The Annals of Statistics
2000, Vol. 28, No. 2, 483–499

STRUCTURAL PROPERTIES AND CONVERGENCE RESULTS FOR CONTOURS OF SAMPLE STATISTICAL DEPTH FUNCTIONS

BY YIJUN ZUO AND ROBERT SERFLING[1]

Arizona State University and University of Texas

Statistical depth functions have become increasingly used in nonparametric inference for multivariate data. Here the contours of such functions are studied. Structural properties of the regions enclosed by contours, such as affine equivariance, nestedness, connectedness and compactness, and almost sure convergence results for sample depth contours, are established. Also, specialized results are established for some popular depth functions, including halfspace depth, and for the case of elliptical distributions. Finally, some needed foundational results on almost sure convergence of sample depth functions are provided.

1. Introduction. Statistical depth functions have become increasingly used in nonparametric inference for multivariate data. Given a distribution F on $\mathbf{R}^d$, a corresponding depth function is any function $D(x; F)$ which provides an F-based center-outward ordering of points $x \in \mathbf{R}^d$. For broad treatments of depth functions, see Liu, Parelius and Singh (1999) and Zuo and Serfling (2000b).

The present paper focuses on depth function *contours*. For $\alpha > 0$, the boundary of the set $\{x \in \mathbf{R}^d : D(x; F) \geq \alpha\}$ denotes a corresponding *α-depth contour*. Besides exhibiting the structure of underlying multivariate distributions and revealing the shape of multivariate datasets, contours play a natural role in generalizing univariate L-statistics and R-statistics to the multivariate setting. It is of fundamental interest, therefore, to establish the properties and convergence behavior of contours for *general* depth functions $D(\cdot; F)$ and allowing *arbitrary* distributions F. Results to date, however, have been somewhat limited. Structural properties and convergence of sample depth contours for the *halfspace* depth function (defined in Section 2) have been studied by Eddy (1985), Nolan (1992), Donoho and Gasko (1992) and Massé and Theodorescu (1994). Under some broad assumptions on the depth function, but primarily for *elliptical* distributions F, He and Wang (1997) also provide a result on convergence.

In Section 3, under various general conditions on depth functions, but not restricting underlying distributions, we establish *structural properties* for the regions enclosed by contours, such as affine equivariance, nestedness, connectedness and compactness. We also treat specifically some particular depth

Received December 1998; revised December 1999.

[1]Supported by NSF Grant DMS-97-05209.

AMS 1991 *subject classifications*. Primary 62H12; secondary 62G20.

Key words and phrases. Statistical depth functions, depth contours, sample depth contours, almost sure convergence, halfspace depth, elliptical distributions.

functions. The results cover sample depth functions as well. (See Theorems 3.1–3.2.) For the case of *elliptical* distributions, the contours exhibit ellipsoidal structure and satisfy some related properties. (See Theorems 3.3–3.5.)

In Section 4, under broad conditions on the underlying depth function and assuming convergence of the associated sample depth function, we establish almost sure convergence results for sample depth contours (Theorem 4.1). For the case of *elliptical* underlying distributions, and for some particular depth functions, we establish a.s. convergence of sample contours to ellipsoidal shapes (Theorem 4.2).

The conditions on depth functions required in these results are based on four properties ideally satisfied by any statistical depth function. These are listed in Section 2, along with examples of particular depth functions about which results are proved in Sections 3 and 4.

Proofs are deferred to Appendix A. Results on convergence of sample depth functions are treated briefly in Appendix B.

2. Key properties of statistical depth functions. For a depth function $D(\cdot;\cdot)$ to serve most effectively as a tool providing an F-based center-outward ordering of points in $\mathbf{R}^d$, it is argued in Zuo and Serfling (2000b) that the following properties are favorable. [We note that these properties were first introduced by Liu (1990) in studying the simplicial depth function.] Denote the class of distributions on $\mathbf{R}^d$ by $\mathscr{F}$ and for any random vector X its cdf by F_X.

P1. *Affine Invariance.* $D(Ax+b; F_{AX+b}) = D(x; F_X)$ for any random vector X in $\mathbf{R}^d$, any $d \times d$ nonsingular matrix A and any d-vector b.

P2. *Maximality at Center.* For any $F \in \mathscr{F}$ having "center" θ (e.g., the point of symmetry relative to some notion of symmetry), $D(\theta; F) = \sup_{x\in\mathbf{R}^d} D(x; F)$.

P3. *Monotonicity Relative to Deepest Point.* For any $F \in \mathscr{F}$ having deepest point θ (i.e., point of maximal depth), $D(x; F) \le D(\theta + \alpha(x-\theta); F)$, $\alpha \in [0, 1]$.

P4. *Vanishing at Infinity.* $D(x; F) \to 0$ as $\|x\| \to \infty$, for each $F \in \mathscr{F}$.

With respect to property P2, various notions of multivariate symmetry may be considered. Here we mention two. A random vector X in $\mathbf{R}^d$ (or its distribution P) is *centrally symmetric* about θ if $X - \theta \stackrel{d}{=} \theta - X$, where "$\stackrel{d}{=}$" denotes "equal in distribution" or, equivalently, if $P(S) = P(S^{(\theta)})$ for any Borel set S and its reflection $S^{(\theta)}$ about θ. This is a classical nonparametric notion of multivariate symmetry. More generally, as introduced and discussed in Zuo and Serfling (2000c, 2000b), we define X to be *halfspace symmetric* about θ if $P(X \in H) \ge 1/2$ for every closed halfspace H containing θ. Since C-symmetry $\Rightarrow$ H-symmetry, the preferred manifestation of property P2 is that maximality at center should hold for $D(\cdot; P)$ for every H-symmetric P. A similar remark holds regarding property P3.

In Zuo and Serfling (2000b), general types and particular cases of depth function are examined in detail relative to properties P1–P4. A number of these are recalled below for present consideration. Denote by ∂C, C^c, C° and $\overline{C}$, respectively, the *boundary*, *complement*, *interior* and *closure* of a set C.

EXAMPLE 2.1. *Type D depth functions* are of the form

$$D(x; P, \mathscr{C}) \equiv \inf_C \{ P(C) \mid x \in C \in \mathscr{C} \},$$

where $\mathscr{C}$ is a given class of *closed* subsets of $\mathbf{R}^d$ satisfying:

D1. If $C \in \mathscr{C}$, then $\overline{C^c} \in \mathscr{C}$.

D2. If $C \in \mathscr{C}$ and $x \in C^\circ$, then there exists $C_1 \in \mathscr{C}$ with $C_1 \subset C^\circ$ and $x \in \partial C_1$.

Thus the $\mathscr{C}$-depth of a point x in $\mathbf{R}^d$ with respect to a probability measure P on $\mathbf{R}^d$ is defined to be the minimum probability mass carried by a set C in $\mathscr{C}$ that contains x.

For $\mathscr{C}$ the class $\mathscr{H}$ of *closed halfspaces* in $\mathbf{R}^d$, we obtain the leading and most studied depth function, the *halfspace depth* introduced by Tukey (1975):

$$HD(x; P) = \inf\{P(H) : H \text{ a closed halfspace, } x \in H\}, \ \ x \in \mathbf{R}^d,$$

for which all of properties P1–P4 hold, including P2 with respect to H-symmetry. □

The further examples that we consider are of other types.

EXAMPLE 2.2 (Simplicial depth [Liu (1990)]).

$$SD(x; P) = P(x \in S[X_1, \dots, X_{d+1}]), \quad x \in \mathbf{R}^d, \tag{1}$$

where $r = d + 1$, $h(x; x_1, \dots, x_{d+1}) = \mathbf{I}\{ x \in S[x_1, \dots, x_{d+1}]\}$, $X_1, \dots, X_{d+1}$ is a random sample from P, and $S[x_1, \dots, x_{d+1}]$ denotes the d-dimensional simplex with vertices $x_1, \dots, x_{d+1}$, i.e, the set of convex combinations of $x_1, \dots, x_{d+1}$. Here P1 and P4 hold in general but P2 and P3 fail in some discrete cases. □

EXAMPLE 2.3 (Majority depth [Singh (1991)]). For given points $x_1, \dots, x_d$ in $\mathbf{R}^d$ which determine a unique hyperplane containing themselves, denote by $H^P_{x_1,\dots,x_d}$ the halfspace with this hyperplane as boundary which carries probability mass $\geq 1/2$ under the distribution P on $\mathbf{R}^d$, and define

$$MJD(x; P) = P\left(x \in H^P_{X_1,\dots,X_d}\right), \quad x \in \mathbf{R}^d, \tag{2}$$

where $X_1, \dots, X_d$ is a random sample from P. Properties P1–P3 hold, including P2 with respect to H-symmetry, but P4 fails. □

EXAMPLE 2.4 (Simplicial volume depth).

$$SVD^\alpha(x; F) \equiv \left(1 + E\left[\left(\frac{\Delta(S[x, X_1, \dots, X_d])}{\sqrt{\det(\Sigma)}}\right)^\alpha\right]\right)^{-1}, \quad x \in \mathbf{R}^d, \tag{3}$$

where $\Delta(S[x; x_1, \dots, x_d])$ denotes the volume of the d-dimensional simplex $S[x, x_1, \dots, x_d]$, $\alpha > 0$, and Σ is the covariance matrix of F. All of P1–P4 hold

(P2 for C-symmetry and $p \geq 1$, P3 for $p \geq 1$). [Like related versions in the literature which, however, are not affine invariant, it is inspired by the notion of multidimensional median given by Oja (1983).] □

EXAMPLE 2.5 (L^p depth).

$$L^pD(x;F) \equiv \left(1 + E\|x - X\|_p\right)^{-1}, \quad x \in \mathbf{R}^d, \tag{4}$$

where $\|\cdot\|_p$ is the usual L^p norm and $p > 0$. While not fully affine invariant, this depth function satisfies P2 for C-symmetry and $p \geq 1$, P3 for $p \geq 1$, and P4 in general.

Although for $p = 2$ this depth function is at least rigid-body invariant, a modification in this case yields a fully affine invariant version:

$$\widetilde{L}^2D(x;F) \equiv (1 + E\,[\,\|x - X\|_{\Sigma^{-1}}\,])^{-1}, \tag{5}$$

where Σ is the covariance matrix of F and $\|x\|_M \equiv \sqrt{x'Mx}$, $x \in \mathbf{R}^d$, a norm introduced by Rao (1988) for positive definite $d \times d$ matrix M. □

EXAMPLE 2.6 (Projection depth). Consider the "outlyingness measure"

$$O(x;F) \equiv \sup_{\|u\|=1} \frac{|\,u'x - \mathrm{Med}(u'X)\,|}{\mathrm{MAD}(u'X)}, \quad x \in \mathbf{R}^d, \tag{6}$$

where F is the distribution of X, Med denotes the univariate median, and MAD denotes the univariate median absolute deviation $\mathrm{MAD}(Y) = \mathrm{Med}(|Y - \mathrm{Med}(Y)|)$. Then define

$$PD(x;F) \equiv (1 + O(x;F))^{-1}, \quad x \in \mathbf{R}^d. \tag{7}$$

This satisfies all of P1–P4, including P2 with respect to H-symmetry. □

EXAMPLE 2.7 (Mahalanobis depth).

$$MHD(x;F) = \Big(1 + d^2_{\Sigma(F)}(x, \mu(F))\Big)^{-1}, \quad x \in \mathbf{R}^d, \tag{8}$$

where $\mu(F)$ and $\Sigma(F)$ are location and covariance measures, respectively, defined on distributions F, and $d^2_M(x, y) = (x-y)'M^{-1}(x-y)$ is the *Mahalanobis distance* [Mahalanobis (1936)] between two points x and y in $\mathbf{R}^d$ with respect to a positive definite $d \times d$ matrix M. This depth function can satisfy all of P1–P4. In particular, P2 is satisfied with respect to any symmetry notion if μ and Σ are affine equivariant and $\mu(F)$ agrees with the point of symmetry. The case of $\mu(F)$ and $\Sigma(F)$ the mean vector and covariance matrix of F, respectively, is considered in Liu (1992) and Liu and Singh (1993). □

For further details on the foregoing depth functions and other varieties as well, see Liu, Parelius and Singh (1999), Zuo and Serfling (2000b) and references cited therein.

For any F-based depth function, we also consider its *sample* version. With $\widehat{F}_n$ the usual empirical distribution placing mass $1/n$ on each observation

$X_1, \dots, X_n$ from some distribution $F \in \mathscr{F}$, a sample version of $D(\cdot; F)$ is given by $D_n(\cdot) = D(\cdot; \widehat{F}_n)$. Of course, *ad hoc* versions of $D_n(x)$ can also be formulated.

3. Structural properties of depth contours. We mentioned briefly in Section 1 the depth contour concept.

DEFINITION 3.1. For a given depth function $D(\cdot; F)$ and for $\alpha > 0$, we call

$$D^\alpha(F) \equiv \{x \in \mathbf{R}^d \mid D(x; F) \geq \alpha\}$$

the corresponding *α-trimmed region* and its boundary $\partial D^\alpha(F)$ the corresponding *α-contour*. For a *sample* depth function $D_n(\cdot)$ we use the notation D_n^α and ∂D_n^α for the corresponding *sample α-trimmed region* and *sample α-contour*.

It is convenient to treat depth contours in terms of their corresponding trimmed regions, which usually are desired to be *affine equivariant, nested, connected* and *compact*. [The region $D^\alpha(F_X)$ is *affine equivariant* if $D^\alpha(F_{AX+b}) = AD^\alpha(F_X) + b$ holds for any random vector X in $\mathbf{R}^d$, any $d \times d$ nonsingular matrix A and any d-vector b. A set E in a topological space X is said to be *connected* if E is not the union of two nonempty sets A and B such that $\overline{A} \cap B = \varnothing = A \cap \overline{B}$.]

THEOREM 3.1. *The depth-trimmed regions satisfy the following properties*:

(a) $D^\alpha(\cdot)$ *and* D_n^α *are affine equivariant if* $D(\cdot; \cdot)$ *and* $D_n(\cdot)$, *respectively, satisfy* P1.

(b) $D^\alpha(\cdot)$ *and* D_n^α *are nested*: *for* $\alpha_1 \geq \alpha_2$, $D^{\alpha_1}(F) \subset D^{\alpha_2}(F)$, *any* F, *and* $D_n^{\alpha_1} \subset D_n^{\alpha_2}$.

(c) $D^\alpha(\cdot)$ *is connected if* $D(\cdot; \cdot)$ *satisfies* P3.

(d) $D^\alpha(\cdot)$ *is compact for the depth functions* $SVD^\beta(x; F)$ $(\beta \geq 1)$, $L^p D(x; F)$ $(p \geq 1)$, $PD(x; F)$ *and* $MHD(x; F)$, *and, in the case of continuous* F, *for* $SD(x; F)$ *and all Type D depth functions* [*including* $HD(x; F)$].

REMARKS 3.1. (i) For the *halfspace depth*, for example, P1 is straightforward to prove and thus by (a) of Theorem 3.1 it follows that the corresponding depth-trimmed regions are affine equivariant. This result is also given by Massé and Theodorescu (1994).

(ii) Connectedness and compactness do not hold in general for depth-trimmed regions. For example, those corresponding to the simplicial depth fail to be connected in certain cases of discrete F for which P3 fails to hold [see examples in Remark 2.1 of Zuo and Serfling (2000b)]. Those corresponding to the majority depth are not compact, due to lack of P4 [see examples following Theorem 2.2 of Zuo and Serfling (2000b)], although they can be shown to be closed.

Connectedness and compactness of *sample* depth-trimmed regions can be established for many cases. In particular, for Type D depth we have the following:

THEOREM 3.2. *Let $\mathscr{C}$ be a class of closed and connected Borel sets satisfying* D1 *and* D2 *of Example* 2.1. *Further, assume*:

D3. *If $x \in C \in \mathscr{C}$ and $\widehat{F}_n(C) < \alpha$, then there exists $C_1 \in \mathscr{C}$ with $x \in C_1^\circ$ and $\widehat{F}_n(C_1) < \alpha$, for $\alpha = k/n$, $1 \le k \le n$.*

Then, for the sample depth function $D(\cdot; \widehat{F}_n, \mathscr{C})$, the corresponding sample depth-trimmed regions D_n^α are connected and compact.

For $\mathscr{C} = \mathscr{H}$, we slightly extend Lemma 2.2 of Donoho and Gasko (1992) in the following result: *for halfspace depth, the sample depth-trimmed regions are nested, connected, convex and compact.*

For special distributions such as elliptical, depth contours possess exactly the same shape as the constant density contours. A random vector X in $\mathbf{R}^d$ is said to be *elliptically distributed*, denoted by $X \sim E_d(h; \mu, \Sigma)$, if its density is of the form

$$f(x) = c|\Sigma|^{-1/2} h((x-\mu)'\Sigma^{-1}(x-\mu)).$$

The following result generalizes Lemma 3.1 of Liu and Singh (1993).

THEOREM 3.3. *Suppose that $X \sim E_d(h; \mu, \Sigma)$ and that $D(\cdot; \cdot)$ is affine invariant* (*property* P1) *and attains maximum value at μ* (*property* P2). *Then*:

(i) *$D(x; F_X)$ is of form*

$$D(x; F_X) = g((x-\mu)'\Sigma^{-1}(x-\mu)) \tag{9}$$

for some nonincreasing function g, and D^α is of form

$$D^\alpha = \{x \in \mathbf{R}^d \mid (x-\mu)'\Sigma^{-1}(x-\mu) \le r_\alpha^2\} \tag{10}$$

for some r_α.

(ii) *$D(x; F_X)$ is strictly decreasing on any ray originating from the center if and only if*

$$\{x \in \mathbf{R}^d \mid D(x; F_X) = \alpha\} = \{x \in \mathbf{R}^d \mid (x-\mu)'\Sigma^{-1}(x-\mu) = r_\alpha^2\}. \tag{11}$$

REMARKS 3.2. The maximality at μ condition on $D(x; F_X)$ in Theorem 3.3 may be replaced by a convexity condition on D^α, which also suffices to prove the necessity part of (ii).

THEOREM 3.4. *Suppose $X \sim E_d(h; \mu, \Sigma)$. Then the depth contours of the simplicial depth, majority depth, simplicial volume depth SVD^α ($\alpha \ge 1$), $\widetilde{L}^2$ depth, projection depth, Mahalanobis depth and halfspace depth are surfaces of ellipsoids.*

THEOREM 3.5. *Suppose that $X \sim E_d(h; \mu, \Sigma)$ and $D(x)$ is affine invariant. Then $D(x)$ strictly decreases as x moves away from the center μ along any ray if and only if $D(x) = f\left((x-\mu)'\Sigma^{-1}(x-\mu)\right)$ for some strictly decreasing continuous function f.*

4. Convergence behavior of sample depth contours. We first establish an *almost sure result* about sample depth contours in a very general setting. We then turn to some specific cases of depth function. For convenience denote $D(x;F)$ by $D(x)$ and $D^{\alpha}(F)$ by D^{α} when the cdf F is understood.

THEOREM 4.1. *Let $D(x)$ be any nonnegative depth function and $D_n(x)$ a corresponding sample depth function. Let D^{α} and D_n^{α} be defined as in Definition* 3.1. *Assume*:

(C1) $D(x) \to 0$ *as* $\|x\| \to \infty$ *and*

(C2) $\sup_{x\in S} |D_n(x) - D(x)| \to 0$ *a.s. for any bounded set* $S \subset \mathbf{R}^d$.

Then, for any $\varepsilon > 0,\ \delta < \varepsilon,\ \alpha \geq 0$ *and* $\alpha_n \to \alpha$,

(i) $D^{\alpha+\varepsilon} \subset D_n^{\alpha_n+\delta} \subset D_n^{\alpha_n} \subset D_n^{\alpha_n-\delta} \subset D^{\alpha-\varepsilon}$ *a.s. for all sufficiently large* n (*uniformly if* $\alpha_n \to \alpha \in [0, \alpha_0]$ *uniformly as* $n \to \infty$, *for* $\alpha_0 < 1$).

(ii) $D_n^{\alpha_n} \overset{a.s.}{\to} D^{\alpha}$ *as* $n \to \infty$, *if* $P(\{x \in \mathbf{R}^d | D(x) = \alpha\}) = 0$. *The convergence is uniform in* α *if* $\alpha_n \to \alpha \in [0, \alpha_0]$ *uniformly as* $n \to \infty$, *for* $\alpha_0 < 1$.

Applying the preceding *general* result about depth-trimmed regions to the special case of *elliptical* distributions and *affine invariant* depth functions, we obtain the following corollary.

COROLLARY 4.1. *Let* $X \sim E_d(h;\mu,\Sigma)$. *Suppose that* $D(x)$ *is nonnegative and satisfies* P1 *and* P4, *that* $\sup_{x\in S} |D_n(x) - D(x)| \to 0$ *a.s. as* $n \to \infty$ *for any bounded set* $S \subset \mathbf{R}^d$, *and that* D_n^{α} *is convex and closed. Then*

$$(*) \quad \{x \in \mathbf{R}^d | D(x) = \alpha\} = \{x \in \mathbf{R}^d | (x-\mu)'\Sigma^{-1}(x-\mu) = r_{\alpha}^2\} \quad \text{for some } r_{\alpha},$$

if and only if for any $\alpha \in (0,1)$ *and* $\varepsilon > 0$, *there exists a* $\delta > 0$ *such that for* n *sufficiently large*,

$$(**) \qquad D^{h(q(\alpha-\varepsilon))} \subset D_n^{\beta_n(\alpha)+\delta} \subset D_n^{\beta_n(\alpha)-\delta} \subset D^{h(q(\alpha+\varepsilon))} \quad a.s.,$$

where $h(x)$, $q(\alpha)$ *and* β_n *are defined by* $P(\{x \in \mathbf{R}^d \mid e(x) \leq q(\alpha)\}) = \alpha$, $P_n(\{x \in \mathbf{R}^d \mid D_n(x) \geq \beta_n(\alpha)\}) = \lfloor \alpha n \rfloor / n$ *and* $D(x) = h(e(x))$. *Further* $(*)$ *implies that* $(**)$ *holds uniformly in* $\alpha \in [0, \alpha_0]$ *for* $\alpha_0 < 1$.

REMARKS 4.1. Corollary 4.1 improves and extends the main result of He and Wang (1997) by relaxing their conditions. Indeed, (D5) of He and Wang (1997) seems to be redundant, since convexity of D_n^{α} and convergence of $D_n(x)$ in He and Wang (1997) imply the convexity of D^{α}, which, combined with their condition (D3), our Remark 3.2 and Theorem 3.3, implies condition (D5).

We now examine a.s. convergence of sample depth contours for some specific depth functions: the simplicial depth, the projection depth and the general Type D depth.

THEOREM 4.2. *Theorem 4.2 Suppose that $X \sim E_d(h; \mu, \Sigma)$. Then for the simplicial, projection and general Type D depths,*

$$D_n^{\alpha_n} \to D^\alpha \quad a.s. \text{ as } n \to \infty,$$

for any sequence α_n with $\alpha_n \to \alpha$ as $n \to \infty$, where ∂D^α is an ellipsoid of the same shape as those of the constant density contours of the parent distribution. Further, the convergence is unform in α if $\alpha_n \to \alpha \in [0, \alpha_0]$ uniformly as $n \to \infty$, for $\alpha_0 < 1$.

REMARK 4.1. It is not difficult to see that the contours in Theorem 4.2 satisfy

$$\lim_{n\to\infty} \rho(D_n^{\alpha_n}, D^\alpha) = 0 \quad \text{a.s.},$$

with this convergence holding uniformly in α if $\alpha_n \to \alpha \in [0, \alpha_0]$ uniformly as $n \to \infty$, for $\alpha_0 < 1$, where ρ represents the *Hausdorff distance*, that is, for any sets A and B,

$$\rho(A, B) = \inf\{\varepsilon \mid \varepsilon > 0,\ A \subset B^\varepsilon, B \subset A^\varepsilon\},$$

where $A^\varepsilon = \{x \mid d(x, A) < \varepsilon\}$ and $d(x, A) = \inf\{d(x, y) \mid y \in A\}$.

For Type D depth functions with $\mathscr{C} = \mathscr{H}$, we obtain the following results for sample halfspace depth contours.

COROLLARY 4.2. *Suppose that $X \sim E_d(h; \mu, \Sigma)$. Then the sample depth-trimmed region $D^{\lfloor \alpha n \rfloor / n}$ of the halfspace depth converges almost surely and uniformly for $\alpha \in [0, 1/2]$, as $n \to \infty$, to D^α, an ellipsoid of the same shape as those of the constant density contours of the parent distribution.*

In particular, for the multivariate normal distribution we have:

COROLLARY 4.3. *Suppose that $X \sim N_d(\mu, \Sigma)$. Then, for the halfspace depth,*

$$\text{(i)} \qquad D^{\lfloor \alpha n \rfloor / n} \to D^\alpha = \left\{x \in \mathbf{R}^d \mid (x-\mu)'\Sigma^{-1}(x-\mu) \le \left(\Phi^{-1}(1-\alpha)\right)^2\right\} \quad a.s.$$

and this holds uniformly for $\alpha \in [0, 1/2]$, where $\Phi^{-1}(p)$ denotes the pth quantile of the standard normal distribution. Also

$$\text{(ii)} \qquad P\left(D^{\lfloor \alpha n \rfloor / n}\right) \to 1 - \beta \quad a.s.$$

uniformly for $\alpha \in [0, \frac{1}{2}]$, where β is determined by $\left(\Phi^{-1}(1-\alpha)\right)^2 = \chi_d^2(\beta)$ and $\chi_d^2(p)$ denotes the pth quantile of the chi-square distribution with d degrees of freedom.

REMARKS 4.2. (i) Corollary 4.2 slightly extends Lemma 2.5 of Donoho and Gasko (1992).

(ii) Applying Corollary 4.3 for $X \sim N_d(0, I)$ and Theorem 4, we obtain Theorem 1 of Yeh and Singh (1997).

APPENDIX A: PROOFS

PROOF OF THEOREM 3.1. Part (a) follows immediately from the affine invariance property P1 and the analogously defined property for sample depth functions. Part (b) follows directly from the definitions of D^α and D_n^α. Also, (c) is immediate. We now prove (d).

It is not difficult to show that $SVD^\beta(x;F)$ $(\beta \geq 1)$, $L^pD(x;F)$ $(p \geq 1)$, $PD(x;F)$ and $MHD(x;F)$ are (uniformly) continuous in x, which implies closedness of $D^\alpha(F)$ for these depth functions. Property P4 for these depth functions implies boundedness of their depth-trimmed regions, and thus the desired compactness follows.

For $SD(x;F)$, boundedness of the depth-trimmed region follows from Theorem 1 of Liu (1990) and closedness follows, for absolutely continuous distributions F, from the continuity of $SD(x;F)$ established in Theorem 2 of Liu (1990).

Finally, for Type D depth functions compactness of D^α is shown in Theorem 2.11 of Zuo and Serfling (2000b). □

PROOF OF THEOREM 3.2. For Type D depth functions we have

$$D^\alpha(F) = \cap\{C \mid F(C) > 1-\alpha,\ C \in \mathscr{C}\}, \tag{A.1}$$

as shown in the proof of Theorem 2.11 of Zuo and Serfling (2000b). Thus $D_n^\alpha = \cap\{C \mid \widehat{F}_n(C) > 1-\alpha,\ C \in \mathscr{C}\}$, from which follows that D_n^α is closed and connected. Since $D_n(x;\widehat{F}_n, \mathscr{C})$ satisfies P4, compactness of D_n^α follows. □

PROOF OF THEOREM 3.3. (i) Utilizing an argument similar to that for Lemma 3.1 of Liu and Singh (1993), one obtains (10). Since the points on the boundary of D^α are of equal depth, (9) follows. The monotonicity of g follows from the fact that, for any x_0, $D(\lambda\mu + (1-\lambda)x_0; F_X) \geq D(x_0;F_X)$, since $(\lambda\mu + (1-\lambda)x_0) \in D^{\alpha_0}$, where $\alpha_0 = D(x_0;F_X)$.

(ii) Sufficiency follows directly from Lemma 3.1 of Liu and Singh (1993). For necessity, we need to show that $D(x;F_X)$ is strictly decreasing if (11) holds. Let $y \neq \mu$ be a point in $\mathbf{R}^d$, and put $y_0 = \lambda\mu + (1-\lambda)y$ for some $\lambda \in (0,1)$. Then $y \in \partial D^{\alpha_y}$ and y_0 is in the interior of D^{α_y}, for some α_y such that $(y-\mu)'\Sigma^{-1}(y-\mu) = r^2_{\alpha_y}$. Hence $D(y_0;F_X) > D(y;F_X)$. □

PROOF OF THEOREM 3.4. By Theorem 3.3 and the affine invariance of these depth functions, we need only check the strictly decreasing property under the elliptical distribution assumption.

(i) For the simplicial depth function use an argument similar to that for Theorem 3 of Liu (1990).

(ii) For the majority depth function use

$$P\left(\lambda\mu + (1-\lambda)x \in H^F_{X_1,\dots,X_d}\right) - P\left(x \notin H^F_{X_1,\dots,X_d}\right)$$
$$= P\left(\lambda\mu + (1-\lambda)x \in H^F_{X_1,\dots,X_d},\ x \notin H^F_{X_1,\dots,X_d}\right) > 0,$$

for any $\lambda \in (0,1)$ and $x \neq \mu$ in $\mathbf{R}^d$.

(iii) For the simplicial volume depth function with $(\alpha \geq 1)$, following the proof of Corollary 2.1 of Zuo and Serfling (2000b), we have

$$\Delta(S[x_0; x_1, \dots, x_d]) \leq \lambda\Delta(S[\mu; x_1, \dots, x_d]) + (1-\lambda)\Delta(S[x; x_1, \dots, x_d]),$$

for any $\lambda \in (0,1)$, $x_0 = \lambda\mu + (1-\lambda)x$, $x \neq \mu$ and $x, x_1, \dots, x_d$ in $\mathbf{R}^d$ and

$$P\{\Delta(S[x_0; X_1, \dots, X_d]) < \lambda\Delta(S[\mu; X_1, \dots, X_d]) + (1-\lambda)\Delta(S[x; X_1, \dots, X_d])\} > 0,$$

for a random sample $X_1, \dots, X_d$ from X. The convexity of x^α for $\alpha \geq 1$ and $0 < x < \infty$, and the maximality of $SVD^\alpha(x; F)$ at μ, now imply that

$$\Delta^\alpha(S[x_0; x_1, \dots, x_d]) \leq \Delta^\alpha(S[x; x_1, \dots, x_d])$$

and

$$P\{\Delta^\alpha(S[x_0; X_1, \dots, X_d]) < \Delta^\alpha(S[x; X_1, \dots, X_d])\} > 0,$$

yielding $SVD^\alpha(x; F) < SVD^\alpha(x_0; F)$.

(iv) For $\widetilde{L}^2(x; F)$, the proof of Theorem 2.6 of Zuo and Serfling (2000b) yields

$$P(\|\lambda\mu + (1-\lambda)x - X\|_M < \lambda\|\mu - X\|_M + (1-\lambda)\|x - X\|_M) > 0,$$

for any positive definite matrix M, $\lambda \in (0,1)$ and $x \neq \mu$ in $\mathbf{R}^d$. Maximality of the depth function at μ now implies that

$$\widetilde{L}^2(\lambda\mu + (1-\lambda)x; F) > \widetilde{L}^2(x; F).$$

(v) For the projection depth function, apply the fact that $\mathrm{Med}(u'X) = u'\mu$ for any unit vector u in $\mathbf{R}^d$; see Zuo and Serfling (2000c).

(vi) For the Mahalanobis depth function, following the proof of (c) of Theorem 2.10 of Zuo and Serfling (2000b), we have

$$d_M^2(x_0; X) < \lambda d_M^2(\mu; X) + (1-\lambda) d_M^2(x; X),$$

for any $\lambda \in (0,1)$, $x \neq \mu$ in $\mathbf{R}^d$ and $x_0 = \lambda\mu + (1-\lambda)x$. Then maximality at μ implies $MHD(x_0; F) > MHD(x; F)$.

(vii) For the halfspace depth, let $\lambda \in (0,1)$, $x \neq \mu$ in $\mathbf{R}^d$ and $x_0 = \lambda\mu + (1-\lambda)x$. To consider the depth of points x_0 and x, we need only take the infimum of $P(H)$ over all $H \in \mathscr{H}$ which do not contain the center μ. Now, for any H_{x_0} with x_0 on its boundary, there always exists an H_x with x on its boundary such that $H_x \subset H_{x_0}$ and $P(H_x) + \varepsilon < P(H_{x_0})$ for some $\varepsilon > 0$. It follows that $HD(x; P) < HD(x_0; P)$. □

PROOF OF THEOREM 3.5. For convenience denote $D(x; F_X)$ by $D(x)$. The sufficiency is trivial. We need only prove the necessity.

(a) By Theorem 3.3, there is a function f such that

$$D(x) = f\left((x-\mu)'\Sigma^{-1}(x-\mu)\right).$$

(b) To show that f is strictly decreasing, let $q_2 > q_1 > 0$. Then there is an $x \in \mathbf{R}^d$, an $\alpha \in (0,1)$ and an $x_0 = \alpha\mu + (1-\alpha)x$ such that $q_1 = (x_0-\mu)'\Sigma^{-1}(x_0-\mu)$ and $q_2 = (x-\mu)'\Sigma^{-1}(x-\mu)$. Now $f(q_1) = D(x_0) > D(x) = f(q_2)$, proving that f is strictly decreasing.

(c) To show that f is continuous, we note, by Theorem 3.3, that $D(x)$ is upper and lower semicontinuous. Since $(x-\mu)'\Sigma^{-1}(x-\mu)$ is also continuous, the continuity of f follows from a standard result. □

PROOF OF THEOREM 4. (i) Clearly, $D_n^{\alpha_n+\delta} \subset D_n^{\alpha_n} \subset D_n^{\alpha_n-\delta}$. To show that $D_n^{\alpha_n-\delta} \subset D^{\alpha-\varepsilon}$, assume that $\alpha - \varepsilon > 0$ (the inclusion relation holds trivally when $\alpha - \varepsilon \le 0$, since then $D^{\alpha-\varepsilon} = \mathbf{R}^d$). Since $\alpha_n \to \alpha$, there is an N_1 such that when $n \ge N_1$, $|\alpha_n - \alpha| < (\varepsilon - \delta)/2$. Then, for ε_1 sufficiently small that $\alpha - \varepsilon - \varepsilon_1 > 0$, we have $D^{\alpha-\varepsilon} \subset D^{\alpha-\varepsilon-\varepsilon_1}$, and by (C1) the region $S = D^{\alpha-\varepsilon-\varepsilon_1}$ is bounded. Then, by (C2), there is an $N_2(\ge N_1)$ such that when $n \ge N_2$

$$(*) \qquad \sup_{x\in S} |D_n(x) - D(x)| \le (\varepsilon - \delta)/2 \quad \text{a.s.}$$

Let $x \in D_n^{\alpha_n-\delta} \cap (D^{\alpha-\varepsilon-\varepsilon_1} - D^{\alpha-\varepsilon})$. Then when $n \ge N_2$

$$D_n(x) - D(x) > \alpha_n - \delta - (\alpha - \varepsilon) \ge \alpha - \frac{\varepsilon-\delta}{2} - \delta - (\alpha-\varepsilon) \ge \frac{\varepsilon-\delta}{2},$$

contradicting $(*)$. Thus either $D_n^{\alpha_n-\delta} \subset D^{\alpha-\varepsilon}$ or $D_n^{\alpha_n-\delta} \cap D^{\alpha-\varepsilon-\varepsilon_1} = \varnothing$. But the latter is impossible, for if it is true and we let $x \in D_n^{\alpha_n-\delta}$, then

$$D_n(x) - D(x) > \alpha_n - \delta - (\alpha-\varepsilon-\varepsilon_1) \ge \alpha - \frac{\varepsilon-\delta}{2} - \delta - (\alpha-\varepsilon-\varepsilon_1) \ge \frac{(\varepsilon-\delta)}{2} + \varepsilon_1,$$

for any $n \ge N_2$, violating (C2) for ε_1 taken sufficiently small. Hence $D_n^{\alpha_n-\delta} \subset D^{\alpha-\varepsilon}$. Employing a similar argument as above, one can show that

$$D^{\alpha+\varepsilon} \subset D_n^{\alpha_n+\delta}.$$

(ii) It is easy to see that

$$\{x \in \mathbf{R}^d | D(x) > \alpha\} = \bigcup_{\varepsilon\in Q^+} D^{\alpha+\varepsilon} \subset \bigcap_{\varepsilon\in Q^+} D^{\alpha-\varepsilon} = \{x \in \mathbf{R}^d | D(x) \ge \alpha\},$$

where Q^+ is the set of positive rational numbers. By the result established above, we can show that

$$\bigcup_{\varepsilon\in Q^+} D^{\alpha+\varepsilon} \subset \liminf_{n\to\infty} D_n^{\alpha_n} \subset \limsup_{n\to\infty} D_n^{\alpha_n} \subset \bigcap_{\varepsilon\in Q^+} D^{\alpha-\varepsilon} \quad \text{a.s.}$$

Then $P(\{x \in \mathbf{R}^d | D(x) = \alpha\}) = 0$ implies that

$$\lim_{n\to\infty} D_n^{\alpha_n} = D^\alpha \quad \text{a.s.},$$

completing the proof. □

PROOF OF COROLLARY 4.1 (*Necessity*). Convexity of D_n^α and convergence of $D_n(x)$ imply the convexity of D^α. Remark 3.2 and Theorem 3.3 then yield that $D(x)$ is strictly decreasing as x moves away from μ along any fixed ray. Then Theorem 3.5 implies that $h(x)$ is strictly decreasing and continuous. By Lemma 3 of He and Wang (1997), $\lim_{n\to\infty} \beta_n(\alpha) = h(q(\alpha))$ uniformly in α. The continuity and monotonicity of $q(\alpha)$ and $h(x)$ imply that $h(q(\alpha+\varepsilon)) = h(q(\alpha)) - f_1(\varepsilon)$ and $h(q(\alpha-\varepsilon)) = h(q(\alpha)) + f_2(\varepsilon)$ for some continuous and positive functions $f_1(\varepsilon)$ and $f_2(\varepsilon)$. Necessity now follows from Theorem 4. Also, $(*)$ implies that $(**)$ holds uniformly in $\alpha \in [0, \alpha_0]$ for $\alpha_0 < 1$.

(*Sufficiency*.) By Remark 3.2 and Theorem 3.3, we need only show that $h(x)$ is strictly decreasing as x moves away from μ along any fixed ray. The nonincreasing property of $h(x)$ follows from $(**)$. Assume that $e(x) = q(\alpha+\varepsilon)$, $e(x_0) = q(\alpha-\varepsilon)$, $D_n(y) = \beta_n(\alpha)$ and

$$(***) \qquad D^{h(q(\alpha-\frac{\varepsilon}{2}))} \subset D_n^{\beta_n(\alpha)+\delta} \subset D_n^{\beta_n(\alpha)-\delta} \subset D^{h(q(\alpha+\frac{\varepsilon}{2}))},$$

for sufficiently large n. By convergence of $D_n(x)$ and $(***)$, we have $h(q(\alpha-\varepsilon) = D(x_0) \geq D_n(x_0) - \delta/2 \geq D_n(y) + \delta - \delta/2 \geq D(x) + \delta - \delta/2 = h(q(\alpha+\varepsilon) + \delta/2$, for sufficiently large n. By the continuity and monotonicity of $q(\alpha)$, we obtain that $h(x)$ is strictly decreasing. □

PROOF OF THEOREM 4.2. By results in Section 2 of Zuo and Serfling (2000b), the depth functions $SD(x; P)$, $PD(x; F)$ and $D(x; P, \mathscr{C})$ satisfy (C1) of Theorem 4. In Appendix B below, the a.s. uniform convergence of sample depth functions to population depth functions is established for the sample projection depth function $PD_n(x)$ and sample Type D depth functions $D_n(x; \mathscr{C})$. The same holds for the sample simplicial depth function [see, e.g., Corollary 6.8 of Arcones and Giné (1993)]. Thus, by Theorem 3.4, the conditions of Theorem 4 are satisfied for all three depth functions. □

PROOF OF COROLLARY 4.3. (i) Suppose $Y \in \mathbf{R}^d$ is normally distributed and $Y \sim N_d(0, I)$. By affine invariance of the halfspace depth, the depth contour D^α under Y is a sphere with radius $\Phi^{-1}(1-\alpha)$. Let $X = \Sigma^{\frac{1}{2}} Y + \mu$. Then $X \sim N_d(\mu, \Sigma)$ and affine invariance implies that

$$D^\alpha = \left\{x \in \mathbf{R}^d \mid (x-\mu)'\Sigma^{-1}(x-\mu) \leq \left(\Phi^{-1}(1-\alpha)\right)^2\right\},$$

which, combined with Corollary 4.2, gives (i).

(ii) Since

$$P\left(\liminf_{n\to\infty} D^{\lfloor \alpha n\rfloor/n}\right) \le \liminf_{n\to\infty} P\left(D^{\lfloor \alpha n\rfloor/n}\right)$$
$$\le \limsup_{n\to\infty} P\left(D^{\lfloor \alpha n\rfloor/n}\right) \le P\left(\limsup_{n\to\infty} D^{\lfloor \alpha n\rfloor/n}\right),$$

it follows that

$$\lim_{n\to\infty} P\left(D^{\lfloor \alpha n\rfloor/n}\right) = P\left(\lim_{n\to\infty} D^{\lfloor \alpha n\rfloor/n}\right)$$
$$= P\left((X-\mu)'\Sigma^{-1}(X-\mu) \le \left(\Phi^{-1}(1-\alpha)\right)^2\right).$$

Now since $X \sim N_d(\mu, \Sigma)$, we have $(X-\mu)'\Sigma^{-1}(X-\mu) \sim \chi^2_d$ and thus (ii) follows. □

APPENDIX B: BEHAVIOR OF SAMPLE DEPTH FUNCTIONS

In order to be "reasonable" estimators of population depth functions, the sample versions of depth functions should be *consistent*, that is, we desire that almost surely $[P]$,

$$\text{(B.1)} \qquad \sup_x |D_n(x) - D(x; P)| \to 0, \quad n \to \infty,$$

a condition that has been assumed in the results of Section 4 on convergence of contours. Further perspective is provided in Remark 3 of Appendix A of Zuo and Serfling (2000b). Results on (B.1) have been established for the *half-space depth* by Donoho and Gasko (1992), the *simplicial depth* by Liu (1990), Dümbgen (1990) and Arcones and Giné (1993), and the *majority* and *Mahalanobis* depths by Liu and Singh (1993). Here we prove (B.1) for the *projection* depth and all *Type D* depth functions.

B.1. *Sample projection depth function.* For $\widehat{F}_n$ the sample df, the outlyingness measure given by (6) takes the form

$$O_n(x) = \sup_{\|u\|=1} \frac{|\, u'x - \mathrm{Med}_{1\le i\le n}\{u'X_i\}\,|}{\mathrm{MAD}_{1\le i\le n}\{u'X_i\}}, \quad x \in \mathbf{R}^d,$$

where for a univariate sample $Y_1, \dots, Y_n$ with ordered values $Y_{(1)} \le \cdots \le Y_{(n)}$ we define

$$\mathrm{Med}_{1\le i\le n}\{Y_i\} = \tfrac{1}{2}\left(Y_{(\lfloor \frac{n+1}{2}\rfloor)} + Y_{(\lfloor \frac{n+2}{2}\rfloor)}\right),$$

$$\mathrm{MAD}_{1\le i\le n}\{Y_i\} = \mathrm{Med}_{1\le i\le n}\{|Y_i - \mathrm{Med}_{1\le j\le n}\{Y_j\}|\}.$$

The corresponding sample projection depth is then

$$PD_n(x) \equiv (1 + O_n(x))^{-1}, \quad x \in \mathbf{R}^d.$$

THEOREM B.1. *Assume that F satisfies*

$$\sup_{\|u\|=1} P(u'X \le \operatorname{Med}(u'X) - \varepsilon) < 1/2,$$

$$\sup_{\|u\|=1} P(|u'X - \operatorname{Med}(u'X)| \le \operatorname{MAD}(u'X) - \varepsilon) < 1/2,$$

$$\inf_{\|u\|=1} P(|u'X - \operatorname{Med}(u'X)| \le \operatorname{MAD}(u'X) + \varepsilon) > 1/2$$

and

$$\inf_{\|u\|=1} \operatorname{MAD}(u'X) > 0.$$

Then

$$\sup_{x\in\mathbf{R}^d} \left|PD(x;F) - PD_n(x)\right| \to 0 \qquad a.s.\ as\ n \to \infty.$$

PROOF. The conditions on F ensure that $\operatorname{Med}(u'X)$ and $\operatorname{MAD}(u'X)$ are unique for each unit vector u and also that the sample analogues converge almost surely to these limits uniformly in $\|u\| = 1$ as $n \to \infty$. The arguments are similar to the proof of Theorem 2.3.2 of Serfling (1980) for univariate sample quantiles using probability inequalities of Hoeffding (1963).

Now denote "Med," "MAD," "$\operatorname{Med}_{1\le i\le n}$" and "$\operatorname{MAD}_{1\le i\le n}$" by "$l$," "$s$," "$l_n$" and "$s_n$," respectively. Let $\varepsilon > 0$ be given. Then clearly

$$\begin{aligned}
O_n(x) &\ge \sup_{\|u\|=1} \frac{|\,|u'x| - \operatorname{Med}_{1\le i\le n}\{|u'X_i|\}\,|}{\operatorname{MAD}_{1\le i\le n}\{u'X_i\}} \\
&\ge \frac{|\,|v'x| - \operatorname{Med}_{1\le i\le n}\{|v'X_i|\}\,|}{\operatorname{MAD}_{1\le i\le n}\{v'X_i\}} \\
&= \frac{|\,\|x\| - \operatorname{Med}_{1\le i\le n}\{|v'X_i|\}\,|}{\operatorname{MAD}_{1\le i\le n}\{v'X_i\}} \\
&\ge \frac{\|x\| - \sup_{\|u\|=1} \operatorname{Med}_{1\le i\le n}\{|u'X_i|\}}{\sup_{\|u\|=1} \operatorname{MAD}_{1\le i\le n}\{u'X_i\}} \\
&= \frac{\|x\| - \sup_{\|u\|=1} l_n(u'X_i)}{\sup_{\|u\|=1} s_n(u'X_i)} \\
&\ge \frac{M_\varepsilon - \sup_{\|u\|=1} l_n(u'X_i)}{\sup_{\|u\|=1} s_n(u'X_i)},
\end{aligned}$$

for $\|x\| > M_\varepsilon$. This shows that $O_n(x) \to \infty$ a.s. and hence $PD_n(x) \longrightarrow 0$ a.s. as $\|x\| \to \infty$. More precisely, we have

$$PD_n(x) \le \frac{\sup_{\|u\|=1} s_n(u'X_i)}{M_\varepsilon + \sup_{\|u\|=1} s_n(u'X_i) - \sup_{\|u\|=1} l_n(u'X_i)}$$

$$\to \frac{s(u'X)}{M_\varepsilon + \sup_{\|u\|=1} s(u'X) - \sup_{\|u\|=1} l(u'X)} \quad \text{a.s.}$$

as $n \to \infty$. Now choose M_ε large enough that both $\sup_{\|x\|\le M_\varepsilon} PD(x;F) < \varepsilon$ and

$$\frac{\sup_{\|u\|=1} s(u'X)}{M_\varepsilon + \sup_{\|u\|=1} s(u'X) - \sup_{\|u\|=1} l(u'X)} < \varepsilon$$

are satisfied. Then almost surely $\sup_{\|x\|\le M_\varepsilon} PD_n(x) < 2\varepsilon$ for all sufficiently large n. Thus it suffices to show that for any given M

$$\sup_{\|x\|\le M} \left|PD(x;F) - PD_n(x)\right| \to 0 \quad \text{a.s. as } n \to \infty.$$

Now

$$\left| \sup_{\|u\|=1} \left| \frac{|u'x - l_n\{u'X_i\}|}{s_n\{u'X_i\}} \right| - \sup_{\|u\|=1} \left| \frac{|u'x - l(u'X)|}{s(u'X)} \right| \right|$$

$$\le \sup_{\|u\|=1} \left| \frac{|u'x - l_n\{u'X_i\}|}{s_n\{u'X_i\}} - \frac{|u'x - l(u'X)|}{s(u'X)} \right|$$

$$\le \sup_{\|u\|=1} \left| \frac{u'x - l_n\{u'X_i\}}{s_n\{u'X_i\}} - \frac{u'x - l(u'X)}{s(u'X)} \right|$$

$$\le \sup_{\|u\|=1} \left| \frac{u'x}{s_n\{u'X_i\}} - \frac{u'x}{s(u'X)} \right| + \sup_{\|u\|=1} \left| \frac{l(u'X)}{s(u'X)} - \frac{l_n\{u'X_i\}}{s_n\{u'X_i\}} \right|$$

$$= \mathrm{I} + \mathrm{II}, \quad \text{say.}$$

For $\|x\| \le M$ we have

$$\mathrm{I} = \sup_{\|u\|=1} \frac{\left|u'x\,(s(u'X) - s_n\{u'X_i\})\right|}{s_n\{u'X_i\} s(u'X)}$$

$$\le \frac{M\Delta_s^n}{\inf_{\|u\|=1} s_n\{u'X_i\} \inf_{\|u\|=1} s(u'X)},$$

where $\Delta_s^n = \sup_{\|u\|=1} |s(u'X) - s_n\{u'X_i\}|$ and

$$\mathrm{II} = \sup_{\|u\|=1} \frac{\left|(l(u'X) - l_n\{u'X_i\})s_n\{u'X_i\} + l_n\{u'X_i\}(s_n\{u'X_i\} - s(u'X))\right|}{s_n\{u'X_i\}s(u'X)}$$
$$\le \frac{\Delta_l^n \sup_{\|u\|=1} s_n\{u'X_i\} + \sup_{\|u\|=1} \left|l_n\{u'X_i\}\right| \Delta_s^n}{\inf_{\|u\|=1} s_n\{u'X_i\} \inf_{\|u\|=1} s(u'X)},$$

where $\Delta_l^n = \sup_{\|u\|=1} |l(u'X) - l_n\{u'X_i\}|$. Since $\inf_{\|u\|=1} s(u'X) > 0$, by the above results we thus obtain

$$\sup_{\|x\|\le M} \left| \sup_{\|u\|=1} \left| \frac{|u'x - l_n\{u'X_i\}|}{s_n\{u'X_i\}} \right| - \sup_{\|u\|=1} \left| \frac{|u'x - l(u'X)|}{s(u'X)} \right| \right|$$
$$\le \mathrm{I} + \mathrm{II} \to 0 \quad \text{a.s.}, \; n \to \infty. \qquad \square$$

B.2. *Sample Type D depth functions.* Let $D(x; P, \mathscr{C})$ and $D_n(x; \mathscr{C})$ be defined as in Example 2.1.

THEOREM B.2. *Let $\mathscr{C}$ be a Vapnik-Červonenkis (VC) class of sets in $\mathbf{R}^d$ and P a probability measure on $\mathbf{R}^d$. Suppose that $\sup_{C\in\mathscr{C}} |\widehat{P}_n(C) - P(C)|$ is measurable. Then*

$$\sup_{x\in\mathbf{R}^d} \left|D_n(x; \mathscr{C}) - D(x; P, \mathscr{C})\right| \to 0 \quad \textit{a.s. as } n \to \infty.$$

PROOF. Write

$$\sup_{x\in\mathbf{R}^d} \left|D_n(x; \mathscr{C}) - D(x; P, \mathscr{C})\right| = \sup_{x\in\mathbf{R}^d} \left| \inf_{C_x\in\mathscr{C}} \widehat{P}_n(C_x) - \inf_{C_x\in\mathscr{C}} P(C_x) \right|$$
$$\le \sup_{C\in\mathscr{C}} \left|\widehat{P}_n(C) - P(C)\right|,$$

where C_x is a set C with x on its boundary. Now apply well-known results on almost sure convergence of the empirical measure uniformly on VC classes [e.g., Shorack and Wellner (1986), page 828]. $\square$

For $\mathscr{C} = \mathscr{H}$, we obtain the result of Donoho and Gasko (1992) for the sample halfspace depth function.

Acknowledgments. The authors greatly appreciate the thoughtful and constructive remarks of an Associate Editor and two referees, which led to distinct improvements.

REFERENCES

ARCONES, M. A. and GINÉ, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542.

DONOHO, D. L. and GASKO, M. (1992). Breakdown properties of multivariate location parameters and dispersion matrices. *Ann. Statist.* **20** 1803–1827.

EDDY, W. F. (1985). Ordering of multivariate data. In *Computer Science and Statistics: The Interface* (L. Billard, ed.) 25–30. North-Holland, Amsterdam.

HE, X. and WANG, G. (1997). Convergence of depth contours for multivariate datasets. *Ann. Statist.* **25** 495–504.

HOEFFDING, W. (1963). Probability inequalities for sum of bounded random variables. *J. Amer. Statist. Assoc.* **36** 369–408.

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

LIU, R. Y. (1992). Data depth and multivariate rank tests. In L_1-*Statistics and Related Methods* (Y. Dodge, ed.) 279–294. North-Holland, Amsterdam.

LIU, R. Y., PARELIUS, J. M. and SINGH, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion). *Ann. Statist.* **27** 783–858.

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

MAHALANOBIS, P. C. (1936). On the generalized distance in statistics. *Proc. Nat. Acad. Sci. India* **12** 49–55.

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

MOSTELLER, C. F. and TUKEY, J. W. (1977). *Data Analysis and Regression.* Addison-Wesley, Reading, MA.

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

RAO, C. R. (1988). Methodology based on the L_1 norm in statistical inference. *Sankhyā Ser. A* **50** 289–313.

ROUSSEEUW, P. J. and HUBERT, M. (1999). Regression depth (with discussion). *J. Amer. Statist. Assoc.* **94** 388–433.

ROUSSEEUW, P. J. and RUTS, I. (1996). Bivariate location depth. *J. Roy. Statist. Soc. Ser. C* **45** 516–526.

ROUSSEEUW, P. J. and STRUYF, A. (1998). Computing location depth and regression depth in higher dimensions. *Statist. Comput.* **8** 193–203.

RUTS, I. and ROUSSEEUW, P. J. (1996). Computing depth contours of bivariate point clouds. *Comput. Statist. Data Analysis* **23** 153–168

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SERFLING, R. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.

SINGH, K. (1991). A notion of majority depth. Unpublished manuscript.

TUKEY, J. W. (1975). Mathematics and picturing data. In *Proceedings of the International Congress on Mathematics* (R. D. James, ed.) **2** 523–531. Canadian Math. Congress.

VAPNIK, V. N. and ČERVONENKIS, A. YA. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.* **16** 264–280.

YEH, A. B. and SINGH, K. (1997). Balanced confidence regions based on Tukey's depth and the bootstrap. *J. Roy. Statist. Soc. Ser. B* **59** 639–652.

ZUO, Y. and SERFLING, R. (2000a). Nonparametric notions of multivariate "scatter measure" and "more scattered" based on statistical depth functions. *J. Multivariate Anal.* To appear.

ZUO, Y. and SERFLING, R. (2000b). General notions of statistical depth function. *Ann. Statist.* **28** 461–482.

ZUO, Y. and SERFLING, R. (2000c). On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry. *J. Statist. Plann. Inference* **84** 55–79.

DEPARTMENT OF MATHEMATICS
ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85287-1804
E-MAIL: zuoyijun@math.la.asu.edu

PROGRAMS IN MATHEMATICAL SCIENCES
UNIVERSITY OF TEXAS AT DALLAS
RICHARDSON, TEXAS 75083-0688
E-MAIL: serfling@utdallas.edu