Sociedad Española de Estadística e Investigación Operativa

Test

Volume 15, Number 1. June 2006

Empirical Depth Processes

Miguel A. Arcones
Department of Mathematical Sciences
Binghamton University, USA
Hengjian Cui
Department of Mathematics
Beijing Normal University, China
Yijun Zuo
Department of Statistics and Probability
Michigan State University, USA

Sociedad de Estadística e Investigación Operativa
Test (2006) *Vol.* **15**, *No.* 1, pp. 151–177

Sociedad de Estadística e Investigación Operativa
Test (2006) *Vol.* **15**, *No.* 1, pp. 151–177

Empirical Depth Processes

Miguel A. Arcones*
Department of Mathematical Sciences
Binghamton University, USA

Hengjian Cui†
Department of Mathematics
Beijing Normal University, China

Yijun Zuo‡
Department of Statistics and Probability
Michigan State University, USA

Abstract

We consider the asymptotics of two depth processes, the projection depth process and certain generalizations of the Tukey halfspace depth process, and give natural conditions for the uniform convergence of these processes over certain subsets of $\mathbb{R}^d$. In general, these processes do not converge uniformly over $\mathbb{R}^d$.

Key Words: Projection depth, empirical projection depth process.

AMS subject classification: Primary 62E20; Secondary 62G35, 62G20.

1 Introduction

An interesting statistical problem is the construction of robust multivariate location and scatter estimators. Many of these estimators in the literature are obtained through depth functions. Reviews in this area are Small (1990), Liu et al. (1999), and Zuo and Serfling (2000). The key idea of depth functions is to provide a *center-outward* ordering of multivariate observations with respect to the underlying data set (or distribution). Points

*Correspondence to: Miguel A. Arcones. Department of Mathematical Sciences, Binghamton University. Binghamton, NY 13902. USA. E-mail: arcones@math.binghamton.edu

† Research partially supported by NSFC (No:10231030) and the EYTP of China.
‡ Research partially supported by NSF Grants DMS-0071976 and DMS-0134628.

Received: October 2003; Accepted: June 2004

deep inside the data cloud get higher depth and those on the outskirts get lower depth. Depth-induceding ordering provides promising new tools in multivariate data analysis and inference, just like order and rank statistics do for univariate data.

There are a number of notions of data depth, including Mahalanobis depth, halfspace depth (Tukey, 1975) and the simplicial depth (Liu, 1990). One favorable depth function considered in the literature is the projection depth function. Projection based location and scatter estimators possess very desirable properties, such as high robustness and high efficiency, and behave very well overall compared with their competitors. Consequently, they represent very favorable choices. The large sample properties of these estimators, such as asymptotic distribution, asymptotic efficiency, and convergence rate, are essential for statistical inference in practice, and are dictated by the asymptotic behavior of the empirical depth processes. The latter has not yet been thoroughly studied in the literature and is the focus of this paper. Results obtained in this paper can be utilized in the study of the aforesaid depth induced estimators (see Zuo and Cui, 2005; Zuo et al., 2004) as well as depth induced procedures such as testing and confidence region estimation.

The paper is organized as follows. In Section 2, we present several results on the weak convergence of supremum of empirical processes. These results are of independent interest and utilized in later sections. Many authors have used empirical processes to obtain asymptotic properties of depth statistics (see e.g. Arcones et al., 1994).

In Section 3, we study the weak convergence of the empirical projection depth process

$$\{G_n(x) := n^{1/2}(PD(x, F_n) - PD(x, F)) : x \in \mathbb{R}^d\}, \tag{1.1}$$

where $PD(\cdot, \cdot)$ is the "projection depth" defined in (3.3). We use $PD(x, F_n)$ to estimate $PD(x, F)$. The asymptotics of the projection depth process are used in the study of certain statistical properties of the projection depth. We see that in general this process does not converge when indexed over the whole space $\mathbb{R}^d$. For example, in the symmetric distribution case, we may have to exclude a neighborhood of the center of symmetry. The process then converges uniformly over the resulting space in $\mathbb{R}^d$. Projection depth is defined based on univariate location and scale estimators (see (3.3)). We show that our results are applied to a general class of location and scale estimators. In particular, our results apply to location estimators such as the mean, the median, and the Hodges–Lehmann estimator and scale estimators such as the sample variance, the median of the absolute deviations (MAD) and the interquantile range.

Section 4 deals with the weak convergence of the so-called halfspace depth process

$$\{H_n(x) := n^{1/2}(HD(x, F_n) - HD(x, F)) : x \in \mathbb{R}^d\}, \tag{1.2}$$

where *HD* stands for "halfspace depth" which is defined in (4.1). In an attempt to generalize the univariate median to the multivariate setting, Tukey (1975) introduced the notion of halfspace depth. Further discussions of halfspace depth were given in Donoho and Gasko (1992) and Rousseeuw and Ruts (1999). As in the projection depth process case, we show that the halfspace depth process can converge uniformly over the space $\mathbb{R}^d$ excluding a small neighborhood containing some "irregular" point(s) such as the center of symmetry of a symmetric distribution.

2 Convergence of certain supremum of stochastic processes

In this section, we present several results on the weak convergence of supremum of stochastic processes. We will use the general definition of weak convergence of stochastic processes in van der Vaart and Wellner (1996) and Dudley (1999), i.e. we consider stochastic processes as elements of $l_\infty(T)$, where T is an index set. $l_\infty(T)$ is the Banach space consisting of the bounded functions defined in T with the norm $\|x\|_\infty = \sup_{t\in T} |x(t)|$. To avoid measurability problems we will use outer measures when needed. To obtain the weak convergence of the process in (3.3), we will use the following theorem.

Theorem 2.1. *Let $\{Y_n(x, u) : x \in K, u \in T\}$ be a sequence of stochastic processes and $Y : K \times T \to \mathbb{R}$ be a function, where K is a set and (T, d) is a metric space. Suppose that:*

(i) For each $x \in K$, $u(x) = \{u \in T : Y(x, u) = L(x)\} \neq \emptyset$, where $L(x) := \sup_{u\in T} Y(x, u)$.

(ii) $\{n^{1/2}(Y_n(x, u) - Y(x, u)) : x \in K, u \in T\}$ converges weakly to a stochastic process $\{W(x, u) : x \in K, u \in T\}$.

(iii) For each $\delta > 0$,

$$\eta(\delta) := \inf_{x\in K} \inf_{u\notin u(x,\delta)} (L(x) - Y(x,u)) > 0,$$

where $u(x,\delta) = \{u \in T : d(u, u(x)) \le \delta\}$.

(iv) For each $\eta > 0$,

$$\lim_{\delta\to 0} \Pr{}^*\{\sup_{x\in K} |\sup_{u\in u(x,\delta)} W(x,u) - \sup_{u\in u(x)} W(x,u)| \ge \eta\} = 0.$$

Then,

$$\{n^{1/2}(\sup_{u\in T} Y_n(x,u) - \sup_{u\in T} Y(x,u)) : x \in K\} \xrightarrow{w} \{\sup_{u\in u(x)} W(x,u) : x \in K\}.$$

Asymptotics of the supremum of a empirical process were considered by Massé (2004) in the particular case of the Tukey depth process.

In order to get the asymptotic linear representation, we will use the following variant of the previous theorem.

Theorem 2.2. *With the notation above, let* $\{Z_n(x,u) : x \in K, u \in T\}$ *be another stochastic process. Suppose that:*

(i) For each $x \in K$, $u(x) = \{u \in T : Y(x,u) = L(x)\} \neq \emptyset$, *where* $L(x) := \sup_{u\in T} Y(x,u)$.

(ii) For each $M > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \Pr{}^*\{\sup_{x\in K} \sup_{u\notin u(x,\delta)} n^{1/2}(Y_n(x,u) - Y(x,u)) \ge M\} = 0.$$

(iii) For each $M > 0$,

$$\lim_{n\to\infty} \Pr{}^* \left\{ \sup_{x\in K} \left(- \sup_{u\notin u(x)} (Z_n(x,u) \right) \ge M \right\} = 0.$$

(iv) For each $\delta > 0$,

$$\eta(\delta) := \inf_{x\in K} \inf_{u\notin u(x,\delta)} (L(x) - Y(x,u)) > 0.$$

(v) *For each* $\eta > 0$,

$$\lim_{n\to\infty} \Pr{}^*\left\{ \sup_{x\in K} \sup_{u\in u(x,\delta)} |n^{1/2}(Y_n(x,u) - L(x)) - Z_n(x,u)| \geq \eta \right\}$$

converges to 0 as $\delta \to 0$.

(vi) *For each* $\eta > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \Pr{}^*\{\sup_{x\in K}(\sup_{u\in u(x,\delta)} Z_n(x,u) - \sup_{u\in u(x)} Z_n(x,u)) \geq \eta\} = 0.$$

Then, for each $\eta > 0$,

$$\Pr{}^*\left\{ \sup_{x\in K} |n^{1/2}(\sup_{u\in T} Y_n(x,u) - \sup_{u\in T} Y(x,u)) - \sup_{u\in u(x)} Z_n(x,u)| \geq \eta \right\}$$

converges to 0 as $n \to \infty$.

The proof of the previous theorem is omitted, since it is similar to that of Theorem 2.1.

The measurability conditions in the previous theorems are satisfied by the considered stochastic processes, because these stochastic process are determined by their values in a fixed countable subset of the index set.

3 Projection depth processes

One of the depth function considered in the literature is the projection depth function. In this section, we will study the properties of the projection depth process. Let μ and σ be location and scale functionals in $\mathbb{R}$. We assume that μ is *translation* and *scale equivariant* and σ is *scale equivariant* and *translation invariant*, that is, $\mu(F_{sY+c}) = s\mu(F_Y) + c$ and $\sigma(F_{sY+c}) = |s|\sigma(F_Y)$ respectively for any scalars s and c and random variable $Y \in \mathbb{R}$. The outlyingness of a point $x \in \mathbb{R}^d$ with respect to a given distribution function F of a $\mathbb{R}^d$–valued random vector X is defined as

$$O(x,F) = \sup_{\|u\|=1} \frac{|u'x - \mu(F_u)|}{\sigma(F_u)} = \sup_{\|u\|=1} g(x,u,F), \tag{3.1}$$

where F_u is the distribution of $u'X$, $u \in \mathbb{R}^d$, and

$$g(x, u, F) = \frac{\mu(F_u) - u'x}{\sigma(F_u)}. \tag{3.2}$$

The projection depth of a point $x \in \mathbb{R}^d$ with respect to the distribution F is then defined as

$$PD(x, F) = \frac{1}{1 + O(x, F)}. \tag{3.3}$$

When $\mu(F)$ is the median of F and $\sigma(F)$ is the median of the absolute deviations (MAD) of F, the projection function above is used in the construction of the Stahel–Donoho estimator (Donoho, 1982; Stahel, 1981). This estimator possesses a high breakdown point and has been studied by Donoho and Gasko (1992); Tyler (1994) and Maronna and Yohai (1995). In the generality above, the considered projection depth appeared in Zuo (2003).

Throughout our discussions, μ and σ are assumed to exist for all related univariate distributions. The empirical versions of $g(x, u, F)$, $O(x, F)$, and $PD(x, F)$ shall be denoted by $g(x, u, F_n)$, $O(x, F_n)$, and $PD(x, F_n)$ respectively, where F_n is the empirical distribution function. They are obtained by replacing F with its empirical version F_n in the corresponding formulae.

Since the projection depth function is based on a univariate location and scale functional, conditions on μ and σ are given first. We use F_{nu} as the empirical distribution function of $\{u'X_i,\ i = 1, \ldots, n\}$ for any $u \in \mathbb{R}^d$.

(C1) : $\sup_{\|u\|=1} |\mu(F_u)| < \infty, \quad \sup_{\|u\|=1} \sigma(F_u) < \infty, \text{ and } \inf_{\|u\|=1} \sigma(F_u) > 0.$

(C1)' : $\mu(F_u)$ and $\sigma(F_u)$ are continuous in u and $\sigma(F_u) > 0$.

(C2) : $\sup_{\|u\|=1} |\mu(F_{nu}) - \mu(F_u)| = o_P(1), \quad \sup_{\|u\|=1} |\sigma(F_{nu}) - \sigma(F_u)| = o_P(1).$

(C3) : $\sup_{\|u\|=1} |\mu(F_{nu}) - \mu(F_u)| = o(1)$ a.s., $\sup_{\|u\|=1} |\sigma(F_{nu}) - \sigma(F_u)| = o(1)$ a.s..

(C4) : $\sup_{\|u\|=1} \sqrt{n}|\mu(F_{nu}) - \mu(F_u)| = O_P(1), \quad \sup_{\|u\|=1} \sqrt{n}|\sigma(F_{nu}) - \sigma(F_u)| = O_P(1).$

(C4)' :

$$\{n^{1/2}\left(\mu(F_{nu})-\mu(F_u),\sigma(F_{nv})-\sigma(F_v)\right):\|u\|=1,\|v\|=1\}$$
$$\xrightarrow{\text{W}}\{(Z_\mu(u),Z_\sigma(v)):\|u\|=1,\|v\|=1\}$$

with $(Z_\mu(u),\, Z_\sigma(v))$ having continuous sample paths.

(C4)" : There are stochastic processes $\{Z_{\mu,n}(u):\|u\|=1\}$ and $\{Z_{\sigma,n}(u):\|u\|=1\}$ such that

$$\sup_{\|u\|=1}|\sqrt{n}(\mu(F_u)-\mu(F_u))-Z_{\mu,n}(u)|\xrightarrow{\text{Pr}}0,$$

$$\sup_{\|u\|=1}|\sqrt{n}(\sigma(F_u)-\sigma(F_u))-Z_{\sigma,n}(u)|\xrightarrow{\text{Pr}}0$$

and $\{(Z_{\mu,n}(u),Z_{\sigma,n}(v)):\|u\|=1,\|v\|=1\}\xrightarrow{\text{W}}\{(Z_\mu(u),Z_\sigma(v)):\|u\|=1,\|v\|=1\}$ with $(Z_\mu(u),\, Z_\sigma(v))$ having continuous sample paths.

In most cases, $Z_{\mu,n}(u)$ and $Z_{\sigma,n}(u)$ satisfy the above conditions.

We now consider the asymptotic behavior of $G_n(x)$ defined in (1.1). The following result is utilized later.

Theorem 3.1. *Under* **(C1)** *we have*

(1) $\sup_{x\in\mathbb{R}^d}(1+\|x\|)\,|PD(x,F_n)-PD(x,F)|=o_P(1)$ *or* $o(1)$ *a.s. if* **(C2)** *or* **(C3)** *holds,*

(2) $\sup_{x\in\mathbb{R}^d}(1+\|x\|)\sqrt{n}\,|PD(x,F_n)-PD(x,F)|=O_P(1)$ *if* **(C4)** *holds.*

For any x, let $u(x)$ be the set of directions satisfying $O(x,F)=g(x,u,F)$. If $u(x)$ is a singleton, we also use $u(x)$ as the unique direction. If X is a continuous random variable, nonuniqueness of $u(x)$ may occur at finitely many points. Under minimal conditions, it is possible to get the asymptotic normality of $G_n(x)$ for a fix x.

Theorem 3.2. *Assume* **(C1)'** *and* **(C4)'**. *Then, for each* $x \in \mathbb{R}^d$,

$$\sqrt{n}\,(PD(x,F_n) - PD(x,F)) \xrightarrow{d} -\sup_{u\in u(x)} Z(x,u)$$

where

$$Z(x,u) = \frac{Z_\mu(u) - O(x,F)Z_\sigma(u)}{\sigma(F_u)(1+O(x,F))^2}.$$

Remark 3.1. *For a given* x *where* $u(x)$ *is a singleton, the distribution of* $Z(x,u(x))$ *is typically Gaussian. Cui and Tian (1994) established the pointwise limiting distribution of* $G_n(x)$ *for a special case when* μ *and* σ *are the median and MAD functionals respectively.*

It is not possible to get the weak convergence of the projection depth process over the whole $\mathbb{R}^d$. Points x with $u(x)$ different from a singleton present a problem. The following theorems provide sufficient conditions for the weak convergence of the projection depth process over certain subsets of $\mathbb{R}^d$.

Theorem 3.3. *Assume that* **(C1)'** *and* **(C4)'**. *Let* $T \subset \mathbb{R}^d$ *be a set such that for each* $M > 0$ *and each* $\delta > 0$,

$$\inf_{\substack{x\in T\\ \|x\|\le M}} \inf_{u\notin u(x,\delta)} (O(x,F) - g(x,u,F)) > 0, \tag{3.4}$$

where $u(x,\delta) = \{u \in \mathbb{R}^d : \|u\| = 1, d(u,u(x)) \le \delta\}$. *Then,*

$$\{\sqrt{n}\,(PD(x,F_n) - PD(x,F)) : x \in T\} \xrightarrow{\mathrm{w}} \{-\sup_{u\in u(x)} Z(x,u) : x \in T\}.$$

Corollary 3.1. *Assume that* **(C1)'** *and* **(C4)'**. *Suppose that* $u(x)$ *consists of a singleton except for finitely many points* $\{y_1,\ldots,y_m\}$. *Then, for each* $\delta > 0$,

$$\left\{\sqrt{n}\,(PD(x,F_n) - PD(x,F)) : x \in \mathbb{R}^d - \bigcup_{j=1}^m B(y_j,\delta)\right\} \xrightarrow{\mathrm{w}} \left\{-Z(x,u(x)) : x \in \mathbb{R}^d - \bigcup_{j=1}^m B(y_j,\delta)\right\},$$

where $B(y,\delta) = \{x \in \mathbb{R}^d : \|y-x\| < \delta\}$.

It is possible to see in simplest cases that it is not possible to get the weak convergence of the projection depth process over the whole $\mathbb{R}^d$. We present the following:

Proposition 3.1. *Let $d = 1$, $\mu(F) = \int_{-\infty}^{\infty} x\, F(dx)$, and $\sigma(F) = \left(\int_{-\infty}^{\infty}(x - \mu(F))^2\, F(dx)\right)^{1/2}$. Suppose that X has finite fourth moment. Then, the finite dimensional distributions of $\{n^{1/2}(O(x, F_n) - O(x, F)) : x \in \mathbb{R}\}$ converge, but the process does not converge weakly.*

The proof of the previous proposition indicates that the process $\{n^{1/2} \cdot(O(x, F_n) - O(x, F)) : x \in I\}$ does not converge weakly for neither $I = (\mu, \mu + \delta)$ nor $I = (\mu - \delta, \mu)$ for any $\delta > 0$.

Using Theorem 2.2 and arguments in the proof of the previous theorems, it is possible to prove that projection depth process is asymptotically linear. The proof of the next theorem is omitted.

Theorem 3.4. *Assume that* **(C1)'** *and* **(C4)"**. *Suppose that $u(x)$ consists of a singleton except for finitely many points $\{y_1, \dots, y_m\}$ with $\Pr\{X = y_j\} = 0$. Then, for each $\delta > 0$,*

$$\sup_{x \in \mathbb{R}^d - \cup_{j=1}^m B(y_j, \delta)} \Big| \sqrt{n}\,(PD(x, F_n) - PD(x, F)) + \frac{Z_{\mu,n}(u(x)) - O(x, F) Z_{\sigma,n}(u(x))}{\sigma(F_{u(x)})(1 + O(x, F))^2} \Big| \xrightarrow{\Pr} 0.$$

We will consider location and scale M-estimators defined through U-statistics. Given a function $h : \mathbb{R}^m \to \mathbb{R}$, the U-statistic with kernel h is defined by

$$\frac{(n-m)!}{n!} \sum_{(i_1,\dots,i_k) \in I_m^n} h(X_{i_1}, \dots, X_{i_k}),$$

where $I_k^n = \{(i_1, \dots, i_k) : 1 \le i_j \le n \text{ and } i_l \ne i_j \text{ for } j \ne l\}$.

The previous theorem reduces the problem to the delta method holding uniformly on $\{u \in \mathbb{R}^d : \|u\| = 1\}$. The following theorem gives sufficient conditions for the first order expansion of a class of projection M-estimators to go to zero in probability.

Theorem 3.5. *Let $\psi : \mathbb{R} \to \mathbb{R}$ be a nonincreasing function. Let k be a positive integer. Suppose that:*

(i) *For each* $\|u\| = 1$*, there exists a* $\theta_0(u) \in \mathbb{R}$ *such that* $H(\theta_0(u), u) = 0$*, where* $H(\theta, u) = E[\psi(k^{-1}u'(X_1 + \cdots + X_k) - \theta)]$.

(ii) $0 < \inf_{\|u\|=1} H'(\theta_0(u), u) = 0$*, where* $H'(\theta, u)$ *is the derivative with respect to* θ *at* $\theta_0(u)$.

(iii)

$$\lim_{h\to 0} \sup_{\|u\|=1} |h^{-1}(H(\theta_0(u) + h), u) - H(\theta_0(u), u)) - H'(\theta_0(u), u)| = 0.$$

(iv)

$$\lim_{h\to 0} \sup_{\|u\|=1} \operatorname{Var}(\psi(k^{-1}u'(X_1 + \cdots + X_k) - \theta_0(u) + h) \\ - \psi(k^{-1}u'(X_1 + \cdots + X_k) - \theta_0(u))) = 0.$$

(v) *There exists a* $\delta > 0$ *such that*

$$E[\sup_{\|u\|=1} \sup_{|\theta-\theta_0(u)|\le\delta} |\psi(k^{-1}u'(X_1 + \cdots + X_k) - \theta)|^2] < \infty.$$

Then,

$$\sup_{\|u\|=1} |n^{1/2}(\hat{\theta}_n(u) - \theta_0(u)) + (H'(\theta_0(u), u))^{-1} n^{1/2} H_n(\theta_0(u), u)| \overset{\Pr}{\to} 0,$$

where

$$H_n(\theta, u) = \frac{(n-k)!}{n!} \sum_{(i_1,\ldots,i_k)\in I_k^n} \psi(k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) - \theta),$$

$\hat{\theta}_n^*(u) = \sup\{t : H_n(t, u) < 0\}$, $\hat{\theta}_n^{**}(u) = \inf\{t : H_n(t, u) > 0\}$, *and* $\hat{\theta}_n(u) = (\hat{\theta}_n^*(u) + \hat{\theta}_n^{**}(u))/2$.

For some popular location and scale estimator including generalized medians and generalized medians of the absolute deviations, we check **(C4)”** in Appendix B.

4 Generalized Tukey halfspace depth processes

In this section, we study the properties of a generalization of the Tukey halfspace depth process. Tukey (1975) generalized the median to several dimensions, using the smallest probability of halfspace containing a point. Given $x \in \mathbb{R}^d$ and $\|u\| = 1$, let $H(x,u) = \{y \in \mathbb{R}^d : u'(y-x) \geq 0\}$. The Tukey halfspace depth is defined by

$$HD(x,F) = \inf_{\|u\|=1} F(H(x,u)), \tag{4.1}$$

where $F(H(x,u))$ means the probability of $H(x,u)$ with respect to the distribution determined by F. Further properties of the halfspace depth function was proved by Rousseeuw and Ruts (1999) and Massé (2002) proved the asymptotic distribution of the Tukey's halfspace median. Massé (2004) study the weak convergence of the process

$$\{H_n(x) := n^{1/2}(HD(x,F_n) - HD(x,F)) : x \in \mathbb{R}^d\}. \tag{4.2}$$

and gave applications to L–statistics over this depth function.

We will consider generalizations of the Tukey halfspace depth function. A way to generalize the Tukey halfspace depth is:

$$GHD(x,h,F) = \inf_{\|u\|=1} \int h(u'(y-x))dF(y) \tag{4.3}$$

where h is a nondecreasing nonnegative function with $h(x) = 0$ for $x < 0$ and $\mathrm{E}h(\|Y\|) < +\infty$. If $h(x) = I\{x \geq 0\}$, $GHD(x,h,F)$ is just the Tukey halfspace depth. If X has spherical distribution with center μ, then $GHD(\mu,h,F) = \max_{x\in\mathbb{R}^d} GHD(x,h,F)$ and $\lim_{\|x\|\to\infty} GHD(x,h,F) = 0$.

Sufficient conditions for the weak convergence of

$$\{n^{1/2}(GHD(x,h,F_n) - GHD(x,h,F)) : x \in K\} \tag{4.4}$$

follow from the results in Section 2. For any x, let $u(x)$ be the set of directions satisfying $\int h(u'(y-x))dF(y) = GHD(x,h,F)$.

We consider the following conditions for a function h and a set K of $\mathbb{R}^d$:

(D1) : $E[\sup_{\substack{x\in K\\ \|u\|=1}} |h(u'(X-x))|^2] < \infty.$

(D2) : $\lim_{\delta\to 0}\sup_{x\in K}\sup_{\substack{\|u_1-u_2\|\le\delta\\ \|u_1\|=\|u_2\|=1}} \mathrm{Var}(h(u_1'(X-x))-h(u_2'(X-x)))=0.$

(D3) : $\lim_{M\to\infty}\sup_{\substack{x\in K\\ \|x\|\ge M}}\sup_{\|u\|=1}\mathrm{Var}(h(u'(X-x)))=0.$

(D4) : For each $M>0$ and each $\delta>0$,

$$\inf_{\substack{x\in K\\ \|x\|\le M}}\inf_{u\notin u(x,\delta)}\left(\int h(u'(y-x))dF(y)-GHD(x,h,F)\right)>0,$$

where $u(x,\delta)=\{u\in\mathbb{R}^d:\|u\|=1,d(u,u(x))\le\delta\}$.

For the Tukey halfspace depth ($h(x)=I(x\ge 0)$) conditions **(D1)**–**(D3)** are satisfied for $K=\mathbb{R}^d$ if for each $x\in\mathbb{R}^d$ and each $\|u\|=1$, $P(u'(X-x)=0)=0$. Under the previous condition, **(D4)** is also satisfied for a closed set K consisting of points x such that $u(x)$ is a singleton.

Condition *(i)* in Theorem 2.1 follows from the following lemma:

Lemma 4.1. *Let h be a real function and let K be a set of $\mathbb{R}^d$ satisfying condition* **(D1)**. *Then,*

$$\{n^{-1/2}\sum_{j=1}^{n}(h(u'(X_j-x))-E[h(u'(X-x))]):x\in K,\|u\|=1\}$$
$$\xrightarrow{\mathrm{w}}\{V_h(x,u):x\in K,\|u\|=1\},$$

where $\{V_h(x,u):x\in K,\|u\|=1\}$ is a Gaussian process with mean zero and covariance given by

$$E[V_h(x_1,u_1)V_h(x_2,u_2)]=\mathrm{Cov}(h(u_1'(X-x_1)),h(u_2'(X-x_2))).$$

If $h(x)=x$ or $h(x)=x^2$ is not possible to obtain the weak convergence of the stochastic process in the previous lemma for unbounded sets K.

Theorem 4.1. *Let K be a set of $\mathbb{R}^d$ satisfying* **(D1)**–**(D4)**. *Then,*

$$\{\sqrt{n}\,(GHD(x,h,F_n)-GHD(x,h,F)):x\in T\}\xrightarrow{\mathrm{w}}\{\inf_{u\in u(x)}V_h(x,u):x\in T\}.$$

Corollary 4.1. *Assume that* **(D1)**–**(D3)**. *Suppose that* $u(x)$ *consists of a singleton except for finitely many points* $\{y_1, \dots, y_m\}$. *Then, for each* $\delta > 0$,

$$\left\{ \sqrt{n}\,(GHD(x,h,F_n) - GHD(x,h,F)) : x \in \mathbb{R}^d - \bigcup_{j=1}^{m} B(y_j,\delta) \right\}$$
$$\xrightarrow{\mathrm{w}} \left\{ V_h(x,u(x)) : x \in \mathbb{R}^d - \bigcup_{j=1}^{m} B(y_j,\delta) \right\},$$

where $B(y,\delta) = \{x \in \mathbb{R}^d : \|y - x\| < \delta\}$.

Theorem 4.2. *Assume that* **(D1)**–**(D3)**. *Suppose that* $u(x)$ *consists of a singleton except for finitely many points* $\{y_1, \dots, y_m\}$ *with* $\Pr\{X = y_j\} = 0$. *Then, for each* $\delta > 0$,

$$\sup_{x \in \mathbb{R}^d - \cup_{j=1}^{m} B(y_j,\delta)} n^{1/2} \left| (GHD(x,h,F_n) - GHD(x,h,F)) - (F_n - F)h(u(x)'(\cdot - x)) \right| \xrightarrow{\Pr} 0.$$

Another way to generalize the Tukey halfspace depth median is through outlyingness. Given a monotone function h and one dimensional data, an estimator is defined as a value $\hat{\theta}_n$ such that $n^{-1}\sum_{j=1}^{n} h(X_j - \hat{\theta}_n)$ is approximately zero. Equivalently, $\hat{\theta}_n$ is a value which minimizes $|n^{-1}\sum_{j=1}^{n} h(X_j - x)|$, $x \in \mathbb{R}$. In the multivariate case, we may define the Tukey outlyingness with respect to the function h as

$$TO(x,h,F) = \sup_{\|u\|=1} \int h(u'(y-x))\,dF(y). \tag{4.5}$$

The empirical Tukey halfspace outlyingness is defined as $TO(x,h,F_n)$. We call a value which minimizes $TO(x,h,F_n), x \in \mathbb{R}^d$, a generalized Tukey halfspace median with respect to the function h. If $h(x) = 1 - 2I(x \geq 0)$, then $TO(x,h,F) = 1 - 2HD(x,F)$ and we have the usual Tukey halfspace median. If $h(x) = x$, then $TO(x,h,F_n) = |\bar{X} - x|$ and the generalized Tukey halfspace median is the mean. As in the one dimensional case, monotone functions h which grow as $|x|$ goes to infinity slower than $|x|$ are prefered. One such function is $h(x) = x$, for $|x| \leq k$ and $h(x) = k\mathrm{sign}(x)$, for $|x| > k$.

Another possible function is $h(x) = |x|^{p-1}\text{sign}(x)$, where $1 \le p < 2$. More functions like that are in Chapter 7 in Serfling (1980).

Massé (2004) also obtained similar results for the Tukey halfspace depth process via a different approach. Other ways to generalized the Tukey halfspace depth process are in Zhang (2002).

The process convergence results in Sections 3 and 4 can be utilized immediately for the study of the asymptotic behavior of estimators induced from the corresponding depth functions, in particular in the study of L–statistics over depth functions (see Zuo et al., 2004).

Appendix A: Proofs

Proof of Theorem 2.1. By a representation theorem (see for example Theorem 3.5.1 in Dudley, 1999), we have a version of the stochastic processes converging a.s. So, we may assume that

$$\sup_{x\in K, u\in T} \left| n^{1/2}(Y_n(x,u) - Y(x,u)) - W(x,u) \right| \to 0 \text{ a.s.}$$

and the process W satisfies *(iv)*. We have that

$$\begin{aligned} & n^{1/2}\left(\sup_{u\in T} Y_n(x,u) - \sup_{u\in T} Y(x,u)\right) \\ & = n^{1/2}\max\left(\sup_{u\in u(x,\delta)}(Y_n(x,u) - L(x)), \sup_{u\notin u(x,\delta)}(Y_n(x,u) - L(x))\right). \end{aligned}$$

Now,

$$\begin{aligned} & n^{1/2}\sup_{u\notin u(x,\delta)}(Y_n(x,u) - L(x)) \\ & \qquad \le n^{1/2}\sup_{u\notin u(x,\delta)}(Y_n(x,u) - Y(x,u) - \eta(\delta)) \overset{\Pr}{\to} -\infty, \end{aligned}$$

uniformly in $x \in K$. We also have that

$$\begin{aligned} & \sup_{u\in u(x)} W(x,u) \overset{\text{a.s.}}{\longleftarrow} n^{1/2}\sup_{u\in u(x)}(Y_n(x,u) - Y(x,u)) \\ & = n^{1/2}\sup_{u\in u(x)}(Y_n(x,u) - L(x)) \le n^{1/2}\sup_{u\in u(x,\delta)}(Y_n(x,u) - L(x)) \\ & \le n^{1/2}\sup_{u\in u(x,\delta)}(Y_n(x,u) - Y(x,u)) \overset{\text{a.s.}}{\longrightarrow} \sup_{u\in u(x,\delta)} W(x,u), \end{aligned}$$

uniformly on $x \in K$. Hence, the claim follows. □

Proof of Theorem 3.1. We prove only part (2). The proof for part (1) is similar and thus omitted. It is observed that

$$\left| \inf_{\|u\|=1} \sigma(F_{nu}) - \inf_{\|u\|=1} \sigma(F_u) \right| \quad \le \quad \sup_{\|u\|=1} |\sigma(F_{nu}) - \sigma(F_u)|.$$

By condition **(C4)**, $\inf_{\|u\|=1} \sigma(F_{nu}) \to \inf_{\|u\|=1} \sigma(F_u)$ in probability as $n \to \infty$. Consequently, $\inf_{\|u\|=1} \sigma(F_{nu})$ is bounded below from 0 in probability as $n \to \infty$. Write $l_n(u) = \mu(F_{nu}) - \mu(F_u)$ and $s_n(u) = \sigma(F_{nu}) - \sigma(F_u)$. Note that

$$|O(x, F_n) - O(x, F)| \le \sup_{\|u\|=1} \frac{|u'x||s_n(u)| + |\mu(F_u)||s_n(u)| + \sigma(F_u)|l_n(u)|}{\sigma(F_{nu})\sigma(F_u)}$$

$$\le \|x\| Q_n + R_n,$$

where

$$Q_n = \frac{\sup_{\|u\|=1} |s_n(u)|}{\inf_{\|u\|=1} (\sigma(F_{nu})\sigma(F_u))}$$

and

$$R_n = \frac{\sup_{\|u\|=1} |\mu(F_u)| \sup_{\|u\|=1} |s_n(u)| + \sup_{\|u\|=1} \sigma(F_u) \sup_{\|u\|=1} |l_n(u)|}{\inf_{\|u\|=1} (\sigma(F_{nu})\sigma(F_u))}.$$

By conditions **(C1)** and **(C4)** and the boundedness of $\inf_{\|u\|=1} \sigma(F_{nu})$ away from 0, it is readily seen that $|O(x, \hat{F}_n) - O(x, F)| = o_P(1)$ uniformly for bounded x and that $\sqrt{n} Q_n$ and $\sqrt{n} R_n$ are $O_P(1)$. Note further that

$$|PD(x, F_n) - PD(x, F)| \quad = \quad \frac{|O(x, F_n) - O(x, F)|}{(1 + O(x, F_n))(1 + O(x, F))}$$

$$\le \quad |O(x, F_n) - O(x, F)|.$$

For any fixed $M > 0$, we have $\sup_{\|x\| \le M} |G_n(x)| = O_P(1)$. Take $M > \sup_{\|u\|=1} |\mu(F_u)|$. Then, we see that for sufficiently large n

$$|G_n(x)| \quad \le \quad \sup_{\|u\|=1} \sigma(F_{nu}) \sup_{\|u\|=1} \sigma(F_u) \times \frac{\sqrt{n}(\|x\| Q_n + R_n)}{\left(\|x\| - \sup_{\|u\|=1} |\mu(F_{nu})|\right) \left(\|x\| - \sup_{\|u\|=1} |\mu(F_u)|\right)}. \tag{A.1}$$

From **(C1)**, **(C4)** and the boundedness of $\sqrt{n}Q_n$ and $\sqrt{n}R_n$ it follows that $\sup_{\|x\|>M} \|x\||G_n(x)| = O_P(1)$. Hence, the claims follows. □

Proof of Theorem 3.2. Since

$$n^{1/2}\left(\frac{\mu(F_{nu})-u'x}{\sigma(F_{nu})}-\frac{\mu(F_u)-u'x}{\sigma(F_u)}\right)$$
$$=\frac{n^{1/2}(\mu(F_{nu})-\mu(F_u))\sigma(F_u)+n^{1/2}(u'x-\mu(F_u))(\sigma(F_{nu})-\sigma(F_u))}{\sigma(F_u)\sigma(F_{nu})},$$

we have that

$$\{n^{1/2}(g(x,u,F_n)-g(x,u,F)) : \|u\|=1\}$$
$$\overset{\mathrm{w}}{\to}\left\{\frac{\sigma(F_u)Z_\mu(u)+(u'x-\mu(F_u))Z_\sigma(u)}{\sigma^2(F_u)} : \|u\|=1\right\}.$$

So, by Theorem 2.1, for each $x \in \mathbb{R}^d$,

$$n^{1/2}(O(x,F_n)-O(x,F)) \overset{\mathrm{d}}{\to} \sup_{u\in u(x)}\left\{\frac{Z_\mu(u)\sigma(F_u)+(u'x-\mu(F_u))Z_\sigma(u)}{\sigma^2(F_u)}\right\}.$$

This implies that for each $x \in \mathbb{R}^d$,

$$\sqrt{n}\,(PD(x,F_n)-PD(x,F)) = \frac{-n^{1/2}(O(x,F_n)-O(x,F))}{(1+O(x,F))(1+O(x,F_n))}$$
$$\overset{d}{\longrightarrow} -\sup_{u\in u(x)} Z(x,u)$$

where

$$Z(x,u)=\frac{Z_\mu(u)-O(x,F)Z_\sigma(u)}{\sigma(F_u)(1+O(x,F))^2}.$$

□

Proof of Theorem 3.3. By (A.1), for each $\epsilon > 0$,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\Pr\{\sup_{\|x\|\geq M}|n^{1/2}(PD(x,F_n)-PD(x,F))|\geq\epsilon\}=0.$$

So, it suffices to show that for each $0 < M < \infty$,

$$\{n^{1/2}(PD(x,F_n) - PD(x,F)) : x \in T, \|x\| \le M\} \\ \xrightarrow{w} \{-\sup_{u \in u(x)} Z(x,u) : x \in T, \|x\| \le M\}.$$

This follows from the fact that for each $M > 0$,

$$\{n^{1/2}(O(x,F_n) - O(x,F)) : x \in T, \|x\| \le M\} \\ \xrightarrow{\mathrm{w}} \left\{ \sup_{u \in u(x)} \frac{\sigma(F_u)Z_\mu(u) + (u'x - \mu(F_u))Z_\sigma(u)}{\sigma^2(F_u)} : x \in T, \|x\| \le M \right\}.$$

To get this, we apply Theorem 2.1. It is obvious that for each $M > 0$,

$$\{n^{1/2}(g(x,u,F_n) - g(x,u,F)) : \|x\| \le M, \|u\| = 1\} \\ \xrightarrow{\mathrm{w}} \left\{ \frac{\sigma(F_u)Z_\mu(u) + (u'x - \mu(F_u))Z_\sigma(u)}{\sigma^2(F_u)} : \|x\| \le M, \|u\| = 1 \right\}.$$

The rest of the conditions in Theorem 2.1 hold trivially. □

Proof of Corollary 3.1. We apply Theorem 3.3. First, we note $O(x,F)$ is continuous, because

$$|O(x_1,F) - O(x_2,F)| \le \sup_{\|u\|=1} \left| \frac{\mu(F_u) - u'x_1}{\sigma(F_u)} - \frac{\mu(F_u) - u'x_2}{\sigma(F_u)} \right| \\ \le \frac{\|x_1 - x_2\|}{\inf_{\|u\|=1} \sigma(F_u)}.$$

We claim that u is a continuous function in $\mathbb{R}^d - \cup_{j=1}^m \{y_j\}$. Take $x \in \mathbb{R}^d - \cup_{j=1}^m \{y_j\}$. If $u(x)$ is not continuous at x, then there exists a sequence $x_n \to x$ such that $u(x_n) \not\to u(x)$. Since $\|u(x_n)\| = 1$, we may assume that $u(x_n) \to u_0 \ne u(x)$. Since $O(x_n,F) \to O(x,F)$,

$$\frac{\mu(F_{u(x_n)}) - (u(x_n))'x_n}{\sigma(F_{u(x_n)})} \to \frac{\mu(F_{u(x)}) - (u(x))'x}{\sigma(F_{u(x)})}.$$

But,

$$\frac{\mu(F_{u(x_n)}) - (u(x_n))'x_n}{\sigma(F_{u(x_n)})} \to \frac{\mu(F_{u_0}) - u_0'x}{\sigma(F_{u_0})},$$

in contradiction. The continuity of $u(x)$ implies that condition (3.4) holds. □

Proof of Proposition 3.1. We have that $O(x,F) = \frac{|x-\mu|}{\sigma}$ and $O(x,F_n) = \frac{|x-\bar{x}|}{\sigma_n}$, where $\mu = E[X]$, $\sigma^2 = E[(X-\mu)^2]$, $\bar{x} = n^{-1}\sum_{i=1}^n X_i$ and $\sigma_n^2 = n^{-1}\sum_{i=1}^n (X_i - \bar{x})^2$. By the Central Limit Theorem

$$(n^{1/2}(\bar{x}-\mu), n^{1/2}(\sigma_n^2 - \sigma^2)) \xrightarrow{d} (Z_1, Z_2)$$

where (Z_1, Z_2) is a Gaussian random vector with zero means and covariances given by

$$E[Z_1^2] = \sigma^2,\ E[Z_1 Z_2] = E[(X-\mu)((X-\mu)^2 - \sigma^2)] \ \text{ and}$$
$$E[Z_2^2] = E[((X-\mu)^2 - \sigma^2)^2].$$

It is easy to see that the finite dimensional distributions of

$$\left\{ L_n(x) := n^{1/2}\left(\frac{|x-\bar{x}|}{\sigma_n} - \frac{|x-\mu|}{\sigma}\right) : x \in \mathbb{R} \right\}$$

converges to those of $\{W(x) : x \in \mathbb{R}\}$, where $W(x) = -\sigma^{-1}Z_1 - 2^{-1}\sigma^{-3}(x-\mu)Z_2$, for $x > \mu$; $W(x) = \sigma^{-1}Z_1 + 2^{-1}\sigma^{-3}(x-\mu)Z_2$, for $x < \mu$; and $W(\mu) = \sigma^{-1}|Z_1|$. If we had weak convergence of the process $n^{1/2}\left(\frac{|x-\bar{x}|}{\sigma_n} - \frac{|x-\mu|}{\sigma}\right)$, then for each $\eta > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \Pr\{\sup_{\substack{d(t_1,t_2)\le\delta \\ t_1,t_2\in\mathbb{R}}} |L_n(t_1) - L_n(t_2)| \ge \eta\} = 0,$$

where $d(t_1,t_2) = E[\min(|W(t_1) - W(t_2)|, 1)]$ (see Theorem 3 in Arcones, 1996). This implies that for each $\eta > 0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \Pr\{\sup_{\substack{\mu < t_1,t_2 \le \mu+1 \\ |t_1-t_2|\le\delta}} |L_n(t_1) - L_n(t_2)| \ge \eta\} = 0. \tag{A.2}$$

However, this condition does not hold. Take $0 < \delta < 1/2$. Suppose that $\mu \le \bar{x} \le \mu + 2\delta$, take $t_1 = \bar{x}$ and $t_2 = \frac{\bar{x}+\mu}{2}$. Then, $\mu \le t_1, t_2 \le \mu + 1$, $|t_2 - t_1| \le \delta$, $L_n(\frac{\bar{x}+\mu}{2}) - L_n(\bar{x}) = \frac{n^{1/2}|\bar{x}-\mu|(\sigma_n+\sigma)}{2\sigma_n\sigma}$. So,

$$\Pr\{\sup_{\substack{\mu < t_1,t_2 \le \mu+1 \\ |t_1-t_2|\le\delta}} |L_n(t_1) - L_n(t_2)| \ge \eta\}$$
$$\ge \Pr\{\mu \le \bar{x} \le \mu + 2\delta, \frac{n^{1/2}|\bar{x}-\mu|(\sigma_n+\sigma)}{2\sigma_n\sigma} \ge \eta\} \to \Pr\{Z_1 \ge \eta\sigma\} > 0,$$

which contradicts (A.2). ☐

Proof of Theorem 4.1. Since the class of functions $g_{x,u}(y) = u'(y-x)$ is a $d+1$ dimensional vector space of functions, by Theorem 4.2.1 in Dudley (1999), the class of functions, $\{g_{x,u} : x \in \mathbb{R}^d, \|u\| = 1\}$ is a VC–subgraph class of functions. By Theorem 4.2.3 in Dudley (1999), so is the class of functions $\{h \circ g_{x,u} : x \in \mathbb{R}^d, \|u\| = 1\}$. Hence, the claim follows from the Pollard's central limit theorem (see e.g. Theorem 6.3.1 in Dudley, 1999). □

Proof of Theorem 4.2. We claim that for each $\epsilon > 0$,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \Pr\{ \sup_{\substack{x\in K \\ \|x\|\geq M}} |n^{1/2}(GHD(x,h,F_n) - GHD(x,h,F))| \geq \epsilon\} = 0. \quad \text{(A.3)}$$

Note that

$$\begin{aligned} &\sup_{\|x\|\geq M} |n^{1/2}(GHD(x,h,F_n) - GHD(x,h,F))| \\ &\leq \sup_{\substack{x\in K \\ \|x\|\geq M}} \sup_{\|u\|=1} |n^{1/2}(F_n - F)(h(u'(\cdot - x))| \\ &\xrightarrow{\text{w}} \sup_{\substack{x\in K \\ \|x\|\geq M}} \sup_{\|u\|=1} |V_h(x,u)| \end{aligned}$$

Now, $\{V(x,u) : x \in \mathbb{R}^d, \|u\| = 1\}$ has a version which is bounded and uniformly continuous with the respect to the distance

$$d((x_1,u_1),(x_2,u_2)) = \mathrm{Var}(h(u_1'(X - x_1)) - h(u_2'(X - x_2))).$$

By **(D3)**, have that

$$\sup_{\substack{x\in K \\ \|x\|\geq M}} \sup_{\|u\|=1} E[(V_h(x,u))^2] = \sup_{\substack{x\in K \\ \|x\|\geq M}} \sup_{\|u\|=1} \mathrm{Var}(h(u'(X-x)))$$

which tends to zero as $M \to \infty$. So, (A.3) holds.

By (A.3), we may assume that K is a bounded set.

We apply Theorem 2.1 with $Y_n(x,u) = -F_n h(u'(\cdot - x))$ and $Y(x,u) = -Fh(u'(\cdot - x))$ and $W(x,u) = -V(x,u)$. Condition *(i)* in Theorem 2.1

follows from Lemma 4.1. Condition *(ii)* in Theorem 2.1 is assumed. By **(D2)**, with probability one,

$$\lim_{\delta \to 0} \sup_{\substack{x \in K \\ \|x\| \le M}} \sup_{\|u_1 - u_2\| \le \delta} |V_h(x, u_2) - V_h(x, u_1)| = 0.$$

□

Corollary 4.1 follows from Theorem 4.1 directly.

Appendix B: Uniform delta method for projection estimators

To prove that the considered projection location parameters are asymptotically linear uniformly on the projection parameter, we use the following theorem:

Theorem B.1. *Let $\{Z_n(\theta, u) : \theta \in \mathbb{R}, u \in T\}$ be a sequence of stochastic processes, where T is an index set. Let $\theta_0 : T \to \mathbb{R}$ and $b : T \to \mathbb{R}$ be two functions. Suppose that for each $u \in T$ and each n, $Z_n(\theta, u)$ is nondecreasing function in θ. Let $\{a_n\}$ be a sequence of real numbers converging to infinity. Let $\hat{\theta}_n^*(u) = \sup\{t : Z_n(t, u) < 0\}$, $\hat{\theta}_n^{**}(u) = \inf\{t : Z_n(t, u) > 0\}$, and $\hat{\theta}_n(u) = (\hat{\theta}_n^*(u) + \hat{\theta}_n^{**}(u))/2$. Assume that:*

(i) $\inf_{u \in T} b(u) > 0$.

(ii) *For each* $0 < M < \infty$,

$$\sup_{u \in T} \sup_{|\tau| \le M} |a_n(Z_n(\theta_0(u) + a_n^{-1}\tau, u) - Z_n(\theta_0(u), u)) - b(u)\tau| \xrightarrow{\Pr} 0.$$

(iii) $\sup_{u \in T} a_n |Z_n(\theta_0(u), u)| = O_P(1)$.

Then,

$$\sup_{u \in T} |a_n(\hat{\theta}_n(u) - \theta_0(u)) + (b(u))^{-1} a_n Z_n(\theta_0(u), u)| \xrightarrow{\Pr} 0.$$

Proof. Given $\tau > 0$, we prove that

$$\Pr\{b(u)a_n(\hat{\theta}_n(u) - \theta_0(u)) + a_n Z_n(\theta_0(u), u) < -\tau, \text{ for some } u \in T\} \to 0 \tag{B.1}$$

and

$$\Pr\{b(u)a_n(\hat{\theta}_n(u) - \theta_0(u)) + a_n Z_n(\theta_0(u), u) \le \tau \text{ for some } u \in T\} \to 1. \tag{B.2}$$

For each $t \in \mathbb{R}$ and $u \in T$, $\{\hat{\theta}_n(u) < t\} \subset \{Z_n(t, u) > 0\}$. So,

$$\begin{aligned} &\{b(u)a_n(\hat{\theta}_n(u) - \theta_0(u)) + a_n Z_n(\theta_0, u) < -\tau \text{ for some } u \in T\} \\ &= \{\hat{\theta}_n(u) < \theta_0(u) - a_n^{-1}(b(u))^{-1}(\tau + a_n Z_n(\theta_0(u), u)) \text{ for some } u \in T\} \\ &\subset \{Z_n(\theta_0(u) - a_n^{-1}(b(u))^{-1}(\tau + a_n Z_n(\theta_0(u), u)), u) > 0 \text{ for some } u \in T\}. \end{aligned} \tag{B.3}$$

By conditions *(ii)* and *(iii)*

$$\begin{aligned} &\sup_{u \in T} |a_n(Z_n(\theta_0(u) - a_n^{-1}(b(u))^{-1} \\ &\times (\tau + a_n Z_n(\theta_0(u), u)), u) - Z_n(\theta_0(u), u)) \\ &+ (\tau + a_n Z_n(\theta_0(u), u))| \xrightarrow{\Pr} 0. \end{aligned}$$

Thus,

$$\sup_{u \in T} |a_n Z_n(\theta_0(u) - a_n^{-1}(b(u))^{-1}(\tau + a_n Z_n(\theta_0(u), u))) + \tau| \xrightarrow{\Pr} 0.$$

This and (B.3) imply (B.1). (B.2) follows similarly. □

It is easy to see that $\hat{\theta}_n^*(u) \le \hat{\theta}_n^{**}(u)$. $\hat{\theta}_n^*(u)$ and $\hat{\theta}_n^{**}(u)$ estimate the smallest and biggest solutions to $Z_n(\theta, u) = 0$.

Proof of Theorem 3.5. We apply Theorem B.1 with $Z_n(\theta, u) = H_n(\theta, u)$, $b(u) = H'(\theta_0(u), u)$ and $a_n = n^{1/2}$. Condition *(i)* in Theorem B.1 is assumed. Condition *(iii)* implies that for each $0 < M < \infty$,

$$\lim_{n \to \infty} \sup_{\|u\|=1} \sup_{|\tau| \le M} |n^{1/2} E[H_n(\theta_0 + n^{-1/2}\tau, u) - H_n(\theta_0, u)] - \tau H'(\theta_0(u), u)| = 0.$$

Hence, to check condition *(ii)* in Theorem B.1, we need to prove that for each $0 < M < \infty$,

$$\begin{aligned} &\sup_{\|u\|=1} \sup_{|\tau| \le M} |n^{1/2}(H_n(\theta_0 + n^{-1/2}\tau, u) - H_n(\theta_0, u) \\ &-H(\theta_0(u) + n^{-1/2}\tau, u) + H(\theta_0(u), u))| \xrightarrow{\Pr} 0. \end{aligned} \tag{B.4}$$

We claim that by the CLT for U–processes indexed by a VC class of functions (Theorem 4.9 in Arcones and Giné, 1993)

$$\{n^{1/2}(H_n(\theta,u)-H(\theta,u)):\|u\|=1,|\theta-\theta_0(u)|\le\delta\} \tag{B.5}$$

converges weakly. We claim that $\{\psi(k^{-1}u'(x_1+\cdots+x_k)-\theta):\theta\in\mathbb{R},\|u\|=1\}$ is VC subgraph class of functions in the sense of Dudley (1999, p. 159). Observe that the subgraph of $\psi(k^{-1}u'(x_1+\cdots+x_k)-\theta)$ is

$$\begin{aligned}&\{(x_1,\ldots,x_k,t):0\le t,\psi^{-1}(t)\le k^{-1}u'(x_1+\cdots+x_k)-\theta\}\\ \cup&\{(x_1,\ldots,x_k,t):0\ge t,\psi^{-1}(t)\ge k^{-1}u'(x_1+\cdots+x_k)-\theta\}\end{aligned}$$

Since the class of sets where a finite dimensional set of functions is nonnegative is a VC class of sets (Theorem 4.2.1 in Dudley, 1999), the class of functions $\{\psi(k^{-1}u'(x_1+\cdots+x_k)-\theta):\theta\in\mathbb{R},\|u\|=1\}$ is VC subgraph class of functions. Condition *(v)* implies that the subclass $\{\psi(k^{-1}u'(x_1+\cdots+x_k)-\theta):|\theta-\theta_0(u)|\le\delta,\|u\|=1\}$ has an envelope with finite second moment. So, the stochastic process in (B.5) converges weakly. So, condition *(iv)* implies (B.4). Condition *(iii)* in Theorem B.1 follows from the weak convergence of the process in (B.5). □

When $\psi(x)=I(x\le 0)-2^{-1}$, Theorem 3.5 gives the following:

Corollary B.1. *For $\|u\|=1$, let $F_u^*(t)=\Pr\{k^{-1}u'(X_1+\cdots+X_k)\le t\}$. Suppose that:*

(i) *For each $\|u\|=1$, there exists $m(u)$ such that $F_u^*(m(u))=2^{-1}$.*

(ii) *For each $\|u\|=1$, $F_u^{*\prime}(m(u))$ exists.*

(iii) $0<\inf_{\|u\|=1}F_u^{*\prime}(m(u))<\infty$.

(iv) $\lim_{h\to0}\sup_{\|u\|=1}|h^{-1}(F_u^*(m(u)+h)-F_u^*(m(u)))-F_u^{*\prime}(m(u))|=0$.

(v) $\lim_{h\to0}\sup_{\|u\|=1}|F_u^*(m(u)+h)-F_u^*(m(u))|=0$.

Then,

$$\begin{aligned}&\sup_{\|u\|=1}|n^{1/2}(\hat m_n(u)-m(u))+(F_u^{*\prime}(m(u)))^{-1}n^{1/2}\frac{(n-k)!}{n!}\\ &\times\sum_{(i_1,\ldots,i_k)\in I_k^n}(I(k^{-1}u'(X_{i_1}+\cdots+X_{i_k})\le\theta_0(m))-2^{-1})|\stackrel{\Pr}{\to}0,\end{aligned}$$

where $\hat m_n(u)=\mathrm{med}\{k^{-1}u'(X_{i_1}+\cdots+X_{i_k}):1\le i_1<\cdots<i_k\le n\}$.

The estimator in the previous theorem is the median if $k = 1$ and the Hodges–Lehmann estimator if $k = 2$. The asymptotics of the projection median was considered by Cui and Tian (1994). Their conditions are slightly stronger than above and the conclusions are also stronger.

The next theorem considers a general class of scale parameters generalizing the median of the absolute deviations.

Theorem B.2. *Under the notation and assumptions in Corollary B.1, suppose that:*

(i) *For each* $\|u\| = 1$, *there exists* $\mathrm{MAD}(u)$ *such that*

$$F_u^*(m(u) + \mathrm{MAD}(u)) - F_u^*(m(u) - \mathrm{MAD}(u)) = 2^{-1}.$$

(ii) *For each* $\|u\| = 1$, $F_u^{*\prime}(m(u) + \mathrm{MAD}(u))$ *and* $F_u^{*\prime}(m(u) - \mathrm{MAD}(u))$ *exist.*

(iii) $0 < \inf_{\|u\|=1} (F_u^{*\prime}(m(u) + \mathrm{MAD}(u))) + F_u^{*\prime}(m(u) - \mathrm{MAD}(u)))$.

(iv) $\lim_{h\to 0} \sup_{\|u\|=1} |h^{-1}(F_u(m(u) + \mathrm{MAD}(u) + h) - F_u(m(u) + \mathrm{MAD}(u))) - F_u'(m(u) + \mathrm{MAD}(u))| = 0$

and

$\lim_{h\to 0} \sup_{\|u\|=1} |h^{-1}(F_u(m(u) - \mathrm{MAD}(u) + h) - F_u(m(u) - \mathrm{MAD}(u))) - F_u'(m(u) - \mathrm{MAD}(u))| = 0.$

(v)

$$\lim_{h\to 0} \sup_{\|u\|=1} |F_u^*(m(u) + \mathrm{MAD}(u) + h) - F_u^*(m(u) + \mathrm{MAD}(u))| = 0$$

and

$$\lim_{h\to 0} \sup_{\|u\|=1} |F_u^*(m(u) - \mathrm{MAD}(u) + h) - F_u^*(m(u) - \mathrm{MAD}(u))| = 0.$$

Then,

$$\sup_{\|u\|=1} |n^{1/2}(\widehat{\mathrm{MAD}}_n(u) - \mathrm{MAD}(u)) + (b(u))^{-1} n^{1/2} \frac{(n-k)!}{n!}$$
$$\times \sum_{(i_1,\dots,i_k)\in I_k^n} (I(|k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) - m(u)| \le \mathrm{MAD}(u)) - 2^{-1})| \stackrel{\Pr}{\to} 0,$$

where $b(u) = F_u^{*\prime}(m(u) + \mathrm{MAD}(u)) + F_u^{*\prime}(m(u) - \mathrm{MAD}(u))$,

$$\widehat{\mathrm{MAD}}_n^{**} = \inf\{t \in \mathbb{R} : \frac{(n-k)!}{n!} \sum_{(i_1,\dots,i_k)\in I_k^n} I(|k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) - \hat{m}_n(u)| \le t) > 1/2\},$$
$$\widehat{\mathrm{MAD}}_n^{*} = \sup\{t \in \mathbb{R} : \frac{(n-k)!}{n!} \sum_{(i_1,\dots,i_k)\in I_k^n} I(|k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) - \hat{m}_n(u)| \le t) < 1/2\},$$

and $\widehat{\mathrm{MAD}}_n = 2^{-1}(\widehat{\mathrm{MAD}}_n^{*} + \widehat{\mathrm{MAD}}_n^{**})$.

Proof. Let

$$G_n(t,v,u) = \frac{(n-k)!}{n!} \sum_{(i_1,\dots,i_k)\in I_k^n} I(|k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) - v| \le t).$$

We apply Theorem B.1 with $Z_n(\theta,u) = G_n(\theta, m_n(u), u)$ and $a_n = n^{1/2}$. By the arguments in the Theorem 3.5, we have that

$$\{n^{1/2}(G_n(t,v,u) - G(t,v,u)) : t \ge 0, r \in \mathbb{R}, \|u\| = 1\}$$

converges weakly, where $G(t,v,u) = \Pr\{|k^{-1}u'(X_1 + \cdots + X_k) - v| \le t\}$. By condition *(v)*, for each $0 < M < \infty$, we have that

$$\sup_{\|u\|=1} \sup_{|\tau|\le M} n^{1/2}|G_n(\mathrm{MAD}(u) + n^{-1/2}\tau, m_n(u), u) - G(\mathrm{MAD}(u) + n^{-1/2}\tau, m_n(u), u) - G_n(\mathrm{MAD}(u), m(u), u) + G(\mathrm{MAD}(u), m(u), u)| \overset{\Pr}{\to} 0.$$

So, it suffices to prove that

$$\sup_{\|u\|=1} \sup_{|\tau|\le M} |n^{1/2}(G(\mathrm{MAD}(u) + n^{-1/2}\tau, m(u), u) - G(\mathrm{MAD}(u), m(u), u)) - \tau b(u)| \to 0.$$

This follows from *(iv)*. □

When $k = 1$, the estimator in the previous theorem is the median of the absolute deviations. For $k \ge 2$, medians in the previous are understood in the sense of Hodges–Lehmann.

The next theorem considers a general class of scale parameters generalizing the interquantile range. Given $0 < p < 1$, we consider the p–th quantile range $QR = F^{-1}(1 - 2^{-1}p) - F^{-1}(2^{-1}p)$.

Theorem B.3. *Let $0 < p < 1$. Suppose that:*

(i) For each $\|u\| = 1$, there exists $q_1(u)$ and $q_2(u)$ such that $F_u(q_1(u)) = 2^{-1}p$ and $F_u(q_2(u)) = 1 - 2^{-1}p$.

(ii) For each $\|u\| = 1$ and each $i = 1, 2$, $F_u^{*\prime}(q_i(u))$ exists and $0 < \inf_{\|u\|=1} F_u^{*\prime}(q_i(u))$.

(iii) For each $i = 1, 2$,

$$\lim_{h\to 0} \sup_{\|u\|=1} |h^{-1}(F_u^*(q_i(u) + h) - F_u^*(q_i(u))) - F_u^{*\prime}(q_i(u))| = 0.$$

(iv) For each $i = 1, 2$,

$$\lim_{h\to 0} \sup_{\|u\|=1} |F_u^*(q_i(u) + h) - F_u^*(q_i(u))| = 0.$$

Then,

$$\begin{aligned} &\sup_{\|u\|=1} |n^{1/2}(\widehat{\mathrm{QR}}_n(u) - \mathrm{QR}(u)) \\ &+ n^{1/2}(F_u^{*\prime}(q_2(u)))^{-1}(F_{n,u}^*(q_2(u)) - F_u^*(q_2(u))) \\ &- n^{1/2}(F_u^{*\prime}(q_1(u)))^{-1}(F_{n,u}^*(q_1(u)) - F_u^*(q_1(u)))| \overset{\Pr}{\to} 0, \end{aligned}$$

where $\mathrm{QR}(u) = q_2(u) - q_1(u)$,

$$F_{n,u}^*(t) = \frac{(n-k)!}{n!} \sum_{(i_1,\ldots,i_k)\in I_k^n} I(k^{-1}u'(X_{i_1} + \cdots + X_{i_k}) \le t)$$

$\hat{q}_{n,2}^{**} = \inf\{t \in \mathbb{R} : F_{n,u}^*(t) > 1 - (p/2)\}$, $\hat{q}_{n,2}^{*} = \sup\{t \in \mathbb{R} : F_{n,u}^*(t) < 1 - (p/2)\}$, $\hat{q}_{n,2} = 2^{-1}(\hat{q}_2^{**} + \hat{q}_2^{*})$, $\hat{q}_{n,1}^{**} = \inf\{t \in \mathbb{R} : F_{n,u}^*(t) > p/2\}$, $\hat{q}_{n,1}^{*} = \sup\{t \in \mathbb{R} : F_{n,u}^*(t) < p/2\}$, $\hat{q}_{n,1} = 2^{-1}(\hat{q}_{n,1}^{**} + \hat{q}_{n,1}^{*})$ *and* $\widehat{\mathrm{QR}}_n(u) = \hat{q}_{n,2}(u) - \hat{q}_{n,1}(u)$.

When $p = 1/2$ and $k = 1$, the estimator QR in the previous theorem is the interquantile range. The previous theorem follows similarly to Theorems 3.5 and B.2.

References

Arcones, M. A. (1996). Some remarks on the uniform weak convergence of stochastic processes. *Statistics & Probability Letters*, 28:41–49.

Arcones, M. A., Chen, Z., and Giné, E. (1994). Estimators related to U–processes with applications to multivariate medians: asymptotic normality. *The Annals of Statistics*, 22:1460–1477.

Arcones, M. A. and Giné, E. (1993). Limit theorems for U–processes. *The Annals of Probability*, 21:1494–1542.

Cui, H. and Tian, Y. (1994). Estimation of the projection absolute median deviation and its application. *Journal of Systems Science and Mathematical Sciences*, 14:63–72. In Chinese.

Donoho, D. L. (1982). *Breakdown properties of multivariate location estimators*. PhD Thesis, Department of Statistics, Harvard University. Qualifying Paper.

Donoho, D. L. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, 20:1803–1827.

Dudley, R. M. (1999). *Uniform Central Limit Theorems*. Cambridge University Press, Cambridge.

Liu, R. Y. (1990). On a notion of data depth based on random simplices. *The Annals of Statistics*, 18:405–414.

Liu, R. Y., Parelius, J. M., and Singh, K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference (with discussion). *The Annals of Statistics*, 27:783–858.

Maronna, R. A. and Yohai, V. J. (1995). The behavior of the Stahel–Donoho robust multivariate estimator. *Journal of the American Statistical Association*, 90:330–341.

Massé, J.-C. (2002). Asymptotics for the Tukey median. *Journal of Multivariate Analysis*, 81:286–300.

Massé, J.-C. (2004). Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean. *Bernoulli*, 10(3):397–419.

ROUSSEEUW, P. J. and RUTS, I. (1999). The depth function of a population distribution. *Metrika*, 49:213–244.

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

SMALL, C. G. (1990). A survey of multidimensional medians. *International Statistical Review*, 58:263–277.

STAHEL, W. A. (1981). *Robust Estimation: Infinitesimal Optimality and Covariance Matrix Estimators*. PhD Thesis, ETH, Zurich. In German.

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

TYLER, D. E. (1994). Finite sample breakdown points of projection based multivariate location and scatter statistics. *The Annals of Statistics*, 22:1024–1044.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer–Verlag, New York.

ZHANG, J. (2002). Some extensions of Tukey's depth function. *Journal of Multivariate Analysis*, 82:134–165.

ZUO, Y. (2003). Projection based depth functions and associated medians. *The Annals of Statistics*, 31:1460–1490.

ZUO, Y. and CUI, H. (2005). Depth weighted scatter estimators. *The Annals of Statistics*, 33(1):381–413.

ZUO, Y., CUI, H., and HE, X. (2004). On the Stahel–Donoho estimator and depth-weighted means of multivariate data. *The Annals of Statistics*, 32:167–188.

ZUO, Y. and SERFLING, R. (2000). General notions of statistical depth function. *The Annals of Statistics*, 28:461–482.