The Annals of Statistics
2004, Vol. 32, No. 1, 189–218

INFLUENCE FUNCTION AND MAXIMUM BIAS OF PROJECTION DEPTH BASED ESTIMATORS

BY YIJUN ZUO,[1] HENGJIAN CUI[2] AND DENNIS YOUNG

Michigan State University, Beijing Normal University and Arizona State University

Location estimators induced from depth functions increasingly have been pursued and studied in the literature. Among them are those induced from projection depth functions. These projection depth based estimators have favorable properties among their competitors. In particular, they possess the best possible finite sample breakdown point robustness. However, robustness of estimators cannot be revealed by the finite sample breakdown point alone. The influence function, gross error sensitivity, maximum bias and contamination sensitivity are also important aspects of robustness. In this article, we study these other robustness aspects of two types of projection depth based estimators: projection medians and projection depth weighted means. The latter includes the Stahel–Donoho estimator as a special case. Exact maximum bias, the influence function, and contamination and gross error sensitivity are derived and studied for both types of estimators. Sharp upper bounds for the maximum bias and the influence functions are established. Comparisons based on these robustness criteria reveal that the projection depth based estimators enjoy desirable local as well as global robustness and are very competitive among their competitors.

1. Introduction. Depth induced location estimators, especially depth medians, increasingly have been pursued and studied in the literature. Among them are the half-space median [Tukey (1975)], the simplicial median [Liu (1990) and Liu, Parelius and Singh (1999)], the projection median [Tyler (1994), Zuo and Serfling (2000a) and Zuo (2003)] and the projection depth weighted mean, which includes the Stahel–Donoho estimator as a special case [Zuo, Cui and He (2004)]. The half-space and projection medians and the projection depth weighted mean are implementations of the projection pursuit methodology. They have been shown to possess high finite sample breakdown points [Donoho and Gasko (1992), Tyler (1994), Maronna and Yohai (1995), Zuo (2003) and Zuo, Cui and He (2004)] and are favorable as robust location estimators among their competitors. The finite

---

Received October 2001; revised December 2002.

[1]Supported in part by NSF Grants DMS-00-71976 and DMS-01-34628 and RIA at Arizona State University.

[2]Supported in part by NSFC 10231030 and the EYTP of China.

AMS 2000 subject classifications. Primary 62E20; secondary 62G35, 62G20.

Key words and phrases. Maximum bias, influence function, contamination sensitivity, gross error sensitivity, projection depth, projection median, projection depth weighted mean, robustness, breakdown point.

sample breakdown point alone, however, does not depict the entire robustness picture of an estimator. The influence function, gross error sensitivity, maximum bias and contamination sensitivity, as local and global measures of estimator robustness, also play indispensable roles in assessing robustness of estimators. Deriving these robustness measures (especially maximum bias and contamination sensitivity) for multivariate location estimators, however, is usually very challenging and difficult. This is true even at spherically symmetric distributions. The robustness picture of the half-space median for very general distributions was completed by Chen and Tyler (2002). Adrover and Yohai (2002) obtained important results on the maximum bias and gross error sensitivity of the projection median for spherically symmetric distributions. The corresponding results under weaker assumptions of symmetry, however, are yet to be explored. Further, the influence function of the projection median is unknown. As for the projection depth weighted mean, its robustness features beyond the finite sample breakdown point have not yet been characterized.

In this article, we focus on the influence function, gross error sensitivity, maximum bias and contamination sensitivity of the projection median and projection depth weighted mean. Sharp upper bounds for the maximum bias and influence function are established, and exact influence functions, gross error sensitivities, maximum bias functions and contamination sensitivities are also derived for nonelliptically as well as elliptically symmetric distributions. It turns out that the influence functions of these estimators are bounded and so are their maximum bias functions for contaminations under 50%. Comparisons with the half-space median indicate that these estimators share desirable local and global robustness properties and are very favorable overall. Our results here fill the gap in the robustness study of the projection depth weighted mean and complement the work of Adrover and Yohai (2002) (AY02) on the projection median. Specifically, (a) we study the robustness of both the projection depth weighted mean and the projection median, while AY02 studied only the robustness of the latter, and (b) we drop, in the robustness study of the projection median, the spherical symmetry and other assumptions required in AY02 (consequently our results are more general, but proofs become truly multivariate and technically more demanding): we provide a sharp upper bound of gross error sensitivity and an influence function that are not given in AY02, establish a maximum bias upper bound that is sharper than that in AY02 and prove a conjecture in AY02.

The rest of this article is organized as follows. Formal definitions of projection depth, maximum bias, contamination sensitivity, the influence function, and gross error sensitivity and related notation are presented in Section 2. The main results of the paper are given in Sections 3 and 4. In Section 3, for the projection depth weighted mean we (a) establish upper bounds for its maximum bias and influence function, (b) derive its exact influence function and gross error sensitivity for very general as well as elliptical distributions and (c) obtain the exact maximum bias and contamination sensitivity (under point-mass contamination) for elliptical distributions. Section 4 is devoted to the projection median, where (a) sharp upper bounds for the maximum bias and influence function are established, (b) exact maximum bias and contamination sensitivity are derived for "d-version symmetric" distributions (including as special cases elliptical distributions and distributions generated from i.i.d. Cauchy marginals), and are shown to be dimension-dependent for nonspherical distributions (proving a conjecture in AY02) and (c) exact the influence function and gross error sensitivity are obtained for elliptical distributions. Simulation studies and comparisons of projection depth estimators with Tukey's half-space median are undertaken in Section 5. A special weight function is used for the weighted mean. The asymptotic breakdown point for both projection estimators is $1/2$, whereas it is only $1/3$ for the half-space median. The maximum bias of the projection weighted mean is the smallest when compared with those of the two medians at standard normal models in $\mathbb{R}^d$ for $d \le 5$. The maximum bias of the half-space median is slightly smaller than that of the projection median for contamination close to or less than 0.3, but jumps to infinity as the contamination approaches $1/3$. The contamination sensitivities of the two medians are the same at elliptical models, but larger than that of the projection weighted mean for the weight function we choose. The influence function and gross error sensitivity of the projection median are twice those of the half-space median at elliptical models and the influence function norm of the projection weighted mean is smaller than those of the two medians for most points at standard normal models. The dimension-free property of the maximum bias, contamination and gross error sensitivity of the two medians at spherical models disappears at other models (in fact, they are of order $\sqrt{d}$ at some other models for all three estimators). Selected (sketches of) proofs and auxiliary lemmas are presented in Appendix A.

2. Definitions and notation. For a given distribution F in $\mathbb{R}^d$ and an $\varepsilon > 0$, the version of F contaminated by an ε amount of an arbitrary distribution G is denoted by $F(\varepsilon, G) = (1-\varepsilon)F + \varepsilon G$. The *maximum bias* of a given location functional T under an ε amount of contamination at F is defined as [Hampel, Ronchetti, Rousseeuw and Stahel (1986)]

$$B(\varepsilon; T, F) = \sup_G \|T(F(\varepsilon, G)) - T(F)\|,$$

where (and hereafter) $\|\cdot\|$ stands for Euclidean norm. The *contamination sensitivity* of T at F [He and Simpson (1993)] is defined as

$$\gamma(T, F) = \lim_{\varepsilon \to 0+} B(\varepsilon; T, F)/\varepsilon,$$

where $B(\varepsilon; T, F)$ is the maximum deviation (bias) of T under an ε amount of contamination at F and it mainly measures the global robustness of T. The notation $\gamma(T, F)$ indicates the maximum relative effect on T of an infinitesimal contamination at F and measures the local as well as global robustness of T.

The minimum amount ε^* of contamination at F which leads to an unbounded $B(\varepsilon; T, F)$ is called the (asymptotic) *breakdown point* (BP) of T at F, that is, $\varepsilon^* = \min\{\varepsilon : B(\varepsilon; T, F) = \infty\}$.

The *influence function* (IF) of T at a given point $x \in \mathbb{R}^d$ for a given F is defined as

$$\mathrm{IF}(x; T, F) = \lim_{\varepsilon \to 0+} \big(T(F(\varepsilon, \delta_x)) - T(F)\big)/\varepsilon,$$

where δ_x is the point-mass probability measure at $x \in \mathbb{R}^d$, and the *gross error sensitivity* of T at F is then defined as [Hampel, Ronchetti, Rousseeuw and Stahel (1986)]

$$\gamma^*(T, F) = \sup_{x \in \mathbb{R}^d} \|\mathrm{IF}(x; T, F)\|.$$

The function $\mathrm{IF}(x; T, F)$ describes the relative effect (influence) on T of an infinitesimal point-mass contamination at x and measures the local robustness of T. The function $\gamma^*(T, F)$ is the maximum relative effect on T of an infinitesimal point-mass contamination and measures the global as well as local robustness of T.

Let $\mu(\cdot)$ and $\sigma(\cdot)$ be univariate location and scale functionals, respectively. The projection depth (PD) of a point $x \in \mathbb{R}^d$ for a given distribution F of the random vector $X \in \mathbb{R}^d$ is defined as

$$\mathrm{PD}(x, F) = 1/\big(1 + O(x, F)\big),$$

where the outlyingness $O(x, F) = \sup_{\|u\|=1} g(x, u, F)$ and $g(x, u, F) = (u'x - \mu(F_u))/\sigma(F_u)$, and F_u is the distribution of $u'X$ [Liu (1992) and Zuo and Serfling (2000b)]. Throughout our discussions, μ and σ are assumed to exist for the univariate distributions considered. We also 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 $Y \in \mathbb{R}$. Further, we assume that $\sup_{\|u\|=1} \mu(F_u) < \infty$ and $0 < \inf_{\|u\|=1} \sigma(F_u) \le \sup_{\|u\|=1} \sigma(F_u) < \infty$. Denote the projected distribution of $F(\varepsilon, G)$ to a unit vector u by $F_u(\varepsilon, G)$. Then $F_u(\varepsilon, G) = (1 - \varepsilon)F_u + \varepsilon G_u$. Define for any $\varepsilon > 0$, any unit vector u and any distribution G,

$$B(\varepsilon; \mu, F) = \sup_{u, G} \big|\mu\big(F_u(\varepsilon, G)\big) - \mu(F_u)\big|,$$

$$B(\varepsilon; \sigma, F) = \sup_{u, G} \big|\sigma\big(F_u(\varepsilon, G)\big) - \sigma(F_u)\big|,$$

$$B^*(\varepsilon; \mu, F) = \sup_{u, G} \mu\big(F_u(\varepsilon, G)\big),$$

$$B^*(\varepsilon; \sigma, F) = \sup_{u, G} \sigma\big(F_u(\varepsilon, G)\big),$$

$$B_*(\varepsilon; \sigma, F) = \inf_{u, G} \sigma\big(F_u(\varepsilon, G)\big),$$

and

$$\varepsilon^*(\mu, F) = \min\{\varepsilon : B(\varepsilon; \mu, F) = \infty\},$$
$$\varepsilon^*(\sigma, F) = \min\bigl\{\varepsilon : B(\varepsilon; \sigma, F) + B_*(\varepsilon; \sigma, F)^{-1} = \infty\bigr\}.$$

3. Projection depth weighted means and Stahel–Donoho estimator. Depth induced multivariate estimators, called *DL* statistics, were first considered in Liu (1990) and then elaborated further in Liu, Parelius and Singh (1999). The *DL* statistics are weighted means with weights that typically depend on the depth-induced rank (not depth) of points. This section studies the robustness of projection *depth* weighted means with weights that depend directly on the depth of points. Specifically, a projection depth weighted mean (PWM) at distribution F is defined as [Zuo, Cui and He (2004)]

$$\mathrm{PWM}(F) = \int x w\bigl(\mathrm{PD}(x, F)\bigr)\, dF(x) \Big/ \int w\bigl(\mathrm{PD}(x, F)\bigr)\, dF(x), \tag{3.1}$$

where $w(\cdot)$ is a weight function. If PD is replaced with a general depth function [see Liu, Parelius and Singh (1999) and Zuo and Serfling (2000b)], a general depth weighted mean is obtained. In the following discussion we confine our attention to PD and the weight function $w(\cdot)$ on [0, 1] that satisfies

$$\begin{aligned} & w(0) = 0, \\ & w_*(s) = \inf_{t \ge s} w(t) > 0 \qquad \forall\, 0 < s \le 1, \\ & w^0 = \sup_{s,\, t \in [0,1],\, s \ne t} |w(s) - w(t)|/|s - t| < \infty. \end{aligned} \tag{3.2}$$

That is, w is positive (except with value 0 at 0) and Lipschitz continuous (with the smallest Lipschitz constant w^0). One such w, which was utilized in Zuo, Cui and He (2004) and is used herein in Section 5 (with parameters $k > 0$ and $0 < c \le 1$), is

$$\begin{aligned} w_1(r) = {} & \bigl(\exp\bigl(-k(1 - r/c)^2\bigr) - \exp(-k)\bigr)\big/\bigl(1 - \exp(-k)\bigr) I(r < c) \\ & + I(r \ge c). \end{aligned} \tag{3.3}$$

Taking $w(\cdot) = W_{\mathrm{SD}}(1/(1 + \cdot))$, where $W_{\mathrm{SD}}(\cdot)$ is the weight function in the Stahel–Donoho (SD) estimator [Donoho and Gasko (1992)], we then have the SD estimator, a special case of PWM. Now let

$$\mu_{\sup}(\varepsilon, F) = \sup_{\|u\|=1,\, G} \bigl|\mu\bigl(F_u(\varepsilon, G)\bigr) - u'\,\mathrm{PWM}(F)\bigr|$$

and

$$A(\varepsilon, F) = \sup_{a > 0} P\{\|X\| \le a\} w_*\bigl(B_*(\varepsilon; \sigma, F)\big/\bigl(B_*(\varepsilon; \sigma, F) + B^*(\varepsilon; \mu, F) + a\bigr)\bigr).$$

We have the following upper bound for the maximum bias of PWM.

THEOREM 3.1. *For w in* (3.2) *and any given distribution F and $\varepsilon > 0$*:

$$\text{(i)}\quad B\big(\varepsilon; \text{PWM}, F\big) \le \frac{w^0 \max\{\mu_{\sup}(\varepsilon, F), B^*(\varepsilon; \sigma, F)\}}{A(\varepsilon, F)} \times \left((1-\varepsilon)\frac{\max\{B(\varepsilon; \mu, F), B(\varepsilon; \sigma, F)\}}{B_*(\varepsilon; \sigma, F)} + \varepsilon\right);$$

$$\text{(ii)}\quad \varepsilon^*(\text{PWM}, F) \ge \min\{\varepsilon^*(\mu, F), \varepsilon^*(\sigma, F)\}.$$

The upper bound of the maximum bias and the lower bound of the breakdown point of PWM are determined by those of μ and σ, respectively. When μ and σ are the median (Med) and the median absolute deviation (MAD), respectively, it is seen that the maximum bias of PWM is bounded for contamination less than 50%, and the breakdown point of PWM is $1/2$ in this case.

Define the gross error sensitivity of μ and σ for a given F in $\mathbb{R}^d$ as

$$\gamma^*(\mu, F) = \sup_{\|u\|=1,\, x\in\mathbb{R}^d} \text{IF}(u'x; \mu, F_u) \quad \text{and}$$

$$\gamma^*(\sigma, F) = \sup_{\|u\|=1,\, x\in\mathbb{R}^d} \text{IF}(u'x; \sigma, F_u).$$

Taking $G = \delta_x$ and following the proof of Theorem 3.1, we obtain an upper bound for $\gamma^*(\text{PWM}, F)$.

THEOREM 3.2. *Assume that w is given by* (3.2) *and*:

(i) $\mu(F_u(\varepsilon, \delta_x)) \to \mu(F_u)$ *uniformly in u as $\varepsilon \to 0$ for any $x \in \mathbb{R}^d$*;
(ii) $\sigma(F_u(\varepsilon, \delta_x)) \to \sigma(F_u)$ *uniformly in u as $\varepsilon \to 0$ for any $x \in \mathbb{R}^d$*.

Then

$$\gamma^*(\text{PWM}, F) \le w^0 \max\{\mu_{\sup}(0, F), B^*(0; \sigma, F)\}/A(0, F) \times \big(\max\{\gamma^*(\mu, F), \gamma^*(\sigma, F)\}/B_*(0; \sigma, F) + 1\big).$$

The upper bound of the gross error sensitivity of PWM is determined by those of μ and σ, and is finite provided that $\max\{\gamma^*(\mu, F), \gamma^*(\sigma, F)\} < \infty$.

Deriving the exact expressions of the maximum bias, contamination and gross error sensitivity, and the influence function of PWM turns out to be very challenging and difficult. First we establish the influence function of PWM. Let $u(x) = \{u : \|u\| = 1, g(u, x, F) = O(x, F)\}$. For a given $y \in \mathbb{R}^d$, define $S(y)$ to be the set of x such that $\text{IF}(u'y; \sigma, F_u)$ and $\text{IF}(u'y; \mu, F_u)$ are discontinuous at $u(x)$. We have the following theorem.

THEOREM 3.3. *Assume that the conditions of Theorem* 3.2 *hold. Additionally assume*:

(i) $\mu(F_u)$ *and* $\sigma(F_u)$ *are continuous in u*;

(ii) *the derivative* $w^{(1)}$ *of* w *is continuous*;
(iii) $u(x)$ *is a singleton a.s.*;
(iv) $\mathrm{IF}(u'y; \sigma, F_u)$ *and* $\mathrm{IF}(u'y; \mu, F_u)$ *are bounded* (*in* u) *for a given* $y \in \mathbb{R}^d$.

Then for the given y *with* $P(S(y)) = 0$,

$$\begin{aligned}
&\mathrm{IF}(y; \mathrm{PWM}, F)\\
&\quad = \left(\int h(x, y)\, dF(x) + \big(y - \mathrm{PWM}(F)\big) w\big(\mathrm{PD}(y, F)\big) \right)\\
&\qquad \times \left[\int w\big(\mathrm{PD}(x, F)\big)\, dF(x) \right]^{-1},
\end{aligned}$$

where

$$\begin{aligned}
h(x, y) = {}& \big(x - \mathrm{PWM}(F)\big) w^{(1)}\big(\mathrm{PD}(x, F)\big)\\
& \times \frac{(O(x, F)\mathrm{IF}(u(x)'y; \sigma, F_{u(x)}) + \mathrm{IF}(u(x)'y; \mu, F_{u(x)}))}{\sigma(F_{u(x)})(1 + O(x, F))^2}.
\end{aligned}$$

The above result can be extended to a very general setting. For a general depth weighted mean L [replacing PD in (3.1) with a general depth D] and general F, under mild conditions on w, $\mathrm{IF}(x; L, F) = (\int (y - L) w^{(1)}(D(y, F))\mathrm{IF}(y; D(x, F), F) F(dy) + (x - L) w(D(x, F)))/\int w(D(x, F)) F(dx)$, provided that $\mathrm{IF}(x; D, F)$ exists. The latter is true for depth functions such as the half-space depth function [Romanazzi (2001)] and the projection depth function.

The influence function in Theorem 3.3 takes a much simpler form for special distributions like elliptically symmetric ones and for $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD})$. Let $X \sim F$ in $\mathbb{R}^d$ such that $u'X \stackrel{d}{=} a(u)Z$ with $a(u) = a(-u) > 0$ for $\|u\| = 1$, $Z \stackrel{d}{=} -Z$, $Z \in \mathbb{R}^1$, where $\stackrel{d}{=}$ stands for equal in distribution. The distribution F (or X) is called *elliptically symmetric* about the origin if $a(u) = \sqrt{u'\Sigma u}$, where Σ is a positive definite matrix. Elliptically symmetric about an arbitrary $\theta \in \mathbb{R}^d$ can be defined similarly. In the following discussion, we assume without loss of generality (w.l.o.g.) that $\theta = 0$ and $\mathrm{MAD}(Z) = 1$. Let λ_1 be the largest eigenvalue of Σ and let $\Sigma^{-1/2}X \sim F_0$ (note that F_0 is spherically symmetric). For a positive definite matrix A, define $x/\|Ax\| = 0$ when $x = 0$.

THEOREM 3.4. *For* $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD})$ *and elliptically symmetric* F *of* $X \in \mathbb{R}^d$, *assume that the density* $h(z)$ *of* Z *is continuous*, $h(0)h(1) > 0$, $w^{(1)}(\cdot)$ *is continuous*. *Then*:

(i) $\mathrm{IF}(y; \mathrm{PWM}, F)$

$$= \frac{(k_0 + \|\Sigma^{-1/2}y\| w(1/(1 + \|\Sigma^{-1/2}y\|)))y/\|\Sigma^{-1/2}y\|}{\int w((1 + \|x\|)^{-1})\, dF_0(x)} \qquad \textit{for any } y \in \mathbb{R}^d,$$

(ii) $\gamma^*(\mathrm{PWM}, F)$

$$= \sqrt{\lambda_1} \sup_{r \ge 0} \big|k_0 + r w\big((1+r)^{-1}\big)\big| \Big/ \int w\big((1+\|x\|)^{-1}\big)\, dF_0(x),$$

where $k_0 = 1/(2h(0)) \int |x_1| w^{(1)}((1+\|x\|)^{-1})/(1+\|x\|)^2\, dF_0(x)$ *and* $x' = (x_1, \dots, x_d)$.

The theorem implies that for the weight function w such that $rw(1/(1+r))$ is bounded for $r \ge 0$, the influence function of PWM is bounded.

Now we derive the maximum bias of the projection depth weighted mean for elliptically symmetric F and $(\mu, \sigma) = (\text{Med, MAD})$. In many cases, the maximum bias is attained by a point-mass distribution, that is, $B(\varepsilon; T, F) = \sup_{x \in \mathbb{R}^d} \|T(F(\varepsilon, \delta_x)) - T(F)\|$ [see Huber (1964), Martin, Yohai and Zamar (1989), Chen and Tyler (2002) and Adrover and Yohai (2002)]. In the following discussion, we derive the maximum bias and contamination sensitivity of PWM under point-mass contamination. We conjecture that our results hold for general contamination. Note that under point-mass contamination the only difference between the contamination sensitivity $\gamma(T, F)$ and the gross error sensitivity $\gamma^*(T, F)$ is the order in which the suprema and the limits are taken in their respective definitions. This might tempt one to believe that these two sensitivities are the same if it is taken for granted that the order in which the supremum and the limit are taken is interchangeable. Unfortunately, this is not always the case [see, e.g., Chen and Tyler (2002), where $\gamma(T, F) = 2\gamma^*(T, F)$]. In the following text, we prove that for PWM, the order *is* interchangeable and the contamination sensitivity *is* the same as the gross error sensitivity. The proof and the derivation of the following result, given in Appendix A, is rather technically demanding.

For the random variable Z introduced before Theorem 3.4 and any $\varepsilon \ge 0$ and $c \in \mathbb{R}$, let $d_1(\varepsilon)$ and $m_i(c, \varepsilon)$, $i = 1, 2$, be the quantiles of Z and $|Z - c|$, respectively, such that $P(Z \le d_1(\varepsilon)) = 1/(2(1-\varepsilon))$, $P(|Z - c| \le m_1(c, \varepsilon)) = (1 - 2\varepsilon)/(2(1-\varepsilon))$ and $P(|Z - c| \le m_2(c, \varepsilon)) = 1/(2(1-\varepsilon))$.

THEOREM 3.5. *Assume that the conditions in Theorem* 3.4 *hold. Then*:

(i)

$B(\varepsilon; \mathrm{PWM}, F)$

$$= \sqrt{\lambda_1} \sup_{r \ge 0} \frac{|(1-\varepsilon) \int x_1 w(1/(1 + f_1(x, r, \varepsilon)))\, dF_0 + \varepsilon r w(1/(1 + f_2(r, \varepsilon)))|}{(1-\varepsilon) \int w(1/(1 + f_1(x, r, \varepsilon)))\, dF_0 + \varepsilon w(1/(1 + f_2(r, \varepsilon)))},$$

(ii) $$\gamma(\mathrm{PWM}, F) = \gamma^*(\mathrm{PWM}, F) = \frac{\sqrt{\lambda_1} \sup_{r \ge 0} |k_0 + r w(1/(1+r))|}{\int w(1/(1 + \|x\|))\, dF_0(x)},$$

where

$$f_1(x,r,\varepsilon) = \sup_{0\le u_1\le 1} \frac{\sqrt{1-u_1^2}\|x_2\| + |u_1x_1 - f_4(u_1,r,d_1)|}{f_3(u_1,r,d_1)},$$

$$f_2(r,\varepsilon) = \sup_{0\le u_1\le 1} \frac{|u_1r - f_4(u_1,r,d_1)|}{f_3(u_1,r,d_1)},$$

where $f_3(u_1,r,d_1)$ *is the median of* $\{m_1(f_4(u_1,r,d_1),\varepsilon), |u_1r - f_4(u_1,r,d_1)|, m_2(f_4(u_1,r,d_1),\varepsilon)\}$ *and* $f_4(u_1,r,d_1)$ *is the median of* $\{-d_1, u_1r, d_1\}$, $x' = (x_1, x_2')$ *and* $u' = (u_1, u_2')$.

4. Projection medians. For a given depth function, the point with the maximum depth (the deepest point) is called the depth median. For a given projection depth PD, the deepest point is called the projection median, denoted by $\mathrm{PM}(F)$, that is, $\mathrm{PM}(F) = \arg\max_{x\in\mathbb{R}^d} \mathrm{PD}(x,F)$. Since $\mathrm{PM}(F)$ is affine equivariant, we assume, w.l.o.g., $\mathrm{PM}(F) = 0$ throughout our discussion. Define

$$C^*(\varepsilon;\sigma,F) = \frac{B^*(\varepsilon;\sigma,F)}{B_*(\varepsilon;\sigma,F)} = \frac{\sup_{\|u\|=1,G}\ \sigma(F_u(\varepsilon,G))}{\inf_{\|u\|=1,G}\ \sigma(F_u(\varepsilon,G))}.$$

THEOREM 4.1. *For any given distribution* F *in* $\mathbb{R}^d$ *and* $\varepsilon > 0$,

$$\begin{aligned}(4.1)\quad B(\varepsilon;\mathrm{PM},F) &\le \sup_G \sup_{\|u\|=1}\left(\mu\big(F_u(\varepsilon,G)\big) + \sigma\big(F_u(\varepsilon,G)\big)\sup_{\|u\|=1}\frac{\mu(F_u(\varepsilon,G))}{\sigma(F_u(\varepsilon,G))}\right)\\ &\le B^*(\varepsilon;\mu,F)\big(1 + C^*(\varepsilon;\sigma,F)\big),\end{aligned}$$

$$(4.2)\quad \varepsilon^*(\mathrm{PM},F) \ge \min\{\varepsilon^*(\mu,F), \varepsilon^*(\sigma,F)\}.$$

The upper bound of the maximum bias and the lower bound of the breakdown point of PM are determined by those of μ and σ, respectively. When $(\mu,\sigma) =$(Med, MAD), it is seen that the maximum bias of PM is bounded for contamination less than 50%, that is, the breakdown point of PM is $1/2$. Note that in many cases the upper bound in (4.1) is strictly less than that in (4.2). For centrally symmetric F, Adrover and Yohai (2002) obtained the upper bound of $B(\varepsilon;\mathrm{PM},F)$ in (4.2).

A distribution F in $\mathbb{R}^d$ is called *μ-symmetric* about a point θ if $\mu(F_u) = u'\theta$ for any $\|u\| = 1$ [Zuo (2003)]. In the following discussion, we assume w.l.o.g. that $\theta = 0$. Note that any standard symmetric F is μ-symmetric with μ being the median functional. Any distribution F of $X \in \mathbb{R}^d$ is μ-symmetric about $E(X)$ with μ being the mean functional, provided that $E(X)$ exists.

THEOREM 4.2. *Assume that F is μ-symmetric about the origin and the conditions in Theorem* 3.2 *hold*. *Then*

$$\gamma^*(\mathrm{PM}, F) \le \sup_x \sup_{\|u\|=1} \left(\mathrm{IF}(u'x; \mu, F_u) + \sigma(F_u) \sup_{\|u\|=1} \frac{\mathrm{IF}(u'x; \mu, F_u)}{\sigma(F_u)} \right) \tag{4.3}$$

$$\le \gamma^*(\mu, F)\big(1 + C^*(0; \sigma, F)\big). \tag{4.4}$$

The gross error sensitivity upper bound of PM is determined by that of μ and $C^*(0; \sigma, F)$ for μ-symmetric F and is finite provided that $\gamma^*(\mu, F) < \infty$. Note that in many cases the upper bound in (4.3) is strictly less than that in (4.4). Neither of these bounds was given by Adrover and Yohai (2002).

Now we derive the exact expressions of the maximum bias function, the contamination and gross error sensitivities, and the influence function of PM. The derivations are again technically very challenging. In the definition of elliptical symmetry before Theorem 3.4, if $a(u)$ is a very general function (not necessarily $\sqrt{u'\Sigma u}$), then F (or X) is called *d-version symmetric* about the origin [Eaton (1981)]. In a similar manner, d-version symmetric about an arbitrary $\theta \in \mathbb{R}^d$ can be defined, but again we assume w.l.o.g. that $\theta = 0$. Clearly, d-version symmetric F includes a wide class of distributions such as elliptically and α-symmetric F $(0 < \alpha \le 2)$ [Fang, Kotz and Ng (1990)]. The latter corresponds to $a(u) = (\sum_{i=1}^d a_i |u_i|^\alpha)^{1/\alpha}$ with $a_i > 0$. The F generated from i.i.d. Cauchy marginals is α-symmetric with $\alpha = 1$ [Eaton (1981)]. Define for any $\varepsilon \ge 0$, $d_2 = d_2(\varepsilon) = m_2(d_1, \varepsilon)$, $d_3 = d_3(\varepsilon) = m_1(d_1, \varepsilon)$ and $d_0 = d_0(\varepsilon) = \sup_{0 \le c \le d_1} c/m_1(c, \varepsilon)$. Let $u^0 \in \arg\max_{\|u\|=1} a(u)$.

THEOREM 4.3. *Let $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD})$, let $X \sim F$ in $\mathbb{R}^d$ $(d > 1)$ be d-version symmetric about the origin and let $a(u)/a(u^0) \ge u'u^0$. Let the density $h(x)$ of Z be nonincreasing in $\|x\|$ and positive and continuous in small neighborhoods of* $\mathrm{Med}(Z)$ *and* $\mathrm{MAD}(Z)$. *Then*:

(i) $B(\varepsilon; \mathrm{PM}, F) = a(u^0)(d_1 + d_0 d_2)$;

(ii) $\gamma(\mathrm{PM}, F) = a(u^0)/h(0)$;

(iii) $\varepsilon^* = 1/2$.

The theorem indicates that the maximum bias of PM is bounded for d-version symmetric distributions as long as the contamination is less than 50%. Note that the condition $a(u)/a(u^0) \ge u'u^0$ is satisfied by a wide class of distributions including elliptically and α-symmetric distributions with $1 \le \alpha \le 2$. The detailed proof for the latter case can be obtained from the authors upon request. The behavior of $a(u)$ corresponding to α-symmetry with respect to various α's and $d = 2$ is illustrated in Appendix B.

For spherically symmetric F, Adrover and Yohai (2002) obtained $B(\varepsilon; \mathrm{PM}, F)$. Under the assumption that $c/m_1(c, \varepsilon)$ is nondecreasing for all small $\varepsilon > 0$, they also obtained $\gamma(\mathrm{PM}, F)$ for the same type of F. By Theorem 4.3, $B(\varepsilon; \mathrm{PM}, F)$ is dimension-free for spherically symmetric F. Adrover and Yohai (2002) conjectured that the dimension-free property of $B(\varepsilon; \mathrm{PM}, F)$ no longer holds for F generated from d independently and identically distributed Cauchy marginals. Theorem 4.3 proves this conjecture since for this F, $a(u^0) = \sqrt{d}$. Furthermore, by Theorem 4.3 it can be shown that $B(\varepsilon; \mathrm{PM}, F)$ and $\gamma(\mathrm{PM}, F)$ increase at a rate $d^{(2-\alpha)/(2\alpha)}$ as $d \to \infty$ when $a(u) = (\sum_{i=1}^{d} |u_i|^{\alpha})^{1/\alpha}$, $1 \le \alpha \le 2$. For the F generated from i.i.d. Cauchy marginals, Chen and Tyler (2002) showed that the maximum bias and contamination sensitivity of the half-space median also increase at rate $\sqrt{d}$.

Note that $B(\varepsilon; \mathrm{PM}, F)$ and $\gamma(\mathrm{PM}, F)$ can be dimension-free even for elliptically symmetric F if we adopt the norm $\|x\|_{\Sigma} = \sqrt{x'\Sigma^{-1}x}$ in the definition of $B(\varepsilon; \mathrm{PM}, F)$. Here the dimension-dependent factor is absorbed into the original $B(\varepsilon; \mathrm{PM}, F)$.

From the proof of Theorem 4.3 it is seen that the maximum bias of PM attains the upper bound in (4.1). Hence (4.1) is a sharp upper bound. On the other hand, by Lemmas A.2 and A.3 in Appendix A, $B^*(\varepsilon; \mu, F)(1 + C^*(\varepsilon; \sigma, F)) \ge a(u^0)d_1(1 + \sup_{0 \le c \le d_1}(d_2/(m_1(c, \varepsilon)))) > a(u^0)(d_1 + d_2 d_0)$, indicating that the upper bound in (4.1) is sharper than that in (4.2).

Now we present the exact influence function and gross error sensitivity expressions of PM.

THEOREM 4.4. *Let* $(\mu, \sigma) = (\mathrm{Med}, \mathrm{MAD})$ *and* $X \sim F$ *in* $\mathbb{R}^d$ $(d > 1)$ *be elliptically symmetric about the origin*, $h(x)$ *be nonincreasing in* $\|x\|$, *and positive and continuous in small neighborhoods of* $\mathrm{Med}(Z)$ *and* $\mathrm{MAD}(Z)$. *Then for any* $x \in \mathbb{R}^d$:

(i) $$\mathrm{IF}(x; \mathrm{PM}, F) = \frac{1}{h(0)} \frac{x}{\sqrt{x'\Sigma^{-1}x}};$$

(ii) $$\gamma^*(\mathrm{PM}, F) = \frac{\sqrt{\lambda_1}}{h(0)}.$$

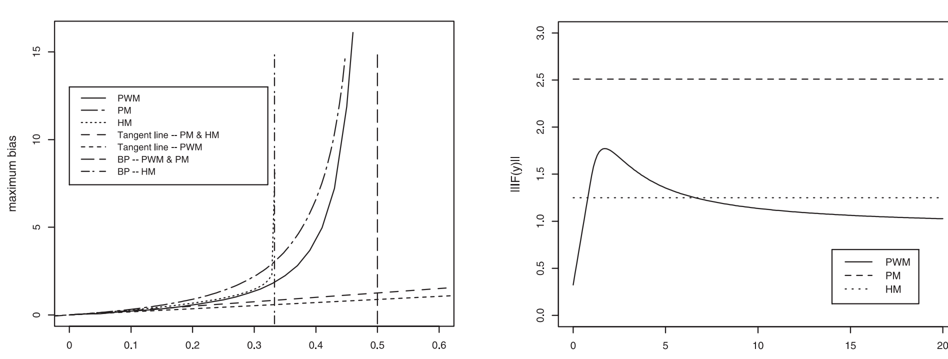
The theorem says that the influence function of PM is bounded [with its norm $\le \sqrt{\lambda_1}/h(0)$]. Note that the above gross error sensitivity of PM attains the upper bound in (4.3). To see this, note that by Lemmas A.2 and A.3 in Appendix A, the upper bound in (4.3) is greater than or equal to $\sqrt{\lambda_1}/h(0)$. Hence (4.3) is a sharp upper bound. Denote by λ_d the smallest eigenvalue of Σ associated with an elliptically symmetric F. The upper bound in (4.4) is greater than or equal to $\sqrt{\lambda_1}(1 + \sqrt{\lambda_1/\lambda_d})/(2h(0)) > \sqrt{\lambda_1}/h(0)$ if $\lambda_1 > \lambda_d$, indicating that the upper bound in (4.3) is sharper than that in (4.4).

Under spherical symmetry and other assumptions, Adrover and Yohai (2002) obtained $\gamma^*(\mathrm{PM}, F)$ but not $\mathrm{IF}(x; \mathrm{PM}, F)$.

5. Simulation and comparison. In this section, the behavior of the maximum bias and influence function of the two types of projection depth estimators is examined and compared with that of the half-space median (HM), a very popular competitor for robust and nonparametric estimation of multivariate location. The weight function w_1 in (3.3) is utilized here.

5.1. *Maximum bias, contamination sensitivity and breakdown point.* By Theorems 3.1 and 4.1, both the projection depth based estimators can have breakdown point $1/2$, substantially higher than that of the half-space median $1/3$ [Chen and Tyler (2002)]. The maximum bias functions $B(\varepsilon; \mathrm{PWM}, F)$ and $B(\varepsilon; \mathrm{PM}, F)$ are given in Theorems 3.5 and 4.3 for elliptically and d-version symmetric F, respectively. For elliptically symmetric F, $B(\varepsilon; \mathrm{HM}, F) = \sqrt{\lambda_1}\, F_Z^{-1}(\frac{1+\varepsilon}{2(1-\varepsilon)})$ [Chen and Tyler (2002)]. These three functions are plotted in Figure 1(a) for $N_2(\mathbf{0}, \mathbf{I}_2)$. The figure indicates that HM breaks down at $1/3$ while PM and PWM do not break down for any contamination less than 50%. The maximum bias of PWM is lower than that of HM and PM for any $\varepsilon < 1/2$. This actually is true for $d \le 5$ [but not true for $d > 5$ since $B(\varepsilon; \mathrm{PWM}, F)$ increases at a rate of $\sqrt{d}$ as d increases while $B(\varepsilon; \mathrm{PM}, F)$ and $B(\varepsilon; \mathrm{HM}, F)$ are dimension-free for normal models]. The simulation in Table 1 confirms this fact. The function $B(\varepsilon; \mathrm{PM}, F)$ is slightly higher than $B(\varepsilon; \mathrm{HM}, F)$ for ε close to or less than 0.3. The latter jumps to infinity when $\varepsilon \to 1/3$ while the former remains finite for any $\varepsilon < 1/2$.

The slope of the tangent line of $B(\varepsilon; T, F)$ at $(0, 0)$ is the contamination sensitivity $\gamma(T, F)$. For elliptically symmetric F, $\gamma(\mathrm{HM}, F) = \gamma(\mathrm{PM}, F) = \sqrt{\lambda_1}/h(0)$, which is $\sqrt{2\pi} \approx 2.507 > \gamma(\mathrm{PWM}, F) \approx 1.772$ for $N_2(\mathbf{0}, \mathbf{I}_2)$. This fact is reflected in Figure 1(a). The tangent line for PWM is slightly lower than that for HM and PM, indicating that the maximum bias function of PWM increases slightly more slowly than that of HM and PM for small $\varepsilon > 0$.



(a) (b)

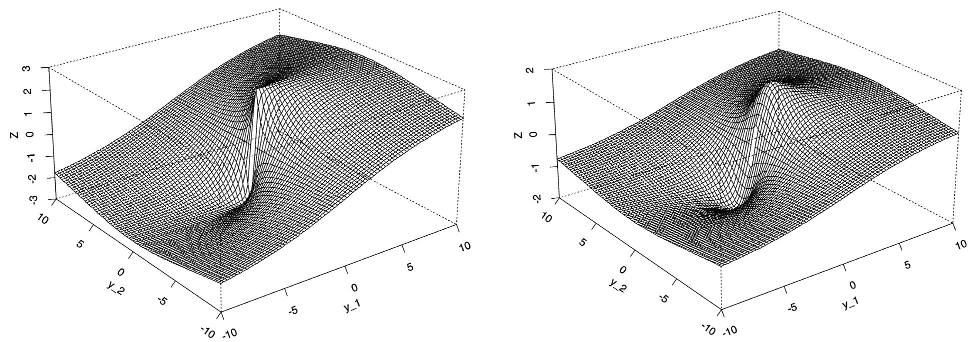
FIG. 1. (a) *Maximum bias functions, contamination sensitivities (the slopes of the tangent lines) and breakdown points of PM, PWM and HM.* (b) *Influence function norms of PM, PWM and HM.*

TABLE 1
Maximum biases of PWM and PM

d	$\mathbf{PWM_0}$	$\mathbf{PWM_1}$	**PM**	$\mathbf{PWM_0}$	$\mathbf{PWM_1}$	**PM**
	$\varepsilon = 0.05$			$\varepsilon = 0.10$		
2	0.076	0.072	0.141	0.196	0.199	0.321
3	0.108	0.092	0.141	0.231	0.226	0.321
4	0.120	0.116	0.141	0.276	0.292	0.321
5	0.131	0.121	0.141	0.310	0.312	0.321
6	0.146	0.141	0.141	0.329	0.334	0.321
7	0.157	0.158	0.141	0.355	0.356	0.321
8	0.165	0.172	0.141	0.392	0.395	0.321
9	0.176	0.183	0.141	0.399	0.399	0.321
10	0.186	0.190	0.141	0.423	0.430	0.321
15	0.236	0.243	0.141	0.531	0.538	0.321
20	0.262	0.258	0.141	0.620	0.613	0.321
	$\varepsilon = 0.20$			$\varepsilon = 0.30$		
2	0.536	0.568	0.897	1.408	1.354	2.201
3	0.659	0.680	0.897	1.848	1.679	2.201
4	0.771	0.800	0.897	2.248	1.973	2.201
5	0.899	0.889	0.897	2.570	2.198	2.201
6	1.030	0.964	0.897	2.994	2.426	2.201
7	1.089	1.041	0.897	3.256	2.637	2.201
8	1.195	1.113	0.897	3.552	2.816	2.201
9	1.256	1.172	0.897	3.789	2.991	2.201
10	1.304	1.249	0.897	4.005	3.185	2.201
15	1.644	1.525	0.897	5.036	3.935	2.201
20	1.964	1.782	0.897	6.035	4.609	2.201

5.2. *Influence function and gross error sensitivity.* By Theorem 4.4 and Chen and Tyler (2002), $\mathrm{IF}(y; \mathrm{PM}, F) = 2\mathrm{IF}(y; \mathrm{HM}, F) = 1/(h(0))y/\|\Sigma^{-1/2}y\|$ for elliptically symmetric F. For specific w and F, $\mathrm{IF}(y; \mathrm{PWM}, F)$ in Theorem 3.4 takes a much simpler form. For example, for $N_d(\mathbf{0}, \mathbf{I}_d)$ and $w_0(r) = r/(1-r)$ (Theorem 3.4 holds with this w_0), $\mathrm{IF}(y; \mathrm{PWM}, F) = k_1/(h(0))y_0$ with $y_0 = \frac{y}{\|y\|}$ and $k_1 = \frac{\Gamma(d/2)}{2\sqrt{\pi}}(1/\Gamma(\frac{d+1}{2}) + 2/\Gamma(\frac{d-1}{2}))$. For $N_2(\mathbf{0}, \mathbf{I}_2)$, $k_1 = 2/\pi$, corresponding to a constant $\|\mathrm{IF}(y; \mathrm{PWM}, F)\|$. In Figure 1(b), the norms of the three influence functions are plotted against $\|y\|$ at $N_2(\mathbf{0}, \mathbf{I}_2)$ and $w_1(r)$ given in (3.3) with $c = 1/(1+\sqrt{d})$ and $k = 3$ is used for PWM. $\|\mathrm{IF}(y; \mathrm{HM}, F)\|$ and $\|\mathrm{IF}(y; \mathrm{PM}, F)\|$ are constants $\sqrt{2\pi}/2$ (≈ 1.253) and $\sqrt{2\pi}$ (≈ 2.507), respectively. The function $\|\mathrm{IF}(y; \mathrm{PWM}, F)\|$ (as a function of $\|y\|$) increases from 0 to its maximum value $\gamma^*(\mathrm{PWM}, F) = 1.772$ and then decreases and is much smaller than that of PM and HM for most y.

In Figures 2(a) and (b), the first coordinates of the influence functions of PM and PWM are plotted as functions of y [the graphs of the second coordinates are similar



(a) (b)

FIG. 2. (a) *The first coordinate of the influence function of PM*. (b) *The first coordinate of the influence function of PWM*.

and the graphs of $\mathrm{IF}(y, \mathrm{HM}, F)$ are similar to those of $\mathrm{IF}(y, \mathrm{PM}, F)$]. In both Figures 2(a) and (b), $y = 0$ is a discontinuous (jump) point. When $y(\neq 0) \to 0$, the first coordinates of the influence functions approach $\gamma^*\mathrm{sign}(y_1)$, the extreme points.

5.3. *Dimension-free property.* The maximum bias and sensitivity functions of HM and PM are dimension-free for spherically symmetric F. This dimension-free property, however, does not hold for other F's such as the F from i.i.d. Cauchy marginals. Functions $B(\varepsilon; \mathrm{PWM}, F)$ and $\gamma^*(\mathrm{PWM}, F)$ are dimension-dependent. Based on an empirical argument and simulation studies (see Table 1), we conclude that they increase at a rate of $\sqrt{d}$ as $d \to \infty$ for fixed ε (this is also confirmed by the fact that $k_1/\sqrt{d}$ has a limit as $d \to \infty$). Table 1 lists maximum biases of PWM and PM for different d's and ε's and $X \sim N_d(\mathbf{0}, \mathbf{I}_d)$, and w_0 and w_1 are employed for PWM (called PWM_0 and PWM_1, respectively). The maximum biases of PM are constants for different d's and fixed ε. The maximum biases of PWM_0 are very competitive with those of PWM_1 for small ε and d; they become consistently larger than those of PWM_1 for large ε or d. Adrover and Yohai (2002) also reported the maximum biases of PWM_0, which are consistent with our results for small ε and d. In our computation, the exact maximum bias formula in Theorem 3.5 is utilized, which leads to a one-dimensional computational problem (otherwise a d-dimensional computational problem is encountered). For small ε, the results for PWM_0 in the table are confirmed by using $\varepsilon(k_1/(h(0)))$.

5.4. *Weight function.* The idea behind w_1 is [see Zuo, Cui and He (2004)] that with proper k and c, weight 1 is given to the half of the points with higher depth, while the weight given to the other half of the points decreases rapidly (exponentially) as the depth of these points decreases, leading to a desired balance between efficiency and robustness of the resulting depth weighted mean. The weight w_1 is shown to result in a highly efficient PWM_1 (which is more efficient than PWM_0). It is seen that PWM_1 also outperforms PWM_0 with respect to bias and sensitivity.

5.5. *Computing issue.* The formula for $B(\varepsilon; \mathrm{PWM}, F)$ given in Theorem 3.5 seems somewhat awkward. Computationally, however, it is very feasible. For the $N_d(\mathbf{0}, \mathbf{I}_d)$ distribution the computation is rather straightforward because one needs only to generate independently $N(0, 1)$ points ($x_1 \sim N(0, 1)$) and $\chi^2(d-1)$ points ($\|x_2\|^2 \sim \chi^2(d-1)$). The problem then becomes one-dimensional.

5.6. *Conclusion.* In summary, PM and PWM have breakdown point $1/2$, which is much higher than the $1/3$ value for HM. For spherically symmetric F, HM and PM have bounded and dimension-free influence and maximum bias functions, while those of PWM are bounded but dimension-dependent (with a rate $\sqrt{d}$ as $d \to \infty$). The dimension-free advantage of HM and PM, however, disappears for other distributions. Furthermore, there seems to be a trade-off between the dimension-free property and high efficiency. The two cannot always work in tandem. Note that PWM_1 possesses an extremely high efficiency (94%) in $\mathbb{R}^2$ and approaches 100% rapidly as d increases [Zuo, Cui and He (2004)], while the finite sample efficiencies of HM and PM are approximately 77 and 79%, respectively [Rousseeuw and Ruts (1998) and Zuo (2003)].

Taking all the findings into account, we conclude that projection depth based location estimators appear to be very competitive and represent favorable choices for location estimators when compared to their competitors.

APPENDIX A

Selected (sketches of) proofs and auxiliary lemmas.

PROOF OF THEOREM 3.1. (i) Write $\mathrm{PWM}(F(\varepsilon, G)) - \mathrm{PWM}(F) = (I_1 + I_2)/I_3$ with

$$\begin{aligned} I_1 &= I_1(\varepsilon, G) \\ &= (1-\varepsilon)\int \big(x - \mathrm{PWM}(F)\big)\big[w\big(\mathrm{PD}(x, F(\varepsilon, G))\big) - w\big(\mathrm{PD}(x, F)\big)\big]\, dF(x), \\ I_2 &= I_2(\varepsilon, G) = \varepsilon \int \big(x - \mathrm{PWM}(F)\big) w\big(\mathrm{PD}(x, F(\varepsilon, G))\big)\, dG(x), \\ I_3 &= I_3(\varepsilon, G) \\ &= (1-\varepsilon)\int w\big(\mathrm{PD}(x, F(\varepsilon, G))\big)\, dF(x) + \varepsilon \int w\big(\mathrm{PD}(x, F(\varepsilon, G))\big)\, dG(x) \end{aligned}$$

and

$$J_1(\varepsilon, G) = \sup_{\|u\|=1} \max\{|\mu(F_u(\varepsilon, G)) - u'\mathrm{PWM}(F)|, \sigma(F_u(\varepsilon, G))\},$$

$$J_2(\varepsilon, G) = \sup_{\|u\|=1} \max\{|\sigma(F_u(\varepsilon, G)) - \sigma(F_u)|, |\mu(F_u(\varepsilon, G)) - \mu(F_u)|\}/\sigma(F_u(\varepsilon, G)).$$

It follows that $\|x - \mathrm{PWM}(F)\|/(1 + O(x, F(\varepsilon, G))) \le J_1(\varepsilon, G)$ and

$$\begin{aligned}
&|\mathrm{PD}(x, F(\varepsilon, G)) - \mathrm{PD}(x, F)| \\
&\quad = \frac{|O(x, F(\varepsilon, G)) - O(x, F)|}{(1 + O(x, F(\varepsilon, G)))(1 + O(x, F))} \\
&\quad \le \sup_{\|u\|=1} \frac{O(x, F)|\sigma(F_u(\varepsilon, G)) - \sigma(F_u)| + |\mu(F_u(\varepsilon, G)) - \mu(F_u)|}{\sigma(F_u(\varepsilon, G))(1 + O(x, F(\varepsilon, G)))(1 + O(x, F))} \\
&\quad \le \frac{J_2(\varepsilon, G)}{1 + O(x, F(\varepsilon, G))}.
\end{aligned}$$

Thus

$$\begin{aligned}
\|I_1\| &\le w^0(1-\varepsilon) \int \|x - \mathrm{PWM}(F)\| |\mathrm{PD}(x, F(\varepsilon, G)) - \mathrm{PD}(x, F)| \, dF(x) \\
&\le w^0(1-\varepsilon) J_2(\varepsilon, G) \int \|x - \mathrm{PWM}(F)\| / (1 + O(x, F(\varepsilon, G))) \, dF(x) \\
&\le w^0(1-\varepsilon) J_1(\varepsilon, G) J_2(\varepsilon, G), \\
\|I_2\| &\le w^0 \varepsilon \int \|x - \mathrm{PWM}(F)\| \mathrm{PD}(x, F(\varepsilon, G)) \, dG(x) \le w_0 \varepsilon J_1(\varepsilon, G).
\end{aligned}$$

Observe that $\|I_3\| \ge A(\varepsilon, F)$. The desired result in part 1 follows.

(ii) Let $\varepsilon_0 = \min\{\varepsilon^*(\mu, F), \varepsilon^*(\sigma, F)\}$. Then $\mu_{\sup}(\varepsilon_0, F) \le B(\varepsilon_0; \mu, F) + \sup_{\|u\|} \mu(F_u) + \|\mathrm{PWM}(F)\| < \infty$, $B^*(\varepsilon_0; \sigma, F) \le B(\varepsilon_0; \sigma, F) + \sup_{\|u\|} \sigma(F_u) < \infty$ and $B_*(\varepsilon_0; \sigma, F) > 0$. Hence $A(\varepsilon_0, F) > 0$. Thus $B(\varepsilon_0; \mathrm{PWM}, F) < \infty$. □

Proof of Theorem 3.3.

LEMMA A.1. *Assume that*:

(i) *$u(x)$ is a singleton for a given x*,
(ii) *the conditions* 1 *and* 4 *of Theorem* 3.3 *hold, and*
(iii) $\mathrm{IF}(u(x)'y; \sigma, F_{u(x)})$ *and* $\mathrm{IF}(u(x)'y; \mu, F_{u(x)})$ *are continuous at $u(x)$ for a given y.*

Then

$$\mathrm{IF}(y; O(x, F), F) = -\big(O(x, F)\mathrm{IF}(u(x)'y; \sigma, F_{u(x)}) + \mathrm{IF}(u(x)'y; \mu, F_{u(x)})\big)/\sigma(F_{u(x)}).$$

PROOF. For any $\varepsilon > 0$ and the given $x, y \in \mathbb{R}^d$,

$$-\big(O\big(x, F(\varepsilon, \delta_y)\big) - O(x, F)\big) = \min\bigg\{\inf_{u \in u(x)} -g\big(x, u, F(\varepsilon, \delta_y)\big) + O(x, F),$$
$$\inf_{u \in u(x)^C} -g\big(x, u, F(\varepsilon, \delta_y)\big) + O(x, F)\bigg\},$$

where S^C denotes the complement of the set S. Define $u(x, \tau) = \{u : \|u\| = 1, d(u, u(x)) \le \tau\}$ for any $\tau > 0$, where $d(x, S) = \inf\{\|x - y\| : y \in S\}$ for any set S. Denote $\tilde{g}(u, x, y, \varepsilon, F) = -g(x, u, F(\varepsilon, \delta_y)) + O(x, F)$ and write

$$\inf_{u \in u(x)^C} \tilde{g}(u, x, y, \varepsilon, F)$$
$$= \min\bigg\{\inf_{u \in u(x)^C \cap u(x,\tau)^C} \tilde{g}(u, x, y, \varepsilon, F), \inf_{u \in u(x)^C \cap u(x,\tau)} \tilde{g}(u, x, y, \varepsilon, F)\bigg\}.$$

Observe that as $\varepsilon \to 0^+$,

$$\big(-g\big(x, u, F(\varepsilon, \delta_y)\big) + g(x, u, F)\big)/\varepsilon$$
$$= \frac{(u'x - \mu(F_u))(\sigma(F_u(\varepsilon, \delta_y)) - \sigma(F_u)) + \sigma(F_u)(\mu(F_u(\varepsilon, \delta_y)) - \mu(F_u))}{\varepsilon\sigma(F_u(\varepsilon, \delta_y))\sigma(F_u)}$$
$$\to \big(g(x, u, F)\mathrm{IF}(u'y; \sigma, F_u) + \mathrm{IF}(u'y; \mu, F_u)\big)/\sigma(F_u),$$

which is uniformly bounded in u for given x and y. By this and the given conditions,

$$\inf_{u \in u(x)^C \cap u(x,\tau)^C} \tilde{g}(u, x, y, \varepsilon, F)/\varepsilon$$
$$= \inf_{u \in u(x)^C \cap u(x,\tau)^C} \Big(\big(-g\big(x, u, F(\varepsilon, \delta_y)\big) + g(x, u, F)\big)/\varepsilon$$
$$+ \big(O(x, F) - g(x, u, F)\big)/\varepsilon\Big)$$
$$\ge \inf_{u \in u(x)^C \cap u(x,\tau)^C} \big(O(x, F) - g(x, u, F)\big)/(2\varepsilon) = O(1/\varepsilon)$$

for any given $\tau > 0$ and sufficiently small ε. Also

$$O(1) = \inf_{u \in u(x)^C \cap u(x,\tau)} \tilde{g}(u, x, y, \varepsilon, F)/\varepsilon$$
$$= \inf_{u \in u(x)^C \cap u(x,\tau)} \Big(\big(-g\big(x, u, F(\varepsilon, \delta_y)\big) + g(x, u, F)\big)/\varepsilon$$
$$+ \big(O(x, F) - g(x, u, F)\big)/\varepsilon\Big)$$
$$\ge \inf_{u \in u(x)^C \cap u(x,\tau)} \big(-g\big(x, u, F(\varepsilon, \delta_y)\big) + g(x, u, F)\big)/\varepsilon$$
$$\ge \inf_{u \in u(x,\tau)} \big(-g\big(x, u, F(\varepsilon, \delta_y)\big) + g(x, u, F)\big)/\varepsilon.$$

Thus for sufficiently small $\varepsilon > 0$ and any given $\tau > 0$,

$$\begin{aligned}
&\inf_{u\in u(x)^C} \bigl(-g(x,u,F_t) + O(x,F)\bigr)/\varepsilon \\
&\qquad = \inf_{u\in u(x)^C\cap u(x,\tau)} \bigl(-g\bigl(x,u,F(\varepsilon,\delta_y)\bigr) + O(x,F)\bigr)/\varepsilon \\
&\qquad \ge \inf_{u\in u(x,\tau)} \bigl(-g\bigl(x,u,F(\varepsilon,\delta_y)\bigr) + g(x,u,F)\bigr)/\varepsilon.
\end{aligned}$$

Hence, we have, for any given $\tau > 0$,

$$\begin{aligned}
&\inf_{u\in u(x,\tau)} \bigl(-g\bigl(x,u,F(\varepsilon,\delta_y)\bigr) + g(x,u,F)\bigr)/\varepsilon \\
&\qquad \le -\bigl(O\bigl(x,F(\varepsilon,\delta_y)\bigr) - O(x,F)\bigr)/\varepsilon \\
&\qquad \le \inf_{u\in u(x)} \bigl(-g\bigl(x,u,F(\varepsilon,\delta_y)\bigr) + g(x,u,F)\bigr)/\varepsilon.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\inf_{u\in u(x,\tau)} g(x,u,F)\mathrm{IF}(u'y;\sigma,F_u) + \mathrm{IF}(u'y;\mu,F_u)/\sigma(F_u) \\
&\qquad \le \lim_{\varepsilon\to 0^+} -\bigl(O\bigl(x,F(\varepsilon,\delta_y)\bigr) - O(x,F)\bigr)/\varepsilon \\
&\qquad \le \inf_{u\in u(x)} \bigl(g(x,u,F)\mathrm{IF}(u'y;\sigma,F_u) + \mathrm{IF}(u'y;\mu,F_u)\bigr)/\sigma(F_u).
\end{aligned}$$

Let $\tau \to 0^+$. By the given conditions 2 and 3, we obtain

$$\begin{aligned}
&\lim_{\varepsilon\to 0^+} -\bigl(O\bigl(x,F(\varepsilon,\delta_y)\bigr) - O(x,F)\bigr)/\varepsilon \\
&\qquad = \inf_{u\in u(x)} \bigl(O(x,F)\mathrm{IF}(u'y;\sigma,F_u) + \mathrm{IF}(u'y;\mu,F_u)\bigr)/\sigma(F_u).
\end{aligned}$$

Therefore, for point x with $u(x)$ a singleton, the desired result follows. □

PROOF OF THEOREM 3.3. By equivariance, assume, without loss of generality, that $\mathrm{PWM}(F) = 0$. By the given conditions and following the proof of Theorem 2.2 of Zuo (2003), it follows that $\mathrm{PD}(x, F(\varepsilon,\delta_y)) \to \mathrm{PD}(x,F)$ as $\varepsilon \to 0$. Lebesgue's dominated convergence theorem leads to

$$\int w\bigl(\mathrm{PD}\bigl(x,F(\varepsilon,\delta_y)\bigr)\bigr)F(\varepsilon,\delta_y)(dx) \to \int w\bigl(\mathrm{PD}(x,F)\bigr)F(dx) \qquad \text{as } \varepsilon \to 0.$$

The mean value theorem gives

$$\begin{aligned}
&\int xw\bigl(\mathrm{PD}\bigl(x,F(\varepsilon,\delta_y)\bigr)\bigr)\,dF\bigl(x;F(\varepsilon,\delta_y)\bigr) \\
&\qquad = (1-\varepsilon)\int xw^{(1)}(\xi)\bigl(\mathrm{PD}(x,F(\varepsilon,\delta_y)) - \mathrm{PD}(x,F)\bigr)\,dF(x) \\
&\qquad\quad + \varepsilon yw\bigl(\mathrm{PD}\bigl(y,F(\varepsilon,\delta_y)\bigr)\bigr)
\end{aligned}$$

$$= (1-\varepsilon)\int x w^{(1)}(\xi)\frac{O(x,F(x,F)) - O(x,F(\varepsilon,\delta_y))}{(1+O(x,F))(1+O(x,F(\varepsilon,\delta_y)))}\,dF(x)$$
$$+ \varepsilon y w\big(\mathrm{PD}\big(y, F(\varepsilon,\delta_y)\big)\big),$$

where ξ is a point between $\mathrm{PD}(x, F(\varepsilon,\delta_y))$ and $\mathrm{PD}(x,F)$. By the given conditions and applying again Lebesgue's dominated convergence theorem, we have

$$\frac{1}{\varepsilon}\int x w\big(\mathrm{PD}\big(x, F(\varepsilon,\delta_y)\big)\big)F(\varepsilon,\delta_y)(dx)$$
$$\to -\int x w^{(1)}(\mathrm{PD}(x,F))\mathrm{IF}\big(y, O(x,F), F\big)/\big(1+O(x,F)\big)^2\,dF(x)$$
$$+ y w\big(\mathrm{PD}(y,F)\big) \qquad \text{as } \varepsilon \to 0.$$

By Lemma A.1, the desired result now follows. □

PROOF OF THEOREM 3.4. (i) For the given elliptically symmetric F, it follows [see Example 2.1 of Zuo (2003)] that

$$O(x,F) = \|\Sigma^{-1/2}x\|, \qquad u(x) = \Sigma^{-1}x/\|\Sigma^{-1}x\| \qquad (x \neq 0)$$

and

$$\mathrm{PD}(x,F) = 1/\big(1+\|\Sigma^{-1/2}x\|\big).$$

From Cui and Tian (1994) it follows that

$$\mathrm{IF}\big(u(x)'y, \mathrm{Med}, F_{u(x)}\big) = \|\Sigma^{-1/2}x\|/\big(2h(0)\|\Sigma^{-1}x\|\big)\ \mathrm{sign}(x'\Sigma^{-1}y)$$

and

$$\mathrm{IF}\big(u(x)'y, \mathrm{MAD}, F_{u(x)}\big)$$
$$= \|\Sigma^{-1/2}x\|/\big(4h(1)\|\Sigma^{-1}x\|\big)\ \mathrm{sign}\big(|x'\Sigma^{-1}y| - \|\Sigma^{-1/2}x\|\big).$$

Note that the conditions in Theorem 3.3 are satisfied and $P(S(y)) = 0$ for $y \in \mathbb{R}^d$. Thus we obtain the corresponding influence function. Since $\mathrm{MAD}(F_{u(x)}) = \|\Sigma^{-1/2}x\|/\|\Sigma^{-1}x\|$ is even in x, we have

$$\int x w^{(1)}\big(\mathrm{PD}(x,F)\big)O(x,F)$$
$$\times\ \mathrm{IF}\big(u(x)'y; \mathrm{MAD}, F_{u(x)}\big)/\big(\mathrm{MAD}\big(F_{u(x)}\big)\big(1+O(x,F)\big)^2\big)\,dF(x) = 0.$$

On the other hand,

$$\int x w^{(1)}(\mathrm{PD}(x,F))\mathrm{IF}\big(u(x)'y; \mathrm{Med}, F_{u(x)}\big)/\big(\mathrm{MAD}\big(F_{u(x)}\big)\big(1+O(x,F)\big)^2\big)\,dF(x)$$

$$= 1/h(0)\int x w^{(1)}\frac{((1+\|\Sigma^{-1/2}x\|)^{-1})(1/2 - I\{x'\Sigma^{-1}y \le 0\})}{(1+\|\Sigma^{-1/2}x\|)^2}\,dF(x)$$

$$= \Sigma^{1/2}/h(0) \int x w^{(1)} \frac{((1+\|x\|)^{-1})(1/2 - I\{x'\Sigma^{-1/2}y \le 0\})}{(1+\|x\|)^2}\, dF_0(x)$$

$$= y/\big(2h(0)\|\Sigma^{-1/2}y\|\big) \int |x_1| w^{(1)}\big((1+\|x\|)^{-1}\big)/(1+\|x\|)^2\, dF_0(x).$$

By Theorem 3.3, the desired result follows.

(ii) It follows from part (i) that

$$\gamma^*(\mathrm{PWM}, F)$$

$$= \sup_{y\in\mathbb{R}^d} \frac{\|(k_0\Sigma^{1/2} + \Sigma^{1/2}\|\Sigma^{-1/2}y\| w(1/(1+\|\Sigma^{-1/2}y\|))\Sigma^{-1/2}y/\|\Sigma^{-1/2}y\|\|}{\int w(1/(1+\|x\|))\, dF_0(x)}$$

$$= \sup_{\substack{r\ge 0,\\ \|u\|=1}} \frac{|k_0 + r w(1/(1+r))|\|\Sigma^{1/2}u\|}{\int w(1/(1+\|x\|))\, dF_0(x)} = \frac{\sqrt{\lambda_1}\sup_{r\ge 0}|k_0 + r w(1/(1+r))|}{\int w(1/(1+\|x\|))\, dF_0(x)}.$$

This completes the proof. □

Proof of Theorem 3.5.

LEMMA A.2. *Suppose that* $a = F^{-1}(\frac{1-2\varepsilon}{2(1-\varepsilon)})$ *and* $b = F^{-1}(\frac{1}{2(1-\varepsilon)})$ *exist for some* F *in* $\mathbb{R}^1$ *and* $0 \le \varepsilon < 1$. *Then for any distribution* G *and point* x *in* $\mathbb{R}^1$, (i) $a \le \mathrm{Med}(F(\varepsilon, G)) \le b$, (ii) $\mathrm{Med}(F(\varepsilon, \delta_x)) = \mathrm{Med}\{a, b, x\}$, (iii) $m_1(\mathrm{Med}(F(\varepsilon, G)), \varepsilon) \le \mathrm{MAD}(F(\varepsilon, G)) \le m_2(\mathrm{Med}(F(\varepsilon, G)), \varepsilon)$ *and* (iv) $\mathrm{MAD}(F(\varepsilon, \delta_x)) = \mathrm{Med}\{m_1(\mathrm{Med}(F(\varepsilon, \delta_x)), \varepsilon), |x - \mathrm{Med}(F(\varepsilon, \delta_x))|, m_2(\mathrm{Med}(F(\varepsilon, \delta_x)), \varepsilon)\}$.

LEMMA A.3. *Suppose that* $X \sim F$ *is* d-*version symmetric about the origin. Then* (i) $\mathrm{Med}(F_u(\varepsilon, \delta_x)) = \mathrm{Med}\{-a(u)d_1, u'x, a(u)d_1\}$ *and* (ii) $\mathrm{MAD}(F_u(\varepsilon, \delta_x)) = \mathrm{Med}\{a(u)m_1(\mathrm{Med}(F_u(\varepsilon, \delta_x))/a(u), \varepsilon), |u'x - \mathrm{Med}(F_u(\varepsilon, \delta_x))|, a(u) \times m_2(\mathrm{Med}(F_u(\varepsilon, \delta_x))/a(u), \varepsilon)\}$.

LEMMA A.4. *Let* τ (> 0) *be sufficiently small*, $|s - t| \le \tau$ *and* $r \ge 0$. *Then* (i) $|\mathrm{Med}\{a_1, c, b_1\} - \mathrm{Med}\{a_2, c, b_2\}| \le |a_2 - a_1| + |b_2 - b_1|$ *for any* c *and* $a_i \le b_i, i = 1, 2$, (ii) $|\mathrm{Med}\{-a, sr, a\} - \mathrm{Med}\{-b, tr, b\}| \le |a-b| + 2(a+b)(\sqrt{\tau} + I\{|s| \le \sqrt{\tau}\})$ *for* $a, b > 0$ *and* (iii) $|\mathrm{Med}\{a, |sr - c_1|, b\} - \mathrm{Med}\{a, |tr - c_2|, b\}| \le 2m(I\{|s| \le \sqrt{\tau}\} + I\{|sr - c_0| \le c\sqrt{\tau}\} + I\{|sr + c_0| \le c\sqrt{\tau}\})$ *for* $c_0 > 0$ *a fixed constant*, $c = 2c_0 + 1$, $a \le c_0 \le b$, $m = \max\{|a - c_0|, |b - c_0|\} \le \tau$ *and* $|c_i| \le \tau$, $i = 1, 2$. *Here* $I\{\cdot\}$ *stands for the indicator function.*

LEMMA A.5. *If the density* h *of* Z *is continuous in a small neighborhood of* 1 $[= \mathrm{MAD}(Z)]$, *then for small* $\varepsilon > 0$, $0 \le d_1(\varepsilon) \le A_1\varepsilon$, $|m_i(0, \varepsilon) - 1| \le A_2\varepsilon$ *and* $\sup_{0\le|c|\le A_3\varepsilon} |m_i(c, \varepsilon) - m_i(0, \varepsilon)| = o(\varepsilon)$ *for some positive constants* A_1, A_2, A_3 *and* $i = 1, 2$.

LEMMA A.6. *Let* $(\mu,\sigma) = (\mathrm{Med}, \mathrm{MAD})$, *let* $X \sim F$ *in* $\mathbb{R}^d$ *be elliptically symmetric about the origin with associated* Σ *and let the density* h *of* Z *be continuous and* $h(0)h(1) > 0$. *Then*

$$E\biggl[\sup_{\|u-u(X)\|\le\tau} \frac{1}{\varepsilon}\biggl(\biggl|\frac{\tilde\mu(F_u,\varepsilon,y)}{\sigma(F_u)} - \frac{\tilde\mu(F_{u(X)},\varepsilon,y)}{\sigma(F_{u(x)})}\biggr| + \biggl|\frac{\tilde\sigma(F_u,\varepsilon,y)}{\sigma(F_u)} - \frac{\tilde\sigma(F_{u(X)},\varepsilon,y)}{\sigma(F_{u(x)})}\biggr|\biggr)\biggr] \to 0$$

uniformly in $y \in \mathbb{R}^d$ *as* $\varepsilon \to 0$, *where* $u(x) = \Sigma^{-1}x/\|\Sigma^{-1}x\|$, $\tau = \tau(\varepsilon)$ $(\ge \varepsilon > 0) \to 0$ *as* $\varepsilon \to 0$ *and* $\tilde\mu(F_v,\varepsilon,y) = \mu(F_v(\varepsilon,\delta_y)) - \mu(F_v)$ *and* $\tilde\sigma(F_v,\varepsilon,y) = \sigma(F_v(\varepsilon,\delta_y)) - \sigma(F_v)$ *for a unit vector* v.

PROOF OF THEOREM 3.5. (i) By Lemma A.3, for any $y \in \mathbb{R}^d$, we have that

$$\begin{aligned}
\frac{\mu(F_u(\varepsilon,\delta_y))}{a(u)} &= \frac{\mathrm{Med}\{-a(u)d_1, u'y, a(u)d_1\}}{a(u)} \\
&= \mathrm{Med}\biggl\{-d_1, (\Sigma^{1/2}u)'\frac{\Sigma^{-1/2}y}{a(u)}, d_1\biggr\}, \\
\frac{\sigma(F_u(\varepsilon,\delta_y))}{a(u)} &= \mathrm{Med}\biggl\{m_1\biggl(\frac{\mu(F_u(\varepsilon,\delta_y))}{a(u)},\varepsilon\biggr), \\
&\qquad \biggl|\frac{(\Sigma^{1/2}u)'}{a(u)}\Sigma^{-1/2}y - \frac{\mu(F_u(\varepsilon,\delta_y))}{a(u)}\biggr|, m_2\biggl(\frac{\mu(F_u(\varepsilon,\delta_y))}{a(u)},\varepsilon\biggr)\biggr\}.
\end{aligned}$$

Let $v = \Sigma^{1/2}u/a(u)$, $\tilde y = \Sigma^{-1/2}y$ and $\tilde x = \Sigma^{-1/2}x$. Then all the mappings are one-to-one and $\|v\| = 1$. Denote $f_5(u,x,d_1) = \mathrm{Med}\{-d_1, u'x, d_1\}$. Then

$$\begin{aligned}
&O(x, F(\varepsilon,\delta_y)) \\
&\quad = \sup_{\|v\|=1} \frac{v'\tilde x - f_5(v,\tilde y,d_1)}{\mathrm{Med}\{m_1(f_5(v,\tilde y,d_1),\varepsilon), |v'\tilde y - f_5(v,\tilde y,d_1)|, m_2(f_5(v,\tilde y,d_1),\varepsilon)\}}.
\end{aligned}$$

Let U be an orthogonal matrix with $\tilde y/\|\tilde y\|$ as its first column and let $U'v = \tilde v$. Then $f_5(v,\tilde y,d_1) = \mathrm{Med}\{-d_1, \tilde v_1\|\tilde y\|, d_1\} = f_4(\tilde v_1, \|\tilde y\|, d_1)$ and $O(x, F(\varepsilon,\delta_y))$ becomes

$$\begin{aligned}
&\sup_{\|\tilde v\|=1}\Bigl[\bigl(\tilde v'U'\tilde x - f_4(\tilde v_1,\|\tilde y\|,d_1)\bigr) \\
&\qquad \times \Bigl(\mathrm{Med}\Bigl\{m_1\bigl(f_4(\tilde v_1,\|\tilde y\|,d_1),\varepsilon\bigr), \\
&\qquad\qquad |\tilde v_1\|\tilde y\| - f_4(\tilde v_1,\|\tilde y\|,d_1)|, m_2\bigl(f_4(\tilde v_1,\|\tilde y\|,d_1),\varepsilon\bigr)\Bigr\}\Bigr)^{-1}\Bigr] \\
&\quad = \sup_{\|\tilde v\|=1} \tilde v'U'\tilde x - \frac{f_4(\tilde v_1,\|\tilde y\|,d_1)}{f_3(\tilde v_1,\|\tilde y\|,d_1)}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\int xw\big(\mathrm{PD}(x, F(\varepsilon, \delta_y))\big)\,dF \\
&\qquad = \int \Sigma^{1/2}\tilde{x}w\bigg(1 + \sup_{\|\tilde{v}\|=1} \tilde{v}'U'\tilde{x} - f_4(\tilde{v}_1, \|\tilde{y}\|, d_1)/f_3(\tilde{v}_1, \|\tilde{y}\|, d_1)\bigg)^{-1}\,dF \\
&\qquad = \int \Sigma^{1/2}Uxw\bigg(1 + \sup_{\|u\|=1} u'x - f_4(u_1, \|\tilde{y}\|, d_1)/f_3(u_1, \|\tilde{y}\|, d_1)\bigg)^{-1}\,dF_0.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\sup_{\|u\|=1} \frac{u'x - f_4(u_1, \|\tilde{y}\|, d_1)}{f_3(u_1, \|\tilde{y}\|, d_1)} \\
&\qquad = \sup_{-1\le u_1\le 1}\ \sup_{\|u_2\|=\sqrt{1-u_1^2}} \frac{u_2'x_2 + u_1x_1 - f_4(u_1, \|\tilde{y}\|, d_1)}{f_3(u_1, \|\tilde{y}\|, d_1)} \\
&\qquad = \sup_{0\le u_1\le 1} \frac{\sqrt{1-u_1^2}\|x_2\| + |u_1x_1 - f_4(u_1, \|\tilde{y}\|, d_1)|}{f_3(u_1, \|\tilde{y}\|, d_1)} = f_1(x, \|\tilde{y}\|, \varepsilon),
\end{aligned}$$

an even function of x_2. Hence $\int xw(\mathrm{PD}(x, F(\varepsilon, \delta_y)))\,dF = \int y/\|\tilde{y}\|x_1w(1 + f_1(x, \|\tilde{y}\|, \varepsilon))^{-1}\,dF_0$. Therefore

$$\begin{aligned}
&B(\varepsilon; \mathrm{PWM}, F) \\
&\qquad = \sup_{y\in\mathbb{R}^d}\Bigg[\bigg\{\bigg|(1-\varepsilon)\int (y/\|\tilde{y}\|)x_1w\big(1 + f_1(x, \|\tilde{y}\|, \varepsilon)\big)^{-1}\,dF_0 \\
&\qquad\qquad + \varepsilon yw\big(1 + f_2(\|\tilde{y}\|, \varepsilon)\big)^{-1}\bigg|\bigg\} \\
&\qquad\qquad \times \bigg\{(1-\varepsilon)\int w\big(1 + f_1(x, \|\tilde{y}\|, \varepsilon)\big)^{-1}\,dF_0 \\
&\qquad\qquad + \varepsilon w\big(1 + f_2(\|\tilde{y}\|, \varepsilon)\big)^{-1}\bigg\}^{-1}\Bigg] \\
&\qquad = \sup_{\substack{r\ge 0,\\ \|u\|=1}}\Bigg[\|\Sigma^{1/2}u\| \\
&\qquad\qquad \times \frac{|(1-\varepsilon)\int x_1w(1 + f_1(x, r, \varepsilon))^{-1}\,dF_0 + \varepsilon rw(1 + f_2(r, \varepsilon))^{-1}|}{(1-\varepsilon)\int w(1 + f_1(x, r, \varepsilon))^{-1}\,dF_0 + \varepsilon w(1 + f_2(r, \varepsilon))^{-1}}\Bigg] \\
&\qquad = \sqrt{\lambda_1}\sup_{r\ge 0} \frac{|(1-\varepsilon)\int x_1w(1 + f_1(x, r, \varepsilon))^{-1}\,dF_0 + \varepsilon rw(1 + f_2(r, \varepsilon))^{-1}|}{(1-\varepsilon)\int w(1 + f_1(x, r, \varepsilon))^{-1}\,dF_0 + \varepsilon w(1 + f_2(r, \varepsilon))^{-1}}.
\end{aligned}$$

(ii) Since $\gamma(\mathrm{PWM}, F) \geq \gamma^*(\mathrm{PWM}, F)$, we now show that $\gamma(\mathrm{PWM}, F) \leq \gamma^*(\mathrm{PWM}, F)$. For any fixed constant $M > 0$, by Lemma 4.1 of Zuo, Cui and He (2004), there is a function of ε, $\tau = \tau(\varepsilon)$ $(\geq \varepsilon > 0)$, such that $\tau \to 0$ and $\inf_{\|u-u(x)\|\geq\tau,\, 1/M\leq\|x\|\leq M} (O(x, F) - g(x, u, F))/\varepsilon \to \infty$ as $\varepsilon \to 0$. By the uniform boundedness of the influence functions of Med and MAD, assume that there is a constant $C_1 > 0$ such that for small $\varepsilon > 0$, $\sup_{y\in\mathbb{R}^d,\, \|u\|=1} \max\{|\mu(F_u(\varepsilon, \delta_y)) - \mu(F_u)|, |\sigma(F_u(\varepsilon, \delta_y)) - \sigma(F_u)|\} \leq C_1\varepsilon$. Denote $\tilde{\mu}(F_v, \varepsilon, y) = \mu(F_v(\varepsilon, \delta_y)) - \mu(F_v)$ and $\tilde{\sigma}(F_v, \varepsilon, y) = \sigma(F_v(\varepsilon, \delta_y)) - \sigma(F_v)$ for any unit vector v. Then there is a constant C_2 such that

$$
\begin{aligned}
&\sup_{y\in\mathbb{R}^d} \big|\mathrm{PD}(x, F(\varepsilon, \delta_y)) - \mathrm{PD}(x, F)\big| \\
(*)\qquad &= \big|O(x, F) - O\big(x, F(\varepsilon, \delta_y)\big)\big| \big/ \Big(\big(1 + O\big(x, F(\varepsilon, \delta_y)\big)\big)\big(1 + O(x, F)\big)\Big) \\
&\leq C_2 \sup_{y\in\mathbb{R}^d, \|u\|=1} \max\big\{|\tilde{\mu}(F_u, \varepsilon, y|, |\tilde{\sigma}(F_u, \varepsilon, y)|\big\} = O(\varepsilon).
\end{aligned}
$$

Define $S(x, M) = \{x : 1/M \leq \|\Sigma^{-1/2}x\| \leq M\}$. From the proof of Lemma A.1, it follows that, for $x \in S(x, M)$ and small ε,

$$
\begin{aligned}
&\Big|\big\{\big(O(x, F) - O\big(x, F(\varepsilon, \delta_y)\big)\big) - \big(g\big(x, u(x), F\big) - g\big(x, u(x), F(\varepsilon, \delta_y)\big)\big)\big\}\Big| \big/ \varepsilon \\
&\leq \sup_{\|u-u(x)\|\leq\tau} \Big|\big(g\big(x, u(x), F\big) - g\big(x, u(x), F(\varepsilon, \delta_y)\big)\big) \\
&\qquad\qquad - \big(g(x, u, F) - g\big(x, u, F(\varepsilon, \delta_y)\big)\big)\Big| \big/ \varepsilon \\
&= \sup_{\|u-u(x)\|\leq\tau} \Big|\Big(\big(u(x)'x - \mu\big(F_{u(x)}\big)\big)\tilde{\sigma}\big(F_{u(x)}, \varepsilon, y\big) \\
&\qquad\qquad + \sigma\big(F_{u(x)}\big)\tilde{\mu}\big(F_{u(x)}, \varepsilon, y\big)\Big) \big/ \big(\varepsilon\sigma\big(F_{u(x)}(\varepsilon, \delta_y)\big)\sigma\big(F_{u(x)}\big)\big) \\
&\qquad\qquad - \Big(\big(u'x - \mu(F_u)\big)\tilde{\sigma}(F_u, \varepsilon, y) \\
&\qquad\qquad\qquad + \sigma(F_u)\tilde{\mu}(F_u, \varepsilon, y)\Big) \big/ \big(\varepsilon\sigma\big(F_u(\varepsilon, \delta_y)\big)\sigma(F_u)\big)\Big| \\
&\leq C_3 \sup_{\|u-u(x)\|\leq\tau} 1/\varepsilon\big(|\tilde{\mu}\big(F_{u(x)}, \varepsilon, y\big) - \tilde{\mu}\big(F_u, \varepsilon, y\big)| \\
&\qquad\qquad + \big|\tilde{\sigma}\big(F_{u(x)}, \varepsilon, y\big) - \tilde{\sigma}(F_u, \varepsilon, y)\big|\big) + o(1),
\end{aligned}
$$

where $C_3 = C_3(M) > 0$ is a constant and $o(\cdot)$ is uniform in y. Call the first term on the right-hand side of the last inequality $I_4 = I_4(x, y, \varepsilon)$. By Lemma A.6, for $(\mu, \sigma) = (\text{Med}, \text{MAD})$, we have, for any bounded function $K(x) \geq 0$,

$$E\Big[K(X) \sup_{y\in\mathbb{R}^d,\, \|u-u(X)\|\leq\tau} 1/\varepsilon\Big(\big|\tilde{\mu}\big(F_{u(X)}, \varepsilon, y\big) - \tilde{\mu}(F_u, \varepsilon, y)\big| + \big|\tilde{\sigma}\big(F_{u(X)}, \varepsilon, y\big) - \tilde{\sigma}(F_u, \varepsilon, y)\big|\Big)\Big] \to 0$$

as $\varepsilon \to 0$, which, in conjunction with $(*)$ and Lebesgue's dominated convergence theorem, leads to

$$\begin{aligned}
&\frac{\mathrm{PWM}(F(\varepsilon, \delta_y))}{\varepsilon} \\
&\quad = \Big[\frac{1-\varepsilon}{\varepsilon}\int x w^{(1)}(\xi)\frac{O(x, F) - O(x, F(\varepsilon, \delta_y))}{(1 + O(x, F(\varepsilon, \delta_y)))(1 + O(x, F))}\, dF(x) + yw\big(\mathrm{PD}\big(y, F(\varepsilon, \delta_y)\big)\big)\Big] \\
&\qquad \times \Big[\int w\big(\mathrm{PD}\big(x, F(\varepsilon, \delta_y)\big)\big)\, dF(x; \varepsilon, \delta_y)\Big]^{-1} \\
&\quad = \Big[\int x w^{(1)}\big(\mathrm{PD}(x, F)\big)\frac{O(x, F) - O(x, F(\varepsilon, \delta_y))}{\varepsilon(1 + O(x, F))^2}\, dF(x) + yw\big(\mathrm{PD}(y, F)\big)\Big] \\
&\qquad \times \Big[\int w\big(\mathrm{PD}(x, F)\big)\, dF(x)\Big]^{-1} + o(1) \\
&\quad = \Big[\int_{S(x,M)} x w^{(1)}\big(\mathrm{PD}(x, F)\big)\frac{O(x, F) - O(x, F(\varepsilon, \delta_y))}{\varepsilon(1 + O(x, F))^2}\, dF(x) + yw\big(\mathrm{PD}(y, F)\big)\Big] \\
&\qquad \times \Big[\int w\big(\mathrm{PD}(x, F)\big)\, dF(x)\Big]^{-1} + I_5(M, y, \varepsilon) + o(1) \\
&\quad = \Big[\int_{S(x,M)} x w^{(1)}\big(\mathrm{PD}(x, F)\big)\frac{g(x, u(x), F) - g(x, u(x), F(\varepsilon, \delta_y))}{\varepsilon(1 + O(x, F))^2}\, dF(x) + yw\big(\mathrm{PD}(y, F)\big)\Big] \\
&\qquad \times \Big[\int w\big(\mathrm{PD}(x, F)\big)\, dF(x)\Big]^{-1} + I_5(M, y, \varepsilon) + o(1),
\end{aligned}$$

where $\sup_{y\in\mathbb{R}^d,\, \varepsilon<0.5} \|I_5(M, y, \varepsilon)\| \to 0$ as $M \to \infty$ and $o(\cdot)$ is uniform in $y \in \mathbb{R}^d$. Note that $u(x) = \Sigma^{-1}x/\|\Sigma^{-1}x\|$ for $x \neq 0$, $\mu(F_{u(x)}) = 0$ and

$\sigma(F_{u(x)}) = \|\Sigma^{-1/2}x\|/\|\Sigma^{-1}x\|$. Hence

$$\Big(\big(u(x)'x - \mu(F_{u(x)})\big)\big(\sigma(F_{u(x)}(\varepsilon,\delta_y)) - \sigma(F_{u(x)})\big)\Big)/\big(\varepsilon\sigma(F_{u(x)}(\varepsilon,\delta_y))\sigma(F_{u(x)})\big)$$

is an even function of x. Thus

$$\begin{aligned}
&\frac{\mathrm{PWM}(F(\varepsilon,\delta_y))}{\varepsilon}\\
&\quad= \bigg[\int_{S(x,M)} x w^{(1)}\big(\mathrm{PD}(x,F)\big)\frac{\mu(F_{u(x)}(\varepsilon,\delta_y))}{\varepsilon\sigma(F_{u(x)})(1+O(x,F))^2}\,dF(x)\\
&\qquad\qquad + y w\big(\mathrm{PD}(y,F)\big)\bigg]\\
&\qquad \times \bigg[\int w\big(\mathrm{PD}(x,F)\big)\,dF(x)\bigg]^{-1} + I_5(M,y,\varepsilon) + o(1)\\
&\quad= \bigg[\int_{S(x,M)} \frac{x w^{(1)}(\mathrm{PD}(x,F))}{\varepsilon(1+O(x,F))^2}\mathrm{Med}\bigg\{-d_1, \frac{x'\Sigma^{-1}y}{\|\Sigma^{-1/2}x\|}, d_1\bigg\}\,dF(x)\\
&\qquad\qquad + y w\big(\mathrm{PD}(y,F)\big)\bigg]\\
&\qquad \times \bigg[\int w\big(\mathrm{PD}(x,F)\big)\,dF(x)\bigg]^{-1} + I_5(M,y,\varepsilon) + o(1)
\end{aligned}$$

by Lemma A.3. Call the first term on the right-hand side of the last equality $I_6 = I_6(M,y,\varepsilon)$. Then

$$\mathrm{PWM}\big(F(\varepsilon,\delta_y)\big)/\varepsilon = I_6(M,y,\varepsilon) + I_5(M,y,\varepsilon) + o(1).$$

By changing variables ($\Sigma^{-1/2}x \to x$) and then taking an orthogonal transformation (using an orthogonal matrix U with $\Sigma^{-1/2}y/\|\Sigma^{-1/2}y\|$ as its first column), we have

$$\begin{aligned}
I_6 &= \bigg[\Sigma^{1/2}\int_{1/M\le\|x\|\le M} \frac{x w^{(1)}(1/(1+\|x\|))}{\varepsilon(1+\|x\|)^2}\\
&\qquad\quad \times \mathrm{Med}\bigg\{-d_1(\varepsilon), \frac{x'\Sigma^{-1/2}y}{\|x\|}, d_1(\varepsilon)\bigg\}\,dF_0(x) + y w\bigg(\frac{1}{1+\|\Sigma^{-1/2}y\|}\bigg)\bigg]\\
&\quad \times \bigg[\int w\big(\mathrm{PD}(x,F)\big)\,dF(x)\bigg]^{-1}\\
&= \bigg[\int_{1/M\le\|x\|\le M} \frac{y x_1 w^{(1)}(1/(1+\|x\|))}{\|\Sigma^{-1/2}y\|\varepsilon(1+\|x\|)^2}\\
&\qquad \times \mathrm{Med}\bigg\{-d_1(\varepsilon), \frac{x_1\|\Sigma^{-1/2}y\|}{\|x\|}, d_1(\varepsilon)\bigg\}\,dF_0(x) + y w\bigg(\frac{1}{1+\|\Sigma^{-1/2}y\|}\bigg)\bigg]\\
&\quad \times \bigg[\int w\big(\mathrm{PD}(x,F)\big)\,dF(x)\bigg]^{-1}.
\end{aligned}$$

Since

$$0 \le x_1 \mathrm{Med}\{-d_1(\varepsilon),\, x_1\|\Sigma^{-1/2}y\|/\|x\|,\, d_1(\varepsilon)\} \le |x_1| d_1(\varepsilon),$$

we have

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^d} \|I_6(M,y,\varepsilon)\| \\
&\quad \le \Bigg(\sup_{\|u\|=1,\, r\ge 0}\Bigg[\|\Sigma^{1/2}u\| \\
&\qquad\qquad \times \Bigg|\int_{1/M\le\|x\|\le M} \frac{|x_1| w^{(1)}(1/(1+\|x\|))}{(1+\|x\|)^2}\frac{d_1(\varepsilon)}{\varepsilon}\, dF_0(x) \\
&\qquad\qquad\qquad + r w\Big(\frac{1}{1+r}\Big)\Bigg|\Bigg]\Bigg) \\
&\qquad \times \Big[\int w\big(\mathrm{PD}(x,F)\big)\, dF(x)\Big]^{-1}.
\end{aligned}$$

By Lebesgue's dominated convergence theorem, it follows that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\sup_{y\in\mathbb{R}^d} \|I_6(M,y,\varepsilon)\| \\
&\quad \le \Bigg[\sqrt{\lambda_1}\sup_{r\ge 0}\Bigg|\int_{1/M\le\|x\|\le M}\frac{|x_1| w^{(1)}(1/(1+\|x\|))}{2h(0)(1+\|x\|)^2}\, dF_0(x) + r w\Big(\frac{1}{1+r}\Big)\Bigg|\Bigg] \\
&\qquad \times \Big[\int w\big(1/(1+\|x\|)\big)\, dF_0(x)\Big]^{-1}.
\end{aligned}$$

Since

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^d} \big\|\mathrm{PWM}\big(F(\varepsilon,\delta_y)\big)\big\|/\varepsilon \\
&\qquad \le \sup_{y\in\mathbb{R}^d}\|I_6(M,y,\varepsilon)\| + \sup_{y\in\mathbb{R}^d}\|I_5(M,y,\varepsilon)\| + o(1),
\end{aligned}$$

letting $\varepsilon \to 0$ and then letting $M \to \infty$, we get

$$\begin{aligned}
\gamma(\mathrm{PWM},F) &= \lim_{\varepsilon\to 0}\sup_{y\in\mathbb{R}^d}\frac{\|\mathrm{PWM}(F(\varepsilon,\delta_y))\|}{\varepsilon} \\
&\le \frac{\sqrt{\lambda_1}\sup_{r\ge 0}|k_0 + r w(1/(1+r))|}{\int w(1/(1+\|x\|))\, dF_0(x)} \\
&= \gamma^*(\mathrm{PWM},F).
\end{aligned}$$

The desired result now follows. □

Proof of Theorems 4.3 and 4.4.

LEMMA A.7. *Let $X \sim F$ be d-version symmetric about the origin. Let $\varepsilon > 0$, $x \in \mathbb{R}^d$ and $s_x = \sup_{\|u\|=1} u'x_u$. Then for $0 \le j \le 3$, $O(t_j x, F(\varepsilon, \delta_x)) = \min_{-\infty<t<\infty} O(tx, F(\varepsilon, \delta_x)) = d_0(1 - t_j), x \in S_j$, where $S_0 = \{x : 0 \le s_x < d_1\}$, $S_1 = \{x : d_1 \le s_x < d_1 + d_3\}$, $S_2 = \{x : d_1 + d_3 \le s_x < d_1 + d_2\}$, $S_3 = \{x : s_x \ge d_1 + d_2\}$ and $t_0 = 1$, $t_1 = (d_1 + d_0 d_3)/(s_x + d_0 d_3)$, $t_2 = (d_1 + d_0(s_x - d_1))/(s_x + d_0(s_x - d_1))$, $t_3 = (d_1 + d_0 d_2)/(s_x + d_0 d_2)$.*

LEMMA A.8. *Let $X \sim F$ be d-version symmetric about the origin with $a(u)/a(u^0) \ge u'u^0$. Then $\mathrm{PM}(F(\varepsilon, \delta_x)) = t_3 x$ for $x = \kappa u^0$, where $\kappa \ge a(u^0) \times \max\{d_1 + d_2, d_1 a(u^0)/a(u_0)\}$ and $u_0 = \arg\min_{\|u\|=1} a(u)$.*

By Lemmas A.7 and A.8, we have the following lemmas.

LEMMA A.9. *Let $X \sim F$ be d-version symmetric about 0 and let the density h of Z be continuous and positive in a small neighborhood of 1. Then for small $\varepsilon > 0$, $d_0 = \sup_{0 \le c \le d_1}(c(m_1(c, \varepsilon))) = d_1/d_3$.*

LEMMA A.10. *Let $X \sim F$ be elliptically symmetric about the origin with Σ associated. Then $\mathrm{PM}(F(\varepsilon, \delta_x)) = t_i x\, I(s_x \in S_i)$, $0 \le i \le 3$, for S_i in Lemma A.8, where $t_0 = 1$ and $s_x = \sqrt{x'\Sigma^{-1}x}$.*

PROOF OF THEOREM 4.3. (i) First we show that $B(\varepsilon; \mathrm{PM}, F) \le a(u^0)(d_1 + d_0 d_2)$. By Lemmas A.2 and A.3, it is readily seen that $|\mathrm{Med}(F_u(\varepsilon, G))| \le a(u)d_1$ for any $\|u\| = 1$. The nonincreasingness of $h(x)$ in $|x|$ implies that $m_i(x, \varepsilon)$, $i = 1, 2$, are nondecreasing in $\|x\|$. Hence $m_1(\mathrm{Med}(F_u(\varepsilon, G)), \varepsilon) \le \mathrm{MAD}(F_u(\varepsilon, G)) \le m_2(\mathrm{Med}(F_u(\varepsilon, G)), \varepsilon) \le a(u)d_2$ by Lemmas A.2 and A.3. By (4.1) in Theorem 4.1 and Lemmas A.2 and A.3, it is seen that

$$\begin{aligned}
B(\varepsilon; \mathrm{PM}, F) &\le \sup_{G, \|u\|=1} \mu\big(F_u(\varepsilon, G)\big) \\
&\quad + \sup_{G, \|u\|=1} \sigma\big(F_u(\varepsilon, G)\big) \sup_{G, \|u\|=1} \mu\big(F_u(\varepsilon, G)\big)/\sigma\big(F_u(\varepsilon, G)\big) \\
&\le a(u^0)\Big(d_1 + d_2 \sup_{G, \|u\|=1} \mu\big(F_u(\varepsilon, G)\big)/\sigma\big(F_u(\varepsilon, G)\big)\Big) \\
&\le a(u^0)\Big(d_1 + d_2 \sup_{G, \|u\|=1} \mu\big(F_u(\varepsilon, G)\big)/m_1\big(\mathrm{Med}\big(F_u(\varepsilon, G)\big), \varepsilon\big)\Big) \\
&\le a(u^0)(d_1 + d_2 d_0).
\end{aligned}$$

Now we show that $B(\varepsilon; \mathrm{PM}, F) \geq a(u^0)(d_1 + d_0 d_2)$. By Lemma A.8, we have

$$B(\varepsilon; \mathrm{PM}, F) \geq \sup_{\substack{x=\kappa u^0 \\ \kappa > k_2}} t_3 \|x\| = \sup_{\kappa > k_2} (d_1 + d_0 d_2)\kappa \big/ \big(\kappa / a(u^0) + d_0 d_2\big)$$

$$= a(u^0)(d_1 + d_0 d_2),$$

where $k_2 = a(u^0) \max\{d_1 + d_2, d_1 a(u^0)/a(u_0)\}$.

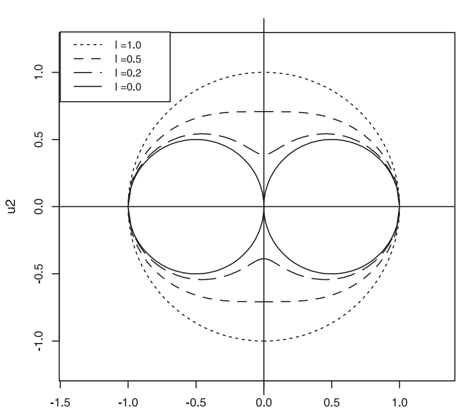
(ii) By Lemma A.9, for small $\varepsilon > 0$, $d_0 = d_1/d_3$. On the other hand, $d_1(\varepsilon)/\varepsilon \to 1/(2h(0))$ and $d_2/d_3 \to 1$ as $\varepsilon \to 0$. Thus we have $\gamma(\mathrm{PM}, F) = \lim_{\varepsilon \to 0} B(\varepsilon; \mathrm{PM}, F)/\varepsilon = a(u^0)/h(0)$. □

PROOF OF THEOREM 4.4. (i) We need only consider the nontrivial case $x \neq 0$. The result follows directly from Lemmas A.7 and A.10, and the facts that $s_x \geq d_1$ and $d_0 = d_1/d_3$ for small $\varepsilon > 0$ and $d_1 \to 0$, $d_2/d_3 \to 1$ and $d_1/\varepsilon \to 1/(2h(0))$ as $\varepsilon \to 0$.

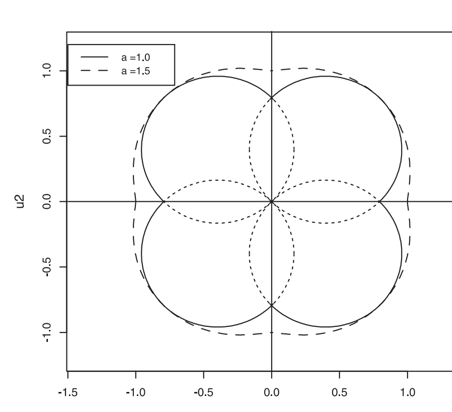
(ii) This follows in a straightforward fashion from part (i). □

APPENDIX B

Behavior of $a(u)$ in Theorem 3.3 corresponding to α-symmetric F. In Figure 3(a), an elliptically symmetric F with $\Sigma = \mathrm{diag}(1, l)$ and $0 \leq l \leq 1$ is considered. Here $\alpha = 2$, $u_0 = (1, 0)'$ and $a(u_0) = 1$. The innermost solid curve (two circles) corresponds to $l = 0$, representing the extreme case $a(u)/a(u_0) = u'u_0$. For other l's, the curve of $a(u)$ is inscribed by the innermost solid curve. Hence $a(u)/a(u_0) \geq u'u_0$ holds true for any $\|u\| = 1$. In Figure 3(b), the innermost



(a)



(b)

FIG. 3. (a) *The behavior of* $a(u) = \sqrt{u_1^2 + l u_2^2}$ *corresponding to elliptical symmetry with different* l*'s.* (b) *The behavior of* $a(u) = (|u_1|^\alpha + |u_2|^\alpha)^{1/\alpha}$ *corresponding to* α*-symmetry with different* α*'s*, $\alpha = 1.0, 1.5$.

solid curve (formed by four half circles), corresponding to the extreme case $a(u)/a(u_0) = u'u_0$, represents the shape of the $a(u)$ with $\alpha = 1$, where $u_0 = (1/\sqrt{2}, 1\sqrt{2})'$. For other α's, the curve of $a(u)$ is inscribed by the innermost solid curve [the outermost dashed curve in Figure 3(b) corresponds to $a(u)$ with $\alpha = 1.5$]. Again $a(u)/a(u_0) \geq u'u_0$ holds true for any $\|u\| = 1$.

Acknowledgments. We greatly appreciate the insightful and constructive comments of an Associate Editor and two referees, which led to distinct improvements in the article. We thank Ricardo Maronna for helpful suggestions.

REFERENCES

ADROVER, J. and YOHAI, V. J. (2002). Projection estimates of multivariate location. *Ann. Statist.* **30** 1760–1781.

CHEN, Z. and TYLER, D. E. (2002). The influence function and maximum bias of Tukey's median. *Ann. Statist.* **30** 1737–1759.

CUI, H. and TIAN, Y. (1994). Estimation of the projection absolute median deviation and its application. *J. Systems Sci. Math. Sci.* **14** 63–72 (in Chinese).

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

EATON, M. L. (1981). On the projections of isotropic distributions. *Ann. Statist.* **9** 391–400.

FANG, K. T., KOTZ, S. and NG, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

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

HE, X. and SIMPSON, D. G. (1993). Lower bounds for contamination bias: Globally minimax versus locally linear estimation. *Ann. Statist.* **21** 314–337.

HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

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

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

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

MARONNA, R. A. and YOHAI, V. J. (1995). The behavior of the Stahel–Donoho robust multivariate estimator. *J. Amer. Statist. Assoc.* **90** 330–341.

MARTIN, D. R., YOHAI, V. J. and ZAMAR, R. H. (1989). Min–max bias robust regression. *Ann. Statist.* **17** 1608–1630.

ROMANAZZI, M. (2001). Influence function of halfspace depth. *J. Multivariate Anal.* **77** 138–161.

ROUSSEEUW, P. J. and RUTS, I. (1998). Constructing the bivariate Tukey median. *Statist. Sinica* **8** 828–839.

TUKEY, J. W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians* **2** 523–531. Canadian Math. Congress, Montreal.

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

ZUO, Y. (2003). Projection-based depth functions and associated medians. *Ann. Statist.* **31** 1460–1490.

ZUO, Y., CUI, H. and HE, X. (2004). On the Stahel–Donoho estimator and depth-weighted means for multivariate data. *Ann. Statist.* **32** 169–190.

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

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

Y. ZUO
DEPARTMENT OF STATISTICS
AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MICHIGAN 48824
USA
E-MAIL: zuo@msu.edu

H. CUI
DEPARTMENT OF MATHEMATICS
BEIJING NORMAL UNIVERSITY
BEIJING 100875
CHINA
E-MAIL: hjcui@bnu.edu.cn

D. YOUNG
DEPARTMENT OF MATHEMATICS
AND STATISTICS
ARIZONA STATE UNIVERSITY
TEMPE, ARIZONA 85287-1804
USA
E-MAIL: dennis.young@asu.edu