Journal of Nonparametric Statistics
Vol. 00, No. 00, January 2008, 1–20

RESEARCH ARTICLE

Influence functions of some depth functions, and application to depth-weighted L-statistics

Xin Dang[a*], Robert Serfling[b] and Weihua Zhou[c]

[a]*Department of Mathematics, University of Mississippi, University, MS 38677, USA*; [b]*Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75083-0688, USA*; [c]*Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA*

(*received 1 June 2007; final version received 24 March 2008*)

Depth functions are increasingly being used in building nonparametric outlier detectors and in constructing useful nonparametric statistics such as depth-weighted L-statistics. Robustness of a depth function is an essential property for such applications. Here, robustness of three key depth functions, the spatial, simplicial, and generalized Tukey, is explored via the influence function (IF) approach. For all three depths, the IFs are derived and found to be bounded, an important robustness property, and are applied to evaluate two other robustness features, gross error sensitivity and local shift sensitivity. These IFs are also used as components of the IFs of associated depth-weighted L-statistics, for which through a standard approach consistency and asymptotic normality are then derived. In turn, the asymptotic normality is applied to obtain asymptotic relative efficiencies. For spatial depth, two forms of weight function suggested in the recent literature are considered and asymptotic relative efficiencies in comparison with the mean obtained. For all three depths and one of these weight functions, finite sample relative efficiencies are obtained by simulation under normal, contaminated normal, and heavy-tailed t distributions. As a technical tool of general interest, needed here with the simplicial depth, the IF of a general U-statistic is derived.

Keywords: depth function; influence function; L-statistic; relative efficiency; robustness; U-statistic

AMS Subject Classification: Primary 62G99; Secondary 60F15

1. Introduction

Depth functions are playing increasingly important roles in robust and nonparametric multivariate analysis and inference. Compensating for lack of linear order in dimensions higher than one, they provide center-outward orderings for points in $\mathbb{R}^d$ with respect to a distribution or a data cloud. Popular types include halfspace depth (Tukey, 1975), simplicial depth (Liu, 1988, 1990), projection depth (Liu ,1992, Zuo and Serfling, 2000a,b, and Zuo, 2003), spatial depth (Vardi and Zhang, 2000), and a generalized halfspace depth (Zhang, 2002).

A desirable and necessary property of depth functions and especially of depth-based estimators or outlier detectors is robustness. Key quantitative measures of robustness – boundedess, gross error sensitivity, and local shift sensitivity – may be investigated using the influence function (IF) approach (Hampel, 1968, 1974).

*Corresponding author. Email: xdang@olemiss.edu

ISSN: 1048-5252 print/ISSN 1029-0311 online

DOI: 10.1080/1048525YYxxxxxxxx
http://www.informaworld.com

In this paper we develop and study the IFs of the spatial, simplicial, and generalized Tukey depths. Previously, the IF of Tukey depth has been investigated by Romanazzi (2001) and, in a more general setting, by Wang and Serfling (2006), and the IF of projection depth has been studied in Zuo, Cui, and Young (2004).

A direct application of depth functions is to induce location estimators such as depth-weighted L-statistics (Liu, 1990, and Liu, Parelius, and Singh, 1999). Asymptotics of such "DL-statistics" are developed by Zuo, Cui, and He (2004), who provide sufficient conditions under which such estimators are asymptotically normal and focus on projection DL-statistics including the Stahel-Donoho estimator as a special case. In the present paper, specializing their conditions to the spatial, simplicial, and generalized Tukey depths, we show by an IF approach similar to theirs that the corresponding DL-statistics have limiting normal distributions.

Section 2 develops the IFs for the spatial, simplicial, and generalized Tukey depths and provides sample representations, comparisons, and illustrations. As a technical tool of general interest, needed here with the simplicial depth, the IF of a general U-statistic is derived. Section 3 treats corresponding DL-statistics, establishing consistency and asymptotic normality of sample versions, and develops asymptotic relative efficiencies. The Appendix provides selected proofs.

2. IFs of selected depth functions

The IF introduced by Hampel (1968, 1974) is now a standard heuristic tool in robust statistics for measuring effects on estimators due to infinitesimal perturbations of sample distribution functions. For a cdf F on $\mathbb{R}^d$ and a functional $T : F \mapsto T(F) \in \mathbb{R}^m$ with $m \geq 1$, the IF of T at F may be expressed as

$$\mathrm{IF}(\boldsymbol{z}; T, F) = \lim_{\epsilon \downarrow 0} \frac{T((1-\epsilon)F + \epsilon\delta_{\boldsymbol{z}}) - T(F)}{\epsilon}, \quad \boldsymbol{z} \in \mathbb{R}^d,$$

where $\delta_{\boldsymbol{z}}$ denotes the point mass distribution at $\boldsymbol{z}$. Under regularity conditions on T (see Hampel *et al.*, 1986, or Serfling, 1980, for details), we have $E_F\{\mathrm{IF}(\boldsymbol{X}; T, F)\} = 0$ and the von Mises expansion

$$T(F_n) - T(F) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{IF}(\boldsymbol{X}_i; T, F) + o_p(n^{-1/2}), \tag{1}$$

where F_n denotes the sample df based on sample $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n$. This representation shows the connection of the IF with robustness of T, observation by observation, and the relevance of measures such as *gross error sensitivity*

$$\gamma^* = \sup_{\boldsymbol{z}} |\mathrm{IF}(\boldsymbol{z}; T, F)|$$

and *local shift sensitivity*

$$\lambda^* = \sup_{\boldsymbol{x} \neq \boldsymbol{y}} \frac{\|\mathrm{IF}(\boldsymbol{x}; T, F) - \mathrm{IF}(\boldsymbol{y}; T, F)\|}{\|\boldsymbol{x} - \boldsymbol{y}\|}.$$

Further, (1) yields asymptotic m-variate normality of $T(F_n)$,

$$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{d} N(\mathbf{0}, E_F(\mathrm{IF}(\boldsymbol{X}; T, F)\mathrm{IF}(\boldsymbol{X}; T, F)^T). \tag{2}$$

2.1. *The IF of spatial depth*

The spatial depth (Vardi and Zhang, 2000) is closely related to spatial quantiles, introduced by Dudley and Koltchinskii (1992) and Chaudhuri (1996) and further developed by Serfling (2002, 2004). The *spatial depth* of $\boldsymbol{x}$ with respect to a given cdf F for random variable $\boldsymbol{X}$ may be defined as

$$D_s(\boldsymbol{x}, F) = 1 - \|E_F(S(\boldsymbol{x} - \boldsymbol{X}))\|,$$

where

$$S(\boldsymbol{x}) = \begin{cases} \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}, & \boldsymbol{x} \in \mathbb{R}^d, \boldsymbol{x} \neq \boldsymbol{0}, \\ \boldsymbol{0}, & \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

is the vector *sign function* in $\mathbb{R}^d$. The point of maximal spatial depth coincides with the so-called "*spatial median*", which has a long history (Small, 1990). The corresponding *sample spatial depth* is $D_s(\boldsymbol{x}, F_n) = 1 - \left\|n^{-1} \sum_{i=1}^n S(\boldsymbol{x} - \boldsymbol{X}_i)\right\|$.

Theorem 2.1: *Let $\boldsymbol{\mu}_F$ be the spatial median of F.*
(i) If $\boldsymbol{x} \neq \boldsymbol{\mu}_F$, and $\boldsymbol{z} \neq \boldsymbol{x}$, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 1 - D_s(\boldsymbol{x}, F) - \cos\theta$, *with θ being the angle between $\boldsymbol{u}_{\boldsymbol{x}}$ and $\boldsymbol{z}_{\boldsymbol{x}}$, where $\boldsymbol{u}_{\boldsymbol{x}} = E_F \frac{\boldsymbol{x}-\boldsymbol{X}}{\|\boldsymbol{x}-\boldsymbol{X}\|}$ and $\boldsymbol{z}_{\boldsymbol{x}} = \frac{\boldsymbol{x}-\boldsymbol{z}}{\|\boldsymbol{x}-\boldsymbol{z}\|}$.*
(ii) If $\boldsymbol{x} \neq \boldsymbol{\mu}_F$, and $\boldsymbol{z} = \boldsymbol{x}$, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 1 - D_s(\boldsymbol{x}, F)$.
(iii) If $\boldsymbol{x} = \boldsymbol{\mu}_F$, and $\boldsymbol{z} \neq \boldsymbol{x}$, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = -1$.
(iv) If $\boldsymbol{z} = \boldsymbol{x} = \boldsymbol{\mu}_F$, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 0$.

Note that:

(1) $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ is bounded, with $-1 \leq \mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) \leq 2$ for $\forall\, \boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^d$.
(2) For Case (i), $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ is a function of θ and $D_s(\boldsymbol{x}, F)$. If $\boldsymbol{x}$ is fixed, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ is determined by the position of $\boldsymbol{z}$ in the sense of the angle between $\boldsymbol{z} - \boldsymbol{x}$ and $\boldsymbol{u}_{\boldsymbol{x}}$. Some special cases are:
   a) If $\boldsymbol{z}_{\boldsymbol{x}} = -\boldsymbol{u}_{\boldsymbol{x}}$, then $\theta = \pi$, and $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 2 - D_s(\boldsymbol{x}, F)$, which reaches the upper bound for $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ and is always positive.
   b) If $\boldsymbol{z}_{\boldsymbol{x}} = \boldsymbol{u}_{\boldsymbol{x}}$, then $\theta = 0$, and $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = -D_s(\boldsymbol{x}, F)$, which attains the lower bound and is always negative. When additional infinitesimal mass is placed along the direction $\boldsymbol{u}_{\boldsymbol{x}}$, the centrality of the point $\boldsymbol{x}$ is reduced.
   c) If $\boldsymbol{z}_{\boldsymbol{x}} \perp \boldsymbol{u}_{\boldsymbol{x}}$, then $\theta = \pi/2$, and $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 1 - D_s(\boldsymbol{x}, F)$, which is always positive.

   If θ is fixed, then $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ is a decreasing function of $D_s(\boldsymbol{x}, F)$, so that the greater the outlyingness, the greater the influence of contamination at the point $\boldsymbol{z}$.
(3) For Case (ii), $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = 1 - D_s(\boldsymbol{x}, F) > 0$. Putting additional infinitesimal mass at point $\boldsymbol{z} = \boldsymbol{x}$ has a positive influence on its depth, that is, the degree of centrality of $\boldsymbol{z} = \boldsymbol{x}$ increases.
(4) For Cases (iii) and (iv), $\boldsymbol{x} = \boldsymbol{\mu}_F$, and contamination on any other point will have the same negative influence on the depth of the median, that is, the degree of centrality of the spatial median will decrease. Additional infinitesimal mass at the spatial median has no influence on the depth of the spatial median.
(5) Although $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F)$ is a continuous function of θ, it is not continuous in $\boldsymbol{z}$, and hence the local shift sensitivity of $D_s(\boldsymbol{x}, F)$ is infinity.

A related but quite different result is the IF of the spatial quantile function, given by Koltchinskii (1997).

In deriving asymptotic normality of spatial DL-statistics, we will apply asymptotic normality of the sample spatial depth. This follows via the approach of (1) and (2) via

$$D_s(\boldsymbol{x}, F_n) - D_s(\boldsymbol{x}, F) = n^{-1} \sum_{i=1}^{n} \mathrm{IF}(\boldsymbol{X}_i; D_s(\boldsymbol{x}, F), F) + R_{1n},$$

provided that the remainder term R_{1n} satisfies $\sqrt{n}R_{1n} = o_p(1)$. This indeed holds, as per the following lemma.

Lemma 2.2: *For any subset $K \subset \mathbb{R}^d$,*

$$D_s(\boldsymbol{x}, F_n) - D_s(\boldsymbol{x}, F) = n^{-1} \sum_{i=1}^{n} IF(\boldsymbol{X}_i; D_s(\boldsymbol{x}, F), F) + o_p(n^{-1/2})$$

holds uniformly for $\boldsymbol{x} \in K$.

Comparison with IF of halfspace depth

The IF of halfspace (Tukey) depth at $\boldsymbol{x}$ takes two values, either $1 - D_h(\boldsymbol{x}, F)$ or $-D_h(\boldsymbol{x}, F)$, depending on whether the contaminant at $\boldsymbol{z}$ belongs to a special "optimal" halfspace or not, i.e.,

$$\mathrm{IF}(\boldsymbol{z}; D_h(\boldsymbol{x}, F), F) = -D_h(\boldsymbol{x}, F) + I_{(\boldsymbol{z} \in \mathcal{H}^*_{\boldsymbol{x}})},$$

where $I_{(\cdot)}$ denotes the indicator function, the halfspace depth $D_h(\boldsymbol{x}, F) = \inf_{\mathcal{H}_{\boldsymbol{x}}} P_F(\mathcal{H}_{\boldsymbol{x}})$, and the "optimal" halfspace $\mathcal{H}^*_{\boldsymbol{x}} = \arg\inf_{\mathcal{H}_{\boldsymbol{x}}} P_F(\mathcal{H}_{\boldsymbol{x}})$, with $\mathcal{H}_{\boldsymbol{x}}$ ranging through closed halfspaces containing $\boldsymbol{x}$. (See Romanazzi, 2001, and Wang and Serfling, 2006.) We compare the IFs for spatial and halfspace depths using a heavy tailed distribution, the bivariate Pareto density

$$f(\boldsymbol{x}; \alpha) = \alpha(\alpha + 1)(x_1 + x_2 - 1)^{-(\alpha+2)} \; x_1 > 1, x_2 > 1,$$

where $\alpha > 0$. Depth values and contours of $D(\boldsymbol{x}, F)$ are plotted in Figure 1 and IFs of $D((3, 3), F)$ and $D((3, 5), F)$ in Figure 2. It is seen that both IFs are bounded, limiting the effects of outliers, but each has discontinuities, making the local shift sensitivity infinite. For the Tukey depth, the IF is discontinuous at the boundary separating the optimal and non-optimal halfspaces. For the spatial depth, the IF is discontinuous at two points, $\boldsymbol{z} = (3, 3)$ and $(3, 5)$.

2.2. *The IF of simplicial depth*

The *simplicial depth* (Liu, 1988, 1990) of $\boldsymbol{x}$ with respect to cdf F in $\mathbb{R}^d$ is defined as the probability that $\boldsymbol{x}$ belongs to the closed simplex $\Delta[\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{d+1}]$ in $\mathbb{R}^d$ formed by $d + 1$ random observations $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{d+1}$ from F,

$$D_\Delta(\boldsymbol{x}, F) = P_F\{\boldsymbol{x} \in \Delta[\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{d+1}]\}, \; \boldsymbol{x} \in \mathbb{R}^d.$$

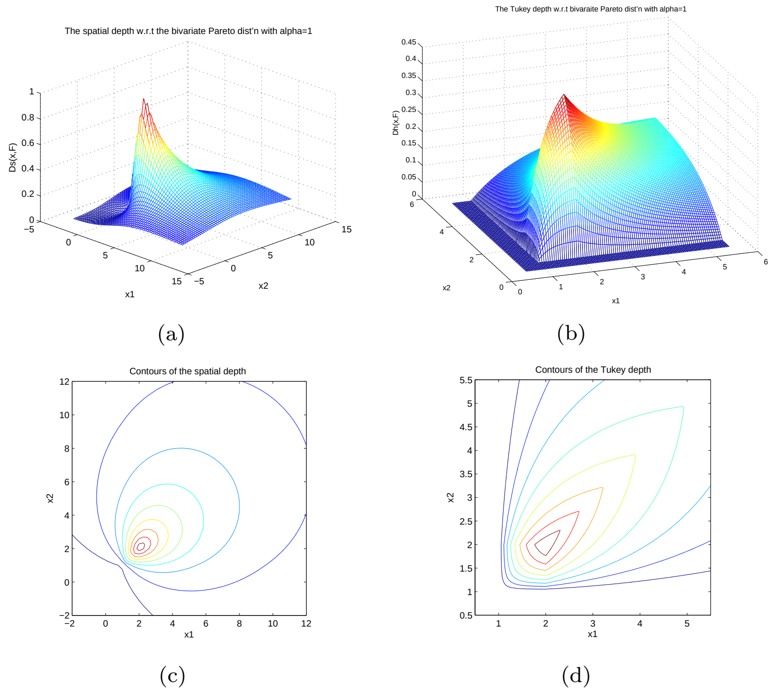


Figure 1. 3D-plots and contour plots of depths w.r.t. the bivariate Pareto distribution (left panel (a) & (c) for the spatial depth; right panel (b) & (d) for the Tukey depth).

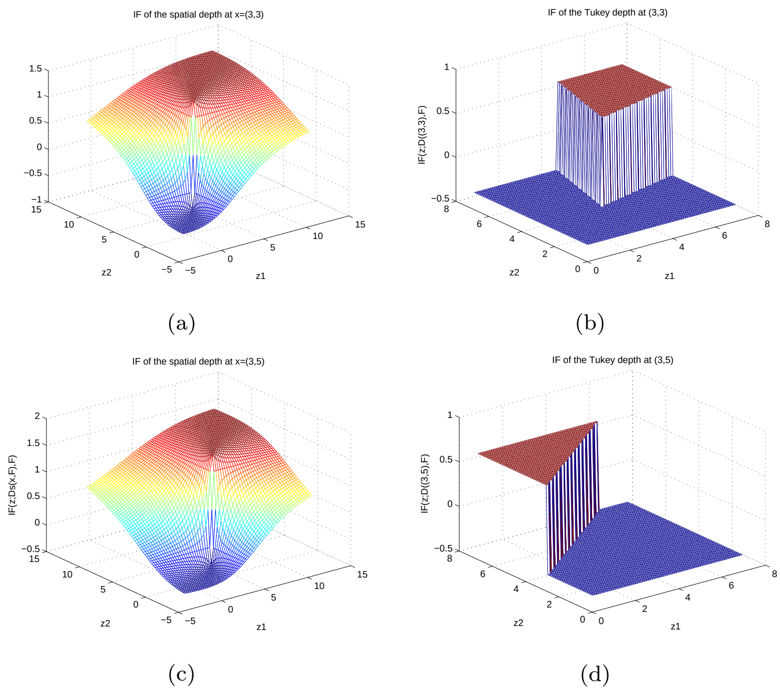


Figure 2. IF of depths at points (3,3) and (3,5) w.r.t. the bivariate Pareto distribution (left panel (a) & (c) for the spatial depth; right panel (b) & (d) for the Tukey depth).

The sample version is the fraction of sample random simplices containing $\boldsymbol{x}$,

$$D_{\Delta}(\boldsymbol{x}, F_n) = \binom{n}{d+1}^{-1} \sum I_{(\boldsymbol{x} \in \Delta[\boldsymbol{X}_{i_1}, \ldots, \boldsymbol{X}_{i_{d+1}}])}, \ \boldsymbol{x} \in \mathbb{R}^d,$$

where the summation is over the $\binom{n}{d+1}$ combinations of $d+1$ distinct elements $\{i_1, ...i_{d+1}\}$ from $\{1, ..., n\}$. We note that $D_\Delta(\boldsymbol{x}, F_n)$ is a U-statistic in form, say U_n, with the kernel function $h_{\boldsymbol{x}}(\boldsymbol{x}_1, ..., \boldsymbol{x}_{d+1}) = I_{(\boldsymbol{x} \in \Delta[\boldsymbol{x}_1, ..., \boldsymbol{x}_{d+1}])}$. Then $D_\Delta(\boldsymbol{x}, F)$ $= E_F h_{\boldsymbol{x}}(\boldsymbol{X}_1, ..., \boldsymbol{X}_{d+1}) = E_F U_n$ and we thus obtain the IF for simplicial depth as a special case of that for a general U-statistic functional, as provided by the following lemma.

Lemma 2.3: *For a kernel $h(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_m)$ symmetric in its arguments and a corresponding U-statistic $U_n(h)$, the* IF *of the functional $T(F) = E_F(U_n(h))$ $= E_F(h(\boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_m))$ is*

$$\mathrm{IF}(\boldsymbol{z}; T(F), F) = mT(F \,|\, \boldsymbol{X_1} = \boldsymbol{z}) - mT(F), \;\; \boldsymbol{z} \in \mathbb{R}^d.$$

A sample version of this IF is the U-statistic

$$\mathrm{IF}(\boldsymbol{z}; U_n(h), F) = m \binom{n}{m-1}^{-1} \sum h(\boldsymbol{z}, \boldsymbol{X}_{i_1}, ..., \boldsymbol{X}_{i_{m-1}}) - mU_n(h),$$

where the summation is over the $\binom{n}{m-1}$ combinations of $m-1$ distinct elements $\{i_1, ...i_{m-1}\}$ from $\{1, ..., n\}$. It is unbiased for $\mathrm{IF}(\boldsymbol{z}; T(F), F)$:

Lemma 2.4: $E_F\{\mathrm{IF}(\boldsymbol{z}; U_n(h), F)\} = \mathrm{IF}(\boldsymbol{z}; T(F), F)$.

The above lemmas immediately yield the IF of simplicial depth:

Theorem 2.5: *For fixed $\boldsymbol{x}$,*

$$\mathrm{IF}(\boldsymbol{z}; D_\Delta(\boldsymbol{x}, F), F) = (d+1) \int ... \int h_{\boldsymbol{x}}(\boldsymbol{z}, \boldsymbol{x}_1, ...\boldsymbol{x}_d) \prod_{i=1}^{d} dF(\boldsymbol{x}_i) - (d+1) D_\Delta(\boldsymbol{x}, F).$$

Boundedness follows since the function inside the integration is bounded.

As for the sample spatial depth, we have asymptotic normality of the sample simplicial depth. Results of Dümbgen (1992) and Arcones and Giné (1993) yield weak convergence of the simplicial depth process, from which we obtain

Lemma 2.6: *For any subset $K \subset \mathbb{R}^d \setminus \{\boldsymbol{x} : D_\Delta(\boldsymbol{x}, F) = 0\}$,*

$$D_\Delta(\boldsymbol{x}, F_n) - D_\Delta(\boldsymbol{x}, F) = n^{-1} \sum_{i=1}^{n} \mathrm{IF}(\boldsymbol{X_i}; D_\Delta(\boldsymbol{x}, F), F) + o_p(n^{-1/2})$$

holds uniformly in K.

2.3. *The IFs of some generalized Tukey depths*

Certain generalized versions of Tukey depth were proposed and studied by Zhang (2002). For a continuous bounded function $g(\cdot)$, the *generalized Tukey depth* of $\boldsymbol{x}$ relative to cdf F for random variable $\boldsymbol{X} \in \mathbb{R}^d$ is defined as

$$D_g(\boldsymbol{x}, F) = [1 + O_g(\boldsymbol{x}, F)]^{-1},$$

in terms of the outlyingness function

$$O_g(\boldsymbol{x}, F) = \sup_{\|\boldsymbol{u}\|=1} E_F g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{X}))$$

with

$$Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{X}) = \frac{\boldsymbol{u}^T(\boldsymbol{x} - \boldsymbol{X})}{\sigma(F_{\boldsymbol{u}})},$$

for $F_{\boldsymbol{u}}$ the distribution of $\boldsymbol{u}^T\boldsymbol{X}$ and $\sigma(\cdot)$ a scale functional. Important special cases include:

(1) Tukey depth: $g(y) = I_{(y \geq 0)}$.
(2) A Huber type generalized Tukey depth: $g(y) = yI_{(|y| \leq c)} + \text{sign}(y)cI_{(|y|>c)}$, with a tuning constant $c > 0$.
(3) A smoothed version of Tukey depth: $g(y) = (1 - e^{-cy})/(1 + e^{-cy})$, with a large tuning constant $c > 0$.

IF of generalized Tukey depth

The general theorem below shows how the IF of generalized Tukey depth depends on the IF of the scale estimator σ and the choice of function g.

Theorem 2.7: *Define* $\boldsymbol{u}_{\boldsymbol{x}} = \arg\sup_{\|\boldsymbol{u}\|=1} E_F g(Y_F(\boldsymbol{u}, \boldsymbol{x}, X))$. *If*
(A1) $\boldsymbol{u}_{\boldsymbol{x}}$ *is a singleton almost surely,*
(A2) the function g is differentiable, and
(A3) $\text{IF}(\boldsymbol{u}_{\boldsymbol{x}}^T\boldsymbol{z}; \sigma, F_{\boldsymbol{u}_{\boldsymbol{x}}})$ *is continuous at* $\boldsymbol{u}_{\boldsymbol{x}}$ *for a given* $\boldsymbol{z}$,
then

$$\text{IF}(\boldsymbol{z}; D_g(\boldsymbol{x}, F), F) = -O_g(\boldsymbol{x}, F) + g(Y_F(\boldsymbol{u}_{\boldsymbol{x}}, \boldsymbol{x}, \boldsymbol{z})) - A\frac{\text{IF}(\boldsymbol{u}_{\boldsymbol{x}}^T\boldsymbol{z}; \sigma, F_{\boldsymbol{u}_{\boldsymbol{x}}})}{\sigma(F_{\boldsymbol{u}_{\boldsymbol{x}}})},$$

where $A = \int g\prime(Y_F(\boldsymbol{u}_{\boldsymbol{x}}, \boldsymbol{x}, X))Y_F(\boldsymbol{u}_{\boldsymbol{x}}, \boldsymbol{x}, X)\, dF(X)$. *Accordingly,*

$$\text{IF}(\boldsymbol{z}; D_g(\boldsymbol{x}, F), F) = -\frac{\text{IF}(\boldsymbol{z}; O_g(\boldsymbol{x}, F), F)}{(1 + O_g(\boldsymbol{x}, F))^2}.$$

For g bounded and the scale estimator σ robust, such as the median of absolute deviations (MAD), then the IF of generalized Tukey depth is bounded.

For an expansion of form (1) for generalized Tukey depth, we need:
(B1) $\sigma(F_{\boldsymbol{u}})$ is continuous in $\boldsymbol{u}$, with $\inf_{\|\boldsymbol{u}\|=1} \sigma(F_{\boldsymbol{u}}) > 0$;
(B2) The representation $\sigma(F_{n\boldsymbol{u}}) - \sigma(F_{\boldsymbol{u}}) = n^{-1}\sum_{i=1}^n \text{IF}(\boldsymbol{u}^T\boldsymbol{X}_i; \sigma, F) + o_p(n^{-1/2})$ holds uniformly in $\boldsymbol{u}$;
(B3) $g(\cdot)$ is differentiable, with $g'(.)$ continuous;
(B4) $E[\sup_{\|\boldsymbol{u}\|=1} |g(\cdot)|^2] < \infty$.

Lemma 2.8: *Assume (B1)-(B4) and suppose that* $\boldsymbol{u}_{\boldsymbol{x}}$ *consists of a singleton for all* $\boldsymbol{x}$ *except finitely many points* $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$. *Then, for each* $\delta > 0$,

$$D_g(\boldsymbol{x}, F_n) - D_g(\boldsymbol{x}, F) = n^{-1}\sum_{i=1}^{n} \text{IF}(\boldsymbol{X}_i; D_g(\boldsymbol{x}, F), F) + o_p(n^{-1/2})$$

holds uniformly for $\boldsymbol{x} \in K = \mathbb{R}^d - \cup_{j=1}^m B(\boldsymbol{y}_j, \delta)$, *with* $B(\boldsymbol{y}, \delta) = \{x \in \mathbb{R}^d : \|\boldsymbol{y} - \boldsymbol{x}\| < \delta\}$.

Illustration and comparison with the projection depth

Both generalized Tukey depth and projection depth are affine invariant projection pursuit approaches and coincide if $g(x) = x$ in the generalized Tukey depth and $\mu(F_{\boldsymbol{u}}) = \int \boldsymbol{u}^T \boldsymbol{x}\, dF$ in the projection depth. The robustness of projection depth depends on the robustness of both the location and scale estimators, whereas the robustness of generalized Tukey depth is determined by the robustness of the scale estimator and the choice of function g. The IFs of generalized Tukey depth for different g functions are provided below and compared with the IF of projection depth for $(\mu, \sigma) =$ (Med,MAD) with respect to the bivariate normal distribution $N(\mathbf{0}, \Sigma)$, with $\Sigma = \left(\begin{smallmatrix} 1 & -.8 \\ -.8 & 2 \end{smallmatrix}\right)$. From Zuo (2003), this is given by

$$\mathrm{IF}(\boldsymbol{z}, D_P(\boldsymbol{x}, F), F) = -(\mathrm{IF}(\boldsymbol{z}, O_P(\boldsymbol{x}, F), F)/(1 + O_P(x, F))^2,$$

with

$$\mathrm{IF}(\boldsymbol{z}, O_P(\boldsymbol{x}, F), F) = \mathrm{sign}(|\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{z}| - \|\Sigma^{-1/2}\boldsymbol{x}\|)\frac{\|\Sigma^{-1/2}\boldsymbol{x}\|}{4h(1)} + \mathrm{sign}(\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{z})\frac{1}{2h(0)},$$

where $h(y) = (2\pi)^{-1}\exp(-y/2)$. For the generalized Tukey depth, we consider two $g(\cdot)$ functions. The first g is the above"Huber" with $c = 2$, i.e., $g(y) = yI_{(|y|\le 2)} + 2\mathrm{sign}(y)I_{(|y|>2)}$; the second g is the above smoothed function $g(y) = (1-e^{-cy})/(1+e^{-cy})$ with $c = 2$. General explicit expressions for the IFs of generalized Tukey depths are elusive, but we can obtain results numerically.

Figure 3 illustrates that the IFs for projection depth and generalized Tukey depth are bounded but discontinuous. The projection depth IF has three jumps due to the IFs of Med and MAD, which are sign functions. The generalized Tukey depth IFs include the IF of the scale estimator, MAD, so they also have jumps but are smoother than the projection depth IF. Unlike the three jumps in the projection depth IF, the smoother (than the Huber) g function smooths out two jumps, leaving only one jump in the generalized Tukey depth IF.

3. Application to depth-weighted L-statistics

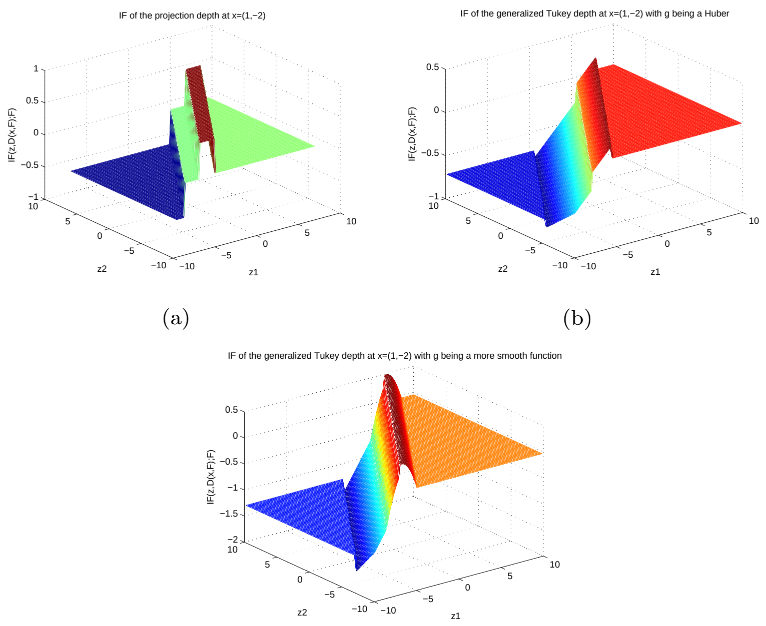
Depth-weighted L-statistics (DL-statistics) are treated in Liu (1990), Liu, Parelius, and Singh (1999), and Zuo, Cui, and He (2004). For multivariate cdf F, depth function $D(\boldsymbol{x}, F)$, and weight function W, and a random sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ from F, a DL-statistic is the sample version $L(F_n)$ of the DL-functional

$$L(F) = \frac{\int \boldsymbol{x}\, W(D(\boldsymbol{x}, F))\, F(d\boldsymbol{x})}{\int W(D(\boldsymbol{x}, F))\, F(d\boldsymbol{x})}. \tag{3}$$

Zuo, Cui, and He (2004) give general results on asymptotics with focus on the projection depth case, while Gao (2003) studies the spatial case using a V-statistic approach. Here we develop the IF of DL-statistics for the spatial, simplicial and generalized Tukey depths and then apply these to obtain asymptotic results.

3.1. *IFs and asymptotics of DL-statistics*

The following general result is given by Zuo, Cui, and Young (2004).



(a) (b)

(c)

Figure 3. For the bivariate normal distribution with $\mu = (0, 0)^T, \sigma_{11} = 1, \sigma_{22} = 2, \sigma_{12} = \sigma_{21} = -0.8$, at point $\boldsymbol{x} = (1, -2)^T$, (a) IF of projection depth; (b) IF of generalized Tukey depth with g being a Huber; (c) IF of generalized depth with g being a smoother function.

Lemma 3.1: *Assume that depth $D(\boldsymbol{x}, F)$ and weight function W satisfy*
(C1) $\int W(D(\boldsymbol{x}, F))\, F(d\boldsymbol{x}) > 0$;
(C2) $\int \|\boldsymbol{x}\|\, W(D(\boldsymbol{x}, F))\, F(d\boldsymbol{x}) < \infty$;
(C3) W has continuous derivative and $W(r) = 0$ for $r \in [0, \alpha r_0]$ for some $\alpha > 1$ and $r_0 \geq 0$;
(C4) $\mathrm{IF}(\boldsymbol{z}; D(\boldsymbol{x}, F), F)$ *exists.*
Then the IF of the DL-estimator is

$$\mathrm{IF}(\boldsymbol{z}; L(F), F) = \frac{\int (\boldsymbol{x} - L)\, W'(D(\boldsymbol{x}, F))\, \mathrm{IF}(\boldsymbol{z}; D(\boldsymbol{x}, F), F)\, dF(\boldsymbol{x}) + (\boldsymbol{z} - L)\, W(D(\boldsymbol{z}, F))}{\int W(D(\boldsymbol{x}, F))\, dF(\boldsymbol{x})}.$$

The first two conditions ensure a well-defined L-statistic: (C1) holds in typical cases and (C2) becomes trivial if $E\|\boldsymbol{X}\| < \infty$ or if $W(D(\boldsymbol{x}, F))$ is 0 outside some bounded set. Condition (C3) requires a sufficiently smooth weight function which is 0 for r in a neighborhood of 0 and ensures a bounded IF for the depth weighted L-statistic, provided that $\mathrm{IF}(\boldsymbol{z}; D(\boldsymbol{x}, F), F)$ is bounded. Condition (C4) is straightforward for the spatial and simplicial depths and also, resorting to conditions (A1)-(A3) in Theorem 2.7, for the generalized Tukey depth. Even for spherically symmetric distributions and for simple L-statistics based on the spatial depth, it is extremely difficult to derive an explicit expression for the IF of DL-estimator because of the complexity of the formula in terms of the weight function and its derivative and the depth function and its IF which also depends on the position of point contamination z. In practice, therefore, we use numerical integration to evaluate (3.1).

We adopt the sufficient conditions of Zuo, Cui, and He (2004) for approximation of $L(F_n)$ as in (1). Put $\nu_n = \sqrt{n}(F_n - F)$, $H_n(\boldsymbol{x}) = \sqrt{n}(D(\boldsymbol{x}, F_n) - D(\boldsymbol{x}, F))$, $\|H_n\|_\infty = \sup_{\boldsymbol{x} \in \mathbb{R}^d} \|H_n(\boldsymbol{x})\|$, and $D_r = \{\boldsymbol{x} : D(\boldsymbol{x}, F) \geq r\}$ for $r \geq 0$. The needed assumptions are:

(D1) $\|H_n\|_\infty = O_p(1)$ and $\sup_{\boldsymbol{x}\in D_{r_0}} \|x\||H_n(\boldsymbol{x})| = O_p(1)$ for some $r_0 \geq 0$;
(D2) $\int \|\boldsymbol{x}\|^2\,(W(D(\boldsymbol{x},F)))^2\,F(d\boldsymbol{x}) < \infty$;
(D3) $\mathrm{IF}(\boldsymbol{x}; D(\boldsymbol{x},F),F)$ exists, and $H_n(\boldsymbol{x}) = \int \mathrm{IF}(\boldsymbol{x}, D(\boldsymbol{y},F),F)\,\nu_n\,(d\boldsymbol{y}) + o_p(1)$ uniformly on $S_n \subset D_{r_0}$ with $P(D_{r_0} - S_n) = o_p(1)$;
(D4) $\int [\int \|\boldsymbol{y}\|\,|W'(D(\boldsymbol{y},F))\mathrm{IF}(\boldsymbol{x}, D(\boldsymbol{y},F),F)|\,F(d\boldsymbol{y})]^2\,F(d\boldsymbol{x}) < \infty$.

Condition (D3) requires (C4) in Lemma 3.1 and conditions such that $D(\boldsymbol{x},F_n) - D(\boldsymbol{x},F) = n^{-1}\sum_{i=1}^n \mathrm{IF}(\boldsymbol{X}_i; D(\boldsymbol{x},F),F) + o_p(n^{-1/2})$ holds uniformly on some set. As a direct result of Theorem 2.1, Lemma 2.2, Theorem 2.5, and Lemma 2.6, (D3) automatically holds for the spatial and simpilicial depths. In the case of the generalized Tukey depth, conditions (A1)-(A3) and (B1)-(B4) imply (D3).

Theorem 3.2 : *Under conditions (C1)-(C3) and (D1)-(D4), we have*

$$L(F_n) - L(F) = n^{-1}\sum_{i=1}^n IF(\boldsymbol{X}_i; L(F),F) + o_p(n^{-1/2}).$$

We also need the following result, easily checked.

Lemma 3.3: *For $L(F)$ defined by (3) with $D(\boldsymbol{x},F)$ the spatial, simplicial, and generalized Tukey depths, we have $E_F(IF(\boldsymbol{X}; L(F),F)) = 0$.*

Consequently, for the spatial, simplicial, and generalized Tukey depths under conditions (C1)-(C3) and (D1)-(D4) we have asymptotic normality of the corresponding DL-statistics:

$$\sqrt{n}(L(F_n) - L(F)) \xrightarrow{d} N(\boldsymbol{0}, E_F\{\mathrm{IF}(\boldsymbol{X}; L(F),F)\mathrm{IF}(\boldsymbol{X}; L(F),F)^T\}). \quad (4)$$

We note, however, that (D1)-(D4) fail for the spatial depth and simplicial depth if $r_0 = 0$. For these depths, asymptotic normality of the corresponding DL-statistics is established only for weight functions that trim observations of sufficiently low depth. Below we will consider typical weight functions in the literature which do not trim in this fashion, so strictly speaking the asymptotic normality results apply only to truncated versions of these weight functions. However, for purposes of comparison of DL-estimators with the sample mean, it will be convenient to use the simple approximate approach of deriving numerical results for the untruncated versions of our weight functions rather than for arbitrary truncated versions.

3.2. *Efficiency of DL-statistics*

Besides robustness and large sample behavior of DL-statistics, we are concerned also with efficiency. In formula (4), let us denote $E_F\{\mathrm{IF}(\boldsymbol{X}; L(F),F)\mathrm{IF}(\boldsymbol{X}; L(F),F)^T\}$ by $V(L,F)$. Then the asymptotic covariance of $L(F_n)$ is $n^{-1}V(L,F)$. For Gaussian models, the most efficient location estimator is the sample mean $\bar{\boldsymbol{X}}$s with covariance matrix $n^{-1}\Sigma$, where Σ is the population covariance. Therefore, for the DL-statistic, we have

$$ARE(L(F_n), \bar{\boldsymbol{X}}) = \frac{[\det(\Sigma)]^{1/d}}{[\det(V(L,F))]^{1/d}},$$

where $\det(A)$ denotes the determinant of the matrix A. This definition of ARE is a standard and classical one as discussed, for example, in Serfling (1980), but

Table 1. ARE of spatial DL-statistic relative to sample mean, for two weight functions, at d-variate normal models.

d	2	3	5	10	20
W_e	0.893	0.957	0.997	0.999	1.000
W_g	0.624	0.782	0.885	0.958	0.983

we acknowledge that other measures could be chosen to be used instead or in conjunction with this version in order to obtain a larger perspective.

Obviously, the ARE depends on the dimension d and the weight function W. Maronna and Yohai (1995) suggested the Huber-type weight function $W_h(r;c) = r^2/c^2 I_{(r<c)} + I_{(r\geq c)}$. However, it is not smooth enough to satisfy condition (C3) which requires a continuous derivative. Two other weight functions are considered. Zuo, Cui, and He (2004) recommend the weight function $W_e(r;c,K) = (\exp(-K(1-r/c)^2) - \exp(-K))/(1-\exp(-K))I_{(r<c)} + I_{(r\geq c)}$. Gervini (2002) proposes a Gaussian weight function $W_g(r;c) = \varphi((1/r-1)^2/(1/c-1)^2)/\varphi(1)$, where $\varphi(\cdot)$ is the standard normal density. The parameter c in the weight functions balances efficiency and robustness of the DL-statistic. For standard normal distributions of dimensions $d = 2, 3, 5, 10, 20$, the cutoff values of c in each weight function are 0.36, 0.33, 0.32, 0.30, 0.30, respectively, which are the medians of the spatial depth. Both $W_e(\cdot)$ and $W_g(\cdot)$ assign exponentially deceasing weight to the lower depth half of points; for the other points, $W_e(\cdot)$ assigns weight 1, while $W_g(\cdot)$ assigns weight greater than 1. The parameter K in $W_e(\cdot)$ controls the rate of the decreasing weight for the points with low depth. Here $K = 3$.

Because of the difficulty of obtaining an explicit expression for $\text{IF}(\boldsymbol{X}; L(F), F)$, numerical integration is used. Also, here we only consider the spatial DL-statistic due to its simplicity. In this case, as noted above, the use of $V(L,F)$ from (4) presupposes the use of truncations of the above weight functions. However, a suitable approximate ARE benchmark is provided by using the untruncated versions instead of introducing arbitrary truncations. Values of ARE are provided in Table 1. We see that ARE increases with the dimension d and that $W_e(\cdot)$ is preferred. For $d \geq 10$, the spatial DL-statistics attain ARE = 1 while still possessing robustness, such as a bounded influence function. The ARE tending to 1 is linked to the straightforward property for Gaussian variables that as dimension increases the data points lie near the surface of an expanding sphere.

A companion approach is to explore finite sample relative efficiency through simulation. For each of sample sizes $n = 20, 30, 50, 100$, we generated $m = 1000$ random samples from the bivariate standard normal, the (heavy-tailed) t distribution with 3 degrees of freedom, and the contaminated normal model $(1-\varepsilon)N((0,0)^T, I_{2\times 2}) + \varepsilon N((10,10)^T, I_{2\times 2})$, with $\varepsilon = 0.1$. For each estimator we calculated the empirical mean squared error, $EMSE = \frac{1}{m}\sum_{j=1}^{m} \|L(F_n)_j - \mu\|^2$, with $m = 1000$, $\mu = (0,0)^T$, and $L(F_n)_j$ the estimate for j-th sample. The RE of $L(F_n)$ is then ratio of the EMSE of the sample mean to that of $L(F_n)$.

Three DL-estimators based on spatial depth, projection depth, and generalized Tukey depth are considered and denoted as SL, PL, and GL, respectively, using the weight function $W_e(\cdot)$ with $K = 3$. Under the bivariate standard normal model, the cutoff values of c are 0.36, 0.36, 0.45, the medians of the spatial, projection, and generalized Tukey depths, respectively. Table 2 gives finite sample REs of these DL-statistics with respect to the sample mean. All REs are above 90% under the Gaussian model, significantly higher than those for the halfspace median (Rousseeuw and Ruts, 1998), the projection median (Zuo, 2003), and the spatial median (Gao, 2003) at about 76%, 78%, 78%, respectively. Under T_3 and contam-

Table 2. Empirical mean squared error and relative efficiency of DL-statistics.

		n=20		n=30		n=50		n=100	
		EMSE	RE	EMSE	RE	EMSE	RE	EMSE	RE
	$\bar{X}$	0.0972	1.000	0.0684	1.000	0.0416	1.000	0.0205	1.000
	SL	0.1062	0.915	0.0751	0.911	0.0456	0.912	0.0226	0.907
Normal	PL	0.1031	0.943	0.0722	0.947	0.0430	0.967	0.0212	0.967
	GL	0.1078	0.902	0.0759	0.901	0.0460	0.904	0.0227	0.903
	$\bar{X}$	0.3045	1.000	0.1813	1.000	0.1163	1.000	0.0626	1.000
	SL	0.1584	1.922	0.1009	1.797	0.0624	1.866	0.0316	1.982
T_3	PL	0.1573	1.936	0.1026	1.768	0.0646	1.802	0.0326	1.917
	GL	0.1604	1.898	0.1031	1.758	0.0634	1.835	0.0316	1.980
	$\bar{X}$	2.0119	1.000	2.0365	1.000	2.0481	1.000	2.0125	1.000
Mixed	SL	0.1948	10.88	0.1777	11.46	0.1714	11.95	0.1692	11.89
normal	PL	0.1326	15.98	0.0952	21.38	0.0695	29.48	0.0471	42.71
	GL	0.1661	12.75	0.1191	17.10	0.0888	23.07	0.0603	33.38

inated normal models, all depth-weighted means are much better than the sample mean with projection depth-weighted mean slightly better. However, for the projection and generalized Tukey depths, there is a computational burden of order $O(n^{d+1})$ for dimension d.

4. Appendix: Proofs

Proof: (Proof of Theorem 2.1). For Case (i),

$$D_s(\boldsymbol{x}, F) = 1 - \left\| -E_F \left\{ \frac{\boldsymbol{X} - \boldsymbol{x}}{\|\boldsymbol{X} - \boldsymbol{x}\|} \right\} \right\| = 1 - \|\boldsymbol{u_x}\|.$$

Let $\widetilde{F} = (1-\varepsilon)F + \varepsilon\delta_{\boldsymbol{z}}$. Then,

$$\begin{aligned} D_s(\boldsymbol{x}, \widetilde{F}) &= 1 - \left\| -E_{\widetilde{F}} \left\{ \frac{\boldsymbol{X} - \boldsymbol{x}}{\|\boldsymbol{X} - \boldsymbol{x}\|} \right\} \right\| \\ &= 1 - \left\| (1-\varepsilon) E_F \left\{ \frac{\boldsymbol{X} - \boldsymbol{x}}{\|\boldsymbol{X} - \boldsymbol{x}\|} \right\} + \varepsilon \frac{\boldsymbol{z} - \boldsymbol{x}}{\|\boldsymbol{z} - \boldsymbol{x}\|} \right\| \\ &= 1 - \|(1-\varepsilon)\boldsymbol{u_x} + \varepsilon \boldsymbol{z_x}\|. \end{aligned}$$

Hence

$$\begin{aligned} \mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) &= \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \left(1 - \|(1-\varepsilon)\boldsymbol{u_x} + \varepsilon \boldsymbol{z_x}\| - (1 - \|\boldsymbol{u_x}\|)\right) \\ &= \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \left(\|\boldsymbol{u_x}\| - \sqrt{(1-\varepsilon)^2 \|\boldsymbol{u_x}\|^2 + \varepsilon^2 + 2(1-\varepsilon)\varepsilon \|\boldsymbol{u_x}\| \cos\theta}\,\right) \\ &= (\|\boldsymbol{u_x}\|^2 - \|\boldsymbol{u_x}\| \cos\theta\,)/\|\boldsymbol{u_x}\| \\ &= \|\boldsymbol{u_x}\| - \cos\theta \\ &= 1 - D_s(\boldsymbol{x}, F) - \cos\theta. \end{aligned}$$

For Case (ii), $\boldsymbol{z} = \boldsymbol{x}$, then $\boldsymbol{x} - \boldsymbol{z} = \boldsymbol{0}$ and $\frac{\boldsymbol{x}-\boldsymbol{z}}{\|\boldsymbol{x}-\boldsymbol{z}\|} = \boldsymbol{0}$. Hence $D_s(\boldsymbol{x}, \widetilde{F}) = 1 - \|(1-\varepsilon)\boldsymbol{u_x} + \varepsilon\boldsymbol{0}\| = 1 - (1-\varepsilon)\|\boldsymbol{u_x}\|$ and $\mathrm{IF}(\boldsymbol{x}; D_s(\boldsymbol{x}, F), F) = \|\boldsymbol{u_x}\| = 1 - D_s(\boldsymbol{x}, F) \geq 0$.

For Case (iii), $\boldsymbol{x} = \boldsymbol{\mu}_F$, then $E_F\left\{\frac{\boldsymbol{x}-\boldsymbol{X}}{\|\boldsymbol{x}-\boldsymbol{X}\|}\right\} = \boldsymbol{0}$ and $D_s(\boldsymbol{x}, F) = 1$. If $\boldsymbol{z} \neq \boldsymbol{x}$, then

$D_s(\boldsymbol{x}, \widetilde{F}) = 1 - \|(1-\varepsilon)\boldsymbol{0} + \varepsilon \boldsymbol{z_x}\| = 1 - \varepsilon$. Hence $\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{\mu}_F, F), F) = -1$.
For Case (iv), $\boldsymbol{z} = \boldsymbol{x} = \boldsymbol{\mu}_F$ and $D_s(\boldsymbol{x}, \widetilde{F}) = D_s(\boldsymbol{x}, F) = 1$, so $\mathrm{IF}(\boldsymbol{\mu}_F; D_s(\boldsymbol{\mu}_F, F), F) = 0$. □

Proof: (Proof of Lemma 2.2). Put $\nu_n = \sqrt{n}(F_n - F)$ and $h_n(\boldsymbol{x}) = n^{-1/2} \int \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|} \nu_n(d\boldsymbol{y})$. From the proof of Theorem 1, it is straightforward to obtain

$$\mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) = \begin{cases} \frac{\langle \boldsymbol{u_x}, \boldsymbol{u_x} - \boldsymbol{z_x} \rangle}{\|\boldsymbol{x}\|}, & \boldsymbol{x} \in \mathbb{R}^d \setminus \{\boldsymbol{x} : \boldsymbol{u_x} = \boldsymbol{0}\}, \\ -\boldsymbol{z_x}, & \boldsymbol{x} \in \{\boldsymbol{x} : \boldsymbol{u_x} = \boldsymbol{0}\}. \end{cases}$$

For $\boldsymbol{x} \in \mathbb{R}^d \setminus \{\boldsymbol{x} : \boldsymbol{u_x} = \boldsymbol{0}\}$, Taylor expansion yields

$$\|\boldsymbol{u_x} + h_n(\boldsymbol{x})\| - \|\boldsymbol{u_x}\| = \frac{\langle \boldsymbol{u_x}, h_n(\boldsymbol{x}) \rangle}{\|\boldsymbol{u_x}\|} + O(\|h_n\|_\infty^2).$$

Thus

$$\begin{aligned} D_s(\boldsymbol{x}, F_n) - D_s(\boldsymbol{x}, F) &= \left\| \int \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|} F(d\boldsymbol{y}) \right\| - \left\| \int \frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|} F_n(d\boldsymbol{y}) \right\| \\ &= -(\|\boldsymbol{u_x} + h_n(\boldsymbol{x})\| - \|\boldsymbol{u_x}\|) \\ &= -\left[\frac{\langle \boldsymbol{u_x}, h_n(\boldsymbol{x}) \rangle}{\|\boldsymbol{u_x}\|} + O(\|h_n\|_\infty^2) \right] \\ &= -\left[\frac{\langle \boldsymbol{u_x}, \int \boldsymbol{z_x} F_n(d\boldsymbol{z}) - \boldsymbol{u_x} \rangle}{\|\boldsymbol{u_x}\|} + O(\|h_n\|_\infty^2) \right] \\ &= \int \mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) F_n(d\boldsymbol{z}) + o_p(n^{-1/2}) \end{aligned}$$

holds uniformly in $\mathbb{R}^d \setminus \{\boldsymbol{x} : \boldsymbol{u_x} = \boldsymbol{0}\}$, where the last equality comes from $\|h_n\|_\infty^2 = O(n^{-1})$. Also, the above equality holds uniformly over the set $\{\boldsymbol{x} : \boldsymbol{u_x} = \boldsymbol{0}\}$: $D_s(\boldsymbol{x}, F_n) - D_s(\boldsymbol{x}, F) = -(\|\boldsymbol{u_x} + h_n(\boldsymbol{x})\| - \|\boldsymbol{u_x}\|) = -\|h_n(\boldsymbol{x})\| = -\int \boldsymbol{z_x} F_n(d\boldsymbol{z})$ $= \int \mathrm{IF}(\boldsymbol{z}; D_s(\boldsymbol{x}, F), F) F_n(d\boldsymbol{z})$. □

Proof: (Proof of Lemma 2.3). Consider the functional $T(F) = \int ... \int h(\boldsymbol{x}_1, ..., \boldsymbol{x}_m)\, dF(\boldsymbol{x}_1) \cdots dF(\boldsymbol{x}_m)$. Writing

$$\begin{aligned} T(F + \varepsilon(G - F)) = \sum_{j=0}^{m} \binom{m}{j} \varepsilon^{m-j} \int \cdots \int h(\boldsymbol{x}_1, ..., \boldsymbol{x}_m) \prod_{i=1}^{j} dF(\boldsymbol{x}_i) \\ \times \prod_{i=j+1}^{m} d(G(\boldsymbol{x}_i) - F(\boldsymbol{x}_i)), \end{aligned}$$

for $k = 1, ..., m$, we have

$$
\begin{aligned}
&\frac{d^k}{d\varepsilon^k}T(F+\varepsilon(G-F))|_{\varepsilon=0^+}\\
&\quad= \sum_{j=0}^{m-k}\binom{m}{j}(m-j)\cdots(m-j-k+1)\varepsilon^{m-j-k}\int\cdots\int h(\boldsymbol{x}_1,...,\boldsymbol{x}_m)\\
&\qquad\times\prod_{i=1}^{j}dF(\boldsymbol{x}_i)\prod_{i=j+1}^{m}d(G(\boldsymbol{x}_i)-F(\boldsymbol{x}_i))|_{\varepsilon=0^+}\\
&\quad= m(m-1)\cdots(m-k+1)\int\cdots\int h(\boldsymbol{x}_1,...,\boldsymbol{x}_k,\boldsymbol{y}_1,...,\boldsymbol{y}_{m-k})\\
&\qquad\times\prod_{i=1}^{m-k}dF(\boldsymbol{y}_i)\prod_{i=1}^{k}d(G(\boldsymbol{x}_i)-F(\boldsymbol{x}_i))
\end{aligned}
$$

and for $k > m$,

$$\frac{d^k}{d\varepsilon^k}T(F+\epsilon(G-F))|_{\varepsilon=0^+} = 0.$$

The IF of $T(F)$ is just the special case of this Gâteaux differential of T at F in the direction of G with $G = \delta_{\boldsymbol{z}}$ and $k = 1$ (for details, see Serfling, 1980). We thus obtain

$$
\begin{aligned}
\mathrm{IF}(\boldsymbol{z};T(F),F) &= m\int\cdots\int h(\boldsymbol{x}_1,\boldsymbol{y}_1,...,\boldsymbol{y}_m)\prod_{i=1}^{m-1}dF(\boldsymbol{y}_i)d(G(\boldsymbol{x}_1)-F(\boldsymbol{x}_1))\\
&= m(\int\cdots\int h(\boldsymbol{z},\boldsymbol{y}_1,...,\boldsymbol{y}_{m-1})\prod_{i=1}^{m-1}dF(\boldsymbol{y}_i)-T(F))\\
&= m(T(F\,|\,\boldsymbol{x}_1=\boldsymbol{z})-T(F)).
\end{aligned}
$$

□

Proof: (Proof of Lemma 2.4).

$$
\begin{aligned}
E_F\{\mathrm{IF}(\boldsymbol{z};U_n(h),F)\} &= m\binom{n}{m-1}^{-1}\sum_C E_F h(\boldsymbol{z},\boldsymbol{x}_{i_1},...,\boldsymbol{x}_{i_{m-1}}) - mE_F U_n(h)\\
&= m\binom{n}{m-1}^{-1}\binom{n}{m-1}\int...\int h(\boldsymbol{z},\boldsymbol{x}_{i_1},...,\boldsymbol{x}_{i_{m-1}})dF(\boldsymbol{x}_{i_1})\\
&\quad\times dF(\boldsymbol{x}_{i_2})\cdots dF(\boldsymbol{x}_{i_{m-1}}) - mE_F U_n(h)\\
&= m(T(F\,|\,\boldsymbol{x}_1=\boldsymbol{z})-T(F))\\
&= \mathrm{IF}(\boldsymbol{z};T(F),F).
\end{aligned}
$$

□

Proof: (Proof of Theorem 2.7). For any $\varepsilon > 0$ and given $\boldsymbol{x},\boldsymbol{z} \in \mathbb{R}^d$, let $\widetilde{F} = (1-\varepsilon)F + \varepsilon\delta_{\boldsymbol{z}}$, denote $\boldsymbol{u}_{\boldsymbol{x}}(\eta) = \{\boldsymbol{u} : \|\boldsymbol{u}\| = 1 \cap \|\boldsymbol{u}-\boldsymbol{u}_{\boldsymbol{x}}\| \le \eta \text{ for any } \eta > 0\}$, and write

$$h(\boldsymbol{u},\widetilde{F}) = -E_{\widetilde{F}}\{g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})\} + O_g(\boldsymbol{x},F).$$

Then

$$
\begin{aligned}
-(O_g(\boldsymbol{x},\widetilde{F})-O_g(\boldsymbol{x},F))/\varepsilon &= -(\sup_{\|\boldsymbol{u}\|=1} E_{\widetilde{F}}\{g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},X)\} - O_g(\boldsymbol{x},F))/\varepsilon \\
&= \inf_{\|\boldsymbol{u}\|=1} (-E_{\widetilde{F}}\{g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},X)\} + O_g(\boldsymbol{x},F))/\epsilon \\
&= \inf_{\|\boldsymbol{u}\|=1} h(\boldsymbol{u},\widetilde{F})/\varepsilon \\
&= \min\left\{ \inf_{\boldsymbol{u}=\boldsymbol{u_x}} h(\boldsymbol{u},\widetilde{F})/\varepsilon, \inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C} h(\boldsymbol{u},\widetilde{F})/\varepsilon \right\} \\
&= \min\left\{ \inf_{\boldsymbol{u}=\boldsymbol{u_x}} h(\boldsymbol{u},\widetilde{F})/\varepsilon, \inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C\cap\boldsymbol{u_x}(\eta)} h(\boldsymbol{u},\widetilde{F})/\varepsilon, \right. \\
&\qquad \left. \inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C\cap\boldsymbol{u_x}(\eta)^C} h(\boldsymbol{u},\widetilde{F})/\varepsilon \right\}. \qquad (5)
\end{aligned}
$$

Now, as $\epsilon \downarrow 0$,
$(-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\epsilon$

$$
\begin{aligned}
&= \frac{-(1-\epsilon)\int g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\,dF - \epsilon\int g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\,d\delta_{\boldsymbol{z}} + \int g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\,dF}{\epsilon} \\
&\to \int g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\,dF + \int g\prime(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\frac{Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})}{\sigma(F_{\boldsymbol{u}})}\mathrm{IF}(\boldsymbol{u}^T\boldsymbol{z};\sigma,F_{\boldsymbol{u}})\,dF \\
&\quad -g(Y_F(\boldsymbol{u_x},\boldsymbol{x},\boldsymbol{z})) \\
&= E_F(g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X}))) - g(Y_F(\boldsymbol{u_x},\boldsymbol{x},\boldsymbol{z})) \\
&\quad + \int g\prime(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})\,dF \times \frac{\mathrm{IF}(\boldsymbol{u}^T\boldsymbol{z};\sigma,F_{\boldsymbol{u}})}{\sigma(F_{\boldsymbol{u}})}.
\end{aligned}
$$

By the given assumptions, this is uniformly bounded in $\boldsymbol{u}$ for the given $\boldsymbol{x}$ and $\boldsymbol{z}$. Hence the third term in (5) satisfies
$\inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C\cap\boldsymbol{u_x}(\eta)^C} h(\boldsymbol{u},\widetilde{F})/\epsilon$

$$
\begin{aligned}
&= \inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C\cap\boldsymbol{u_x}(\eta)^C} ((-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon + (O_g(\boldsymbol{x},F) \\
&\quad -E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon) \\
&\geq \inf_{\boldsymbol{u}\in\boldsymbol{u_x}^C\cap\boldsymbol{u_x}(\eta)^C} (O_g(\boldsymbol{x},F) - E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/2\varepsilon \qquad \text{(for sufficiently small } \varepsilon) \\
&= O(1/\varepsilon),
\end{aligned}
$$

and the second term in (5) satisfies

$$
\begin{aligned}
&\inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}{}^C\cap \boldsymbol{u}_{\boldsymbol{x}}(\eta)} h(\boldsymbol{u},\widetilde{F})/\epsilon \\
&\quad = \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}{}^C\cap \boldsymbol{u}_{\boldsymbol{x}}(\eta)} ((-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon + (O_g(\boldsymbol{x},F) \\
&\qquad - E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon) \\
&\quad \geq \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}{}^C\cap \boldsymbol{u}_{\boldsymbol{x}}(\eta)} (-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon \\
&\quad \geq \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}(\eta)} (-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon.
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}(\eta)} (-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon \\
&\qquad \leq \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}(\eta)} (-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + O_g(\boldsymbol{x},F))/\varepsilon \\
&\qquad \leq (-O_g(\boldsymbol{x},\widetilde{F}) + O_g(\boldsymbol{x},F))/\varepsilon \\
&\qquad \leq \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}} (-E_{\widetilde{F}} g(Y_{\widetilde{F}}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})) + E_F g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X})))/\varepsilon.
\end{aligned}
$$

Letting $\varepsilon \downarrow 0$, we have

$$
\begin{aligned}
&\inf\nolimits_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}(\eta)} \Big(E_F(g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X}))) - g(Y_F(\boldsymbol{u}_{\boldsymbol{x}},\boldsymbol{x},\boldsymbol{X})) \\
&\qquad + \textstyle\int g\prime(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})\, dF \times \frac{\mathrm{IF}(\boldsymbol{u}^T\boldsymbol{z};\sigma,F_{\boldsymbol{u}})}{\sigma(F_{\boldsymbol{u}})} \Big) \\
&\qquad \leq \lim_{\varepsilon\downarrow 0} (-O_g(\boldsymbol{x},\widetilde{F}) + O_g(\boldsymbol{x},F))/\varepsilon \\
&\qquad \leq \inf_{\boldsymbol{u}\in \boldsymbol{u}_{\boldsymbol{x}}} (E_F(g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{X}))) - g(Y_F(\boldsymbol{u}_{\boldsymbol{x}},\boldsymbol{x},\boldsymbol{z})) \\
&\qquad\quad + \int g\prime(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})\, dF \times \frac{\mathrm{IF}(\boldsymbol{u}^T\boldsymbol{z};\sigma,F_{\boldsymbol{u}})}{\sigma(F_{\boldsymbol{u}})}).
\end{aligned}
$$

Letting $\eta \to 0$, by the condition that $\boldsymbol{u}_{\boldsymbol{x}}$ is singleton almost surely, we obtain

$$
\mathrm{IF}(\boldsymbol{z};O_g(\boldsymbol{x},F),F) = -\left[O_g(\boldsymbol{x},F) - g(Y_F(\boldsymbol{u}_{\boldsymbol{x}},\boldsymbol{x},\boldsymbol{z})) + A \times \frac{\mathrm{IF}(\boldsymbol{u}_{\boldsymbol{x}}^T\boldsymbol{z};\sigma,F_{\boldsymbol{u}_{\boldsymbol{x}}})}{\sigma(F_{\boldsymbol{u}_{\boldsymbol{x}}})} \right],
$$

where $A = \int g\prime(Y_F(\boldsymbol{u}_{\boldsymbol{x}},\boldsymbol{x},\boldsymbol{y}))Y_F(\boldsymbol{u}_{\boldsymbol{x}},\boldsymbol{x},\boldsymbol{y})\, dF$. The remaining result follows from the definition of the depth function. □

Proof: (Proof of Lemma 2.8). First we observe that

$$
\begin{aligned}
&\sqrt{n}\left\{ \int g(Y_{F_n}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})F_n(d\boldsymbol{y}) - \int g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))F(d\boldsymbol{y}) \right\} \\
&\quad = \sqrt{n}\int \left\{ g(Y_{F_n}(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})) - g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\right\} F_n(d\boldsymbol{X}) + \int g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\nu_n(d\boldsymbol{y}) \\
&\quad = \int g'(\theta(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))T_n(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y})F_n(d\boldsymbol{y}) + \int g(Y_F(\boldsymbol{u},\boldsymbol{x},\boldsymbol{y}))\nu_n(d\boldsymbol{y}),
\end{aligned}
$$

where $\theta(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})$ is a point between $Y_{F_n}(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})$ and $Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})$ with

$$\sup |\theta(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) - Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})| = O_p(1/\sqrt{n}),$$

and

$$\begin{aligned} T_n(\boldsymbol{y}, \boldsymbol{x}, \boldsymbol{y}) &= \sqrt{n}(Y_{F_n}(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) - Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) \\ &= \frac{-Y_F(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y})}{\sigma(F_{\boldsymbol{u_x}})} \int \mathrm{IF}(\boldsymbol{u}_{\boldsymbol{x}}^T \boldsymbol{z}; \sigma, F) F_n(d\boldsymbol{z}) + o_p(1) = O_p(1). \end{aligned}$$

On one hand, we have

$$\Big| \int (g'(\theta(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) - g'(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}))) T_n(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) F_n(d\boldsymbol{y}) \Big| = o_p(1),$$

since $g'(\cdot)$ is a continuous function. On the other hand
$\Big| \int g'(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) T_n(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})(F_n - F)(d\boldsymbol{y}) \Big|$

$$= \frac{1}{\sqrt{n}} \Big| \int g'(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) T_n(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) \nu_n(d\boldsymbol{y}) \Big| = o_p(1).$$

Therefore we obtain
$\sqrt{n} \Big\{ \int g(Y_{F_n}(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) F_n(d\boldsymbol{X}) - \int g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) F(d\boldsymbol{y}) \Big\}$

$$= \int g'(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) T_n(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) F(d\boldsymbol{y}) + \int g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) \nu_n(d\boldsymbol{y}) + o_p(1).$$

Writing

$$S_n(\boldsymbol{u}, \boldsymbol{x}) = \sqrt{n} \Big\{ \int g(Y_{F_n}(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) F_n(d\boldsymbol{y}) - \int g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) F(d\boldsymbol{y}) \Big\}$$

and

$$S(\boldsymbol{u}, \boldsymbol{x}) = \int g'(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) T_n(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{X}) F(d\boldsymbol{y}) + \int g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) \nu_n(d\boldsymbol{y}),$$

we have

$$\sup_{\|\boldsymbol{u}\|=1} |S_n(\boldsymbol{u}, \boldsymbol{x}) - S(\boldsymbol{u}, \boldsymbol{x})| \longrightarrow 0, \quad \text{a.s..}$$

Letting $V_{1n} = \int g(Y_{F_n}(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y}) F_n(d\boldsymbol{y})$, $V_1 = \int g(Y_F(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{y})) F(d\boldsymbol{y})$, $S_{1n} = S_n$, $S_1 = S$, $\Delta = \{\boldsymbol{u} : \|\boldsymbol{u}\| = 1, \boldsymbol{u} \in \mathbb{R}^d\}$, and $B_1 = \boldsymbol{u_x}$, and invoking Lemma 6.1 in Zhang (2002), we obtain that as $n \to \infty$,

$$\big\{ \sqrt{n}(O_g(\boldsymbol{x}, F_n) - O_g(\boldsymbol{x}, F_n)) : \boldsymbol{x} \in K \big\} \xrightarrow{d} \{ \sup_{\boldsymbol{u} \in \boldsymbol{u_x}} S(\boldsymbol{u}, \boldsymbol{x}) : \boldsymbol{x} \in K \}.$$

Noting that

$$n^{-1/2}T_n(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y}) = -\frac{Y_F(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y})}{\sigma(F_{\boldsymbol{u_x}})} \int \mathrm{IF}(\boldsymbol{u_x}^T \boldsymbol{z}; \sigma, F) F_n(d\boldsymbol{z}) + O_p(1)$$

and $\boldsymbol{u_x}$ is a singleton for $\boldsymbol{x} \in K$, we have
$\{\sup_{\boldsymbol{u} \in \boldsymbol{u_x}} S(\boldsymbol{u}, \boldsymbol{x}) : \boldsymbol{x} \in K\}$

$$\begin{aligned}
&= Z(\boldsymbol{u_x}, \boldsymbol{x}) \\
&= \int g'(Y_F(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y})) T_n(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y}) F(d\boldsymbol{y}) + \int g(Y_F(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{y})) \nu_n(d\boldsymbol{y}), \\
&= \sqrt{n} \int [g(Y_F(\boldsymbol{u_x}, \boldsymbol{x}, \boldsymbol{z})) - O_g(\boldsymbol{x}, F) - \frac{A}{\sigma(F_{\boldsymbol{u_x}})} \int \mathrm{IF}(\boldsymbol{u_x}^T \boldsymbol{z}; \sigma, F)] F_n(d\boldsymbol{z}) + O_p(1) \\
&= \sqrt{n} \int \mathrm{IF}(\boldsymbol{z}; O_g(\boldsymbol{x}, F), F) F_n(d\boldsymbol{z}) + O_p(1).
\end{aligned}$$

The remaining statements to be proved follow from the definition of the depth. □

Proof: (Proof of Lemma 3.3). For the simplicial depth, the result is straightforward by integration of the equation in Theorem 2.5. We prove the other two cases. To prove $E_F(\mathrm{IF}(\boldsymbol{X}; L(F), F)) = \int \mathrm{IF}(\boldsymbol{X}; L(F), F) dF(\boldsymbol{X}) = 0$, it suffices to show $E_F(\mathrm{IF}(\boldsymbol{X}; D(\boldsymbol{y}, F), F) = 0$, since we can switch the order in the double integral by Fubini's Theorem and $\int (\boldsymbol{x} - L) W(D(\boldsymbol{x}, F)) dF(\boldsymbol{x}) = 0$. For the spatial depth, we have $\mathrm{IF}(\boldsymbol{x}; D_S(\boldsymbol{y}, F), F) = 1 - D_S(\boldsymbol{y}, F) - \cos(\theta)$, with θ being the angle between $\boldsymbol{u_y} = E_F\{(\boldsymbol{X} - \boldsymbol{y})/\|\boldsymbol{X} - \boldsymbol{y}\|\}$ and the unit vector $\boldsymbol{x_y} = (\boldsymbol{x} - \boldsymbol{y})/\|\boldsymbol{x} - \boldsymbol{y}\|$ if $\boldsymbol{x} \neq \boldsymbol{y}$. Hence

$$\begin{aligned}
E_F(\mathrm{IF}(\boldsymbol{X}; D_S(\boldsymbol{y}, F), F)) &= \int (1 - D_S(\boldsymbol{y}, F) - \cos(\theta)) dF(\boldsymbol{x}) \\
&= 1 - D_S(\boldsymbol{y}, F) - \int \frac{\boldsymbol{u_y} \cdot \boldsymbol{x_y}}{\|\boldsymbol{u_y}\|} dF(\boldsymbol{x}) \\
&= 1 - D_S(\boldsymbol{y}, F) - \frac{\boldsymbol{u_y} \cdot \int \boldsymbol{x_y} dF(\boldsymbol{x})}{\|\boldsymbol{u_y}\|} \\
&= 1 - D_S(\boldsymbol{y}, F) - \|\boldsymbol{u_y}\| \\
&= 0.
\end{aligned}$$

For the generalized half-space depth,

$$\mathrm{IF}(\boldsymbol{x}; D_g(\boldsymbol{y}, F), F) = -\mathrm{IF}(\boldsymbol{x}; O(\boldsymbol{y}, F), F)/(1 + O(\boldsymbol{y}, F))^2.$$

Thus we need only show $E_F(\mathrm{IF}(\boldsymbol{x}; O_g(\boldsymbol{x}, F), F)) = 0$. Here

$$\mathrm{IF}(\boldsymbol{x}; O_g(\boldsymbol{y}, F), F) = O_g(\boldsymbol{y}, F) - g(Y_F(\boldsymbol{u_y}, (\boldsymbol{y} - \boldsymbol{x}))) + A \frac{\mathrm{IF}(\boldsymbol{u_y}^T \boldsymbol{y}; \sigma, F_{\boldsymbol{u_y}})}{\sigma(F_{\boldsymbol{u_y}})}$$

where $A = \int g'(Y_F(\boldsymbol{u_y}), \boldsymbol{y}) Y_F(\boldsymbol{u_y}, \boldsymbol{y})\, dF$. So we have

$$E_F(\mathrm{IF}(\boldsymbol{X}; O_g(\boldsymbol{y}, F), F)) = O_g(\boldsymbol{y}, F) - \int g(Y_F(\boldsymbol{u_y}, (\boldsymbol{y} - \boldsymbol{x})))dF(\boldsymbol{x}) + A\frac{\int \mathrm{IF}(\boldsymbol{u_y}^T\boldsymbol{y}; \sigma, F_{\boldsymbol{u_y}})dF(\boldsymbol{x})}{\sigma(F_{\boldsymbol{u_y}})}. \quad (6)$$

From the definition of the outlyingness function, the second term in (6) satisfies $\int g(Y_F(\boldsymbol{u_y}, (\boldsymbol{y} - \boldsymbol{x})))\, dF(\boldsymbol{x}) = O_g(\boldsymbol{y}, F)$. Noting that $\int \mathrm{IF}(\boldsymbol{u_y}^T\boldsymbol{y}; \sigma, F_{\boldsymbol{u_y}})\, dF(\boldsymbol{x})$ is the expectation of the IF for a scale estimator in a univariate case and equals 0, the proof is complete. □

Proof: (Proof of Theorem 3.2). Without loss of generality, we assume that $\theta = L(F) = 0$. According to the von Mises expansion,

$$\begin{aligned}\sqrt{n}R_{1n} &= \sqrt{n}(L(F_n) - 0) - \int \mathrm{IF}(\boldsymbol{z}; L(\boldsymbol{x}, F), F)\nu_n(d\boldsymbol{z}) \\ &= \frac{\sqrt{n}\int \boldsymbol{x} W(D(\boldsymbol{x}, F_n))F_n(d\boldsymbol{x})}{\int W(D(\boldsymbol{x}, F_n))F_n(d\boldsymbol{x})} \\ &\quad - \frac{\int\{\int W'(D(\boldsymbol{x}, F))\mathrm{IF}(\boldsymbol{x}; D(\boldsymbol{y}, F))F(d\boldsymbol{y}) + xW(D(\boldsymbol{x}, F))\}\nu_n(d\boldsymbol{x})}{\int W(D(\boldsymbol{y}, F))F(d\boldsymbol{y})}.\end{aligned}$$

We note that $\mathrm{IF}(\boldsymbol{x}; D(\boldsymbol{y}, F))$ is exactly $h(\boldsymbol{x}, \boldsymbol{y})$, which satisfies condition (A4) in Zuo, Cui, and He (2004). Then, following the same steps as in the proof of Theorem 2.1 in Zuo, Cui, and He (2004), we have

$$\sqrt{n}\int \boldsymbol{x} W(D(\boldsymbol{x}, F_n))F_n(d\boldsymbol{x})$$

$$= \int \left\{ \int W'(D(\boldsymbol{x}, F))\mathrm{IF}(\boldsymbol{x}; D(\boldsymbol{y}, F))F(d\boldsymbol{y}) + \boldsymbol{x}W(D(\boldsymbol{x}, F)) \right\} \nu_n(d\boldsymbol{x}) + o_p(1)$$

and

$$\int W(D(\boldsymbol{x}, F_n))F_n(d\boldsymbol{x}) = \int W(D(\boldsymbol{x}, F))F(d\boldsymbol{x}) + O_p(1/\sqrt{n}).$$

Therefore $\sqrt{n}R_{1n} = o_p(1)$. □

Acknowledgments

Very helpful comments and suggestions by referees are much appreciated. The support of NSF Grants DMS-0103698 and CCF-0430366 is gratefully acknowledged. The results on influence functions are in part from doctoral dissertation work (Dang, 2005) under direction of Robert Serfling.

References

[1] M.A. Arcones and E. Giné, *Limit theorems for U-processes*, Ann. Prob. 21 (1993), pp. 1494–1542.
[2] P. Chaudhuri, *On a geometric notion of quantiles for multivariate data.* J. Am. Stat. Asso. 91 (1996), pp. 1038–1044.
[3] X. Dang, *Nonparametric Multivariate Outlier Detection Methods, with Applications*, Ph.D. diss., University of Texas at Dallas, 2005.

[4] R.M. Dudley and V.I. Koltchinskii, *The spatial quantiles*, Unpublished manuscript, (1992).
[5] L. Dümbgen, *Limit theorem for the simplicial depth*, Stat. Probabil. Lett. 14 (1992), pp. 119–128.
[6] Y. Gao, *Data depth based on spatial rank*, Stat. Probabil. Lett. 65 (2003), pp. 217–225.
[7] D. Gervini, *The influence function of the Stahel-Donoho estimator of multivariate location and scatter*, Stat. Probabil. Lett. 60 (2002), pp. 425–435.
[8] F.R. Hampel, *Contributions to the theroy of robust estimation*, Ph.D diss., University of California at Berkeley, 1968.
[9] F.R. Hampel, *The influence curve and its role in robust estimation*, J. Am. Stat. Assoc. 69 (1974), pp. 383–393.
[10] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel, *Robust Statistics: The approach based on influence functions*, Wiley, New York, 1986.
[11] V.I. Koltchinskii, *M-estimation, convexity and quantiles*, Ann. Stat., 25 (1997), pp. 435–477.
[12] R.Y. Liu, *On a notion of simplicial depth, in Proc. Nat. Acad. Sci.* 85 (1988), pp. 1732–1734.
[13] ———, *On a notion of data depth based on random simplices* Ann. Stat. 18 (1990), pp. 405–414.
[14] ———, *Data depth and multivariate rank tests*, In L_1*-statistical Analysis and Related Methods (Y. Dodge, ed.)*, 1992, pp. 279–294.
[15] R.Y. Liu, J.M. Parelius, and K. Singh, *Multivariate analysis by data depth: Descriptive statistics, graphics and inference (with discussion)*. Ann. Stat. 27 (1997), pp. 783–858.
[16] R. Maronna and V. Yohai, *The behavior of the Stahel-Donoho robust multivaraite estimator*, J. Am. Stat. Assoc. 90 (1995), pp. 330–343.
[17] M. Romanazzi, *Influence function of halfspace depth*, J. Multivariate Anal. 77 (2001), pp. 138–161.
[18] P.J. Rousseeuw and I. Ruts, *Constructing the bivariate Tukey median*, Stat. Sinica 8 (1998), pp. 828–839.
[19] R. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.
[20] ———, *A depth function and a scale curve based on spatial quantiles*, In *Statistical Data Analysis Based on the L1-Norm and Related Methods (Y. Dodge, ed.)*, 2002, pp. 25-28, Birkhaeuser.
[21] ———, *Nonparametric multivariate descriptive measures based on spatial quantiles*, J. Stat. Plan. Infer. 123 (2004), pp. 259-278.
[22] C.G. Small, C. G. (1990). *A survey of multidimensional medians*, Int. Stat. Rev. 58 (1990), pp. 263–277.
[23] J.W. Tukey, *Mathematics and picturing of data*, in *Proc. Int. Congr. Math.*, Vancouver, 2 (1975), pp. 523–531.
[24] Y. Vardi and C. Zhang, *The multivariate* L_1*-median and associated data depth, in Proc. Nat. Acad. Sci.* 97 (2000), pp. 1423–1426.
[25] J. Wang and R. Serfling, *Influence functions for a general class of depth-based generalized quantile functions*, J. Multivariate Anal. 97 (2006), pp. 810–826.
[26] J. Zhang, *Some extensions of Tukey's depth function*, J. Multivariate Anal. 82 (2002), pp. 134–165.
[27] Y. Zuo, *Projection-based depth functions and associated medians*, Ann. Stat. 31 (2003), pp. 1460–1490.
[28] Y. Zuo, H. Cui, and X. He, *On the Stahel-Donoho estimator and depth-weighted means of multivariate data*, Ann. Stat. 32 (2004), pp. 167–188.
[29] Y. Zuo, H. Cui, and D. Young, *Influence function an maximum bias of projection depth based on estimators*, Ann. Stat. 32 (2004), pp. 189–218.
[30] Y. Zuo and R. Serfling, *General notions of statistical depth function*, Ann. Stat. 28 (2000) pp. 461–482.
[31] ———, *Structural properties and convergence results for contours of sample statistical depth functions*, Ann. Stat. 28 (2000), pp. 483–499.