Geometric representations of multivariate skewed elliptically contoured distributions

Wolf-Dieter Richter* and John Venz*

* University of Rostock, Institute of Mathematics, Ulmenstraße 69, Haus 3, 18057 Rostock, Germany

Abstract

We derive a wide class of geometric representation formulas for multivariate skewed elliptically contoured distributions and show in a unified geometric way how some of them are related to stochastic representations known in the literature. Furthermore, we make use of the geometric measure representation to explore independence between collections of components of accordingly distributed random vectors, and to investigate contour plots of skewed normal densities from a geometric viewpoint.

Key words: skew-elliptical distribution; stochastic representation; geometric representation; density contour plot

2010 Mathematics Subject Classification: 60E05; 62H05; 60D05; 28A05; 60E20; 49G03; 49F10

1 Introduction

Over the last decade, the field of multivariate skewed distributions was a very vibrant research area. The first well studied type of such distributions is the multivariate skewed normal distribution that is considered in [6] and [14]. Later on, many authors tackled different approaches to generalize this distribution. A very important generalization for the purposes of the present paper is the class of multivariate skewed elliptical distributions introduced in [10]. Because of the vast development of the area of skewed distributions, several authors as those in [1],[2] and [4] put a lot of effort into finding as general and systematic approaches to it as possible.

Recently, the authors of [13] provided an approach to the univariate skewed normal and univariate skewed elliptical distributions that unifies several known representations of these distributions from a certain geometric point of view. At the same time, this viewpoint makes it possible to establish a whole class of new such representations. The aim of the present paper consists in extending this geometric way of dealing with skewed distributions to the multivariate case, and in drawing first consequences from it.

To be more concrete, in the following, we denote by $\Phi_k(\cdot; g^{(k)})$ a continuous spherical distribution on the Borel σ-field $\mathfrak{B}^{(k)}$ in the Euclidean space $\mathbb{R}^k$ having the density generator $g^{(k)}$ and by $SE_k(\boldsymbol{\xi}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$ a member of the class of continuous skewed elliptical distributions on $\mathfrak{B}^{(k)}$ as it was introduced in [10]. The results in [13] show the following. If $Z \sim SE_1(0, 1, \delta; g^{(2)})$, then its cumulative distribution function (cdf) allows each of the representations

$$P(Z < z) = 2\Phi_2(C(a, b, c, d, e); g^{(2)}), \tag{1}$$

where the cone

$$C(a, b, c, d, e) = H_1(a, b) \cap H_2(c, d, e)$$

is the intersection of two half spaces of $\mathbb{R}^2$,

$$H_1(a, b) = \{(x, y)^T \in \mathbb{R}^2 : ax + by < 0\}$$

and

$$H_2(c, d, e) = \{(x, y)^T \in \mathbb{R}^2 : cx + dy < e\},$$

and where the parameters a, b, c, d and e fulfill the equations

$$z = \frac{e}{\sqrt{c^2 + d^2}} \tag{2}$$

and

$$\delta = -\frac{ac + bd}{\sqrt{a^2 + b^2}\sqrt{c^2 + d^2}}. \tag{3}$$

In other words, if $(X, Y) \sim \Phi_2(\cdot; g^{(2)})$ then for all parameters satisfying equations (2) and (3),

$$P(Z < z) = 2P(aX + bY < 0,\ cX + dY < e),$$

and

$$P(Z < z) = P(cX + dY < e \mid aX + bY < 0).$$

Notice that for every given pair (δ, z), there are uncountably many solutions (a, b, c, d, e) of equations (2) and (3) corresponding to orthogonally transformed cones $C(a,b,c,d,e)$, hence each giving rise to its own representation formula of the cdf of a one-dimensional skewed elliptically distributed random variable. Let

$$\mathfrak{C}_2(\delta,z)=\{C(a,b,c,d,e)\ :\ a,b,c,d,e \text{ satisfy (2) and (3)}\}$$

be the class of cones whose two-dimensional spherical measure coincides with $(1/2)\cdot P(Z<z)$ where $Z\sim SE_1(0,1,\delta;g^{(2)})$. Then $\mathfrak{C}_2(\delta_1,z_1)\cap\mathfrak{C}_2(\delta_2,z_2)=\emptyset$ if $(\delta_1,z_1)\neq(\delta_2,z_2)$. The value of the parameter δ in (3) is equal to that of the cosine of the angle between the vectors $(-a,-b)^T$ and $(c,d)^T$. This angle can be considered as the opening angle of the cone $C(a,b,c,d,e)$. Furthermore, the absolute value of the parameter z in (2) is equal to that of the distance from the line $\partial H_2(c,d,e)$ to the origin. The origin always belongs to the boundary of $H_1(a,b)$. Note that we can also write $H_2(c,d,e)=\{(x,y)^T\in\mathbb{R}^2 : \frac{c}{\sqrt{c^2+d^2}}x+\frac{d}{\sqrt{c^2+d^2}}y<z\}$ and therefore

$$C(a,b,c,d,e)=\{(x,y)^T\in\mathbb{R}^2 : \boldsymbol{a}_0^T\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)<0,\ \boldsymbol{a}_1^T\left(\begin{smallmatrix}x\\y\end{smallmatrix}\right)<z\},\qquad(4)$$

where $\boldsymbol{a}_0^T=(a,b)$ and $\boldsymbol{a}_1^T=(\frac{c}{\sqrt{c^2+d^2}},\frac{d}{\sqrt{c^2+d^2}})$ is a normalized vector.

The proof of the above statement is immediate from the proofs of Theorems 1 and 2 in [13] and the comments given there at the end of Section 4. The results of [13] can now be considered as special cases of the present formulation by suitably choosing $g^{(2)}$, and $\delta=\nu/\sqrt{1+\nu^2}$.

To give a first impression of how the representation formula (1) can be used, we derive with its help a specific representation for skewed elliptically distributed random variables based on the maximum of two jointly elliptically contoured distributed random variables.

It was demonstrated in [13] that different representations of skewed elliptically contoured distributions can be derived in the same unified geometric way and it will be shown later in the present paper that this unified way of proving stochastic representations works in higher dimensions, too.

Before we go further, we recall that a continuous k-dimensional random vector $\boldsymbol{Y}$ is called elliptically contoured distributed with location parameter $\boldsymbol{\mu}\in\mathbb{R}^k$, symmetric regular form or scale parameter matrix $\Sigma\in\mathbb{R}^{k\times k}$ and density generator $g^{(k)}:\mathbb{R}^+\to\mathbb{R}^+$ if it has the density

$$f(\boldsymbol{y};\boldsymbol{\mu},\Sigma,g^{(k)})=|\Sigma|^{-1/2}g^{(k)}((\boldsymbol{y}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{y}-\boldsymbol{\mu})),\quad \boldsymbol{y}\in\mathbb{R}^k.$$

We write $\boldsymbol{Y}\sim EC_k(\boldsymbol{\mu},\Sigma;g^{(k)})$ for short. For a treatment of elliptically contoured distributions, we refer to [11].

In the following, we show that if $(X,Y)^T \sim EC_2(\mathbf{0}_2, \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right); g^{(2)})$ then $\max\{X,Y\} \sim SE_1(0,1,\{1/2(1-\rho)\}^{1/2}; g^{(2)})$ by making use of the cdf representation (1) of the univariate skewed elliptical distribution. To this end, we define the diagonal matrix $D = diag(1/\sqrt{1+\rho}, 1/\sqrt{1-\rho})$ and the orthogonal matrix $O = (1/\sqrt{2})\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ and note that the transformed vector $(U,V)^T = DO(X,Y)^T$ satisfies

$$(U,V)^T \sim \Phi_2(\cdot; g^{(2)}).$$

Since, first, $P(\max\{X,Y\} < t)$ can be written as $P((X,Y)^T \in B(t))$ with $B(t) = \{(x,y)^T \in \mathbb{R}^2 : x < t, y < t\}$, and next, $(X,Y)^T \in B(t)$ holds iff $(U,V)^T \in DOB(t)$, we obtain

$$P(\max\{X,Y\} < t) = \Phi_2(DOB(t); g^{(2)}), t \in \mathbb{R}. \tag{5}$$

Note that $DOB(t)$

$$= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : (\frac{1+\rho}{2})^{1/2}x + (\frac{1-\rho}{2})^{1/2}y < t,\ (\frac{1+\rho}{2})^{1/2}x - (\frac{1-\rho}{2})^{1/2}y < t \right\}$$

is a cone in $\mathbb{R}^2$ that is symmetric w.r.t. the x-axis. With the notation

$$\tilde{B}(t) = \left\{ (x,y)^T \in \mathbb{R}^2 : y > 0, (\frac{1+\rho}{2})^{1/2}x + (\frac{1-\rho}{2})^{1/2}y < t \right\},$$

we have $DOB(t) = \overline{\tilde{B}(t)} \cup \left[\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{B}(t)\right]$ and $\tilde{B}(t) \cap \left[\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{B}(t)\right] = \emptyset$, where the origin belongs to the topological boundary of the cone $\tilde{B}(t)$ and $\overline{\tilde{B}(t)}$ denotes the closure of $\tilde{B}(t)$. Thus

$$\Phi_2(DOB(t); g^{(2)}) = \Phi_2(\tilde{B}(t); g^{(2)}) + \Phi_2(\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\tilde{B}(t); g^{(2)}).$$

Therefore,

$$\Phi_2(DOB(t); g^{(2)}) = 2\Phi_2(\tilde{B}(t); g^{(2)}). \tag{6}$$

Here, $\tilde{B}(t)$ is the cone $C(a,b,c,d,e)$ with parameters $a = 0,\ b = -1,\ c = (\frac{1+\rho}{2})^{1/2},\ d = (\frac{1-\rho}{2})^{1/2}$ and $e = t$. Inserting these values into (2) and (3), we obtain $z = t$ and $\delta = \{1/2(1-\rho)\}^{1/2}$ and from representation (1) then follows that

$$2\Phi_2(\tilde{B}(t); g^{(2)}) = P(Z < t), \tag{7}$$

where $Z \sim SE_1(0, 1, \{1/2(1-\rho)\}^{1/2}, g^{(2)})$. Hence, on combining (5), (6) and (7), we observe that

$$P(\max\{X, Y\} < t) = P(Z < t), \quad t \in \mathbb{R},$$

i.e. the maximum statistic follows an univariate skewed elliptical distribution with parameters $(\xi, \Omega) = (0, 1)$ and skewing parameter $\delta = \{1/2(1-\rho)\}^{1/2}$ if the two-dimensional elliptically contoured sample distribution has location parameter $\boldsymbol{\mu} = \mathbf{0}_2$ and its scale parameter matrix Σ is actually a correlation matrix, $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$. This explicit result may be also derived from Proposition 10 in [8] concerning classes of distributions, and its proof given there. Corresponding results assigning the one-dimensional skewed normal or skewed spherical distribution to the maximum distribution of two-dimensional Gaussian or spherical vectors are due to [16], [3], and [15]. For certain generalizations of such results that are based upon a representation of skewed $l_{n,p}$-symmetric distributions in [5], we refer to [9] and [17].

In the first proof of the main result of the present paper, we follow the approach in [12] of defining a skewed elliptically contoured distribution by stating its density. We say that a k-dimensional random vector $\boldsymbol{Z}$ is distributed according to the skewed elliptically contoured distribution $SE_k(\boldsymbol{\xi}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$, where $\boldsymbol{\xi} \in \mathbb{R}^k$, $\boldsymbol{\delta} \in \mathbb{R}^k$, Ω is a symmetric and positive definite (s.p.d.) $k \times k$ matrix, $\boldsymbol{\delta}$ and Ω fulfill $\boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta} < 1$, and $g^{(k+1)} : \mathbb{R}^+ \to \mathbb{R}^+$ is the density generator of an elliptically contoured distribution in $\mathbb{R}^{k+1}$, if it has the density

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = 2|\Omega|^{-1/2} \int_{-\infty}^{\boldsymbol{\lambda}^T(\boldsymbol{z}-\boldsymbol{\xi})} g^{(k+1)}(s^2 + (\boldsymbol{z}-\boldsymbol{\xi})^T \Omega^{-1} (\boldsymbol{z}-\boldsymbol{\xi}))\, ds, \qquad (8)$$

where

$$\boldsymbol{\lambda} = (1 - \boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta})^{-1/2} \Omega^{-1} \boldsymbol{\delta}. \qquad (9)$$

We mention here that the distribution $SE_k(\boldsymbol{\xi}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$ was originally introduced in [10] in another way and refer for details to Proposition 1 in Section 2. Additionally, throughout the rest of the paper, we assume that the matrix Ω is a correlation matrix. Note that our assumptions imply that $\left(\begin{smallmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{smallmatrix}\right)$ is positive definite and that $\boldsymbol{\delta} \in (-1, 1)^{\times k}$. We will prove this in Appendix A.

The main aim of the present paper is to establish a multivariate generalization of the results in [13] as they were stated in the first part of this section. For this purpose, we generalize the class of cones $\mathfrak{C}_2(\delta, z)$ to a proper multivariate version. This will be done in Section 2. Following this line, we present then in Theorem 1 a multivariate extension of the representation (1). In Section 3, we investigate relations between the geometric measure representations proved in the present paper and some stochastic representations of skewed elliptically contoured distributed random vectors already known from the literature. Actually, we show that the latter may be derived in a unified geometric way from the new representation in Theorem 1. In Section 4, we further exploit the geometric representation formula in Theorem 1. First, we formulate geometrically stated conditions for the independence of collections of components of a skewed normal random vector. Next, the application of Theorem 1 will be extended to a greater class of cones through symmetrization. Finally, we give some new interpretations for density contour plots of two-dimensional skewed normal vectors.

2 Main result

As announced, we now introduce more general classes of cones than the one considered in Section 1. The cones studied in this section are intersections of $k+1$ half spaces from $\mathbb{R}^{k+1}$, where at least one of them contains the origin in its boundary. Each of the classes of cones $C(\boldsymbol{z})$ will be used to represent the cdf of the k-dimensional skewed elliptically contoured distribution $SE_k(\mathbf{0}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$ by the values $2\Phi_{k+1}(C(\boldsymbol{z}); g^{(k+1)}),\ \boldsymbol{z} \in \mathbb{R}^k$. To this end, we specify the cones that are needed to formulate a suitable multivariate generalization of formula (1). The half space

$$H_0(\boldsymbol{a}_0) = \{\boldsymbol{y} \in \mathbb{R}^{k+1} \ : \ \boldsymbol{a}_0^T \boldsymbol{y} < 0\}, \quad \boldsymbol{a}_0 \in \mathbb{R}^{k+1},$$

contains the origin in its boundary while the boundaries of the half spaces

$$H_i(\boldsymbol{a}_i, z_i) = \{\boldsymbol{y} \in \mathbb{R}^{k+1} \ : \ \boldsymbol{a}_i^T \boldsymbol{y} < z_i\}, \quad z_i \in \mathbb{R},\ \boldsymbol{a}_i \in \mathbb{R}^{k+1},\ i = 1, ..., k,$$

do not contain the origin, in general. The vectors $\boldsymbol{a}_0, ..., \boldsymbol{a}_k$ are assumed to be linearly independent and the vectors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_k$ are assumed to satisfy the normalization assumption

$$||\boldsymbol{a}_i|| = 1, \quad i = 1, ..., k.$$

One of the consequences is that in the case $k = 1$ equation (2) reads as $z = e$. The cones of interest are now

$$C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}) = H_0(\boldsymbol{a}_0) \cap (\bigcap_{i=1}^{k} H_i(\boldsymbol{a}_i, z_i)), \boldsymbol{z} \in \mathbb{R}^k.$$

We say that a cone $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ belongs to the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$, where the parameters $\boldsymbol{\delta} = (\delta_1, ..., \delta_k)$ and $\boldsymbol{z}$ belong to $\mathbb{R}^k$ and $\Omega = (\omega_{i,j})_{i,j=1,...,k}$ is a $k \times k$ s.p.d. correlation matrix, and $\boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta} < 1$, if the vectors $\boldsymbol{a}_0, ..., \boldsymbol{a}_k$ satisfy the equations

$$\begin{aligned} \delta_i &= \frac{-\boldsymbol{a}_0^T \boldsymbol{a}_i}{||\boldsymbol{a}_0||}, \qquad\qquad i = 1, ..., k, \\ \omega_{i,j} = \omega_{j,i} &= \boldsymbol{a}_i^T \boldsymbol{a}_j, \quad i < j,\ i, j = 1, ..., k. \end{aligned} \tag{10}$$

If $\phi_{0,i} \in (0, \pi)$ denotes the angle between $\partial H_0(\boldsymbol{a}_0)$ and $\partial H_i(\boldsymbol{a}_i, z_i)$, $i = 1, ..., k$, and $\phi_{i,j} \in (0, \pi)$ is the angle between $\partial H_i(\boldsymbol{a}_i, z_i)$ and $\partial H_j(\boldsymbol{a}_j, z_j)$, $i, j = 1, ..., k$ then (10) means

$$\delta_i = \cos(\phi_{0,i}) \text{ and } \omega_{i,j} = \omega_{j,i} = -\cos(\phi_{i,j}), i < j,\ i, j = 1, ..., k.$$

In case of $k = 1$, $\Omega = 1$ is the only admissible value for Ω. Note that $\mathfrak{C}_2(1, \delta, z)$ is equal to $\mathfrak{C}_2(\delta, z)$ from Section 1.

In the following, we generalize representation (1) to the multivariate setting. The inequality $\boldsymbol{u} < \boldsymbol{v}$, where $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^k$, is to be read componentwise.

Theorem 1. *If $\boldsymbol{Z} \sim SE_k(\boldsymbol{0}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$ then, for all cones $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$, the cdf of $\boldsymbol{Z}$ allows the representation*

$$P(\boldsymbol{Z} < \boldsymbol{z}) = 2\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}), \qquad \boldsymbol{z} \in \mathbb{R}^k. \tag{11}$$

One may say that the parameters of the cdf of the skewed elliptically contoured distribution are expressed in this theorem in terms of geometric parameters of the cones $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$, and vice versa. Similarly to equation (4), we can write

$$C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}) = \{\boldsymbol{y} \in \mathbb{R}^{k+1} : \boldsymbol{a}_0^T \boldsymbol{y} < 0,\ \boldsymbol{a}_1^T \boldsymbol{y} < z_1, ..., \boldsymbol{a}_k^T \boldsymbol{y} < z_k\}. \tag{12}$$

The absolute value of z_i is the distance of $\partial H_i(\boldsymbol{a}_i, z_i)$ from the origin, $i = 1, ..., k$. Here, $\min\{\|\boldsymbol{w}\| : \boldsymbol{a}^T \boldsymbol{w} = z\} = |z|/\|\boldsymbol{a}\|$ is the distance of the hyperplane $\{\boldsymbol{w} \in \mathbb{R}^{k+1} : \boldsymbol{a}^T \boldsymbol{w} = z\}$ from the origin. Furthermore, we have the above stated relations following from (10), between the distribution parameters $\boldsymbol{\delta}$ and Ω and the angles between the hyperplanes that are boundaries of the cone.

Classes $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$ corresponding to different parameters $(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$ are disjoint. For $SE_k(\mathbf{0}, \Omega, \boldsymbol{\delta}; g^{(k+1)})$ and $\boldsymbol{z}$ being fixed, there are uncountably many cones generated by orthogonal cone transformations which satisfy the representation (11).

Lemma 1. *If $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ and $C(\boldsymbol{a}_0^*, \boldsymbol{a}_1^*, ..., \boldsymbol{a}_k^*; \boldsymbol{z})$ are $(k+1)$-dimensional cones which are elements of the same class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$ then there is an orthogonal transformation O such that*

$$C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}) = OC(\boldsymbol{a}_0^*, \boldsymbol{a}_1^*, ..., \boldsymbol{a}_k^*; \boldsymbol{z}).$$

Proof. We define $\tilde{\boldsymbol{a}}_0 := \boldsymbol{a}_0/||\boldsymbol{a}_0||$ and $\tilde{\boldsymbol{a}}_0^* := \boldsymbol{a}_0^*/||\boldsymbol{a}_0^*||$ and note that the sets $\{\tilde{\boldsymbol{a}}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k\}$ and $\{\tilde{\boldsymbol{a}}_0^*, \boldsymbol{a}_1^*, ..., \boldsymbol{a}_k^*\}$ are both bases of $\mathbb{R}^{k+1}$. Therefore, there exists a unique linear map $f : \mathbb{R}^{k+1} \to \mathbb{R}^{k+1}$ such that $\tilde{\boldsymbol{a}}_0 = f(\tilde{\boldsymbol{a}}_0^*)$ and $\boldsymbol{a}_i = f(\boldsymbol{a}_i^*),\ i = 1, ..., k$. By using the assumptions of Lemma 1, it can be shown that the map f is orthogonal, i.e. for any pair of vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{k+1}$ holds $< f(\boldsymbol{v}), f(\boldsymbol{w}) >=< \boldsymbol{v}, \boldsymbol{w} >$ where $< \cdot, \cdot >$ denotes the standard scalar product in $\mathbb{R}^{k+1}$. Hence, there is an orthogonal matrix O such that $\tilde{\boldsymbol{a}}_0 = O\tilde{\boldsymbol{a}}_0^*$ and $\boldsymbol{a}_i = O\boldsymbol{a}_i^*,\ i = 1, ..., k$. It follows that $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}) = OC(\boldsymbol{a}_0^*, \boldsymbol{a}_1^*, ..., \boldsymbol{a}_k^*; \boldsymbol{z})$. □

It is one of the aims of the present note to make as clear as possible the relations between the well established techniques from the area of skewed distributions and the new techniques from the geometric approach. In this sense, we present in the following two alternative proofs of Theorem 1 in order to show different aspects of these relations. The first proof uses essentially the orthogonal invariance property in Lemma 1. The second one shows that one can also start from the stochastic representation of skewed elliptical random vectors that was basically used in [10].

Proof 1 of Theorem 1. It was shown in [7] that if Ω is a correlation matrix then (9) is equivalent to

$$\boldsymbol{\delta} = \frac{\Omega\boldsymbol{\lambda}}{(1 + \boldsymbol{\lambda}^T\Omega\boldsymbol{\lambda})^{1/2}}. \tag{13}$$

Let us consider a random vector

$$\begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix} \sim EC_{k+1}(\boldsymbol{0}_{k+1}, \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & \Omega \end{pmatrix}; g^{(k+1)}) \tag{14}$$

where X_0 and $\boldsymbol{Y}$ take values in $\mathbb{R}$ and $\mathbb{R}^k$, respectively. It follows from (8), (14), and $\begin{vmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & \Omega \end{vmatrix} = |\Omega|$ that

$$\begin{aligned} P(\boldsymbol{Z} < \boldsymbol{z}) &= 2|\Omega|^{-1/2} \int_{\boldsymbol{y}<\boldsymbol{z}} \int_{-\infty}^{\boldsymbol{\lambda}^T\boldsymbol{y}} g^{(k+1)}(s^2 + \boldsymbol{y}^T\Omega^{-1}\boldsymbol{y}) \; ds \; d\boldsymbol{y} \\ &= 2|\Omega|^{-1/2} \int_{\boldsymbol{y}<\boldsymbol{z}} \int_{-\infty}^{\boldsymbol{\lambda}^T\boldsymbol{y}} g^{(k+1)}((s, \boldsymbol{y}^T) \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & \Omega \end{pmatrix}^{-1} \begin{pmatrix} s \\ \boldsymbol{y} \end{pmatrix}) \; ds \; d\boldsymbol{y} \\ &= 2P(\boldsymbol{Y} < \boldsymbol{z}, X_0 < \boldsymbol{\lambda}^T\boldsymbol{Y}). \end{aligned}$$

Because Ω is regular, there is a regular $k \times k$-matrix C such that

$$\Omega = CC^T. \tag{15}$$

It follows from the properties of elliptically contoured distributions that if $\boldsymbol{Y}^* := C^{-1}\boldsymbol{Y}$ then

$$\begin{pmatrix} X_0 \\ \boldsymbol{Y}^* \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & C^{-1} \end{pmatrix} \begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix} \sim \Phi_{k+1}(\cdot; g^{(k+1)}).$$

Let

$$A_0^*(\boldsymbol{z}) := \{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \; : \; x_0 < \boldsymbol{\lambda}^T\boldsymbol{y}, \; \boldsymbol{y} < \boldsymbol{z}\}, \quad \boldsymbol{z} \in \mathbb{R}^k,$$

then

$$\begin{aligned} 2P(\boldsymbol{Y} < \boldsymbol{z}, \; X_0 < \boldsymbol{\lambda}^T\boldsymbol{Y}) &= 2P((X_0, \boldsymbol{Y}^T)^T \in A_0^*(\boldsymbol{z})) \\ &= 2P\left((X_0, \boldsymbol{Y}^{*T})^T \in A_0(\boldsymbol{z})\right) \end{aligned}$$

where

$$A_0(\boldsymbol{z}) = \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & C^{-1} \end{pmatrix} A_0^*(\boldsymbol{z}) = \left\{ \begin{pmatrix} x_0 \\ \boldsymbol{y}^* \end{pmatrix} \in \mathbb{R}^{k+1} \; : \; x_0 < \boldsymbol{\lambda}^T C\boldsymbol{y}^*, \; C\boldsymbol{y}^* < \boldsymbol{z} \right\}.$$

Hence,

$$P(\boldsymbol{Z} < \boldsymbol{z}) = 2\Phi_{k+1}(A_0(\boldsymbol{z}); g^{(k+1)}). \tag{16}$$

The set $A_0(\boldsymbol{z})$ allows the representation

$$A_0(\boldsymbol{z}) = \left\{\boldsymbol{y} \in \mathbb{R}^{k+1} \ : \ \boldsymbol{a}_0^{*T}\boldsymbol{y} < 0, \ \boldsymbol{a}_1^{*T}\boldsymbol{y} < z_1, \ ..., \ \boldsymbol{a}_k^{*T}\boldsymbol{y} < z_k\right\}$$

where $\boldsymbol{a}_0^{*T} = (1, \ -\boldsymbol{\lambda}^T C), \ \boldsymbol{a}_i^{*T} = (0, \ \boldsymbol{e}_i^T C), \ i = 1, ..., k$, and $\boldsymbol{e}_i$ denotes the ith usual unit vector of $\mathbb{R}^k$. The vectors $\boldsymbol{a}_0^*, \boldsymbol{a}_1^*, ..., \boldsymbol{a}_k^*$ are linearly independent because C is a full-rank matrix. Furthermore, we have the following equations which together with (13), prove that the vectors $\boldsymbol{a}_i^*, \ i = 0, 1, ..., k$, satisfy (10):

$$\frac{-\boldsymbol{a}_0^{*T}\boldsymbol{a}_i^*}{||\boldsymbol{a}_0^*||} = \frac{\boldsymbol{\lambda}^T C C^T \boldsymbol{e}_i}{(1 + \boldsymbol{\lambda}^T C C^T \boldsymbol{\lambda})^{1/2}} = \frac{\boldsymbol{\lambda}^T \Omega \boldsymbol{e}_i}{(1 + \boldsymbol{\lambda}^T \Omega \boldsymbol{\lambda})^{1/2}} = \delta_i, \quad i = 1, ..., k,$$
$$\boldsymbol{a}_i^{*T}\boldsymbol{a}_j^* = \boldsymbol{e}_i^T C C^T \boldsymbol{e}_j = \boldsymbol{e}_i^T \Omega \boldsymbol{e}_j = \omega_{i,j}, \quad i, j = 1, ..., k,$$

thus

$$||\boldsymbol{a}_i^*||^2 = \omega_{i,i} = 1, \qquad i = 1, ..., k.$$

Hence, the cone $A_0(\boldsymbol{z})$ belongs to the class $\mathfrak{C}_{k+1}(\boldsymbol{\delta}, \Omega, \boldsymbol{z})$. Because of Lemma 1, every cone $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ from the same class can be mapped orthogonally onto $A_0(\boldsymbol{z})$. Because of the orthogonal invariance of the spherical distribution $\Phi_{k+1}(\cdot; g^{(k+1)})$,

$$\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}) = \Phi_{k+1}(A_0(\boldsymbol{z}); g^{(k+1)}). \tag{17}$$

The claim of Theorem 1 now follows on combining (16) and (17). □

Before we present the announced alternative proof of Theorem 1, we recall the stochastic representation of skewed elliptically contoured random vectors that was basically used in [10]. Here, $\mathfrak{L}(.|.)$ denotes the conditional probability law.

Proposition 1. *If* $\begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix} \sim El_{k+1}(\begin{pmatrix} 0 \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix}; g^{(k+1)})$ *then the skewed elliptically contoured distribution allows the conditional distribution representation*

$$SE_k(\boldsymbol{\xi}, \Omega, \boldsymbol{\delta}; g^{(k+1)}) = \mathfrak{L}(\boldsymbol{Y}|X_0 > 0).$$

In [10], the statement of Proposition 1 was actually the definition for the skewed elliptical distribution, and the density representation (8) was derived from it. The following second proof of Theorem 1 makes use of Proposition 1.

Proof 2 of Theorem 1. Let $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ be a cone from the class $\boldsymbol{C}_{k+1}(\boldsymbol{\delta}, \Omega, \boldsymbol{z})$. We define the $(k+1) \times k$-matrix A by $A = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_k \end{pmatrix}$ and assume that $\begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} \sim \Phi_{k+1}(\cdot; g^{(k+1)})$ where U_0 and $\boldsymbol{U}$ take values in $\mathbb{R}$ and $\mathbb{R}^k$, respectively. Note that

$$\begin{aligned}
\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}) &= P(A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < \boldsymbol{z}, \boldsymbol{a}_0^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < 0) \\
&= \frac{1}{2} P(A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < \boldsymbol{z} \,|\, \boldsymbol{a}_0^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < 0) \\
&= \frac{1}{2} P(A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < \boldsymbol{z} \,|\, -\frac{\boldsymbol{a}_0^T}{||\boldsymbol{a}_0||} \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} > 0).
\end{aligned}$$

Hence, with the notation $\tilde{\boldsymbol{a}}_0 = -\frac{\boldsymbol{a}_0}{||\boldsymbol{a}_0||}$,

$$2\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}) = P(A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} < \boldsymbol{z} \,|\, \tilde{\boldsymbol{a}}_0^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} > 0). \quad (18)$$

It follows from the equations (10) that $\boldsymbol{\delta} = A^T \tilde{\boldsymbol{a}}_0$ and $\Omega = A^T A$. The properties of elliptically contoured distributions ensure that

$$\begin{pmatrix} \tilde{\boldsymbol{a}}_0^T \\ A^T \end{pmatrix} \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} \sim El_{k+1}(\boldsymbol{0}_{k+1}, \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix}; g^{(k+1)}).$$

Finally, with $X_0 = \tilde{\boldsymbol{a}}_0^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix}$ and $\boldsymbol{Y} = A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix}$, Proposition 1 yields

$$\mathfrak{L}\left(A^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} \,|\, \tilde{\boldsymbol{a}}_0^T \begin{pmatrix} U_0 \\ \boldsymbol{U} \end{pmatrix} > 0\right) = SE_k(\boldsymbol{0}_k, \Omega, \boldsymbol{\delta}, g^{(k+1)}). \quad (19)$$

The claim of Theorem 1 follows on combining (18) and (19). □

3 Stochastic representations

It is known from the second proof of Theorem 1 that representation formula (11) can be derived from the conditional distribution representation in Proposition 1. In this section, we demonstrate how, vice versa, Theorem 1 can be used to derive this and other representations. This way, we reprove in a unified geometric way several representations of the skewed elliptically contoured distribution, including that of Proposition 1.

3.1 Representation based upon selection mechanism

In this subsection, we re-prove Proposition 1 by using Theorem 1. Thereby, we restrict us to the special case that $\boldsymbol{\xi} = \mathbf{0}_k$ and Ω is a correlation matrix. The case of arbitrary $\boldsymbol{\xi}$ can be treated similarly by changing $\boldsymbol{Y}$ with $\boldsymbol{Y} - \boldsymbol{\xi}$. The matrix $\begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix}$ is assumed to be s.p.d., hence we can write $\begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix} = BB^T$ with a certain non-singular $(k+1) \times (k+1)$-matrix B. We define $\begin{pmatrix} X_0^* \\ \boldsymbol{Y}^* \end{pmatrix} := B^{-1} \begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix}$, where $\begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix}$ satisfies the assumptions of Proposition 1. Because of the properties of elliptical distributions,

$$\begin{pmatrix} X_0^* \\ \boldsymbol{Y}^* \end{pmatrix} \sim \Phi_{k+1}(\cdot; g^{(k+1)}).$$

Let

$$A_1^*(\boldsymbol{z}) := \{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ x_0 > 0, \ \boldsymbol{y} < \boldsymbol{z}\}, \quad \boldsymbol{z} \in \mathbb{R}^k.$$

We have

$$\begin{aligned} P(\boldsymbol{Y} < \boldsymbol{z} | X_0 > 0) &= 2P(\boldsymbol{Y} < \boldsymbol{z}, X_0 > 0) = 2P((X_0, \boldsymbol{Y}^T)^T \in A_1^*(\boldsymbol{z})) \\ &= 2P\left((X_0^*, \boldsymbol{Y}^{*T})^T \in A_1(\boldsymbol{z})\right), \end{aligned}$$

where

$$\begin{aligned} A_1(\boldsymbol{z}) = B^{-1} A_1^*(\boldsymbol{z}) &= \left\{ B^{-1} \begin{pmatrix} x_0 \\ \boldsymbol{y} \end{pmatrix} \in \mathbb{R}^{k+1} \ : \ -x_0 < 0, \ \boldsymbol{y} < \boldsymbol{z} \right\} \\ &= \left\{ \begin{pmatrix} x_0^* \\ \boldsymbol{y}^* \end{pmatrix} \in \mathbb{R}^{k+1} \ : \ \begin{pmatrix} -1 & \mathbf{0}_k^T \\ \mathbf{0}_k & I_k \end{pmatrix} B \begin{pmatrix} x_0^* \\ \boldsymbol{y}^* \end{pmatrix} < \begin{pmatrix} 0 \\ \boldsymbol{z} \end{pmatrix} \right\}. \end{aligned}$$

We get the following intermediate result:

$$P(\boldsymbol{Y} < \boldsymbol{z} | X_0 > 0) = 2\Phi_{k+1}(A_1(\boldsymbol{z}); g^{(k+1)}). \tag{20}$$

The set $A_1(z)$ allows the representation

$$A_1(\boldsymbol{z}) = \left\{ \boldsymbol{y} \in \mathbb{R}^{k+1} \ : \ \boldsymbol{a}_0^T \boldsymbol{y} < 0, \ \boldsymbol{a}_1^T \boldsymbol{y} < z_1, \ ..., \ \boldsymbol{a}_k^T \boldsymbol{y} < z_k \right\}$$

where $\boldsymbol{a}_0^T = -\boldsymbol{e}_1^T B, \ \boldsymbol{a}_i^T = \boldsymbol{e}_{i+1}^T B, \ i = 1, ..., k$, and $\boldsymbol{e}_i$ denotes now the ith unit vector of $\mathbb{R}^{k+1}$. The vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k$ are linearly independent because the matrix B is of full rank. Furthermore, the following equations show that the parameter vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k$ satisfy (10):

$$\frac{-\boldsymbol{a}_0^T\boldsymbol{a}_i}{||\boldsymbol{a}_0||} = \frac{\boldsymbol{e}_1^T BB^T \boldsymbol{e}_{i+1}}{(\boldsymbol{e}_1^T BB^T \boldsymbol{e}_1)^{1/2}} = \frac{\boldsymbol{e}_1^T \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix} \boldsymbol{e}_{i+1}}{\left(\boldsymbol{e}_1^T \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix} \boldsymbol{e}_1\right)^{1/2}} = \frac{\delta_i}{1}, \quad i = 1, ..., k,$$

$$\boldsymbol{a}_i^T\boldsymbol{a}_j = \boldsymbol{e}_{i+1}^T BB^T \boldsymbol{e}_{j+1} = \boldsymbol{e}_{i+1}^T \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix} \boldsymbol{e}_{j+1} = \omega_{i,j}, \quad i, j = 1, ..., k,$$

thus

$$||\boldsymbol{a}_i||^2 = \omega_{i,i} = 1, \quad i = 1, ..., k.$$

Therefore, $A_1(\boldsymbol{z})$ is a cone from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$. Proposition 1 now follows from (20) and Theorem 1.

3.2 Representation based upon linear combination

The following stochastic representation of a skewed elliptically contoured random vector was originally derived in another way in [8] and [12], and in slightly different notation. In this subsection, we re-prove this result using Theorem 1.

Proposition 2. *Let* $\begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix} \sim EC_{k+1}(\boldsymbol{0}_{k+1}, \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & \Psi \end{pmatrix}; g^{(k+1)})$*, where* Ψ *is a s.p.d.* $k \times k$ *correlation matrix. Further, let* $Z_j = \delta_j|X_0| + (1 - \delta_j^2)^{1/2} Y_j,\ j = 1, ..., k,$ *where* $\boldsymbol{\delta} = (\delta_1, ..., \delta_k)^T \in (-1, 1)^k$*, and* $\Delta = diag(\delta_1, ..., \delta_k)$*, and* $\boldsymbol{Z} := (Z_1, ..., Z_k)^T = |X_0|\boldsymbol{\delta} + (I_k - \Delta^2)^{1/2}\boldsymbol{Y}$*. Then*

$$\boldsymbol{Z} \sim SE_k(\boldsymbol{0}_k, \Omega, \boldsymbol{\delta}; g^{(k+1)}),$$

where

$$\Omega = \boldsymbol{\delta}\boldsymbol{\delta}^T + (I_k - \Delta^2)^{1/2}\Psi(I_k - \Delta^2)^{1/2}. \tag{21}$$

To prove Proposition 2, we use that the matrix Ψ is s.p.d., so that $\Psi = CC^T$ with certain regular $k \times k$ matrix C. Let $\boldsymbol{Y}^* := C^{-1}\boldsymbol{Y}$. It follows from the properties of elliptically contoured distributions that

$$\begin{pmatrix} X_0 \\ \boldsymbol{Y}^* \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & C^{-1} \end{pmatrix} \begin{pmatrix} X_0 \\ \boldsymbol{Y} \end{pmatrix} \sim \Phi_{k+1}(\ \cdot\ ; g^{(k+1)}).$$

With

$$A_2^*(\boldsymbol{z}) := \{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ \boldsymbol{\delta}|x_0| + (I_k - \Delta^2)^{1/2}\boldsymbol{y} < \boldsymbol{z}\}, \quad \boldsymbol{z} \in \mathbb{R}^k,$$

we observe that

$$\begin{aligned} P(\boldsymbol{\delta}|X_0| + (I_k - \Delta^2)^{1/2}\boldsymbol{Y} < \boldsymbol{z}) &= P((X_0, \boldsymbol{Y}^T)^T \in A_2^*(\boldsymbol{z})) \\ &= P\left((X_0, \boldsymbol{Y}^{*T})^T \in \tilde{A}_2(\boldsymbol{z})\right), \end{aligned}$$

where

$$\begin{aligned} \tilde{A}_2(\boldsymbol{z}) &= \begin{pmatrix} 1 & \boldsymbol{0}_k^T \\ \boldsymbol{0}_k & C^{-1} \end{pmatrix} A_2^*(\boldsymbol{z}) \\ &= \left\{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ \boldsymbol{\delta}|x_0| + (I_k - \Delta^2)^{1/2}C\boldsymbol{y} < \boldsymbol{z}\right\} \\ &= \left\{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ x_0 \geq 0, \ \boldsymbol{\delta}x_0 + (I_k - \Delta^2)^{1/2}C\boldsymbol{y} < \boldsymbol{z}\right\} \\ &\quad \cup \{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ x_0 < 0, \ -\boldsymbol{\delta}x_0 + (I_k - \Delta^2)^{1/2}C\boldsymbol{y} < \boldsymbol{z}\}. \end{aligned} \tag{22}$$

If

$$A_2(\boldsymbol{z}) = \left\{(x_0, \boldsymbol{y}^T)^T \in \mathbb{R}^{k+1} \ : \ x_0 > 0, \ \boldsymbol{\delta}x_0 + (I_k - \Delta^2)^{1/2}C\boldsymbol{y} < \boldsymbol{z}\right\},$$

we have $\tilde{A}_2(\boldsymbol{z}) = \overline{A_2(\boldsymbol{z})} \cup \left[\left(\begin{smallmatrix} -1 & 0 \\ 0 & I_k \end{smallmatrix}\right) A_2(\boldsymbol{z})\right]$ and $A_2(\boldsymbol{z}) \cap \left[\left(\begin{smallmatrix} -1 & 0 \\ 0 & I_k \end{smallmatrix}\right) A_2(\boldsymbol{z})\right] = \emptyset$. Therefore,

$$\begin{aligned} P\left((X_0, \boldsymbol{Y}^{*T})^T \in \tilde{A}_2(\boldsymbol{z})\right) &= \Phi_{k+1}(\tilde{A}_2(\boldsymbol{z})) \\ &= \Phi_{k+1}(A_2(\boldsymbol{z})) + \Phi_{k+1}(\left(\begin{smallmatrix} -1 & 0 \\ 0 & I_k \end{smallmatrix}\right) A_2(\boldsymbol{z})) \\ &= 2\Phi_{k+1}(A_2(\boldsymbol{z})). \end{aligned}$$

Hence,

$$P(\boldsymbol{\delta}|X_0| + (I_k - \Delta^2)^{1/2}\boldsymbol{Y} < \boldsymbol{z}) = 2\Phi_{k+1}(A_2(\boldsymbol{z}); g^{(k+1)}). \tag{23}$$

We can represent $A_2(\boldsymbol{z})$ as

$$A_2(\boldsymbol{z}) = \left\{\boldsymbol{y} \in \mathbb{R}^{k+1} \ : \ \boldsymbol{a}_0^T\boldsymbol{y} < 0, \ \boldsymbol{a}_1^T\boldsymbol{y} < z_1, \ ..., \ \boldsymbol{a}_k^T\boldsymbol{y} < z_k\right\}$$

where $\boldsymbol{a}_0^T = (-1, \boldsymbol{0}_k^T), \ \boldsymbol{a}_i^T = (\delta_i, (1 - \delta_i^2)^{1/2}\boldsymbol{e}_i^T C), \ i = 1, ..., k$, and $\boldsymbol{e}_i$ denotes here the ith unit vector of $\mathbb{R}^k$. The vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k$ are linearly independent because the matrix C is a full-rank matrix. Furthermore, we have the following equations which together with (21) show that the parameter vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k$ satisfy (10):

$$\frac{-\boldsymbol{a}_0^T \boldsymbol{a}_i}{||\boldsymbol{a}_0||} = \frac{\delta_i}{1} = \delta_i, \quad i = 1, ..., k,$$

$$\begin{aligned}\boldsymbol{a}_i^T \boldsymbol{a}_j &= \delta_i \delta_j + (1-\delta_i^2)^{1/2} \boldsymbol{e}_i^T C C^T \boldsymbol{e}_j (1-\delta_j^2)^{1/2} \\ &= \delta_i \delta_j + (1-\delta_i^2)^{1/2} \boldsymbol{e}_i^T \Psi \boldsymbol{e}_j (1-\delta_j^2)^{1/2} \\ &= \delta_i \delta_j + (1-\delta_i^2)^{1/2} \psi_{i,j} (1-\delta_j^2)^{1/2}, \qquad i, j = 1, ..., k,\end{aligned}$$

and

$$||\boldsymbol{a}_i||^2 = \delta_i^2 + (1-\delta_i^2)\psi_{i,i} = \delta_i^2 + (1-\delta_i^2) = 1, \quad i = 1, ..., k.$$

Hence, $A_2(\boldsymbol{z})$ is a cone from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$, where Ω is given by (21). Proposition 2 now follows from (23) and Theorem 1.

3.3 Discussion

It was shown so far in this section that some of the known representations of skewed elliptically contoured distributions can be derived in a unified geometric way from Theorem 1. We want to add a few more words on the opposite direction, that is how Theorem 1 can be derived from any of these known representations, too. The first proof of Theorem 1 essentially makes use of representation (16) for the cdf of the skewed elliptically contoured distribution. As a matter of fact, one can also use Proposition 1 together with (20) in order to get

$$P(\boldsymbol{Z} < \boldsymbol{z}) = 2\Phi_{k+1}(A_1(\boldsymbol{z}); g^{(k+1)}),$$

and can use this last equation instead of (16) in the first proof of Theorem 1. Then, for proving the claim of Theorem 1 in the same way as in the first proof of Theorem 1, one can use of the fact that $A_1(\boldsymbol{z})$ is a special cone from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$. After that, Lemma 1 and the orthogonal invariance property of $\Phi_{k+1}(\cdot; g^{(k+1)})$ apply.

Similarly, one can prove Theorem 1 starting from Proposition 2. This proposition together with (23) applies to show $P(\boldsymbol{Z} < \boldsymbol{z}) = 2\Phi_{k+1}(A_2(\boldsymbol{z}); g^{(k+1)})$. One can use now this equation instead of (16) in the first proof of Theorem 1, and can perform then the same reasoning as above by exploiting the fact that $A_2(z)$ is a special cone from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$.

Actually, one can use any (possibly yet even unknown) representation of the skewed elliptical distribution that implies

$$P(\boldsymbol{Z} < \boldsymbol{z}) = 2\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)})$$

for just one special cone $C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z})$ from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$. Then, the extension of the claim of Theorem 1 to any cone from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$ follows from Lemma 1.

4 Applications and examples

4.1 Describing independence

General conditions ensuring that sub-vectors of multivariate skewed normal vectors are mutually independent, follow from Proposition 6 in [7]. Here, we show that formula (11) applies to derive geometrically stated conditions under which such independence relations hold.

Remark 1. *Let the cdf of the k-dimensional random vector $\boldsymbol{Z}$ satisfy representation* (11) *with $g^{(k+1)}$ being the density generator of the normal distribution, and let $\boldsymbol{Z}$ be partitioned as $\boldsymbol{Z} = (\boldsymbol{Y}_1^T, ..., \boldsymbol{Y}_h^T)^T$ where the sub-vectors have dimensions $m_1, ..., m_h$, respectively, $m_1 + ... + m_h = k$. If the linear spaces $\mathfrak{L}(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_{m_1})$, $\mathfrak{L}(\boldsymbol{a}_{m_1+1}, \boldsymbol{a}_{m_1+2}, ..., \boldsymbol{a}_{m_1+m_2})$, $\mathfrak{L}(\boldsymbol{a}_{m_1+m_2+1}, ..., \boldsymbol{a}_{m_1+m_2+m_3})$, ..., $\mathfrak{L}(\boldsymbol{a}_{m_1+...+m_{h-1}+1}, \boldsymbol{a}_{m_1+...+m_{h-1}+2}, ..., \boldsymbol{a}_{m_1+...+m_h})$ spanned up by the vectors in parentheses are orthogonal to each other then $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_h$ are independent. Furthermore, in this case, $\boldsymbol{Y}_1$ will be skewed normally distributed, whereas $\boldsymbol{Y}_2, ..., \boldsymbol{Y}_h$ are normal random vectors.*

Proof. For simplicity, and without loss of generality, we consider the case $h = 3$. We suppose that $\boldsymbol{z} \in \mathbb{R}^{k+1}$ is partitioned as $\boldsymbol{z} = (\boldsymbol{y}_1^T, \boldsymbol{y}_2^T, \boldsymbol{y}_3^T)^T$ where $\boldsymbol{y}_1,\ \boldsymbol{y}_2,\ \boldsymbol{y}_3$ have dimensions m_1, m_2, m_3, respectively. We define $A_1 = \begin{pmatrix}\boldsymbol{a}_0 & \boldsymbol{a}_1 & ... & \boldsymbol{a}_{m_1}\end{pmatrix}$, $A_2 = \begin{pmatrix}\boldsymbol{a}_{m_1+1} & ... & \boldsymbol{a}_{m_1+m_2}\end{pmatrix}$, and $A_3 = \begin{pmatrix}\boldsymbol{a}_{m_1+m_2+1} & ... & \boldsymbol{a}_k\end{pmatrix}$, and moreover $\boldsymbol{X} \sim N_{k+1}(\boldsymbol{0}_{k+1}, I_{k+1})$. Then

$$\begin{aligned}
P(\boldsymbol{Z} < \boldsymbol{z}) &= 2\Phi_{k+1}(C(\boldsymbol{a}_0, \boldsymbol{a}_1, ..., \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}) \\
&= 2P(\boldsymbol{a}_0^T\boldsymbol{X} < 0,\ \boldsymbol{a}_1^T\boldsymbol{X} < z_1,\ ..., \boldsymbol{a}_k^T\boldsymbol{X} < z_k) \\
&= 2P(A_1^T\boldsymbol{X} < (0, \boldsymbol{y}_1^T)^T,\ A_2^T\boldsymbol{X} < \boldsymbol{y}_2,\ A_3^T\boldsymbol{X} < \boldsymbol{y}_3) \\
&= 2P(A_1^T\boldsymbol{X} < (0, \boldsymbol{y}_1^T)^T) \cdot P(A_2^T\boldsymbol{X} < \boldsymbol{y}_2) \cdot P(A_3^T\boldsymbol{X} < \boldsymbol{y}_3).
\end{aligned}$$

In the last equation, the independence of $A_1^T \boldsymbol{X}$, $A_2^T \boldsymbol{X}$, and $A_3^T \boldsymbol{X}$ was used. This property follows by considering the distribution of $\begin{pmatrix} A_1 & A_2 & A_3 \end{pmatrix}^T \boldsymbol{X}$ and exploiting the orthogonality condition assumed in the Remark 1. Furthermore,

$$P(\boldsymbol{Y}_1 < \boldsymbol{y}_1) = \lim_{z_{m_1+1}\to\infty} \dots \lim_{z_k\to\infty} P(\boldsymbol{Z} < \boldsymbol{z}) = 2P(A_1^T \boldsymbol{X} < (0, \boldsymbol{y_1}^T)^T),$$

analogously $P(\boldsymbol{Y}_2 < \boldsymbol{y}_2) = P(A_2^T \boldsymbol{X} < \boldsymbol{y}_2)$ and $P(\boldsymbol{Y}_3 < \boldsymbol{y}_3) = P(A_3^T \boldsymbol{X} < \boldsymbol{y}_3)$. Indeed, $\boldsymbol{Y}_2$ and $\boldsymbol{Y}_3$ are normally distributed. To prove that $\boldsymbol{Y}_1$ is skewed normal, one can find an orthogonal $(k+1)\times(k+1)$-matrix such that $OA_1 = \begin{pmatrix} \tilde{\boldsymbol{a}}_0 & \tilde{\boldsymbol{a}}_1 & \dots & \tilde{\boldsymbol{a}}_{m_1} \\ \boldsymbol{0}_{k-(m_1+1)} & \boldsymbol{0}_{k-(m_1+1)} & \dots & \boldsymbol{0}_{k-(m_1+1)} \end{pmatrix}$ where $\tilde{\boldsymbol{a}}_i \in \mathbb{R}^{m_1+1}$, $i = 0, 1, \dots, m_1$. Hence, because of $O^T \boldsymbol{X} \stackrel{d}{=} \boldsymbol{X}$,

$$\begin{aligned} P(\boldsymbol{Y}_1 < \boldsymbol{y}_1) &= 2P(A_1^T \boldsymbol{X} < (0, \boldsymbol{y_1}^T)^T) = 2P(A_1^T O^T \boldsymbol{X} < (0, \boldsymbol{y_1}^T)^T) \\ &= 2P((OA_1)^T \boldsymbol{X} < (0, \boldsymbol{y_1}^T)^T) \\ &= 2\Phi_{m_1+1}(C(\tilde{\boldsymbol{a}}_0, \tilde{\boldsymbol{a}}_1, \dots, \tilde{\boldsymbol{a}}_{m_1}; \boldsymbol{y}_1); g^{(k+1)}) \end{aligned}$$

and thus, $\boldsymbol{Y}_1$ is a m_1-dimensional skewed normal random vector, becauce of Theorem 1. □

Notice that the conditions for independence, which follow from Proposition 6 in [7], are met if the orthogonality condition from Remark 1 is satisfied. This follows by considering the equations (10).

4.2 Deriving representations through symmetrization

The present subsection illustrates that one has not necessarily to restrict considerations of the geometric measure representations for skewed distributions to cones which include the origin in at least one bounding hyperplane. To be specific, we derive here a representation of $P(\boldsymbol{Z} < \boldsymbol{z})$ in terms of Φ_{k+1}-values of sets derived from sets of the type $C(\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_k; \boldsymbol{z})$ through symmetrization.

Remark 2. *If $\boldsymbol{Z} \sim SE_k(\boldsymbol{0}, \Omega, \boldsymbol{\delta}, g^{(k+1)})$ then, for every cone $C(\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_k; \boldsymbol{z})$ from the class $\mathfrak{C}_{k+1}(\Omega, \boldsymbol{\delta}, \boldsymbol{z})$, the cdf of $\boldsymbol{Z}$ allows the representation*

$$\begin{aligned} P(\boldsymbol{Z} < \boldsymbol{z}) = \Phi_{k+1}(&C(\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_k; \boldsymbol{z}) \\ &\cup (I_{k+1} - \frac{2}{\boldsymbol{a}_0^T \boldsymbol{a}_0} \boldsymbol{a}_0 \boldsymbol{a}_0^T) C(\boldsymbol{a}_0, \boldsymbol{a}_1, \dots, \boldsymbol{a}_k; \boldsymbol{z}); g^{(k+1)}). \quad (24) \end{aligned}$$

Proof. The Householder matrix $(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)$ mirrors $C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})$ on the bounding hyperplane $\{\boldsymbol{y}\in\mathbb{R}^{k+1}:\boldsymbol{a}_0^T\boldsymbol{y}=0\}$ which contains the origin. Therefore, $C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})$ and $(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})$ are disjoint. Furthermore, $(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)$ is orthogonal. Hence,

$$\Phi_{k+1}((I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z}))=\Phi_{k+1}(C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z}))$$

and thus

$$\begin{aligned}&\Phi_{k+1}(C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})\cup(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z});g^{(k+1)})\\&=2\Phi_{k+1}(C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z}))=P(\boldsymbol{Z}<\boldsymbol{z}),\end{aligned}$$

where the last equation follows from Theorem 1. □

Example 1. *Let us recall that the set $\tilde{B}(t)$ from Section 1 allows the representation $\tilde{B}(t)=C(\boldsymbol{a}_0,\boldsymbol{a}_1;z)$ with $\boldsymbol{a}_0=(0,-1)^T$, $\boldsymbol{a}_1=((\frac{1+\rho}{2})^{1/2},(\frac{1-\rho}{2})^{1/2})^T$, and $\boldsymbol{z}=t$. The corresponding Householder matrix is therefore $(I_2-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)=\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$. Hence, the set $DOB(t)$ from Section 1 allows the representation $DOB(t)=C(\boldsymbol{a}_0,\boldsymbol{a}_1;\boldsymbol{z})\cup(I_2-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)C(\boldsymbol{a}_0,\boldsymbol{a}_1;\boldsymbol{z})$.*

Example 2. *The set $A_2(\boldsymbol{z})$ in the proof of Proposition 2 allows the representation $A_2(\boldsymbol{z})=C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})$ with $\boldsymbol{a}_0=(-1,\boldsymbol{0}_k^T)^T$. The corresponding Householder matrix is thus $(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)=\begin{pmatrix}-1&\boldsymbol{0}^T\\\boldsymbol{0}&I_k\end{pmatrix}$. After symmetrization of the cone $C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})$, we get the set*

$$\begin{aligned}&C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})\cup(I_{k+1}-\frac{2}{\boldsymbol{a}_0^T\boldsymbol{a}_0}\boldsymbol{a}_0\boldsymbol{a}_0^T)C(\boldsymbol{a}_0,\boldsymbol{a}_1,...,\boldsymbol{a}_k;\boldsymbol{z})\\&=A_2(\boldsymbol{z})\cup\left(\begin{smallmatrix}-1&0\\0&I_k\end{smallmatrix}\right)A_2(\boldsymbol{z})\\&=\left\{(x_0,\boldsymbol{y}^T)^T\in\mathbb{R}^{k+1}\ :\ x_0\geq 0,\ \boldsymbol{\delta}x_0+(I_k-\Delta^2)^{1/2}C\boldsymbol{y}<\boldsymbol{z}\right\}\\&\quad\cup\{(x_0,\boldsymbol{y}^T)^T\in\mathbb{R}^{k+1}\ :\ x_0<0,\ -\boldsymbol{\delta}x_0+(I_k-\Delta^2)^{1/2}C\boldsymbol{y}<\boldsymbol{z}\}\\&=\left\{(x_0,\boldsymbol{y}^T)^T\in\mathbb{R}^{k+1}\ :\ \boldsymbol{\delta}|x_0|+(I_k-\Delta^2)^{1/2}C\boldsymbol{y}<\boldsymbol{z}\right\},\end{aligned}$$

which coincides with (22).

4.3 Contour plots

The consideration in this subsection is restricted to the case $k = 2$. We observe by systematically changing certain parameters how the shape of the density level sets of a two-dimensional skewed normal vector (Z_1, Z_2) depends on the linearly independent vectors $\boldsymbol{a}_0, \boldsymbol{a}_1$ and $\boldsymbol{a}_2$. These vectors are normal to the boundary-planes of that cone in $\mathbb{R}^3$ which is used to express the cdf of (Z_1, Z_2) according to (11). In comparison with (12), we slightly modify the notation for this cone and put

$$C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2) := \left\{ \boldsymbol{x} \in \mathbb{R}^3 : \boldsymbol{a}_0^T \boldsymbol{x} < 0, \ \frac{\boldsymbol{a}_1^T}{||\boldsymbol{a}_1||} \boldsymbol{x} < z_1, \frac{\boldsymbol{a}_2^T}{||\boldsymbol{a}_2||} \boldsymbol{x} < z_2 \right\}.$$

This set depends only on the directions of the vectors $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ but not on their norms what is essentially the same in all previous sections where we restrict vectors $\boldsymbol{a}_1, ..., \boldsymbol{a}_k$ to be normalized. Furthermore, in this subsection, we use $\Phi_3(\cdot)$ to denote the three-dimensional standard Gaussian measure.

Figure 1 shows density contour plots of two-dimensional skewed normal distributions having different parameter vectors a_0, a_1 and a_2. For more figures reflecting the effects of different changes in the vectors a_0, a_1 and a_2, we refer to Appendix B. Similar contour plots reflecting effects of varying a correlation coefficient or Madia's skewness measure are to be found in [14] and [18], respectively.

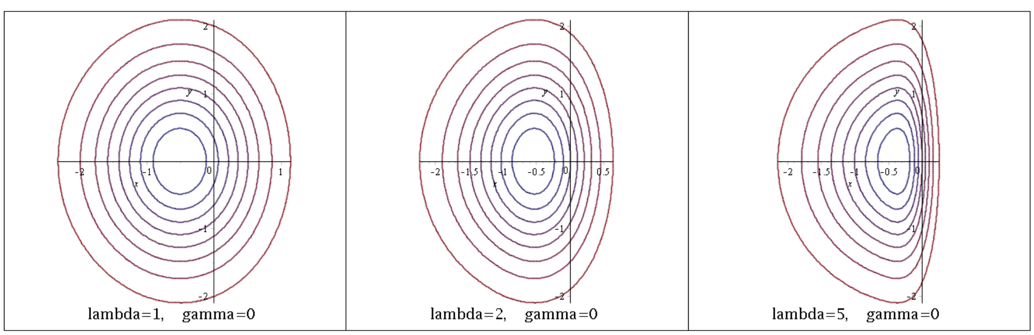

Figure 1: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$ with $\boldsymbol{a}_0 = (1, \lambda, \gamma)^T$, $\boldsymbol{a}_1 = (0, 1, 0)^T$, $\boldsymbol{a}_2 = (0, 0, 1)^T$. Here, $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$ and changes are only made in λ.

In the final remarks we describe some observations which can be made when considering these figures.

Remark 3. *If* $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$*, as it is the case in Figures 1 and 2, then the density contour plot of* (Z_1, Z_2) *is always symmetric w.r.t., and the density looks "skewed into the direction" of the vector* $(\delta_1, \delta_2) = -(\cos(\angle(\boldsymbol{a}_0, \boldsymbol{a}_1)), \cos(\angle(\boldsymbol{a}_0, \boldsymbol{a}_2)))$*. From Figure 2 , one may get the impression that "skewing to the left" increases as the angle between* $\boldsymbol{a}_0$ *and* $\boldsymbol{a}_1$ *becomes sharper, i.e. as* λ *increases. Moreover, one may argue that "skewing downwards" increases as the angle between* $\boldsymbol{a}_0$ *and* $\boldsymbol{a}_2$ *becomes sharper, i.e. as* γ *increases. We let it here an open problem to give these statements a precise mathematical sense in a future work.*

Remark 4. *If there are no restrictions upon the vectors* $a_i, i = 1, 2, 3$ *then it is not as easy to uniquely detect general rules on their skewing effects. Still, the angle between* $\boldsymbol{a}_0$ *and* $\boldsymbol{a}_1$ *seems to be related to skewing to the left or right, and the angle between* $\boldsymbol{a}_0$ *and* $\boldsymbol{a}_2$ *seems to be related to skewing downwards or upwards. Besides, we also obtain an effect of higher or lower concentration of contour lines which seems to be essentially influenced by the angle between* $\boldsymbol{a}_1$ *and* $\boldsymbol{a}_2$*. Having a closer look onto the contour plots, however, one gets a visual impression of how all the three effects superimpose. Figures 3 to 7 indicate the great variety of skewing two-dimensional normal densities. Moreover, if we would change both the signs of* λ *and* γ *in Figure 3 and Figure 4 , then the contour plots would mirror along the x-axis. If we would choose* $\boldsymbol{a}_2 = (0, 0, -1)^T$ *instead of* $\boldsymbol{a}_2 = (0, 0, 1)^T$*, then the contour plots would mirror along the y-axis. In Figure 6 , one can observe ëxtremeplots where in the left figure, the angle* $\angle(\boldsymbol{a}_0, \boldsymbol{a}_1)$ *is very sharp, in the central figure,* $\angle(\boldsymbol{a}_0, \boldsymbol{a}_2)$ *is very sharp, and in the right figure,* $\angle(\boldsymbol{a}_1, \boldsymbol{a}_2)$ *is very sharp. In Figure 7 , the value of all three angles decrease when turning from the left to the right.*

References

[1] Arellano-Valle, R. B., Azzalini, A. (2006). On a unification of families of skew-normal distributions. *Scand. J. Stat.* **33,** 561–574.

[2] Arellano-Valle, R. B., Branco, M.D., Genton, M. G. (2006). A unified view on skewed distributions arizing from selections. *Can. J. Stat.* **34,** 581–601.

[3] Arellano-Valle, R. B., Genton, M. G. (2008). On the exact distribution of the maximum of absolutely continuous dependent random variables. *Stat. Probab. Lett.*, **78,** 28–35.

[4] Arellano-Valle, R. B., Genton, M. G. (2010). Multivariate unified skew-elliptical distributions. *Chil. J. Stat.*, **1,** 17–33.

[5] Arellano-Valle, R.B., Richter, W.-D. (2012). On skewed continuous $l_{n,p}$-symmetric distributions. *Chil. J. Stat.* **3,** 193–212.

[6] Azzalini, A., Dalla Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83,** 715–726.

[7] Azzalini, A., Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *J. R. Statist. Soc.* **B, 61,** 579–602.

[8] Azzalini, A., Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Statist. Soc.* **B, 65,** 367–389.

[9] Batún-Cutz, J., González-Farías, G. , Richter, W.-D. (2012). Maximum distributions for $l_{2,p}$-symmetric vectors are skewed $l_{1,p}$-symmetric distributions. *Stat. Probab. Lett.* **83,** 10, 2260–2268. DOI 10.1016/j.spl.2013.06.022.

[10] Branco, M. D., Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. J. Multivar. Anal. **79,** 99–113.

[11] Fang, K.-T., Kotz, S., Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman & Hall, London.

[12] Fang, B. Q. (2003). The skew elliptical distributions and their quadratic forms. J. Multivar. Anal. **87,** 298–314.

[13] Günzel, T., Richter, W.-D., Scheutzow, S., Schicker, K., Venz, J. (2012). Geometric approach to the skewed normal distribution. *J. Stat. Plan. Infer.* **142,** 3209–3224.

[14] Gupta, A.K., González-Farías, G., Domínguez-Molina, J.A. (2004). A multivariate skew normal distribution. J. Multivar. Anal. **89,** 1, 181–190.

[15] Jamalizadeh, A., Balakrishnan, N. (2010). Distributions of order statistics and linear combinations of order statistics from an elliptical distribution as mixtures of unified skew-elliptical distributions. J. Multivar. Anal. **101,** 1412–1427.

[16] Loperfido, N., Navarro, J., Ruiz, J. M., Sandoval, C. J. (2007). Some Relationships Between Skew-Normal Distributions and Order Statistics from Exchangeable Normal Random Vectors. *Commun. Stat. A-Theor.* **36,** 1719–1733.

[17] Müller, K., Richter, W.-D. (2014). Exact extreme value, product and ratio distributions under non-standard assumptions. *AStA Advances in Statist. Anal.,* accepted for print.

[18] Sahu, S. K., Dey, D. K., Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *Can. J. Stat.* **31,** 129–150.

A

In the following, we prove that if Ω is a symmetric and positive definite $k \times k$ matrix and $\boldsymbol{\delta} \in \mathbb{R}^k$ fulfills $\boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta} < 1$, then $\left(\begin{smallmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{smallmatrix}\right)$ is positive definite. To this end, we denote with $\Omega^{1/2}$ the positive definite square root of Ω and, furthermore, note that $1 - \boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta} > 0$ holds. Therefore,

$$B = \begin{pmatrix} (1 - \boldsymbol{\delta}^T \Omega^{-1} \boldsymbol{\delta})^{1/2} & \boldsymbol{\delta}^T \Omega^{-1/2} \\ \mathbf{0}_k & \Omega^{1/2} \end{pmatrix}$$

is a regular $(k+1) \times (k+1)$ matrix and

$$BB^T = \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{pmatrix},$$

which implies that $\left(\begin{smallmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{smallmatrix}\right)$ is positive definite.

Furthermore, if Ω is a correlation matrix, then all diagonal elements of Ω are equal to 1 and the positive definiteness of $\left(\begin{smallmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \Omega \end{smallmatrix}\right)$ implies that $\boldsymbol{\delta} \in (-1, 1)^{\times k}$

B Figures

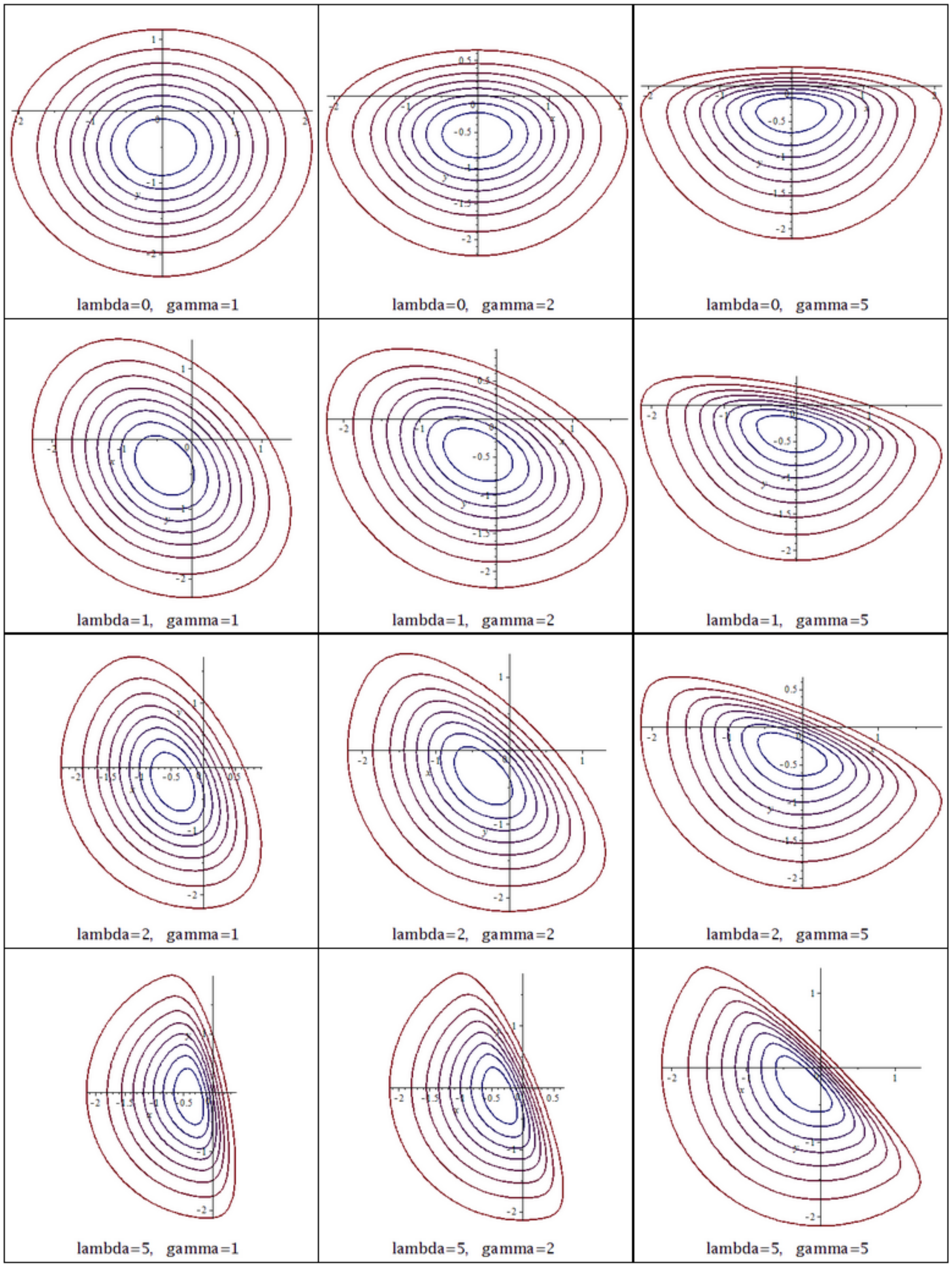

Figure 2: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$ with $\boldsymbol{a}_0 = (1, \lambda, \gamma)^T$, $\boldsymbol{a}_1 = (0, 1, 0)^T$, $\boldsymbol{a}_2 = (0, 0, 1)^T$. Here, $\boldsymbol{a}_1 \perp \boldsymbol{a}_2$ and changes are only made in $\boldsymbol{a}_0$. In the first row, we have additionally $\boldsymbol{a}_0 \perp \boldsymbol{a}_1$.

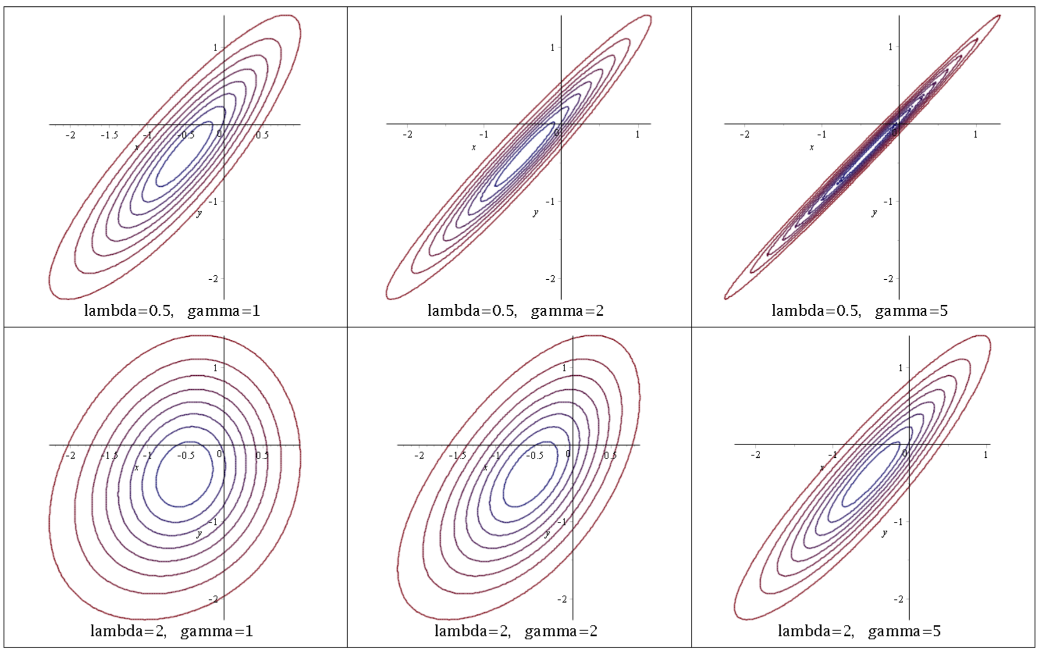

Figure 3: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$ with $\boldsymbol{a}_0 = (1, 1, 1)^T$, $\boldsymbol{a}_1 = (0, \lambda, \gamma)^T$, $\boldsymbol{a}_2 = (0, 0, 1)^T$. Changes are only in $\boldsymbol{a}_1$ and $\lambda > 0, \gamma > 0$.

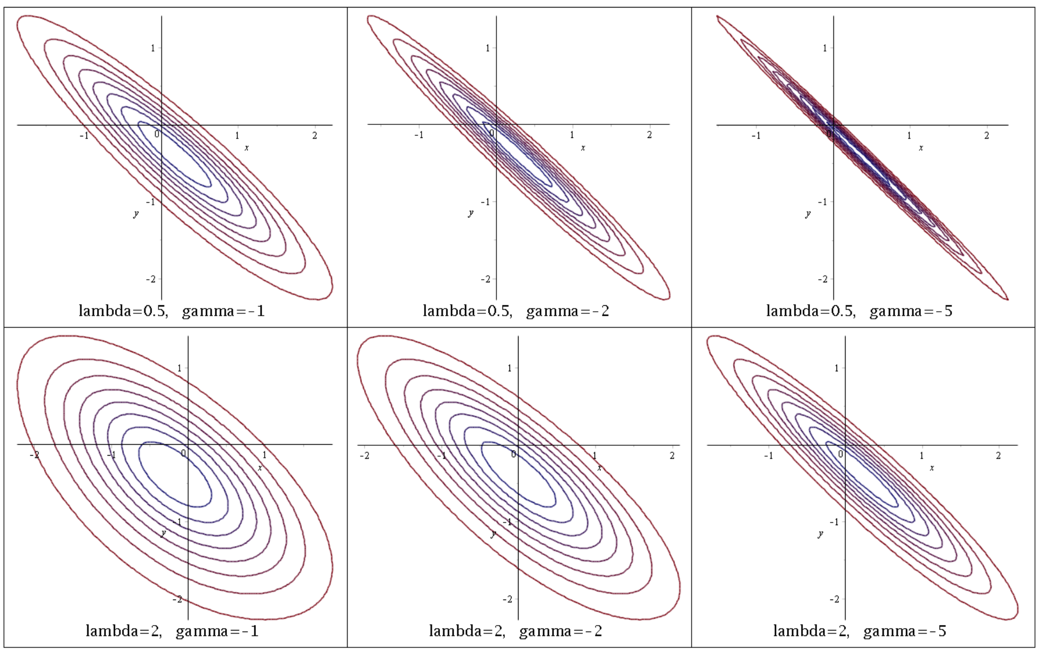

Figure 4: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$ with $\boldsymbol{a}_0 = (1,1,1)^T$, $\boldsymbol{a}_1 = (0, \lambda, \gamma)^T$, $\boldsymbol{a}_2 = (0,0,1)^T$. Changes are only in $\boldsymbol{a}_1$ and $\lambda > 0$, $\gamma < 0$.

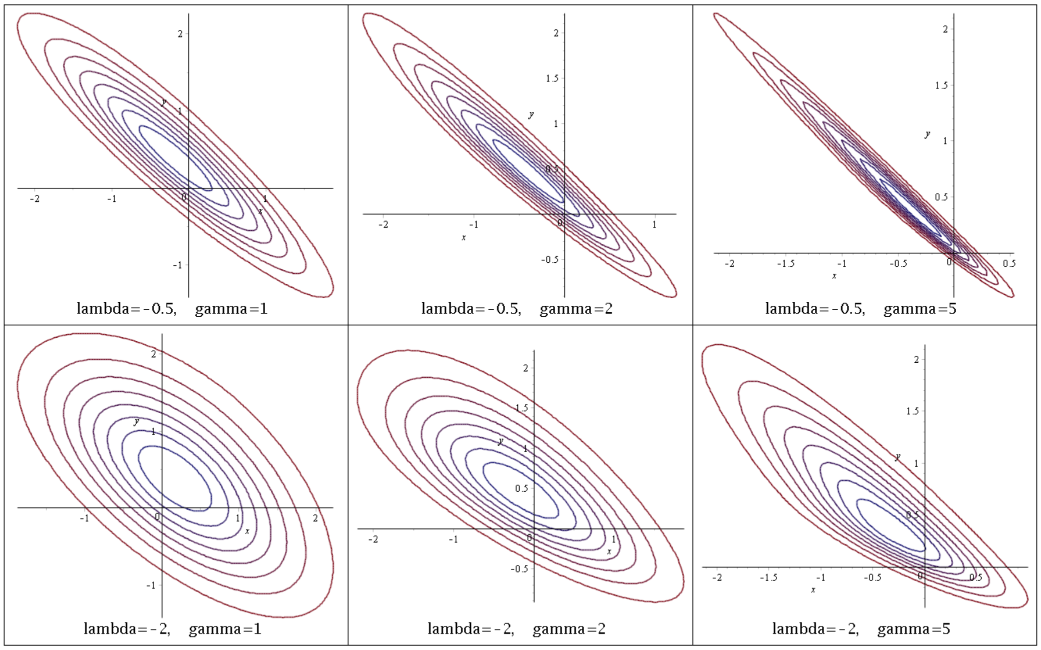

Figure 5: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$ with $\boldsymbol{a}_0 = (1, 1, \gamma)^T$, $\boldsymbol{a}_1 = (0, \lambda, \gamma)^T$, $\boldsymbol{a}_2 = (0, 0, -1)^T$. Changes are made in $\boldsymbol{a}_0$ and $\boldsymbol{a}_1$.

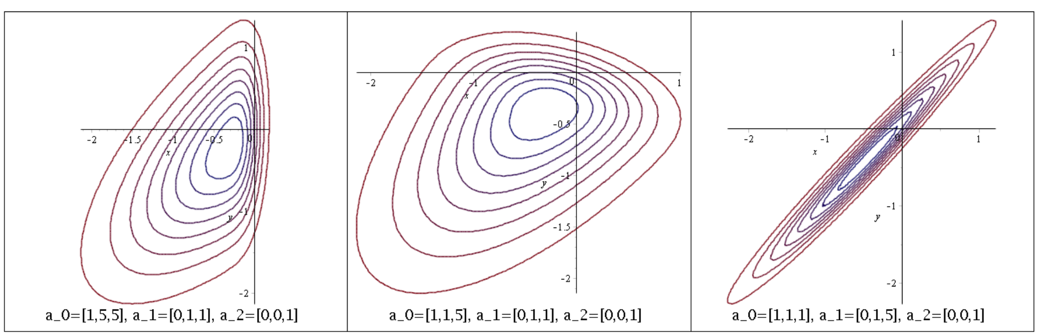

Figure 6: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$. One of the three angles is chosen particularly sharp. It is $\angle(\boldsymbol{a}_0, \boldsymbol{a}_1)$ in the left figure, $\angle(\boldsymbol{a}_0, \boldsymbol{a}_2)$ in the center figure, and $\angle(\boldsymbol{a}_1, \boldsymbol{a}_2)$ in the right figure.

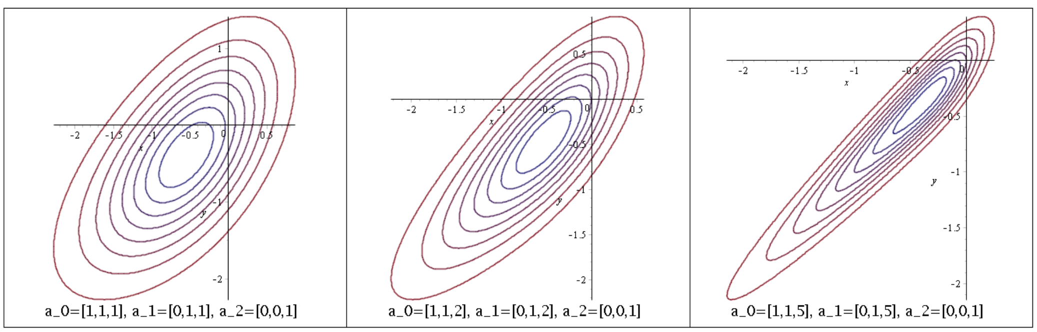

Figure 7: Contour plot of the density of (Z_1, Z_2) where $P((Z_1, Z_2) < (z_1, z_2)) = 2\Phi_3(C_3(\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, z_1, z_2))$. All three angles between $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2$ decrease from the left to the right.