameterizes the dependence in the limit, 0 and ∞ corresponding to complete dependence and asymptotic independence, respectively. In fact, Kabluchko *et al.* [20] provide a simple argument that these are also the only possible limit points for such triangular arrays.

More generally, Hüsler and Reiss [17] consider triangular arrays with i.i.d. entries of d-variate zero-mean, unit-variance normal random vectors with correlation matrix Σ_n in the nth row satisfying

$$\lim_{n\to\infty} \log n\big(\mathbf{1}\mathbf{1}^\top - \Sigma_n\big) = \Lambda \in [0, \infty)^{d\times d}, \tag{3}$$

where $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^d$ and $^\top$ denotes the transpose sign. Under this assumption, the row-wise maxima converge to the d-variate, max-stable Hüsler–Reiss distribution whose dependence structure is fully characterized by the matrix Λ. Note that condition (3) implies that all off-diagonal entries of Σ_n converge to 1 as $n \to \infty$. A slightly more general representation is given in Kabluchko [19] in terms of Poisson point processes and negative definite kernels.

In fact, it turns out that these distributions not only attract Gaussian arrays but also classes of related distributions. For instance, Hashorva [13] shows, that the convergence of maxima holds for triangular arrays of general bivariate elliptical distributions, if the random radius is in the domain of attraction of the Gumbel distribution. The generalization to multivariate elliptical distributions can be found in Hashorva [14]. Moreover, Hashorva *et al.* [15] prove that also some non-elliptical distributions are in the domain of attraction of the Hüsler–Reiss distribution, for instance multivariate χ^2-distributions.

Apart from being one of the few known parametric families of multivariate extreme value distributions, the Hüsler–Reiss distributions play a prominent role in modeling spatial extremes since they are the finite-dimensional distributions of Brown–Resnick processes [6,20].

Recently, Hashorva and Weng [16] analyzed maxima of stationary Gaussian triangular arrays where the variables in each row are identically distributed but not necessarily independent. They show that weak dependence is asymptotically negligible, whereas stronger dependence may influence the max-limit distribution.

In this paper, we consider independent triangular arrays $\mathbf{X}_{i,n} = (X_{i,n}^{(1)}, \dots, X_{i,n}^{(d)})$, $n \in \mathbb{N}$ and $1 \le i \le n$, where $\mathbf{X}_{i,n}$ is a zero-mean, unit-variance Gaussian random vector with correlation matrix $\Sigma_{i,n}$. Thus, in each row the random variables are independent, but may have different dependence structures. Letting $\mathbf{M}_n = (M_n^{(1)}, \dots, M_n^{(d)})$ denote the vector consisting of the componentwise maxima $M_n^{(j)} = \max_{i=1,\dots,n} X_{i,n}^{(j)}$, $j \in \{1, \dots, d\}$, we are interested in the convergence of the rescaled, row-wise maximum

$$b_n(\mathbf{M}_n - b_n), \tag{4}$$

as $n \to \infty$, and the respective limit distributions.

In Section 2, we start with bivariate triangular arrays. For this purpose, we introduce a sequence of counting measures which capture the dependence structure in each row and which is used to state necessary and sufficient conditions for the convergence of (4). Moreover, the limits turn out to be new max-stable distributions that generalize (2). The results on triangular arrays are used to completely characterize the max-limits of independent, but not necessarily identically distributed sequences of bivariate Gaussian vectors. Explicit examples for the bivariate limit distributions are given at the end of Section 2. The multivariate case is treated in Section 3, giving rise to a class of d-dimensional max-stable distributions. In Section 4, we show how these distributions arise as finite-dimensional margins of the new class of max-mixtures of Brown–Resnick processes. Furthermore, it is shown that these processes offer a large variety of extremal correlation functions which makes them interesting for modeling dependencies in spatial extremes. Finally, Section 5 comprises the proofs of the main theorems.

2. The bivariate case

Before we start with bivariate triangular arrays, let us note that even the case of univariate sequences of independent yet non-identically distributed Gaussian random variables is not trivial. In fact, many different distributions for the max-limits may arise, which are not necessarily max-stable (see Example 2 below). In the sequel, we will therefore restrict to the case that the variances of the univariate margins are close to some constant, which can be assumed to be 1 without loss of generality, and we fix the normalization in (4). Later, for the sake of simplicity, we will always consider margins with unit variance.

In order to state the main results in the bivariate case, we need probability measures on the extended positive half-line $[0,\infty]$. Endowed with the metric $d(x,y) = |\mathrm{e}^{-x} - \mathrm{e}^{-y}|$, the space $[0,\infty]$ becomes compact. A function $g:[0,\infty] \to \mathbb{R}$ is continuous iff it is continuous in the usual topology on $[0,\infty)$ and the limit $\lim_{x\to\infty} g(x)$ exists and equals $g(\infty)$.

2.1. Limit theorems

Consider a triangular array of independent bivariate Gaussian random vectors $\mathbf{X}_{i,n} = (X_{i,n}^{(1)}, X_{i,n}^{(2)})$, $n \in \mathbb{N}$ and $1 \le i \le n$, with zero expectation and covariance matrix

$$\operatorname{Cov}(\mathbf{X}_{i,n}) = \begin{pmatrix} \sigma_{i,n,1}^2 & \sigma_{i,n,1,2} \\ \sigma_{i,n,1,2} & \sigma_{i,n,2}^2 \end{pmatrix},$$

with $\sigma_{i,n,j}^2 > 0$ for all $n \in \mathbb{N}$, $1 \le i \le n$ and $j \in \{1,2\}$. Further, denote by $\rho_{i,n} = \sigma_{i,n,1,2}/(\sigma_{i,n,1}\sigma_{i,n,2})$ the correlation of $\mathbf{X}_{i,n}$. For $n \in \mathbb{N}$, we define a probability measure η_n on $[0,\infty] \times \mathbb{R}^2$ by

$$\eta_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(\sqrt{b_n^2(1-\rho_{i,n})/2},\, b_n^2(1-1/\sigma_{i,n,1}),\, b_n^2(1-1/\sigma_{i,n,2}))} \tag{5}$$

which encodes the suitably normalized variances and correlations in the nth row. More precisely, it maps the rate with which the variances and correlations converge to 1. Here, for any measurable space $(S, \mathcal{S})$ and $a \in S$, δ_a denotes the Dirac measure on the point a.

In this general situation, the next theorem gives a sufficient condition in terms of η_n for the convergence of row-wise maxima of this triangular array.

Theorem 2.1. *For $n \in \mathbb{N}$ and $1 \le i \le n$, let $\mathbf{X}_{i,n}$ and η_n be defined as above. Further suppose that for some $\varepsilon > 0$ the measures $(\eta_n)_{n\in\mathbb{N}}$ satisfy the integrability condition*

$$\sup_{n\in\mathbb{N}} \int_{[0,\infty]\times\mathbb{R}^2} \left[\mathrm{e}^{\theta(1+\varepsilon)} + \mathrm{e}^{\gamma(1+\varepsilon)}\right] \eta_n\big(\mathrm{d}(\lambda,\theta,\gamma)\big) < \infty. \tag{6}$$

If for $n \to \infty$, η_n converges weakly to some probability measure η on $[0,\infty] \times \mathbb{R}^2$, that is, $\eta_n \Rightarrow \eta$, then

$$\max_{i=1,\dots,n} b_n(\mathbf{X}_{i,n} - b_n) \tag{7}$$

converges in distribution to a random vector with distribution function F_η given by

$$\begin{aligned} -\log F_\eta(x,y) = \int_{[0,\infty]\times\mathbb{R}^2} & \Phi\left(\lambda + \frac{y - x + \theta - \gamma}{2\lambda}\right)\mathrm{e}^{-(x-\theta)} \\ & + \Phi\left(\lambda - \frac{y - x + \theta - \gamma}{2\lambda}\right)\mathrm{e}^{-(y-\gamma)} \eta\big(\mathrm{d}(\lambda,\theta,\gamma)\big), \end{aligned} \tag{8}$$

for $x, y \in \mathbb{R}$.

Remark 2.2. Condition (6) implies

$$\sup_{n\in\mathbb{N}, 1\le i\le n} \frac{1}{n}\left(\mathrm{e}^{b_n^2(1-1/\sigma_{i,n,1})(1+\varepsilon)} + \mathrm{e}^{b_n^2(1-1/\sigma_{i,n,2})(1+\varepsilon)}\right) < \infty.$$

Since $b_n^2 \sim 2\log n$ for n large, it follows that the variances of both components are uniformly bounded. Thus, the single random variables in each row satisfy the uniform asymptotical negligibility condition (see, for instance, [1])

$$\max_{i=1,\dots,n} \mathbb{P}\big(b_n\big(X_{i,n}^{(j)} - b_n\big) > x\big) \to 0, \qquad n \to \infty, \tag{9}$$

for $j = 1, 2$ and any $x \in \mathbb{R}$.

Remark 2.3. In fact, one can extend the distribution F_η to mixture measures η taking infinite mass at negative infinity. The only condition which needs to be satisfied is

$$\int_{[0,\infty]\times\mathbb{R}^2} \left[\mathrm{e}^{\theta} + \mathrm{e}^{\gamma}\right] \eta\big(\mathrm{d}(\lambda,\theta,\gamma)\big) < \infty.$$

Remark 2.4. Random variables with variances bounded away from 1 from above do not influence the maximum in the limit of (7). This can easily be seen by allowing weak convergence of η_n to η on the extended space $[0,\infty] \times [-\infty,\infty)^2$.

Note that the one-dimensional marginals of F_η are Gumbel distributed with certain location parameters, for instance,

$$-\log F_\eta(x,\infty) = \exp\left[-x + \log \int_{[0,\infty]\times\mathbb{R}^2} \mathrm{e}^{\theta}\eta\big(\mathrm{d}(\lambda,\theta,\gamma)\big)\right].$$

Moreover, F_η is a max-stable distribution since

$$F_\eta^n(x+\log n, y+\log n) = F_\eta(x,y),$$

for all $n \in \mathbb{N}$. This is a remarkable fact, since, in general, limits of row-wise maxima of triangular arrays are not max-stable, not even if the random variables in each row are identically distributed.

The idea of constructing extreme value distributions as in (8) is not new. Indeed, it is well known that any mixture of spectral measures is again a spectral measure. In our case, however, these mixture distributions also arise naturally as the max-limits of independent Gaussian triangular arrays.

If we assume that the margins have variance 1, that is, $\sigma_{i,n,1} = \sigma_{i,n,2} = 1$, we can obtain a necessary and sufficient condition for the convergence of maxima. We denote by $\mathcal{M}_1([0,\infty])$ the space of all probability measures on $[0,\infty]$ endowed with the topology of weak convergence. By Helly's theorem, this space is compact.

Theorem 2.5. *Consider a triangular array of independent bivariate Gaussian random vectors* $\mathbf{X}_{i,n} = (X_{i,n}^{(1)}, X_{i,n}^{(2)})$, $n \in \mathbb{N}$ *and* $1 \le i \le n$, *where* $X_{i,n}^{(1)}$ *and* $X_{i,n}^{(2)}$ *are standard normal random variables. Denote by* $\rho_{i,n}$ *the correlation of* $\mathbf{X}_{i,n}$. *Let*

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\sqrt{b_n^2(1-\rho_{i,n})/2}} \tag{10}$$

be a probability measure on $[0,\infty]$. *For* $n \to \infty$,

$$\max_{i=1,\dots,n} b_n(\mathbf{X}_{i,n} - b_n) \tag{11}$$

converges in distribution if and only if ν_n *converges weakly to some probability measure* ν *on* $[0,\infty]$, *that is*, $\nu_n \Rightarrow \nu$. *In this case, the limit of* (11) *has distribution function* F_ν *given by*

$$-\log F_\nu(x,y) = \int_0^\infty \left[\Phi\left(\lambda + \frac{y-x}{2\lambda}\right)\mathrm{e}^{-x} + \Phi\left(\lambda + \frac{x-y}{2\lambda}\right)\mathrm{e}^{-y}\right]\nu(\mathrm{d}\lambda), \tag{12}$$

$x, y \in \mathbb{R}$. *The distribution in* (12) *uniquely determines the measure* ν, *that is, for two probability measures* $\nu, \tilde{\nu} \in \mathcal{M}_1([0,\infty])$ *with* $\nu \ne \tilde{\nu}$ *it follows that* $F_\nu \ne F_{\tilde{\nu}}$. *Furthermore,* F_ν *depends continuously on* ν, *in the sense that if* $\nu_n \Rightarrow \nu$, *as* $n \to \infty$, *and* $\nu_n, \nu \in \mathcal{M}_1([0,\infty])$, *then* F_{ν_n} *converges pointwise to* F_ν.

Remark 2.6. If ν is a probability measure on $[0,\infty)$, an alternative construction of the distribution F_ν is the following [19], Section 3: Let $\sum_{i=1}^{\infty}\delta_{U_i}$ be a Poisson point process on $\mathbb{R}$ with intensity $\mathrm{e}^{-u}\,\mathrm{d}u$ and suppose that B has the normal distribution $N(-2S^2, 4S^2)$ with random mean and variance, where S is ν-distributed. Then, for a sequence $(B_i)_{i\in\mathbb{N}}$ of i.i.d. copies of B, the bivariate random vector $\max_{i\in\mathbb{N}}(U_i, U_i + B_i)$ has distribution F_ν.

Example 1. For an arbitrary probability measure $\nu \in \mathcal{M}_1([0,\infty])$, let $(R_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. samples of ν. Putting $\rho_{i,n} = \max(1 - 2R_i^2/b_n^2, -1)$ in Theorem 2.5 yields

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{\min(R_i,b_n)} \Rightarrow \nu, \qquad \text{a.s.},$$

by the law of large numbers. Hence, (11) converges a.s. in distribution to F_ν.

The above theorem can be applied to completely characterize the distribution of the maxima of a *sequence* of independent, but not necessarily identically distributed bivariate Gaussian random vectors with unit variance.

Corollary 2.7. *Suppose that* $\mathbf{X}_i = (X_i^{(1)}, X_i^{(2)})$, $n \in \mathbb{N}$ *and* $1 \le i \le n$, *is a sequence of independent bivariate Gaussian random vectors where* $X_i^{(1)}$ *and* $X_i^{(2)}$ *are standard normal random variables. Denote by* ρ_i *the correlation of* $\mathbf{X}_i$ *and let*

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{\sqrt{b_n^2(1-\rho_i)/2}}$$

be a probability measure on $[0,\infty]$. *For* $n \to \infty$,

$$\max_{i=1,\dots,n} b_n(\mathbf{X}_i - b_n) \tag{13}$$

converges in distribution if and only if ν_n *converges weakly to some probability measure* ν *on* $[0,\infty]$. *In this case, the limit of* (13) *has distribution function* F_ν *as in* (12). *Furthermore, for all* $\nu \in \mathcal{M}_1([0,\infty])$, F_ν *is attained as a limit of* (13) *for a suitable sequence* $(\mathbf{X}_i)_{i\in\mathbb{N}}$.

Remark 2.8. It is worthwhile to note that, in general, the class of max-selfdecomposable distributions (cf. Mejzler [22], de Haan and Ferreira [8], Theorem 5.6.1), that is, the max-limits of *sequences* of independent (not necessarily identically distributed) random variables, is a proper subclass of max-infinitely-divisible distributions, that is, the max-limits of *triangular arrays* with i.i.d. random variables in each row. The latter coincides with the class of max-limits of *triangular arrays*, where the rows are merely independent but not identically distributed [1,10]. In the (bivariate) Gaussian case, the above shows that the max-limits of i.i.d. *triangular arrays*, namely the Hüsler–Reiss distributions in (2), are a proper subclass of max-limits of independent *triangular arrays*, namely the distributions in (12), which, on the other hand, coincide with the max-limits of independent *sequences*.

Example 2. The following example shows that without any assumptions on the variances, even the univariate case is not trivial. Let X_i, $i \in \mathbb{N}$, be a sequence standard normal distribution. Define the sequence of maxima

$$M_n = \max_{i=1,\dots,n} X_i/i, \qquad n \in \mathbb{N}.$$

Clearly, M_n converges almost surely to the non-degenerate random variable $\max_{i\in\mathbb{N}} X_i/i$, which however is not an extreme value distribution.

2.2. Examples

In multivariate extreme value theory, it is important to have flexible and tractable models for dependencies of extremal events. The max-stable distributions F_ν in Theorem 2.5 for $\nu \in \mathcal{M}_1([0,\infty])$ are max-mixtures of Hüsler–Reiss distributions with different dependency parameters. They constitute a large class of new bivariate max-stable distributions. We derive two of them explicitly by evaluating the integral in (12).

Example 3 (Rayleigh distributed ν***).*** The Rayleigh distribution has density

$$f_\sigma(\lambda) = \frac{\lambda}{\sigma^2} \mathrm{e}^{-\lambda^2/(2\sigma^2)}, \qquad \lambda \geq 0, \tag{14}$$

for $\sigma > 0$. Choosing the dependence parameter λ according to the Rayleigh distribution ν_σ, we obtain the bivariate distribution function

$$-\log F_{\nu_\sigma}(x,y) = \int_0^\infty \left[\Phi\left(\lambda + \frac{y-x}{2\lambda}\right)\mathrm{e}^{-x} + \Phi\left(\lambda + \frac{x-y}{2\lambda}\right)\mathrm{e}^{-y}\right] \frac{\lambda}{\sigma^2} \mathrm{e}^{-\lambda^2/(2\sigma^2)}\,\mathrm{d}\lambda, \tag{15}$$

for $x, y \in \mathbb{R}$. In order to evaluate this integral, we apply partial integration and use formulae 3.471.9 and 3.472.3 in Gradshteyn and Ryzhik [12]. Equation (15) then simplifies to

$$F_{\nu_\sigma}(x,y) = \exp\left[-\mathrm{e}^{-\min(x,y)} - \frac{1}{\eta}\mathrm{e}^{-(y+x)/2}\mathrm{e}^{-|y-x|\eta/2}\right], \qquad x, y \in \mathbb{R}, \tag{16}$$

where $\eta = \sqrt{1 + 1/\sigma^2} \in (1, \infty)$. Note that σ parameterizes the dependence of F_{ν_σ}. As σ goes to 0 (i.e., η goes to ∞) the margins become equal. On the other hand, as σ goes to ∞ (i.e., η goes to 1) the margins become completely independent. The corresponding Pickands' dependence function is given by

$$\begin{aligned} A_{\nu_\sigma}(\omega) &= -\log F_{\nu_\sigma}\bigl(-\log\omega, -\log(1-\omega)\bigr) \\ &= \max(\omega, 1-\omega) + \frac{1}{\eta}\sqrt{\omega(1-\omega)}\max\left\{\frac{\omega}{1-\omega}, \frac{1-\omega}{\omega}\right\}^{-\eta/2}, \qquad \omega \in [0,1]. \end{aligned}$$

Example 4 (Type-2 Gumbel distributed ν). The Type-2 Gumbel distribution has density

$$f_b(\lambda) = 2b\lambda^{-3}\mathrm{e}^{-b/\lambda^2}, \qquad \lambda \geq 0,$$

for $b > 0$. With similar arguments as for the Rayleigh distribution the distribution function F_{ν_b}, where ν_b has density f_b, is given by

$$F_{\nu_b}(x, y) = \exp\bigl[-\mathrm{e}^{-x} - \mathrm{e}^{-y} + \mathrm{e}^{-(y+x)/2}\mathrm{e}^{-\sqrt{((y-x)/2)^2+2b}}\bigr], \qquad x, y \in \mathbb{R}.$$

In this case, the parameter $b \in (0, \infty)$ interpolates between complete independence and complete dependence of the bivariate distribution. In particular, if $b \to 0$, then the margins are equal and, on the other hand, if $b \to \infty$ then the margins are independent. Here, Pickands' dependence function is

$$A_{\nu_b}(\omega) = 1 - \sqrt{\omega(1-\omega)}\exp\left(-\sqrt{\left(\frac{\log(\omega/(1-\omega))}{2}\right)^2 + 2b}\right), \qquad \omega \in [0, 1].$$

Every multivariate max-stable distribution admits a spectral representation [23], Chapter 5, where the spectral measure contains all information about the extremal dependence. Recently, Cooley *et al.* [7] and Ballani and Schlather [2] constructed new parametric models for spectral measures. For the bivariate Hüsler–Reiss distribution, de Haan and Pereira [9] give an explicit form of its spectral density on the positive sphere $S^1_+ = \{(x_1, x_2) \in [0, \infty)^2, x_1^2 + x_2^2 = 1\}$. More precisely, they show that for $\lambda \in (0, \infty)$

$$-\log F_\lambda(x, y) = \int_0^{\pi/2} \max\bigl\{\mathrm{e}^{-x}\sin\theta, \mathrm{e}^{-y}\cos\theta\bigr\} s_\lambda(\theta)\,\mathrm{d}\theta, \qquad x, y \in \mathbb{R},$$

and give a rather complicated expression for s_λ. Using the equation

$$\phi\left(\lambda - \frac{\log\tan\theta}{2\lambda}\right) = \frac{\sin\theta}{\cos\theta}\phi\left(\lambda + \frac{\log\tan\theta}{2\lambda}\right), \qquad \lambda \in (0, \infty), \theta \in [0, \pi/2],$$

their expression can be considerably simplified and the spectral density becomes

$$s_\lambda(\theta) = \frac{1}{2\lambda\sin\theta\cos^2\theta}\phi\left(\lambda + \frac{\log(\tan\theta)}{2\lambda}\right), \qquad \theta \in [0, \pi/2].$$

For the spectral density s_ν of the Hüsler–Reiss mixture distribution F_ν as in (12), where ν does neither have an atom at 0 nor at ∞, we have the relation

$$s_\nu(\theta) = \int_0^\infty s_\lambda(\theta)\nu(\mathrm{d}\lambda), \qquad \theta \in [0, \pi/2].$$

For the two examples above, we can compute the corresponding spectral densities.

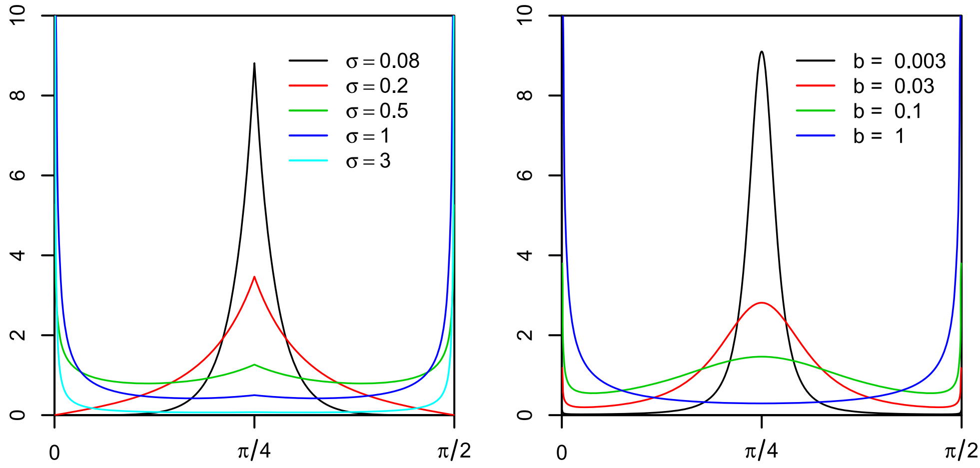

Figure 1. Spectral densities of the Rayleigh (left) and Type-2 Gumbel (right) mixture distribution for different parameters σ and b, respectively.

Example 3 (continued). For the Rayleigh distribution, s_{ν_σ} is given by

$$s_{\nu_\sigma}(\theta) = \frac{\mathrm{e}^{-(1/\sqrt{2})|\log\tan\theta|\sqrt{1+1/\sigma^2}}}{4\sqrt{\sigma^4+\sigma^2}(\sin\theta\cos\theta)^{3/2}}, \qquad \theta \in [0, \pi/2].$$

Example 4 (continued). For the Type-2 Gumbel distribution with parameter $b > 0$, the spectral density has the form

$$s_{\nu_b}(\theta) = \frac{\mathrm{e}^{-u_b(\theta)}}{4(\sin\theta\cos\theta)^{3/2}}\left(1 - \frac{(\log\tan\theta)^2}{4u_b(\theta)^2}\right)\left(1 + \frac{1}{u_b(\theta)}\right), \qquad \theta \in [0, \pi/2],$$

with $u_b(\theta) = \sqrt{(\log\tan\theta)^2/4 + 2b}$.

Figure 1 illustrates how these spectral measures interpolate between complete independence and complete dependence for different parameters.

3. The multivariate case

Similarly as in Hüsler and Reiss [17], the results for standard bivariate Gaussian random vectors can be generalized to d-dimensional random vectors. To this end, define a triangular array of independent d-dimensional Gaussian random vectors $\mathbf{X}_{i,n} = (X_{i,n}^{(1)}, \dots, X_{i,n}^{(d)})$, $n, d \in \mathbb{N}$ and $1 \le i \le n$, where $X_{i,n}^{(j)}$, $j \in \{1, \dots, d\}$, are standard normal random variables. Denote by $\Sigma_{i,n} =$

$(\rho_{j,k}(i,n))_{1\le j,k\le d}$ the correlation matrix of $\mathbf{X}_{i,n}$. Let $\mathbf{1} = (1,\ldots,1)^\top \in \mathbb{R}^d$ and

$$\eta_n = \frac{1}{n}\sum_{i=1}^n \delta_{\sqrt{b_n^2(\mathbf{1}\mathbf{1}^\top - \Sigma_{i,n})/2}} \tag{17}$$

be a probability measure on the metric space $[0,\infty)^{d\times d}$, equipped with the Euclidean distance. Throughout this paper, squares and square roots of matrices are to be understood componentwise. For a measure τ on $[0,\infty)^{d\times d}$, we will denote by τ^2 the image measure of τ under the transformation $[0,\infty)^{d\times d} \to [0,\infty)^{d\times d}$, $\Lambda \mapsto \Lambda^2$. Further, let $D \subset [0,\infty)^{d\times d}$ be the subspace of conditionally negative definite matrices which are symmetric and have zeros on the diagonal, that is,

$$D := \Bigg\{ (a_{j,k})_{1\le j,k\le d} = A \in [0,\infty)^{d\times d} \colon \mathbf{x}^\top A\mathbf{x} \le 0 \text{ for all } \mathbf{x} \in \mathbb{R}^d \setminus \{0\} \text{ s.t.} \\ \sum_{i=1}^d x_i = 0, a_{j,k} = a_{k,j}, a_{j,j} = 0 \text{ for all } 1 \le j,k \le d \Bigg\},$$

and let $D_0 \subset D$ be the space of strictly conditionally negative definite matrices, that is, where $\mathbf{x}^\top A\mathbf{x} < 0$ holds in the above definition. In particular, note that D_0 is open in D and that D is a suitable subspace for the measures η_n^2 since $\eta_n^2(D) = 1$ for all $n \in \mathbb{N}$. For $\Lambda = (\lambda_{j,k})_{1\le j,k\le d} \in [0,\infty)^{d\times d}$, define a family of transformed matrices by

$$\Gamma_{l,m}(\Lambda) = 2\big(\lambda^2_{m_j,m_l} + \lambda^2_{m_k,m_l} - \lambda^2_{m_j,m_k}\big)_{1\le j,k\le l-1},$$

where $2 \le l \le d$ and $m = (m_1,\ldots,m_l)$ with $1 \le m_1 < \cdots < m_l \le d$. It follows from the proof of Lemma 2.1 in Berg *et al.* [3] that if $\Lambda \in D_0$, then $\Gamma_{l,m}(\sqrt{\Lambda})$ is a (strictly) positive definite matrix.

Denote by $S(\cdot|\Psi)$ the so-called survivor function of an l-dimensional normal random vector with mean vector $\mathbf{0}$ and covariance matrix Ψ. That is, if $\mathbf{X} \sim N(\mathbf{0},\Psi)$ and $\mathbf{x} \in \mathbb{R}^l$, then $S(\mathbf{x}|\Psi) = \mathbb{P}(X_1 > x_1,\ldots,X_l > x_l)$. If Ψ is not a covariance matrix, we put $S(\mathbf{x}|\Psi) = 0$.

For a fixed $\Lambda = (\lambda_{j,k})_{1\le j,k\le d} \in [0,\infty)^{d\times d}$, let

$$H_\Lambda(\mathbf{x}) = \exp\Bigg(\sum_{l=1}^d (-1)^l \sum_{m:1\le m_1<\cdots<m_l\le d} h_{l,m,\Lambda}(x_{m_1},\ldots,x_{m_l})\Bigg),$$

where

$$h_{l,m,\Lambda}(y_1,\ldots,y_l) = \int_{y_l}^\infty S\big((y_i - z + 2\lambda^2_{m_i,m_l})_{i=1,\ldots,l-1}\big|\Gamma_{l,m}(\Lambda)\big)\mathrm{e}^{-z}\,\mathrm{d}z,$$

for $2 \le l \le d$ and $h_{1,m,\Lambda}(y) = \mathrm{e}^{-y}$ for $m = 1,\ldots,d$. For alternative representations of the multivariate Hüsler–Reiss distribution H_Λ, see Joe [18] and Kabluchko [19]. With this notation, we are now in a position to state the following theorem.

Theorem 3.1. *Consider a triangular array of independent d-dimensional Gaussian random vectors as above. If for $n \to \infty$ the measure η_n in* (17) *converges weakly to some probability measure η on $[0,\infty)^{d\times d}$, i.e., $\eta_n \Rightarrow \eta$, s.t. $\eta^2(D_0) = 1$, then*

$$\max_{i=1,\dots,n} b_n(\mathbf{X}_{i,n} - b_n)$$

converges in distribution to a random vector with distribution function H_η given by

$$H_\eta(x_1,\dots,x_d) = \exp\biggl(\int_{[0,\infty)^{d\times d}} \log H_\Lambda(x)\eta(\mathrm{d}\Lambda)\biggr), \qquad x \in \mathbb{R}^d. \tag{18}$$

Remark 3.2. Similarly to Remark 2.6, we can give an alternative construction of the distribution H_η in terms of Poisson point processes. Let $\sum_{i=1}^{\infty} \delta_{U_i}$ be a Poisson point process on $\mathbb{R}$ with intensity $\mathrm{e}^{-u}\,\mathrm{d}u$ and suppose that $\mathbf{B}$ has the multivariate normal distribution $N(-\operatorname{diag}(\Gamma_{d,(1,\dots,d)}(\Lambda))/2, \Gamma_{d,(1,\dots,d)}(\Lambda))$ with random mean and variance, where Λ is η-distributed. Then, for a sequence $(\mathbf{B}_i)_{i\in\mathbb{N}}$ of i.i.d. copies of $\mathbf{B}$, the random vector $\max_{i\in\mathbb{N}}(U_i, U_i + \mathbf{B}_i)$ has distribution H_η.

Remark 3.3. We believe that the above theorem also holds in the case when η has positive measure on *non-strictly* conditionally negative definite matrices, i.e., $\eta^2(D \setminus D_0) > 0$. Our proof of this theorem however breaks down in this situation such that another technique might be necessary.

Remark 3.4. It is an open question if, similarly to the bivariate case, the distribution H_η uniquely determines the mixture measure η. By Remark 3.2, this problem is equivalent to the question if the distribution of normal mixtures $N(-\operatorname{diag}(\Gamma_{d,(1,\dots,d)}(\Lambda))/2, \Gamma_{d,(1,\dots,d)}(\Lambda))$, where Λ is η-distributed, determines the measure η. The solution of this problem is crucial to show that in Theorem 3.1 the weak convergence $\eta_n \Rightarrow \eta$ is also necessary for the convergence of the maxima.

4. Application to Brown–Resnick processes

The d-dimensional Hüsler–Reiss distributions arise in the theory of maxima of Gaussian random fields as the finite-dimensional distributions of the Brown–Resnick process [6] and its generalizations [20]. In this section, we introduce a new class of max-stable processes with finite-dimensional distributions given by (18) for suitable measures η.

Let us briefly recall the definition of the processes introduced in Kabluchko *et al.* [20]. For a zero-mean Gaussian process $\{W(t), t \in \mathbb{R}^d\}$ with stationary increments, variance $\sigma^2(t)$ and variogram $\gamma(t) = \mathbb{E}(W(t) - W(0))^2$, consider i.i.d. copies $\{W_i, i \in \mathbb{N}\}$ of W and a Poisson point process $\sum_{i=1}^{\infty} \delta_{U_i}$ on $\mathbb{R}$ with intensity $\mathrm{e}^{-u}\,\mathrm{d}u$, independent of the family $W_i, i \in \mathbb{N}$. Kabluchko *et al.* [20] showed that the Brown–Resnick process

$$\xi(t) = \max_{i\in\mathbb{N}}\bigl(U_i + W_i(t) - \sigma(t)^2/2\bigr), \qquad t \in \mathbb{R}^d, \tag{19}$$

is max-stable and stationary with standard Gumbel margins and that its law depends only on the variogram γ.

We generalize this construction by allowing the variogram of the Gaussian processes W_i to be random. In fact, this defines a new class of max-stable processes whose finite-dimensional distributions are of the form (18).

Definition 4.1. *Let V_d be the measurable space of all variograms on $\mathbb{R}^d$, i.e., conditionally negative definite functions γ on $\mathbb{R}^d$ with $\gamma(0) = 0$, equipped with the product σ-algebra. Further, let $\mathbb{Q}$ be an arbitrary probability measure on this space and γ_i, $i \in \mathbb{N}$, be an i.i.d. sequence of random variables with distribution $\mathbb{Q}$. For each $i \in \mathbb{N}$, let W_i be a random field such that, conditionally on γ_i, W_i is a zero-mean Gaussian process with stationary increments, variogram $4\gamma_i$ and $W_i(0) = 0$ a.s. Further, let $\sum_{i=1}^{\infty} \delta_{U_i}$ be a Poisson point process on $\mathbb{R}$ with intensity $\mathrm{e}^{-u}\,\mathrm{d}u$ which is independent of the W_i, $i \in \mathbb{N}$. Then, the process $\xi_{\mathbb{Q}}$ given by*

$$\xi_{\mathbb{Q}}(t) = \max_{i\in\mathbb{N}} \bigl(U_i + W_i(t) - 2\gamma_i(t)\bigr), \qquad t \in \mathbb{R}^d,$$

is called a max-mixture of Brown–Resnick processes w.r.t. the mixture measure $\mathbb{Q}$.

Note that a different kind of process can be defined through a hierarchical or Bayesian approach, which is not considered here and which does not lead to a max-stable process, in general: first, exactly one realization of the variogram is drawn from $\mathbb{Q}$. Then, conditionally on this realization, a Brown–Resnick process is simulated. Obviously, the resulting process must lie in the max-domain of attraction of the process given in Definition 4.1, with the same law $\mathbb{Q}$ for the variograms. This implies immediately the following proposition; a direct proof is given in Section 5.

Proposition 4.2. *The process $\xi_{\mathbb{Q}}$ is max-stable and stationary and has finite-dimensional distributions given by* (18) *with η induced by $\mathbb{Q}$.*

This new class of processes thus also realize a large variety of extremal dependence structures, which can for instance be measured by the *extremal correlation function* ρ [24,26]. For a stationary, max-stable random field $\{X(t), t \in \mathbb{R}^d\}$ with Gumbel margins, ρ is a natural approach to measure bivariate extremal dependencies and for $h \in \mathbb{R}^d$ it is determined by

$$\mathbb{P}\bigl(X(0) \le x, X(h) \le x\bigr) = \mathbb{P}\bigl(X(0) \le x\bigr)^{2-\rho(h)},$$

for some (and hence all) $x \in \mathbb{R}$. For instance, for the process in (19) it is given by

$$\rho_\gamma(h) = 2\bigl(1 - \Phi\bigl(\sqrt{\gamma(h)}/2\bigr)\bigr), \qquad h \in \mathbb{R}^d.$$

The processes introduced in Definition 4.1 extend this class of extremal correlation functions. Indeed, for an arbitrary variogram γ and mixture measure ν on $(0, \infty)$, let the measure $\mathbb{Q}$ in

Definition 4.1 be the law of the scale mixture $S^2\gamma$, where S is ν-distributed. The corresponding process $\xi_{\mathbb{Q}}$ possesses the extremal correlation function

$$\rho_{\gamma,\nu}(h) = \int_0^\infty 2\bigl(1 - \Phi\bigl(s\sqrt{\gamma(h)}\bigr)\bigr)\nu(\mathrm{d}s), \qquad h \in \mathbb{R}^d. \tag{20}$$

Moreover, from the construction it is obvious that processes with this dependence structure can be simulated easily as max-mixtures of Brown–Resnick processes. Gneiting [11] analyzes this kind of scale mixtures of the complementary error function in a more general framework. The following corollary is a consequence of Theorems 3.7 and 3.8 therein.

Corollary 4.3. *For a fixed variogram* γ *the class of extremal correlation functions in* (20) *is given by all functions* $\varphi(\sqrt{\gamma(h)})$, $h \in \mathbb{R}^d$, *where* $\varphi : [0, \infty) \to \mathbb{R}$ *is a continuous function with* $\varphi(0) = 1$, $\lim_{h\to\infty} \varphi(h) = 0$, *and the function*

$$(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}h^k}\bigl[-\varphi'(\sqrt{h})\bigr] \tag{21}$$

is nonnegative for infinitely many positive integers k, *i.e.*, $-\varphi'(\sqrt{h})$ *is completely monotone* (*cf. the paragraph after Theorem* 3.8 *in Gneiting* [11]).

For instance, if ν_1 is the Rayleigh distribution (14) with density f_1, we obtain

$$\rho_{\gamma,\nu_1}(h) = 2\left(1 - \int_0^\infty \Phi(\lambda) f_{\sqrt{\gamma(h)}}(\lambda)\,\mathrm{d}\lambda\right) = 1 - \left(\frac{\gamma(h)}{\gamma(h)+1}\right)^{1/2}, \qquad h \in \mathbb{R}^d,$$

immediately from equation (16). In fact, $\rho_{\gamma,\nu_1}(h) = \psi(\gamma(h))$, where $\psi(x) = 1 - (x/(x+1))^{1/2}$ is a completely monotone member of the Dagum family [4]. However, it is interesting to note that when writing $\rho_{\gamma,\nu_1}(h) = \varphi(\sqrt{\gamma(h)})$ with $\varphi(x) = 1 - (x^2/(x^2+1))^{1/2}$ as in Corollary 4.3, the function φ merely satisfies (21) but is not completely monotone.

Similarly, for the Type-2 Gumbel distribution with $b = 1$, the extremal correlation function is given by $\rho(h) = \exp(-\sqrt{2\gamma(h)})$. In particular, it follows that for any variogram γ and any $r > 0$ the function

$$\rho(h) = \exp\bigl(-r\sqrt{\gamma(h)}\bigr), \qquad h \in \mathbb{R}^d,$$

is an extremal correlation function. Since this class of extremal correlation functions is closed under the operation of mixing with respect to probability measures, this implies that for any measure $\mu \in \mathcal{M}_1((0,\infty))$ the Laplace transform $\mathcal{L}\mu$ yields an extremal correlation function

$$\rho_\mu(h) = \mathcal{L}\mu\bigl(\sqrt{\gamma(h)}\bigr) = \int_0^\infty \mathrm{e}^{-r\sqrt{\gamma(h)}}\mu(\mathrm{d}r), \qquad h \in \mathbb{R}^d.$$

Equivalently, for any completely monotone function ψ with $\psi(0) = 1$, the function $\psi(\sqrt{\gamma(h)})$ is an extremal correlation function. A corresponding max-stable, stationary random field is given by a max-mixture of Brown–Resnick processes with suitable $\nu \in \mathcal{M}_1((0,\infty))$.

5. Proofs

Proof of Theorem 2.1. Let $x, y \in \mathbb{R}$ and put $u_n(z) = b_n + z/b_n$, for $z \in \mathbb{R}$.

$$\begin{aligned}
&\log \mathbb{P}\Big(\max_{i=1,\dots,n} X_{i,n}^{(1)} \le u_n(x), \max_{i=1,\dots,n} X_{i,n}^{(2)} \le u_n(y)\Big) \\
&\quad = \sum_{i=1}^{n} \log\big(1 - \big[\mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) + \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big) - \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x), X_{i,n}^{(2)} > u_n(y)\big)\big]\big) \\
&\quad = -\sum_{i=1}^{n} \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) - \sum_{i=1}^{n} \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big) \\
&\qquad + \sum_{i=1}^{n} \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x), X_{i,n}^{(2)} > u_n(y)\big) + R_n,
\end{aligned} \tag{22}$$

where R_n is a remainder term from the Taylor expansion of $\log(1-z) = -z - z^2/2 + \mathrm{o}(z^2)$, as $z \to 0$. Thus, by (9) there is an $n_0 \in \mathbb{N}$ s.t. for all $n \ge n_0$ we have

$$\begin{aligned}
|R_n| &\le \sum_{i=1}^{n} \big[\mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) + \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big)\big]^2 \\
&\le \max_{i=1,\dots,n} \big[\mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) + \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big)\big] \\
&\quad \cdot \sum_{i=1}^{n} \big[\mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) + \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big)\big].
\end{aligned} \tag{23}$$

For the one-dimensional margins, we observe

$$\begin{aligned}
-\sum_{i=1}^{n} \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) &= -\sum_{i=1}^{n} \int_{u_n(x)/\sigma_{i,n,1}}^{\infty} \phi(z)\,\mathrm{d}z \\
&= -\sum_{i=1}^{n} \int_{x/\sigma_{i,n,1} - b_n^2(1-1/\sigma_{i,n,1})}^{\infty} \frac{1}{b_n} \phi\big(u_n(z)\big)\,\mathrm{d}z \\
&= -\int_{[0,\infty]\times\mathbb{R}^2} \int_{(1-\theta/b_n^2)x-\theta}^{\infty} \mathrm{e}^{-z-z^2/(2b_n^2)}\,\mathrm{d}z\,\eta_n\big(\mathrm{d}(\lambda,\theta,\gamma)\big),
\end{aligned}$$

where for the last equation we used $b_n = n\phi(b_n)$ and the definition of the measure η_n in (5) to replace the sum by the integral. For $n \in \mathbb{N}$, let

$$h_n(\theta) = \int_{(1-\theta/b_n^2)x-\theta}^{\infty} \mathrm{e}^{-z-z^2/(2b_n^2)}\,\mathrm{d}z, \qquad \theta \in \mathbb{R}.$$

Clearly, as $n \to \infty$, h_n converges uniformly on compact sets to the function $h(\theta) = \exp(\theta - x)$. Note that h and h_n are continuous functions on $\mathbb{R}$. Put $\omega = (\lambda, \theta, \gamma)$ and observe for $K > 0$ that

$$\begin{aligned}
&\left|\int_{[0,\infty]\times\mathbb{R}^2} h_n(\theta)\eta_n(\mathrm{d}\omega) - \int_{[0,\infty]\times\mathbb{R}^2} h(\theta)\eta(\mathrm{d}\omega)\right| \\
&\quad\le \left|\int_{[0,\infty]\times\mathbb{R}^2} h_n(\theta)\mathbf{1}_{h_n>K}\eta_n(\mathrm{d}\omega) - \int_{[0,\infty]\times\mathbb{R}^2} h(\theta)\mathbf{1}_{h>K}\eta(\mathrm{d}\omega)\right| \qquad (24)\\
&\qquad + \left|\int_{[0,\infty]\times\mathbb{R}^2} h_n(\theta)\mathbf{1}_{h_n<K}\eta_n(\mathrm{d}\omega) - \int_{[0,\infty]\times\mathbb{R}^2} h(\theta)\mathbf{1}_{h<K}\eta(\mathrm{d}\omega)\right|.
\end{aligned}$$

By Theorem 5.5 in Billingsley [5] (see also the remark after the theorem), $\eta_n h_n^{-1}$ converges weakly to ηh^{-1}. Moreover, since $h\mathbf{1}_{h<K}$ and the $h_n\mathbf{1}_{h_n<K}$ are uniformly bounded in n, the second summand in (24) converges to 0 as $n \to \infty$, for arbitrary $K > 0$. By the integrability condition (6) and Fatou's lemma, we have $\int_{[0,\infty]\times\mathbb{R}^2} h(\theta)\eta(\mathrm{d}\omega) < \infty$ and hence, also the first summand in (24) tends to zero as $K, n \to \infty$. Consequently,

$$-\sum_{i=1}^n \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x)\big) \to -\int_{[0,\infty]\times\mathbb{R}^2} \exp\big[-(x-\theta)\big]\eta(\mathrm{d}\omega). \qquad (25)$$

Similarly, we get

$$-\sum_{i=1}^n \mathbb{P}\big(X_{i,n}^{(2)} > u_n(y)\big) \to -\int_{[0,\infty]\times\mathbb{R}^2} \exp\big[-(y-\gamma)\big]\eta(\mathrm{d}\omega). \qquad (26)$$

It now also follows from (9), (23), (25) and (26) that the remainder term R_n converges to zero as $n \to \infty$.

We now turn to the third term in (22).

$$\begin{aligned}
&\sum_{i=1}^n \mathbb{P}\big(X_{i,n}^{(1)}/\sigma_{i,n,1} > u_n(x)/\sigma_{i,n,1}, X_{i,n}^{(2)}/\sigma_{i,n,2} > u_n(y)/\sigma_{i,n,2}\big) \\
&\quad = \sum_{i=1}^n \int_{u_n(y)/\sigma_{i,n,2}}^\infty \left[1 - \Phi\left(\frac{u_n(x)/\sigma_{i,n,1} - \rho_{i,n}z}{(1-\rho_{i,n}^2)^{1/2}}\right)\right]\phi(z)\,\mathrm{d}z \\
&\quad = \frac{1}{n}\sum_{i=1}^n \int_{y/\sigma_{i,n,2} - b_n^2(1-1/\sigma_{i,n,2})}^\infty \left[1 - \Phi\left(\frac{u_n(x)/\sigma_{i,n,1} - \rho_{i,n}u_n(z)}{(1-\rho_{i,n}^2)^{1/2}}\right)\right]\mathrm{e}^{-z-z^2/(2b_n^2)}\,\mathrm{d}z \\
&\quad = \int_{[0,\infty]\times\mathbb{R}^2} \int_{(1-\gamma/b_n^2)y-\gamma}^\infty \big[1 - \Phi\big(s_n(\lambda,\theta,z,x)\big)\big]\mathrm{e}^{-z-z^2/(2b_n^2)}\,\mathrm{d}z\,\eta_n(\mathrm{d}\omega),
\end{aligned}$$

where we used $b_n = n\phi(b_n)$ for the second last equation and s_n is defined by

$$s_n(\lambda,\theta,z,x) := \frac{\lambda}{(1-\lambda^2/b_n^2)^{1/2}} + \frac{(1-\theta/b_n^2)x - z - \theta}{(1-\lambda^2/b_n^2)^{1/2}2\lambda} + \frac{\lambda z}{(1-\lambda^2/b_n^2)^{1/2}b_n^2}.$$

For the last equation, we replaced the sum by the integral w.r.t. the empirical measure η_n as in (5). Note that for $i \in \{1, \dots, n\}$, in fact a short computation yields

$$s_n\Big(\sqrt{b_n^2(1-\rho_{i,n})/2}, b_n^2(1-1/\sigma_{i,n,1}), z, x\Big) = \frac{u_n(x)/\sigma_{i,n,1} - \rho_{i,n}u_n(z)}{(1-\rho_{i,n}^2)^{1/2}}.$$

For $n \in \mathbb{N}$, let

$$g_n(\lambda, \theta, \gamma) = \mathbf{1}_{\lambda \le b_n} \int_{(1-\gamma/b_n^2)y-\gamma}^{\infty} \big[1 - \Phi\big(s_n(\lambda, \theta, z, x)\big)\big] \mathrm{e}^{-z-z^2/(2b_n^2)}\, \mathrm{d}z$$

be a measurable function on $[0, \infty] \times \mathbb{R}^2$. It is easy to see, that as $n \to \infty$, g_n converges pointwise to the function

$$g(\lambda, \theta, \gamma) = \int_{y-\gamma}^{\infty} \big[1 - \Phi\big(s(\lambda, \theta, z, x)\big)\big] \mathrm{e}^{-z}\, \mathrm{d}z,$$

with

$$s(\lambda, \theta, z, x) := \lambda + \frac{x - z - \theta}{2\lambda}.$$

Note that g is a continuous function on $[0, \infty] \times \mathbb{R}^2$ and $g(0, \theta, \gamma) = g_n(0, \theta, \gamma) = \exp(-\max(x-\theta, y-\gamma))$ and $g(\infty, \theta, \gamma) = g_n(\infty, \theta, \gamma) = 0$, for any $(\theta, \gamma) \in \mathbb{R}^2$ and n sufficiently large. Here, the values are understood as the limits as $\lambda \to 0$ and $\lambda \to \infty$ (using dominated convergence), respectively, for example, $\lim_{\lambda \to 0} g(\lambda, \theta, \gamma) = \int_{y-\gamma}^{\infty} \mathbf{1}_{z > x-\theta} \mathrm{e}^{-z}\, \mathrm{d}z = \exp(-\max(x-\theta, y-\gamma))$. In order to establish the weak convergence $\eta_n g_n^{-1} \Rightarrow \eta g^{-1}$, we show that g_n converges uniformly on compact sets to g as $n \to \infty$. To this end, let $C = [0, \infty] \times [\theta_0, \theta_1] \times [\gamma_0, \gamma_1]$ be an arbitrary compact set in $[0, \infty] \times \mathbb{R}^2$ and let $\varepsilon > 0$ be given. First, note that instead of g_n it suffices to consider the function $\tilde{g}_n$, defined as

$$\tilde{g}_n(\lambda, \theta, \gamma) = \mathbf{1}_{\lambda \le b_n} \int_{(1-\gamma/b_n^2)y-\gamma}^{\infty} \big[1 - \Phi\big(s_n(\lambda, \theta, z, x)\big)\big] \mathrm{e}^{-z}\, \mathrm{d}z,$$

since for n large enough

$$\sup_{(\lambda, \theta, \gamma) \in C} |g_n(\lambda, \theta, \gamma) - \tilde{g}_n(\lambda, \theta, \gamma)| \le \mathbf{1}_{\lambda \le b_n} \int_{-2|y|-\gamma_1}^{\infty} \mathrm{e}^{-z}\big(1 - \mathrm{e}^{-z^2/(2b_n^2)}\big)\, \mathrm{d}z \to 0,$$

as $n \to \infty$, by dominated convergence. Further, for any $\varepsilon > 0$, let $z_1 > -\log \varepsilon$ which implies $\int_{z_1}^{\infty} \mathrm{e}^{-z}\, \mathrm{d}z < \varepsilon$. We note that for n large enough

$$\begin{aligned} s_n(\lambda, \theta, z, x) &\ge \big(1 - \lambda^2/b_n^2\big)^{-1/2}\bigg(\lambda\bigg(1 + \frac{-2|y| - \gamma_1}{b_n^2}\bigg) + \frac{-2|x| - z_1 - \theta_1}{2\lambda}\bigg) \\ &\ge \bigg(\frac{\lambda}{2} + \frac{-2|x| - z_1 - \theta_1}{2\lambda}\bigg), \end{aligned}$$

for all $\lambda \le b_n$, $-2|y| - \gamma_1 \le z \le z_1$ and $(\lambda, \theta, \gamma) \in C$, independently of $n \in \mathbb{N}$. Hence, there is a $\lambda_1 > 0$ s.t. for all $\lambda_1 \le \lambda \le b_n$

$$1 - \Phi\bigl(s_n(\lambda, \theta, z, x)\bigr) < \varepsilon \mathrm{e}^{-2|y|-\gamma_1}.$$

Thus, for all $n \in \mathbb{N}$ large enough,

$$\sup_{(\lambda,\theta,\gamma)\in C, \lambda\ge\lambda_1} \tilde{g}_n(\lambda, \theta, \gamma) \le \mathbf{1}_{\lambda\le b_n}\left(\int_{-2|y|-\gamma_1}^{z_1} \varepsilon \mathrm{e}^{-2|y|-\gamma_1}\mathrm{e}^{-z}\,\mathrm{d}z + \int_{z_1}^{\infty} \mathrm{e}^{-z}\,\mathrm{d}z\right) \le 2\varepsilon,$$

and in the same manner, $\sup_{(\lambda,\theta,\gamma)\in C,\lambda\ge\lambda_1} g(\lambda, \theta, \gamma) \le 2\varepsilon$. Furthermore, we observe

$$\lim_{\lambda\to 0} \Phi\bigl(s_n(\lambda, \theta, z, x)\bigr) = \mathbf{1}_{z<(1-\theta/b_n^2)x-\theta} \quad \text{and} \quad \lim_{\lambda\to 0} \Phi\bigl(s(\lambda, \theta, z, x)\bigr) = \mathbf{1}_{z<x-\theta}.$$

Choose $n_0 \in \mathbb{N}$ such that for all $n > n_0$ and all $\theta \in [\theta_0, \theta_1]$ we find an open interval (a_θ, b_θ) of size $\varepsilon/2$ that contains $\{(1 - \theta/b_n^2)x - \theta, x - \theta\}$. Put $I_\theta = (a_\theta - \varepsilon/4, b_\theta + \varepsilon/4)$, then we find a $\lambda_0 > 0$, s.t. for all $(\lambda, \theta, \gamma) \in C$, $\lambda \le \lambda_0$, $z \in I_\theta$ and $n > n_0$, we have $|\Phi(s_n(\lambda, \theta, z, x)) - \Phi(s(\lambda, \theta, z, x))| \le \varepsilon$. Consequently,

$$\begin{aligned}
&\sup_{(\lambda,\theta,\gamma)\in C,\lambda\le\lambda_0} \bigl|\tilde{g}_n(\lambda, \theta, \gamma) - g(\lambda, \theta, \gamma)\bigr| \\
&\quad\le \sup_{(\lambda,\theta,\gamma)\in C,\lambda\le\lambda_0} \int_{-2|y|-\gamma_1}^{\infty} (\mathbf{1}_{z\in I_\theta} + \varepsilon \mathbf{1}_{z\in\mathbb{R}\setminus I_\theta})\mathrm{e}^{-z}\,\mathrm{d}z \le 2\varepsilon \mathrm{e}^{2|y|+\gamma_1}.
\end{aligned}$$

Choose $n_1 \in \mathbb{N}$, s.t. $b_{n_1} > \lambda_1$. For $\lambda_0 \le \lambda \le \lambda_1$ and $n > n_1$,

$$\begin{aligned}
&\bigl|s_n(\lambda, \theta, z, x) - s(\lambda, \theta, z, x)\bigr| \\
&\quad= \left|\left(\lambda + \frac{x - z - \theta}{2\lambda}\right)\left(1 - \frac{1}{(1 - \lambda_1^2/b_n^2)^{1/2}}\right) - \frac{\lambda^2 z - \theta}{(1 - \lambda_1^2/b_n^2)^{1/2} b_n^2 2\lambda}\right| \qquad (27)\\
&\quad\le M_1\left|1 - \frac{1}{(1 - \lambda_0^2/b_n^2)^{1/2}}\right| + \frac{M_2}{(1 - \lambda_1^2/b_n^2)^{1/2} b_n^2} \to 0
\end{aligned}$$

for $n \to \infty$, uniformly in $z \in [-2|y| - \gamma_1, z_1]$ and $(\lambda, \theta, \gamma) \in C$ with $\lambda_0 \le \lambda \le \lambda_1$. Here, M_1 and M_2 are positive constants that only depend on $x, y, \lambda_0, \lambda_1, \theta_0, \theta_1, \gamma_1$. Let $n_2 \in \mathbb{N}$, s.t. for all $n > \max(n_1, n_2)$ the difference in (27) is less than or equal to $\varepsilon \mathrm{e}^{-2|y|-\gamma_1}$. By the Lipschitz continuity of Φ, we obtain for all $\lambda_0 \le \lambda \le \lambda_1$ and $(\lambda, \theta, \gamma) \in C$,

$$\begin{aligned}
&\int_{-2|y|-\gamma_1}^{\infty} \bigl|\Phi\bigl(s_n(\lambda, \theta, z, x)\bigr) - \Phi\bigl(s(\lambda, \theta, z, x)\bigr)\bigr|\mathrm{e}^{-z}\,\mathrm{d}z \\
&\quad\le \int_{-2|y|-\gamma_1}^{z_1} \bigl|s_n(\lambda, \theta, z, x) - s(\lambda, \theta, z, x)\bigr|\mathrm{e}^{-z}\,\mathrm{d}z + \int_{z_1}^{\infty} \mathrm{e}^{-z}\,\mathrm{d}z \\
&\quad\le \int_{-2|y|-\gamma_1}^{z_1} \varepsilon \mathrm{e}^{-2|y|-\gamma_1}\mathrm{e}^{-z}\,\mathrm{d}z + \int_{z_1}^{\infty} \mathrm{e}^{-z}\,\mathrm{d}z \le 2\varepsilon.
\end{aligned}$$

Putting the parts together yields

$$\lim_{n\to\infty} \sup_{(\lambda,\theta,\gamma)\in C} \left|\tilde{g}_n(\lambda,\theta,\gamma) - g(\lambda,\theta,\gamma)\right| = 0.$$

The assumptions of Theorem 5.5 in Billingsley [5] are satisfied and therefore $\eta_n g_n^{-1}$ converges weakly to ηg^{-1}. By a similar argument as in (24) together with the integrability condition (6), we obtain for $n \to \infty$

$$\sum_{i=1}^{n} \mathbb{P}\big(X_{i,n}^{(1)} > u_n(x), X_{i,n}^{(2)} > u_n(y)\big) \to \int_{[0,\infty]\times\mathbb{R}^2} g(\lambda,\theta,\gamma)\eta\big(\mathrm{d}(\lambda,\theta,\gamma)\big).$$

Finally, partial integration gives

$$\begin{aligned} g(\lambda,\theta,\gamma) = \mathrm{e}^{-(y-\gamma)} + \mathrm{e}^{-(x-\theta)} - \Phi\left(\lambda + \frac{y-x+\theta-\gamma}{2\lambda}\right)\mathrm{e}^{-(x-\theta)} \\ - \Phi\left(\lambda - \frac{y-x+\theta-\gamma}{2\lambda}\right)\mathrm{e}^{-(y-\gamma)}. \end{aligned}$$

Together with (22), (25), (26) and the fact that R_n converges to zero, this implies the desired result. □

Proof of Theorem 2.5. Sufficiency is a simple consequence of Theorem 2.1, where the covariance matrix of $\mathbf{X}_{i,n}$ is given by

$$\begin{pmatrix} 1 & \rho_{i,n} \\ \rho_{i,n} & 1 \end{pmatrix}.$$

For necessity, suppose that the sequence $(\max_{i=1,\dots,n} b_n(\mathbf{X}_{i,n} - b_n))_{n\in\mathbb{N}}$ of bivariate random vectors converges in distribution to some random vector Y. Let the ν_n, $n \in \mathbb{N}$, be defined as in (10) and assume that the sequence $(\nu_n)_{n\in\mathbb{N}} \subset \mathcal{M}_1([0,\infty])$ does not converge. Then, by sequential compactness, it has at least two different accumulation points $\nu, \tilde{\nu} \in \mathcal{M}_1([0,\infty])$. By the first part of this theorem, $(\max_{i=1,\dots,n} b_n(\mathbf{X}_{i,n} - b_n))_{n\in\mathbb{N}}$ converges in distribution to $F_\nu \equiv F_{\tilde{\nu}}$. It now suffices to show that $F_\nu \equiv F_{\tilde{\nu}}$ implies $\nu \equiv \tilde{\nu}$ to conclude that $(\nu_n)_{n\in\mathbb{N}} \subset \mathcal{M}_1([0,\infty])$ converges to some measure ν and that Y has distribution F_ν.

The fact that there is a one-to-one correspondence between Hüsler–Reiss distributions F_λ and the dependence parameter $\lambda \in [0,\infty]$ is straightforward [20]. Showing a similar result in our case, however, requires more effort.

To this end, for two measures $\nu_1, \nu_2 \in \mathcal{M}_1([0,\infty])$ define random variables Y_1 and Y_2 with distribution F_{ν_1} and F_{ν_2}, respectively. First, suppose that $\nu_1(\{\infty\}) = \nu_2(\{\infty\}) = 0$. For $j = 1, 2$, by Remark 2.6 we have the stochastic representation $Y_j = \max_{i\in\mathbb{N}}(U_{i,j}, U_{i,j} + B_{i,j})$, where $\sum_{i=1}^{\infty} \delta_{U_{i,j}}$ are Poisson point process on $\mathbb{R}$ with intensity $\mathrm{e}^{-u}\,\mathrm{d}u$ and the $(B_{i,j})_{i\in\mathbb{N}}$ are i.i.d. copies of the random variable B_j with normal distribution $N(-2S_j^2, 4S_j^2)$, where S_j is ν_j-distributed. Assume that

$$F_{\nu_1}(x,y) = F_{\nu_2}(x,y), \qquad \text{for all } x, y \in \mathbb{R}, \tag{28}$$

that is, the max-stable distributions of Y_1 and Y_2 are equal. Since a Poisson point process is determined by its intensity on a generating system of the σ-algebra, it follows that the point processes $\Pi_1 = \sum_{i=1}^{\infty} \delta_{(U_{i,1}, U_{i,1}+B_{i,1})}$ and $\Pi_2 = \sum_{i=1}^{\infty} \delta_{(U_{i,2}, U_{i,2}+B_{i,2})}$ are equal in distribution. Therefore, the measurable mapping

$$h : \mathbb{R}^2 \to \mathbb{R}^2, \qquad (x_1, x_2) \mapsto (x_1, x_2 - x_1)$$

induces two Poisson point processes $h(\Pi_1)$ and $h(\Pi_2)$ on $\mathbb{R}^2$ with coinciding intensity measures $\mathrm{e}^{-u}\,\mathrm{d}u\mathbb{P}_{B_1}(\mathrm{d}x)$ and $\mathrm{e}^{-u}\,\mathrm{d}u\mathbb{P}_{B_2}(\mathrm{d}x)$, respectively. Hence, B_1 and B_2 have the same distribution. Denote by ψ_j the Laplace transform of the Gaussian mixture B_j, $j = 1, 2$. A straightforward calculation yields for $u \in (0, 1)$

$$\psi_j(u) = \mathbb{E}\exp(uB_j) = \int_{[0,\infty)} \exp\bigl(-2\lambda^2\bigl(u - u^2\bigr)\bigr)\nu_j(\mathrm{d}\lambda), \qquad j = 1, 2.$$

By Lemma 7 in Kabluchko *et al.* [20], this implies the equality of measures $\nu_1^2(\mathrm{d}\lambda) = \nu_2^2(\mathrm{d}\lambda)$, where ν_j^2 is the image measure of ν_j under the transformation $[0, \infty] \to [0, \infty]$, $\lambda \mapsto \lambda^2$, for $j = 1, 2$. Hence, it also holds that $\nu_1 \equiv \nu_2$.

For arbitrary $\nu_1, \nu_2 \in \mathcal{M}_1([0, \infty])$, we first need to show that $\nu_1(\{\infty\}) = \nu_2(\{\infty\})$. For $j = 1, 2$, observe that for $n \in \mathbb{N}$

$$\begin{aligned} &-\log F_{\nu_j}(-n, 0) + \log F_{\nu_j}(-n, n) \\ &\quad = \int_{[0,\infty)} \Phi\left(\lambda + \frac{n}{2\lambda}\right)\mathrm{e}^n + \Phi\left(\lambda - \frac{n}{2\lambda}\right) - \Phi\left(\lambda + \frac{n}{\lambda}\right)\mathrm{e}^n - \Phi\left(\lambda - \frac{n}{\lambda}\right)\mathrm{e}^{-n}\nu_j(\mathrm{d}\lambda) \\ &\qquad + \bigl(1 - \mathrm{e}^{-n}\bigr)\nu_j\bigl(\{\infty\}\bigr). \end{aligned}$$

Since the second derivative of Φ is negative on the positive real line, we have the estimate

$$\mathrm{e}^n \left|\Phi\left(\lambda + \frac{n}{2\lambda}\right) - \Phi\left(\lambda + \frac{n}{\lambda}\right)\right| \le \frac{n}{2\lambda\sqrt{2\pi}}\mathrm{e}^n \mathrm{e}^{-(\lambda + n/(2\lambda))^2/2},$$

where the latter term converges pointwise to zero as $n \to \infty$. Moreover, it is uniformly bounded in $n \in \mathbb{N}$ and $\lambda \in [0, \infty)$ by a constant and hence, by dominated convergence

$$\lim_{n\to\infty} -\log F_{\nu_j}(-n, 0) + \log F_{\nu_j}(-n, n) = \nu_j\bigl(\{\infty\}\bigr), \qquad j = 1, 2.$$

It therefore follows from (28) that $\nu_1(\{\infty\}) = \nu_2(\{\infty\})$. If $\nu_1(\{\infty\}) < 1$, we apply the above to the restricted probability measures $\nu_j(\cdot \cap [0, \infty))/(1 - \nu_j(\{\infty\}))$ on $[0, \infty)$, $j = 1, 2$, to obtain $\nu_1 \equiv \nu_2$.

The last claim of the theorem follows from the fact that the integrand in (12) is bounded and continuous in λ for fixed $x, y \in \mathbb{R}$, and hence, for $\nu, \nu_n \in \mathcal{M}_1([0, \infty])$, $n \in \mathbb{N}$, weak convergence of ν_n to ν ensures the pointwise convergence of the distribution functions. □

Proof of Corollary 2.7. The first statement is a consequence of Theorem 2.5, because every sequence of random vectors can be understood as a triangular array where the columns contain equal random vectors.

For the second claim, let $\nu \in \mathcal{M}_1([0,\infty])$ be an arbitrary probability measure. Similarly as in Example 1, define an i.i.d. sequence $(R_i)_{i\in\mathbb{N}}$ of samples of ν. Choosing $\rho_i = \max(1 - 2R_i^2/b_i^2, -1)$ as correlation of $\mathbf{X}_i$ yields

$$\nu_n = \frac{1}{n}\sum_{i=1}^n \delta_{\mathbf{1}_{R_i<b_i} R_i b_n/b_i + \mathbf{1}_{R_i>b_i} b_n}.$$

First, consider the measures $\tilde{\nu}_n = \frac{1}{n}\sum_{i=1}^n \delta_{R_i b_n/b_i}$, for $n \in \mathbb{N}$. For $y \in [0,\infty]$ with $\nu(\{y\}) = 0$ we observe

$$\tilde{\nu}_n\big([0,y]\big) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{[0,y]}(R_i b_n/b_i). \tag{29}$$

Fix $\varepsilon > 0$ and recall from (1) that $b_n/\sqrt{2\log n} \to 1$ as $n \to \infty$. Hence, choose n large enough such that $i > n^{1/(1+\varepsilon)^2}$ implies $b_n/b_i < 1+\varepsilon$. Let n_ε denote the smallest integer which is strictly larger than $n^{1/(1+\varepsilon)^2}$, then (29) yields

$$\begin{aligned}\left|\tilde{\nu}_n\big([0,y]\big) - \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{[0,y]}(R_i)\right| &\le \frac{n_\varepsilon}{n} + \frac{1}{n}\left|\sum_{i=n_\varepsilon}^n \mathbf{1}_{[0,y]}(R_i b_n/b_i) - \sum_{i=n_\varepsilon}^n \mathbf{1}_{[0,y]}(R_i)\right| \\ &\le \frac{n_\varepsilon}{n} + \frac{1}{n}\sum_{i=n_\varepsilon}^n \mathbf{1}_{(y/(1+\varepsilon),y]}(R_i).\end{aligned}$$

Letting $n \to \infty$ gives

$$\lim_{n\to\infty}\left|\tilde{\nu}_n\big([0,y]\big) - \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{[0,y]}(R_i)\right| \le \nu\big((y/(1+\varepsilon),y]\big), \qquad \text{a.s.}$$

Since ε was arbitrary and $\nu(\{y\}) = 0$, it follows from the law of large numbers that $\tilde{\nu}_n$ converges a.s. weakly to ν, as $n \to \infty$. Similarly, one can see that the sequence $(\nu_n)_{n\in\mathbb{N}}$ has a.s. the same limit as $(\tilde{\nu}_n)_{n\in\mathbb{N}}$, as $n \to \infty$. □

Proof of Theorem 3.1. Let $u_n(z) = b_n + z/b_n$ for $z \in \mathbb{R}$, $u_n(\mathbf{x}) = (u_n(x_1),\dots,u_n(x_d))^\top$ for $\mathbf{x} \in \mathbb{R}^d$ and for $\mathbf{x},\mathbf{y} \in \mathbb{R}^d$ write $\mathbf{x} > \mathbf{y}$ if $x_i > y_i$ for all $1 \le i \le d$.

Let $\mathbf{x} = (x_1,\dots,x_d)^\top \in \mathbb{R}^d$ be a fixed vector and $A_{i,n}^l = \{X_{i,n}^{(l)} \le u_n(x_l)\}$ for $n \in \mathbb{N}$, $1 \le i \le n$ and $1 \le l \le d$.

$$\begin{aligned}&\log \mathbb{P}\Big(\max_{i=1,\dots,n} X_{i,n}^{(1)} \le u_n(x_1),\dots,\max_{i=1,\dots,n} X_{i,n}^{(d)} \le u_n(x_d)\Big) \\ &= \sum_{i=1}^n \log \mathbb{P}\left[\bigcap_{l=1}^d A_{i,n}^l\right] = -\sum_{i=1}^n \mathbb{P}\left[\bigcup_{l=1}^d \big(A_{i,n}^l\big)^C\right] + R_n,\end{aligned} \tag{30}$$

where R_n is a remainder term from the Taylor expansion of log. Using the same arguments as for the remainder term in (23), we conclude that R_n converges to zero as $n \to \infty$. By the additivity formula we have

$$-\mathbb{P}\left[\bigcup_{l=1}^{d}\left(A_{i,n}^{l}\right)^{C}\right] = \sum_{l=1}^{d}(-1)^{l}\sum_{m:1\le m_1<\cdots<m_l\le d}\mathbb{P}\left[\bigcap_{k=1}^{l}\left(A_{i,n}^{m_k}\right)^{C}\right]. \tag{31}$$

Consequently, by (30) and (31) it suffices to show that

$$\lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{P}\big(\mathbf{X}_{i,n} > u_n(\mathbf{x})\big) = \int_{[0,\infty)^{d\times d}} h_{d,(1,\ldots,d),\Lambda}(x_1,\ldots,x_d)\eta(\mathrm{d}\Lambda). \tag{32}$$

Let $\mathbf{Z} = (Z_1,\ldots,Z_d)$ be a standard normal random vector with independent margins and let $K = \{1,\ldots,d-1\}$. For a vector $\mathbf{x}\in\mathbb{R}^d$ let $\mathbf{x}_K = (x_1,\ldots,x_{d-1})$. If $A = (a_{j,k})_{1\le j,k\le d}\in\mathbb{R}^{d\times d}$ is a matrix, let $A_{d,K} = (a_{d,1},\ldots,a_{d,d-1})$, $A_{K,d} = (a_{1,d},\ldots,a_{d-1,d})$ and $A_{K,K} = (a_{j,k})_{j,k\in K}$.

We first assume that all $\mathbf{X}_{i,n}$ are non-degenerate, that is, $\eta_n^2(D_0) = 1$, for all $n\in\mathbb{N}$. Then, similarly as in the proof of Theorem 1.1 in Hashorva *et al.* [15], we define a new matrix $B_{i,n}\in\mathbb{R}^{(d-1)\times(d-1)}$ by

$$B_{i,n}B_{i,n}^{\top} = (\Sigma_{i,n})_{K,K} - \boldsymbol{\sigma}_{i,n}\boldsymbol{\sigma}_{i,n}^{\top}, \qquad \boldsymbol{\sigma}_{i,n} = (\Sigma_{i,n})_{K,d}, \tag{33}$$

which is well-defined since $(\Sigma_{i,n})_{K,K} - \boldsymbol{\sigma}_{i,n}\boldsymbol{\sigma}_{i,n}^{\top}$ is positive definite as the Schur complement of $(\Sigma_{i,n})_{d,d}$ in the positive definite matrix $\Sigma_{i,n}$. This enables us to write the vector $\mathbf{X}_{i,n}$ as the joint stochastic representation

$$\big(X_{i,n}^{(1)},\ldots,X_{i,n}^{(d-1)}\big) \stackrel{d}{=} B_{i,n}\mathbf{Z}_K + Z_d\boldsymbol{\sigma}_{i,n}, \qquad X_{i,n}^{(d)} \stackrel{d}{=} Z_d.$$

Therefore, since Z_d is independent of $\mathbf{Z}_K$,

$$\begin{aligned}
\mathbb{P}\big(\mathbf{X}_{i,n} > u_n(\mathbf{x})\big) &= \mathbb{P}\big(B_{i,n}\mathbf{Z}_K + Z_d\boldsymbol{\sigma}_{i,n} > u_n(\mathbf{x}_K), Z_d > u_n(x_d)\big)\\
&= \int_{x_d}^{\infty}\mathbb{P}\big(B_{i,n}\mathbf{Z}_K + u_n(s)\boldsymbol{\sigma}_{i,n} > u_n(\mathbf{x}_K)\big)b_n^{-1}\phi(b_n)\mathrm{e}^{-s-s^2/(2b_n^2)}\,\mathrm{d}s\\
&= \frac{1}{n}\int_{x_d}^{\infty} S\Big(\big(b_n^2\big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)\big)_{K,d} + x_K - s\mathbf{1}\\
&\qquad + sb_n^{-2}\big(b_n^2\big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)\big)_{K,d}\big|b_n^2B_{i,n}B_{i,n}^{\top}\Big)\\
&\qquad \times \mathrm{e}^{-s-s^2/(2b_n^2)}\,\mathrm{d}s.
\end{aligned} \tag{34}$$

It follows from the definition of $B_{i,n}$ in equation (33) that

$$\begin{aligned}
B_{i,n}B_{i,n}^{\top} &= \big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)_{K,d}\mathbf{1}^{\top} + \mathbf{1}\big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)_{d,K} - \big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)_{K,K}\\
&\quad - \big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)_{K,d}\big(\mathbf{1}\mathbf{1}^{\top} - \Sigma_{i,n}\big)_{d,K}.
\end{aligned}$$

Together with (34) and the definition of η_n this yields

$$\sum_{i=1}^{n} \mathbb{P}\big(\mathbf{X}_{i,n} > u_n(\mathbf{x})\big) = \int_{D_0} p_n(A)\eta_n^2(\mathrm{d}A),$$

where p_n is a measurable function from D_0 to $[0,\infty)$ given by

$$p_n(A) = \int_{x_d}^{\infty} S\big(2A_{K,d} + x_K - s\mathbf{1} + 2b_n^{-2} s A_{K,d} | \Gamma_{d,(1,\dots,d)}(\sqrt{A}) - 4b_n^{-2} A_{K,d} A_{d,K}\big) \times \mathrm{e}^{-s-s^2/(2b_n^2)}\,\mathrm{d}s.$$

Further, let p be the measurable function from D_0 to $[0,\infty)$

$$p(A) = \int_{x_d}^{\infty} S\big(2A_{K,d} + x_K - s\mathbf{1} | \Gamma_{d,(1,\dots,d)}(\sqrt{A})\big)\mathrm{e}^{-s}\,\mathrm{d}s.$$

Note that $\eta_n \Rightarrow \eta$ if and only if $\eta_n^2 \Rightarrow \eta^2$. In view of (32) it suffices to show that

$$\lim_{n\to\infty} \int_{D_0} p_n(A)\eta_n^2(\mathrm{d}A) = \int_{D_0} p(A)\eta^2(\mathrm{d}A). \tag{35}$$

To this end, let $A_0 \in D_0$ and $\{A_n, n \in \mathbb{N}\}$ be a sequence in D_0 that converges to A_0. We will show that $p_n(A_n) \to p(A_0)$ as $n \to \infty$. By dominated convergence, it is sufficient to show the convergence of the survivor functions. Since A_0 is in D_0, recall that $\Gamma_{d,(1,\dots,d)}(\sqrt{A_0})$ is in the space $\mathcal{M}_{(d-1)}$ of $(d-1)$-dimensional, non-degenerate covariance matrices. Moreover, since $\mathcal{M}_{(d-1)} \subset \mathbb{R}^{(d-1)\times(d-1)}$ is open and $\Gamma_{d,(1,\dots,d)}(\sqrt{A_n}) - b_n^{-2}4(A_n)_{K,d}(A_n)_{d,K}$ converges to $\Gamma_{d,(1,\dots,d)}(\sqrt{A_0})$, there is an $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$ we have $\Gamma_{d,(1,\dots,d)}(\sqrt{A_n}) - b_n^{-2}4(A_n)_{K,d}(A_n)_{d,K} \in \mathcal{M}_{(d-1)}$. Since also $2(A_n)_{K,d} + x_K - s\mathbf{1} + b_n^{-2}s2(A_n)_{K,d}$ converges to $2(A_0)_{K,d} + x_K - s\mathbf{1}$ as $n \to \infty$, we conclude that the survivor functions converge and consequently $p_n(A_n) \to p(A_0)$. Applying Theorem 5.5 in Billingsley [5] yields (35).

If not all random vectors $\mathbf{X}_{i,n}$ are non-degenerate, then it follows from the weak convergence $\eta_n^2 \Rightarrow \eta^2$ that $\eta_n^2(D \setminus D_0) \to \eta^2(D \setminus D_0) = 0$, as $n \to \infty$. Indeed, since $D \setminus D_0$ is closed in D, we have that $\eta^2(\partial(D \setminus D_0)) = 0$. Thus, the degenerate random vectors in (32) are negligible. This concludes the proof. □

Proof of Proposition 4.2. Let $t_1,\dots,t_m \in \mathbb{R}^d$ and $x_1,\dots,x_m \in \mathbb{R}$ be fixed. It follows from formula (19) in Kabluchko [19] that for a fixed variogram $\gamma_0 \in V_d$, the finite dimensional distribution $(\xi(t_1),\dots,\xi(t_m))$ of the corresponding Brown–Resnick process in (19) is given by $H_{\Lambda_{\gamma_0}}$ with $\Lambda_{\gamma_0} = (\sqrt{\gamma_0(t_j - t_k)}/4)_{1\le j,k\le m}$.

For the max-mixture of Brown–Resnick processes w.r.t. the mixture measure $\mathbb{Q}$, we obtain via void probabilities of Poisson point processes

$$\begin{aligned} -\log \mathbb{P}\big(\xi_{\mathbb{Q}}(t_1) \le x_1,\dots,\xi_{\mathbb{Q}}(t_m) \le x_m\big) &= \int_{\mathbb{R}} \mathrm{e}^{-u}\mathbb{P}\Big(u > \min_{i=1,\dots,m} x_i - W_\gamma(t_i) + 2\gamma(t_i)\Big) \\ &= \mathbb{E} \max_{i=1,\dots,m} \exp\big(W_\gamma(t_i) - 2\gamma(t_i) - x_i\big), \end{aligned}$$

where γ has distribution $\mathbb{Q}$ and the process W_γ, conditional on γ, is a zero-mean Gaussian process with stationary increments, variogram 4γ and $W_\gamma(0) = 0$ a.s. By conditioning on the variogram we get

$$\begin{aligned} &-\log \mathbb{P}\bigl(\xi_{\mathbb{Q}}(t_1) \le x_1, \dots, \xi_{\mathbb{Q}}(t_m) \le x_m\bigr) \\ &\quad = \int_{V_d} \mathbb{E} \max_{i=1,\dots,m} \exp\bigl(W_{\gamma_0}(t_i) - 2\gamma_0(t_i) - x_i\bigr) \mathbb{Q}(\mathrm{d}\gamma_0) \\ &\quad = \int_{V_d} -\log H_{\Lambda_{\gamma_0}}(x_1, \dots, x_m) \mathbb{Q}(\mathrm{d}\gamma_0). \end{aligned} \tag{36}$$

Thus, comparing with (18), the finite dimensional distributions of $\xi_{\mathbb{Q}}$ are given by the max-mixtures of Brown–Resnick processes w.r.t. the mixture measure $\mathbb{Q}$. Consequently, $\xi_{\mathbb{Q}}$ is max-stable and by (36), stationarity is preserved under max-mixing. □

Acknowledgements

The authors are grateful to Kirstin Strokorb for helpful discussions on extremal correlation functions and to two unknown referees who helped to considerably improve the paper. S. Engelke has been financially supported by Deutsche Telekom Stiftung and the Swiss National Science Foundation project 200021-134785. M. Schlather has been financially supported by Volkswagen Stiftung within the project 'WEX-MOP'.

References

[1] Balkema, A.A. and Resnick, S.I. (1977). Max-infinite divisibility. *J. Appl. Probab.* **14** 309–319. MR0438425

[2] Ballani, F. and Schlather, M. (2011). A construction principle for multivariate extreme value distributions. *Biometrika* **98** 633–645. MR2836411

[3] Berg, C., Christensen, J.P.R. and Ressel, P. (1984). *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions. Graduate Texts in Mathematics* **100**. New York: Springer. MR0747302

[4] Berg, C., Mateu, J. and Porcu, E. (2008). The Dagum family of isotropic correlation functions. *Bernoulli* **14** 1134–1149. MR2543589

[5] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396

[6] Brown, B.M. and Resnick, S.I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14** 732–739. MR0517438

[7] Cooley, D., Davis, R.A. and Naveau, P. (2010). The pairwise beta distribution: A flexible parametric multivariate model for extremes. *J. Multivariate Anal.* **101** 2103–2117. MR2671204

[8] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156

[9] de Haan, L. and Pereira, T.T. (2006). Spatial extremes: Models for the stationary case. *Ann. Statist.* **34** 146–168. MR2275238

[10] Gerritse, G. (1986). Supremum self-decomposable random vectors. *Probab. Theory Related Fields* **72** 17–33. MR0835157

[11] Gneiting, T. (1999). Radial positive definite functions generated by Euclid's hat. *J. Multivariate Anal.* **69** 88–119. MR1701408
[12] Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Table of Integrals, Series, and Products*, 6th ed. San Diego, CA: Academic Press. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger. MR1773820
[13] Hashorva, E. (2005). Elliptical triangular arrays in the max-domain of attraction of Hüsler–Reiss distribution. *Statist. Probab. Lett.* **72** 125–135. MR2137118
[14] Hashorva, E. (2006). On the multivariate Hüsler–Reiss distribution attracting the maxima of elliptical triangular arrays. *Statist. Probab. Lett.* **76** 2027–2035. MR2329248
[15] Hashorva, E., Kabluchko, Z. and Wübker, A. (2012). Extremes of independent chi-square random vectors. *Extremes* **15** 35–42. MR2891308
[16] Hashorva, E. and Weng, Z. (2013). Limit laws for extremes of dependent stationary Gaussian arrays. *Statist. Probab. Lett.* **83** 320–330. MR2998759
[17] Hüsler, J. and Reiss, R.-D. (1989). Maxima of normal random vectors: Between independence and complete dependence. *Statist. Probab. Lett.* **7** 283–286. MR0980699
[18] Joe, H. (1994). Multivariate extreme-value distributions with applications to environmental data. *Canad. J. Statist.* **22** 47–64. MR1271444
[19] Kabluchko, Z. (2011). Extremes of independent Gaussian processes. *Extremes* **14** 285–310. MR2824498
[20] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440
[21] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes. Springer Series in Statistics*. New York: Springer. MR0691492
[22] Meĭzler, D.G. (1956). On the question of limit distributions for the maximal term of a variational series. *L'vov. Politehn. Inst. Naučn. Zap. Ser. Fiz.-Mat. Vyp.* **38** 90–109. MR0256445
[23] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes. Springer Series in Operations Research and Financial Engineering*. New York: Springer. Reprint of the 1987 original. MR2364939
[24] Schlather, M. and Tawn, J.A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* **90** 139–156. MR1966556
[25] Sibuya, M. (1960). Bivariate extreme statistics. I. *Ann. Inst. Statist. Math. Tokyo* **11** 195–210. MR0115241
[26] Strokorb, K. and Schlather, M. (2014). An exceptional max-stable process fully parameterized by its extremal coefficients. *Bernoulli*. To appear.

Received May 2012 and revised August 2013