



Research Article

Open Access

Eckhard Liebscher* and Wolf-Dieter Richter

Modelling with star-shaped distributions

<https://doi.org/10.1515/demo-2020-0003>

Received November 19, 2019; accepted February 11, 2020

Abstract: We prove and describe in great detail a general method for constructing a wide range of multivariate probability density functions. We introduce probabilistic models for a large variety of clouds of multivariate data points. In the present paper, the focus is on star-shaped distributions of an arbitrary dimension, where in case of spherical distributions dependence is modeled by a non-Gaussian density generating function.

Keywords: convex data cloud, radially concave data cloud, star contoured data cloud, data cloud oriented model, norm sphere, antinorm sphere, star sphere, global shape approximation, locally refined shape approximation, direction dependent model refinement, star-shaped density, estimation, moments, simulation, star generalized trigonometric functions, star generalized coordinates, multivariate p -generalized ellipsoidal coordinates

MSC: 60E05, 62E17

1 Introduction

The components of a random vector following a spherical or even elliptically contoured distribution are only independent if the density generating function is a Gaussian one. In order to generalize the class of elliptically contoured distributions, Fernandez et al. [4] introduced the star-shaped distributions. Inference problems in linear models with star-shaped mixtures of Gaussian errors were considered by Jensen [5]. In the paper by Kamiya et al. [6], star-shaped distributions are studied within the framework of orbital decomposition and global cross-sections. Later Balkema et al. [2] examined the limit star shape of scaled sample clouds and the related distributions. Yang and Kotz [23] describe an alternative approach by considering center similar distributions. Recently, Kamiya [7] investigated the estimation of the shape of density level sets from star-shaped distributions. In the last decade, star-shaped distributions have become popular because of their flexibility in shape. Deviating from the shape of ellipses/ellipsoids, a wide range of shapes of the density contours are possible among the star-shaped distributions.

In scatterplots of real datasets one can frequently see that the shapes of the contour lines/surfaces are not ellipses/ellipsoids. This is a reason for looking for generalizations. Here star-shaped distributions come into focus. The detailed theory of star-shaped distributions including stochastic and geometric representations is developed by Richter in [16]. For understanding normalizing constants of density generating functions as ball numbers (even in more general cases), we refer to [21]. Semiparametric and parametric estimation methods for density generators, generalized radius distributions and the star-shaped densities are examined in the paper [11]. The latter three papers contain references to a lot of other papers on star-shaped distributions and ball numbers. The star-shaped distributions represent a rather general and flexible class of distributions including convex as well as non-convex shapes of contours. Thus this class is appropriate for describing the

*Corresponding Author: Eckhard Liebscher: University of Applied Sciences Merseburg, Department of Engineering and Natural Sciences, 06217 Merseburg, Germany, E-mail: eckhard.liebscher@hs-merseburg.de

Wolf-Dieter Richter: University of Rostock, Institute of Mathematics, Ulmenstr. 69, Haus 3, 18057 Rostock, Germany E-mail: wolf-dieter.richter@uni-rostock.de

underlying distribution of a large variety of datasets or data clouds. However, flexible parametric classes involve a lot of parameters which in turn carries the risk of overfitting.

The main goal of this paper is to introduce useful specific model classes and to examine their use in the statistical framework. In Section 2 of this paper, the continuous star-shaped distributions are introduced and their basic properties are studied. Here we use rather general coordinates for establishing models for the function defining the contour (level sets). We derive formulas for the first two moments. In establishing proper model classes, identifiability is a very important issue. If identifiability is not given, we cannot expect to get consistent estimators in the framework of statistical inference. In Section 2.3 we provide a sufficient condition for the identifiability. The definitions of model classes for the generating function are presented in Section 3. In Section 4 we introduce a lot of models for the Minkowski functional which determines the shape of the level sets of the density in the two-dimensional case. Section 5 provides such models in the higher-dimensional case. Section 6 is dedicated to the simulation of random vectors with a specified star-shaped distribution. For establishing the simulation procedure we use a representation of the distribution by a generalized radius and spherical coordinates given in Section 2. The maximum likelihood method for the estimation of the parameters is briefly discussed in Section 7. In Section 8 the reader finds a discussion on model checks. Real data examples and its fitted distributions are presented in Section 9. The proofs of the statements can be found in Section 10.

2 Continuous star-shaped distributions

2.1 Introduction and general properties

We assume that $K \subset \mathbb{R}^d$ is a star body (a bounded set with the property $x \in K \Rightarrow \lambda x \in K$ for $0 \leq \lambda \leq 1$) having the origin in its interior. The Minkowski functional of K ,

$$h_K(x) = \inf\{\lambda > 0 : x \in \lambda K\},$$

is defined for every K under consideration and may particularly be any norm or antinorm. Although, for mathematical reasons discussed for a particular case in [19], and more generally in [22], it is not trivial to further assume that the Minkowski functional h_K of K is positively homogeneous of degree one. This restriction made here might be considered not to be too restrictive in many applied situations. A function $f : \mathbb{R}^d \rightarrow [0, \infty)$ is called positively homogeneous of degree k if $f(\lambda x) = \lambda^k f(x)$ holds for $x \in \mathbb{R}^d, \lambda > 0$. The set $K(r) = rK$ and its boundary $S(r) = rS$ are called the star ball and star sphere of star radius $r > 0$, respectively. Notice that $S = \{x : h_K(x) = 1\}$. This star sphere S corresponds to the shape in [7]. Star balls may thus be convex or radially concave subsets of the sample space, for example p -generalized ellipsoids with $p \geq 1$ or $0 < p \leq 1$, respectively.

Let a function $g : [0, \infty) \rightarrow [0, \infty)$ satisfy $0 < I(d, g) < \infty$, where $I(d, g) = \int_0^\infty r^{d-1} g(r) dr$. Such a function is called a density generating function,

$$\phi_{g,K}(x) = C(g, K)g(h_K(x)), \quad x \in \mathbb{R}^d \quad (1)$$

a star-shaped density of a random vector $\mathbf{X} = (X^{(1)}, \dots, X^{(d)})$. The star body K defines the contour. It is thus adapted to the shape of the data cloud. The corresponding probability measure is denoted $\Phi_{g,K}$ and the normalizing constant allows the representation

$$C(g, K) = \frac{1}{\mathfrak{D}_S(S)I(d, g)},$$

where $\mathfrak{D}_S(S)$ means the star-generalized surface content of S , see [16]. If the additional assumption $C(g, K) = 1$ is satisfied then g is called a density generator. A random vector U following the distribution

$$\omega_S(A) = \frac{\mathfrak{D}_S(A)}{\mathfrak{D}_S(S)}, \quad A \in \mathfrak{B}^d \cap S$$

is said to be star-uniformly distributed. The geometric measure representation of a star-shaped distribution law reads

$$\Phi_{g,K}(B) = \frac{1}{I(d,g)} \int_0^\infty r^{d-1} g(r) \mathfrak{F}_S(B, r) dr, \quad B \in \mathfrak{B}^d.$$

Here \mathfrak{B}^d is the σ -algebra of Borel sets of \mathbb{R}^d , and

$$\mathfrak{F}_S(B, r) = \frac{\mathfrak{O}_S([\frac{1}{r}B] \cap S)}{\mathfrak{O}_S(S)} = \omega_S([\frac{1}{r}B] \cap S)$$

is the star sphere intersection proportion function of the set B . The random variable $R = h_K(\mathbf{X})$ is called the star radius of the observation vector \mathbf{X} . This random vector \mathbf{X} satisfies the stochastic representation

$$\mathbf{X} \stackrel{d}{=} R \cdot U \quad (2)$$

with R and U being independent (symbol $\stackrel{d}{=}$ means that the random variables on both sides have the same distribution law). Let, moreover, $\|\cdot\|$ denote any norm or antinorm in \mathbb{R}^d and $B = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ and its boundary \mathcal{S}_B the corresponding unit ball and sphere, respectively. Because of the homogeneity property of h_K , $\phi_{g,K}$ allows the representation

$$\phi_{g,K}(x) = C(g, K) g(\|x\| h_K(\frac{x}{\|x\|})), \quad x \in \mathbb{R}^d, \quad (3)$$

where the point $\frac{x}{\|x\|}$ belongs to the norm or antinorm unit sphere \mathcal{S}_B . One may imagine that there is a first idea for describing a shape in the data cloud by the level sets of the functional $\|\cdot\|$ and afterwards there appears the wish to correct or modify this functional by a suitable direction dependent or locally acting factor $h_K(x/\|x\|)$. For a moment, let us consider the case of the Euclidean norm $\|\cdot\| = \|\cdot\|_2$. In this case, Yang and Kotz [23] studied the class of distributions with density (3), where our function h_K corresponds to b^{-1} in their paper (b is the so-called bound function), and our function g corresponds to function G . Here G is the antiderivative of $r \rightsquigarrow -g(r) \cdot r^{-d}$, where $G(\infty) = 0$ and g is defined as in Lemma 2.3 of Yang and Kotz's paper. Yang and Kotz call this class center-similar distributions and emphasize that $\mathbf{X}/\|\mathbf{X}\|_2$ does not follow a uniform but an arbitrary distribution.

In the next step we incorporate location and scale in the formula for the density. To ensure identifiability, we assume in the following that

$$I(d, g) = \int_0^\infty r^{d-1} g(r) dr = 1. \quad (4)$$

Then $\kappa^{-1} = \mathfrak{O}_S(S)^{-1}$ is the normalizing constant of the density. A tractable formula for κ is given below.

We consider a random vector \mathbf{X} having the density $\phi_{g,h_K,\mu,\Sigma}$

$$\phi_{g,h,\mu,\Sigma}(x) = (\kappa \det(\Sigma))^{-1} g(h_K(\Sigma^{-1}(x - \mu))) \text{ for } x \in \mathbb{R}^d. \quad (5)$$

This distribution of \mathbf{X} is referred to as a *continuous star-shaped distribution*. In this formula function $h = h_K$ determines the contour of the density. $\mu \in \mathbb{R}^d$ is the location parameter. The diagonal matrix $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$ with $\sigma_i > 0$ for $i = 1, \dots, d$ consists of scale parameters of the distribution. κ is a suitable constant.

In view of (2), \mathbf{X} also allows the representation

$$\mathbf{X} \stackrel{d}{=} \mu + R \cdot \Sigma U. \quad (6)$$

For any $b \in \mathbb{R}^d$ and any diagonal matrix $A \in \mathbb{R}^{d \times d}$ with positive diagonal entries, it can be easily shown that $A\mathbf{X} + b$ has the density $\phi_{g,h,A\mu+b,\Sigma A}$. The density $\phi_{g,h,0,I}$ is called standard star-shaped.

Next we develop a representation of h_K using general coordinates $r, \psi_1, \dots, \psi_{d-1}$, where $r \in [0, \infty)$, $(\psi_1, \dots, \psi_{d-1}) \in A_d \subset \mathbb{R}^{d-1}$. Later models for h will be based on this representation. Let $\|\cdot\|$ be a given

norm or antinorm. Further let a transformation from certain d -dimensional coordinates $r, \psi_1, \dots, \psi_{d-1}$ to the Cartesian ones x_j be defined by functions $I_1, \dots, I_d : A_d \rightarrow \mathbb{R}$ as

$$x_j = r I_j(\psi_1, \dots, \psi_{d-1}) \quad (7)$$

for $j = 1, \dots, d$, where $r = \|x\|$. Assume that there is a bijective mapping $T : [0, \infty) \times A_d \rightarrow M$, $M \in \mathfrak{B}^d$, such that $(r, \psi_1, \dots, \psi_{d-1}) \rightsquigarrow (r I_1(\psi_1, \dots, \psi_{d-1}), \dots, r I_d(\psi_1, \dots, \psi_{d-1}))^T$ and $\mathbb{R}^d \setminus M$ has measure zero. Define $\tilde{T} : M \rightarrow A_d$ such that $x \rightsquigarrow T^{-1}(\|x\|^{-1} x)$ and T^{-1} is the inverse mapping to T . Hence $\tilde{T}(x) = \psi = (\psi_1, \dots, \psi_{d-1})^T$. We introduce the Jacobian of $I_1 \dots I_d$ by J . The Jacobian determinant of the transformation is given by

$$\Delta(r, \psi_1, \dots, \psi_{d-1}) = \left| \begin{array}{c} I_1 \\ I_2 \\ \vdots \\ I_d \end{array} \left[r J \right] \right| = r^{d-1} \bar{\Delta}(\psi_1, \dots, \psi_{d-1}) \quad (8)$$

such that $dx = \Delta(r, \psi_1, \dots, \psi_{d-1}) dr d\psi_1 \dots d\psi_{d-1}$, $\bar{\Delta} = \Delta r^{1-d}$. Now we provide two important examples of coordinates.

EXAMPLE 1: *Spherical coordinates* $(r, \alpha_1, \dots, \alpha_{d-1})$: $r = \|x\|_2$,

$$I_j(\alpha_1, \dots, \alpha_{d-1}) = \cos \alpha_j \prod_{k=1}^{j-1} \sin \alpha_k$$

for $j = 1, \dots, d$, where $\alpha_d = 0$, $\alpha = (\alpha_1, \dots, \alpha_{d-1}) \in (0, \pi)^{d-2} \times [0, 2\pi) =: A_d$, $r \geq 0$, $M := \{x \in \mathbb{R}^d : (x_{d-1}, x_d) \neq (0, 0)\}$. If the upper limit of the product is smaller than the lower one, then we define the product to be 1. Then the mapping \tilde{T} assigns to vector x the corresponding spherical angles $\alpha_1, \dots, \alpha_{d-1}$ from A_d . The following Jacobian determinant is well-known:

$$\Delta(r, \alpha_1, \dots, \alpha_{d-1}) = r^{d-1} \prod_{k=1}^{d-2} \sin^{d-k-1} \alpha_k.$$

EXAMPLE 2: *spherical l_p -coordinates* $(r, \alpha_1, \dots, \alpha_{d-1})$ for $p \in (0, \infty)$: These coordinates are introduced in [13]. Define $N_p(\beta) = (|\sin \beta|^p + |\cos \beta|^p)^{1/p}$, and

$$\sin_p(\beta) = \frac{\sin \beta}{N_p(\beta)}, \quad \cos_p(\beta) = \frac{\cos \beta}{N_p(\beta)}.$$

The transformation from these coordinates $r, \alpha_1, \dots, \alpha_{d-1}$ to the Cartesian ones x_j is given by (7) and

$$r = \|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{1/p}, \quad I_j(\alpha_1, \dots, \alpha_{d-1}) = \cos_p(\alpha_j) \prod_{k=1}^{j-1} \sin_p(\alpha_k)$$

for $j = 1 \dots d$, where $\alpha_d = 0$, $\alpha = (\alpha_1, \dots, \alpha_{d-1}) \in (0, \pi)^{d-2} \times [0, 2\pi) =: A_d$, $r \geq 0$. According to Theorem 2 of [13], we can provide the formula for the Jacobian determinant:

$$\Delta(r, \alpha_1, \dots, \alpha_{d-1}) = r^{d-1} \prod_{k=1}^{d-2} \sin^{d-k-1} \alpha_k N_p(\alpha_k)^{k-d-1}. \quad \square$$

The interested reader finds even more general coordinates in papers from the references at the end of this note.

In many particular cases of these coordinates (including Examples 1 and 2), $\psi_1, \dots, \psi_{d-1}$ play the role of generalized spherical angles. We introduce a function $H : A_d \rightarrow [0, \infty)$ such that

$$H(\psi) = h_K(T((1, \psi))) \quad (9)$$

for $\psi \in A_d$. Function h_K can then be written as

$$h_K(x) = \|x\| H(\tilde{T}(x)) = rH(\psi) \quad (10)$$

for $x \in M$, and corresponding $r > 0$, $\psi \in A_d$. The idea behind this approach is to separate the dependence of the density on r and the ψ -coordinates. Here h_K can be written as a product of r and a function H of ψ , and $H(\tilde{T}(\cdot))$ is a positively homogeneous function of degree zero.

The following lemma deals with the computation of probabilities and the constant κ :

Lemma 2.1. *Let \mathbf{X} be a random vector distributed as a continuous star-shaped distribution with the density function (5). We introduce $\bar{A} = \{x = T(r, \psi) : \psi = (\psi_1, \dots, \psi_{d-1})^T \in [\underline{\psi}_1, \bar{\psi}_1] \times \dots \times [\underline{\psi}_{d-1}, \bar{\psi}_{d-1}], r \in [r/H(\psi), \bar{r}/H(\psi)]\}$, and $\Sigma\bar{A} = \{\Sigma x : x \in \bar{A}\}$. Then*

$$\mathbb{P}\{\mathbf{X} \in \mu + \Sigma\bar{A}\} = \kappa^{-1} \int_{\underline{\psi}_{d-1}}^{\bar{\psi}_{d-1}} \dots \int_{\underline{\psi}_1}^{\bar{\psi}_1} H(\psi)^{-d} \bar{\Delta}(\psi) d\psi \int_{\underline{r}}^{\bar{r}} g(r) r^{d-1} dr.$$

Moreover

$$\kappa = \int_{A_d} H(\psi)^{-d} \bar{\Delta}(\psi) d\psi.$$

In this lemma, the sets \bar{A} are slices of \mathbb{R}^d with rectangular range of the coordinates ψ_i , the radius depending on the ψ -coordinates in a special way. Now we have the following statement on the first two moments (\mathbb{V} is the symbol for the variance):

Lemma 2.2. *Let \mathbf{X} be a random vector distributed as a continuous star-shaped distribution with the density function (5). Assume that $G_{d+1} < +\infty$, where $G_j = \int_0^\infty g(q) q^j dq$. Then*

$$\begin{aligned} \mathbb{E}(\mathbf{X}) &= \mu + \Sigma \tilde{\mu}, \quad \tilde{\mu}_j = \kappa^{-1} G_d \int_{A_d} I_j(\psi) \bar{\Delta}(\psi) H(\psi)^{-d-1} d\psi, \\ \mathbb{V}(X^{(j)}) &= \sigma_j^2 \left(\kappa^{-1} G_{d+1} \int_{A_d} I_j^2(\psi) \bar{\Delta}(\psi) H(\psi)^{-d-2} d\psi - \tilde{\mu}_j^2 \right) \end{aligned}$$

for $j = 1 \dots d$.

2.2 Random radius and ψ -coordinates

Remember that the star-generalized radius can be computed by $R = h_K(\Sigma^{-1}(\mathbf{X} - \mu))$. Define $\Psi = \tilde{T}(\Sigma^{-1}(\mathbf{X} - \mu)) \in A_d$. Vector Ψ consists of the ψ -coordinates (angles in the spherical coordinates case) representation of $\Sigma^{-1}(\mathbf{X} - \mu)$. Lemma 2.3 gives the joint distribution of R and Ψ .

Lemma 2.3. *The random variables R and Ψ are independent with densities*

$$f(r) = g(r) r^{d-1}, \quad (11)$$

$$\phi_\Psi(\psi) = \kappa^{-1} H(\psi)^{-d} \bar{\Delta}(\psi). \quad (12)$$

This lemma shows that the distribution of \mathbf{X} can be represented by two independent components. Moreover, the generating function g is directly related to the density of R .

In the following we discuss the identifiability of star-shaped distribution classes.

2.3 Identifiability

Identifiability of model classes for distributions ensures that two different sets of parameters lead to different probability measures of the random vector \mathbf{X} . We pose assumptions which turn out to be crucial for the verification of identifiability.

ASSUMPTION \mathcal{A}_g : \mathcal{G} is the set of right continuous functions $g : [0, \infty) \rightarrow [0, +\infty)$ satisfying (4) and the following two conditions:

(i) For $g_1, g_2 \in \mathcal{G}$, $g_1 \equiv g_2$, there is no $\gamma > 0$ such that

$$g_2(z) = \gamma^d g_1(\gamma z) \text{ for } z > 0.$$

(ii) There are $z_1, z_2 > 0$, $z_1 \neq z_2$ such that $g(z_1), g(z_2), g(0)$ are different values. \square

ASSUMPTION \mathcal{A}_h : Let \mathcal{H} be a set of positively homogeneous functions $\mathbb{R}^d \rightarrow [0, +\infty)$ being symmetric with respect to the origin such that for two different functions $h_1, h_2 \in \mathcal{H}$, there is no diagonal matrix $\tilde{\Sigma}$ satisfying $h_1(x) = h_2(\tilde{\Sigma}x)$ for every $x \in \mathbb{R}^d$. \square

Roughly speaking, it is excluded in Assumptions \mathcal{A}_g and \mathcal{A}_h that scale transformations of $g \in \mathcal{G}$ lead to another element of \mathcal{G} , and scale transformations of $h \in \mathcal{H}$ lead to another element of \mathcal{H} . The special form of scale condition (i) is reasoned by the fact that g_1, g_2 have to fulfil (4). The next Theorem 2.4 provides the identifiability result.

Theorem 2.4. Assume that Assumptions \mathcal{A}_g and \mathcal{A}_h are satisfied. Then the model class $\{\varphi_{g,h,\mu,\Sigma} : g \in \mathcal{G}, h \in \mathcal{H}, \mu \in \mathbb{R}^d, \sigma \in (0, \infty)^d\}$ is identifiable.

Now the task is to check carefully the conditions of Theorem 2.4 in the case of specific model classes.

3 Modelling the density generating function

Certain aspects of dependence modeling on using p -generalized non-Gaussian density generating functions for jointly $l_{n,p}$ -symmetrically distributed random variables are studied in [12]. We consider here a model family $\{g_\theta : \theta \in \Theta_1\}$ of generating functions satisfying \mathcal{A}_g . $\Theta_1 \subset \mathbb{R}^{d_1}$ is the corresponding parameter space. The model leads to a density with finite local maximum at zero if

$$\lim_{r \rightarrow 0+0} g(r) < +\infty \text{ and either } \lim_{r \rightarrow 0+0} g'(r) < 0 \text{ or } \lim_{r \rightarrow 0+0} g'(r) = 0 \wedge \lim_{r \rightarrow 0+0} g''(r) < 0. \quad (13)$$

For many applications, it is beneficial to have this property. Let us introduce three models for g_θ .

(1) KOTZ-TYPE DISTRIBUTION: parameter $\theta = (s, t) \in \Theta_1 = (-d, +\infty) \times (0, +\infty)$,

$$g_{(s,t)}(r) = \frac{t}{\Gamma(\frac{s+d}{t})} r^s e^{-r^t} \text{ for } r > 0.$$

Condition (13) is fulfilled for $s = 0$. The generalized radius R has a generalized Gamma distribution. In the case $s = 1 - d$, R has a Weibull distribution.

(2) MODIFIED EXPONENTIAL MODEL. $\theta = (a, b) \in \Theta_1 = \{(a, b) : b > 0, a > -2\sqrt{b}\}$,

$$g_{(a,b)}(r) = \frac{1 + ar + br^2}{(d-1)! + a d! + b(d+1)!} e^{-r} \text{ for } r > 0.$$

Condition (13) is satisfied provided that $(a = 1 \wedge b \leq \frac{1}{2}) \vee (a < 1)$. The distribution of R is a mixture of Erlang distributions.

(3) MODIFIED PEARSON VII MODEL: parameter $\theta = (s, m) \in \Theta_1 = \{(s, m) : s > -d, m > (s+d)/2\}$,

$$g_{(s,m)}(r) = \frac{2\Gamma(m)r^s}{\Gamma((s+d)/2)\Gamma(m-(s+d)/2)(1+r^2)^m} \text{ for } r > 0.$$

Condition (13) is fulfilled for $s = 0$. In the case $s = 1 - d$, R has a scale-transformed one-sided student- t -distribution. \square

These model classes fulfil Assumption \mathcal{A}_g . Notice that a scaling parameter cannot be incorporated for the reason explained in the previous section.

4 Modelling function H in the two-dimensional case

4.1 Using Euclidean norm

Let $d = 2$ and $\varphi(x, y) \in [0, 2\pi)$ be the angle in radians between the positive x -axis and the line from the point (x, y) to the origin: $\varphi(x, y) = \arctan(y/x)$ for $x > 0$, $y \geq 0$, $\varphi(x, y) = \arctan(y/x) + \pi$ for $x < 0$, $\varphi(x, y) = \arctan(y/x) + 2\pi$ for $x > 0$ and $y < 0$, $\varphi(0, y) = \frac{\pi}{2} \cdot (2 - \operatorname{sgn}(y))$ for $y \neq 0$. In the following we propose model classes $\{H_\eta : \eta \in \Theta_2\}$ for the function H introduced in (9). Here $\Theta_2 \subset \mathbb{R}^{q_2}$ is the parameter space of H . In view of (10), we have

$$h_\eta(x) = \|x\|_2 H_\eta(\tilde{T}(x)) = \|x\|_2 H_\eta(\varphi(x_1, x_2)), \quad (14)$$

for $x = (x_1, x_2)^T \in \mathbb{R}^2 \setminus (0, 0)^T$ and

$$h_\eta((\cos \varphi, \sin \varphi)^T) = H_\eta(\varphi)$$

for $\eta \in \Theta_2$, $\varphi \in [0, 2\pi)$. Here $H_\eta(\varphi(\cdot))$ is positively homogeneous of degree zero. This leads to the density

$$\phi_{g,h,\mu,\Sigma}(x) = (\kappa \det(\Sigma))^{-1} g\left(\left\|\Sigma^{-1}(x - \mu)\right\|_2 H_\eta(\varphi(\Sigma^{-1}(x - \mu)))\right) \text{ for } x \in \mathbb{R}^d.$$

We state the important principle for Section 4:

Construction principle: Assume that for empirical or theoretical reasons it seems that, globally viewed, the sample cloud reflects the shape of the sphere \mathbb{S}_B or some shape being close to it. Choose then a direction dependent or locally acting function $\varphi \rightsquigarrow H_\eta(\varphi)$ for modelling deviations from this shape and with $H_\eta(\varphi(x))$ being homogeneous of degree zero, such that the complete shape is best in some sense approximated by the level sets of the function (14).

The following model classes are introduced according to this principle. Now we introduce two model classes for H_η satisfying the following Assumption \mathcal{H}_2 :

ASSUMPTION \mathcal{H}_2 : $H_\eta : [0, 2\pi) \rightarrow (0, \infty)$ is a bounded, continuously differentiable function,

$$\lim_{t \rightarrow 2\pi-0} H_\eta(t) = H_\eta(0), \quad H_\eta(\pi + \varphi) = H_\eta(\varphi) \text{ for } \varphi \in [0, \pi),$$

$$\inf_{\varphi \in [0, 2\pi)} H_\eta(\varphi) > 0. \quad \square$$

Here the constant κ can be evaluated by using the formula

$$\kappa = 2 \int_0^\pi H_\eta(\varphi)^{-2} d\varphi.$$

A rather simple class of model functions H_η is given by

MODEL CLASS 1:

$$H_\eta(\varphi) = \left(\frac{1 + a \sin^2(\varphi - \beta)}{1 + b \sin^2(\varphi - \beta)} \right)^c \text{ for } \varphi \in [0, 2\pi),$$

where $\eta = (a, b, c, \beta)^T$, $a \in (-1, \infty)$, $b \in (-1, \infty)$, $a \neq b$, $\beta \in [0, \pi)$ are the parameters of this function. \square

Changing the parameter β induces a rotation of the contour curves. Changing a , b , c leads to another shape of the contour lines. In the case $c = 1$, we have

$$\kappa = \frac{a^3 + b^2 \left(-2 - 3a + 2(a+1)^{3/2} \right) + 2a^2(b+1)}{a^2(a+1)^{3/2}} \pi.$$

Model class 1 includes elliptical contours for $c = \frac{1}{2}$, $a = A^2 B^{-2} - 1$, $b = 0$, where A, B are the lengths of the semi axes of the ellipse. In this case the resulting density ϕ is an elliptically contoured one. The resulting function h of model class 1 is given by

$$h_{\eta}(x) = \sqrt{x_1^2 + x_2^2} \left(\frac{x_1^2 (1 + (a+1) \tan^2 \beta) - 2ax_1x_2 \tan \beta + x_2^2 (\tan^2 \beta + a + 1)}{x_1^2 (1 + (b+1) \tan^2 \beta) - 2bx_1x_2 \tan \beta + x_2^2 (\tan^2 \beta + b + 1)} \right)^c \quad (15)$$

To illustrate the shape of the star-shaped density, we depict contour curves (level sets) of the density which are curves $\{(r, \varphi) : r = C_0 H(\varphi)^{-1}, \varphi \in [0, 2\pi)\}$ in polar coordinates with certain values of $C_0 > 0$. Figures 1 and 2 show the variety of shapes of contours within model class 1. In the figures the curves are coloured in blue, orange and green in this order. In Figure 2 left the contour of the density is depicted in a special case to give an impression of the resulting contours.

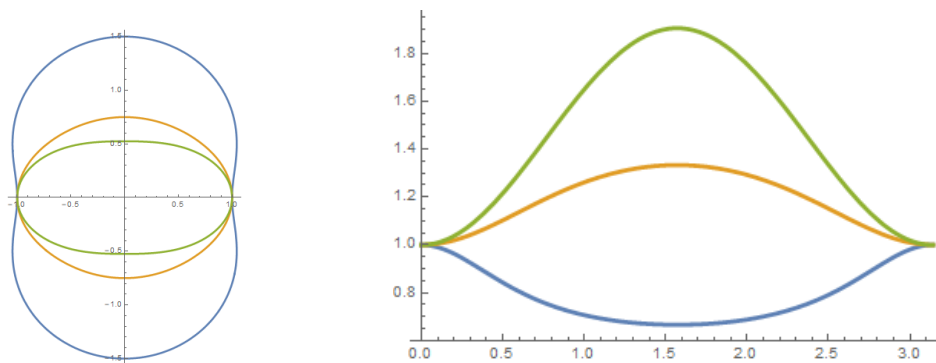


Figure 1: model class 1. contour curves of the density and H as a function of φ in the cases $a = 1$, $b = 2, 0.5, 0.05$, $c = 1$, $\beta = 0$

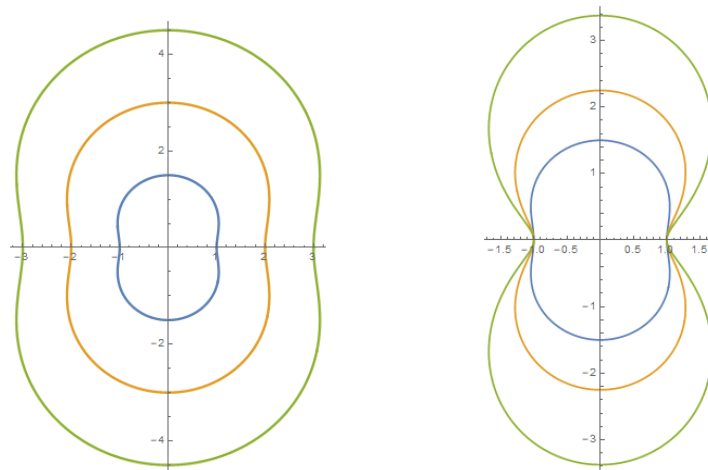


Figure 2: model class 1. left: several level sets of the density in the case $a = 1$, $b = 2$, $c = 1$, $\beta = 0$. right: contour curves of the density in the cases $a = 1$, $b = 2$, $c = 1, 2, 3$, $\beta = 0$

A further possibility is to define function H_{η} as a polynomial of $\sin^2(\varphi - \beta)$. This is done in the next definition.

MODEL CLASS 2:

$$H_{\eta}(\varphi) = \left(1 + a \sin^2(\varphi - \beta) + b \sin^4(\varphi - \beta) \right)^c \text{ for } \varphi \in [0, 2\pi),$$

$\eta = (a, b, c, \beta)^T$. The parameters $a, b, c \in \mathbb{R}, b \neq 0, \beta \in [0, \pi)$ fulfil the condition: $\inf_{x \in [0,1]} (1 + ax + bx^2) > 0$ which is equivalent to $(-1 - a < b) \wedge ((\frac{a}{2b} \geq 0) \vee (\frac{a}{2b} \leq -1) \vee (1 - \frac{a^2}{4b} > 0))$. \square

Figure 3 show the contour curves of the model class 2 in order to illustrate the variety of shapes in this class.

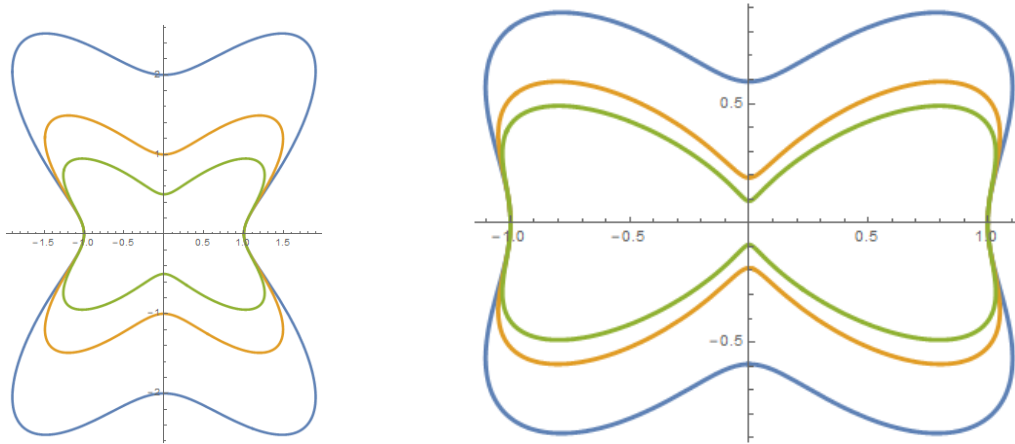


Figure 3: model class 2. left: contour curves of the density in the cases $a = -2, b = 1.5, 2, 3, c = 1, \beta = 0$. right: contour curves in the cases $a = -0.7, b = 1, 2, 3, c = 2, \beta = 0$

From the definition formulas, a statement on the identifiability can be derived immediately.

Proposition 4.1. *With the restriction $c \neq \frac{1}{2} \vee (c = \frac{1}{2} \wedge \beta \neq 0)$, Model class 1 fulfils Assumption \mathcal{A}_h . Model class 2 satisfies Assumption \mathcal{A}_h .*

4.2 Using a general norm

For further dealing with representation (1) we shall make use of suitable coordinates, especially for describing B and \mathcal{S}_B . It turns out that easiest use of suitable coordinates is in dimension two, see [17]. Let, according to [14] and [18], the $\|\cdot\|$ -norm or antinorm or star-generalized polar coordinate transformation $Pol_{\mathcal{S}_B} : [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2$ be defined by $x_1 = r \cos_{\mathcal{S}_B}(\varphi), x_2 = r \sin_{\mathcal{S}_B}(\varphi)$ for $r \in [0, \infty), \varphi \in [0, 2\pi)$:

$$\cos_{\mathcal{S}_B}(\varphi) = \frac{\cos(\varphi)}{\|(\cos \varphi, \sin \varphi)^T\|}, \quad \sin_{\mathcal{S}_B}(\varphi) = \frac{\sin(\varphi)}{\|(\cos \varphi, \sin \varphi)^T\|}. \quad (16)$$

The inverse of this transformation is defined by

$$r = r(x_1, x_2) = \|(x_1, x_2)^T\| \text{ and } \varphi = \varphi(x_1, x_2) = \varphi((x_1, x_2)^T). \quad (17)$$

Here $\varphi(\cdot, \cdot)$ is defined as in Section 4.1. Because of the relation $Pol_{\mathcal{S}_B}^{-1}(\frac{x}{\|x\|}) = (1, \varphi(x))$ and the definition $H_\eta(\varphi) = h_K(Pol_{\mathcal{S}_B}(1, \varphi))$, the following general representation formula for star-shaped densities is proved. For arbitrary norm or antinorm $\|\cdot\|$, the star-shaped density allows the representation

$$\phi_{g,h,0,I}(x) = \kappa^{-1} g(\|x\| H_\eta(\varphi(x))), \quad x \in \mathbb{R}^2,$$

where the function $x \rightsquigarrow H(\varphi(x))$ is homogeneous of degree zero. We recall that the argument of function g is assumed to be positively homogeneous of degree one.

Let us consider the following model class for H .

MODEL CLASS 3: Function H is given by

$$H_\eta(\varphi) = (|\cos \varphi|^p + |\sin \varphi|^p)^{1/p} = \left\| \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right\|_p, \quad \varphi \in [0, 2\pi), \quad (18)$$

where $\eta = p > 0$ and $\|x\|_p$ denotes the $l_{2,p}$ -norm or antinorm of x if respectively $p \geq 1$ or $0 < p < 1$. \square

Using function H_η in (18), it may be checked immediately that the function

$$H_\eta(\varphi(x)) = \frac{\|x\|_p}{\|x\|_2}, \quad x \in \mathbb{R}^2 \quad (19)$$

is positively homogeneous of degree zero. If the shape of the sample cloud deviates from the contour or the level sets of the functional $\|\cdot\|$ in a similar manner as the contour or level sets of the norm $\|\cdot\|_p$ deviates from that of the functional $\|\cdot\|_2$, then the (standard) probability density function

$$\phi_{g,h,0,I}(x) = \kappa^{-1} g(\|x\| \frac{\|x\|_p}{\|x\|_2}), \quad x \in \mathbb{R}^2 \quad (20)$$

could be suitable to (globally and locally) model the data. This is what one might call data cloud oriented modelling. In the present case, the shape defining star body K has Minkowski functional

$$h_K(x) = \|x\| \frac{\|x\|_p}{\|x\|_2}, \quad x \in \mathbb{R}^2. \quad (21)$$

In the particular case that $\|\cdot\| = \|\cdot\|_2$, this means

$$\phi_{g,h,0,I}(x) = \kappa^{-1} g(\|x\|_p), \quad x \in \mathbb{R}^2. \quad (22)$$

In other words, in dependence of the direction determined by angle φ , function H_η in (18) describes the deviation of the $l_{2,p}$ -unit sphere (that is the $l_{2,p}$ -unit circle) from the Euclidean unit sphere.

It may seem not to be trivial to explicitly describe the star uniform distribution on S in this way, in general. Nevertheless, modelling data clouds by primary (or global) approximation with norm $\|\cdot\|$ and secondary (or directionally correcting or local) approximation with function H_η may be successful. Clearly, if one just starts from an ansatz like (21) then, vice versa, one can derive the coordinate representation of H_η in (18) for the purposes of statistical analysis, see Sections 7 and 9.

The following Figure 4 shows the level sets of the function $x \rightsquigarrow H_\eta(\varphi(x))$ and the Minkowski functional h_K for different values of p and norms $\|\cdot\|$, respectively.

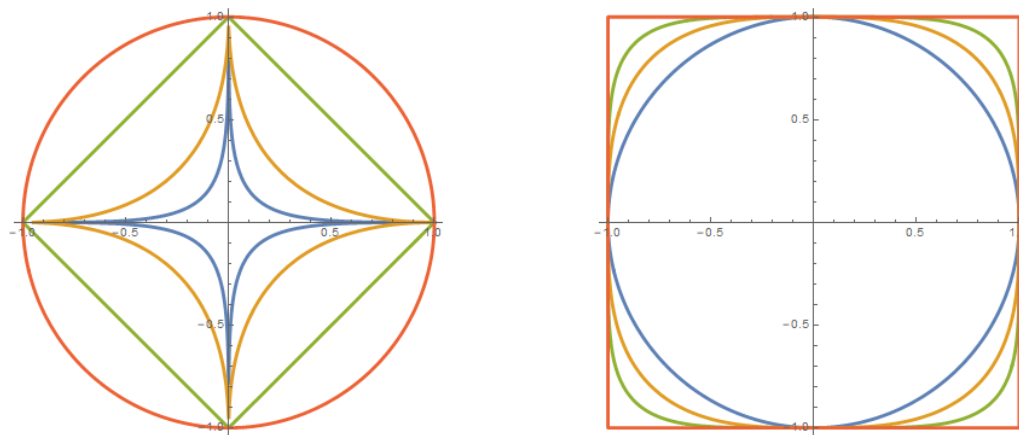


Figure 4: model class 3. left: contour curves of h for value 1 and $p \in \{\frac{1}{3}, \frac{1}{2}, 1, 2\}$, right: $p \in \{2, 3, 4, 5\}$.

If we are given H_η of model class 1 with $\beta = 0$, $a \neq b$, then it follows immediately that $H_\eta(\varphi(x))$ allows the representation (see (15))

$$H_\eta(\varphi(x)) = \frac{\|x\|_{(1,1/\sqrt{a+1})}^2}{\|x\|_{(1,1/\sqrt{b+1})}^2} \quad (23)$$

with $\|x\|_{(u,v)}$ denoting the norm $\|x\|_{(u,v)} = ((\frac{x_1}{u})^2 + (\frac{x_2}{v})^2)^{1/2}$. The function $H_\eta(\varphi(x))$ is homogeneous of degree zero. For several values of A and B , and different norms $\|\cdot\|$, Figure 5 shows the level sets of the function h_K with

$$h_K(x) = \|x\| \frac{x_1^2 + Ax_2^2}{x_1^2 + Bx_2^2}, \quad x \in \mathbb{R}^2. \quad (24)$$

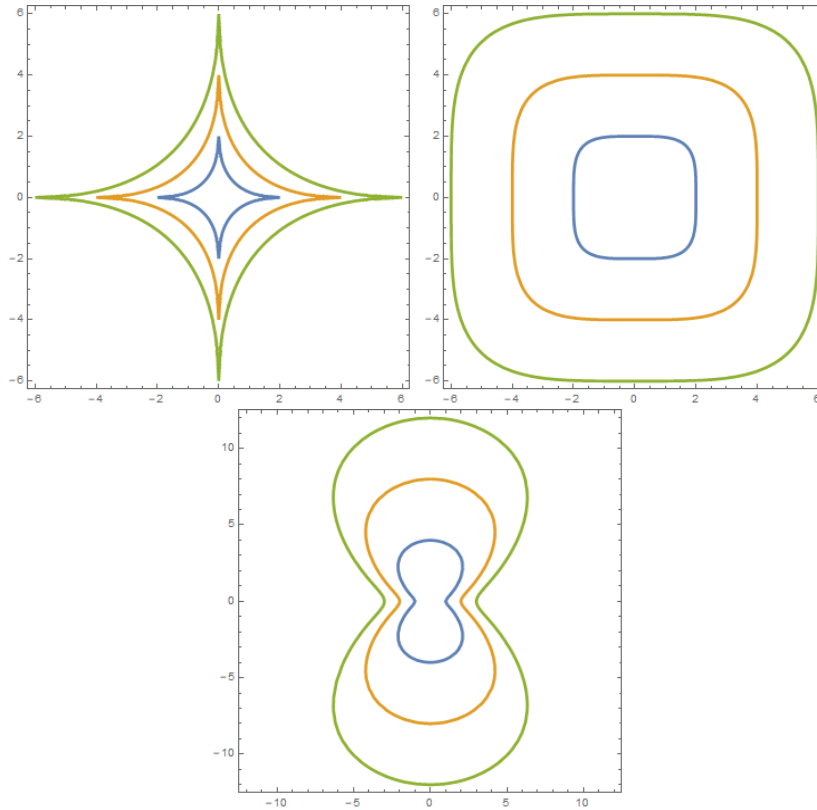


Figure 5: model class 1, h_K according to (24). left: contour curves of h for $p = \frac{1}{2}$, $A, B = 0$, right: for $p = 4$, $A, B = 0$, bottom: for $p = 4$, $A = 1$, $B = 4$

Similarly, one can apply model class 2 of Section 4.1 in this context.

The following class is of particular interest because it combines global and local approximation of data clouds in a very specific way and can be generalized and modified in various directions.

MODEL CLASS 4: Let $a \in (0, 1)$ and $p > 0$ be real numbers ($\eta = (a, p)^T$) and

$$H_\eta(\varphi) = 1 + a \sin_p(\varphi), \quad \varphi \in [0, 2\pi),$$

where the p -generalized sine function \sin_p defined in Example 2 of Section 2.1 is just $\sin_{\mathcal{S}_{\mathcal{B}}}$ in (16) for the case $\|\cdot\| = \|\cdot\|_p$. \square

Then

$$H_\eta(\varphi(x)) = 1 + a \frac{x_2}{\|x\|_p}$$

is positive homogeneous of degree zero and

$$\phi_{g,h,0,I}(x) = \kappa^{-1} g(\|x\| (1 + a \frac{x_2}{\|x\|_p})).$$

The following Figure 6 illustrates the shape of density level sets for model class 4.

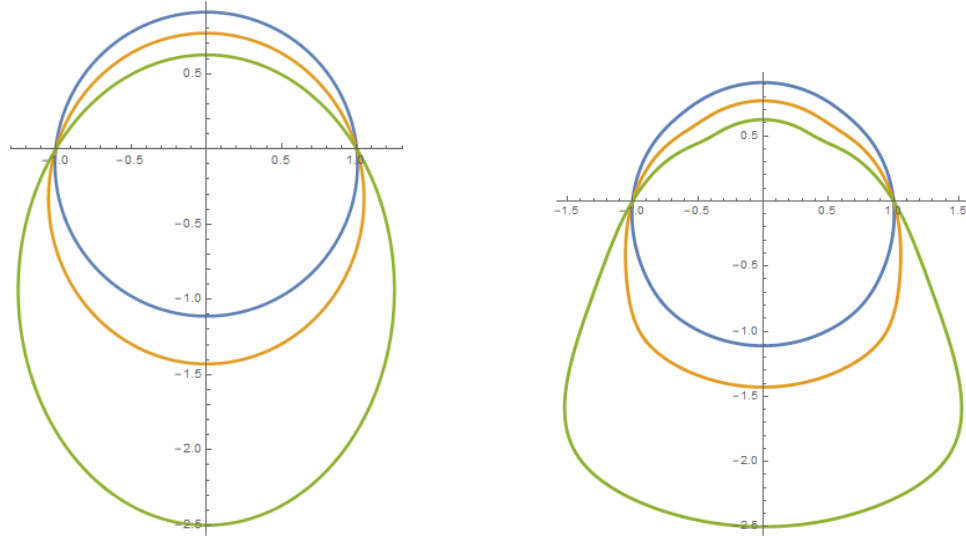


Figure 6: model class 4. left: contour curves of h for value 1 and $p = 2$, $a = 0.1, 0.3, 0.6$, right: the same for $p = 6$

As explained above, an additional parameter β causes a rotation of the contour of h_K if φ is replaced by $\varphi - \beta$ in the formulas for H_η . We do not discuss this aspect here in more detail. The next section is devoted to higher-dimensional models.

5 Modelling function H in higher-dimensional cases

5.1 Using Euclidean norm

In this section, we use the d -dimensional spherical angles $\alpha_1, \dots, \alpha_{d-2} \in (0, \pi)$, $\alpha_{d-1} \in [0, 2\pi)$ according to Example 1 in Section 2.1, combined as vector $\alpha \in A_d = (0, \pi)^{d-2} \times [0, 2\pi)$. It is reasonable to pose a symmetry assumption $h_K(-x) = h_K(x)$ for all $x \in \mathbb{R}^d$ which is equivalent to

$$H_\eta((\pi - \alpha_1, \dots, \pi - \alpha_{d-2}, \pm\pi + \alpha_{d-1})^T) = H_\eta(\alpha) \quad (25)$$

for $\alpha \in A_d$, where $\pm\pi = \pi$ for $\alpha_{d-1} \leq \pi$, $\pm\pi = -\pi$ otherwise. Then H_η is homogeneous of degree zero. We consider model functions $H_\eta : A_d \rightarrow (0, \infty)$ with parameters $\eta \in \Theta_2 \subset \mathbb{R}^{q_2}$, and pose the following assumption on $H = H_\eta$:

ASSUMPTION \mathcal{H}_d : $H : A_d \rightarrow (0, \infty)$ is continuously differentiable, bounded,

$$\lim_{t \rightarrow 2\pi-0} H_\eta(\alpha_1, \dots, \alpha_{d-2}, t) = H_\eta(\alpha_1, \dots, \alpha_{d-2}, 0), \quad (26)$$

$$\lim_{\alpha_j \rightarrow 0+0} H_\eta(\alpha) = \check{H}_j(\alpha_1, \dots, \alpha_{j-1}), \quad (27)$$

$$\lim_{\alpha_j \rightarrow \pi-0} H_\eta(\alpha) = \check{H}_j(\alpha_1, \dots, \alpha_{j-1}) \quad (28)$$

for $j = 1, \dots, d-2$ with constants $\check{H}_1, \check{H}_1 > 0$ and appropriate bounded functions $\check{H}_j : (0, \pi)^{j-1} \rightarrow (0, \infty)$ ($j \geq 2$), (25) holds and

$$\inf_{\alpha \in [0, 2\pi)} H_\eta(\alpha) > 0. \quad \square$$

Identities (26)-(28) describe the continuity of H at boundaries of A_d . For any $j \in \{1, \dots, d-2\}$, $r > 0$, $\alpha_1, \dots, \alpha_{j-1} \in (0, \pi)$, the points $\{(r, \alpha_1, \dots, \alpha_{j-1}, 0, \alpha_{j+1}, \dots, \alpha_{d-1}) : \alpha_{j+1}, \dots, \alpha_{d-2} \in (0, \pi), \alpha_{d-1} \in [0, 2\pi)\}$ in spherical coordinates coincide with one point $(x_1, \dots, x_j, 0, \dots, 0)$ in Cartesian coordinates where

$x_j > 0$. This leads to assumption (27). For any $j \in \{1, \dots, d-2\}$, $r > 0$, $\alpha_1, \dots, \alpha_{j-1} \in (0, \pi)$, the points $\{(r, \alpha_1, \dots, \alpha_{j-1}, \pi, \alpha_{j+1}, \dots, \alpha_{d-1}) : \alpha_{j+1}, \dots, \alpha_{d-2} \in (0, \pi), \alpha_{d-1} \in [0, 2\pi)\}$ in spherical coordinates coincide with one point $(x_1, \dots, x_j, 0, \dots, 0)$ in Cartesian coordinates where $x_j < 0$. Thus condition (28) should be required. Moreover, for any $r > 0$, $\alpha_1, \dots, \alpha_{d-2} \in (0, \pi)$, the points $\{(r, \alpha_1, \dots, \alpha_{d-1}) : \alpha_{d-1} \in \{0, 2\pi)\}$ in spherical coordinates coincide with one point $(x_1, \dots, x_{d-1}, 0)$ in Cartesian coordinates. Therefore (26) is a reasonable requirement.

Let Q_η be an orthogonal matrix describing the rotation of the density in terms of η . Then the formula for the contour defining function is as follows:

$$h_K(x) = \|x\|_2 H_\eta(\tilde{T}(Q_\eta x)).$$

By means of Q_η , we rotate the contour determined by the prototype model H_η . It has to be ensured that a rotation of one element H_{η_1} does not lead to another element H_{η_2} ($\eta_2 \neq \eta_1$) of the model class. To simplify the representation, we do not consider an additional rotation in the following. Incorporating this rotation is left to the reader. Next we introduce the model

MODEL CLASS 5:

$$H_\eta(\alpha) = \left(1 + \sum_{j=1}^{d-1} a_j \prod_{k=1}^j \sin^2 \alpha_k \right)^{-1},$$

$\eta = (a_1, \dots, a_{d-1})^T$. a_1, \dots, a_{d-1} are the parameters of this function, where not all a_j are equal to zero, and $1 + \sum_{j=1}^m a_j > 0$ for all $m = 1, \dots, d-1$. \square

Obviously, (26) and (25) are fulfilled for model class 5 and H_η is bounded away from zero. Lemma 10.1 in Section 10 shows that model class 5 satisfies the remaining conditions of Assumption \mathcal{H}_d . Now we can state an interesting rotational property.

Rotational properties of model class 5: H and therefore the density is invariant w.r.t. rotations about the x_1 -axis if $a_j = 0$ for $j \leq d-2$. H is invariant w.r.t. rotations about the x_d -axis if $a_j = 0$ for $j \geq 2$. Specifically, for $d = 3$, x_1 and x_d correspond to z and y in the common notations.

Let us denote the (marginal) densities of $\Psi_{k_1}, \dots, \Psi_{k_m}$ by ϕ_{k_1, \dots, k_m} . For model class 4 we derive the following lemma on marginal densities of Ψ :

Lemma 5.1. For model class 5 and different numbers $k_1, \dots, k_m \in \{1, \dots, d-1\}$, we have

$$\phi_{k_1, \dots, k_m}(\alpha_{k_1}, \dots, \alpha_{k_m}) = \kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{1 \leq k \leq d-1, k \notin \{k_1, \dots, k_m\}} v_k S(\bar{n}_k) \prod_{k \in \{k_1, \dots, k_m\}} \sin^{\bar{n}_k} \alpha_k,$$

where $\bar{n}_k = 2(n_k + \dots + n_{d-1}) + d - k - 1$, $v_k = 1$ for $k < d-1$, $v_k = 2$ for $k = d-1$, $\alpha_k \in (0, v_k \pi)$,

$$S(m) = \sqrt{\pi} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m+2}{2}\right)}.$$

In particular, the one-dimensional marginal densities can be evaluated by

$$\phi_l(\bar{\alpha}) = \kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=l}^{d-1} a_j^{n_j} \prod_{1 \leq k \leq d-1, k \neq l} v_k S(\bar{n}_k) \sin^{\bar{n}_l} \bar{\alpha}, \quad \bar{\alpha} \in (0, v_l \pi).$$

With this lemma, easy-to-handle formulas for the marginal densities of Ψ of every dimension are provided. S gives integrals of powers of the sine function. As a byproduct of the proof of the previous lemma, we obtain a formula for κ .

Lemma 5.2. For model functions H of class 5, the following identity holds:

$$\kappa = \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{k=1}^{d-1} v_k S(\bar{n}_k),$$

S , \bar{n}_k and v_k as in the previous lemma.

Figure 7 illustrates the shape of level sets of densities of model class 5.

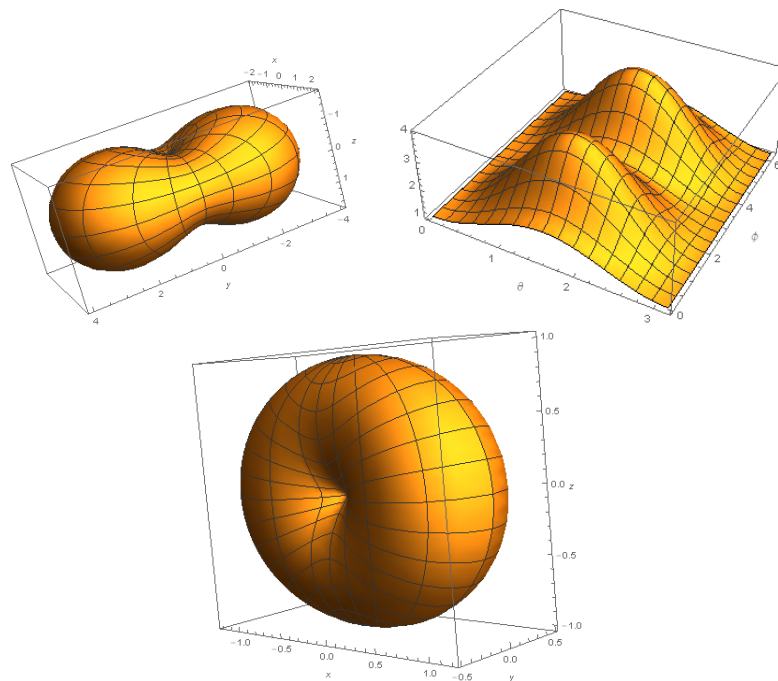


Figure 7: model class 5. left, right: contour surface of the density and function H ($\alpha_1 = \theta$, $\alpha_2 = \varphi$) in the case $a = 2$, $b = 1$. bottom: contour surface in the case $a = -1$, $b = 0.2$

Class 5 can be generalized as follows:

MODEL CLASS 6:

$$H_{\eta}(\alpha) = \left(1 + a_{d-1}(1 - \bar{H}_{\eta}(\alpha_{d-1})^{-1}) \prod_{k=1}^{d-2} \sin^2 \alpha_k + \sum_{j=1}^{d-2} a_j \prod_{k=1}^j \sin^2 \alpha_k \right)^{-1}.$$

Let $c := \min(0, \inf_{\varphi \in [0, 2\pi)} a_{d-1}(1 - \bar{H}_{\eta}(\varphi)^{-1})$. In this model a_1, \dots, a_{d-1} are the parameters of function H_{η} , where not all a_j are equal to zero, and $1 + c + \sum_{j=1}^m a_j > 0$ for all $m = 1 \dots d-2$. \bar{H}_{η} is a function from model class 1 or 2. \square

If $\alpha_1, \dots, \alpha_{d-2} = \frac{\pi}{2}$ and $a_1, \dots, a_{d-2} = 0$, $a_{d-1} = -1$ then $H_{\eta}(\alpha) = \bar{H}_{\eta}(\alpha_{d-1})$. Model class 6 can be treated in a similar way as class 5.

5.2 Using other norms

Let $\|x\|$ be any norm or antinorm of \mathbb{R}^d . In this section the general model for function h is given by

$$h_{\eta}(x) = \|x\| H_{\eta}(\tilde{T}(Q_{\eta}x)),$$

where Q_{η} is an orthogonal matrix describing a rotation. As above we omit this rotation in the following to simplify the presentation.

From a technical point of view, the situation in higher dimensional cases still differs essentially from that in dimension two because the properties of the simple transformation (17) are not being dominant in the multivariate case. More advanced coordinate systems and techniques might be useful to construct functions H_{η} being positively homogeneous of degree zero.

However, on using the common multivariate polar or spherical coordinate transformation and its well known inverse, we can introduce another model class.

MODEL CLASS 7: For $\eta = p > 0$,

$$H_\eta(\varphi) = (|\cos \varphi_1|^p + |\sin \varphi_1 \cos \varphi_2|^p + \dots + |\sin \varphi_1 \dots \sin \varphi_{d-2} \cos \varphi_{d-1}|^p + |\sin \varphi_1 \dots \sin \varphi_{d-2} \sin \varphi_{d-1}|^p)^{1/p}. \quad (29)$$

Figure 8 shows two contour surfaces of densities in Model class 7.

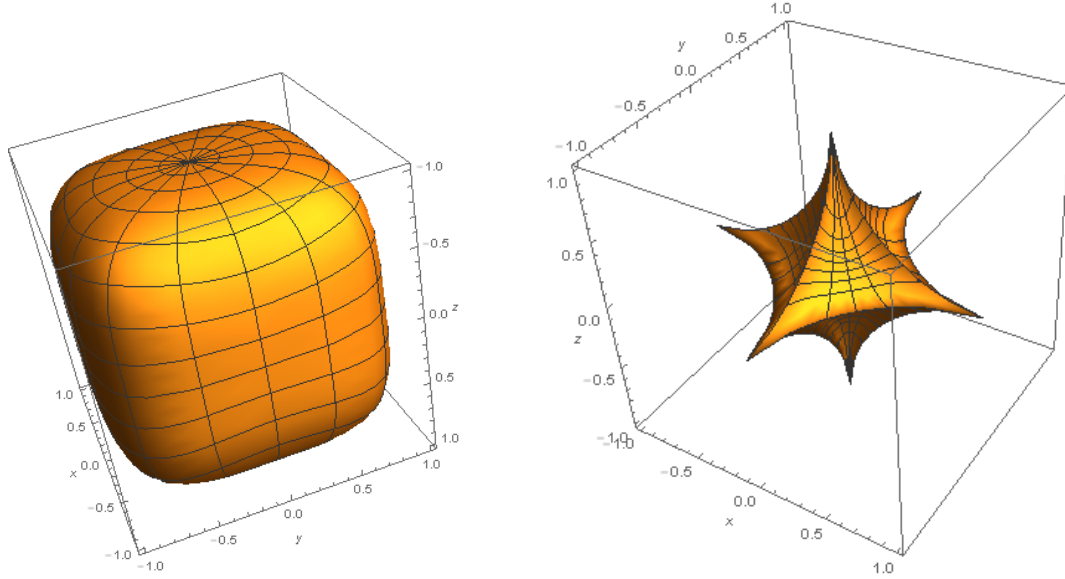


Figure 8: model class 7. left: contour surface of the density in the case $p = 4$. right: contour surface in the case $p = 0.6$

In case that $B = \{x \in \mathbb{R}^d : \|x\|_2 \leq 1\}$, Model class 3 may directly be generalized to the multivariate case, meaning that equations (19) up to (22) are valid also for $x \in \mathbb{R}^d$ with the norms correspondingly defined there.

Now we intend to generalize function H from (29) using ellipsoidal coordinates on the basis of $E_{(a,b)}$ -generalized trigonometric functions being defined in [15] for positive values of a, b as

$$\cos_{(a,b)}(\varphi) = \frac{(\cos \varphi)/a}{N_{a,b}(\varphi)} \text{ and } \sin_{(a,b)}(\varphi) = \frac{(\sin \varphi)/b}{N_{a,b}(\varphi)}, \quad \varphi \in [0, 2\pi),$$

where $N_{a,b}(\varphi) = (((\cos \varphi)/a)^2 + ((\sin \varphi)/b)^2)^{1/2}$. Let the ellipsoidal coordinate transformation

$$T_a^E : M_d \rightarrow \mathbb{R}^d \text{ with } M_d = [0, \infty) \times A_d, \quad A_d = (0, \pi)^{(d-2)} \times [0, 2\pi),$$

where $a = (a_1, \dots, a_d)^T$ consists of positive real numbers, be defined as in [15]. The map $T_{(a,b)}^E$ is almost one-to-one, and with

$$x_i^* = \frac{x_i}{a_i} \text{ and } r_j^* = \left(\sum_{i=j}^d x_i^{*2} \right)^{1/2},$$

its inverse is given a.e. by

$$r = \left(\sum_{i=1}^d x_i^{*2} \right)^{1/2}, \quad \alpha_j = \arccos_{(a_j, a_{j+1})} \left(\frac{x_j^*}{r_j^*} \right), \quad j = 1, \dots, d-2,$$

and, if $x_{d-1} \neq 0$,

$$\arctan \left| \frac{x_d}{x_{d-1}} \right| = \alpha_{d-1} \text{ if } (x_{d-1}, x_d) \in Q_1, = \pi - \alpha_{d-1} \text{ in } Q_2,$$

$$\arctan \left| \frac{x_d}{x_{d-1}} \right| = -\pi + \alpha_{d-1} \text{ if } (x_{d-1}, x_d) \in Q_3, = 2\pi - \alpha_{d-1} \text{ in } Q_4.$$

Here, $\arccos_{(a_j, a_{j+1})}$ denotes the function inverse to $\cos_{(a_j, a_{j+1})}$ and Q_1 up to Q_4 denote anti-clockwise enumerated quadrants from \mathbb{R}^2 .

Let $B = \{x \in \mathbb{R}^d : \|x\|_{a,2} \leq 1\}$ and, for arbitrary $p > 0$, the following model class can be defined:

MODEL CLASS 8: For $\eta = (p, a_1, \dots, a_d)^T$, $p, a_j > 0$,

$$\begin{aligned} H_\eta(\alpha) &= (|\cos_{(a_1, a_2)}(\alpha_1)|^p + |\sin_{(a_1, a_2)}(\alpha_1) \cos_{(a_2, a_3)}(\alpha_2)|^p + \dots \\ &\quad + |\sin_{(a_1, a_2)}(\alpha_1) \dots \sin_{(a_{d-2}, a_{d-1})}(\alpha_{d-2}) \cos_{(a_{d-1}, a_d)}(\alpha_{d-1})|^p \\ &\quad + |\sin_{(a_1, a_2)}(\alpha_1) \dots \sin_{(a_{d-2}, a_{d-1})}(\alpha_{d-2}) \sin_{(a_{d-1}, a_d)}(\alpha_{d-1})|^p)^{1/p}. \quad \square \end{aligned}$$

With the abbreviations $x_j^* = \frac{x_j}{a_j}$ and $r_j^* = (x_j^{*2} + \dots + x_d^{*2})^{1/2}$, for $j = 1, \dots, d$, and $r_1^* = \|x\|_{a,2}$, for $j = 1, \dots, d-2$, the following equations are valid:

$$\cos_{(a_j, a_{j+1})}(\alpha_j) = \frac{x_j^*}{r_j^*} \text{ and } \sin_{(a_j, a_{j+1})}(\alpha_j) = \frac{r_{j+1}^*}{r_j^*}.$$

Moreover,

$$\cos_{(a_{d-1}, a_d)}(\alpha_j) = \frac{|x_{d-1}^*|}{(x_{d-1}^{*2} + x_d^{*2})^{1/2}}, \quad \sin_{(a_{d-1}, a_d)}(\alpha_j) = \frac{|x_d^*|}{(x_{d-1}^{*2} + x_d^{*2})^{1/2}}$$

and

$$N_{a_{d-1}, a_d}(\arctan \left| \frac{x_d}{x_{d-1}} \right|) = \frac{|(x_{d-1}, x_d)|_{(a_{d-1}, a_d)}}{|(x_{d-1}, x_d)|_2}.$$

It follows that, further generalizing equation (19), there holds

$$H_\eta(\tilde{T}(x)) = \frac{\|x\|_{a,p}}{\|x\|_{a,2}}$$

proving that $H_\eta(\tilde{T}(\cdot))$ is positive homogeneous of degree zero. Because of the relations

$$T_{a,p}^{E-1}\left(\frac{x}{\|x\|}\right) = (1, \tilde{T}(x)) \text{ and } H_\eta(\alpha) = h_K(\text{Pol}_{\mathbb{S}_{\mathbb{B}}}(1, \alpha)),$$

we have finally that

$$\phi_{g,h,0,I}(x) = C(g, K)g(\|x\|_{a,p}), \quad x \in \mathbb{R}^d.$$

For the stochastic representation of a vector \mathbf{X} following this distribution, see Section 1. A principal component representation of $\phi_{g,h,0,I}$ has been dealt with in [20].

6 Simulation

Simulations of the distribution can be based on Lemma 2.3. Let $\tilde{\mathbf{X}} = \Sigma^{-1}(\mathbf{X} - \mu)$, and $\Psi = \tilde{T}(\Sigma^{-1}(\mathbf{X} - \mu))$ be defined as in Section 2.2. Remember that $R = h(\tilde{\mathbf{X}})$. By (9), we have

$$R = \|\tilde{\mathbf{X}}\|_2 h(T(1, \Psi)) = \|\tilde{\mathbf{X}}\|_2 H(\Psi).$$

This implies

$$\begin{aligned} \tilde{\mathbf{X}} &= T(\|\tilde{\mathbf{X}}\|_2, \Psi) \\ &= \|\tilde{\mathbf{X}}\|_2 H(\Psi) T(H^{-1}(\Psi), \Psi) \\ &= R T(H^{-1}(\Psi), \Psi). \end{aligned}$$

In view of (6), we have $U \stackrel{d}{=} T(H(\Psi)^{-1}, \Psi)$. Random vector U has the properties described in Section 2.1. Moreover, by (9),

$$h(U) = H(\Psi)^{-1} h(T(1, \Psi)) = H(\Psi)^{-1} H(\Psi) = 1 \text{ a.s.}$$

Let model functions g, H be given. Then we can apply the following algorithm to simulate \mathbf{X} :

- 1) Generate R with density f from Lemma 2.3.
- 2) Generate Ψ (independently of R) with density ϕ_Ψ from Lemma 2.3.
- 3) Evaluate $U = T(H^{-1}(\Psi), \Psi)$ and \mathbf{X} according to (6)

Step 2) is realised in several partial steps: First simulate Ψ_1 with density ϕ_1 , then Ψ_2 with density $\phi_{1,2}(\Psi_1, \cdot)/\phi_1(\Psi_1)$ and so forth until Ψ_{d-1} with density $\phi_{1,\dots,d-1}(\Psi_1, \dots, \Psi_{d-2}, \cdot)/\phi_{1,\dots,d-2}(\Psi_1, \dots, \Psi_{d-2})$. For model class 4, the evaluation of the densities $\phi_{1,\dots,k}$ can be done using Lemma 5.1.

7 Estimation

Throughout this section, let $\mathbf{X}_1, \dots, \mathbf{X}_n$ with $\mathbf{X}_i = (X_{i1}, \dots, X_{id})^T$ be a sample of independent random vectors having the density $\varphi_{g,h,\mu,\Sigma}$ according to (5) and (10). Suppose that function g belongs to the model class $\{g_\theta : \theta \in \Theta_1\}$ with compact $\Theta_1 \subset \mathbb{R}^{q_1}$ as described above. Moreover, it is required that function H depends on a parameter η such that H belongs to the model class $\{H_\eta : \eta \in \Theta_2\}$ with compact $\Theta_2 \subset \mathbb{R}^{q_2}$ as described above. Let $\Theta_3 \subset \mathbb{R}^d$ and $\Theta_4 \subset (0, \infty)^d$ be compact sets.

In this section the aim is to fit the specific parametric model for the density $\varphi_{g,h,\mu,\Sigma}$ (cf. formula (5)) with parameters $\mu, \sigma = (\sigma_1, \dots, \sigma_d)^T, \eta$ and θ , where $\Sigma = \text{diag}(\sigma)$. The log likelihood function reads as follows

$$\ln L(\mu, \sigma, \eta, \theta) = \sum_{i=1}^n \ln g_\theta(h_\eta(\Sigma^{-1}(\mathbf{X}_i - \mu))) - n \sum_{j=1}^d \ln \sigma_j - n \ln(\kappa(\eta)).$$

Here $\kappa(\eta)$ is the normalizing constant of the density which is determined by the formula in Lemma 2.1. A maximization of the likelihood function w.r.t. the parameters leads to

$$L(\hat{\mu}_n, \hat{\sigma}_n, \hat{\eta}_n, \hat{\theta}_n) = \max_{\bar{\mu} \in \Theta_3, \bar{\sigma} \in \Theta_4, \bar{\eta} \in \Theta_2, \bar{\theta} \in \Theta_1} L(\bar{\mu}, \bar{\sigma}, \bar{\eta}, \bar{\theta}),$$

where $\hat{\mu}_n, \hat{\sigma}_n, \hat{\eta}_n, \hat{\theta}_n$ are the maximum likelihood estimators. In applications, estimation of the parameters corresponds to solving an optimization problem by using a computer program (for instance, by using R). For this purpose, one needs good starting points for the parameters in view of the rather high dimension of the problem.

Under appropriate assumptions, maximum likelihood estimators are asymptotically normally distributed (cf. Theorem 5.1 in [8] by Lehmann and Casella, p. 463)

$$(\hat{\mu}_n, \hat{\sigma}_n, \hat{\eta}_n, \hat{\theta}_n) \xrightarrow{d} \mathcal{N}((\mu, \sigma, \eta, \theta), I(\mu, \sigma, \eta, \theta)^{-1}) \text{ for } n \rightarrow \infty,$$

where \xrightarrow{d} is the symbol for convergence in distribution. Here $I(\mu, \sigma, \eta, \theta) = (I_{ij}(\delta))_{i,j=1\dots d+q}$ with $\delta^T = (\delta_1, \dots, \delta_{2d+q_1+q_2}) = (\mu^T, \sigma^T, \eta^T, \theta^T)$ denotes the information matrix which is given by

$$I_{ij}(\delta) = -\mathbb{E} \left(\frac{\partial}{\partial \delta_i \partial \delta_j} \left(\ln g_\theta(h_\eta(\Sigma^{-1}(\mathbf{X}_i - \mu))) - \sum_{j=1}^d \ln \sigma_j - \ln(\kappa(\eta)) \right) \right).$$

8 Model checks

8.1 Checking the model for the distribution of the star-generalized radius

Let $\hat{\mu}, \hat{\sigma}, \hat{\eta}, \hat{\theta}$ be the maximum-Likelihood estimators for the parameters as explained in the previous section. $\mathbf{X}_1, \dots, \mathbf{X}_n$ is the sample as in the previous section. Let $\mathbf{Y}_i = \hat{\Sigma}^{-1}(\mathbf{X}_i - \hat{\mu})$. Here we consider the pseudo-sample $\mathbf{Y}_1, \dots, \mathbf{Y}_n$, and introduce

$$R_i = \|\mathbf{Y}_i\|_2 H_{\hat{\eta}}(\tilde{T}(\mathbf{Y}_i))$$

for $i = 1, \dots, n$. The corresponding empirical distribution function of R_i is given by

$$F_{nR}(r) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{R_i \leq r\}.$$

The order statistics of R_i are denoted by $R_{(1)}, \dots, R_{(n)}$ ($R_{(j-1)} \leq R_{(j)}$). Based on a model function g_θ ($\theta \in \Theta_1$) for the generating function, the model distribution function for the generalized radius can be evaluated by (formula in Lemma 2.3)

$$F_R(r|\theta) = \int_0^r g_\theta(s) s^{d-1} ds.$$

In the following we consider the Anderson-Darling statistic which measures the discrepancy between the model distribution and empirical distribution coming from the sample. This statistic is calculated by

$$\begin{aligned} A_n &= \int_0^{+\infty} \frac{(F_{nR}(r) - F_R(r|\hat{\theta}))^2}{F_R(r|\hat{\theta})(1 - F_R(r|\hat{\theta}))} dF_R(r|\hat{\theta}) \\ &= -1 - \sum_{i=1}^n \frac{2i-1}{n^2} \left(\ln F_R(R_{(i)}|\hat{\theta}) + \ln(1 - F_R(R_{(n+1-i)}|\hat{\theta})) \right). \end{aligned}$$

Let $F_R^{\text{exp}}(x|\lambda) = 1 - e^{-\lambda x}$ be the distribution function of the exponential distribution. The exponential distribution is regarded as a reference distribution here. For comparisons we can calculate the approximation coefficient:

$$\hat{\rho} = 1 - \frac{A(F_n, F_R(\cdot|\hat{\theta}))}{A(F_n, F_R^{\text{exp}}(\cdot|\hat{\lambda}))} \quad (30)$$

($\hat{\rho} \leq 1$). A detailed study of this coefficient can be found in [10]. Here we compare the actual model distribution of R with the exponential distribution as the simplest choice. If $\hat{\rho}$ is large enough, ideally close to 1, the distribution of R can be considered as well-approximated. The application of goodness-of-fit tests like the Kolmogorov test is straightforward and is omitted here. In the framework of elliptical distributions such goodness-of-fit tests are considered in Batsidis and Zografos [3].

8.2 Checking the distribution of Ψ

Let

$$\hat{\Psi}_i = \tilde{T}(\hat{\Sigma}^{-1}(\mathbf{X}_i - \hat{\mu}))$$

for $i = 1, \dots, n$. $\hat{\Psi}_i$ is the vector of spherical angles of the normalized sample item. The goodness-of-approximation of the one-dimensional distributions of $\Psi_1, \dots, \Psi_{d-1}$ (marginal distributions of Ψ) can be checked in the same way as described in the previous section with the uniform distribution as the reference distribution.

Next we want to discuss briefly a measure for the goodness-of-approximation of the copula of Ψ . Here we pursue the approach using the Cramér-von Mises divergence. Concerning this approach and theoretical properties of this divergence, we refer to [9]. Let $F_\Psi(\cdot|\eta)$ and $F_j(\cdot|\eta)$ be the distribution functions of Ψ and Ψ_j depending on the model parameter η of function H . The corresponding density of Ψ is provided in Lemma 2.3. We introduce the model copula of Ψ :

$$C_\eta(\psi) = F_\Psi(F_1^{-1}(\psi_1|\eta), \dots, F_{d-1}^{-1}(\psi_{d-1}|\eta)|\eta)$$

for $\psi \in A_d$. We denote the empirical joint distribution function of the pseudo-sample $\hat{\Psi}_1, \dots, \hat{\Psi}_n$ by \hat{H}_n . Let $\bar{F}_n(x) = (F_{1n}(x_1), \dots, F_{d-1,n}(x_{d-1}))^T$ be the vector of the marginal empirical distribution functions of the pseudo sample $\hat{\Psi}_1, \dots, \hat{\Psi}_n$ for $x = (x_1, \dots, x_{d-1})^T$. The estimated Cramér-von Mises divergence is given by

$$\hat{D}_n(C) = \frac{1}{n} \sum_{i=1}^n \left(\hat{H}_n(\hat{\Psi}_i) - C(\bar{F}_n(\hat{\Psi}_i)) \right)^2,$$

which measures the discrepancy between the empirical copula and the model copula C . Let C_η and C_0 be the copulas of the model and the independent copula (which serves as reference copula), respectively. From [9] we can take the coefficient of goodness-of-approximation:

$$\hat{\rho} = 1 - \frac{\hat{D}_n(C_{\hat{\eta}})}{\hat{D}_n(C_0)},$$

where $\hat{\eta}$ is the estimator for the parameter η , and $\hat{\rho} \leq 1$. The larger the coefficient $\hat{\rho}$ the better is the approximation of the distribution.

9 Real data examples

9.1 Example 1

In this section we consider the dataset 5 of Andrews and Herzberg [1]. The yield of grain and straw are the two variables. Assuming the Pearson VII model for g and the model 1 for H , we achieved an approximation coefficient of 0.9846. The following Figures 9 and 10 show the data and the estimated density.

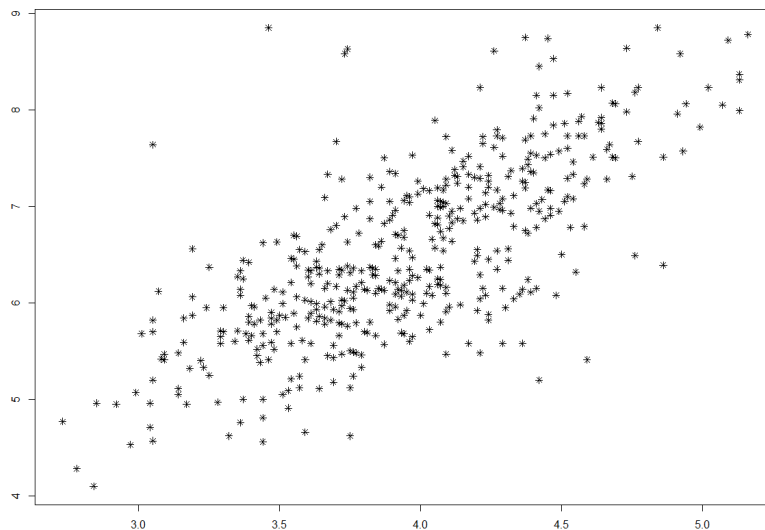


Figure 9: scatter plot of the data

The contour curves in Figure 10 are far away from being ellipses. Thus modelling with star-shaped distributions makes sense.

9.2 Example 2

This example should show that the above described methods work even for economic data. Here we consider weekly index data from Morgan Stanley Capital International of the world and LPX 50 index data for the period April 2003 to December 2016. The LPX 50 Index is a global equity index covering the 50 largest listed private equity companies which fulfil certain liquidity constraints. The index is well diversified across regions, investment- and financing styles, and vintage years. We computed the index values as the ratio of subsequent values minus one. Let F_Ψ be the distribution function of Ψ . For the function H , we used class 1. Fitting the distribution of the random vector by the maximum likelihood method, the following results were

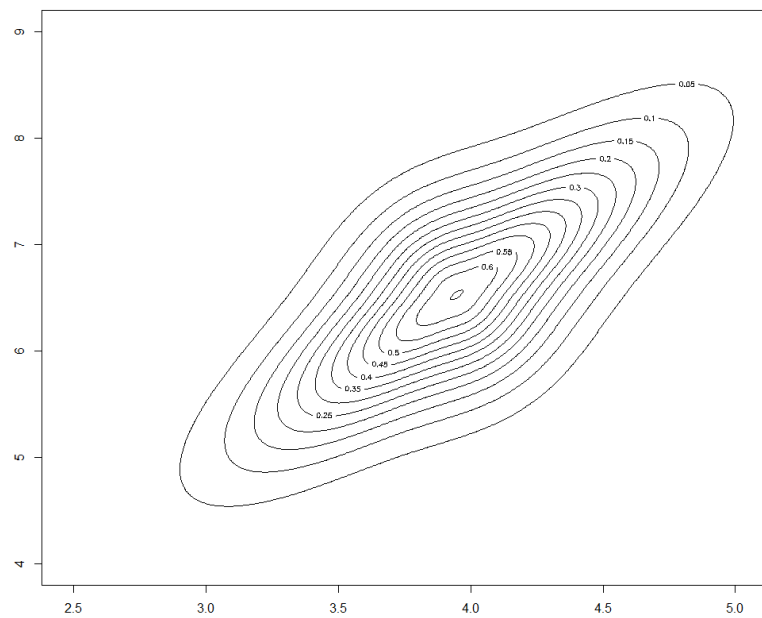


Figure 10: contour plot of the density.

obtained ($\hat{\rho}$ according to (30)):

model for g	$\hat{\rho}$ for F_R	$\hat{\rho}$ for F_Ψ
Kotz type with $s = 0$	0.939	0.977
modified exponential	0.757	0.989
Pearson VII with $s = 0$	0.990	0.992

parameter estimates for the last case: $\hat{\mu} = (0.003138, 0.003586)^T$, $\hat{\sigma} = (0.033375, 0.097245)^T$, $\hat{m} = 2.95742$, $\hat{a} = 8.17546$, $\hat{b} = 2.15740$, $\hat{\beta} = 0.35080$.

The model "Pearson VII" turns out to be the best one. The density and the data are depicted in Figures 11 and 12. The subsequent figures show the distribution functions of R and Ψ .

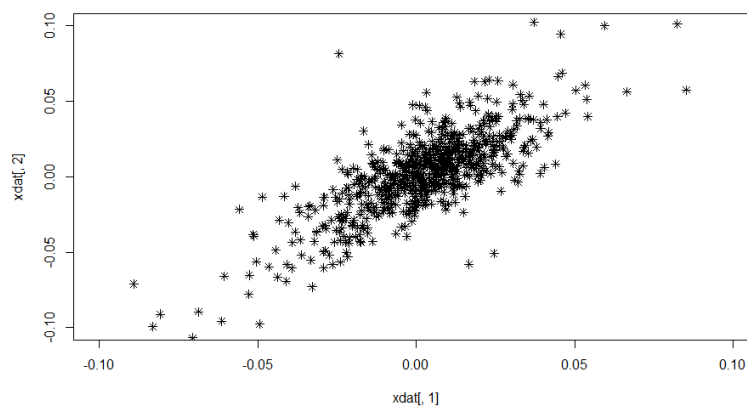


Figure 11: scatter plot of the data

These figures and the $\hat{\rho}$ values show that the fitting was successful.

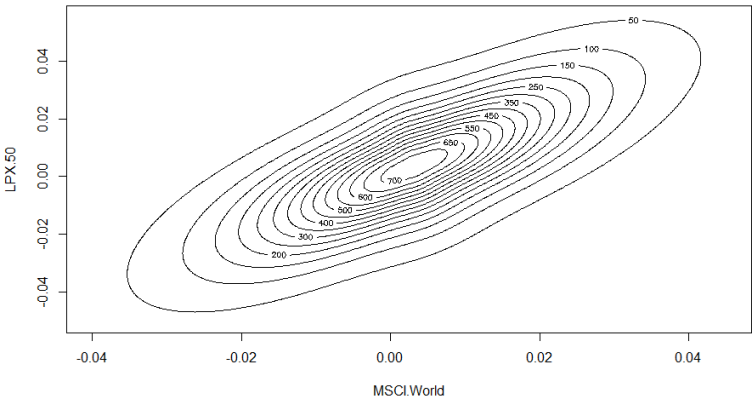


Figure 12: contour plot of the density

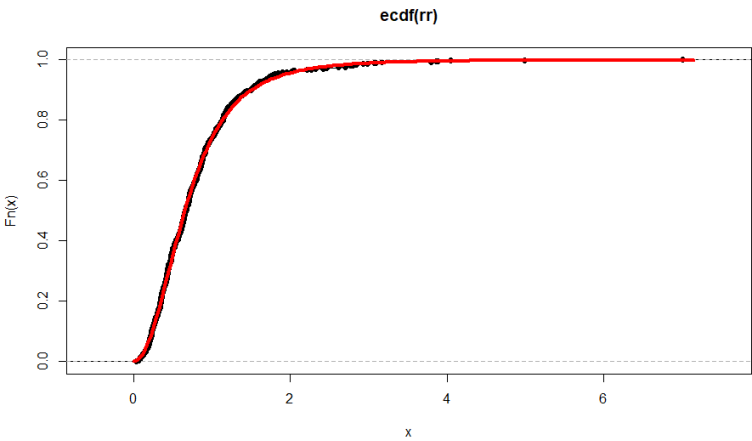


Figure 13: empirical distribution function of R (black) and that of the model (red).

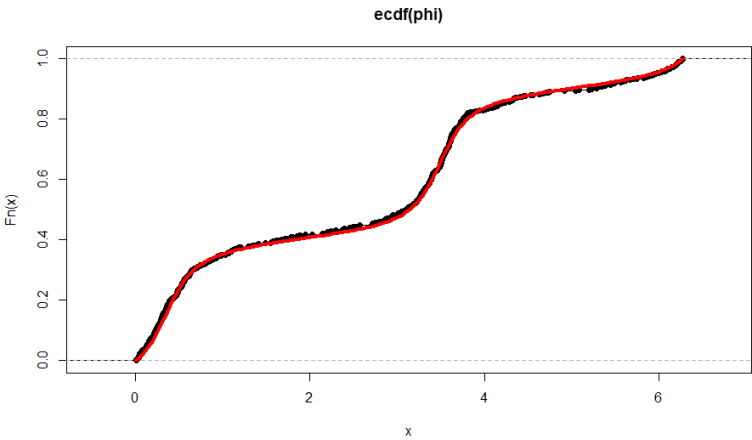


Figure 14: empirical distribution function of Ψ (black) and that of the model (red).

10 Proofs

PROOF OF LEMMA 2.1: Obviously, $\Sigma^{-1}(\mathbf{X} - \mu)$ has density $\phi_{g,h,0,I}$. Applying the transformation (7) with transformation determinant (8) and by $q = rH(\psi)$, we obtain

$$\begin{aligned} \mathbb{P}\{\mathbf{X} \in \mu + \Sigma \bar{A}\} &= \int_{\bar{A}} \phi_{g,h,0,I}(x) \, dx = \kappa^{-1} \int_{\bar{A}} g(h_K(x)) \, dx \\ &= \kappa^{-1} \int_{[\underline{\psi}_1, \bar{\psi}_1] \times \dots \times [\underline{\psi}_{d-1}, \bar{\psi}_{d-1}] \times [\underline{r}/H(\psi), \bar{r}/H(\psi)]} g(rH(\psi)) r^{d-1} \bar{\Delta}(\psi) \, d\psi dr \\ &= \kappa^{-1} \int_{[\underline{\psi}_1, \bar{\psi}_1] \times \dots \times [\underline{\psi}_{d-1}, \bar{\psi}_{d-1}]} \int_{\underline{r}}^{\bar{r}} g(q) q^{d-1} dq H(\psi)^{-d} \bar{\Delta}(\psi) \, d\psi \\ &= \kappa^{-1} \int_{\underline{\psi}_{d-1}}^{\bar{\psi}_{d-1}} \dots \int_{\underline{\psi}_1}^{\bar{\psi}_1} H(\alpha)^{-d} \bar{\Delta}(\psi) \, d\psi_1 \dots d\psi_{d-1} \int_{\underline{r}}^{\bar{r}} g(q) q^{d-1} dq. \end{aligned} \quad (31)$$

This identity yields the first formula of the lemma. Moreover, $\int_{\mathbb{R}^d} \phi_{g,h,0,I}(x) dx = 1$ implies the second assertion of the lemma. \square

PROOF OF LEMMA 2.2: Define $\tilde{\mathbf{X}} = \Sigma^{-1}(\mathbf{X} - \mu) = (\tilde{X}^{(1)}, \dots, \tilde{X}^{(d)})^T$. Then we have

$$\begin{aligned} \mathbb{E}\mathbf{X} &= \mu + \Sigma \mathbb{E}\tilde{\mathbf{X}}, \\ \mathbb{V}(X^{(j)}) &= \sigma_j^2 \mathbb{V}(\tilde{X}^{(j)}) \text{ for } j = 1 \dots d. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E}\tilde{X}^{(j)} &= \kappa^{-1} \int_{\mathbb{R}^d} x_j g(h(x)) \, dx \\ &= \kappa^{-1} \int_0^\infty \int_{A_d} I_j(\psi) g(rH(\psi)) r^d \bar{\Delta}(\psi) \, d\psi dr \\ &= \kappa^{-1} G_d \int_{A_d} I_j(\psi) \bar{\Delta}(\psi) H(\psi)^{-d-1} \, d\psi. \end{aligned}$$

Analogously,

$$\begin{aligned} \mathbb{V}(\tilde{X}^{(j)}) &= \kappa^{-1} \int_{\mathbb{R}^d} x_j^2 g(h(x)) \, dx - \left(\mathbb{E}\tilde{X}^{(j)} \right)^2 \\ &= \kappa^{-1} \int_0^\infty \int_{A_d} I_j^2(\psi) g(rH(\psi)) r^{d+1} \bar{\Delta}(\psi) \, d\psi dr - \left(\mathbb{E}\tilde{X}^{(j)} \right)^2 \\ &= \kappa^{-1} G_{d+1} \int_{A_d} I_j^2(\psi) \bar{\Delta}(\psi) H(\psi)^{-d-2} \, d\psi - \left(\mathbb{E}\tilde{X}^{(j)} \right)^2. \end{aligned}$$

Combining the above formulas, we obtain the lemma. \square

PROOF OF LEMMA 2.3: From (31), we see that (R, Ψ) has the density

$$\phi_{(R,\Psi)}(q, \psi) = H(\psi)^{-d} \bar{\Delta}(\psi) g(q) q^{d-1}.$$

Hence R and Ψ are independent. Their densities can be gathered from this equation. \square

PROOF OF THEOREM 2.4: Suppose that there are two different quadruples of parameters $(g_1, h_1, \sigma_1, \mu_1)$ and $(g_2, h_2, \sigma_2, \mu_2) \in \mathcal{G} \times \mathcal{H} \times (0, \infty)^d \times \mathbb{R}^d$ with $\Sigma_j = \text{diag}(\sigma_j)$ such that the corresponding densities of the random vectors in (5) coincide

$$(\kappa_1 \det(\Sigma_1))^{-1} g_1(h_1(\Sigma_1^{-1}(x - \mu_1))) = (\kappa_2 \det(\Sigma_2))^{-1} g_2(h_2(\Sigma_2^{-1}(x - \mu_2))) \quad (32)$$

for all $x \in \mathbb{R}^d$. κ_1 and κ_2 denote the constant κ for the two densities. Let $\Gamma^{(1)} = \{x : h_1(x) = 1\}$ and $\Gamma^{(2)} = \{x : h_2(x) = 1\}$ be specific level sets of h_1, h_2 , respectively. There are density values $v_1, v_2 \geq 0$ such that $v_1 \neq v_2, v_j(\kappa_1 \det(\Sigma_1)) \neq g_1(0)$ and $v_j(\kappa_2 \det(\Sigma_2)) \neq g_2(0)$. By assumption $\mathcal{A}_g(\text{ii})$, there are $\gamma_1, \gamma_2 \in (0, \infty)$ such that $g_1(\gamma_1) < g_1(\gamma_2), \gamma_j = \min\{\gamma \geq 0 : g_1(\gamma) = g_1(\gamma_j)\}$ for $j = 1, 2$, and there are $\gamma_3, \gamma_4 \in (0, \infty)$ such that $g_2(\gamma_3) < g_2(\gamma_4), \gamma_j = \min\{\gamma \geq 0 : g_2(\gamma) = g_2(\gamma_j)\}$ for $j = 3, 4$, and

$$v_j = (\kappa_1 \det(\Sigma_1))^{-1} g_1(\gamma_j) = (\kappa_2 \det(\Sigma_2))^{-1} g_2(\gamma_{2+j}) \text{ for } j = 1, 2. \quad (33)$$

Note that for every $\gamma > 0$, $h_1(\Sigma_1^{-1}(x - \mu_1)) = \gamma \Leftrightarrow x \in \mu_1 + \gamma \Sigma_1 \Gamma^{(1)}$. Obviously in view of (32), the level sets for a density value equal to v_1 and v_2 have to coincide, respectively. Hence $\mu_1 + \gamma_j \Sigma_1 \Gamma^{(1)} = \mu_2 + \gamma_{j+2} \Sigma_2 \Gamma^{(2)}$ (inner boundaries of the level sets), which is equivalent to $\Gamma^{(1)} = \gamma_j^{-1} \Sigma_1^{-1}(\mu_2 - \mu_1) + \gamma_j^{-1} \gamma_{j+2} \Sigma_1^{-1} \Sigma_2 \Gamma^{(2)}$ for $j = 1, 2$. Note that $\Gamma^{(1)}$ and $\Gamma^{(2)}$ are bounded sets. This implies immediately $\gamma_j^{-1} \gamma_{j+2} = C_1$ for $j = 1, 2$ with a constant C_1 . If $\Sigma_1^{-1}(\mu_2 - \mu_1) \neq 0$ holds, then $\gamma_1 = \gamma_2$ follows which is a contradiction to $g_1(\gamma_1) < g_1(\gamma_2)$. Hence $\Sigma_1^{-1}(\mu_2 - \mu_1) = 0$, and $\mu_1 = \mu_2 = \mu$. Moreover, we have $\Gamma^{(1)} = C_1 \Sigma_0 \Gamma^{(2)}$ with the diagonal matrix $\Sigma_0 = \Sigma_1^{-1} \Sigma_2$. Now let $y \in \mathbb{R}^d$ be arbitrary. Then homogeneity of h_1 implies $h_1(y)^{-1} y \in \Gamma^{(1)}$, and therefore $C_1^{-1} h_1(y)^{-1} \Sigma_0^{-1} y \in \Gamma^{(2)}$. Further

$$\begin{aligned} 1 &= h_2 \left(C_1^{-1} h_1(y)^{-1} \Sigma_0^{-1} y \right), \\ h_1(y) &= h_2(C_1^{-1} \Sigma_0^{-1} y) \end{aligned}$$

using homogeneity of h_2 . Therefore $C_1 \Sigma_0 = I$ and $h_1 = h_2 = h$ in view of Assumption \mathcal{A}_h .

Next we consider the set $\bar{M} = \{\mu + t \cdot e_1 : t \geq 0\}$, $e_1 = (1, 0, \dots, 0)^T$. Let $\sigma_{11} = (\Sigma_1)_{1,1}$, $\sigma_{21} = (\Sigma_2)_{1,1}$. From (32) it follows for $x \in \bar{M}$ that

$$\begin{aligned} g_1(th(e_1)\sigma_{11}^{-1}) &= \kappa_1 \det(\Sigma_1)(\kappa_2 \det(\Sigma_2))^{-1} g_2(th(e_1)\sigma_{21}^{-1}) \text{ for } t \geq 0, \\ g_1(u) &= C_2 g_2(u C_3) \text{ for } u \geq 0 \end{aligned}$$

with $C_2 = \kappa_1 \det(\Sigma_1)(\kappa_2 \det(\Sigma_2))^{-1}$, $C_3 = \sigma_{21}^{-1} \sigma_{11}$ and $C_2 = C_3^d$. Applying Assumption \mathcal{A}_g , one obtains that $C_2 = C_3 = 1$, and hence $g = g_1 = g_2$. Further by (33), we have $g(\gamma_j) = g(\gamma_{j+2})$ for $j = 1, 2$, which implies $\gamma_1 = \gamma_3, \gamma_2 = \gamma_4$, and $C_1 = 1$. So we can conclude $\Sigma_0 = I$, $\Sigma_1 = \Sigma_2$ and $\kappa_1 = \kappa_2$ from the definition of C_2 . This is a contradiction to the assumption that two different quadruples of parameters lead to the same density. \square

Lemma 10.1. Function H of model class 4 fulfils (27), and $\sup_{\alpha \in A_d} H_\eta(\alpha) < +\infty$.

PROOF: For $j = 1 \dots d - 2$, we obtain

$$\lim_{\alpha_j \rightarrow 0+0} H_\eta(\alpha)^{-1} = \left(1 + \sum_{l=1}^{j-1} a_l \prod_{k=1}^l \sin^2 \alpha_k \right).$$

This quantity depends only on $\alpha_1, \dots, \alpha_{j-1}$ which was required in (27). Further

$$\begin{aligned} H_\eta(\alpha)^{-1} &= 1 + \sum_{j=1}^{d-1} a_j \prod_{k=1}^j \sin^2 \alpha_k \\ &\geq 1 + \min_{m=1, \dots, d-1} \sum_{j=1}^m a_j > 0 \end{aligned}$$

for $\alpha \in A_d$, which proves the lemma. \square

PROOF OF LEMMA 5.1: Using the multinomial theorem, we obtain

$$\begin{aligned} H_{\eta}^{-d}(\alpha) &= \left(1 + \sum_{j=1}^{d-1} a_j \prod_{k=1}^j \sin^2 \alpha_k \right)^d \\ &= \sum_{n_1, \dots, n_d: n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{k=1}^j \sin^{2n_j} \alpha_k \end{aligned}$$

for $\alpha \in A_d$. Observe that for integers $m \geq 0$,

$$S(m) = \int_0^{\pi} \sin^m x \, dx = \sqrt{\pi} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m+2}{2})}.$$

Let $k_1, \dots, k_m \in \{1, \dots, d-1\}$ be different numbers, $J_1 = \{k_1, \dots, k_m\}$ and $J_2 = \{1, \dots, d-1\} \setminus J_1$. Define $\bar{A}_m = [0, \pi]^{d-1-m}$ if $d-1 \notin J_2$, and $\bar{A}_m = [0, \pi]^{d-2-m} \times [0, 2\pi]$ otherwise. Further by Lemma 2.3, we have

$$\begin{aligned} \phi_{k_1, \dots, k_m}(\alpha_{k_1}, \dots, \alpha_{k_m}) &= \int_{\bar{A}_m} \phi_{\Psi}(\alpha) \prod_{j \in J_2} d\alpha_j \\ &= \kappa^{-1} \int_{\bar{A}_m} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} \left(a_j^{n_j} \prod_{k=1}^j \sin^{2n_j} \alpha_k \right) \\ &\quad \left(\prod_{k=1}^{d-2} \sin^{d-k-1} \alpha_k \right) \prod_{j \in J_2} d\alpha_j \\ &= \kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \\ &\quad \int_{\bar{A}_m} \left(\prod_{k=1}^{d-1} \prod_{j=k}^{d-1} \sin^{2n_j} \alpha_k \right) \left(\prod_{k=1}^{d-2} \sin^{d-k-1} \alpha_k \right) \prod_{j \in J_2} d\alpha_j \\ &= \kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{k \in J_1} \sin^{2(n_k + \dots + n_{d-1}) + d - k - 1} \alpha_k \\ &\quad \prod_{k \in J_2} \int_{[0, v_k \pi]} \sin^{2(n_k + \dots + n_{d-1}) + d - k - 1} \bar{\alpha} d\bar{\alpha} \\ &= \kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{k \in J_1} \sin^{\bar{n}_k} \alpha_k \prod_{k \in J_2} v_k S(\bar{n}_k). \end{aligned} \quad (34)$$

for $(\alpha_{k_1}, \dots, \alpha_{k_m}) \in \bar{A}_m$, v_k and \bar{n}_k as above in the lemma. This completes the proof. \square

PROOF OF LEMMA 5.2: Considering the case $J_2 = \{1, \dots, d-1\}$, and $\bar{A}_m = A_d$, identity (34) yields

$$\kappa^{-1} \sum_{n_1 + \dots + n_d = d} \frac{d!}{n_1! \dots n_d!} \prod_{j=1}^{d-1} a_j^{n_j} \prod_{k=1}^{d-1} v_k S(\bar{n}_k) = 1. \quad \square$$

References

- [1] Andrews, D. F. and A. M. Herzberg, (1985). *Data: A Collection of Problems from many Fields for the Student and Research Worker*. Springer, New York.
- [2] Balkema, A. A., P. Embrechts and N. Nolde (2010). Meta densities and the shape of their sample clouds. *J. Multivariate Anal.* 101(7), 1738–1754.

- [3] Batsidis, A. and K. Zografos (2013). A necessary test of fit of specific elliptical distributions based on an estimator of Song's measure. *J. Multivariate Anal.* 113, 91–105.
- [4] Fernandez, C., J. Osiewalski and M.F.J. Steel (1995). Modeling and inference with v -spherical distributions. *J. Amer. Stat. Assoc.* 90(432), 1331–1340.
- [5] Jensen, D.R. (1996). Straight-line models in star-shaped mixtures. *Metrika* 44(2), 101–147.
- [6] Kamiya, H., A. Takemura and S. Kuriki (2008). Star-shaped distributions and their generalization. *J. Statist. Plann. Inference* 138(11), 3429–3447.
- [7] Kamiya, H. (2019). A note on estimation of the shape of density level sets of star-shaped distributions. *Comm. Statist. Theory Methods* 48(12), 3092–3104.
- [8] Lehmann, E. L. and G. Casella (1998). *Theory of Point Estimation*. Second edition. Springer, New York.
- [9] Liebscher, E. (2015). Goodness-of-approximation of copulas by a parametric family. In A. Steland, E. Rafajłowicz and K. Szajowski (Eds.), *Stochastic Models, Statistics and their Applications*, pp. 101–109. Springer, Cham.
- [10] Liebscher, E. (2016). Approximation of distributions by using the Anderson Darling statistic. *Comm. Statist. Theory Methods* 45(22), 6732–6745.
- [11] Liebscher, E. and W.-D. Richter (2016). Estimation of star-shaped distributions. *Risks* 4(4), 44.
- [12] Müller, K. and W.-D. Richter (2016). Extreme value distributions for dependent jointly $l_{n,p}$ -symmetrically distributed random variables. *Depend. Model.* 4, 30–62.
- [13] Richter, W.-D. (2007). Generalized spherical and simplicial coordinates. *J. Math. Anal. Appl.* 336(2), 1187–1202.
- [14] Richter, W.-D. (2011a). Circle numbers for star discs. *ISRN Geometry 2011*, Article ID 479262, 16 pages.
- [15] Richter, W.-D. (2011b). Ellipses numbers and geometric measure representations. *J. Appl. Anal.* 17(2), 165–179.
- [16] Richter, W.-D. (2014). Geometric disintegration and star-shaped distributions. *J. Stat. Distrib. Appl.* 1, 20.
- [17] Richter, W.-D. (2015a). Norm Contoured Distributions in \mathbb{R}^2 . S.I.M., Potenza.
- [18] Richter, W.-D. (2015b). Convex and radially concave contoured distributions. *J. Probab. Stat.* 2015, Article ID 165468, 12 pages.
- [19] Richter, W.-D. (2017). The class of $(p; q)$ -spherical distributions with an extension of the sector and circle number functions. *Risks* 5(3), 40.
- [20] Richter, W.-D. (2019a). High-dimensional star-shaped distributions. *J. Stat. Distrib. Appl.* 6, 5.
- [21] Richter, W.-D. (2019b). On ball numbers. *Mathematics* 7(8), 738.
- [22] Richter, W.-D. (2019c). On (p_1, \dots, p_k) -spherical distributions. *J. Stat. Distrib. Appl.* 6, 9.
- [23] Yang, Zh. and S. Kotz (2003). Center-similar distributions with applications in multivariate analysis. *Statist. Probab. Lett.* 64(4), 335–345.