Chilean Journal of Statistics
Vol. 1, No. 1, April 2010, 17–33

Special Issue "In Memory of Pilar Loreto Iglesias Zuazola"
Research Paper

Multivariate unified skew-elliptical distributions

Reinaldo B. Arellano-Valle[1] and Marc G. Genton[2,*]

[1]Departamento de Estadística, Pontificia Universidad Católica de Chile, Santiago, Chile,
[2]Department of Statistics, Texas A&M University, College Station, USA.

To Pili With Love

(Received: 25 January 2010 · Accepted in final form: 08 March 2010)

Abstract

In this article, a class of multivariate unified skew-elliptical (SUE) distributions is introduced and studied in detail. In particular, three stochastic representations, the cumulative distribution function, marginal and conditional distributions, linear transformations, additivity, quadratic forms, and moments of SUE distributions are presented. The paper ends with a discussion of different but equivalent parameterizations for the density-based definition of SUE distributions.

Keywords: Elliptically contoured · Kurtosis · Multivariate · Non-normality · Skewness

Mathematics Subject Classification: Primary 62E15 · Secondary 62H05.

1. Introduction

In recent years, the quest for flexible multivariate parametric distributions has been very intensive as witnessed by the edited book of Genton (2004), the review article by Azzalini (2005), and various subsequent publications. In this paper, we define and study multivariate unified skew-elliptical distributions, a family of distributions that is mathematically tractable while being very general and flexible in its possible shapes. Although its origin is rooted in elliptically contoured (EC) distributions (Fang et al., 1990), it allows for asymmetric distributional forms. We start with a formal definition.

Definition 1.1 [Unified skew-elliptical distribution] A continuous p-dimensional random vector Y has a multivariate unified skew-elliptical (SUE) distribution, denoted by $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$, if its density function at $y \in \mathbb{R}^p$ is

$$\frac{1}{F_q(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})} f_p(y; \xi, \Omega, h^{(p)}) \, F_q\left(\Lambda z + \tau; \Gamma, h^{(q)}_{Q(z)}\right), \tag{1}$$

*Email: genton@stat.tamu.edu, reivalle@mat.puc.cl

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

where $z = \omega^{-1}(y-\xi)$, $Q(z) = z^\top \bar{\Omega}^{-1} z$, Λ is a $q \times p$ real matrix controlling shape, $\tau \in \mathbb{R}^q$ is the extension parameter, and Γ is a $q \times q$ positive definite correlation matrix. Here $f_p(y;\xi,\Omega,h^{(p)}) = |\Omega|^{-1/2} h^{(p)}(Q(z))$ denotes the density function of an elliptically contoured distribution with location $\xi \in \mathbb{R}^p$, positive definite $p \times p$ dispersion matrix Ω, with $p \times p$ scale and correlation matrices $\omega = \mathrm{diag}(\Omega)^{1/2}$ and $\bar{\Omega} = \omega^{-1}\Omega\omega^{-1}$, respectively, and density generator $h^{(p)}$. The function $F_r(x;\Sigma,h^{(r)})$ denotes the r-dimensional centered elliptical cumulative distribution with $r \times r$ dispersion matrix Σ and density generator $h^{(r)}$, and $h^{(q)}_{Q(z)}(u) = h^{(p+q)}\{u+Q(z)\}/h^{(p)}\{Q(z)\}$.

The SUE distribution was also considered in Arellano-Valle and Azzalini (2006) with a different but equivalent parameterization (see also Arellano-Valle et al., 2006a), albeit without a systematic study of its theoretical properties. This paper aims to fill this gap. Except for the factorization of the dispersion matrix Ω as $\omega^{-1}\bar{\Omega}\omega^{-1}$, we use here a similar parameterization as that used by González-Farías et al. (2004) for the closed skew-normal distribution. The SUE distribution reduces to the unified skew-normal (SUN) distribution of Arellano-Valle and Azzalini (2006), up to an equivalent parameterization, when $h^{(p)} = \phi^{(p)}$, the p-variate normal density generator function

$$\phi^{(p)}(u) = (2\pi)^{-p/2} \exp\left(-\frac{u}{2}\right), \quad u > 0, \tag{2}$$

and to the unified skew-t (SUT) distribution when $h^{(p)} = t^{(p)}_\nu$, the p-variate Student-t density generator function

$$t^{(p)}_\nu(u) = c(\nu,p)\left\{1+\frac{u}{\nu}\right\}^{-(\nu+p)/2}, \quad u > 0, \quad \text{with} \quad c(a,b) = \frac{\Gamma\{(a+b)/2\}}{\Gamma(a/2)(\pi a)^{b/2}}. \tag{3}$$

The p-variate probability density functions of the normal and Student-t distributions with location ξ and dispersion matrix Ω are defined through their corresponding density generator functions by $\phi_p(y;\xi,\Omega) = \phi^{(p)}\{(y-\xi)^\top\Omega^{-1}(y-\xi)\}$ and $t_p(y;\xi,\Omega,\nu) = t^{(p)}_\nu\{(y-\xi)^\top\Omega^{-1}(y-\xi)\}$, respectively.

When $q = 1$, the SUE distributions were called extended skew-t (EST) distributions by Arellano-Valle and Genton (2010a) who gave a systematic study of their properties, see also Adcock (2010) in terms of a different but equivalent parameterization. Similarly, the SUN distributions reduce to the extended skew-normal distributions (ESN) when $q = 1$. By analogy, we name extended skew-elliptical (ESE) distributions the SUE distributions with $q = 1$.

The following scheme summarizes the relationships among those various multivariate distributions:

$$\begin{array}{ccccc}
\mathrm{SUE}_{p,q} & \xrightarrow{t^{(p)}_\nu} & \mathrm{SUT}_{p,q} & \xrightarrow{\nu\to\infty} & \mathrm{SUN}_{p,q} \\
\downarrow & & \downarrow & & \downarrow \\
\mathrm{ESE}_{p,1} & \xrightarrow{t^{(p)}_\nu} & \mathrm{EST}_{p,1} & \xrightarrow{\nu\to\infty} & \mathrm{ESN}_{p,1}
\end{array}$$

In this paper, we present the probabilistic properties of the SUE distributions. Their proofs are given in the Appendix. As can be seen from the above scheme, the properties of the five subclasses (ESE, SUT, EST, SUN, ESN) are directly obtained as particular cases of the results given in this paper. If, in addition to $q = 1$, we set also $\tau = 0$, then we obtain the results for skew-elliptical (SE), skew-t (ST), and skew-normal (SN) distributions, see for example Branco and Dey (2001), Azzalini and Capitanio (2003), and Azzalini and Dalla Valle (1996), respectively.

The organization of this article is as follows. In Section 2, we present three stochastic representations of the SUE distributions, as well as their cumulative distribution functions. In Section 3, we derive the marginal and conditional distributions of the SUE family. In Section 4, we study linear transformations, additivity, and quadratic forms of random vectors with SUE distributions. In Section 5, we provide expressions to compute moments of SUE distributions. We end the paper with a discussion of different but equivalent parameterizations for the density-based definition of SUE distributions.

2. Stochastic Representations

Following Arellano-Valle et al. (2006a), the SUE distribution can be introduced as the distribution of a p-dimensional selection random vector defined by

$$Y - \xi \stackrel{d}{=} (W|W_0 < \Lambda W + \tau), \tag{4}$$

where "$\stackrel{d}{=}$" denotes equality in distribution and

$$\begin{pmatrix} W \\ W_0 \end{pmatrix} \sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega & 0 \\ 0 & \Gamma \end{pmatrix}, \varphi^{(p+q)}\right),$$

and $\varphi^{(p+q)}$ denotes the characteristic generator of the $(p+q)$-dimensional EC distribution. This definition allows to incorporate the singular SUE distribution that arises when Ω does not have full rank p; see Arellano-Valle and Azzalini (2006) for a complete discussion of the singular SUN distribution, as well as Rao (1965) and Díaz et al. (2002) for singular EC distributions. If Ω has full rank p, then the distribution of Y is nonsingular and its density can be computed as in Arellano-Valle et al. (2002):

$$f_Y(y) = \frac{1}{\mathrm{P}(W_0 - \Lambda W < \tau)} f_V(y - \xi)\mathrm{P}(W_0 < \Lambda(y - \xi) + \tau | W = y - \xi). \tag{5}$$

In what follows, we assume for simplicity that both Ω and Γ have full rank, so we are adopting Definition 1.1 for SUE distributions. Therefore, by using the decomposition $\Omega = \omega\bar{\Omega}\omega$ considered in Definition 1.1 and writing $W = \omega X$, $W_0 = X_0$, we formalize the connection between Equation (4) and Definition 1.1 in the following proposition.

Proposition 2.1 [Selection representation of SUE distributions] Let $Y = \xi + \omega Z$, where

$$Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau) \tag{6}$$

and

$$\begin{pmatrix} X \\ X_0 \end{pmatrix} \sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} & 0 \\ 0 & \Gamma \end{pmatrix}, h^{(p+q)}\right).$$

Then $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$.

The particular case of Proposition 2.1 for EST distributions is given in Arellano-Valle and Genton (2010a). Next, two equivalent representations (parameterizations) of Z are obtained as follows.

First, let $\tilde{X}_0 = \gamma^{-1}(\Lambda X - X_0)$, where $\gamma = \mathrm{diag}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)^{1/2}$. Then

$$Z \stackrel{d}{=} (X|\tilde{X}_0 + \bar{\tau} > 0), \quad \bar{\tau} = \gamma^{-1}\tau, \tag{7}$$

with

$$\begin{pmatrix} X \\ \tilde{X}_0 \end{pmatrix} \sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} & \Delta \\ \Delta^\top & \bar{\Gamma} \end{pmatrix}, h^{(p+q)}\right), \quad \Delta = \bar{\Omega}\Lambda^\top\gamma^{-1}, \quad \bar{\Gamma} = \gamma^{-1}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)\gamma^{-1}.$$

Second, let $\tilde{X} = X - \Delta\bar{\Gamma}^{-1}\tilde{X}_0$. Then,

$$Z \stackrel{d}{=} (\Delta\bar{\Gamma}^{-1}\tilde{X}_0 + \tilde{X}|\tilde{X}_0 + \bar{\tau} > 0), \tag{8}$$

with

$$\begin{pmatrix} \tilde{X} \\ \tilde{X}_0 \end{pmatrix} \sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top & 0 \\ 0 & \bar{\Gamma} \end{pmatrix}, h^{(p+q)}\right).$$

For example, for the SUN random vector $Y = \xi + \omega Z \sim \mathrm{SUN}_{p,q}(\xi, \Omega, \Lambda, \tau, \Gamma)$, we have by Equation (6) that $Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau)$ with

$$\begin{pmatrix} X \\ X_0 \end{pmatrix} \sim \mathrm{N}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} & 0 \\ 0 & \Gamma \end{pmatrix}\right),$$

while Equation (8) yields the well-known stochastic representation $Z \stackrel{d}{=} \Delta\bar{\Gamma}^{-1}\tilde{X}_* + \tilde{X}_{**}$ with $\tilde{X}_* \stackrel{d}{=} (\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)$, where $\tilde{X}_0 \sim \mathrm{N}_q(0, \bar{\Gamma})$, $\tilde{X}_{**} \sim \mathrm{N}_p\left(0, \bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top\right)$, and they are independent random vectors.

Similarly, for the SUT random vector $Y = \xi + \omega Z \sim \mathrm{SUT}_{p,q}(\xi, \Omega, \Lambda, \nu, \tau, \Gamma)$, we have by Equation (6) that $Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau)$ with

$$\begin{pmatrix} X \\ X_0 \end{pmatrix} \sim \mathrm{t}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} & 0 \\ 0 & \Gamma \end{pmatrix}, \nu\right).$$

In this case, we can show also from Equation (8) that

$$Z \stackrel{d}{=} \Delta\bar{\Gamma}^{-1}\tilde{X}_* + \sqrt{\frac{\nu + \tilde{Q}_*}{\nu + q}}\,\tilde{X}_{**}, \quad \tilde{Q}_* = \tilde{X}_*^\top\bar{\Gamma}^{-1}\tilde{X}_*, \quad \tilde{X}_* \stackrel{d}{=} (\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0), \tag{9}$$

where $\tilde{X}_0 \sim \mathrm{t}_q(0, \bar{\Gamma}, \nu)$, $\tilde{X}_{**} \sim \mathrm{t}_p\left(0, \bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top, \nu + q\right)$, and they are independent random vectors.

We note from Equation (6) that the cumulative distribution function of SUE distributions can be computed easily as

$$\mathrm{P}(Y \le y) = \frac{\mathrm{P}(X \le z, X_0 - \Lambda X < \tau)}{\mathrm{P}(X_0 - \Lambda X < \tau)} = \frac{F_{p+q}\left\{\begin{pmatrix} z \\ \tau \end{pmatrix}; \begin{pmatrix} \bar{\Omega} & -\bar{\Omega}\Lambda^\top \\ -\Lambda\bar{\Omega} & \Gamma + \Lambda\bar{\Omega}\Lambda^\top \end{pmatrix}, h^{(p+q)}\right\}}{F_q\left(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)}\right)},$$

where $z = \omega^{-1}(y - \xi)$.

Next, we provide a SUN-scale mixture representation of SUE distributions. It is a generalization of the well-known representation of ST distributions as a scale mixture of the skew-normal distribution.

Proposition 2.2 [SUN-scale mixture representation of SUE distributions] Let the random vector $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$, where for any positive integer k,

$$h^{(k)}(u) = \int_0^\infty v^{k/2}\phi^{(k)}(\sqrt{v}u)dG_0(v), \tag{10}$$

for some cumulative distribution function G_0 such that $G_0(0) = 0$, which does not depend on k, and where $\phi^{(k)}$ is the k-dimensional normal density generator defined by (2). If $\tau = 0$, then $Y \overset{d}{=} \xi + \omega V_0^{-1/2} Z_0$, where $V_0 \sim G_0$ is independent of $Z_0 \sim \mathrm{SUN}_{p,q}(0, \bar{\Omega}, \Lambda, 0, \Gamma)$.

For example, the SUT random vector $Y \sim \mathrm{SUT}_{p,q}(\xi, \Omega, \Lambda, \nu, \tau, \Gamma)$ can be represented, when $\tau = 0$, as $Y \overset{d}{=} \xi + \omega V_0^{-1/2} Z_0$, where $V_0 \sim \mathrm{Gamma}(\nu/2, \nu/2)$ and V_0 is independent of $Z_0 \sim \mathrm{SUN}_{p,q}(0, \bar{\Omega}, \Lambda, 0, \Gamma)$, since in this case we have in Equation (10) that $h^{(k)}(u) = t_\nu^{(k)}(u)$, the k-dimensional Student-t density generator defined by Equation (3). For the SUN-scale mixture, we have $Z \overset{d}{=} V_0^{-1/2} Z_0$ and so its moments can be computed easily in terms of the V_0-moments (provided they exist) and SUN-moments. For instance, $\mathrm{E}(Z) = \mathrm{E}(V_0^{-1/2})\mathrm{E}(Z_0)$ and $\mathrm{Var}(Z) = \mathrm{E}(V_0^{-1})\mathrm{Var}(Z_0) + \mathrm{Var}(V_0^{-1/2})\mathrm{E}(Z_0)\mathrm{E}(Z_0)^\top$. However, if we consider the hierarchical representation

$$(Y|V_\tau = v) \sim \mathrm{SUN}_{p,q}(\xi, v^{-1}\Omega, \Lambda, \sqrt{v}\,\tau, \Gamma) \quad \text{and} \quad V_\tau \sim G_\tau,$$

where G_τ has density function g_τ related to g_0, the density function of $V_0 \sim G_0$, by

$$g_\tau(v) = \frac{\Phi_q\left(\sqrt{v}\,\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top\right)}{F_q\left(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)}\right)} g_0(v), \tag{11}$$

where $\Phi_q(y; \Sigma)$ denotes the q-dimensional centered normal cumulative distribution function with $q \times q$ covariance matrix Σ, then $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$ with $h^{(p+q)}$ belonging to the representable class (10). The proof of this result is straightforward by considering the fact that $\mathrm{E}\{\Phi_k(\sqrt{V_0}\,x; A)\} = F_k(x; A, h^{(k)})$ for any representable generator $h^{(k)}$. In particular, the $\mathrm{SUT}_{p,q}(\xi, \Omega, \Lambda, \nu, \tau, \Gamma)$ family can be represented as in Equation (11) by assuming that g_0 is the density function of $V_0 \sim \mathrm{Gamma}(\nu/2, \nu/2)$.

3. Marginal and Conditional Distributions

We show in this section that the marginal and the conditional distributions of SUE distributions remain in that family.

Proposition 3.1 [Marginal distribution of SUE distributions] Let the random vector $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$. Consider the partition $Y^\top = (Y_1^\top, Y_2^\top)$ with $\dim(Y_1) = p_1$, $\dim(Y_2) = p_2$, $p_1 + p_2 = p$, and the corresponding partition of the parameters (ξ, Ω, Λ). Then, $Y_i \sim \mathrm{SUE}_{p_i,q}(\xi_i, \Omega_{ii}, \Lambda_{i(j)}, h^{(p_i+q)}, \bar{\tau}_{i(j)}, \bar{\Gamma}_{i(j)})$ for $i = 1, 2$, where

$$\Lambda_{i(j)} = \gamma_{i(j)}^{-1}(\Lambda_i + \Lambda_j\bar{\Omega}_{ji}\bar{\Omega}_{ii}^{-1}), \quad \bar{\tau}_{i(j)} = \gamma_i^{-1}\tau, \quad \bar{\Gamma}_{i(j)} = \gamma_{i(j)}^{-1}(\Gamma + \Lambda_j\tilde{\Omega}_{jj\cdot i}\Lambda_j^\top)\gamma_{i(j)}^{-1},$$

$$\tilde{\Omega}_{jj\cdot i} = \bar{\Omega}_{jj} - \bar{\Omega}_{ji}\bar{\Omega}_{ii}^{-1}\bar{\Omega}_{ij}, \quad \gamma_{i(j)} = \mathrm{diag}(\Gamma + \Lambda_j\tilde{\Omega}_{jj\cdot i}\Lambda_j^\top)^{1/2}, \quad i, j = 1, 2.$$

Note that $\Lambda_1 = 0$ with $\Lambda_2 \neq 0$ does not imply symmetry in the marginal distribution of Y_1. In fact, a necessary and sufficient condition to obtain $\Lambda_{1(2)} = 0$ is that $\Lambda_1 = \Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}$. Similar conditions are necessary to obtain symmetry in the marginal distribution of Y_2. These facts are directly related with the parameterization used in Definition 1.1.

Proposition 3.2 [Conditional distribution of SUE distributions] Let the random vector $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$. Consider the partition $Y^\top = (Y_1^\top, Y_2^\top)$ with $\dim(Y_1) = p_1$, $\dim(Y_2) = p_2$, $p_1 + p_2 = p$, and the corresponding partition of the parameters. Then, $(Y_1|Y_2 = y_2) \sim \mathrm{SUE}_{p_1,q}(\xi_{1\cdot 2}, \Omega_{11\cdot 2}, \Lambda_{1\cdot 2}, h^{(p_1+q)}_{Q_2(z_2)}, \tau_{1\cdot 2}, \Gamma)$, where $Q_2(z_2) = z_2^\top \bar{\Omega}_{22}^{-1} z_2$, with $z_2 = \omega_2^{-1}(y_2 - \xi_2)$, $\xi_{1\cdot 2} = \xi_1 + \Omega_{12}\Omega_{22}^{-1}(y_2 - \xi_2)$, $\Omega_{11\cdot 2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$, $\Lambda_{1\cdot 2} = \Lambda_1\omega_1^{-1}\omega_{1\cdot 2}$, and $\tau_{1\cdot 2} = (\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \Lambda_2)z_2 + \tau$.

Unlike the marginal distributions, we note here for the conditional distribution of $(Y_1|Y_2 = y_2)$ that $\Lambda_1 = 0$ implies symmetry. The particular cases of Propositions 3.1 and 3.2 for EST distributions are given in Arellano-Valle and Genton (2010a). Their application to perturbation of numerical confidential databases via ST distributions has been studied by Lee et al. (2010). The results above suggest that perturbations based on SUE distributions are possible as well.

For the $\mathrm{SUT}_{p,q}(\xi, \Omega, \nu, \tau, \Gamma)$ distribution, we have by Equation (3) the following conditional density generator (see also Arellano-Valle and Bolfarine, 1995; Arellano-Valle et al., 1994, 2006b):

$$
\begin{aligned}
h^{(p_1+q)}_{Q_2(z_2)}(u) &= \frac{h^{(p_1+p_2+q)}\{Q_2(z_2) + u\}}{h^{(p_2)}\{Q_2(z_2)\}} \\
&= \frac{c(\nu, p_1+p_2+q)[1 + \{Q_2(z_2) + u\}/\nu]^{-(\nu+p_1+p_2+q)/2}}{c(\nu, p_2)\{1 + Q_2(z_2)/\nu\}^{-(\nu+p_2)/2}} \\
&= c(\nu + p_2, p_1 + q)\left(\frac{\nu + p_2}{\nu + Q_2(z_2)}\right)^{\frac{p_1+q}{2}} \left\{1 + \left(\frac{\nu + p_2}{\nu + Q_2(z_2)}\right)\frac{u}{\nu + p_2}\right\}^{-\frac{\nu+p_2+p_1+q}{2}} \\
&= \left(\frac{\nu + p_2}{\nu + Q_2(z_2)}\right)^{\frac{p_1+q}{2}} h^{(p_1+q)}\left(\sqrt{\frac{\nu + p_2}{\nu + Q_2(z_2)}}\,u\right).
\end{aligned}
$$

In other words, for the SUT distribution any k-dimensional conditional generator has the form $h_a^{(k)}(u) = \alpha_a^{-k/2} h^{(k)}(\alpha_a^{-1/2}\,u)$, where $h^{(k)}$ is the corresponding unconditional generator and α_a is a scale factor induced by the quadratic form in the conditioning variable. Hence, for the SUT family the conditional density function of $Y_1|Y_2 = y_2$ is

$$
\begin{aligned}
f_{Y_1|Y_2=y_2}(y_1) = {} & \frac{1}{T_q\left(\alpha_{Q_2}^{-1/2}\tau_{1\cdot 2}; \Gamma + \Lambda_1\bar{\Omega}_{11\cdot 2}\Lambda_1^\top, \nu + p_2\right)} t_{p_1}\left(y_1; \xi_{1\cdot 2}, \alpha_{Q_2}^2\Omega_{11\cdot 2}, \nu + p_2\right) \\
& \times T_q\left\{\left(\frac{\nu + p_2 + p_1}{\nu + Q_{1\cdot 2}(z_{1\cdot 2})}\right)^{1/2}(\Lambda_{1\cdot 2}z_{1\cdot 2} + \alpha_{Q_2}^{-1/2}\tau_{1\cdot 2}); \Gamma, \nu + p_2 + p_1\right\}, \quad (12)
\end{aligned}
$$

where $\alpha_{Q_2} = \{\nu + Q_2(z_2)\}/(\nu + p_2)$, $z_{1\cdot 2} = \alpha_{Q_2}^{-1}\omega_{1\cdot 2}^{-1}(y_1 - \xi_{1\cdot 2})$, and $Q_{1\cdot 2}(z_{1\cdot 2}) = z_{1\cdot 2}^\top \bar{\Omega}_{11\cdot 2}^{-1} z_{1\cdot 2}$. Consequently, we will use the following notation for the conditional SUT distribution:

$$
(Y_1|Y_2 = y_2) \sim \mathrm{SUT}_{p_1,q}\left(\xi_{1\cdot 2}, \left(\frac{\nu + Q_2(z_2)}{\nu + p_2}\right)\Omega_{11\cdot 2}, \Lambda_{1\cdot 2}, \nu + p_2, \left(\frac{\nu + p_2}{\nu + Q_2(z_2)}\right)^{1/2}\tau_{1\cdot 2}, \Gamma\right). \quad (13)
$$

4. Linear Transformations, Additivity, and Quadratic Forms

We start by describing linear transformations of SUE distributions.

Proposition 4.1 [Linear transformation of SUE distributions] Let the random vector $Y \sim \mathrm{SUE}_{p,q}(\xi, \Omega, \Lambda, h^{(p+q)}, \tau, \Gamma)$. Then $AY + b \sim \mathrm{SUE}_{r,q}(\xi_A, \Omega_A, \bar{\Lambda}_A, h^{(r+q)}, \bar{\tau}_A, \bar{\Gamma}_A)$ for any $r \times p$ matrix A of rank $r \leq p$ and $r \times 1$ vector b, where $\xi_A = A\xi + b$, $\Omega_A = A\Omega A^\top$, with scale and correlation matrices $\omega_A = \mathrm{diag}(\Omega_A)^{1/2}$ and $\bar{\Omega}_A = \omega_A^{-1}\Omega_A\omega_A^{-1}$, respectively, and

$$\begin{aligned}
\bar{\Lambda}_A &= \gamma_A^{-1}\Lambda_A, \quad \Lambda_A = \Lambda\bar{\Omega}\omega A^\top\omega_A^{-1}\bar{\Omega}_A^{-1} = \Lambda\bar{\Omega}\omega A^\top\Omega_A^{-1}\omega_A, \quad \bar{\tau}_A = \gamma_A^{-1}\tau, \\
\gamma_A &= \mathrm{diag}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top)^{1/2}, \quad \bar{\Gamma}_A = \gamma_A^{-1}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top)\gamma_A^{-1}.
\end{aligned}$$

In particular, if $r = p$, i.e. A is a $p \times p$ non-singular matrix, then $AY + b \sim \mathrm{SUE}_{p,q}(A\xi + b, A\Omega A^\top, \Lambda\omega^{-1}A^{-1}\omega_A, h^{(r+q)}, \tau, \Gamma)$.

Here, once again, the particular case of Proposition 4.1 for EST distributions is given in Arellano-Valle and Genton (2010a).

We note that if, in Proposition 4.1, A is an $r \times p$ matrix of rank $p < r$, then by using the fact that the matrix $A^\top A$ is invertible and by letting $\Lambda_A = \Lambda(A^\top A)^{-1}A^\top$, we have also by Equation (4) that $A(Y - \xi) \stackrel{d}{=} (AW | W_0 < \Lambda_A AW + \tau)$. Therefore, $AY \sim \mathrm{SUE}_{r,q}(A\xi, A\Omega A^\top, \Lambda_A, \varphi^{(r+q)}, \tau, \Gamma)$ defines a singular SUE distribution.

An important special case follows when $A\Omega A^\top = I_p$ and $A\xi + b = 0$, which occurs, for example, when considering the decomposition $\Omega = \Omega^{1/2}\Omega^{1/2}$ and by letting $A = \Omega^{-1/2} = (\Omega^{-1})^{1/2}$ and $b = -\Omega^{-1/2}\xi$. Thus, we are considering the standardized random vector $Z_0 = \Omega^{-1/2}(Y - \xi) \sim \mathrm{SUE}_{p,q}(0, I_p, \bar{\Lambda}, h^{(p+q)}, \tau, \Gamma)$, where $\bar{\Lambda} = \Lambda\omega^{-1}\Omega^{1/2} = \Lambda\bar{\Omega}^{1/2}$. Suppose now that $q \leq p$ and that the shape matrix Λ has rank q. Hence, by applying Proposition 4.1 to the SUE random vector Z_0, but with $A = \bar{\Lambda}$, we obtain $\bar{\Lambda}Z_0 \sim \mathrm{SUE}_{q,q}(0, \Lambda\bar{\Omega}\Lambda^\top, \mathrm{diag}(\Lambda\bar{\Omega}\Lambda^\top)^{1/2}, h^{(2q)}, \tau, \Gamma)$. Moreover, denoting $\Omega_\Lambda = \Lambda\bar{\Omega}\Lambda^\top$, $\omega_\Lambda = \mathrm{diag}(\Lambda\bar{\Omega}\Lambda^\top)^{1/2}$ and $\bar{\Omega}_\Lambda = \omega_\Lambda^{-1}\Omega_\Lambda\omega_\Lambda^{-1}$, we obtain the following canonical representation of the SUE distribution:

$$Z_c = \omega_\Lambda^{-1}\bar{\Lambda}Z_0 \sim \mathrm{SUE}_{q,q}(0, \bar{\Omega}_\Lambda, \omega_\Lambda, h^{(2q)}, \tau, \Gamma). \tag{14}$$

The canonical representation given in Equation (14) reduces (summarizes) the original shape parameters in Λ (qp parameters) to the shape parameters given by the square root of the diagonal elements of the matrix $\Lambda\Omega\Lambda^\top$ (q parameters).

Another interesting example is related to the sum of two marginal SUE random vectors, which is described in the next proposition.

Proposition 4.2 [Sum of SUE distributions] Let Y_1 and Y_2 be two random vectors of dimensions $p_1 \times 1$ and $p_2 \times 1$, respectively, such that

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim \mathrm{SUE}_{p_1+p_2, q_1+q_2}\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix}, \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, h^{(p_1+p_2+q_1+q_2)}, \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix}, \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \Gamma_2 \end{pmatrix} \right),$$

where ξ_i is $p_i \times 1$, ω_i is $p_i \times p_i$, Λ_i is $q_i \times p_i$, τ_i is $q_i \times 1$ and Γ_i is $q_i \times q_i$, for $i = 1, 2$. Then,

$$Y_1 \sim \mathrm{SUE}_{p_1, q_1+q_2}\left(\xi_1, \Omega_1, \begin{pmatrix} \Lambda_1 \\ 0 \end{pmatrix}, h^{(p_1+q_1+q_2)}, \begin{pmatrix} \tau_1 \\ \gamma_2^{-1}\tau_2 \end{pmatrix}, \begin{pmatrix} \Gamma_1 & 0 \\ 0 & \gamma_2^{-1}\Gamma_2\gamma_2^{-1} \end{pmatrix} \right)$$

and

$$Y_2 \sim \mathrm{SUE}_{p_2,q_1+q_2}\left(\xi_2, \Omega_2, \begin{pmatrix} 0 \\ \Lambda_2 \end{pmatrix}, h^{(p_2+q_1+q_2)}, \begin{pmatrix} \gamma_1^{-1}\tau_1 \\ \tau_2 \end{pmatrix}, \begin{pmatrix} \gamma_1^{-1}\Gamma_1\gamma_1^{-1} & 0 \\ 0 & \Gamma_2 \end{pmatrix}\right),$$

and they are uncorrelated, where $\gamma_i = \Gamma_i - \Lambda_i\bar{\Omega}_i\Lambda_i^\top$, $i = 1, 2$. Moreover, for $p_1 = p_2 = r$, we have $Y_1 + Y_2 \sim \mathrm{SUE}_{r,q_1+q_2}(\xi_+, \Omega_+, \bar{\Lambda}_+, h^{(r+q)}, \bar{\tau}_+, \bar{\Gamma}_+)$, where $\xi_+ = \xi_1 + \xi_2$, $\Omega_+ = \Omega_1 + \Omega_2$, and

$$\begin{aligned}
\bar{\Lambda}_+ &= \gamma_+^{-1}\Lambda_+ = \begin{pmatrix} \gamma_1^{-1}\Lambda_{1+} \\ \gamma_2^{-1}\Lambda_{2+} \end{pmatrix}, \quad \Lambda_+ = \begin{pmatrix} \Lambda_{1+} \\ \Lambda_{2+} \end{pmatrix} = \begin{pmatrix} \Lambda_1\bar{\Omega}_1\omega_1\Omega_+^{-1}\omega_+ \\ \Lambda_2\bar{\Omega}_2\omega_2\Omega_+^{-1}\omega_+ \end{pmatrix}, \\
\bar{\tau}_+ &= \gamma_+^{-1}\tau = \begin{pmatrix} \gamma_1^{-1}\tau_1 \\ \gamma_2^{-1}\tau_2 \end{pmatrix}, \quad \bar{\Gamma}_+ = \gamma_+^{-1}\Gamma_+\gamma_+^{-1}, \quad \gamma_+ = \begin{pmatrix} \gamma_{1+} & 0 \\ 0 & \gamma_{2+} \end{pmatrix}, \\
\gamma_{i+} &= \mathrm{diag}(\Gamma_i + \Lambda_i\bar{\Omega}_i\Lambda_i^\top - \Lambda_{i+}\bar{\Omega}_+\Lambda_{i+}^\top)^{1/2}, \quad i = 1, 2, \\
\Gamma_+ &= \begin{pmatrix} \Gamma_1 + \Lambda_1\bar{\Omega}_1\Lambda_1^\top - \Lambda_{1+}\bar{\Omega}_+\Lambda_{1+}^\top & -\Lambda_{1+}\bar{\Omega}_+\Lambda_{2+}^\top \\ -\Lambda_{2+}\bar{\Omega}_+\Lambda_{1+}^\top & \Gamma_2 + \Lambda_2\bar{\Omega}_1\Lambda_2^\top - \Lambda_{2+}\bar{\Omega}_+\Lambda_{2+}^\top \end{pmatrix}.
\end{aligned}$$

As in the elliptically contoured case, here Y_1 and Y_2 are independent random vectors if and only if h is the normal density generator. In other words, if in Proposition 4.2 we consider the SUN distribution, then Y_1 and Y_2 are independent SUN random vectors, and the reciprocal is also true. So within the SUE class with the structure considered in Proposition 4.2, independence characterizes the SUN distribution. This means that within the SUE family, only the SUN model has the additivity property.

Distribution of quadratic forms were studied by Arellano-Valle and Genton (2010b) for the elliptical selection random vectors defined by Arellano-Valle et al. (2006a). For instance, from $Y = \xi + \omega Z$, the distribution of $Q_Y = (Y - \xi)^\top\Omega^{-1}(Y - \xi) = Z^\top\bar{\Omega}^{-1}Z = Q(Z)$, where $Z \sim \mathrm{SUE}_{p,q}(0, \bar{\Omega}, \Lambda, h^{(p+q)}, \tau, \Gamma)$, can be obtained by noting, for example from Equation (6), that $Q_Y \stackrel{d}{=} (Q_X|X_0 < \Lambda X + \tau)$, where $Q_X = X^\top\bar{\Omega}^{-1}X$. Thus, the density function of Q_Y is given by (see property (P2) in Arellano-Valle and Genton, 2010b)

$$f_{Q_Y}(v) = f_{Q_X}(v)\frac{\mathrm{P}(X_0 < \Lambda X + \tau|Q_X = v)}{\mathrm{P}(X_0 < \Lambda X + \tau)},$$

where $\mathrm{P}(X_0 < \Lambda X + \tau) = F_q(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})$ since $\tilde{X}_0 = \gamma^{-1}(\Lambda X - X_0) \sim \mathrm{EC}_q(0, \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})$, and so $\mathrm{P}(\tilde{X}_0 + \bar{\tau} > 0) = \mathrm{P}(\Lambda X - X_0 + \tau > 0)$. To compute the conditional probability $\mathrm{P}(X_0 < \Lambda X + \tau > 0|Q_X = v)$, we note that X_0 is independent of X given Q_X and $(X_0|Q_X = v) \sim \mathrm{EC}_q(0, \Gamma, h_v^{(q)})$. Thus, by using the well-known fact that a uniform random vector $U = Q_X^{-1/2}\bar{\Omega}^{-1/2}X$ is independent of Q_X and also of $T_0 = Q_X^{-1/2}X_0$, where $(T_0|Q_X = v) \sim \mathrm{EC}_q(0, v^{-1}\Gamma, h_v^{(q)})$, we have

$$\begin{aligned}
\mathrm{P}(X_0 < \Lambda X + \tau|Q_X = v) &= \mathrm{E}\{\mathrm{P}(T_0 < \bar{\Lambda}U + \sqrt{v}\,\tau|U, Q_X = v)\}, \\
&= \mathrm{E}\{F_q(\sqrt{v}\,\bar{\Lambda}U + v; \Gamma, h_v^{(q)})\},
\end{aligned}$$

where $\bar{\Lambda} = \Lambda\bar{\Omega}^{1/2}$ and the expectation is taken over U. Hence, we obtain

$$f_{Q_Y}(v) = f_{Q_X}(v)\frac{\mathrm{E}\{F_q(\sqrt{v}\,\bar{\Lambda}U + \tau; \Gamma, h_v^{(q)})\}}{F_q(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})}. \tag{15}$$

It follows that if $\bar{\Lambda} = 0$ and $\tau_v = 0$, that is, $\Lambda\bar{\Omega} = 0$ and $\tau = 0$, then $\mathrm{E}\{F_q(\sqrt{v}\,\bar{\Lambda}U + \tau; \Gamma, h_v^{(q)})\} = F_q(0; \Gamma, h_v^{(q)}) = F_q(0; \Gamma, h^{(q)}) = \Phi_q(0; \Gamma)$, and so $Q_Y \stackrel{d}{=} Q_X$. But the conditions $\Lambda\bar{\Omega} = 0$ and $\tau = 0$ are equivalent to the conditions $\Delta = 0$ and $\bar{\tau} = 0$ in Equation (7), under which $Y \stackrel{d}{=} X \sim \mathrm{EC}_p(\xi, \Omega, h^{(p)})$.

For the ESE distribution (that is, for $q = 1$), when $\tau = 0$, it is well-known that the quadratic forms $(Y-\xi)^\top\Omega^{-1}(Y-\xi) \stackrel{d}{=} (X-\xi)^\top\Omega^{-1}(X-\xi)$ for Y from the full skew-elliptical class of distributions, where $X \sim \mathrm{EC}_p(0, I_p, h^{(p)})$, an elliptically contoured distribution. In particular for an $\mathrm{EST}_p(\xi, \Omega, \nu, \lambda, 0)$ random vector Y we have $(Y-\xi)^\top\Omega^{-1}(Y-\xi) \sim pF_{p,\nu}$. Various particular cases of this invariance property have been studied by Azzalini (1986), Azzalini and Capitanio (1999), Genton et al. (2001), Ma and Genton (2004), Wang et al. (2004a,b) and Genton and Loperfido (2005). For $\tau \neq 0$ and $\psi(Y-c) = (Y-c)^\top C(Y-c)$, letting $C = A^\top A$ and $c = \xi$, Proposition 4.1 with $h^{(p+q)} = t_\nu^{(p+q)}$ implies $\psi(Y-c) = \|A\omega Z\|^2$, yielding invariance when $\lambda_A = 0$, that is, when $A\omega\bar{\Omega}\lambda = 0$, which means $\lambda = 0$ when $p = 1$. In particular, the invariance property does not hold when $p = 1$ and $\tau \neq 0$. For example, if $Y \sim \mathrm{ESN}_1(0, 1, \lambda, \tau)$, then by Equation (15), the density of $Q_Y = Y^2$ is given by

$$f_{Q_Y}(v) = \frac{1}{2\sqrt{v}\,\Phi_1(\tau/\sqrt{1+\lambda^2}; 1)}\phi_1(\sqrt{v})\left\{\Phi_1(\lambda\sqrt{v} + \tau; 1) + \Phi_1(-\lambda\sqrt{v} + \tau; 1)\right\}, \qquad (16)$$

for $v > 0$. It reduces to 0 as $\tau \to -\infty$, and to $\phi_1(\sqrt{v})/\sqrt{v}$ for $\lambda = 0$; or for $\lambda \to \pm\infty$; or for $\tau \to +\infty$. Thus, only in the cases $\lambda = 0$, $\lambda \to \pm\infty$ and $\tau \to +\infty$, we have $Q_Y \sim \chi_1^2$. However, a graphical analysis of Equation (16) indicates that departures from $Q_Y \sim \chi_1^2$ are fairly minor.

5. Moments

Moments of a SUE random vector involve truncated multivariate moments. Consider $Z \sim \mathrm{SUE}_{p,q}(0, \bar{\Omega}, \Lambda, h^{(p+q)}, \tau, \Gamma)$. By Equation (7), we have $g(Z) \stackrel{d}{=} \{g(\tilde{X}) | \tilde{X}_0 + \bar{\tau}\}$ for any Borel function g. Also, if g is integrable, then it is straightforward to show from this last relation that

$$\mathrm{E}\{g(Z)\} = \mathrm{E}[\mathrm{E}\{g(\tilde{X})|\tilde{X}_0\}|\tilde{X}_0 + \bar{\tau} > 0], \qquad (17)$$

where $\tilde{X}|\tilde{X}_0 = \tilde{x}_0 \sim \mathrm{EC}_p(\Delta\bar{\Gamma}^{-1}\tilde{x}_0, \bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top, h_{\tilde{Q}_0}^{(p)})$, with $\tilde{Q}_0 = \tilde{X}_0^\top\bar{\Gamma}^{-1}\tilde{X}_0$. In particular, considering (17), we obtain for the mean vector and covariance matrix (provided they exist) of $Z \sim \mathrm{SUE}_{p,q}(0, \bar{\Omega}, \Lambda, h^{(p+q)}, \tau, \Gamma)$ that

$$\mathrm{E}(Z) = \Delta\bar{\Gamma}^{-1}\mathrm{E}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)$$

and

$$\mathrm{Var}(Z) = \mathrm{E}(\alpha_{\tilde{Q}_0}|\tilde{X}_0 + \bar{\tau} > 0)(\bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top) + \Delta\bar{\Gamma}^{-1}\mathrm{Var}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)\bar{\Gamma}^{-1}\Delta^\top,$$

where $\alpha_{\tilde{Q}_0} = p^{-1}\mathrm{E}\{(\tilde{X} - \Delta\bar{\Gamma}^{-1}\tilde{X}_0)^\top(\bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top)^{-1}(\tilde{X} - \Delta\bar{\Gamma}^{-1}\tilde{X}_0)|\tilde{Q}_0\}$, that is, $\alpha_{\tilde{Q}_0}$ is the common marginal variance parameter associated with the conditional EC distribution of $\{(\bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top)^{-1/2}(\tilde{X} - \Delta\bar{\Gamma}^{-1}\tilde{X}_0)|\tilde{X}_0\} \stackrel{d}{=} \{(\bar{\Omega} - \Delta\bar{\Gamma}^{-1}\Delta^\top)^{-1/2}(\tilde{X} - \Delta\bar{\Gamma}^{-1}\tilde{X}_0)|\tilde{Q}_0\} \sim \mathrm{EC}_p(0, I_p, h_{\tilde{Q}_0}^{(p)})$. Expressions for the truncated mean vector $\mathrm{E}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)$ and truncated covariance matrix $\mathrm{Var}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)$ are given in González-Farías et al. (2004) for the multivariate normal distribution. For an arbitrary multivariate distribution, see Castro et al. (2010a).

In terms of the original parameterization, we obtain the following expressions for the mean vector and covariance of Z:

$$\mathrm{E}(Z) = \bar{\Omega}\Lambda^\top \bar{\mu}_{\bar{\tau}}, \quad \text{with} \quad \bar{\mu}_{\bar{\tau}} = \gamma^{-1}\mathrm{E}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0),$$

and

$$\mathrm{Var}(Z) = a_{\bar{\tau}}\{\bar{\Omega} - \bar{\Omega}\Lambda^\top(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)^{-1}\Lambda\bar{\Omega}\} + \bar{\Omega}\Lambda^\top(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)^{-1}\bar{\Sigma}_{\bar{\tau}}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)^{-1}\Lambda\bar{\Omega},$$

with $a_{\bar{\tau}} = \mathrm{E}(\alpha_{\tilde{Q}_0}|\tilde{X}_0 + \bar{\tau} > 0)$ and $\bar{\Sigma}_{\bar{\tau}} = \gamma^{-1}\mathrm{Var}(\tilde{X}_0|\tilde{X}_0 + \bar{\tau} > 0)\gamma^{-1}$. In particular, when $Z \sim \mathrm{SUT}_{p,q}(0, \bar{\Omega}, \Lambda, \nu, \tau, \Gamma)$ we have by (9) that

$$a_{\bar{\tau}} = \frac{\nu + \mathrm{E}(\tilde{Q}_0|\tilde{X}_0 + \bar{\tau} > 0)}{\nu + q - 2}.$$

Mardia's measures of multivariate skewness and kurtosis (Mardia, 1970) can also be computed following similar results as in Arellano-Valle and Genton (2010a). Substantial simplifications occur when $q = 1$, where for identifiability we should take $\Gamma = 1$. Such simplifications are due to the fact that for $q = 1$ the truncated condition $\tilde{X}_0 + \bar{\tau} > 0$ is a simple event on the real line.

6. Two Alternative Parameterizations

As mentioned in Arellano-Valle and Azzalini (2006), there are different but equivalent parameterizations to define the SUN distribution. A discussion about identifiability of the different parameterizations is given by Castro et al. (2010b) for the ESN distribution. Since these parameterizations are induced by the selection representations given in Equations (6)-(8), they can be considered also in the definition of the SUE class. Next, we present these parameterizations in terms of the density function of the random vector $Z = \omega^{-1}(Y - \xi)$. In fact, the definition of the SUE density given in Equation (1) is induced by the selection representation given in Equation (6), for which the density function of Z, called SUE, is

$$f_Z(z) = \frac{1}{F_q(\bar{\tau}; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})} f_p(z; 0, \bar{\Omega}, h^{(p)}) F_q(\Lambda z + \tau; \Gamma, h^{(q)}_{Q(z)}). \tag{18}$$

Now, if we use Equation (7), we obtain the density function derived by Arellano-Valle and Azzalini (2006), called SUE-1:

$$f_Z(z) = \frac{1}{F_q(\bar{\tau}; \bar{\Gamma}, h^{(q)})} f_p(z; 0, \Omega, h^{(p)}) F_q(\Delta^\top\bar{\Omega}^{-1}z + \bar{\tau}; \bar{\Gamma} - \Delta^\top\bar{\Omega}^{-1}\Delta, h^{(q)}_{Q(z)}). \tag{19}$$

Under this parameterization, we need the condition that $\bar{\Gamma} - \Delta^\top\bar{\Omega}^{-1}\Delta > 0$. This condition holds by construction if we assume that

$$\Omega_* = \begin{pmatrix} \bar{\Omega} & \Delta \\ \Delta^\top & \bar{\Gamma} \end{pmatrix}$$

is a nonsingular correlation (covariance) matrix. Finally, let $\bar{\Omega} = \Psi + \Delta\bar{\Gamma}^{-1}\Delta^\top = \Psi + \Upsilon\bar{\Gamma}\Upsilon^\top$, where $\Upsilon = \bar{\Gamma}^{-1}\Delta$. This parameterization simplifies Equation (7) as $Z \stackrel{d}{=} \{\Upsilon\tilde{X}_0 + \tilde{X}_1|\tilde{X}_0 + \bar{\tau} > 0\}$, where $(\tilde{X}_1^\top, \tilde{X}_0)^\top \sim \mathrm{EC}_{p+q}(0, \mathrm{diag}(\Psi, \bar{\Gamma}), h^{(p+q)})$ and yields the following SUE density function, called SUE-2:

$$f_Z(z) = \frac{1}{F_q(\bar{\tau}; \bar{\Gamma}, h^{(q)})} f_p(z; 0, \bar{\Omega}, h^{(p)}) F_q(\bar{\Gamma}\Upsilon^\top\bar{\Omega}^{-1}z + \bar{\tau}; \bar{\Gamma} - \bar{\Gamma}\Upsilon^\top\bar{\Omega}^{-1}\Upsilon\bar{\Gamma}, h^{(q)}_{Q(z)}), \tag{20}$$

with $\bar{\Omega} = \Psi + \Upsilon\bar{\Gamma}\Upsilon^\top$, and where we note that $\bar{\Gamma} - \bar{\Gamma}\Upsilon^\top\bar{\Omega}^{-1}\Upsilon\bar{\Gamma} = (\bar{\Gamma}^{-1} + \Upsilon^\top\bar{\Omega}^{-1}\Upsilon)^{-1}$.

A schematic relation between the above three parameterizations comes from the correlation matrix Ω_* as follows:

$$\underbrace{\begin{pmatrix} \bar{\Omega} & \bar{\Omega}\Lambda_\gamma^\top \\ \Lambda_\gamma\bar{\Omega} & \Gamma_\gamma + \Lambda_\gamma\bar{\Omega}\Lambda_\gamma^\top \end{pmatrix}}_{\text{SUE density (18)}} \longleftarrow \underbrace{\begin{pmatrix} \bar{\Omega} & \Delta \\ \Delta^\top & \bar{\Gamma} \end{pmatrix}}_{\text{SUE-1 density (19)}} \longrightarrow \underbrace{\begin{pmatrix} \Psi + \Upsilon\bar{\Gamma}\Upsilon^\top & \Upsilon\bar{\Gamma} \\ \bar{\Gamma}\Upsilon^\top & \bar{\Gamma} \end{pmatrix}}_{\text{SUE-2 density (20)}},$$

where $\Lambda_\gamma = \gamma^{-1}\Lambda$ and $\Gamma_\gamma = \gamma^{-1}\Gamma\gamma^{-1}$. Finally, note that $\Lambda_\gamma\bar{\Omega} = 0 \Leftrightarrow \Delta = 0 \Leftrightarrow \Upsilon\bar{\Gamma} = 0$.

Acknowledgements

Arellano-Valle's research was partially supported by grant FONDECYT 1085241-Chile. Genton's research was supported in part by NSF grant CMG ATM-0620624, and by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). The authors thank the editors and a referee for their helpful comments and suggestions.

Appendix: Proofs

Proof of Proposition 2.1 Note first that $f_Y(y) = |\omega|^{-1}f_Z(z)$, where $z = \omega^{-1}(y - \xi)$. Now, the density of $Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau)$ is given by (see, e.g., Arellano-Valle et al., 2002)

$$f_Z(z) = \frac{1}{\mathrm{P}(X_0 - \Lambda X < \tau)} f_X(z)\mathrm{P}(X_0 < \Lambda z + \tau|X = z). \tag{A.1}$$

Thus the proof follows from $(X_0|X = z) \sim \mathrm{EC}_q\left(0, \Gamma, h^{(q)}_{Q(z)}\right)$, where $Q(z) = z^\top\bar{\Omega}^{-1}z$, $X \sim \mathrm{EC}_p(0, \bar{\Omega}, h^{(p)})$ and $X_0 - \Lambda X \sim \mathrm{EC}_q(0, \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})$. ∎

Proof of Proposition 2.2 Let V be a non-negative random variable such that conditionally on $V = v$, we have

$$\left\{\begin{pmatrix} X \\ X_0 \end{pmatrix} \middle| V = v\right\} \stackrel{d}{=} v^{-1/2}\begin{pmatrix} N \\ N_0 \end{pmatrix},$$

where $V \sim G$ and V is independent of

$$\begin{pmatrix} N \\ N_0 \end{pmatrix} \sim \mathrm{N}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega} & 0 \\ 0 & \Gamma \end{pmatrix}\right).$$

Consequently, we obtain $(Z|V=v) \stackrel{d}{=} v^{-1/2}(N|N_0 < \Lambda N + v^{1/2}\tau)$, with the following conditional SUN density:

$$f_{Z|V=v}(z) = \frac{1}{\Phi_q\{\tau; v^{-1}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top)\}}\phi_p(z; v^{-1}\bar{\Omega})\Phi_q(\Lambda z + \tau; v^{-1}\Gamma),$$

from where the proof follows. ■

Proof of Proposition 3.1 Because $Y = \xi + \omega Z$, where $Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau)$, and considering the partitions

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \longrightarrow Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \longrightarrow X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

and

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}, \quad \omega = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}, \quad \bar{\Omega} = \begin{pmatrix} \bar{\Omega}_{11} & \bar{\Omega}_{12} \\ \bar{\Omega}_{21} & \bar{\Omega}_{22} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \Lambda_1 & \Lambda_2 \end{pmatrix},$$

where $\omega_i = \mathrm{diag}(\Omega_{ii})^{1/2}$ and $\bar{\Omega}_{ij} = \omega_i^{-1}\Omega_{ij}\omega_j^{-1}$ for $i,j = 1,2$, it is clear that $Y_1 = \xi_1 + \omega_1 Z_1$, where

$$Z_1 \stackrel{d}{=} (X_1|X_0 < \Lambda X + \tau) = \{X_1|X_{01} < (\Lambda_1 + \Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})X_1 + \tau\},$$

and $X_{01} = X_0 - \Lambda_2 X_{2\cdot1}$ and $X_{2\cdot1} = X_2 - \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}X_1$. Thus, to obtain the density of Z_1, we can apply Equation (A.1) to $\{X_1|X_{01} < (\Lambda_1 + \Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})X_1 + \tau\}$ by noting that

$$\begin{pmatrix} X_1 \\ X_{01} \end{pmatrix} = \begin{pmatrix} X_1 \\ X_0 - \Lambda_2 X_{2\cdot1} \end{pmatrix} \sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \bar{\Omega}_{11} & 0 \\ 0 & \Gamma + \Lambda_2\tilde{\Omega}_{22\cdot1}\Lambda_2^\top \end{pmatrix}, h^{(p_1+q)}\right),$$

and so $(X_{01}|X_1 = z_1) \sim \mathrm{EC}_q(0, \Gamma + \Lambda_2\bar{\Omega}_{22\cdot1}\Lambda_2^\top, h^{(q)}_{Q_1(z_1)})$, where $Q_1(z_1) = z_1^\top\bar{\Omega}_{11}z_1$ and $\tilde{\Omega}_{22\cdot1} = \bar{\Omega}_{22} - \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}\bar{\Omega}_{12}$. Note here that the dispersion matrix $\Gamma + \Lambda_2\tilde{\Omega}_{22\cdot1}\Lambda_2^\top$ induces the correlation matrix $\bar{\Gamma}_{1(2)} = \gamma_{1(2)}^{-1}(\Gamma + \Lambda_2\tilde{\Omega}_{22\cdot1}\Lambda_2^\top)\gamma_{1(2)}^{-1}$, where $\gamma_{1(2)} = \mathrm{diag}(\Gamma + \Lambda_2\tilde{\Omega}_{22\cdot1}\Lambda_2^\top)$ is the corresponding scale matrix. Hence, by noting also that $\Gamma + \Lambda\bar{\Omega}\Lambda^\top = \gamma_1(\bar{\Gamma}_1 + \bar{\Lambda}_1\bar{\Omega}_{11}\bar{\Lambda}_1^\top)\gamma_1$, where $\Lambda_{1(2)} = \gamma_{1(2)}^{-1}(\Lambda_1 + \Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})$, we obtain by using Equation (A.1) the following density for Z_1:

$$f_{Z_1}(z_1) = f_{p_1}\left(z_1; \bar{\Omega}_{11}, h^{(p_1)}\right)\frac{F_q\left(\Lambda_{1(2)}z_1 + \bar{\tau}_{1(2)}; \bar{\Gamma}_{1(2)}, h^{(q)}_{Q_1(z_1)}\right)}{F_q\left(\bar{\tau}_{1(2)}; \bar{\Gamma}_{1(2)} + \Lambda_{1(2)}\bar{\Omega}_{11}\Lambda_{1(2)}^\top\right)},$$

where $\bar{\tau}_{1(2)} = \gamma_{1(2)}^{-1}\tau$. Finally, by applying that $f_{Y_1}(y_1) = |\omega_1|^{-1}f_{Z_1}(z_1)$, where $z_1 = w_1^{-1}(Y_1 - \xi_1)$, we finish the proof for $i = 1$. The proof for $i = 2$ is analogous. ■

Alternative proof of Proposition 3.1 From the properties of elliptically contoured distributions (see, e.g., Fang et al., 1990) we have

$$\begin{aligned} f_p(y;\xi,\Omega,h^{(p)}) &= f_{p_1}(y_1;\xi_1,\Omega_{11},h^{(p_1)}) f_{p_2}(y_2;\xi_{2\cdot 1},\Omega_{22\cdot 1},h^{(q)}_{Q_1(z_1)}) \\ &= |\Omega_{11}|^{-1/2} h^{(p_1)}\{Q_1(z_1)\} |\Omega_{22\cdot 1}|^{-1/2} h^{(p_2)}_{Q_1(z_1)}\{Q_{2\cdot 1}(z_{2\cdot 1})\} \\ &= |\Omega_{11}|^{-1/2} h^{(p_1)}\{Q_1(z_1)\} |\Omega_{22\cdot 1}|^{-1/2} h^{(p_2)}_{Q_1(z_1)}\{\tilde{Q}_{2\cdot 1}(\tilde{z}_2)\}, \end{aligned}$$

where $z_i = \omega_i^{-1}(y_i - \xi_i)$, $i = 1, 2$, $z_{2\cdot 1} = \omega_{2\cdot 1}^{-1}\omega_2(z_2 - \xi_{2\cdot 1})$ and $\tilde{z}_2 = z_2 - \tilde{\xi}_{2\cdot 1}$, with

$$\begin{aligned} \xi_{2\cdot 1} &= \xi_2 + \omega_2\tilde{\xi}_{2\cdot 1}, \quad \tilde{\xi}_{2\cdot 1} = \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1} z_1, \\ \bar{\Omega}_{22\cdot 1} &= \omega_{2\cdot 1}^{-1}\Omega_{22\cdot 1}\omega_{2\cdot 1}^{-1}, \quad \omega_{2\cdot 1} = \operatorname{diag}(\Omega_{22\cdot 1})^{1/2}, \quad \Omega_{22\cdot 1} = \Omega_{22} - \Omega_{21}\Omega_{11}^{-1}\Omega_{12}, \\ \tilde{\Omega}_{22\cdot 1} &= \omega_2^{-1}\Omega_{22\cdot 1}\omega_2^{-1} = \omega_2^{-1}\omega_{2\cdot 1}\bar{\Omega}_{22\cdot 1}\omega_{2\cdot 1}\omega_2^{-1} = \bar{\Omega}_{22} - \bar{\Omega}_{21}\bar{\Omega}_{11}^{-1}\bar{\Omega}_{12}, \\ Q_1(z_1) &= z_1^\top \bar{\Omega}_{11}^{-1} z_2, \\ Q_{2\cdot 1}(z_{2\cdot 1}) &= z_{2\cdot 1}^\top \bar{\Omega}_{22\cdot 1}^{-1} z_{2\cdot 1} = (z_2 - \tilde{\xi}_{2\cdot 1})^\top \tilde{\Omega}_{22\cdot 1}^{-1}(z_2 - \tilde{\xi}_{2\cdot 1}) = \tilde{Q}_{2\cdot 1}(z_2 - \tilde{\xi}_{2\cdot 1}). \end{aligned}$$

On the other hand, since $Q(z) = Q_1(z_1) + Q_{2\cdot 1}(z_{2\cdot 1}) = Q_1(z_1) + \tilde{Q}_{2\cdot 1}(\tilde{z})$ and $\Lambda z = \Lambda_1 z_1 + \Lambda_2 z_2 = (\Lambda_1 + \Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1 + \Lambda_2\tilde{z}_2$, we have after some straightforward algebra that

$$\begin{aligned} & F_q(\Lambda z + \tau; \Gamma, h^{(q)}_{Q(z)}) \\ &= \frac{\int_{u<\Lambda z+\tau} |\Gamma|^{-1/2} h^{(p+q)}\{Q(z) + u^\top\Gamma^{-1}u\} du}{h^{(p)}\{Q(z)\}} \\ &= \frac{\int_{u<(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1+\tau} |\Gamma|^{-1/2} h^{(p_2+q)}_{Q(z_1)}\{\tilde{Q}_{2\cdot 1}(\tilde{z}_2) + (u+\Lambda_2\tilde{z}_2)^\top\Gamma^{-1}(u+\Lambda_2\tilde{z}_2)\} du}{h^{(p_2)}_{Q_1(z_1)}\{\tilde{Q}_{2\cdot 1}(\tilde{z}_2)\}}, \end{aligned}$$

where we note that

$$\begin{aligned} \tilde{Q}_{2\cdot 1}(\tilde{z}_2) + (u+\Lambda_2\tilde{z}_2)^\top\Gamma^{-1}(u+\Lambda_2\tilde{z}_2) &= (\tilde{z}_2 + A\Lambda_2^\top\Gamma^{-1}u)^\top A^{-1}(\tilde{z}_2 + A\Lambda^\top\Gamma^{-1}u) + u^\top B^{-1}u \\ &= \hat{z}_2^\top\hat{z}_2 + u^\top B^{-1}u, \end{aligned}$$

with $\hat{z}_2 = A^{-1/2}(\tilde{z}_2 + A\Lambda_2^\top\Gamma^{-1}u)$, $A = (\tilde{\Omega}_{22\cdot 1}^{-1} + \Lambda_2^\top\Gamma^{-1}\Lambda_2)^{-1}$ and $B = (\Gamma^{-1} - \Gamma^{-1}\Lambda_2 A\Lambda_2^\top\Gamma^{-1})^{-1}$. Thus, by replacing the above results given in Equation (1), we have for the marginal density of Y_1 that

$$\begin{aligned} f_{Y_1}(y_1) &= \frac{1}{F_q(\tau;\Gamma+\Lambda\bar{\Omega}\Lambda^\top, h^{(q)})}\int_{y_2} f_p(y;\xi,\Omega,h^{(p)}) F_q(\Lambda z+\tau;\Gamma,h^{(q)}_{Q(z)}) dy_2 \\ &= \frac{1}{F_q(\tau;\Gamma+\Lambda\bar{\Omega}\lambda^\top, h^{(q)})} f_{p_1}(y_1;\xi_{11},\Omega_{11},h^{(p_1)}) |\tilde{\Omega}_{22\cdot 1}|^{-1/2} |\Gamma|^{-1/2} \\ &\quad \times \int_{u<(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1+\tau} \int_{\tilde{z}_2} h^{(p_2+q)}_{Q(z_1)}\{\tilde{Q}_{2\cdot 1}(\tilde{z}_2) + (u+\Lambda_2\tilde{z}_2)^\top\Gamma^{-1}(u+\Lambda_2\tilde{z}_2)\} d\tilde{z}_2 du \end{aligned}$$

$$= \frac{1}{\mathrm{F}_q(\tau;\Gamma+\Lambda\bar{\Omega}\lambda^\top,h^{(q)})} f_{p_1}(y_1;\xi_{11},\Omega_{11},h^{(p_1)})|\tilde{\Omega}_{22\cdot 1}|^{-1/2}|A|^{1/2}|\Gamma|^{-1/2}|B|^{1/2}$$
$$\times \int_{u<(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1+\tau} \int_{\tilde{z}_2} |B|^{-1/2} h^{(p_2+q)}_{Q(z_1)}(\hat{z}_2^\top \hat{z}_2 + u^\top B^{-1}u)d\tilde{z}_2 du$$
$$= \frac{1}{F_q(\tau;\Gamma+\Lambda\bar{\Omega}\lambda^\top,h^{(q)})} f_{p_1}(y_1;\xi_{11},\Omega_{11},h^{(p_1)})|\tilde{\Omega}_{22\cdot 1}|^{-1/2}|A|^{1/2}|\Gamma|^{-1/2}|B|^{1/2}$$
$$\times \int_{u<(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1+\tau} \int_{\tilde{z}_2} |B|^{-1/2} h^{(q)}_{Q(z_1)}(u^\top B^{-1}u)du$$
$$= \frac{1}{F_q(\tau;\Gamma+\Lambda\bar{\Omega}\Lambda^\top,h^{(q)})} f_{p_1}(y_1;\xi_{11},\Omega_{11},h^{(p_1)})$$
$$\times F_q\{(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})z_1+\tau;\Gamma+\Lambda_2\tilde{\Omega}_{22\cdot 1}\Lambda_2^\top, h^{(q)}_{Q_1(z_1)}\},$$

where we used that $|\tilde{\Omega}_{22\cdot 1}|^{-1/2}|A|^{1/2}|\Gamma|^{-1/2}|B|^{1/2} = 1$, since $B = \Gamma + \Lambda_2\tilde{\Omega}_{22\cdot 1}\Lambda_2^\top$ and $|A| = |\tilde{\Omega}_{22\cdot 1}||\Gamma||B|^{-1}$. By noting now that the dispersion matrix $\Gamma+\Lambda_2\tilde{\Omega}_{22\cdot 1}\Lambda_2^\top$ induces the correlation $\Gamma_{1(2)} = \gamma_{1(2)}^{-1}(\Gamma+\Lambda_2\tilde{\Omega}_{22\cdot 1}\Lambda_2^\top)\gamma_{1(2)}^{-1}$, where $\gamma_{1(2)} = \mathrm{diag}(\Gamma+\Lambda_2\tilde{\Omega}_{22\cdot 1}\Lambda_2^\top)^{1/2}$, and also that $\Gamma+\Lambda\bar{\Omega}\Lambda^\top = \gamma_{1(2)}(\bar{\Gamma}_{1(2)}+\Lambda_{1(2)}\bar{\Omega}_{11}\Lambda_{1(2)})\gamma_{1(2)}$, where $\Lambda_{1(2)} = \gamma_{1(2)}^{-1}(\Lambda_1+\Lambda_2\bar{\Omega}_{21}\bar{\Omega}_{11}^{-1})$, we obtain finally that

$$f_{Y_1}(y_1) = f_{p_1}(y_1;\xi_{11},\Omega_{11},h^{(p_1)}) \frac{F_q(\Lambda_{1(2)}z_1+\bar{\tau}_{1(2)};\bar{\Gamma}_{1(2)},h^{(q)}_{Q_1(z_1)})}{F_q(\bar{\tau}_{1(2)};\bar{\Gamma}_{1(2)}+\Lambda_{1(2)}\bar{\Omega}_{11}\Lambda_{1(2)}^\top,h^{(q)})},$$

where $\bar{\tau}_{1(2)} = \gamma_{1(2)}^{-1}\tau$, therefore ending the proof for the marginal distribution of Y_1. The proof for Y_2 is analogous. ■

Proof of Proposition 3.2 Considering that $Y = \xi + \omega Z$, where $Z \overset{d}{=} (X|X_0 < \Lambda X + \tau)$ and the partition given in the proof of Proposition 3.1, we obtain

$$(Y_1|Y_2 = y_2) = \{\xi_1 + w_1 Z_1 | Z_2 = \omega_2^{-1}(y_2-\xi_2)\} = \xi_{1\cdot 2} + \omega_{1\cdot 2}(Z_{1\cdot 2}|Z_2 = z_2),$$

where $\xi_{1\cdot 2} = \xi_1 + \omega_1\tilde{\xi}_{1\cdot 2}z_2$, $\tilde{\xi}_{1\cdot 2} = \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}z_2$, $z_2 = \omega_2^{-1}(y_2-\xi_2)$, $Z_{1\cdot 2} = \omega_{1\cdot 2}^{-1}\omega_1(Z_1 - \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}Z_2)$, $Z_2 = \omega_2^{-1}(Y_2-\xi_2)$. We note that $\bar{\Omega}_{11\cdot 2} = \bar{\Omega}_{11} - \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}\bar{\Omega}_{21} = \omega_1^{-1}\omega_{1\cdot 2}\bar{\Omega}_{11\cdot 2}\omega_{1\cdot 2}\omega_1^{-1}$, where $\omega_{1\cdot 2}$ and $\bar{\Omega}_{11\cdot 2}$ are the scale and correlation matrices induces by $\Omega_{11\cdot 2}$, respectively. Thus, we have

$$f_{Y_1|Y_2=y_2}(y_1) = |\omega_{1\cdot 2}|^{-1} f_{Z_{1\cdot 2}|Z_2=z_2}(z_{1\cdot 2}), \tag{A.2}$$

where $z_{1\cdot 2} = w_{1\cdot 2}^{-1}\omega_1(z_1 - \bar{\xi}_{1\cdot 2})$, with $z_1 = \omega_1^{-1}(y_1-\xi_1)$. But, since $(Z_{1\cdot 2}|Z_2 = z_2) \overset{d}{=} (X_{1\cdot 2}|X_{0\cdot 2} < \Lambda_1\omega_1^{-1}\omega_{1\cdot 2}X_{1\cdot 2}+\tau, X_2 = z_2)$, where $X_{1\cdot 2} = \omega_{1\cdot 2}^{-1}\omega_1(X_1 - \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}X_2)$ and $X_{0\cdot 2} = X_0 - (\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}+\Lambda_2)X_2$, the conditional density of $(Z_{1\cdot 2}|Z_2 = z_2)$ can be computed as

$$f_{Z_{1\cdot 2}|Z_2=z_2}(z_{1\cdot 2}) = \frac{1}{\mathrm{P}(X_{0\cdot 2} < \Lambda_1 X_{1\cdot 2}+\tau|X_2 = z_2)} f_{X_{1\cdot 2}|X_2=z_2}(z_{1\cdot 2})$$
$$\times \mathrm{P}(X_0 < \Lambda_1 z_{1\cdot 2}+\tau|X_2 = z_2, X_{1\cdot 2} = z_{1\cdot 2}),$$

by using that

$$\left\{\begin{pmatrix} X_{1\cdot 2} \\ X_{0\cdot 2}\end{pmatrix} \Big| X_2 = z_2\right\} = \left\{\begin{pmatrix} \omega_{1\cdot 2}^{-1}\omega_1(X_1 - \bar{\Omega}_{12}\bar{\Omega}_{22}^{-1}X_2) \\ X_0 - (\Lambda_2 + \Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1})X_2\end{pmatrix} \Big| X_2 = z_2\right\}$$
$$\sim \mathrm{EC}_{p+q}\left(\begin{pmatrix} 0 \\ -(\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \Lambda_2)z_2\end{pmatrix}, \begin{pmatrix}\bar{\Omega}_{11\cdot 2} & 0 \\ 0 & \Gamma\end{pmatrix}, h^{(p_1+q)}_{Q_2(z_2)}\right),$$

where $Q_2(z_2) = z_2^\top\bar{\Omega}_{22}^{-1}z_2$. Thus, since this last result implies $(X_{0\cdot 2}|X_2 = z_2, X_{1\cdot 2} = z_{1\cdot 2}) \sim \mathrm{EC}_q(0, \Gamma, h^{(q)}_{Q_2(z_2)+Q_{1\cdot 2}(z_{1\cdot 2})})$, where $Q_2(z_2) + Q_{1\cdot 2}(z_{1\cdot 2}) = z_2^\top\bar{\Omega}_{22}^{-1}z_2 + z_{1\cdot 2}^\top\bar{\Omega}_{11\cdot 2}^{-1}z_{1\cdot 2} = z^\top\bar{\Omega}^{-1}z = Q(z)$, $(X_{1\cdot 2}|X_2 = z_2) \sim \mathrm{EC}_{p_1}(0, \bar{\Omega}_{11\cdot 2}, h^{(p_1)}_{Q_2(z_2)})$ and $\{(X_{0\cdot 2} - \Lambda_1X_{1\cdot 2})|X_2 = z_2\} \sim \mathrm{EC}_q(-(\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \Lambda_2)z_2, \Gamma + \Lambda_1\bar{\Omega}_{11\cdot 2}\Lambda_1^\top, h^{(q)}_{Q_2(z_2)})$, we obtain

$$f_{Z_{1\cdot 2}|Z_2=z_2}(z_{1\cdot 2}) = \frac{1}{F_q\left(\tau_{1\cdot 2}; \Gamma + \Lambda_1\bar{\Omega}_{11\cdot 2}\Lambda_1^\top, h^{(q)}_{Q_2(z_2)}\right)} f_{p_1}\left(z_{1\cdot 2}; 0, \bar{\Omega}_{11\cdot 2}, h^{(p_1)}_{Q_2(z_2)}\right)$$
$$\times F_q\left(\Lambda_1\omega_1^{-1}\omega_{1\cdot 2}z_{1\cdot 2} + \tau_{1\cdot 2}; \Gamma, h^{(q)}_{Q_2(z_2)}\right), \tag{A.3}$$

where $\tau_{1\cdot 2} = (\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \Lambda_2)z_2 + \tau$. The proof follows by replacing Equation (A.3) in Equation (A.2). ■

Alternative proof of Proposition 3.2 Following the alternative proof of Proposition 3.1 and a similar notation as used there, we note that

$$\frac{f_p(y; \xi, \Omega, h^{(p)})}{f_{p_2}(y_2; \xi_2, \Omega_{22}, h^{(p_2)})} = f_{p_1}(y_1; \xi_{1\cdot 2}, \Omega_{11\cdot 2}, h^{(p_1)}_{Q_2(z_2)}),$$

and

$$f_{Y_2}(y_2) = \frac{1}{F_q(\tau; \Gamma + \Lambda\bar{\Omega}\Lambda^\top, h^{(q)})} f_{p_2}(y_2; \xi_{22}, \Omega_{22}, h^{(p_2)})F_q(\Lambda_{2(1)}z_2 + \bar{\tau}_{2(1)}; \bar{\Gamma}_{2(1)}, h^{(q)}_{Q_2(z_2)}),$$

implying that

$$f_{Y_1|Y_2=y_2}(y_1) = \frac{1}{F_q(\Lambda_{2(1)}z_2 + \bar{\tau}_{2(1)}; \bar{\Gamma}_{2(1)}, h^{(q)}_{Q_2(z_2)})} f_{p_1}(y_1; \xi_{1\cdot 2}, \Omega_{11\cdot 2}, h^{(p_1)}_{Q_2(z_2)})F_q(\Lambda z + \tau; \Gamma, h^{(q)}).$$

Thus, because $F_q(\Lambda_{2(1)}z_2 + \bar{\tau}_{2(1)}; \bar{\Gamma}_{2(1)}, h^{(q)}_{Q_2(z_2)}) = F_q\left(\tau_{1\cdot 2}; \Gamma + \Lambda_1\bar{\Omega}_{11\cdot 2}\Lambda_1^\top, h^{(q)}_{Q_2(z_2)}\right)$ and also $F_q(\Lambda z + \tau; \Gamma, h^{(q)}) = F_q\left(\Lambda_1\omega_1^{-1}\omega_{1\cdot 2}z_{1\cdot 2} + \tau_{1\cdot 2}; \Gamma, h^{(q)}_{Q_2(z_2)}\right)$, where $z_{1\cdot 2} = \omega_{1\cdot 2}^{-1}(y_1 - \xi_{1\cdot 2})$ and $\tau_{1\cdot 2} = (\Lambda_1\bar{\Omega}_{12}\bar{\Omega}_{22}^{-1} + \Lambda_2)z_2 + \tau$, the proof follows. ■

Proof of Proposition 4.1 Let $X_A = \omega_A^{-1}A\omega X$ and $X_{0\cdot A} = X_0 - \Lambda X + \Lambda_AX_A$, where $\omega_A = \mathrm{diag}(\Omega_A)^{1/2}$, with $\Omega_A = A\Omega A^\top$, and $\Lambda_A = \Lambda\bar{\Omega}\omega A^\top\omega_A^{-1}\bar{\Omega}_A^{-1}$. From the properties of EC distributions, we have after some algebraic manipulations that

$$\begin{pmatrix} X_A \\ X_{0A}\end{pmatrix} \sim \mathrm{EC}_{r+q}\left(\begin{pmatrix} 0 \\ 0\end{pmatrix}, \begin{pmatrix}\bar{\Omega}_A & 0 \\ 0 & \Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top\end{pmatrix}, h^{(r+q)}\right),$$

where $\bar{\Omega}_A = \omega_A^{-1}\Omega_A\omega_A^{-1}$. Thus, by letting $Z_A = \omega_A^{-1}A\omega Z$, where $Z \stackrel{d}{=} (X|X_0 < \Lambda X + \tau)$, we obtain that $Z_A \stackrel{d}{=} (X_A|X_{0\cdot A} < \Lambda_A X_A + \tau) \stackrel{d}{=} (X_A|\bar{X}_{0\cdot A} < \bar{\Lambda}_A X_A + \bar{\tau}_A)$, where $\bar{X}_{0\cdot A} = \gamma_A^{-1}X_{0\cdot A}$, $\bar{\Lambda}_A = \gamma_A^{-1}\Lambda_A$, $\bar{\tau}_A = \gamma_A^{-1}\tau$ and $\gamma_A = \mathrm{diag}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top)^{1/2}$. From Equation (A.1), we have

$$f_{Z_A}(z) = \frac{1}{F_q(\bar{\tau}_A; \bar{\Gamma}_A + \bar{\Lambda}_A\bar{\Omega}_A\bar{\Lambda}_A^\top, h^{(q)})} f_r(z; 0, \bar{\Omega}_A, h^{(r)}) F_q(\bar{\Lambda}_A z + \bar{\tau}_A; \bar{\Gamma}_A, h^{(q)}_{Q_A(z)}),$$

where $\bar{\Gamma}_A = \gamma_A^{-1}(\Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top)\gamma_A^{-1}$ is the correlation matrix associated with $\Gamma + \Lambda\bar{\Omega}\Lambda^\top - \Lambda_A\bar{\Omega}_A\Lambda_A^\top$. Thus the proof follows by noting that $AY + b = \xi_A + \omega_A Z_A$, where $\xi_A = A\xi + b$. ■

Proof of Proposition 4.2 Let $Y = (Y_1^\top, Y_2^\top)^\top$ and consider the matrices $A_1 = (I_{p_1}, 0)$ and $A_2 = (0, I_{p_2})$. Since $Y_i = A_i Y$, $i = 1, 2$, where

$$Y \sim \mathrm{SUE}_{p,q}((\xi_1^\top, \xi_2^\top)^\top, \mathrm{diag}(\Omega_1, \Omega_2), \mathrm{diag}(\Lambda_1, \Lambda_2), h^{(p+q)}, (\tau_1^\top, \tau_2^\top)^\top, \mathrm{diag}(\Gamma_1, \Gamma_2)),$$

with $p = p_1 + p_2$, $q = q_1 + q_2$, the result for the marginal distribution of Y_i follows by applying the first part of Proposition 4.1. From that part, we obtain also the distribution of the sum of Y_1 and Y_2 when $p_1 = p_2 = r$, since $Y_1 + Y_2 = AY$, where $A = (I_r, I_r)$. ■

References

Adcock, C.J., 2010. Asset pricing and portfolio selection based on the multivariate extended skew-Student-t distribution. Annals of Operations Research, 176, 221-234.

Arellano-Valle, R.B., Azzalini, A., 2006. On the unification of families of skew-normal distributions. Scandinavian Journal of Statistics, 33, 561-574.

Arellano-Valle, R.B., Bolfarine, H., 1995. On some characterizations of the t-distribution. Statistical and Probability Letters, 25, 79-85.

Arellano-Valle, R.B., Bolfarine, H., Iglesias, P.L., 1994. A predictivistic interpretation of the multivariate t distribution. Test, 3, 221-236.

Arellano-Valle, R.B., Branco, M.D., Genton, M.G., 2006a. A unified view on skewed distributions arising from selections. Canadian Journal of Statistics, 34, 581-601.

Arellano-Valle, R.B., del Pino, G., Iglesias, P., 2006b. Bayesian inference in spherical linear models: robustness and conjugate analysis. Journal of Multivariate Analysis, 97, 179-197.

Arellano-Valle, R.B., del Pino, G., San Martín, E., 2002. Definition and probabilistic properties of skew distributions. Statistical and Probability Letters, 58, 111-121.

Arellano-Valle, R.B., Genton, M.G., 2010a. Multivariate extended skew-t distributions and related families. Metron – special issue on "Skew-symmetric and flexible distributions" (in press).

Arellano-Valle, R.B., Genton, M.G., 2010b. An invariance property of quadratic forms in random vectors with a selection distribution, with application to sample variogram and covariogram estimators. Annals of the Institute of Statistical Mathematics, 62, 363-381.

Azzalini, A., 1986. Further results on a class of distributions which includes the normal ones. Statistica, 46, 199-208.

Azzalini, A., 2005. The skew-normal distribution and related multivariate families. Scandinavian Journal of Statistics, 32, 159-200.

Azzalini, A., Capitanio, A., 1999. Statistical applications of the multivariate skew-normal distribution. Journal of The Royal Statistical Society Series B – Statistical Methodology, 61, 579-602.

Azzalini, A., Capitanio, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. Journal of The Royal Statistical Society Series B – Statistical Methodology, 65, 367-389.

Azzalini, A., Dalla Valle, A., 1996. The multivariate skew-normal distribution. Biometrika, 83, 715-726.

Branco, M.D., Dey, D.K., 2001. A general class of multivariate skew-elliptical distributions. Journal of Multivariate Analysis, 79, 99-113.

Castro, L.M., Loschi, R.H., Arellano-Valle, R.B., 2010a. Bayesian inference for the skew-normal shape parameter: an application to change point problems. Submitted.

Castro, L.M., San Martín, E., Arellano-Valle, R.B., 2010b. A note on the parameterization of multivariate skewed normal distributions. Submitted.

Díaz, J.A., Leiva, V., Galea, M., 2002. Singular elliptical distribution: density and applications. Communications in Statistics – Theory and Methods, 31, 665-681.

Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric Multivariate and Related Distributions. Chapman & Hall, London.

Genton, M.G., 2004. Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. Chapman & Hall/CRC, Florida.

Genton, M.G., He, L., Liu, X., 2001. Moments of skew-normal random vectors and their quadratic forms. Statistical and Probability Letters, 51, 319-325.

Genton, M.G., Loperfido, N., 2005. Generalized skew-elliptical distributions and their quadratic forms. Annals of the Institute of Statistical Mathematics, 57, 389-401.

González-Farías, G., Domínguez, J.A., Gupta, A.K., 2004. The closed skew-normal distribution. In Genton, M.G., (ed.). Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality. M.Chapman & Hall/CRC, FL, USA, pp. 25-41.

Lee, S., Genton, M.G., Arellano-Valle, R.B., 2010. Perturbation of numerical confidential data via skew-t distributions. Management Science, 56, 318-333.

Ma, Y., Genton, M.G., 2004. A flexible class of skew-symmetric distributions. Scandinavian Journal of Statistics, 31, 459-468.

Mardia, K.V., 1970. Measures of multivariate skewness and kurtosis with applications. Biometrika, 36, 519-530.

Rao, C.R., 1965. Linear Statistical Inference and its Applications. Wiley, New York.

Wang, J., Boyer, J., Genton, M.G., 2004a. A skew-symmetric representation of multivariate distributions. Statistica Sinica, 14, 1259-1270.

Wang, J., Boyer, J., Genton, M.G., 2004b. A note on an equivalence between chi-square and generalized skew-normal distributions. Statistical and Probability Letters, 66, 395-398.