The Annals of Statistics
1992, Vol. 20, No. 3, 1468–1484

ORDERING DIRECTIONAL DATA: CONCEPTS OF DATA DEPTH ON CIRCLES AND SPHERES[1]

By Regina Y. Liu and Kesar Singh

Rutgers University

Three notions of depth for directional data, *angular simplicial depth* (ASD), *angular Tukey's depth* (ATD) and *arc distance depth* (ADD), are developed and studied. The empirical versions of these depths give rise to center-outward rankings of angular data which may be regarded as extensions of the usual center-outward ranking on the line. Three medians derived from these depths are examined and compared. Applications in nonparametric classification and in implementing the bootstrap to construct confidence regions for directional parameters are briefly discussed.

1. Introduction. The purpose of this article is to develop three concepts of data depth for directional data, namely, angular simplicial depth (ASD), angular Tukey's depth (ATD) and arc distance depth (ADD). ASD extends the notion of simplicial depth (SD) in Liu (1988, 1990) from $\mathbb{R}^p$ to circles and spheres. ATD is an analog of Tukey's depth (TD) [Tukey (1975)] on $\mathbb{R}^p$ for populations and data on circles and spheres. A notion equivalent to ATD has been introduced by Small (1987). L_1 distance in the Euclidean space gives rise to the notion of ADD for spheres and circles.

The concept of depth on spheres leads to a proper notion of center (or median) and a ranking of directional data in the order of centrality. In particular, such ranking leads to detection of "extreme" data values, a natural definition of interquartile range (on the circle) and analogs of linear combinations of order statistics of directional data in general. The rankings derived from ASD and ATD can be justified as natural extensions of the usual linear ranking by the following argument. When the entire distribution is concentrated on a semicircle, the distribution could be regarded as being on the line segment $[-\pi/2, \pi/2]$. In such a case one would naturally expect the angular depths (being zero throughout the other semicircle) to coincide with their parent notions of depth on the line. Both ASD and ATD possess this *consistency property*. As a result, the center-outward rankings based on the decreasing values of these depths completely agree with that based on the usual order statistics on the line. As an illustrative example, suppose the angular data (in degrees) are 62, 73, 85, 96, 97; then the ranking in the order of centrality

Received March 1989; revised June 1991.
[1]Research supported by NSF Grants DMS-88-02558 and DMS-90-04658.
AMS 1980 *subject classifications*. Primary 62H05, 62H12; secondary 60D05.
Key words and phrases. Angular simplicial depth, angular Tukey's depth, arc distance depth, median, directional data, center-outward ordering, von Mises class of distributions.

provided by the linear ranking, ASD or ATD is 85, (73, 96), (62, 97), where a pair (,) indicates a tie.

The center-outward ranking given by ASD or ATD has several interesting applications. Specifically, we mention the following two:

1. *On classification problems:* Suppose two training samples $(X_1, \dots, X_m)$ and $(Y_1, \dots, Y_n)$ from two different spherical populations are given. Consider the problem of classifying a new data point Z to one of the two populations. First, compute, respectively, the center-outward ranks of Z with respect to X_i's and Y_i's. Let these ranks be denoted by r_X and r_Y. The proposed rule is to classify Z to the X population if $r_X/m < r_Y/n$, and to the Y population otherwise. This classification rule is studied in Gross and Liu (1989).
2. *On implementing the bootstrap:* Let θ be the parameter of interest on S_d, a d-dimensional unit sphere, and let $\hat{\theta}_n$ be its associated estimate. A center-outward ranking is essential for implementing the percentile method to form a bootstrap confidence region for θ [see Efron (1979) for the percentile method]. The procedure is as follows: First, obtain a certain number of bootstrap replicas of $\hat{\theta}_n$; second, assign the center-outward rank (according to ASD or ATD) to each replica; finally, delete $100\alpha\%$ of the "outmost" replicas. The smallest convex patch on S_d containing the remaining replicas is then a $(1 - \alpha)$ bootstrap confidence region of θ. Properties of this bootstrap confidence region will be reported elsewhere.

The constancy of ASD and ATD throughout a circle or a sphere presents an interesting situation. Of course, their parent depths are never constant on $\mathbb{R}^d$. We have fully studied the constancy, and the main results are distributional characterizations in terms of constant depth (cf. Sections 3 and 4).

All three notions of angular depth give rise to medians on S_d. Some detailed comparisons of those medians are presented later. The definition of a median on a circle given in Mardia [(1972), page 28] is in spirit the same as the median derived from ADD, although in some unusual cases the definition can lead to only a local maximum of ADD if the definition is followed literally.

For simplicity we restrict ourselves to continuous distributions on the unit circle and absolutely continuous distributions on the unit sphere; in each case we take the origin (denoted by O and $\mathbf{O}$, respectively) as the center. Throughout this article, $-\theta$ is used to indicate the diametrically opposite point of θ.

Section 2 contains basic definitions and notation.

In Section 3 we present the following properties of ASD w.r.t. a spherical distribution:

1. *Computational simplicity of ASD:* Checking that a point belongs to a spherical triangle is equivalent to solving a 3×3 linear system of equations.
2. *A differential formula for ASD and its applications:* The derivative of ASD($\cdot$) on a circle has a simple explicitly formula [see (3.1)] which yields many interesting properties of ASD. These properties include a monotonic-

ity property of the ASD and a characterization of an antipodally symmetric distribution on the circle as having a constant ASD value, 1/4.

3. *An equation connecting ASD and SD and its applications:* Equation (3.4) [(3.6)] connects ASD on a circle (a sphere) with the SD on a line (a plane). This equation can be used to characterize antipodally symmetric distributions on the sphere by the constant value of ASD = 1/8 throughout the sphere.

We show in Section 4 that ATD($\cdot$) has the appropriate properties to justify itself as a notion of data depth. The maximum value for this depth does not exceed 1/2 and it is attained, for instance, at the mode of any member of the von Mises class of distributions.

Section 5 discusses the robustness aspect of Mardia's median and medians given by ASD, ATD and ADD in terms of influence functions and a "breakdown" concept. Some illustrative examples are also given.

Some concluding remarks are made in Section 6.

2. Definitions and notation.

Angular simplicial depth. The angular simplicial depth which we propose in this article is a natural analog for directional data of the simplicial depth for data on euclidean spaces, introduced in Liu (1988, 1990), which we now describe briefly.

In $\mathbb{R}^d$, a simplex $\diamond(x_1, \dots, x_{d+1})$ with $(d+1)$ vertices $x_1, \dots, x_{d+1}$ is defined to be the closed convex hull with extremities at these points. Let $F(\cdot)$ be a distribution and x a point in $\mathbb{R}^d$. The simplicial depth of x w.r.t. F, SD(x), is then defined to be the probability that x be in a simplex $\diamond(X_1, \dots, X_{d+1})$, where $X_1, \dots, X_{d+1}$ are $(d+1)$ i.i.d. observations from F. In $\mathbb{R}^1$, $\diamond(X_1, X_2)$ is simply the closed line segment joining X_1 and X_2, say $\overline{X_1X_2}$, and $\mathrm{SD}(x) \equiv P_F(x \in \overline{X_1X_2})$ [$= 2F(x)(1 - F(x))$, assuming that F is continuous]. In $\mathbb{R}^2$, $\diamond(X_1, X_2, X_3)$ is the closed triangle with vertices X_1, X_2 and X_3, say $\Delta(X_1, X_2, X_3)$, and $\mathrm{SD}(x) \equiv P_F(x \in \Delta(X_1, X_2, X_3))$. The simplicial median is then the point which maximizes SD($\cdot$) (or the average of such points if there are many). Note that in $\mathbb{R}^1$ the simplicial median divides the line into two half-lines of equal probabilities and it agrees with the "usual" median. In Liu (1990) it is argued that SD($\cdot$) can be viewed as a measure of data depth, and that the simplicial median possesses many desirable features of a notion of median.

The edges of a simplex in $\mathbb{R}^d$ are the line segments connecting pairs of points (vertices). When we move to the sphere, it is natural to replace such a line segment by the "shortest curve" joining a pair of points on the sphere. Let p_1 and p_2 be two points on a sphere. It is known that such a shortest curve is the short arc joining p_1 and p_2 on the circle which passes through p_1 and p_2 and has the same center as the sphere. (Such a circle is referred to as a great circle.) Evidently this shortest curve can be generalized to spheres of any dimension and is ambiguous only in the nongeneric case where p_1 and p_2 are diametrically opposite each other. The idea of the shortest curve allows us to generalize the notion of simplex to the spherical case. We discuss only the cases of the circle and the two-dimensional sphere, although it will be clear that the definition extends inductively to any dimension. For any two points p_1 and p_2 on a circle the corresponding simplex is the short arc joining p_1 and p_2 [denoted by $\mathrm{arc}(p_1, p_2)$], and for three points q_1, q_2 and q_3 on a sphere it is the spherical triangle [denoted by $\Delta_s(q_1, q_2, q_3)$] bounded by the three short arcs $\mathrm{arc}(q_1, q_2)$, $\mathrm{arc}(q_1, q_3)$ and $\mathrm{arc}(q_2, q_3)$. We now define the angular simplicial depth to be

$$\mathrm{ASD}(p) \equiv P_H(p \in \mathrm{arc}(W_1, W_2)) \tag{2.1}$$

if p is a point and H a distribution on a circle, and W_1 and W_2 are i.i.d. observations from H;

$$\mathrm{ASD}(\mathbf{p}) = P_H(\mathbf{p} \in \Delta_S(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3)) \tag{2.2}$$

if $\mathbf{p}$ is a point and H a distribution on a sphere and $\mathbf{W}_1$, $\mathbf{W}_2$ and $\mathbf{W}_3$ are i.i.d. observations from H. Note that if H is continuous on a circle and absolutely continuous on a sphere, then the ambiguous simplicies occur with probability zero. A maximum point of $\mathrm{ASD}(\cdot)$ is defined to be an *angular simplicial median* (ASM). Evidently this median is *rotation invariant*.

We define the empirical version of $\mathrm{ASD}(\cdot)$ as

$$\mathrm{ASD}_n(p) \equiv \binom{n}{2}^{-1} \sum_{*} I\big(p \in \mathrm{arc}(W_{i_1}, W_{i_2})\big) \tag{2.3}$$

for a point p on the circle, where $W_1, \dots, W_n$ is a random sample from a circular distribution and $\sum_*$ runs over all possible pairs of (W_{i_1}, W_{i_2}), and as

$$\mathrm{ASD}_n(\mathbf{p}) = \binom{n}{3}^{-1} \sum_{**} I\big(\mathbf{p} \in \Delta_s(\mathbf{W}_{i_1}, \mathbf{W}_{i_2}, \mathbf{W}_{i_3})\big) \tag{2.4}$$

for a point $\mathbf{p}$ on the sphere, where $\mathbf{W}_1, \dots, \mathbf{W}_n$ is a random sample from a spherical distribution and $\sum_{**}$ runs over all possible triplets $(\mathbf{W}_{i_1}, \mathbf{W}_{i_2}, \mathbf{W}_{i_3})$.

Angular Tukey's depth. Following Small (1987), we define the *angular Tukey's depth* for a given spherical distribution H as follows:

$$\mathrm{ATD}_H(\boldsymbol{\theta}) = \inf_{\{S:\, \boldsymbol{\theta} \in S\}} \{P_H(S)\}, \tag{2.5}$$

where the infimum is taken over the set of all closed hemispheres S containing θ in their boundaries *or* in their interiors.

We call a maximum point of $\mathrm{ATD}(\cdot)$ an *angular Tukey's median* (ATM). Note that ATM is also *rotation invariant*. See Example 4.4.4 in Small (1987) for more invariance properties of ATM.

In defining the empirical version of $\mathrm{ATD}(\cdot)$, $\mathrm{ATD}_n(\cdot)$, we replace $P_H(\cdot)$ in (2.5) by its corresponding empirical probability.

Arc distance depth. We define ADD of a point θ on the sphere S_d as

$$\text{ADD}(\theta) = \pi - \int l(\theta, \varphi)\, dH(\varphi),$$

where $l(\theta, \varphi)$ is the Riemannian distance between θ and φ; that is, the length of the short arc joining θ and φ on the great circle determined by θ and φ. Again, the empirical version of ADD is defined by replacing $H(\cdot)$ by $H_n(\cdot)$. A maximum point of ADD$(\cdot)$ is referred to as an *arc distance median* (ADM). This idea of minimizing L_1 distance was used by Gower (1974) to define a generalized median in $\mathbb{R}^d$. Its extension to circles was given in Mardia (1972) and to spheres in Fisher (1985).

3. Properties of angular simplicial depth.

Computational simplicity of ASD. We first point out that determining whether or not a point on a circle (a sphere) lies on the short arc joining two data points (the spherical triangle three data points) can be reduced to solving a simple system of linear equations. This shows that computing ASD$(\cdot)$ is quite straightforward. Let $H(\cdot)$ be the population distribution defined on the unit circle centered at the origin O. Given a point θ on the circle and any two data points W_1 and W_2 from $H(\cdot)$, θ lies on the short arc arc(W_1, W_2) if and only if the line segments $\overline{O\boldsymbol{\theta}}$ and $\overline{W_1W_2}$ intersect. In other words, $\theta \in$ arc(W_1, W_2) if and only if there exist α and β such that $0 \le \alpha, \beta \le 1$ and $\alpha\theta^c = \beta W_1^c + (1-\beta)W_2^c$. Here the notation $*^c$ stands for the Euclidean coordinates of the point $*$. For the spherical case this observation becomes the following: $\boldsymbol{\theta}$ is on the spherical triangle $\Delta_s(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3)$ if and only if $\overline{O\boldsymbol{\theta}}$ intersects the Euclidean triangle $\Delta(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3)$. This is equivalent to $\alpha\theta^c = \beta\mathbf{W}_1^c + \gamma\mathbf{W}_2^c + (1-\beta-\gamma)\mathbf{W}_3^c$ for some α, β and γ such that $0 \le \alpha, \beta, \gamma \le 1$ and $0 \le \beta + \gamma \le 1$. The same observation holds for any general S_d. This computational simplicity of ASD should greatly enhance its applicability.

A differential formula for ASD and its applications. Let $H(\cdot)$ be the distribution on the unit circle and let $h(\cdot)$ be its density if it exists. Here θ can be simply expressed as an angle between 0 and 2π.

PROPOSITION 3.1. *Suppose that $h(\cdot)$ exists and is continuous at θ. Then*

$$\frac{d}{d\theta}\text{ASD}(\theta) = 2(A_\theta - C_\theta)h(\theta), \tag{3.1}$$

where A_θ and C_θ stand for the probabilities of the semicircles joining θ and $-\theta$ in the counterclockwise and clockwise directions, respectively.

PROOF. For a θ and a positive increment $\delta\theta$, the difference [ASD$(\theta + \delta\theta)$ − ASD(θ)] involves only those pairs of observations $\{W_1, W_2\}$ from $H(\cdot)$ which have the following property:

$$\{\theta \in \mathrm{arc}(W_1, W_2) \text{ and } (\theta + \delta\theta) \notin \mathrm{arc}(W_1, W_2)\}$$

or

$$\{\theta \notin \mathrm{arc}(W_1, W_2) \text{ and } (\theta + \delta\theta) \in \mathrm{arc}(W_1, W_2)\}.$$

These two situations will occur if and only if either W_1 or W_2 lies on $\mathrm{arc}(\theta, \theta + \delta\theta)$. Using this fact and the equality

$$P(E_1) - P(E_2) = P(E_1 - E_2) - P(E_2 - E_1),$$

for any two events E_1 and E_2, we obtain

$$\mathrm{ASD}(\theta + \delta\theta) - \mathrm{ASD}(\theta) = 2(A_\theta - C_\theta)\int_\theta^{\theta+\delta\theta} h(a)\, da + o(\delta\theta). \tag{3.2}$$

The proposition follows. □

We define a point θ to be *regular* w.r.t. a distribution $H(\cdot)$ if $H(\cdot)$ has a continuous density in a neighborhood of θ. We also define a point to be a *median axis* on the circle if the diameter passing through θ and $-\theta$ divides the circle into two semicircles with equal probabilities. The following proposition asserts that ASM is always a median axis.

PROPOSITION 3.2. *If θ_0 is a median axis with $h(\theta_0) > h(-\theta_0)$ and the points θ_0 and $-\theta_0$ are regular, then θ_0 is a local maximum of* ASD. *Conversely, if θ_0 is a local maximum of* ASD *and θ_0 and $-\theta_0$ are regular with $h(\theta_0) > 0$, then θ_0 is a median axis and $h(\theta_0) \geq h(-\theta_0)$.*

REMARK 3.1. (i) On the circle ADD$(\cdot)$ also allows a simple differential equation, namely,

$$\frac{d}{d\theta}\mathrm{ADD}(\theta) = (A_\theta - C_\theta)$$

provided that $h(\cdot)$ exists at θ and $-\theta$. The proof is given in Mardia [(1972), page 31].

(ii) The equation in (i) immediately implies that statements similar to Proposition 3.2 hold for ADD.

COROLLARY 3.1 [Monotonicity of ASD$(\cdot)$]. *Suppose $h(\cdot)$ is symmetric about its maximum point θ_0 and decreases monotonically on both sides of θ_0 until its diametrically opposite point $-\theta_0$. Then* ASD$(\cdot)$ *is also monotonic nonincreasing in both directions from θ_0 to $-\theta_0$. In particular, θ_0 is a maximum point of* ASD$(\cdot)$.

The next property of ASD$(\cdot)$ on a circle will allow us to characterize antipodally symmetric distributions. These are defined as follows.

DEFINITION. Let H be the distribution of a random variable W on d-dimensional sphere S_d. H is said to be *antipodally symmetric* (about the origin) if the distribution of $(-W)$ is also H, where $(-W)$ stands for the diametrically opposite point of W. If H has a continuous density, then antipodal symmetry is equivalent to $h(\theta) = h(-\theta)$ for all θ on S_d.

PROPOSITION 3.3. *Assume that $h(\cdot)$ is continuous. Then* ASD(θ) = c, *for a positive constant c and all $\theta \in [0, 2\pi)$, if and only if $h(\theta) = h(-\theta)$ for all θ. Moreover, the constant c must then be* 1/4.

PROOF. ($\Rightarrow$) Suppose ASD(θ) = c throughout. Then by Proposition 3.1, either $A_\theta = C_\theta$ or $h(\theta) = h(-\theta) = 0$ holds for every θ. Thus, $h(\theta) = h(-\theta)$ for all θ in view of the continuity of $h(\cdot)$.

($\Leftarrow$) If $h(\theta) = h(-\theta)$ for all θ, then $A_\theta = C_\theta$ and d ASD(θ)/$d\theta$ = 0 for all θ, which implies ASD(θ) = c for some constant c.

To show that c must be 1/4, we need only show that ASD(0) = 1/4 since the same argument applies to a general point θ after a rotation of the axes by θ. Due to antipodal symmetry [i.e., $h(\theta) = h(-\theta)$ for all θ], we have

$$\mathrm{ASD}(0) = 2\int_0^{\pi} \left(\tfrac{1}{2} - H(a)\right) h(a)\, da.$$

It suffices to check that

$$\int_0^{\pi} H(a) h(a)\, da = \tfrac{1}{8}. \tag{3.3}$$

This is done by letting $H(a) = y$ and converting (2.3) into $\int_0^{1/2} y\, dy$. □

Note that an alternative proof of Proposition 3.3 can be given [in fact under the weaker condition that $H(\cdot)$ is continuous only] using the connecting equation in Proposition 3.4.

REMARK 3.2. On a circle this characterization by the constancy of ASD($\cdot$) also holds for ADD($\cdot$). The constant c there is $\pi/2$.

REMARK 3.3. It may be of interest to note that if the underlying distribution has its probability mass concentrated on a semicircle only, then the depth ASD($\cdot$) is zero for all points on the complementary semicircle.

An equation connecting ASD with SD and its applications. Equation (3.1) for the rate of change of the ASD was the main tool in our study for the circular case. For the sphere, there does not seem to be any such simple equation. Instead we focus on a reduction of the sphere to the plane tangent to the sphere at a given point. Such a process is often called an exponential map. This will allow us to apply some known properties of SD on $\mathbb{R}^2$. To make the discussion clearer, we begin with the construction in the case of circles.

Let $H(\cdot)$ be the distribution and θ some fixed point on the unit circle. There is a natural length-preserving mapping g_θ from the circle (without the point $-\theta$) to the segment $(-\pi, \pi)$ of the tangent line L_θ at θ. For a point φ on the unit circle ($\varphi \neq -\theta$), the absolute value $|g_\theta(\varphi)|$ is simply the length of $\mathrm{arc}(\theta, \varphi)$, and the sign of $g_\theta(\varphi)$ being $-$ or $+$ depends on whether the direction in going from θ to φ is counterclockwise or clockwise. $H_\theta(\cdot)$ is used to represent the resulting distribution on the tangent line L_θ with its entire probability mass on $(-\pi, \pi)$. Let $\mathrm{SD}_\theta(\cdot)$ be the simplicial depth on the tangent line L_θ w.r.t. the distribution $H_\theta(\cdot)$. The depth functions $\mathrm{ASD}(\cdot)$ and $\mathrm{SD}_\theta(\cdot)$ are connected as follows.

PROPOSITION 3.4 (On the circle). *If $H(\cdot)$ is continuous, then for all θ on the circle*

$$\mathrm{ASD}(\theta) + \mathrm{ASD}(-\theta) = \mathrm{SD}_\theta(0). \tag{3.4}$$

PROOF. Let $\{W_1, W_2\}$ be a random sample from $H(\cdot)$. Except for a null set the following three events are equivalent:

$$\left\{0 \in \overline{g_\theta(W_1) g_\theta(W_2)}\right\},$$

$\{W_1$ and W_2 are on two different sides of the diagonal joining θ and $-\theta\}$

and

$$\{\theta \in \mathrm{arc}(W_1, W_2)\} \cup \{-\theta \in \mathrm{arc}(W_1, W_2)\}. \tag{3.5}$$

The proposition follows from the fact that the intersection of the two events in (3.5) has probability 0. □

We now turn to the two-dimensional sphere. Similarly we let $H(\cdot)$ be a distribution and $\boldsymbol{\theta}$ a fixed point on the unit sphere. Let $P_{\boldsymbol{\theta}}$ be the tangent plane to the sphere at $\boldsymbol{\theta}$. For a point φ ($\varphi \neq \boldsymbol{\theta}, -\boldsymbol{\theta}$), consider the great circle which passes through $\boldsymbol{\theta}$ and $\boldsymbol{\varphi}$. The plane of this circle cuts $P_{\boldsymbol{\theta}}$ along a line $L_{\boldsymbol{\theta}, \varphi}$ which is just the tangent line to the circle at $\boldsymbol{\theta}$. This means that if we restrict our attention to this circle and line $L_{\boldsymbol{\theta}, \varphi}$ we are exactly in the situation of the circle discussed in the previous paragraph. In particular, the construction described there applies and $\boldsymbol{\varphi}$ can be mapped into a point $g_{\boldsymbol{\theta}}(\boldsymbol{\varphi})$ on the line $L_{\boldsymbol{\theta}, \varphi}$ between $(-\pi, \pi)$. As $\boldsymbol{\varphi}$ moves sideways on the sphere, it is evident that the line $L_{\boldsymbol{\theta}, \varphi}$ will rotate on the plane $P_{\boldsymbol{\theta}}$. The sphere without $-\boldsymbol{\theta}$ will be mapped by $g_{\boldsymbol{\theta}}$ into a disc in $P_{\boldsymbol{\theta}}$ with center $\boldsymbol{\theta}$ (which is the origin of $P_{\boldsymbol{\theta}}$ now). As before, we use $H_{\boldsymbol{\theta}}(\cdot)$ to denote the resulting distribution on $P_{\boldsymbol{\theta}}$, which has its total probability mass on the disc centered at the origin O with radius π. The analog of (3.4) for the sphere can now be stated as follows.

PROPOSITION 3.5 (On the sphere). *Assume that $H(\cdot)$ is absolutely continuous. Then*

$$\mathrm{ASD}(\boldsymbol{\theta}) + \mathrm{ASD}(-\boldsymbol{\theta}) = \mathrm{SD}_{\boldsymbol{\theta}}(O), \tag{3.6}$$

where $\mathrm{SD}_{\boldsymbol{\theta}}(\cdot)$ is the simplicial depth on the plane corresponding to $H_{\boldsymbol{\theta}}(\cdot)$.

PROOF. Fix $\mathbf{W}_1$ and $\mathbf{W}_2$. The great circles joining $\{\mathbf{W}_1, \boldsymbol{\theta}\}$ and $\{\mathbf{W}_2, \boldsymbol{\theta}\}$, respectively, split the sphere into four pieces. Let $S(\mathbf{W}_1, \mathbf{W}_2, \boldsymbol{\theta})$ be the smaller piece which does not contain $\mathbf{W}_1$ and $\mathbf{W}_2$. (This exists with probability 1.) For any $\mathbf{W}_3$,

$$\mathbf{O} \in \Delta(g_{\boldsymbol{\theta}}(\mathbf{W}_1), g_{\boldsymbol{\theta}}(\mathbf{W}_2), g_{\boldsymbol{\theta}}(\mathbf{W}_3)) \quad \text{if and only if} \quad \mathbf{W}_3 \in S(\mathbf{W}_1, \mathbf{W}_2, \boldsymbol{\theta}).$$

The event $\mathbf{W}_3 \in S(\mathbf{W}_1, \mathbf{W}_2, \boldsymbol{\theta})$ can be divided into the following two mutually exclusive events:

(I) Both $\boldsymbol{\theta}$ and $\mathbf{W}_3$ lie on one of the two hemispheres determined by the great circle joining $\mathbf{W}_1$ and $\mathbf{W}_2$.

(II) Both $-\boldsymbol{\theta}$ and $\mathbf{W}_3$ lie on one of the two hemispheres determined by the great circle joining $\mathbf{W}_1$ and $\mathbf{W}_2$.

Note that (I) is equivalent to the event that $\boldsymbol{\theta} \in \Delta_s(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3)$ and (II) to $-\boldsymbol{\theta} \in \Delta_s(\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3)$. This proves the assertion. □

We now derive the following simple characterization of antipodally symmetric distributions on the sphere.

PROPOSITION 3.6 (On the sphere). *Assume that $h(\cdot)$ is continuous. Then* $\mathrm{ASD}(\boldsymbol{\theta}) = 1/8$ *for all $\boldsymbol{\theta}$ if and only if $h(\boldsymbol{\theta}) = h(-\boldsymbol{\theta})$ for all $\boldsymbol{\theta}$.*

PROOF. ($\Leftarrow$) If $H(\cdot)$ is antipodally symmetric [i.e., $h(\boldsymbol{\theta}) = h(-\boldsymbol{\theta})$], then $\mathrm{ASD}(\boldsymbol{\theta}) = \mathrm{ASD}(-\boldsymbol{\theta})$. It is clearly so because for an observation $\mathbf{W}$ from such an $H(\cdot)$ the random variables $\mathbf{W}$ and $-\mathbf{W}$ have the same distribution. The induced distribution $H_{\boldsymbol{\theta}}(\cdot)$ is symmetric about the origin O on the tangent plane $P_{\boldsymbol{\theta}}$. As a result [see Theorem 4 of Liu (1990)],

$$\mathrm{SD}_{\boldsymbol{\theta}}(O) = \tfrac{1}{4},$$

and the result follows.

($\Rightarrow$) Suppose $\mathrm{ASD}(\boldsymbol{\theta}) = 1/8$ throughout. Then by Proposition 3.5 we have $\mathrm{SD}_{\boldsymbol{\theta}}(O) = 1/4$ for all $\boldsymbol{\theta}$. However, it was shown in Liu (1990) that if the $\mathrm{SD}(\cdot)$ is equal to $1/4$ on the plane at some point, then the distribution $H_{\boldsymbol{\theta}}(\cdot)$ is symmetric around that point. Since this is true for all $\boldsymbol{\theta}$, the distribution $H(\cdot)$ assigns probability $1/2$ to each hemisphere. Therefore $h(\boldsymbol{\theta}) = h(-\boldsymbol{\theta})$ for all $\boldsymbol{\theta}$. □

The corollaries below follow from Propositions 3.4 and 3.5, and they provide upper bounds for $\mathrm{ASD}(\cdot)$ on the circle and on the sphere, respectively.

COROLLARY 3.2 (On the circle). *Under the conditions of Proposition* 3.4, $\mathrm{ASD}(\theta) \le 1/2$ *for every θ on the circle. The equality holds at a point θ_0 if and only if the entire probability mass is on a semicircle and $H(\theta_0) = 1/2$.*

The claim is based on the fact that $\mathrm{SD}_\theta(\varphi) \le 1/2$, which holds since $\mathrm{SD}_\theta(\varphi) = 2H_\theta(\varphi)[1 - H_\theta(\varphi)]$.

COROLLARY 3.3 (On the sphere). *Under the conditions of Proposition* 3.5, $\mathrm{ASD}(\boldsymbol{\theta}) \le 1/4$ *for any* $\boldsymbol{\theta}$ *on the sphere. The equality holds at a point* $\boldsymbol{\theta}_0$ *if the entire distribution is concentrated on a hemisphere containing* $\boldsymbol{\theta}_0$ *and the induced distribution* $H_{\boldsymbol{\theta}_0}(\cdot)$ *is symmetric around the origin.*

Corollary 3.3 is obtained by combining Proposition 3.5 with Theorem 4 of Liu (1990).

Statistical applications of Propositions 3.3 *and* 3.6. The result of Proposition 3.3 (Proposition 3.6) gives rise to a simple test of whether a given circular (spherical) distribution is antipodally symmetric about the center of the circle (sphere). We may use

$$\sup_{\theta} \left| \mathrm{ASD}_n(\theta) - \tfrac{1}{4} \right| \tag{3.7}$$

on the circle and

$$\sup_{\theta} \left| \mathrm{ASD}_n(\boldsymbol{\theta}) - \tfrac{1}{8} \right| \tag{3.8}$$

on the sphere as test statistics. Large values of (3.7) and (3.8) indicate that the distribution is unlikely to be antipodally symmetric. Needless to say, the actual implementation of these testing ideas would require knowledge of the exact or approximate sampling distributions of these test statistics. Perhaps an approximation of the sampling distribution can be obtained from some resampling procedures, for example, the bootstrap method. A different method has been suggested by Fisher (1989) for obtaining the sampling distribution under the null hypothesis of antipodal symmetry: Use random reflection (through the origin) of the original data set to produce 2^n new samples (each of size n) and compute 2^n values for test statistic (3.7) [or (3.8)]. The histogram based on these 2^n values is an approximation of the desired sampling distribution. A detailed study of the proposed tests (3.7) and (3.8) and their comparison with Ajne's test [Ajne (1968)] shall be reported elsewhere.

4. Properties of angular Tukey's depth. It may be instructive to consider first the case where the underlying distribution H is supported only on a semicircle (hemisphere). In this case we can easily relate $\mathrm{ATD}(\cdot)$ to Tukey's depth on the line (plane).

PROPOSITION 4.1. *If* $H(\cdot)$ *is a distribution concentrated on a semicircle, say from* 0 *to* π, *then* $\mathrm{ATD}_H(\cdot)$ *assumes the familiar form of Tukey's depth on the line, namely, for* $0 \le \theta \le 2\pi$,

$$\mathrm{ATD}_H(\theta) = \min\{H(\theta), 1 - H(\theta)\}. \tag{4.1}$$

Clearly $\mathrm{ATD}_H(\cdot)$ *vanishes outside the interval* $[0, \pi]$. *Furthermore*,

$$\mathrm{ATD}_n(\theta) = \min\{H_n(\theta), 1 - H_n(\theta -)\}.$$

Consider the center-outward ranking of data points based on decreasing $\mathrm{ATD}_n(\cdot)$ values. Proposition 4.1 implies that the above ranking coincides with the center-outward ranking based upon the usual order statistics on the line, if the distribution is supported on a semicircle.

In the spherical case we assume that $H(\cdot)$ is supported on a hemisphere S_0. Given any point $\boldsymbol{\theta}$ on S_0, we consider the stereographic projection, with pole at $-\theta$, from the sphere to the tangent plane at $\boldsymbol{\theta}$. The distribution $H(\cdot)$ on the sphere then induces a distribution on the tangent plane which we denote by $H_{\boldsymbol{\theta}}(\cdot)$. Now, given any hemisphere S containing $\boldsymbol{\theta}$ in its interior we can find another hemisphere, say S', containing $\boldsymbol{\theta}$ in its boundary satisfying $P(S') \leq P(S)$. We can visualize S' as follows. Let L be the line of intersection between the planes supporting the boundaries of S and that of S_0. Rotate the boundary of S around L as axis until it passes through $\boldsymbol{\theta}$. One of the two hemispheres thus obtained will have probability less than or equal to $P(S)$, and this is the one we take as S'. This implies the following proposition.

PROPOSITION 4.2. *Let* $H(\cdot)$ *be a distribution supported on the hemisphere* S_0. *Then the following hold*: (a) $\mathrm{ATD}_H(\boldsymbol{\theta}) = 0$ *for all* $\boldsymbol{\theta} \notin S_0$. (b) *For any* $\boldsymbol{\theta}$ *in* S_0, $\mathrm{ATD}_H(\boldsymbol{\theta})$ *agrees with Tukey's depth* (2.5) *taken w.r.t. the distribution* $H_{\boldsymbol{\theta}}(\cdot)$ *induced on the tangent plane at* θ *by stereographic projection from* $-\boldsymbol{\theta}$.

We discuss now some more general properties of $\mathrm{ATD}(\cdot)$.

PROPOSITION 4.3. *On the circle as well as on the sphere*, $\mathrm{ATD}(\cdot)$ *is bounded above by* $1/2$. *The value* $1/2$ *is achieved at a point* θ *on a circle* ($\boldsymbol{\theta}$ *on a sphere*) *if and only if each semicircle* (*hemisphere*) *containing* θ ($\boldsymbol{\theta}$) *has probability greater than or equal to* $1/2$.

In particular, the bound $1/2$ is achieved at the mode of any member of the von Mises class of distributions, and everywhere if the distribution is uniform.

PROPOSITION 4.4. *On the circle* (*sphere*) $\mathrm{ATD}(\cdot)$ *has the constant value* $1/2$ *throughout if and only if any semicircle* (*hemisphere*) *has probability* $1/2$.

Since the property that each semicircle (hemisphere) has probability $1/2$ can be viewed as an alternative definition of an antipodally symmetric distribution, Proposition 4.4 is a characterization of antipodally symmetric distributions.

These observations may suggest that a distribution with constant ATD would have to be antipodally symmetric and that the maximum ATD for any distribution is always $1/2$. Neither statement is true, as is shown in the following example.

EXAMPLE 4.1. Let $\varepsilon = 1/28$, and let $H(\cdot)$ be a distribution on the unit circle with the density function $h(\cdot)$ defined as

$$h(\theta) = \begin{cases} (\frac{1}{4} + 3\varepsilon)(\pi/2)^{-1}, & \text{for } 0 < \theta \le \frac{1}{2}\pi, \\ \varepsilon(\pi/4)^{-1}, & \text{for } \frac{1}{2}\pi < \theta \le \frac{3}{4}\pi, \\ (\frac{1}{4} - 2\varepsilon)(\pi/4)^{-1}, & \text{for } \frac{3}{4}\pi < \theta \le \pi, \\ (\frac{1}{4} - \varepsilon)(\pi/2)^{-1}, & \text{for } \pi < \theta \le \frac{3}{2}\pi, \\ (\frac{1}{4} - 2\varepsilon)(\pi/4)^{-1}, & \text{for } \frac{3}{2}\pi < \theta \le \frac{7}{4}\pi, \\ \varepsilon(\pi/4)^{-1}, & \text{for } \frac{7}{4}\pi < \theta \le 2\pi. \end{cases}$$

For this distribution $\mathrm{ATD}_H(\theta) = 1/2 - 1/14$ for all θ, $0 < \theta \le 2\pi$. To check this we note that each one of these three semicircles: from $\pi/2$ to $(3/2)\pi$, π to 2π and $(7/4)\pi$ to $(3/4)\pi$ has probability $1/2 - 2\varepsilon$. This also turns out to be the minimum probability over all semicircles.

This example contrasts with the following fact related to ASD: The ASD is constant on a circle if and only if the distribution is antipodally symmetric and the value of the constant is always 1/4 (cf. Proposition 3.3).

We now state the key monotonicity property of ATD($\cdot$).

PROPOSITION 4.5. *Let $\boldsymbol{\theta}_0$ be a point on the sphere. We introduce for each point $\boldsymbol{\theta}$ the Euler angle ϕ, $0 \le \phi \le \pi$, which is the angle between $\mathbf{O}\boldsymbol{\theta}_0$ and $\mathbf{O}\boldsymbol{\theta}$, in other words the latitude of $\boldsymbol{\theta}$. If we fix a meridian, the position of $\boldsymbol{\theta}$ will be characterized by ϕ and its longitude η, $0 \le \eta \le 2\pi$, and a density h on the sphere is just a function of (ϕ, η). Assume we have a distribution with density $h(\phi, \eta)$ which decreases in ϕ for each η, and satisfies $h(\phi, \eta) = h(\phi, \eta + \pi)$. Then*

$$\mathrm{ATD}_H(\boldsymbol{\theta}) = \mathrm{ATD}_H((\phi, \eta))$$

is a monotonically nonincreasing function of ϕ for each η. In particular, it attains its maximum 1/2 at $\boldsymbol{\theta}_0$.

PROOF. The argument relies upon the following observation: If a hemisphere contains $\boldsymbol{\theta}_0$ then it has probability greater than or equal to 1/2; if it does not then it has probability less than or equal to 1/2.

Fix a longitude η. Consider two latitudes ϕ_1 and ϕ_2 such that $0 < \phi_1 < \phi_2 \le \pi$. We claim that

$$\mathrm{ATD}_H((\phi_1, \eta)) \le \mathrm{ATD}_H((\phi_2, \eta)).$$

To show this we consider any hemisphere S^+ such that it contains (ϕ_1, η) but not (ϕ_2, η). Let S^- represent the complement which obviously contains (ϕ_2, η). Evidently S^+ also contains the mode $\boldsymbol{\theta}_0$. In view of the above observation we obtain $P(S^+) \ge P(S^-)$. The proposition follows. □

We omit the discussion of the similar monotonicity property in the circle case.

Next, we point out a simple but peculiar property of ATD, which also contrasts with the general strict monotonicity of ASD.

PROPOSITION 4.6. *For any distribution on the sphere (circle), there exists at least one hemisphere (semicircle) where the* ATD($\cdot$) *is constant and the value of the constant is its minimum value.*

In fact if we let $\tilde{S}$ be a hemisphere with the smallest probability, then ATD($\boldsymbol{\theta}$) will be equal to $P(\tilde{S})$ for any point $\boldsymbol{\theta}$ in $\tilde{S}$. This observation will be useful for the comparisons in Section 5.

Evidently, Proposition 4.6 applies to the empirical version of ATD also. This property of ATD suggests a natural trimming of angular data, namely, to trim off all the data points with the minimum ATD.

Finally, we state without the proof the connection between ATD and median axis in the next proposition, which can be seen as the counterpart of Proposition 3.2 for ATD.

PROPOSITION 4.7. *Proposition* 3.2 *holds with ASD replaced by ATD.*

5. Aspects of robustness. The notion of breakdown described in Hampel, Ronchetti, Rousseeuw and Stahel (1986) is an intuitive measure of robustness in $\mathbb{R}^p$. The following definition is a natural adaption of breakdown for directional functions. Under this definition the breakdown is nonzero for all three medians.

DEFINITION. Let H be a distribution on S_d with θ_0 as a median. We define the *breakdown* of this median as the infimum of ε such that the median of $H_\varepsilon = (1 - \varepsilon)H + \varepsilon G$ is $-\theta_0$ for some contaminating distribution G.

Under this definition the following proposition establishes some lower bounds of the breakdown for the three medians.

PROPOSITION 5.1. *On a circle, for any distributions H and G, we have*

$$\text{(i)} \qquad \left|\mathrm{ASD}_{H_\varepsilon}(\theta) - \mathrm{ASD}_H(\theta)\right| \le 2\varepsilon,$$

$$\text{(ii)} \qquad \left|\mathrm{ATD}_{H_\varepsilon}(\theta) - \mathrm{ATD}_H(\theta)\right| \le \varepsilon$$

and

$$\text{(iii)} \qquad \left|\mathrm{ADD}_{H_\varepsilon}(\theta) - \mathrm{ADD}_H(\theta)\right| \le \varepsilon\pi,$$

where $H_\varepsilon = (1 - \varepsilon)H + \varepsilon G$.

The inequality (i) implies the following: if θ_0 is an ASM on the circle and the depth of θ_0 is strictly higher than that of $-\theta_0$, then the breakdown of this median is nonzero. The same statement applies to ATM and ADM. In fact, Proposition 5.1 implies

$$\text{breakdown of ASM} \geq \frac{\mathrm{ASD}(\theta_0) - \mathrm{ASD}(-\theta_0)}{4},$$

$$\text{breakdown of ATM} \geq \frac{\mathrm{ATD}(\theta_0) - \mathrm{ATD}(-\theta_0)}{2}$$

and

$$\text{breakdown of ADM} \geq \frac{\mathrm{ADD}(\theta_0) - \mathrm{ADD}(-\theta_0)}{2\pi}.$$

PROOF. (i) Let $\{W_1, W_2\}$ be a random sample from H; $\{Z_1, Z_2\}$ from G; and $\{\eta_1, \eta_2\}$ from a Bernoulli distribution with $P(\eta_i = 1) = 1 - \varepsilon$ and $P(\eta_i = 0) = \varepsilon$. We assume that the W_i's, Z_i's and η_i's are all independent random variables. Define

$$W_i^* = \begin{cases} W_i, & \text{if } \eta_i = 1, \\ Z_i, & \text{if } \eta_i = 0. \end{cases}$$

We obtain

$$\begin{aligned} \mathrm{ASD}_{H_\varepsilon}(\theta) &= P(\theta \in \mathrm{arc}(W_1^*, W_2^*)) \\ &= P(\theta \in \mathrm{arc}(W_1^*, W_2^*) \cap (\eta_1 = 1, \eta_2 = 1)) \\ &\quad + P\big(\theta \in \mathrm{arc}(W_1^*, W_2^*) \cap (\eta_1 = 1, \eta_2 = 1)^c\big) \\ &= (1 - \varepsilon)^2 \mathrm{ASD}_H(\theta) + R, \end{aligned}$$

where $0 \leq R \leq 2\varepsilon$. The result follows from the fact that $1 - (1 - \varepsilon)^2 \leq 2\varepsilon$.

(ii) The inequality on ATD is easily deduced from the observation that for any semicircle S, $|P_{H_\varepsilon}(S) - P_H(S)| \leq \varepsilon$.

The proof of (iii) is straightforward and is thus omitted. □

REMARK 5.1. In the case of a sphere, the bound for ATD and ADD in Proposition 5.1 remains the same, and it becomes 3ε for ASD.

Besides the notion of breakdown, the influence function is another commonly used tool for the study of robustness. See Hampel, Ronchetti, Rousseeuw and Stahel (1986) for the description of the influence function of a statistic. Since the influence functions of most circular location estimators are bounded, Ko and Guttorp (1988) proposed to divide the influence function by a measure of scale and then take the supremum over the circle and a reasonable class of distributions. If the supremum is bounded, then the estimator is considered to be *standardized-bias robust* (SB-robust). Ko and Guttorp (1988) show that the directional mean is not SB-robust; however, on the von Mises class of distributions, Mardia's median (or ADD) is SB-robust. For a symmetric unimodal circular distribution with the modal angle μ_0 the influence function (IF) of Mardia's median is

$$\text{IF}(\theta; \text{Mardia's median}) = \frac{1}{2}\frac{\text{sign}(\theta - \mu_0)}{f(\mu_0) - f(-\mu_0)}, \tag{5.1}$$

where $\text{sign}(x) = 1$, 0 or -1 as $x > 0$, $x = 0$ or $x < 0$. The result (5.1) is obtained in Wehrly and Shine (1981). It is easy to show that the three medians ASM, ATM and ADM, when they are uniquely defined, all have the same form of influence function as (5.1). Thus they are typically SB-robust. In view of Propositions 3.2 and 4.7, the three medians may be uniquely defined even if there are multiple median axes and multiple modes (see also Example 6.1).

It is not surprising that all the circular medians discussed here give the same influence function. After all, the Euclidean versions of the three depth functions in this article give rise to the same median on the real line. Clearly these circular medians coincide under unimodal symmetric distributions. However, outside this class of distributions it is not known to us if they still coincide as they do on the real line.

Finally, we demonstrate through an example a comparison related to a qualitative robustness aspect of the center-outward ranking of data points based on the decreasing $\text{ADD}_n(\cdot)$ and $\text{ASD}_n(\cdot)$ [or $\text{ATD}_n(\cdot)$] value.

EXAMPLE 5.1 (On a unit circle). Let the data set be $0, 5, 10, 15, 100, 105, 110, 115, 120$, (in degrees). The center-outward ADD ordering is $100, 105, 110, 115, 120, 15, 5, 0$. However, if the data point 100 in the data set is replaced by 20, then the new ordering will be 20, 15, 10, 5, 0, 105, 110, 115, 120. Note that the ordering is drastically altered even though only one data point is changed. In fact, either of the two data situations is likely to arise if "both heaps" of the underlying bimodal distribution have 50% probability. This drastic alteration in ordering is due to the fact that ADD ordering depends on distance and is clearly absent in the case of the ordering based on $\text{ASD}_n(\cdot)$ or $\text{ATD}_n(\cdot)$ [which are the same as the ordering based on the usual order statistics on the line, namely, 100, (15, 105), (10, 110), (5, 115), (0, 120)].

6. Concluding remarks.

REMARK 6.1. The achievable upper bound for ASD is $1/2$ and $1/4$ in the circular case and the spherical case, respectively. For ATD it is $1/2$ in both cases. For antipodally symmetric distributions, ASD equals $1/4$ throughout the circle, $1/8$ throughout the sphere, while ATD is $1/2$ throughout a circle as well as a sphere. In general the upper bound for ATD is more likely to be attained than that of ASD.

REMARK 6.2. Consider the densities which are monotonically decreasing from θ_0 to $-\theta_0$ in a symmetric way (e.g., the von Mises class). In such a case, ASD decreases monotonically from θ_0 to $-\theta_0$ in each direction; whereas ATD decreases up to halfway in each direction and becomes a constant afterwards. The value of ATD at the mode of a unimodal distribution equals $1/2$, whereas that for ASD at the mode (on a circle) ranges from $1/4$ to $1/2$. The value $1/2$ for ASD at the mode occurs if and only if the entire distribution is concentrated on a semicircle. Thus we note that the *maximum value of ATD at the mode is the same as its constant value in the case of the uniform distribution* ($= 1/2$), while *the notion of ASD makes a clear distinction between the two situations*.

REMARK 6.3. For symmetric unimodal distributions (e.g., the von Mises class), ADM, ASM and ATM all coincide with the mode. The example given below further shows that ASD, ATD and ADD may help one pick out the "centralmost" point in the presence of multiple median axes and multiple modes.

EXAMPLE 6.1. Let $h(\cdot)$ be a density function on the unit circle defined as follows:

$$h(\theta) = \begin{cases} \dfrac{\theta}{\pi/2}\dfrac{1}{\pi}, & \text{for } 0 \le \theta \le \dfrac{\pi}{2}, \\ \dfrac{\pi - \theta}{\pi/2}\dfrac{1}{\pi}, & \text{for } \dfrac{\pi}{2} < \theta \le \pi, \\ \dfrac{1}{2\pi}, & \text{for } \pi < \theta < 2\pi. \end{cases}$$

In other words, the distribution is triangular on $[0, \pi]$ and uniform on $(\pi, 2\pi)$. It is easy to check that there are two perpendicular median axes along the two axes. On the other hand, there is a unique maximum point of ASD($\cdot$), namely, the point $\pi/2$. This seems more sensible because among the four median candidates 0, $\pi/2$, π and $3\pi/2$ suggested by median axes, $\pi/2$ stands out as the point with the highest probability concentration *around* it. The claim of unique maximization at $\pi/2$ can be verified by using Proposition 3.1 and the following fact:

$$A_\theta - C_\theta > 0 \quad \text{for } \theta \in \left(0, \frac{\pi}{2}\right) \cup \left(\frac{3\pi}{2}, 2\pi\right)$$

and

$$A_\theta - C_\theta < 0 \quad \text{for } \theta \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right).$$

Thus, ASD and ADD decrease monotonically in the ranges $\pi/2$ to $3\pi/2$ clockwise as well as $\pi/2$ to $3\pi/2$ counterclockwise. Similarly, ATD decreases monotonically on both sides of $\pi/2$ with strict monotonicity between $\pi/2$ and $\pi/2 \pm \pi/4$. Beyond this range, ATD stays constant and assumes its minimum value.

Now, in this example a new mode can actually be created at π or 0 by altering the density locally, thus creating another mode keeping the maximum point of ASD($\cdot$), ATD($\cdot$), ADD($\cdot$) and the two median axes unaffected.

REMARK 6.4. In comparing the rankings derived from three depths, we note that the ranking based on ADD is not even consistent with the linear ranking when the distribution is on a half-circle (cf. Example 5.1). As for ASD and ATD, even though they both satisfy this consistency property, ATD is unable to distinguish all points lying on the hemisphere with the smallest probability (cf. Proposition 4.6 and Remark 6.2). In conclusion, *ASD seems to provide a finer ranking of data points in the order of centrality*, which is particularly useful in detecting outliers [see Collett (1980) for a discussion on outliers in circular data]. On the other hand, ATD may be expected to be superior in terms of the robustness of the associated "center." Another advantage of ASD is that, in general, ASD seems easier to compute than ATD, especially for S_d, $d \geq 2$.

REFERENCES

AJNE, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika* **55** 343–354.

BROWN, B. M. (1983). Statistical uses of the spatial median. *J. Roy. Statist. Soc. Ser. B* **45** 25–30.

COLLETT, D. (1980). Outliers in circular data. *J. Roy. Statist. Soc. Ser. C* **29** 50–57.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

FISHER, N. I. (1985). Spherical medians. *J. Roy. Statist. Soc. Ser. B* **47** 342–348.

FISHER, N. I. (1989). Personal communication.

FISHER, N. I., LEWIS, T. and EMBLETON, B. J. J. (1987). *Statistical Analysis of Spherical Data*. Cambridge Univ. Press.

GOWER, J. C. (1974). The median centre. *J. Roy. Statist. Soc. Ser. C* **23** 466–470.

GROSS, S. and LIU, R. (1989). Classification rules based on concepts of data depth. Unpublished manuscript.

HAMPEL, F. R., RONCHETTI, E. M., ROUSSEEUW, P. J. and STAHEL, W. A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.

HE, X. and SIMPSON, D. G. (1992). Robust direction estimation. *Ann. Statist.* **20** 351–369.

KO, D. and GUTTORP, P. (1988). Robustness of estimators for directional data. *Ann. Statist.* **16** 609–618.

LIU, R. (1988). On a notion of simplicial depth. *Proc. Nat. Acad. Sci. U.S.A.* **85** 1732–1734.

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

MARDIA, K. V. (1972). *Statistics of Directional Data*. Academic, New York.

SMALL, C. G. (1987). Measures of centrality for multivariate and directional distribution. *Canad. J. Statist.* **5** 31–39.

TUKEY, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians, Vancouver, 1974* (R. D. James, ed.) **2** 523–531. Canadian Mathematical Congress.

WATSON, G. S. (1983). *Statistics on Spheres*. Wiley, New York.

WEHRLY, T. E. and SHINE, E. P. (1981). Influence curves of estimators for directional data. *Biometrika* **68** 334–335.

DEPARTMENT OF STATISTICS
HILL CENTER, BUSCH CAMPUS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903