Reweighted based estimators for the common principal components model: Influence funtions and Monte Carlo study*

By Graciela Boente
Universidad de Buenos Aires and CONICET, Argentina
gboente@mate.dm.uba.ar
Ana M. Pires and Isabel M. Rodrigues
Departamento de Matemática and CEMAT, Instituto Superior Técnico, Lisboa, Portugal
ana.pires@math.ist.utl.pt isabel.rodrigues@math.ist.utl.pt

Abstract

The common principal components model for several groups of multivariate observations assumes equal principal axes but different variances along these axes among the groups. Under a common principal components model, plug–in estimators and generalized projection–pursuit estimators have been defined by using score functions on the dispersion measure considered. As it is well known, reweighted estimators allow to improve the asymptotic efficiency of the initial estimators. We will consider plug–in estimators obtained by plugging a reweighted estimator of the scatter matrices into the maximum likelihood equations defining the principal axes. The weights considered penalize observations with large values of the influence measures defined by Boente, Pires and Rodrigues (2002).

Some key words: Common principal components; Outlier detection; Projection–Pursuit; Robust estimation; Reweighted estimators.

Running Head: Reweighted based estimators for the CPC model

Corresponding Author
Graciela Boente
Moldes 1855, 3^o A
Buenos Aires, C1428CRA, Argentina
email: gboente@mate.dm.uba.ar
PHONE AND FAX 54-11-45763375

*This research was partially supported by Grant X-094 from the Universidad de Buenos Aires and Grant 13900-6 of the Fundación Antorchas at Buenos Aires, Argentina and also by the Center for Mathematics and its Applications, Lisbon, Portugal. This research was partially developed while Graciela Boente was visiting the Departamento de Matemática at the Instituto Superior Técnico.

1 Introduction

Several authors, as Flury (1988), have studied models for common structure dispersion. As it is well known, those models have been introduced to overcome the problem of an excessive number of parameters, when dealing with several populations, in multivariate analysis. One such basic common structure assumes that the k covariance matrices have different eigenvalues but identical eigenvectors, i.e.,

$$\mathbf{\Sigma}_i = \boldsymbol{\beta}\mathbf{\Lambda}_i\boldsymbol{\beta}^{\mathrm{T}} \,, \quad 1 \le i \le k \,, \tag{1}$$

where $\mathbf{\Lambda}_i$ are diagonal matrices, $\boldsymbol{\beta}$ is the orthogonal matrix of the common eigenvectors and $\mathbf{\Sigma}_i$ is the covariance matrix of the $i-$th population. The more restrictive proportionality model assumes that the scatter matrices are equal up to a proportionality constant, i.e.,

$$\mathbf{\Sigma}_i = \rho_i \mathbf{\Sigma}_1 \,, \text{ for } 1 \le i \le k \text{ and } \rho_1 = 1 \,. \tag{2}$$

Model (1) was proposed in Flury (1984) and became known as the *Common Principal Components* (CPC) model. The maximum likelihood estimators of $\boldsymbol{\beta}$ and $\mathbf{\Lambda}_i$ are derived in Flury (1984), assuming multivariate normality of the original variables. In Flury (1988) a unified study of the maximum likelihood estimators under a CPC model and under a proportionality model is given.

Let $(\mathbf{x}_{ij})_{1\le j\le n_i, 1\le i\le k}$ be independent observations from k independent samples in $I\!\!R^p$ with location parameter $\boldsymbol{\mu}_i$ and scatter matrix $\mathbf{\Sigma}_i$. Let $N = \sum_{i=1}^{k} n_i$, $\tau_i = \dfrac{n_i}{N}$ and $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i})$. For the sake of simplicity and without loss of generality, we will assume that $\boldsymbol{\mu}_i = \mathbf{0}_p$.

It is well known that, in practice, the classical CPC analysis can be affected by the existence of outliers in a sample. In order to obtain robust estimators, Boente and Orellana (2001) extended the plug–in approach studied in Croux and Haesbroeck (2000) to several populations by considering robust affine equivariant estimators of the covariance matrices $\mathbf{\Sigma}_i$, $1 \le i \le k$. On the other hand, also in the one population setting, i.e., $k = 1$, Croux and Ruiz–Gazen (2005) studied the influence function of the projection–pursuit approach introduced by Li and Chen (1985). Boente and Orellana (2001) also generalized to the common principal components model the projection–pursuit estimates i.e., the estimator of $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p\right)$ is defined as the solution of

$$\widehat{\boldsymbol{\beta}}_1 = \underset{\|\mathbf{b}\|=1}{\operatorname{argmax}} \sum_{i=1}^{k} \tau_i \, s^2(\mathbf{X}_i^{\mathrm{T}}\mathbf{b}) \qquad \widehat{\boldsymbol{\beta}}_j = \underset{\mathbf{b}\in\mathcal{B}_j}{\operatorname{argmax}} \sum_{i=1}^{k} \tau_i s^2(\mathbf{X}_i^{\mathrm{T}}\mathbf{b}) \quad 2 \le j \le p \,, \tag{3}$$

where $\mathcal{B}_j = \{\mathbf{b} : \|\mathbf{b}\| = 1, \mathbf{b}^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_m = 0 \text{ for } 1 \le m \le j-1\}$ and s is a univariate scale estimator. The partial influence functions of the functionals related to both classes of estimators were studied in Boente, Pires and Rodrigues (2002). A more general approach which consists on applying a score function to the scale estimator was considered by Boente, Pires and Rodrigues (2005). This proposal considers a general increasing score function $f : I\!\!R^+ \to I\!\!R$ and estimates the common directions as

$$\widehat{\boldsymbol{\beta}}_1 = \underset{\|\mathbf{b}\|=1}{\operatorname{argmax}} \sum_{i=1}^{k} \tau_i \, f\left\{s^2(\mathbf{X}_i^{\mathrm{T}}\mathbf{b})\right\} \qquad \widehat{\boldsymbol{\beta}}_j = \underset{\mathbf{b}\in\mathcal{B}_j}{\operatorname{argmax}} \sum_{i=1}^{k} \tau_i f\left\{s^2(\mathbf{X}_i^{\mathrm{T}}\mathbf{b})\right\} \quad 2 \le j \le p \,. \tag{4}$$

The estimators of the eigenvalues of the i-th population are then computed as $\widehat{\lambda}_{ij} = s^2(\mathbf{X}_i^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_j)$ for $1 \le j \le p$.

In this paper, we consider a reweighted estimator of the scatter matrices of each population, where the weights do not depend on the Mahalanobis distance as usually (see, for instance, Lopuhaä, 1999), but on the outlier detection measures defined in Boente, Pires and Rodrigues (2002). The paper is organized as follows. In Section 2, we motivate and introduce our proposal while in Sections 3 and 4 the partial influence functions and the asymptotic variances are derived, respectively. In Section 5, through a simulation study, the proposed estimators are compared with those defined through (4) for normal and contaminated samples. All proofs are given in the Appendix.

2 The estimators

Let $\mathcal{O}(p)$ be the group of orthogonal matrices of order p. Let $\mathbf{S}_i$ be the sample covariance matrix of the $i-$th population. Flury (1988) defined maximum likelihood estimates (ML) for normal data as the values solving the system

$$\begin{aligned}
\widehat{\boldsymbol{\Lambda}}_i &= \operatorname{diag}\left(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\mathbf{S}_i\widehat{\boldsymbol{\beta}}\right) \\
\widehat{\boldsymbol{\beta}}_m^{\mathrm{T}}\left[\sum_{i=1}^{k} n_i \frac{\widehat{\lambda}_{im}-\widehat{\lambda}_{ij}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}}\mathbf{S}_i\right]\widehat{\boldsymbol{\beta}}_j &= 0 \qquad \text{for } m \neq j \\
\widehat{\boldsymbol{\beta}}_m^{\mathrm{T}}\widehat{\boldsymbol{\beta}}_j &= \delta_{mj}\ .
\end{aligned} \tag{5}$$

In order to obtain robust alternatives, Boente and Orellana (2001) replaced the sample covariance matrices by robust scatter matrices asymptotically normally distributed and spherically equivariant. Under these conditions, they derived the asymptotic behavior of the robust plug–in estimates for the common principal axis and for their size. As with maximum likelihood estimation, a solution for (5) where the sample matrix is replaced by a robust one, always exists, since $\mathcal{O}(p)$ is compact. Uniqueness conditions are similar to those given in Flury (1988) for the maximum likelihood estimators.

For one population, several authors, such as Critchley (1985), Jaupi and Saporta (1993), Shi (1997), Croux and Haesbroeck (1999), Pison et al. (2000) and Croux and Ruiz–Gazen (2005), have suggested statistical diagnostics and graphical displays for detecting outliers in multivariate analysis, such as side–by–side boxplots of the scores obtained from a robust principal component analysis and index plots based on empirical influence functions. Under a common principal components model, partial influence functions can also be used to detect influential observations in a sample. We will remind the definition of the outlier detection measures introduced by Boente, Pires and Rodrigues (2002). Given an observation $\mathbf{x}$ from the ith population, let

$$IML_i^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda}) = \sum_{r=1}^{p} \frac{\mathrm{PIF}_i(\mathbf{x},\lambda_{\mathbf{S},ir},F)^2}{v_{ir}(\boldsymbol{\beta},\boldsymbol{\lambda})} \tag{6}$$

$$IMB_i^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda}) = \sum_{r=1}^{p} \left\{\mathrm{PIF}_i(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},\boldsymbol{\beta}_{\mathbf{S},r}^{(r)},F_0)\right\}^{\mathrm{T}} \mathbf{A}_{ir}^{-1}(\boldsymbol{\beta},\boldsymbol{\lambda}) \left\{\mathrm{PIF}_i(\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},\boldsymbol{\beta}_{\mathbf{S},r}^{(r)},F_0)\right\}\ , \tag{7}$$

where $\mathrm{PIF}_i(\mathbf{x},\lambda_{\mathbf{S},ir},F)$ denotes the partial influence functions, $\lambda_{\mathbf{S},i\ell}$ and $\boldsymbol{\beta}_{\mathbf{S},r}$ indicate the functionals related to the classical functional estimators $\mathbf{S}$ of the scatter matrix, $\boldsymbol{\beta}$ and $\boldsymbol{\lambda} = (\lambda_{11},\ldots,\lambda_{1p},\ldots,\lambda_{k1},\ldots,\lambda_{kp})^{\mathrm{T}}$ are the unknown parameters, $\mathbf{z}^{(r)}$ the vector $\mathbf{z}$ without the rth component and

$$\mathbf{A}_{ir}(\boldsymbol{\beta},\boldsymbol{\lambda}) = E_{F_i}\left\{\mathrm{PIF}_i(\mathbf{u},\boldsymbol{\beta}_{\mathbf{S},r}^{(r)},F_0)\mathrm{PIF}_i(\mathbf{u},\boldsymbol{\beta}_{\mathbf{S},r}^{(r)},F_0)^{\mathrm{T}}\right\}$$

$$v_{ir}(\boldsymbol{\beta},\boldsymbol{\lambda}) \;\; = \;\; E_{F_i}\left\{\text{PIF}_i(\mathbf{u},\lambda_{\mathbf{S},ir},F)\right\}^2 \; .$$

From now on, the index $\mathbf{S}$ or $\mathbf{V}$ indicates the scatter estimates, classical or robust one, used to compute the plug–in estimates of the common directions and their size. The rth coordinate is not included in the expression for $IMB_i(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda})$, since both its partial influence function and its variance are equal to zero when transforming the data to the diagonal case. When $F_{i,0} = N(\mathbf{0}_p,\boldsymbol{\Lambda}_i)$, expressions (6) and (7) simplify to

$$IML^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i) = IML_i^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda}) \;\; = \;\; \sum_{r=1}^{p} \frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2 - \lambda_{ir}\right\}^2}{2\lambda_{ir}^2} \tag{8}$$

$$IMB^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i) = IMB_i^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda}) \;\; = \;\; \sum_{r=1}^{p}\sum_{s\neq r} \frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)\right\}^2}{\lambda_{ir}\lambda_{is}} \; . \tag{9}$$

The outlier detection measures were defined as $IML(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) = IML_i(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\lambda}})$ and $IMB(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) = IMB_i(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\lambda}})$, where the 'hat' denotes replacement of the unknown parameters by their robust estimators $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\Lambda}}_i = \text{diag}(\widehat{\lambda}_{i1},\ldots,\widehat{\lambda}_{ip})$. This proposal is analogous to the one considered by Pison et al. (2000) for principal factor analysis in order to avoid the masking effect. As those authors mentioned, if one computes $IML_i(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda})$ and $IMB_i(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\lambda})$ using the partial influence function of a robust functional and then the diagnostics measures $IML_i(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\lambda}})$ and $IMB_i(\mathbf{x},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\lambda}})$ at robust estimators, one will not achieve the desired property of detecting influential points.

An estimate of $\boldsymbol{\Sigma}_i$ can be defined as

$$\widehat{\boldsymbol{\Sigma}}_i \;\; = \;\; \kappa_i \, \frac{\displaystyle\sum_{j=1}^{n_i} w\left(IML^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i), IMB^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i)\right)\mathbf{x}_{ij}\,\mathbf{x}_{ij}^{\mathrm{T}}}{\displaystyle\sum_{j=1}^{n_i} w\left(IML^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i), IMB^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i)\right)} \tag{10}$$

since we have assumed $\boldsymbol{\mu}_i = \mathbf{0}_p$. κ_i is a normalizing constant in order to attain asymptotically unbiased estimators of the scatter matrices $\boldsymbol{\Sigma}_i$ at the normal distribution and w is a weight function which bounds the effect of outlying observations. Now, the estimators of the principal axes and of the eigenvalues can be defined by plugging–in these estimators into the system of equations (5), defining the maximum likelihood estimators for normal data.

Our proposal, considers a three–step procedure

- **Step 1**: Obtain initial estimates of the common directions and their size either by considering a plug–in approach or by using generalized projection–pursuit estimates. Denote $\widehat{\boldsymbol{\beta}}_{\mathrm{I}}$ and $\widehat{\boldsymbol{\lambda}}_{\mathrm{I}}$ the resulting estimates.
- **Step 2**: Define $\widehat{\boldsymbol{\Sigma}}_i$ through (10) using $\widehat{\boldsymbol{\beta}}_{\mathrm{I}}$ and $\widehat{\boldsymbol{\lambda}}_{\mathrm{I}}$.
- **Step 3**: Replace $\widehat{\boldsymbol{\Sigma}}_i$ in (5) and solve. Denote $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\lambda}}$ the final estimates.

It should be noticed that the reweighted matrices $\widehat{\boldsymbol{\Sigma}}_i$ are rotationally equivariant, but need not to be affine equivariant. However, this allow us to construct equivariant estimates for the common directions.

3 Influence Functions for the Reweighted Estimators

As mentioned in Section 2, the estimators are obtained by plugging the reweighted robust scatter matrices into the equations defining the maximum likelihood estimators for normal data. They are defined as the solution of (5) where $\mathbf{S}_i$ is replaced by $\widehat{\boldsymbol{\Sigma}}_i$. In order to define the functional related to the estimation procedures, let

$$\begin{aligned} W_1(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}) &= w\left(IML^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}), IMB^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda})\right) & (11)\\ \Psi_1(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}) &= w\left(IML^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}), IMB^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda})\right) \mathbf{x}\mathbf{x}^{\mathrm{T}} & (12)\end{aligned}$$

where $IML(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda})$ and $IMB(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda})$ are defined in (8) and (9), respectively. For a given distribution $F = F_1 \times \ldots \times F_k$, let $\mathbf{V}_i = \mathbf{V}_i(F)$ be the robust scatter functional related to $\widehat{\boldsymbol{\Sigma}}_i$ evaluated at the distribution F, i.e.,

$$\mathbf{V}_i(F) = \kappa_i \frac{E_{F_i}\Psi_1\left(\mathbf{x}_i, \boldsymbol{\beta}_{\mathrm{I}}(F), \boldsymbol{\Lambda}_{\mathrm{I},\,i}(F)\right)}{E_{F_i}W_1\left(\mathbf{x}_i, \boldsymbol{\beta}_{\mathrm{I}}(F), \boldsymbol{\Lambda}_{\mathrm{I},\,i}(F)\right)}\,, \qquad (13)$$

where $\boldsymbol{\beta}_{\mathrm{I}}(F)$ and $\boldsymbol{\Lambda}_{\mathrm{I},\,i}(F)$ are the functionals related to the initial estimates.

We will thus define the functionals $\boldsymbol{\beta}_{\mathbf{V}}(F)$, $\boldsymbol{\Lambda}_{\mathbf{V},i}(F)$, $1 \le i \le k$, related to $\mathbf{V} = (\mathbf{V}_1, \ldots, \mathbf{V}_k)$, which are in fact functions of the initial functionals $\boldsymbol{\beta}_{\mathrm{I}}$ and $\boldsymbol{\Lambda}_{\mathrm{I},\,i}$, as the solution of

$$\begin{aligned} \operatorname{diag}\left\{\boldsymbol{\beta}_{\mathbf{V}}(F)^{\mathrm{T}}\mathbf{V}_i(F)\boldsymbol{\beta}_{\mathbf{V}}(F)\right\} &= \boldsymbol{\Lambda}_{\mathbf{V},i}(F) & (14)\\ \boldsymbol{\beta}_{\mathbf{V},m}(F)^{\mathrm{T}}\left\{\sum_{i=1}^k \tau_i \frac{\lambda_{\mathbf{V},im}(F) - \lambda_{\mathbf{V},ij}(F)}{\lambda_{\mathbf{V},im}(F)\lambda_{\mathbf{V},ij}(F)}\mathbf{V}_i(F)\right\}\boldsymbol{\beta}_{\mathbf{V},j}(F) &= 0 \qquad \text{for } m \neq j & (15)\\ \boldsymbol{\beta}_{\mathbf{V},m}(F)^{\mathrm{T}}\boldsymbol{\beta}_{\mathbf{V},j}(F) &= \delta_{mj}\,. & (16)\end{aligned}$$

When $\boldsymbol{\beta}_{\mathrm{I}}$ and $\boldsymbol{\Lambda}_{\mathrm{I}}$ provide Fisher–consistent estimators, $\mathbf{V}_i$ will provide Fisher–consistent estimators, thus, the solution $(\boldsymbol{\Lambda}_{\mathbf{V},\,i}(F), \boldsymbol{\beta}_{\mathbf{V}}(F))$ will be Fisher–consistent for $(\boldsymbol{\Lambda}_i, \boldsymbol{\beta})$. The following Theorem gives the values of the partial influence functions for the plug–in functionals defined through (14) to (16).

Theorem 3.1. *Let $\mathbf{V}_i(F)$ be the reweighted scatter functional such that $\mathbf{V}_i(F) = \boldsymbol{\Sigma}_i$. Denote by $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p$, $\lambda_{i1}, \ldots, \lambda_{ip}$ the common eigenvectors and the eigenvalues of $\boldsymbol{\Sigma}_i$. Assume that the partial influence functions $PIF_i\left(\mathbf{x}, \mathbf{V}_\ell, F\right)$ exist and that $\lambda_{11} > \ldots > \lambda_{1p}$. Then the partial influence functions of the solution $\boldsymbol{\beta}_{\mathbf{V}}(F)$, $\boldsymbol{\Lambda}_{\mathbf{V},i}(F)$, $1 \le i \le k$, of (14) to (16) are given by*

$$\begin{aligned} PIF_i(\mathbf{x}, \lambda_{\mathbf{V},\,\ell j}, F) &= \boldsymbol{\beta}_j^{\mathrm{T}} PIF_i\left(\mathbf{x}, \mathbf{V}_\ell, F\right)\boldsymbol{\beta}_j & (17)\\ PIF_i(\mathbf{x}, \boldsymbol{\beta}_{\mathbf{V},\,j}, F) &= \sum_{m\neq j}\left\{\sum_{\ell=1}^k \tau_\ell \frac{(\lambda_{\ell m} - \lambda_{\ell j})^2}{\lambda_{\ell m}\lambda_{\ell j}}\right\}^{-1}\left\{\sum_{\ell=1}^k \tau_\ell \frac{\lambda_{\ell j} - \lambda_{\ell m}}{\lambda_{\ell m}\lambda_{\ell j}}\boldsymbol{\beta}_j^{\mathrm{T}} PIF_i\left(\mathbf{x}, \mathbf{V}_\ell, F\right)\boldsymbol{\beta}_m\right\}\boldsymbol{\beta}_m\,. & (18)\end{aligned}$$

Remark 3.1. It is worth noticing that in this case, without any assumptions on the underlying distribution, the partial influence function $\mathrm{PIF}_i(\mathbf{x}, \lambda_{\mathbf{V},\,\ell j}, F)$ is not equal to 0 for $\ell \neq i$ as it is for the plug–in and for the projection pursuit estimates studied in Boente, Pires and Rodrigues (2002, 2005). However, as it will be shown latter, when all the populations have the same elliptical distribution except for changes in the scatter, $\mathrm{PIF}_i(\mathbf{x}, \lambda_{\mathbf{V},\,\ell j}, F) = 0$ for $\ell \neq i$. This happens for instance, if we are interested in computing the partial influence functions for normal distributions.

On the other hand, the $\mathrm{PIF}_i(\mathbf{x}, \boldsymbol{\beta}_{\mathbf{V},\,j}, F)$ depends on the partial influence functions of all the scatter matrices, since now the scatter matrices of the different populations are not independent.

The following Theorem gives the partial influence functions of the reweighted scatter functional for differentiable weight functions.

Theorem 3.2. *Let $\mathbf{V}_i(F)$ be the reweighted scatter functional defined in (13), such that $\mathbf{V}_i(F) = \boldsymbol{\Sigma}_i$. Denote by $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p$, $\lambda_{\ell 1}, \ldots, \lambda_{\ell p}$ the common eigenvectors and the eigenvalues of $\boldsymbol{\Sigma}_\ell$. Let $\boldsymbol{\beta}_{\mathrm{I}}(F)$ and $\boldsymbol{\Lambda}_{\mathrm{I},\,i}(F)$ be Fisher–consistent functionals related to the initial estimates of the common eigenvectors and of the eigenvalues of the $i-th$ population. Assume that the function $w : I\!R^2 \to I\!R$ is differentiable and that the partial influence functions $PIF_i\left(\mathbf{x}, \boldsymbol{\beta}_{\mathrm{I}}, F\right)$ and $PIF_i\left(\mathbf{x}, \boldsymbol{\Lambda}_{\mathrm{I},\,\ell}, F\right)$ exists and that $PIF_i\left(\mathbf{x}, \boldsymbol{\Lambda}_{\mathrm{I},\,\ell}, F\right) = 0$ for $\ell \neq i$. Denote $w_1(u, v) = \dfrac{\partial w(u,v)}{\partial u}$ and $w_2(u, v) = \dfrac{\partial w(u,v)}{\partial v}$ and assume that we can differentiate under the integral with respect to ϵ the functions*

$$
\begin{aligned}
&E_{F_\ell}\Psi_1\left(\mathbf{x}, \boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}), \boldsymbol{\Lambda}_{\mathrm{I},\,i}(F_{\epsilon,\mathbf{x}_o,i})\right)\\
&E_{F_\ell}W_1\left(\mathbf{x}, \boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}), \boldsymbol{\Lambda}_{\mathrm{I},\,i}(F_{\epsilon,\mathbf{x}_o,i})\right)\ ,
\end{aligned}
$$

where $F_{\epsilon,\mathbf{x},i} = F_1 \times \ldots \times F_{i-1} \times F_{i,\epsilon,\mathbf{x}} \times F_{i+1} \times \ldots \times F_k$ and $F_{i,\epsilon,\mathbf{x}} = (1-\epsilon)F_i + \epsilon\delta_{\mathbf{x}}$, where $\delta_{\mathbf{x}}$ denotes the point mass at $\mathbf{x}$. Let $\varphi_j^{(\ell)}(\mathbf{x}) = w_j\left(IML^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}_\ell), IMB^2(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}_\ell)\right)$ for $j = 1, 2$. Then, the partial influence functions of $\mathbf{V}_\ell(F)$ are given by

$$
PIF_i(\mathbf{y}, \mathbf{V}_\ell, F) \;=\; D_\ell^{-1}(F)\,\kappa_\ell\,\boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{y}, F) - D_\ell^{-1}(F)\Delta_{i,\,\ell}(\mathbf{y}, F)\mathbf{V}_\ell(F)\ .
$$

where

$$
\begin{aligned}
D_\ell(F) \;&=\; E_{F_\ell}W_1\left(\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\Lambda}_\ell\right)\\
\Delta_{i,\,i}(\mathbf{y}, F) \;&=\; -D_i(F) + W_1(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\Lambda}_i) + \\
&\quad + \sum_{r=1}^{p} \frac{1}{\lambda_{ir}^3} E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2 - \lambda_{ir}\right]\left[2\,\lambda_{ir}\,\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}} PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,r}, F\right) - \right.\right.\\
&\quad \left.\left. - \left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2 PIF_i\left(\mathbf{y}, \lambda_{\mathrm{I},\,ir}, F\right)\right]\right\} + \\
&\quad + 2\sum_{r=1}^{p}\sum_{s\neq r} E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left[PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,r}, F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s + PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,s}, F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_r - \right.\right.\\
&\quad \left.\left. - \frac{1}{2}\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left(PIF_i\left(\mathbf{y}, \lambda_{\mathrm{I},\,ir}, F\right)\lambda_{is} + PIF_i\left(\mathbf{y}, \lambda_{\mathrm{I},\,is}, F\right)\lambda_{ir}\right)\right]\right\}\\
\Delta_{i,\,\ell}(\mathbf{y}, F) \;&=\; 2\sum_{r=1}^{p}\boldsymbol{\beta}_r^{\mathrm{T}} E_{F_\ell}\left\{\varphi_1^{(\ell)}(\mathbf{x})\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2 - \lambda_{\ell r}\right\}}{\lambda_{\ell r}^2}\mathbf{x}\mathbf{x}^{\mathrm{T}}\right\} PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,r}, F\right) + \\
&\quad + 2\sum_{r=1}^{p}\sum_{s\neq r} E_{F_\ell}\left\{\varphi_2^{(\ell)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\left[PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,r}, F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s + PIF_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,s}, F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_r\right]\right\}\\
&\qquad \text{for } \ell \neq i\\
\boldsymbol{\Upsilon}_{i,\,i}(\mathbf{y}, F) \;&=\; \Psi_1(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\Lambda}_i) - \mathbf{N}_i(F) +
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{r=1}^{p}\frac{1}{\lambda_{ir}^{3}}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)^2-\lambda_{ir}\right]\left[2\,\lambda_{ir}\,\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\,\mathbf{x}^{\mathrm T} PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)-\right.\right.\\
&\left.\left.-\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)^2 PIF_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)\right]\mathbf{x}\,\mathbf{x}^{\mathrm T}\right\}+\\
&+2\sum_{r=1}^{p}\sum_{s\neq r}E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm T}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left[PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm T}\mathbf{x}\,\mathbf{x}^{\mathrm T}\boldsymbol{\beta}_s+PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm T}\mathbf{x}\,\mathbf{x}^{\mathrm T}\boldsymbol{\beta}_r-\right.\right.\\
&\left.\left.-\frac{1}{2}\frac{\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm T}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left[PIF_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+PIF_i\left(\mathbf{y},\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right]\right]\mathbf{x}\,\mathbf{x}^{\mathrm T}\right\}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{y},F) &= 2\sum_{r=1}^{p}\frac{1}{\lambda_{\ell r}^{2}}E_{F_\ell}\left\{\varphi_1^{(\ell)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)^2-\lambda_{\ell r}\right]\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\mathbf{x}^{\mathrm T}PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)\mathbf{x}\mathbf{x}^{\mathrm T}\right\}+\\
&+2\sum_{r=1}^{p}\sum_{s\neq r}E_{F_\ell}\left\{\varphi_2^{(\ell)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm T}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm T}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\left[PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm T}\mathbf{x}\,\mathbf{x}^{\mathrm T}\boldsymbol{\beta}_s+PIF_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm T}\mathbf{x}\,\mathbf{x}^{\mathrm T}\boldsymbol{\beta}_r\right]\mathbf{x}\,\mathbf{x}^{\mathrm T}\right\}\\
&\qquad\text{for } \ell\neq i\ ,
\end{aligned}
$$

with $\mathbf{N}_\ell(F)=E_{F_\ell}\Psi_1\left(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_\ell\right)$.

The expression for the partial influence function of the scatter matrices involves, as expected, the partial influence functions of the eigenvalues and eigenvectors functionals computed in the first step, so that the usual influence function of a scatter matrix with nonrandom weights, given by $\Psi_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)-\mathbf{N}_i(F)$ is corrected due to the initial estimation procedure.

From Theorems 3.1 and 3.2, using that $\Delta_{i,\,\ell}(\mathbf{y},F)=0$ and $\boldsymbol{\beta}_j^{\mathrm T}\boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{y},F)\boldsymbol{\beta}_j=0$, if F_i is an ellipsoidal distribution, for $1\le i\le k$, we get easily the following Corollary.

Corollary 3.1. *Under the conditions of Theorem 3.2, denote by* $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_p$, $\lambda_{i1},\dots,\lambda_{ip}$ *the common eigenvectors and the eigenvalues of* $\boldsymbol{\Sigma}_i$. *Assume that* F_i *are ellipsoidal distributions, for* $1\le i\le k$. *Then, the partial influence functions of the solution* $\boldsymbol{\beta}_{\mathbf{V}}(F)$, $\boldsymbol{\Lambda}_{\mathbf{V},i}(F)$, $1\le i\le k$, *of (14) to (16) are given by*

$$
\begin{aligned}
PIF_i(\mathbf{x},\lambda_{\mathbf{V},\,\ell j},F) &= \delta_{\ell i}\,D_i^{-1}(F)\left\{\kappa_i\,\boldsymbol{\beta}_j^{\mathrm T}\boldsymbol{\Upsilon}_{i,\,i}(\mathbf{x},F)\boldsymbol{\beta}_j-\Delta_{i,\,i}(\mathbf{x},F)\lambda_{ij}\right\}\\
PIF_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},\,j},F) &= \sum_{m\neq j}\left\{\sum_{\ell=1}^{k}\tau_\ell\frac{(\lambda_{\ell m}-\lambda_{\ell j})^2}{\lambda_{\ell m}\lambda_{\ell j}}\right\}^{-1}\left\{\sum_{\ell=1}^{k}\tau_\ell\,\frac{\lambda_{\ell j}-\lambda_{\ell m}}{\lambda_{\ell m}\lambda_{\ell j}}\kappa_\ell\,D_\ell^{-1}(F)\boldsymbol{\beta}_j^{\mathrm T}\boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{x},F)\boldsymbol{\beta}_m\right\}\boldsymbol{\beta}_m\ .
\end{aligned}
$$

Note that, as mentioned above, under elliptical distributions, $\mathrm{PIF}_i(\mathbf{x},\lambda_{\mathbf{V},\,\ell j},F)=0$ for $\ell\neq i$. Moreover, straightforward calculations (given in the Appendix) allow to show that, if all the populations have ellipsoidal distributions, $\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},\,j},F)$ depends only on the partial influence functions of the initial eigenvectors while $\mathrm{PIF}_i(\mathbf{x},\lambda_{\mathbf{V},\,ij},F)$ depends only on those of the initial eigenvalues.

Corollary 3.2. *Under the conditions of Theorem 3.2, denote by* $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_p$, $\lambda_{i1},\dots,\lambda_{ip}$ *the common eigenvectors and the eigenvalues of* $\boldsymbol{\Sigma}_i$. *Moreover, assume that* $\boldsymbol{\Sigma}_i=diag\,(\lambda_{i1},\dots,\lambda_{ip})$, *i.e.,* $\boldsymbol{\beta}=\mathbf{I}_p$ *and that* $\boldsymbol{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{i1}=\mathbf{z}_i$ *have the same spherical distribution* G *for all* $1\le i\le k$. *Denote by* κ *the constant* *that satisfies*

$$\mathbf{I}_p = \kappa \, \frac{E_G W_1(\mathbf{z},\mathbf{I},\mathbf{I})\,\mathbf{z}\mathbf{z}^{\mathrm{T}}}{E_G W_1(\mathbf{z},\mathbf{I},\mathbf{I})} \,.$$

Then, $PIF_i(\mathbf{x}, \lambda_{\mathbf{V},\,\ell j}, F) = 0$ *for* $\ell \neq i$ *and*

$$\begin{aligned} PIF_i(\mathbf{x}, \lambda_{\mathbf{V},\,ij}, F) &= \frac{\kappa\, \Psi_1(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)_{jj} - \lambda_{ij}\, W_1(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i) - \alpha_1(G) PIF_i\left(\mathbf{x}, \lambda_{\mathrm{I},\,ij}, F\right) - \alpha_2(G)\lambda_{ij} A_j(\mathbf{x},F)}{D(G)} \\ &= \frac{W_1(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)\left(\kappa\, x_j^2 - \lambda_{ij}\right) - \alpha_1(G) PIF_i\left(\mathbf{x}, \lambda_{\mathrm{I},\,ij}, F\right) - \alpha_2(G)\lambda_{ij} A_j(\mathbf{x},F)}{D(G)} \end{aligned} \qquad (19)$$

and

$$PIF_i(\mathbf{x}, \boldsymbol{\beta}_{\mathbf{V},\,j}, F) = \frac{\kappa}{D(G)} \left\{ \sum_{m\neq j} \frac{\tau_i \dfrac{\lambda_{ij}-\lambda_{im}}{\lambda_{im}\lambda_{ij}} \Psi_1(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)_{jm}}{\displaystyle\sum_{\ell=1}^{k} \tau_\ell \frac{(\lambda_{\ell m}-\lambda_{\ell j})^2}{\lambda_{\ell m}\lambda_{\ell j}}} \boldsymbol{\beta}_m - \alpha_3(G) \sum_{m\neq j} PIF_i\left(\mathbf{x}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m \boldsymbol{\beta}_m \right\} , \qquad (20)$$

where $A_j(\mathbf{x},F) = \sum_{r\neq j} \frac{1}{\lambda_{ir}} PIF_i\left(\mathbf{x}, \lambda_{\mathrm{I},\,ir}, F\right)$ *and* $\alpha_1(G) = \kappa\, d_{11}(\varphi_1) - c_{11}(\varphi_1) + 2(p-1) f_{12}(\varphi_2)$, $\alpha_2(G) = \kappa\, d_{12}(\varphi_1) - c_{11}(\varphi_1) + 2 f_{12}(\varphi_2) + 2(p-2) f_{123}(\varphi_2)$, $\alpha_3(G) = 4\left[b_{12}(\varphi_2) + (p-2) b_{123}(\varphi_2)\right] + 2 d_{12}(\varphi_1)$, *with*

$$\begin{aligned} D(G) &= E_G W_1(\mathbf{z},\mathbf{I},\mathbf{I}) & a_{12}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\, z_1^2\, z_2^2\right\} \\ b_{12}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\, z_1^4\, z_2^2\right\} & b_{123}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\, z_1^2\, z_2^2\, z_3^2\right\}\delta_{p>2} \\ c_{11}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\left(z_1^2-1\right)\, z_1^2\right\} & & \\ d_{11}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\left(z_1^2-1\right)\, z_1^4\right\} & d_{12}(\varphi) &= E_G\left\{\varphi(\mathbf{z})\left(z_1^2-1\right)\, z_1^2\, z_2^2\right\} \\ f_{12}(\varphi) &= \kappa\, b_{12}(\varphi) - a_{12}(\varphi) & f_{123}(\varphi) &= \kappa\, b_{123}(\varphi) - a_{12}(\varphi) \end{aligned}$$

and $\varphi_j(\mathbf{z}) = w_j\left(IML^2(\mathbf{z},\mathbf{I},\mathbf{I}), IMB^2(\mathbf{z},\mathbf{I},\mathbf{I})\right)$.

Remark 3.2. Corollary 3.2 gives the partial influence functions assuming differentiability of the weight function. Similar expressions can be obtained when the function w is not differentiable by requiring differentiability to the density of the common distribution G. However, since the weight function is choosen by the practioner, it seems more natural to require smoothness on it than on the underlying distribution.

Remark 3.3. Expressions (19) and (20) show that the partial influence functions of the reweighted functionals are those of the reweighted functional computed with the true parameters corrected by the partial influence functions of the initial functionals. Figures 1 and 2 show the plots of the partial influence functions when $F = F_1 \times F_2$ with $F_1 = N\left(\mathbf{0}, \mathrm{diag}(2,1)\right)$ and $F_2 = N\left(\mathbf{0}, 4\mathrm{diag}(2,1)\right)$. We have considered as initial estimates, the plug–in estimators computed with an S–estimator using as ρ function the biweight Tukey's function calibrated to attain 25% breakdown point and the projection–pursuit estimate computed with an M–estimator of scale using Huber's function calibrated to attain 50% breakdown. The weight function was taken as $w\left(IML^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i), IMB^2(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i)\right) = w_{\text{IML}}(\mathbf{x}_{ij}) \times w_{\text{IMB}}(\mathbf{x}_{ij})$, with

$$w_{\text{IML}}(\mathbf{x}_{ij}) = \begin{cases} 1 & IML(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) < a_L \\ \exp\left[-\dfrac{\left(IML(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) - a_L\right)^2}{2c_L^2}\right] & IML(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) \geq a_L \end{cases}$$

$$w_{\text{IMB}}(\mathbf{x}_{ij}) = \begin{cases} 1 & IMB(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) < a_B \\ \exp\left[-\dfrac{\left(IMB(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) - a_B\right)^2}{2c_B^2}\right] & IMB(\mathbf{x}_{ij},\widehat{\boldsymbol{\beta}},\widehat{\boldsymbol{\Lambda}}_i) \geq a_B \end{cases} \tag{21}$$

with $a_L = G_\lambda^{-1}(0.95)$, $c_L = \dfrac{b_L - a_L}{p+1}$ and $b_L = G_\lambda^{-1}(0.975)$ while $a_B = G_{\boldsymbol{\beta}}^{-1}(0.95)$, $c_B = \dfrac{b_B - a_B}{p+1}$ and $b_B = G_{\boldsymbol{\beta}}^{-1}(0.975)$. The functions G_λ and $G_{\boldsymbol{\beta}}$ are the distribution functions of the random variables γ_λ and $\gamma_{\boldsymbol{\beta}}$, defined in Boente, Pires and Rodrigues (2002) in order to detect influential observations, i.e.,

$$\begin{aligned} \gamma_\lambda &= \left\{\sum_{r=1}^p \frac{\left(z_r^2 - 1\right)^2}{2}\right\}^{\frac{1}{2}} \\ \gamma_{\boldsymbol{\beta}} &= \left(\sum_{r=1}^p \sum_{s\neq r} z_r^2 z_s^2\right)^{\frac{1}{2}} = \left\{\left(\sum_{r=1}^p z_r^2\right)^2 - \sum_{r=1}^p z_r^4\right\}^{\frac{1}{2}}, \end{aligned}$$

where $z_1, \ldots, z_p$ are independent and identically distributed $N(0,1)$ random variables. The weight function is plotted in Figure 3.

It is worth noticing that the plots corresponding to the reweighted projection–pursuit estimates of the largest eigenvalue look like the corresponding ones in a proportional model (see, Boente, Critchley and Orellana, 2004).

An adaptive outlier detection procedure can be designed with the detection measures IML and IMB, as it was done for the Mahalanobis distance by Filzmoser (2004) and Filzmoser, Reimann and Garrett (2005), in order to take into account the sample size. This procedure can be combined with reweighting to provide adaptive reweighted type estimators for the common axes.

4 Asymptotic variances for ellipsoidal distributions

In this section, for the sake of simplicity, we will assume that $\boldsymbol{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{i1} = \mathbf{z}_i$ have the same spherical distribution G for all $1 \leq i \leq k$.

We will derive the asymptotic variances of the estimates defined through (5) when the matrices $\widehat{\boldsymbol{\Sigma}}_i$ are defined in (10) using as initial estimates the plug–in or the projection pursuit estimates defined in Boente and Orellana (2001) and in Boente, Pires and Rodrigues (2005), respectively. One of the main disadvantage of projection–pursuit estimators is their low efficiency. In fact, they provide more resistant estimators than plug–in methods at the cost of some loss of efficiency. Our procedure overcomes this problem, since it improves the efficiency of the initial estimates when dealing with projection–pursuit techniques preserving the robustness of the initial estimates. In other settings, such as regression models, it is well known that reweighted estimates have the same breakdown point as the initial estimates considered. The influence functions given in Corollary 3.2 suggest that the same result holds in this setting. In this Section, the asymptotic variances derived allow to show that the low efficiency of the initial projection–pursuit estimates can be improved by appropiately choosing the tuning constant of the weight function.

4.1 Initial Plug–in Estimators

When considering as initial estimates the plug–in estimates defined through an initial scatter matrix $\mathbf{V}_{i,o}$, Boente, Pires and Rodrigues (2002) have shown that if the influence function $\text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i)$ exists and $\lambda_{11} > \ldots > \lambda_{1p}$, then the partial influence functions of the initial plug–in estimators $\boldsymbol{\beta}_{\text{I}}(F)$, $\boldsymbol{\Lambda}_{\text{I},i}(F)$, $1 \le i \le k$,

$$\text{PIF}_i(\mathbf{x}, \lambda_{\text{I},\,\ell j}, F) = \delta_{\ell i}\, \boldsymbol{\beta}_j^{\text{T}} \text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i)\, \boldsymbol{\beta}_j \tag{22}$$

$$\text{PIF}_i(\mathbf{x}, \boldsymbol{\beta}_{\text{I},\,j}, F) = \tau_i \sum_{m\neq j} \frac{\lambda_{ij}-\lambda_{im}}{\lambda_{im}\lambda_{ij}} \left\{ \sum_{\ell=1}^{k} \tau_\ell \frac{(\lambda_{\ell m}-\lambda_{\ell j})^2}{\lambda_{\ell m}\lambda_{\ell j}} \right\}^{-1} \left\{ \boldsymbol{\beta}_j^{\text{T}} \text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i)\, \boldsymbol{\beta}_m \right\} \boldsymbol{\beta}_m\,. \tag{23}$$

Moreover, if $\mathbf{V}_{i,o}$ is an affine equivariant scatter matrix functional, there exists two functions $\alpha^{(i)}$ and $\gamma^{(i)} : [0, \infty) \to I\!R$ such that $\text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i) = \alpha^{(i)}\,(d_i(\mathbf{x}))\, \mathbf{x}\mathbf{x}^{\text{T}} - \gamma^{(i)}\,(d_i(\mathbf{x}))\, \boldsymbol{\Sigma}_i$, where $d_i^2(\mathbf{x}) = \mathbf{x}^{\text{T}} \boldsymbol{\Sigma}_i^{-1} \mathbf{x}$. From now on, we will assume that the initial plug–in estimates are evaluated using affine equivariant scatter estimates.

It is well known that if the initial estimates of $\boldsymbol{\Sigma}_i$ are asymptotically normally distributed and spherically equivariant, i.e.,

A1. $\sqrt{n_i}\,(\mathbf{V}_{i,o} - \boldsymbol{\Sigma}_i) \xrightarrow{\mathcal{D}} \mathbf{Z}_i$ where $\text{vec}\,(\mathbf{Z}_i)$ has a multivariate normal distribution with zero mean and covariance matrix $\mathbf{C}_i$ and $\xrightarrow{\mathcal{D}}$ stands for weak convergence.

A2. For any $\mathbf{G}$ such that $\mathbf{G}\,\mathbf{G}^{\text{T}} = \boldsymbol{\Sigma}_i^{-1}$, the distribution of $\mathbf{G}\mathbf{Z}_i\mathbf{G}^{\text{T}}$ is invariant under orthogonal transformations.

then, $\text{vec}\,(\mathbf{Z}_i) \sim N(\mathbf{0}_p, \mathbf{C}_i)$ where

$$\mathbf{C}_i = \sigma_{i,\,1}\,(\mathbf{I} + K_{pp})\,(\boldsymbol{\Sigma}_i \otimes \boldsymbol{\Sigma}_i) + \sigma_{i,\,2}\ \text{vec}\,(\boldsymbol{\Sigma}_i)\ \text{vec}\,(\boldsymbol{\Sigma}_i)^{\text{T}}\ , \tag{24}$$

with K_{pp} the $p^2 \times p^2$ block matrix with the $(l, m)-$block equal to a $p \times p$ matrix with a 1 at entry (l, m) and 0 everywhere else. Under mild regularity conditions, $E\left(\text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i)\, \text{IF}\,(\mathbf{x}, \mathbf{V}_{i,o}, F_i)^{\text{T}}\right) = \mathbf{C}_i$. This entails that

$$\begin{aligned}
\text{ASVAR}\left(\widehat{\lambda}_{\text{I},\,ij}\right) &= (2\sigma_{i,\,1} + \sigma_{i,\,2})\lambda_{ij}^2\,, \quad 1 \le j \le p \\
\text{ASCOV}\left(\widehat{\lambda}_{\text{I},\,ij}, \widehat{\lambda}_{\text{I},\,im}\right) &= \sigma_{i,\,2}\lambda_{ij}\lambda_{im} \qquad \text{for } m \neq j \\
\text{ASVAR}\left(\widehat{\beta}_{\text{I},\,jm}\right) &= \left[\sum_{i=1}^{k} \tau_i \sigma_{i,\,1} \frac{(\lambda_{ij}-\lambda_{im})^2}{\lambda_{ij}\lambda_{im}}\right] \left[\sum_{i=1}^{k} \tau_i \frac{(\lambda_{ij}-\lambda_{im})^2}{\lambda_{ij}\lambda_{im}}\right]^{-2} \qquad \text{for } m \neq j\,.
\end{aligned}$$

Standard calculations lead to the asymptotic variances of the reweighted based estimates with initial plug–in estimates which are stated in the following Theorem.

Theorem 4.1. *Let $\mathbf{x}_{i1}, \dots, \mathbf{x}_{in_i}$, $1 \le i \le k$, be independent observations from k independent samples with scatter matrix $\mathbf{\Sigma}_i = \mathbf{\Lambda}_i = \text{diag}(\lambda_{i1}, \dots, \lambda_{ip})$ such that $\mathbf{\Lambda}_i^{-\frac{1}{2}} \mathbf{x}_{i1} = \mathbf{z}_i$ have the same spherical distribution G for all $1 \le i \le k$. Assume that $\lambda_{11} > \dots > \lambda_{1p}$, $n_i = \tau_i N$ with $0 < \tau_i < 1$ fixed numbers such that $\sum_{i=1}^k \tau_i = 1$. Let $\mathbf{V}_{i,o}$ be robust affine equivariant estimates of the scatter matrices $\mathbf{\Sigma}_i$, satisfying **A1** and **A2**, where $\mathbf{C}_i$ can be written as in (24). Denote by $\widehat{\boldsymbol{\beta}}_{\text{I}}$ and $\widehat{\boldsymbol{\lambda}}_{\text{I}}$ the plug–in estimates obtained from $\mathbf{V}_{i,o}$ and by $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\lambda}}$ the final estimates computed from the reweighted matrix defined in (10) using $\widehat{\boldsymbol{\beta}}_{\text{I}}$ and $\widehat{\boldsymbol{\lambda}}_{\text{I}}$. Then, we have that*

$$\begin{aligned}
\text{asvar}\left(\widehat{\lambda}_{ij}\right) &= \frac{\lambda_{ij}^2}{\tau_i\, D^2(G)} \Big\{ E_G\left(\eta_1^2(\mathbf{z})\left(\kappa\, z_1^2 - 1\right)^2\right) + \alpha_1^2(G)\left(2\sigma_{i,1} + \sigma_{i,2}\right) + 2(p-1)\alpha_1(G)\alpha_2(G)\sigma_{i,2} + \\
&\qquad + (p-1)\alpha_2^2(G)\left(2\sigma_{i,1} + (p-1)\sigma_{i,2}\right) - 2\alpha_1(G)c_{i,11} - 2(p-1)\alpha_2(G)c_{i,12} \Big\} \\
\text{asvar}\left(\widehat{\beta}_{jm}\right) &= \frac{\kappa^2}{D^2(G)}\pi_{jm}\left\{a_{12}\left(\eta_1^2\right) + \pi_{jm}\left[\sum_{i=1}^k \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{ij}\lambda_{im}}\left(\alpha_3^2(G)\sigma_{i,1} - 2\alpha_3(G)d_{i,12}\right)\right]\right\} \\
\text{ascov}\left(\widehat{\beta}_{jm}, \widehat{\beta}_{js}\right) &= 0 \qquad \text{for} \qquad m \ne s
\end{aligned}$$

where κ and $a_{12}(\varphi)$ are defined in Corollary 3.2, $\eta_1(\mathbf{z}) = W_1(\mathbf{z}, \mathbf{I}, \mathbf{I})$ and

$$\begin{aligned}
\pi_{jm} &= \left[\sum_{\ell=1}^k \tau_\ell \frac{(\lambda_{\ell m} - \lambda_{\ell j})^2}{\lambda_{\ell j}\lambda_{\ell m}}\right]^{-1} \\
c_{i,11} &= E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right) IF(\mathbf{z}, \mathbf{V}_{i,o}, G)_{11}\right) = E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right)\left[\alpha^{(i)}\left(\|\mathbf{z}\|\right) z_1^2 - \gamma^{(i)}\left(\|\mathbf{z}\|\right)\right]\right) \\
c_{i,12} &= E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right) IF(\mathbf{z}, \mathbf{V}_{i,o}, G)_{22}\right) = E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right)\left[\alpha^{(i)}\left(\|\mathbf{z}\|\right) z_2^2 - \gamma^{(i)}\left(\|\mathbf{z}\|\right)\right]\right) \\
d_{i,12} &= E\left(\eta_1(\mathbf{z}) z_1\, z_2 IF(\mathbf{z}, \mathbf{V}_{i,o}, G)_{12}\right) = E\left(\eta_1(\mathbf{z})\alpha^{(i)}\left(\|\mathbf{z}\|\right) z_1^2\, z_2^2\right) .
\end{aligned}$$

Note that if the same family of scatter matrices is considered for all populations, then $\text{asvar}\left(\widehat{\beta}_{jm}\right) = \pi_{jm}\, \kappa^2\, D^{-2}(G)\left\{a_{12}\left(\eta_1^2\right) + \alpha_3^2(G)\sigma_{1,1} - 2\alpha_3(G)d_{1,12}\right\}$ which entails that the efficiency of $\widehat{\beta}_{jm}$ is given by $\kappa^2\, D^{-2}(G)\left\{a_{12}\left(\eta_1^2\right) + \alpha_3^2(G)\sigma_{1,1} - 2\alpha_3(G)d_{1,12}\right\}$.

4.2 Initial Projection–Pursuit Estimators

We will now consider as initial estimates the projection–pursuit estimates defined in (3) through an univariate scale estimator s. Let G_0 be the distribution of z_{11}, where it will be assumed that $\mathbf{\Sigma}_i^{-\frac{1}{2}} \mathbf{x}_{i1} = \mathbf{z}_i$ have the same spherical distribution G for all $1 \le i \le k$. Boente, Pires and Rodrigues (2002) have shown that if the function $(\epsilon, y) \to \sigma\left\{(1-\epsilon)G_0 + \epsilon\delta_y\right\}$ is twice continuously differentiable at $(0, y)$, the partial influence functions of the eigenvalues and eigenvectors are given by

$$\text{PIF}_i(\mathbf{x}, \lambda_{\text{I},\ell j}, F) = 2\,\delta_{\ell i}\,\lambda_{ij}\,\text{IF}\left(\frac{\mathbf{x}^{\text{T}}\boldsymbol{\beta}_j}{\lambda_{ij}^{\frac{1}{2}}}, \sigma, G_0\right) \tag{25}$$

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I},j},F) &= \tau_i\lambda_{ij}^{\frac{1}{2}}\ \mathrm{DIF}\left(\frac{\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_j}{\lambda_{ij}^{\frac{1}{2}}},\sigma,G_0\right)\sum_{s=j+1}^{p}\frac{1}{\nu_j-\nu_s}\boldsymbol{\beta}_s\left(\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s\right)+\\ &+ \tau_i\sum_{s=1}^{j-1}\frac{1}{\nu_j-\nu_s}\boldsymbol{\beta}_s\lambda_{is}^{\frac{1}{2}}\ \mathrm{DIF}\left(\frac{\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s}{\lambda_{is}^{\frac{1}{2}}},\sigma,G_0\right)\left(\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_j\right)\ , \qquad (26)\end{aligned}$$

where $\mathrm{DIF}(y,\sigma,G)$ denotes the derivative of the influence function of the scale functional σ, $\mathrm{IF}(y,\sigma,G)$, with respect to y. The same authors have also shown that the corresponding asymptotic variances are given by

$$\begin{aligned}\mathrm{asvar}\left(\widehat{\lambda}_{\mathrm{I},\,ij}\right) &= \frac{1}{\tau_i}\,\lambda_{ij}^2\,\sigma_{11}\\ \mathrm{ascov}\left(\widehat{\lambda}_{\mathrm{I},ij},\widehat{\lambda}_{\mathrm{I},im}\right) &= \frac{1}{\tau_i}\,\lambda_{ij}\lambda_{im}\,\sigma_{12}\\ \mathrm{asvar}\left(\widehat{\beta}_{\mathrm{I},jm}\right) &= \frac{\sum_{i=1}^{k}\tau_i\lambda_{ij}\lambda_{im}}{(\nu_j-\nu_m)^2}\,\upsilon_{12}\ , \qquad \text{for } m\neq j\ ,\end{aligned}$$

with $\sigma_{11}=4\,\mathrm{asvar}\,(\sigma,G_0)$, $\sigma_{12}=4\,\mathrm{cov}_G\,(\mathrm{IF}(z_1,\sigma,G_0),\mathrm{IF}(z_2,\sigma,G_0))$, $\upsilon_{12}=E_G\left\{[\mathrm{DIF}\,(z_1,\sigma,G_0)\,z_2]^2\right\}$ and $\nu_j=\sum_{i=1}^{k}\tau_i\lambda_{ij}$.

On the other hand, if, in addition, f is twice continuously differentiable and we consider the general projection–pursuit estimates defined through (4), we have that

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\lambda_{\mathrm{I},\,\ell j},F) &= 2\,\delta_{\ell i}\,\lambda_{ij}\ \mathrm{IF}\left(\frac{\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}},\sigma,G_0\right) \qquad (27)\\ \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I},j},F) &= \tau_i\ \boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}\sum_{s=1}^{j-1}\frac{1}{\nu_{sj}-\nu_{ss}}\sqrt{\lambda_{is}}f'(\lambda_{is})\ \mathrm{DIF}\left(\frac{\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}}{\sqrt{\lambda_{is}}},\sigma,G_0\right)\ \boldsymbol{\beta}_s+\\ &+ \tau_i\sqrt{\lambda_{ij}}f'(\lambda_{ij})\,\mathrm{DIF}\left(\frac{\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}}{\sqrt{\lambda_{ij}}},\sigma,G_0\right)\sum_{s=j+1}^{p}\frac{1}{\nu_{jj}-\nu_{js}}\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\,\boldsymbol{\beta}_s \qquad (28)\end{aligned}$$

where $\nu_{js}=\sum_{i=1}^{k}\tau_i f'(\lambda_{ij})\,\lambda_{is}$ and $\nu_{js}\neq\nu_{jj}$ for $s\neq j$ (see, Boente, Pires and Rodrigues, 2005). Thus, the partial influence function of the eigenvalues is independent of the score function f. Moreover, the asymptotic variances of the eigenvectors are given by

$$\mathrm{ASVAR}\left(\widehat{\beta}_{jm}\right) = \sum_{i=1}^{k}\tau_i\lambda_{ij}\lambda_{im}\left\{\frac{\delta_{m>j}\,\{f'(\lambda_{ij})\}^2}{(\nu_{jj}-\nu_{jm})^2}+\frac{\delta_{m<j}\,\{f'(\lambda_{im})\}^2}{(\nu_{mj}-\nu_{mm})^2}\right\}\upsilon_{12}.$$

Straightforward calculations lead to the asymptotic variances of the proposed estimates with initial projection–pursuit estimates given in the following Theorem. This result can be applied only when the initial projection–pursuit functionals are Fisher–consistent. Conditions under which this assumption holds are given in Boente, Pires and Rodrigues (2005).

Theorem 4.2. *Let $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}$, $1 \le i \le k$, be independent observations from k independent samples with scatter matrix $\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}_i = \text{diag}(\lambda_{i1}, \ldots, \lambda_{ip})$ such that $\boldsymbol{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{i1} = \mathbf{z}_i$ have the same spherical distribution G for all $1 \le i \le k$. Assume that $\lambda_{11} > \ldots > \lambda_{1p}$, $n_i = \tau_i N$ with $0 < \tau_i < 1$ fixed numbers such that $\sum_{i=1}^k \tau_i = 1$. Let $s(\cdot)$ be a univariate robust scale statistic related to the functional $\sigma(F)$ and assume that $\sigma(G_0) = 1$, where G_0 is the distribution of z_{11}. Moreover, assume that the function $(\epsilon, y) \to \sigma\left\{(1-\epsilon)G_0 + \epsilon\delta_y\right\}$ is twice continuously differentiable at $(0, y)$.*
Let κ, $\alpha_1(G)$, $\alpha_2(G)$, $\alpha_3(G)$ and $a_{12}(\varphi)$ be defined in Corollary 3.2, $\eta_1(\mathbf{z}) = W_1(\mathbf{z}, \mathbf{I}, \mathbf{I})$ and

$$\begin{aligned}
\pi_{jm} &= \left[\sum_{\ell=1}^k \tau_\ell \frac{(\lambda_{\ell m} - \lambda_{\ell j})^2}{\lambda_{\ell j}\lambda_{\ell m}}\right]^{-1} \\
c_{\sigma,11} &= E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right) IF(z_1, \sigma, G_0)\right) \\
c_{\sigma,12} &= E\left(\eta_1(\mathbf{z})\left(\kappa\, z_1^2 - 1\right) IF(z_2, \sigma, G_0)\right) \\
d_{\sigma,12} &= E\left(\eta_1(\mathbf{z}) z_1^2\, z_2 DIF(z_2, \sigma, G_0)\right) .
\end{aligned}$$

a) Denote by $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$ the projection–pursuit estimates obtained through (3) or (4) and by $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\lambda}}$ the final estimates computed from the reweighted matrix defined in (10) using $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$. Then, we have that

$$\begin{aligned}
\text{asvar}\left(\widehat{\lambda}_{ij}\right) &= \frac{\lambda_{ij}^2}{\tau_i\, D^2(G)}\left\{E_G\left(\eta_1^2(\mathbf{z})\left(\kappa\, z_1^2 - 1\right)^2\right) + \alpha_1^2(G)\sigma_{11} + 2(p-1)\alpha_1(G)\alpha_2(G)\sigma_{12} + \right. \\
&\quad \left. +(p-1)\alpha_2^2(G)\left(\sigma_{11} + (p-1)\sigma_{12}\right) - 4\alpha_1(G)c_{\sigma,11} - 4(p-1)\alpha_2(G)c_{\sigma,12}\right\}
\end{aligned}$$

b) When $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$ are the projection–pursuit estimates obtained from s through (3) and $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\lambda}}$ are the final estimates computed from the reweighted matrix defined in (10) using $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$, if $\nu_1 > \nu_2 > \ldots > \nu_p$, with $\nu_j = \sum_{i=1}^k \tau_i \lambda_{ij}$, we have that

$$\text{asvar}\left(\widehat{\beta}_{jm}\right) = \frac{\kappa^2\, \pi_{jm}}{D^2(G)}\left\{a_{12}\left(\eta_1^2\right) - 2\alpha_3(G)d_{\sigma,12} + \alpha_3^2(G)\, \upsilon_{12}\pi_{jm}^{-1}\left[\sum_{i=1}^k \tau_i \frac{\lambda_{ij}\lambda_{im}}{(\nu_j - \nu_m)^2}\right]\right\}$$

c) When $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$ are the projection–pursuit estimates obtained from s through (4) and $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\lambda}}$ are the final estimates computed from the reweighted matrix defined in (10) using $\widehat{\boldsymbol{\beta}}_{\rm I}$ and $\widehat{\boldsymbol{\lambda}}_{\rm I}$, if $\nu_{js} \neq \nu_{jj}$ for $s \neq j$, with $\nu_{js} = \sum_{i=1}^k \tau_i f'\left(\lambda_{ij}\right)\lambda_{is}$, we have that

$$\begin{aligned}
\text{asvar}\left(\widehat{\beta}_{jm}\right) &= \frac{\kappa^2\, \pi_{jm}}{D^2(G)}\left\{a_{12}\left(\eta_1^2\right) - 2\alpha_3(G)d_{\sigma,12}\left[\sum_{i=1}^k \tau_i\left(\lambda_{ij} - \lambda_{im}\right)\left\{\frac{\delta_{m>j} f'\left(\lambda_{ij}\right)}{\nu_{jj} - \nu_{jm}} + \frac{\delta_{m<j} f'\left(\lambda_{im}\right)}{\nu_{mj} - \nu_{mm}}\right\}\right] + \right. \\
&\quad \left. + \alpha_3^2(G)\, \upsilon_{12}\pi_{jm}^{-1}\left[\sum_{i=1}^k \tau_i \lambda_{ij}\lambda_{im}\left\{\frac{\delta_{m>j}\left\{f'\left(\lambda_{ij}\right)\right\}^2}{\left(\nu_{jj} - \nu_{jm}\right)^2} + \frac{\delta_{m<j}\left\{f'\left(\lambda_{im}\right)\right\}^2}{\left(\nu_{mj} - \nu_{mm}\right)^2}\right\}\right]\right\} .
\end{aligned}$$

Moreover, when $G = N(\mathbf{0}_p, \mathbf{I}_p)$, $\text{ascov}\left(\widehat{\beta}_{jm}, \widehat{\beta}_{js}\right) = 0$, for $m \neq s$, in both cases.

Note that when $w \equiv 1$, we get the asymptotic variances of the classical estimates, showing that by calibrating the tuning constant of the weight function, we can obtain more efficient estimates of the common directions, as mentioned above and showed in the simulation study discussed in Section 5.

5 Monte Carlo Study

We performed a simulation study in dimension 4 and another one in dimension 2. In both situations, we evaluated the estimators defined in Section 2 using as initial estimates the projection–pursuit estimators based on $f(t) = \ln(t)$ and the plug–in estimates based on the Donoho–Stahel scatter matrices. We have also computed the scatter reweighted estimates based on the Mahalanobis distance evaluated with the Donoho–Stahel scatter matrices and the plug–in estimates derived from them, which will be denoted by WPI_2. For the reweighted estimates based on the Mahalanobis distance, we chose as weight function $w_{\mathrm{MD}}(t)$

$$w_{\mathrm{MD}}(t) = \begin{cases} 1 & t < a_M \\ \exp\left[-\dfrac{(t-a_M)^2}{2c_M^2}\right] & t \geq a_M\,, \end{cases}$$

where $a_M^2 = \chi^2_{p,0.95}$ and $b_M^2 = \chi^2_{p,0.975}$ are the percentiles 0.95 and 0.975 of a χ^2 distribution with p degrees of freedom, respectively, and $c_M = \dfrac{b_M - a_M}{p+1}$.

For the projection–pursuit estimator, an $M-$scale estimator with score function $\chi(t) = \min\left(\frac{t^2}{c^2}, 1\right) - \frac{1}{2}$ and $c = 1.041$ was used, while for the reweighted estimates we have considered as weight functions

$$\begin{aligned} w_{\mathrm{m}}\left(IML^2(\mathbf{x}_{ij}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Lambda}}_i), IMB^2(\mathbf{x}_{ij}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Lambda}}_i)\right) &= \min(w_{\mathrm{IML}}(\mathbf{x}_{ij}), w_{\mathrm{IMB}}(\mathbf{x}_{ij})) \\ w_{\mathrm{p}}\left(IML^2(\mathbf{x}_{ij}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Lambda}}_i), IMB^2(\mathbf{x}_{ij}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\Lambda}}_i)\right) &= w_{\mathrm{IML}}(\mathbf{x}_{ij}) \times w_{\mathrm{IMB}}(\mathbf{x}_{ij})\,, \end{aligned}$$

where w_{IML} and w_{IMB} are defined in (21). In all Tables LPP will denote the projection–pursuit estimates while, in Tables 2 and 3, $\mathrm{W_m}$LPP and $\mathrm{W_p}$LPP correspond to the reweighted estimates related to them using the function w_{m} and w_{p}, respectively. Similarly, PI_1, $\mathrm{W_m PI}_1$ and $\mathrm{W_p PI}_1$ denote the plug–in estimates obtained using the Donoho–Stahel matrix and the two reweighted estimates related to them, respectively. As it will be shown for the estimation of the eigenvectors, the results for both weight functions are similar. Therefore, when reporting all the results in dimension 2 and for the remaining tables in dimension 4, the notation WLPP and WPI_1 will be used to indicate the estimators related to w_{m}.

5.1 Simulation Conditions in Dimension 4

In dimension $p = 4$, we have considered $k = 2$ populations with $\boldsymbol{\Sigma}_1 = \mathrm{diag}(16, 8, 2, 1)$ and $\boldsymbol{\Sigma}_2 = 4\,\boldsymbol{\Sigma}_1$.

In all models, we performed 500 replications generating k independent samples of size $n_i = n = 100$. The eigenvectors were ordered according to a decreasing order of the eigenvalues of the first population and so, $\boldsymbol{\beta}_j = \mathbf{e}_j$.

The results for normal data sets will be indicated by C_0 in the Tables, while C_1 and C_2 will denote the following two contaminations.

- C_1: $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in}$ are i.i.d. $(1-\epsilon)N(\mathbf{0}, \boldsymbol{\Sigma}_i) + \epsilon N(\boldsymbol{\mu}, \boldsymbol{\Sigma}_i)$ with $\boldsymbol{\mu} = 10\,\mathbf{e}_4 = (0,0,0,10)^{\mathrm{T}}$. We present the results for $\epsilon = 0.10$. This case corresponds to contaminating both populations in the direction of the smallest eigenvalue. The aim is to study changes in the estimation of the principal directions.

- C_2: $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. $(1-\epsilon)N(\mathbf{0}, \boldsymbol{\Sigma}) + \frac{\epsilon}{2}N(\boldsymbol{\mu}, 0.01\boldsymbol{\Sigma}) + \frac{\epsilon}{2}N(-\boldsymbol{\mu}, 0.01\boldsymbol{\Sigma})$ where $\mathbf{x}_j = \left(\mathbf{x}_{1j}^{\mathrm{T}}, \mathbf{x}_{2j}^{\mathrm{T}}\right)^{\mathrm{T}}$, $\boldsymbol{\Sigma} = \mathrm{diag}\,(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}}\right)^{\mathrm{T}}$ and $\boldsymbol{\mu}_i = \boldsymbol{\Sigma}_i^{\frac{1}{2}}\sqrt{7}\,\mathbf{e}_4$. We present the results for $\epsilon = 0.10$. This case corresponds to contaminating both populations in the direction of the smallest eigenvalue with a mild outlier. The aim is to study changes in the estimation of the principal directions between the outlier detection measures we propose to use in the reweighting step and the Mahalanobis distance.

For simplicity, we present only the results corresponding to the common eigenvectors and to the eigenvalues of the first population. The projection–pursuit estimators were computed as in Boente and Orellana (2001).

Tables 2 and 3 give respectively the means and medians, over the replications, of the square distance between the j-th estimated and target eigenvector, $\|\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j\|^2$. This measure determines the angle $\widehat{\theta}_j$ between the j-th estimated and true direction by $\cos(\widehat{\theta}_j) = 1 - \|\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j\|^2/2$.

Table 4 gives summary measures for the eigenvalue estimates. Mean values, standard deviations and mean square errors of $\log\left(\frac{\widehat{\lambda}_j}{\lambda_j}\right)$ are reported.

5.2 Simulation in Dimension 2

We performed a simulation under the CPC model, in dimension $p = 2$ with $n_1 = n_2 = n = 50$ and 500 replications, in order to evaluate the behavior of the estimators considered above in more detail. In dimension 2, it is possible to compute the projection–pursuit estimates without using resampling techniques by maximizing over 1000 fixed and equally spaced directions. The common direction estimates were ordered according to decreasing values of the eigenvalues.

To avoid consistency problems with the projection–pursuit estimates, we have considered, as in Rodrigues (2003), the covariance matrices $\boldsymbol{\Sigma}_1 = \mathrm{diag}(14, 4)$ and $\boldsymbol{\Sigma}_2 = \mathrm{diag}(12, 2)$ which have well separated eigenvalues.

We generated normally and contaminated distributed samples, denoted respectively as C_0, $C_{1,\epsilon}$ and C_2 in Tables 5 to 7. The contaminated samples correspond to

- $C_{1,\epsilon}$: $(\mathbf{x}_{1j})_{1\le j\le n}$ i.i.d. with distribution $(1-\epsilon)N(\mathbf{0}, \boldsymbol{\Sigma}_1) + \frac{\epsilon}{2}\delta_{\boldsymbol{\mu}} + \frac{\epsilon}{2}\delta_{-\boldsymbol{\mu}}$ where $\boldsymbol{\mu} = (-0.042; 4.72)^{\mathrm{T}}$ and $(\mathbf{x}_{2j})_{1\le j\le n}$ i.i.d. with non–contaminated distribution $N(\mathbf{0}, \boldsymbol{\Sigma}_2)$.

- C_2: $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are i.i.d. $0.8\,N(\mathbf{0}, \boldsymbol{\Sigma}) + 0.1\delta_{\boldsymbol{\mu}} + 0.1\delta_{-\boldsymbol{\mu}}$ where $\mathbf{x}_j = \left(\mathbf{x}_{1j}^{\mathrm{T}}, \mathbf{x}_{2j}^{\mathrm{T}}\right)^{\mathrm{T}}$, $\boldsymbol{\Sigma} = \mathrm{diag}\,(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$, $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_1^{\mathrm{T}}, \boldsymbol{\mu}_2^{\mathrm{T}}\right)^{\mathrm{T}}$ and $\boldsymbol{\mu}_i = \boldsymbol{\Sigma}_i^{\frac{1}{2}}(0, 2.4)^{\mathrm{T}}$. Now, both populations are contaminated at the same time.

In this case, also the mean and the median of square distance between $\|\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1\|^2$ are considerably different, showing that in some replications large values are obtained, as shown in Table 5. Therefore, we also report in Table 6 the number of times that the absolute value of the angle between $\boldsymbol{\beta}_1$ and $\widehat{\boldsymbol{\beta}}_1$ is greater than 45 (N_{45}), 60 (N_{60}) and 80 (N_{80}) degrees. Table 7 gives summary measures for the eigenvalue estimates. Mean values, standard deviations and mean square errors of $\log\left(\frac{\widehat{\lambda}_j}{\lambda_j}\right)$ are reported.

5.3 Results

5.3.1 Behavior of the eigenvector estimates

In dimension 4, under normal data, the weighting procedure leads to direction estimates with smaller median and mean of square distance, when the initial estimate is the projection–pursuit one, see Tables 2 and 3. This reduction is of order 60% for the median square error and 50% for the mean square error. When the initial estimates are the plug–in ones, the reweighted–based estimates have median square error of the same order, while mean square errors are now about twice those of the initial estimates and are quite similar to those of the Mahalanobis reweighted estimates. All the reweighted procedures have similar mean and median square errors, showing that the reweighting step improves the efficiency of the projection–pursuit estimators.

Similar results are observed in dimension 2, with respect to the median square error (Table 5), where the reduction is of order 75% for the projection–pursuit estimates and only a small reduction is obtained for the plug–in ones. With respect to the mean square error, we observe now a small reduction. Indeed, the weighted estimates based on the projection–pursuit show their advantage with respect to the two other competitors with a much smaller mean and median square error.

Under contamination, in dimension 4, the weighting procedure performs better than the initial estimates under C_1 which is a contamination with an extreme outlier in the direction of the smaller eigenvalue. All reweighted estimators show median square errors of the same order under C_1, while higher mean square errors are observed for $\mathrm{W_m}$LPP and $\mathrm{W_p}$LPP. Under C_2, which is a contamination with a mild outlier, difficult to detect, an improvement is observed in the directions of the two largest eigenvalues when the initial estimate is the projection–pursuit one, while similar square errors are obtained for the two smaller ones. The results for the other reweighted procedures are slightly worse than for the plug–in estimate based on the Donoho-Stahel scatter matrices. This shows that this mild contamination produces a rotation in the principal axis related to the two smallest eigenvalues, which is worse than the extreme contamination C_1, even if it does not produce breakdown. Since both weight functions lead to similar performance, from now on, we only report the results corresponding to w_{m}.

In dimension 2, the best performance under contamination is attained with the reweighted estimates based on the projection–pursuit estimators, in all cases, even if some extreme values are obtained, leading to possible breakdown as shown in Table 6. However, these estimates show mean and median square errors which are about a half of those obtained by the initial estimates and which are comparable or smaller than those of the reweighted estimates derived using the Mahalanobis distance as an outlier detection procedure. Note that under $C_{1,0.2}$ and C_2 the plug–in estimates, lead to very large angles in about half of the replications. In this case, the resulting directions will not allow to detect influential observations explaining the bad performance of the reweighted estimates.

In order to explain the similar behavior of the reweighted estimates based on the plug–in estimators and of those based on the Mahalanobis distance, denote $d_2(\mathbf{z}) = \sum_{i=1}^{p} z_i^2$ and $d_4(\mathbf{z}) = \sum_{i=1}^{p} z_i^4$. Then, when

$\mathbf{\Sigma} = \mathbf{I}$, we have

$$\begin{aligned} IML^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) &= d_4(\mathbf{z}) - 2d_2(\mathbf{z}) + 1 \\ IMB^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) &= d_2^2(\mathbf{z}) - d_4(\mathbf{z}) \end{aligned}$$

Thus, if $a_{L,1-\alpha}$ denotes the percentile $1-\alpha$ of $IML(\mathbf{z}, \mathbf{I}, \mathbf{I})$ and $a_{B,1-\alpha}$ that of $IMB(\mathbf{z}, \mathbf{I}, \mathbf{I})$, we have that $IML^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{L,1-\alpha}^2$ and $IMB^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{B,1-\alpha}^2$, entailing that $d_2^2(\mathbf{z}) \le a_{B,1-\alpha}^2 + 2d_2(\mathbf{z}) + a_{L,1-\alpha}^2 - 1$ and so $d_2(\mathbf{z}) \le 1 + \sqrt{a_{B,1-\alpha}^2 + a_{L,1-\alpha}^2}$.

On the other hand, if $a_{M,1-\alpha}$ denotes the percentile $1-\alpha$ of the Mahalanobis distance and $d_2(\mathbf{z}) \le a_{M,1-\alpha}^2$ we have that $IMB^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{M,1-\alpha}^4$ and $IML^2(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le d_2^2(\mathbf{z}) - 2d_2(\mathbf{z}) + 1 = (d_2(\mathbf{z}) - 1)^2 \le \max\left(a_{M,1-\alpha}^2 - 1, 1\right)$ using that $d_4(\mathbf{z}) \le d_2^2(\mathbf{z})$. These values are reported in Table 1, when $\alpha = 0.05$ and $p = 2, 4$.

Figure 4 illustrates that, in dimension 2, if $IML(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{L,0.95}$ and $IMB(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{B,0.95}$ then $d_2(\mathbf{z}) \le a_{M,0.95}^2$. On the other hand, if $d_2(\mathbf{z}) \le a_{M,0.95}^2$ then $IML(\mathbf{z}, \mathbf{I}, \mathbf{I}) \le a_{L,0.975}$. These inequalities explain the similar behavior observed under the selected contaminations for the reweighted estimators.

5.3.2 Behavior of the eigenvalue estimates

With respect to the estimation of the eigenvalues, all estimates present similar biases under normal data, except LPP which shows a quite smaller one specially for larger eigenvalues. With respect to the standard deviation, as expected, reweighting improves the efficiency of the projection–pursuit estimates. All reweighting procedures show similar efficiencies in dimension 4.

In dimension 4, under contamination, the reweighted estimates perform similarly as their initial ones, for the larger eigenvalues. The same happens in dimension 2, except for the LPP which is better than WLPP under C_2. On the other hand, in $p = 4$, reweighting improves both bias and standard deviation for the smaller eigenvalues.

According to the simulation study, the best performance is obtained with the reweighted estimates based on our outlier detection measures with initial projection–pursuit estimators.

6 Conclusions

We have introduced reweighted–based estimators of the paramters under a CPC model, using the outlier detection measures defined by Boente, Pires and Rodrigues (2002) and we have obtained their partial influence functions.

The partial influence functions turn out to be that of the plug–in estimates computed using reweighted scatter matrices with the true parameters corrected by those of the initial estimates used, when all the populations have the same elliptical distribution except for changes on the scatter.

The corresponding asymptotic variances were derived heuristically and allow to calibrate the estimates in order to improve the efficiency of the initial estimates.

Our procedure has performed better under the contaminations considered in the simulation study.

A Appendix

Note that

$$\begin{aligned} IML(\mathbf{x},\mathbf{I},\mathbf{I}) &= \left[\sum_{r=1}^{p} \frac{\{x_r^2-1\}^2}{2}\right]^{\frac{1}{2}} \\ IMB(\mathbf{x},\mathbf{I},\mathbf{I}) &= \left[\sum_{r=1}^{p}\sum_{s\neq r} x_r^2\, x_s^2\right]^{\frac{1}{2}} , \end{aligned}$$

and that $IML(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}) = IML(\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},\mathbf{I},\mathbf{I})$ and similarly $IMB(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}) = IMB(\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x},\mathbf{I},\mathbf{I})$. Therefore, if $\mathbf{x}$ is spherically distributed, we have that $E\left[W_1\left(\mathbf{x},\mathbf{I},\mathbf{I}\right)x_r x_s\right]=0$, if $r\neq s$.

Proof of Theorem 3.1. The proof follows the same steps as those given in the proof of Theorem 1 in Boente, Pires and Rodrigues (2002).

Let $F_{i,\epsilon,\mathbf{x}} = (1-\epsilon)F_i + \epsilon\delta_{\mathbf{x}}$ and let $F_{\epsilon,\mathbf{x},i} = F_1\times\ldots\times F_{i-1}\times F_{i,\epsilon,\mathbf{x}}\times F_{i+1}\times\ldots\times F_k$. Let $\boldsymbol{\beta}_{j,\epsilon,i} = \boldsymbol{\beta}_{\mathbf{V},j}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\lambda_{\ell j,\epsilon,i} = \lambda_{\mathbf{V},\ell j}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\mathbf{V}_{\ell,\epsilon,i} = \mathbf{V}_{\ell}\left(F_{\epsilon,\mathbf{x},i}\right)$. Then we have that for $1\leq \ell\leq k$, $1\leq m,j\leq p$

$$\begin{aligned} \lambda_{\ell j,\epsilon,i} &= \boldsymbol{\beta}_{j,\epsilon,i}^{\mathrm{T}}\mathbf{V}_{\ell,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i} && \text{(A.1)}\\ 0 &= \boldsymbol{\beta}_{m,\epsilon,i}^{\mathrm{T}}\left(\sum_{\ell=1}^{k}\tau_\ell\frac{\lambda_{\ell m,\epsilon,i}-\lambda_{\ell j,\epsilon,i}}{\lambda_{\ell m,\epsilon,i}\,\lambda_{\ell j,\epsilon,i}}\mathbf{V}_{\ell,\epsilon,i}\right)\boldsymbol{\beta}_{j,\epsilon,i} \qquad m\neq j && \text{(A.2)}\\ \delta_{mj} &= \boldsymbol{\beta}_{m,\epsilon,i}^{\mathrm{T}}\,\boldsymbol{\beta}_{j,\epsilon,i}\,. && \text{(A.3)} \end{aligned}$$

Therefore, differentiating (A.3) with respect to ϵ, we get that

$$\begin{aligned} \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},m},F)^{\mathrm{T}}\boldsymbol{\beta}_m &= 0 && \text{(A.4)}\\ \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},m},F)^{\mathrm{T}}\boldsymbol{\beta}_j + \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F)^{\mathrm{T}}\boldsymbol{\beta}_m &= 0\,. && \text{(A.5)} \end{aligned}$$

- *Partial influence functions for the eigenvalues.*

Differentiating (A.1), we get

$$\begin{aligned} \mathrm{PIF}_i(\mathbf{x},\lambda_{\mathbf{V},\ell j},F) &= 2\lambda_{\ell j}\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F)^{\mathrm{T}}\boldsymbol{\beta}_j + \boldsymbol{\beta}_j^{\mathrm{T}}\mathrm{PIF}_i(\mathbf{x},\mathbf{V}_\ell,F)\boldsymbol{\beta}_j\\ &= \boldsymbol{\beta}_j^{\mathrm{T}}\mathrm{PIF}_i(\mathbf{x},\mathbf{V}_\ell,F)\boldsymbol{\beta}_j\,, \end{aligned}$$

since (A.4) holds and $\mathbf{V}_\ell\boldsymbol{\beta}_j = \boldsymbol{\Sigma}_\ell\boldsymbol{\beta}_j = \lambda_{\ell j}\boldsymbol{\beta}_j$, which entails (17).

- *Partial influence functions for the eigenvectors.*

It remains to show (18). Differentiating (A.2) leads to

$$\begin{aligned} 0 \;=\; &\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},m},F)^{\mathrm{T}}\left(\sum_{\ell=1}^{k}\tau_\ell\frac{\lambda_{\ell m}-\lambda_{\ell j}}{\lambda_{\ell m}\lambda_{\ell j}}\mathbf{V}_\ell\right)\boldsymbol{\beta}_j + \boldsymbol{\beta}_m^{\mathrm{T}}\left(\sum_{\ell=1}^{k}\tau_\ell\frac{\lambda_{\ell m}-\lambda_{\ell j}}{\lambda_{\ell m}\lambda_{\ell j}}\mathbf{V}_\ell\right)\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F) +\\ &+\boldsymbol{\beta}_m^{\mathrm{T}}\left\{\sum_{\ell=1}^{k}\tau_\ell\frac{\partial}{\partial\epsilon}\left(\frac{\lambda_{\ell m,\epsilon,i}-\lambda_{\ell j,\epsilon,i}}{\lambda_{\ell m,\epsilon,i}\lambda_{\ell j,\epsilon,i}}\right)\Big|_{\epsilon=0}\mathbf{V}_\ell + \sum_{\ell=1}^{k}\tau_\ell\frac{\lambda_{\ell m}-\lambda_{\ell j}}{\lambda_{\ell m}\lambda_{\ell j}}\mathrm{PIF}_i(\mathbf{x},\mathbf{V}_\ell,F)\right\}\boldsymbol{\beta}_j \;\text{ for } m\neq j\,. \qquad \text{(A.6)} \end{aligned}$$

Using again the fact that $\mathbf{\Sigma}_\ell\boldsymbol{\beta}_j=\lambda_{\ell j}\boldsymbol{\beta}_j$ in (A.6), the orthogonality condition $\boldsymbol{\beta}_m^{\mathrm{T}}\boldsymbol{\beta}_j=0$ for $m\neq j$ and (A.5), we obtain, after some algebra,

$$\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F)^{\mathrm{T}}\boldsymbol{\beta}_m=\left\{\sum_{\ell=1}^{k}\tau_\ell\frac{(\lambda_{\ell m}-\lambda_{\ell j})^2}{\lambda_{\ell m}\lambda_{\ell j}}\right\}^{-1}\sum_{\ell=1}^{k}\tau_\ell\,\frac{\lambda_{\ell j}-\lambda_{\ell m}}{\lambda_{\ell m}\lambda_{\ell j}}\left\{\boldsymbol{\beta}_j^{\mathrm{T}}\mathrm{PIF}_i\left(\mathbf{x},\mathbf{V}_\ell,F\right)\boldsymbol{\beta}_m\right\}\quad\text{for } m\neq j\ .\ \square$$

Lemma A.3. *Denote by $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_p$, $\lambda_{i1},\dots,\lambda_{ip}$ the common eigenvectors and the eigenvalues of $\mathbf{\Sigma}_i$. Let $\boldsymbol{\beta}_{\mathrm{I}}(F)$ and $\mathbf{\Lambda}_{\mathrm{I},\,i}(F)$ be Fisher–consistent functionals related to the initial estimates of the common eigenvectors and of the eigenvalues of the $i-$th population. Assume that the partial influence functions $PIF_i\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},F\right)$ and $PIF_i\left(\mathbf{x},\mathbf{\Lambda}_{\mathrm{I},\,\ell},F\right)$ exists and satisfy $PIF_i\left(\mathbf{x},\mathbf{\Lambda}_{\mathrm{I},\,\ell},F\right)=0$ for $\ell\neq i$. Then, the partial influence functions of $IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F),\mathbf{\Lambda}_{\mathrm{I},\,\ell}(F))$ and $IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F),\mathbf{\Lambda}_{\mathrm{I},\,\ell}(F))$ defined through (8) and (9), are given by*

$$\begin{aligned}
PIF_i\left(\mathbf{x}_o,IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\mathbf{\Lambda}_{\mathrm{I},\,\ell}),F\right) &= 2\sum_{r=1}^{p}\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{\ell r}\right\}}{\lambda_{\ell r}^2}\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\quad \text{for } \ell\neq i\\
PIF_i\left(\mathbf{x}_o,IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\mathbf{\Lambda}_{\mathrm{I},\,i}),F\right) &= \sum_{r=1}^{p}\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{ir}\right\}}{\lambda_{ir}^3}\left[2\,\lambda_{ir}\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}-\right.\\
&\quad\left.-\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2 PIF_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,ir},F\right)\right]\\
PIF_i\left(\mathbf{x}_o,IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\mathbf{\Lambda}_{\mathrm{I},\,\ell}),F\right) &= 2\sum_{r=1}^{p}\sum_{s\neq r}\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\left[\left(PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)+\right.\\
&\quad\left.+\left(PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\right]\quad \text{for } \ell\neq i\\
PIF_i\left(\mathbf{x}_o,IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\mathbf{\Lambda}_{\mathrm{I},\,i}),F\right) &= 2\sum_{r=1}^{p}\sum_{s\neq r}\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left[\left(PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)+\right.\\
&\quad\left.+\left(PIF_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\right]-\\
&\quad-\sum_{r=1}^{p}\sum_{s\neq r}\left[\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\right]^2\left[PIF_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+\right.\\
&\quad\left.+PIF_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right]\ .
\end{aligned}$$

Proof. Let $F_{i,\epsilon,\mathbf{x}_o}=(1-\epsilon)F_i+\epsilon\delta_{\mathbf{x}_o}$ and let $F_{\epsilon,\mathbf{x}_o,i}=F_1\times\ldots\times F_{i-1}\times F_{i,\epsilon,\mathbf{x}_o}\times F_{i+1}\times\ldots\times F_k$. Let $\boldsymbol{\beta}_{j,\epsilon,i}=\boldsymbol{\beta}_{\mathrm{I},\,j}\left(F_{\epsilon,\mathbf{x}_o,i}\right)$, $\lambda_{\ell j,\epsilon,i}=\lambda_{\mathrm{I},\,\ell j}\left(F_{\epsilon,\mathbf{x}_o,i}\right)$, $IMB^2_{\ell,\epsilon,i}(\mathbf{x})=IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}((F_{\epsilon,\mathbf{x}_o,i})),\mathbf{\Lambda}_{\mathrm{I},\,\ell}((F_{\epsilon,\mathbf{x}_o,i})))$ and $IML^2_{\ell,\epsilon,i}(\mathbf{x})=IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}((F_{\epsilon,\mathbf{x}_o,i})),\mathbf{\Lambda}_{\mathrm{I},\,\ell}((F_{\epsilon,\mathbf{x}_o,i})))$. Then we have that for $1\leq\ell\leq k$

$$IML^2_{\ell,\epsilon,i}(\mathbf{x})\ =\ \frac{1}{2}\sum_{r=1}^{p}\left\{\frac{\left(\boldsymbol{\beta}_{r,\epsilon,i}^{\mathrm{T}}\mathbf{x}\right)^2}{\lambda_{\ell r,\,\epsilon,i}}-1\right\}^2$$

$$IMB^2_{\ell,\epsilon,i}(\mathbf{x}) = \sum_{r=1}^{p}\sum_{s\neq r}\frac{\left\{\left(\boldsymbol{\beta}_{r,\epsilon,i}^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_{s,\epsilon,i}^{\mathrm{T}}\mathbf{x}\right)\right\}^2}{\lambda_{\ell r,\epsilon,i}\lambda_{\ell s,\epsilon,i}}\,.$$

Differentiating with respect to ϵ, we get

$$\begin{aligned}
\mathrm{PIF}_i\left(\mathbf{x}_o, IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\ell}),F\right) &= \sum_{r=1}^{p}\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{\ell r}\right\}}{\lambda_{\ell r}^3}\left[2\,\lambda_{\ell r}\,\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},r},F\right)^{\mathrm{T}}\mathbf{x}-\right.\\
&\quad\left.-\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\ell r},F\right)\right]\\
\mathrm{PIF}_i\left(\mathbf{x}_o, IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\ell}),F\right) &= \sum_{r=1}^{p}\sum_{s\neq r}2\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\left[\left(\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},r},F\right)^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)+\right.\\
&\quad\left.+\left(\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},s},F\right)^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\right]-\\
&\quad-\left[\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\,\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\right]^2\left[\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\ell r},F\right)\lambda_{\ell s}+\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\ell s},F\right)\lambda_{\ell r}\right]
\end{aligned}$$

Using that $\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\Lambda}_{\mathrm{I},\ell},F\right)=0$, we conclude the proof. □

Proof of Theorem 3.2. Let $F_{i,\epsilon,\mathbf{x}_o}=(1-\epsilon)F_i+\epsilon\delta_{\mathbf{x}_o}$ and let $F_{\epsilon,\mathbf{x}_o,i}=F_1\times\ldots\times F_{i-1}\times F_{i,\epsilon,\mathbf{x}_o}\times F_{i+1}\times\ldots\times F_k$ and denote $\mathbf{V}_{\ell,\epsilon,i}=\mathbf{V}_\ell\left(F_{\epsilon,\mathbf{x}_o,i}\right)$. Moreover, for $\ell\neq i$ define

$$\begin{aligned}
\mathbf{N}_{\ell,\epsilon,i}=\mathbf{N}_\ell(F_{\epsilon,\mathbf{x}_o,i}) &= E_{F_\ell}\Psi_1\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\ell}(F_{\epsilon,\mathbf{x}_o,i})\right)\\
\mathbf{N}_{i,\epsilon,i}=\mathbf{N}_i(F_{\epsilon,\mathbf{x}_o,i}) &= E_{F_{i,\epsilon,\mathbf{x}_o}}\Psi_1\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},i}(F_{\epsilon,\mathbf{x}_o,i})\right)\\
D_{\ell,\epsilon,i}=D_\ell(F_{\epsilon,\mathbf{x}_o,i}) &= E_{F_\ell}W_1\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\ell}(F_{\epsilon,\mathbf{x}_o,i})\right)\\
D_{i,\epsilon,i}=D_i(F_{\epsilon,\mathbf{x}_o,i}) &= E_{F_{i,\epsilon,\mathbf{x}_o}}W_1\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},i}(F_{\epsilon,\mathbf{x}_o,i})\right)\ ,
\end{aligned}$$

where W_1 and Ψ_1 are defined in (11) and (12) respectively. Since $\mathbf{V}_{\ell,\epsilon,i}=\kappa_i\,D_{\ell,\epsilon,i}^{-1}\mathbf{N}_{\ell,\epsilon,i}$ we get

$$\begin{aligned}
\left.\frac{\partial\mathbf{V}_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0} &= \kappa_i\,D_\ell^{-1}(F)\left.\frac{\partial\mathbf{N}_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0}-\kappa_i\,D_\ell^{-2}(F)\left.\frac{\partial D_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0}\mathbf{N}_\ell(F)\\
&= \kappa_i\,D_\ell^{-1}(F)\left.\frac{\partial\mathbf{N}_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0}-D_\ell^{-1}(F)\left.\frac{\partial D_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0}\mathbf{V}_\ell(F)\ .
\end{aligned}$$

We begin by considering the case $\ell\neq i$. Differentiating $\mathbf{N}_{\ell,\epsilon,i}$ and $D_{\ell,\epsilon,i}$ with respect to ϵ, we obtain

$$\begin{aligned}
\left.\frac{\partial D_{\ell,\epsilon,i}}{\partial\epsilon}\right|_{\epsilon=0} &= E_{F_\ell}\left.\frac{\partial w\left(IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\ell}(F_{\epsilon,\mathbf{x}_o,i})),IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\ell}(F_{\epsilon,\mathbf{x}_o,i}))\right)}{\partial\epsilon}\right|_{\epsilon=0}\\
&= E_{F_\ell}\left\{w_1\left(IML^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_\ell),IMB^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_\ell)\right)\mathrm{PIF}_i\left(\mathbf{x}_o,IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\ell}),F\right)+\right.\\
&\quad\left.+w_2\left(IML^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_\ell),IMB^2(\mathbf{x},\boldsymbol{\beta},\boldsymbol{\Lambda}_\ell)\right)\mathrm{PIF}_i\left(\mathbf{x}_o,IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\ell}),F\right)\right\}\\
&= 2\sum_{r=1}^{p}\boldsymbol{\beta}_r^{\mathrm{T}}E_{F_\ell}\left\{\varphi_1^{(\ell)}(\mathbf{x})\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{\ell r}\right\}}{\lambda_{\ell r}^2}\mathbf{x}\mathbf{x}^{\mathrm{T}}\right\}\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},r},F\right)+
\end{aligned}$$

$$+2\sum_{r=1}^{p}\sum_{s\neq r}E_{F_\ell}\left\{\varphi_2^{(\ell)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\left[\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s+\right.\right.$$
$$\left.\left.+\,\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_r\right]\right\}$$

Similarly, we get

$$\begin{aligned}\frac{\partial\mathbf{N}_{\ell,\epsilon,i}}{\partial\epsilon}\Big|_{\epsilon=0} &= 2\sum_{r=1}^{p}E_{F_\ell}\left\{\varphi_1^{(\ell)}(\mathbf{x})\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{\ell r}\right\}}{\lambda_{\ell r}^2}\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)\mathbf{x}\mathbf{x}^{\mathrm{T}}\right\}+\\ &\quad+2\sum_{r=1}^{p}\sum_{s\neq r}E_{F_\ell}\left\{\varphi_2^{(\ell)}(\mathbf{x})\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{\ell r}\lambda_{\ell s}}\mathbf{x}\mathbf{x}^{\mathrm{T}}\left[\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s+\right.\right.\\ &\quad\left.\left.+\,\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_r\right]\right\}\end{aligned}$$

which entails the desired result.

We consider now the case $\ell = i$. Differentiating $D_{i,\epsilon,i}$ with respect to ϵ, we obtain

$$\begin{aligned}\frac{\partial D_{i,\epsilon,i}}{\partial\epsilon}\Big|_{\epsilon=0} &= \frac{\partial}{\partial\epsilon}\left[(1-\epsilon)E_{F_i}W_1\left(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\,i}(F_{\epsilon,\mathbf{x}_o,i})\right)+\epsilon\,W_1\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I}}(F_{\epsilon,\mathbf{x}_o,i}),\boldsymbol{\Lambda}_{\mathrm{I},\,i}(F_{\epsilon,\mathbf{x}_o,i})\right)\right]\Big|_{\epsilon=0}\\ &= -D_i(F)+W_1(\mathbf{x}_o,\boldsymbol{\beta},\boldsymbol{\Lambda}_i)+E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\mathrm{PIF}_i\left(\mathbf{x}_o,IML^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\,i}),F\right)\right\}+\\ &\quad+E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\mathrm{PIF}_i\left(\mathbf{x}_o,IMB^2(\mathbf{x},\boldsymbol{\beta}_{\mathrm{I}},\boldsymbol{\Lambda}_{\mathrm{I},\,i}),F\right)\right\}\\ &= -D_i(F)+W_1(\mathbf{x}_o,\boldsymbol{\beta},\boldsymbol{\Lambda}_i)+\\ &\quad+\sum_{r=1}^{p}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\frac{\left\{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{ir}\right\}}{\lambda_{ir}^3}\left[2\,\lambda_{ir}\,\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)-\right.\right.\\ &\quad\left.\left.-\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,ir},F\right)\right]\right\}+\\ &\quad+\sum_{r=1}^{p}\sum_{s\neq r}E_{F_i}\varphi_2^{(i)}(\mathbf{x})\left\{2\,\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\left[\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_s+\right.\right.\\ &\quad\left.+\,\mathrm{PIF}_i\left(\mathbf{x}_o,\boldsymbol{\beta}_{\mathrm{I},\,s},F\right)^{\mathrm{T}}\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\boldsymbol{\beta}_r\right]-\\ &\quad\left.-\left[\frac{\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)}{\lambda_{ir}\lambda_{is}}\right]^2\left[\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+\mathrm{PIF}_i\left(\mathbf{x}_o,\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right]\right\}\end{aligned}$$

In an analogous way we can derive $\frac{\partial\mathbf{N}_{\ell,\epsilon,i}}{\partial\epsilon}\Big|_{\epsilon=0}$, which concludes the proof. □

Proof of Corollary 3.2. When F_i are ellipsoidal distributions and $\boldsymbol{\beta}=\mathbf{I}_p$, i.e., $\boldsymbol{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{ij}$ is spherically distributed, for $1\le i\le k$, using that $\boldsymbol{\beta}_r^{\mathrm{T}}\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)=0$, we get that

$$\boldsymbol{\beta}_r^{\mathrm{T}}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{ir}\right]\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right) = 0$$

$$
\begin{aligned}
\boldsymbol{\beta}_r^{\mathrm{T}} E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left(\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}\right)^2\left[\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{ir}\right]\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right) &= 0 \\
\boldsymbol{\beta}_s^{\mathrm{T}} E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right) &= 0 \qquad \text{for} \quad s\neq r \\
\boldsymbol{\beta}_s^{\mathrm{T}} E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left(\boldsymbol{\beta}_j^{\mathrm{T}}\mathbf{x}\right)^2\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right) &= 0 \qquad \text{for} \quad s\neq r \\
\boldsymbol{\beta}_j^{\mathrm{T}} E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2-\lambda_{ir}\right]\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)^2\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\boldsymbol{\beta}_m &= 0 \qquad \text{for} \quad j\neq m \\
\boldsymbol{\beta}_j^{\mathrm{T}} E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left[\left(\boldsymbol{\beta}_r^{\mathrm{T}}\mathbf{x}\right)\left(\boldsymbol{\beta}_s^{\mathrm{T}}\mathbf{x}\right)\right]^2\mathbf{x}\,\mathbf{x}^{\mathrm{T}}\right\}\boldsymbol{\beta}_m &= 0 \qquad \text{for} \quad j\neq m\,.
\end{aligned}
$$

Thus, for $\ell\neq i$, $\Delta_{i,\,\ell}(\mathbf{y},F)=0$ and

$$
\begin{aligned}
\Delta_{i,\,i}(\mathbf{y},F) &= -D_i(F)+W_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)-\sum_{r=1}^{p}\frac{1}{\lambda_{ir}^3}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[x_r^2-\lambda_{ir}\right]x_r^2\right\}\ \mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)- \\
&\quad -\ \sum_{r=1}^{p}\sum_{s\neq r}E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left[\frac{x_r\,x_s}{\lambda_{ir}\lambda_{is}}\right]^2\right\}\left[\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right]
\end{aligned}
$$

On the other hand, since $\mathbf{N}_i(F)_{jm}=0$ for $j\neq m$, we have

$$
\begin{aligned}
\boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{\Upsilon}_{i,\,i}(\mathbf{y},F)\boldsymbol{\beta}_j &= \left(\Psi_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)-\mathbf{N}_i(F)\right)_{jj}+ \\
&\quad -\sum_{r=1}^{p}\frac{1}{\lambda_{ir}^3}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[x_r^2-\lambda_{ir}\right]x_r^2\,x_j^2\right\}\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)- \\
&\quad -\ \sum_{r=1}^{p}\sum_{s\neq r}E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left[\frac{x_r\,x_s}{\lambda_{ir}\lambda_{is}}\right]^2\ x_j^2\right\}\left[\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right] \\
\boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{\Upsilon}_{i,\,i}(\mathbf{y},F)\boldsymbol{\beta}_m &= \left(\Psi_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)\right)_{jm}+ \\
&\quad +2\sum_{r=j,m}\frac{1}{\lambda_{ir}^2}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[x_r^2-\lambda_{ir}\right]x_r\,x_j\,x_m\ \left(\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\right\}+ \\
&\quad +4\sum_{r=j,m}\sum_{s\neq r}E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\frac{x_r\,x_s^2}{\lambda_{ir}\lambda_{is}}\,x_j\,x_m\ \left(\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\right\} \\
\boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{y},F)\boldsymbol{\beta}_m &= 2\sum_{r=j,m}\frac{1}{\lambda_{\ell r}^2}E_{F_\ell}\left\{\varphi_1^{(\ell)}(\mathbf{x})\left[x_r^2-\lambda_{\ell r}\right]x_r\,x_j\,x_m\ \left(\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\right\}+ \\
&\quad +4\sum_{r=j,m}\sum_{s\neq r}E_{F_\ell}\left\{\varphi_2^{(\ell)}(\mathbf{x})\frac{x_r\,x_s^2}{\lambda_{\ell r}\lambda_{\ell s}}\,x_j\,x_m\ \left(\mathrm{PIF}_i\left(\mathbf{y},\boldsymbol{\beta}_{\mathrm{I},\,r},F\right)^{\mathrm{T}}\mathbf{x}\right)\right\}\ .
\end{aligned}
$$

Then, using that $\kappa_i\,D_i^{-1}(F)\mathbf{N}_i(F)=\mathbf{V}_i(F)=\boldsymbol{\Sigma}_i$, we get easily

$$
\begin{aligned}
\mathrm{PIF}_i(\mathbf{y},\lambda_{\mathbf{V},\,\ell j},F) &= 0 \qquad \text{for} \quad \ell\neq i \\
\mathrm{PIF}_i(\mathbf{y},\lambda_{\mathbf{V},\,ij},F) &= D_i^{-1}(F)\kappa_i\left(\Psi_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)\right)_{jj}-D_i^{-1}(F)W_1(\mathbf{y},\boldsymbol{\beta},\boldsymbol{\Lambda}_i)) \\
&\quad -D_i^{-1}(F)\sum_{r=1}^{p}\frac{1}{\lambda_{ir}^3}E_{F_i}\left\{\varphi_1^{(i)}(\mathbf{x})\left[x_r^2-\lambda_{ir}\right]x_r^2\ \left[\kappa_i\,x_j^2-\lambda_{ij}\right]\right\}\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)- \\
&\quad -D_i^{-1}(F)\ \sum_{r=1}^{p}\sum_{s\neq r}E_{F_i}\left\{\varphi_2^{(i)}(\mathbf{x})\left[\frac{x_r\,x_s}{\lambda_{ir}\lambda_{is}}\right]^2\ \left[\kappa_i\,x_j^2-\lambda_{ij}\right]\right\}\ \left[\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,ir},F\right)\lambda_{is}+\right. \\
&\quad \left.+\mathrm{PIF}_i\left(\mathbf{y},\lambda_{\mathrm{I},\,is},F\right)\lambda_{ir}\right]
\end{aligned}
$$

In particular, if $\mathbf{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{i1} = \mathbf{z}_i$ has the same spherical distribution G for all $1 \le i \le k$, straightforward calculations lead to the given expression of $\mathrm{PIF}_i(\mathbf{x}, \lambda_{\mathbf{V},\,\ell j}, F)$. Moreover, using that $\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,m}, F\right)_j = -\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m$, we get

$$
\begin{aligned}
\boldsymbol{\beta}_j^{\mathrm{T}} \boldsymbol{\Upsilon}_{i,\,i}(\mathbf{y}, F)\boldsymbol{\beta}_m &= \Psi_1(\mathbf{y}, \boldsymbol{\beta}, \mathbf{\Lambda}_i)_{jm} + \alpha_3(G)\left[\lambda_{im}\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m + \lambda_{ij}\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,m}, F\right)_j\right] \\
&= \Psi_1(\mathbf{y}, \boldsymbol{\beta}, \mathbf{\Lambda}_i)_{jm} + \alpha_3(G)\left(\lambda_{im} - \lambda_{ij}\right)\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m \\
\boldsymbol{\beta}_j^{\mathrm{T}} \boldsymbol{\Upsilon}_{i,\,\ell}(\mathbf{y}, F)\boldsymbol{\beta}_m &= \alpha_3(G)\left[\lambda_{\ell m}\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m + \lambda_{\ell j}\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,m}, F\right)_j\right] \\
&= \alpha_3(G)\left(\lambda_{\ell m} - \lambda_{\ell j}\right)\,\mathrm{PIF}_i\left(\mathbf{y}, \boldsymbol{\beta}_{\mathrm{I},\,j}, F\right)_m\ ,
\end{aligned}
$$

which entails (20). □

References

Boente, G., Critchley, F and Orellana, L. (2004). Influence functions for robust estimators under proportional scatter matrices. http://www.ic.fcen.uba.ar/preprints/boecriore.pdf

Boente, G. and Orellana, L. (2001). A robust approach to common principal components. In *Statistics in Genetics and in the Environmental Sciences*, eds. L. T. Fernholz, S. Morgenthaler, and W. Stahel, pp. 117-147. Basel: Birkhauser.

Boente, G., Pires, A. M. and Rodrigues, I. M. (2002). Influence functions and outlier detection under the common principal components model: A robust approach. *Biometrika*, **89**, 861-875.

Boente, G., Pires, A. M. and Rodrigues, I. M. (2005). General projection–pursuit estimators for the common principal components model: Influence functions and Monte Carlo study. *Journal of Multivariate Analysis*, to appear.

Critchley, F. (1985). Influence in principal components analysis. *Biometrika*, **72**, 627-36.

Croux, C. and Haesbroeck, G. (2000). Principal component analysis based on robust estimators of the covariance or correlation matrix: Influence functions and efficiencies. *Biometrika*, **87**, 603-18.

Croux, C., and Ruiz-Gazen, A. (2005). High breakdown estimators for principal components: the projection-pursuit approach revisited. *Journal of Multivariate Analysis*, to appear.

Filzmoser, P., Reimann, C. and Garrett, R. G. (2005). Multivariate outlier detection in exploration geochemistry. *Computers and Geosciences*, to appear.

Filzmoser, P. (2004). A multivariate outlier detection method. In *Proceedings of the Seventh International Conference on Computer Data Analysis and Modeling*(S. Aivazian, P. Filzmoser, and Yu. Kharin, editors). Minsk: Belarusian State University, Vol. 1, 18-22.

Flury, B. K. (1984). Common principal components in k groups. *Journal of the American Statistical Association,* **79**, 892-898.

Flury, B. K. (1988). *Common Principal Components and Related Multivariate Models.* New York: John Wiley.

Jaupi, L. and Saporta, G. (1993). Using the influence function in robust principal components analysis. *New Directions in Statistical Data Analysis and Robustness*, eds. S. Morgenthaler, E. Ronchetti and W. Stahel, pp. 147-56. Basel: Birkhauser.

Li, G. and Chen, Z. (1985). Projection–pursuit approach to robust dispersion matrices and principal components: primary theory and Monte Carlo. *Journal of the American Statistical Association*, **80**, 759-766.

Lopuhaä, H. P. (1999) Asymptotics of reweighted estimators of multivariate location and scatter. *Annals of Statistics*, **27**, 1-28.

Pires, A. M. and Branco, J. (2002). Partial influence functions. *Journal of Multivariate Analysis*, **83**, 451-468.

Pison, G., Rousseeuw, P. J., Filzmoser, P. and Croux, C. (2000). A robust version of principal factor analysis. In *Compstat: Proceedings in Computational Statistics,* eds. J. Bethlehem and P. van der Heijden, Heidelberg: Physica–Verlag, 385-390.

Rodrigues, I. M. (2003). *Métodos Robustos em Análise de Componentes Principais Comuns.* Unpublished PhD Thesis (in portuguese), Universidade Técnica de Lisboa. Available on `http://www.math.ist.utl.pt/~apires/phd.html`.

Shi, L. (1997). Local influence in principal components analysis. *Biometrika*, **84**, 175-86.

p	a_L	a_B	a_M	$1+\sqrt{a_B^2+a_L^2}$	$\max\left(a_M^2-1,1\right)$	a_M^4
2	2.927	3.088	2.448	5.2548	4.991	35.898
4	3.957	6.809	3.080	8.8753	8.488	90.017

Table 1: Percentiles 0.95 of the Mahalanobis distance and of the outlier detection measures $IML\left(\mathbf{z},\mathbf{I},\mathbf{I}\right)$ and $IMB\left(\mathbf{z},\mathbf{I},\mathbf{I}\right)$ and bounds giving the relation between them.

λ		LPP	PI_1	$\mathrm{W_mLPP}$	$\mathrm{W_pLPP}$	$\mathrm{W_mPI}_1$	$\mathrm{W_pPI}_1$	WPI_2
16	$\boldsymbol{C_0}$	0.0319	0.0063	0.0096	0.0096	0.0078	0.0080	0.0078
	$\boldsymbol{C_1}$	0.0460	0.0072	0.0090	0.0090	0.0078	0.0080	0.0080
	$\boldsymbol{C_2}$	0.0432	0.0078	0.0109	0.0110	0.0089	0.0090	0.0083
8	$\boldsymbol{C_0}$	0.0429	0.0089	0.0126	0.0130	0.0114	0.0114	0.0113
	$\boldsymbol{C_1}$	0.0955	0.0103	0.0132	0.0134	0.0119	0.0118	0.0121
	$\boldsymbol{C_2}$	0.1110	0.0113	0.0164	0.0164	0.0141	0.0142	0.0134
2	$\boldsymbol{C_0}$	0.0447	0.0099	0.0126	0.0126	0.0130	0.0130	0.0120
	$\boldsymbol{C_1}$	0.3377	0.8805	0.0170	0.0168	0.0152	0.0150	0.0144
	$\boldsymbol{C_2}$	0.5463	0.4156	0.5448	0.5566	0.8752	0.8752	0.6466
1	$\boldsymbol{C_0}$	0.0329	0.0073	0.0095	0.0096	0.0093	0.0094	0.0086
	$\boldsymbol{C_1}$	0.3498	0.8842	0.0139	0.0140	0.0126	0.0124	0.0116
	$\boldsymbol{C_2}$	0.5467	0.4141	0.5441	0.5532	0.8742	0.8742	0.6468

Table 2: Median of the square distance between the estimated common principal directions and the true principal axes related to the eigenvalue λ, under a proportional model in dimension 4.

λ		LPP	PI_1	$\mathrm{W_mLPP}$	$\mathrm{W_pLPP}$	$\mathrm{W_mPI}_1$	$\mathrm{W_pPI}_1$	WPI_2
16	$\boldsymbol{C_0}$	0.0609	0.0130	0.0299	0.0298	0.0278	0.0278	0.0216
	$\boldsymbol{C_1}$	0.0874	0.0133	0.0313	0.0314	0.0240	0.0280	0.0235
	$\boldsymbol{C_2}$	0.0930	0.0204	0.0510	0.0512	0.0338	0.0340	0.0266
8	$\boldsymbol{C_0}$	0.0683	0.0151	0.0325	0.0326	0.0304	0.0304	0.0242
	$\boldsymbol{C_1}$	0.1426	0.0163	0.0381	0.0382	0.0270	0.0310	0.0264
	$\boldsymbol{C_2}$	0.1710	0.0234	0.0556	0.0558	0.0377	0.0378	0.0303
2	$\boldsymbol{C_0}$	0.0798	0.0156	0.0217	0.0218	0.0236	0.0236	0.0196
	$\boldsymbol{C_1}$	0.5667	0.9106	0.0638	0.0640	0.0398	0.0428	0.0392
	$\boldsymbol{C_2}$	0.7188	0.7380	0.7642	0.7688	0.8959	0.8978	0.8115
1	$\boldsymbol{C_0}$	0.0702	0.0137	0.0190	0.0192	0.0212	0.0212	0.0173
	$\boldsymbol{C_1}$	0.5787	0.9103	0.0624	0.0626	0.0373	0.0404	0.0369
	$\boldsymbol{C_2}$	0.7089	0.7367	0.7622	0.7670	0.8941	0.8960	0.8098

Table 3: Mean of the square distance between the estimated common principal directions and the true principal axes related to the eigenvalue λ, under a proportional model in dimension 4.

$\lambda_{14} = 16$

	LPP	PI_1	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_1}$		$\boldsymbol{C_2}$		$\boldsymbol{C_0}$			$\boldsymbol{C_1}$			$\boldsymbol{C_2}$		
Mean	0.0020	-0.0618	0.0259	-0.0473	-0.2567	-0.2037	-0.0607	-0.0529	-0.0486	-0.0312	-0.0258	-0.0214	-0.2377	-0.1978	-0.1840
SD	0.1961	0.1585	0.1900	0.1618	0.2320	0.1784	0.1785	0.1750	0.1699	0.1783	0.1799	0.1720	0.2139	0.1982	0.1918
MSE	0.0385	0.0289	0.0368	0.0284	0.1198	0.0734	0.0356	0.0334	0.0312	0.0327	0.0330	0.0301	0.1024	0.0785	0.0707

$\lambda_{13} = 8$

	LPP	PI_1	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_1}$		$\boldsymbol{C_2}$		$\boldsymbol{C_0}$			$\boldsymbol{C_1}$			$\boldsymbol{C_2}$		
Mean	0.0174	-0.0595	0.0955	-0.0491	-0.2178	-0.2087	-0.0442	-0.0504	-0.0475	-0.0164	-0.0300	-0.0273	-0.2202	-0.2043	-0.1916
SD	0.2024	0.1516	0.1934	0.1597	0.2172	0.1800	0.1720	0.1726	0.1693	0.1725	0.1813	0.1739	0.2056	0.2030	0.1982
MSE	0.0413	0.0265	0.0465	0.0279	0.0947	0.0760	0.0316	0.0323	0.0309	0.0300	0.0338	0.0310	0.0908	0.0830	0.0761

$\lambda_{12} = 2$

	LPP	PI_1	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_1}$		$\boldsymbol{C_2}$		$\boldsymbol{C_0}$			$\boldsymbol{C_1}$			$\boldsymbol{C_2}$		
Mean	0.0633	-0.0574	0.4267	0.1256	0.2753	-0.0939	-0.0321	-0.0494	-0.0432	0.0236	-0.0261	-0.0198	-0.0271	-0.0518	-0.0621
SD	0.1955	0.1514	0.2854	0.2109	0.2535	0.1256	0.1721	0.1714	0.1679	0.1794	0.1765	0.1695	0.1320	0.1351	0.1324
MSE	0.0423	0.0262	0.2639	0.0603	0.1402	0.0246	0.0306	0.0318	0.0301	0.0327	0.0318	0.0291	0.0182	0.0209	0.0214

$\lambda_{11} = 1$

	LPP	PI_1	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_1}$		$\boldsymbol{C_2}$		$\boldsymbol{C_0}$			$\boldsymbol{C_1}$			$\boldsymbol{C_2}$		
Mean	0.0465	-0.0831	0.6657	0.5190	0.5558	0.3428	-0.0576	-0.0812	-0.0724	0.0210	0.0145	-0.0040	0.4061	0.3660	0.3686
SD	0.2208	0.1633	0.2862	0.1750	0.2369	0.1377	0.1802	0.1778	0.1771	0.1900	0.1775	0.1732	0.1464	0.1567	0.1408
MSE	0.0509	0.0336	0.5260	0.3005	0.3656	0.1367	0.0358	0.0382	0.0366	0.0365	0.0317	0.0300	0.1867	0.1587	0.1559

Table 4: Estimation of $\log\left(\frac{\widehat{\lambda}}{\lambda}\right)$ of $\boldsymbol{\Sigma}_1$, under a proportional model in dimension 4.

	Median		Mean		Median			Mean		
	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
$\boldsymbol{C_0}$	0.0080	0.0032	0.0221	0.0105	0.0023	0.0030	0.0028	0.0046	0.0072	0.0065
$\boldsymbol{C_{1,0.1}}$	0.0644	0.0040	0.1755	0.1810	0.0027	0.0042	0.0038	0.0516	0.1982	0.1920
$\boldsymbol{C_{1,0.2}}$	0.3371	0.0372	0.4731	0.8984	0.0049	0.0676	0.0241	0.3380	0.9163	0.8633
$\boldsymbol{C_2}$	0.6797	0.2594	0.6589	0.8626	0.0135	0.3416	0.1263	0.3489	0.8541	0.8081

Table 5: Median and mean of $\|\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1\|^2$, in dimension 2.

	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_{1,0.1}}$		$\boldsymbol{C_0}$			$\boldsymbol{C_{1,0.1}}$		
N_{45}	1	1	38	47	0	0	0	12	52	50
N_{60}	1	1	18	47	0	0	0	12	52	50
N_{80}	0	1	2	44	0	0	0	12	45	46
	$\boldsymbol{C_{1,0.2}}$		$\boldsymbol{C_2}$		$\boldsymbol{C_{1,0.2}}$			$\boldsymbol{C_2}$		
N_{45}	178	239	336	241	88	245	230	99	245	232
N_{60}	53	239	31	234	88	245	230	94	234	218
N_{80}	1	230	0	183	88	231	220	62	168	159

Table 6: Estimation of the common principal directions. N_α denotes the number of times that the absolute value of the angle between $\mathbf{e}_1$ and the estimated first principal direction is greater than α degrees, in dimension 2.

Estimates of $\lambda = 14$

	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_{1,0.1}}$		$\boldsymbol{C_0}$			$\boldsymbol{C_{1,0.1}}$		
Mean	-0.0260	-0.1462	-0.1704	-0.2634	0.0426	-0.1093	-0.1056	-0.0850	-0.2232	-0.2305
SD	0.2750	0.2654	0.2387	0.2800	0.2363	0.2695	0.2583	0.2431	0.2738	0.2694
MSE	0.0763	0.0919	0.0861	0.1479	0.0576	0.0846	0.0778	0.0663	0.1249	0.1258
	$\boldsymbol{C_{1,0.2}}$		$\boldsymbol{C_2}$		$\boldsymbol{C_{1,0.2}}$			$\boldsymbol{C_2}$		
Mean	-0.1198	-0.3072	-0.0028	-0.2891	-0.1506	-0.2829	-0.3030	-0.1235	-0.2608	-0.2833
SD	0.1803	0.2408	0.1722	0.2417	0.1972	0.1920	0.1952	0.2021	0.1903	0.1926
MSE	0.0469	0.1526	0.0296	0.1422	0.0616	0.1171	0.1301	0.0561	0.1044	0.1175

Estimates of $\lambda = 4$

	LPP	PI_1	LPP	PI_1	WLPP	WPI_1	WPI_2	WLPP	WPI_1	WPI_2
	$\boldsymbol{C_0}$		$\boldsymbol{C_{1,0.1}}$		$\boldsymbol{C_0}$			$\boldsymbol{C_{1,0.1}}$		
Mean	0.0011	-0.1213	0.3963	0.2512	0.0515	-0.0911	-0.0911	0.4737	0.3722	0.3730
SD	0.2946	0.2701	0.3566	0.2844	0.2385	0.2788	0.2692	0.2371	0.3096	0.2766
MSE	0.0868	0.0877	0.2845	0.1441	0.0596	0.0860	0.0808	0.2811	0.2347	0.2159
	$\boldsymbol{C_{1,0.2}}$		$\boldsymbol{C_2}$		$\boldsymbol{C_{1,0.2}}$			$\boldsymbol{C_2}$		
Mean	0.8303	0.4110	0.9454	0.3963	0.7439	0.5228	0.5178	0.7451	0.5119	0.5152
SD	0.3070	0.3481	0.2589	0.3500	0.1682	0.3562	0.3336	0.1698	0.3649	0.3390
MSE	0.7850	0.2905	0.9626	0.2799	0.5828	0.4008	0.3800	0.5850	0.3958	0.3809

Table 7: Estimation of $\log\left(\frac{\widehat{\lambda}}{\lambda}\right)$ of $\boldsymbol{\Sigma}_1$, in dimension 2.

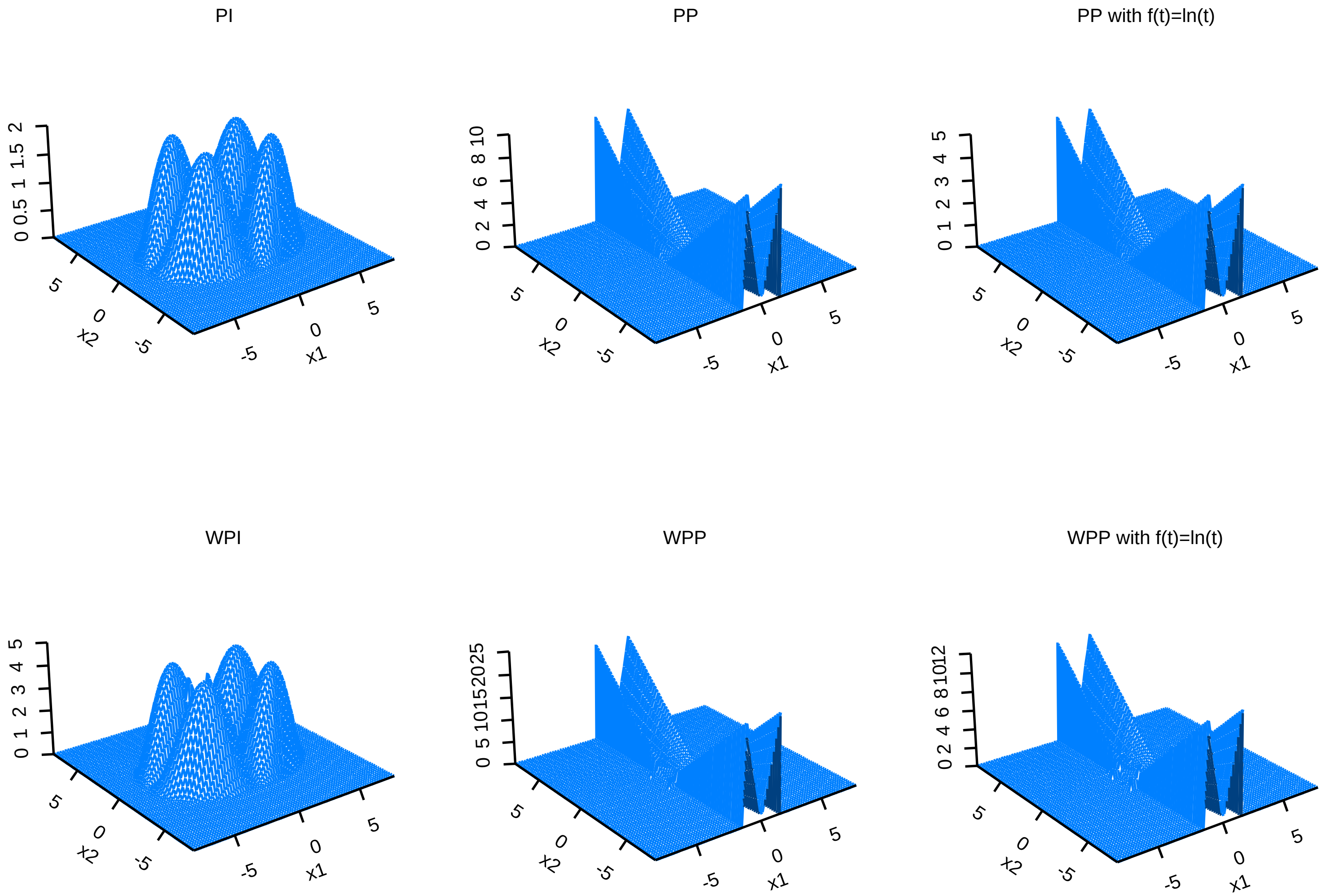

Figure 1: $\|\mathrm{PIF}_1(\mathbf{x}, \boldsymbol{\beta}_1, F)\|$ at $F = F_1 \times F_2$ with $F_1 = N\left(\mathbf{0}, \mathrm{diag}(2,1)\right)$ and $F_2 = N\left(\mathbf{0}, 4\mathrm{diag}(2,1)\right)$.

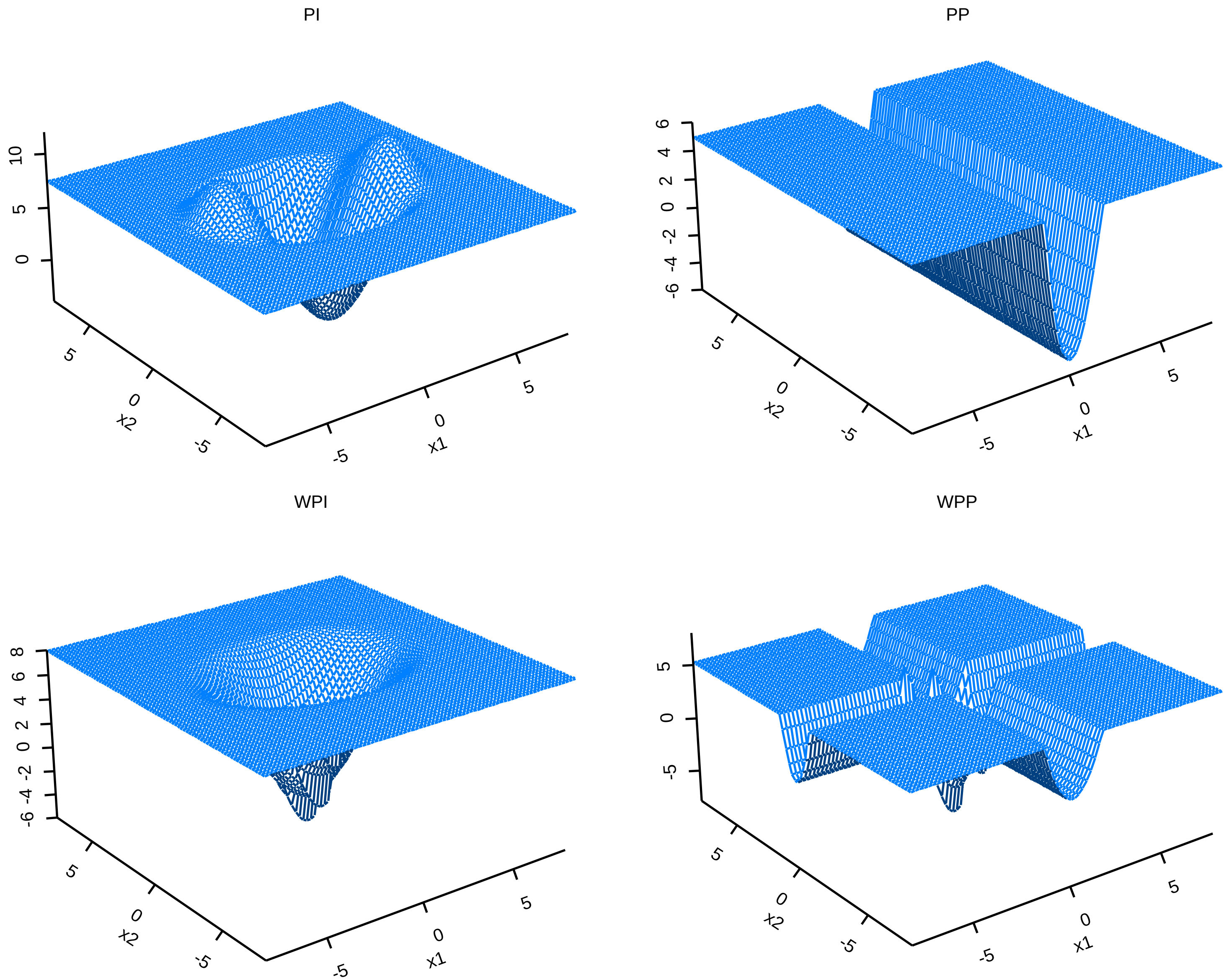

Figure 2: $\|\mathrm{PIF}_1(\mathbf{x}, \lambda_{11}, F)\|$ at $F = F_1 \times F_2$ with $F_1 = N(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N(\mathbf{0}, 4\mathrm{diag}(2,1))$.

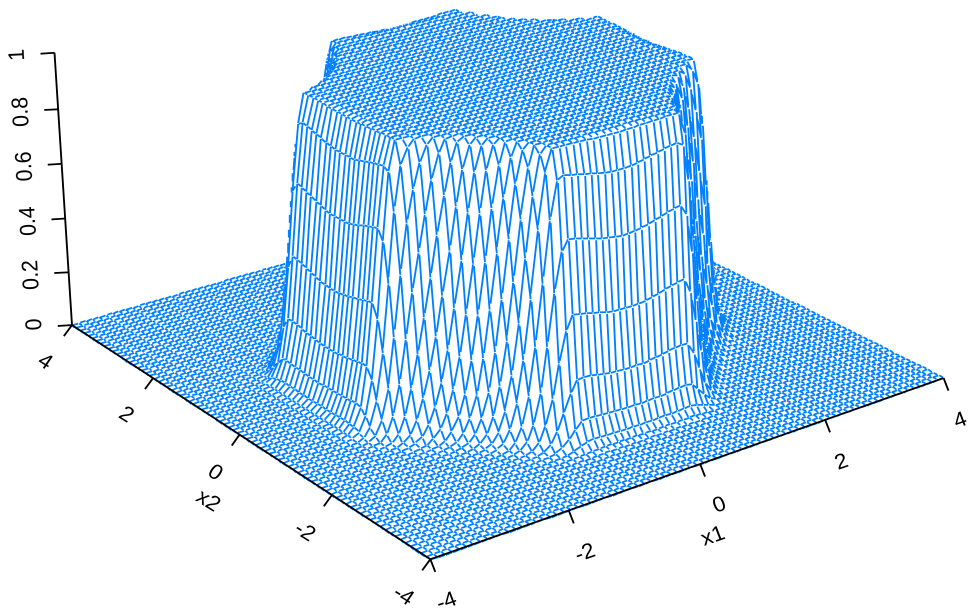

Figure 3: Weight Function used to compute the partial influence functions.

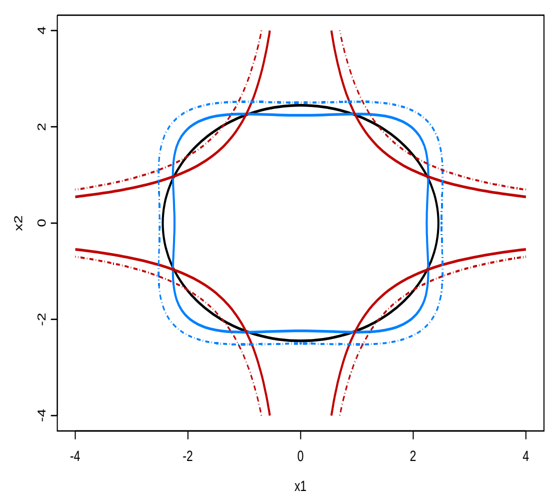

Figure 4: Detection regions obtained with Mahalanobis distance (ellipse), IML_1 (closed curves) and IMB_1 (open curves). The 95% detection limits are the solid curves and the 97.5% detection limits are the dashed curves.