The Annals of Statistics
1997, Vol. 25, No. 2, 495–504

CONVERGENCE OF DEPTH CONTOURS FOR MULTIVARIATE DATASETS

BY XUMING HE[1] AND GANG WANG[2]

University of Illinois and DePaul University

Contours of depth often provide a good geometrical understanding of the structure of a multivariate dataset. They are also useful in robust statistics in connection with generalized medians and data ordering. If the data constitute a random sample from a spherical or elliptic distribution, the depth contours are generally required to converge to spherical or elliptical shapes. We consider contour constructions based on a notion of data depth and prove a uniform contour convergence theorem under verifiable conditions on the depth measure. Applications to several existing depth measures discussed in the literature are also considered.

1. Introduction. The notion of data depth is not new. It is indeed one of the fundamental issues in multivariate data analysis. Tukey (1975) put forward a measure of data ordering as follows. The depth of a point x in a one-dimensional dataset $\{x_1, x_2, \dots, x_n\}$ is the minimum of the number of data points on one side of x. The higher-dimensional depth of a point is the smallest depth in any one-dimensional projection of the dataset. The main idea of data depth is to provide an ordering of all points from a center outward. Contours of depth are often used to reveal the shape and structure of a multivariate dataset. These contours are analogous to quantiles in the univariate case, and they permit computation of L-estimators of location-scatter parameters (such as trimmed means). Recent studies of data depth that are affine invariant include Liu (1990), Donoho and Gasko (1992), Nolan (1992), Liu and Singh (1993), Massé and Theodorescu (1994) and Koshevoy and Mosler (1996) among others. Unlike the univariate case, multivariate ordering can be defined in different ways, but the following two requirements are usually desirable. First, for samples from certain class of distributions such as the elliptic ones, the depth contours should track the contours of the underlying model. Second, the contours should not be greatly influenced by outliers in the dataset. For example, contours based on the distance $(x - \bar{X}_n)' S_n^{-1} (x - \bar{X}_n)$, where $\bar{X}_n$ and S_n are the sample mean and covariance matrix, are usually not reliable as they can be fooled easily by a single outlier. The zonoid depth of Koshevoy and Mosler (1996) centers around $\bar{X}_n$, and therefore has the same pitfall.

There have also been studies of multivariate median and related notation of depth without imposing affine invariance. For example, the spatial median

Received July 1992; revised May 1996.

[1]Research partially supported by NSA Grant MDA904-96-1-0011.

[2]Research partially supported by a grant from the College of LAS of DePaul University.

AMS 1991 *subject classifications*. Primary 62H12; secondary 62F35, 62H05, 60H05.

Key words and phrases. Convergence, contour, data depth, elliptic distributions, location-scatter, M-estimator, multivariate dataset, robustness.

is only rotation equivariant. The confidence balls considered in Beran (1996) can be viewed as a type of depth contours that are defined on specific scales. However, we take the view that if a depth measure is to be constructed for a general purpose in multivariate data analysis, it is desirable that it can track the natural elliptic contours if the data constitute a random sample from an elliptic distribution.

It is intuitively clear that the contour convergence follows from convergence of depth. The technicalities involved however are not always trivial. The purpose of the present paper is to rigorously formalize such connections. We show that uniform convergence of depth contours is indeed verifiable by uniform convergence of the depth measure on any compact set.

The rest of the paper is organized as follows. In Section 2 we introduce a rather general notion of data depth and investigate the analytic and geometric properties for various depth related objects, all of which are preliminary to our main result on contour convergence. Convergence to elliptic contours is emphasized throughout. Section 3 considers applications to several useful depth measures. The Appendix provides detailed proofs.

2. Preliminaries and main results. Let $\{x_1, x_2, \dots, x_n\}$ be a sample of size n from a p-variate distribution F. We consider any depth measure $D_n(x; x_1, x_2, \dots, x_n) \equiv D_n(x)$ for $x \in R^p$ that satisfies the following conditions (D1) and (D2). A relaxation of the convexity condition in (D1) will be discussed in Example 3.2.

(D1) The set $O_c^n = \{x\colon D_n(x) \geq c\}$ is convex and closed almost surely for any c and n.

(D2) $\lim_{n\to\infty} D_n(x) = D(x)$ almost surely for each x.

For the convenience of presentation, we assume the following.

(D3) The contours of $D(x)$ are in the form of $\{x\colon e(x) = c\}$ for some $e(x)$.

For each n, let $x_1^n, \dots, x_n^n$ be ordered by the corresponding depths $D_n(x_1^n) \geq \cdots \geq D_n(x_n^n)$, Then, the αth depth contour can be constructed by

$$\delta^*(\alpha) = \{x\colon D_n(x) = D_n(x^n_{[\alpha n]})\}. \tag{2.1}$$

Let $\hat{\mu}_n$ be any point with the largest depth, that is, $D_n(\hat{\mu}_n) = \max_x D_n(x)$. The set

$$S_n(\alpha) = \{x : D_n(x) \geq D_n(x^n_{[\alpha n]})\} \tag{2.2}$$

contains $[\alpha n]$ data points closest to $\hat{\mu}_n$.

For any $e(x)$ specified in (D3), define

$$\delta_\alpha = \{x\colon e(x) = q(\alpha)\}, \qquad S_\alpha = \{x\colon e(x) \leq q(\alpha)\} \tag{2.3}$$

for each $\alpha \in (0, 1)$, where $q(\alpha)$ is determined by $P\{S_\alpha\} = \alpha$.

Depth contours are often used to understand or discover some underlying features from the data. Elliptic symmetry is one such feature that we can look for, as it would allow for more convenient modeling. If F is an elliptic distribution with location-scatter parameter (μ, Σ), that is, it has density function of the form

$$f(x) = |\Sigma|^{-1/2} g((x-\mu)'\Sigma^{-1}(x-\mu)), \tag{2.4}$$

where $\mu \in R^p$, Σ is a $p \times p$ positive definite matrix and g is a univariate function, we have natural elliptic contours for the distribution. To be more specific, take

$$e(x) = (x-\mu)'\Sigma^{-1}(x-\mu). \tag{2.5}$$

In this case, δ_α is the αth elliptic contour, and then the conditions (D2) and (D3) provide the basis for elliptic contour convergence.

If depth measures that are of only rotation invariance are of interest, we expect convergence of $\delta^*(\alpha)$ to spherical contours under spherically symmetric distributions with $e(x) = ||x-\mu||$. In the rest of the paper, we focus on the uniform convergence to elliptic contours under the models (2.4) and with $e(x)$ specified in (2.5) above. We note that only minor modifications are needed if convergence to spherical or other classes of contours is considered.

REMARK 1. The depth contour $\delta^*(\alpha)$ as defined by (2.1) can be a p-dimensional region in R^p, whereas the true contour δ_α is a $(p-1)$-dimensional surface. However, if we take the depth contour to be the boundary of $S_n(\alpha)$, which is also $(p-1)$-dimensional, all our results will remain valid.

To prove contour convergence, we need to strengthen the conditions (D2) and (D3) by the following conditions (D4) and (D5), respectively.

(D4) On any compact set $C \subset R^p$,

$$\lim_{n\to\infty} \sup_{x\in C} |D_n(x) - D(x)| = 0 \quad \text{a.s.} \tag{2.6}$$

(D5) $D(x)$ is a strictly monotone function of $e(x)$, which implies that for any $c > 0$,

$$P\{x: D(x) = c\} = 0. \tag{2.7}$$

In some applications, the convergence of $D_n(x)$ is actually uniform on R^p. But (2.6) is often easier to verify. For the rest of the section, we also assume without loss of generality that $D_n(x)$ is nonnegative for each x.

Let $O_c = \{x: D(x) \geq c\}$. First, we need to understand some basic properties of $D(x)$ and O_c by the following lemmas. Note that the closedness of O_c^n for any c is equivalent to upper semi-continuity of $D_n(x)$, namely

$$\lim_{r\to 0} \sup_{y\in B(x,r)} D_n(y) = D_n(x) \tag{2.8}$$

for any x, where $B(x, r) = \{y: |x-y| < r\}$.

LEMMA 1. *Under the conditions* (D1)–(D3), O_c *is convex, and* $D(x) = h(e(x))$ *for some nonincreasing function* h. *Furthermore, with condition* (D4), O_c *is closed for any* c, *and* h *is left continuous.*

Under (D5), we have $S_\alpha = O_{h(q(\alpha))}$. We now show that $S_n(\alpha)$ approximates S_α as $n \to \infty$. Denote by Ω the sample space.

LEMMA 2. *Suppose that* a_n *and* b_n *are two sequences of random variables such that for some random variable* a *taking values on* $[0, \infty]$, $a_n \to a$ *and* $b_n \to a$ *on a positive measure subset of the sample space* $S \subset \Omega$ *as* $n \to \infty$. *Then under the conditions* (2.6) *and* (2.7), $P\{O^n_{a_n} \triangle O_{b_n}\} \to 0$ *almost surely on the set* S, *where* $A \triangle B = (A \cup B) \setminus (A \cap B)$, *the symmetric difference of two sets.*

LEMMA 3. *Suppose that* (D1), (D4) *and* (D5) *are satisfied and the density function* $f(x)$ *is everywhere positive. Then* $\lim_{n\to\infty} D_n(x^n_{[\alpha n]}) = h(q(\alpha))$ *uniformly in* $\alpha \in [0, 1]$.

LEMMA 4. *Under the conditions of Lemma* 3, *for any* $c > c' > 0$, *almost surely there exists* $N = N(c, c')$ *such that* $\cup_{n \geq N} O^n_c \subset O_{c'}$. *In particular, the above sets are bounded almost surely.*

Our main result is now given as follows.

THEOREM 1. *Under the conditions* (D1), (D3) *and* (D4), (D5) *holds if and only if for any* $\alpha \in (0, 1)$ *and* $\varepsilon > 0$, *there exists* $\delta > 0$ *such that as* $n \to \infty$,

$$S_{\alpha-\varepsilon} \subset O^n_{D_n(x^n_{[\alpha n]})+\delta} \subset O^n_{D_n(x^n_{[\alpha n]})-\delta} \subset S_{\alpha+\varepsilon} \quad a.s. \tag{2.9}$$

Furthermore, (D5) *implies that the convergence in* (2.9) *is uniform in* $\alpha \in [0, \alpha_0]$ *for any* $\alpha_0 < 1$.

As a consequence, we have for every $\varepsilon > 0$ and $\alpha_0 < 1$, as $n \to \infty$,

$$\delta^*(\alpha) \subset S_{\alpha+\varepsilon} \setminus S_{\alpha-\varepsilon} \quad \text{a.s.} \tag{2.10}$$

uniformly in $\alpha \in [0, \alpha_0]$ under conditions (D1), (D4) and (D5). It also implies that the deepest point $\hat{\mu}_n$ is an consistent estimator of μ.

REMARK 2. Looking dizzy at first glance, (2.9) simply means that the depth contour which contains $[\alpha n]$ deepest data points contains $S_{\alpha-\varepsilon}$ even when it is made a little smaller, and is contained by $S_{\alpha+\varepsilon}$ even when it is made a little larger. Given here as a necessary and sufficient condition for (D5), (2.9) is just to indicate that the former is almost necessary for contour convergence. But examples show that (D5) is not necessary for (2.10). Note that the sample contours are invariant under any monotone transformation of $D_n(x)$. For some versions of D_n, the limiting function h satisfying $D(x) = h(e(x))$ can be flat on $e(x) \in (c_1, c_2)$.

REMARK 3. For contour approximations (2.10) to be uniform in all $\alpha \in [0, 1)$, condition (D4) needs to be strengthened to uniform convergence of depth in the whole space R^p. Consider an example where $D_n(x)$ is non-stochastic. It is equal to $D(x) = \{1 + e(x)\}^{-1}$ everywhere except on the strips $\{x: i \le e(x) \le i + a(n, i)\}$ on which it equals the value of $D(x)$ when $e(x) = i$ $(i = 1, 2, \dots)$. We choose $a(n, i)$ to be such that $a(n, i) \to 0$ as $n \to \infty$ for each fixed i, but the convergence is more and more slowly as i increases. In this case, conditions (D1)–(D5) are met, but (2.10) is not uniform for larger α.

3. Applications and examples. In this section, several types of affine invariant measures of depth are considered. Some are generated by location-scatter estimates, while the others are based on some specific depth measures found in the literature. In most cases, the uniform consistency of $D_n(x)$ is the hardest to check. It has to be done for each depth in consideration. The main contribution of Theorem 1 is to make convergence of contours directly verifiable through the depth.

EXAMPLE 3.1. Contours generated by location-scatter estimates. Based on an elliptic model distribution (2.4), any consistent and affine equivariant location-scatter estimate $(\hat{\mu}_n, \hat{\Sigma}_n)$ of (μ, Σ) will automatically provide an ordering of points by $r_n^2(x) = (x - \hat{\mu}_n)'\hat{\Sigma}_n^{-1}(x - \hat{\mu})$. Let $D_n(x) = \{1 + r_n^2(x)\}^{-1}$. It is then straightforward to verify that conditions (D1)–(D5) are satisfied with $D(x) = \{1 + e(x)\}^{-1}$. Equivalently, the resulting contour $\hat{\delta}_n(\alpha) = \{x: r_n^2(x) = r_{n,[\alpha n]}^2\}$ converges to the αth elliptic contour, where $r_{n,1}^2 \le r_{n,2}^2 \le \cdots \le r_{n,n}^2$ are the ordered sequence of $\{r_n^2(x_i)\}$.

Note that using different scatter matrix estimators that differ only by a multiplicative constant has no effects on the depth contours here. Therefore we have Theorem 2.

THEOREM 2. *If* $(\hat{\mu}_n, \hat{\Sigma}) \to (\mu, c\Sigma)$ *almost surely as* $n \to \infty$ *for some constant* $c > 0$, *then for any* $\varepsilon \in (0, \alpha)$, $\hat{\delta}_n(\alpha) \subset S_{\alpha+\varepsilon} \setminus S_{\alpha-\varepsilon}$ *for sufficiently large* n.

Even when the form of g in (2.4) is unknown, a class of M-estimators can be utilized to give desirable location-scatter estimates. An M-estimator of (μ, Σ) is the solution of

$$\sum_{i=1}^{n} v_1(e_i)(x_i - \mu) = 0,$$

$$\sum_{i=1}^{n} \{v_2(e_i)(x_i - \mu)(x_i - \mu)' - \Sigma\} = 0,$$

where v_1 and v_2 are properly chosen real-value functions on $[0, \infty)$ and $e_i = (x_i - \mu)'\Sigma^{-1}(x_i - \mu)$. By choosing the v_i $(i = 1, 2)$ functions for consistency at a fixed distribution, say the multivariate normal, the resulting estimate satisfies the assumption of Theorem 2 at any elliptic distribution; see Maronna (1976) for details.

The M-estimators are little affected by a small number of gross errors in the dataset. Parameter estimates that are highly robust against multiple outliers are also available; see Rousseeuw and Leroy (1987) and Davies (1987) for a class of S-estimators.

EXAMPLE 3.2 (Data depth based on U-statistics). A class of depth measures can be defined through U-statistics of the form

$$D_n(x) = \text{average}\{h(x; x_{i_1}, \dots, x_{i_m})\}, \tag{3.1}$$

where the average is taken over all subsets of size m from $(1, 2, \dots, n)$, and m is a fixed integer.

Oja (1983) considered a special case where h is the volume of the simplex formed by x and $m = p$ of the data points. More generally, if h is a continuous and convex function of x, then condition (D1) is satisfied. Arcones and Giné (1992) considered the uniform convergence of D_n. As a result, contour convergence can be easily verified.

One notable exception to the convexity requirement of (D1) is the simplicial depth of Liu (1990) with h taken to be the indicator function of $\{x \in S[x_{i_1}, \dots, x_{i_{p+1}}]\}$, where $S[x_{i_1}, \dots, x_{i_{p+1}}]$ is the (closed) simplex formed by the $p+1$ data points. Liu (1990) showed that the convergence of $D_n(x)$ to $D(x)$ is uniform in x, but the set O_c^n is not necessarily convex. However, if a point x has depth $D_n(x) = a$, then there must be a convex set around x as intersections of simplices over which D_n remains constant at a. This means that O_c^n must be a union of convex sets formed by halfspaces. By a well known result of Vapnik and Červonenkis [(1971), page 266]; see also Lemma 6.6 of Arcones and Giné (1992)], we know that the total number of convex sets formed by halfspaces is less than $(p+1)n^{p^2}$. The same proof in the Appendix shows that (D4) and (D5) imply (2.9). This is because (A.2) holds uniformly over all sets in the form of O_c^n [see Theorem 2 of Vapnik and Červonenkis (1971)].

EXAMPLE 3.3 (Projection depth). Donoho and Gasko (1992) discussed Tukey's notion of depth based on a one-dimensional projection. In our notation, the depth measure corresponds to $nD_n(x) = \min\{\#\{i\colon u'x_i \le u'x\}\colon |u| = 1\}$. Using the half-space metric, Donoho and Gasko (1992) showed that $D_n(x)$ converges uniformly to $\Pi(x) = \inf_u P(H_{u,x})$ where $H_{u,x}$ is the half space $\{y\colon u'y \le u'x\}$. They also verified our condition (D1). Contour convergence for this depth measure has been studied by Nolan (1992). Uniform contour convergence would follow from Theorem 1.

The contour construction generated by location-scatter estimates is essentially based on a continuous "outlyingness" measure. It produces exact ellipsoids as depth contours. They are specially geared for datasets with ellipsoidal symmetry. Depth measures like the simplicial depth and projection depth are discrete for a given sample of size n, and they are nonparametric in nature. Such depth contours track the shape of the datasets and remain informative in a broader range of problems.

It is of special interest that the deepest point, $\hat{\mu}_n$, serves as a generalized version of a multivariate median. There is a vast literature in this direction; see Rousseeuw and Leroy (1987), Arcones, Chen and Giné (1994), Bose and Chaudhuri (1993), just to mention a few. A global measure of insensitivity of the median estimate against data contamination is the so-called breakdown point. Write $D(x; F)$ as the limiting depth measure if the underlying distribution is F, and $\mu(F)$ as the maximizer of $D(x; F)$ over $x \in R^p$. Then the breakdown point of $\mu(F)$, the estimating functional, at the model distribution F_0 can be formulated as

$$\varepsilon^* = \sup\{\varepsilon > 0, \quad \sup_G |\mu((1-\varepsilon)F_0 + \varepsilon G)| < \infty\}.$$

The breakdown point of the projection method (Example 3.3) is limited to 1/3, since $\Pi(x)$ achieves its maximum at $x = x_0$ when $F = 2/3F_0 + 1/3G$ with G putting point mass at the point x_0 that can be arbitrarily remote. Donoho and Gasko (1992) gave the same bound for a finite sample version of the breakdown point. The breakdown point of the center estimate based on the simplicial depth is still unknown but believed to be positive and dimension dependent. An argument was given in Niinimaa, Oja and Tableman (1990), which indicates that the Oja median has a functional breakdown point of zero. The breakdown points of the location-scatter parameter estimates as we mentioned in Example 3.1, have been well studied in the robust statistics literature and will not be pursued here.

APPENDIX

Proofs. We prove the results in Sections 2 when F and $e(x)$ are in the forms of (2.4) and (2.5) respectively. The ideas of the proofs extend to more general settings.

Proof of Lemma 1. First, we show that O_c is convex. For any $x_1, x_2 \in O_c$, consider $x_0 = \lambda x_1 + (1-\lambda)x_2$ for $0 < \lambda < 1$. By (D2), for any $\varepsilon > 0$, we have $x_i \in O^n_{c-\varepsilon}$, $i = 1, 2$ for sufficiently large n. By convexity of $O^n_{c-\varepsilon}$, $x_0 \in O^n_{c-\varepsilon}$, and therefore $D(x_0) \geq c - 2\varepsilon$. Letting $\varepsilon \to 0$ yields $x_0 \in O_c$.

Under the additional assumption of (D4), similar arguments show that O_c is closed.

By (D3), $D(x) = h(e(x))$ for some function h. To see that h is nonincreasing, let $e(x_1) > e(x_2)$. Because $\{x\colon e(x) = e(x_1)\} \subset O_{D(x_1)}$ and O_c is convex for any c, we have $x_2 \in \{x\colon e(x) \leq e(x_1)\} \subset O_{D(x_1)}$. Hence, $h(e(x_2)) = D(x_2) \geq D(x_1) = h(e(x_1))$.

To show h is left continuous, let $y_n \uparrow y$. Then $h(y_n) \geq h(y-)$. Choose a sequence $x_n \in R^p$ such that $e(x_n) = y_n$ and $x_n \to x$, where $e(x) = y$. Then, $x_n \in O_{h(y-)}$ for all n. Because $O_{h(y-)}$ is closed, we have $x \in O_{h(y-)}$, and thus, $h(y) = D(x) \geq h(y-)$. Since h is nonincreasing, it follows that $h(y) = h(y-)$. $\square$

PROOF OF LEMMA 2. Observe by (2.6) that

$$\begin{aligned}\bigcap_{n=1}^{\infty}\bigcup_{N=n}^{\infty} O_{a_N}^N \Delta O_{b_N} &= \{x\colon x \in O_{a_n}^n \Delta O_{b_n},\ n, \text{ infinitely often}\}\\ &= \{x\colon (D_n(x) \ge a_n)\Delta(D(x) \ge b_n), n, \text{infinitely often}\}\\ &= \{x\colon (D_n(x) \ge a_n, D(x) < b_n) \cup (D_n(x) < a_n, D(x) \ge b_n),\\ &\qquad\qquad n, \text{infinitely often}\}\\ &\subset \{x\colon D(x) = a\}\end{aligned}$$

for almost all $\omega \in S$. Therefore, $P(\cup_{N=n}^{\infty} O_{a_N}^N \Delta O_{b_N}) \to 0$ on S almost surely by (2.7). □

PROOF OF LEMMA 3. First note that $q(\alpha)$ is a strictly increasing and continuous function of α since $P\{x\colon e(x) \le \alpha\}$ is. We will show for any convergent sequence $\alpha_n \to \alpha$ and $\varepsilon > 0$,

$$\liminf D_n(x_{[\alpha_n n]}^n) \ge h(q(\alpha + \varepsilon)) \quad \text{a.s.} \tag{A.1}$$

By a well-known theorem of Vapnik and Červonenkis (1971),

$$\sup_{S \text{ convex}} |P_n(S) - P(S)| \to 0 \quad \text{a.s.}, \tag{A.2}$$

where P_n is the empirical distribution given by $P_n(A) = (1/n)\#\{i\colon X_i \in A\}$.

Let $c_n = D_n(x_{[\alpha_n n]}^n)$, and $S = \{\omega \in \Omega\colon c_n < h(q(\alpha_n + \varepsilon))$, for infinitely many $n\}$. It suffices to show $P(S) = 0$. We only need to consider α, $\alpha_n \in [0, 1 - \varepsilon)$. If $P(S) > 0$, we have by (D1) and (A.2) that $P(O_{c_n}^n) \to \alpha$, since $P_n(O_{c_n}^n) = [\alpha_n n]/n \to \alpha$ almost surely on S.

For each $\omega \in S$, we can choose a subsequence n_k such that c_{n_k} is monotonely convergent to $c_\infty \le h(q(\alpha+\varepsilon))$. By Lemma 2, $P(O_{c_{n_k}}^{n_k} \Delta O_{c_{n_k}}) \to 0$, which implies $P\{x\colon D(x) \ge c_\infty\} = \alpha$ almost surely on S. But this cannot be true as we have $P\{x\colon D(x) \ge c_\infty\} \ge P\{x\colon e(x) \le q(\alpha + \varepsilon)\} = \alpha + \varepsilon$. The proof of (A.1) is complete.

Similarly, we can show that $\limsup c_n \le h(q(\alpha - \varepsilon))$ almost surely. Lemma 3 follows. □

PROOF OF LEMMA 4. It suffices to consider $c < D(\mu)$. By Lemma 1, $O_{c'}$ is a compact set. Thus by (D4), for $\varepsilon = \min\{D(\mu) - c, (c - c')/3\}$, there exists an integer N such that $n \ge N$ implies

$$\sup_{O_{c'}} |D_n(x) - D(x)| \le \varepsilon. \tag{A.3}$$

If $x \in O_c^n \cap (O_{c'} \setminus O_{c-2\varepsilon})$ for some $n \ge N$, then $D_n(x) - D(x) \ge c - (c - 2\varepsilon) > \varepsilon$, contradicting (A.3). So by convexity and by the fact that $\mu \in O_c^n \cap O_{c-2\varepsilon}$ for all $n \ge N$, we obtain $O_c^n \subset O_{c-2\varepsilon} \subset O_{c'}$, for $n \ge N$. □

PROOF OF THEOREM 1. First, we prove (2.9) under (D1)–(D5). Let

$$\delta = \frac{1}{2}\min\left\{h(q(\alpha-\varepsilon)) - h(q(\alpha - \frac{\varepsilon}{2})), h\left(q\left(\alpha+\frac{\varepsilon}{2}\right)\right) - h\left(q\left(\alpha+\varepsilon\right)\right)\right\}.$$

By (D5), we have $\delta > 0$. As in the proof of Lemma 3, write $c_n = D_n(x^n_{[\alpha n]})$. By (D4) and Lemma 3, there exists an integer N such that as $n \geq N$,

$$\sup_{S_\alpha} \left|D_n(x) - D(x)\right| \leq \delta, \tag{A.4}$$

and

$$h(q(\alpha+\varepsilon/2)) \leq c_n \leq h(q(\alpha-\varepsilon/2)) \tag{A.5}$$

uniformly in $\alpha \in [0, \alpha_0]$.

To show that $S_{\alpha-\varepsilon} \subset O^n_{c_n+\delta}$ for $n \geq N$, let $x \in S_{\alpha-\varepsilon} \subset S_\alpha$. Then $D(x) \geq h(q(\alpha-\varepsilon))$ and by (A.4) and (A.5),

$$\begin{aligned} D_n(x) \geq D(x) - \delta &\geq h(q(\alpha-\varepsilon)) - \frac{1}{2}\left(h(q(\alpha-\varepsilon)) - h\left(q\left(\alpha - \frac{\varepsilon}{2}\right)\right)\right) \\ &\geq c_n + \frac{1}{2}\left(h(q(\alpha-\varepsilon)) - h\left(q\left(\alpha-\frac{\varepsilon}{2}\right)\right)\right) \\ &\geq c_n + \delta, \end{aligned}$$

which means that $x \in O^n_{c_n+\delta}$.

To show $O^n_{c_n-\delta} \subset S_{\alpha+\varepsilon}$ as $n \to \infty$, we only need to consider $\alpha + \varepsilon < 1$. By Lemma 4, we can find an integer N' such that for $n \geq N'$, (A.5) holds and $\sup_{Q_\varepsilon} \left|D_n(x) - D(x)\right| \leq \delta$, where $Q_\varepsilon = \cup_{n\geq N'} O^n_{h(q(\alpha+\varepsilon/2))-\delta}$. Thus by (A.5), for $n \geq N'$,

$$\left|D_n(x) - D(x)\right| \leq \delta. \tag{A.6}$$

for all $x \in O^n_{c_n-\delta}$.

Let $n \geq N'$ and $x \in O^n_{c_n-\delta}$, then (A.6) implies that $D(x) \geq h(q(\alpha+\varepsilon))$, or $e(x) \leq q(\alpha+\varepsilon)$. Therefore, $O^n_{c_n-\delta} \subset S_{\alpha+\varepsilon}$.

Now, (2.9) implies that as $n \to \infty$,

$$\{x\colon D_n(x) = c_n\} \subset \{x\colon c_n - \delta \leq D_n(x) < c_n + \delta\} \subset S_{\alpha+\varepsilon} \setminus S_{\alpha-\varepsilon} \quad \text{a.s.}, \tag{A.7}$$

and therefore, $\lim_{n\to\infty} P\{\delta^*(\alpha) \subset S_{\alpha+\varepsilon} \setminus S_{\alpha-\varepsilon}\} = 1$ almost surely.

Next, we shall show that (D5) is necessary for (2.9).

Assume that h is a constant C on $[q(\alpha-2\varepsilon), q(\alpha+2\varepsilon)]$ for some α and $\varepsilon > 0$. As $n \to \infty$, (D4) implies

$$\sup_{S_{\alpha+2\varepsilon}} \left|D_n(x) - D(x)\right| \leq \frac{\delta}{2} \quad \text{a.s.} \tag{A.8}$$

By (A.7), $q(\alpha-\varepsilon) \leq e(x^n_{[\alpha n]}) \leq q(\alpha+\varepsilon)$ and thus $D(x^n_{[\alpha n]}) = C$ for sufficiently large n. Moreover, we have $O^n_{c_n-\delta} \setminus O^n_{c_n+\delta} \subset S_{\alpha+2\varepsilon}$. Thus, (A.8) implies that $|C - c_n| < \delta/2$. Consequently, for $x \in \{x\colon q(\alpha-2\varepsilon) \leq e(x) \leq q(\alpha+2\varepsilon)\} \subset S_{\alpha+2\varepsilon}$, we have $D(x) = C$ and $|D_n(x) - c_n| \le |D_n(x) - D(x)| + |D(x) - c_n| < \delta$ for sufficiently large n, which, together with (A.7), implies

$$\begin{aligned} S_{\alpha+2\varepsilon} \setminus S_{\alpha-2\varepsilon} &= \{x\colon q(\alpha - 2\varepsilon) < e(x) \le q(\alpha + 2\varepsilon)\} \\ &\subset \{x\colon c_n - \delta \le D_n(x) < c_n + \delta\} \subset S_{\alpha+\varepsilon} \setminus S_{\alpha-\varepsilon}, \end{aligned}$$

an obvious contradiction. The proof is then complete. □

Acknowledgments. The authors are grateful to the editor Professor Lawrence Brown, an Associate Editor and a referee for their assistance and helpful suggestions on an earlier version of the manuscript which was held up for nearly three years by the previous editorial board. We also thank Miguel Arcones and Regina Liu for helpful discussions and Rudy Beran and Gleb Koshevoy for sending us their preprints.

REFERENCES

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

ARCONES, M. A., CHEN, Z. and GINÉ, E. (1994). Estimators related to U-processes with applications to multivariate medians: asymptotic normality. *Ann. Statist.* **22** 1460–1477.

BERAN, R. (1996). Confidence sets centered at C_p-estimators. *Ann. Inst. Statist. Math.* **48**. To appear.

BOSE, A. and CHAUDHURI, P. (1993). On the dispersion of multivariate median. *Ann. Inst. Statist. Math.* **45** 541–550.

DAVIES, P. L. (1987). Asymptotic behavior of S-estimates of multivariate location parameters and dispersion matrices. *Ann. Statist.* **15** 1269–1292.

DONOHO, D. L. and GASKO, M. (1992). Breakdown properties of location estimates based on half-space depth and projected outlyingness. *Ann. Statist.* **20** 1803–1827.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

KOSHEVOY, G. and MOSLER, K. (1996). Zonoid trimming for multivariate distributions. Preprint.

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

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.

MARONNA, R. A. (1976). Robust M-estimators of multivariate location and scatter. *Ann. Statist.* **4** 51–67.

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

NIINIMAA, A., OJA, H. and TABLEMAN, M. (1990). The finite-sample breakdown point of the Oja bivariate median and of the corresponding half-samples version. *Statist. Probab. Lett.* **10** 325–328.

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

OJA, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.* **1** 327–332.

ROUSSEEUW, P. J. and LEROY, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.

TUKEY, J. W. (1975). Mathematics and picturing data. *Proceedings of International Congress of Mathematicians, Vancouver* **2** 523–531.

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.

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
CHAMPAIGN, ILLINOIS 61820
E-MAIL: he@stat.uiuc.edu

DEPARTMENT OF MATHEMATICS
DEPAUL UNIVERSITY
CHICAGO, ILLINOIS 60614
E-MAIL: gwang@condor.depaul.edu