risks

MDPI

Article

Estimation of Star-Shaped Distributions

Eckhard Liebscher [1,*] and Wolf-Dieter Richter [2]

[1] Department of Engineering and Natural Sciences, University of Applied Sciences Merseburg, 06217 Merseburg, Germany

[2] Institute of Mathematics, University of Rostock, Ulmenstraße 69, Haus 3, 18057 Rostock, Germany; wolf-dieter.richter@uni-rostock.de

* Correspondence: eckhard.liebscher@hs-merseburg.de; Tel.: +49-346-146-2860

Academic Editor: Meelis Käärik

Received: 2 September 2016; Accepted: 18 November 2016; Published: 30 November 2016

Abstract: Scatter plots of multivariate data sets motivate modeling of star-shaped distributions beyond elliptically contoured ones. We study properties of estimators for the density generator function, the star-generalized radius distribution and the density in a star-shaped distribution model. For the generator function and the star-generalized radius density, we consider a non-parametric kernel-type estimator. This estimator is combined with a parametric estimator for the contours which are assumed to follow a parametric model. Therefore, the semiparametric procedure features the flexibility of nonparametric estimators and the simple estimation and interpretation of parametric estimators. Alternatively, we consider pure parametric estimators for the density. For the semiparametric density estimator, we prove rates of uniform, almost sure convergence which coincide with the corresponding rates of one-dimensional kernel density estimators when excluding the center of the distribution. We show that the standardized density estimator is asymptotically normally distributed. Moreover, the almost sure convergence rate of the estimated distribution function of the star-generalized radius is derived. A particular new two-dimensional distribution class is adapted here to agricultural and financial data sets.

Keywords: star-shaped distributions; antinorm contoured distributions; norm contoured distributions; non-concentric elliptically contoured distributions; kernel density estimators

AMS Subject Classification: 60E05; 62G07; 62H12

1. Introduction

The classes of multivariate Gaussian and elliptically contoured distributions have served as the probabilistic basis of many multivariate statistical models over a period of several decades. Accounts of the theory of elliptically contoured distributions may be found in [1–3]. The book [4] by Fang and Anderson contains a big chapter about statistical inference of elliptically contoured distributions. The theory of elliptically contoured distributions including applications to portfolio theory is presented in the monograph by [5] Gupta et al. On combining the advantages of several estimators, semiparametric density estimators for elliptical distributions were derived in papers [6–8] by Stute and Werner, by Cui and He and by Liebscher. In [9] Battey and Linton considered a density estimator for elliptical distributions based on Gaussian mixture sieves. The performance of their estimators heavily depends on how the density can be approximated by a mixture of normal distributions. Scatter plots of multivariate data sets, however, motivate modeling of star-shaped distributions beyond elliptically contoured ones.

The more flexible star-shaped densities were studied in [10] and later in [11]. The general structure of their normalizing constant given a density generating function was discovered, a geometric measure representation and, based upon it a stochastic representation were derived, and a survey of applications

of such densities was given in [12]. Moreover, two-dimensional non-concentric elliptically contoured distributions are introduced there and, based upon two-dimensional star-shaped densities, a universal star-shaped generalization of the univariate von Mises density is derived. These results are further studied in detail in [13] for several particular classes. The big classes of norm and antinorm contoured distributions, being particular cases of star-shaped distributions, are considered in [14,15] for dimension two and for arbitrary finite dimension, respectively. In this paper we study several of those classes of distributions for arbitrary finite dimension and introduce a particular new class of distributions for dimension two. The rather general class of distributions considered in the present paper covers distributions with convex as well as such with non-convex contours.

The main goal of this paper is to develop estimation procedures for fitting multivariate generalized star-shaped distributions. The semiparametric procedure combines the flexibility of nonparametric estimators and the simple estimation and interpretation of parametric estimators. Since we apply nonparametric estimation to a univariate function, we avoid the disadvantages of nonparametric estimators in connection with the curse of dimensionality. The semiparametric approach of this paper is based on that in the earlier paper [7] on elliptical distributions but uses partially weaker assumptions. Alternatively, we consider a pure parametric method. In both cases, a parametric model is assumed for the density contours given by the star body and the Minkowski functional of it. For the semiparametric method, we assume that the contours are smooth, more precisely, that the Minkowski functional is continuously differentiable. The parameters are estimated using a method of moments. The star generalized radius density is estimated nonparametrically by use of a kernel density estimator, or parametrically.

The paper is structured as follows. The class of continuous star-shaped distributions and several of its subclasses are considered in Section 2. Section 3 deals with the estimation of the density and the star-generalized radius distribution. In Section 3, the statements on convergence rates and on the asymptotic normality of the density estimator as well as on the convergence rate of the estimated distribution function of the generalized radius are provided. First the case of a given star body is considered, later the more general case of a parametrized star body is taken into consideration. The particular Section 3.4 surveys on the one hand to a certain extent examples where different subclasses of star-shaped distributions appear in practice and deals on the other hand with applications of the methods developed here to the analysis of two-dimensional agricultural and financial data. The proofs can be found in Section 4.

2. Continuous Star-Shaped Distributions

2.1. The General Distribution Class

Throughout this paper, $K \subset \mathbb{R}^d$ denotes a star body, i.e., a non-empty star-shaped set that is compact and is equal to the closure of its interior, having the origin as an interior point. The Minkowski functional of K is defined by

$$h_K(x) = \inf\{\lambda \geq 0 : x \in \lambda K\} \text{ for } x \in \mathbb{R}^d.$$

The boundary of K is just the set $\{(u,v)^T : h_K((u,v)^T) = 1\}$. Further we find a ball $\{y \in \mathbb{R}^d : \|y\| \leq r\}$ which covers K where $||.||$ denotes Euclidean norm. Hence $h_K(x\,\|x\|^{-1}) \geq 1/r$ and

$$h_K(x) \geq \frac{1}{r}\,\|x\|\,. \tag{1}$$

The function h_K is assumed to be homogeneous of degree one,

$$h_K(\lambda x) = \lambda h_K(x) \text{ for } x \in \mathbb{R}^d, \lambda > 0,$$

and to satisfy a further assumption.

A countable collection $\mathfrak{F} = \{C_1, C_2, ...\}$ of pairwise disjoint sectors (closed convex cones C_j containing no half-space, with non-empty interior and vertex being the origin 0_d) such that $\mathbb{R}^d = \bigcup_j C_j$ will be called a fan. By $\mathfrak{B}_d$ we denote the Borel-σ-field in $\mathbb{R}^d$ and by S, the boundary of K. We denote $S_j = S \cap C_j$, $S_j \cap \mathfrak{B}_d = \mathfrak{B}_{S,j}$ and $\mathfrak{B}_S = \sigma\{\mathfrak{B}_{S,1}, \mathfrak{B}_{S,2}, ...\}$. We shall consider only star bodies K and sets $A \in \mathfrak{B}_S$ satisfying the following condition.

Assumption 1. *The star body K and the set $A \in \mathfrak{B}_S$ are chosen such that for every j the set*

$$G(A \cap S_j) = \{\vartheta \in \mathbb{R}^{d-1} : \exists \eta \text{ with } \theta = (\vartheta^T, \eta)^T \in A \cap S_j\}$$

is well defined and such that for every $\vartheta = (\vartheta_1, ..., \vartheta_{d-1})^T \in G(A \cap S_j)$ there is a uniquely determined $\eta > 0$ satisfying $h_K((\vartheta_1, ..., \vartheta_{d-1}, \eta)^T) = 1$.

A star body K satisfying this assumption will be called for short an $A1$-star-body.

Let $g : [0, +\infty) \to [0, +\infty)$ be a nonnegative function which fulfills the condition

$$0 < \int_{\mathbb{R}^d} g(h_K(x))\mathrm{d}x < \infty.$$

Such function is called a density generating function (dgf).

We consider the class of continuous star-shaped distributions of random vectors X taking values in $\mathbb{R}^d$:

$$CStSh^{(d)} \quad = \quad \{\Phi_{g,K,\mu}\colon \mu \in \mathbb{R}^d, K \text{ is an A1-star-body with } 0_d \in int\, K, g \text{ is a dgf}\}$$

where *int* K means the interior of K. Suppose that the distribution law $\Phi_{g,K,\mu}$ has the density

$$\varphi_{g,K,\mu}(x) = C(g,K)\, g\left(h_K(x-\mu)\right) \text{ for } x \in \mathbb{R}^d, \tag{2}$$

where $C(g,K)$ is a suitable normalizing constant. Moreover, K is called the contour defining star body of $\varphi_{g,K,\mu}$.

We consider the random vector X having the density (2) (in symbols $X \sim \Phi_{g,K,\mu}$). According to Theorem 8 in [12], this random vector has the representation

$$X - \mu \stackrel{d}{=} RU, \tag{3}$$

where the star-generalized radius variable $R = h_K(X-\mu)$ and the star-generalized uniform basis vector $U = \frac{1}{R}(X-\mu)$ are independent.

Moreover, R has the density

$$f_R(r) = \frac{1}{I(g,d)} r^{d-1} g(r) \tag{4}$$

with $I(g,d) = \int_0^\infty r^{d-1} g(r) dr$, and U has a star-generalized uniform probability distribution on the boundary of K, i.e., $P(U \in A) = \mathcal{O}_S(A)/\mathcal{O}_S(S)$ for $A \in \mathcal{B}_S$. According to [12], $\mathcal{O}_S$ means the star-generalized surface measure which is a non-Euclidean one unless for S being the Euclidean sphere, and which is well defined if Assumption 1 is fulfilled. Note that $\mathcal{O}_S(S) = d \cdot vol(K)$. If $\lim_{r \to 0+0} g(r) > 0$ is finite, then in view of (4), R takes values in the neighbourhood of zero with a rather small probability. This behaviour is called the volcano effect and is the stronger the higher the dimension is. The density (2) may be written as

$$\varphi_{g,K,\mu}(x) = \frac{1}{\mathcal{O}_S(S)} \left(h_K(x-\mu)\right)^{1-d} f_R\left(h_K(x-\mu)\right), \; x \in \mathbb{R}^d. \tag{5}$$

Estimating such density may be studied under various assumptions concerning the degree of knowledge on the groups of parameters $K, h_K, \mathcal{O}_S(S)$ and f_R, as well as μ. Let $X = (X^{(1)}, \ldots, X^{(d)})^T$ and $U = (U_1, \ldots, U_d)^T$. The next lemma gives helpful information about the mean and the covariances. Here and in what follows $\mathbf{1}\{A\}$ denotes the indicator function of an event A.

Lemma 1. *If* $\mathbb{E}R^2 < +\infty$ *and* K *is symmetric w.r.t. the origin, then*

$$\mathbb{E}X = \mu, \text{ and } \mathbb{C}\mathbf{ov}(X^{(j)}, X^{(k)}) = \mathbb{E}U_jU_k\,\mathbb{E}R^2 \text{ for } j,k = 1,\ldots,d.$$

Proof. In view of (3), we have first to show that $\mathbb{E}U = 0$. Because of the symmetry of K, U has the same distribution as $-U$, and $U_j = 0$ with probability 0 for each j. Thus we obtain

$$\begin{aligned} \mathbb{E}X\text{-}\mu &= \mathbb{E}R\;\mathbb{E}U = \mathbb{E}R\;\left(\mathbb{E}\left(U\,\mathbf{1}\left\{U_1 > 0\right\}\right) + \mathbb{E}\left(U\,\mathbf{1}\left\{U_1 < 0\right\}\right)\right) \\ &= \mathbb{E}R\;\left(\mathbb{E}\left(U\,\mathbf{1}\left\{U_1 > 0\right\}\right) + \mathbb{E}\left(-U\,\mathbf{1}\left\{U_1 < 0\right\}\right)\right) = 0. \end{aligned}$$

Moreover, it follows that

$$\mathbb{C}\mathbf{ov}(X^{(j)}, X^{(k)}) = \mathbb{E}\left(R^2U_jU_k\right) = \mathbb{E}U_jU_k\,\mathbb{E}R^2.$$

□

The general approach followed here includes non-convex bodies which can occur in applications. Obviously, $h_K(U) = 1$ and the distribution of the random vector U is concentrated on the set $\{u \in \mathbb{R}^d : h_K(u) = 1\}$.

2.2. A Class of Two-Dimensional Distributions Whose Contour Defining Star Bodies Are Squared Sine Transformed Euclidean Circles

We define $\alpha(u,v) \in (-\pi, \pi]$ to be the angle in radians between the positive x-axis and the line through the point (u,v) and the origin: $\alpha(u,v) = \arctan(v/u)$ for $u > 0$, $\alpha(u,v) = \arctan(v/u) + \operatorname{sgn}(v)\pi$ for $u < 0$, $\alpha(0,v) = \frac{\pi}{2} \cdot \operatorname{sgn}(v)$. The Minkowski functional of any two-dimensional star body K can then be written as

$$h_K((u,v)^T) = \sqrt{u^2+v^2}H(\alpha(u,v)),$$

where $H : (-\pi, \pi] \to (0, \infty)$ is a bounded function. In the following examples we consider two-dimensional star bodies with smooth boundaries; i.e., H is differentiable. Here, the following generator function is used

$$g_0(r) = \frac{1+r}{3}e^{-r}$$

which corresponds to the star-generalized radius density of mixed Erlang type

$$f_{0R}(r) = \frac{r+r^2}{3}e^{-r}.$$

Example 1. *Here we consider the Minkowski functional*

$$h_K((u,v)^T) = \sqrt{u^2+v^2}\left(1 + a\frac{v^2}{u^2+v^2}\right) = \sqrt{u^2+v^2}\left(1 + a\sin^2\alpha(u,v)\right)$$

where $a \in (-1, +\infty)$ *is a parameter. The Figures 1–4 show the contour lines of the boundary of the body for several values of a and the resulting density for one choice of a. These figures show that the distribution class includes densities with convex as well as with non-convex contours.*

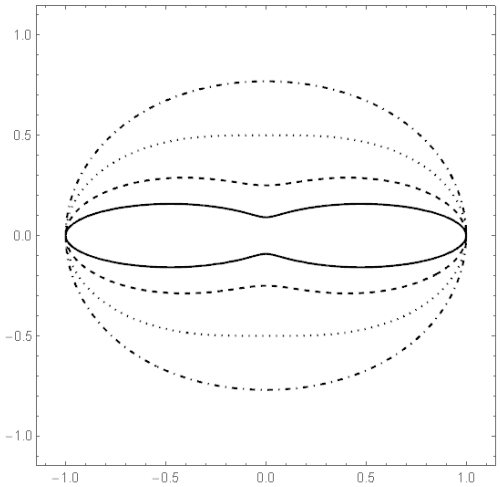

Figure 1. Plot of the boundaries of K in the cases $a = 10$ (solid line), $a = 3$ (dashed line), $a = 1$ (dotted line) and $a = 0.3$ (dashed/dotted line).

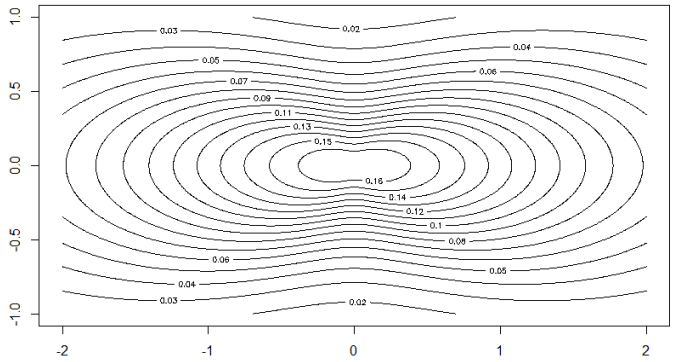

Figure 2. Plot of the boundaries of K in the cases $a = 10$ (solid line), $a = 3$ (dashed line), $a = 1$ (dotted line) and $a = 0.3$ (dashed/dotted line).

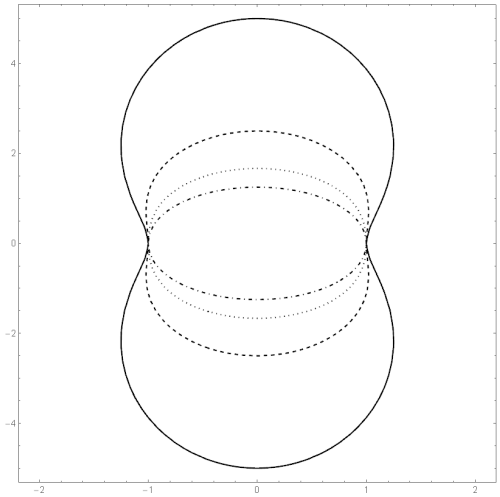

Figure 3. Plot of the boundaries of K in the cases $a = -0.8$ (solid line), $a = -0.6$ (dashed line), $a = -0.4$ (dotted line), $a = -0.2$ (dashed/dotted line).

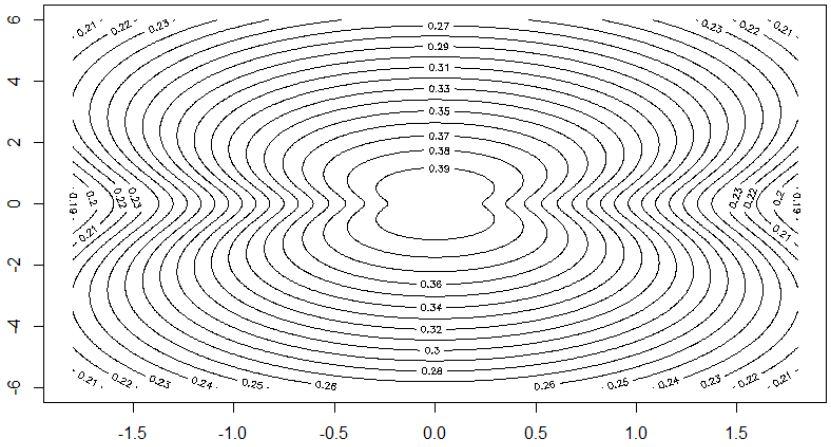

Figure 4. Contour plot of the density of $\varphi_{g,K,\mu}$ for $a = -0.8$ and levels $0.39, 0.38, \cdots 0.19$.

Example 2. *We consider the star body K with Minkowski functional*

$$h_K((u,v)^T) = \sqrt{u^2+v^2}\left(1 + a_1 \frac{(u - a_2 v)^2}{u^2+v^2}\right),$$

where a_1 and $a_2 \in \mathbb{R}$ are parameters such that $1 + a_1\left(1 + a_2^2\right) > 0$, and

$$H(\alpha) = 1 + a_1 \frac{(1 - a_2 \tan\alpha)^2}{1 + \tan^2\alpha}.$$

This star body arises from a rotation of K in Example 1 by an angle α_0 where $a_2 = 1/\tan\alpha_0$. In Figures 5 and 6 the boundaries of (multiples of) K are depicted.

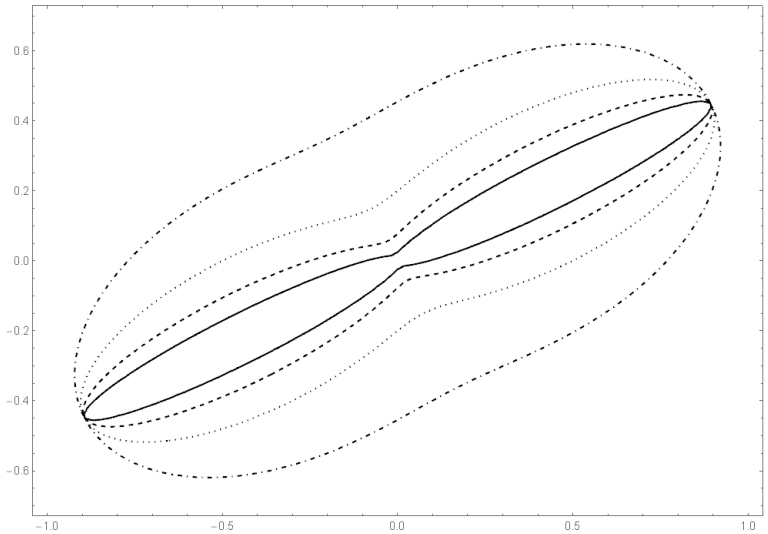

Figure 5. Plot of the boundaries of K for $a_2 = 2$ in the cases $a_1 = 10$ (solid line), $a_1 = 3$ (dashed line), $a_1 = 1$ (dotted line) and $a_1 = 0.3$ (dashed/dotted line).

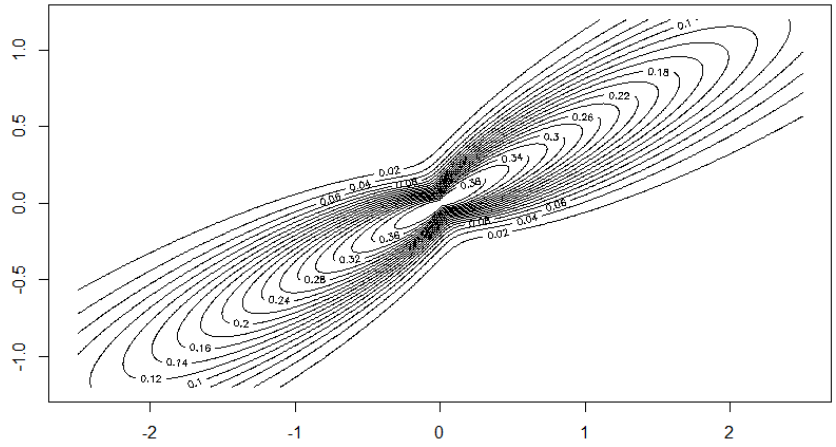

Figure 6. Contour plot of the density $\varphi_{g,K,\mu}$ for $a_1 = 3$ and levels as in Figure 4.

A specific motivation for considering a star body K as in Example 2 arises when studying the dataset 5 of [38], see Figure 10 below.

2.3. Norm-Contoured Distributions

Specific norm-contoured distributions were studied in several papers which are in part surveyed in Richter [14]. A geometric measure representation of arbitrary norm contoured distributions is proved in [15]. The class of all norm-contoured distributions is denoted, according to these papers, by $\mathcal{NC}$. The class $\mathcal{NC}$ is a subclass of the class *StSh* of star-shaped distributions. Here, we consider the subclass of continuous norm-contoured distributions *CNC*.

It is well known that there is a one-to-one correspondence between the class of convex bodies being symmetric w.r.t. the origin, where $x \in K$ implies $-x \in K$, and the class of norms in $\mathbb{R}^d$. If K is any such symmetric convex body then $h_K(x) = \|x\|$ where $\|.\|$ is the uniquely determined norm. On the other hand, if $\|.\|$ is any norm, then $K = \{x : \|x\| \leq 1\}$ is the corresponding convex symmetric body having the origin as interior point, and $\|x\| = h_K(x)$.

Throughout this section, let $\|.\|$ be any norm and $K = \{x : \|x\| \leq 1\}$, and let the density of a norm-contoured distribution be

$$\varphi(x; g, \|.\|, O, \mu) = C(g, \|.\|)\, g\left(\left\|O^{-1}(x-\mu)\right\|\right) \quad \text{for } x \in \mathbb{R}^d$$

where O is any orthogonal $d \times d$-matrix. Because any rotated or mirrored norm-ball is again a norm ball we shall restrict our attention to the case O being the unit matrix and will write then $X \sim CNC(g, \|.\|, \mu)$. In the present situation, $S = \{x : \|x\| = 1\}$ is considered to be the unit sphere in the Minkowski space $(\mathbb{R}^d, \|.\|)$.

In the following we consider several specific cases of norms and the corresponding norm-contoured distributions.

Example 3. *If K is the Euclidean unit ball then h_K is the Euclidean norm and (2) defines a shifted spherical distribution.*

Example 4. *Let $A \in \mathbb{R}^{d,d}$ be a $d \times d$-matrix satisfying* $\det(A) > 0$, $\|.\|^\circ$ *any norm, and*

$$\|x\| = \|Ax\|^\circ$$

another norm. If X is $\|.\|^\circ$-contoured distributed and A is a symmetric and positive definite then we call the distribution of AX an elliptically generalized $\|.\|^\circ$-contoured distribution.

Example 5. *Let $a = (a_1, \ldots, a_d)^T$ be a vector with $a_i > 0, i = 1, ..., d$ and $A = \mathrm{diag}(1/a_1, \ldots, 1/a_d)$. If, in Example 4, $\|.\|^\circ$ is the p-norm, $p \geq 1$, then the corresponding norm $\|.\|$ is*

$$\|x\| = \left(\sum_{i=1}^{d} \left| \frac{x_i}{a_i} \right|^p \right)^{1/p}.$$

The distribution of X is called an *axes aligned p-generalized elliptically contoured distribution*.

Example 6. *Assume that data are grouped into k groups, and let $k > 1, n = n_1 + \ldots + n_k, n_i \geq 1, p_i \geq 1$, $i = 1 \ldots k$ and*

$$h_K(x) = \sum_{j=1}^{k} \left(\sum_{i=1}^{d} \left| \frac{x_i}{a_i} \right|^{p_j} \right)^{1/p_j}.$$

K may be called then an $(a, p_1, \ldots, p_k)$-generalized axis-aligned ellipsoid and we will say that X follows a *grouped $(a, p_1, \ldots, p_k)$-generalized axis-aligned elliptically contoured distribution* in $\mathbb{R}^n$.

Example 7. *In the case of two-dimensional observations, let P_n denote the polygon having the n vertices $I_{n,i} = \left(\cos\left(\frac{2\pi}{n}(i-1)\right), \sin\left(\frac{2\pi}{n}(i-1)\right)\right)^T$, $i = 1, \ldots, n, n \geq 3$. The convex body which is circumscribed by P_n will be denoted by K. Then h_K is a norm defined in $\mathbb{R}^2$ and $\varphi_{g,K,0}$ a polygonally contoured density which was used implicitly in [13] to construct a corresponding geometric generalization of the von Mises density. For the more general class of multivariate polyhedral star-shaped distributions, see [16].*

Example 8. *Given a homogeneous polynomial p of degree k with $p(|x_1|, \ldots, |x_d|) \geq 0$, the function $N(x) := (p(|x_1|, \ldots, |x_d|))^{1/k}$ defines a norm in $\mathbb{R}^d$ if it is subadditive. An example for a homogeneous polynomial of degree 3 and $d = 2$ is $p(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 x_2$.*

2.4. Antinorm-Contoured Distributions

A function $g : \mathbb{R}^n \to [0, \infty)$ which is continuous, positively homogeneous, non-degenerate and superadditive in some fan is called an antinorm in [17]. Thereby, g is called superadditive in a sector C or in the fan $\mathfrak{F}$ if it satisfies the reverse triangle inequality in C or in every sector of the fan $\mathfrak{F}$, respectively.

Example 9. *If the (a, p)-functional $|.|_{a,p}$ is defined as $\|.\|$ in Example 5 but with $p \in (0, 1)$ then it is an antinorm.*

For geometric measure representations of elements from a big class of continuous antinorm contoured distributions we refer to [14,15]. For figures of two-dimensional antinorm balls, see [17]).

2.5. Continuous Non-Concentric Elliptically Contoured Distributions

Let $a = (a_1, \ldots, a_d)^T, a_i > 0$; $i = 1 \ldots n$ and $K_a = \{x \in \mathbb{R}^d : \sum_{i=1}^{d} (x_i/a_i)^2 \leq 1\}$. If $e = (e_1, ..., e_d)^T$ satisfies $\sum_{i=1}^{d} (e_i/a_i)^2 < 1$ then $K_{a,e} = K_a - e$ is a star body having the origin as an interior point, and

$$rK_{a,e} = \left\{ x \in \mathbb{R}^d : \sum_{i=1}^{d} \left(\frac{x_i + re_i}{ra_i} \right)^2 \leq 1 \right\} = K_{ra,re}.$$

Moreover, $r_1 K_{a,e} \subset r_2 K_{a,e}$ for $r_1 \leq r_2$. A Minkowski functional $h_{K_{a,e}}$ which is homogeneous of degree one will be called a non-concentric elliptically contoured function and $\varphi_{g,K_{a,e},\mu}$ a non-concentric elliptically contoured density. If $\mathcal{O} : \mathbb{R}^d \to \mathbb{R}^d$ denotes an arbitrary orthogonal transformation then

$h_{\mathcal{O}K_{a,e}}$ is also a non-concentric elliptically contoured function which is homogeneous of degree one. For the special case of $d = 2$ see [12,13].

3. Estimation for Continuous Star-Shaped Distributions

3.1. Parametric Estimators

Let $X_1, \ldots, X_n$ be a sample of independent random vectors, where $X_i \sim \Phi_{g,K,\mu}$ and $X_i = (X_{i1}, \ldots, X_{id})^T$. Assume that the star body K is given and Assumption 1 is satisfied. From now on, we suppose that K is symmetric w.r.t. the origin. We consider a model family $\{f_\theta : \theta \in \Theta_1\}$ of continuously differentiable densities for the star-generalized radius R on $[0, \infty)$, see (4). $\Theta = \Theta_1 \times \Theta_2, \Theta_1 \subset \mathbb{R}^q, \Theta_2 \subset \mathbb{R}^d$ is the parameter space which is assumed to be compact. Suppose that $h_K(.)$ is a continuous function.

Next we give two reasonable model classes for f_θ:

(1) Modified exponential model. $\theta = \tau \in (0, +\infty)$,

$$f_\tau(r) = \frac{1}{(d+1)(d-1)!}\tau^{-d} r^{d-1}\left(1 + \frac{r}{\tau}\right) e^{-r/\tau} \text{ for } r > 0$$

with the expectation

$$\int_0^\infty r f_\theta(r)\, \mathrm{d}r = \frac{d(d+2)\tau}{d+1}. \tag{6}$$

(2) Weibull model. $\theta = (\tau, a) \in (0, +\infty) \times (1, +\infty)$,

$$f_\theta(r) = \frac{a}{\tau^d \Gamma(d/a)} r^{d-1} e^{-(r/\tau)^a} \text{ for } r > 0$$

with the expectation

$$\int_0^\infty r f_\theta(r)\, \mathrm{d}r = \frac{\Gamma(\frac{d+1}{a})\tau}{\Gamma(\frac{d}{a})}.$$

Let $f_R \in \{f_\theta : \theta \in \Theta_1\}$. In this section the aim is to fit the specific parametric model for the density $\varphi_{g,K,\mu}$ to the data by estimating the parameters θ and μ where $\varphi_{g,K,\mu}$ is given according to (5) and (4) with $f_R = f_\theta$. Therefore, the two models [1] and [2] fulfill the condition $\lim_{r \to 0+0} g'(r) = 0$ which ensures the differentiability of the density $\varphi_{g,K,\mu}$ at zero.

For the statistical analysis we suppose that the data $X_1, \ldots, X_n$ are given and these data comprise independent random vectors having density $\varphi_{g,K,\mu}$. Suppose that θ and μ are interior points of Θ_1 and Θ_2, respectively. The concentrated log likelihood function (constant addends can be omitted) reads as follows

$$L(\theta, \mu) = \sum_{i=1}^n \left(\ln f_\theta(h_K(X_i - \mu)) + (1-d)\ln h_K(X_i - \mu)\right).$$

We introduce the maximum likelihood estimators $\hat{\theta}_n, \hat{\mu}_n$ of θ and μ as joint maximizers of the likelihood function:

$$L(\hat{\theta}_n, \hat{\mu}_n) = \max_{(\theta,\mu)\in\Theta} L(\theta, \mu).$$

Under appropriate assumptions, the maximum-likelihood-estimator are asymptotically normally distributed (cf. Theorem 5.1 in [18], p. 463)

$$\sqrt{n}(\hat{\theta}_n - \theta, \hat{\mu}_n - \mu)^T \xrightarrow{d} \mathcal{N}(0, I(\theta, \mu)^{-1}) \text{ for } n \to \infty,$$

where $\xrightarrow{d}$ is the symbol for convergence in distribution and the information matrix is given by $I(\theta,\mu) = (I_{ij}(\delta))_{i,j=1\ldots d+q}$ with $\delta^T = (\delta_1,\ldots,\delta_{d+q}) = (\theta^T,\mu^T)$ and

$$I_{ij}(\delta) = -\mathbb{E}\left(\frac{\partial}{\partial\delta_i\partial\delta_j}\left(\ln f_\theta(h_K(X_k-\mu)) + (1-d)\ln h_K(X_k-\mu)\right)\right).$$

3.2. Nonparametric Estimators without Scale Fit

In the present section we deal with nonparametric estimators in the context of star-shaped distributions. This type of estimators is of special interest if no suitable parametric model can be found. The cdf of R will be denoted by F_R.

3.2.1. Estimating μ and F_R

Let $X_1,\ldots,X_n$ be the sample as in Section 3.1. In the following the focus is on the estimation of the parameter μ and the distribution function of the generalized radius R.

First we choose an estimator for μ. For this purpose we assume that $\mathbb{E}|X| < +\infty$. In view of Lemma 1,

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{7}$$

is an unbiased estimator for the unknown parameter μ. Define $R_i = h_K(X_i-\mu)$ and $\hat{R}_i = h_K(X_i-\hat{\mu}_n)$ for $i = 1,\ldots,n$. Using this definition, an estimator for the cdf of $R = h_K(X-\mu)$ (cf. (3)) is given by the formula

$$\hat{F}_n^R(r) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{\hat{R}_i \le r\right\} \tag{8}$$

for $r \ge 0$. At a first glance, $\hat{F}_n^R(r)$ just approximates the empirical distribution function

$$F_n^R(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{R_i \le r\right\}$$

which is not available from the data because of the unknown μ. We can prove that $\hat{F}_n^R$ converges to F_R *a.s.*, in fact at the same rate as every common empirical distribution function converges to the cdf. This is the assertion of the following theorem.

Theorem 1. *Suppose that Assumption 1 is satisfied, $h_K(.)$ is Lipschitz-continuous on $\mathbb{R}^d$, and*

$$\int_0^\infty r^{d+1}g(r)\,dr < +\infty. \tag{9}$$

If further f is bounded on $[0,+\infty)$, then, for $n\to\infty$,

$$\sup_{r\ge 0}\left|\hat{F}_n^R(r) - F_R(r)\right| = O\left(\sqrt{\frac{\ln\ln n}{n}}\right) \quad a.s.$$

Here the condition (9) ensures that $\mathbb{E}R^2 < +\infty$ which in turn is an assumption for the law of iterated logarithm of $\hat{\mu}_n$.

3.2.2. Density Estimation

In the remainder of Section 3.2, we establish an estimator for the density $\varphi_{g,K,\mu}$ in the case of a bounded generator function g, and provide statements on convergence properties of the estimator. An estimator for μ is available by Formula (7), the estimation of g is still an open problem. If we

want to estimate g, then it is necessary that this function is identifiable. In (2), however, function g is determined up to a constant factor. Therefore, we require $I(g,d) = 1$ to obtain the uniqueness and identifiability. As a consequence, we get, according to [12]

$$C(g,K) = \frac{1}{\mathcal{O}_S(S)}.$$

In the following we adopt the approach introduced in Section 2 of [7] to the present much more general situation. This approach combines the advantages of two estimators and avoids their disadvantages. Let $\psi : [0,\infty) \to [0,+\infty)$ be a function having a derivative ψ' with $\psi'(y) > 0$ for $y \geq 0$, and the property $\psi(0) = 0$. We introduce the random variable $Y = \psi(h_K(X-\mu))$ and denote the inverse function of ψ by Ψ. The transformation using ψ is applied to adjust the volcano effect described above. In view of (4), the density χ of $Y = \psi(R)$ is given by

$$\chi(y) = \Psi(y)^{d-1} g(\Psi(y)) \cdot \Psi'(y)$$

for $y \geq 0$. This equation implies the following formula for g:

$$g(z) = z^{1-d}\ \psi'(z)\ \chi(\psi(z)).$$

The next step is to establish the estimator for χ. Nonparametric estimators have the advantage that they are flexible and there is no need to assume a specific model. Let us consider the transformed sample $Y_{1n},\ldots,Y_{nn}$ with $Y_{in} = \psi(\hat{R}_i)$. Further we apply the following kernel density estimator for χ:

$$\hat{\chi}_n(y) = n^{-1}b^{-1}\sum_{i=1}^{n}\left(k\left((y-Y_{in})b^{-1}\right) + k\left((y+Y_{in})b^{-1}\right)\right) \quad \text{for } y \geq 0, \tag{10}$$

where $b = b(n)$ is the bandwidth and k the kernel function. Note that $\hat{\chi}_n$ represents the usual kernel density estimator for χ based upon the Y_{in}'s and including a boundary correction at zero (the second addend in the outer parentheses of (10)). The mirror rule is used as a simple boundary correction. Other more elegant corrections can be applied at the price of a higher technical effort. The properties of $\hat{\chi}_n$ are essentially influenced by the bandwidth b. Since the kernel estimator shows reasonable properties only in the case of bounded χ, we have to guarantee by suitable assumptions that $\lim_{z\to 0+} z^{1-d}\ \psi'(z) > 0$ in order to get the boundedness of χ (see below). On the basis of $\hat{\chi}_n$, we can establish the following estimator for $\varphi_{g,K,\mu}$:

$$\hat{\varphi}_n(x) = \mathcal{O}_S(S)^{-1}\ \hat{g}_n\left(h_K(x-\hat{\mu}_n)\right),$$

where

$$\hat{g}_n(z) = z^{1-d}\ \psi'(z)\ \hat{\chi}_n(\psi(z)). \tag{11}$$

This approach has the property that the theory of kernel density estimators applies here (cf. [19]). The kernel estimators are a very popular type of nonparametric density estimators because of their comparatively simple structure. In the literature the reader can find a lot of hints concerning the choice of the bandwidth.

Let us add here some words to the comparison between this paper and [7]. Although the main idea for the construction of estimators is the same, there is a difference in the definition of the generator functions (say g and g_L). Considering the special case $h_K(x) = \left\|\Sigma^{-1/2}(x-\mu)\right\|$, identity $g(t) = g_L(t^2)$ can be established for $t \geq 0$. This causes some changes in the formulas. For more details in a particular case, see Section 3 in [20].

3.2.3. Assumptions Ensuring Convergence Properties of Estimators

Next we provide the assumptions for the theorems below. Assumption 2 concerns the parameter $b = b(n)$ and the function k of the kernel estimator whereas Assumption 3 is posed on function ψ.

Assumption 2. *(a) We assume that*

$$\lim_{n\to\infty} b \ln\ln n = 0 \quad \text{and } \bar{b} \geq b \geq C_1 \cdot n^{-1/5}$$

with constants $\bar{b}, C_1 > 0$.

(b) Suppose that the kernel function $k : \mathbb{R} \longrightarrow \mathbb{R}$ *is continuous and vanishes outside the interval* $[-1, 1]$, *and has a Lipschitz continuous derivative on* $[-1, 1]$. *Moreover, assume that* $k(-t) = k(t)$ *holds for* $t \in [-1, 1]$,

$$\int_{-1}^{1} k(t)\, dt = 1 \text{ and } \int_{-1}^{1} t^j k(t)\, dt = 0 \tag{12}$$

for even $j : 0 < j < p$, *where* $p \geq 2$ *is an integer.*

Note that continuity of the derivative at an enclosed boundary point means that the one-sided derivative exists and is the limit of the derivatives in a neighbourhood of this point. Symmetric kernel functions k satisfying (12) are called kernels of order p. Assumption 2 ensures that the bias of the density estimator $\hat{\chi}_n$ converges to zero at a certain rate. Under Assumption 2 with $p = 2$ and $k(t) \geq 0$, the estimator $\hat{\chi}_n$ is indeed a density. The case $p > 2$ is added to complete the presentation and is of minor practical importance unless we have a very large sample size (cf. the discussion in [21]). From the asymptotic theory for density estimators, it is known that the Epanechnikov kernel

$$k_{\text{epa}}(t) = \begin{cases} \frac{3}{4}\left(1 - t^2\right) & \text{for } t \in [-1, 1], \\ 0 \text{ otherwise} \end{cases}$$

is an optimal kernel of order 2 (i.e., in the case $p = 2$ in Assumption 2) with respect to the asymptotic mean square error (cf. [19]). This kernel function is simple in structure and leads to fast computations. The consideration of optimal kernels can be extended to higher-order kernels. It turned out that their use is advantageous only in the case of sufficiently large sample sizes (for instance, for a size greater than 1000).

Assumption 3. *The* $(p+1)$*-th order derivative of* Ψ^d *exists and is continuous on* $[0, \infty)$, ψ' *is positive and bounded on* $(0, +\infty)$ *for some integer* $p \geq 2$, *and* ψ'' *is bounded on* $(0, +\infty)$. *The functions* $z \rightsquigarrow z^{d-1}\psi'(z)^{-1}$ *and* $z \rightsquigarrow z^{-1}\psi'(z)$ *have bounded derivatives on* $[0, M_1]$ *with some* $M_1 > 0$. *Moreover,*

$$\lim_{t\to 0+} \left(\Psi^d(t)\right)' > 0. \tag{13}$$

Notice that in Assumption 3 we require that the right-sided limit of the $(p+1)$-th order derivative of Ψ^d is finite at zero. Hence Assumption 3 implies that

$$\begin{aligned} \Psi^d(t) &= (\Psi^d(\tilde{t}))' t, \, \tilde{t} \in (0, t), \\ \lim_{t\to 0+} t^{-1/d}\Psi(t) &= C_2, \end{aligned} \tag{14}$$

and

$$\lim_{z\to 0+} z^{-d}\psi(z) = C_2^{-d}$$

with a finite constant $C_2 > 0$. On the other hand, it follows from (13) that

$$\lim_{z\to 0+} z^{1-d}\,\psi'(z) = \lim_{t\to 0+} \Psi(t)^{1-d}\Psi'(t)^{-1} = d \lim_{t\to 0+} \left(\left(\Psi^d(t)\right)'\right)^{-1} = C_3$$

with a finite constant $C_3 > 0$. Therefore, χ is bounded under Assumption 3.

Example 10. *Let*

$$\psi(z) = -a + \left(a^d + z^d\right)^{1/d}$$

with a constant $a > 0$. Then $\Psi^d(t) = (t+a)^d - a^d$, $\left(\Psi^d(t)\right)' = d\,(t+a)^{d-1}$,

$$\begin{aligned} z^{-1}\psi'(z) &= z^{d-2}\left(a^d + z^d\right)^{-1+1/d},\ z^{d-1}\psi'(z)^{-1} = \left(a^d + z^d\right)^{1-1/d}, \\ \left(z^{-1}\psi'(z)\right)' &= \left(a^d + z^d\right)^{1/d-2} z^{d-3}\left(a^d d - 2a^d - z^d\right), \quad \textit{and} \\ \left(z^{d-1}\psi'(z)^{-1}\right)' &= (d-1)\,z^{d-1}\left(a^d + z^d\right)^{-1/d}. \end{aligned}$$

Hence, Assumption 3 is satisfied for every $p \geq 2$.

Another condition is required now for h_K.

Assumption 4. *For any bounded subset Q of $\mathbb{R}^d, 0 \notin Q$, the partial derivatives $G_1, \ldots, G_d$ of $h_K(.)$ exist and are bounded on Q, and $x \rightsquigarrow \psi'(h_K(x))G_j(x)$ is Hölder continuous of order $\alpha > 0.2$ on Q for each $j \in \{1, \ldots, d\}$.*

Assumption 5. *For any bounded subset Q of $\mathbb{R}^d, 0 \notin Q$, h_K is Hölder continuous of order $\bar{\alpha} > 0.2$.*

If Assumption 3 is fulfilled, the function h_K has second order derivatives $\frac{\partial^2}{\partial x_j \partial x_J} h_K(x) = G_{jJ}(x)$, and these are bounded on bounded subsets of $\mathbb{R}^d$, then the Assumption 4 is satisfied.

Example 11. *We consider the q-norm/antinorm: $h_K(x) = \|x\|_q$,*

$$\begin{aligned} G_j(x) &= \frac{x_j |x_j|^{q-2}}{\|x\|_q^{q-1}} \ \textit{for } j = 1, \ldots, d, \\ G_{jJ}(x) &= \begin{cases} \frac{(1-q)x_j|x_j|^{q-2}x_J|x_J|^{q-2}}{\|x\|_q^{2q-1}} \ \textit{for } j \neq J, \\ \frac{(q-1)}{\|x\|_q^{2q-1}}|x_j|^{q-2}\sum_{\nu\neq j}|x_\nu|^q \ \textit{for } j = J. \end{cases} \end{aligned}$$

Therefore, Assumption 4 is fulfilled in the case $q > 1.2$, and Assumption 5 is fulfilled in the case $q > 0.2$.

Examples 1 and 2: (Continued) One can show that $G_j(x)$ exists for $x \neq 0, j \in \{1, \ldots, d\}$, and is bounded. Moreover, $\psi'(h_K(.))G_j$ is Lipschitz continuous on $\mathbb{R}^d \backslash \{0\}$. Hence, Assumption 4 is satisfied.

3.2.4. Properties of the Density Estimator

First we provide the result on strong convergence of the density estimator.

Theorem 2. *Suppose that the p-th order derivative $g^{(p)}$ of g exists and is bounded on $[0, \infty)$ for some even integer $p \geq 2$. Moreover, assume that condition (9) as well as Assumptions 1 to 3 are satisfied for the given p.*

Let Assumption 4 or Assumption 5 be satisfied. In the first case define $r_n := \sqrt{\ln n}(nb)^{-1/2}$, *in the latter case* $r_n := n^{-\bar{\alpha}/2}b^{-1}$. *Then, for any compact set* D *with* $\mu \notin D$ *and* $n \to \infty$,

$$\sup_{x\in D} \left|\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x)\right| = O\left(r_n + b^p\right) \quad a.s. \tag{15}$$

For any compact set D *with* $\mu \in D$ *and* $n \to \infty$,

$$\sup_{x\in D} \left|\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x)\right| = O\left(r_n + b^{1/d}\right) \quad a.s.$$

Theorem 2 applies in particular to the Euclidean case $h_K = \|.\|_2$. Since Assumption 2 is weaker than the corresponding assumption on the kernel in [7], Theorem 2 extends Theorem 3.1 in [7] even in the case of $h_K = \|.\|_2$. The convergence rate in (15) is the same as that known for one-dimensional kernel density estimators and cannot be improved under the assumptions posed here (cf. [22]).

The next theorem represents the result about the asymptotic normality of the estimator $\hat{\varphi}_n$.

Theorem 3. *Suppose that the assumptions of Theorem 2 and Assumption 4 are satisfied. Let* $x \in \mathbb{R}^d, x \neq \mu$ *such that* $g^{(p)}$ *is continuous at* $\tilde{x} := h_K(x - \mu)$.

(i) Define

$$\begin{aligned} \bar{\sigma}^2(\tilde{x}) &= \mathcal{O}_S(S)^{-2}\tilde{x}^{1-d}\psi'(\tilde{x})g(\tilde{x})\int_{-1}^{1} k^2(t)\,dt, \\ \Lambda(\tilde{x}) &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\frac{1}{p!}\chi^{(p)}(\psi(\tilde{x}))\int_{-1}^{1} t^p k(t)\,dt. \end{aligned}$$

Then

$$\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x) = Z_n + e_n,$$

where $e_n = \Lambda(\tilde{x})\, b^p + o(b^p)$,

$$\sqrt{nb}Z_n \xrightarrow{d} \mathcal{N}(0, \bar{\sigma}^2(\tilde{x})) \text{ for } n \to \infty.$$

(ii) If additionally $\lim_{n\to\infty} n^{1/(2p+1)}b = C_4$ *holds true with a constant* $C_4 \geq 0$, *then, for* $n \to \infty$,

$$\sqrt{nb}\left(\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x)\right) \xrightarrow{d} \mathcal{N}(C_4^{(2p+1)/2}\Lambda(\tilde{x}), \bar{\sigma}^2(\tilde{x})).$$

The assertion of Theorem 3 can be used to construct an asymptotic confidence region for $\varphi_{g,K,\mu}(x)$. Term e_n describes the asymptotic behaviour of the the bias of the estimator $\hat{\varphi}_n$ whereas the fluctuations of the estimator are represented by Z_n. In view of Theorem 3, $\sqrt{nb}Z_n$ converges in distribution to $Z \sim \mathcal{N}(0, \bar{\sigma}^2(\tilde{x}))$. The mean squared deviation of the leading terms in the asymptotic expansion of $\hat{\varphi}_n$ is thus given by

$$\mathbb{E}\left(n^{-1/2}b^{-1/2}Z + \Lambda(\tilde{x})\, b^p\right)^2 = n^{-1}b^{-1}\bar{\sigma}^2(\tilde{x}) + \Lambda^2(\tilde{x})\, b^{2p}.$$

The minimization of this function w.r.t. b leads to the asymptotically optimal bandwidth

$$b^* = \left(\frac{\bar{\sigma}^2(\tilde{x})}{2p\Lambda^2(\tilde{x})n}\right)^{1/(2p+1)}. \tag{16}$$

The bandwidth b^* converges at rate $n^{-1/(2p+1)}$ to zero. Under the conditions of Theorem 3(ii), $\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x)$ has the convergence rate $n^{-p/(2p+1)}$. This convergence rate of $\hat{\varphi}_n$ is better than that of a nonparametric density estimator but slower than the usual rate $n^{-1/2}$ for parametric estimators. In principle, Formula (16) could be used for the optimal choice of the bandwidth. However, one would need then an estimator for $\chi^{(p)}$ and typically, estimators of derivatives of densities do not exhibit a

good performance unless n is very large. As a resort, one can consider a bandwidth which makes reference to a specific radius distribution.

To illustrate how the estimators work in practice, we simulated data from a q-norm distribution with $q = 1.3$ and the radius distribution to be the modified exponential distribution with $\tau = 1.1$. The Figures 7 and 8 include graphs of the underlying function g and its estimator in two cases.

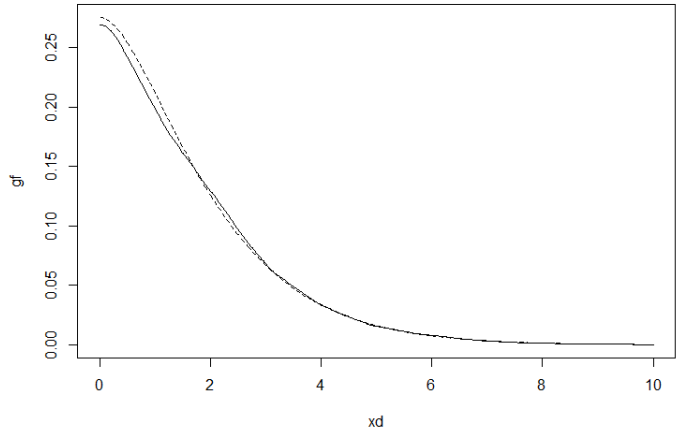

Figure 7. Estimator of g (solid line) and the model function (dashed line) for $n = 1000$.

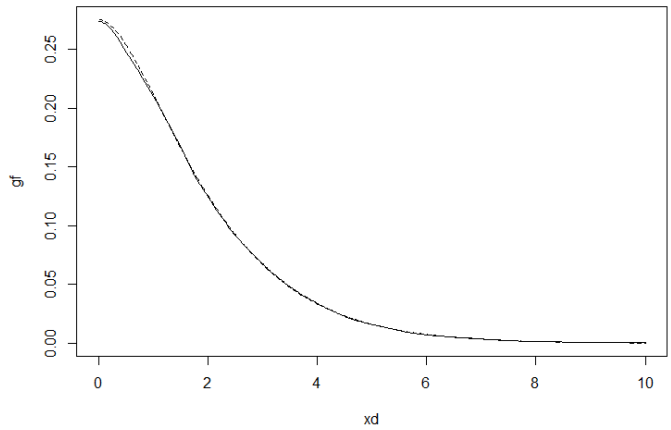

Figure 8. Estimator of g (solid line) and the model function (dashed line) for $n = 10,000$.

3.2.5. Reference Bandwidth

Let us consider an estimator $\hat{\varphi}_n(x)$ with Epanechnikov kernel, function ψ as in Example 10, and modified exponential radius density in the case $p = 2$. According to (16), the reference bandwidth is then

$$b^* = \left(15(d+1)(d-1)!e^{\tilde{x}/\tau}\tilde{x}^{4d}\tau^{5/d}(1+\tilde{x}^d)^{-1+5/d}(\tilde{x}+\tau)\ \left(\mathcal{O}_S(S)\,D^2 n\right)^{-1}\right)^{1/5},$$

where

$$\begin{aligned} D &= \tilde{x}^2(\tilde{x}+(d-2)\tau)+\tilde{x}^{d+2}(2\tilde{x}-(d+1)\tau) \\ &\quad +\tilde{x}^{2d}\left(\tilde{x}^3+(d-2)(d-1)(\tilde{x}+\tau)\tau^2+\tilde{x}^2\tau(1-2d)\right). \end{aligned}$$

This formula was generated using the computer algebra system Mathematica. The parameter τ can be estimated by utilizing the above Formula (6) for the expectation of the radius.

3.3. Semiparametric Estimators Involving a Scale and a Parameter Fit

In this section we consider the situation where the contour of the body K depends on scale parameters $\sigma_1, \ldots, \sigma_d$. Suppose that $I(g,d) = 1$. We introduce the diagonal matrix $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_d)$ and a master body K_0, which is symmetric w.r.t. the origin. Define

$$K = \Sigma K_0 := \{\Sigma x : x \in K_0\},$$

and $\tilde{U} = \Sigma^{-1} U$. The distribution of $\tilde{U}$ is concentrated on the boundary S_0 of K_0. We assume that K_0 is given such that

$$\mathbb{E}\tilde{U}_j^2 = 1 \text{ for } j = 1, \ldots, d. \tag{17}$$

Otherwise, K_0 is rescaled. Suppose that h_{K_0} depends on a further parameter vector $\theta \in \Theta$ where the parameter space $\Theta \subset \mathbb{R}^q$ is a compact set. Then

$$h_K(x) = h_{K_0}(\theta, \Sigma^{-1} x) \text{ for } x \in \mathbb{R}^d.$$

The parameter vector θ is able to describe the shape of the boundary of body K, see Examples 1 and 2 (parameters a_1 and a_2). From Lemma 1, we obtain

$$\begin{aligned} \mathbb{V}(X^{(j)}) &= \sigma_j^2 \mathbb{E}\tilde{U}_j^2 \, \mathbb{E}R^2 = \sigma_j^2 \, \mathbb{E}R^2, \\ \mathbb{C}\mathbf{ov}(X^{(j)}, X^{(k)}) &= \sigma_j \sigma_k \mathbb{E}\tilde{U}_j \tilde{U}_k \, \mathbb{E}R^2, \end{aligned}$$

and

$$\rho_{jk} = \mathbb{C}\mathbf{orr}(X^{(j)}, X^{(k)}) = \mathbb{E}\tilde{U}_j \tilde{U}_k \text{ for } j,k = 1, \ldots, d.$$

Here we see that (17) results in $\mathbb{V}(X^{(j)}) = \sigma_j^2 \, \mathbb{E}R^2$. The density is given by

$$\varphi_{g,K,\mu}(x) = \mathcal{O}_{S_0}(S_0)^{-1} \det(\Sigma)^{-1} g\left(h_{K_0}(\theta, \Sigma^{-1}(x-\mu))\right), x \in \mathbb{R}^d.$$

In this context, a scaling problem occurs concerning g. Assume that g is a suitably given generator function satisfying $I(g,d) = 1$. Then $x \rightsquigarrow g_t^*(x) := t^{d-1} g(tx)$ is a modified generator for every $t \in \mathbb{R}$ with $I(g_t^*, d) = 1$. For any $t \in \mathbb{R}$, we obtain the same model when g is replaced by g_t^* and σ_j is replaced by $\sigma_j t$ for $j = 1, \ldots, d$. To get uniqueness, we choose t such that

$$\mathbb{E}R^2 = \int_0^\infty r^{d+1} g_t^*(r) \, dr = t^{-2} \int_0^\infty r^{d+1} g(r) \, dr = 1.$$

Let $\hat{\mu}_n = (\hat{\mu}_{n1}, \ldots, \hat{\mu}_{nd})^T$ as above. Then σ_j^2 represents the variance of the j-th component of X. Based on this property, the sample variances of the components of X can be used as estimators for σ_j^2:

$$\hat{\sigma}_{nj}^2 = \frac{1}{n-1} \sum_{i=1}^n \left(X_{ij} - \hat{\mu}_{nj}\right)^2 \text{ for } j = 1, \ldots, d.$$

Moreover, we have the sample correlations

$$\hat{\rho}_{njk} = \hat{\sigma}_j^{-1} \hat{\sigma}_k^{-1} \, \frac{1}{n-1} \sum_{i=1}^n \left(X_{ij} - \hat{\mu}_{nj}\right) \left(X_{ik} - \hat{\mu}_{nk}\right) \text{ for } j,k = 1, \ldots, d.$$

In the following we use the notation $\hat{\Sigma}_n = \mathrm{diag}(\hat{\sigma}_{n1}, \ldots, \hat{\sigma}_{nd})$. If θ is unknown, we consider moment estimators based on the correlations. For this we need the following assumptions.

Assumption 6. *Let I be a subset of $\{(j,k) : j,k = 1,\dots,d, j < k\}$ with cardinality q. There is a vector $\rho = (\rho_{jk})_{(j,k)\in I} \in \mathbb{R}^q$ such that for $l = 1,\dots,q$,*

$$\theta_l = \gamma_l(\rho).$$

Assume that $\gamma_l : \mathcal{C} \to \Theta$ has bounded partial derivatives, $\rho \in \mathcal{C}$ and θ is an interior point of Θ.

Assumption 7. *For any bounded subset Q of $\mathbb{R}^d, 0 \notin Q$, the partial derivatives $\tilde{G}_1,\dots,\tilde{G}_q$, $G_1,\dots,G_d$ of $(\theta,x) \rightsquigarrow h_{K_0}(\theta,x)$ exist, are bounded for $\theta \in \Theta, x \in Q$, and $(\theta,x) \rightsquigarrow \psi'(h_{K_0}(\theta,x))\, G_j(\theta,x)$, $\psi'(h_{K_0}(\theta,x))\tilde{G}_j(\theta,x)$ is Hölder continuous of order $\alpha > 0.2$ on $\Theta \times Q$ for each $j \in \{1,\dots,d\}$.*

Assumption 8. *For any bounded subset Q of $\mathbb{R}^d, 0 \notin Q$, h_{K_0} is Hölder continuous of order $\bar{\alpha} > 0.2$.*

Examples 1 and 2: (Continued) Similarly as above, it can be proven that Assumption 7 is fulfilled. Let $\hat{\rho}$ be the sample version of ρ. Then

$$\hat{\theta}_{nl} = \gamma_l(\hat{\rho}) \text{ for } l = 1,\dots,q$$

is the estimator for θ, $\hat{\theta}_n = (\hat{\theta}_{nl})_{l=1\dots q}$. Define $\hat{R}_i = \hat{\Sigma}^{-1}(X_i - \hat{\mu}_n)$. With this definition, $\hat{F}_n^R$ is determined according to Formula (8). The following result on the convergence rate of $\hat{F}_n^R$ can be proven:

Theorem 4. *Suppose that Assumptions 1 and 6 are satisfied, and*

$$\int_0^\infty r^{d+3} g(r)\, dr < +\infty. \tag{18}$$

Let $r \rightsquigarrow r f_R(r)$ be bounded on $[0,+\infty)$. Then, for $n \to \infty$,

$$\sup_{r\geq 0} \left| \hat{F}_n^R(r) - F_R(r) \right| = O\left(\sqrt{\frac{\ln\ln(n)}{n}} \right) \quad a.s.$$

In this section the transformed sample $Y_{1n},\dots,Y_{nn}$ is given by $Y_{in} = \psi(h_{K_0}(\hat{\theta}_n, \hat{R}_i))$ with ψ as in Section 3.2. The estimator $\hat{g}_n$ for the generator g is calculated using Formulas (10) and (11) from the previous section. The following estimator for the density has thus been established:

$$\hat{\varphi}_n(x) = \mathcal{O}_{S_0}(S_0)^{-1} \det(\hat{\Sigma}_n)^{-1} \hat{g}_n \left(h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(x - \hat{\mu}_n)) \right) \tag{19}$$

The next two theorems show the results concerning strong convergence and asymptotic normality of the density estimator:

Theorem 5. *Suppose that the p-th order derivative $g^{(p)}$ of g exists and is bounded on $[0,\infty)$ for some even integer $p \geq 2$. Where needed, with this p, assume further that Assumptions 1, 2, 3, 6, (1) and (18) are satisfied. Let Assumption 7 or Assumption 8 be satisfied, and define in the first case $r_n := \sqrt{\ln n}(nb)^{-1/2}$ and in the latter case $r_n := n^{-\bar{\alpha}/2} b^{-1}$.Then the claim of Theorem 2 holds true for estimator $\hat{\varphi}_n$ defined in (19).*

Theorem 6. *Suppose that the assumptions of Theorem 5 are satisfied. Let $x \in \mathbb{R}^d, x \neq \mu$ such that $g^{(p)}$ is continuous at $\tilde{x} := h_{K_0}(\theta, \Sigma^{-1}(x-\mu))$. Assume that $\lim_{n\to\infty} n^{1/(2p+1)} b = C_4$ holds true with a constant $C_4 \geq 0$. Then for $n \to \infty$,*

$$\sqrt{nb}\left(\hat{\varphi}_n(x) - \varphi_{g,K,\mu}(x)\right) \xrightarrow{d} \mathcal{N}(C_4^{(2p+1)/2}\Lambda(\tilde{x}), \bar{\sigma}^2(\tilde{x})),$$

where $\hat{\varphi}_n$ is defined in (19),

$$\begin{aligned}\bar{\sigma}^2(\tilde{x}) &= \mathcal{O}_{S_0}(S_0)^{-2}\det(\Sigma)^{-2}\tilde{x}^{1-d}\psi'(\tilde{x})g(\tilde{x})\int_{-1}^{1}k^2(t)\,dt,\\ \Lambda(\tilde{x}) &= \mathcal{O}_{S_0}(S_0)^{-1}\det(\Sigma)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\frac{1}{p!}\chi^{(p)}(\psi(\tilde{x}))\int_{-1}^{1}t^p k(t)\,dt.\end{aligned}$$

The remarks following Theorems 2 and 3 are valid similarly.

3.4. Applications

For many decades, statistical applications of multivariate distribution theory were manly based upon Gaussian and elliptically contoured distributions. Studies using non-elliptically contoured star-shaped distributions were basically made during the last two decades and are dealing in most cases with p-generalized normal distributions. Such distributions are convex or radially concave contoured if $p \geq 1$ or $0 < p \leq 1$, respectively, and are also called power exponential distributions. Moreover, common elliptically contoured power exponential (ecpe) distributions build a particular class of the wide class of star-shaped distributions that allows modeling much more flexible contours than elliptically ones.

The class of ecpe distributions is used in a crossover trial on insulin applied to rabbits in [23], in image denoising in [24] and in colour texture retrieval in [25]. Applications of multivariate g-and-h distributions to jointly modeling body mass index and lean body mass are demonstrated in [26] and accompanied by star-shaped contoured density illustrations. The $l_{n,p}$-elliptically contoured distributions build another big class of star-shaped distributions and are used in [27] to explore to which extent orientation selectivity and contrast gain control can be used to model the statistics of natural images. Mixtures of ecpe distributions are considered for bioinformation data sets in [28]. Texture retrieval using the p-generalized Gaussian densities is studied in [29]. A random vector modeling data from quantitative genetics presented in [30] are shown in [31] to be more likely to have a power exponential distribution different from a normal one. The reconstruction of the signal induced by cosmic strings in the cosmic microwave background, from radio-interferometric data, is made in [32] based upon generalized Gaussian distributions. These distributions are also used in [33] for voice detection.

More recently, the considerations in [11] opened a new field of financial applications of more general star-shaped asymptotic distributions, where suitably scaled sample clouds converge onto a deterministic set.

Figure 3 in [34] represents a sample cloud which might be modeled with a density being star-shaped w.r.t. a fan having six cones that include sample points and other cones that do not. Note, however, that Figure 1 d-f in the same paper do not reflect a homogeneous density but might be compared in some sense to the level sets of the characteristic functions of certain polyhedral star-shaped distributions in [16], Figure 5.2.

When modeling Lymphoma data, [35] analyze sample clouds of points, see Figures 2 and 3, which might be interpreted as mixtures of densities having contours in part looking similar like that in [36] where flow cytometric data, Australian Institute of Sport data and Iris data are analyzed, or like that of certain skewed densities as they were (analytically derived and) drawn in [37]. In a similar manner, Figures 2 and 5 in [20] indicate that mixtures of different types of star-shaped distributions might be suitable for modeling residuals of certain stock exchange indices. It could be of interest to closer study in future work more possible connections of all the models behind.

The following numerical examples of the present section are aimed to illustrate the agricultural and financial application of the estimators described in this paper. To this end, we make use of the new particular non-elliptically contoured but star-shaped distribution class introduced in Section 2.2 of the present paper.

Example 12. *Example 2 of Section 2.2 continued. We consider the class of bodies K of Example 2. Let* $a_2 = 1$. *Figure 9 shows the dependence of the correlation on the parameter* a_1.

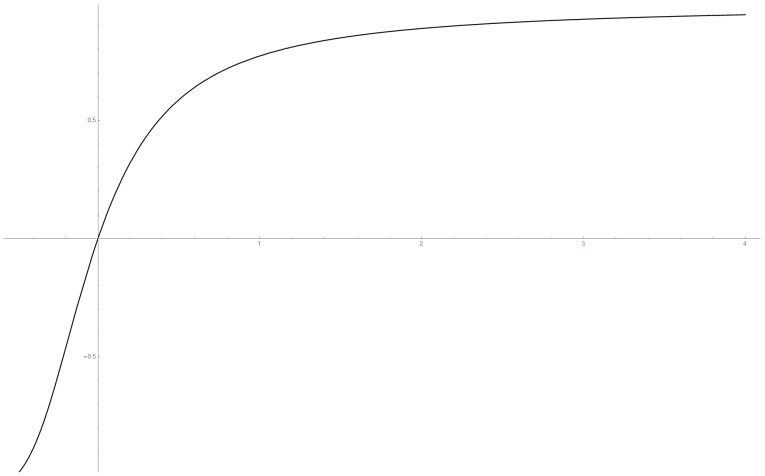

Figure 9. Function γ_1^{-1} with a_1 on the x-axis, ρ on the y-axis; γ_1 is defined in Assumption 6.

Here we apply the above described methods to the dataset 5 of [38]. The yield of grain and straw are the two variables. The sample correlation is 0.73077. Starting from that value, we can calculate the moment estimator for a_1: $\hat{a}_1 = 0.83641$. Moreover, we obtain

$$\begin{aligned} \mathcal{O}_{S_0}(S_0) &= 2.6406, \hat{\mu}_1 = 3.9480, \hat{\mu}_2 = 6.5148 \\ \hat{\sigma}_1 &= 0.45798, \hat{\sigma}_2 = 0.89831. \end{aligned}$$

The data and the shape of the estimated multivariate density are depicted in Figures 10 and 11.

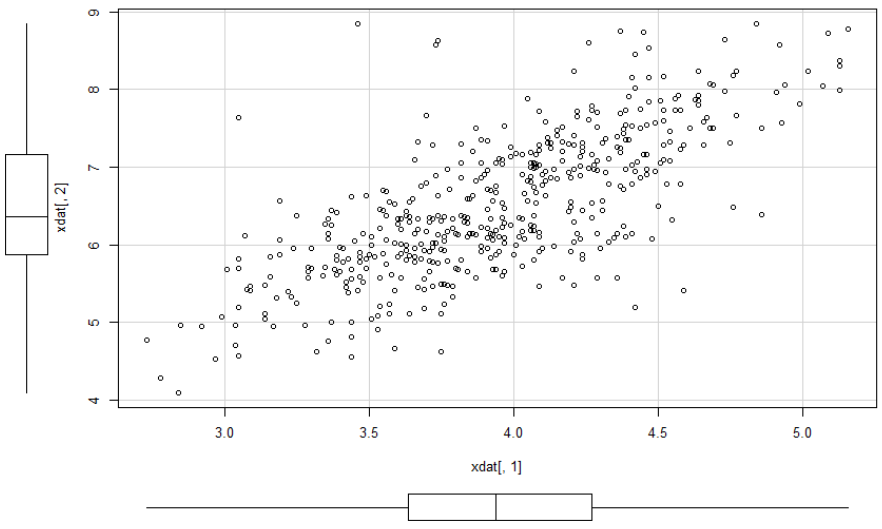

Figure 10. Scatter plot of the dataset 5 of [38].

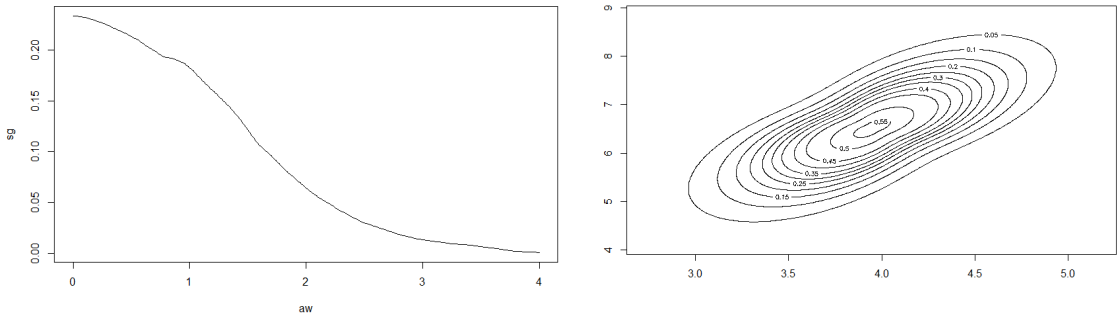

Figure 11. Estimated generator function g at **left** (bandwidth $b = 0.5$) and contour plot of $\hat{\varphi}$ at **right**.

Example 13. *We want to illustrate the potential of our approach for applications to financial data and consider daily index data from Morgan Stanley Capital International of the countries Germany and UK for the period August 2011 to June 2016. The data indicate the continuous daily return values computed as logarithm of the ratio of two subsequent index values. The modelling of MSCI data using elliptical models is considered in [5]. The data are depicted in Figure 12. A visual inspection seems to give some preference for our model from Section 2.2 compared to the elliptically contoured model. Figures 13 and 14 show the estimated model for the data. The basic numerical results are:*

$$\begin{aligned} \mathcal{O}_{S_0}(S_0) &= 2.1966, \hat{a}_1 = 1.2088, \hat{\mu}_1 = 0.00026725, \hat{\mu}_2 = 0.00026855 \\ \hat{\sigma}_1 &= 0.013519, \hat{\sigma}_2 = 0.011715. \end{aligned}$$

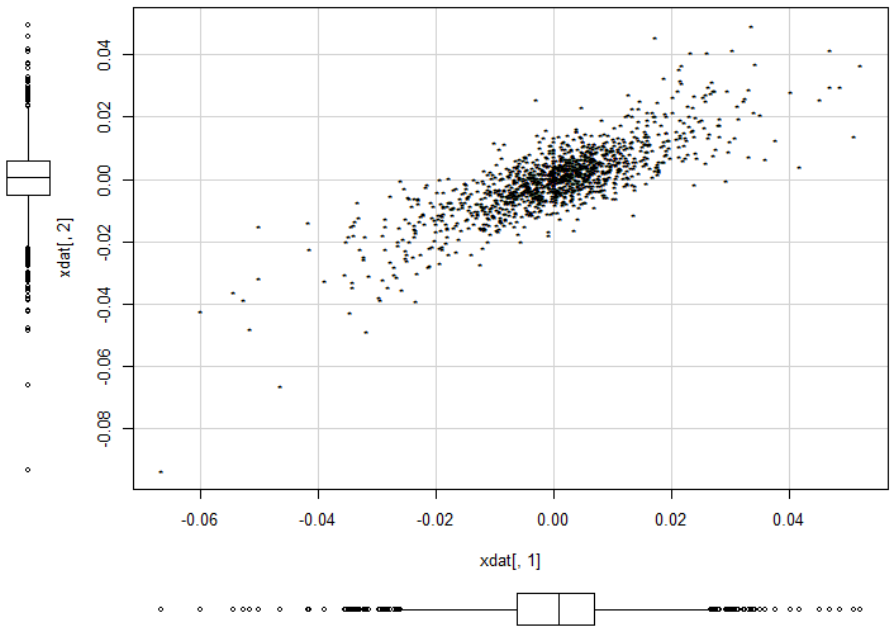

Figure 12. Scatter plot of the MSCI data.

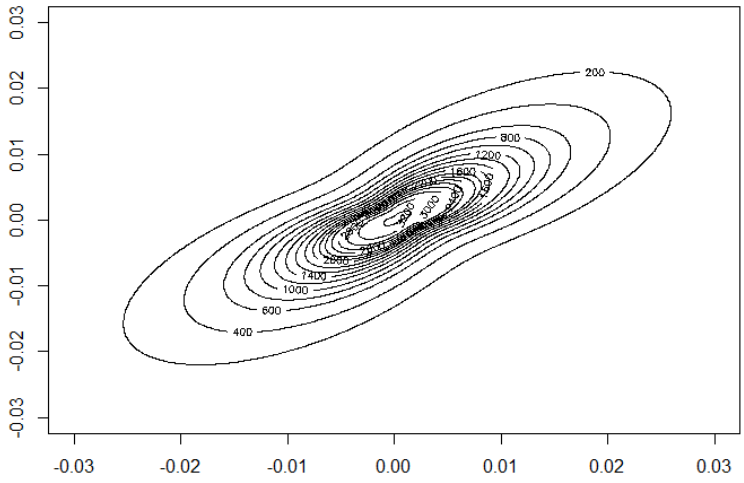

Figure 13. Contour plot of the estimated density.

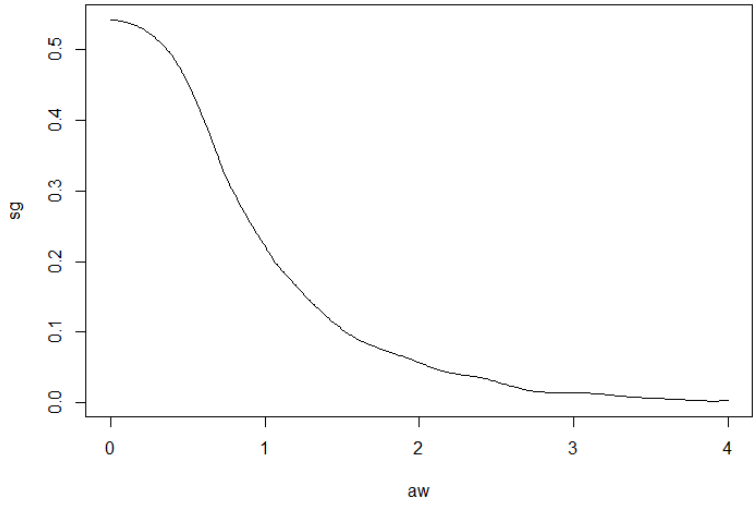

Figure 14. Estimated generator function g (bandwidth $b = 0.25$).

Further we proceed with proving the results.

4. Proofs

Throughout the remainder of the paper, suppose that Assumptions 1–3 are satisfied for some integer $p \geq 2$. First, we prove auxiliary statements which are used in the proof of strong convergence of $\hat{\varphi}_n$ and later.

4.1. Proof of Auxiliary Statements

The following Lemma 2 clarifies the asymptotic behaviour of χ in the neighbourhood of zero.

Lemma 2. *Suppose that g' exists and is bounded. Then*

$$\sup_{t,u\in[0,\bar{M}]} |\chi(t) - \chi(u)|\,|t-u|^{-1/d} < +\infty$$

for any $\bar{M} > 0$.

Proof. Observe that by the Lipschitz continuity of the functions g and $z \rightsquigarrow z^{d-1}\psi'(z)^{-1}$ in view of Assumption 3,

$$\begin{aligned} & |\chi(t) - \chi(u)| \\ \leq\ & C \cdot \left(\frac{\Psi(t)^{d-1}}{\psi'(\Psi(t))} \, |g(\Psi(t)) - g(\Psi(u))| + \left| \frac{\Psi(t)^{d-1}}{\psi'(\Psi(t))} - \frac{\Psi(u)^{d-1}}{\psi'(\Psi(u))} \right| g(\Psi(u)) \right) \\ \leq\ & C \cdot |\Psi(t) - \Psi(u)| \end{aligned} \tag{20}$$

uniformly for $t, v \in [0, \bar{M}]$. Here and in the following C is a generic constant which may differ from formula to formula. By assumption, we have

$$\left| \Psi^d(t) - \Psi^d(u) \right| \leq C \cdot |t - u|.$$

On the other hand,

$$\sup_{t,u \geq 0} \frac{|t^{1/d} - u^{1/d}|}{|t - u|^{1/d}} < +\infty.$$

Hence

$$|\Psi(t) - \Psi(u)| \leq C \cdot \left| \Psi^d(t) - \Psi^d(u) \right|^{1/d} \leq C \cdot |t - u|^{1/d}.$$

In view of (20), the proof is complete. □

Lemma 3. *Let (9) be fulfilled. Then, for $n \to \infty$,*

$$\|\hat{\mu}_n - \mu\|_2 = O\left(\sqrt{\frac{\ln \ln(n)}{n}} \right) \quad a.s.$$

If in addition (18) is satisfied then, for $n \to \infty$,

$$\max_{j=1,\ldots,d} |\hat{\sigma}_{nj} - \sigma_j| = O\left(\sqrt{\frac{\ln \ln(n)}{n}} \right) \quad a.s.$$

Proof. Because of (9), the law of iterated logarithm applies and leads to

$$\limsup_{n \to \infty} \sqrt{\frac{n}{\ln \ln n}} \, |\hat{\mu}_{nj} - \mu_j| = \sqrt{2 \, \mathbb{V} X_j} \quad a.s. \tag{21}$$

Since there is a constant $C > 0$ such that $\|y\|_2 \leq C \|y\|_\infty$ for all $y \in \mathbb{R}^d$ in view of the norm equivalence property, the first part of the lemma follows from (21). The second part can be shown similarly. □

In several places, we will use the following property:

Lemma 4. *Suppose that $\Lambda : \mathbb{R}^d \to \mathbb{R}$ is a measurable function with $\Lambda(-x) = -\Lambda(x)$. Then*

$$\mathbb{E}\left(\Lambda(X - \mu)\right) = 0$$

Proof. Since $-X + \mu$ has the same distribution as $X - \mu$, we have

$$\begin{aligned} \mathbb{E}\left(\Lambda(X - \mu)\right) &= \mathbb{E}\left(\Lambda(X - \mu)\mathbf{1}\{X_1 - \mu_1 < 0\}\right) + \mathbb{E}\left(\Lambda(X - \mu)\mathbf{1}\{X_1 - \mu_1 > 0\}\right) \\ &= \mathbb{E}\left(\Lambda(-X + \mu)\mathbf{1}\{X_1 - \mu_1 > 0\}\right) + \mathbb{E}\left(\Lambda(X - \mu)\mathbf{1}\{X_1 - \mu_1 > 0\}\right) = 0 \end{aligned}$$

which proves the lemma. □

4.2. Proving Convergence of $\hat{F}_n^R$

In this section we prove Theorem 1. The law of the iterated logarithm for the empirical process says (cf. [39], p. 268, for example)

$$\Delta_n := \sup_{r\in\mathbb{R}} \left| F_n^R(r) - F_R(r) \right| = O\left(\sqrt{\frac{\ln\ln(n)}{n}}\right) \ a.s.$$

By Lipschitz-continuity of h_K and Lemma 3,

$$\begin{aligned} \bar{\Delta}_n &: = \sup_{x\in\mathbb{R}^d} |h_K(x-\hat{\mu}_n) - h_K(x-\mu)| \\ &\leq C \left\|\hat{\mu}_n - \mu\right\|_2 = O\left(\sqrt{\frac{\ln\ln(n)}{n}}\right) \ a.s. \end{aligned}$$

Moreover,

$$\begin{aligned} \hat{F}_n^R(r) - F_R(r) &= \frac{1}{n}\sum_{i=1}^n \left(\mathbf{1}\left\{h_K(X_i-\hat{\mu}_n)\leq r\right\} - \mathbf{1}\left\{h_K(X_i-\mu)\leq r\right\}\right) + F_n^R(r) - F_R(r) \\ &\leq \frac{1}{n}\sum_{i=1}^n \left(\mathbf{1}\left\{h_K(X_i-\hat{\mu}_n)\leq r, h_K(X_i-\mu) > r\right\}\right. \\ &\qquad \left. -\mathbf{1}\left\{h_K(X_i-\mu)\leq r, h_K(X_i-\hat{\mu}_n) > r\right\}\right) + F_n^R(r) - F_R(r) \\ &\leq \frac{1}{n}\sum_{i=1}^n \mathbf{1}\left\{r < h_K(X_i-\mu)\leq r+\bar{\Delta}_n\right\} + \Delta_n, \end{aligned}$$

and

$$\hat{F}_n^R(r) - F_R(r) \geq -\frac{1}{n}\sum_{i=1}^n \mathbf{1}\left\{r-\bar{\Delta}_n < h_K(X_i-\mu)\leq r\right\} - \Delta_n.$$

Hence, by the boundedness of f_R,

$$\begin{aligned} &\sup_{r\geq 0}\left|\hat{F}_n^R(r) - F_R(r)\right| \\ \leq\ & \frac{1}{n}\sup_{r\geq 0}\sum_{i=1}^n \mathbf{1}\left\{r-\bar{\Delta}_n < h_K(X_i-\mu)\leq r+\bar{\Delta}_n\right\} + \Delta_n \\ =\ & \sup_{r\geq 0}\left(F_n^R(r+\bar{\Delta}_n) - F_n^R(r-\bar{\Delta}_n)\right) + \Delta_n \\ \leq\ & \sup_{r\geq 0}\left(F_R(r+\bar{\Delta}_n) - F_R(r-\bar{\Delta}_n)\right) + 3\Delta_n \\ \leq\ & 2\sup_{r\geq 0} f_R(r)\,\bar{\Delta}_n + O\left(\sqrt{\frac{\ln\ln n}{n}}\right) = O\left(\sqrt{\frac{\ln\ln n}{n}}\right) \ a.s., \end{aligned}$$

which leads to the theorem. □

4.3. Proving Strong Convergence of the Density Estimator

Let $\tilde{Y}_i = \psi\left(h_K(X_i-\mu)\right)$ for $i = 1,\ldots,n$, $K_b(y,t) = k((y-t)/b) + k((y+t)/b)$ for $y,t \geq 0$, $Y_{in} = \psi(h_K(X_i-\hat{\mu}_n))$, and

$$\tilde{\chi}_n(y) = \frac{1}{nb}\sum_{i=1}^n K_b(y,\tilde{Y}_i)$$

for $y \geq 0$. Then (cf. (10))

$$\hat{\chi}_n(y) = \frac{1}{nb} \sum_{i=1}^{n} K_b(y, Y_{in}).$$

Next we prove strong convergence rates for $\hat{\chi}_n$ and later for $\hat{\varphi}_n$. Throughout this section we suppose that Assumptions 1 to 3 and (9) are fulfilled for some even integer $p \geq 2$. The compact set $[m, M]$ with arbitrary m and M, $0 \leq m < M$ can be covered with closed intervals $U_1, \ldots, U_n$ having sides of length $(M-m)n^{-1}$ and centres $u_1, \ldots, u_n$ such that $\bigcup_{i=1}^{n} U_i = [m, M]$. The constants m and M will be specified later. Note that

$$\begin{aligned} \sup_{y \in [m,M]} |\hat{\chi}_n(y) - \chi(y)| \quad \leq \quad & \max_{l=1,\ldots,n} \left(\sup_{y \in U_l} |\hat{\chi}_n(y) - \hat{\chi}_n(u_l)| + |\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)| \right. \\ & \left. + |\tilde{\chi}_n(u_l) - \chi(u_l)| + \sup_{y \in U_l} |\chi(u_l) - \chi(y)| \right). \end{aligned} \tag{22}$$

The asymptotic behaviour of the right hand side in (22) as $n \to \infty$ is analyzed term by term in the next lemmas.

Lemma 5. *Assume that the p-th order derivative $\chi^{(p)}$ of χ exists for some even integer $p \geq 2$ and is bounded on every finite closed subinterval of $(0, \infty)$. Let g' be bounded. Then*

$$\max_{l=1,\ldots,n} |\tilde{\chi}_n(u_l) - \chi(u_l)| = O\left(\sqrt{\ln n}(nb)^{-1/2} + \beta_n\right) \quad a.s.$$

where $\beta_n = b^p$ if $m > 0$ and $\beta_n = b^{1/d}$ if $m = 0$.

The proof of this lemma is omitted since, with minor changes, this lemma can be proven in the same way as Lemma 4.4 in [7]. The following lemma is used later several times in proofs of almost sure convergence rates. We provide it without proof. The proof is almost identical to that of Lemma 4.6 in [7].

Lemma 6. *Assume that χ is bounded. Let $\bar{k}, \lambda : \mathbb{R} \to \mathbb{R}$ be bounded measurable functions with $\bar{k}(t) = 0$ for $t : |t| > 1$. Then*

$$\begin{aligned} \max_{l=1,\ldots,n} \left| \sum_{i=1}^{n} (U_{nil} - \mathbb{E}U_{nil}) \right| &= O\left(\sqrt{nb\ln(n)}\right) \quad a.s., \\ \max_{l=1,\ldots,n} \left| \sum_{i=1}^{n} (\bar{U}_{nil} - \mathbb{E}\bar{U}_{nil}) \right| &= O\left(\sqrt{nb\ln(n)}\right) \quad a.s., \\ \max_{l=1,\ldots,n} \left| \sum_{i=1}^{n} (V_{nil} - \mathbb{E}V_{nil}) \right| &= O\left(\sqrt{nb\ln(n)}\right) \quad a.s. \end{aligned}$$

where $U_{nil} := \bar{k}((u_l - \tilde{Y}_i)/b)\mathbf{1}\{|\tilde{Y}_i - u_l| \leq b - w_n\}\lambda(X_i)$,
$\bar{U}_{nil} := \bar{k}((u_l + \tilde{Y}_i)/b)\mathbf{1}\{|\tilde{Y}_i - u_l| \leq b - w_n\}\lambda(X_i)$,
$V_{inl} := \mathbf{1}\{b - w_n < |\tilde{Y}_i - u_l| < b + w_n\}$.

We proceed with proving convergence rates of the terms in (22).

Lemma 7. *Suppose that Assumption 4 or 5 is satisfied. Then, as* $n \to \infty$,

$$\max_{l=1,\ldots,n} \left|\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)\right| = \begin{cases} o\left((nb)^{-1/2}\right) \text{ under Assumption 4,} \\ O(b^{-1}\left(n^{-1}\ln\ln n\right)^{\bar{\alpha}/2}) \text{ under Assumption 5} \end{cases} \quad a.s.$$

Proof. (a) Let Assumption 4 be satisfied. In view of Lemma 3, we obtain

$$\begin{aligned} \left|Y_{in} - \tilde{Y}_i\right| &\leq \sup_{t\in[0,\infty)} \left|\psi'(t)\right| \left|h_K(X_i - \hat{\mu}_n) - h_K(X_i - \mu)\right| \\ &\leq C_5 \cdot \sqrt{\frac{\ln\ln n}{n}} =: w_n \end{aligned}$$

with a suitable constant $C_5 > 0$ for $n \geq n_1(\omega)$. We introduce

$$\kappa_n(u,t) = \left(k'((u-t)/b) - k'((u+t)/b)\right) \mathbf{1}\left(|u-t| \leq b - w_n\right).$$

Let $\bar{\psi}(z) := z^{-1}\psi'(z)$. Observe that k' is bounded and Lipschitz continuous on $[-1,1]$, ψ', $\bar{\psi}$ and $\bar{\psi}'$ are bounded on $[0,+\infty)$, functions G_j are bounded, and functions $\psi'(h_K(.))G_j$ are Hölder continuous of order $\alpha > 0.2$. We have then by Taylor expansion

$$\begin{aligned} & k\left(\frac{u - Y_{in}}{b}\right) + k\left(\frac{u + Y_{in}}{b}\right) - k\left(\frac{u - \tilde{Y}_i}{b}\right) - k\left(\frac{u + \tilde{Y}_i}{b}\right) \\ = & -\frac{1}{b}\left(k'((u-\tilde{Y}_i)/b) - k'((u+\tilde{Y}_i)/b)\right)\psi'(h_K(\bar{X}_i)) \\ & \times \sum_{j=1}^{d} G_j(\bar{X}_i)\ \left(\hat{\mu}_{nj} - \mu_j\right) \mathbf{1}\left\{\left|\tilde{Y}_i - u\right| \leq b - w_n\right\} + W_{ni}(u), \end{aligned}$$

where

$$\begin{aligned} |W_{ni}(u)| &\leq C\left(b^{-2}w_n^2 + b^{-1}w_n^{1+\alpha}\right)\mathbf{1}\left\{\left|\tilde{Y}_i - u\right| \leq b - w_n\right\} \\ &\quad + Cb^{-1}w_n \mathbf{1}\left\{b - w_n < \left|\tilde{Y}_i - u\right| < b + w_n\right\}\ a.s. \end{aligned}$$

uniformly w.r.t. $u \in [0,M]$. Here we have used Assumption 4 and Lipschitz continuity of k on $\mathbb{R}$. This leads to

$$\max_{l=1,\ldots,n} \left|\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)\right| \leq C\sum_{j=1}^{d} B_{1nj}\left|\hat{\mu}_{nj} - \mu_j\right| + B_{2n} + B_{3n}, \tag{23}$$

where

$$\begin{aligned} B_{1nj} &= n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\kappa_n(u_l,\tilde{Y}_i)\psi'(h_K(\bar{X}_i))G_j(\bar{X}_i)\right|, \\ B_{2n} &= C\left(b^{-3}w_n^2 + b^{-2}w_n^{1+\alpha}\right)\max_{l=1,\ldots,n}\sum_{i=1}^{n}\mathbf{1}\left\{\left|\tilde{Y}_i - u_l\right| \leq b - w_n\right\}, \\ B_{3n} &= Cn^{-1}b^{-2}w_n\max_{l=1,\ldots,n}\sum_{i=1}^{n}\mathbf{1}\left\{b - w_n < \left|\tilde{Y}_i - u_l\right| < b + w_n\right\}\ a.s. \end{aligned}$$

Note that

$$\begin{aligned} B_{1nj} &= n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\kappa_n(u_l,\tilde{Y}_i)\psi'(h_K(\bar{X}_i))G_j(\bar{X}_i)\right. \\ &\quad \left. - \sum_{i=1}^{n}\mathbb{E}\left(\kappa_n(u_l,\psi(h_K(\bar{X}_i)))\psi'(h_K(\bar{X}_i))G_j(\bar{X}_i)\right)\right| \end{aligned}$$

since the expectation in the last term is zero in view of Lemma 4 ($G_j(-x) = -G_j(x)$ holds for all $x \in \mathbb{R}^d$). Applying Lemma 6, it follows that

$$\sum_{j=1}^{d} B_{1nj} \left|\hat{\mu}_{nj} - \mu_j\right| \leq Cn^{-3/2}b^{-2}\sqrt{\ln\ln n} \cdot \sqrt{nb\ln n} = o\left((nb)^{-1/2}\right). \tag{24}$$

On the other hand, we obtain

$$\begin{aligned} B_{2n} &\leq C\frac{\ln\ln n}{n^2b^3}\left(1 + bn^{(1-\alpha)/2}\right) \max_{l=1,\ldots,n} \sum_{i=1}^{n} I\left(\left|\tilde{Y}_i - u_l\right| \leq b - w_n\right) \leq C\frac{\ln\ln n}{n^2b^3}\left(1 + bn^{(1-\alpha)/2}\right) \\ &\quad \left(\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\left(\mathbf{1}\left\{\left|\tilde{Y}_i - u_l\right| \leq b - w_n\right\} - \mathbb{P}\left\{\left|\tilde{Y}_i - u_l\right| \leq b - w_n\right\}\right)\right|\right. \\ &\quad \left. + n \cdot \sup_{v\in[0,\infty)} \mathbb{P}\left\{v \leq \tilde{Y}_i \leq v + 2b - 2w_n\right\}\right) \\ &\leq C\frac{\ln\ln n}{n^2b^3}\left(1 + n^{(1-\alpha)/2}b\right)\left(\sqrt{nb\ln n} + nb\right) \\ &\leq C\ln\ln n\left(n^{-1}b^{-2} + n^{-(1+\alpha)/2}b^{-1}\right) = o\left((nb)^{-1/2}\right) \quad a.s., \end{aligned} \tag{25}$$

by utilizing Lemma 6 and taking $\alpha > 0.2$ into account. Similarly, it follows that

$$\begin{aligned} B_{3n} &\leq C\sqrt{\frac{\ln\ln n}{n^3b^4}} \max_{l=1,\ldots,n} \sum_{i=1}^{n} \mathbf{1}\left\{b - w_n \leq \left|\tilde{Y}_i - u_l\right| \leq b + w_n\right\} \\ &\leq C\sqrt{\frac{\ln\ln n}{n^3b^4}} \max_{l=1,\ldots,n}\left(\left|\sum_{i=1}^{n}\left(\mathbf{1}\left\{b - w_n \leq \left|\tilde{Y}_i - u_l\right| \leq b + w_n\right\}\right.\right.\right. \\ &\quad \left.\left.\left. - \mathbb{P}\left\{b - w_n \leq \left|\tilde{Y}_i - u_l\right| \leq b + w_n\right\}\right)\right| + n \cdot \mathbb{P}\left\{u_l + b - w_n < \tilde{Y}_i \leq u_l + b + w_n\right\}\right. \\ &\quad \left. + n \cdot \mathbb{P}\left\{u_l - b - w_n < \tilde{Y}_i \leq u_l - b + w_n\right\}\right) \\ &= C\sqrt{\frac{\ln\ln n}{n^3b^4}}\left(\sqrt{nb\ln n} + nw_n\right) \\ &= o\left((nb)^{-1/2}\right) \quad a.s. \end{aligned} \tag{26}$$

Therefore, an application of (23)–(26) leads to the lemma under Assumption 4.
(b) Let Assumption 5 be satisfied. We obtain

$$\begin{aligned} \left|Y_{in} - \tilde{Y}_i\right| &\leq \sup_{t\in[0,\infty)} \left|\psi'(t)\right| \left|h_K(X_i - \hat{\mu}_n) - h_K(X_i - \mu)\right| \\ &\leq C_5 \cdot \left(\frac{\ln\ln n}{n}\right)^{\bar{\alpha}} =: w_n. \end{aligned}$$

Further, by Lipschitz continuity of k,

$$\left|k\left(\frac{u - Y_{in}}{b}\right) + k\left(\frac{u + Y_{in}}{b}\right) - k\left(\frac{u - \tilde{Y}_i}{b}\right) - k\left(\frac{u + \tilde{Y}_i}{b}\right)\right| \leq Cb^{-1}w_n\mathbf{1}\left\{\left|\tilde{Y}_i - u\right| \leq b + w_n\right\}$$

uniformly w.r.t. $u \in [0, M]$. Hence

$$\begin{aligned}
&\max_{l=1,\ldots,n} |\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)| \\
\leq\ & Cn^{-1}b^{-2}w_n \max_{l=1,\ldots,n} \left(\sum_{i=1}^{n} \left(\mathbf{1}\left\{ \left|\tilde{Y}_i - u_l\right| \leq b + w_n \right\} - \mathbb{P}\left\{ \left|\tilde{Y}_i - u_l\right| \leq b + w_n \right\} \right) \right. \\
& \left. + n \cdot \mathbb{P}\left\{ u_l - b - w_n \leq \tilde{Y}_i \leq u_l + b + w_n \right\} \right) \\
=\ & n^{-1}b^{-2}w_n \left(\sqrt{n(b + w_n) \ln n} + n(b + w_n) \right) = O(w_n b^{-1}) \ a.s.
\end{aligned}$$

□

Lemma 8. *Suppose that the assumptions of Lemma 5 are satisfied. Then*

$$\begin{aligned}
&(a) \max_{l=1,\ldots,n} \sup_{y \in U_l} |\hat{\chi}_n(y) - \hat{\chi}_n(u_l)| &&= O(n^{-1}b^{-2}) \ a.s., \\
&(b) \max_{l=1,\ldots,n} \sup_{y \in U_l} |\tilde{\chi}_n(y) - \tilde{\chi}_n(u_l)| &&= O(n^{-1}b^{-2}) \ a.s., \\
&(c) \sup_{y \in [m,M]} |\hat{\chi}_n(y) - \chi(y)| &&= O\left(\sqrt{\ln n}(nb)^{-1/2} + \beta_n \right) \quad a.s.,
\end{aligned}$$

β_n as in Lemma 5.

Proof. In view of Lemma 2, we have

$$\max_{l=1,\ldots,n} \sup_{y \in U_l} |\chi(u_l) - \chi(y)| = \begin{cases} O(n^{-1}) \text{ if } m \neq 0, \\ O(n^{-1/d}) \text{ if } m = 0. \end{cases}$$

Moreover, by the Lipschitz continuity of k, we obtain

$$\begin{aligned}
&\max_{l=1,\ldots,n} \sup_{y \in U_l} |\hat{\chi}_n(y) - \hat{\chi}_n(u_l)| \\
=\ & n^{-1}b^{-1} \max_{l=1,\ldots,n} \sup_{y \in U_l} \left| \sum_{i=1}^{n} \left(K_b(y, Y_{in}) - K_b(u_l, Y_{in}) \right) \right| \\
\leq\ & Cb^{-2} \max_{l=1,\ldots,n} \sup_{y \in U_l} |y - u_l| \\
=\ & O(n^{-1}b^{-2}) \ a.s.
\end{aligned}$$

which proves assertion (a). Analogously, the validity of assertion (b) can be shown. In view of (22), the lemma follows by Lemma 5 and 7. □

We are now in a position to prove the result on strong convergence of $\hat{\varphi}_n$.

Proof of Theorem 2: (i) Case $\mu \notin D$. By Lemma 3, there are $M_0 > m_0 > 0$ such that $h_K(x-\hat{\mu}_n) \in [m_0, M_0]$ for $x \in D$, $n \geq n_2(\omega)$. In view of (11), we obtain

$$\begin{aligned}
\sup_{x\in D}|\hat{\varphi}_n(x)-\varphi(x)| &\leq \mathcal{O}_S(S)^{-1}\left(\sup_{x\in D}|\hat{g}_n\left(h_K(x-\hat{\mu}_n)\right)-g\left(h_K(x-\hat{\mu}_n)\right)| \right.\\
&\quad \left. +\sup_{z\geq 0}|g'(z)|\sup_{x\in D}|h_K(x-\hat{\mu}_n)-h_K(x-\mu)|\right)\\
&\leq \mathcal{O}_S(S)^{-1}\sup_{z\in[m_0,M_0]}|\hat{g}_n(z)-g(z)|+C\sqrt{\frac{\ln\ln n}{n}}\\
&\leq \mathcal{O}_S(S)^{-1}\sup_{z\geq 0}\left(z^{1-d}\,\psi'(z)\right)\sup_{z\in[\psi(m_0),\psi(M_0)]}|\hat{\chi}_n(z)-\chi(z)|+C\sqrt{\frac{\ln\ln n}{n}}\quad a.s.
\end{aligned}$$

for $n \geq n_3(\omega)$. Lemma 8 applies to complete the proof of part (i).

(ii) Case $\mu \in D$. The proof can be done analogously to part (i) taking $m_0 = 0$ into account. □

4.4. Proving Asymptotic Normality of $\hat{\varphi}_n(x)$

Throughout this subsection, assume that Assumptions 1–3 and (9) are fulfilled for some integer $p \geq 2$. First, an auxiliary result is proven. Define $\hat{x} := h_K(x-\hat{\mu}_n)$ and $\tilde{x} := h_K(x-\mu)$.

Lemma 9. *Under Assumption 4, we have*

$$\begin{aligned}
&(a)\ |\hat{\chi}_n(\psi(\tilde{x}))-\tilde{\chi}_n(\psi(\tilde{x}))| &&= o\left((nb)^{-1/2}\right) \quad a.s.\\
&(b)\ |\hat{\chi}_n(\psi(\hat{x}))-\hat{\chi}_n(\psi(\tilde{x}))| &&= o\left((nb)^{-1/2}\right) \quad a.s.\ as\ n\to\infty.
\end{aligned}$$

Proof. Let U_l and u_l as in Section 4.3. We can choose $m, M : 0 < m < M$ such that $\psi(\hat{x})$, $\psi(\tilde{x}) \in [m, M]$ for $n \geq n_4(\omega)$. By Lemmas 7 and 8,

$$\begin{aligned}
\sup_{y\in[m,M]}|\hat{\chi}_n(y)-\tilde{\chi}_n(y)| &\leq \max_{l=1,\ldots,n}\sup_{y\in U_l}|\hat{\chi}_n(y)-\hat{\chi}_n(u_l)|+\max_{l=1,\ldots,n}|\hat{\chi}_n(u_l)-\tilde{\chi}_n(u_l)|\\
&\quad +\max_{l=1,\ldots,n}\sup_{y\in U_l}|\tilde{\chi}_n(y)-\tilde{\chi}_n(u_l)|\\
&\leq Cn^{-1}b^{-2}+o\left((nb)^{-1/2}\right)=o\left((nb)^{-1/2}\right)\quad a.s. \qquad (27)
\end{aligned}$$

which yields immediately assertion (a). Since

$$|\hat{x}-\tilde{x}| = O(w_n)\quad a.s.$$

by Lemma 3, we obtain the inequality

$$|\tilde{\chi}_n\left(\psi(\hat{x})\right)-\tilde{\chi}_n\left(\psi(\tilde{x})\right)| \leq D_{1n}+D_{2n}+D_{3n}$$

by Taylor expansion, where

$$\begin{aligned} D_{1n} &= \left| n^{-1}b^{-2}\sum_{i=1}^{n}\sum_{j=1}^{d} k'((\psi(\tilde{x}))-\tilde{Y}_i)b^{-1})\mathbf{1}\left\{\left|\psi(\tilde{x})-\tilde{Y}_i\right|\le b-w_n\right\}\ \psi'(\tilde{x})G_j(x-\mu)\ \ (\hat{\mu}_{nj}-\mu_j)\right|, \\ D_{2n} &\le Cn^{-1}b^{-3}\cdot w_n^2\sum_{i=1}^{n}\mathbf{1}\left\{\left|\psi(\tilde{x})-\tilde{Y}_i\right|\le b-w_n\right\}, \\ D_{3n} &\le Cn^{-1}b^{-2}\cdot w_n\sum_{i=1}^{n}\mathbf{1}\left\{b-w_n<\left|\psi(\tilde{x})-\tilde{Y}_i\right|<b+w_n\right\} \end{aligned}$$

a.s. Observe that $k'(-t)=-k'(t)$ and

$$\begin{aligned} &\mathbb{E}k'((\psi(\tilde{x})-\tilde{Y}_i)b^{-1})\mathbf{1}\left\{\left|\psi(\tilde{x})-\tilde{Y}_i\right|\le b-w_n\right\} = b\int_{-1+w_nb^{-1}}^{1-w_nb^{-1}} k'(t)\chi(\psi(\tilde{x})-tb)\,\mathrm{d}t \\ =\ & b\int_{0}^{1-w_nb^{-1}} k'(t)\left(\chi(\psi(\tilde{x})-tb)-\chi(\psi(\tilde{x})+tb)\right)\,\mathrm{d}t \\ =\ & O(b^2). \end{aligned}$$

Analogously to Lemma 6, we can deduce

$$\begin{aligned} D_{1n} &\le Cn^{-1}b^{-2}\cdot w_n\cdot\sum_{j=1}^{d}\left|\sum_{i=1}^{n}k'((\psi(\tilde{x})-\tilde{Y}_i)b^{-1})\ \mathbf{1}\left\{\left|\psi(\tilde{x})-\tilde{Y}_i\right|\le b-w_n\right\}\psi'(\tilde{x})G_j(x-\mu)\right| \\ &\le Cn^{-3/2}b^{-2}\sqrt{\ln\ln n}\cdot\left(\sqrt{nb\ln n}+nb^2\right) \\ &= o\left((nb)^{-1/2}\right)\ \ a.s. \end{aligned}$$

since the expectation is zero due to Lemma 4. Analogously to the examination of B_{3n} and B_{4n} in Lemma 7, we obtain

$$D_{2n}+D_{3n}=o\left((nb)^{-1/2}\right)\ \ a.s.$$

Hence

$$\left|\tilde{\chi}_n\left(\psi(\hat{x})\right)-\tilde{\chi}_n\left(\psi(\tilde{x})\right)\right|=o\left((nb)^{-1/2}\right)\ \ a.s.$$

In view of (27), the proof of part b) is complete. □

From kernel density estimation theory, we can take the following lemma, see [40]. Subsequently, we prove asymptotic normality of $\hat{\varphi}_n$.

Lemma 10. *Suppose that χ is continuous at y. Then*

$$\sqrt{nb}\left(\tilde{\chi}_n(y)-\mathbb{E}\tilde{\chi}_n(y)\right)\xrightarrow{d}\mathcal{N}(0,\sigma_1^2),\ \sigma_1^2=\chi(y)\int_{-1}^{1}k^2(t)\,dt.$$

Proof of Theorem 3. Note that $z\rightsquigarrow z^{1-d}\psi'(z)$ has a bounded derivative on every finite subinterval of $(0,\infty)$. By Lemmas 3 and 9, we obtain

$$h_K(x-\hat{\mu}_n)^{1-d}\psi'(h_K(x-\hat{\mu}_n))-\tilde{x}^{1-d}\psi'(\tilde{x})=O\left(\sqrt{\frac{\ln\ln n}{n}}\right)$$

and hence

$$\begin{aligned}\hat{\varphi}_n(x)-\varphi(x) &= \mathcal{O}_S(S)^{-1}h_K(x-\hat{\mu}_n)^{1-d}\,\psi'(h_K(x-\hat{\mu}_n))\,(\hat{\chi}_n(\psi(h_K(x-\hat{\mu}_n)))-\hat{\chi}_n(\psi(\tilde{x})))\\ &\quad+\mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\;(\hat{\chi}_n(\psi(\tilde{x}))-\chi(\psi(\tilde{x})))\\ &\quad+\left(h_K(x-\hat{\mu}_n)^{1-d}\psi'(h_K(x-\hat{\mu}_n))-\tilde{x}^{1-d}\psi'(\tilde{x})\right)\mathcal{O}_S(S)^{-1}\,\hat{\chi}_n(\psi(\tilde{x}))\\ &= o\left((nb)^{-1/2}\right)+\left(\mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})+o((nb)^{-1/2})\right)\;(\hat{\chi}_n(\psi(\tilde{x}))-\chi(\psi(\tilde{x})))\\ &= Z_n+e_n+o\left((nb)^{-1/2}\right)\;\;a.s.,\end{aligned}$$

where

$$\begin{aligned}Z_n &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\,(\hat{\chi}_n(\psi(\tilde{x}))-\mathbb{E}\hat{\chi}_n(\psi(\tilde{x}))),\\ e_n &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\,(\mathbb{E}\hat{\chi}_n(\psi(\tilde{x}))-\chi(\psi(\tilde{x}))).\end{aligned}$$

Using Lemma 10, we have

$$\sqrt{nb}Z_n \xrightarrow{d} \mathcal{N}(0,\bar{\sigma}^2)$$

where $\bar{\sigma}^2(\tilde{x})=\mathcal{O}_S(S)^{-2}\tilde{x}^{2-2d}\;\psi'(\tilde{x})^2\sigma_1^2$. By Taylor expansion, we obtain

$$\begin{aligned}e_n &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\;\psi'(\tilde{x})\left(\int_0^\infty k\left(\frac{\psi(\tilde{x})-y}{b}\right)\chi(y)\,\mathrm{d}y-\chi(\psi(\tilde{x}))\right)\\ &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\;\psi'(\tilde{x})\int_{-1}^1 k(t)\;(\chi(\psi(\tilde{x})-tb)-\chi(\psi(\tilde{x})))\,\mathrm{d}t\\ &= \mathcal{O}_S(S)^{-1}\tilde{x}^{1-d}\;\psi'(\tilde{x})\cdot\int_{-1}^1 k(t)\;\left(\sum_{j=1}^{p-1}(j!)^{-1}\chi^{(j)}(\psi(\tilde{x}))\,(-tb)^j+(p!)^{-1}\chi^{(p)}(\tilde{x}_n)\;t^pb^p\right)\mathrm{d}t\\ &= \mathcal{O}_S(S)^{-1}(p!)^{-1}\tilde{x}^{1-d}\psi'(\tilde{x})\;\chi^{(p)}(\psi(\tilde{x}))\int_{-1}^1 t^pk(t)\,\mathrm{d}t\;b^p+o(b^p),\end{aligned}$$

and $\tilde{x}_n$ lies between $\psi(\tilde{x})-tb$ and $\psi(\tilde{x})$. This completes the proof. □

4.5. Proofs When Additional Scale Fit Is Involved

When proving Theorem 4 we shall make use of the following lemma.

Lemma 11. *Let (18) be satisfied. Then*

$$\hat{\theta}_{nl}-\theta_l=O\left(\sqrt{\frac{\ln\ln n}{n}}\right)\;\;a.s.\;for\;l=1,\ldots,q.$$

Proof. By the law of iterated logarithm and Lemma 3, we obtain

$$\begin{aligned}\hat{\rho}_{njk}-\rho_{jk} &= \hat{\sigma}_{nj}^{-1}\hat{\sigma}_{nk}^{-1}\left(\frac{1}{n-1}\sum_{i=1}^n X_{ij}X_{ik}-\mathbb{E}X_jX_k\right.\\ &\quad\left.-\frac{n}{n-1}\left((\hat{\mu}_{nj}-\mu_j)\,\hat{\mu}_{nk}+\mu_j\,(\hat{\mu}_{nk}-\mu_k)\right)+O(n^{-1})\right)\\ &\quad+\mathbb{E}X_jX_k\;\left(\hat{\sigma}_{nj}\hat{\sigma}_{nk}\sigma_j\sigma_k\right)^{-1}\left(\sigma_j\,(\sigma_k-\hat{\sigma}_{nk})+\hat{\sigma}_{nk}\,(\sigma_j-\hat{\sigma}_{nj})\right)\\ &= O\left(\sqrt{\frac{\ln\ln n}{n}}\right)\;\;a.s.\end{aligned}$$

for all $j,k=1,\ldots,d$. Since the partial derivatives of functions γ_l are bounded, it follows that

$$\begin{aligned}\left|\hat{\theta}_{nl}-\theta_l\right| &\leq C\max_{j,k=1,\ldots,d}\left|\hat{\rho}_{njk}-\rho_{jk}\right| \\ &= O\left(\sqrt{\frac{\ln\ln n}{n}}\right) \ a.s.\end{aligned}$$

□

Proof of Theorem 4. Let $u(x)=h_{K_0}(\theta,\Sigma^{-1}(x-\mu))$. Note that $u(X_i)=R_i$. By Lipschitz-continuity of h_{K_0}, (1) and Lemma 3, we have

$$\begin{aligned}\bar{\Delta}_n &:= \sup_{x\in\mathbb{R}^d}\left|h_{K_0}(\hat{\theta}_n,\hat{\Sigma}_n^{-1}(x-\hat{\mu}_n))-u(x)\right|(1+u(x))^{-1} \\ &\leq C\sup_{x\in\mathbb{R}^d}\left(\left\|\hat{\theta}_n-\theta\right\|+\hat{\Sigma}_n^{-1}\left\|\hat{\mu}_n-\mu\right\|+\sum_{j=1}^{d}\left|\hat{\sigma}_{nj}^{-1}-\sigma_j^{-1}\right|\left|x_j-\mu_j\right|\right)\cdot(1+u(x))^{-1} \\ &\leq C\sqrt{\frac{\ln\ln n}{n}}\cdot\sup_{x\in\mathbb{R}^d}(1+\|x-\mu\|)\cdot(1+u(x))^{-1} \\ &= O\left(\sqrt{\frac{\ln\ln n}{n}}\right) \ a.s.\end{aligned}$$

We obtain the result as follows:

$$\begin{aligned}&\sup_{r\geq 0}\left|\hat{F}_n^R(r)-F_R(r)\right| \\ \leq\ & \frac{1}{n}\sup_{r\geq 0}\sum_{i=1}^{n}\mathbf{1}\left\{r-\bar{\Delta}_n(1+u(X_i))\leq u(X_i)\leq r+\bar{\Delta}_n(1+u(X_i))\right\}+\Delta_n \\ =\ & \frac{1}{n}\sup_{r\geq 0}\sum_{i=1}^{n}\mathbf{1}\left\{\frac{r-\bar{\Delta}_n}{1+\bar{\Delta}_n}\leq R_i\leq\frac{r+\bar{\Delta}_n}{1-\bar{\Delta}_n}\right\}+\Delta_n \\ =\ & \sup_{r\geq 0}\left(F_n^R((r+\bar{\Delta}_n)/(1-\bar{\Delta}_n))-F_n^R((r-\bar{\Delta}_n)/(1+\bar{\Delta}_n))\right)+\Delta_n \\ \leq\ & \sup_{r\geq 0}\left(F_R((r+\bar{\Delta}_n)/(1-\bar{\Delta}_n))-F_R((r-\bar{\Delta}_n)/(1+\bar{\Delta}_n))\right)+3\Delta_n \\ \leq\ & 2\bar{\Delta}_n\sup_{r\geq 0}rf_R(r)\ (1-\bar{\Delta}_n^2)^{-1}+C\sqrt{\frac{\ln\ln n}{n}}=O\left(\sqrt{\frac{\ln\ln n}{n}}\right) \ a.s.\end{aligned}$$

□

We proceed with proving Theorem 5. The next two lemmas are used in the proof of this theorem. Here we define $Y_{in}=\psi(h_{K_0}(\hat{\theta}_n,\hat{\Sigma}_n^{-1}(X_i-\hat{\mu}_n)))$ and $\tilde{Y}_i=\psi(h_{K_0}(\theta,\Sigma^{-1}(X_i-\mu)))$. Notice that $\tilde{Y}_i$ has density χ, and Lemmas 5 and 6 hold true for the modified Y_{in} and $\tilde{Y}_i$, too.

Lemma 12. *Let g' be bounded. Then we have*

$$\sup_{u\in[0,M]}\left|\mathbb{E}\kappa_n(u,\psi(R_i))\psi'(R_i)R_i\right|=O(b^2).$$

Proof. Note that

$$\int_{-1+w_n/b}^{1-w_n/b}k'(t)\,dt=k(1-w_n/b)-k(-1+w_n/b)=0.$$

For $u > b$, we obtain

$$\begin{aligned} & \mathbb{E}k'((u-\psi(R_i))/b)\psi'(R_i)R_i\,\mathbf{1}\{|u-\psi(R_i)| \le b - w_n\} \\ = & \int_0^\infty k'\left(\frac{u-\psi(t)}{b}\right)\psi'(t)t^d g(t)\,\mathbf{1}\left(|u-\psi(t)| \le b-w_n\right)\mathrm{d}t \\ = & -b\int_{-1+w_n/b}^{1-w_n/b} k'(t)\left(\Psi^d(u-tb)g(\Psi(u-tb)) - \Psi^d(u)g(\Psi(u))\right)\mathrm{d}t. \end{aligned}$$

Since g, g', Ψ^d and $(\Psi^d)'$ are bounded, we can conclude

$$\sup_{u\in(b,M]} \left|\mathbb{E}\kappa_n(u,\psi(R_i))\psi'(R_i)R_i\right| = O(b^2).$$

Further we obtain

$$\begin{aligned} & \sup_{u\in[0,b]} \Bigg| \mathbb{E}\left(k'\left(\frac{u-\psi(R_i)}{b}\right) - k'\left(\frac{u+\psi(R_i)}{b}\right)\right)\psi'(R_i)R_i \\ & \qquad \mathbf{1}\{|u-\psi(R_i)| \le b-w_n\}\Bigg| \\ \le & C\sup_{u\in[-1,1]}|k'(u)|\sup_{u\ge 0}|g(u)|\int_0^{\Psi(2b)} t^d\psi'(t)\,\mathrm{d}t \\ \le & C\int_0^{2b}\Psi^d(t)\,\mathrm{d}t \le C\int_0^{2b} t\,\mathrm{d}t = O(b^2), \end{aligned}$$

by (14). This inequality completes the proof. □

Lemma 13. *Suppose that Assumption 7 or Assumption 8 is satisfied. Then*

$$\max_{l=1,\dots,n}|\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)| = \begin{cases} o\left((nb)^{-1/2}\right) \text{ under Assumption 7}, \\ O(b^{-1}\left(n^{-1}\ln\ln n\right)^{\bar{\alpha}/2}) \text{ under Assumption 8} \end{cases} \quad a.s.$$

Proof. We prove the Lemma under Assumption 7, the proof of the other part is analogous to that of Lemma 7(b). As we see later, we need only to involve data vectors X_i with $Y_{in} = \psi(h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(X_i - \hat{\mu}_n))) \le M + \bar{b}$ or $\tilde{Y}_i = \psi(h_{K_0}(\theta, \Sigma^{-1}(X_i - \mu))) \le M + \bar{b}$. This implies $\min\{\|X_i - \hat{\mu}_n\|, \|X_i - \mu\|\} \le C\Psi(M+\bar{b})$ by (1), and therefore $\|X_i - \mu\| \le C_5$ for $n \ge n_6(\omega)$ with a constant $C_5 > 0$. In view of Lemma 3 and by Assumption 7, we obtain

$$\begin{aligned} \left|Y_{in} - \tilde{Y}_i\right| & \le \sup_{t\in[0,\infty)}\left|\psi'(t)\right|\left|h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(X_i - \hat{\mu}_n)) - h_{K_0}(\theta, \Sigma^{-1}(X_i - \mu))\right| \\ & \le C\left(\left\|\hat{\theta}_n - \theta\right\| + \sum_{j=1}^d\left(\hat{\sigma}_{nj}^{-1}\left|\mu_j - \hat{\mu}_{nj}\right| + \left|X_{ij} - \mu_j\right|\left|\hat{\sigma}_{nj}^{-1} - \sigma_j^{-1}\right|\right)\right) \\ & \le C_6 \cdot \sqrt{\frac{\ln\ln n}{n}} =: w_n \end{aligned}$$

($X_i = (X_{i1}, \dots, X_{id})^T$) with a suitable constant $C_6 > 0$ for $n \ge n_7(\omega)$. We introduce

$$\kappa_n(u,t) = \left(k'((u-t)/b) - k'((u+t)/b)\right)\mathbf{1}\left(|u-t| \le b - w_n\right).$$

Let $\bar{\psi}(z) := z^{-1}\psi'(z), \bar{X}_i := \Sigma^{-1}(X_i - \mu) = (\bar{X}_{i1}, \dots, \bar{X}_{id})^T$. Observe that k' is bounded and Lipschitz continuous on $[-1,1]$, $\psi', \bar{\psi}$ and $\bar{\psi}'$ are bounded on $[0,+\infty)$, functions $G_j, \tilde{G}_j$ are bounded,

and functions $\psi'(h_K(.))G_j$ are Hölder continuous of order $\alpha > 0.2$. By Taylor expansion, Assumption 7 and Lipschitz continuity of k on $\mathbb{R}$, we have

$$\begin{aligned} & k\left(\frac{u-Y_{in}}{b}\right)+k\left(\frac{u+Y_{in}}{b}\right)-k\left(\frac{u-\tilde{Y}_i}{b}\right)-k\left(\frac{u+\tilde{Y}_i}{b}\right) \\ = & -\frac{1}{b}\left(k'((u-\tilde{Y}_i)/b)-k'((u+\tilde{Y}_i)/b)\right)\psi'(h_{K_0}(\theta,\bar{X}_i))\mathbf{1}\left\{\left|\tilde{Y}_i-u\right|\leq b-w_n\right\} \\ & \cdot\left(\sum_{j=1}^{d}G_j(\theta,\bar{X}_i)\ \left(\hat{\sigma}_{nj}^{-1}\left(\mu_j-\hat{\mu}_{nj}\right)+\left(X_{ij}-\mu_j\right)\left(\hat{\sigma}_{nj}^{-1}-\sigma_j^{-1}\right)\right)\right. \\ & \left.+\sum_{j=1}^{q}\tilde{G}_j(\theta,\bar{X}_i)\ \left(\hat{\theta}_{nj}-\theta_j\right)\right)+W_{ni}(u), \end{aligned}$$

where, uniformly for $u \in [0, M]$,

$$\begin{aligned} |W_{ni}(u)| \ \leq \ & Cb^{-2}w_n^2\mathbf{1}\left\{\left|\tilde{Y}_i-u\right|\leq b-w_n\right\} \\ & +Cb^{-1}w_n\mathbf{1}\left\{b-w_n<\left|\tilde{Y}_i-u\right|<b+w_n\right\}\ a.s. \end{aligned}$$

This leads to

$$\begin{aligned} & \max_{l=1,\ldots,n}\left|\hat{\chi}_n(u_l)-\tilde{\chi}_n(u_l)\right| \\ \leq \ & C\left(\sum_{j=1}^{d}\left(B_{1nj}\hat{\sigma}_{nj}^{-1}\left|\hat{\mu}_{nj}-\mu_j\right|+B_{4nj}\left|\hat{\sigma}_{nj}^{-1}-\sigma_j\right|\right)+\sum_{j=1}^{q}B_{5nj}\left|\hat{\theta}_{nj}-\theta_j\right|\right) \\ & +B_{2n}+B_{3n}, \end{aligned}$$

where B_{2n}, B_{3n} are as above, and

$$\begin{aligned} B_{1nj} &= n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\kappa_n(u_l,\tilde{Y}_i)\psi'(h_{K_0}(\theta,\bar{X}_i))G_j(\theta,\bar{X}_i)\right|, \\ B_{4nj} &= n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\kappa_n(u_l,\tilde{Y}_i)\psi'(h_{K_0}(\theta,\bar{X}_i))G_j(\theta,\bar{X}_i)\bar{X}_i\right|, \\ B_{5nj} &= n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\kappa_n(u_l,\tilde{Y}_i)\psi'(h_{K_0}(\theta,\bar{X}_i))\tilde{G}_j(\theta,\bar{X}_i)\right|. \end{aligned}$$

Analogously to Lemma 7, we obtain

$$\sum_{j=1}^{d}B_{1nj}\left|\hat{\mu}_{nj}-\mu_j\right|\leq Cn^{-3/2}b^{-2}\sqrt{\ln\ln n}\cdot\sqrt{nb\ln n}=o\left((nb)^{-1/2}\right).$$

Observe that $\bar{X}_i = R_i\bar{U}_i$, $G_j(\theta, R_i\bar{U}_i) = G_j(\theta,\bar{U}_i)$ and $\tilde{G}_j(\theta, R_i\bar{U}_i) = R_i\tilde{G}_j(\theta,\bar{U}_i)$ *a.s.* ($h_{K_0}(\theta,.)$ is a homogeneous function). Applying Lemmas 6 and 12, we can derive

$$\begin{aligned} B_{4nj} \ \leq \ & n^{-1}b^{-2}\max_{l=1,\ldots,n}\left|\sum_{i=1}^{n}\left(\kappa_n(u_l,\tilde{Y}_i)\psi'(h_{K_0}(\theta,\bar{X}_i))G_j(\theta,\bar{X}_i)\bar{X}_{ij}\right.\right. \\ & \left.\left.-\mathbb{E}\left(\kappa_n(u_l,\psi(h_{K_0}(\theta,\bar{X}_i)))\psi'(h_{K_0}(\theta,\bar{X}_i))G_j(\theta,\bar{X}_i)\right)\bar{X}_{ij}\right)\right| \\ & +b^{-2}\max_{l=1,\ldots,n}\left|\mathbb{E}\left(\kappa_n(u_l,\psi(R_i))\psi'(R_i)R_i\right)\mathbb{E}\bar{U}_{ij}G_j(\theta,\bar{U}_i)\right| \\ = \ & O\left(\sqrt{n^{-1}b^{-3}\ln n}+1\right) \end{aligned}$$

($\bar{U}_i = (\bar{U}_{i1}, \ldots, \bar{U}_{id})^T$) which implies

$$\sum_{j=1}^{d} B_{4nj} \left| \hat{\sigma}_{nj}^{-1} - \sigma_j \right| = o\left((nb)^{-1/2} \right).$$

Further, by Lemma 12,

$$\begin{aligned} B_{5nj} &\leq n^{-1} b^{-2} \max_{l=1,\ldots,n} \Bigg| \sum_{i=1}^{n} \left(\kappa_n(u_l, \tilde{Y}_i) \psi'(h_{K_0}(\theta, \bar{X}_i)) \tilde{G}_j(\theta, \bar{X}_i) \right. \\ &\quad \left. - \mathbb{E}\left(\kappa_n(u_l, \psi(h_{K_0}(\theta, \bar{X}_i))) \psi'(h_{K_0}(\theta, \bar{X}_i)) \tilde{G}_j(\theta, \bar{X}_i) \right) \right) \Bigg| \\ &\quad + b^{-2} \max_{l=1,\ldots,n} \left| \mathbb{E}\left(\kappa_n(u_l, \psi(R_i)) \psi'(R_i) R_i \right) \mathbb{E} \tilde{G}_j(\theta, \bar{U}_i) \right) \Big| \\ &= O\left(\sqrt{n^{-1} b^{-3} \ln n} + 1 \right) \end{aligned}$$

and

$$\sum_{j=1}^{q} B_{5nj} \left| \hat{\theta}_{nj} - \theta_j \right| = o\left((nb)^{-1/2} \right).$$

This completes the proof. □

Proof of Theorem 5. We consider only the case $\mu \notin D$, the proof in the other case is similar. By Lemma 3, there are $M_0 > m_0 > 0$ such that $h_{K_0}(\hat{\theta}_n, \hat{\Sigma}^{-1}(x - \hat{\mu}_n)) \in [m_0, M_0]$ for $x \in D$, $n \geq n_8(\omega)$. In view of (11), we obtain

$$\begin{aligned} & \sup_{x \in D} |\hat{\varphi}_n(x) - \varphi(x)| \\ \leq\ & \mathcal{O}_S(S)^{-1} \det(\hat{\Sigma}_n)^{-1} \\ & \Bigg(\sup_{x \in D} \left| \hat{g}_n\left(h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(x - \hat{\mu}_n)) \right) - g\left(h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(x - \hat{\mu}_n)) \right) \right| \\ & \quad + \sup_{z \geq 0} |g'(z)| \sup_{x \in D} \left| h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(x - \hat{\mu}_n)) - h_{K_0}(\theta, \Sigma^{-1}(x - \mu)) \right| \Bigg) \\ \leq\ & C \left(\sup_{z \in [m_0, M_0]} |\hat{g}_n(z) - g(z)| + \sqrt{\frac{\ln \ln n}{n}} \right) \\ \leq\ & C \left(\sup_{z \geq 0} \left(z^{1-d}\, \psi'(z) \right) \sup_{z \in [\psi(m_0), \psi(M_0)]} |\hat{\chi}_n(z) - \chi(z)| + \sqrt{\frac{\ln \ln n}{n}} \right) \ a.s. \end{aligned}$$

for $n \geq n_9(\omega)$. An application of Lemma 8 completes the proof. □

In last part of this section, we prove the result on asymptotic normality. Define $\hat{x} := h_{K_0}(\hat{\theta}_n, \hat{\Sigma}_n^{-1}(x - \hat{\mu}_n))$ and $\tilde{x} := h_{K_0}(\theta, \Sigma^{-1}(x - \mu))$.

Lemma 14. *Under Assumption 7, we have*

$$\begin{aligned} a)\ |\hat{\chi}_n(\psi(\tilde{x})) - \tilde{\chi}_n(\psi(\tilde{x}))| &= o\left((nb)^{-1/2} \right) \quad a.s. \\ b)\ |\hat{\chi}_n(\psi(\hat{x})) - \hat{\chi}_n(\psi(\tilde{x}))| &= o\left((nb)^{-1/2} \right) \quad a.s. \end{aligned}$$

Proof. Let U_l and u_l as in Section 4.3. We can choose $m, M : 0 < m < M$ such that $\psi(\hat{x}), \psi(\tilde{x}) \in [m, M]$ for $n \geq n_{10}(\omega)$. By Lemmas 7 and 8,

$$\begin{aligned}\sup_{y\in[m,M]} |\hat{\chi}_n(y) - \tilde{\chi}_n(y)| &\leq \max_{l=1,\ldots,n} \sup_{y\in U_l} |\hat{\chi}_n(y) - \hat{\chi}_n(u_l)| + \max_{l=1,\ldots,n} |\hat{\chi}_n(u_l) - \tilde{\chi}_n(u_l)| \\ &\quad + \max_{l=1,\ldots,n} \sup_{y\in U_l} |\tilde{\chi}_n(y) - \tilde{\chi}_n(u_l)| \\ &\leq Cn^{-1}b^{-2} + o\left((nb)^{-1/2}\right) = o\left((nb)^{-1/2}\right) \quad a.s. \end{aligned} \tag{28}$$

which yields immediately assertion a). Since

$$|\hat{x} - \tilde{x}| = O(w_n) \quad a.s.$$

by Lemma 3, we obtain the inequality

$$|\tilde{\chi}_n(\psi(\hat{x})) - \tilde{\chi}_n(\psi(\tilde{x}))| \leq D_{1n} + D_{2n} + D_{3n}$$

by Taylor expansion, where

$$\begin{aligned} D_{1n} &= \Bigg| n^{-1}b^{-2} \sum_{i=1}^{n} k'((\psi(\tilde{x})) - \tilde{Y}_i)b^{-1})\psi'(\tilde{x})\mathbf{1}\left\{\left|\psi(\tilde{x}) - \tilde{Y}_i\right| \leq b - w_n\right\} \\ &\quad \left(\sum_{j=1}^{d} G_j(\theta, \tilde{x}) \left(\hat{\sigma}_{nj}^{-1}\left(\mu_j - \hat{\mu}_{nj}\right) + \left(X_{ij} - \mu_j\right)\left(\hat{\sigma}_{nj}^{-1} - \sigma_j^{-1}\right)\right)\right. \\ &\quad \left. + \sum_{j=1}^{q} \tilde{G}_j(\theta, \tilde{x}) \left(\hat{\theta}_{nj} - \theta_j\right)\right) \Bigg|, \end{aligned}$$

D_{2n} and D_{3n} are as in Section 4.3. Analogously to Lemma 6, we can deduce that

$$\begin{aligned} D_{1n} &\leq Cn^{-1}b^{-2}w_n \cdot \left(\left| \sum_{i=1}^{n} k'((\psi(\tilde{x}) - \tilde{Y}_i)b^{-1})\mathbf{1}\left\{\left|\psi(\tilde{x}) - \tilde{Y}_i\right| \leq b - w_n\right\} \right| \right. \\ &\quad \left. \psi'(\tilde{x}) \left(\sum_{j=1}^{d} \left|G_j(\theta, \tilde{x})\right| + \sum_{j=1}^{q} \left|\tilde{G}_j(\theta, \tilde{x})\right| \right) \right| \\ &\leq Cn^{-3/2}b^{-2}\sqrt{\ln\ln n} \cdot \left(\sqrt{nb\ln n} \right. \\ &\quad \left. + n \left| \mathbb{E}\left(k'((\psi(\tilde{x}) - \tilde{Y}_i)b^{-1})\mathbf{1}\left\{\left|\psi(\tilde{x}) - \tilde{Y}_i\right| \leq b - w_n\right\}\right) \right| \right) \\ &= o\left((nb)^{-1/2}\right) \quad a.s. \end{aligned}$$

The remainder of the proof is done as in the proof of Lemma 9. □

Proof of Theorem 6. By Lemmas 3 and 14, and analogously to the proof of Theorem 3, we obtain

$$\begin{aligned} \hat{\varphi}_n(x) - \varphi(x) &= \left(\mathcal{O}_S(S)^{-1} \det(\Sigma)^{-1} \tilde{x}^{1-d}\, \psi'(\tilde{x}) + o((nb)^{-1/2})\right) (\tilde{\chi}_n(\psi(\tilde{x})) - \chi(\psi(\tilde{x}))) \\ &\quad + o((nb)^{-1/2})\ a.s. \end{aligned}$$

The remainder of the proof can be done in the same manner as in the proof of Theorem 3. □

Acknowledgments: The authors are grateful to the referees who provided valuable suggestions on improving the paper and additional references.

Author Contributions: The authors contributed equally to this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bilodeau, M.; Brenner, D. *Theory of Multivariate Statistics*; Springer: New York, NY, USA, 1999.
2. Fang, K.-T.; Kotz, S.; Ng, K. *Symmetric Multivariate and Related Distributions*; Chapman & Hall: London, UK, 1990.
3. Fang, K.-T.; Zhang, Y. *Generalized Multivariate Analysis*; Springer: New York, NY, USA, 1990.
4. Fang, K.-T.; Anderson, T.W. (Eds.) *Statistical Inference in Elliptically Contoured and Related Distributions*; Allerton Press; New York, NY, USA, 1990.
5. Gupta, A.K.; Varga, T.; Bodnar, T. *Elliptically Contoured Models in Statistics and Portfolio Theory*; Springer-Verlag: New York, NY, USA, 2013.
6. Cui, H.; He, X. The consistence of semiparametric estimation of elliptic densities. *Acta Math. Sin. Engl. Ser.* **1995**, *11*, 44–58.
7. Liebscher, E. A semiparametric density estimator based on elliptical distributions. *J. Multivar. Anal.* **2005**, *92*, 205–225.
8. Stute, W.; Werner, U. Nonparametric estimation of elliptically contoured densities. In *Nonparametric Functional Estimation and Related Topics (Spetses, 1990)*; NATO Science Series C: Mathematical and Physical Sciences; Kluwer Academic Publisher: Dordrecht, The Netherlands, 1991; pp. 173–190.
9. Battey, H.; Linton, O. Nonparametric estimation of multivariate elliptic densities via finite mixture sieves. *J. Multivar. Anal.* **2014**, *123*, 43–67.
10. Fernandez, C.; Osiewalski, J.; Steel, M.F.J. Modeling and inference with ν-spherical distributions. *J. Am. Stat. Assoc.* **1995**, *90*, 1331–1340.
11. Balkema, A.A.; Embrechts, P.; Nolde, N. Meta densities and the shape of their sample clouds. *J. Multivar. Anal.* **2010**, *101*, 1738–1754.
12. Richter, W.-D. Geometric disintegration and star-shaped distributions. *J. Stat. Distrib. Appl.* **2014**, *1*, doi:10.1186/s40488-014-0020-6.
13. Dietrich, T.; Richter, W.-D. Classes of geometrically generalized von Mises distributions. *Sankhya B* **2016**, doi:10.1007/s13571-016-0118-6.
14. Richter,W.-D. Norm contoured distributions in R^2. *Lect. Notes Semin. Interdiscip. Mat.* **2015**, *12*, 179–199.
15. Richter, W.-D. Convex and radially concave contoured distributions. *J. Prob. Stat.* **2015**, *2015*, 165468.
16. Richter, W.-D.; Schicker, K. Polyhedral star-shaped distributions. Representations, properties and applications. *J. Prob. Stat.* **2016**, in press.
17. Moszyńska, M.; Richter, W.-D. Reverse triangle inequality. Antinorms and semi-antinorms. *Stud. Sci. Math. Hung.* **2012**, *49*, 120–138.
18. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*, 2nd ed.; Springer: New York, NY, USA, 1998.
19. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman & Hall: London, UK, 1986.
20. Müller, K.; Richter, W.-D. Exact distributions of order statistics of dependent random variables from $l_{n,p}$-symmetric sample distributions. *Depend. Model.* **2016**, *4*, doi:10.1515/demo-2016-0001.
21. Marron, J.S.; Wand, M.P. Exact mean integrated squared error. *Ann. Stat.* **1992**, *20*, 712–736.
22. Stute, W. A law of the logarithm for kernel density estimators. *Ann. Prob.* **1982**, *10*, 414–422.
23. Lindsey, J.K. Multivariate elliptically contoured distributions for repeated measurements. *Biometrics* **1999**, *55*, 1277–1280.
24. Cho, D.; Bin, T.D. Multivariate statistical modeling for image denoising using wavelet transforms. *Signal Process. Image Commun.* **2005**, *20*, 77–89.
25. Verdoolaege, G.; de Becker, S.; Scheunders, P. Multiscale colour texture retrieval using the geodesic distance between multivariate generalized models. In Proceedings of the 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008; pp. 169–172.
26. Field, C., Genton, M.C. The multivariate g-and h- distribution. *Technometrics* **2006**, *48*, 104–111.
27. Sinz, F.; Bethge, M. The conjoint effect of divisive normalization and orientation selectivity on redundancy reduction. In Proceedings of the 2008 Conference on Advances in Neural Information Processing Systems 21, Vancouver, BC, Canada, 8–10 December 2008; pp. 1521–1628.

28. Dang, U.J.; Browne, R.P.; McNicholas, P.D. Mixtures of multivariate power exponential distributions. *Biometrics* **2015**, *71*, 1081–1089.
29. Do, M.N.; Vetterli, M. Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler Distance. *IEEE Trans. Image Process.* **2002**, *11*, 146–158.
30. Santiago, E.; Albornoz, J.; Domínguez, A.; Toro, M.A.; López-Fanjul, C. The distribution of spontaneous mutations on quantitative traits and fitness in Drosophilamelanogaster. *Genetics* **1992**, *132*, 771–781.
31. Gómez, E.; Gomez-Viilegas, M.A.; Marín, J.M. A multivariate generalization of the power exponential family of distributions. *Commun. Stat. Theory Methods* **1998**, *27*, 589–600.
32. Wiaux, Y.; Puy, G.; Vandergheynst, P. Compressed sensing reconstruction of a string signal from interferometric observations of the cosmic microwave background. *Mon. Not. R. Astron. Soc.* **2010**, *402*, 2626–2636.
33. Chang, J.-H.; Shin, J.W.; Kim, N.S. Voice activity detector employing generalized Gaussian distribution. *Electron. Lett.* **2004**, *40*, 24.
34. Forbes, F.; Wraith, D. A new family of multivariate heavy-tailed distributions with variable marginal amounts of tail weight: Application to robust clustering. *Stat. Comput.* **2014**, *24*, 971–984.
35. Wraith, D.; Forbes, F. Location and scale mixtures of Gaussians with flexible tail behaviour: Properties, inference and application to multivariate clustering. *Comput. Stat. Data Anal.* **2015**, *90*, 61–73.
36. Lee, S.X.; McLachlan, G.J. Finite mixtures of canonical fundamental skew *t*-distributions. The unification of the restricted and unrestricted skew *t*-mixture models. *Stat. Comput.* **2016**, *26*, 573–589.
37. Richter, W.-D.; Venz, J. Geometric representations of multivariate skewed elliptically contoured distributions. *Chil. J. Stat.* **2014**, *5*, 71–90.
38. Andrews, D.F.; Herzberg, A.M. *Data: A Collection of Problems from Many Fields for the Student and Research Worker*; Springer: New York, NY, USA, 1985.
39. Van der Vaart, A.W. *Asymptotic Statistics*; Cambridge University Press: Cambridge, UK, 1998.
40. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076.

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).