The version of record of this article, first published in Mathematica Slovaca, is available online at Publisher's website: https://doi.org/10.1515/ms-2023-0039

NEW INSIGHTS ON THE MULTIVARIATE SKEW EXPONENTIAL POWER DISTRIBUTION

JORGE M AREVALILLO* AND HILARIO NAVARRO**

Abstract. The multivariate exponential power is a useful distribution for modeling departures from normality in data by means of a tail weight scalar parameter that regulates the non-normality of the model. The incorporation of a shape asymmetry vector into the model serves to account for potential asymmetries and gives rise to the multivariate skew exponential power distribution. This work is aimed at revisiting the skew exponential power distribution taking as a starting point its formulation as a scale mixture of skew-normal distributions. The paper provides some highlights and theoretical insights on the role played by its parameters to assess two complementary aspects of the multivariate non-normality such as directional asymmetry and tail weight behavior regardless of the asymmetry; the intuition behind both issues relies on well-known mathematical ideas about skewness maximization and convex transform stochastic orderings.

1. Introduction

The exponential power (EP) distribution has an ancient history in statistics which dates back to Subbotin's pioneer work [42] with an early multivariate generalization introduced by De Simoni [41], later on quoted by the follow-up [24]; it can also be considered as a particular case of the Kotz-type family of distributions [31] as highlighted in [19, Section 3.2]. Unlike other non-normal distributions, like for example the t distribution, the EP family has the advantage of including distributions having heavier and lighter tails than the normal one with the normal model being an intermediate distribution within the family. Due to its flexibility and wide scope, a lot of research has been carried out since the aforementioned foundational works with contributions that show its usefulness both in univariate [13, 14, 46, 1, 18, 25, 39, 45] and multivariate settings [32, 22, 23, 35, 20], a matrix generalization of the model [40], its study in the context of the general class of elliptically contoured distributions [21, 3, 30, 38], some caveats about its limitation to remain closed under marginalization [27], as well as related variants to assess asymmetry [6, 2, 12] and some other models under which this family can be considered [37, 38].

This paper adopts the approach of several of the aforementioned works [24, 19, 27, 21] to define the multivariate EP distribution as follows: we say that an input vector $\boldsymbol{X}$ follows a p-dimensional exponential power distribution with location vector $\boldsymbol{\xi}$ and full rank scale matrix $\boldsymbol{\Omega}$ if its probability density function (pdf) is given by

$$f(\boldsymbol{x};\boldsymbol{\xi},\boldsymbol{\Omega},\beta)=\frac{\beta\Gamma(\frac{p}{2})}{\pi^{p/2}\Gamma(\frac{p}{2\beta})2^{p/2\beta}}\mid\boldsymbol{\Omega}\mid^{-1/2}\exp\left\{-\frac{1}{2}\left[(\boldsymbol{x}-\boldsymbol{\xi})^{\top}\boldsymbol{\Omega}^{-1}(\boldsymbol{x}-\boldsymbol{\xi})\right]^{\beta}\right\},\qquad(1.1)$$

$\boldsymbol{x}\in\mathbb{R}^p$, where $\beta>0$ is a tail weight parameter accounting for the peakedness of the model. The notation $\boldsymbol{X}\sim EP_p(\boldsymbol{\xi},\boldsymbol{\Omega},\beta)$ is used to indicate that $\boldsymbol{X}$ has a distribution with pdf in (1.1). Note that when $\beta=1$ the EP vector has a multivariate normal distribution $N_p(\boldsymbol{\xi},\boldsymbol{\Omega})$. Figure 1

2010 *Mathematics Subject Classification.* Primary 62H99, 60E15; Secondary 60E99.

Key words and phrases. Skew exponential power, asymmetry, skewness, tail weight, kurtosis.

illustrates the differences of peakedness by visual comparison of the shapes of the bidimensional normal ($\beta = 1$) and double exponential ($\beta = 1/2$) density functions.

FIGURE 1. Probability density functions for the bivariate double exponential (left) and normal (right) variables with location $\boldsymbol{\xi} = (0, 0)$ and scale matrix $\boldsymbol{\Omega} = \boldsymbol{I}_2$.

The skew exponential power (SEP) class arises when asymmetry is injected into a EP distribution so that a skewed model is obtained as a result; some past work has addressed the issue dealing with different variants of skewed EP distributions: see the works [6, 2, 12] or the alternative formulation as a scale mixture of skew-normal (SMSN) distributions [15, Section 3]. This paper examines the multivariate SEP family from the latter representation, from which analytical closed for expressions for moments can be obtained [16, 28], a relevant issue when addressing skewness maximization. The paper is organized as follows: The next section gives some background about the multivariate SN distribution and the SMSN representation of the SEP vector. Section 3 discusses the role played by the shape vector and the tail weight β as parameters that account for the multivariate non-normality of the model; some novel insights regarding the assessment of directional asymmetry, through the direction that yields the maximal skewness projection, as well as the connection of the tail weight parameter with well established multivariate convex transform stochastic orderings, are studied in order to enhance their interpretation in the assessment of the multivariate non-normality. Finally, Section 4 recaps the main findings of the paper suggesting yet unexplored ideas to advance progress towards future research directions.

2. The multivariate skew exponential power distribution

This section deals with some material aimed at presenting the SMSN formulation of the SEP distribution. Firstly, some background about the SN distribution and its extension to the SMSN family is provided; then the representation of the multivariate exponential power distribution as a scale mixture of normal (SMN) distributions is recalled in order to introduce the SMSN representation of the SEP model in a natural way.

2.1. **The multivariate skew-normal and SMSN families.** The multivariate skew-normal (SN) distribution was introduced by [9] to regulate asymmetry departures from normality. Here, it is adopted the notation of the original seminal works [9, 8] to define the pdf of a p-dimensional SN vector with location vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)^\top$ and scale matrix $\boldsymbol{\Omega}$ as follows:

$$f(\boldsymbol{x}; \boldsymbol{\xi}, \boldsymbol{\alpha}, \boldsymbol{\Omega}) = 2\phi_p(\boldsymbol{x} - \boldsymbol{\xi}; \boldsymbol{\Omega})\Phi(\boldsymbol{\alpha}^\top \boldsymbol{\omega}^{-1}(\boldsymbol{x} - \boldsymbol{\xi})) \quad : \quad \boldsymbol{x} \in \mathbb{R}^p, \tag{2.1}$$

where $\phi_p(\cdot;\boldsymbol{\Omega})$ denotes the pdf of a p-dimensional normal variate with zero mean and covariance matrix $\boldsymbol{\Omega}$, Φ is the distribution function of a standard $N(0,1)$ scalar variable, $\boldsymbol{\omega} = diag(\omega_1, \ldots, \omega_p)$ is a scale diagonal matrix with non negative entries such that $\overline{\boldsymbol{\Omega}} = \boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\omega}^{-1}$ is a correlation matrix and $\boldsymbol{\alpha}$ is a p-dimensional shape vector that regulates the asymmetry of the model. If $\boldsymbol{\alpha} = \mathbf{0}$ then the pdf in (2.1) reduces to the density of a normal vector with mean vector $\boldsymbol{\xi}$ and covariance matrix $\boldsymbol{\Omega}$. We use the notation $\boldsymbol{X} \sim SN_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ to indicate that $\boldsymbol{X}$ follows a p-dimensional skew-normal distribution with pdf given by (2.1).

Note that the diagonal matrix $\boldsymbol{\omega}$ can be written as $\boldsymbol{\omega} = (\boldsymbol{\Omega} \odot \boldsymbol{I}_p)^{1/2}$, where the symbol $\odot$ denotes the entry-wise matrix product. It also holds that $\boldsymbol{X} = \boldsymbol{\xi} + \boldsymbol{\omega}\boldsymbol{Z}$, where $\boldsymbol{Z}$ is a normalized multivariate skew-normal variable with pdf given by

$$f(\boldsymbol{z}; \mathbf{0}, \boldsymbol{\alpha}, \overline{\boldsymbol{\Omega}}) = 2\phi_p(\boldsymbol{z}; \overline{\boldsymbol{\Omega}})\Phi(\boldsymbol{\alpha}^\top \boldsymbol{z}). \tag{2.2}$$

The normalized SN variate has a simple pdf and a tractable stochastic representation which enables its extension to the wider SMSN family in a natural way [11, 16, 17] (with reference [17] a posthumous reprint of [16]). This paper adopts the formulation from the aforementioned works to define a SMSN vector as follows:

Definition 1. Let $\boldsymbol{Z}$ be a random vector such that $\boldsymbol{Z} \sim SN_p(\mathbf{0}, \overline{\boldsymbol{\Omega}}, \boldsymbol{\alpha})$, with pdf (2.2), and let S be a non negative scalar variable independent of $\boldsymbol{Z}$. The random vector $\boldsymbol{X} = \boldsymbol{\xi} + \boldsymbol{\omega} S \boldsymbol{Z}$, where $\boldsymbol{\omega}$ is a diagonal matrix with non negative entries, is said to follow a multivariate SMSN distribution.

We put $\boldsymbol{X} \sim SMSN_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, H)$, with H the distribution function of the mixing variable S and $\boldsymbol{\Omega} = \boldsymbol{\omega}\overline{\boldsymbol{\Omega}}\boldsymbol{\omega}$, to indicate that $\boldsymbol{X}$ follows a p-dimensiontal SMSN distribution. Note that when H is degenerate at $S = 1$ then $\boldsymbol{X}$ becomes a $SN_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ vector.

2.2. **The SMSN formulation of the skew exponential power distribution.** The SMSN distribution reduces to a scale mixture of normal distributions when the shape vector is equal to the p-dimensional zero vector, $\boldsymbol{\alpha} = \mathbf{0}_p$, since the vector $\boldsymbol{Z}$ from Definition 1 has a p-dimensional normal distribution, i.e. $\boldsymbol{Z} \sim N_p(\mathbf{0}, \overline{\boldsymbol{\Omega}})$. We put $\boldsymbol{X} \sim SMN_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, H)$ to indicate that $\boldsymbol{X}$ follows a SMN distribution. The multivariate EP distribution with pdf (1.1) admits a SMN formulation as long as the tail weight parameter satisfies that $0 < \beta \leq 1$, as the following proposition shows [23].

Proposition 2.1. *Let $\boldsymbol{X}$ be a p-dimensional vector such that $\boldsymbol{X} \sim EP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \beta)$. Then it holds that $\boldsymbol{X} \sim SMN_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, H)$ if and only if $0 < \beta \leq 1$. When $\beta \in (0,1)$ the distribution H of the mixing variable is absolutely continuous and it has pdf given by*

$$h_\beta(s) = \frac{2^{\frac{p}{2}(1-\frac{1}{\beta})}\Gamma(1+\frac{p}{2})}{\Gamma(1+\frac{p}{2\beta})}\, s^{p-3} S_\beta\left(\frac{1}{2}s^{-2}; 2^{-\frac{1}{\beta}}\right) \quad : \quad s > 0, \tag{2.3}$$

where $S_\beta(\cdot; 2^{-\frac{1}{\beta}})$ denotes the pdf of a non negative stable distribution having characteristic function $\varphi(t) = \exp\left\{-\frac{1}{2}|t|^\beta e^{-i\frac{\pi}{2}\beta sign(t)}\right\}$. On the other hand, when $\beta = 1$ H is the degenerate distribution at $S = 1$.

The result of Proposition 2.1 motivates to define the skew EP distribution as follows.

Definition 2. Let $\boldsymbol{Z}$ be a vector such that $\boldsymbol{Z} \sim SN_p(\mathbf{0}, \overline{\boldsymbol{\Omega}}, \boldsymbol{\alpha})$ with pdf (2.2) and let S be a non negative scalar variable, independent of $\boldsymbol{Z}$, with the density function of Proposition 2.1 for any $\beta \in (0, 1]$. If $\boldsymbol{\omega}$ is a diagonal matrix with non-negative entries then we say that the random vector $\boldsymbol{X} = \boldsymbol{\xi} + \boldsymbol{\omega} S \boldsymbol{Z}$ follows a multivariate skew exponential power distribution.

We write $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$ to indicate that $\boldsymbol{X}$ follows a p-dimensional SEP distribution with location $\boldsymbol{\xi}$, scale matrix $\boldsymbol{\Omega} = \boldsymbol{\omega}\overline{\boldsymbol{\Omega}}\boldsymbol{\omega}$, shape vector $\boldsymbol{\alpha}$ and tail weight parameter $\beta \in (0, 1]$. It is worthwhile noting that the SEP family of distributions belongs to the wider skew-elliptical class of distributions [15] and it is also closed under full-rank affine transformations.

From the stochastic representation of the SEP vector in Definition 2, the mean vector and the covariance matrix of the model are obtained by combining (6.18) of [11] with Lemma 2.1:

Lemma 2.1. *Let $\boldsymbol{X}$ be a vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$. Then the moments of the mixing variable are given by* $E(S^k) = \dfrac{2^{k/2\beta}\Gamma(\frac{p}{2})\Gamma(\frac{p+k}{2\beta})}{2^{k/2}\Gamma(\frac{p}{2\beta})\Gamma(\frac{p+k}{2})}$.

Proof. See the appendix. □

Then it can be shown in an easy way that

$$E(\boldsymbol{X}) = \boldsymbol{\xi} + E(S)\sqrt{\frac{2}{\pi}}\boldsymbol{\gamma} \text{ and } var(\boldsymbol{X}) = E(S^2)\boldsymbol{\Omega} - E(S)^2\frac{2}{\pi}\boldsymbol{\gamma}\boldsymbol{\gamma}'$$

with $E(S) = 2^{\frac{1-\beta}{2\beta}}\dfrac{\Gamma(\frac{p}{2})\Gamma(\frac{p+1}{2\beta})}{\Gamma(\frac{p}{2\beta})\Gamma(\frac{p+1}{2})}$, $E(S^2) = 2^{\frac{1-\beta}{\beta}}\dfrac{2\Gamma(\frac{p+2}{2\beta})}{p\Gamma(\frac{p}{2\beta})}$, $\boldsymbol{\gamma} = \dfrac{\boldsymbol{\Omega}\boldsymbol{\eta}}{\sqrt{1+\boldsymbol{\eta}'\boldsymbol{\Omega}\boldsymbol{\eta}}}$ and $\boldsymbol{\eta} = \boldsymbol{\omega}^{-1}\boldsymbol{\alpha}$.

The stochastic representation also provides a natural strategy for the simulation of observations from a p-dimensional SEP vector. Although the observations from the SN vector $\boldsymbol{\omega}\boldsymbol{Z}$ can be obtained in an easy way using the functionalities of the `sn` R package [7], the non tractability of the density function (2.3) makes the simulation of the mixing variable a difficult task. However, for some specific tail weight parameters the mixing variable has a well-known distribution; in such cases we can simulate observations from S: for example, when $\beta = 1$, the model becomes the SN distribution, whereas $\beta = 1/2$ corresponds to the skewed version of the multivariate double exponential distribution, with the mixing variable S following a generalized gamma distribution [23]. Figure 2 displays the plots of the pdf for the skewed bivariate double exponential variate; it illustrates the effect caused by the parametric vector $\boldsymbol{\alpha}$ on the shape of the densities as well as the contoured plots obtained after injection of asymmetry across different directions.

3. Insights on the multivariate non-normality of the model

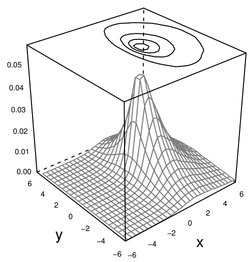
This section examines the role played by the parameters of the SEP model to assess multivariate non-normality; some ideas that elaborate on kurtosis and directional asymmetry are discussed next.

3.1. **Assessment of kurtosis.**

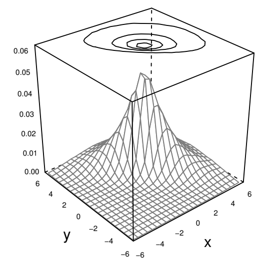
First of all, we introduce some notation: Let $\boldsymbol{X}$ be SEP vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$ and let us denote by $\boldsymbol{X}_0$ the vector obtained when $\boldsymbol{\alpha} = \boldsymbol{0}_p$, where $\boldsymbol{0}_p$ is the p-dimensional zero vector so that $\boldsymbol{X}_0 \sim EP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \beta)$. Let us consider the quadratic forms $\boldsymbol{Q}_X$ and $\boldsymbol{Q}_{X_0}$, associated to $\boldsymbol{X}$ and $\boldsymbol{X}_0$, given by $\boldsymbol{Q}_X = (\boldsymbol{X} - \boldsymbol{\xi})^\top\boldsymbol{\Omega}^{-1}(\boldsymbol{X} - \boldsymbol{\xi})$ and $\boldsymbol{Q}_{X_0} = (\boldsymbol{X}_0 - \boldsymbol{\xi})^\top\boldsymbol{\Omega}^{-1}(\boldsymbol{X}_0 - \boldsymbol{\xi})$. Finally, let us denote by *SEP* the class of skew exponential power distributions having a SMSN representation, which is given by

$$SEP = \{\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta) : \boldsymbol{\xi} \in \mathbb{R}^p, \boldsymbol{\Omega} \in \mathbb{R}^p \times \mathbb{R}^p, \boldsymbol{\alpha} \in \mathbb{R}^p, 0 < \beta \leq 1\}.$$

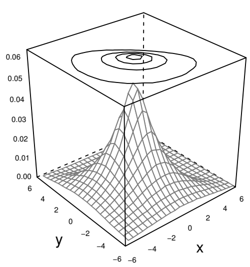
Since $\boldsymbol{X}_0$ follows an elliptical distribution, the quadratic form $\boldsymbol{Q}_{X_0}$ has the same distribution as the so called modular variable, R^2_β, from the stochastic representation of the elliptical vector $\boldsymbol{X}_0$ [21]. On the other hand, as the SEP distribution belongs to the SMSN class —which in turn has the perturbed symmetry structure of [11, Proposition 1.1] (see [12] page 872 or [11] page 172)— then the SEP vector meets the modulation invariance property from [11, Proposition 1.4] or [10, Proposition 2] so we can assert that $\boldsymbol{Q}_X$ and $\boldsymbol{Q}_{X_0}$ are equally distributed, i.e. $\boldsymbol{Q}_X \overset{d}{=} \boldsymbol{Q}_{X_0}$. This



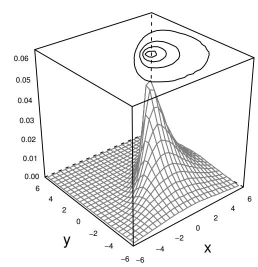
(A) $\boldsymbol{\alpha}^\top = (3, 0)$



(B) $\boldsymbol{\alpha}^\top = (3, 3)$



(C) $\boldsymbol{\alpha}^\top = (-3, -3)$



(D) $\boldsymbol{\alpha}^\top = (3, -3)$

FIGURE 2. Density functions of the bivariate double exponential ($\beta = 1/2$), with location $\boldsymbol{\xi} = (0, 0)$ and scale matrix $\boldsymbol{\Omega} = \boldsymbol{I}_2$, for different shape vectors.

finding could also follow from [11, Corollary 5.9], also a consequence of the modulation invariance property. Therefore, we conclude that $\boldsymbol{Q}_X \overset{d}{=} \boldsymbol{Q}_{X_0} \overset{d}{=} R_\beta^2$.

We now define a stochastic relation between *SEP* vectors as follows:

Definition 3. Let $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be two vectors such that $\boldsymbol{X}_i \sim SEP_p(\boldsymbol{\xi}_i, \boldsymbol{\Omega}_i, \boldsymbol{\alpha}_i, \beta_i) : i = 1, 2$. We say that $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are modular related, and we denote it by $\boldsymbol{X}_1 \mathcal{R} \boldsymbol{X}_2$, if their corresponding quadratic forms $\boldsymbol{Q}_{X_1}$ and $\boldsymbol{Q}_{X_2}$ have the same distribution, i.e. $\boldsymbol{Q}_{X_1} \overset{d}{=} \boldsymbol{Q}_{X_2}$.

It can be shown that $\mathcal{R}$ defines an equivalence relation. The equivalence classes that form the quotient space induced by $\mathcal{R}$ are characterized as follows: for each vector $\boldsymbol{Y}$ such that $\boldsymbol{Y} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta^*)$, its equivalence class is the subset in *SEP* that contain the vectors fulfilling $\boldsymbol{Q}_X \overset{d}{=} \boldsymbol{Q}_Y$; it is given by

$$SEP_{\beta^*} = \{\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta) : \boldsymbol{\xi} \in \mathbb{R}^p, \boldsymbol{\Omega} \in \mathbb{R}^p \times \mathbb{R}^p, \boldsymbol{\alpha} \in \mathbb{R}^p, \beta = \beta^*\}$$

so it holds that $SEP = \bigcup_{\beta \in (0,1]} SEP_\beta$. Hence, *SEP* can be partitioned in accordance to the tail weight parameter, with each equivalence class being characterized by all the p-dimensional SEP distributions having the same tail weight parameter. This is a revealing insight that motivates the kurtosis stochastic comparison of equivalence classes borrowing some ideas of previous works [44, 3] on Van Zwet's convex transform ordering [43]. The next definition formalizes this intuition.

Definition 4 (kurtosis ordering between *SEP* vectors)**.** Let us consider $\boldsymbol{X}$ and $\boldsymbol{Y}$ two SEP vectors such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}_1, \boldsymbol{\Omega}_1, \boldsymbol{\alpha}_1, \beta_1)$ and $\boldsymbol{Y} \sim SEP_p(\boldsymbol{\xi}_2, \boldsymbol{\Omega}_2, \boldsymbol{\alpha}_2, \beta_2)$. We say that $\boldsymbol{X}$ is less than or equal to $\boldsymbol{Y}$ in kurtosis, and we denote it by $\boldsymbol{X} \leq_k \boldsymbol{Y}$, if and only if $F^{-1}_{Q_Y}(F_{Q_X}(r))$ is convex for $r \geq 0$, where F_{Q_X} and F_{Q_Y} are the distribution functions of their corresponding quadratic forms: $\boldsymbol{Q}_X = (\boldsymbol{X} - \boldsymbol{\xi}_1)^\top \boldsymbol{\Omega}_1^{-1} (\boldsymbol{X} - \boldsymbol{\xi}_1)$ and $\boldsymbol{Q}_Y = (\boldsymbol{Y} - \boldsymbol{\xi}_2)^\top \boldsymbol{\Omega}_2^{-1} (\boldsymbol{Y} - \boldsymbol{\xi}_2)$.

The k-ordering handles the comparison of SEP vectors by setting the problem in terms of the convex transform ordering of their quadratic forms $\boldsymbol{Q}_X$ and $\boldsymbol{Q}_Y$. The next theorem shows that the order in Definition 4 is actually a total ordering in the same direction as the tail weight parameter β.

Theorem 3.1. *Let us consider two random vectors* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *such that* $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}_1, \boldsymbol{\Omega}_1, \boldsymbol{\alpha}_1, \beta_1)$ *and* $\boldsymbol{Y} \sim SEP_p(\boldsymbol{\xi}_2, \boldsymbol{\Omega}_2, \boldsymbol{\alpha}_2, \beta_2)$*. If* $0 < \beta_1 \leq \beta_2 \leq 1$ *then* $\boldsymbol{Y} \leq_k \boldsymbol{X}$ *when* $p > 2$*.*

Proof. We must proof the convexity of the function $F^{-1}_{Q_X}(F_{Q_Y}(r))$ for $r \geq 0$.

The proof follows from a simple argument that uses the definition of the relation $\mathcal{R}$ as follows: from the properties of quadratic forms of SN vectors [11] and the properties of the multivariate exponential power distribution [21] it is obtained that $\boldsymbol{Q}_X \stackrel{d}{=} R^2_{\beta_1}$ and $\boldsymbol{Q}_Y \stackrel{d}{=} R^2_{\beta_2}$, where $R^2_{\beta_1}$ and $R^2_{\beta_2}$ are the modular variables associated with exponential power vectors $\boldsymbol{X}_0$ and $\boldsymbol{Y}_0$ such that $\boldsymbol{X}_0 \sim EP_p(\boldsymbol{\xi}_1, \boldsymbol{\Omega}_1, \beta_1)$ and $\boldsymbol{Y}_0 \sim EP_p(\boldsymbol{\xi}_2, \boldsymbol{\Omega}_2, \beta_2)$. Taking into account [3, Proposition 3], we can assert that $F^{-1}_{R^2_{\beta_1}}(F_{R^2_{\beta_2}}(r))$ is a convex function for $r \geq 0$ from which the convexity of the function $F^{-1}_{Q_X}(F_{Q_Y}(r))$ follows. □

Theorem 3.1 provides intriguing theoretical insights. In addition to providing a mathematically sounded result for the stochastic of equivalence classes from the relation $\mathcal{R}$, it also enhances the role played by the tail weight parameter β as an indicator of kurtosis compatible with the multivariate convex transform k-ordering. Thus, the implications are twofold: on the one hand, it serves to carry out multivariate stochastic comparisons between SEP vectors irrespective of their location, scale and shape asymmetry vector by comparing their tail weight parameters; on the other hand, it allows to assess the multivariate kurtosis of wide subclasses within *SEP*, as determined by the equivalence classes SEP_β, using a single parameter β which has a meaningful interpretation in terms of a kurtosis convex transform ordering.

3.2. **Assessment of directional asymmetry.** While the tail weight parameter β regulates the peakedness of the distribution irrespective of its asymmetry, the shape vector $\boldsymbol{\alpha}$ —or its equivalent counterpart $\boldsymbol{\eta} = \boldsymbol{\omega}^{-1}\boldsymbol{\alpha}$— is aimed at accounting for multivariate asymmetry in a directional fashion. In order to delve into its role to handle directional asymmetry, we set out the problem of finding directions that yield maximal skewness projections when the input vector $\boldsymbol{X}$ follows a SEP distribution. The problem can be formally described as follows:

Let $\boldsymbol{X}$ be a p-dimensional input vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$ with $0 < \beta \leq 1$. Our goal is to find the vector $\boldsymbol{c}$ for which the scalar variable $Y = \mathbf{c}^\top \mathbf{X}$ attains the maximal skewness, that is, the vector yielding the maximal skewness proyection. This goal is accomplished by solving the optimization problem: $\max\limits_{\boldsymbol{c} \in \mathbb{R}^p_0} \gamma_1(\mathbf{c}^\top \mathbf{X})$, with γ_1 the standard third moment skewness measure defined by $\gamma_1(Y) = E^2\left(\dfrac{Y - \mu_Y}{\sigma_Y}\right)^3$ and $\mathbb{R}^p_0$ the set of all non null p-dimensional vectors.

Since γ_1 is scale invariant, without loss of generality, we can confine to vectors such that $\boldsymbol{c}^\top \boldsymbol{\Omega} \boldsymbol{c} = 1$; hence, the problem can be formulated as

$$\max_{\boldsymbol{c}\in\mathbb{S}_p} \gamma_1(\mathbf{c}^\top \mathbf{X}) \tag{3.1}$$

where $\mathbb{S}_p = \{\boldsymbol{c} \in \mathbb{R}^p : \boldsymbol{c}^\top \boldsymbol{\Omega} \boldsymbol{c} = 1\}$.

The maximal skewness $\gamma_{1,p}^D = \max_{\boldsymbol{c}\in\mathbb{S}_p} \gamma_1(\mathbf{c}^\top \mathbf{X})$, with the superscript D standing for directional skewness, is a well-known measure that allows to capture the directional nature of the asymmetry [36]. Theorem 3.2 shows that the maximal skewness projection is attained at the direction of the shape vector of the SEP model, regardless of the value of the tail weight parameter. Its proof uses the following auxiliary lemmas.

Lemma 3.1. *Let S be the mixing variable of a scale mixture model with density function (2.3). Then it holds the moment inequality $E(S^3) - E(S)E(S^2) \geq 0$.*

Proof. See the appendix. □

Lemma 3.2. *Let $\boldsymbol{X}$ be a vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$. Then the moments of the mixing variable meet the following inequality: $\dfrac{4}{\pi}E(S)E(S^3) - E(S^2)^2 \geq 0$.*

Proof. See the appendix. □

Lemma 3.3. *Let $\boldsymbol{X}$ be a vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$. Then it holds the following inequality: $E(S^3) + \dfrac{4}{\pi}E(S)^3 - 2E(S)E(S^2) \geq 0$.*

Proof. See the appendix. □

The result on directional asymmetry is now stated in the next theorem.

Theorem 3.2. *Let $\boldsymbol{X}$ be a random vector such that $\boldsymbol{X} \sim SEP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha}, \beta)$ with $0 < \beta \leq 1$. Then the maximum skewness in (3.1) is attained at the direction of the shape vector $\boldsymbol{\eta}^\top = \boldsymbol{\alpha}^\top \boldsymbol{\omega}^{-1}$, specifically at $\boldsymbol{c}_*^\top = \boldsymbol{\alpha}^\top \boldsymbol{\omega}^{-1} / \sqrt{\boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha}} = \boldsymbol{\eta}^\top / \sqrt{\boldsymbol{\eta}^\top \boldsymbol{\Omega} \boldsymbol{\eta}}$.*

Proof. When $\beta = 1$, the input vector $\boldsymbol{X}$ follows a SN; so the proof follows from [33]. In order to prove the statement when $\beta \in (0, 1)$, we use the lemmas in the Appendix in combination with some of the arguments deployed by [33, 4].

Since γ_1 is location invariant, we can assume that $\boldsymbol{\xi} = \boldsymbol{0}$ for the sake of simplicity. Taking into account the SMSN representation of the SEP vector, the properties of the SN under linear transformations —see (5.42) and (5.43) from [11] or alternatively the earlier work [8]— and the assumption $\boldsymbol{c}^\top \boldsymbol{\Omega} \boldsymbol{c} = 1$ for the vector driving the direction, it follows that the scalar variable $Y = \boldsymbol{c}^\top \boldsymbol{X} = SZ$, with $Z = \boldsymbol{c}^\top \boldsymbol{\omega} \boldsymbol{Z}$, has a SN distribution such that $Z \sim SN_1(0, 1, \lambda)$, where the shape scalar parameter λ is given by

$$\lambda = \frac{\boldsymbol{c}^\top \boldsymbol{\omega} \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha} - (\boldsymbol{c}^\top \boldsymbol{\omega} \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha})^2}}.$$

Therefore, Y admits a SNSM formulation so that we can apply Proposition 3 from [16] to get

$$\gamma_1(Y) = E^2 \left(\frac{Y - \mu_Y}{\sigma_Y} \right)^3 = \frac{2}{\pi} \frac{\delta^2 \left(a\delta^2 - 3b \right)^2}{\sigma_Y^6} \ : \ 0 \leq \delta^2 < 1, \tag{3.2}$$

where the quantities appearing in (3.2) are defined by $a = E(S)^3 \dfrac{4}{\pi} - E(S^3)$, $b = E(S)E(S^2) - E(S^3)$, $\sigma_Y^2 = E(S^2) - \dfrac{2}{\pi} E(S)^2 \delta^2$ and $\delta^2 = \dfrac{\lambda^2}{1 + \lambda^2}$. Note that simple calculus using Lemma 2.1

with the expectations involved in $\gamma_1(Y)$ would give an explicit expression for the skewness on any direction.

We now show that $\gamma_1(Y)$ is non decreasing function with respect to δ^2. Its first derivative is given by

$$\frac{\partial \gamma_1(Y)}{\partial \delta^2} = \frac{6}{\pi} \frac{(a\delta^2 - 3b)(a\delta^2 - b)\sigma_Y^2 + E(S)^2(2/\pi)\delta^2(a\delta^2 - 3b)^2}{\sigma_Y^8}. \tag{3.3}$$

The quantities appearing in the expression above can be rearranged to give

$$\frac{\partial \gamma_1(Y)}{\partial \delta^2} = \frac{6}{\pi} \frac{(a\delta^2 - 3b)h(\delta^2)}{\sigma_Y^8},$$

where $h(\delta^2) = c\delta^2 - bE(S^2)$ with $c = E(S^3)[\frac{4}{\pi}E(S)^2 - E(S^2)]$.

Lemmas 3.1 and 3.3 show that $-b \geq 0$ and $a - 2b \geq 0$ which implies that $a - 3b \geq 0$. We now study the sign of the factors $a\delta^2 - 3b$ and $h(\delta^2)$ from the expression above.

If it happened that $a \geq 0$ then we would obtain that $a\delta^2 - 3b \geq 0$ whereas if $a < 0$, it would result that $a\delta^2 - 3b > a - 3b$ so $a\delta^2 - 3b \geq 0$ once again. To study the sign of $h(\delta^2)$, we also distinguish two cases: if $c \geq 0$ then it suffices to invoke Lemma 3.1 once again to get $h(\delta^2) \geq 0$; meanwhile, if it happened that $c < 0$ then, taking into account Lemma 3.2, we obtain that $h(1) = c - bE(S^2) \geq 0$; so we conclude that $h(\delta^2) > h(1) \geq 0$.

The non negativity of $a\delta^2 - 3b$ and $h(\delta^2)$ shows that $\gamma_1(Y)$ is a non decreasing function of δ^2. Therefore, the maximization of the skewness measure $\gamma_1(\boldsymbol{c}^\top \boldsymbol{X})$ is equivalent to the maximization of the quantity

$$\delta^2 = \frac{\lambda^2}{1 + \lambda^2} = \frac{(\boldsymbol{c}^\top \boldsymbol{\omega} \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha})^2}{1 + \boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha}},$$

where $\boldsymbol{c}^\top \boldsymbol{\omega} \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha} = (\boldsymbol{c}^\top \boldsymbol{\Omega}^{1/2})(\boldsymbol{\Omega}^{1/2} \boldsymbol{\omega}^{-1} \boldsymbol{\alpha})$ with $\boldsymbol{\Omega}^{1/2}$ a positive definite symmetric matrix such that $\boldsymbol{\Omega}^{1/2}\boldsymbol{\Omega}^{1/2} = \boldsymbol{\Omega}$. Hence, using an analogous argument as in [33, Proposition 2.2], we can assert that $(\boldsymbol{c}^\top \boldsymbol{\omega} \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha})^2 \leq \boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha}$ from which we conclude that the maximal skewness is attained at the direction of the vector $\boldsymbol{c}_*^\top = \boldsymbol{\alpha}^\top \boldsymbol{\omega}^{-1} / \sqrt{\boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha}}$, as we aimed to prove. □

Note that if we replace δ^2 in expression (3.2) by the quantity $\delta_*^2 = \dfrac{\lambda_*^2}{1 + \lambda_*^2}$, where $\lambda_*^2 = \boldsymbol{\alpha}^\top \overline{\boldsymbol{\Omega}} \boldsymbol{\alpha} = \boldsymbol{\eta}^\top \boldsymbol{\Omega} \boldsymbol{\eta}$ is the value of λ^2 at the maximal skewness direction $\boldsymbol{c}_*^\top$, we would obtain the skewness attained at the maximal skewness projection. It is given by

$$\gamma_{1,p}^D = \frac{2\lambda_*^2 \left[\left(\frac{4}{\pi} E^3(S) - 3E(S)E(S^2) + 2E(S^3) \right) \lambda_*^2 - 3(E(S)E(S^2) - E(S^3)) \right]^2}{\pi \left[E(S^2) + \left(E(S^2) - \frac{2}{\pi} E^2(S) \right) \lambda_*^2 \right]^3} \tag{3.4}$$

where the moments appearing in this expression can be calculated from their general expression in Lemma 2.1. Actually, the quantity above provides an analytical formula for Malkovich-Afifi's measure of skewness [36] under SEP multivariate distributions.

For the specific case of a tail weight parameter $\beta = 1$, the mixing variable is degenerate at $S = 1$ so that $E(S) = E(S^2) = E(S^3) = 1$; hence, the quantity in expression (3.4) becomes

$$\gamma_{1,p}^D = 2(4 - \pi)^2 \left\{ \frac{\boldsymbol{\eta}^\top \boldsymbol{\Omega} \boldsymbol{\eta}}{\pi + (\pi - 2)\boldsymbol{\eta}^\top \boldsymbol{\Omega} \boldsymbol{\eta}} \right\}^3. \tag{3.5}$$

This quantity corresponds to Mardia's and Malkovich-Afifi's measures of multivariate skewness under the multivariate SN distribution [8, 33, 11] as was expected to happen.

Figure 3 displays the heat color surfaces that depict the skewness measure $\gamma_{1,p}^{D}$ as a function of the non-normality pair (β, δ_*^2) with $0 < \beta \leq 1$ and $\delta_*^2 = \dfrac{\lambda_*^2}{1+\lambda_*^2}$ ranging in the interval $[0, 1)$. We can see that, due to the difficulties for computing large gamma functions in the expectations of the mixing variable, some blanks appear in the plots specially for the higher dimensions. Overall, it can be observed that the skewness decreases as the dimension increases. In addition, the curves on the skewness surface that trace the lines with equal tail weight show a dominance relation which means that, as the tail weight parameter β decreases, the corresponding skewness curve appear each one above the preceding one; this dominance relationship is illustrated by the plots displayed in Figure 4.

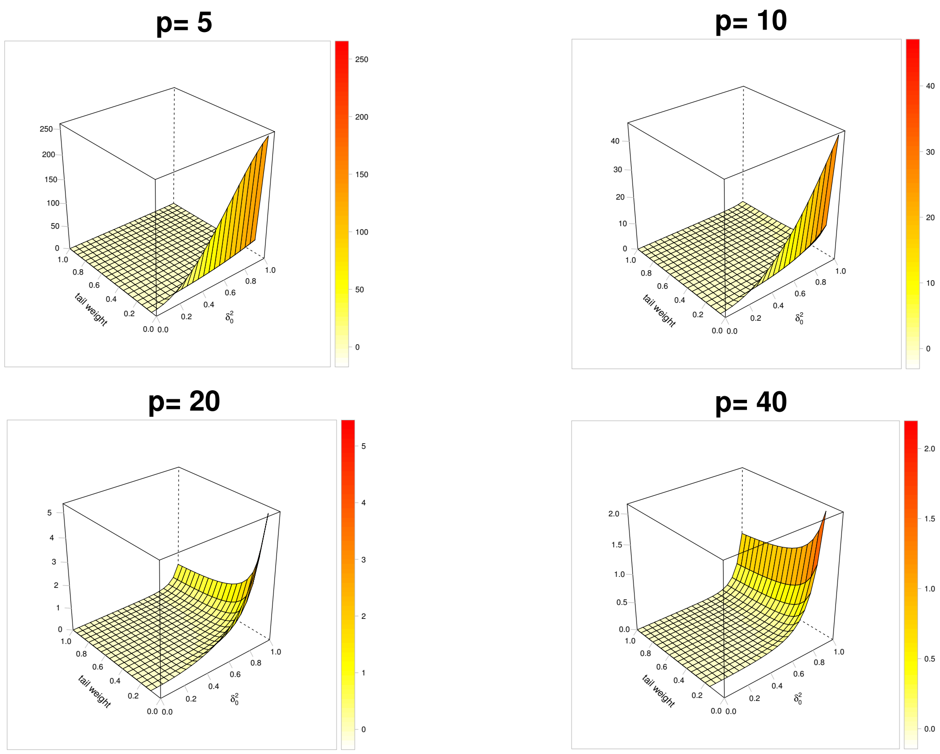


Figure 3. Skewness surfaces with respect to the quantities $\delta_*^2 = \dfrac{\lambda_*^2}{1+\lambda_*^2}$ and the tail weight parameter β for dimensions $p = 5$ (top left), $p = 10$ (top right), $p = 20$ (bottom left) and $p = 40$ (bottom right).

4. Concluding remarks

In this paper we have examined the multivariate skew exponential power distribution and have studied some of its theoretical properties in order to delve into the role played by the parameters of the model to handle the multivariate non-normality. It has been assumed that the tail weight parameter meets the condition $0 < \beta \leq 1$ so that the skew exponential power vector admits a SMSN stochastic representation. The SMSN formulation is useful to show that the tail weight parameter accounts for multivariate kurtosis regardless of the shape asymmetry vector, whereas

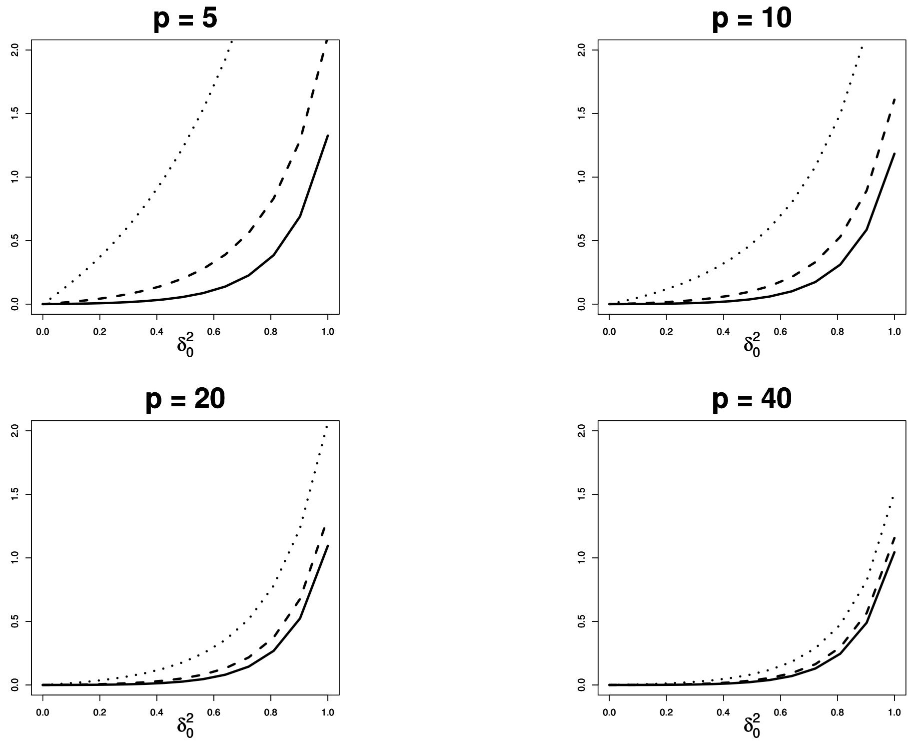


FIGURE 4. Skewness curves for the tail weight parameters: $\beta = 3/4$ (solid line), $\beta = 1/2$ (dashed line), $\beta = 1/4$ (dotted line) and dimensions $p = 5$ (top left), $p = 10$ (top right), $p = 20$ (bottom left) $p = 40$ (bottom right).

the shape vector regulates the multivariate asymmetry in a directional fashion; specifically, it is shown that such a shape vector lies on the direction yielding the maximal skewness projection regardless of the value of the tail weight parameter. When the tail weight equals to $\beta = 1$, the underlying model reduces to the SN distribution and our result on directional skewness agrees with previous work for SN vectors [33]. Summing up, we advocate the unique and singular role of the tail weight and the shape vector of multivariate skew exponential power model to account independently for different facets of the non-normality.

The theoretical results from Theorems 3.1 and 3.2 have implications for research in the statistical practice including but not limited to: skewness-based projection pursuit, the modeling of non-normal asset returns, model-based clustering and the role played by the SEP parameters as well as the sensitivity of gaussian bayesian networks to deviations from normality along the lines of previous work [35]. Moreover, an explicit expression of the pdf of the SEP model would enable inferential work; research in this direction may benefit from the approach followed in Section 4.1 of [10].

Our findings also highlight some other problems for future research: A major advantage of the exponential power distribution is that it can model both light and heavy tails; as the Kotz-type distribution has the same characteristic and additionally it can also model holes [29], we wonder whether the properties studied for SEP distributions would also be valid for skewed generalizations of Kotz-type distributions. Moreover, the consideration of other SEP formulations, like those ones introduced in [2, 12], may deserve further investigation on kurtosis assessment and directional skewness for the case $\beta > 1$. On the other hand, the skew-normal Cauchy model can be represented as a shape-scale mixture of SN distributions whose properties, including third moments and cumulants, have been investigated by [26] and are closely related to those ones of the SEP model; the study of directional asymmetry under shape-scale mixtures of SN vectors is also an open problem. Another issue is concerned with the consideration of other measures of multivariate skewness that depend on third moments [34].

As an afterthought, the results derived in this work also point out further research regarding the stochastic comparison of SMSN vectors: the SMSN representation of the multivariate skew exponential power model suggests looking into the role played by the non-normality parameters of the SMSN class by means of an ad hoc convex transform stochastic ordering; research effort along this line would complement previous related work on the issue [44, 3, 5].

Appendix: proofs of auxiliary lemmas

Proof of Lemma 2.1. The proof relies on a property about quadratic forms of SN vectors. Taking into account [11, Corollary 5.1] or the property appearing in expression (5.7) of [11, Section 5.7], we can assert that the quadratic form $\boldsymbol{Q} = (\boldsymbol{X}-\boldsymbol{\xi})^\top \boldsymbol{\Omega}^{-1}(\boldsymbol{X}-\boldsymbol{\xi}) = S^2\boldsymbol{Y}^\top\boldsymbol{\Omega}^{-1}\boldsymbol{Y}$ —with $\boldsymbol{Y}$ a vector such that $\boldsymbol{Y} \sim SN_p(\boldsymbol{0}_p, \boldsymbol{\Omega}, \boldsymbol{\alpha})$— has the same distribution as $S^2\boldsymbol{Y}_0^\top\boldsymbol{\Omega}^{-1}\boldsymbol{Y}_0$ with $\boldsymbol{Y}_0$ a normal vector such that $\boldsymbol{Y}_0 \sim N_p(\boldsymbol{0}_p, \boldsymbol{\Omega})$. Hence, $\boldsymbol{Q} \stackrel{d}{=} S^2\boldsymbol{Y}_0^\top\boldsymbol{\Omega}^{-1}\boldsymbol{Y}_0 \stackrel{d}{=} S^2G^2$, where G^2 is a scalar variable, independent of the mixing variable S, following a χ^2_p distribution. On the other hand, we also know that the quadratic form $\boldsymbol{Q}$ has the same distribution as the squared of the modular variable R associated to the p-dimensional elliptically contoured vector with distribution $EP_p(\boldsymbol{\xi}, \boldsymbol{\Omega}, \beta)$ [22]. The density function of R has been derived by [21]; it is given by

$$h_R(r) = \frac{p}{\Gamma\left(1+\frac{p}{2\beta}\right)2^{\frac{p}{2\beta}}} r^{p-1}\exp\left\{-\frac{1}{2}r^{2\beta}\right\} I_{(o,\infty)}(r).$$

Since the scalar variables R^2 and $\boldsymbol{Q}$ have the same distribution we can put: $E(R^k) = E(S^k)E(G^k)$, from which we get $E(S^k) = \dfrac{E(R^k)}{E((G^2)^{k/2})}$. The expectations appearing in $E(S^k)$ can be calculated in an easy way due to the tractability of the density functions of R and G^2; simple integral calculus gives

$$E(R^k) = \frac{2^{k/2\beta}\Gamma\left(\frac{p+k}{2\beta}\right)}{\Gamma\left(\frac{p}{2\beta}\right)} \text{ and } E(G^k) = E((G^2)^{k/2}) = \frac{2^{k/2}\Gamma\left(\frac{p+k}{2}\right)}{\Gamma\left(\frac{p}{2}\right)}$$

which in turn leads to the expression of the statement. □

Proof of Lemma 3.1. The result follows from the next property: if f and g are two increasing functions then it holds that $Cov(f(X), g(X)) \geq 0$ provided that the second order moments and cross moments of $f(X)$ and $g(X)$ exist. Since S is a non negative random variable the statement of the lemma will follow putting $f(S) = S^2$ and $g(S) = S$. □

Proof of Lemma 3.2. Using the general expression from Lemma 2.1, we get

$$\frac{4}{\pi}E(S)E(S^3) - E(S^2)^2 = \frac{2^{2/\beta}\Gamma^2(\frac{p}{2})}{2^2\Gamma^2(\frac{p}{2\beta})}\left[\frac{4}{\pi}\frac{\Gamma(\frac{p+1}{2\beta})\Gamma(\frac{p+3}{2\beta})}{\Gamma(\frac{p+1}{2})\Gamma(\frac{p+3}{2})} - \frac{\Gamma^2(\frac{p+2}{2\beta})}{\Gamma^2(\frac{p+2}{2})}\right]$$

$$= \frac{2^{2/\beta}}{p^2} \frac{\Gamma^2(\frac{p+2}{2\beta})}{\Gamma^2(\frac{p}{2\beta})} \left[\frac{4}{\pi} \frac{2}{p+1} \frac{\Gamma^2(\frac{p+2}{2})}{\Gamma^2(\frac{p+1}{2})} \frac{\Gamma(\frac{p+1}{2\beta})\Gamma(\frac{p+3}{2\beta})}{\Gamma^2(\frac{p+2}{2\beta})} - 1 \right]. \tag{4.1}$$

We now define the function: $g_0(x) = g(x + \frac{\omega}{2}) - g(x)$ with $g(x) = x\psi(x) : x \geq 1$, $\omega \geq 1$ and ψ the digamma function. This instrumental function will also be used in the proofs of Lemmas 3.2 and 3.3.

From the series expansion of the digamma function given by

$$\psi(z) = -\gamma - \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+z} \right),$$

where γ is the Euler-Mascheroni constant, we obtain that $g''(x) = 2\psi'(x) + x\psi''(x) = 2\sum_{n=1}^{\infty} \frac{n}{(n+x)^3}$. Hence, g is a convex function which implies that $g'(x)$ and $g_0(x)$ are non decreasing functions, with the latter one being non decreasing because $g_0'(x) = g'(x + \omega/2) - g'(x)$. Therefore,

$$g_0\left(\frac{p}{2}\omega + \omega\right) - g_0\left(\frac{p}{2}\omega + \frac{\omega}{2}\right) = \omega \left[\frac{p+1}{2} \psi\left(\frac{p+1}{2}\omega\right) - (p+2)\psi\left(\frac{p+2}{2}\omega\right) \right.$$

$$\left. + \frac{p+3}{2}\psi\left(\frac{p+3}{2}\omega\right) \right] = \omega(\log f(\omega))' > 0,$$

with f the function defined by $f(\omega) = \frac{\Gamma\left(\frac{p+1}{2}\omega\right)\Gamma\left(\frac{p+3}{2}\omega\right)}{\Gamma^2\left(\frac{p+2}{2}\omega\right)} : \omega \geq 1$ and $(\log f(\omega))'$ its first log derivative. The previous expression shows that f is non decreasing; so putting $\omega = \frac{1}{\beta}$ we get

$$f(\frac{1}{\beta}) = \frac{\Gamma\left(\frac{p+1}{2\beta}\right)\Gamma\left(\frac{p+3}{2\beta}\right)}{\Gamma^2\left(\frac{p+2}{2\beta}\right)} \geq f(1) = \frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(\frac{p+3}{2}\right)}{\Gamma^2\left(\frac{p+2}{2}\right)} = \frac{p+1}{2} \frac{\Gamma^2\left(\frac{p+1}{2}\right)}{\Gamma^2\left(\frac{p+2}{2}\right)},$$

which implies that the quantity (4.1) is non negative, as we aimed to prove. □

Proof of Lemma 3.3. From the general expression of the moments of the mixing variable in Lemma 2.1, we obtain after simple calculations that $E(S^3) + \frac{4}{\pi}E(S)^3 - 2E(S)E(S^2) = 2E(S)^3 \left[\frac{2}{\pi} + f(\omega)\right]$, where $\omega = \frac{1}{\beta}$ and f is a function defined as follows:

$$f(\omega) = \frac{\Gamma^2\left(\frac{p+1}{2}\right)}{\Gamma^2\left(\frac{p}{2}\right)} f_1(\omega)\left[f_2(\omega) - 1\right] : \omega \geq 1$$

with f_1 and f_2 functions given by

$$f_1(\omega) = \frac{2\Gamma\left(\frac{p+2}{2}\omega\right)\Gamma\left(\frac{p}{2}\omega\right)}{p\Gamma^2\left(\frac{p+1}{2}\omega\right)} \text{ and } f_2(\omega) = \frac{p\Gamma\left(\frac{p+3}{2}\omega\right)\Gamma\left(\frac{p}{2}\omega\right)}{2(p+1)\Gamma\left(\frac{p+1}{2}\omega\right)\Gamma\left(\frac{p+2}{2}\omega\right)}.$$

Their first derivatives are

$$f_1'(\omega) = \frac{f_1(\omega)\left[g_0\left(\frac{p+1}{2}\omega\right) - g_0\left(\frac{p}{2}\omega\right)\right]}{\omega} \text{ and } f_2'(\omega) = \frac{f_2(\omega)\left[g_0\left(\frac{p+2}{2}\omega\right) - g_0\left(\frac{p}{2}\omega\right)\right]}{\omega},$$

with g_0 the same function as the one defined in the proof of the previous lemma. Using analogous arguments as before, we can conclude that f_1 and f_2 are non decreasing which implies that $f_1(\omega) \geq \dfrac{\Gamma^2\left(\frac{p}{2}\right)}{\Gamma^2\left(\frac{p+1}{2}\right)}$ and $f_2(\omega) \geq 1/2$. We now calculate the first derivative of f:

$$\begin{aligned} f'(\omega) = \frac{\Gamma^2\left(\frac{p+1}{2}\right)}{\Gamma^2\left(\frac{p}{2}\right)} f_1(\omega) \Bigg\{ & \left[\frac{p+2}{2}\psi\left(\frac{p+2}{2}\omega\right) + \frac{p}{2}\psi\left(\frac{p}{2}\omega\right)\right. \\ & \left. -(p+1)\psi\left(\frac{p+1}{2}\omega\right)\right][f_2(\omega)-1] + f_2(\omega)\left[\frac{p+3}{2}\psi\left(\frac{p+3}{2}\omega\right)\right. \\ & \left. +\frac{p}{2}\psi\left(\frac{p}{2}\omega\right) - \frac{p+1}{2}\psi\left(\frac{p+1}{2}\omega\right) - \frac{p+2}{2}\psi\left(\frac{p+2}{2}\omega\right)\right]\Bigg\}. \end{aligned}$$

Taking into account the lower bound for f_2, we get: $\dfrac{\Gamma^2\left(\frac{p}{2}\right)}{\Gamma^2\left(\frac{p+1}{2}\right)}\dfrac{f'(\omega)}{f_1(\omega)} \geq$

$$\begin{aligned} & -\frac{1}{2}\left[\frac{p+2}{2}\psi\left(\frac{p+2}{2}\omega\right) + \frac{p}{2}\psi\left(\frac{p}{2}\omega\right) - (p+1)\psi\left(\frac{p+1}{2}\omega\right)\right] + \frac{1}{2}\left[\frac{p+3}{2}\cdot\right. \\ & \left.\psi\left(\frac{p+3}{2}\omega\right) + \frac{p}{2}\psi\left(\frac{p}{2}\omega\right) - \frac{p+1}{2}\psi\left(\frac{p+1}{2}\omega\right) - \frac{p+2}{2}\psi\left(\frac{p+2}{2}\omega\right)\right] \\ & = \frac{1}{2}\left\{\frac{p+3}{2}\psi\left(\frac{p+3}{2}\omega\right) - (p+2)\psi\left(\frac{p+2}{2}\omega\right) + \frac{p+1}{2}\psi\left(\frac{p+1}{2}\omega\right)\right\} \\ & = \frac{1}{2\omega}\left\{g_0\left(\frac{p+2}{2}\omega\right) - g_0\left(\frac{p+1}{2}\omega\right)\right\} \geq 0. \end{aligned}$$

Therefore, we can assert that f is also a non decreasing function for $\omega \geq 1$; consequently $f(\omega) \geq f(1) = -1/2$ which implies that

$$E(S^3) + \frac{4}{\pi}E(S)^3 - 2E(S)E(S^2) = 2E(S)^3\left[\frac{2}{\pi} + f(\omega)\right] \geq 2E(S)^3\left[\frac{2}{\pi} - \frac{1}{2}\right] \geq 0$$

as we aimed to prove. □

Acknowledgement. The authors are grateful to the reviewers for their useful recommendations that contributed to improve the draft version of this paper.

References

[1] G. Agrò. Maximum likelihood estimation for the exponential power function parameters. *Communications in Statistics - Simulation and Computation*, 24(2):523–536, 1995.

[2] R. Arellano-Valle and Wolf-Dieter Richter. On skewed continuous $l_{n,p}$-symmetric distributions. *Chilean Journal of Statistics*, 3:193–212, 2012.

[3] Jorge M. Arevalillo and Hilario Navarro. A study of the effect of kurtosis on discriminant analysis under elliptical populations. *Journal of Multivariate Analysis*, 107:53 – 63, 2012.

[4] Jorge M. Arevalillo and Hilario Navarro. A note on the direction maximizing skewness in multivariate skew-t vectors. *Statistics & Probability Letters*, 96:328–332, 2015.

[5] Jorge M. Arevalillo and Hilario Navarro. A stochastic ordering based on the canonical transformation of skew-normal vectors. *TEST*, 28(2):475–498, 2019.

[6] A Azzalini. Further results on a class of distributions which includes the normal ones. *Statistica*, XLVI:199–208, 1986. Reprinted with annotations and corrigenda in vol. 80 (2020), pp. 161–175.

[7] A. Azzalini. *The R package* `sn`*: The Skew-Normal and Related Distributions such as the Skew-t and the SUN (version 2.0.0).* Università di Padova, Italia, 2021.

[8] A. Azzalini and A. Capitanio. Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society Series B*, 61(3):579–602, 1999.

[9] A. Azzalini and A. Dalla Valle. The multivariate skew-normal distribution. *Biometrika*, 83(4):715–726, 1996.

[10] Adelchi Azzalini and Antonella Capitanio. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society Series B*, 65(2):367–389, 2003.
[11] Adelchi Azzalini and Antonella Capitanio. *The Skew-Normal and Related Families*. IMS monographs. Cambridge University Press, 2014.
[12] Adelchi Azzalini and Giuliana Regoli. Some properties of skew-symmetric distributions. *Annals of the Institute of Statistical Mathematics*, 64(4):857–879, 2012.
[13] G. E. P. Box. A note on regions for tests of kurtosis. *Biometrika*, 40(3-4):465–468, 1953.
[14] G.E.P. Box and G.C. Tiao. *Bayesian Inference in Statistical Analysis*. Addison-Wesley Publishing Company, 1973.
[15] Márcia D. Branco and Dipak K. Dey. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, 79(1):99 – 113, 2001.
[16] Antonella Capitanio. On the canonical form of scale mixtures of skew-normal distributions. *arXiv/1207.0797*, 2012.
[17] Antonella Capitanio. On the canonical form of scale mixtures of skew-normal distributions. *Statistica*, 80(2):145–160, 2020.
[18] S. T. B. Choy and A. F. M. Smith. Hierarchical models with scale mixtures of normal distributions. *Test*, 6(1):205–221, 1997.
[19] Kai-Tai Fang, Samuel Kotz, and Kai Wang Ng. *Symmetric multivariate and related distributions*. Monographs on statistics and applied probability. Chapman & Hall, 1990.
[20] M.A. Gómez-Villegas, P. Main, H. Navarro, and R. Susi. Assessing the effect of kurtosis deviations from gaussianity on conditional distributions. *Applied Mathematics and Computation*, 219(21):10499–10505, 2013.
[21] E. Gómez, M.A. Gómez-Villegas, and J.M. Marín. A multivariate generalization of the power exponential family of distributions. *Communications in Statistics - Theory and Methods*, 27(3):589–600, 1998.
[22] E. Gómez-Sánchez-Manzano, M. A. Gómez-Villegas, and J. M. Marín. Sequences of elliptical distributions and mixtures of normal distributions. *Journal of Multivariate Analysis*, 97(2):295–310, 2006.
[23] E Gómez-Sánchez-Manzano, MA Gómez-Villegas, and JM Marín. Multivariate exponential power distributions as mixtures of normal distributions with bayesian applications. *Communications in Statistics - Theory and Methods*, 37(6):972–985, 2008.
[24] Irwin R Goodman and Samuel Kotz. Multivariate θ-generalized normal distributions. *Journal of Multivariate Analysis*, 3(2):204–219, 1973.
[25] Alon Hazan, Zinoviy Landsman, and Udi E Makov. Robustness via a mixture of exponential power distributions. *Computational Statistics & Data Analysis*, 42(1):111–121, 2003.
[26] F. Kahrari, M. Rezaei, F. Yousefzadeh, and R.B. Arellano-Valle. On the multivariate skew-normal-cauchy distribution. *Statistics & Probability Letters*, 117:80–88, 2016.
[27] Y. Kano. Consistency property of elliptic probability density functions. *Journal of Multivariate Analysis*, 51(1):139–147, 1994.
[28] Hyoung-Moon Kim and Chiwhan Kim. Moments of scale mixtures of skew-normal distributions and their quadratic forms. *Communications in Statistics - Theory and Methods*, 46(3):1117–1126, 2017.
[29] Tõnu Kollo and Anu Roos. *Contemporary Multivariate Analysis and Design of Experiments*, chapter On Kotz-Type Elliptical Distributions, pages 159–170. World Scientific.
[30] Victor Korolev. Some properties of univariate and multivariate exponential power distributions and related topics. *Mathematics*, 8(11), 2020.
[31] Samuel Kotz. Multivariate distributions at a cross road. In G. P. Patil, S. Kotz, and J. K. Ord, editors, *A Modern Course on Statistical Distributions in Scientific Work*, pages 247–270, Dordrecht, 1975. Springer Netherlands.
[32] Yoichi Kuwana and Takeaki Kariya. Lbi tests for multivariate normality in exponential power distributions. *Journal of Multivariate Analysis*, 39(1):117–134, 1991.
[33] Nicola Loperfido. Canonical transformations of skew-normal variates. *TEST*, 19(1):146–165, 2010.
[34] Nicola Loperfido. Singular value decomposition of the third multivariate moment. *Linear Algebra and its Applications*, 473:202–216, 2015.
[35] P. Main and H. Navarro. Analyzing the effect of introducing a kurtosis parameter in gaussian bayesian networks. *Reliability Engineering & System Safety*, 94(5):922–926, 2009.
[36] J. F. Malkovich and A. A. Afifi. On tests for multivariate normality. *Journal of the American Statistical Association*, 68(341):176–179, 1973.
[37] Raúl Alejandro Morán-Vásquez and Silvia L. P. Ferrari. Box–cox elliptical distributions with application. *Metrika*, 82(5):547–571, 2019.
[38] Raúl Alejandro Morán-Vásquez and Silvia L. P. Ferrari. New results on truncated elliptical distributions. *Communications in Mathematics and Statistics*, 9(3):299–313, 2021.
[39] Saralees Nadarajah. A generalized normal distribution. *Journal of Applied Statistics*, 32(7):685–694, 2005.

[40] Eusebio Gómez Sánchez-Manzano, Miguel Angel Gómez-Villegas, and Juan-Miguel Marín-Diazaraque. A matrix variate generalization of the power exponential family of distributions. *Communications in Statistics - Theory and Methods*, 31(12):2167–2182, 2002.
[41] S. Simoni. Su una estensione dello schema delle curve normali di ordina r alle variabili doppie. *Statistica*, 28:151–170, 1968.
[42] M.T Subbotin. On the law of frequency of error. *Matematicheskii Sbornik*, 31(2):296–301, 1923.
[43] W R Van Zwet. *Convex Transformations of Random Variables*. Mathematish Centrum, Amsterdam, 1964.
[44] Jin Wang. A family of kurtosis orderings for multivariate distributions. *Journal of Multivariate Analysis*, 100(3):509–517, 2009.
[45] Min Wang and Tao Lu. A matching prior for the shape parameter of the exponential power distribution. *Statistics & Probability Letters*, 97:150–154, 2015.
[46] Mike West. On scale mixtures of normal distributions. *Biometrika*, 74(3):646–648, 1987.

* University Nacional Educación a Distancia (UNED)
Department of Statistics, Operational Research and Numerical Analysis
Juan del Rosal 10
Madrid, 28040
SPAIN
Email address: jmartin@ccia.uned.es

** University Nacional Educación a Distancia (UNED)
Department of Statistics, Operational Research and Numerical Analysis
Juan del Rosal 10
Madrid, 28040
SPAIN
Email address: hnavarro@ccia.uned.es