Multivariate skewing mechanisms: a unified perspective based on the transformation approach

Christophe Ley[a], Davy Paindaveine[*,a]

[a] *Université Libre de Bruxelles, E.C.A.R.E.S., Institut de Recherche en Statistique, and Département de Mathématique, Avenue F.D. Roosevelt, 50, CP114, B-1050, Brussels, Belgium*
Phone: +32-2-650.38.45, Fax: +32-2-650.44.75

Abstract

In recent years, models for (possibly multivariate) skewed distributions have become more and more popular. In the univariate case, Ferreira and Steel (2006) [Ferreira, J.T.A.S., Steel, M.F.J., 2006. A constructive representation of univariate skewed distributions. J. Amer. Statist. Assoc. 101, 823-829] introduced general skewing mechanisms in order to compare existing skewing methods in a common framework and to ease construction of new such methods according to the needs in given situations. In this paper, we make use of the classical transformation approach to define alternative skewing mechanisms for the same purpose. While keeping all the nice features of Ferreira and Steel's skewing mechanisms (flexibility, surjectivity, the possibility of retaining prespecified characteristics of the original symmetric distribution, etc.), our skewing mechanisms, unlike theirs, can easily be extended to the multivariate case. We describe our skewing schemes, investigate their main properties, and illustrate their effects on standard (multi)normal distributions by means of a few examples. Finally, we briefly discuss their relevance in the context of optimal symmetry testing.

Key words: Probability transform, Skewing mechanism, Skew-normal distribution, Tail behavior, Transformation approach
2000 MSC: 60E05, 62H05, 62E10

1. Introduction.

The reasons for the growing interest in models for asymmetric distributions are mainly twofold. First, they of course potentially provide a much better fit for data presenting some strong departure from symmetry. Second, they provide specific alternatives in the construction of tests for symmetry. Most models of asymmetric distributions proposed in the literature allow for a continuous variation from symmetry to asymmetry, obtained by transforming an arbitrary symmetric distribution by means of a *skewing mechanism*. The resulting skewed distributions often share some of the properties of their symmetric antecedent, depending on the nature of the mechanism. We now briefly discuss some well-known models of asymmetric distributions.

Introduced in Azzalini (1985), the skew-normal distributions met a huge success in the subsequent years, thanks to their mathematical tractability and to the fact that they retain some of the properties of the normal distribution. The basic *skew-normal* density is given by

$$x \mapsto f^{\mathrm{SN}}(x;\delta) = 2\phi(x)\Phi(\delta x), \quad x \in \mathbb{R}, \tag{1.1}$$

where ϕ and Φ respectively stand for the probability density function (pdf) and the cumulative distribution function (cdf) of the standard normal distribution, and the parameter $\delta \in \mathbb{R}$ captures asymmetry. Positive (resp., negative) values of δ yield right-skewed (resp., left-skewed) distributions,

[*]Corresponding author
Email addresses: chrisley@ulb.ac.be (Christophe Ley), dpaindav@ulb.ac.be (Davy Paindaveine)

while $\delta = 0$ corresponds to the symmetric case. The pioneering work of Azzalini (1985) led to numerous extensions of the skew-normal model. To cite a few examples, Azzalini and Dalla Valle (1996) defined the multivariate skew-normal densities, Genton and Loperfido (2005) the generalized skew-elliptical distributions, and Wang et al. (2004) the multivariate *skew-symmetric* ones. The latter are associated with densities of the form

$$x \mapsto f^{\rm SS}(x;\delta) = 2f(x)\Pi(\delta' x), \quad x \in \mathbb{R}^k,$$

where f is the k-variate symmetric pdf to be skewed, $\Pi : \mathbb{R} \to [0,1]$ is a function satisfying $\Pi(-x) = 1 - \Pi(x)$ for all x, and $\delta \in \mathbb{R}^k$ stands for the asymmetry parameter. Skew-symmetric densities are the most general multivariate extensions of the univariate density in (1.1). For more details, see the review paper Azzalini (2005).

Univariate symmetric distributions can alternatively be skewed by introducing *different scale factors* on either side of the symmetry center. Starting from a symmetric pdf f (throughout, symmetric pdfs/cdfs refer to pdfs/cdfs of random variables—or random vectors—whose distribution does not change under reflection *about the origin*), the resulting densities typically are of the form

$$x \mapsto f^{\rm DSF}(x;\delta) = \begin{cases} a_1(\delta) f(a_2(\delta)x) & \text{if } x \leq 0 \\ b_1(\delta) f(b_2(\delta)x) & \text{if } x > 0, \end{cases}$$

where $a_2(.)$ and $b_2(.)$ are heterogeneous scaling functions, $a_1(.)$ and $b_1(.)$ are normalizing functions, and $\delta \in \mathbb{R}$ again plays the role of a skewness parameter. This appealing and easily interpretable skewing scheme has been used by several authors, e.g., Fechner (1897), Fernàndez and Steel (1998), or Mudholkar and Hutson (2000), who defined the so-called epsilon-skew-normal distributions. Of particular interest is the inverse scale factors model from Fernàndez and Steel (1998), where $a_2(\delta) = e^\delta = (b_2(\delta))^{-1}$. Ferreira and Steel (2007) proposed a multivariate extension of this model, by introducing different scale factors in each marginal of a k-variate random vector. For more information on the link between different scaling functions, we refer the reader to Jones (2006).

There exist other ways of skewing symmetric distributions, for example by using order statistics (Jones, 2004), by applying Tukey's g-and-h transformations (Hoaglin 1986, or Field and Genton 2006 for the multivariate case) or through the sinh-arcsinh transformation (Jones and Pewsey, 2009). Each of the pre-cited skewing methods has advantages and disadvantages, but they all share a common drawback: most properties of the resulting skewed distributions strongly depend on the choice of the method, and hence follow a pre-defined pattern and cannot be adapted according to the needs in certain situations. Thus, despite their flexibility compared to their symmetric counterparts, those asymmetric models still remain restrictive. Moreover, some models suffer from serious inferential problems. In the vicinity of symmetry, Fisher information matrices happen to be singular in some parametric subclasses of the skew-symmetric models, implying that the optimal tests for symmetry against those alternatives coincide with the trivial test, that is, the test rejecting the null of symmetry at level α whenever an auxiliary Bernoulli variable with parameter α takes the value 1; see Ley and Paindaveine (2010a) for more details.

A solution to the problems mentioned above has been proposed in the univariate setup by Ferreira and Steel (2006), who presented a unified perspective on skewing distributions. Their key idea consists in separating the skewing mechanism from the original symmetric distribution to be skewed. More concretely, Ferreira and Steel (2006)—hereafter referred to as FS06—based their mechanisms on a probability transform turning any symmetric cdf F into a cdf of the form

$$x \mapsto F_L^{\rm FS}(x) = L(F(x)), \tag{1.2}$$

where L is a cdf over $[0,1]$. FS06 showed that $F_L^{\rm FS}$ is symmetric iff L is antisymmetric with respect to $1/2$ (in the sense that $L(1-x) = 1 - L(x)$ for all x). If the emphasis is on skewing, it is therefore natural to define a skewing mechanism $\mathcal{L}^S = \{L\}$ as a collection of cdfs L over $[0,1]$ such that the only cdf in $\mathcal{L}^S$ being antisymmetric with respect to $1/2$ is the identity function I—which of course leaves the symmetric cdf F untouched. It is straightforward to show that, starting from a given cdf F, any (asymmetric) distribution can be obtained using the probability transform in (1.2), provided that its support coincides with (or is included in) the support of F. This *surjectivity* property thus allows to compare different skewing mechanisms in a common framework, as is done in FS06 for some of the models presented above. The decomposition into two distinct components—the skewing mechanism and the original symmetric distribution F—allows to improve on existing skewing methods: on one hand, the freedom of choosing the skewing mechanism independently of F eases the construction of new skewed distributions with pre-defined characteristics, replacing the arbitrariness encountered in the usual skewing methods with a large amount of flexibility. On the other hand, it provides inferential advantages, since very general optimal tests for symmetry can be achieved; see Ley and Paindaveine (2009).

However, nice and appealing as they are, the FS06 skewing mechanisms present a major disadvantage: they do not allow for a satisfactory multivariate extension (see Section 2). The present paper therefore proposes an alternative univariate probability transform by resorting to the well-known transformation approach. Parallel to the FS06 proposal, our transform separates the skewing mechanism from the symmetric distribution to be skewed, but it can be extended to the multivariate case in a highly satisfactory way. In particular, our proposal ensures surjectivity in any dimension, and so all the (ad hoc) multivariate asymmetric models described above are compatible with our general construction, and hence can be compared in a common framework. In other words, we combine the mathematical ease of the classical transformation approach with the FS06 idea of a unified perspective on skewing mechanisms.

The paper is organized as follows. In Section 2, we present this alternative univariate probability transform as well as its multivariate extension, establish the corresponding (multidimensional) surjectivity property, and introduce the resulting skewing mechanisms. The main properties of our skewing scheme are stated in Section 3, while Section 4 contains examples of such skewing mechanisms. Some final remarks are given in Section 5. Finally, the Appendix collects the technical proofs.

2. The proposed skewing mechanisms.

2.1. An alternative probability transform.

Denoting by ℓ the pdf associated with the cdf L, the pdf of the distribution in (1.2) is

$$x \mapsto f_L^{\mathrm{FS}}(x) = \ell(F(x))\, f(x),$$

and hence can be regarded as a weighted version of the pdf $f = F'$, with a weight function given by $x \mapsto \ell(F(x))$. Note that this way of skewing distributions is obtained by weighting the *quantile space*.

In this paper, we make use of the transformation approach to rather propose skewing mechanisms based on a probability transform that acts on the *sample space*. More precisely, we define

$$x \mapsto f_H(x) = f(H(x))\, H'(x), \tag{2.1}$$

where $H : \mathbb{R} \to \mathbb{R}$ is a monotone increasing diffeomorphism of $\mathbb{R}$ (throughout, the terminology *diffeomorphism* of $S \subset \mathbb{R}^k$ onto $T \subset \mathbb{R}^k$ refers to a one-to-one mapping $H : S \to T$ such that both H and its inverse H^{-1} are continuously differentiable). If the random variable X has cdf F, then f_H is the pdf of $H^{-1}(X)$ (which explains the connection with the transformation approach), and the corresponding cdf is given by $x \mapsto F_H(x) := F(H(x))$. Clearly, F_H is symmetric iff H is an odd function, which allows us to define general univariate skewing mechanisms along the same lines as in FS06.

These skewing mechanisms enjoy the same nice properties as their FS06 competitors defined in (1.2): any skewed distribution can be written as a skewed version F_H of F for an appropriate function H (this surjectivity property here remains valid for skewed versions with larger supports than F, which constitutes an improvement on the FS06 proposal), the skewing mechanisms may be defined independently of F, etc. However, the link with the transformation approach makes it easy to extend the probability transform in (2.1) to the multivariate setup: the relevant probability transform then maps the pdf f of the k-variate random vector X onto the pdf of $H^{-1}(X)$ (where H is a diffeomorphism of $\mathbb{R}^k$), that is, onto

$$x \mapsto f_H(x) = f(H(x))\,|DH(x)|, \tag{2.2}$$

where $|DH(x)|$ stands for the absolute value of the determinant of the jacobian matrix $DH(x)$ of H at x. This nicely allows for a stochastic representation of the resulting skewed distribution.

In the sequel, we restrict to diffeomorphisms $H : x = (x_1, \ldots, x_k)' \mapsto (H_1(x), \ldots, H_k(x))'$ of $\mathbb{R}^k$ such that, for all $j = 1, \ldots, k$, (i) $H_j(x)$ does not depend on $x_{j+1}, \ldots, x_k$ and (ii) $h_j^{x_1,\ldots,x_{j-1}}(x_j) := H_j(x)$ is, for any fixed $x_1, \ldots, x_{j-1}$, strictly monotone increasing (and hence, invertible) with respect to x_j. Throughout, we denote by $\mathcal{H}$ the collection of such diffeomorphisms. Clearly, $(\mathcal{H}, \circ)$ is a group. Adopting the notation $\stackrel{\mathcal{D}}{=}$ for equality in distribution, the following result, which is proved in the Appendix, explains why one may restrict to $\mathcal{H}$ and shows that, as in the univariate case, any distribution can be obtained from any other distribution by the probability transform (2.2), hence that surjectivity extends to the multivariate setup.

Theorem 2.1 *For any couple (X, Y) of absolutely continuous k-variate random vectors, there exists a unique $H \in \mathcal{H}$ such that $H(X) \stackrel{\mathcal{D}}{=} Y$.*

It appears difficult (if at all possible) to define a multivariate extension of the FS06 transform that would satisfy the surjectivity property in Theorem 2.1. For instance, defining f_L^{FS} as the k-variate pdf associated with the cdf $x \mapsto F_L^{\mathrm{FS}}(x) := L(F(x))$, where L still stands for a cdf over $[0, 1]$, clearly would not achieve this surjectivity as such a skewing scheme does not allow for skewing marginals individually: if $X = (X_1, \ldots, X_k)'$ has a symmetric cdf F, no L can make the resulting distribution skewed in the first marginal only (more precisely, denoting by Y a k-variate random vector with pdf f_L^{FS}, it is not possible to have $-Y_j \stackrel{\mathcal{D}}{=} Y_j$ iff $j = 2, \ldots, k$). A much more flexible extension of the FS06 transform, proposed by an anonymous referee, maps the symmetric k-variate cdf

$$x \mapsto F(x) = C(F_1(x_1), \ldots, F_k(x_k))$$

(written in terms of the corresponding copula C and marginal cdfs F_j, $j = 1, \ldots, k$) onto

$$x \mapsto F_{L_1,\ldots,L_k}^{\mathrm{FS}}(x) := C(L_1(F_1(x_1)), \ldots, L_k(F_k(x_k))),$$

where each L_j is a cdf over $[0, 1]$. This skewing scheme applies the univariate FS06 skewing method in each marginal independently (which allows us to generate arbitrary marginal distributions), but leaves the copula function untouched. This implies that the surjectivity in Theorem 2.1 cannot hold, and that some classical multivariate skewing schemes cannot be written in this form. For instance, this rules out the Azzalini and Dalla Valle (1996) multivariate skew-normal scheme, which typically turns the standard multinormal distribution into a distribution with dependent marginals. In the multivariate case, the class of skewing mechanisms that we propose in this paper is therefore much richer than (the natural multivariate extensions of) the FS06 one.

One might argue that the class $\mathcal{H}$ of diffeomorphisms above introduces an unpleasant ordering of the components $x_1, \ldots, x_k$. This, however, only fixes an ordering in (the arguments of) the transformations themselves: the surjectivity in Theorem 2.1 ensures that, despite this ordering, any random vector X can be turned into any random vector Y. Moreover, in the bivariate case (similar reasonings are likely to hold in higher dimensions), the ordering can be circumvented by applying our transformation to the polar coordinates instead of the usual cartesian ones.

As a closing remark for this subsection, we point out that if X is a k-variate random vector ($k > 1$) with cdf F and pdf f, then the cdf of $H^{-1}(X)$, $H \in \mathcal{H}$, unlike in the univariate case, in general is *not* $x \mapsto F(H(x))$. This can be seen by noticing that, by definition of $\mathcal{H}$, the jacobian matrix $DH(x)$ is triangular, and hence the pdf in (2.2) reduces to $x \mapsto f_H(x) = f(H(x)) \prod_{j=1}^k \partial_{x_j}(H_j(x))$ and thus differs, for $k > 1$, from the pdf $x \mapsto \partial^k_{x_1,x_2,\ldots,x_k} F(H(x))$ corresponding to the cdf $x \mapsto F(H(x))$. Therefore, in the sequel, we simply define F_H as the cdf associated with the pdf f_H in (2.2).

2.2. The resulting class of (multivariate) skewing mechanisms.

In order to define the general multivariate skewing mechanisms based on the probability transform in (2.2), we need the following result.

Theorem 2.2 *Let f be the pdf of a symmetric absolutely continuous k-variate random vector X and fix $H \in \mathcal{H}$. Then (i) $Y \sim f_H$ iff $Y \stackrel{\mathcal{D}}{=} H^{-1}(X)$; (ii) $f_H = f$ iff $H = I$; (iii) if $Y \sim f_H$, then $-Y \sim f_{\bar{H}}$, where $\bar{H}(x) := -H(-x)$; (iv) f_H is symmetric iff H is an odd function.*

The diffeomorphism $\bar{H}$ will be called the dual of H. In the univariate case, if f_H is skewed to the right (resp., to the left), then $f_{\bar{H}}$ is skewed to the left (resp., to the right) *with the same type and amount of skewness*. Similar comments can be made in the multivariate case. Clearly, $\bar{\bar{H}} = H$ for all $H \in \mathcal{H}$, and the only mappings fixed by this duality operator are the odd ones. In view of Theorem 2.2, it is natural to adopt the following definition.

Definition 2.1 *A* skewing mechanism (SM) *is a subset $\mathcal{H}^S$ of $\mathcal{H}$ satisfying the following conditions: (i) $I \in \mathcal{H}^S$ and (ii) the set $\mathcal{H}^S \setminus \{I\}$ is nonempty but contains no odd mapping. An SM is said to be symmetric iff $\bar{H} \in \mathcal{H}^S$ for all $H \in \mathcal{H}^S$.*

Clearly, for any symmetric k-variate pdf f, the pdf f_H, $H \in \mathcal{H}^S$, will then be skewed, unless, of course, $H = I$—in which case $f_H = f$. The skewing mechanism $\mathcal{H}^S$ therefore can be regarded as a tool for skewing any fixed symmetric density f.

An important special case (especially so when the focus is on statistical inference; see Ley and Paindaveine, 2009, 2010b) occurs when $\mathcal{H}^S$ is indexed by some finite dimensional parameter—a d-dimensional parameter δ, say.

Definition 2.2 *A* parametric skewing mechanism (PSM) *is a skewing mechanism of the form $\mathcal{H}^S = \{H_\delta : \delta \in \mathcal{D} \subset \mathbb{R}^d\}$ for which $H_0 = I$. Moreover, this PSM is said to be a* canonical PSM (CPSM) *iff $\bar{H}_\delta = H_{-\delta}$ for all $\delta \in \mathcal{D}$ such that $\bar{H}_\delta \in \mathcal{H}^S$.*

For any PSM, it is therefore assumed that $\mathcal{D}$ contains the origin of $\mathbb{R}^d$, which is the only (see Definition 2.1) value of the skewness parameter δ that does not skew the underlying symmetric distribution. Theorem 2.2(ii) and (iv) for a PSM then translate into $f_{H_\delta} = f$ iff $\delta = 0$ iff f_{H_δ} is symmetric. This means that a skewness parameter equal to 0 corresponds to letting the pdf f be symmetric (or equivalently, untouched), an idea that appears to be quite natural. In the same spirit, we favor CPSMs over PSMs, since the former ensure that $Y \sim f_{H_\delta}$ iff $-Y \sim f_{H_{-\delta}}$; in other words, if δ is the value associated with some skewness to the right (resp., left) for a univariate CPSM, then the corresponding skewness to the left (resp., right) is obtained for the value $-\delta$ of the parameter. Again, a similar interpretation holds in the multivariate case.

3. Characteristics of the proposed skewing mechanisms.

Part of the flexibility of the proposed SMs is that one can define an SM in such a way the resulting skewed distributions exhibit some specific properties or retain some prespecified features of the original symmetric distribution. In this section, we provide some structural characteristics inherent to the probability transform in (2.2), and hence also to the related SMs. These characteristics determine whether a certain SM is more or less appropriate in certain situations, and allow for a comparison between different SMs. Since FS06 analyzed the univariate skew-normal distribution and the inverse scale factor model on basis of their framework, we use instead the multivariate sinh-arcsinh transform (see Jones and Pewsey 2009) to illustrate the subsequent properties of our general SMs. It is straightforward to check that this transform is associated with our SM $\mathcal{H}^{\mathrm{SA}} := \{H_\delta^{\mathrm{SA}} : \delta = (\delta_1, \ldots, \delta_k)' \in \mathbb{R}^k\} \subset \mathcal{H}$, where $H_\delta^{\mathrm{SA}}(x_1, \ldots, x_k) = (H_{\delta_1}(x_1), \ldots, H_{\delta_k}(x_k))'$, with $H_{\delta_j}(x_j) = \sinh(\operatorname{arcsinh}(x_j) + \delta_j)$ for $j = 1, \ldots, k$.

While FS06 turned their attention to the mode as the location measure for univariate distributions, we focus instead on the median. From an inferential point of view, fixing the median of the original symmetric distribution happens to be more important than the (possibly not unique) mode. Moreover, considering both the median and the mode as functionals over the space of densities, the median, unlike the mode, is continuous with respect to the $\mathcal{L}^\infty$-norm, which is highly desirable when aiming at optimal inference about location. Therefore, we present in Theorem 3.1 a necessary and sufficient condition for a (multivariate) $H \in \mathcal{H}$ to fix the (componentwise) median of the original symmetric distribution.

Theorem 3.1 *Fix $H \in \mathcal{H}$. Then, the componentwise median of f_H coincides, for every symmetric pdf f, with the componentwise median of f iff for all $j \in \{1, \ldots, k\}$, we have that $h_j^{x_1,\ldots,x_{j-1}}(0) = 0$ for all $x_1, \ldots, x_{j-1}$ (for $k = 1$, this can be simply rewritten as $H(0) = h_1(0) = 0$).*

As an immediate consequence, it follows that $H(0_k) = 0_k$ (where 0_k stands for the origin of $\mathbb{R}^k$) is a necessary but not sufficient condition for $H \in \mathcal{H}$ to fix the componentwise median—unless $k = 1$, that is, unless we are in the univariate setup. It should be noted that the condition $h_j^{x_1,\ldots,x_{j-1}}(0) = 0$ for any $x_1, \ldots, x_{j-1}$ is actually an "iff" condition for fixing the median in the jth marginal (see the proof of Theorem 3.1). Since the mapping $x \mapsto \sinh(\mathrm{arcsinh}(x) + \delta)$ for $\delta \neq 0$ does not fix 0, the skewing mechanism $\mathcal{H}^{\mathrm{SA}}$ does not fix the componentwise median (actually, not even the median of any marginal), which may be a serious drawback in some inferential problems; see, e.g., Ley and Paindaveine (2009).

Let us now examine the tail behavior of f_H. To do so, we extend the FS06 definition of the largest right/left moment of a univariate distribution to the largest u-directional moment of a k-variate distribution, where u belongs to $\mathcal{S}^{k-1}$, the unit sphere in $\mathbb{R}^k$.

Definition 3.1 *Letting $H_u^+ := \{x \in \mathbb{R}^k : u'x > 0\}$ for any $u \in \mathcal{S}^{k-1}$, we define the largest u-directional moment of the k-variate pdf g as*

$$M_{u,g} := \sup\left\{ r \in \mathbb{R}^+ : \int_{H_u^+} |u'x|^r \, g(x)\, dx < \infty \right\} \tag{3.1}$$

and the largest moment of g as $M_g := \sup_{u \in \mathcal{S}^{k-1}} M_{u,g}$.

If g is a symmetric pdf, then $M_{u,g} = M_{-u,g}$ for any $u \in \mathcal{S}^{k-1}$, which, in the univariate case, simply means that the largest right moment $M_{1,g}$ coincides with the largest left one $M_{-1,g}$. Theorem 3.2 below links the largest u-directional moments of f_H with those of the original symmetric density f (see the Appendix for a proof).

Theorem 3.2 *Let f be a symmetric k-variate pdf and let $H \in \mathcal{H}$. Fix $q > 0$ and $u \in \mathcal{S}^{k-1}$. Define, for any $D > 0$, $H_u^+(D) := \{x \in \mathbb{R}^k : u'x > D\}$ and $HH_u^+(D) := \{x \in \mathbb{R}^k : u'H^{-1}(x) > D\}$. Then:*
(i) if there exist $D_1, D_2, c > 0$ such that $|u'H^{-1}(x)|/|u'x|^q \leq c$ for all $x \in HH_u^+(D_1) \cap H_u^+(D_2)$ and if $|u'H^{-1}(x)|$ is bounded over $HH_u^+(D_1) \setminus H_u^+(D_2)$, then $M_{u,f_H} \geq M_{u,f}/q$;
(ii) if there exist $D_1, D_2, c_1, c_2 > 0$ such that $c_1 \leq |u'H^{-1}(x)|/|u'x|^q \leq c_2$ for all $x \in HH_u^+(D_1) \cap H_u^+(D_2)$ and if $|u'H^{-1}(x)|$ is bounded over $HH_u^+(D_1) \setminus H_u^+(D_2)$ and $|u'x|$ over $H_u^+(D_2) \setminus HH_u^+(D_1)$, then $M_{u,f_H} = M_{u,f}/q$.
Moreover, for $k = 1$, we have that
(iii) $\lim_{x\to\pm\infty} |x|/|H(x)|^q$ is finite iff $M_{\pm 1,f_H} \geq M_{\pm 1,f}/q$, and
(iv) if $\lim_{x\to\pm\infty} |x|/|H(x)|^q$ is finite and non-zero, then $M_{\pm 1,f_H} = M_{\pm 1,f}/q$.

The complex nature of multidimensional largest u-directional moments explains why the first two points of Theorem 3.2 are not equivalences for $k > 1$: since the integral in (3.1) is taken over the half-space H_u^+, there may be many diverse reasons, difficult to identify, for it not converging. Such issues, however, do not appear in the univariate case ($k = 1$), where, as shown in Theorem 3.2(iii)-(iv), the sufficient conditions involve simple limits only.

This allows for investigating straightforwardly the tail behavior of distributions skewed by means of the skewing mechanism $\mathcal{H}^{\mathrm{SA}}$. First note that $(H_\delta^{\mathrm{SA}})^{-1}(x) = (H_{\delta_1}^{-1}(x_1), \ldots, H_{\delta_k}^{-1}(x_k))'$, where $H_{\delta_j}^{-1}(x_j) := \sinh(\mathrm{arcsinh}(x_j) - \delta_j)$, satisfies $\lim_{x_j \to \pm\infty} H_{\delta_j}^{-1}(x_j)/x_j = \exp(\mp\delta_j)$ for any δ_j. Hence Theorem 3.2(ii) yields that $M_{\pm e_j, f_H} = M_{\pm e_j, f}$ for any $j \in \{1, \ldots, k\}$, where e_j stands for the jth vector of the canonical basis of $\mathbb{R}^k$. In other words, the sinh-arcsinh transform leaves untouched the largest moments in each semi-axial direction. It can be seen that this extends to any direction $u \in \mathcal{S}^{k-1}$.

Theorem 3.2, on one hand, is a useful tool for the evaluation of the moment characteristics of previously defined classes of skewed distributions, and on the other hand it provides the conditions that the probability transforms H of an SM $\mathcal{H}^S = \{H\}$ have to fulfill if one aims at defining novel *moment-preserving* classes of skewed distributions. The latter requirement corresponds to the special case $q = 1$ in Theorem 3.2(ii), and was met above for the skewing mechanism $\mathcal{H}^{\mathrm{SA}}$. Since the tail behavior can be well controlled by the (multivariate) probability transforms that we propose, it seems quite obvious that, by adding a vectorial parameter ν, one could define general skewing-tailing mechanisms on the basis of the framework presented here. This idea is strengthened by the fact that quantiles—on which natural tail weight measures can be based—interact well with transformations (see Gilchrist, 2000). However, this is beyond the scope of the present paper, and we do not pursue that idea here.

Besides these distributional aspects, FS06 also puts much emphasis on the fact that the skewing mechanisms introduced there can be defined independently of the pdf f to be skewed. As already mentioned, this requirement also applies in the context of our SMs.

Definition 3.2 *An SM $\mathcal{H}^S$ is said to be independent of the pdf f to be skewed if none of its members is defined in terms of f.*

Clearly, the skewing mechanism $\mathcal{H}^{\mathrm{SA}}$ is independent of the symmetric distributions to be skewed, which clearly constitutes an advantage of the skewing method described by Jones and Pewsey (2009). Now fix, in the univariate case, an FS06 skewing mechanism $\mathcal{L}^S = \{L\}$, and consider, for any arbitrary symmetric density f, the collection of resulting skewed pdfs $\{f_L^{\mathrm{FS}}\}$, or equivalently, the collection of skewed cdfs $\{F_L^{\mathrm{FS}}\}$. Thanks to the surjectivity of our own SM, there exists, for any $L \in \mathcal{L}^S$, some H belonging to $\mathcal{H}$ such that $F_L^{\mathrm{FS}} = F_H$; in other words, the FS06 skewing mechanism $\mathcal{L}^S = \{L\}$ induces an SM $\mathcal{H}^S = \{H\}$ through the relation

$$(L \circ F =:)\, F_L^{\mathrm{FS}} = F_H \,(:= F \circ H), \text{ equivalently through } H = F^{-1} \circ L \circ F.$$

Most importantly, we want to stress that this clearly shows that, if the original skewing mechanism $\mathcal{L}^S = \{L\}$ is independent of f, then the induced SM $\mathcal{H}^S = \{H = F^{-1} \circ L \circ F\}$ does depend on f. And vice versa, it can be shown that an SM $\mathcal{H}^S$ that does not depend on the pdf f to be skewed induces an FS06 skewing mechanism $\mathcal{L}^S$ that will crucially depend on f. An example for the latter statement is provided by the PSM $\mathcal{H}_1 = \{H_\delta(x) = x - \delta : \delta \in \mathbb{R}\}$. Such a PSM corresponds to a location shift, which, in terms of FS06 parametric skewing mechanisms, entails dependence on the original distribution. We conclude that the skewing mechanisms being independent of the original symmetric distribution is certainly an interesting feature but only makes sense when considered with respect to some fixed type of skewing mechanism (either the FS06 one or ours). This independence, in any case, should therefore be treated with care.

4. Some examples.

In this section, we consider three examples of symmetric PSMs (in fact all of them are CPSMs), investigate their properties and show their effects on standard (multi)normal distributions. The first two examples are univariate, whereas the third one is bivariate. Most properties of those PSMs will be discussed in the light of the results of Section 3. All three examples are chosen in such a way that they meet the independence property requirement of Definition 3.2. For the sake of illustration, Figure 1 provides plots of several functions H_δ belonging to each of the two one-dimensional PSMs considered, along with the corresponding skewed standard normal densities; Figure 2 plots skewed versions of the standard bivariate normal density obtained through the two-dimensional PSM $\mathcal{H}_3^S$ given below.

As a first example, consider the PSM $\mathcal{H}_1^S := \{H_{1\delta} : \delta \in [-1, 1]\}$ defined by

$$H_{1\delta}(x) := x - \delta \cos(x).$$

The set of admissible values for δ (namely, $[-1, 1]$) is here the largest subset of $\mathbb{R}$ over which $H_{1\delta}$ is a diffeomorphism of $\mathbb{R}$. Theorem 3.1 readily shows that this PSM does not fix the median of the original distribution. Moreover, a simple limit calculation reveals (Theorem 3.2(iv)) that $\mathcal{H}_1^S$ preserves the right and left moment structure of the original symmetric distribution to be skewed, so the tail behavior remains unchanged by the PSM.

The second PSM is defined by $\mathcal{H}_2^S := \{H_{2\delta} : \delta \in \mathbb{R}\}$ where

$$H_{2\delta}(x) := x\,\frac{(1 + \exp(\delta x))}{2}.$$

In contrast to $\mathcal{H}_1^S$, this second PSM preserves the median of the original symmetric distribution since $H_{2\delta}(0) = 0$ for any $\delta \in \mathbb{R}$, which makes this PSM more suitable for inferential purposes. Turning our attention to the tails of the resulting skewed distributions, the situation becomes quite interesting since their left and right weights differ. Actually, for $\delta > 0$, $\lim_{x\to-\infty} |x|/|H_{2\delta}(x)| = 2$ and $\lim_{x\to+\infty} |x|/|H_{2\delta}(x)|^q = 0$ for each $q > 0$, implying that the largest left moment of the skewed version is inherited from the original distribution whereas the largest right moment becomes infinite. Intuitively, the exponential term concentrates a high mass of probability in the neighborhood of the origin on the positive real half-line, leading to extremely light tails on the right. Clearly, negative values of the skewness parameter lead to the opposite conclusions.

Our final example involves a two-dimensional PSM, namely $\mathcal{H}_3^S := \{H_{3\delta} : \delta \in \mathbb{R}\}$ defined by

$$H_{3\delta}(x_1, x_2) := (x_1, (2 + \arctan(\delta x_1))x_2/2)'.$$

We see that, thanks to Theorem 3.1, this two-dimensional PSM fixes the componentwise median of the original symmetric distribution. As for the largest u-directional moments, first note that, for each $\delta \in \mathbb{R}$, the inverse of $H_{3\delta}$ is simply $H_{3\delta}^{-1}(x_1, x_2) = (x_1, 2(2 + \arctan(\delta x_1))^{-1}x_2)'$. One immediately sees that the largest $\pm e_1$-directional moments are preserved. The boundedness of $(2 + \arctan(\delta x_1))^{-1}$ leads to the same conclusion for the largest $\pm e_2$-directional moments, showing that the PSM $\mathcal{H}_3^S$ conserves the largest semi-axial moments. Again, it can easily be seen that this extends to any direction u.

5. Final remarks.

When the focus is on testing for symmetry, it is clear that the more flexible the asymmetric alternatives are, the more general the tests become. This latter fact was the main motivation for Ley and Paindaveine (2009) for building Le Cam optimal tests for symmetry against (univariate) skewed distributions generated by the FS06 skewing mechanisms. In the same spirit, we shall, in a future work (Ley and Paindaveine, 2010b), construct Le Cam optimal tests for symmetry against distributions skewed by means of the PSMs proposed in this paper, both in the univariate and multivariate setups.

It should be noted that, for inferential purposes, our PSMs may indeed appear more advantageous than the FS06 ones; in particular, the stochastic representation of the corresponding skewed densities (recall that, for any $H \in \mathcal{H}$, f_H is the pdf of $H^{-1}(X)$ if X admits the original symmetric pdf f) paves the way to natural procedures of estimation of the skewness parameter in any given PSM, or, alternatively, might be useful for discriminating between PSMs.

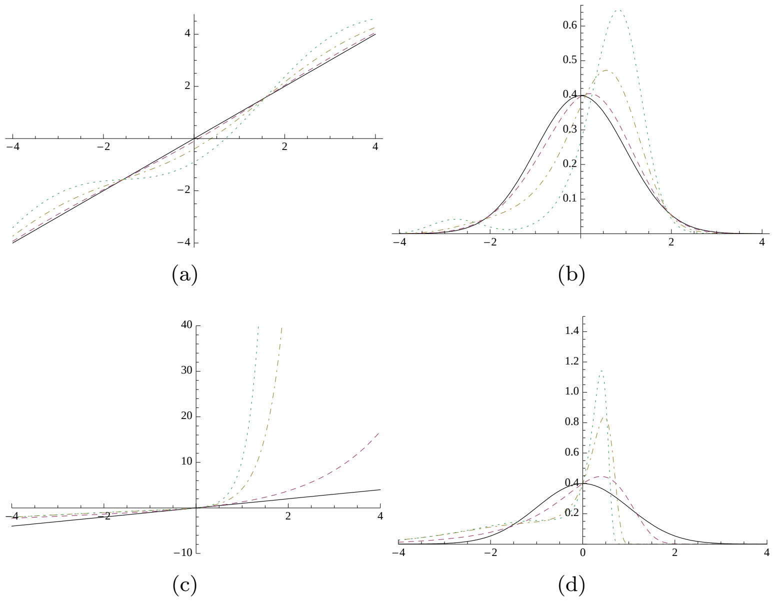

Figure 1: Plots of $H_{1\delta}$ (a) (resp., $H_{2\delta}$ (c)) and of the resulting skewed versions of the standard normal density (b) (resp., (d)) for $\delta = 0$ (solid line), $\delta = 0.1$ (resp., 0.5) (dashed line), $\delta = 0.4$ (resp., 2) (dash-dot line), and $\delta = 0.9$ (resp., 3) (dotted line).

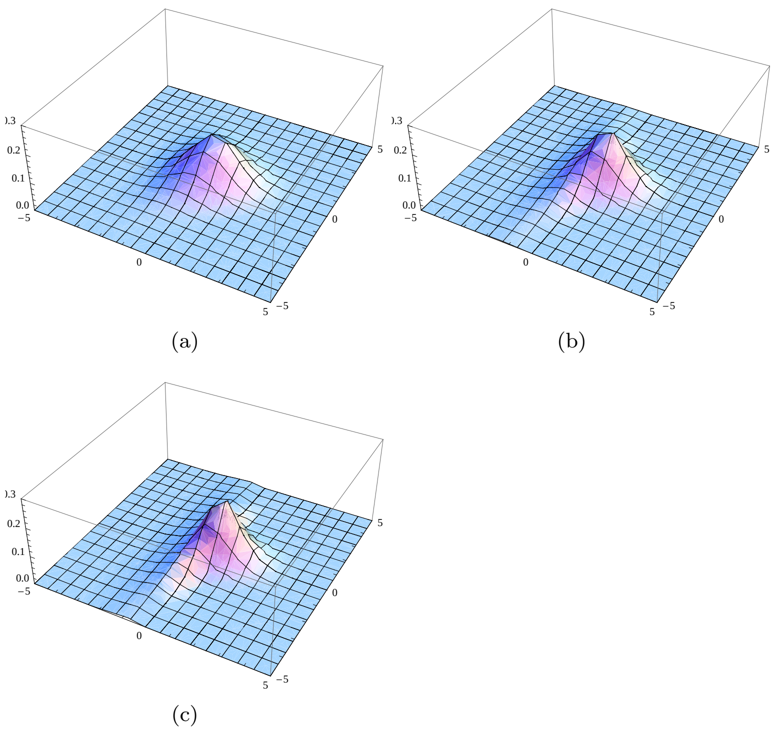

Figure 2: Plots of the $H_{3\delta}$-skewed versions of the standard bivariate normal density for $\delta = 0.5$, $\delta = 2$, and $\delta = 5$.

A. Appendix: Proofs of Theorems 2.1, 2.2, 3.1, and 3.2.

Proof of Theorem 2.1. Let $V = H_{x1}(X)$, where $H_{x1}(x_1,\ldots,x_k) = (F^{X_1}(x_1),\ldots,F^{X_k}(x_k))'$ with $F^{X_j}(.)$ the cdf of the j-th marginal X_j of $X = (X_1,\ldots,X_k)'$. Under our assumption on X, the cdf of V, namely $v \mapsto C(v) = F^X(H_{x1}^{-1}(v))$, is a continuously differentiable copula. Also, for all $j = 2,\ldots,k$, the conditional cdf of V_j given $V_1 = v_1,\ldots,V_{j-1} = v_{j-1}$ corresponds to

$$v_j \mapsto \mathrm{P}[V_j \le v_j \,|\, V_1 = v_1,\ldots,V_{j-1} = v_{j-1}] = \frac{\partial^{j-1}_{1,2,\ldots,j-1} C(\bar{v}_j)}{\partial^{j-1}_{1,2,\ldots,j-1} C(\bar{v}_{j-1})},$$

where $\bar{v}_j := (v_1,\ldots,v_j,1,\ldots,1)' \in \mathbb{R}^k$, $j = 1,\ldots,k-1$, and $\bar{v}_k = v$. It is then easy to show that

$$H_{x2}(v) = \Big(v_1, \frac{\partial_1 C(\bar{v}_2)}{\partial_1 C(\bar{v}_1)}, \frac{\partial^2_{1,2} C(\bar{v}_3)}{\partial^2_{1,2} C(\bar{v}_2)}, \ldots, \frac{\partial^{k-1}_{1,2,\ldots,k-1} C(v)}{\partial^{k-1}_{1,2,\ldots,k-1} C(\bar{v}_{k-1})}\Big)$$

is a diffeomorphism from $[0,1]^k$ to $[0,1]^k$, with the jacobian

$$|DH_{x2}(v)| = \prod_{j=2}^{k} \frac{\partial^{j}_{1,2\ldots,j} C(\bar{v}_j)}{\partial^{j-1}_{1,2,\ldots,j-1} C(\bar{v}_{j-1})} = \partial^k_{1,2,\ldots,k} C(v).$$

Now, defining $H_x := H_{x2} \circ H_{x1}$, the pdf of $H_x^{-1}(U)$, where U is uniformly distributed on $[0,1]^k$, is given by $x \mapsto f^U(H_x(x))\,|DH_x(x)| = |DH_{x2}(H_{x1}(x))||DH_{x1}(x)| = (\prod_{j=1}^k f^{X_j}(x_j))\partial^k_{1,2,\ldots,k} C(H_{x1}(x)) = f^X(x)$, so $H_x^{-1}(U) \stackrel{\mathcal{D}}{=} X$. Similarly, $H_y^{-1}(U) \stackrel{\mathcal{D}}{=} Y$ for some diffeomorphism H_y from $\mathbb{R}^k$ to $[0,1]^k$. Therefore, $H(X) \stackrel{\mathcal{D}}{=} Y$, with $H := H_y^{-1} \circ H_x$. Finally, it is easy to check that $H \in \mathcal{H}$.

As for unicity, assume that $H_a(X) \stackrel{\mathcal{D}}{=} Y \stackrel{\mathcal{D}}{=} H_b(X)$ for some $H_a, H_b \in \mathcal{H}$. Then $H(X) \stackrel{\mathcal{D}}{=} X$, where $H = H_a^{-1} \circ H_b \in \mathcal{H}$. In the first marginal, this implies that $F^{X_1}(h_1^{-1}(z)) = \mathrm{P}[h_1(X_1) \le z] = \mathrm{P}[X_1 \le z] = F^{X_1}(z)$ for all $z \in \mathbb{R}$, and hence that $H_1(x) = h_1(x_1) = x_1$ for all $x \in \mathbb{R}^k$. Now, assume that for all $j = 1,\ldots,m$ $(1 \le m \le k-1)$, we have that $H_j(x) = x_j$ for all $x \in \mathbb{R}^k$. In combination with the fact that $H(X) \stackrel{\mathcal{D}}{=} X$, this yields

$$(X_1,\ldots,X_m, h_{m+1}^{X_1,\ldots,X_m}(X_{m+1}))' \stackrel{\mathcal{D}}{=} (X_1,\ldots,X_m,X_{m+1})',$$

which in turn implies that

$$\begin{aligned}
F^{X_{m+1}|X_1=x_1,\ldots,X_m=x_m}((h_{m+1}^{x_1,\ldots,x_m})^{-1}(z)) &= \mathrm{P}[X_{m+1} \le (h_{m+1}^{x_1,\ldots,x_m})^{-1}(z) \,|\, X_1 = x_1,\ldots,X_m = x_m] \\
&= \mathrm{P}[h_{m+1}^{x_1,\ldots,x_m}(X_{m+1}) \le z \,|\, X_1 = x_1,\ldots,X_m = x_m] \\
&= \mathrm{P}[h_{m+1}^{X_1,\ldots,X_m}(X_{m+1}) \le z \,|\, X_1 = x_1,\ldots,X_m = x_m] \\
&= \mathrm{P}[X_{m+1} \le z \,|\, X_1 = x_1,\ldots,X_m = x_m] \\
&= F^{X_{m+1}|X_1=x_1,\ldots,X_m=x_m}(z),
\end{aligned}$$

for all $z \in \mathbb{R}$. Hence, we have $h_{m+1}^{x_1,\ldots,x_m}(z) = z$ for all $z \in \mathbb{R}$, or equivalently, $H_{m+1}(x) = x_{m+1}$ for all $x \in \mathbb{R}^k$. This establishes that $H_a^{-1} \circ H_b = H = I$. □

Proof of Theorem 2.2. The result in (i) is just the classical jacobian formula. Part (ii) readily follows from the unicity statement in Theorem 2.1. As for Part (iii), the symmetry of X and Part (i) entail $-H^{-1}(-X) \stackrel{\mathcal{D}}{=} -H^{-1}(X) \stackrel{\mathcal{D}}{=} -Y \stackrel{\mathcal{D}}{=} \bar{H}^{-1}(X)$, so the claim follows again from the unicity statement in Theorem 2.1 and from the fact that $x \mapsto H(-x)$ does not belong to $\mathcal{H}$. Eventually, Part (iv) is a direct consequence of (iii). □

Proof of Theorem 3.1. In view of Theorem 2.1 (i), $Y \sim f_H$ is equivalent to $Y \stackrel{\mathcal{D}}{=} H^{-1}(X)$; hence the median of f_H corresponds to the median of $H^{-1}(X)$. Since h_1 is (strictly) monotone increasing, it trivially follows that $\mathrm{P}[(H^{-1}(X))_1 \leq 0] = \mathrm{P}[X_1 \leq h_1(0)]$; clearly, the latter probability equals 1/2 iff $h_1(0) = 0$. For $j \in \{2, ..., k\}$, conditioning with respect to $X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_k$ shows that $\mathrm{Med}[(H^{-1}(X))_j] = 0$ is equivalent to

$$\int_{\mathbb{R}^{k-1}} \mathrm{P}[(H^{-1}(X))_j \leq 0 \,|\, X_1 = x_1, ..., X_{j-1} = x_{j-1}, X_{j+1} = x_{j+1}, \ldots, X_k = x_k]$$
$$f^{(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)'}(x_1, ..., x_{j-1}, x_{j+1}, \ldots, x_k) d_{1,\ldots,j-1}(x) d_{j+1,\ldots,k}(x) = \frac{1}{2}, \quad \text{(A.1)}$$

where we adopt the notation $d_{i_1,\ldots,i_2}(x) := dx_{i_1} \cdots dx_{i_2}$ for any $i_1 < i_2 \in \{1, \ldots, k\}$. Since this conditioning fixes the $j-1$ first components of X, the assumption that $h_j^{x_1,\ldots,x_{j-1}}(.)$ is a (strictly) monotone increasing function allows to rewrite the left-hand side of (A.1) as

$$\int_{\mathbb{R}^{k-1}} \int_{-\infty}^{h_j^{x_1,\ldots,x_{j-1}}(0)} f^{X_j \,|\, X_1=x_1,\ldots,X_{j-1}=x_{j-1},X_{j+1}=x_{j+1},\ldots,X_k=x_k}(x_j) dx_j$$
$$f^{(X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_k)'}(x_1, ..., x_{j-1}, x_{j+1}, \ldots, x_k) d_{1,\ldots,j-1}(x) d_{j+1,\ldots,k}(x)$$
$$= \int_{\mathbb{R}^{k-1}} \int_{-\infty}^{h_j^{x_1,\ldots,x_{j-1}}(0)} f^{(X_1,\ldots,X_k)'}(x_1, ..., x_k) dx_j d_{1,\ldots,j-1}(x) d_{j+1,\ldots,k}(x).$$

Since 0 is the median of X_j, (A.1) is then equivalent to

$$\int_{\mathbb{R}^{k-1}} \int_{0}^{h_j^{x_1,\ldots,x_{j-1}}(0)} f^{(X_1,\ldots,X_{j-1},X_j,X_{j+1},\ldots,X_k)'}(x_1, ..., x_{j-1}, x_j, x_{j+1}, \ldots, x_k) dx_j d_{1,\ldots,j-1}(x) d_{j+1,\ldots,k}(x) = 0,$$

which in turn implies that $h_j^{x_1,\ldots,x_{j-1}}(0) = 0$, as the last equality has to hold for any symmetric absolutely continuous pdf f. □

Proof of Theorem 3.2. Throughout the proof, let u belong to $\mathcal{S}^{k-1}$ and $q > 0$ be some fixed real number. Note that, for any $r \in \mathbb{R}^+$, a simple change of variables allows to rewrite $\int_{H_u^+} |u'x|^{r/q} f_H(x) dx$ as $\int_{u'H^{-1}(x)>0} |u'H^{-1}(x)|^{r/q} f(x) dx$. Hence, for the rest of this proof, all the results are established with respect to the latter expression.

Let us start with Theorem 3.2 (i). The sufficient condition states that there exist some positive constants D_1, D_2, c such that $|u'H^{-1}(x)| \leq c|u'x|^q$ for all $x \in HH_u^+(D_1) \cap H_u^+(D_2)$ and that $|u'H^{-1}(x)|$ is bounded over $HH_u^+(D_1) \setminus H_u^+(D_2)$. It trivially follows that

$$\int_{HH_u^+(D_1)\cap H_u^+(D_2)} |u'H^{-1}(x)|^{r/q} f(x) dx \leq c^{r/q} \int_{HH_u^+(D_1)\cap H_u^+(D_2)} |u'x|^r f(x) dx, \quad \text{(A.2)}$$

where the right-hand side is finite for $r < M_{u,f}$. Moreover, the boundedness of $|u'H^{-1}(x)|$ over $HH_u^+(D_1) \setminus H_u^+(D_2)$ implies that the convergence of the integral $\int_{HH_u^+(D_1)} |u'H^{-1}(x)|^{r/q} f(x) dx$, and hence also of $\int_{u'H^{-1}(x)>0} |u'H^{-1}(x)|^{r/q} f(x) dx$, is entirely determined by the left-hand side of (A.2). We easily deduce that $M_{u,f_H} \geq M_{u,f}/q$, and thus the claim holds.

Now, as regards Theorem 3.2 (ii), the existence of an additional lower bound to $|u'H^{-1}(x)|/|u'x|^q$ allows us to turn inequality (A.2) into a double inequality of the form

$$c_1^{r/q} \int_{HH_u^+(D_1)\cap H_u^+(D_2)} |u'x|^r f(x) dx \quad \leq \quad \int_{HH_u^+(D_1)\cap H_u^+(D_2)} |u'H^{-1}(x)|^{r/q} f(x) dx$$
$$\leq c_2^{r/q} \int_{HH_u^+(D_1)\cap H_u^+(D_2)} |u'x|^r f(x) dx;$$

the same reasoning as above combined with the additional boundedness condition on $|u'x|$ thus leads to the announced result.

In the univariate case, it is clear that the sufficiency conditions of Theorem 3.2 (i) and (ii) can be re-expressed in terms of limits; thus we only provide a proof for the necessity part of Theorem 3.2 (iii). Assume that $\lim_{x\to\pm\infty}|H^{-1}(x)|/|x|^q$ is infinite. This means that there exist some positive real constants C and ϵ such that $|H^{-1}(x)|>|x|^{q+\epsilon}$ if $x\in H^{+}_{\pm1}(C)$, which in terms of integrals implies that

$$\int_{H^{+}_{\pm1}(C)}|x|^{r+r\epsilon/q}f(x)dx<\int_{H^{+}_{\pm1}(C)}|H^{-1}(x)|^{r/q}f(x)dx.$$

Now, for $r=M_{\pm1}$, the left-hand integral becomes infinite, which indicates that $M_{\pm1,f_H}<M_{\pm1,f}/q$. □

Acknowledgements

The research of Christophe Ley is supported by a Mandat d'Aspirant from the Fonds National de la Recherche Scientifique, Communauté française de Belgique, whereas that of Davy Paindaveine (who is also member of ECORE, the association of CORE and ECARES) is supported by a Mandat d'Impulsion Scientifique from the Fonds National de la Recherche Scientifique, Communauté française de Belgique, and by an A.R.C. contract of the Communauté Française de Belgique. The authors thank an anonymous referee for insightful comments and suggestions that led to a significant improvement of the paper.

References

[1] Azzalini, A., 1985. A class of distributions which includes the normal ones. Scand. J. Statist. 12, 171-178.

[2] Azzalini, A., 2005. The skew-normal distribution and related multivariate families (with discussion). Scand. J. Statist. 32, 159-188.

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

[4] Fechner, G.T., 1897. Kollectivmasslehre. Engleman, Leipzig.

[5] Fernández, C., Steel, M.F.J., 1998. On Bayesian modeling of fat tails and skewness. J. Amer. Statist. Assoc. 93, 359-371.

[6] Ferreira, J.T.A.S., Steel, M.F.J., 2006. A constructive representation of univariate skewed distributions. J. Amer. Statist. Assoc. 101, 823-829.

[7] Ferreira, J.T.A.S., Steel, M.F.J., 2007. A new class of skewed multivariate distributions with applications to regression analysis. Statist. Sinica 17, 505-529.

[8] Field, C., Genton, M.G., 2006. The multivariate g-and-h distribution. Technometrics 48, 104-111.

[9] Genton, M.G., Loperfido, N., 2005. Generalized skew-elliptical distributions and their quadratic forms. Ann. Inst. Statist. Math. 57, 389-401.

[10] Gilchrist, W.G., 2000. Statistical modeling with quantile functions. Chapman and Hall/CRC.

[11] Hoaglin, D.C., 1986. Summarizing shape numerically: the g-and-h distributions. In Exploring Data Tables, Trends and Shapes, Eds. D.C. Hoaglin, F. Mosteller & J.W. Tukey, 461-513. New York: Wiley.

[12] Jones, M.C., 2004. Families of distributions arising from distributions of order statistics. Test 13, 1-43.

[13] Jones, M.C., 2006. A note on rescalings, reparametrizations and classes of distributions. J. Statist. Plann. Infer. 136, 3730-3733.

[14] Jones, M.C., Pewsey, A., 2009. Sinh-arcsinh distributions. Biometrika 96, 761-780.

[15] Ley, C., Paindaveine, D., 2009. Le Cam optimal tests for symmetry against Ferreira and Steel's general skewed distributions. J. Nonparam. Statist. 21, 943-967.

[16] Ley, C., Paindaveine, D., 2010a. On the singularity of multivariate skew-symmetric models, J. Multivariate Anal. 101, 1434-1444.

[17] Ley, C., Paindaveine, D., 2010b. Optimal tests for symmetry in independent component models: an approach based on flexible skewing mechanisms, manuscript in preparation.

[18] Mudholkar, G.S., Hutson, A.D., 2000. The epsilon-skew-normal distribution for analyzing near-normal data, J. Statist. Plann. Infer. 83, 291-309.

[19] Wang, J., Boyer, J., Genton, M.G., 2004. A skew-symmetric representation of multivariate distribution. Statist. Sinica 14, 1259-1270.