The Annals of Statistics
1997, Vol. 25, No. 5, 1998–2017

ZONOID TRIMMING FOR MULTIVARIATE DISTRIBUTIONS

BY GLEB KOSHEVOY AND KARL MOSLER

C.E.M.I. and Universität zu Köln

A family of trimmed regions is introduced for a probability distribution in Euclidean d-space. The regions decrease with their parameter α, from the closed convex hull of support (at $\alpha = 0$) to the expectation vector (at $\alpha = 1$). The family determines the underlying distribution uniquely. For every α the region is affine equivariant and continuous with respect to weak convergence of distributions. The behavior under mixture and dilation is studied. A new concept of data depth is introduced and investigated. Finally, a trimming transform is constructed that injectively maps a given distribution to a distribution having a unique median.

1. Introduction. A most fundamental tool of data analysis is ordering. One approach is to order a set of multivariate observations via trimmed regions. For univariate data, a trimmed region is the interval between two properly chosen quantiles. Tukey (1975) and Eddy (1984) suggested a multivariate analogue of the quantile function. Based on such quantiles, concepts of multivariate trimmed regions have been introduced by Nolan (1992) and Massé and Theodorescu (1994). These trimmed regions are generalizations of univariate quantile intervals.

Newey and Powell (1987) defined the *expectiles* of a random variable. An expectile relates to the mean in a way similar to a quantile relating to the median; see also Abdous and Remillard (1995). Intervals between expectiles play the same role with respect to the mean as interquantile intervals do with respect to the median. Such interexpectile intervals have been studied in a more general setting by Averous and Meste (1990) and Breckling and Chambers (1988).

In this paper we propose a new concept of multivariate trimming. Based on the lift zonoid introduced recently in Koshevoy and Mosler (1997), we provide multivariate trimmed regions that are centered about the mean instead of some median. Our trimming regions give rise to a new concept of depth related to the mean. Both the trimming regions and the depth have nice mathematical properties, which are studied in detail. We also consider natural estimators of trimmed regions and investigate their asymptotic behavior.

Given a d-variate probability distribution μ, we define a family of trimmed regions as follows.

Received November 1995; revised November 1996.
AMS 1991 *subject classifications*. Primary 62H05; secondary 52A22, 60F05.
Key words and phrases. Trimmed regions, data depth, expectile, multivariate median, quantile.

DEFINITION 1.1. Let $\mathscr{M}$ denote the set of probability distributions μ on $(\mathbb{R}^d, \mathscr{B}^d)$ that have a finite expectation $E(\mu)$. For $\mu \in \mathscr{M}$, $\alpha \in (0, 1]$, we call

$$D_\alpha(\mu) = \left\{ \int_{\mathbb{R}^d} \mathbf{x} g(\mathbf{x})\, d\mu(\mathbf{x}) : g : \mathbb{R}^d \to \left[0, \frac{1}{\alpha}\right] \text{ measurable and } \int_{\mathbb{R}^d} g(\mathbf{x})\, d\mu(\mathbf{x}) = 1 \right\} \tag{1.1}$$

the zonoid α-trimmed region of μ. For $\alpha = 0$ we define

$$D_0(\mu) = \mathrm{cl}\left(\bigcup_{\alpha \in (0,1]} D_\alpha(\mu) \right) \tag{1.2}$$

where cl denotes closure.

In the univariate case, $d = 1$, the usual quantile trimming is a family of intervals of the form

$$D_\alpha^{qu}(\mu) = [Q(1-\alpha), Q(\alpha)], \qquad \alpha \in \left[\tfrac{1}{2}, 1\right) \tag{1.3}$$

with the usual left-continuous quantile function $Q(s) = \inf\{x \in \mathbb{R}: \mu(]-\infty, x]) \geq s\}$, $s \in (0, 1]$. We call them the *quantile α-trimmed region* of μ.

The zonoid trimmed regions are of the form

$$D_\alpha(\mu) = \left[\frac{1}{\alpha} \int_0^\alpha Q(s)\, ds, \frac{1}{\alpha} \int_{1-\alpha}^1 Q(s)\, ds \right]. \tag{1.4}$$

Equation (1.4) may be seen as an average quantile trimming; see also Section 8.

Newey and Powell (1987) define the τ-expectile of μ as the solution $\xi = \xi(\tau)$ of the equation

$$\frac{\tau}{1-\tau} = \frac{\int_{-\infty}^{\xi} (\xi - x)\, d\mu(x)}{\int_{\xi}^{\infty} (x - \xi)\, d\mu(x)}, \qquad 0 < \tau < 1. \tag{1.5}$$

A relation between the expectile function and the zonoid trimmed intervals is the following: the expectile is a convex combination of the mean and left or right ends of the zonoid trimmed interval.

For example, if μ is the empirical distribution on the points 0 and 1, then the interquantile intervals are $D_\alpha^{qu}(\mu) = [0, 1]$ for $\alpha \in (\frac{1}{2}, 1]$, the interexpectile intervals are $[\tau, 1-\tau]$ for $0 < \tau \leq \frac{1}{2}$ and the zonoid trimmed intervals are

$$D_\alpha(\mu) = \begin{cases} [0, 1], & \text{for } \alpha \in \left[0, \dfrac{1}{2}\right], \\[2ex] \left[1 - \dfrac{1}{2\alpha}, \dfrac{1}{2\alpha}\right], & \text{for } \alpha \in \left[\dfrac{1}{2}, 1\right). \end{cases}$$

If μ is the continuous uniform distribution on $[0, 1]$ we get $D_\alpha(\mu) = D_{1-\alpha/2}^{qu}(\mu) = [\alpha/2, 1 - \alpha/2]$ for every $\alpha \in (0, 1]$, while the interexpectile intervals are of the form $[\sqrt{\tau}/(\sqrt{\tau} + \sqrt{1-\tau}), 1 - \sqrt{\tau}/(\sqrt{\tau} + \sqrt{1-\tau})]$.

In the case of a continuous (with respect to Lebesgue measure) distribution, D_α may be defined in two other equivalent ways. For any Borel set U, let $E_U(\mu) = (\mu(U))^{-1}\int_U \mathbf{x}\, d\mu(\mathbf{x})$ denote the *U-centroid* of μ, provided $\mu(U) > 0$. If μ is continuous and $0 < \alpha \le 1$, then $D_\alpha(\mu)$ equals the set of all U-centroids, $U \in \mathscr{B}^d$, $\mu(U) = \alpha$, and, equivalently, the convex hull of all H-centroids, $H \in \mathscr{H}^d$, $\mu(H) = \alpha$, where $\mathscr{H}^d$ is the set of halfspaces in $\mathbb{R}^d$ (Theorem 3.1).

In view of this, the zoonid trimmed interval $[D_\alpha^-, D_\alpha^+]$ of a univariate continuous distribution has the following probabilistic interpretation: D_α^- and D_α^+ are the gravity centers of the lower respectively upper tails that have probability α.

To derive many important properties of the trimming, we establish a close relation between our α-trimmed regions and the so-called lift-zonoid of μ. This is the reason for naming our trimming the *zonoid trimming*. The zonoid of a probability distribution is a well-known notion; see Bolker (1969) and Goodey and Weil (1993). The lift-zonoid of a distribution $\mu \in \mathscr{M}$ has been introduced in Koshevoy and Mosler (1997) as follows. Let X be a random vector in $\mathbb{R}^d$ whose distribution is μ, and let $\hat{\mu}$ be the distribution of the "lifted" random vector $(1, X)$ in $\mathbb{R}^{d+1}$. The zonoid of $\hat{\mu}$ is called the *lift zonoid* of μ (and of X). Now, for some $\alpha \in (0, 1]$, consider the α-cut of the lift zonoid of μ, that is, its intersection with the hyperplane $\{\mathbf{x}: x_0 = \alpha\}$. The set $D_\alpha(\mu)$ comes out to be the projection of the α-cut on the last d coordinates, enlarged by $1/\alpha$ (Proposition 2.1).

Section 2 introduces the lift-zonoid and its relation to our notion of trimming, while Section 3 gives equivalent definitions of trimmed regions for continuous distributions. In Section 4 we characterize the trimmed regions of an empirical distribution in a way which easily lends itself to a computer code (Theorem 4.1).

Section 5 collects properties of the trimming. Under affine transformations of d-space $D_\alpha(\mu)$ is equivariant and, for $0 < \alpha \le 1$, a continuous function of μ and α. In particular, the α-trimmed region $D_\alpha(\widetilde{\mu}_n)$ of a random empirical distribution $\widetilde{\mu}_n$ converges to the α-trimmed region of the underlying law. As a consequence of this, in case $d = 1$, the usual law of large numbers for order statistics follows. The α-trimmed regions decrease in α; at $\alpha = 0$ we get the convex hull of the distribution's support and at $\alpha = 1$ the expectation vector. The family $\{D_\alpha(\mu)\}$ uniquely determines μ. Finally, we show that the α-trimmed region of a mixture of distributions is the union of certain mixtures of trimmed regions. Several examples and a numerical illustration are presented in Section 6.

In Section 7, we propose a new notion of data depth, the zonoid data depth. We establish its properties and contrast it with other notions of depths. In Section 8, we define an injective mapping, which we call the trimming transform. It maps a given distribution to a distribution that has a unique multivariate median. As a consequence, the zonoid data depth of μ equals twice the Tukey data depth of the trimming transform of μ.

Some notation: $\mathscr{M}_0$ is the set of distributions in $\mathscr{M}$ that are continuous with respect to Lebesgue measure, $\mathscr{M}_c$ the set of distributions in $\mathscr{M}$ that have a compact support, $\mathscr{M}_{0c} = \mathscr{M}_0 \cap \mathscr{M}_c$. Further, $\to_w$ denotes weak convergence of measures with corresponding convergence of the expectations of the norm and $\to_H$ convergence of sets in Hausdorff distance. The term $\mathrm{conv}(\mu)$ is the convex hull of the support of μ. For a set S and $\beta \in \mathbb{R}$, $\mathrm{conv}(S)$ is the convex hull of S, and $\beta \cdot S = \{\beta x: x \in S\}$; $\mathrm{cl}\, S$, $\mathrm{int}\, S$ and ∂S denote the closure, interior and boundary, respectively. Vectors are rows, $\langle \mathbf{x}, \mathbf{y} \rangle$ is the inner product of $\mathbf{x}$ and $\mathbf{y}$ and a prime denotes the transpose. We write S^{d-1} for the unit sphere in $\mathbb{R}^d$. For $\mathbf{p} \in S^{d-1}$ and $t \in \mathbb{R}$, $H(\mathbf{p}, t) = \{\mathbf{x} \in \mathbb{R}^d: \langle \mathbf{p}, \mathbf{x} \rangle \le t\}$, and $\mu_{\mathbf{p}}(t) = \mu(H(\mathbf{p}, t))$. If a random vector X in $\mathbb{R}^d$ is distributed by μ then $\mu_{\mathbf{p}}$ is the distribution function of the random variable $\langle X, \mathbf{p} \rangle$.

Of a point $\mathbf{x} \in \mathbb{R}^{d+1}$, the coordinates are indexed from 0 to d, $\mathbf{x} = (x_0, x_1, \dots, x_d)$. For a set S in $\mathbb{R}^{d+1}$ and $\alpha \in (0, 1]$, $S(\alpha)$ denotes the intersection S with the hyperplane $\{\mathbf{x} \in \mathbb{R}^{d+1}: x_0 = \alpha\}$; $\mathrm{proj}_\alpha(S) \in \mathbb{R}^d$ denotes the projection of $S(\alpha)$ to the last d coordinates.

2. The zonoid trimming.

The lift zonoid of a d-variate probability distribution is a convex set in $\mathbb{R}^{d+1}$ and is defined as follows.

DEFINITION 2.1. Let μ be a probability distribution. For a measurable function $h: \mathbb{R}^d \to [0, 1]$, consider the point $z(\mu, h) = (z_0(\mu, h), \zeta(\mu, h)) \in \mathbb{R}^{d+1}$, where

$$z_0(\mu, h) = \int_{\mathbb{R}^d} h(\mathbf{x})\, d\mu(\mathbf{x}), \qquad \zeta(\mu, h) = \int_{\mathbb{R}^d} \mathbf{x} h(\mathbf{x})\, d\mu(\mathbf{x}). \tag{2.1}$$

The set

$$\widehat{Z}(\mu) = \{z(\mu, h): h: \mathbb{R}^d \to [0, 1] \text{ measurable}\}$$

is called the lift-zonoid of μ. If X is a random vector distributed by μ, then $\widehat{Z}(X) = \widehat{Z}(\mu)$ is called the lift-zonoid of X.

The lift zonoid of a probability distibution may be seen as the set-valued expectation of a random segment.

Recall that a *random convex set* C is a Borel measurable map from a probability space $(\Omega, \mathscr{B}, P)$ to the space of nonempty, compact, convex subsets of $\mathbb{R}^k$. The *set-valued expectation*, $E(C)$, of a random convex set C is the set given implicitly by

$$\phi_{E(C)}(\mathbf{p}) = E(\phi_C(\mathbf{p})), \qquad \mathbf{p} \in \mathbb{R}^k, \tag{2.2}$$

where $\phi_C(\mathbf{p}) = \max_{\mathbf{x} \in C} \langle \mathbf{x}, \mathbf{p} \rangle$, $\mathbf{p} \in S^{k-1}$, is the support function of C. This set-valued expectation has been used in different settings; see Weil and Wieacker (1993). If there exists a finite expectation of the norm of a random set, then its expectation is a compact set.

PROPOSITION 2.1 [Koshevoy and Mosler (1997)]. *Let $\mu \in \mathscr{M}_0$ and X be a random vector distributed by μ. Then $\widehat{Z}(X) = E([\mathbf{0}, X])$.*

Next we state some properties of the lift zonoid which will be of use in the sequel.

THEOREM 2.1 [Koshevoy and Mosler (1997)]. (i) *For $\mu \in \mathscr{M}$, the lift zonoid $\widehat{Z}(\mu)$ is a convex compact set. It contains $\mathbf{0} \in \mathbb{R}^{d+1}$ and is symmetric about $(\frac{1}{2}, \frac{1}{2}E(\mu))$. If the support of μ is in $\mathbb{R}^d_+$, then $\widehat{Z}(\mu)$ is contained in the $(d+1)$-dimensional rectangle between $\mathbf{0}$ and $(1, \mathrm{E}(\mu))$.*

(ii) *A distribution $\mu \in \mathscr{M}$ is uniquely determined by its lift zonoid.*

(iii) *If $\mu_n \to_w \mu$, then $\widehat{Z}(\mu_n) \to_H \widehat{Z}(\mu)$.*

(iv) *Given $\mu \in \mathscr{M}$, there exists a sequence (μ_n) of empirical distributions such that $\mu_n \to_w \mu$, $\widehat{Z}(\mu_n) \to_H \widehat{Z}(\mu)$ and $\widehat{Z}(\mu_n) \subset \widehat{Z}(\mu)$ for all n.*

Proposition 2.2, whose proof is immediate from Definition 2.1, provides the relation between our α-trimmed regions and the lift-zonoid and thus justifies the name "zonoid trimming."

PROPOSITION 2.2. *Let $\alpha \in (0, 1]$. Then*

$$D_\alpha(\mu) = \frac{1}{\alpha} \operatorname{proj}_\alpha(\widehat{Z}(\mu)). \tag{2.3}$$

Let us explain what the zonoid trimmed regions mean in the univariate case and establish their relation with the interexpectile intervals. If $d = 1$, $\widehat{Z}(\mu)$ is the region between the generalized Lorenz curve and its dual, that is, it is the convex hull of the following points in $\mathbb{R}^2$: $(0, 0)$, $(1, E(\mu))$,

$$\left(\int_{(-\infty, y]} \mu(dx), \int_{(-\infty, y]} x\mu(dx)\right), \qquad y \in \mathbb{R} \quad \text{and} \tag{2.4}$$

$$\left(\int_{[y, \infty)} \mu(dx), \int_{[y, \infty)} x\mu(dx)\right), \qquad y \in \mathbb{R}. \tag{2.5}$$

Or, equivalently, $\widehat{Z}(\mu)$ is the convex hull of

$$\left(t, \int_0^t Q(s)\,ds\right), \qquad 0 \le t \le 1 \quad \text{and} \quad \left(t, \int_{1-t}^1 Q(s)\,ds\right), \qquad 0 \le t \le 1.$$

Therefore, in such a case,

$$D_\alpha(\mu) = \left[\frac{1}{\alpha}\int_0^\alpha Q(s)\,ds, \frac{1}{\alpha}\int_{1-\alpha}^1 Q(s)\,ds\right], \tag{2.6}$$

that is, (1.3).

A relation between the expectile function $\xi = \xi(\tau)$ defined by (1.5) and the zonoid trimmed intervals is the following: let μ be a continuous univariate distribution, $E(\mu)$ be its mean. Given τ, denote $\alpha^* = \mu((-\infty, \xi(\tau)])$, and let D_α^-, D_α^+ denote the left and right ends of the interval $D_\alpha(\mu)$. Then

$$\xi(\tau) = \begin{cases} \dfrac{\tau E(\mu) + (1-2\tau)\alpha^* D_{\alpha^*}^-}{\tau + (\alpha^* - 2\tau\alpha^*)}, & \text{if } \tau \le \dfrac{1}{2}, \\[2ex] \dfrac{(1-\tau)E(\mu) + (2\tau-1)(1-\alpha^*)D_{\alpha^*}^+}{(1-\tau) + (2\tau-1)(1-\alpha^*)}, & \text{if } \tau \ge \dfrac{1}{2}. \end{cases} \tag{2.7}$$

In view of (2.7), $\xi(1/2) = E(\mu)$ and the continuity of trimmed intervals, the interexpectile interval $[\xi(\tau), \xi(1-\tau)]$ has the form $[D^-_{\alpha_1}, D^+_{\alpha_2}]$ for some $\alpha_1 > \alpha^*$ and $\alpha_2 < \mu((1-\xi(\tau), \infty))$.

3. Continuous distributions. Recall that, given a Borel set U, $\mu(U) > 0$, a vector $E_U(\mu) = (\mu(U))^{-1} \int_U \mathbf{x}\, d\mu(\mathbf{x})$ is the U-centroid of μ. The U-centroid of μ is the gravity center of U with respect to μ. When we apply our notion, with respect to Lebesgue measure, to continuous distributions, the following holds.

THEOREM 3.1. *Let* $\mu \in \mathscr{M}_0$ *and* $\alpha > 0$. *Then we have the following*:

(i) $D_\alpha(\mu) = \{E_U(\mu): U \in \mathscr{B}^d,\ \mu(U) = \alpha\}$.
(ii) $D_\alpha(\mu) = \mathrm{conv}\{E_H(\mu): H \in \mathscr{H}^d,\ \mu(H) = \alpha\}$.

The theorem says that, for $0 < \alpha \le 1$, the α-trimmed region of a continuous distribution is equal to the set of U-centroids where U has probability α. In other words, a point y belongs to $D_\alpha(\mu)$ iff there is a Borel set having probability α of which y is the gravity center. Note that, in the continuous distribution case, for every α there is some U with $\mu(U) = \alpha$; hence, given α, the set of centroids is nonempty. Equivalently, $D_\alpha(\mu)$ is the convex hull of U-centroids where U is a halfspace having probability α. For $\alpha = 0$, $D_0(\mu)$ is the convex hull of the support of μ.

PROOF. (i) In the case of a continuous μ, the lift zonoid is given by all points of the form (2.1) with functions h that are indicator functions of sets in $\mathscr{B}^d$. This follows from Liapunov's theorem [Bolker (1969)]. With respect to (2.3), we get that every point of $D_\alpha(\mu)$ has the form $(\mu(U))^{-1} \int_U \mathbf{x}\, d\mu(\mathbf{x})$ with some U such that $\mu(U) = \alpha$. This proves (i).

(ii) Extreme points of $\widehat{Z}(\mu)$ have the form (2.1) with h being the indicator function of a halfspace; see Koshevoy and Mosler (1997), Theorem 2.1. Due to the continuity of μ, for every $\alpha > 0$ and every direction in $\mathbb{R}^d$, there is a halfspace whose normal vector equals the chosen direction and which has μ-probability α. Extreme points of a section $x_0 = \alpha$ remain extreme points of its projection on the last d coordinates. That yields the proof of (ii). □

4. Empirical distributions. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be arbitrary points in $\mathbb{R}^d$. By an *empirical distribution* on $\mathbf{x}_1, \dots, \mathbf{x}_n$, we mean a probability distribution that assigns probability $1/n$ to each $\mathbf{x}_i$.

The lift-zonoid of an empirical distribution μ on $\mathbf{x}_1, \dots, \mathbf{x}_n$ [see Koshevoy and Mosler (1997)] is the sum of line segments $[\mathbf{0}, (1/n, \mathbf{x}_i/n)]$,

$$\widehat{Z}(\mu) = \sum_{i=1}^{n} \left[\mathbf{0}, \left(\frac{1}{n}, \frac{\mathbf{x}_i}{n}\right)\right]. \tag{4.1}$$

THEOREM 4.1. *Let* $\alpha \in [k/n, (k+1)/n]$, $k = 1, \dots, n-1$. *Then*

$$D_\alpha(\mu) = \operatorname{conv}\left\{ \frac{1}{\alpha n} \sum_{j=1}^{k} \mathbf{x}_{i_j} + \left(1 - \frac{k}{\alpha n}\right) \mathbf{x}_{i_{k+1}} : \{i_1, \dots, i_{k+1}\} \subset N \right\}, \tag{4.2}$$

where $N = \{1, \dots, n\}$. *For* $\alpha \in [0, 1/n]$,

$$D_\alpha(\mu) = \operatorname{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\} = \bigcap \{H \in \mathscr{H}^d : \mu(H) = 1\}. \tag{4.3}$$

PROOF. It follows from the definition that $\widehat{Z}(\mu)$ is a convex polytope,

$$\begin{aligned} \widehat{Z}(\mu) &= \operatorname{conv}\left\{ \sum_{i=1}^{n} \delta_i (1/n, \mathbf{x}_i/n) : \delta_i \in \{0, 1\}, i \in N \right\} \\ &= \operatorname{conv}\left(\bigcup_{k=0}^{n} \mathscr{V}_k \right), \end{aligned} \tag{4.4}$$

where $\mathscr{V}_k = \{\sum_{j=1}^{k} (1/n, \mathbf{x}_{i_j}/n) : \{i_1, \dots, i_k\} \subset N\}$. The $\mathscr{V}_k$ are pairwise disjoint sets. Denote $\widehat{Z}(\mu, \alpha) = \widehat{Z}(\mu) \cap \{\mathbf{x} : x_0 = \alpha\}$. Then every extreme point z^* of $\widehat{Z}(\mu, \alpha)$ is either an extreme point of $\widehat{Z}(\mu)$ or the intersection of an edge of $\widehat{Z}(\mu)$ with the hyperplane at $x_0 = \alpha$. Note that every edge, that is, one-dimensional face, of $\widehat{Z}(\mu)$ can be written [Shephard (1974)]

$$\left[\mathbf{0}, \left(\frac{1}{n}, \frac{\mathbf{x}_s}{n}\right)\right] + \sum_{i \neq s} \epsilon_i \left(\frac{1}{n}, \frac{\mathbf{x}_i}{n}\right), \tag{4.5}$$

where $\epsilon_i \in \{0, 1\}$, $i \neq s$, $s = 1, \dots, n$. We proceed in three steps.

Step 1. First we prove (4.2) for $\alpha = k/n$, $k = 1, \dots, n$. Obviously, $\operatorname{conv}(\mathscr{V}_k) \subset \widehat{Z}(\mu, k/n)$ holds. Let z^* be an extreme point of $\widehat{Z}(\mu, k/n)$. Then z^* either is an extreme point of $\widehat{Z}(\mu)$, hence in $z^* \in \mathscr{V}_k$, or belongs to an edge of $\widehat{Z}(\mu)$; in the latter case (4.5) implies that $z^* \in \mathscr{V}_k$. We conclude that $\operatorname{conv}(\mathscr{V}_k) = \widehat{Z}(\mu, k/n)$. Then $D_{k/n}(\mu) = n/k \operatorname{proj}_{k/n}(\widehat{Z}(\mu)) = \operatorname{conv}\{\sum_{j=1}^{k} \mathbf{x}_{i_j}/k : \{i_1, \dots, i_k\} \subset N\}$. That yields (4.2) for $\alpha = k/n$.

Step 2. Second, we show (4.2) for $\alpha \in (k/n, (k+1)/n)$, $k = 1, \dots, n-1$. Again, let z^* be an extreme point of $\widehat{Z}(\mu, k/n)$. As z^* cannot be an extreme point of $\widehat{Z}(\mu)$, from (4.5) we get

$$z^* = \left(\alpha - \frac{k}{n}\right)\left(1, \mathbf{x}_{i_{k+1}}\right) + \sum_{j=1}^{k} \frac{1}{n}\left(1, \mathbf{x}_{i_j}\right). \tag{4.6}$$

The convex hull of all z^* is $\widehat{Z}(\mu, \alpha)$, given by (4.6). So,

$$\frac{1}{\alpha} \operatorname{proj}_\alpha(\widehat{Z}(\mu)) = \operatorname{conv}\left\{ \sum_{j=1}^{k} \frac{1}{n\alpha} \mathbf{x}_{i_j} + \left(1 - \frac{k}{n\alpha}\right) \mathbf{x}_{i_{k+1}} : \{i_1, \dots, i_{k+1}\} \subset N \right\}.$$

That yields (4.2) also for $\alpha \in (k/n, (k+1)/n)$.

Step 3. It remains to prove (4.3), the second equality of which is obvious. For $\alpha = 1/n$, $D_\alpha(\mu) = \operatorname{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ follows from (4.2). Let $\alpha \in (0, 1/n)$ and $\mathbf{x} \in \widehat{Z}(\mu, \alpha)$. Then $\mathbf{x}$ is no extreme point of $\widehat{Z}(\mu)$; it is a convex combination of the origin $\mathbf{0}$ and some $\mathbf{y} \in \widehat{Z}(\mu, 1/n)$. As $x_0 = \alpha$, we get $\mathbf{x} = \alpha\mathbf{y}$. Further, since $\widehat{Z}(\mu, 1/n) = \operatorname{conv}(\mathscr{V}_1) = (1/n)\operatorname{conv}\{(1, \mathbf{x}_1), \dots, (1, \mathbf{x}_n)\}$, we conclude that $\widehat{Z}(\mu, \alpha) \subset \alpha \cdot \operatorname{conv}\{(1, \mathbf{x}_1), \dots, (1, \mathbf{x}_n)\}$. As the reverse set inclusion is obvious, we get

$$D_\alpha(\mu) = \frac{1}{\alpha} \operatorname{proj}_\alpha(\widehat{Z}(\mu)) = \operatorname{conv}\{\mathbf{x}_1, \dots, \mathbf{x}_n\};$$

hence (4.3) for $\alpha \in (0, 1/n]$. For $\alpha = 0$, (4.3) follows from this and the definition of $D_0(\mu)$. This completes the proof. □

An immediate consequence of Theorem 4.1 is the monotonicity of D_α for empirical distributions.

COROLLARY 4.1. *Let μ be an empirical distribution. Then, for $0 \le \alpha_1 \le \alpha_2 \le 1$,*

$$D_{\alpha_2}(\mu) \subset D_{\alpha_1}(\mu).$$

5. General properties of zonoid trimmed regions.

Various properties of zonoid trimmed regions are collected in the following theorems. The first two state that $D_\alpha(\mu)$ is a continuous function of α and μ.

THEOREM 5.1. *Let $\mu \in \mathscr{M}$, $\alpha_n \to \alpha$ and $\alpha > 0$. Then*

$$D_{\alpha_n}(\mu) \to_H D_\alpha(\mu).$$

PROOF. Let $\alpha_n \to \alpha > 0$. Then, due to the convexity and compactness of $\widehat{Z}(\mu)$, we get $\operatorname{proj}_{\alpha_n} \widehat{Z}(\mu) \to_H \operatorname{proj}_\alpha \widehat{Z}(\mu)$. Therefore, $\lim_n D_{\alpha_n} = \lim_n (1/\alpha_n) \times \operatorname{proj}_{\alpha_n} \widehat{Z}(\mu) = (1/\alpha) \lim_n \operatorname{proj}_{\alpha_n} \widehat{Z}(\mu) = (1/\alpha)\operatorname{proj}_\alpha \widehat{Z}(\mu) = D_\alpha(\mu)$. □

THEOREM 5.2. *Let $\mu_n \to_w \mu$ in $\mathscr{M}$. Then we have the following*:

(i) $D_\alpha(\mu) = \lim_n D_\alpha(\mu_n)$ *in the Hausdorff distance if* $0 < \alpha \le 1$,
(ii) $D_0(\mu) \subset \lim_n D_0(\mu_n)$.

PROOF. Let $\mu_n \to_w \mu$. Then $\widehat{Z}(\mu_n) \to_H \widehat{Z}(\mu)$ by Theorem 2.1(iii). Therefore, for any $\alpha \in (0, 1]$, $\widehat{Z}(\mu_n, \alpha) \to_H \widehat{Z}(\mu, \alpha)$ holds where $\widehat{Z}(\nu, \alpha) = \widehat{Z}(\nu) \cap \{\mathbf{x}: x_0 = \alpha\}$. So, $D_\alpha(\mu_n) \to_H D_\alpha(\mu)$ holds for $\alpha \in (0, 1]$. The convergence is not uniform in α, therefore $D_0(\mu) = \operatorname{cl}(\bigcup_{\alpha \in (0,1]} D_\alpha(\mu)) = \operatorname{cl}(\bigcup_{\alpha \in (0,1]} \lim_n D_\alpha(\mu_n)) \subset \lim_n \operatorname{cl}(\bigcup_{\alpha \in (0,1]} D_\alpha(\mu_n)) = \lim_n D_0(\mu_n)$. □

A law of large numbers follows from this theorem. Let $X_1, \dots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$ that are distributed with μ, and let $\widetilde{\mu}_n$ be their random empirical distribution. By the Glivenko–Cantelli theorem we know that μ-a.s. $\lim_{n\to\infty} \widetilde{\mu}_n = \mu$, and from Theorem 5.2 we derive the corollary

COROLLARY 5.1. *$\lim_{n\to\infty} D_\alpha(\widetilde{\mu}_n) = D_\alpha(\mu)$ holds μ-a.s. if $0 < \alpha \le 1$.*

In the univariate case we get the following. Let $X_1, \dots, X_n$ be given i.i.d. random variables in $\mathbb{R}$ that are distributed with μ, and let $X_{1:1} \le \cdots \le X_{n:n}$ denote their order statistics. Then (Theorem 4.1)

$$D_\alpha(\widetilde{\mu}_n) = \Bigg[\frac{(1-\{n\alpha\})[n\alpha]}{n\alpha}\left(\frac{\sum_{i=1}^{[n\alpha]} X_{i:n}}{[n\alpha]}\right) + \frac{\{n\alpha\}([n\alpha]+1)}{n\alpha}\left(\frac{\sum_{i=1}^{[n\alpha]+1} X_{i:n}}{[n\alpha]+1}\right),$$
$$\frac{(1-\{n\alpha\})[n\alpha]}{n\alpha}\left(\frac{\sum_{i=n-[n\alpha]+1}^{n} X_{i:n}}{[n\alpha]}\right)$$
$$+ \frac{\{n\alpha\}([n\alpha]+1)}{n\alpha}\left(\frac{\sum_{i=n-[n\alpha]}^{n} X_{i:n}}{[n\alpha]+1}\right)\Bigg].$$

As usual, $\{\beta\}$ and $[\beta]$ are the fraction and the integer parts of a real number β. Obviously,

$$\lim_{n\to\infty} D_\alpha(\widetilde{\mu}_n) = \lim_{n\to\infty}\left[\frac{\sum_{i=1}^{[n\alpha]} X_{i:n}}{n\alpha}, \frac{\sum_{i=n-[n\alpha]+1}^{n} X_{i:n}}{n\alpha}\right]. \tag{5.1}$$

Thus, in the univariate case, from Corollary 5.1 we derive the law of large numbers for order statistics.

$$\left[\frac{\sum_{i=1}^{[n\alpha]} X_{i:n}}{n\alpha}, \frac{\sum_{i=n-[n\alpha]+1}^{n} X_{i:n}}{n\alpha}\right] \to \left[\frac{\int_0^\alpha Q(s)\,ds}{\alpha}, \frac{\int_{1-\alpha}^1 Q(s)\,ds}{\alpha}\right]$$

μ-a.s. if $0 < \alpha \le 1$. From this we get

$$\frac{\sum_{i=1}^{[n\alpha]} X_{i:n}}{n\alpha} \to \frac{\int_0^\alpha Q(s)\,ds}{\alpha}, \quad \mu\text{–a.s.} \tag{5.2}$$

$$\frac{\sum_{i=n-[n\alpha]+1}^{n} X_{i:n}}{n\alpha} \to \frac{\int_{1-\alpha}^1 Q(s)\,ds}{\alpha}, \quad \mu\text{–a.s.} \tag{5.3}$$

For empirical distributions in $\mathbb{R}^d$ we have already shown (Corollary 4.1) that the D_α decrease with α. The result is now extended to general distributions. Moreover, it is shown that D_α is strictly decreasing as soon as it differs from D_0.

THEOREM 5.3. *Let $\mu \in \mathscr{M}$, $0 \le \alpha_1 < \alpha_2 \le 1$. Then:*

(i) $D_{\alpha_2}(\mu) \subset D_{\alpha_1}(\mu) \subset D_0(\mu)$.
(ii) *If $D_{\alpha_1}(\mu) \ne D_0(\mu)$ then $D_{\alpha_2}(\mu) \ne D_{\alpha_1}(\mu)$.*

PROOF. (i) Let μ be an empirical distribution. Then Corollary 4.1 says that $D_\alpha(\mu)$ is monotone decreasing in α. For a general μ, according to Theorem 2.1(iv), there exists a sequence of empirical distributions $\mu_n \to_w \mu$ such that $\widehat{Z}(\mu_n) \to_H \widehat{Z}(\mu)$ for all n. Therefore, for $\alpha_2 \geq \alpha_1 > 0$, because of Theorem 5.2, $D_{\alpha_2}(\mu) = \lim_n D_{\alpha_2}(\mu_n) \subset \lim_n D_{\alpha_1}(\mu_n) = D_{\alpha_1}(\mu)$. For $\alpha_1 = 0$, the proposition follows from the definition of D_0.

(ii) Let $D_{\alpha_1}(\mu) \neq D_0(\mu)$ and assume that $D_{\alpha_2}(\mu) = D_{\alpha_1}(\mu)$. Then

$$\widehat{Z}(\mu) \cap \{x_0 = \alpha_1\} = \frac{\alpha_1}{\alpha_2}\big(\widehat{Z}(\mu) \cap \{x_0 = \alpha_2\}\big). \tag{5.4}$$

As $\widehat{Z}(\mu)$ is convex and $\mathbf{0} \in \widehat{Z}(\mu)$, (5.4) implies that, for any $0 < \alpha \leq \alpha_1$,

$$\widehat{Z}(\mu) \cap \{x_0 = \alpha\} \subset \frac{\alpha}{\alpha_2}\big(\widehat{Z}(\mu) \cap \{x_0 = \alpha_2\}\big), \tag{5.5}$$

hence $D_\alpha(\mu) \subset D_{\alpha_2}(\mu)$. We conclude that $D_0(\mu) = \mathrm{cl}(\bigcup_{\alpha>0} D_\alpha(\mu)) \subset \mathrm{cl}(D_{\alpha_2}(\mu)) = D_{\alpha_2}(\mu)$, which contradicts our assumption. □

The trimming regions have the following geometric properties.

THEOREM 5.4. *Let $\mu \in \mathscr{M}$.*

(i) *$D_\alpha(\mu)$ is convex and closed if $0 \leq \alpha \leq 1$ and, in addition, compact if $0 < \alpha \leq 1$.*

(ii) *If the support of μ is in $\mathbb{R}^d_+$, then $D_\alpha(\mu)$ is contained in the d-dimensional rectangle between $\mathbf{0}$ and $(1/\alpha)\mathrm{E}(\mu)$.*

PROOF. (i) The property is derived from Theorem 2.1(i). For $\alpha = 0$ we use the monotonicity of trimming regions (Theorem 5.3).

(ii) For μ with support in $\mathbb{R}^d_+$, $\widehat{Z}(\mu)$ is a subset of the $(d+1)$-dimensional rectangle between $\mathbf{0}$ and $(1, \mathrm{E}(\mu))$. The d-dimensional rectangle between $\mathbf{0}$ and $\mathrm{E}(\mu)$ is proj_α of this rectangle, so $D_\alpha(\mu)$ belongs to the d-dimensional rectangle between $\mathbf{0}$ and $(1/\alpha)\mathrm{E}(\mu)$. □

The following theorem says that the zonoid 0-trimmed region is the convex hull of the support of the distribution.

THEOREM 5.5. *For $\mu \in \mathscr{M}$, $D_0(\mu) = \bigcap\{H \in \mathscr{H}^d: \mu(H) = 1\}$.*

PROOF. Let $\mu_n \to_w \mu$ be a sequence of empirical distributions according to Theorem 2.1(iv). As $\widehat{Z}(\mu_n) \subset \widehat{Z}(\mu)$, there holds $D_\alpha(\mu_n) \subset D_\alpha(\mu)$ for all $\alpha > 0$, and therefore $D_0(\mu_n) \subset D_0(\mu)$. We conclude that $\lim_n D_0(\mu_n) \subset D_0(\mu)$. Theorem 5.2 yields that $D_0(\mu) \subset \lim_n D_0(\mu_n)$, hence $D_0(\mu) = \lim_n D_0(\mu_n)$. Due to (4.3), $D_0(\mu_n) = \bigcap\{H \in \mathscr{H}^d: \mu_n(H) = 1\}$ holds. When n goes to infinity we get $D_0(\mu) = \lim_n D_0(\mu_n) = \bigcap\{H \in \mathscr{H}^d: \mu(H) = 1\}$. □

Now we show a uniqueness result: the family $\{D_\alpha(\mu)\}$ uniquely determines μ in $\mathscr{M}$.

THEOREM 5.6. *Let* $\mu, \nu \in \mathscr{M}$.

(i) *If* $D_\alpha(\mu) = D_\alpha(\nu)$ *for all* $\alpha \in (0, \frac{1}{2}]$ *then* $\mu = \nu$.
(ii) *If* $D_\alpha(\mu) = D_\alpha(\nu)$ *for all* $\alpha \in [\frac{1}{2}, 1)$ *then* $\mu = \nu$.

PROOF. Due to Theorem 2.1(ii), $D_\alpha(\mu) = D_\alpha(\nu)$ for $0 < \alpha \le 1$ implies $\mu = \nu$. As the lift-zonoid $\widehat{Z}(\mu)$ is symmetric, we get that

$$\mathrm{proj}_\alpha(\widehat{Z}(\mu)) - \alpha \mathrm{E}(\mu) \text{ is centrally symmetric to } \mathrm{proj}_{1-\alpha}(\widehat{Z}(\mu)) - (1-\alpha)\mathrm{E}(\mu).$$

Therefore it is sufficient to have $D_\alpha(\mu) = D_\alpha(\nu)$ either for α in the lower or in the upper half of the interval. □

Theorem 5.7 states that the zoonid trimming is *affine equivariant*.

THEOREM 5.7. *Let A be a regular $d \times d$ matrix, $\mathbf{a}$ be a vector in $\mathbb{R}^d$ and $\mu \in \mathscr{M}$. We denote $\mu^{(A,\mathbf{a})}(\mathbf{x}) = \mu(A^{-1}(\mathbf{x}-\mathbf{a}))$. Then*

$$D_\alpha(\mu) = A D_\alpha(\mu^{(A,\mathbf{a})}) + \mathbf{a}, \qquad \alpha \in [0,1].$$

PROOF. According to Theorems 2.1(iv) and 5.2, we restrict ourselves to the case of an empirical distribution. Let μ be an empirical distribution with support $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. Then $\mu^{(A,\mathbf{a})}$ is an empirical distribution with support $\{A^{-1}(\mathbf{x}_1 - \mathbf{a}), \dots, A^{-1}(\mathbf{x}_k - \mathbf{a})\}$. Therefore, in view of (4.2), $D_\alpha(\mu^{(A,\mathbf{a})}) = A^{-1}(D_\alpha(\mu) - \mathbf{a})$. □

Finally, we prove that the α-trimmed region of a mixture of distributions is the union of certain mixtures of trimmed regions.

THEOREM 5.8. *Let* $\mu, \nu \in \mathscr{M}$, $\beta \in [0,1]$. *Then*

$$\begin{aligned} &D_\alpha(\beta\mu + (1-\beta)\nu) \\ &\quad = \bigcup \left\{ \frac{\beta\delta}{\alpha} D_\delta(\mu) + \frac{(1-\beta)\delta'}{\alpha} D_{\delta'}(\nu) \colon \beta\delta + (1-\beta)\delta' = \alpha \right\}. \end{aligned} \tag{5.6}$$

PROOF. $\widehat{Z}(\beta\mu + (1-\beta)\nu) = \beta\widehat{Z}(\mu) + (1-\beta)\widehat{Z}(\nu)$. Therefore,

$$\begin{aligned} &\widehat{Z}(\beta\mu + (1-\beta)\nu, \alpha) \\ &\quad = \bigcup \left\{ \beta\widehat{Z}(\mu,\delta) + (1-\beta)\widehat{Z}(\nu,\delta') \colon \beta\delta + (1-\beta)\delta' = \alpha \right\}. \end{aligned} \tag{5.7}$$

Thus,

$$\begin{aligned} &\frac{1}{\alpha}\mathrm{proj}_\alpha(\widehat{Z}(\beta\mu + (1-\beta)\nu)) \\ &\quad = \bigcup \left\{ \frac{\beta}{\alpha}\mathrm{proj}_\delta(\widehat{Z}(\mu)) + \frac{1-\beta}{\alpha}\mathrm{proj}_{\delta'}(\widehat{Z}(\nu)) \colon \beta\delta + (1-\beta)\delta' = \alpha \right\} \\ &\quad = \bigcup \left\{ \frac{\beta\delta}{\alpha} D_\delta(\mu) + \frac{(1-\beta)\delta'}{\alpha} D_{\delta'}(\nu) \colon \beta\delta + (1-\beta)\delta' = \alpha \right\}, \end{aligned}$$

that is (5.6). □

6. Examples. Let us consider several examples. The expression $B(\mathbf{x}, r)$ denotes the Euclidean ball about some $\mathbf{x} \in \mathbb{R}^d$ with radius r.

1. We consider the uniform distribution μ on the unit disc $B(\mathbf{x}, 1)$ in $\mathbb{R}^2$. Then

$$D_\alpha(\mu) = B(\mathbf{x}, r_\alpha) \quad \text{with } r_\alpha = \frac{(1 - c^2(\alpha))^{3/2}}{3\pi\alpha},$$

   where $c(\alpha) \in [0, 1]$ is the solution to $2\pi\alpha = \pi/2 - \arcsin c - c\sqrt{1 - c^2}/2$ if $\alpha \le 1/2$ and is the solution to $2\pi\alpha = \pi/2 - \arcsin c + c\sqrt{1 - c^2}/2$ if $\alpha > 1/2$.

2. Let μ be the uniform distribution on the boundary of the unit disc $B(\mathbf{x}, 1)$ in $\mathbb{R}^2$. Then

$$D_\alpha(\mu) = B(\mathbf{x}, r_\alpha) \quad \text{with } r_\alpha = \begin{cases} \dfrac{\sin \pi\alpha}{\pi\alpha}, & \text{if } \alpha > 0, \\ 1, & \text{if } \alpha = 0. \end{cases}$$

3. Let μ be a multivariate standard normal distribution, $\mu = N(\mathbf{0}, I)$. Then

$$D_\alpha(\mu) = B(\mathbf{0}, r_\alpha) \qquad \text{with } r_\alpha = \frac{\exp(-t^2(\alpha)/2)}{\sqrt{2\pi}\alpha}, \qquad \alpha \in [0, 1].$$

   Here $t(\alpha)$ is the unique solution to $\int_t^\infty \exp(-x^2/2)\,dx = \sqrt{2\pi}\alpha$. In view of Theorem 5.7, the trimmed regions of a general nondegenerate multivariate normal $N(\mathbf{a}, \Sigma)$ are given by

$$D_\alpha(N(\mathbf{a}, \Sigma)) = \{\mathbf{x}\colon (\mathbf{x} - \mathbf{a})^T \Sigma^{-1} (\mathbf{x} - \mathbf{a}) \le r_\alpha^2\}. \tag{6.1}$$

Recall that for every $\mu \in \mathscr{M}$ that has a positive definite covariance matrix Σ_μ, the Mahalanobis distance [Mahalanobis (1936)] gives rise to the trimmed regions

$$D_\alpha(\mu) = \{\mathbf{x}\colon (\mathbf{x} - E(\mu))^T \Sigma_\mu^{-1} (\mathbf{x} - E(\mu)) \le \alpha\}, \qquad 0 \le \alpha < \infty. \tag{6.2}$$

These regions parallel the zonoid trimmed regions of the multivariate normal $N(E(\mu), \Sigma_\mu)$.

In the nondegenerate normal case, the zonoid trimmed regions are parallel to the trimmed regions defined by Massé and Theodorescu (1994), and both regions are parallel to the Mahalanobis trimmed regions.

Figure 1 exhibits the trimmed regions, for $\alpha = 0.0, 0.1, 0.2, \ldots, 1.0$, of an empirical distribution on ten given points in the plane.

7. Zonoid data depth. We introduce a new notion of data depth, called zonoid data depth and investigate its properties.

DEFINITION 7.1. The zonoid data depth, $\text{depth}_\mu(\mathbf{x})$, of a point $\mathbf{x} \in \mathbb{R}^d$ is defined by

$$\text{depth}_\mu(\mathbf{x}) = \begin{cases} \sup\{\alpha\colon \mathbf{x} \in D_\alpha(\mu)\}, & \text{if } \mathbf{x} \in D_\alpha(\mu) \text{ for some } \alpha, \\ 0, & \text{otherwise.} \end{cases} \tag{7.1}$$

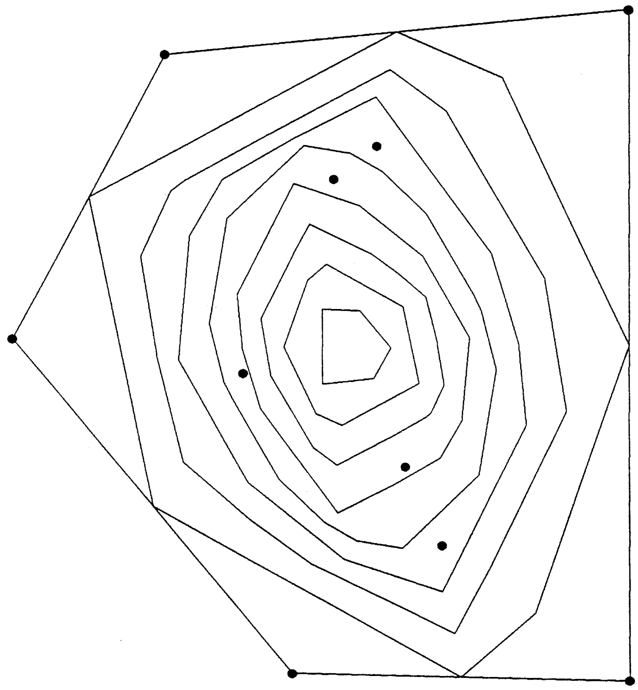

FIG. 1. *Zonoid trimmed regions, for* $\alpha = 0.0, 0.1, 0.2, \dots, 1.0$, *of an empirical distribution on ten points in the plane.*

A geometrical interpretation reads as follows. The data depth of a point $\mathbf{x}$ is the maximal height α at which $\alpha\mathbf{x} \in \mathrm{proj}_\alpha\, \widehat{Z}(\mu)$; see Proposition 2.2.

Several properties of the zonoid data depth are immediate from the definition: the depth of $\mathbf{x}$ equals zero if $\mathbf{x}$ lies outside $D_\alpha(\mu)$ for all α; it equals one if $\mathbf{x}$ is the expectation. If $\alpha > 0$, $D_\alpha(\mu)$ is the set of all points that have data depth greater than or equal to α.

We give an equivalent definition of the zonoid data depth which is easily extended to more general situations. For given $\mu \in \mathscr{M}$ and $\mathbf{x} \in \mathbb{R}^d$, define the set

$$\mathscr{A}(\mu, \mathbf{x}) = \{\nu \in \mathscr{M}\colon \nu \text{ is } \mu\text{–continuous and } E(\mu) = \mathbf{x}\}.$$

The subset of distributions ν that have expectation $\mathbf{x}$ and possess a density g with respect to μ, $\nu = g\mu$, is $\mathscr{A}(\mu, \mathbf{x})$.

PROPOSITION 7.1. *Let* $\mu \in \mathscr{M}$ *and* $\mathbf{x} \in \mathbb{R}^d$ *and assume that* $depth_\mu(\mathbf{x}) > 0$. *Then*

$$depth_\mu(\mathbf{x}) = \sup\left\{\frac{1}{\|g\|_\infty}\colon g\mu \in \mathscr{A}(\mu, \mathbf{x})\right\}. \tag{7.2}$$

PROOF. According to Definition 1.1, $x \in D_\alpha(\mu)$ if and only if there exists a measurable function $g\colon \mathbb{R}^d \to \mathbb{R}$ such that

$$\mathbf{x} = \int_{\mathbb{R}^d} \mathbf{y} g(\mathbf{y})\, d\mu(\mathbf{y}), \qquad \int_{\mathbb{R}^d} g(\mathbf{y})\, d\mu(\mathbf{y}) = 1 \quad \text{and} \tag{7.3}$$
$$0 \le g(\mathbf{y}) \le \frac{1}{\alpha} \qquad \text{for all } \mathbf{y}.$$

Equation (7.3) says that g is a density with $\|g\|_\infty \le 1/\alpha$ and $E(g\mu) = \mathbf{x}$. In other words, $x \in D_\alpha(\mu)$ iff there exists some probability measure ν such that

$$\mathbf{x} = E(\nu), \qquad \nu = g\mu \quad \text{and } \|g\|_\infty \le \frac{1}{\alpha}. \tag{7.4}$$

Then $\text{depth}_\mu(\mathbf{x})$ is the supremum of all α that meet (7.4); hence

$$\begin{aligned} \text{depth}_\mu(\mathbf{x}) &= \sup\left\{\alpha\colon \mathbf{x} = E(\nu),\ \nu = g\mu,\ \|g\|_\infty \le \frac{1}{\alpha}\right\} \\ &\le \sup\left\{\frac{1}{\|g\|_\infty}\colon g\mu \in \mathscr{A}(\mu, \mathbf{x})\right\}. \end{aligned}$$

On the other hand, let $\alpha' > \text{depth}_\mu(\mathbf{x})$. Then $x \notin D_{\alpha'}(\mu)$, and it follows that there exists no $g\mu \in \mathscr{A}(\mu, \mathbf{x})$ with $||g||_\infty < 1/\alpha'$, hence (7.2). □

If μ is an empirical distribution on $\mathbf{x}_1, \dots, \mathbf{x}_n$, (7.1) becomes

$$\text{depth}_\mu(\mathbf{x}) = \sup\left\{\alpha\colon \mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}_i,\ \sum_{i=1}^n \lambda_i = 1, \qquad 0 \le n\lambda_i \le \frac{1}{\alpha}\,\forall\, i\right\}. \tag{7.5}$$

In such a case, μ-continuous distributions form the set of discrete distributions with support $\subset \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. We identify them with their probability vectors and write

$$\mathscr{A}(\mu, \mathbf{x}) = \left\{\lambda \in \mathbb{R}^n\colon \mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}_i, \qquad \sum_{i=1}^n \lambda_i = 1,\ \lambda_i \ge 0\,\forall\, i\right\}. \tag{7.6}$$

The alternative formula (7.2) for the zonoid depth then reads

$$\text{depth}_\mu(\mathbf{x}) = \sup\left\{\frac{1}{n \max_{1\le i\le n} \lambda_i}\colon \mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}_i,\ \sum_{i=1}^n \lambda_i = 1,\ \lambda_i \ge 0\ \forall i\right\}.$$

Again let $\mu \in \mathscr{M}$. For any $\mathbf{x} \in \partial \operatorname{conv} \mu$ and $\alpha \ge 0$, we have

$$\mathbf{x} \in \partial D_\alpha(\mu) \quad \Leftrightarrow \quad \alpha \le \text{depth}_\mu(\mathbf{x}). \tag{7.7}$$

Moreover, for any $\mathbf{x} \in \text{int conv}(\mu)$ holds

$$\mathbf{x} \in \partial D_\alpha(\mu) \quad \Leftrightarrow \quad \alpha = \text{depth}_\mu(\mathbf{x}). \tag{7.8}$$

This is due to the definition (7.1) and Theorems 5.3 and 5.5.

The main properties of the zonoid data depth are summarized in the following theorem.

THEOREM 7.1. (i) (Zero at infinity) *For* $\mu \in \mathscr{M}$, $\sup_{\|\mathbf{x}\| \geq M} \textit{depth}_\mu(\mathbf{x}) \to 0$ *as* $M \to \infty$.

(ii) (Continuous on $\mathbf{x}$) *Let* $\mu \in \mathscr{M}$, $\mathbf{x} \in \text{conv}(\mu)$, $\mathbf{x}_n \to \mathbf{x}$. *Then* $\text{depth}_\mu(\mathbf{x}_n) \to \text{depth}_\mu(\mathbf{x})$.

(iii) (Continuous on μ) *Let* $\mu, \mu_n \in \mathscr{M}$ *with* $\mu_n \to_w \mu$ *and* $\mathbf{x} \in \text{int conv}(\mu)$. *Then* $\text{depth}_{\mu_n}(\mathbf{x}) \to \text{depth}_\mu(\mathbf{x})$.

(iv) (Unity only at expectation) *If* $\mu \in \mathscr{M}$ *and* $\mathbf{x} \neq E(\mu)$, *then* $\text{depth}_\mu(\mathbf{x}) < 1 = \text{depth}_\mu(E(\mu))$.

(v) (Monotone on $\mathbf{x}$) *For every* $\mathbf{x} \in \mathbb{R}^d$, $\text{depth}_\mu(c\mathbf{x} + E(\mu))$ *is monotone decreasing on* $c \geq 0$.

(vi) (Affine equivariant) $\text{depth}_{\mu^{A,a}}(\mathbf{x}) = \text{depth}_\mu(A\mathbf{x} + a)$ *if* A *is a regular* $d \times d$ *matrix and* $a \in \mathbb{R}^d$.

(vii) (Monotone on dilation) $\text{depth}_\mu(\mathbf{x}) \leq \text{depth}_\nu(\mathbf{x})$ *if* ν *is a dilation of* μ.

PROOF. (i) Let $(\mathbf{x}_i)$ be an unbounded sequence in $\mathbb{R}^d$. Assume that there exists some α such that $\text{depth}_\mu(\mathbf{x}_i) \geq \alpha > 0$ holds for all i. Then $\mathbf{x}_i \in D_\alpha(\mu)$ for all i, which contradicts the compactness of $D_\alpha(\mu)$. Therefore

$$\sup_i \text{depth}_\mu(\mathbf{x}_i) = 0.$$

(ii) Let $\mu \in \mathscr{M}$, $\mathbf{x}_i \in \text{conv}(\mu)$ for $i \in \mathbb{N}$, and $\mathbf{x}_i \to \mathbf{x}$. Then $(\text{depth}_\mu(\mathbf{x}_i))_{i \in \mathbb{N}}$ is a bounded sequence. In order to prove that $\lim_i \text{depth}_\mu(\mathbf{x}_i) = \text{depth}_\mu(\mathbf{x})$, we shall show that, whenever a subsequence $(\text{depth}_\mu(\mathbf{x}_{i_j}))_{j \in \mathbb{N}}$ converges, then

$$\lim_j \text{depth}_\mu(\mathbf{x}_{i_j}) = \text{depth}_\mu(\mathbf{x}). \tag{7.9}$$

Let $(\text{depth}_\mu(\mathbf{x}_{i_j}))_{j \in \mathbb{N}}$ be a convergent subsequence. Then, as $D_\alpha(\mu)$ is continuous on α (Theorem 5.1), $\lim_j D_{\text{depth}_\mu(\mathbf{x}_{i_j})}(\mu) = D_{\lim_j \text{depth}_\mu(\mathbf{x}_{i_j})}(\mu)$. Since $\mathbf{x}_{i_j} \to \mathbf{x}$ and $\mathbf{x}_{i_j} \in \partial D_{\text{depth}_\mu(\mathbf{x}_{i_j})}(\mu)$ for any j, we get $\mathbf{x} \in \partial D_{\lim_j \text{depth}_\mu(\mathbf{x}_{i_j})}(\mu)$.

If $\mathbf{x} \in \text{int conv}(\mu)$, then (7.9) follows from (7.8). If $\mathbf{x} \in \partial \text{conv}(\mu)$, we conclude from (7.7) that $\lim_j \text{depth}_\mu(\mathbf{x}_{i_j}) \leq \text{depth}_\mu(\mathbf{x})$. Assume that $\lim_j \text{depth}_\mu(\mathbf{x}_{i_j}) < \text{depth}_\mu(\mathbf{x})$ and let $\gamma = (\lim_j \text{depth}_\mu(\mathbf{x}_{i_j}) + \text{depth}_\mu(\mathbf{x}))/2$. Then $D_\gamma(\mu) \supset D_{\text{depth}_\mu(\mathbf{x})}(\mu)$ and, for all j that are larger than some j_0, $\text{depth}_\mu(\mathbf{x}_{i_j}) < \gamma$ holds, hence $\mathbf{x}_{i_j} \notin D_\gamma(\mu)$ and therefore $\mathbf{x}_{i_j} \notin D_{\text{depth}_\mu(\mathbf{x})}(\mu)$. That contradicts the convergence $\mathbf{x}_{i_j} \to \mathbf{x}$.

(iii) According to Theorem 5.2, we have

$$\lim_k D_{\text{depth}_{\mu_n}(\mathbf{x})}(\mu_k) \to_H D_{\text{depth}_{\mu_n}(\mathbf{x})}(\mu)$$

for every $n \in \mathbb{N}$. The sequence $(\text{depth}_{\mu_n(\mathbf{x})}(\mu_n))$ is bounded. Consider a convergent subsequence $(\text{depth}_{\mu_{n_j}(\mathbf{x})}(\mu_{n_j}))$. Then

$$\lim_j D_{\text{depth}_{\mu_n}(\mathbf{x})}(\mu_{n_j}) \to_H D_{\text{depth}_{\mu_n}(\mathbf{x})}(\mu) \quad \text{for all } n \in \mathbb{N}. \tag{7.10}$$

We select $n = n_j$ in (7.10) and go to the limit $j \to \infty$ on both sides. This yields $\lim_j D_{\text{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu_{n_j}) = \lim_j D_{\text{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu)$. Theorem 5.1 ensures that

$\lim_j D_{\mathrm{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu) = D_{\lim_j \mathrm{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu)$, hence

$$\lim_j D_{\mathrm{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu_{n_j}) = D_{\lim_j \mathrm{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu).$$

As $\mathbf{x} \in \operatorname{int}\operatorname{conv}(\mu)$, due to (7.8) we have $\mathbf{x} \in \partial D_{\mathrm{depth}_{\mu_{n_j}}(\mathbf{x})}(\mu_{n_j})$. Therefore, $\mathbf{x} \in \partial D_{\lim_j \mathrm{depth}_{\mu_{n_j}}}(\mathbf{x})$, and, again with (7.8), $\lim_j \mathrm{depth}_{\mu_{n_j}}(\mathbf{x}) = \mathrm{depth}_{\mu}(\mathbf{x})$. As the last equality holds for every convergent subsequence, it remains true for the original sequence.

(iv) $D_1(\mu) = \mathrm{E}(\mu)$.

(v) For any μ and α, $\mathrm{E}(\mu) \in D_\alpha(\mu)$ holds. The monotonicity on $c \geq 0$ follows from this fact and from the monotonicity of $D_\alpha(\mu)$ on α (Theorem 5.3).

(vi) Follows from the affine equivariance of $D_\alpha(\mu)$ (Theorem 5.7).

(vii) Assume that ν is a dilation of μ. Dilation implies inclusion of the lift zonoids; see Koshevoy and Mosler (1997), Theorem 5.2. Therefore $D_\alpha(\mu) \subset D_\alpha(\nu)$ for every $\alpha \in [0, 1]$ or, equivalently, $\mathrm{depth}_\mu(\mathbf{x}) \leq \mathrm{depth}_\nu(\mathbf{x})$ for every $x \in \mathbb{R}^d$. $\square$

Our notion differs from other depth notions: Mahalanobis's depth, Tukey's depth [Tukey (1975)], simplicial depth [Liu (1990)], majority depth [Singh (1991)]. However, there are some connections.

The Mahalanobis depth of $\mu \in \mathscr{M}$ is given by

$$d_\mu^{Ma}(\mathbf{x}) = \left(1 + (\mathbf{x} - \mathrm{E}(\mu))'\Sigma_\mu^{-1}(\mathbf{x} - E(\mu))\right)^{-1},$$

provided μ has a positive definite covariance matrix Σ_μ. It follows from (6.2) that the Mahalanobis depth of a given μ is a strictly increasing transform of the zonoid depth of the normal distribution $N(\mathrm{E}(\mu), \Sigma_\mu)$.

In the case of an empirical distribution, we have the following relation with the simplicial depth. In (7.6) $\mathscr{A}(\mu, \mathbf{x})$ is a convex polytope. The simplicial depth of the point $\mathbf{x}$ equals a properly counted number of vertices of this polytope.

Theorem 7.1 shows that our notion has many properties in general which other notions have under some restrictions only, see, for example, Liu and Singh (1993) for properties of Tukey's, simplicial and majority depths. As will be clear from the next section, our notion of data depth equals twice Tukey's data depth of a properly transformed distribution.

8. A trimming transform. In this section we shall indicate the dimension d by a superscript. Let $\mathscr{M}_*^d \subset \mathscr{M}^d$ be the subset of distributions that have a *unique median*. That is, for every $\mu \in \mathscr{M}_*^d$, there is some $x(\mu) \in \mathbb{R}^d$, the median, such that any hyperplane passing through $x(\mu)$ divides $\mathbb{R}^d$ in two halfspaces with equal μ-masses.

We shall construct a trimming transform that is an injection $\Psi_d\colon \mathscr{M}^d \to \mathscr{M}_*^d$. Let us start with the univariate case, $\mu \in \mathscr{M}^1$. We define $\Psi_1(\mu)$ via its inverse function.

$$(\Psi_1(\mu))^{-1}(\alpha) = \begin{cases} \dfrac{1}{2\alpha}\displaystyle\int_0^{2\alpha} Q(s)\,ds, & \alpha \in \left(0, \dfrac{1}{2}\right], \\ \dfrac{1}{2(1-\alpha)}\displaystyle\int_{2\alpha-1}^{1} Q(s)\,ds, & \alpha \in \left(\dfrac{1}{2}, 1\right]. \end{cases} \tag{8.1}$$

Obviously $(\Psi_1(\mu))^{-1}$ is a monotone and continuous function. Therefore it may be considered as the quantile of a distribution function $\Psi_1(\mu)$. As this quantile function is continuous, $\mathrm{E}(\mu)$ is the unique median of $\Psi_1(\mu)$; hence, $\Psi_1(\mu) \in \mathscr{M}_*^1$. Observe that the convex hull of the support of μ and that of $\Psi_1(\mu)$ are the same: $\mathrm{conv}(\mu) = \mathrm{conv}(\Psi_1(\mu))$. So, we see that the zonoid α-trimmed region of $\mu \in \mathscr{M}^1$ equals the quantile $(1-(\alpha/2))$-trimmed region of $\Psi_1(\mu)$. Averous and Meste (1990) have constructed a similar relation for expectiles.

Our aim is to extend this relation between the lift zonoid trimming and the quantile trimming to higher dimensions.

To introduce quantile trimmed regions in many dimensions, let $\mu \in \mathscr{M}^d$ and $\mathbf{p} \in S^{d-1}$. Eddy (1984) defines the quantile of μ in the direction $\mathbf{p}$ as the halfspace

$$H_\alpha(\mathbf{p}, \mu) = \{\mathbf{x} \in \mathbb{R}^d\colon \langle \mathbf{x}, \mathbf{p}\rangle \le Q_{\mathbf{p}}(\alpha)\}, \tag{8.2}$$

where $\mu_{\mathbf{p}}(t) = \mu(H(\mathbf{p}, t))$ and $Q_{\mathbf{p}}(\alpha) = \inf\{t\colon \mu_{\mathbf{p}}(t) \ge \alpha\}$ is the quantile function of $\mu_{\mathbf{p}}$. Therefore we rewrite (8.2) as follows.

$$H_\alpha(\mathbf{p}, \mu) = \bigcap\left\{H(\mathbf{p}, t)\colon \int_{H(\mathbf{p},t)} d\mu \ge \alpha\right\}.$$

The following set is named the quantile α-trimmed region of a distribution $\mu \in \mathscr{M}^d$.

$$D_\alpha^{qu}(\mu) = \bigcap_{p \in S^{d-1}} H_\alpha(\mathbf{p}, \mu), \qquad \alpha \in [\tfrac{1}{2}, 1]. \tag{8.3}$$

Obviously, $D_\alpha^{qu}(\mu) = \bigcap\left\{H \in \mathscr{H}^d\colon \mu(H) \ge \alpha\right\}$. Trimmed regions of this form have been considered by Massé and Theodorescu (1994).

Define a map $\Psi_d\colon \mathscr{M}^d \to \mathscr{M}_*^d$ by the following: for every $\alpha \in [0, 1]$, the quantile $(1-(\alpha/2))$-trimmed region of $\Psi_d(\mu)$ is equal to the zonoid α-trimmed region of $\mu \in \mathscr{M}^d$. The main result of this section states that such a map exists for distributions having compact support.

THEOREM 8.1. *An injective map $\Psi_d\colon \mathscr{M}_c^d \to \mathscr{M}_*^d$ exists, such that for every $\mu \in \mathscr{M}_c^d$ and every $\alpha \in [0, 1]$ there holds: the zonoid α-trimmed region of μ equals the quantile $(1-\frac{\alpha}{2})$-trimmed region of $\Psi_d(\mu)$.*

PROOF. Consider some $\mathbf{p} \in S^{d-1}$. An extreme point of $D_\alpha(\mu)$ in direction $\mathbf{p} \in S^{d-1}$ has the form $\alpha^{-1}\int_{H(-\mathbf{p}, p_0)} \mathbf{x}\, d\mu(\mathbf{x})$, where p_0 is such that

$\mu(H(-\mathbf{p}, p_0)) = \alpha$. See Koshevoy and Mosler (1997). By this, the support function of $D_\alpha(\mu)$ is of the form

$$\phi_{D_\alpha(\mu)}(\mathbf{p}) = \frac{1}{\alpha}\int_{H(-\mathbf{p},\, p_0)} \langle \mathbf{x}, \mathbf{p}\rangle \, d\mu(\mathbf{x}). \tag{8.4}$$

In view of (8.4), we have

$$D_\alpha(\mu) = \bigcap_{p\in S^{d-1}} \left\{\mathbf{x}\colon \langle \mathbf{x}, \mathbf{p}\rangle \le \frac{1}{\alpha}\int_{1-\alpha}^{1} Q_{\mathbf{p}}(s)\, ds\right\}. \tag{8.5}$$

It is easy to check that $\int_{1-\alpha}^{1} Q_{-\mathbf{p}}(s)\, ds = -\int_0^\alpha Q_{\mathbf{p}}(s)\, ds$. Therefore we get

$$\left\{\mathbf{x}\colon \langle \mathbf{x}, -\mathbf{p}\rangle \le \frac{1}{\alpha}\int_{1-\alpha}^{1} Q_{-\mathbf{p}}(s)\, ds\right\} = \left\{\mathbf{x}\colon \langle \mathbf{x}, \mathbf{p}\rangle \ge \frac{1}{\alpha}\int_{0}^{\alpha} Q_{\mathbf{p}}(s)\, ds\right\}. \tag{8.6}$$

In view of (8.6), we rewrite (8.5) as follows:

$$D_\alpha(\mu) = \bigcap_{\mathbf{p}\in \mathbb{P}^{d-1}} \left\{\mathbf{x}\colon \frac{1}{\alpha}\int_{0}^{\alpha} Q_{\mathbf{p}}(s)\, ds \le \langle \mathbf{x}, \mathbf{p}\rangle \le \frac{1}{\alpha}\int_{1-\alpha}^{1} Q_{\mathbf{p}}(s)\, ds\right\}, \tag{8.7}$$

where $\mathbb{P}^{d-1} = S^{d-1}/\underline{+}1$ is the factor space of S^{d-1} with respect to reflection. So, to prove the existence of Ψ_d, we have to demonstrate that there exists a random vector $\widetilde{X}$ such that, for every $\mathbf{p} \in \mathbb{P}^{d-1}$, $\Psi_1(\mu_{\mathbf{p}})$ is the distribution of $\langle \widetilde{X}, \mathbf{p}\rangle$.

Let us start with the case when μ has an infinitely differentiable distribution function. Then the existence of Ψ_d is tantamount to the existence of an inverse to the Radon transform. According to the Paley–Wiener theorem [Helgason (1980)], we employ the following conditions:

(i) The distribution function $\Psi_1(\mu_{\mathbf{p}})(t)$ is infinitely differentiable by $\mathbf{p}$ and t and has a compact support.
(ii) The integral $\int_{-\infty}^{\infty} t^k\, d\Psi_1(\mu_{\mathbf{p}})(t)$ is a homogeneous polynomial of degree k on $\mathbf{p}_1, \dots, \mathbf{p}_d$, $k = 0, 1, \dots$.

Here and below, the notation Ψ_d is used for the transform of the respective distribution functions as well.

As the distribution function of μ is infinitely differentiable, $Q_{\mathbf{p}}(t)$ is infinitely differentiable by $\mathbf{p}$ and t, hence $\alpha^{-1}\int_{1-\alpha}^{1} Q_{\mathbf{p}}(s)\, ds$ is infinitely differentiable by $\mathbf{p}$ and α. Moreover, as $conv(\mu)$ is compact, $\mathrm{conv}(\mu_{\mathbf{p}}) = \mathrm{conv}(\Psi_1(\mu_{\mathbf{p}}))$. So, Condition 1 is met.

Recall [e.g., Schweizer and Sklar (1983)] that if F is a d-dimensional distribution function with one-dimensional marginals $F_1, \dots, F_d$ then there exists a d-dimensional copula C such that $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$. Observe that $F_i(x_i) = F_{\mathbf{e}_i}(\langle \mathbf{e}_i, \mathbf{x}\rangle)$, where $\mathbf{e}_i$ is the ith canonical base vector, $i = 1, \dots, d$.

LEMMA 8.1. *Let F be the distribution function of some $\mu \in \mathscr{M}^d$, $A\colon \mathbb{R}^d \to \mathbb{R}^d$ be a regular matrix, $F^A(\mathbf{x}) = F(A^{-1}\mathbf{x})$, C and C^A be the copulas of F and F^A, respectively. Then*

$$C^A(F_{A'\mathbf{e}_1}(\langle A'\mathbf{e}_1, \mathbf{x}\rangle), \dots, F_{A'\mathbf{e}_d}(\langle A'\mathbf{e}_d, \mathbf{x}\rangle) = C(F_1(x_1), \dots, F_d(x_d)). \tag{8.8}$$

PROOF.

$$\begin{aligned} F_i^A(x_i) = F_{\mathbf{e}_i}^A(\langle \mathbf{e}_i, \mathbf{x}\rangle) &= \int_{\{\mathbf{y}:\ \langle \mathbf{e}_i, \mathbf{y}\rangle \le \langle \mathbf{e}_i, \mathbf{x}\rangle\}} dF^A(\mathbf{y}) \\ &= \int_{\{\mathbf{y}:\ \langle \mathbf{e}_i, \mathbf{y}\rangle \le \langle \mathbf{e}_i, \mathbf{x}\rangle\}} dF(A^{-1}\mathbf{y}) = \int_{\{\mathbf{z}:\ \langle \mathbf{e}_i, A\mathbf{z}\rangle \le \langle \mathbf{e}_i, \mathbf{x}\rangle\}} dF(\mathbf{z}) \\ &= \int_{\{\mathbf{z}:\ \langle A'\mathbf{e}_i, \mathbf{z}\rangle \le x_i\}} dF(\mathbf{z}) = F_{A'\mathbf{e}_i}(x_i). \end{aligned}$$

Therefore,

$$\begin{aligned} C(F_1(x_1), \dots, F_d(x_d)) &= F(x) = F^A(A\mathbf{x}) \\ &= C^A(F_{A'\mathbf{e}_1}(\langle A'\mathbf{e}_1, \mathbf{x}\rangle), \dots, F_{A'\mathbf{e}_d}(\langle A'\mathbf{e}_d, \mathbf{x}\rangle)). \qquad \square \end{aligned}$$

Fix some $\mathbf{p} \in S^{d-1}$ and let A be a regular matrix with $A'\mathbf{e}_1 = p$. Then there exists a copula C such that $C(\Psi_1(\mu_{\mathbf{p}})(\langle A'\mathbf{e}_1, \mathbf{x}\rangle), \dots, \Psi_1(\mu_{A'\mathbf{e}_d})(\langle A'\mathbf{e}_d, \mathbf{x}\rangle))$ is a distribution function, e.g., the product copula. Let $G(x_1, \dots, x_d)$ denote this distribution function.

Plugging this C into the right-hand side of (8.8) yields $G_{\mathbf{p}} = G_{A'\mathbf{e}_1} = \Psi_1(\mu_{\mathbf{p}})$. Therefore, $\int_{-\infty}^{\infty} t^k \, dG_{\mathbf{p}}(t) = \int_{\mathbb{R}^d} \langle \mathbf{p}, \mathbf{x}\rangle^k \, dG(\mathbf{x})$ is a homogeneous polynomial, and $\int_{-\infty}^{\infty} t^k d\Psi_1(\mu_{\mathbf{p}})(t)$ is a homogeneous polynomial as well. That yields Condition 2. For a general $\mu \in \mathscr{M}_c^d$, a limit argument completes the proof. $\square$

Acknowledgments. We thank Rainer Dyckerhoff, Alexander Shananin, Radu Theodorescu, the editors and a referee for their valuable remarks and comments.

REFERENCES

ABDOUS, B. and REMILLARD, B. (1995). Relating quantiles and expectiles under weighted symmetry. *Ann. Inst. Statist. Math.* **47** 371–384.

AVEROUS, J. and MESTE, M. (1990). Location, skewness and tailweight in $\mathbb{L}_S$ sense: a coherent approach. *Statistics* **21** 57–74.

BOLKER, E. D. (1969). A class of convex bodies. *Trans. Amer. Math. Soc.* **145** 323–346.

BRECKLING, J. and CHAMBERS, R. (1988). M-quantiles. *Biometrika* **75** 761–771.

EDDY, W. F. (1984). Set-valued ordering for bivariate data. *Stochastic Geometry, Geometric Statistics, Stereology* (R. Ambartzumian and W. Weil, eds.) 79–90. Teubner, Leipzig.

GOODEY, P. and WEIL, W. (1993). Zonoids and generalizations. *Handbook of Convex Geometry* B (P. M. Gruber and J. M. Wills, eds.) 1297–1326. North-Holland, Amsterdam.

HELGASON, S. (1980). *The Radon Transform*. Birkhäuser, Boston.

KOSHEVOY, G. A. and MOSLER, K. (1997). Lift zonoids, random convex hulls and the variability of random vectors. *Bernoulli* **3**. To appear.

LIU, R. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414.

LIU, R. and SINGH, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* **88** 252–260.

MASSÉ, J.-C. and THEODORESCU, R. (1994). Halfplane trimming for bivariate distributions. *J. Multivariate Anal.* **48** 188–202.

NEWEY, W. K. and POWELL, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847.

NOLAN, D. (1992). Asymptotics for multivariate trimming. *Stochastic Process. Appl.* **42** 157–169.

SCHWEIZER, B. and SKLAR, A. (1983). *Probabilistic Metric Spaces.* North-Holland, Amsterdam.

SHEPHARD, G. C. (1974). Combinatorial properties of associated zonotopes. *Canad. J. Math.* **26** 302–321.

SINGH, K. (1991). A notion of majority depth. Technical report, Dept. Statistics, Rutgers Univ.

TUKEY, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians* (*Vancouver, B.C., 1974*) **2** 523–531. Canad. Math. Congress, Montreal.

WEIL, W. and WIEACKER, J. A. (1993). Stochastic geometry. In *Handbook of Convex Geometry* (P. M. Gruber and J. M. Wills, eds.) 1391–1438. North-Holland, Amsterdam.

CENTRAL INSTITUTE OF MATHEMATICS
AND ECONOMICS
RUSSIAN ACADEMY OF SCIENCE
KRASIKOVA 32
MOSCOW 117418
RUSSIA
E-MAIL: gleb.koshevoy@uni-koeln.de

SEMINAR FÜR WIRTSCHAFTS-
UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN
D–50923 KÖLN
GERMANY
E-MAIL: mosler@wiso.uni-koeln.de