# A skew-symmetric representation of multivariate distributions

Jiuzhou Wang, Joseph Boyer, and Marc G. Genton*

January 15, 2004

**Abstract**

This article studies a family of multivariate skew-symmetric distributions. We show that any multivariate probability density function admits a skew-symmetric representation. We derive several characteristics of this representation and establish an invariance property. We present a stochastic representation of skew-symmetric distributions which lends itself readily to simulations. The flexibility of skew-symmetric distributions is illustrated through several graphical examples.

*Some key words*: Elliptical; Kurtosis; Multimodality; Quadratic form; Skewness; Stochastic representation

## 1 INTRODUCTION

During the last decade, there has been a growing interest in the construction of flexible parametric classes of multivariate distributions that exhibit skewness and kurtosis which is different from the normal distribution. The motivation originates from data sets, including environmental, financial, and biomedical ones, which often do not follow the normal law. In order to model departures from normality, a popular approach consists of modifying the probability density function (pdf) of a random vector in a multiplicative fashion. Although this idea has been in the literature for a long time, it is Azzalini (1985, 1986) who thoroughly set the foundations for the univariate normal distribution, yielding the so called skew-normal distribution. An extension to the multivariate setting was then proposed by Azzalini and Dalla Valle (1996). They defined

*Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203, USA. E-mail: {jwang3, jgboyer, genton}@stat.ncsu.edu

an $n$-dimensional random vector $\mathbf{X}$ as having a multivariate skew-normal distribution, denoted by $SN_n(\boldsymbol{\mu}, \Omega, \boldsymbol{\alpha})$, if it is continuous with pdf

$$2\phi_n(\mathbf{x}; \boldsymbol{\mu}, \Omega)\Phi(\boldsymbol{\alpha}^T(\mathbf{x} - \boldsymbol{\mu})), \qquad \mathbf{x} \in \mathbb{R}^n, \tag{1}$$

where $\phi_n(\mathbf{x}; \boldsymbol{\mu}, \Omega)$ is the $n$-dimensional normal pdf with mean $\boldsymbol{\mu}$ and correlation matrix $\Omega$, $\Phi(\cdot)$ is the standard normal cdf $N(0,1)$, and $\boldsymbol{\alpha}$ is an $n$-dimensional shape parameter. When $\boldsymbol{\alpha} = \mathbf{0}$, the pdf (1) reduces to the one of the multivariate normal distribution $N_n(\boldsymbol{\mu}, \Omega)$. Statistical applications of the multivariate skew-normal distribution were emphasized by Azzalini and Capitanio (1999).

Generalizations of these ideas to other distributions have been proposed by many authors. For instance, multivariate distributions such as skew-Cauchy (Arnold and Beaver (2000)), skew-$t$ (Branco and Dey (2001); Azzalini and Capitanio (2003); Sahu, Dey and Branco (2003)), skew-logistic (Wahed and Ali (2001)), and other skew-elliptical ones (Azzalini and Capitanio (1999); Branco and Dey (2001); Sahu *et al.* (2003)) were defined in the literature. Domínguez-Molina, González-Farías and Gupta (2001) also introduced a general skew-normal distribution by replacing the univariate cdf $\Phi$ in (1) by a multivariate one, thus allowing for closure under conditioning and addition. Arnold and Beaver (2002) discuss the construction of skew-elliptical models related to hidden truncations and selective reporting. Arellano-Valle, del Pino and San Martin (2002) present yet another view on skewed distributions, based on conditioning and the sign function. Recently, Genton and Loperfido (2002) defined a class of generalized skew-elliptical (GSE) distributions by pdfs of the form

$$2f(\mathbf{x})\pi(\mathbf{x}), \tag{2}$$

where $f$ is an elliptical pdf and $\pi$ is a *skewing function*, i.e. it satisfies $0 \leq \pi(\mathbf{x}) \leq 1$ and $\pi(-\mathbf{x}) = 1 - \pi(\mathbf{x})$. It is not difficult to see that (2) encompasses many skew-elliptical distributions defined in the papers above with appropriate choices of $f$ and $\pi$. Genton and Loperfido (2002) showed that the distribution of any even function, in particular quadratic forms, in GSE random vectors does not depend on the skewing function $\pi$. Similar results have been derived by Azzalini and Capitanio (2003) in the context of distributions generated by perturbation of symmetry. This has important implications for statistical inference based on quadratic forms as noted by Genton, He and Liu (2001) and Loperfido (2001). Indeed, the sample autocovariance function in time series, the sample variogram in spatial statistics, and the Mahalanobis distance in multivariate analysis, are all based on quadratic forms of the data. We will see that this invariance property still holds for the very general class of multivariate skew-symmetric distributions described in this article.

This paper is set up as follows. Section 2 describes the family of skew-symmetric distributions, introduces its use to represent multivariate distributions, and presents a stochastic representation suitable for simulation. Section 3 studies its important characteristics, and gives visual examples. Section 4 presents a discussion. All proofs of propositions are provided in the Appendix.

## 2 SKEW-SYMMETRIC DISTRIBUTIONS

The class of skew-symmetric distributions and its construction are defined by means of the following proposition.

**Proposition 1 (Skew-Symmetric Construction)** *Consider a function from $\mathbb{R}^n \to \mathbb{R}_+$ of the form*

$$2f(\mathbf{x}-\boldsymbol{\xi})\pi(\mathbf{x}-\boldsymbol{\xi}), \tag{3}$$

*where the pdf $f$ is symmetric around $\mathbf{0}$, i.e. $f(-\mathbf{x}) = f(\mathbf{x})$, $\pi : \mathbb{R}^n \to [0,1]$ is a skewing function, and $\boldsymbol{\xi}$ is any point in $\mathbb{R}^n$. Then (3) is a pdf.*

Intuitively, the skewing function in (3) merely reallocates the density between a point and its polar opposite. We call (3) a *skew-symmetric distribution* with respect to $\boldsymbol{\xi}$ with *symmetric component* $f$ and *skewed component* $\pi$. Azzalini and Capitanio (2003) independently arrived at a pdf of the form

$$2f(\mathbf{x}-\boldsymbol{\xi})G(w(\mathbf{x}-\boldsymbol{\xi})), \tag{4}$$

where the continuous pdf $f$ is symmetric around $\mathbf{0}$, $G : \mathbb{R} \to [0,1]$ is the cdf of a continuous random variable that is symmetric around 0, and $w : \mathbb{R}^n \to \mathbb{R}$ is an odd function. It turns out that (4) describes the same class as (3).

**Proposition 2 (Equivalence of Representations)** *The class of skew-symmetric distributions described by (3) is the same as that described by (4).*

Note that the representation of a skewing function $\pi(\cdot)$ in the form $G(w(\cdot))$ is not unique. A suitable odd function $w$ can be found for any strictly increasing $G$, as can be seen from the proof of Proposition 2 in the Appendix.

The generalized skew-elliptical (GSE) distribution presented in Genton and Loperfido (2002) is the same as (3) except that $f$ is required to be elliptical as well as symmetric. The motivation for (3) is that the skew-symmetric distribution is a generalization of the GSE distribution, which in turn is a generalization of skew-normal (SN) and related skew-elliptical (SE) distributions. The fact that GSE's, SE's, and SN's are pdfs follows directly from Proposition 1. As is shown next, the class of skew-symmetric distributions is completely general, since any pdf has a skew-symmetric representation.

**Proposition 3 (Existence and Uniqueness of Skew-Symmetric Representation)** *Let $g : \mathbb{R}^n \to \mathbb{R}_+$ be a pdf and $\boldsymbol{\xi}$ be any point in $\mathbb{R}^n$. Then*

$$g(\mathbf{x}) = 2 f_{\boldsymbol{\xi}}(\mathbf{x} - \boldsymbol{\xi}) \pi_{\boldsymbol{\xi}}(\mathbf{x} - \boldsymbol{\xi}), \tag{5}$$

*where $f_{\boldsymbol{\xi}}$ is a pdf, symmetric around $\mathbf{0}$, and $\pi_{\boldsymbol{\xi}}$ is a skewing function. This representation is unique for any $\boldsymbol{\xi}$, and*

$$f_{\boldsymbol{\xi}}(\mathbf{s}) = \frac{g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s})}{2} \tag{6}$$

$$\pi_{\boldsymbol{\xi}}(\mathbf{s}) = \frac{g(\boldsymbol{\xi} + \mathbf{s})}{g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s})}. \tag{7}$$

We call (5) the *skew-symmetric representation* of the pdf $g$ with respect to $\boldsymbol{\xi}$.

It turns out that (3) has a convenient stochastic representation. Let $\mathbf{Y}$ be a continuous random vector with pdf $f(\mathbf{y})$. Let $U$ be a uniform random variable on $(0,1)$, independent of $\mathbf{Y}$. A random vector $\mathbf{X}$ with pdf (3) can be simulated via the following representation:

$$\mathbf{X} = \begin{cases} \mathbf{Y} + \boldsymbol{\xi} & \text{if } U < \pi(\mathbf{Y}), \\ -\mathbf{Y} + \boldsymbol{\xi} & \text{if } U > \pi(\mathbf{Y}). \end{cases} \tag{8}$$

Azzalini and Capitanio (2003) present a slightly more complex version of the above approach.

As an example, let $\mathbf{X}$ have a skew-normal pdf of the form (1) and $\mathbf{Y}$ have the pdf $\phi_n(\mathbf{x}; \mathbf{0}, \Omega)$. Applying the probability integral transformation to (8), we arrive at

$$\mathbf{X} = \begin{cases} \mathbf{Y} + \boldsymbol{\mu} & \text{if } W < \boldsymbol{\alpha}^T \mathbf{Y}, \\ -\mathbf{Y} + \boldsymbol{\mu} & \text{if } W > \boldsymbol{\alpha}^T \mathbf{Y}. \end{cases}$$

where $W$ is $N(0,1)$, independent of $\mathbf{Y}$. This is a familiar procedure, provided by Azzalini and Dalla Valle (1996), for simulating the skew-normal distribution.

# 3 PROPERTIES OF THE SKEW-SYMMETRIC REPRESENTATION

From (6) and (7), we can always write

$$f_{\boldsymbol{\xi}}(\mathbf{0}) = g(\boldsymbol{\xi}), \quad \pi_{\boldsymbol{\xi}}(\mathbf{0}) = \frac{1}{2}. \tag{9}$$

We can also make conclusions about continuity and differentiability of $f_{\boldsymbol{\xi}}$ and $\pi_{\boldsymbol{\xi}}$ from the continuity and the gradient $\nabla g$ of $g$.

**Lemma 1 (Continuity and Differentiability)**

1. *Assume $g$ is continuous at $\boldsymbol{\xi}+\mathbf{s}$ and $\boldsymbol{\xi}-\mathbf{s}$. Then, $f_{\boldsymbol{\xi}}(\mathbf{s})$ is continuous in both $\boldsymbol{\xi}$ and $\mathbf{s}$ at the point $(\boldsymbol{\xi}, \mathbf{s})$. If at least one of $g(\boldsymbol{\xi}+\mathbf{s})$ or $g(\boldsymbol{\xi}-\mathbf{s})$ is not zero, then $\pi_{\boldsymbol{\xi}}(\mathbf{s})$ is also continuous in both $\mathbf{s}$ and $\boldsymbol{\xi}$ at the point $(\boldsymbol{\xi}, \mathbf{s})$.*

2. *Assume $g$ is differentiable at $\boldsymbol{\xi}+\mathbf{s}$ and $\boldsymbol{\xi}-\mathbf{s}$. Then $\nabla f_{\boldsymbol{\xi}}(\mathbf{s}) = (\nabla g(\boldsymbol{\xi}+\mathbf{s}) - \nabla g(\boldsymbol{\xi}-\mathbf{s}))/2$. If at least one of $g(\boldsymbol{\xi}+\mathbf{s})$ or $g(\boldsymbol{\xi}-\mathbf{s})$ is not zero, then*

$$\nabla \pi_{\boldsymbol{\xi}}(\mathbf{s}) = \frac{[g(\boldsymbol{\xi}+\mathbf{s}) + g(\boldsymbol{\xi}-\mathbf{s})]\nabla g(\boldsymbol{\xi}+\mathbf{s}) - 2g(\boldsymbol{\xi}+\mathbf{s})\nabla f(\mathbf{s})}{[g(\boldsymbol{\xi}+\mathbf{s}) + g(\boldsymbol{\xi}-\mathbf{s})]^2}. \tag{10}$$

3. *If $g$ is differentiable at $\boldsymbol{\xi}$, then $\nabla f_{\boldsymbol{\xi}}(\mathbf{0}) = \mathbf{0}$.*

4. *If $g$ is twice differentiable at $\boldsymbol{\xi}$, then the Hessian of $f_{\boldsymbol{\xi}}$ at $\mathbf{0}$ equals the Hessian of $g$ at $\boldsymbol{\xi}$.*

We are now able to discuss the modality of $g$ and of the symmetric component $f_{\boldsymbol{\xi}}$.

**Proposition 4 (Modality and Concavity of Symmetric Component)**

1. *If $g$ has a local (global) maximum (minimum) at $\mathbf{m}$, then $f_{\mathbf{m}}$ has a local (global) maximum (minimum) at $\mathbf{0}$. If $g$ is locally (globally) monotonic along every ray from $\mathbf{m}$, $f_{\mathbf{m}}$ is locally (globally) monotonic along every ray from $\mathbf{0}$.*

2. *If $g$ is concave (convex) on an open convex set $N$, then for all $\boldsymbol{\xi} \in N$, $f_{\boldsymbol{\xi}}$ is concave (convex) on $S_{\boldsymbol{\xi}}$ and is maximized (minimized) on $S_{\boldsymbol{\xi}}$ at $\mathbf{0}$, where $S_{\boldsymbol{\xi}} = \{\mathbf{s} | \boldsymbol{\xi}+\mathbf{s} \in N$ and $\boldsymbol{\xi}-\mathbf{s} \in N\}$.*

3. *Conversely, if $f_{\boldsymbol{\xi}}$ has a local maximum (minimum) at $\mathbf{0}$ $\forall$ $\boldsymbol{\xi}$ in an open convex set $N$, then $g$ is concave (convex) on $N$.*

Consider the case of a pdf $g$ which is concave in a neighborhood of the mode $\mathbf{m}$ and convex outside that neighborhood. Then $f_{\boldsymbol{\xi}}$ will be (at least in a neighborhood of $\mathbf{0}$) unimodal for $\boldsymbol{\xi}$ inside the inflection points of $g$. As we move $\boldsymbol{\xi}$ farther from $\mathbf{m}$, the mode of $f_{\boldsymbol{\xi}}$ at $\mathbf{0}$ will become less and less distinct. Finally, at an inflection point of $g$, the mode of $f_{\boldsymbol{\xi}}$ at $\mathbf{0}$ will disappear entirely. If we continue to move $\boldsymbol{\xi}$ farther away from $\mathbf{m}$, $f_{\boldsymbol{\xi}}$ will exhibit multimodality. For instance, take $g(x) = \phi(x)$, the univariate standard normal pdf. Its mode is $m = 0$ and it has inflection points at $\pm 1$. The first column of Figure 1 illustrates the shape of $f_\xi$ in the skew-symmetric representation (5) for $\xi = 0, -0.5, -1, -1.5$.

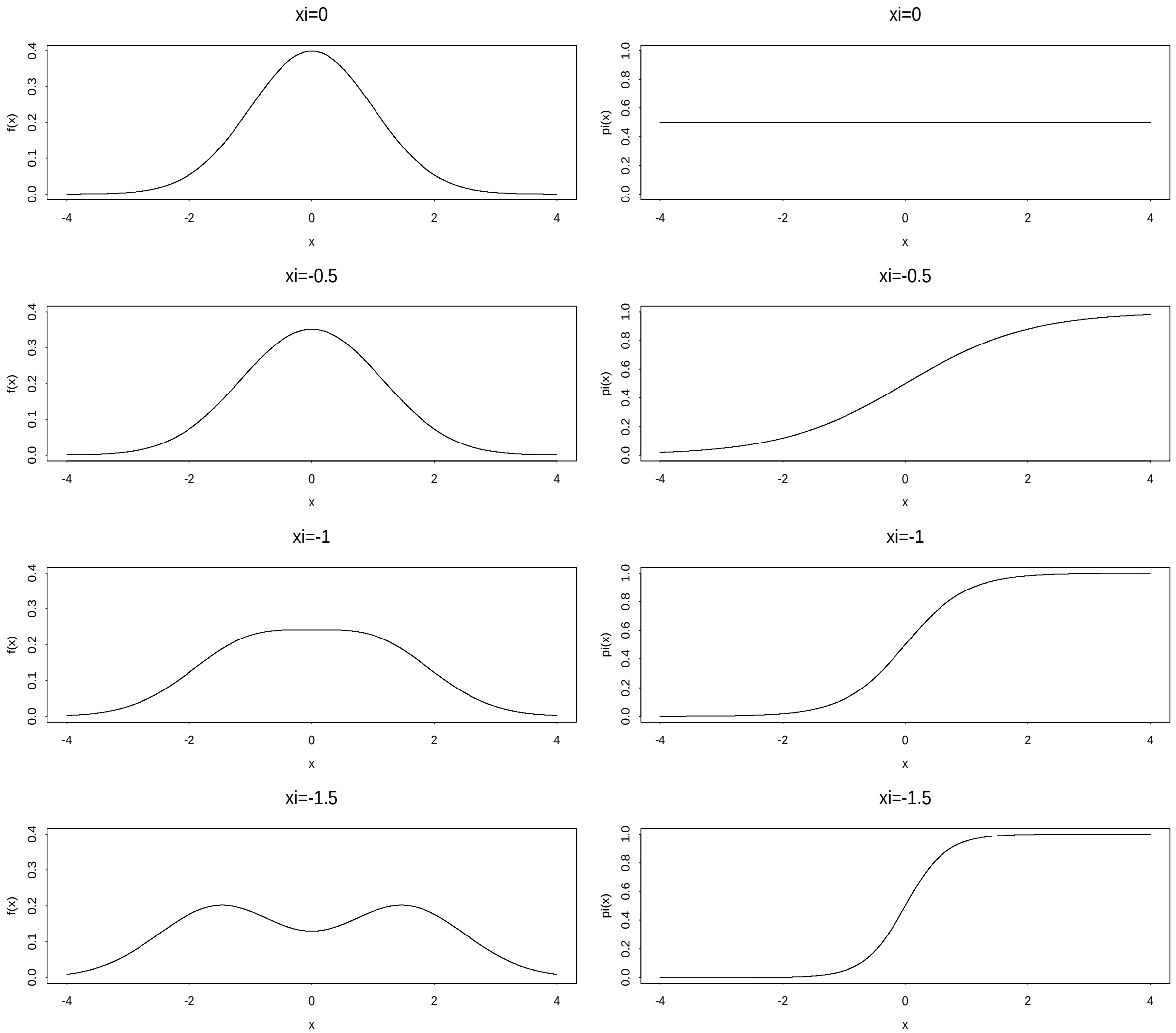


Figure 1: Graphical illustration of the skew-symmetric representation (5) of the univariate standard normal pdf for $\xi = 0, -0.5, -1, -1.5$: symmetric component (column 1), skewed component (column 2).

The following proposition concerns the skewed component.

**Proposition 5 (Monotonicity of the Skewed Component)**

1. $\pi_{\boldsymbol{\xi}}(\mathbf{s}) \gtreqless \pi_{\boldsymbol{\xi}}(-\mathbf{s}) \iff \frac{g(\boldsymbol{\xi}+\mathbf{s})}{g(\boldsymbol{\xi}-\mathbf{s})} \gtreqless 1$.

2. *$\pi_{\boldsymbol{\xi}}(\mathbf{s})$ has the same monotonicity as $\frac{g(\boldsymbol{\xi}+\mathbf{s})}{g(\boldsymbol{\xi}-\mathbf{s})}$.*

3. *If $g$ is differentiable at $\boldsymbol{\xi}$ and $g(\boldsymbol{\xi}) \neq 0$, then $\nabla \pi_{\boldsymbol{\xi}}(\mathbf{0}) = 2\frac{\nabla g(\boldsymbol{\xi})}{g(\boldsymbol{\xi})}$.*

Part 3 of Proposition 5 implies that the skewed component with respect to the mode will be flat at $\mathbf{0}$. Consider, as a special case, a differentiable univariate distribution $g$ on a neighborhood on which it is unimodal. At a point $\xi$ to the left of the local mode $m$, the third result of Proposition 5 tells us that $\pi_\xi$ is increasing in a neighborhood of 0. Then, using (9), we know that in a region to the immediate right of 0, $\pi_\xi(s) > \frac{1}{2}$, and that to the immediate left of 0, $\pi_\xi(s) < \frac{1}{2}$. At $\xi = m$, $\pi_\xi$ is flat at 0 and thus close to $\frac{1}{2}$ in a neighborhood of 0. Whether and to what extent $\pi_\xi$ deviates from $\frac{1}{2}$ as $s$ moves away from 0 depends on the degree of asymmetry in $g$, embodied in $\frac{g(\xi+s)}{g(\xi-s)}$. If $g$ is symmetric around $\xi$, then $\pi_\xi$ is simply the constant function $\frac{1}{2}$. The behavior of $\pi_\xi(s)$ at values of $s$ far from 0 depends on the tail behavior of $\frac{g(\xi+s)}{g(\xi-s)}$, which depends upon both $g$ and $\xi$. Typically, $\lim_{s\to\infty} \pi_\xi(s)$ equals 1, 0, or $\frac{1}{2}$. The second column of Figure 1 illustrates the shape of $\pi_\xi$ in the skew-symmetric representation (5) of $g(x) = \phi(x)$ for $\xi = 0, -0.5, -1, -1.5$. It is more difficult to analyze the behavior of $f_{\boldsymbol{\xi}}$ and $\pi_{\boldsymbol{\xi}}$ if $g$ is a multivariate distribution. However, one can get around the problem of multiple dimensions by characterizing $f_{\boldsymbol{\xi}}$ and $\pi_{\boldsymbol{\xi}}$ along a particular line through the origin.

The distributional invariance property of even functions of GSE random vectors (Genton and Loperfido (2002)) is a powerful tool for evaluating the distributions of quadratic forms, which are even functions. It generalizes the chi-square properties of univariate skew-normal distributions (Azzalini (1985)) and multivariate skew-normal distributions (Azzalini and Dalla Valle (1996)), see also Wang, Boyer and Genton (2004). The skew-symmetric representation allows us to extend an even function distributional invariance result to all random vectors.

**Proposition 6 (Distributional Invariance of Even Functions)** *If a random vector $\mathbf{X}$ has a skew-symmetric representation of the form (5), then the distribution of $\tau(\mathbf{X}-\boldsymbol{\xi})$, where $\tau$ is an even function, does not depend on the skewing function $\pi_{\boldsymbol{\xi}}$.*

Because of our Proposition 2, Proposition 6 is equivalent to Proposition 2 in Azzalini and Capitanio (2003). We give however an alternative proof, see the Appendix. Normal distribution theory can be used in conjunction with Proposition 6 to derive the distribution of quadratic forms of random vectors whose symmetric component is a normal pdf. For instance, if a random vector $\mathbf{X}$ has a skew-symmetric representation of the form (5), $f_{\boldsymbol{\xi}}$ is a multivariate normal pdf with nonsingular covariance matrix $\Omega$ of rank $p$, then $(\mathbf{X}-\boldsymbol{\xi})^T\Omega^{-1}(\mathbf{X}-\boldsymbol{\xi}) \sim \chi^2_p$.

Next, we provide several graphical examples of skew-symmetric distributions. Figure 2 displays the skew-symmetric representation of five different univariate pdfs, each with respect to a certain value of $\xi$. The first row displays the representation for a skew-normal $SN_1(0,1,4)$ distribution with $\xi = 0$. In this case, $f_\xi$ is simply the standard normal distribution, and $\pi_\xi$ is simply $\Phi(4x)$, as can be seen in (1).

Row 2 shows the representation of the $\chi^2_3$ pdf with respect to its mode $\xi = 1$. In cases where $\xi + s$ or $\xi - s$ is outside the support of the original density, $\pi_\xi(s)$ will be either zero or one. Consistent with Propositions 4 and 5, $f_\xi$ is maximized at 0 and $\pi$ is flat at zero.

Row 3 displays the representation with respect to $\xi = -0.25$ of the double exponential pdf with location parameter 0 and scale parameter 1; without concavity in the neighborhood of the mode, there is no guarantee of a neighborhood for $\xi$ around the mode for which $f_\xi$ is locally concave at zero.

In the fourth row the Cauchy$(0,1)$ pdf is represented with respect to $\xi = -4$, illustrating part 2 of Proposition 4. If $\xi$ is far enough out in the tails, $f_\xi$ will have a local minimum at 0, and thus of necessity be bimodal, even if $g$ is unimodal. In such cases, $\pi_\xi$ will often be a cdf which rises steeply in the neighborhood of 0. However, if $g$ has thick enough tails, $\pi_\xi$ may approximate a cdf around 0 but asymptote to limits other than 0 and 1. With the Cauchy, these limits are $\frac{1}{2}$, $\frac{1}{2}$ no matter what the value of $\xi$.

Row 5 displays the extreme value distribution $g(x) = \exp(-x)\exp(-\exp(-x)))$ with $\xi = -0.5$. This representation looks similar to that of the skew-normal, illustrating Propositions 4 and 5, which indicate the existence of a region to the left of the mode such that for all $\xi$ in the region, $f_\xi$ is unimodal near 0 and $\pi_\xi$ is increasing in the neighborhood of 0.

Which $\boldsymbol{\xi}$ should be used for a skew-symmetric representation of a pdf $g$? A natural choice is to take the mean, the median, or the mode of $g$. For each of them, there is a unique skew-symmetric representation (5) as shown in Proposition 3. However, many other choices of $\boldsymbol{\xi}$ are possible and each of them will imply, along with the characteristics of $g$, different shapes and properties of the symmetric and skewed component as shown in Propositions 4 and 5. For

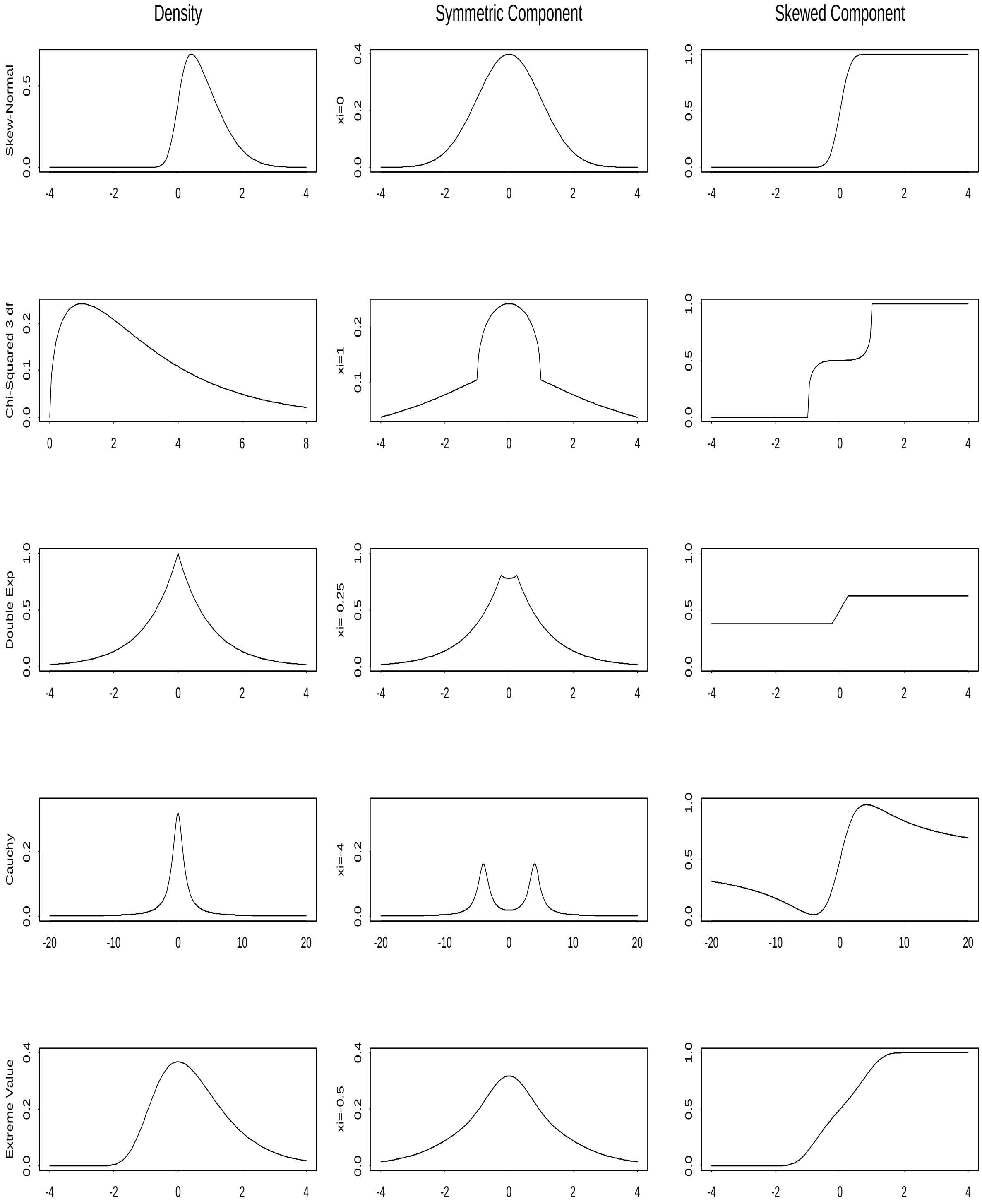


Figure 2: Graphical illustration of the skew-symmetric representation (5) of five univariate pdfs plotted in the first column: skew-normal (row 1), chi-squared (row 2), double-exponential (row 3), Cauchy (row 4), extreme value (row 5).

instance, the choice $\boldsymbol{\xi} = \boldsymbol{\mu}$ in the multivariate skew-normal distribution (1) of Azzalini and Dalla Valle (1996) does not correspond to its mean, median, or mode. The skew-symmetric representation (5) is of theoretical and conceptual interest because it generalizes the ideas of skew-normal and other skew-elliptical distributions. In practice however, a parametric family of symmetric (or elliptical) pdfs $f$ and of skewing functions $\pi$ are chosen, and the skew-symmetric construction (3) of Proposition 1 is used. The vector $\boldsymbol{\xi}$ is then a location parameter which can be estimated from data.

## 4 DISCUSSION

This paper extends the line of work which seeks to generalize the univariate skew-normal distribution of Azzalini (1985). We have examined some properties of skew-symmetric distributions which includes all GSEs, but is very general since any multivariate pdf admits a skew-symmetric representation. We have shown that this family has a succinct mathematical form and can be easily simulated.

Suppose a pdf is smooth (differentiable) and locally concave on some convex neighborhood (as in the neighborhood of a local mode, for instance). Lemma 1 and Proposition 4 then imply that the symmetric component with respect to any point in that neighborhood will be smooth and locally concave. This means that locally, its contours will typically be both convex and smooth. Thus, for an entire range of possible $\boldsymbol{\xi}$'s, the symmetric component can be arguably well-approximated by an elliptical density. It might then be possible, as in the fifth row of Figure 2, to choose such a $\boldsymbol{\xi}$ (or let the data do so) at which the skewing component is well-approximated by a function, such as that in (1), which is practical for empirical analysis. Thus, GSE models as proposed by Genton and Loperfido (2002) may well have wide applicability.

Note that one can use other skew-symmetric representations of a pdf $g$. For instance, one referee suggested replacing the arithmetic mean of $g(\boldsymbol{\xi}+\mathbf{s})$ and $g(\boldsymbol{\xi}-\mathbf{s})$ in (6) by a general mean of the form $\psi^{-1}[(\psi(g(\boldsymbol{\xi}+\mathbf{s}))+\psi(g(\boldsymbol{\xi}-\mathbf{s})))/2]$, where $\psi$ is a suitable function, typically a power function. For instance, the choice $\psi(x)=\log(x)$ yields the geometric mean and therefore yields an alternative skew-symmetric representation $g(\mathbf{x}) = C\tilde{f}_{\boldsymbol{\xi}}(\mathbf{x}-\boldsymbol{\xi})\tilde{\pi}_{\boldsymbol{\xi}}(\mathbf{x}-\boldsymbol{\xi})$, where

$$\tilde{f}_{\boldsymbol{\xi}}(\mathbf{s}) = C^{-1}\sqrt{g(\boldsymbol{\xi}+\mathbf{s})g(\boldsymbol{\xi}-\mathbf{s})}, \qquad \tilde{\pi}_{\boldsymbol{\xi}}(\mathbf{s}) = \sqrt{\frac{g(\boldsymbol{\xi}+\mathbf{s})}{g(\boldsymbol{\xi}-\mathbf{s})}}.$$

The function $\tilde{f}_{\boldsymbol{\xi}}$ is a symmetric pdf once the normalizing constant $C$ has been computed. Note that $C=1$ when $\psi(x)=x$, that is for our skew-symmetric representation (5). The function

$\tilde{\pi}_{\boldsymbol{\xi}}$ has the same monotonicity as $\pi_{\boldsymbol{\xi}}$, but has different characteristics: $0 \leq \tilde{\pi}_{\boldsymbol{\xi}}(\mathbf{x}) < \infty$ and $\tilde{\pi}_{\boldsymbol{\xi}}(-\mathbf{x}) = 1/\tilde{\pi}_{\boldsymbol{\xi}}(\mathbf{x})$. Among all these alternative skew-symmetric representations, our motivation for the one corresponding to $\psi(x) = x$, given in Proposition 3, was that it was analogous to the skew-normal distribution introduced by Azzalini (1985) and all similar generalizations that have since been seen in the literature.


### ACKNOWLEDGMENTS

We would like to thank the Editor, an associate editor, and two anonymous referees for helpful comments that improved this article. We are also grateful to Roger Berger, Nicola Loperfido, Yanyuan Ma, and John Monahan for thoughtful comments on earlier drafts of this paper.


## 5 REFERENCES


Arellano-Valle, R. B., del Pino, G. and San Martin, E. (2002). Definition and probabilistic properties of skew-distributions. *Statistics and Probability Letters* **58**, 111-121.

Arnold, B. C. and Beaver, R. J. (2000). The skew-Cauchy distribution. *Statistics and Probability Letters* **49**, 285-290.

Arnold, B. C. and Beaver, R. J. (2002). Skewed multivariate models related to hidden truncation and/or selective reporting. *Test* **11**, 7-54.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Stat.* **12**, 171-178.

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

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

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

Azzalini, A. and 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.

Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions. *J. Multivariate Analysis* **79**, 99-113.

Domínguez-Molina, J. A., González-Farías, G. and Gupta, A. K. (2001). A general multivariate skew-normal distribution. Technical Report No. 01-09. Department of Mathematics and Statistics, Bowling Green State University.

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

Genton, M. G. and Loperfido, N. (2002). Generalized skew-elliptical distributions and their quadratic forms. *Institute of Statistics Mimeo Series* **#2539**, revised for *Annals of the Institute of Statistical Mathematics.*

Loperfido, N. (2001). Quadratic forms of skew-normal random vectors. *Statistics and Probability Letters* **54**, 381-387.

Sahu, S. K., Dey, D. K. and Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics* **31**, 129-150.

Wahed, A. and Ali, M. M. (2001). The skew-logistic distribution. *Journal of Statistical Research* **35**, 71-80.

Wang, J., Boyer, J. and Genton, M. G. (2004). A note on an equivalence between chi-square and generalized skew-normal distributions. *Statistics and Probability Letters*, in press.

# APPENDIX

**Proof of Proposition 1**

Since $f(\mathbf{x}) \geq 0$ and $\pi(\mathbf{x}) \geq 0$, we need only show that (3) sums or integrates to 1. We prove the result for continuous $f$ with $\boldsymbol{\xi} = \mathbf{0}$. The proof for general $\boldsymbol{\xi}$ follows from the fact that $g(\mathbf{x} - \boldsymbol{\xi})$ is in the same location family as $g(\mathbf{x})$. Let $A^+ = \{(x_1, \ldots, x_n); x_1 \geq 0\}$ and $A^- = \{(x_1, \ldots, x_n); x_1 < 0\}$. Since $A^+ \cup A^- = \mathbb{R}^n$ and $A^+ \cap A^- = \emptyset$,

$$\begin{aligned}\int_{\mathbb{R}^n} 2f(\mathbf{x})\pi(\mathbf{x})d\mathbf{x} &= \int_{A^+} 2f(\mathbf{x})\pi(\mathbf{x})d\mathbf{x} + \int_{A^-} 2f(\mathbf{x})\pi(\mathbf{x})d\mathbf{x} \\ &= 2\int_{A^+} [f(\mathbf{x})\pi(\mathbf{x}) + f(\mathbf{x})(1 - \pi(\mathbf{x}))]d\mathbf{x} \\ &= \int_{\mathbb{R}^n} f(\mathbf{x})d\mathbf{x} = 1,\end{aligned}$$

where we use the properties of $f$ and $\pi$. $\square$

**Proof of Proposition 2**

We show that the class of functions of the form $G(w(\cdot))$ is identical to the class of skewing functions. By the properties of $G$ and $w$, $G(w(-\mathbf{s})) = G(-w(\mathbf{s})) = 1 - G(w(\mathbf{s}))$; so $G(w(\cdot))$ is a skewing function. Conversely, let $H$ be a strictly increasing cdf of a random variable symmetric around 0. We can write, for any skewing function $\pi$, $\pi(\mathbf{s}) = H(k(\mathbf{s}))$, where $k(\mathbf{s}) \equiv H^{-1}(\pi(\mathbf{s}))$. By the properties of $H$ and $\pi$, $H^{-1}(\pi(-\mathbf{s})) = H^{-1}(1 - \pi(\mathbf{s})) = -H^{-1}(\pi(\mathbf{s}))$; so $k(\mathbf{s})$ is an odd function. $\square$

**Proof of Proposition 3**

To prove existence, we can write

$$g(\mathbf{x}) = 2\frac{g(\boldsymbol{\xi} + (\mathbf{x} - \boldsymbol{\xi})) + g(\boldsymbol{\xi} - (\mathbf{x} - \boldsymbol{\xi}))}{2} \frac{g(\boldsymbol{\xi} + (\mathbf{x} - \boldsymbol{\xi}))}{g(\boldsymbol{\xi} + (\mathbf{x} - \boldsymbol{\xi})) + g(\boldsymbol{\xi} - (\mathbf{x} - \boldsymbol{\xi}))}.$$

The required symmetric and skewed components are then

$$f_{\boldsymbol{\xi}}(\mathbf{s}) = \frac{g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s})}{2}, \tag{11}$$

$$\pi_{\boldsymbol{\xi}}(\mathbf{s}) = \frac{g(\boldsymbol{\xi} + \mathbf{s})}{g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s})}. \tag{12}$$

Since $g(\boldsymbol{\xi} + \mathbf{s})$ and $g(\boldsymbol{\xi} - \mathbf{s})$ are members of the same location family as $g$, $f_{\boldsymbol{\xi}}(\mathbf{s})$ is an average of pdfs and thus a pdf. It is also clearly symmetric around $\mathbf{0}$. Since $g$ is a pdf and always greater than zero, $0 \le \pi_{\boldsymbol{\xi}}(\mathbf{s}) \le 1$. Straightforward algebra verifies that $\pi_{\boldsymbol{\xi}}(-\mathbf{s}) = 1 - \pi_{\boldsymbol{\xi}}(\mathbf{s})$. Note here that $\pi_{\boldsymbol{\xi}}(\mathbf{s})$ is not defined if $g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s}) = 0$. Since this implies $f_{\boldsymbol{\xi}}(\mathbf{s}) = 0$, without loss of generality we can simply define $\pi_{\boldsymbol{\xi}}(\mathbf{s}) \equiv \frac{1}{2} \equiv \pi_{\boldsymbol{\xi}}(-\mathbf{s})$ if $g(\boldsymbol{\xi} + \mathbf{s}) + g(\boldsymbol{\xi} - \mathbf{s}) = 0$. In order to prove uniqueness, let $g(\mathbf{x}) = 2f_{\boldsymbol{\xi}}(\mathbf{x} - \boldsymbol{\xi})\pi_{\boldsymbol{\xi}}(\mathbf{x} - \boldsymbol{\xi})$ for some $\boldsymbol{\xi}$, where $f_{\boldsymbol{\xi}}$ is symmetric around $\mathbf{0}$ and $\pi_{\boldsymbol{\xi}}$ is a skewing function. Straightforward algebra, using the properties of $f_{\boldsymbol{\xi}}$ and $\pi_{\boldsymbol{\xi}}$, can be used to verify that $f_{\boldsymbol{\xi}}$ and $\pi_{\boldsymbol{\xi}}$ must satisfy (11) and (12). $\square$

**Proof of Lemma 1**

The first result is evident from (11) and (12). The second and third are obtained by differentiating both sides of (11) and (12). The fourth can be shown by taking two derivatives of (11) $\square$

**Proof of Proposition 4**

We prove the results that pertain to maxima.

1. Let $g(\mathbf{m}) \geq g(\mathbf{x})$ $\forall$ $\mathbf{x}$ in some open ball of radius $\delta$ around $\mathbf{m}$. From (11) and (12), $f_{\mathbf{m}}(\mathbf{0}) = \frac{g(\mathbf{m})+g(\mathbf{m})}{2} \geq \frac{g(\mathbf{m}+\mathbf{s})+g(\mathbf{m}-\mathbf{s})}{2} = f_{\mathbf{m}}(\mathbf{s})$, $\forall$ $\mathbf{s}$ in an open ball of radius $\delta$ around $\mathbf{0}$. If $g$ has a global maximum at $\mathbf{m}$, then $\delta = \infty$. Suppose in addition that $g$ monotonically decreases along any ray in an open ball of radius $\delta$ and $\mathbf{s}_1 > \mathbf{s}_0$. Then by (11), $f_{\mathbf{m}}(\mathbf{s}_0) = \frac{g(\mathbf{m}+\mathbf{s}_0)+g(\mathbf{m}-\mathbf{s}_0)}{2} \geq \frac{g(\mathbf{m}+\mathbf{s}_1)+g(\mathbf{m}-\mathbf{s}_1)}{2} = f_{\mathbf{m}}(\mathbf{s}_1)$ if $\mathbf{s}_0$ and $\mathbf{s}_1$ have length $\leq \delta$. This proves the unimodality of $f_{\mathbf{m}}$.

2. Let $\boldsymbol{\xi} \in N$ and let $\mathbf{s} \in S_{\boldsymbol{\xi}}$. Then $f_{\boldsymbol{\xi}}(\mathbf{0}) = g(\boldsymbol{\xi}) = g(\frac{1}{2}(\boldsymbol{\xi}+\mathbf{s}) + \frac{1}{2}(\boldsymbol{\xi}-\mathbf{s})) \geq \frac{1}{2}g(\boldsymbol{\xi}+\mathbf{s}) + \frac{1}{2}g(\boldsymbol{\xi}-\mathbf{s}) = f_{\boldsymbol{\xi}}(\mathbf{s})$. Let $\mathbf{s}_1$ and $\mathbf{s}_2 \in S$ and let $0 \leq \lambda \leq 1$. Then using (11),

$$\begin{aligned}
f_{\boldsymbol{\xi}}(\lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2) &= \frac{g(\boldsymbol{\xi} + \lambda\mathbf{s}_1 + (1-\lambda)\mathbf{s}_2) + g(\boldsymbol{\xi} - \lambda\mathbf{s}_1 - (1-\lambda)\mathbf{s}_2)}{2} \\
&= \frac{g(\lambda(\boldsymbol{\xi}+\mathbf{s}_1) + (1-\lambda)(\boldsymbol{\xi}+\mathbf{s}_2)) + g(\lambda(\boldsymbol{\xi}-\mathbf{s}_1) + (1-\lambda)(\boldsymbol{\xi}-\mathbf{s}_2))}{2} \\
&\geq \frac{\lambda g(\boldsymbol{\xi}+\mathbf{s}_1) + (1-\lambda)g(\boldsymbol{\xi}+\mathbf{s}_2) + \lambda g(\boldsymbol{\xi}-\mathbf{s}_1) + (1-\lambda)g(\boldsymbol{\xi}-\mathbf{s}_2)}{2} \\
&= \lambda f_{\boldsymbol{\xi}}(\mathbf{s}_1) + (1-\lambda)f_{\boldsymbol{\xi}}(\mathbf{s}_2)
\end{aligned}$$

3. For all $\boldsymbol{\xi} \in N$, $f_{\boldsymbol{\xi}}(\mathbf{0}) = g(\boldsymbol{\xi}) \geq \frac{g(\boldsymbol{\xi}+\mathbf{s})+g(\boldsymbol{\xi}-\mathbf{s})}{2} = f_{\boldsymbol{\xi}}(\mathbf{s})$ $\forall$ $\mathbf{s}$ in some neighborhood of $\mathbf{0}$. By the same line of reasoning that is used to prove that concave functions are continuous, one can use this fact to show that $g$ is continuous on $N$. We can now prove our result by contradiction. Suppose $g$ is not concave on $N$. Then, by the continuity of $g$ and the definition of concavity, there are points $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ in $N$ such that for all $\boldsymbol{\xi}$ on the line segment $L$ connecting $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, $g(\boldsymbol{\xi}) \leq GL(\boldsymbol{\xi})$, with strict inequality on some interval, where $GL$ is the linear function on $L$ such that $GL(\boldsymbol{\xi}_1) = g(\boldsymbol{\xi}_1)$ and $GL(\boldsymbol{\xi}_2) = g(\boldsymbol{\xi}_2)$. By the continuity of $g$, there exists a set of points that minimize $g(\boldsymbol{\xi}) - GL(\boldsymbol{\xi})$ on $L$. For at least one such point $\boldsymbol{\xi}^*$, $f_{\boldsymbol{\xi}^*}$ cannot attain a local maximum at $\mathbf{0}$. (The details of this last point have been left out for brevity.) Here is our contradiction.

□

## Proof of Proposition 5

The first two results can be verified from (12) using straightforward algebra and the two properties of a skewing function. The third follows directly from (10). $\square$

## Proof of Proposition 6

Consider the characteristic function of $\tau(\mathbf{X}-\boldsymbol{\xi})$, $h(t)=Ee^{i\tau(\mathbf{X}-\boldsymbol{\xi})t}$. We have

$$\begin{aligned} h(t) &= \int_{\mathbb{R}^n} 2e^{i\tau(\mathbf{x}-\boldsymbol{\xi})t} f_{\boldsymbol{\xi}}(\mathbf{x}-\boldsymbol{\xi})\pi_{\boldsymbol{\xi}}(\mathbf{x}-\boldsymbol{\xi})d\mathbf{x} \\ &= \int_{A^+} 2e^{i\tau(\mathbf{y})t} f_{\boldsymbol{\xi}}(\mathbf{y})\pi_{\boldsymbol{\xi}}(\mathbf{y})d\mathbf{y} + \int_{A^-} 2e^{i\tau(\mathbf{y})t} f_{\boldsymbol{\xi}}(\mathbf{y})\pi_{\boldsymbol{\xi}}(\mathbf{y})d\mathbf{y} \qquad (\text{let } \mathbf{y}=\mathbf{x}-\boldsymbol{\xi}) \\ &= 2\int_{A^+} e^{i\tau(\mathbf{y})t} f_{\boldsymbol{\xi}}(\mathbf{y})(\pi_{\boldsymbol{\xi}}(\mathbf{y})+1-\pi_{\boldsymbol{\xi}}(\mathbf{y}))d\mathbf{y} \\ &= \int_{\mathbb{R}^n} e^{i\tau(\mathbf{y})t} f_{\boldsymbol{\xi}}(\mathbf{y})d\mathbf{y} \end{aligned}$$

does not depend on $\pi_{\boldsymbol{\xi}}$, where $A^+=\{(y_1,\ldots,y_n),y_1\geq 0\}$ and $A^-=\{(y_1,\ldots,y_n),y_1<0\}$. $\square$