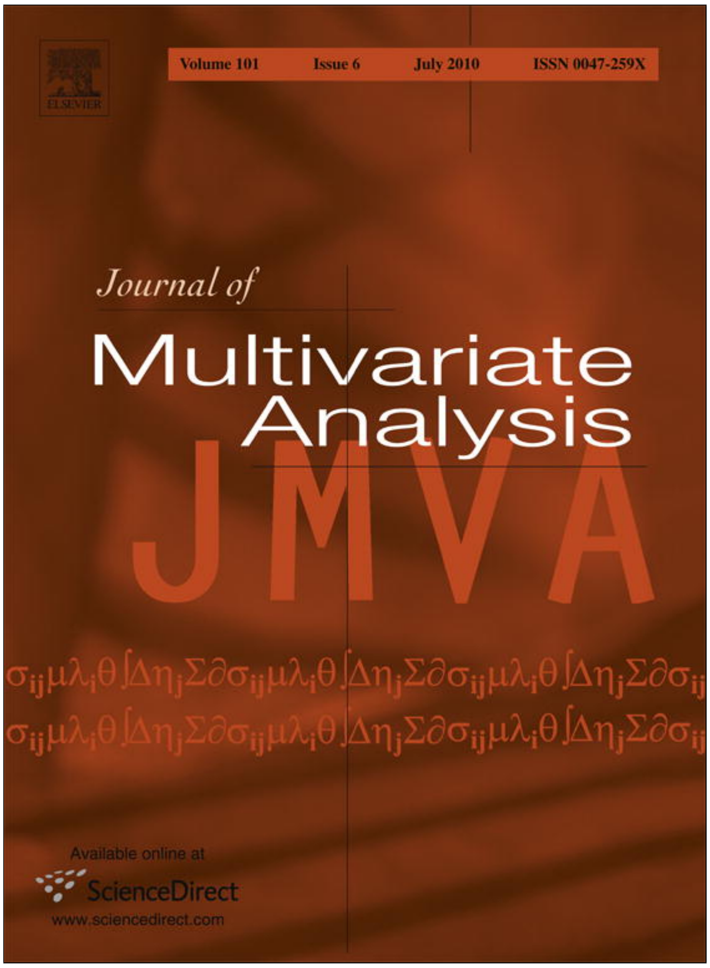
Volume 101
Issue 6
July 2010
ISSN 0047-259X
ELSEVIER
Journal of
Multivariate
Analysis
JMVA
Available online at
ScienceDirect
www.sciencedirect.com

Journal of Multivariate Analysis 101 (2010) 1400–1411

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

A new projection estimate for multivariate location with minimax bias☆

Jorge G. Adrover [a,b,c,∗], Víctor J. Yohai [d,c]

[a] *Universidad Nacional de Córdoba, Argentina*
[b] *CIEM, Argentina*
[c] *CONICET, Argentina*
[d] *Universidad de Buenos Aires, Argentina*

ARTICLE INFO

Article history:
Received 12 April 2009
Available online 11 December 2009

AMS 2010 subject classifications:
62F35
62H12

Keywords:
Projection estimates
Maximum bias
Robust estimates
Multivariate location

ABSTRACT

The maximum asymptotic bias of an estimator is a global robustness measure of its performance. The projection median estimator for multivariate location shows a remarkable behavior regarding asymptotic bias. In this paper we consider a modification of the projection median estimator which renders an estimate with better bias performance for point mass contaminations (the worst situation for the projection median estimator). Moreover, it achieves the lowest bound for an equivariant estimate for point mass contaminations.

1. Introduction

The seminal paper by Huber [1] highlights the median as the most bias robust estimator in the location model since it minimizes the asymptotic bias among the class of translation equivariant estimators. Several proposals tried to extend the median for multidimensional data. A remarkable attempt in this direction is the paper by Tukey [2], who introduces the concept of depth in the multivariate data cloud and the deepest point according to that definition is known as Tukey's median. Stahel [3] and Donoho [4] introduce projection depth based weighted means which are extremely competitive regarding bias (see Zuo et al. [5] for a detailed account). Zuo and Serfling [6] also give a deep insight into the concept of depth.

A good measure of the robustness of an estimate is the maximum bias, which is the maximum asymptotic bias of the estimate caused by a given fraction of contamination. Other measures used to summarize the robustness performance of an estimate, such as the breakdown point (see Hampel [7]) and the gross error sensitivity (see Hampel [8]), can be derived from the maximum bias. Riedel [9] and He and Simpson [10] find lower bounds for the maximum bias of equivariant estimates. Adrover and Yohai [11] derive explicitly this lower bound for the case of elliptical distribution and show that the projection median proposed in Tyler [12] has a maximum bias which is approximately twice this lower bound for small levels of contaminations. Zuo et al. [5] derive the maximum bias of the projection median under weaker assumptions on the underlying model and the maximum bias of projection based weighted means in the case of point mass contaminations.

In this paper we consider a modification in the definition of the projection median estimate which has minimax bias, i.e., its maximum bias (maxbias) function attains the mentioned lower bound for point mass contaminations for an important

☆ This research was partially supported by Grants X-018 from Universidad de Buenos Aires, PID 5505 from CONICET and PICTs 21407 and 00899 from ANPCYT, Argentina.

∗ Corresponding address: FaMAF, Universidad Nacional de Córdoba, Avda. Medina Allende s/n, Ciudad Universitaria, 5000 Córdoba, Argentina.
E-mail addresses: adrover@famaf.unc.edu.ar (J.G. Adrover), vyohai@uolsinectis.com.ar (V.J. Yohai).

0047-259X/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2009.12.007

range of level of contaminations. The basic idea common to the different projection based estimates (P-estimates) is to transform, by means of projections, a multivariate problem into the corresponding univariate problem, which is dealt with using univariate estimates. We define the functional form of the P-estimate $\mathbf{T}_P$ as follows. Let T and S be location and dispersion univariate estimating functionals, i.e., if $y^* = ay + b$ and $a, b \in R$ then

$$T(\mathcal{D}(y^*)) = aT(\mathcal{D}(y)) + b, \qquad S(\mathcal{D}(y^*)) = |a|S(\mathcal{D}(y)), \tag{1}$$

where $\mathcal{D}(x)$ denotes the distribution of x. The P-estimate approach relies on the idea that $\boldsymbol{\zeta} \in R^p$ is a good center of the data, if for any direction $\mathbf{a} \in R^p$, the univariate projected set $\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta})$ is well centered around 0. Then, a standardized measure of how wrongly centered is $\mathbf{X} - \boldsymbol{\zeta}$, is given by

$$v(\boldsymbol{\zeta}, F) = \sup_{\mathbf{a} \neq 0} |h(\boldsymbol{\zeta}, \mathbf{a}, F)|, \tag{2}$$

where

$$h(\boldsymbol{\zeta}, \mathbf{a}, F) = \frac{T(\mathcal{D}(\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta})))}{S(\mathcal{D}(\mathbf{a}'\mathbf{X}))}. \tag{3}$$

An ideal center for a distribution F would be a value $\boldsymbol{\zeta}$ such that $v(\boldsymbol{\zeta}, F) = \mathbf{0}$, i.e., such that all the projected vectors are perfectly centered for any direction $\mathbf{a}$. But in general, for an arbitrary distribution, such a vector $\boldsymbol{\zeta}$ does not exist. Then the functional version of the P-estimate of multivariate location defined by Tyler [12] is given by

$$\mathbf{T}_P(\mathbf{F}) = \arg\min_{\boldsymbol{\zeta} \in R^p} v(\boldsymbol{\zeta}, F). \tag{4}$$

In the rest of the paper we will take as T the median (med) and as S the median absolute deviation around the median (MAD). The corresponding projection estimate $\mathbf{T}_P$ will be called $\mathbf{T}_{MP}$. It may occur that for some $\mathbf{a} \neq \mathbf{0}$ we have $S(\mathcal{D}(\mathbf{a}'\mathbf{X})) = 0$. This may happen, at least in the case that S is the MAD, if $\mathbf{X}$ lies with probability at least 0.5 in a hyperplane of the form $\mathbf{a}'\mathbf{X} = b$, where $\mathbf{a} = (a_1, \ldots, a_p)'$. This implies that any affine equivariant location estimate $\mathbf{T} = (T_1, \ldots, T_p)'$ with breakdown point 0.5 should lie in the same hyperplane too, see Section 6.2.2 of Maronna et al. [13]. Therefore. in this case we can delete any component X_i of $\mathbf{X}$ such that $a_i \neq 0$, and estimate the location of the corresponding $(p-1)$-dimensional observations. If in the $(p-1)$-dimensional space all the linear combinations have a scale different from 0, we can define T_j for $j \neq i$, by (4), and then set $T_i = -\sum_{j \neq i} a_j T_j / a_i$. In case that in the $(p-1)$-dimensional space there are still linear combinations with scale equal to zero we eliminate another variable. We continue reducing the dimension of the problem in this way until all linear combinations have scale different from zero.

Tyler [12] shows that the P-estimates of multivariate location have a finite sample breakdown point close to 0.5, as long as the corresponding univariate estimates of location and dispersion also have this property. It is also shown that they are affine equivariant. The $\sqrt{n}$-rate of convergence and the non-normal asymptotic distribution of the P-estimates are analyzed by Kim and Hwang [14] and Zuo [15].

In other words, we can say that the $\mathbf{T}_{MP}$ estimate is the point in R^p such that when the data are centered around this point, we minimize the maximum absolute value of the standardized median when the centered data are projected along all directions. The new estimator is defined using a similar idea, but, instead of minimizing the maximum absolute value of the standardized median we propose to minimize the maximum difference between the standardized medians corresponding to projecting the centered data along two arbitrary directions. An interesting property of the new proposal is that it will capture a desirable property of the centered data: The median should not change too much when the centered data are projected in different directions.

For the definition of the modified projection estimate we need the following concepts. Given any vector $\mathbf{a} = (a_1, \ldots, a_p)' \in R^p - \{\mathbf{0}\}$, the corresponding half-space $L(\mathbf{a})$ through the origin is defined by

$$L(\mathbf{a}) = \{\mathbf{x} \in R^p : \mathbf{a}'\mathbf{x} \geq 0\}. \tag{5}$$

If it is clear enough from the context we will simply write L instead of $L(\mathbf{a})$ and the set of all half-spaces is denoted by $\mathcal{L}$. Next, we define a new measure to assess the outlyingness of a point $\boldsymbol{\zeta}$. Given $\boldsymbol{\zeta} \in R^p$, $L \in \mathcal{L}$ and a distribution F in R^p, define

$$V(\boldsymbol{\zeta}, L, F) = \sup_{\mathbf{a} \in S^{p-1} \cap L} h(\boldsymbol{\zeta}, \mathbf{a}, F) - \inf_{\mathbf{a} \in S^{p-1} \cap L} h(\boldsymbol{\zeta}, \mathbf{a}, F), \tag{6}$$

where $S^{p-1} = \{\mathbf{a} \in R^p : \|\mathbf{a}\| = 1\}$. Then the modified projection estimate is defined by

$$\mathbf{T}_{MP}^{M}(F) = \arg\min_{\boldsymbol{\zeta} \in R^p} \inf_{L \in \mathcal{L}} V(\boldsymbol{\zeta}, L, F). \tag{7}$$

$V(\boldsymbol{\zeta}, L, F)$ measures the maximum difference between the standardized medians of two projections of the data centered around $\boldsymbol{\zeta}$, when both directions are in the half-space L. Note that $h(\boldsymbol{\zeta}, \mathbf{a}, F) = -h(\boldsymbol{\zeta}, -\mathbf{a}, F)$ and therefore, if instead of taking the two directions in the same half-space we would consider all the differences between two arbitrary directions in S^{p-1}, we would obtain the same outlyingness measure $v(\boldsymbol{\zeta}, F)$ given in (2) which was used to define the projection estimate $\mathbf{T}_{MP}$.

In Section 2 we state the main results concerning the optimal bias behavior of the $\mathbf{T}_{MP}^{M}$ estimate. Section 3 gives an algorithm to compute an approximate version of the estimate. Section 4 contains simulation studies comparing the efficiency performance for the new proposal and some competitors including the P-estimator. Section 5 is an Appendix with the proofs. For shortness sake, some of the proofs are omitted here and can be found in a Technical Report by Adrover and Yohai [16] available on the Web.

2. Bias performance of the modified P-estimate

In the multivariate location model we observe a p-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_p)'$ with distribution $F_{\boldsymbol{\mu}}(\mathbf{x}) = F_{\mathbf{0}}(\mathbf{x} - \boldsymbol{\mu})$, where $F_{\mathbf{0}}$ is symmetric around $\mathbf{0}$, i.e., if $\mathbf{X}$ has distribution $F_{\mathbf{0}}$, then $-\mathbf{X}$ also has distribution $F_{\mathbf{0}}$. An important case is the family of elliptical distributions. We say that $\mathbf{X}$ has an elliptical distribution if it has a density of the form

$$f(\mathbf{x}, \boldsymbol{\mu}, \Sigma) = \frac{1}{(\det \Sigma)^{1/2}} f_0((\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})), \tag{8}$$

where $f_0 : R^+ \to R^+$, and Σ is a $p \times p$ positive definite matrix. If $\mathbf{X}$ has a density $f(\mathbf{x}, \mathbf{0}, I)$, then $\mathbf{a}'\mathbf{X}$ has the same distribution for all $\mathbf{a} \in S^{p-1} = \{\mathbf{a} \in R^p : \|\mathbf{a}\| = 1\}$. This common distribution will be denoted by H_0 and its density by h_0.

All multivariate location estimating functionals $\mathbf{T}$ considered in this paper are affine equivariant, i.e., given a $p \times p$ nonsingular matrix A and $\mathbf{b} \in R^p$,

$$\mathbf{T}(\mathcal{D}(A\mathbf{X} + \mathbf{b})) = A\mathbf{T}(\mathcal{D}(\mathbf{x})) + \mathbf{b}. \tag{9}$$

It is immediate to show that the modified P-estimate introduced in Section 1 is affine equivariant.

An estimating functional $\mathbf{T}$ is Fisher consistent if $\mathbf{T}(F_{\boldsymbol{\mu}}) = \boldsymbol{\mu}$. In the next Theorem we state the Fisher consistency of the estimate $\mathbf{T}_{MP}^{M}$ defined by (7) for elliptical families,

Theorem 1. *Let $\mathbf{X}$ be a random vector with elliptical density given by (8). Then, $\mathbf{T}_{MP}^{M}$ is Fisher consistent estimating functional of $\boldsymbol{\mu}$.*

To study the robustness property of the multivariate location estimate we will consider contamination neighborhoods of the target distribution. Given a fraction of contamination $\varepsilon > 0$, the corresponding contamination neighborhood of $F_{\boldsymbol{\mu}}$ is defined by

$$\mathcal{V}_{\varepsilon}(F_{\boldsymbol{\mu}}) = \{F = (1 - \varepsilon)F_{\boldsymbol{\mu}} + \varepsilon F^* : F^* \text{ any distribution on } R^p\}.$$

All estimates studied here are defined by means of a functional on a subset $\mathcal{F}$ of the space of all the distributions on R^p. We will assume that $\mathcal{F}$ contains the empirical distributions, all distributions belonging to $\mathcal{V}_{\varepsilon}(F_{\boldsymbol{\mu}})$, and that it is closed under affine transformations. If $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is a random sample from some distribution F and $\mathbf{T}$ is a continuous functional in the sense of weak convergence, then $\mathbf{T}(F)$ is the a.s. limit value of $\mathbf{T}_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Then it is natural to require that an estimating functional $\mathbf{T}$ have the Fisher consistency property: $\mathbf{T}(F_{\boldsymbol{\mu}}) = \boldsymbol{\mu}$. In general, given $F \in \mathcal{V}_{\varepsilon}(F_{\boldsymbol{\mu}})$ we will have $\mathbf{T}(F) \neq \boldsymbol{\mu}$. Then, we define the asymptotic bias of $\mathbf{T}$ in F by

$$b(\mathbf{T}, F, \boldsymbol{\mu}) = ((\mathbf{T}(F) - \boldsymbol{\mu})' \Lambda(F_{\boldsymbol{\mu}})^{-1} (\mathbf{T}(F) - \boldsymbol{\mu}))^{1/2}, \tag{10}$$

where Λ is an affine equivariant scatter functional. The maximum asymptotic bias of an estimating functional $\mathbf{T}$ for a fraction of contamination ε is defined by

$$B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) = \sup_{F \in \mathcal{V}_{\varepsilon}(F_{\boldsymbol{\mu}})} b(\mathbf{T}, F, \boldsymbol{\mu}). \tag{11}$$

The inclusion of the scatter matrix $\Lambda(F_{\mathbf{0}})$ in (10) yields a definition of maximum asymptotic bias which is invariant by affine transformations when applied to an equivariant functional. Therefore, if the functional $\mathbf{T}$ is affine equivariant, the maximum bias does not depend on $\boldsymbol{\mu}$, i.e., $B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) = B(\mathbf{T}, \varepsilon, F_{\mathbf{0}})$. In the elliptical case, we will assume that the scatter matrix Λ used in (10) is Fisher consistent for the shape of Σ, i.e., $\Lambda(F_{\boldsymbol{\mu}}) = \lambda \Sigma$, where λ is an scalar. In this case, if $\mathbf{T}$ is affine equivariant then the maximum bias is also independent of Σ.

He and Simpson [10] introduced the *contamination sensitivity* of an estimate $\mathbf{T}$ as

$$\gamma(\mathbf{T}, F_{\boldsymbol{\mu}}) = \left. \frac{\partial B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}})}{\partial \varepsilon} \right|_{\varepsilon=0}.$$

Observe that $\gamma(\mathbf{T}, F_{\boldsymbol{\mu}}) = \gamma(\mathbf{T}, F_{\mathbf{0}})$ because of the invariance of the bias. For small ε, the maximum bias can be approximated by

$$B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) \approx \varepsilon \gamma(\mathbf{T}, F_{\boldsymbol{\mu}}). \tag{12}$$

The contamination sensitivity $\gamma(\mathbf{T}, F_{\boldsymbol{\mu}})$ is closely related to the *gross error sensitivity* $\gamma^*(\mathbf{T}, F_{\boldsymbol{\mu}})$ defined in Hampel [7]. In fact, it is easy to show that always

$$\gamma(\mathbf{T}, F_{\boldsymbol{\mu}}) \geq \gamma^*(\mathbf{T}, F_{\boldsymbol{\mu}}),$$

where

$$\gamma^*(\mathbf{T}, F_{\boldsymbol{\mu}}) = \sup_{\mathbf{c} \in R^p} \left\| \lim_{\varepsilon \to 0} \frac{\mathbf{T}((1 - \varepsilon)F_{\boldsymbol{\mu}} + \varepsilon \delta_{\mathbf{c}}) - \mathbf{T}(F_{\boldsymbol{\mu}})}{\varepsilon} \right\|,$$

and $\delta_{\mathbf{c}}$ stands for a point mass contamination at $\mathbf{c}$. Under very general regularity conditions $\gamma^*(\mathbf{T}, F_{\boldsymbol{\mu}}) = \gamma(\mathbf{T}, F_{\boldsymbol{\mu}})$. Another relevant concept associated with the maximum bias is the *asymptotic breakdown point* which measures the least level of contamination for which the bias is unbounded and then noninformative. More precisely,

$$\epsilon^*(\mathbf{T}, F_{\boldsymbol{\mu}}) = \arg\inf\{\varepsilon > 0 : B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) = \infty\}.$$

In many situations $B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}})$ is extremely difficult to be calculated while $\epsilon^*(\mathbf{T}, F_{\boldsymbol{\mu}})$ is much easier to handle since the definition does not rely on the actual form of $B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}})$.

Huber [1] proved that if L_0 is a univariate symmetric distribution with unimodal density l_0 and $L_\mu(x) = L_0(x - \mu)$, the median estimating functional T_M is minimax among the translation equivariant estimates, i.e., if T is another translation equivariant estimating functional, then

$$B(T, \varepsilon, L_\mu) \ \geq B(T_M, \varepsilon, L_\mu) = L_0^{-1}\left(\frac{1}{2(1-\varepsilon)}\right) = d_1(\varepsilon, L_0). \tag{13}$$

He and Simpson [10] obtained a lower bound for the maximum bias of equivariant estimates. Using this result Adrover and Yohai [11] proved that $d_1(\varepsilon, H_0)$ is a lower bound for any equivariant multivariate location estimator when the central model is elliptical, with H_0 the univariate marginal distribution for $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = I$. More precisely, if $\mathbf{X}$ has a distribution with density given by (8), where f_0 is nonincreasing, then, for any affine equivariant estimate $\mathbf{T}$ of multivariate location we have

$$B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) \geq d_1(\varepsilon, H_0) \tag{14}$$

and

$$\gamma^*(\mathbf{T}, F_{\boldsymbol{\mu}}) \geq \frac{1}{2h_0(0)}. \tag{15}$$

A restricted neighborhood of $\mathcal{V}_\varepsilon(F)$ of special importance is defined to be

$$\mathcal{V}_\varepsilon^R(F_{\boldsymbol{\mu}}) = \{(1-\varepsilon)F_{\boldsymbol{\mu}} + \varepsilon\delta_{\mathbf{c}}, \mathbf{c} \in R^p\}. \tag{16}$$

Analogously, we can have a maximum bias function restricted to this set,

$$B^R(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) = \sup_{G \in \mathcal{V}_\varepsilon^R(F_{\boldsymbol{\mu}})} b(\mathbf{T}, G, \boldsymbol{\mu}).$$

Actually, in most of the cases $B(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}}) = B^R(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}})$.

Similarly we can define the restricted *contamination sensitivity* of an estimate $\mathbf{T}$ as

$$\gamma^R(\mathbf{T}, F_{\boldsymbol{\mu}}) = \left.\frac{\partial B^R(\mathbf{T}, \varepsilon, F_{\boldsymbol{\mu}})}{\partial \varepsilon}\right|_{\varepsilon=0}.$$

If f_0 in (8) is a decreasing function, it can be proved (see Adrover and Yohai [11] and Zuo et al. [5]) that

$$B(\mathbf{T}_{MP}, \varepsilon, F_{\boldsymbol{\mu}}) = B^R(\mathbf{T}_{MP}, \varepsilon, F_{\boldsymbol{\mu}}).$$

To give the expression for the maxbias of $\mathbf{T}_{MP}$ and $\mathbf{T}_{MP}^M$ we need to introduce the following notation:

$$m_1(c) = \arg\min\left\{d : P(|X - c| \leq d) \geq \frac{1-2\varepsilon}{2(1-\varepsilon)}\right\}, \tag{17}$$

$$m_2(c) = \arg\min\left\{d : P(|X - c| \leq d) \geq \frac{1}{2(1-\varepsilon)}\right\}$$

and

$$k(c) = \frac{c}{m_1(c)}. \tag{18}$$

Moreover, put

$$d_1 = H_0^{-1}\left(\frac{1}{2(1-\varepsilon)}\right),$$

$$d_0 = \sup_{c \in [0, d_1]} k(c),$$

$$d_2 = m_2(d_1)$$

and

$$d_3 = m_1(d_1).$$

Adrover and Yohai [11] show that the maxbias of $\mathbf{T}_{MP}$ is given by

$$B(\mathbf{T}_{MP}, \varepsilon, F_{\boldsymbol{\mu}}) = d_1 + d_0 d_2. \tag{19}$$

Table 1
Maximum biases for $\varepsilon = 0.05$ and $\varepsilon = 0.10$.

p	$\varepsilon = 0.05$						$\varepsilon = 0.10$					
	Estimator						Estimator					
	SD_0	SD_{90}	MVE	MCD	MP	MPM	SD_0	SD_{90}	MVE	MCD	MP	MPM
2	0.08	0.13	0.39	0.28	0.14	0.066	0.16	0.29	0.62	0.54	0.32	0.14
3	0.09	0.15	0.39	0.31	0.14	0.066	0.20	0.33	0.69	0.65	0.32	0.14
4	0.12	0.19	0.39	0.34	0.14	0.066	0.27	0.44	0.72	0.79	0.32	0.14
5	0.16	0.22	0.39	0.37	0.14	0.066	0.36	0.52	0.73	0.93	0.32	0.14
6	0.20	0.25	0.39	0.41	0.14	0.066	0.48	0.61	0.74	1.09	0.32	0.14
7	0.25	0.30	0.39	0.44	0.14	0.066	0.62	0.72	0.75	1.28	0.32	0.14
8	0.31	0.35	0.39	0.48	0.14	0.066	0.76	0.84	0.75	1.48	0.32	0.14
9	0.37	0.40	0.39	0.52	0.14	0.066	0.89	0.95	0.75	1.70	0.32	0.14
10	0.42	0.44	0.39	0.56	0.14	0.066	1.02	1.07	0.75	1.96	0.32	0.14
15	0.71	0.72	0.39	0.78	0.14	0.066	1.75	1.75	0.76	3.85	0.32	0.14
20	1.01	1.01	0.39	1.12	0.14	0.066	2.47	2.47	0.77	7.00	0.32	0.14

Remark 1. When $k(x)$ is non-decreasing for $x \in [0, d_1]$, it turns out that $d_0 = d_1/d_3$. If $\mathbf{X}$ is multivariate normal, then H_0 is $N(0, 1)$. Numerical computations show that in this case $k(x)$ is increasing for $x \leq d_1$ provided $\varepsilon < 0.4088$, and $B(\mathbf{T}_{MP}, \varepsilon, F_{\mathbf{0}}) = d_1(1 + d_2/d_3)$. If $\varepsilon > 0.4088$, then $\partial k(x, \varepsilon)/\partial x|_{x=d_1} < 0$ and the maxbias is given by (19).

Theorem 2 below shows that $\mathbf{T}_{MP}^M$, as its original counterpart $\mathbf{T}_{MP}$, also reaches the maximal breakdown point for an equivariant estimator.

Theorem 2. *The asymptotic breakdown point of $\mathbf{T}_{MP}^M$ is $\epsilon^*(\mathbf{T}_{MP}^M, F_{\boldsymbol{\mu}}) = 0.5$.*

To get a deeper insight into the robust behavior of the modified PM estimate, the following result gives the restricted maxbias curve of $\mathbf{T}_{MP}^M$.

Theorem 3. *Let $F_{\boldsymbol{\mu}}$ with density given by (8) and f_0 decreasing. Assume that $\Lambda(F_{\boldsymbol{\mu}}) = \Sigma$. Then,*

(i) *The maximum bias of $\mathbf{T}_{MP}^M$ restricted to contaminations in the neighborhood $\mathcal{V}_\varepsilon^R(F_{\boldsymbol{\mu}})$ is given by*

$$B^R(\mathbf{T}_{MP}^M, \varepsilon, F_{\boldsymbol{\mu}}) = \begin{cases} d_1 & \text{if } d_2(d_0 - 1) < d_1 \\ d_0 d_2 & \text{if } d_2(d_0 - 1) \geq d_1. \end{cases}$$

(ii) *There exists $\varepsilon_0 > 0$ such that $d_2(d_0 - 1) < d_1$ holds for $\varepsilon < \varepsilon_0$. Therefore T_{MP}^M is bias minimax for $\varepsilon < \varepsilon_0$ for the restricted neighborhood.*

(iii) *The restricted contamination sensitivity is*

$$\gamma^R(\mathbf{T}_{MP}^M, \varepsilon, F_{\boldsymbol{\mu}}) = \left.\frac{\partial d_1(\varepsilon)}{\partial \varepsilon}\right|_{\varepsilon=0},$$

which is the lowest bound attainable.

Remark 2. In the normal case the condition $d_2(d_1 - d_3) < d_1 d_3$ holds provided $\varepsilon < 0.3144$ and therefore the lowest bound for maximum bias is attainable. It is worth noticing that even when $\varepsilon > 0.3144$ the maximum bias of $\mathbf{T}_{MP}^M$ is smaller than the bias of $\mathbf{T}_{MP}$ and according to the expression given in Theorem 3 (i) and (19) the difference between both is d_1.

In Tables 1 and 2 we compare the maximum biases for pointwise contamination of the estimates T_{MP}^M (MPM), T_{MP} (MP), the minimum volume ellipsoid (MVE), the minimum covariance determinant (MCD) and two Stahel–Donoho estimates (SD_0 and SD_{90}). The Stahel–Donoho estimates are of the form

$$\widehat{\mu} = \frac{\sum_{i=1}^{n} w(\nu(\mathbf{x}_i, F_n))\mathbf{x}_i}{\sum_{i=1}^{n} w(\nu(\mathbf{x}_i, F_n))}, \tag{20}$$

where F_n is the empirical distribution, ν is defined in (2), the weight function w is equal to $w(u) = 1/u$ for SD_0 and

$$w(u) = \min\left(\frac{1}{u}, \frac{1}{c}\right)$$

and $c = \left(\chi^2_{0.90,p}\right)^{1/2}$, where $\chi^2_{\alpha,p}$ is the α-quantile of the χ^2-distribution with p degrees of freedom for SD_{90}. We have also tried other intermediate values of c, and the maximum biases in all cases were an increasing function of c.

The values for MVE, MCD, SD_0 and SD_{90} were taken from Tables 1 and 2 from [11]. We note that the new estimate MPM outperforms all the other estimates for all values of p and ε. When p increases, the advantage of the MPM estimate becomes more notorious.

Table 2
Maximum biases for $\varepsilon = 0.15$ and $\varepsilon = 0.20$.

p	$\varepsilon = 0.15$						$\varepsilon = 0.20$					
	Estimator						Estimator					
	SD_0	SD_{90}	MVE	MCD	MP	MPM	SD_0	SD_{90}	MVE	MCD	MP	MPM
2	0.27	0.51	0.91	0.90	0.56	0.23	0.40	0.81	1.29	1.49	0.90	0.32
3	0.36	0.56	1.03	1.20	0.56	0.23	0.57	0.98	1.48	2.22	0.90	0.32
4	0.48	0.79	1.08	1.58	0.56	0.23	0.84	1.31	1.57	3.19	0.90	0.32
5	0.67	0.95	1.12	2.04	0.56	0.23	1.18	1.59	1.65	4.46	0.90	0.32
6	0.90	1.14	1.14	2.58	0.56	0.23	1.58	1.91	1.71	6.17	0.90	0.32
7	1.13	1.33	1.16	3.23	0.56	0.23	1.98	2.25	1.78	8.53	0.90	0.32
8	1.41	1.54	1.18	4.07	0.56	0.23	2.44	2.62	1.84	11.69	0.90	0.32
9	1.67	1.77	1.20	5.05	0.56	0.23	2.91	3.02	1.90	15.92	0.90	0.32
10	1.91	1.99	1.22	6.28	0.56	0.23	3.41	3.48	1.97	21.68	0.90	0.32
15	3.27	3.28	1.31	18.00	0.56	0.23	5.76	5.77	2.30	116.8	0.90	0.32
20	4.63	4.63	1.41	49.92	0.56	0.23	8.12	8.11	2.65	412.0	0.90	0.32

3. Computing algorithm

We compute an approximate modified MP estimate as follows. Consider a sample $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in R^p. First, we compute the approximate outlyingness $\upsilon_n(\mathbf{x}_i)$ for each observation of the sample as in Section 4.1 of [17]. According to (4), the location MP estimate is the value in R^p with the smallest outlyingness. A set of candidates to minimize the outlyingness is generated as follows. We draw M random subsamples J of size $p+1$ from the set $\{1, \ldots, n\}$ and we compute the mean and covariance matrix for those subsamples $\mu_J = \text{ave}_{i\in J}(\mathbf{x}_i)$ and $\Sigma_J = \text{ave}_{i\in J}(\mathbf{x}_i - \boldsymbol{\mu}_J)'(\mathbf{x}_i - \boldsymbol{\mu}_J)$, where ave stands for average. Let $h = [n/2]$. We perform two concentration steps as proposed by Rousseeuw and Van Driessen [18]. The concentration steps are as follows: given the Mahalanobis distances $d_i(\boldsymbol{\mu}_J, \Sigma_J) = \sqrt{(\mathbf{x}_i - \boldsymbol{\mu}_J)\Sigma_J^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_J)'}$, $i = 1, \ldots, n$ we construct a h-subset by sorting the Mahalanobis distances

$$d_{1:n}(\boldsymbol{\mu}_J, \Sigma_J) \le d_{2:n}(\boldsymbol{\mu}_J, \Sigma_J) \le \cdots \le d_{\dot{n}:n}(\boldsymbol{\mu}_J, \Sigma_J)$$

and keeping the indexes $H_1 = H_1(J) = \{\pi_1(1), \ldots, \pi_1(h)\}$, with π_1 the permutation which gives the ordered sample $d_{i:n}(\boldsymbol{\mu}_J, \Sigma_J) = d_{\pi(i)}(\boldsymbol{\mu}_J, \Sigma_J)$. Then we compute $\boldsymbol{\mu}_{H_1} = \text{ave}_{i\in H_1}(\mathbf{x}_i)$ and $\Sigma_{H_1} = \text{ave}_{i\in H_1}(\mathbf{x}_i - \boldsymbol{\mu}_{H_1})'(\mathbf{x}_i - \boldsymbol{\mu}_{H_1})$. Next, we carry out another concentration step by computing the Mahalanobis distances $d_i(\boldsymbol{\mu}_{H_1}, \Sigma_{H_1})$, $i = 1, \ldots, n$ and we get another subset $H_2 = H_2(J) = \{\pi_2(1), \ldots, \pi_2(h)\}$ through the ordered sample $d_{i:n}(\boldsymbol{\mu}_{H_1}, \Sigma_{H_1}) = d_{\pi(i)}(\boldsymbol{\mu}_{H_1}, \Sigma_{H_1})$, $i = 1, \ldots, n$, where π_2 is the permutation which gives the ordered sample. The resulting mean is $\boldsymbol{\mu}_{H_2} = \text{ave}_{i\in H_2}(\mathbf{x}_i)$. Then the set of M candidates to minimize the modified outlyingness is given by $U = \{\boldsymbol{\mu}_{H_2(J)} : J \text{ a random subsample}\}$. We consider $M(M-1)/2$ half-spaces generated as $L(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = \{\mathbf{x} : (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\mathbf{x} \ge 0\}$, with, $\boldsymbol{\mu}_1 \ne \boldsymbol{\mu}_2$, $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ in U. The set of directions $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_\mathbf{N}\}$ is generated through a random sample $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ from a multivariate normal distribution $N_p(\mathbf{0}, I)$. Then, an approximate MP estimate $\widehat{\boldsymbol{\mu}}_n$ and an approximate modified MP estimate $\widehat{\boldsymbol{\mu}}_n^M$ are computed through the following scheme:

$$h_n(\boldsymbol{\mu}, \mathbf{a}) = \frac{\text{med}_{1\le i\le n}(\mathbf{a}'(\mathbf{x}_i - \boldsymbol{\mu}))}{\text{med}_{1\le i\le n}|\mathbf{a}'\mathbf{x}_i - \text{med}_{1\le j\le n}(\mathbf{a}'\mathbf{x}_j)|}, \quad \mathbf{a} \in A, \boldsymbol{\mu} \in U$$

$$\nu_n(\boldsymbol{\mu}) = \max_{\mathbf{a}\in A} |h_n(\boldsymbol{\mu}, \mathbf{a})|, \quad \boldsymbol{\mu} \in U$$

$$V_n(\boldsymbol{\mu}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2) = \max_{\mathbf{a}\in L(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2)\cap A} h_n(\boldsymbol{\mu}, \mathbf{a}) - \min_{\mathbf{a}\in L(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2)\cap A} h_n(\boldsymbol{\mu}, \mathbf{a})$$

$$\widehat{\boldsymbol{\mu}}_n = \arg\min_{\boldsymbol{\mu}\in U} \nu_n(\boldsymbol{\mu})$$

$$\widehat{\boldsymbol{\mu}}_n^M = \arg \min_{\boldsymbol{\mu}\in U, \boldsymbol{\mu}_1\in U, \boldsymbol{\mu}_2\in U} V_n(\boldsymbol{\mu}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2).$$

The sample mean was also included in the set of candidates to improve the efficiency of the estimate. There are some small differences between the efficiencies computed in Adrover and Yohai [11] since the approximate algorithm used in this paper differs from the procedure in [11].

4. Monte Carlo efficiencies

We perform a Monte Carlo study to compare the efficiencies under multivariate normal distribution for finite sample size of the estimates considered in Sections 2 and 3. Since all the estimates are equivariant we consider without loss of generality only the case of zero mean and identity covariance matrix. We also include in this study the sample mean which is optimal in the normal case. We take $p = 2$–10, 15 and 20. The sample size n was chosen as equal to 100. The number of replications was 500. For each estimate $\mathbf{T}$ we compute the mean square error (MSE) defined by

$$\frac{1}{500}\sum_{i=1}^{500} \|\mathbf{T}_i\|^2,$$

Table 3
MSE of the Mean and Relative Efficiencies (RE) of robust estimates for the Gaussian Distribution and $n = 100$.

p	MEAN	SD_0	SD_{90}	MVE	MCD	MP	MPM
	MSE	RE	RE	RE	RE	RE	RE
2	0.019	0.80	0.97	0.17	0.20	0.72	0.64
3	0.030	0.83	0.98	0.14	0.22	0.77	0.75
4	0.039	0.89	0.98	0.13	0.28	0.76	0.75
5	0.049	0.89	0.98	0.13	0.30	0.77	0.78
6	0.058	0.91	0.98	0.13	0.32	0.76	0.78
7	0.069	0.91	0.98	0.13	0.34	0.79	0.77
8	0.081	0.92	0.99	0.13	0.35	0.83	0.78
9	0.090	0.93	0.99	0.13	0.38	0.82	0.79
10	0.099	0.94	0.99	0.14	0.39	0.86	0.80
15	0.148	0.95	0.99	0.18	0.43	0.94	0.81
20	0.203	0.95	0.99	0.22	0.45	0.97	0.84

where $\mathbf{T}_i$ is the value of the estimate for the ith sample. To compute the $\mathbf{T}_{MP}$ and the $\mathbf{T}_{MP}^{M}$ estimators we use the algorithm described in Section 3 with $M = 500$ and $N = 500$.

In Table 3 we show the MSE of the mean and the relative efficiencies with respect to the mean of SD_0, SD_{90}, MVE, MCD, MP and MPM for different values of p. The results for the estimates SD_0, SD_{90}, MVE and MCD were taken from Table 4 of [11]. The most efficient estimates are both SD estimates followed by the MP estimate. The new proposal MPM ranks a little less efficient than the MP estimate but much more efficient than MVE and MCD estimates. The MVE estimate was computed using subsampling as explained as in Subsection 6.7.3 of Maronna et al. [13]. To compute SD_0 and SD_{90}, $\nu(\mathbf{x}_i, F_n)$ in (20) was approximated using 500 directions. Each of these directions is orthogonal to the hyperplane determined by a random subsample of size p. The MCD estimate was computed using the fast algorithm proposed in Rousseeuw and Van Driessen [18] with 500 subsamples and two concentration steps.

5. Concluding remarks

A modification of the projection based estimators introduced by Tyler [12] was considered. We show that this new estimator is bias minimax in the restricted neighborhood of point mass contaminations for a range $[0, \varepsilon_0]$, $\varepsilon_0 > 0$ of levels of contaminations. The value ε_0 depends on the central distribution and it is equal to 0.3044 in the normal case. Even we cannot prove that this estimate is bias minimax in the subset $[\varepsilon_0, 0.5)$, it has smaller bias than any other known robust estimate including the projection estimate. The main shortcoming of the new estimate is its computational complexity. Nevertheless, we describe an algorithm based on subsampling which seems to be at the present the best computational approximation to our proposal.

The multivariate problem usually requires the estimation of a center of the data and a scatter matrix. Projection estimates allow for the estimation of the location without using any dispersion matrix. With the help of this estimator we can have a better center of the data and in this way a more accurate estimation of the dispersion matrix can be performed.

Acknowledgments

We are very grateful to two anonymous referees whose comments have contributed greatly to improve the paper.

Appendix

Because of the affine equivariance of the P-estimate, without loss of generality we will assume in this Appendix that the true parameters are $\boldsymbol{\mu} = \mathbf{0}$ and $\Sigma = I$. We need some notation and definitions to deal with the proofs of the main results. $\mathbf{e}_1, \ldots, \mathbf{e}_p$ will denote the canonical basis in R^p, that is $\mathbf{e}_j$ stands for the vector with a 1 in the jth coordinate and 0's elsewhere. Given a set $A \subset R^p$ and $\boldsymbol{\gamma} \in R^p$, we define

$$A - \boldsymbol{\gamma} = \{\mathbf{b} \in R^p : \mathbf{b} + \boldsymbol{\gamma} \in A\},$$
$$-A = \{\mathbf{b} \in R^p : -\mathbf{b} \in A\},$$
$$A^{\perp} = \{\mathbf{n} \in \mathbf{R}^p : \mathbf{a}'\mathbf{n} = 0 \quad \text{for all } \mathbf{a} \in A\}.$$

When $A = \{\mathbf{a}\}$ we will denote $\mathbf{a}^{\perp} = \{\mathbf{a}\}^{\perp}$. The set of affine subspaces is denoted by

$$\mathcal{P} = \{l = \mathbf{a}^{\perp} + \mathbf{v} : \mathbf{a} \in R^p, \mathbf{a} \neq \mathbf{0}, \mathbf{v} \in R^p\}.$$

Given the ball $B(\mathbf{0}, d_1)$ its boundary is denoted by $C(\mathbf{0}, d_1)$. Let $d(Q, \mathbf{v})$ be the Euclidean distance from the point $\mathbf{v}$ to the subset Q. Let $d > 0$ and $l \in \mathcal{P}$ such that $\mathbf{0} \notin l$, then

$$\mathcal{T}(d) = \{\pi \in \mathcal{P} : d(\pi, \mathbf{0}) \leq d\},$$

$$\mathcal{T}_b(d) = \{\pi \in \mathcal{P} : d(\pi, \mathbf{0}) = d\},$$
$$Q_l(d) = \left\{ \mathbf{v} \in R^p : \mathbf{v} = d\frac{(1-\theta)\mathbf{a} + \theta\mathbf{b}}{\|(1-\theta)\mathbf{a} + \theta\mathbf{b}\|}, \mathbf{a}, \mathbf{b} \in l \cap S^{p-1}, 0 \le \theta \le 1 \right\}.$$

We next show the Fisher consistency and the breakdown point of the MP-estimator $\mathbf{T}^M_{MP}$.

Proof of Theorem 1. Without loss of generality we can assume that $F_\mathbf{0}$ is spherically symmetric around $\mathbf{0}$. Suppose first $\boldsymbol{\zeta} = \mathbf{0}$. Since $h(\boldsymbol{\zeta}, \mathbf{a}, F_\mathbf{0}) = 0$, then, for any half-space L we have

$$V(\boldsymbol{\zeta}, L, F_\mathbf{0}) = \sup_{\mathbf{a} \in S^{p-1} \cap L} h(\boldsymbol{\zeta}, \mathbf{a}, F_\mathbf{0}) - \inf_{\mathbf{a} \in S^{p-1} \cap L} h(\boldsymbol{\zeta}, \mathbf{a}, F_\mathbf{0}) = 0.$$

We show now that given $\boldsymbol{\zeta} \neq \mathbf{0}$, we obtain $\min_{L \in \mathcal{L}} V(\boldsymbol{\zeta}, L, F_\mathbf{0}) > 0$. In fact, we have

$$\frac{\mathrm{med}_{F_\mathbf{0}}(\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta}))}{\mathrm{MAD}_{F_\mathbf{0}}(\mathbf{a}'\mathbf{X})} = -\frac{\mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_{F_\mathbf{0}}(\mathbf{a}'\mathbf{X})} = -c\mathbf{a}'\boldsymbol{\zeta},$$

since $\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})$ is constant for any $\mathbf{a} \in S^{p-1}$. Then there exists $0 \le \phi_0 \le \pi/2$ such that

$$V(\boldsymbol{\zeta}, L, F_\mathbf{0}) = c\left[\sup_{\mathbf{a} \in S^{p-1} \cap L} \mathbf{a}'\boldsymbol{\zeta} - \inf_{\mathbf{a} \in S^{p-1} \cap L} \mathbf{a}'\boldsymbol{\zeta}\right]$$
$$= \begin{cases} c\|\boldsymbol{\zeta}\|[\cos\phi_0 + 1] & \text{if } \boldsymbol{\zeta} \notin L, \\ c\|\boldsymbol{\zeta}\|(1 + \sin\phi_0) & \text{if } \boldsymbol{\zeta} \in L. \end{cases}$$

Since the $\min_{0 \le \phi_0 \le \pi/2}(\cos\phi_0 + 1, 1 + \sin\phi_0) = 1$ we get $\min_L V(\boldsymbol{\zeta}, L, F_\mathbf{0}) > c\|\boldsymbol{\zeta}\|$. □

Proof of Theorem 2. Let $\varepsilon < 0.5$. Suppose that there exists a sequence of contaminated distributions H_n such that putting $\mathbf{u}_n = \mathbf{T}^M_{MP}(H_n)$, we have $\lim_{n\to\infty} \|\mathbf{u}_n\| = \infty$. Call L_n the half-space which gives the minimum of

$$V(\mathbf{u}_n, L, H_n) = \sup_{\mathbf{a} \in S^{p-1} \cap L} h(\mathbf{u}_n, \mathbf{a}, H_n) - \inf_{\mathbf{a} \in S^{p-1} \cap L} h(\mathbf{u}_n, \mathbf{a}, H_n)$$

among the half-spaces $L \in \mathcal{L}$. Without loss of generality we can assume that $\mathbf{u}_n \in L_n$. Let $\mathbf{w}_n \in L_n$ such that $\mathbf{u}_n'\mathbf{w}_n = 0$. Then, for n large enough it holds that

$$V(\mathbf{u}_n, L_n, H_n) \ge |h(\mathbf{u}_n, \mathbf{w}_n, H_n) - h(\mathbf{u}_n, \mathbf{u}_n/\|\mathbf{u}_n\|, H_n)| \tag{21}$$

and the right-hand side of the inequality converges to ∞. On the other hand, since the univariate location and dispersion estimates have breakdown point of 0.5 we have

$$V(\mathbf{0}, L_n, H_n) = \sup_{\mathbf{a} \in L_n} \frac{\mathrm{med}_F(\mathbf{a}'\mathbf{X})}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} - \inf_{\mathbf{a} \in L_n} \frac{\mathrm{med}_F(\mathbf{a}'\mathbf{X})}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} < \infty. \tag{22}$$

(21) and (22) contradict the fact that $\mathbf{u}_n = \mathbf{T}^M_{MP}(H_n)$. Therefore, the estimate cannot break down for $\epsilon < 0.5$. □

We next introduce some extra lemmas, rather technical in nature but which are needed to derive the bias performance of the modified P-estimate. The proofs of the lemmas can be found in [16], except for the proof of Lemma A.1, which is quite straightforward, and then it is omitted.

Lemma A.1. (i) *If $l \in T_b(d)$ then there exists a unique point $\mathbf{t} = l \cap C(0, d)$.*
(ii) *Given $\mathbf{c} \in R^p$, $\|\mathbf{c}\| > d_1$ there exists $l^*(\mathbf{c}) \in \mathcal{P}$ such that for each vector $\mathbf{x} \in l^*(\mathbf{c}) \cap C(0, d_1)$ there exists $l^+(\mathbf{x}) \in T_b(d_1)$ such that $\{\mathbf{c}, \mathbf{x}\} \in l^+(\mathbf{x})$ and $l^+(\mathbf{x}) \cap C(0, d_1) = \{\mathbf{x}\}$ for every $\mathbf{x} \in l^*(\mathbf{c}) \cap C(0, d_1)$.*
(iii) *Let $l^*(\mathbf{c})$ be as in* (ii). *Then, for every $\mathbf{x} \in l^*(\mathbf{c}) \cap C(0, d_1)$, $\mathbf{c}'\mathbf{x} = d_1^2$. Moreover, $l^*(\mathbf{c}) = \mathbf{c}^\perp + d_1^2\mathbf{c}/\|\mathbf{c}\|^2$.*

Lemma A.2 below summarizes Lemma 3 and 4 from Adrover and Yohai [11]. This technical result is crucial in the derivation of the main results since it calculates the median and MAD of projections when the central distribution is contaminated by point mass distributions.

Lemma A.2. *Let $F = (1-\varepsilon)F_\mathbf{0} + \varepsilon\delta_\mathbf{c}$ for $\mathbf{c} \in R^p$. If $\|\mathbf{a}\| = 1$ then*

$$\mathrm{med}_F(\mathbf{a}'\mathbf{X}) = \max\{-d_1, \min(\mathbf{a}'\mathbf{c}, d_1)\}$$

and

$$\mathrm{MAD}_F(\mathbf{a}'\mathbf{X}) = m_1(|\mathbf{a}'\mathbf{c}|)1_{(0,d_1)}(|\mathbf{a}'\mathbf{c}|) + d_3 1_{(d_1,d_1+d_3)}(|\mathbf{a}'\mathbf{c}|) + |\mathbf{a}'\mathbf{c}| - d_1|1_{(d_1+d_3,d_1+d_2)}(|\mathbf{a}'\mathbf{c}|) + d_2 1_{(d_1+d_3,d_1+d_2)}(|\mathbf{a}'\mathbf{c}|).$$

Lemma A.2 gives an analytic formulation for the median and the MAD of $\mathbf{a}'\mathbf{X}$. A more geometrical insight of $\mathrm{med}_F(\mathbf{a}'\mathbf{X})$ is provided by the following lemma.

Lemma A.3. *Let* $F = (1-\epsilon)F_{\mathbf{0}} + \epsilon\delta_{\mathbf{c}}$ *with* $\mathbf{c} = (c, 0, \ldots, 0)'$. *Let* $L(\mathbf{c})$ *be defined as in* (5). *Let* $l^*(\mathbf{c}) \in \mathcal{P}$ *be defined as in Lemma A.1* (ii). *Then,*

(i) *If* $0 \le c \le d_1$, $\|\mathbf{a}\| = 1$ *and* $0 \le \mathbf{a}'\mathbf{c}$ *then* $\mathrm{med}_F(\mathbf{a}'\mathbf{X}) = d(\mathbf{a}^{\perp} + \mathbf{c}, 0) = \mathbf{a}'\mathbf{c}$.
(ii) *If* $\|\mathbf{a}\| = 1$, $d_1\mathbf{a} \in Q_{l^*(\mathbf{c})}(d_1)$ *and* $c \ge d_1$ *then* $\mathrm{med}_F(\mathbf{a}'\mathbf{X}) = d_1$.
(iii) *If* $\|\mathbf{a}\| = 1$, $d_1\mathbf{a} \in C(0, d_1) \cap \left(Q_{l^*(\mathbf{c})}(d_1)\right)^c \cap L(\mathbf{c})$,*and* $c \ge d_1$ *then* $\mathrm{med}_F(\mathbf{a}'\mathbf{X}) = d(\mathbf{a}^{\perp} + \mathbf{c}, 0) = \mathbf{a}'\mathbf{c}$.
(iv) *If* $\boldsymbol{\zeta} = \zeta\mathbf{e}_1$ *with* $d_1 < \zeta \le c$, *then there exists* $l^*(\boldsymbol{\zeta}) \in \mathcal{P}$ *such that* $\mathrm{med}_F(\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta})) \le 0$ *if* $\mathbf{a} \in Q_{l^*(\boldsymbol{\zeta})}(d_1)$ *and* $\mathrm{med}_F(\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta})) \ge 0$ *if* $\mathbf{a} \in C(0, d_1) \cap \left(Q_{l^*(\boldsymbol{\zeta})}(d_1)\right)^c \cap L(\mathbf{c})$. *Moreover,* $l^*(\boldsymbol{\zeta}) = \boldsymbol{\zeta}^{\perp} + d_1^2\boldsymbol{\zeta}/\|\boldsymbol{\zeta}\|^2$.

The following lemma shows that the outlyingness measure $\inf_{L\in\mathcal{L}} V(\boldsymbol{\zeta}, L, (1-\varepsilon)F_{\mathbf{0}} + \varepsilon\delta_{\mathbf{c}})$ given in (5) is obtained for the half-space $L(\mathbf{c})$, which is needed in the derivation of the maximum bias.

Lemma A.4. *Let* $\mathbf{X} \sim (1-\varepsilon)F_{\mathbf{0}} + \varepsilon\delta_{\mathbf{c}}$ *for* $\mathbf{c} = (c, 0, \ldots, 0)'$. *If* $\mathbf{a} = (a_1, \ldots, a_p)'$, $\mathbf{b} = D\mathbf{a}$ *where* $D = (d_{ij})$ *is a diagonal matrix such that* $d_{11} = 1$ *and* d_{jj} *is 1 or* -1 *for* $j > 1$. *Put* $\boldsymbol{\zeta} = (\zeta, 0, \ldots, 0)'$, $0 \le \zeta \le c$, *then* $\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta})$ *and* $\mathbf{b}'(\mathbf{X} - \boldsymbol{\zeta})$ *have the same distribution, and therefore*

$$\frac{\mathrm{med}_F(\mathbf{a}'(\mathbf{X} - \boldsymbol{\zeta}))}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{\mathrm{med}_F(\mathbf{b}'(\mathbf{X} - \boldsymbol{\zeta}))}{\mathrm{MAD}_F(\mathbf{b}'\mathbf{X})}. \tag{23}$$

Moreover, let $L(\mathbf{c}) = L^* = \left\{\mathbf{x} : \mathbf{x} = (x_1, x_2, \ldots, x_p)' : x_1 \ge 0\right\}$. *Then, for all* $L \in \mathcal{L}$,

$$V(\boldsymbol{\zeta}, L^*, F) \le V(\boldsymbol{\zeta}, L, F).$$

Proof of Theorem 3. (i) Since $\mathbf{c} = c\mathbf{e}_1$, it can be shown that $\mathbf{T}_{MP}^M$ has also the last coordinates equal to 0. Take $\boldsymbol{\zeta} = \mathbf{d}_1 = d_1\mathbf{e}_1$ and L^* as in Lemma A.4. Let $m_1(c)$ be defined as in (17). Let $c \ge d_1$. Then, if $\mathbf{a}'\mathbf{c} \le d_1$, by using Lemma A.2 we get

$$\begin{aligned} 0 &\le \frac{\mathrm{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{\mathbf{a}'\mathbf{c} - d_1\mathbf{a}'\mathbf{e}_1}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{(c-d_1)}{c}\frac{\mathbf{a}'\mathbf{c}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} \\ &= \frac{(c-d_1)}{c}\frac{\mathbf{a}'\mathbf{c}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{(c-d_1)}{c}\frac{\mathbf{a}'\mathbf{c}}{m_1(|\mathbf{a}'\mathbf{c}|)} \\ &\le \frac{(c-d_1)}{c}d_0. \end{aligned}$$

On the contrary, if $\mathbf{a}'\mathbf{c} > d_1$, by using Lemma A.2 we get

$$\begin{aligned} 0 &\le \frac{\mathrm{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1 - d_1\mathbf{a}'\mathbf{e}_1}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1(1 - \mathbf{a}'\mathbf{e}_1)}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} \\ &\le \frac{d_1(1 - d_1/c)}{d_3} \le \frac{d_1}{d_3}\frac{(c-d_1)}{c}. \end{aligned}$$

If $\mathbf{a} = \mathbf{e}_p'$ then med $(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta} = 0$ and then $h(\boldsymbol{\zeta}, \mathbf{a}, F) = 0$. Therefore,

$$\begin{aligned} V(\mathbf{d}_1, L^*, F) &= \sup_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) - \inf_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) \\ &= d_0\frac{(c-d_1)}{c}. \end{aligned}$$

Let us take now $d_1 < \zeta < c$. To calculate $V(\boldsymbol{\zeta}, L^*, F)$ we proceed as follows. If $\mathbf{a}'\mathbf{c} \le d_1$, then

$$\begin{aligned} 0 &\le \frac{\mathrm{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{\mathbf{a}'\mathbf{c} - \zeta\mathbf{a}'\mathbf{e}_1}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{(c-\zeta)}{c}\frac{\mathbf{a}'\mathbf{c}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} \\ &= \frac{(c-\zeta)}{c}\frac{\mathbf{a}'\mathbf{c}}{m_1(|\mathbf{a}'\mathbf{c}|)} \le \frac{(c-\zeta)}{c}d_0. \end{aligned}$$

If $\mathbf{a}'\mathbf{c} > d_1$ we have

$$\frac{\mathrm{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1 - \zeta\mathbf{a}'\mathbf{e}_1}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} \le \frac{d_1\left(1 - \frac{\zeta}{c}\right)}{d_3} = \frac{d_1}{d_3}\frac{(c-\zeta)}{c}.$$

In brief,

$$\sup_{\mathbf{a}\in S^{p-1}\cap L^*} \frac{\mathrm{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\mathrm{MAD}_F(\mathbf{a}'\mathbf{X})} = d_0\frac{(c-\zeta)}{c}.$$

Let us compute the infimum. If $\mathbf{a}'\mathbf{c} \le d_1$, it holds that

$$\frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{\mathbf{a}'\mathbf{c} - \zeta\mathbf{a}'\mathbf{e}_1}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{(c-\zeta)}{c}\frac{\mathbf{a}'\mathbf{c}}{m_1(|\mathbf{a}'\mathbf{c}|)} \ge 0. \tag{24}$$

If $d_1 < c \le d_1 + d_3$ then, either $\mathbf{a}'\mathbf{c} \le d_1$ or $d_1 < \mathbf{a}'\mathbf{c} \le d_1 + d_3$. Because of (24), we just have to deal with the case $d_1 < \mathbf{a}'\mathbf{c} \le d_1 + d_3$. Thus,

$$\frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1 - \zeta\mathbf{a}'\mathbf{e}_1}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} \ge \frac{d_1-\zeta}{d_3}.$$

Therefore, we have

$$\begin{aligned} V(\boldsymbol{\zeta}, L^*, F) &= \sup_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) - \inf_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) \\ &= d_0\frac{(c-\zeta)}{c} + \frac{\zeta - d_1}{d_3} \\ &= d_0\frac{(c-d_1)}{c} + (\zeta - d_1)\left(\frac{1}{d_3} - \frac{d_0}{c}\right) \\ &> d_0\frac{(c-d_1)}{c} = V(\mathbf{d}_1, L^*, F), \end{aligned}$$

if and only if

$$c \ge d_0 d_3.$$

Then, if $d_1 \le c \le d_1 + d_3$ we get that

$$\mathbf{T}^M_{MP}((1-\epsilon)F_{\mathbf{0}} + \epsilon\delta_{\mathbf{c}}) = \begin{cases} \mathbf{d}_1 & \text{if } d_0 d_3 \le c \le d_1 + d_3, \\ \mathbf{c} & \text{if } c \le \min\{d_1 + d_3, d_0 d_3\}. \end{cases} \tag{25}$$

Let us take $d_1 + d_3 < c \le d_1 + d_2$. Since (24) holds, we only have to consider two cases: If $d_1 < \mathbf{a}'\mathbf{c} \le d_1 + d_3$ we have

$$\frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1 - \zeta\mathbf{a}'\mathbf{e}_1}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} \ge \frac{d_1 - \zeta(d_1+d_3)/c}{d_3} = I_1(\zeta).$$

If $d_1 + d_3 < \mathbf{a}'\mathbf{c} \le d_1 + d_2$, since the function $g(x) = x/(cx - d_1)$ is decreasing in $[0, 1]$ and $g(\mathbf{a}'\mathbf{e}_1) \ge g(1)$, then we have

$$\begin{aligned} \frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} &= \frac{d_1 - (\zeta - c)\mathbf{a}'\mathbf{e}_1 - c\mathbf{a}'\mathbf{e}_1}{\mathbf{a}'\mathbf{c} - d_1} \\ &= -1 + (c-\zeta)\frac{\mathbf{a}'\mathbf{e}_1}{\mathbf{a}'\mathbf{c} - d_1} \\ &= -1 + \frac{(c-\zeta)\mathbf{a}'\mathbf{e}_1}{c\mathbf{a}'\mathbf{e}_1 - d_1} \ge -1 + \frac{(c-\zeta)}{c-d_1} \\ &= \frac{d_1-\zeta}{c-d_1} = I_2(\zeta). \end{aligned}$$

Let us note that

$$I_2(d_1) = 0 < d_1\frac{c-(d_1+d_3)}{cd_3} = I_1(d_1),$$
$$I_2(c) = -1 = I_1(c),$$

and then

$$I_2(\zeta) < I_1(\zeta) \quad \text{for } \zeta \in [d_1, c).$$

Then,

$$\begin{aligned} V(\boldsymbol{\zeta}, L^*, F) &= \sup_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) - \inf_{\mathbf{a}\in S^{p-1}\cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) \\ &= d_0\frac{(c-\zeta)}{c} + \frac{\zeta-d_1}{c-d_1} \\ &= \left[\frac{1}{c-d_1} - \frac{d_0}{c}\right]\zeta + d_0 - \frac{d_1}{c-d_1}. \end{aligned}$$

If,

$$\text{either} \quad c < d_1 \frac{d_0}{d_0 - 1} \quad \text{or} \quad d_0 - 1 \leq 0$$

then

$$\frac{1}{c - d_1} - \frac{d_0}{c} > 0,$$

and

$$V(\boldsymbol{\zeta}, L^*, F) > d_0 \left(\frac{c - d_1}{c} \right) = V(\mathbf{d}_1, L^*, F).$$

On the contrary, if

$$d_0 - 1 > 0 \quad \text{and} \quad d_1 \frac{d_0}{d_0 - 1} \leq c < d_1 + d_2$$

then

$$\frac{1}{c - d_1} - \frac{d_0}{c} \leq 0,$$

and we get that

$$\begin{aligned} V(\mathbf{c}, L^*, F) &= \sup_{\mathbf{a} \in S^{p-1} \cap L^*} h(\mathbf{c}, \mathbf{a}, F) - \inf_{\mathbf{a} \in L^*} h(\mathbf{c}, \mathbf{a}, F) = 1 \\ &\leq \frac{d_0 (c - d_1)}{c} = V(\mathbf{d}_1, L^*, F). \end{aligned}$$

Then, if $d_1 + d_3 < c \leq d_1 + d_2$ we have

$$\mathbf{T}^M_{MP}((1-\epsilon)F_{\mathbf{0}} + \epsilon \delta_{\mathbf{c}}) = \begin{cases} \mathbf{d}_1 & \text{if } \begin{array}{l} d_0 - 1 \leq 0 \quad \text{or} \\ c \leq d_1 d_0/(d_0 - 1), \end{array} \\ \mathbf{c} & \text{if } \begin{array}{l} d_0 - 1 > 0 \quad \text{and} \\ d_1 d_0/(d_0 - 1) \leq c \leq d_1 + d_2. \end{array} \end{cases} \tag{26}$$

Let us take now $c > d_1 + d_2$. Then

$$\frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} = \frac{d_1 - \zeta \mathbf{a}'\mathbf{e}_1}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} \geq \frac{d_1 - \zeta}{d_2} = J_1(\zeta).$$

On the other hand, if $d_1 + d_3 < \mathbf{a}'\mathbf{c} \leq d_1 + d_2$,

$$\begin{aligned} \frac{\text{med}(\mathbf{a}'\mathbf{X}) - \mathbf{a}'\boldsymbol{\zeta}}{\text{MAD}_F(\mathbf{a}'\mathbf{X})} &= \frac{d_1 - (\zeta - c)\mathbf{a}'\mathbf{e}_1 - c\mathbf{a}'\mathbf{e}_1}{\mathbf{a}'\mathbf{c} - d_1} \\ &= -1 + (c - \zeta) \frac{\mathbf{a}'\mathbf{e}_1}{\mathbf{a}'\mathbf{c} - d_1} \\ &= -1 + \frac{(c - \zeta)\mathbf{a}'\mathbf{e}_1}{c\mathbf{a}'\mathbf{e}_1 - d_1} \geq -1 + \frac{(c - \zeta)(d_1 + d_2)}{cd_2} \\ &= \frac{cd_1 - \zeta(d_1 + d_2)}{cd_2} = J_2(\zeta). \end{aligned}$$

Since $c > d_1 + d_2$ then

$$\begin{aligned} & J_1(\zeta) < J_2(\zeta) \quad \text{for } \zeta \in [d_1, c), \\ & J_1(\zeta) = \frac{d_1 - \zeta}{d_2} < \frac{d_1 - \zeta}{c - d_1} = I_2(\zeta) < I_1(\zeta) < 0 \quad \text{for } \zeta \in [d_1, c), \end{aligned}$$

and we have

$$\begin{aligned} V(\boldsymbol{\zeta}, L^*, F) &= \sup_{\mathbf{a} \in S^{p-1} \cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) - \inf_{\mathbf{a} \in S^{p-1} \cap L^*} h(\boldsymbol{\zeta}, \mathbf{a}, F) \\ &= d_0 \frac{(c - \zeta)}{c} + \frac{\zeta - d_1}{d_2} \\ &= \left[\frac{1}{d_2} - \frac{d_0}{c} \right] \zeta + \left[d_0 - \frac{d_1}{d_2} \right]. \end{aligned}$$

Thus,

$$\frac{1}{d_2} - \frac{d_0}{c} < 0$$

if and only if

$$c < d_0 d_2$$

and

$$V(\mathbf{c}, L^*, F) = \sup_{\mathbf{a} \in S^{p-1} \cap L^*} h(\mathbf{c}, \mathbf{a}, F) - \inf_{\mathbf{a} \in S^{p-1} \cap L^*} h(\mathbf{c}, \mathbf{a}, F) = \frac{c - d_1}{d_2}$$
$$< d_0 \frac{(c - d_1)}{c} = V(\mathbf{d}_1, L^*, F).$$

Then, if $d_1 + d_2 < c$ we have

$$\mathbf{T}_{MP}^{M}((1-\epsilon)F_{\mathbf{0}} + \epsilon\delta_{\mathbf{c}}) = \begin{cases} \mathbf{d}_1 & \text{if } d_2 d_0 \le d_1 + d_2 \quad \text{or} \quad c > d_0 d_2, \\ \mathbf{c} & \text{if } \begin{matrix} d_1 + d_2 < d_2 d_0, \quad \text{and} \\ d_1 + d_2 < c \le d_0 d_2. \end{matrix} \end{cases} \tag{27}$$

Observe that (25)–(27) are connected one another, since $d_3 d_0 < d_1 + d_3$ if and only if, either $d_0 - 1 \le 0$ or $d_1 + d_3 < d_1 d_0/(d_0 - 1)$. On the other hand, $d_0 - 1 > 0$ and $d_1 d_0/(d_0 - 1) \le d_1 + d_2$ if and only if $d_1 + d_2 \le d_2 d_0$. Summing up, if $d_3 d_0 < d_1 + d_3$, from (25)–(27) we obtain that the estimator is given by

$$\mathbf{T}_{MP}^{M}((1-\epsilon)F_{\mathbf{0}} + \epsilon\delta_{\mathbf{c}}) = \begin{cases} \mathbf{d}_1 & \text{if } c \in \left[d_1, d_1 \dfrac{d_0}{d_0 - 1}\right], \\ \mathbf{c} & \text{if } c \in \left[d_1 \dfrac{d_0}{d_0 - 1}, d_0 d_2\right], \\ \mathbf{d}_1 & \text{if } c \in (d_0 d_2, \infty). \end{cases}$$

If $d_1 + d_3 \le d_3 d_0$, then the estimator is given by

$$\mathbf{T}_{MP}^{M}((1-\epsilon)F_{\mathbf{0}} + \epsilon\delta_{\mathbf{c}}) = \begin{cases} \mathbf{c} & \text{if } c \in [d_1, d_0 d_2], \\ \mathbf{d}_1 & \text{if } c \in (d_0 d_2, \infty), \end{cases}$$

and the point mass maximum bias is

$$B^R(\mathbf{T}_{MP}^{M}, \varepsilon, F_{\mathbf{0}}) = \begin{cases} d_1 & \text{if } d_2(d_0 - 1) < d_1, \\ d_2 d_0 & \text{if } d_2(d_0 - 1) \ge d_1. \end{cases}$$

(ii) Note that $\lim_{\varepsilon \to 0^+} d_1 = 0$, $\lim_{\varepsilon \to 0^+} d_0 = 0$ and $\lim_{\varepsilon \to 0^+} d_2 > 0$. Then, $\lim_{\varepsilon \to 0^+} d_2(d_0 - 1) < 0 = \lim_{\varepsilon \to 0^+} d_1$ and the statement follows.

(iii) It follows easily from (i) and (ii). □

References

[1] P.J. Huber, Robust estimation of a location parameter, Ann. Math. Statist. 35 (1964) 73–101.
[2] J.W. Tukey, Mathematics and the picturing of data, in: Proceedings of the International Congress of Mathematicians, in: Canadian Math. Congress Montreal, vol. 2, 1975, pp. 523–531.
[3] W.A. Stahel, Breakdown of covariance estimators, Research Report 31, Fachgruppe für Statistik, ETH, Zürich (1981).
[4] D.L. Donoho, Breakdown properties of multivariate location estimators, Ph.D. Thesis, Dept. of Statistics, Harvard University, 1982.
[5] Y. Zuo, H. Cui, D. Young, Influence function and maximum bias of projection depth based estimators, Ann. Statist. 32 (2004) 189–218.
[6] Y. Zuo, R. Serfling, General notions of statistical depth function, Ann. Statist. 28 (2000) 461–482.
[7] F.R. Hampel, A general qualitative definition of robustness, Ann. Math. Statist. 42 (1971) 1887–1896.
[8] F.R. Hampel, The influence curve and its role in robust estimation, J. Amer. Statist. Assoc. 69 (1974) 383–393.
[9] M. Riedel, Bias robustness in parametric models generated by groups, Statistics 22 (1991) 559–578.
[10] X. He, D.G. Simpson, Lower bounds for contamination bias: Globally minimax versus locally linear estimation, Ann. Statist. 21 (1993) 314–337.
[11] J.G. Adrover, V.J. Yohai, Projection estimates of multivariate location, Ann. Statist. 30 (2002) 1760–1781.
[12] D.E. Tyler, Finite sample breakdown points of projection based multivariate location and scatter statistics, Ann. Statist. 22 (1994) 1024–1044.
[13] R.A. Maronna, R.D. Martin, V.J. Yohai, Robust Statistics, Theory and Methods, Wiley, Chichester, 2006.
[14] J. Kim, J. Hwang, Asymptotic properties of location estimators based on projection depth, Statist. Probab. Lett. 55 (2001) 293–299.
[15] Y. Zuo, Projection-based depth functions and associated medians, Ann. Statist. 31 (2003) 1460–1490.
[16] J.G. Adrover, V.J. Yohai, A new projection estimate for multivariate location with minimax bias, Technical Report, Instituto de Calculo, Universidad de Buenos Aires (2009). Available at http://mate.dm.uba.ar/~vyohai/TechRepProjection.pdf.
[17] R.A. Maronna, V.J. Yohai, The behavior of Stahel–Donoho robust multivariate estimator, J. Amer. Statist. Assoc. 90 (1995) 330–341.
[18] P.J. Rousseeuw, K. Van Driessen, A fast algorithm for the minimum covariance estimator, Technometrics 41 (1999) 212–223.