Statistics and Probability Letters 147 (2019) 105–114

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Extremal properties of the multivariate extended skew-normal distribution, Part B

B. Beranger[a,*], S.A. Padoan[b], Y. Xu[a], S.A. Sisson[a]

[a] School of Mathematics and Statistics, University of New South Wales, Sydney, Australia
[b] Department of Decision Sciences, Bocconi University of Milan, Italy

ARTICLE INFO

Article history:
Received 8 May 2018
Received in revised form 4 September 2018
Accepted 25 September 2018
Available online 7 December 2018

Keywords:
Asymptotic independence
Coefficient of upper-tail dependence
Pickands dependence function
Multivariate extreme-value distribution
Stable-tail dependence function

ABSTRACT

The skew-normal and related families are flexible and asymmetric parametric models suitable for modelling a diverse range of systems. We show that the multivariate maximum of a high-dimensional extended skew-normal random sample has asymptotically independent components and derive the speed of convergence of the joint tail. To describe the possible dependence among the components of the multivariate maximum, we show that under appropriate conditions an approximate multivariate extreme-value distribution that leads to a rich dependence structure can be derived.

1. Introduction

The skew-normal and related families, such as the more flexible extended skew-normal and extended skew-*t* distributions (Arellano-Valle and Genton, 2010; Azzalini and Capitanio, 2014, Ch 5.), are suitable for data that exhibit an asymmetric distribution, while still providing relatively simple probabilistic models. For risk analysis in the fields of insurance (credit risk management, loss ratios), climatology (floods, heat waves, storms) and health (influenza mortality), it is of particular interest to study the tail behaviour of the skew-normal and its related families (e.g. Peng et al., 2016; Fung and Seneta, 2014; Liao et al., 2014; Azzalini and Capitanio, 2014, Ch. 4). As a consequence, a number of results on the limiting extreme-value distribution for the extremes of skew-normal and skew-*t* samples have been obtained (e.g. Chang and Genton, 2007; Lysenko et al., 2009; Padoan, 2011; Beranger et al., 2017). However, while the extremal properties of skew-normal and skew-*t* distributions have been extensively studied, those of the more flexible extended skew-normal distribution have not yet been investigated.

In this contribution we derive the extremal properties of the multivariate extended skew-normal distribution. Recall that a d-dimensional random vector $\boldsymbol{X}$ follows an extended skew-normal distribution (Arellano-Valle and Genton, 2010), denoted as $\boldsymbol{X} \sim ESN_d(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \tau)$, if its probability density function (pdf) is given by

$$\phi_d(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \tau) = \frac{\phi_d(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Omega})}{\Phi\left(\tau/\sqrt{1+Q_{\bar{\Omega}}(\boldsymbol{\alpha})}\right)} \Phi(\boldsymbol{\alpha}^\top \boldsymbol{z} + \tau), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{1}$$

where $\phi_d(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Omega})$ is a d-dimensional normal pdf with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and $d \times d$ covariance matrix $\boldsymbol{\Omega}$, $\boldsymbol{z} = \boldsymbol{\omega}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$, $\boldsymbol{\omega} = \operatorname{diag}(\boldsymbol{\Omega})^{1/2}$, $\bar{\boldsymbol{\Omega}} = \boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\omega}^{-1}$, $Q_{\bar{\Omega}}(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^\top \bar{\boldsymbol{\Omega}} \boldsymbol{\alpha}$, $\Phi(\cdot)$ is the standard univariate normal cumulative distribution function (cdf) and $\boldsymbol{\alpha} \in \mathbb{R}^d$ and $\tau \in \mathbb{R}$ are the slant and extension parameters, respectively, which control the nature of density

DOI of original article: https://doi.org/10.1016/j.spl.2018.09.018.
* Corresponding author.
E-mail address: B.Beranger@unsw.edu.au (B. Beranger).

https://doi.org/10.1016/j.spl.2018.11.031
0167-7152/

deviations away from normality. When $\tau = 0$ or $\tau = 0$ and $\boldsymbol{\alpha} = \mathbf{0}$ the extended skew-normal distribution reduces to the skew-normal $SN_d(\boldsymbol{\mu}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ or the normal $N_d(\boldsymbol{\mu}, \boldsymbol{\Omega})$ distribution. Without loss of generality, we work with location and scale standardised distributions throughout, so that $ESN_d(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$ and $\Phi_d(\boldsymbol{x}; \bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$ refer to the d-dimensional extended skew-normal distribution and extended skew-normal cdf with location $\boldsymbol{\mu} = \mathbf{0}$ and correlation matrix $\bar{\boldsymbol{\Omega}}$, respectively. Finally, in the univariate setting, for brevity, we write the distributional parameters in the subscript of the pdf and cdf so that $\phi(x; \alpha, \tau) = \phi_{\alpha,\tau}(x)$ and $\Phi(x; \alpha, \tau) = \Phi_{\alpha,\tau}(x)$.

In this paper we establish that the multivariate maximum of a high-dimensional extended skew-normal random sample has asymptotically independent components. In particular, in the bivariate case we derive the speed of convergence of the joint upper tail. To describe the possible dependence between the components of the multivariate maximum, we consider a similar approach to that introduced in Hüsler and Reiss (1989). We compute a multivariate maximum over a triangular array of extended skew-normal random vectors and, under suitable conditions, derive an approximate multivariate extreme-value distribution, for large sample sizes. This leads to a model with a rich extremal dependence structure, of which we illustrate several features.

The paper is organised as follows. In Section 2 we briefly review basic notions of multivariate extreme-value theory. In Section 3 we show that the multivariate sample maximum has asymptotically independent components and for the bivariate case deduce the convergence speed of the joint tail. We complete the section by deriving an approximate multivariate extreme-value distribution and discuss some features of its extremal dependence structure. All proofs are provided in the Appendix A.

2. Extreme-value theory background

Let $I = \{1, \dots, d\}$ be an index set denoting variables of interest. Let $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$, be a series of *iid* d-dimensional random vectors, where $\boldsymbol{X}_i = (X_{i,1}, \dots, X_{i,d})^\top$ for $i = 1, \dots, n$, with a continuous joint distribution function F defined on $\mathbb{R}^d$, with marginal distributions $F_j, j \in I$. The vector of (n-partial) sample maxima is defined componentwise as $\boldsymbol{M}_n = (M_{n,1}, \dots, M_{n,d})^\top$ with $M_{n,j} = \max_{i=1,\dots,n} X_{i,j}, \quad j \in I$. As with the univariate setting, if there is a sequence of normalising constants $\boldsymbol{a}_n = (a_{n,1}, \dots, a_{n,d})^\top > \mathbf{0} = (0, \dots, 0)^\top$ and $\boldsymbol{b}_n = (b_{n,1}, \dots, b_{n,d})^\top \in \mathbb{R}^d$ such that

$$\lim_{n\to\infty} \Pr\left(\frac{\boldsymbol{M}_n - \boldsymbol{b}_n}{\boldsymbol{a}_n} \le \boldsymbol{x}\right) = \lim_{n\to\infty} F^n(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n) = G(\boldsymbol{x}), \tag{2}$$

for all continuity points $\boldsymbol{x} = (x_1, \dots, x_d)^\top \in \mathbb{R}^d$ of $G(\boldsymbol{x})$, and where $\boldsymbol{a}_n\boldsymbol{x}$ denotes componentwise multiplication, then if G is a distribution function with nondegenerate margins it is called a multivariate extreme-value distribution (e.g. Beirlant et al., 2004, Ch. 6). Specifically, G takes the form $G(\boldsymbol{x}) = C\{G_1(x_1), \dots, G_d(x_d)\}$, $\boldsymbol{x} \in \mathbb{R}^d$, where its univariate margins $G_j, j \in I$, are members of the GEV family (e.g. Beirlant et al., 2004, p. 47) and C is an extreme-value copula with expression

$$C(\boldsymbol{u}) = \exp\{-L(-\ln u_1, \dots, -\ln u_d)\}, \quad \boldsymbol{u} \in (0, 1]^d,$$

where $\boldsymbol{u} = (u_1, \dots, u_d)^\top$ and where $L : [0, \infty)^d \mapsto [0, \infty)$ is the stable dependence function (e.g. Beirlant et al., 2004, Section 8.2.2). Specifically,

$$L(\boldsymbol{z}) = d \int_{\mathcal{S}_d} \max(z_1 w_1, \dots, z_d w_d)\, H(\mathrm{d}\boldsymbol{w}), \quad \boldsymbol{z} \in [0, \infty)^d, \tag{3}$$

where $\boldsymbol{w} = (w_1, \dots, w_d)^\top$, $\boldsymbol{z} = (z_1, \dots, z_d)^\top$ and where the angular measure H is a probability measure defined on the d-dimensional unit simplex $\mathcal{S}_d := \left\{(v_1, \dots, v_d) \in [0, 1]^d : v_1 + \cdots + v_d = 1\right\}$ satisfying the mean constraint $\int_{\mathcal{S}_d} w_j H(\mathrm{d}\boldsymbol{w}) = 1/d$ for all $j \in I$. By the homogeneity property of L it follows that

$$L(\boldsymbol{z}) = (z_1 + \cdots + z_d) A(\boldsymbol{t}), \qquad \boldsymbol{z} \in [0, \infty)^d,$$

where $\boldsymbol{t} = (t_1, \dots, t_d)^\top$ with $t_j = z_j/(z_1 + \cdots + z_d)$ for $j = 1, \dots, d-1$, $t_d = 1 - t_1 - \cdots - t_{d-1}$, where A is Pickands dependence function (e.g. Beirlant et al., 2004, Section 8.2.5), which is the restriction of L on $\mathcal{S}_d$. It quantifies the level of dependence between the extremes, and satisfies the condition $1/d \le \max(t_1, \dots, t_d) \le A(\boldsymbol{t}) \le 1$ for all $\boldsymbol{t} \in \mathcal{S}_d$, with the lower and upper bounds representing complete dependence and independence, respectively.

An important and useful summary of extremal dependence is the coefficient of upper-tail dependence, denoted by χ (Li, 2009; Joe, 1997, Ch. 2). In the bivariate case, it is constructed as the probability that X_i and X_j, $i \neq j \in I$, are jointly extreme. Explicitly, $\chi := \lim_{u\to 0^+} \chi(u)$, where

$$\chi(u) = \frac{\Pr(F_i(X_i) \ge 1-u, F_j(X_j) \ge 1-u)}{u}, \quad u \in (0, 1], \tag{4}$$

where $0 \le \chi \le 1$. The variables (X_i, X_j) are said to be asymptotically independent in the upper-tail when $\chi = 0$ and are asymptotically dependent when $\chi > 0$. The case where $\chi = 1$ represents complete dependence between X_i and X_j. On the basis of the speed of convergence of $\chi(u)$ to zero as $u \to 0^+$, Ledford and Tawn (1996) proposed an approach to describe the sub-asymptotic, upper-tail dependence in the case of asymptotic independence. Specifically, they assumed that the upper-tail dependence function $\chi(u)$ (4) behaves as $\chi(u) = u^{1/\eta-1}\mathscr{L}(1/u)$, as $u \to 0^+$, where $\eta \in (0, 1]$ is the coefficient of tail dependence and $\mathscr{L}(1/u)$ is a slowly varying function, such that $\mathscr{L}(a/u)/\mathscr{L}(u) \to 1$ as $u \to 0^+$, for fixed $a > 0$. Considering $\mathscr{L}$ as a constant, at extreme levels margins are negatively associated when $\eta < 1/2$, independent when $\eta = 1/2$ and positively associated when $1/2 < \eta < 1$. When $\eta = 1$ and $\mathscr{L}(1/u) \nrightarrow 0$ asymptotic dependence is obtained.

3. Extremes of extended skew-normal random samples

It is well known that the components of both normal and skew-normal random vectors are asymptotically independent. That is, the limit distribution of the normalised vector of componentwise maxima given by (2) is equal to the product of its marginal distributions (e.g Lysenko et al., 2009; Beirlant et al., 2004, pp. 285–87). However, Beranger et al. (2017) showed that for the skew-normal case, the rate of convergence to zero of the upper-tail dependence function $\chi(u)$ in (4) depends on the slant parameters $\boldsymbol{\alpha}$, and depending on the sign of the elements of $\boldsymbol{\alpha}$, this can occur at a faster or slower rate than that of the normal case. Accordingly, from both theoretical and applied perspectives, it is important to understand whether these results also hold for the tail behaviour of the extended skew-normal distribution, in which the extension parameter τ also plays a part in the speed of convergence. We first consider the question of asymptotic dependence or asymptotic independence.

Proposition 3.1 (*Asymptotic Independence*). *Let $\boldsymbol{X} \sim ESN_d(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$. Let $\chi(u)$ with $u \in (0, 1]$ be the joint probability in (4). Then, for every bivariate pair (X_i, X_j) with $1 \le i < j \le d$ we have that $\chi = 0$.*

That is, it follows from Proposition 3.1 that regardless of the degree of sub-asymptotic dependence, the components of the multivariate extended skew-normal distribution are asymptotically independent, and so the asymptotic distribution is a product of univariate standard Gumbel distributions. We now examine the rate of convergence of $\chi(u) \to 0$ in the extended skew-normal case. Here the primary aim is to evaluate the effect on the rate of convergence of the extension parameter τ.

Proposition 3.2 (*Bivariate Tail Convergence*). *Let $(X_1, X_2) \sim ESN_2(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$, where the off-diagonal term of $\bar{\boldsymbol{\Omega}}$ is $\omega \in [0, 1)$, $\boldsymbol{\alpha} \in \mathbb{R}^2$ and $\tau \in \mathbb{R}$. Set $K = \Phi(\tau/\sqrt{1+\alpha_1^2+\alpha_2^2+2\omega\alpha_1\alpha_2})$, $\bar{\alpha}_j = (1+\alpha_j^{*2})^{1/2}$, $\alpha_j^* = (\alpha_j + \omega\alpha_{3-j})/\{1+\alpha_{3-j}(1-\omega^2)\}^{1/2}$ for j = 1,2. Then, $\chi(u) \approx u^{1/\eta-1}\mathscr{L}(1/u)$ as $u \to 0^+$, where*

(i) when either $\alpha_1, \alpha_2 \ge 0$ or $\omega > 0$ and $\alpha_j \le 0$ and $\omega\alpha_{3-j} + \alpha_j \ge 0$ for j = 1,2, then

$$\eta = (1+\omega)/2, \quad \mathscr{L}(1/u) = (1+\omega)K/(1-\omega)(4\pi \ln(1/u))^{-\frac{\omega}{1+\omega}}.$$

(ii) when $\omega > 0$, $\alpha_j \le 0$ and $-\alpha_j\omega \le \alpha_{3-j} < -\alpha_j/\omega$ for j = 1,2, then

(iia) if $\alpha_{3-j} > -\alpha_j/\bar{\alpha}_j$, then

$$\eta = \frac{(1-\omega^2)\bar{\alpha}_j{}^2}{1-\omega^2+(\bar{\alpha}_j^2-\omega)^2}, \quad \mathscr{L}(1/u) = \frac{\bar{\alpha}_j^2(1-\omega^2)K^{1/\eta-1/\{\bar{\alpha}_j(1+\omega)\}}}{(\bar{\alpha}_j-\omega)(1-\omega\bar{\alpha}_j)}(4\pi \ln(1/u))^{1/2\eta-1}.$$

(iib) if $\alpha_{3-j} < -\alpha_j/\bar{\alpha}_j$, then

$$\eta = \left[\{1-\omega^2+(\bar{\alpha}_j^2-\omega)^2\}/\{(1-\omega)^2\bar{\alpha}_j{}^2\} + \left(\alpha_{3-j}-\alpha_j/\bar{\alpha}_j\right)^2\right]^{-1},$$

$$\mathscr{L}(1/u) = \frac{e^{-\tau^2/2}\bar{\alpha}_j^2(1-\omega^2)(\alpha_{3-j}-\alpha_j/\bar{\alpha}_j)^{-1}K^{1/\eta-1}}{(\bar{\alpha}_j-\omega)\{1-\omega\bar{\alpha}_j+\alpha_{3-j}\alpha_j\bar{\alpha}_j(1-\omega^2)\}}(4\pi \ln(1/u))^{1/2\eta-3/2}.$$

(iii) when either $\alpha_1, \alpha_2 < 0$ or $\omega > 0$, $\alpha_j < 0$ and $0 < \alpha_{3-j} < -\omega\alpha_j$ for j = 1,2, then

$$\eta = (1-\omega^2)\left\{\frac{\alpha_{3-j}^2(1-\omega^2)+1}{\bar{\alpha}_{3-j}^2} + \frac{\alpha_j^2(1-\omega^2)+1}{\bar{\alpha}_j^2} + \frac{2(\alpha_1\alpha_2(1-\omega^2)-\omega)}{\bar{\alpha}_1\bar{\alpha}_2}\right\}^{-1},$$

$$\mathscr{L}(1/u) = \frac{\bar{\alpha}_j^3\bar{\alpha}_{3-j}(1-\omega^2)\left[\{1-\omega\bar{\alpha}_{3-j}/\bar{\alpha}_j\}/(1-\omega^2)+\alpha_j\left(\alpha_j+\alpha_{3-j}\bar{\alpha}_{3-j}/\bar{\alpha}_j\right)\right]^{-1}e^{-\tau^2/2}K^{1/\eta-1}}{[\{1+\alpha_{3-j}^2(1-\omega)\}\{\bar{\alpha}_j-\omega\bar{\alpha}_{3-j}\}+\alpha_1\alpha_2(1-\omega^2)^{3/2}(1-\omega)\bar{\alpha}_{3-j}](\alpha_{3-j}\bar{\alpha}_j+\alpha_j\bar{\alpha}_{3-j})} \times (4\pi \ln(1/u))^{1/2\eta-3/2}.$$

From Proposition 3.2 we see that the contribution of the extension parameter τ to the rate of tail convergence is contained in the K^ψ term, where the power ψ is independent of τ and changes depending on the value of $\boldsymbol{\alpha}$. For a bivariate skew-normal distribution, Fig. 1 illustrates the behaviour of $\chi(1-v)$ against $v \to 1^-$, where $\chi(u)$ is the upper-tail dependence function (4), for different values of the model parameters $\omega, \alpha_1, \alpha_2$ and τ. In each panel, for fixed ω and τ, the speed of convergence of $\chi(1-v)$ to 0 as $v \to 1^-$ is fastest when both slant parameters (α_1, α_2) are negative. It is slower in any other case, with

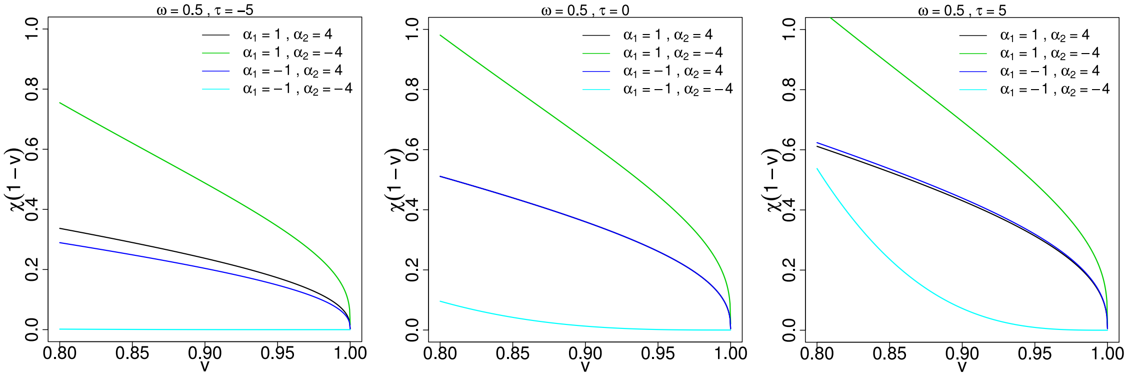

Fig. 1. The behaviour of $\chi(1-v)$ versus v for different values of the parameters $\omega, \alpha_1, \alpha_2$ and τ, for the bivariate skew-normal distribution. From left to right, the panels illustrate the effect of negative, zero and positive values of τ, respectively.

the slowest convergence rate depending on both the sign and magnitude of the slant parameters. However the effect of τ on the rate of convergence is more straightforward. While fixing all other parameters, for lower values of τ (left panel) the rate of convergence is faster than for higher values (right panel) (see Fig. 1).

Proposition 3.1 states that the marginal (componentwise) maxima $M_{n,1}, \dots, M_{n,d}$ are asymptotically independent, thereby determining an extremal framework that only permits independence among observed sample maxima. However, for data following Gaussian-type distributions, Hüsler and Reiss (1989) developed an approach by which, under suitable conditions, an alternative non-independence asymptotic distribution for componentwise maxima may be formulated. This allows an extremal dependence structure possessing a rich class of asymptotic behaviour, ranging from independence to complete dependence, to be derived. We now develop this alternative asymptotic distribution for the extended skew-normal class.

Precisely, for $n = 1, 2, \dots$ let $\boldsymbol{X}_{n,i}, i = 1, \dots, n$, be a triangular array of random vectors, where $\boldsymbol{X}_{n,i} = (X_{n,i;1}, \dots, X_{n,i;d})^\top$. Following Hüsler and Reiss (1989), for each n, assume that $\boldsymbol{X}_{n,1}, \dots, \boldsymbol{X}_{n,n}$ are independent random vectors, where $\boldsymbol{X}_{n,i} \sim ESN_d(\bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$. Here, the dependence structure and asymmetry of the extended skew-normal distribution, as measured through $\bar{\boldsymbol{\Omega}}_n$ and $\boldsymbol{\alpha}_n$, changes as the sample size n increases. In particular it is assumed that the strength of dependence and asymmetry increase with n at an appropriate rate. We formalise this as follows.

Condition 1. *For all $j \in I$, the elements of $\boldsymbol{\alpha}_n = (\alpha_{n;1}, \dots, \alpha_{n;d})^\top$ satisfy $\alpha_{n;j} \to \pm\infty$ as $n \to \infty$ and*

$$\alpha_j^\circ = \lim_{n\to\infty} \alpha_{n;j} (\ln n)^{-1/2} \in \mathbb{R},$$

with $\alpha_1^\circ + \dots + \alpha_d^\circ = 0$. For every $i, j \in I$, the correlations $\omega_{n;i,j}$ of the d-dimensional matrix $\bar{\boldsymbol{\Omega}}_n$ satisfy

$$\lambda_{i,j}^2 = \lim_{n\to\infty} (1 - \omega_{n;i,j}) \ln n \in (0, \infty].$$

Under the assumptions in Condition 1, we are now able to establish Hüsler and Reiss (1989)'s alternative extremal limit in the case of the extended skew-normal distribution.

Theorem 3.1. *Consider a triangular array of extended skew-normal random vectors $\boldsymbol{X}_{1,n}, \dots, \boldsymbol{X}_{n,n}$, $n = 1, 2, \dots$. Let $\boldsymbol{M}_{n,n} = (M_{n,n;1}, \dots, M_{n,n;d})^\top$ where $M_{n,n;j} = \max(X_{n,1;j}, X_{n,2;j}, \dots, X_{n,n;j})$, $j \in I$. Under the assumptions in Condition 1 there are sequences of normalising constants $\boldsymbol{a}_n > \boldsymbol{0}$ and $\boldsymbol{b}_n \in \mathbb{R}^d$ such that $\Phi_d^n(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n; \bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau) \to G(\boldsymbol{x})$ as $n \to \infty$, where the univariate margins of G are standard Gumbel distributions, i.e. $G(x) = e^{-e^{-x}}$ with $x \in \mathbb{R}$, and*

$$L(\boldsymbol{z}) = \sum_{j=1}^{d} z_j \, \Phi_{d-1} \left\{ \left(\lambda_{ij} + \frac{1}{2\lambda_{ij}} \log \frac{\tilde{z}_j}{\tilde{z}_i}, i \in I_j \right)^\top ; \bar{\boldsymbol{\Lambda}}_j, \tilde{\boldsymbol{\alpha}}_j, \tilde{\tau}_j \right\}, \; \boldsymbol{z} \in [0, \infty)^d, \tag{5}$$

and where $\bar{\boldsymbol{\Lambda}}_j$ is a $(d-1) \times (d-1)$ correlation matrix with upper diagonal entries $\frac{\lambda_{ij}^2 + \lambda_{kj}^2 - \lambda_{ik}^2}{2\lambda_{ij}\lambda_{kj}}$, $j \in I$, $i, k \in I_j = I \backslash \{j\}$ $\tilde{\boldsymbol{\alpha}}_j = (\sqrt{2}\alpha_i^\circ \lambda_{i,j}, i \in I_j)^\top$, $\tilde{\tau}_j = \tau - \sum_{i \in I_j} \sqrt{2}\alpha_i^\circ \lambda_{i,j}$ and

$$\tilde{z}_i = z_i \Phi \left(\frac{\tau - \sum_{k \in I_j} \sqrt{2}\lambda_{k,j}^2 \alpha_k^\circ}{\sqrt{1 + \sum_{k,m \in I_j} \alpha_k^\circ \alpha_m^\circ (\lambda_{k,j}^2 + \lambda_{m,j}^2 - \lambda_{k,m}^2)}} \right) \Big/ \Phi \left(\frac{\tau}{\sqrt{1 + \sum_{k,m \in I_j} \alpha_k^\circ \alpha_m^\circ (\lambda_{k,j}^2 + \lambda_{m,j}^2 - \lambda_{k,m}^2)}} \right)$$

and $\tilde{z}_j$ is defined as $\tilde{z}_i$ but where the index i is replaced by j and vice versa.

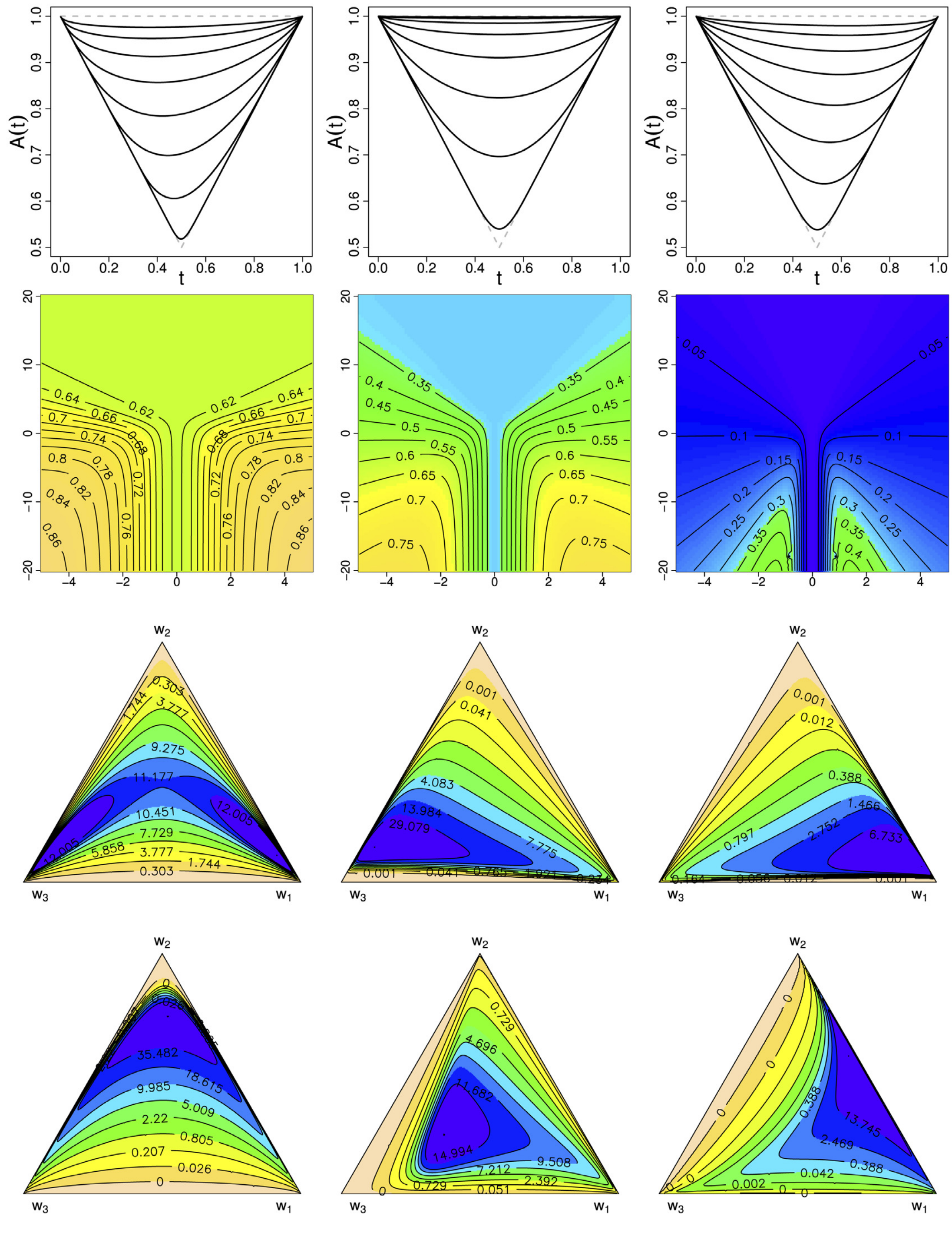

Fig. 2. Extremal dependence for the extended skew-normal distribution: (Top panels) Pickands dependence function, $A(\boldsymbol{t})$, (second row) the coefficient of upper-tail dependence χ and (bottom two rows) the angular density, $h(\boldsymbol{w})$, for different values of α°, τ and λ (see main text for details).

For the resulting multivariate extreme-value distribution in Theorem 3.1 we may derive representations of the extremal dependence. In particular, from (5) we may construct Pickands dependence function as

$$A(\boldsymbol{t}) = \sum_{j=1}^{d} t_j \,\Phi_{d-1}\left\{\left(\lambda_{ij} + \frac{1}{2\lambda_{ij}}\log\frac{\tilde{t}_j}{\tilde{t}_i},\, i \in I_j\right)^{\top};\, \bar{\boldsymbol{\Lambda}}_j, \tilde{\boldsymbol{\alpha}}_j, \tilde{\tau}_j\right\}, \tag{6}$$

for $\boldsymbol{t} = (t_1, \ldots, t_d)^{\top}$, where $\tilde{t}_j$ and $\tilde{t}_i$ are defined as $\tilde{z}_j$ and $\tilde{z}_i$.

By exploiting the method described in Beirlant et al. (e.g. 2004, pp. 263–264, 292-293) the angular measure H (defined through (3)) relative to (5) may be derived. Specifically, H places mass only in the interior of the simplex and so the angular density on $\mathcal{S}_d$ may be expressed as

$$h(\boldsymbol{w}) = \frac{\phi_{d-1}\left\{\left(\lambda_{i1} + \frac{1}{2\lambda_{i1}}\log\frac{\tilde{w}_i}{\tilde{w}_1}, i \in I_1\right)^\top ; \bar{\boldsymbol{\Lambda}}_1, \tilde{\boldsymbol{\alpha}}_1, \tilde{\tau}_1\right\}}{2\, w_1^2 \prod_{i=2}^d w_i \lambda_{i,1}}, \quad \boldsymbol{w} \in \mathcal{S}_d, \tag{7}$$

where $\tilde{w}_j$ and $\tilde{w}_i$ are defined as $\tilde{z}_j$ and $\tilde{z}_i$. Finally, for a bivariate random vector (Z_1, Z_2) with distribution given in Theorem 3.1, the coefficient upper-tail dependence in (4) is

$$\begin{aligned}\chi &= 1 - \Phi\left(\lambda_{1,2} + \frac{1}{2\lambda_{1,2}}\log\frac{\Phi\left(\frac{\tau - \sqrt{2}\lambda_{1,2}^2\alpha_2^\circ}{1+2\lambda_{1,2}^2\alpha_1^\circ}\right)}{\Phi\left(\frac{\tau + \sqrt{2}\lambda_{1,2}^2\alpha_1^\circ}{1+2\lambda_{1,2}^2\alpha_2^\circ}\right)}; -\sqrt{2}\lambda_{1,2}\alpha_2^\circ, \tau + \sqrt{2}\lambda_{1,2}^2\alpha_2^\circ\right)\\ &= 1 - \Phi\left(\lambda_{1,2} + \frac{1}{2\lambda_{1,2}}\log\frac{\Phi\left(\frac{\tau + \sqrt{2}\lambda_{1,2}^2\alpha_1^\circ}{1+2\lambda_{1,2}^2\alpha_2^\circ}\right)}{\Phi\left(\frac{\tau - \sqrt{2}\lambda_{1,2}^2\alpha_2^\circ}{1+2\lambda_{1,2}^2\alpha_1^\circ}\right)}; \sqrt{2}\lambda_{1,2}\alpha_1^\circ, \tau - \sqrt{2}\lambda_{1,2}^2\alpha_1^\circ\right).\end{aligned} \tag{8}$$

Fig. 2 graphically illustrates a range of extremal dependence structures in terms of Pickands dependence function $A(\boldsymbol{t})$ (6), the angular density $h(\boldsymbol{w})$ (7) and the coefficient of upper tail dependence χ (8). Each bivariate Pickands dependence function (top row) is constructed with λ taking eight equally spaced values between 0.1 and 3. Left to right, the top panels illustrate left-skewed (with $\alpha^\circ = -20$ and $\tau = -6$), symmetric ($\alpha^\circ = \tau = 0$) and right-skewed ($\alpha^\circ = 20$, $\tau = 6$) dependence functions. The bivariate coefficient of upper-tail dependence (second row), is illustrated for different values of $\alpha^\circ \in [-5, 5]$ and $\tau \in [-20, 20]$ and, from left to right, with $\lambda = 0.5$, 1 and 2.5. It is apparent that for fixed values of α° and τ, χ decreases for increasing values of the dependence parameter λ. Similarly, for fixed values of λ and α°, χ increases for decreasing values of τ. Finally, for fixed values of λ and τ, χ increases for increasing values of $|\alpha^\circ|$.

The bottom two rows illustrate trivariate angular densities for $\boldsymbol{\lambda}^\top = (\lambda_{12}, \lambda_{13}, \lambda_{23}) = (0.52, 0.71, 0.52)$ and different values of the parameters $\boldsymbol{\alpha}^\circ = (\alpha_1^\circ, \alpha_2^\circ, \alpha_3^\circ)^\top$ and τ. From left to right and top to bottom, the plots are produced with slant parameters $\boldsymbol{\alpha}$ given by $(0, 0, 0), (0, 5, -5), (-5, 5, 0), (4, -7, 3), (6, 0, -6)$ and $(6, -3, 3)$ and extension parameters 0, 0, 0, 3, 5 and −5, respectively. The mass in the left panel of the third row concentrates around the centre of the simplex meaning that there is strong dependence among all variables. In the middle (and right) panels of the third row, the mass is concentrated on the bottom left (right) corner and left (right) edge. This means that two variables are themselves mildly dependent, and weakly dependent of the third. In the bottom row, the mass in the left panel is concentrated on one corner and two edges, meaning that one variable is mildly dependent on the other two, and these are weakly dependent themselves. In the centre panel of the fourth row, the mass concentration in the centre panel is in the centre and on two edges, meaning that one variable is strongly dependent on the others, and these are themselves weakly dependent. Finally, in the right panel the mass concentrates on one edge, so that two variables are strongly dependent but they are each weakly dependent on the third.

4. Summary

The success of the multivariate skew-normal family is also due to its stochastic representations which motivate its use as stochastic model for data. For instance, sampling from the multivariate extended skew-normal distribution can be achieved through the distribution of the first d components of a $(d + 1)$-dimensional Gaussian random vector, conditionally that the $(d + 1)$th component satisfies a certain condition (Arellano-Valle and Genton, 2010; Azzalini and Capitanio, 2014, Ch. 5.1.3, 5.3.3). We have studied the extremal behaviour of extended skew-normal random vectors. Although their multivariate sample maximum has asymptotically independent components, we have showed that the slant and extension parameters affect the speed of convergence of the joint upper tail for each of its bivariate components. Furthermore, we have derived the asymptotic distribution for the sample maximum of a triangular array of independent extended skew-normal random vectors, under appropriate conditions on the correlations and the slant parameters. This produces a skewed version of the well-known Hüsler–Reiss model (Hüsler and Reiss, 1989), where the skewness of such a distribution is affected by the extension parameter. Hashorva and Ling (2016) have investigated the asymptotic distribution for the sample maximum of a triangular array of independent bivariate skew-elliptical triangular arrays. They have also found that a modified version of the Hüsler–Reiss model emerges as possible asymptotic distribution under an appropriate condition on the random radius relative to the elliptical random vectors. Their result differs from our result. In future would be interesting to investigate the extremal properties of a multivariate skew-elliptical distributions (Azzalini and Capitanio, 2014, Ch. 6) and study the relation with the triangular array type of approach investigated in Hashorva and Ling (2016).

Acknowledgements

We thank two anonymous referees for having carefully read the first version of this manuscript and for their helpful comments that have contributed to improving the presentation of this article. BB and SAS are supported by the Australian Centre of Excellence in Mathematical and Statistical Frontiers (ACEMS; CE140100049) and by the Australian Research Council Discovery Projects Scheme (FT170100079). SAP is supported by the Bocconi Institute for Data Science and Analytics (BIDSA), Italy.

Appendix A. Proofs

Auxiliary proofs and details of the results in Lemmas 1–3 are provided in the Supplementary Material.

A.1. Proof of Proposition 3.1

For $1 \le i < j \le d$ we define

$$\alpha_i^* = (\alpha_i + \omega_{i,j}\alpha_j)/\{1 + \alpha_j^2(1 - \omega_{i,j}^2)\}^{1/2}, \quad \alpha_j^* = (\alpha_j + \omega_{i,j}\alpha_i)/\{1 + \alpha_i^2(1 - \omega_{i,j}^2)\}^{1/2}. \tag{9}$$

We analyse the following four possible scenarios: (a) $0 < \alpha_i^* < \alpha_j^*$, (b) $\alpha_i, \alpha_j < 0$ and assume $\alpha_i^* < \alpha_j^*$ with $\alpha_i^* < 0$, (c) $\alpha_i^*, \alpha_j^* < 0$ and assume that $\alpha_i < 0$ and $\alpha_j \ge 0$ and (d) $\alpha_i^* < 0$ and $\alpha_j^* \ge 0$. Interchanging α_i^* with α_j^* produces the same results. For brevity we set $\bar{\alpha}_i = (1 + \alpha_i^{*2})^{1/2}$ and $\bar{\alpha}_j = (1 + \alpha_j^{*2})^{1/2}$. We first need the following result.

Lemma 1. *Let $\boldsymbol{X} \sim ESN_d(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$. for every pair (X_i, X_j) with $1 \le i < j \le d$ we have that under the scenarios (a) and (b)*

$$\lim_{x\to\infty} \frac{\Pr(X_j \ge x, X_i \ge x)}{\Pr(X_i \ge x)} = 0.$$

While under the scenario (c) and (d) we respectively have

$$\lim_{x\to\infty} \frac{\Pr\left(X_j \ge x, X_i \ge x\bar{\alpha}_j/\bar{\alpha}_i\right)}{\Pr\left(X_i \ge x\bar{\alpha}_j/\bar{\alpha}_i\right)} = 0 \quad \textit{and} \quad \lim_{x\to\infty} \frac{\Pr\left(X_j \ge x, X_i \ge x/\bar{\alpha}_i\right)}{\Pr\left(X_i \ge x/\bar{\alpha}_i\right)} = 0.$$

Consider the case (a) $0 < \alpha_i^* < \alpha_j^*$. Using definition (9) this assumption implies the inequality

$$\{1 + \alpha_i^2(1 - \omega_{i,j}^2)\}(\alpha_i + \omega_{i,j}\alpha_j)^2 < (\alpha_j + \omega_{i,j}\alpha_i)^2\{1 + \alpha_j^2(1 - \omega_{i,j}^2)\}$$

and from this with elementary computations we obtain $\alpha_i^2 < \alpha_j^2$ and

$$\tau_j^* = \tau/\{1 + \alpha_i^2(1 - \omega_{i,j}^2)\}^{1/2} > \tau/\{1 + \alpha_j^2(1 - \omega_{i,j}^2)\}^{1/2} = \tau_i^*.$$

Therefore, as $x \to \infty$,

$$\frac{\phi(x)\Phi(\alpha_i^* x + \tau_i^*)}{\Phi\left(\tau_i^*/\bar{\alpha}_i\right)} < \frac{\phi(x)\Phi(\alpha_j^* x + \tau_j^*)}{\Phi\left(\tau_j^*/\bar{\alpha}_j\right)},$$

which implies that $1 - \Phi_{\alpha_i^*,\tau_i^*}(x) < 1 - \Phi_{\alpha_j^*,\tau_j^*}(x)$ and $\Phi_{\alpha_i^*,\tau_i^*}(x) > \Phi_{\alpha_j^*,\tau_j^*}(x)$ as $x \to \infty$. Then

$$\begin{aligned}
\chi &= \lim_{u\to 0^+} \Pr(\Phi_{\alpha_j^*,\tau_j^*}(X_j) \ge 1 - u | \Phi_{\alpha_i^*,\tau_i^*}(X_i) \ge 1 - u) \\
&= \lim_{x\to\infty} \Pr(\Phi_{\alpha_j^*,\tau_j^*}(X_j) \ge \Phi_{\alpha_i^*,\tau_i^*}(x) | X_i \ge x) \\
&\le \lim_{x\to\infty} \Pr(\Phi_{\alpha_j^*,\tau_j^*}(X_j) \ge \Phi_{\alpha_j^*,\tau_j^*}(x) | X_i \ge x) = \lim_{x\to\infty} \frac{\Pr(X_j \ge x, X_i \ge x)}{\Pr(X_i \ge x)}.
\end{aligned}$$

By Lemma 1 the last limit is equal to zero and therefore $\chi = 0$.

Consider case (b) $\alpha_i, \alpha_j < 0$ and assume $\alpha_i^* < \alpha_j^*$ with $\alpha_i^* < 0$. Using similar arguments we obtain

$$\chi < \lim_{x\to\infty} \frac{\Pr(X_j \ge x, X_i \ge x)}{\Pr(X_i \ge x)}.$$

Then, by applying Lemma 1 we obtain $\chi = 0$.

Consider case (c) $\alpha_i^*, \alpha_j^* < 0$ and assume that $\alpha_i < 0$ and $\alpha_j \geq 0$, which implies that $\alpha_i^* < \alpha_j^*$. Applying Proposition 2.2 from Beranger et al. (2019) to $1 - \Phi_{\alpha_i^*,\tau_i^*}(x)$ and Mill's ratio (Mills, 1926) to $\Phi(\alpha_i^* x + \tau_i^*)$ we obtain

$$\begin{aligned} 1 - \Phi_{\alpha_i^*,\tau_i^*}(x) &\approx \frac{\phi(x)\Phi(\alpha_i^* x + \tau_i^*)}{\Phi\left(\tau_i^*/\bar{\alpha}_i\right)\{\bar{\alpha}_i^2 x + \alpha_i^*\tau_i^*\}} \quad \text{as } x \to \infty \\ &\approx \frac{\phi(x)\phi(\alpha_i^* x + \tau_i^*)}{\Phi\left(\tau_i^*/\bar{\alpha}_i\right)\{\bar{\alpha}_i^2 x + \alpha_i^*\tau_i^*\}\{-(\alpha_i^* x + \tau_i^*)\}} \quad \text{as } x \to \infty \\ &\approx \frac{\phi\left(x\bar{\alpha}_i\right)}{\Phi\left(\tau_i^*/\bar{\alpha}_i\right)\bar{\alpha}_i^2(-\alpha_i^*)\sqrt{2\pi}x^2} \quad \text{as } x \to \infty. \end{aligned}$$

Now note that

$$\phi\left(x(\bar{\alpha}_j/\bar{\alpha}_i)\bar{\alpha}_i\right) = \phi\left(x\bar{\alpha}_j\right)$$

$$\frac{\phi\left(x(\bar{\alpha}_j/\bar{\alpha}_i)\bar{\alpha}_i\right)}{\Phi\left(\tau_j^*/\bar{\alpha}_j\right)(\bar{\alpha}_j^2/\bar{\alpha}_i^2)\bar{\alpha}_i^2(-\alpha_i^*)\sqrt{2\pi}x^2} = \frac{\phi\left(x\bar{\alpha}_j\right)}{\Phi\left(\tau_j^*/\bar{\alpha}_j\right)\bar{\alpha}_j^2(-\alpha_i^*)\sqrt{2\pi}x^2}.$$

Since $\alpha_i^* < \alpha_j^*$ then $-1/\alpha_i^* < -1/\alpha_j^*$ and it follows that

$$\frac{\phi\left(x(\bar{\alpha}_j/\bar{\alpha}_i)\bar{\alpha}_i\right)}{\Phi\left(\tau_j^*/\bar{\alpha}_j\right)(\bar{\alpha}_j^2/\bar{\alpha}_i^2)\bar{\alpha}_i^2(-\alpha_i^*)\sqrt{2\pi}x^2} < \frac{\phi\left(x\bar{\alpha}_j\right)}{\Phi\left(\tau_j^*/\bar{\alpha}_j\right)\bar{\alpha}_j^2(-\alpha_j^*)\sqrt{2\pi}x^2}.$$

Therefore, $1 - \Phi_{\alpha_i^*,\tau_i^*}(x\bar{\alpha}_j/\bar{\alpha}_i) < 1 - \Phi_{\alpha_j^*,\tau_j^*}(x)$ and $\Phi_{\alpha_j^*,\tau_j^*}(x) < \Phi_{\alpha_i^*,\tau_i^*}(x\bar{\alpha}_j/\bar{\alpha}_i)$ and $\Phi_{\alpha_i^*,\tau_i^*}(x\bar{\alpha}_j/\bar{\alpha}_i) < \Phi_{\alpha_i^*,\tau_i^*}(x)$. From this, with some manipulation we may obtain

$$\chi \leq \lim_{x\to\infty} \frac{\Pr\left(X_j \geq x, X_i \geq x\bar{\alpha}_j/\bar{\alpha}_i\right)}{\Pr\left(X_i \geq x\bar{\alpha}_j/\bar{\alpha}_i\right)}.$$

Now, applying Lemma 1 we obtain $\chi = 0$.

Finally, consider case (d) $\alpha_i^* < 0$ and $\alpha_j^* \geq 0$. Note that as $x \to \infty$ we have

$$\Phi\left(\alpha_j^*\bar{\alpha}_i x + \tau^*\right) > \frac{1}{\sqrt{2\pi}(-\alpha_i^*)\bar{\alpha}_j x}$$

which implies that $1 - \Phi_{x\alpha_j^*,\tau_j^*}(\bar{\alpha}_i) > 1 - \Phi_{\alpha_i^*,\tau_i^*}(x)$ and $\Phi_{\alpha_j^*,\tau_j^*}(x\bar{\alpha}_i) < \Phi_{\alpha_i^*,\tau_i^*}(x)$ as $x \to \infty$. These results imply that

$$\chi \leq \lim_{x\to\infty} \frac{\Pr\left(X_j \geq x, X_i \geq x/\bar{\alpha}_i\right)}{\Pr\left(X_i \geq x/\bar{\alpha}_i\right)}.$$

Then, by applying Lemma 1 we obtain $\chi = 0$. Since $\chi = 0$ for all $1 \leq i < j \leq d$ then by Resnick (1987, Proposition 5.27) we have that $\boldsymbol{X} \sim ESN_d(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$ has asymptotically independent components.

A.2. Proof of Proposition 3.2

From Arellano-Valle and Genton (2010), recall that if $\boldsymbol{X} \sim ESN_2(\bar{\boldsymbol{\Omega}}, \boldsymbol{\alpha}, \tau)$ then for $j = 1,2$ we have

$$X_j \sim ESN(\alpha_j^*, \tau_j^*), \quad \alpha_j^* = \frac{\alpha_j + \omega\alpha_{3-j}}{\sqrt{1 + \alpha_{3-j}^2(1-\omega^2)}}, \quad \tau_j^* = \frac{\tau}{\sqrt{1 + \alpha_{3-j}^2(1-\omega^2)}},$$

$$X_j | X_{3-j} \sim ESN\left(\omega x_{3-j}, \sqrt{1-\omega^2}, \alpha_{j\cdot 3-j}, \tau_{j\cdot 3-j}\right), \quad \alpha_{j\cdot 3-j} = \alpha_j\sqrt{1-\omega^2}, \quad \tau_{j\cdot 3-j} = (1-\omega)\alpha_{3-j}x_{3-j} + \tau.$$

Define $x_j(u) = \Phi^{\leftarrow}(1-u; \alpha_j^*, \tau_j^*)$, for any $u \in [0, 1]$, where $\Phi^{\leftarrow}(\cdot; \alpha_j^*, \tau_j^*)$ is the inverse of the marginal distribution function $\Phi(\cdot; \alpha_j^*, \tau_j^*)$, for $j = 1,2$. The asymptotic behaviour of $x_j(u)$ as $u \to 0$ is

$$x_j(u) = \begin{cases} x(u), & \text{if } \alpha_j^* \geq 0 \\ \frac{x(u)}{\bar{\alpha}_j} - \frac{\alpha_j^*\tau_j^*}{\bar{\alpha}_j^2} - \frac{\ln(2\sqrt{\pi}) + \ln(|\alpha_j^*|) + 1/2\ln\ln(1/u) + \alpha_j^{*2}/2}{\ell_{1/u,\alpha_j^*}}, & \text{if } \alpha_j^* < 0 \end{cases} \tag{10}$$

for $j = 1,2$, where $\bar{\alpha}_j = \{1 + \alpha_j^{*2}\}^{1/2}$, $x(u) \approx \ell_{1/u,0} - \{\ln(2\sqrt{\pi}) + 1/2\ln\ln(1/u) + \ln\Phi(\tau_j^*/\bar{\alpha}_j)\}/\ell_{1/u,0}$ and $\ell_{1/u,a} = \sqrt{2\ln(1/u)(1+a^2)}$ for any $a \in \mathbb{R}$. We denote the asymptotic joint survivor function of the bivariate extended skew-normal distribution by $p(u) = \mathbb{P}\{X_1 > x_1(u), X_2 > x_2(u)\}$ for $u \to 0$.

For case (i), when $\alpha_1, \alpha_2 > 0$ then $x_1(u) = x_2(u) = x(u)$. Set $K = \Phi(\tau/\sqrt{1+\alpha_1^2+\alpha_2^2+2\omega\alpha_1\alpha_2})$. Then, the joint upper tail $p(u)$ behaves as $u \to 0$ as

$$
\begin{aligned}
p(u) &= \int_{x(u)}^{\infty} \left\{1 - \Phi\left(\frac{y(u)-\omega v}{\sqrt{1-\omega^2}}; \alpha_{1\cdot 2}, \tau_{1\cdot 2}\right)\right\} \phi(v; \alpha_2^*, \tau_2^*)\mathrm{d}v \\
&\approx \frac{\sqrt{1-\omega^2}}{x(u)} \int_0^{\infty} \frac{\phi(x(u), x(u)+t/x(u); \bar{\Omega}, \boldsymbol{\alpha}, \tau)}{x(u)(1-\omega)-\omega t/x(u)}\mathrm{d}t \\
&\approx \frac{K^{-1}e^{-\frac{x^2(u)}{1+\omega}}}{2\pi(1-\omega)x^2(u)} \left(\int_0^{\infty} e^{-\frac{t}{1+\omega}}\mathrm{d}t - \frac{e^{-\frac{x^2(u)(\alpha_1+\alpha_2)^2}{2}-x(u)(\alpha_1+\alpha_2)\tau}}{\sqrt{2\pi}(\alpha_1+\alpha_2)x(u)} \int_0^{\infty} e^{-t\{\frac{1}{1+\omega}+\alpha_2(\alpha_1+\alpha_2)\}}\mathrm{d}t\right) \\
&= \frac{e^{-x^2(u)/(1+\omega)}(1+\omega)}{2\pi K(1-\omega)x(u)^2} \left(1 - \frac{e^{-x^2(u)(\alpha_1+\alpha_2)^2/2-x(u)(\alpha_1+\alpha_2)\tau}}{\sqrt{2\pi}(\alpha_1+\alpha_2)\{1+\alpha_2(\alpha_1+\alpha_2)(1+\omega)\}x(u)}\right).
\end{aligned}
\tag{11}
$$

The first approximation is obtained using Proposition 2.2 from Beranger et al. (2019). The second approximation uses Mills' ratio approximation. Substituting $x(u)$ into (11) we obtain the approximation $p(u) \approx u^{1/\eta}\mathscr{L}(1/u)$ as $u \to 0^+$, where $\eta = (1+\omega)/2$ and

$$
\mathscr{L}(x) = \frac{(1+\omega)K^{\frac{1-\omega}{1+\omega}}}{(1-\omega)(4\pi \ln 1/u)^{\frac{\omega}{1+\omega}}} \left(1 - \frac{(4\pi \ln 1/u)^{\frac{(\alpha_1+\alpha_2)^2-1}{2}} u^{(\alpha_1+\alpha_2)^2} K^{(\alpha_1+\alpha_2)^2} e^{-\frac{\tau}{2}}}{(1+\omega)^{-1}(1-\omega)(\alpha_1+\alpha_2)\{1+\alpha_2(\alpha_1+\alpha_2)(1+\omega)\}}\right). \tag{12}
$$

As the second term in the parentheses in (12) is $o(u^{(\alpha_1+\alpha_2)^2})$ for $u \to 0^+$, then the quantity inside the parentheses $\to 1$ rapidly as $u \to 0^+$, and so $\mathscr{L}(1/u)$ is well approximated by the first term in (12). When $\alpha_2 < 0$ and $\alpha_1 \geq -\alpha_2/\omega$, then $\alpha_1^*, \alpha_2^* > 0$ and we obtain the same outcome.

For case (ii), when $\alpha_2 < 0$ and $-\omega, \alpha_2 \leq \alpha_1 < -\omega^{-1}\alpha_2$, then $\alpha_1^* \geq 0$ and $\alpha_2^* < 0$ and hence $x_1(u) = x(u)$ and $x_2(u)$ is given as in the second line of (10). For the case (iia), i.e. when $\alpha_1 > -\bar{\alpha}_2\alpha_2$, then following a similar derivation to that of (11), we obtain that

$$
p(u) \approx \frac{\bar{\alpha}_2^2(1-\omega^2)(1-\omega\bar{\alpha}_2)^{-1}}{2\pi^2K(\bar{\alpha}_2-\omega)x^2(u)} \exp\left[-\frac{x^2(u)}{2}\left\{\frac{1-\omega^2+(\bar{\alpha}_2-\omega)^2}{(1-\omega^2)\bar{\alpha}_2^2}\right\}\right], \quad u \to 0.
$$

Similarly, for the case (iib), i.e. when $\alpha_1 < -\bar{\alpha}_2\alpha_2$, by applying Mills' ratio we obtain

$$
p(u) \approx \frac{-\bar{\alpha}_2^2\{1-\omega\bar{\alpha}_2+\alpha_2(\alpha_2+\alpha_1\bar{\alpha}_2)(1-\omega^2)\}^{-1}}{\pi^{3/2}K(\bar{\alpha}_2-\omega)(1-\omega^2)^{-1}(\alpha_1+\alpha_2/\bar{\alpha}_2)x^3(u)} e^{-\frac{x^2(u)}{2}\left\{\frac{1-\omega^2+(\bar{\alpha}_2-\omega)^2}{(1-\omega^2)\bar{\alpha}_2^2}+\left(\alpha_1+\frac{\alpha_2}{\bar{\alpha}_2}\right)^2-\frac{\tau^2}{2}\right\}}, \quad u \to 0.
$$

For case (iii), when $\alpha_2 < 0$ and $0 < \alpha_1 < -\omega\alpha_2$, then $\alpha_1^*, \alpha_2^* < 0$ and hence $x_1(u)$ and $x_2(u)$ are given as in the second line of (10). Then, by Proposition 2.2 from Beranger et al. (2019) we obtain

$$
\begin{aligned}
p(u) \approx\; & \frac{-\bar{\alpha}_2^{3/2}\bar{\alpha}_1^2(1-\omega^2)(\bar{\alpha}_2-\omega\bar{\alpha}_1)^{-1}(\alpha_1\bar{\alpha}_2+\alpha_2\bar{\alpha}_1)^{-1}}{(2\pi)^{3/2}K\{1-\omega\bar{\alpha}_2+\alpha_2(\alpha_2+\alpha_1\bar{\alpha}_2/\bar{\alpha}_1)(1-\omega^2)\}x^3(u)} \\
& \times \exp\left[-\frac{x^2(u)}{2(1-\omega^2)}\left(\frac{\alpha_1^2(1-\omega^2)+1}{\bar{\alpha}_1^2} + \frac{\alpha_2^2(1-\omega^2)+1}{\bar{\alpha}_2^2} + \frac{2(\alpha_1\alpha_2(1-\omega^2)-\omega)}{\bar{\alpha}_1\bar{\alpha}_2}\right) - \frac{\tau^2}{2}\right] \quad u \to 0.
\end{aligned}
$$

When $\alpha_1, \alpha_2 < 0$ and $\omega_2^{-1}\alpha_2 \leq \alpha_1 < 0$ the same argument holds. Finally, interchanging α_1 with α_2 produces the same results but where α_j and $\bar{\alpha}_j$ are substituted in the above with α_{3-j} and $\bar{\alpha}_{3-j}$ respectively, for $j = 1, 2$.

A.3. Proof of Theorem 3.1

Let $\boldsymbol{X}_{n,m} \sim ESN_d(\bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$, $n \in \mathbb{N}$ and $m = 1, \ldots, n$, where $\bar{\boldsymbol{\Omega}}_n$ and $\boldsymbol{\alpha}_n$ are defined in Condition 1 and $\tau \in \mathbb{R}$. We want to derive norming constants $\boldsymbol{a}_n > \boldsymbol{0}$ and $\boldsymbol{b}_n \in \mathbb{R}^d$ such that we can derive a non-trivial limit distribution for $\Phi_d^n(\boldsymbol{a}_n\boldsymbol{x}+\boldsymbol{b}_n; \bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$. Recall that from Arellano-Valle and Genton (2010) we have that for all $j \in I$, $X_{n,m;j} \sim ESN(\alpha_{n;j}^*, \tau_{n;j}^*)$, where $\alpha_{n;j}^*$ and $\tau_{n;j}^*$ are appropriate slant and extension marginal parameters. Then, we may state the following result.

Lemma 2. *For all $j \in I$ define the normalising constants $a_{n;j} = \ell_n^{-1}$,*

$$
b_{n;j} = \ell_n - \frac{\ln(2\sqrt{\pi}) + (1/2)\ln\ln n + \ln\Phi\left(\tau_{n;j}^*/\bar{\alpha}_{n;j}\right) - \ln\Phi\left(\alpha_{n;j}^*\ell_n + \tau_{n;j}^*\right)}{\ell_n}, \qquad \text{if } \alpha_{n;j}^* \geq 0,
$$

$$
b_{n;j} = \ell_n - \frac{\ln\sqrt{2\pi} + \ln\Phi\left(\tau_{n;j}^*/\bar{\alpha}_{n;j}\right) - \ln\Phi\left(\alpha_{n;j}^*\ell_n + \tau_{n;j}^*\right)}{\ell_n} - \frac{\ln\Phi\left(\bar{\alpha}_{n;j}^2\ell_n + \alpha_{n;j}^*\tau_{n;j}^*\right)}{\ell_n}, \quad \text{if } \alpha_{n;j}^* < 0,
$$

where $\bar{\alpha}_{n;j} = \{1+\alpha_{n;j}^{*2}\}^{1/2}$, $\ell_n = \sqrt{2\ln n}$. *Then, for all* $j \in I$,

$$\lim_{n\to\infty} \Phi^n_{\alpha^*_{n;j},\tau^*_{n;j}}(a_{n;j}x_j + b_{n;j}) = e^{-e^{-x_j}}, \quad x_j \in \mathbb{R}.$$

Since for all $j \in I$, $e^{-e^{-x_j}}$ is continuous then the weak convergence of $ESN_d(\bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$ is equivalent to weak convergence of the marginal distributions functions and the copula function (e.g. Beirlant et al., 2004, Section 8.3.2). It remains to derive the limiting form of the copula function of $ESN_d(\bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$. We complete the proof deriving the stable-tail dependence function L, since an extreme-value copula is of the form $C(\boldsymbol{u}) = \exp\{-L(-\ln u_1, \ldots, -\ln u_d)\}$ (see Section 2).

Lemma 3. *The stable-tail dependence function associated with the limit distribution of* $\Phi_d^n(\boldsymbol{a}_n\boldsymbol{x} + \boldsymbol{b}_n; \bar{\boldsymbol{\Omega}}_n, \boldsymbol{\alpha}_n, \tau)$ *is*

$$\begin{aligned} L(\boldsymbol{z}) &= \lim_{n\to\infty} n\left\{1 - \Pr\left(\Phi_{\alpha^*_{n;j},\tau^*_{n;j}}(X_j) \le 1 - \frac{z_j}{n}, j = 1, \ldots, d\right)\right\}, \quad \boldsymbol{z} \in [0,\infty)^d \\ &= \sum_{j=1}^d z_j\, \Phi_{d-1}\left\{\left(\lambda_{ij} + \frac{1}{2\lambda_{ij}} \log\frac{\tilde{z}_j}{\tilde{z}_i}, i \in I_j\right)^\top; \bar{\boldsymbol{\Lambda}}_j, \tilde{\boldsymbol{\alpha}}_j, \tilde{\tau}_j\right\}, \end{aligned}$$

where for all $j \in I$, $\bar{\boldsymbol{\Lambda}}_j$, $\tilde{\boldsymbol{\alpha}}_j$ *and* $\tilde{\tau}_j$ *are given in statement of the theorem.*

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.spl.2018.11.031.

References

Arellano-Valle, R.B., Genton, M.G., 2010. Multivariate extended skew-*t* distributions and related families. Metron 68 (3), 201–234.
Azzalini, A., Capitanio, A., 2014. The Skew-Normal and Related Families. University Press, Cambridge, Cambridge.
Beirlant, J., Goegebeur, Y., Segers, J., 2004. Statistics of Extremes: Theory and Applications. John Wiley & Sons, Ltd, Chichester.
Beranger, B., Padoan, S.A., Sisson, S.A., 2017. Models for extremal dependence derived from skew-symmetric families. Scand. J. Stat. 44 (1), 21–45. http://dx.doi.org/10.1111/sjos.12240.
Beranger, B., Padoan, S.A., Xu, Y., Sisson, S.A., 2019. Extremal properties of the univariate extended skew-normal distribution, Part A. Statist. Probab. Lett. 147, 73–82. http://dx.doi.org/10.1016/j.spl.2018.09.018.
Chang, S.-M., Genton, M.G., 2007. Extreme value distributions for the skew-symmetric family of distributions. Comm. Statist. Theory Methods 36 (9), 1705–1717.
Fung, T., Seneta, E., 2014. Convergence rate to a lower tail dependence coefficient of a skew-*t* distribution. J. Multivariate Anal. 128, 62–72.
Hashorva, E., Ling, C., 2016. Maxima of skew elliptical triangular arrays. Comm. Statist. Theory Methods 45, 3692–3705.
Hüsler, J., Reiss, R.-D., 1989. Maxima of normal random vectors: between independence and complete dependence. Statist. Probab. Lett. 7 (4), 283–286.
Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall.
Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83 (1), 169–187.
Li, H., 2009. Orthant tail dependence of multivariate extreme value distributions. J. Multivariate Anal. 100, 243–256.
Liao, X., Peng, Z., Wang, X., 2014. Rates of convergence of extremes from skew-normal samples. Statist. Probab. Lett. 84, 40–47.
Lysenko, N., Roy, P., Waeber, R., 2009. Multivariate extremes of generalized skew-normal distributions. Statist. Probab. Lett 79 (4), 525–533.
Mills, J.P., 1926. Table of the ratio: area to bounding ordinate, for any portion of normal curve. Biometrika 18 (3/4), 395–400.
Padoan, S.A., 2011. Multivariate extreme models based on underlying skew-*t* and skew-normal distributions. J. Multivariate Anal. 102 (5), 977–991.
Peng, Z., Li, C., Nadarajah, S., 2016. Extremal properties of the skew-t distribution. Statist. Probab. Lett. 112, 10–19.
Resnick, S.I., 1987. Extreme Values, Regular Variation, and Point Processes. Springer-Verlag.