Influence functions for robust estimators under proportional scatter matrices*

Graciela Boente[†] Frank Critchley[‡] Liliana Orellana[§]

Abstract

In this paper, under a proportional model, two families of robust estimates for the proportionality constants, the common principal axes and their size are discussed. The first approach is obtained by plugging–in robust scatter matrices on the maximum likelihood equations for normal data. A projection–pursuit and a modified projection-pursuit approach, adapted to the proportional setting, are also considered. For all families of estimates estimates, partial influence functions are obtained and asymptotic variances are derived from them. Through a Monte Carlo study the performance of the estimates is compared.

MSC: Primary 62F35, Secondary 62H25.

Key words and phrases: Robust Estimation, Robust Scatter Matrices, Projection–Pursuit, Common Principal Components, Proportional Models, Partial Influence Curves.

Running Title: Influence functions under proportionality

*This research was partially supported by Grant PICT # 03-00000-006277 from ANPCYT and a Grant from the Fundación Antorchas at Buenos Aires, Argentina. The research of Graciela Boente was begin while she was granted with a Guggenheim fellowship.

[†]Universidad de Buenos Aires and CONICET, Argentina

[‡]The Open University, United Kingdom

[§]Universidad de Buenos Aires, Argentina.

1 Introduction

In many situations involving several populations, in multivariate analysis, such as discriminant analysis or MANOVA, equality of covariance matrices is often assumed. If this assumption does not hold, an alternative to model the structure of the related scatter matrices of several populations, is to assume that they differ just in a proportionality constant while a weaker assumption is that they are conmutable. In both situations, the scatter matrices of the different populations have the same eigenvectors.

Assume that we have k populations in R^p, with covariance matrices $\mathbf{\Sigma}_1, ..., \mathbf{\Sigma}_k$. The common principal components (CPC) model (Flury (1984)) states that $\mathbf{\Sigma}_i = \boldsymbol{\beta}\mathbf{\Lambda}_i\boldsymbol{\beta}'$, $1 \leq i \leq k$, where $\mathbf{\Lambda}_i$ is a diagonal matrix and $\boldsymbol{\beta}$ is an orthogonal matrix, while the proportional model assumes that

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

Estimation for proportional covariance matrices in the two–sample case and for normal populations have been studied by Khatri(1967) and by Pillai, Al–Ani and Jouris (1969) who studied the ratios of the characteristic roots of $\mathbf{S}_1\mathbf{S}_2^{-1}$ where $\mathbf{S}_i$ are the sample covariance matrices. An extensive study of proportionality between covariance matrices was done by Kim (1971) and the solution for $k = 2$ populations was later published by Guttman, Kim and Olkin (1985). Similar results were published independently by Rao (1983). The case of several groups, has been considered independently by several authors. 0wen (1984) used proportional scatter matrices in classification and gave an algorithm to find the maximum likelihood estimators. Essentially the same algorithm was proposed by Manly and Rainer (1987) and by Eriksen (1987), who, in addition, proved the convergence of the algorithm and the uniqueness of maximum likelihood estimates. Flury (1986) using a parametrization based on the eigenvalues and eigenvectors of $\mathbf{\Sigma}_1$, obtained a system of equations defining the maximum likelihood estimates and gave an algorithm to solve it. The asymptotic behavior of the proportionality constants was first obtained by Kim (1971) and can be seen in Flury (1988). Proportional matrices were used in discriminant analysis by Flury and Schmid (1992) and by Flury, Schmid and Narayanan (1994) where better rates of correct classification are obtained if the most parsimonious among all correct models is used for discrimination.

In many situations robust estimators of the proportionality constants and of the eigenvectors and eigenvalues, under proportionality of the scatter matrices, are desirable. As in the one–population setting and in the CPC model, two approaches will be considered in this paper. The first one compute the principal directions by plugging in robust scatter matrices in the equations defining the maximum likelihood estimates for normal data. Robust plug–in estimates under the CPC model were proposed by Novi Inverardi and Flury (1992) and studied by Boente and Orellana (2001). The second approach is a projection–based method, which defines the principal axes by maximizing (or minimizing) a robust scale of the projected observations. Boente and Orellana (2001) also considered a projection–pursuit approach under a CPC model while the partial influence functions were computed by Boente, Pires and Rodrigues (2002) for both proposals.

Robust plug–in estimators for the principal axes, the proportionality constants and the eigenvalues were proposed by Boente and Orellana (2004). These authors also considered three proposals for estimating the proportionality constants, derived the asymptotic behavior of the robust estimates and proposed a test for equality against proportionality. Let $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}$, $1 \leq i \leq k$, be independent observations from k independent samples in R^p, with location parameter $\boldsymbol{\mu}_i$ and scatter matrix $\mathbf{\Sigma}_i$. We are interested in estimating robustly the common eigenvectors $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p\right)$ of $\mathbf{\Sigma}_i$, the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_p$ of $\mathbf{\Sigma}_1$ and the proportionality constants ρ_i under the proportionality model (1). In order to identify uniquely the axes we will assume that the characteristic roots of $\mathbf{\Sigma}_1$ are distinct and that the the columns of $\boldsymbol{\beta}$ are arranged according to decreasing values of the eigenvalues of $\mathbf{\Sigma}_1$. Let $\mathbf{V}_i$ be a robust affine equivariant estimate of the scatter matrix of the $i-$th population. The so–called **Proposal 1** in Boente and Orellana (2004), estimates the parameters under a proportional model by plugging–in the matrices $\mathbf{V}_i$ in the equations defining the maximum likelihood estimates for normal data, which lead to the following system of equations

$$\rho_i = \frac{1}{p}\sum_{j=1}^{p}\frac{\boldsymbol{\beta}_j'\mathbf{V}_i\boldsymbol{\beta}_j}{\lambda_j} \qquad i=2,\ldots,k \tag{2}$$

$$\lambda_j = \frac{1}{N}\sum_{i=1}^{k}\frac{n_i}{\rho_i}\boldsymbol{\beta}_j'\mathbf{V}_i\boldsymbol{\beta}_j \tag{3}$$

$$0 = \boldsymbol{\beta}_l'\left(\sum_{i=1}^{k}\frac{n_i}{\rho_i}\mathbf{V}_i\right)\boldsymbol{\beta}_j \qquad l\neq j \tag{4}$$

$$\delta_{mj} = \boldsymbol{\beta}_m'\boldsymbol{\beta}_j\ , \tag{5}$$

where $N=\sum_{i=1}^{k} n_i$, $\delta_{mj}=1$ when $m=j$ and $\delta_{mj}=0$ when $m\neq j$. On the other hand, in their **Proposal 2** and **3**, the proportionality constants are estimated, respectively, as

$$\widehat{\rho}_i^{(2)} = \left[\frac{|\mathbf{V}_i|}{|\mathbf{V}_1|}\right]^{\frac{1}{p}} \qquad \widehat{\rho}_1^{(2)}=1\ , \tag{6}$$

$$\widehat{\rho}_i^{(3)} = \frac{\operatorname{tr}(\mathbf{V}_1^{-1}\mathbf{V}_i)}{p} \qquad \widehat{\rho}_1^{(3)}=1\ . \tag{7}$$

Once the estimates of the proportionality constants have been obtained, using (6) or (7) another family of estimates for the eigenvectors $\boldsymbol{\beta}$ and the eigenvalues λ_i can be obtained by solving (3) and (4).

As mentioned above, Boente and Orellana (2001) considered, for the CPC model, a projection–pursuit approach to estimate the common directions. This method can be also used under a proportional model. Let $s(.)$ be a univariate robust scale statistic, $\tau_i=\frac{n_i}{N}$ for $1\le i\le k$ and $\underline{\mathbf{X}}_i=(\mathbf{x}_{i1},\ldots,\mathbf{x}_{in_i})$. The estimates of the common principal axes defined in Boente and Orellana (2001) are obtained by solving iteratively

$$\begin{cases}\widehat{\boldsymbol{\beta}}_1 & =\underset{\|\mathbf{b}\|=1}{\operatorname{argmax}}\sum_{i=1}^{k}\tau_i\, s^2(\underline{\mathbf{X}}_i'\mathbf{b})\\ \widehat{\boldsymbol{\beta}}_j & =\underset{\mathbf{b}\in\mathcal{B}_j}{\operatorname{argmax}}\sum_{i=1}^{k}\tau_i s^2(\underline{\mathbf{X}}_i'\mathbf{b}) \quad 2\le j\le p\ ,\end{cases} \tag{8}$$

where $\mathcal{B}_j=\{\mathbf{b}:\|\mathbf{b}\|=1, \mathbf{b}'\widehat{\boldsymbol{\beta}}_m=0 \text{ for } 1\le m\le j-1\}$ and we include the weights τ_i to adapt for different sample sizes. Once the common direction estimates have been obtained, the eigenvalues and the covariance matrix of the $i-$th population can be defined as

$$\widehat{\lambda}_{ij}=s^2(\underline{\mathbf{X}}_i'\widehat{\boldsymbol{\beta}}_j) \quad \text{for } 1\le j\le p \qquad \mathbf{V}_i=\sum_{j=1}^{p}\widehat{\lambda}_{ij}\widehat{\boldsymbol{\beta}}_j\widehat{\boldsymbol{\beta}}_j'\ . \tag{9}$$

Under proportionality of the scatter matrices, the estimates of the eigenvalues of the first population and of the proportionality constants can be computed in two different ways:

a) Solve (2) and (3) with the estimated directions obtained in (8) and the matrices $\mathbf{V}_i$ defined in (9).

b) Define the estimators of the proportionality constants through **Proposal 2** or **Proposal 3**, given in (6) and (7) respectively. Then, define the estimate of $j-$th eigenvalue for the $i-$th population as $\widehat{\rho}_i\widehat{\lambda}_{1j}$.

The axes can be reordered in such a way that $\widehat{\lambda}_1>\ldots>\widehat{\lambda}_p$. In Boente and Orellana (2001), an algorithm similar to that given by Croux and Ruiz–Gazen (1996), is described in order to obtain the estimates defined through (8).

The projection pursuit estimates are rotational equivariant if the same rotation is applied to all the populations in order to preserve the proportional structure. The eigenvectors are also Fisher–consistent if all the populations have the same elliptical distribution except for changes in the scatter matrices, if the scale functional related to the scale estimate is scale equivariant and if the eigenvalues of $\mathbf{\Sigma}_1$ are distinct, as was shown in Proposition 1 of Boente and Orellana (2001). Note that the scale functional need to be calibrated in order to obtain also Fisher consistent estimates of the eigenvalues.

If a preliminary estimate, $\widehat{\rho}_i$, of ρ_i is available, a new class of projection-pursuit estimates can be defined. More precisely, in this paper, we will consider a generalization of (8) where the common principal axes are estimated by solving

$$\begin{cases} \widehat{\boldsymbol{\beta}}_1 & = \underset{\|\mathbf{b}\|=1}{\operatorname{argmax}} \sum_{i=1}^{k} \frac{\tau_i}{\widehat{\rho}_i} s^2(\underline{\mathbf{X}}_i'\mathbf{b}) \\ \widehat{\boldsymbol{\beta}}_j & = \underset{\mathbf{b}\in\mathcal{B}_j}{\operatorname{argmax}} \sum_{i=1}^{k} \frac{\tau_i}{\widehat{\rho}_i} s^2(\underline{\mathbf{X}}_i'\mathbf{b}) \quad 2 \leq j \leq p \,, \end{cases} \tag{10}$$

The eigenvalues of the first population and of the proportionality constants can then, be estimated as decribed above. This new class of estimators has an advantage over the one defined in (8). As it will be shown in Section 3, the projection–pursuit estimates of the common directions defined in (8), have an efficiencie which depend on the relative size among populations. However, by introducing the preliminary estimates $\widehat{\rho}_i$, we will solve this problem and we will attain the same efficiencies as in the one–population setting.

Influence functions are a measure of robustness with respect to single outliers. When dealing with one population, the influence function is essentially the first derivative of the functional version of the estimator. When dealing with several populations, partial influence functions should be considered as a way to measure resistance towards pointwise contaminations at each population.

This paper focus on how observations belonging to each population affect the estimation of the parameters under a proportional model. An approach based on partial influence functions will be followed to quantified this effect. Besides being of theoretical interest and helpful to calibrate the efficiency of the robust estimates measuring the influence of an observation on the classical estimates can be used as a diagnostic tool to detect influential observations, as was done under a CPC model by Boente, Pires and Rodrigues (2002). The paper is organized as follows. In Section 2, we will derive, partial influence functions for the estimates obtained by plugging in robust scatter matrices in the equations defining the maximum likelihood estimates and for those defined through a projection–pursuit approach. From the partial influence functions, asymptotic variances are obtained heuristically in Section 3. Finally, in Section 4, through a simulation study the performance of the different proposals is compared. Proofs are given in the Appendix due to their tedious computations.

2 Partial Influence functions

As is well known, influence functions are a measure of robustness with respect to single outliers. The importance of the influence function lies in its heuristic interpretation. It describes the effect of an infinitesimal contamination at a single point on the estimate, standarized by the amount of contamination. When dealing with one population, the influence function is essentially the first derivative of the functional version of the estimate. When dealing with several populations, partial influence functions were introduced by Pires and Branco (2002), as a way to measure resistance towards pointwise contaminations at each population. We remind its definition. For a given distribution $F = F_1 \times \ldots \times F_k$ and a functional $T(F)$, partial influence functions of are defined as $\text{PIF}_{i_0}(\mathbf{x}, T, F) = \lim_{\epsilon\to 0} \frac{T(F_{\epsilon,\mathbf{x},i_0}) - T(F)}{\epsilon}$, where $F_{\epsilon,\mathbf{x},i_0} = F_1 \times \ldots \times F_{i_0-1} \times F_{i_0,\epsilon,\mathbf{x}} \times F_{i_0+1} \times \ldots F_k$ and $F_{i,\epsilon,\mathbf{x}} = (1-\epsilon)F_i + \epsilon\delta_{\mathbf{x}}$. In Pires and Branco (2002), it is shown that the asymptotic variance of the estimates can be evaluated as

$$\text{ASVAR}\,(T,F) = \sum_{i=1}^{k} {\tau_i}^{-1} E_{F_i}\left(\text{PIF}_i\,(\mathbf{x}_i,T,F)\,\text{PIF}_i\,(\mathbf{x}_i,T,F)'\right)\ . \tag{11}$$

Therefore, from the partial influence functions, the asymptotic variances for the estimates of the common eigenvectors, of the proportionality constants and of the eigenvalues of the first population can be derived using the above heuristic formula. A proof of the asymptotic normality of the plug–in estimates can be found in Boente and Orellana (2004) while the asymptotic distribution of the projection–pursuit estimates is beyond the scope of this paper.

2.1 Partial influence functions for the Plug–in estimates

As described in the Introduction, these estimates are obtained by plugging–in robust scatter matrices in the equation defining the maximum likelihood estimates for normal data. For a given distribution $F = F_1 \times \ldots \times F_k$, and a robust scatter functional $\mathbf{V}_i = \mathbf{V}_i(F_i)$, the functionals related to $\mathbf{V} = (\mathbf{V}_1, \ldots, \mathbf{V}_k)$, are defined as the solution of

$$\rho^{(1)}_{\mathbf{V},i}(F) = \frac{1}{p}\sum_{j=1}^{p} \frac{\boldsymbol{\beta}_{\mathbf{V},j}(F)'\,\mathbf{V}_i(F_i)\,\boldsymbol{\beta}_{\mathbf{V},j}(F)}{\lambda_{\mathbf{V},j}(F)} \qquad i=2,\ldots,k \qquad \rho^{(1)}_{\mathbf{V},1}(F) = 1 \tag{12}$$

$$\lambda_{\mathbf{V},j}(F) = \sum_{i=1}^{k} \frac{\tau_i}{\rho^{(1)}_{\mathbf{V},i}(F)} \boldsymbol{\beta}_{\mathbf{V},j}(F)'\,\mathbf{V}_i(F_i)\,\boldsymbol{\beta}_{\mathbf{V},j}(F) \qquad 1 \le j \le p \tag{13}$$

$$0 = \boldsymbol{\beta}_{\mathbf{V},m}(F)' \left[\sum_{i=1}^{k} \frac{\tau_i}{\rho^{(1)}_{\mathbf{V},i}(F)} \mathbf{V}_i(F_i)\right] \boldsymbol{\beta}_{\mathbf{V},j}(F) \qquad m \neq j \tag{14}$$

$$\delta_{mj} = \boldsymbol{\beta}_{\mathbf{V},m}(F)'\boldsymbol{\beta}_{\mathbf{V},j}(F)\ . \tag{15}$$

The following Theorem gives the values of the partial influence functions for the plug–in estimates under proportionality of the scatter matrices.

Theorem 2.1. *Let $\mathbf{V}_i(F)$ be a scatter functional such that $\mathbf{V}_i(F_i) = \boldsymbol{\Sigma}_i = \rho_i \boldsymbol{\Sigma}_1$, $\rho_1 = 1$. Denote by $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p, \lambda_1, \ldots, \lambda_p$ the eigenvectors and eigenvalues of $\boldsymbol{\Sigma}_1$. Assume that $\lambda_1 > \ldots > \lambda_p$ and that $IF(\mathbf{x}, \mathbf{V}_i, F_i)$ exists. Denote by $A_i(\mathbf{x}) = \frac{1}{p}\sum_{j=1}^{p} \frac{\boldsymbol{\beta}_j' IF(\mathbf{x}, \mathbf{V}_i, F_i)\,\boldsymbol{\beta}_j}{\lambda_j}$. Then, the partial influence functions of the solution $(\rho^{(1)}_{\mathbf{V},i}(F), \lambda_{\mathbf{V},j}(F), \boldsymbol{\beta}_{\mathbf{V},j}(F))$ of (12) to (15) are given by*

$$PIF_i(\mathbf{x}, \lambda_{\mathbf{V},j}, F) = \frac{\tau_i}{\rho_i} \boldsymbol{\beta}_j' IF(\mathbf{x}, \mathbf{V}_i, F_i)\,\boldsymbol{\beta}_j - \frac{\tau_i}{\rho_i}\lambda_j A_i(\mathbf{x}) + \lambda_j A_1(\mathbf{x})\delta_{i1} \qquad 1 \le j \le p \tag{16}$$

$$PIF_i(\mathbf{x}, \rho^{(1)}_{\mathbf{V},\ell}, F) = A_i(\mathbf{x})\ (1-\delta_{i1})\,\delta_{\ell i} - \rho_\ell A_1(\mathbf{x})\delta_{i1}\,(1-\delta_{\ell i}) \qquad 2 \le \ell \le k \tag{17}$$

$$PIF_i(\mathbf{x}, \boldsymbol{\beta}_{\mathbf{V},j}, F) = \frac{\tau_i}{\rho_i} \sum_{m \neq j} \frac{1}{\lambda_j - \lambda_m} \left[\boldsymbol{\beta}_j' IF(\mathbf{x}, \mathbf{V}_i, F_i)\,\boldsymbol{\beta}_m\right] \boldsymbol{\beta}_m \qquad 1 \le j \le p\ . \tag{18}$$

Remark 2.1. Except for the constant τ_i, the partial influence functions of the plug–in common eigenvectors are equal to the influence function of the eigenvectors of $\mathbf{V}_i$, which was obtained by Croux and Haesbroeck (2000). On the other hand, the first term of partial influence function of the eigenvalues of the first population equals the influence function of the eigenvalues of $\mathbf{V}_i$, up to a constant. An extra term is added when several populations are present. The expression $A_i(\mathbf{x})$ can be viewed as weighted average of the influence functions of the eigenvalues of $\mathbf{V}_i$, the weights been inversely proportional to the size of the direction.

Remark 2.2. Influence functions at the normal distribution can easily be derived from Lemma 1 from Croux and Haesbroek (2000) which gives an expression for the influence function of a robust scatter functional. According to it, if $F = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $d^2(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$, then, for any affine equivariant scatter matrix functional $\mathbf{V}$ possesing an influence function, there exists two functions $\alpha_{\mathbf{V}}$ and $\gamma_{\mathbf{V}} : [0,\infty) \to R$ such that $\text{IF}(\mathbf{x}, \mathbf{V}, F) = \alpha_{\mathbf{V}}(d(\mathbf{x}))(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})' - \gamma_{\mathbf{V}}(d(\mathbf{x}))\boldsymbol{\Sigma}$. So that under a proportional scatter model, we have that $\text{PIF}_i(\mathbf{x}, \boldsymbol{\beta}_{\mathbf{V},j}, F) = \frac{\tau_i}{\rho_i}\alpha_{\mathbf{V}}(d_i(\mathbf{x}))\sum_{m\neq j}\frac{z_{ij}z_{im}}{\lambda_j-\lambda_m}\boldsymbol{\beta}_m$, for $1 \le j \le p$, where $z_{ij} = \boldsymbol{\beta}_j'(\mathbf{x}-\boldsymbol{\mu}_i)$ and $d_i^2(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu}_i)'\boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) = \frac{(\mathbf{x}-\boldsymbol{\mu}_i)'\boldsymbol{\Sigma}_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)}{\rho_i}$. As in the one population setting, the function $\alpha_{\mathbf{V}}$ can be interpreted as a downweighting function, which bounds the influence of the classical estimates for the common directions since both functions $\alpha_{\mathbf{V}}$ and $\gamma_{\mathbf{V}}$ are equal to 1 for the sample covariance matrix. Thus, the maximum likelihood estimates have partial influence functions given by

$$\begin{aligned}
\text{PIF}_i(\mathbf{x}, \lambda_{\text{ML},j}, F) &= \frac{\tau_i}{\rho_i}\left[z_{ij}^2 - \rho_i\lambda_j\right] - \frac{\tau_i}{\rho_i}\lambda_j A_{\text{ML},i}(\mathbf{x}) + \lambda_j A_{\text{ML},1}(\mathbf{x})\delta_{i1} \qquad 1 \le j \le p \\
\text{PIF}_i(\mathbf{x}, \rho^{(1)}_{\text{ML},\ell}, F) &= A_{\text{ML},i}(\mathbf{x})\,(1-\delta_{i1})\,\delta_{\ell i} - \rho_\ell A_{\text{ML},1}(\mathbf{x})\delta_{i1}\,(1-\delta_{\ell i}) \qquad 2 \le \ell \le k \\
\text{PIF}_i(\mathbf{x}, \boldsymbol{\beta}_{\text{ML},j}, F) &= \frac{\tau_i}{\rho_i}\sum_{m\neq j}\frac{z_{ij}z_{im}}{\lambda_j-\lambda_m}\boldsymbol{\beta}_m \qquad 1 \le j \le p\,,
\end{aligned}$$

where $A_{\text{ML},i}(\mathbf{x}) = \frac{1}{p}\sum_{j=1}^{p}\frac{z_{ij}^2}{\lambda_j} - \rho_i = \rho_i\left[\frac{1}{p}d_i^2(\mathbf{x}) - 1\right]$ are unbounded functions. Based on the partial influence functions of the classical functionals, outlier detection measures can be defined as in Boente, Pires and Rodrigues (2002).

Theorem 2.2 gives the partial influence functions for the estimates of the proportionality constants given by **Proposal 2** and **3**. Its proof relies on Lemma 3 of Croux and Haesbroek (2000) which gives the expression for the influence functions of the eigenvalues of a robust scatter functional.

Theorem 2.2. *Let $\mathbf{V}_i(F)$ be a scatter functional such that $\mathbf{V}_i(F_i) = \boldsymbol{\Sigma}_i = \rho_i\boldsymbol{\Sigma}_1$, $\rho_1 = 1$. Denote by $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_p, \lambda_1, \ldots, \lambda_p$ the eigenvectors and eigenvalues of $\boldsymbol{\Sigma}_1$ and by $\lambda_{\mathbf{V}_i,j}(F_i)$ the eigenvalues of $\mathbf{V}_i(F_i)$. Assume that $\lambda_1 > \ldots > \lambda_p$ and that $IF(\mathbf{x}, \mathbf{V}_i, F_i)$ exists. Then, if $A_i(\mathbf{x}) = \frac{1}{p}\sum_{j=1}^{p}\frac{\boldsymbol{\beta}_j' IF(\mathbf{x}, \mathbf{V}_i, F_i)\boldsymbol{\beta}_j}{\lambda_j}$, the partial influence functions of the functionals $\rho^{(2)}_{\mathbf{V},i}(F)$ and $\rho^{(3)}_{\mathbf{V},i}(F)$ defined through*

$$\rho^{(2)}_{\mathbf{V},i}(F) = \left[\frac{|\mathbf{V}_i(F_i)|}{|\mathbf{V}_1(F_1)|}\right]^{\frac{1}{p}} \qquad \rho^{(2)}_{\mathbf{V},1}(F) = 1 \tag{19}$$

$$\rho^{(3)}_{\mathbf{V},i}(F) = \frac{tr(\mathbf{V}_1(F_1)^{-1}\mathbf{V}_i(F_i))}{p} \qquad \rho^{(3)}_{\mathbf{V},1}(F) = 1 \tag{20}$$

are given by

$$PIF_i(\mathbf{x}, \rho^{(s)}_{\mathbf{V},\ell}, F) = A_i(\mathbf{x})\,(1-\delta_{i1})\,\delta_{\ell i} - \rho_\ell A_1(\mathbf{x})\delta_{i1}\,(1-\delta_{\ell i}) \qquad s = 2,3 \tag{21}$$

From Theorem 2.2, we conclude that the three estimates of the proportionality constants have the same partial influence functions and therefore, they will have the same asymptotic variance.

From the expressions given for the partial influence functions in Theorem 2.1 and using (11), one obtains the expressions for the asymptotic variance of the plug–in estimates solution of (2) to (5), which were derived in Boente and Orellana (2004), when the estimates of the scatter matrix $\boldsymbol{\Sigma}_i$ are asymptotically normally distributed and spherically invariant.

2.2 Partial influence functions for the Projection–Pursuit estimates

Let $\mathbf{x}_i$ be independent vectors such that $\mathbf{x}_i \sim F_i$, where F_i has location parameter $\boldsymbol{\mu}_i$ and scatter matrix $\boldsymbol{\Sigma}_i = \mathbf{C}_i\mathbf{C}_i'$ satisfying (1). Denote by $F_i[\mathbf{b}]$ the distribution of $\mathbf{b}'(\mathbf{x}_i - \boldsymbol{\mu}_i)$ and by F the product measure, $F = F_1 \times \ldots \times F_k$. For the sake of simplicity and without loss of generality, we will assume $\boldsymbol{\mu}_i = 0$ when computing influence functions. Two projection pursuit functionals will be studied, one related to the proposal given by Boente and Orellana (2001) under a CPC model and another one adapted to the proportional model as described in the Introduction.

Let $\varsigma(\mathbf{b}) = \sum_{i=1}^{k} \tau_i\, \sigma^2(F_i[\mathbf{b}])$ and $\varsigma_\rho(\mathbf{b}) = \sum_{i=1}^{k} \frac{\tau_i}{\rho_i(F)}\, \sigma^2(F_i[\mathbf{b}])$, where $\sigma(\cdot)$ is a univariate scale estimator, which is equivariant under scale transformations and $\rho_i(F)$ are functionals related to preliminary estimates of the proportionality constant.

The functionals related to the three proposals for the proportionality constants can be obtained as follows. First, the projection–pursuit functional for the common directions $\mathbf{B}_\sigma(F) = (\boldsymbol{\beta}_{\sigma,1}(F), \ldots, \boldsymbol{\beta}_{\sigma,p}(F))$ are derived as the solution of

$$\begin{cases} \boldsymbol{\beta}_{\sigma,1}(F) & = \underset{\|\mathbf{b}\|=1}{\operatorname{argmax}}\, \varsigma(\mathbf{b}) \\ \boldsymbol{\beta}_{\sigma,j}(F) & = \underset{\mathbf{b}\in\mathcal{B}_j}{\operatorname{argmax}}\, \varsigma(\mathbf{b}) \quad 2 \le j \le p\,, \end{cases} \tag{22}$$

with $\mathcal{B}_j = \{\mathbf{b} : \|\mathbf{b}\| = 1, \mathbf{b}'\boldsymbol{\beta}_{\sigma,m}(F) = 0 \text{ for } 1 \le m \le j-1\}$. The eigenvalues and the covariance matrices functionals are defined as $\lambda_{\sigma,ij}(F) = \sigma^2\left(F_i\left[\boldsymbol{\beta}_{\sigma,j}(F)\right]\right)$ and $\mathbf{V}_{\sigma,i}(F) = \sum_{j=1}^{p} \lambda_{\sigma,ij}(F)\boldsymbol{\beta}_{\sigma,j}(F)\boldsymbol{\beta}_{\sigma,j}(F)'$ respectively. Boente, Pires and Rodrigues (2002) derived the PIF of the principal directions defined by this procedure, which will be reminded in the formula (29) in Theorem 2.3.

The functionals related to the estimates for the proportionality constants and for the eigenvalues of the first population can be obtained by solving

$$\rho^{(1)}_{\sigma,i}(F) = \frac{1}{p}\sum_{j=1}^{p} \frac{\boldsymbol{\beta}_{\sigma,j}(F)'\,\mathbf{V}_{\sigma,i}(F)\,\boldsymbol{\beta}_{\sigma,j}(F)}{\lambda_{\sigma,j}(F)} = \frac{1}{p}\sum_{j=1}^{p}\frac{\lambda_{\sigma,ij}(F)}{\lambda_{\sigma,j}(F)} \quad i = 2, \ldots, k \quad \rho_{\sigma,1}(F) = 1 \tag{23}$$

$$\lambda_{\sigma,j}(F) = \sum_{i=1}^{k} \frac{\tau_i}{\rho^{(1)}_{\sigma,i}(F)}\boldsymbol{\beta}_{\sigma,j}(F)'\,\mathbf{V}_{\sigma,i}(F)\,\boldsymbol{\beta}_{\sigma,j}(F) = \frac{1}{p}\sum_{i=1}^{k}\frac{\tau_i}{\rho^{(1)}_{\sigma,i}(F)}\lambda_{\sigma,ij}(F) \tag{24}$$

Also, **Proposals 2** and **3** can be considered using the scatter matrices $\mathbf{V}_{\sigma,i}(F)$ defined above. More precisely, the related functionals for the proportionality constants are

$$\rho^{(2)}_{\sigma,i}(F) = \left[\frac{|\mathbf{V}_{\sigma,i}(F)|}{|\mathbf{V}_{\sigma,1}(F)|}\right]^{\frac{1}{p}} = \left[\prod_{j=1}^{p}\frac{\lambda_{\sigma,ij}(F)}{\lambda_{\sigma,1j}(F)}\right]^{\frac{1}{p}} \qquad \rho^{(2)}_{\sigma,1}(F) = 1 \tag{25}$$

$$\rho^{(3)}_{\sigma,i}(F) = \frac{1}{p}\operatorname{tr}(\mathbf{V}_{\sigma,1}(F)^{-1}\mathbf{V}_{\sigma,i}(F)) = \frac{1}{p}\sum_{j=1}^{p}\frac{\lambda_{\sigma,ij}(F)}{\lambda_{\sigma,1j}(F)} \qquad \rho^{(3)}_{\sigma,1}(F) = 1 \tag{26}$$

The following Theorem gives the partial influence functions, under proportionality of the scatter matrices, for the common eigenvectors, the eigenvalues of the first populations and for each proposal for the proportionality constants.

Theorem 2.3. *Let $\mathbf{x}_i$ be independent random vectors with ellipsoidal distribution F_i, with location parameters $\boldsymbol{\mu}_i = 0$ and scatter matrices $\boldsymbol{\Sigma}_i = \mathbf{C}_i\mathbf{C}_i'$ satisfying (1) and such that $\mathbf{C}_i^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_i) = \mathbf{z}_i$ has the same* *spherical distribution G for all $1 \le i \le k$. Assume that $\sigma(G_0) = 1$ where G_0 is the distribution of z_{11} and that $\boldsymbol{\Sigma} = \boldsymbol{\beta} \text{diag}\left(\xi_1, \ldots, \xi_p\right) \boldsymbol{\beta}'$ where $\xi_j = \lambda_j \sum_{i=1}^k \tau_i \rho_i$ and $\lambda_1 > \lambda_2 > \ldots > \lambda_p$.*

Then, if the function $(\epsilon, y) \to \sigma\left((1-\epsilon) G_0 + \epsilon \delta_y\right)$ is twice continuously differentiable at $(0, y)$, we have that for any $\mathbf{x}$ the partial influence functions of the projection–pursuit estimates for the common directions defined through (22) and for the solution $(\rho^{(1)}_{\sigma,\,i}(F), \lambda_{\sigma,\,j}(F))$ of (23) and (24) are given by

$$\begin{aligned} PIF_i(\mathbf{x}, \lambda_{\sigma,\,j}, F) &= \frac{\tau_i}{\rho_i} PIF_i\left(\mathbf{x}, \lambda_{\sigma,\,ij}, F\right) - \frac{\tau_i}{\rho_i} \lambda_j A_i(\mathbf{x}) + \lambda_j A_1(\mathbf{x}) \delta_{i1} \qquad 1 \le j \le p & (27)\\ PIF_i(\mathbf{x}, \rho^{(1)}_{\sigma,\,\ell}, F) &= A_i(\mathbf{x})\ (1-\delta_{i1})\, \delta_{\ell i} - \rho_\ell A_1(\mathbf{x}) \delta_{i1} \left(1 - \delta_{\ell i}\right) \qquad 2 \le \ell \le k & (28)\\ PIF_i(\mathbf{x}, \boldsymbol{\beta}_{\sigma,\,j}, F) &= \frac{\tau_i \sqrt{\rho_i}}{\sum_{\ell=1}^k \tau_\ell \rho_\ell} \left[\sqrt{\lambda_j}\ DIF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\rho_i \lambda_j}}; \sigma, G_0\right) \sum_{s=j+1}^p \frac{1}{\lambda_j - \lambda_s} \left(\mathbf{x}'\boldsymbol{\beta}_s\right) \boldsymbol{\beta}_s + \right. \\ &+ \left. \left(\mathbf{x}'\boldsymbol{\beta}_j\right) \sum_{s=1}^{j-1} \frac{\sqrt{\lambda_s}}{\lambda_j - \lambda_s}\ DIF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_s}{\sqrt{\rho_i \lambda_s}}; \sigma, G_0\right) \boldsymbol{\beta}_s \right] \qquad 1 \le j \le p & (29) \end{aligned}$$

where $A_i(\mathbf{x}) = \frac{1}{p} \sum_{j=1}^p \frac{PIF_i\left(\mathbf{x}, \lambda_{\sigma,\,ij}, F\right)}{\lambda_j}$ and $PIF_i\left(\mathbf{x}, \lambda_{\sigma,\,ij}, F\right) = 2\rho_i \lambda_j IF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\rho_i \lambda_j}}, \sigma, G_0\right)$. Moreover, the partial influence functions of $\rho^{(2)}_{\sigma,\,i}(F)$ and $\rho^{(3)}_{\sigma,\,i}(F)$ defined in (25) and (26), are also given by (28).

Let us now consider a preliminary Fisher–consistent functional for the proportionality constants and the functional $\varsigma_\rho(\mathbf{b})$. The projection–pursuit functional for the common directions $\mathbf{B}_{\sigma,\rho}(F) = (\boldsymbol{\beta}_{\sigma,\rho,1}(F), \ldots, \boldsymbol{\beta}_{\sigma,\rho,p}(F))$ are obtained as the solution of

$$\left\{ \begin{array}{ll} \boldsymbol{\beta}_{\sigma,\rho,\,1}(F) &= \underset{\|\mathbf{b}\|=1}{\text{argmax}}\ \varsigma_\rho(\mathbf{b}) \\ \boldsymbol{\beta}_{\sigma,\rho,\,j}(F) &= \underset{\mathbf{b} \in \mathcal{B}_j}{\text{argmax}}\ \varsigma_\rho(\mathbf{b}) \quad 2 \le j \le p\ , \end{array} \right. \tag{30}$$

while the functionals for the eigenvalues and the covariance matrices functionals are $\lambda_{\sigma,\rho,\,ij}(F) = \sigma^2\left(F_i\left[\boldsymbol{\beta}_{\sigma,\rho,\,j}(F)\right]\right)$ and $\mathbf{V}_{\sigma,\,i}(F) = \sum_{j=1}^p \lambda_{\sigma,\rho,\,ij}(F) \boldsymbol{\beta}_{\sigma,\rho,\,j}(F) \boldsymbol{\beta}_{\sigma,\rho,\,j}(F)'$, respectively. As above, the functionals related to the estimates for the proportionality constants and for the eigenvalues of the first population solve

$$\begin{aligned} \rho^{(1)}_{\sigma,\rho,\,i}(F) &= \frac{1}{p} \sum_{j=1}^p \frac{\lambda_{\sigma,\rho,\,ij}(F)}{\lambda_{\sigma,\rho,\,j}(F)} \qquad i = 2, \ldots, k \qquad \rho_{\sigma,\rho,\,1}(F) = 1 & (31)\\ \lambda_{\sigma,\rho,\,j}(F) &= \frac{1}{p} \sum_{i=1}^k \frac{\tau_i}{\rho^{(1)}_{\sigma,\rho,\,i}(F)} \lambda_{\sigma,\rho,\,ij}(F) & (32) \end{aligned}$$

Also, **Proposals 2** and **3** can be considered using the scatter matrices $\mathbf{V}_{\sigma,\rho,\,i}(F)$ defined above, leading to functionals $\rho^{(2)}_{\sigma,\rho,\,i}(F)$ and $\rho^{(3)}_{\sigma,\rho,\,i}(F)$.

The following Theorem shows that the preliminary estimate of the proportionality constant, modifies only the partial influence functions of the common direction functionals, while those related to the eigenvalues of the first population and to the proportionality constants remain the same.

Theorem 2.4. *Let $\mathbf{x}_i$ be independent random vectors with ellipsoidal distribution F_i, with location parameters $\boldsymbol{\mu}_i = 0$ and scatter matrices $\boldsymbol{\Sigma}_i = \mathbf{C}_i \mathbf{C}_i'$ satisfying (1) and such that $\mathbf{C}_i^{-1}\left(\mathbf{x}_i - \boldsymbol{\mu}_i\right) = \mathbf{z}_i$ has the* *same spherical distribution G for all $1 \leq i \leq k$. Assume that $\sigma(G_0) = 1$ where G_0 is the distribution of z_{11}, $\rho_i(F) = \rho_i$ and that $\mathbf{\Sigma}_1 = \boldsymbol{\beta}\text{diag}\,(\lambda_1, \ldots, \lambda_p)\,\boldsymbol{\beta}'$ where $\lambda_1 > \lambda_2 > \ldots > \lambda_p$. Moreover, assume that $PIF_i\,(\mathbf{x}, \rho_\ell, F)$ exists.*

Then, if the function $(\epsilon, y) \to \sigma\left((1-\epsilon)G_0 + \epsilon\delta_y\right)$ is twice continuously differentiable at $(0, y)$, we have that for any $\mathbf{x}$ the partial influence functions of the projection–pursuit estimates for the common directions defined by (30) and for the solution $(\rho^{(1)}_{\sigma,\rho,\,i}(F), \lambda_{\sigma,\rho,\,j}(F))$ of (31) and (32) are given by

$$PIF_i(\mathbf{x}, \lambda_{\sigma,\rho,\,j}, F) = \frac{\tau_i}{\rho_i} PIF_i\,(\mathbf{x}, \lambda_{\sigma,\rho,\,ij}, F) - \frac{\tau_i}{\rho_i}\lambda_j A_i(\mathbf{x}) + \lambda_j A_1(\mathbf{x})\delta_{i1} \qquad 1 \leq j \leq p \tag{33}$$

$$PIF_i(\mathbf{x}, \rho^{(1)}_{\sigma,\rho,\,\ell}, F) = A_i(\mathbf{x})\,\left(1 - \delta_{i1}\right)\delta_{\ell i} - \rho_\ell A_1(\mathbf{x})\delta_{i1}\left(1 - \delta_{\ell i}\right) \qquad 2 \leq \ell \leq k \tag{34}$$

$$\begin{aligned} PIF_i(\mathbf{x}, \boldsymbol{\beta}_{\sigma,\rho,\,j}, F) &= \frac{\tau_i}{\sqrt{\rho_i}}\left[\sqrt{\lambda_j}\; DIF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\rho_i\lambda_j}}; \sigma, G_0\right) \sum_{s=j+1}^{p} \frac{1}{\lambda_j - \lambda_s}\left(\mathbf{x}'\boldsymbol{\beta}_s\right)\boldsymbol{\beta}_s + \right. \\ &+ \left. \left(\mathbf{x}'\boldsymbol{\beta}_j\right)\sum_{s=1}^{j-1}\frac{\sqrt{\lambda_s}}{\lambda_j - \lambda_s}\; DIF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_s}{\sqrt{\rho_i\lambda_s}}; \sigma, G_0\right)\boldsymbol{\beta}_s\right] \qquad 1 \leq j \leq p \end{aligned} \tag{35}$$

where $A_i(\mathbf{x}) = \dfrac{1}{p}\displaystyle\sum_{j=1}^{p} \frac{PIF_i\,(\mathbf{x}, \lambda_{\sigma,\rho,\,ij}, F)}{\lambda_j}$ and $PIF_i\,(\mathbf{x}, \lambda_{\sigma,\rho,\,ij}, F) = 2\rho_i\lambda_j IF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\rho_i\lambda_j}}, \sigma, G_0\right)$. Moreover, the partial influence functions of $\rho^{(2)}_{\sigma,\rho,\,i}(F)$ and $\rho^{(3)}_{\sigma,\rho,\,i}(F)$ defined as in (25) and (26) respectively, are also given by (34).

Remark 2.3. As for the plug–in estimates, Theorem 2.3 and 2.4 show that the partial influence functions for the eigenvalues add an extra term to the influence function of the eigenvalues for the one population case studied by Croux and Ruiz–Gazen (2000). On the other hand, the partial influence functions for the proportionality constants can again be viewed as a weighted average of the partial influence functions of the eigenvalues of $\mathbf{V}_{\sigma,\,i}$. As in the one–population case, by using a scale estimator with bounded influence, we get bounded influences for the proportionality constants and for the eigenvalues. However, as mentioned by Boente, Pires and Rodrigues (2002) the partial influence functions of the eigenvectors may be unbounded.

Remark 2.4. When $\sigma^2(F) = \text{VAR}(F)$, Theorem 2.3 entails that the partial influence functions of the projection–pursuit estimates of the eigenvalues of the first population and of the proportionality constants are those given in Section 2.1 for the maximum likelihood estimates. On the other hand, as noted by Boente, Pires and Rodrigues (2002) the partial influence functions of the common directions using the variance are not those of the maximum likelihood, but they are the influence functions for the eigenvalues of the pooled matrix, except for the factor τ_i.

2.3 Partial influence functions of the proportion of the variance explained

A useful index in principal component analysis is the proportion of the variance explained by the $j-$th component, $\nu_j = \lambda_j \left\{\sum_{\ell=1}^{p} \lambda_\ell\right\}^{-1}$. From Theorems 2.1, 2.3 and 2.4, we easily obtain the partial influence function of ν_j.

Proposition 2.1. *Denote $\nu_{\mathbf{V},j}(F) = \lambda_{\mathbf{V},j}(F)\left\{\sum_{\ell=1}^{p} \lambda_{\mathbf{V},\ell}(F)\right\}^{-1}$, $\nu_{\sigma,\,j}(F) = \lambda_{\sigma,\,j}(F)\left\{\sum_{\ell=1}^{p} \lambda_{\sigma,\,\ell}(F))\right\}^{-1}$ and $\nu_{\sigma,\rho,\,j}(F) = \lambda_{\sigma,\rho,\,j}(F)\left\{\sum_{\ell=1}^{p} \lambda_{\sigma,\rho,\,\ell}(F))\right\}^{-1}$. Moreover, let $B_{ij}(\mathbf{x}, \mathbf{V}) = \boldsymbol{\beta}_j' IF(\mathbf{x}, \mathbf{V}_i, F_i)\boldsymbol{\beta}_j$ and $B_{ij}(\mathbf{x}, \sigma) =$*

$2\rho_i\lambda_j IF\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\rho_i\lambda_j}}, \sigma, G_0\right)$. *Then,*

(a) *Under the conditions of Theorem 2.1, the partial influence function of $\nu_{\mathbf{V},j}(F)$ is given by*

$$PIF_i(\mathbf{x}, \nu_{\mathbf{V},j}, F) = \frac{\tau_i}{\rho_i}\left\{B_{ij}(\mathbf{x},\mathbf{V})\sum_{\ell\neq j}\lambda_\ell - \lambda_j\sum_{\ell\neq j}B_{i\ell}(\mathbf{x},\mathbf{V})\right\}\left\{\sum_{\ell=1}^{p}\lambda_\ell\right\}^{-2}.$$

(b) *Under the conditions of Theorems 2.2 or 2.3 the partial influence function of $\nu_{\sigma,j}(F)$ or $\nu_{\sigma,\rho,j}(F)$ are both given by*

$$PIF_i(\mathbf{x}, \nu_{\sigma,j}, F) = PIF_i(\mathbf{x}, \nu_{\sigma,\rho,j}, F)\frac{\tau_i}{\rho_i}\left\{B_{ij}(\mathbf{x},\sigma)\sum_{\ell\neq j}\lambda_\ell - \lambda_j\sum_{\ell\neq j}B_{i\ell}(\mathbf{x},\sigma)\right\}\left\{\sum_{\ell=1}^{p}\lambda_\ell\right\}^{-2}.$$

Remark 2.5. By using a scatter matrix with bounded influence or a scale estimator with bounded influence, we get bounded influences for the proportion of the variance explained by the $j-$th component. On the other hand, for the maximum likelihood estimates we get unbounded partial influence functions since

$$\mathrm{PIF}_i(\mathbf{x}, \nu_{\mathrm{ML},j}, F) = \frac{\tau_i}{\rho_i}\left\{z_{ij}^2\sum_{\ell=1}^{p}\lambda_\ell - \lambda_j\|\mathbf{z}_i\|^2\right\}\left\{\sum_{\ell=1}^{p}\lambda_\ell\right\}^{-2},$$

where $z_{ij} = \boldsymbol{\beta}_j'(\mathbf{x}-\boldsymbol{\mu}_i)$ and $\mathbf{z}_i = (z_{i1},\dots,z_{ip})'$. It is worthwhile noticing that even for large values of $\mathbf{x}$, along some directions the maximum likelihood estimators will lead to values of $\mathrm{PIF}_i(\mathbf{x}, \nu_{\mathrm{ML},j}, F)$ equal to 0 and so outliers along these directions do not influence the percentage of variance explained by the $j-$th component.

Figures 1 to 4 give the plots of the norm of the partial influence function PIF_1 of the first eigenvector and the partial influence function PIF_1 for the proportionality constant and the eigenvalues when $p = 2$ 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 5 shows the partial influence function PIF_1 for the explained proportion $\nu_1 = \frac{\lambda_1}{\lambda_1+\lambda_2}$. We have considered the maximum likelihood estimators, 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. In all cases, as in the one–population case, the shape of the partial influence functions of the robust estimates is comparable to that of their classical relatives at the center of the distribution, while the influence at points further away is downweighted for the robust estimates while it is much more larger for the classical ones. Figures 2, 3 and 4 confirm the boundedness of the partial influence functions of both the plug–in and the projection–pursuit estimates. On the other hand, for the first eigenvector, the norm of the partial influence function of the plug–in estimate, as in the case $k = 1$, is largest along the bisectors while, as noted before, the projection–pursuit estimates can still attain large values but only for smaller values of x_1 combined with large values of x_2.

Figures 1 to 5 around here.

3 Asymptotic variances for the Projection–Pursuit estimates

When all the populations have a Gaussian distribution, the asymptotic variances of these estimates turn out to be particularly simple. From Corollary 1 in Boente, Pires and Rodrigues (2002) we obtain that under the conditions of Proposition 3.1, the asymptotic variance of the projection–pursuit estimates of the common eigenvectors is given by

$$\text{ASVAR}\left(\widehat{\boldsymbol{\beta}}_{jm}\right) = \frac{\sum_{i=1}^{k} \tau_i \rho_i^2}{\left(\sum_{i=1}^{k} \tau_i \rho_i\right)^2} \frac{\lambda_j \lambda_m}{(\lambda_j - \lambda_m)^2} E_G \left[\text{DIF}\left(z_{1j}; \sigma, G_0\right) z_{1m}\right]^2 \qquad \text{for } m \neq j \ . \tag{36}$$

In particular, when $G = N(\mathbf{0}, \mathbf{I}_p)$, we have that $\text{ASCOV}\left(\widehat{\boldsymbol{\beta}}_{jm}, \widehat{\boldsymbol{\beta}}_{jr}\right) = 0$ for $m \neq j$, $m \neq r$, $r \neq j$ and

$$\text{ASVAR}\left(\widehat{\boldsymbol{\beta}}_{jm}\right) = \frac{\sum_{i=1}^{k} \tau_i \rho_i^2}{\left(\sum_{i=1}^{k} \tau_i \rho_i\right)^2} \frac{\lambda_j \lambda_m}{(\lambda_j - \lambda_m)^2} E_\Phi \left[\text{DIF}\left(Y; \sigma, \Phi\right)\right]^2 \qquad \text{for } m \neq j$$

The asymptotic variance of the projection–pursuit estimates of the eigenvalues of the first population and of the proportionality constants can be obtained heuristically using (11) and Theorem 2.3.

Proposition 3.1. *Let $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}$, $1 \leq i \leq k$, be independent observations from k independent samples with distribution F_i, location parameter $\boldsymbol{\mu}_i = 0$ and scatter matrix $\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}_i = \rho_i \boldsymbol{\Lambda}_1$, i.e., $\boldsymbol{\Sigma}_i$ satisfies (1) with $\boldsymbol{\beta} = \mathbf{I}_p$. Let $n_i = \tau_i N$ with $0 < \tau_i < 1$ and $\sum_{i=1}^{k} \tau_i = 1$.*

Moreover, assume that $\boldsymbol{\Lambda}_i^{-\frac{1}{2}} \mathbf{x}_{i1} = \mathbf{z}_i$ has the same spherical distribution G for all $1 \leq i \leq k$. Assume that $\sigma(G_0) = 1$ where G_0 is the distribution of z_{11} and that $\boldsymbol{\Sigma} = diag(\xi_1, \ldots, \xi_p)$ where $\xi_1 > \xi_2 > \ldots > \xi_p$.

Let $s(\cdot)$ be a univariate robust scale statistic and $\underline{\mathbf{X}}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i})$, for $1 \leq i \leq k$. Define the estimates $\widehat{\boldsymbol{\beta}}_j$ of the common principal axes by solving iteratively (8).

Let the estimates of the eigenvalues and the covariance matrix of the i-th population be

$$\widehat{\lambda}_{ij} = s^2(\underline{\mathbf{X}}_i' \widehat{\boldsymbol{\beta}}_j) \quad \text{for } 1 \leq j \leq p \qquad \mathbf{V}_i = \sum_{j=1}^{p} \widehat{\lambda}_{ij} \widehat{\boldsymbol{\beta}}_j \widehat{\boldsymbol{\beta}}_j' \ .$$

Define the following estimates for the proportionality constants and for the eigenvalues of the first population

$$\widehat{\rho}_i = \frac{1}{p} \sum_{j=1}^{p} \frac{\widehat{\lambda}_{ij}}{\widehat{\lambda}_j} \quad i = 2, \ldots, k \quad \widehat{\rho}_1 = 1$$

$$\widehat{\lambda}_j = \frac{1}{N} \sum_{i=1}^{k} \frac{n_i}{\widehat{\rho}_i} \widehat{\lambda}_{ij}$$

and the following two other estimates of the proportionality constants

$$\widehat{\rho}_i^{(2)} = \left[\frac{|\mathbf{V}_i|}{|\mathbf{V}_1|}\right]^{\frac{1}{p}} = \left[\prod_{j=1}^{p} \frac{\widehat{\lambda}_{ij}}{\widehat{\lambda}_{1j}}\right]^{\frac{1}{p}} \qquad \widehat{\rho}_1^{(2)} = 1$$

$$\widehat{\rho}_i^{(3)} = \frac{tr(\mathbf{V}_1^{-1} \mathbf{V}_i)}{p} = \frac{1}{p} \sum_{j=1}^{p} \frac{\widehat{\lambda}_{ij}}{\widehat{\lambda}_{1j}} \qquad \widehat{\rho}_1^{(3)} = 1 \ .$$

Then, the asymptotic variances of the projection–pursuit estimates of the proportionality constants and of the eigenvalues of the first population are given by

$$
\begin{aligned}
ASCOV\left(\widehat{\lambda}_j,\widehat{\lambda}_s\right) &= 4\lambda_j\lambda_s\left\{\frac{1}{\tau_1}\frac{\gamma}{p}-\frac{\kappa}{p}+\kappa\delta_{js}\right\} \qquad 1\le j\ ,\ s\le p\\
ASVAR\left(\widehat{\rho}_i\right)=ASVAR\left(\widehat{\rho}_i^{(2)}\right) &= ASVAR\left(\widehat{\rho}_i^{(3)}\right)=\frac{4\gamma}{p}\left(\frac{1}{\tau_1}+\frac{1}{\tau_i}\right)\rho_i^2 \qquad 2\le i\le k\ ,
\end{aligned}
$$

where $\kappa = ASVAR\left(\sigma,G_0\right)-COV_G\left(IF\left(z_{11},\sigma,G_0\right),IF\left(z_{12},\sigma,G_0\right)\right)$ *and* $\gamma=\kappa+p\ COV_G\left(IF\left(z_{11},\sigma,G_0\right),IF\left(z_{12},\sigma,G_0\right)\right)$. *Moreover, the asymptotic variance of the estimates,* $\widehat{\nu}_j=\widehat{\lambda}_j\left\{\sum_{\ell=1}^p\widehat{\lambda}_\ell\right\}^{-1}$, *of the proportion of variance explained by the* $j-th$ *component is given by*

$$
ASCOV\left(\widehat{\nu}_j,\widehat{\nu}_s\right) = 4\,\kappa\,\nu_j\nu_s\left(\sum_{\ell=1}^p\nu_\ell^2-\nu_j-\nu_s+\delta_{js}\right) \qquad 1\le j\ ,\ s\le p\ .
$$

In particular, when $G=N\left(\mathbf{0},\mathbf{I}_p\right)$, *we have that*

$$
\begin{aligned}
ASVAR\left(\widehat{\lambda}_j\right) &= 4ASVAR\left(\sigma,G_0\right)\left(1-\frac{1}{p}+\frac{1}{p\tau_1}\right)\lambda_j^2 \qquad 1\le j\le p\\
ASVAR\left(\widehat{\nu}_j\right) &= 4\,ASVAR\left(\sigma,G_0\right)\nu_j^2\left[\left(1-\nu_j\right)^2+\sum_{\ell\ne j}\nu_\ell^2\right] \qquad 1\le j\le p\\
ASVAR\left(\widehat{\rho}_i\right) &= ASVAR\left(\widehat{\rho}_i^{(2)}\right)=ASVAR\left(\widehat{\rho}_i^{(3)}\right)=\frac{4}{p}ASVAR\left(\sigma,G_0\right)\left(\frac{1}{\tau_1}+\frac{1}{\tau_i}\right)\rho_i^2 \qquad i=2,\dots,k\ ,
\end{aligned}
$$

Remark 3.1. As mentioned above, Boente and Orellana (2004) derived the asymptotic distribution of the plug–in estimates of the eigenvalues of the first population, when the estimates of the scatter matrix $\mathbf{\Sigma}_i$ are asymptotically normally distributed and spherically invariant. If all populations have the same ellipsoidal distribution except for changes in the scatter matrices and we consider the same family of robust scatter estimators for all of them, using the expressions of the asymptotic covariances they have obtained or the heuristic expression (11) for the estimates of ν_j, we easily obtain that their asymptotic covariances are analogous to those given in Proposition 3.1 replacing $\kappa=\text{ASVAR}\left(\sigma,G_0\right)-\text{COV}_G\left(\text{IF}\left(z_{11},\sigma,G_0\right),\text{IF}\left(z_{12},\sigma,G_0\right)\right)$ by $\kappa=\frac{\sigma_1}{2}$, where σ_1 gives the efficiencies of the off-diagonal elements of the matrices $\mathbf{V}_i$. Moreover, from these expressions for the asymptotic covariances, one can easily derive the asymptotic variance of the estimates of the total amount of variance explained by considering the first q common directions as

$$
\text{ASVAR}\left(\sum_{j=1}^q\widehat{\nu}_j\right)=4\kappa\left\{\sum_{\ell=1}^p\nu_\ell^2\left(\sum_{\ell=1}^q\nu_\ell\right)^2+\sum_{\ell=1}^q\nu_\ell^2\left(1-2\sum_{\ell=1}^q\nu_\ell\right)\right\}\ ,
$$

which would allow to make inference on the number of common directions to be selected.

Proposition 3.2. *Let* $\mathbf{x}_{i1},\dots,\mathbf{x}_{in_i}$, $1\le i\le k$, *be independent observations from* k *independent samples with distribution* F_i, *location parameter* $\boldsymbol{\mu}_i=0$ *and scatter matrix* $\mathbf{\Sigma}_i=\mathbf{\Lambda}_i=\rho_i\mathbf{\Lambda}_1$, *i.e.,* $\mathbf{\Sigma}_i$ *satisfies (1) with* $\boldsymbol{\beta}=\mathbf{I}_p$. *Let* $n_i=\tau_iN$ *with* $0<\tau_i<1$ *and* $\sum_{i=1}^k\tau_i=1$.

Moreover, assume that $\mathbf{\Lambda}_i^{-\frac{1}{2}}\mathbf{x}_{i1}=\mathbf{z}_i$ *has the same spherical distribution* G *for all* $1\le i\le k$. *Assume that* $\sigma(G_0)=1$ *where* G_0 *is the distribution of* z_{11} *and that* $\mathbf{\Sigma}_1=diag\left(\lambda_1,\dots,\lambda_p\right)$ *where* $\lambda_1>\dots>\lambda_p$.

Let $s(\cdot)$ be a univariate robust scale statistic and $\underline{\mathbf{X}}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i})$, for $1 \le i \le k$. Define the estimates $\widehat{\boldsymbol{\beta}}_j$ of common principal axes by solving iteratively (10).

Let the estimates of the eigenvalues and the covariance matrix of the i-th population be

$$\widehat{\lambda}_{ij} = s^2(\underline{\mathbf{X}}_i' \widehat{\boldsymbol{\beta}}_j) \quad \text{for } 1 \le j \le p \qquad \mathbf{V}_i = \sum_{j=1}^{p} \widehat{\lambda}_{ij} \widehat{\boldsymbol{\beta}}_j \widehat{\boldsymbol{\beta}}_j' \ .$$

Define the estimates for the proportionality constants and for the eigenvalues of the first population as in Proposition 3.1.

Then, the asymptotic variances of the projection–pursuit estimates of the proportionality constants and of the eigenvalues of the first population are those given in Proposition 3.1, while

$$\mathit{ASVAR}\left(\widehat{\boldsymbol{\beta}}_{jm}\right) = \frac{\lambda_j \lambda_m}{(\lambda_j - \lambda_m)^2} E_G \left[\mathit{DIF}\left(z_{1j}; \sigma, G_0\right) z_{1m}\right]^2 \qquad \text{for } m \ne j \ .$$

Remark 3.2. Note that, when $G = N(\mathbf{0}, \mathbf{I}_p)$, the efficiencies of both the eigenvalues and the proportionality constants estimates with respect to the maximum likelihood estimates are given by $\dfrac{1}{2\mathrm{ASVAR}\,(\sigma, G_0)}$, which is the efficiency of the scale estimator used in the projection pursuit procedure. Moreover, it is worthwhile noticing that for the projection–pursuit estimates the efficiency of the eigenvalue estimates do not depend on the relative size among populations as it does for the plug–in ones. On the other hand, the plug–in estimates of the principal directions have the same efficiencies as the off–diagonal elements of the scatter matrices used while, as shown by (36) for the projection–pursuit estimates defined in (8), the efficiencies do depend on the relative size among populations. Note that by introducing the preliminary estimates $\widehat{\rho}_i$, we solve this problem and we attain the same efficiencies as in the one–population setting.

4 Monte Carlo Study

We performed a simulation study in order to compare the behavior of the different proposals for small sample sizes by considering $k = 2$ populations in dimension $p = 4$. The behavior of the following estimates was studied: the plug–in estimates using the Donoho–Stahel matrix (Donoho (1982)–Stahel (1981)) or the M-scatter estimate (Maronna (1976)) and two projection–pursuit estimates using different scale estimates. The performance of these estimates is compared with that of maximum likelihood estimates. In all the Tables and Figures MLE denotes the maximum likelihood estimates, ME and DSE the estimates obtained by plugging–in an M-scatter matrix (with Huber score function with tuning constant 7.6176) and the Donoho–Stahel matrix (with tuning constant $\sqrt{\chi^2_p(0.95)} = 3.0803$) respectively in (2) to (4). Meanwhile, PPE$_i$ denotes the projection pursuit estimates. PPE$_1$ was obtained using the MAD, PPE$_2$ using an scale M–estimate with score function $\chi(t) = \min\left(\frac{t^2}{c^2}, 1\right) - \frac{1}{2}$ and $c = 1.041$ which gives a scale estimate with breakdown point $\frac{1}{2}$ and efficiency 0.509.

We have performed 1000 replications generating two independent samples of size $n_1 = n_2 = n = 100$ with covariance matrices $\mathbf{\Sigma}_2 = \rho \mathbf{\Sigma}_1$ with $\rho = 4$ and $\mathbf{\Sigma}_1 = \mathrm{diag}(4, 3, 2, 1)$. The results for normal data sets will be indicated by C_0 in the tables, while $C_{1,\epsilon}$ and by $C_{2,\epsilon}$ will denote the following two contaminations.

- $C_{1,\epsilon}$: For $i = 1, 2$, $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in}$, are i.i.d. $(1 - \epsilon) N(\mathbf{0}, \mathbf{\Sigma}_i) + \epsilon N(\mathbf{0}, 9\mathbf{\Sigma}_i)$. We present the results for $\epsilon = 0.1$. This contamination corresponds to inflating the covariance matrix and thus, will only affect the variance of the axes at each population but not their direction.
- $C_{2,\epsilon}$: For $i = 1, 2$, $\mathbf{x}_{i1}, \ldots, \mathbf{x}_{in}$, are i.i.d. $(1 - \epsilon) N(\mathbf{0}, \mathbf{\Sigma}_i) + \epsilon N(\boldsymbol{\mu}, \mathbf{\Sigma}_i)$ with $\boldsymbol{\mu} = (0, 0, 0, 10)'$. We present the results for $\epsilon = 0.05$ and $\epsilon = 0.1$. This contamination corresponds to contaminating the first population in the direction of the smaller eigenvalue with the aim to study changes in the directions.

Tables 1, 2 and Figures 6 and 7 summarize the results of the simulations for the common eigenvectors and the three estimates of the proportionality constants proposed in Section 2. Since the results for the eigenvalues of the first population are similar to those described for the CPC model in Boente and Orellana (2001), they are not reported here. Let $\widehat{\theta}_j$ be the angle between the j-th estimated and true direction. In Figure 6, the density estimates of $cos(\widehat{\theta}_4)$, evaluated using the normal kernel with a bandwidth equal to 0.3, are plotted. The plots given in black correspond to the densities of $cos(\widehat{\theta}_4)$ evaluated over the 1000 normally distributed samples, while those in light blue and in red correspond to the asymmetric contaminated samples generated according to $C_{2,0.05}$ and $C_{2,0.1}$, respectively. On the other hand, Table 1 gives as a summary measure for the eigenvectors estimation, the median over the replications of the distance between the estimated and the target $j-$th eigenvector, $\|\widehat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j\|^2$. Since $\boldsymbol{\Sigma}_1 = \text{diag}(4, 3, 2, 1)$, the common directions are the vectors of the canonical basis, $\mathbf{e}_j$, $j = 1, \ldots, 4$. In all cases we ordered the eigenvectors according to a decreasing order of the eigenvalues of the first population.

Table 2 gives means, standard deviations and mean square errors of the estimates for the proportionality constants, while Figure 7 shows the boxplots of $\log(\widehat{\rho}^{(j)}) - \log(\rho)$, for $j = 1$ to 3.

Some Comments. As was expected the maximum likelihood estimates perform poorly in the presence of outliers. With regard to eigenvectors estimation, Figure 6 shows, not only the poor behavior of the maximum likelihood estimates and of the estimates obtained using an M–scatter estimate in the presence of asymmetric outliers, but also the sensitivity of those obtained by using the Donoho–Stahel matrix. Moreover, as can be seen also in Table 1, under $C_{2,0.1}$, maximum likelihood and M–estimates interchange all the axes, while the estimates based on the Donoho–Stahel scatter matrix and the projection–pursuit estimates are less sensitive. With respect to the plug–in estimates based on the Donoho–Stahel matrix, the axis corresponding to the larger eigenvalues are not modified but the two other ones are rotated with a median rotation absolute angle around 58^o. Note that the direction of the smaller eigenvalue was the one we have choosen to contaminate the samples.

Table 1 and Figure 6

around here.

For the projection estimates, under the considered contaminations, the largest median rotation absolute angle is 51^o. $\boldsymbol{\Sigma}_1 = \text{diag}(4, 3, 2, 1)$ and thus even the projection–pursuit estimates do not break down by using the MAD scale estimate they will breakdown with a 13% of contamination (see Boente and Orellana (2001) for bias computations). Moreover, since the eigenvalues of $\boldsymbol{\Sigma}_1$ are quite close, the sample size does not allow to distinct between them and thus the projection pursuit estimates produce a systematic bias for normal data, specially when estimating the second and third direction.

Thus, when the aim is to obtain principal directions, under the model and contaminations considered, the plug–in procedure using the Donoho Stahel estimators was the best procedure for mild contaminations while the projection pursuit method seem to be the less sensitive for the higher ones.

Table 2 and Figure 7

around here.

With respect to the estimation of the proportionality constants, **Proposal 3** shows the largest bias under normality. In all situations, **Proposal 1** and **2** for estimating ρ performed similarly.

The main effect observed with maximum likelihood estimates is a negative bias in C_2, which can be explained since this contamination increases the smallest eigenvalue of $\boldsymbol{\Sigma}_1$ and then the relation between the two scatter matrices is modified. On the other hand, for normal and contaminated data, the projection–pursuit estimates based on the MAD are more biased than the other projection–procedures.

Under normal data sets, the statistic $\widehat{\rho}^{(3)}$ presents a slightly larger variance than $\widehat{\rho}^{(1)}$ and $\widehat{\rho}^{(2)}$, even if the asymptotic variance is the same for all of them. The larger variances observed for the projection pursuit estimates are due to the lack of efficiency of the scale–estimates used.

Under C_1 and $C_{2,0.05}$ the estimates of the proportionality constant obtained using the Donoho–Stahel scatter and those using the M–estimates perform similarly. Only a larger bias and standard deviation can be observed for the M–estimate in $C_{2,0.1}$.

Acknowledgment

The authors would like to thank David Tyler for stimulating discussions during the preparation of this paper.

- Graciela Boente
  Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales
  Ciudad Universitaria, Pabellón 2
  Buenos Aires, C1428EHA
  Argentina
  email: gboente@dm.uba.ar

A Appendix

From now on, denote by $F_{i,\epsilon,\mathbf{x}} = (1-\epsilon)F_i + \epsilon\delta_{\mathbf{x}}$ and by $F_{\epsilon,\mathbf{x},i} = F_1 \times \ldots \times F_{i-1} \times F_{i,\epsilon,\mathbf{x}} \times F_{i+1} \times \ldots \times F_k$. The influence functions for the eigenvalues of the first population and the proportionality constants solution of (12) and (13) or (23) and (24) will be derived from the following Lemma.

Lemma A.1. *Let $\psi_{ij}(G)$ be functionals such that $\psi_{ij}(F) = \rho_i\lambda_i$, where $F = F_1 \times \ldots \times F_k$. Assume that $\lambda_1 > \ldots > \lambda_p$ and that $PIF_i(\mathbf{x}, \psi_{\ell j}, F)$ exists and $PIF_i(\mathbf{x}, \psi_{\ell j}, F) = 0$, when $\ell \neq i$.*

Then, the partial influence functions of the solution $(\rho_i(G), \lambda_j(G))$ of

$$\rho_i(G) = \frac{1}{p}\sum_{j=1}^{p}\frac{\psi_{ij}(G)}{\lambda_j(G)} \qquad i = 2, \ldots, k \qquad \rho_1^{(1)}(G) = 1 \tag{A.1}$$

$$\lambda_j(G) = \sum_{i=1}^{k}\frac{\tau_i}{\rho_i(G)}\psi_{ij}(G) \qquad 1 \leq j \leq p \tag{A.2}$$

are given by

$$PIF_i(\mathbf{x}, \lambda_j, F) = \frac{\tau_i}{\rho_i}PIF_i(\mathbf{x}, \psi_{ij}, F) - \frac{\tau_i}{\rho_i}\lambda_j A_i(\mathbf{x}) + \lambda_j A_1(\mathbf{x})\delta_{i1} \qquad 1 \leq j \leq p \tag{A.3}$$

$$PIF_i(\mathbf{x}, \rho_\ell, F) = A_i(\mathbf{x})\,(1-\delta_{i1})\,\delta_{\ell i} - \rho_\ell A_1(\mathbf{x})\delta_{i1}\,(1-\delta_{\ell i}) \qquad 2 \leq \ell \leq k \tag{A.4}$$

where $A_i(\mathbf{x}) = \dfrac{1}{p}\displaystyle\sum_{j=1}^{p}\frac{PIF_i(\mathbf{x}, \psi_{ij}, F_i)}{\lambda_j}$ and $\delta_{\ell i} = 1$ when $i = \ell$ and $\delta_{\ell i} = 0$ when $i \neq \ell$.

Proof. Let $\rho_{\ell,\epsilon,i} = \rho_\ell(F_{\epsilon,\mathbf{x},i})$, $\psi_{\ell j,\epsilon,i} = \psi_{\ell j}(F_{\epsilon,\mathbf{x},i})$ and $\lambda_{j,\epsilon,i} = \lambda_j(F_{\epsilon,\mathbf{x},i})$. We have that

$$\rho_{\ell,\epsilon,i} = \frac{1}{p}\sum_{j=1}^{p}\frac{\psi_{\ell j,\epsilon,i}}{\lambda_{j,\epsilon,i}} \qquad \ell = 2, \ldots, k \qquad \ell \neq i \qquad \rho_{1,\epsilon,i} = 1 \tag{A.5}$$

$$\lambda_{j,\epsilon,i} = \sum_{\ell=1}^{k}\frac{\tau_\ell}{\rho_{\ell,\epsilon,i}}\psi_{\ell j,\epsilon,i}\;. \tag{A.6}$$

Deriving (A.5) and (A.6), we obtain $\mathrm{PIF}_i(\mathbf{x},\rho_1,F)=0$ and

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F) &= \frac{1}{p}\sum_{j=1}^{p}\frac{1}{\lambda_j}\left[\mathrm{PIF}_i(\mathbf{x},\psi_{\ell j},F)-\rho_\ell\mathrm{PIF}_i(\mathbf{x},\lambda_j,F)\right] \qquad \text{for } \ell\neq 1\\ \mathrm{PIF}_i(\mathbf{x},\lambda_j,F) &= \sum_{\ell=1}^{k}\frac{\tau_\ell}{\rho_\ell}\left[\mathrm{PIF}_i(\mathbf{x},\psi_{\ell j},F)-\lambda_j\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)\right]\end{aligned}$$

Since, $\mathrm{PIF}_i(\mathbf{x},\psi_{\ell j},F)=0$, when $\ell\neq i$, we get

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F) &= -\rho_\ell\frac{1}{p}\sum_{j=1}^{p}\frac{\mathrm{PIF}_i(\mathbf{x},\lambda_j,F)}{\lambda_j} \qquad \text{for } \ell\neq 1 \text{ and } \ell\neq i\\ \mathrm{PIF}_i(\mathbf{x},\rho_i,F) &= -\rho_i\frac{1}{p}\sum_{j=1}^{p}\frac{\mathrm{PIF}_i(\mathbf{x},\lambda_j,F)}{\lambda_j}+\frac{1}{p}\sum_{j=1}^{p}\frac{\mathrm{PIF}_i(\mathbf{x},\psi_{ij},F)}{\lambda_j} \qquad \text{for } i\neq 1\\ \mathrm{PIF}_i(\mathbf{x},\lambda_j,F) &= -\lambda_j\sum_{\ell=2}^{k}\frac{\tau_\ell}{\rho_\ell}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)+\frac{\tau_i}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\psi_{ij},F)\ .\end{aligned}$$

Denote $B_i(\mathbf{x})=\frac{1}{p}\sum_{j=1}^{p}\frac{\mathrm{PIF}_i(\mathbf{x},\lambda_j,F)}{\lambda_j}$. Then, we have that

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F) &= -\rho_\ell B_i(\mathbf{x}) \qquad \ell\neq 1 \text{ and } \ell\neq i &\text{(A.7)}\\ \mathrm{PIF}_i(\mathbf{x},\rho_i,F) &= \left[-\rho_i B_i(\mathbf{x})+A_i(\mathbf{x})\right](1-\delta_{i1}) &\text{(A.8)}\\ \mathrm{PIF}_i(\mathbf{x},\lambda_j,F) &= -\lambda_j\sum_{\ell=2}^{k}\frac{\tau_\ell}{\rho_\ell}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)+\frac{\tau_i}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\psi_{ij},F)\ . &\text{(A.9)}\end{aligned}$$

Let us first consider the case when $i\neq 1$. In this case, from (A.7) and (A.8) it follows that

$$\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)=\frac{\rho_\ell}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\rho_i,F)-\frac{\rho_\ell}{\rho_i}A_i(\mathbf{x}) \qquad \text{for } \ell\neq 1 \text{ and } \ell\neq i\ .$$

Therefore, $\sum_{\ell=2}^{k}\frac{\tau_\ell}{\rho_\ell}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)=\frac{1-\tau_1}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\rho_i,F)-\frac{1-\tau_1-\tau_i}{\rho_i}A_i(\mathbf{x})$, which entails that

$$\mathrm{PIF}_i(\mathbf{x},\lambda_j,F)=-\lambda_j\frac{1-\tau_1}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\rho_i,F)+\frac{1-\tau_1-\tau_i}{\rho_i}\lambda_jA_i(\mathbf{x})+\frac{\tau_i}{\rho_i}\mathrm{PIF}_i(\mathbf{x},\psi_{ij},F)\ .$$

Thus, $B_i(\mathbf{x})=\frac{1-\tau_1}{\rho_i}\left[A_i(\mathbf{x})-\mathrm{PIF}_i(\mathbf{x},\rho_i,F)\right]$, and replacing in (A.7) to (A.9), we get the desired result.

Let us now study the case when $i=1$. We have the equations

$$\begin{aligned}\mathrm{PIF}_1(\mathbf{x},\rho_\ell,F) &= -\rho_\ell B_1(\mathbf{x}) \qquad \ell\neq 1\\ \mathrm{PIF}_1(\mathbf{x},\rho_1,F) &= 0\\ \mathrm{PIF}_1(\mathbf{x},\lambda_j,F) &= -\lambda_j\sum_{\ell=2}^{k}\frac{\tau_\ell}{\rho_\ell}\mathrm{PIF}_1(\mathbf{x},\rho_\ell,F)+\tau_1\mathrm{PIF}_1(\mathbf{x},\psi_{1j},F)\ ,\end{aligned}$$

which led to $\sum_{\ell=2}^{k}\frac{\tau_\ell}{\rho_\ell}\mathrm{PIF}_1(\mathbf{x},\rho_\ell,F)=-(1-\tau_1)B_1(\mathbf{x})$ and thus, $\mathrm{PIF}_1(\mathbf{x},\lambda_j,F)=(1-\tau_1)\lambda_jB_1(\mathbf{x})+\tau_1\mathrm{PIF}_1(\mathbf{x},\psi_{1j},F)$. Therefore, $B_1(\mathbf{x})=(1-\tau_1)B_1(\mathbf{x})+\tau_1A_1(\mathbf{x})$ and so $A_1(\mathbf{x})=B_1(\mathbf{x})$, which proves (A.3) and (A.4). $\square$

Proof of Theorem 2.1. The derivation of the partial influence function of the eigenvectors follows the same steps as those used in the proof of Lemma 3 in Croux and Haesbroeck (2000) which can be found in Croux and Haesbroeck (1999). For the sake of simplicity, we will denote $\rho_{\mathbf{V},i}(F) = \rho^{(1)}_{\mathbf{V},i}(F)$.

Let $\boldsymbol{\beta}_{j,\epsilon,i} = \boldsymbol{\beta}_{\mathbf{V},j}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\lambda_{j,\epsilon,i} = \lambda_{\mathbf{V},j}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\rho_{\ell,\epsilon,i} = \rho_{\mathbf{V},\ell}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\mathbf{V}_{i,\epsilon} = \mathbf{V}_i\left(F_{i,\epsilon,\mathbf{x}}\right)$ and $\mathbf{V}_\ell = \mathbf{V}_\ell\left(F_\ell\right)$. Then, we have that

$$\begin{aligned} 0 &= \boldsymbol{\beta}'_{m,\epsilon,i}\left[\sum_{\ell=1,\ell\neq i}^{k} \frac{\tau_\ell}{\rho_{\ell,\epsilon,i}}\mathbf{V}_\ell\right]\boldsymbol{\beta}_{j,\epsilon,i} + \frac{\tau_i}{\rho_{i,\epsilon,i}}\boldsymbol{\beta}'_{m,\epsilon,i}\mathbf{V}_{i,\epsilon}\boldsymbol{\beta}_{j,\epsilon,i} \qquad m\neq j \qquad (A.10)\\ \delta_{mj} &= \boldsymbol{\beta}'_{m,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i}\,. \qquad (A.11)\end{aligned}$$

Therefore, deriving (A.11) with respect to ϵ, we get that

$$\begin{aligned} \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},m},F)'\boldsymbol{\beta}_m &= 0 \qquad (A.12)\\ \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},m},F)'\boldsymbol{\beta}_j + \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F)'\boldsymbol{\beta}_m &= 0\,. \qquad (A.13)\end{aligned}$$

Deriving (A.10) and using (A.13) and $\boldsymbol{\beta}'_m\boldsymbol{\beta}_j = 0$ for $m\neq j$, after some algebra, we obtain

$$\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\mathbf{V},j},F)'\boldsymbol{\beta}_m = \frac{\tau_i}{\rho_i}\frac{1}{\lambda_j-\lambda_m}\boldsymbol{\beta}'_j\mathrm{IF}\left(\mathbf{x},\mathbf{V}_i,F_i\right)\boldsymbol{\beta}_m \qquad \text{for } m\neq j$$

which together with (A.12) entail (18).

Let $\psi_{ij}(G) = \boldsymbol{\beta}_j(G)'\mathbf{V}_i(G_i)\boldsymbol{\beta}_j(G)$. From (A.12) and $\mathbf{V}_\ell\boldsymbol{\beta}_j = \boldsymbol{\Sigma}_\ell\boldsymbol{\beta}_j = \rho_\ell\lambda_j\boldsymbol{\beta}_j$ we get that $\mathrm{PIF}_i(\mathbf{x},\psi_{ij},F) = \boldsymbol{\beta}'_j\mathrm{IF}(\mathbf{x},\mathbf{V}_i,F_i)\boldsymbol{\beta}_j$. Now, using Lemma A.1, we obtain (16) and (17). □

Proof of Theorem 2.2. Denote by $\lambda_{\mathbf{V}_i,1}(F_i) > \ldots > \lambda_{\mathbf{V}_i,p}(F_i)$ the eigenvalues of $\mathbf{V}_i(F_i)$. According to Lemma 3 in Croux and Haesbroeck (2000), the influence functions of $\lambda_{\mathbf{V}_i,j}(F_i)$ at a distribution F_i such that $\mathbf{V}_i(F_i) = \boldsymbol{\Sigma}_i = \boldsymbol{\beta}'\boldsymbol{\Lambda}_i\boldsymbol{\beta}$, where $\boldsymbol{\Lambda}_i = \mathrm{diag}\left(\lambda_{i1},\ldots,\lambda_{ip}\right)$, is given by

$$\mathrm{IF}\left(\mathbf{x},\lambda_{\mathbf{V}_i,j},F_i\right) = \boldsymbol{\beta}'_j\mathrm{IF}\left(\mathbf{x},\mathbf{V}_i,F_i\right)\boldsymbol{\beta}_j\,. \qquad (A.14)$$

As in Theorem 2.1, let $\rho^{(s)}_{\ell,\epsilon,i} = \rho^{(s)}_{\mathbf{V},\ell}\left(F_{\epsilon,\mathbf{x},i}\right)$, for $s = 2,3$, $\mathbf{V}_{i,\epsilon} = \mathbf{V}_i\left(F_{i,\epsilon,\mathbf{x}}\right)$ and $\mathbf{V}_\ell = \mathbf{V}_\ell\left(F_\ell\right)$. Denote by $\phi_{\mathbf{V},\ell}(F) = \ln\left(\rho^{(2)}_{\mathbf{V},\ell}(F)\right)$ and $\phi_{\ell,\epsilon,i} = \ln\left(\rho^{(2)}_{\ell,\epsilon,i}\right)$. Then,

$$\mathrm{PIF}_i(\mathbf{x},\phi_{\mathbf{V},\ell},F) = \frac{1}{\rho^{(2)}_{\mathbf{V},\ell}}\mathrm{PIF}_i(\mathbf{x},\rho^{(2)}_{\mathbf{V},\ell},F)\,. \qquad (A.15)$$

Let us first consider the case when $i\neq 1$. If $\ell\neq i$ $\rho^{(s)}_{\ell,\epsilon,i} = \rho^{(s)}_\ell$ and thus, $\mathrm{PIF}_i(\mathbf{x},\rho^{(s)}_{\mathbf{V},\ell},F) = 0$ for $s = 2,3$. When $\ell = i$, we have,

$$\begin{aligned} \phi_{i,\epsilon,i} = \ln\left(\rho^{(2)}_{i,\epsilon,i}\right) &= \frac{1}{p}\sum_{j=1}^{p}\left[\ln\left(\lambda_{\mathbf{V}_i,j}\left(F_{i,\epsilon,\mathbf{x}}\right)\right) - \ln\left(\lambda_{\mathbf{V}_1,j}\left(F_1\right)\right)\right] \qquad (A.16)\\ \rho^{(3)}_{i,\epsilon,i} &= \frac{1}{p}\mathrm{tr}(\mathbf{V}_1(F_1)^{-1}\mathbf{V}_{i,\epsilon})\,. \qquad (A.17)\end{aligned}$$

From (A.14) and (A.16), we get $\mathrm{PIF}_i\left(\mathbf{x},\phi_{\mathbf{V},i},F\right) = \frac{1}{p}\sum_{j=1}^{p}\frac{1}{\rho_i\lambda_j}\boldsymbol{\beta}'_j\mathrm{IF}\left(\mathbf{x},\mathbf{V}_i,F_i\right)\boldsymbol{\beta}_j$ which together with (A.15) entails (21).

Using that $\mathbf{V}_1 = \boldsymbol{\beta}'\boldsymbol{\Lambda}_1\boldsymbol{\beta}$ and (A.17), we obtain

$$\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(3)}_{\mathbf{V},i}, F\right) = \frac{1}{p}\mathrm{tr}\left[\mathbf{V}_1^{-1}\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_i, F_i\right)\right] = \frac{1}{p}\mathrm{tr}\left[\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\beta}'\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_i, F_i\right)\boldsymbol{\beta}\right] = \frac{1}{p}\sum_{j=1}^{p}\frac{1}{\lambda_j}\boldsymbol{\beta}_j'\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_i, F_i\right)\boldsymbol{\beta}_j$$

which concludes the proof when $i \neq 1$.

Let us now consider the case when $i = 1$. Since $\ell \neq 1$, we have,

$$\phi_{\ell,\epsilon,1} = \frac{1}{p}\sum_{j=1}^{p}\left[\ln\left(\lambda_{\mathbf{V}_\ell,j}\left(F_\ell\right)\right) - \ln\left(\lambda_{\mathbf{V}_1,j}\left(F_{1,\epsilon,\mathbf{x}}\right)\right)\right] \tag{A.18}$$

$$\rho^{(3)}_{\ell,\epsilon,1} = \frac{1}{p}\mathrm{tr}(\mathbf{V}_1(F_{1,\epsilon,\mathbf{x}})^{-1}\mathbf{V}_\ell)\,. \tag{A.19}$$

Now (21) follows deriving (A.18) with respect to ϵ and using (A.14) and (A.15).

Finally, from (A.17), using that $\mathbf{V}_1 = \boldsymbol{\beta}'\boldsymbol{\Lambda}_1\boldsymbol{\beta}$, $\mathbf{V}_\ell = \rho_\ell\mathbf{V}_1$ and $\frac{\partial}{\partial\epsilon}\left[\mathbf{V}_1(F_{1,\epsilon,\mathbf{x}})^{-1}\right]\Big|_{\epsilon=0} = -\mathbf{V}_1^{-1}\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_1, F\right)\mathbf{V}_1^{-1}$, we get

$$\begin{aligned}\mathrm{PIF}_1\left(\mathbf{x}, \rho^{(3)}_{\mathbf{V},\ell}, F\right) &= \frac{1}{p}\mathrm{tr}\left[\frac{\partial}{\partial\epsilon}\left[\mathbf{V}_1^{-1}\right]|_{\epsilon=0}\mathbf{V}_\ell\right] = -\frac{1}{p}\mathrm{tr}\left[\mathbf{V}_1^{-1}\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_1, F\right)\mathbf{V}_1^{-1}\mathbf{V}_\ell\right]\\ &= -\rho_\ell\frac{1}{p}\mathrm{tr}\left[\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\beta}'\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_1, F\right)\boldsymbol{\beta}\right] = -\rho_\ell\frac{1}{p}\sum_{j=1}^{p}\frac{1}{\lambda_j}\boldsymbol{\beta}_j'\mathrm{IF}\left(\mathbf{x}, \mathbf{V}_1, F_1\right)\boldsymbol{\beta}_j\end{aligned}$$

which concludes the proof. □

Proof of Theorem 2.3. The partial influence functions for the eigenvectors are obtained from Theorem 2 of Boente, Pires and Rodrigues (2002). The proof for the eigenvalues and for the proportionality constants follows from Lemma A.1 taking $\psi_{ij}(F) = \lambda_{\sigma,ij}(F)$ and noticing that from Theorem 2 of Boente, Pires and Rodrigues (2002),

$$\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,\ell j}, F) = 0 \qquad \text{when } \ell \neq i \tag{A.20}$$

$$\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,ij}, F) = 2\lambda_{ij}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}}; \sigma, G_0\right)\,. \tag{A.21}$$

On the other hand, (A.20) and (A.21) entail that, for $\ell \neq i$,

$$\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(3)}_\ell, F\right) = \frac{1}{p}\sum_{j=1}^{p}\left[\frac{\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,\ell j}, F)}{\lambda_{1j}} - \frac{\lambda_{\ell j}}{\lambda_{1j}^2}\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,1j}, F)\right] = -\delta_{i1}\frac{2}{p}\rho_\ell\sum_{j=1}^{p}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_j}}; \sigma, G_0\right)$$

and, for $i \neq 1$,

$$\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(3)}_i, F\right) = \frac{1}{p}\sum_{j=1}^{p}\left[\frac{\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,ij}, F)}{\lambda_{1j}} - \frac{\lambda_{ij}}{\lambda_{1j}^2}\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,1j}, F)\right] = \frac{2}{p}\rho_i\sum_{j=1}^{p}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}}; \sigma, G_0\right)\,.$$

As in Theorem 2.2, since $\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(2)}_\ell, F\right) = \rho_\ell\mathrm{PIF}_i\left(\mathbf{x}, \phi_\ell, F\right) = \rho_\ell\frac{1}{p}\sum_{j=1}^{p}\left[\frac{\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,\ell j}, F)}{\lambda_{\ell j}} - \frac{\mathrm{PIF}_i(\mathbf{x}, \lambda_{\sigma,1j}, F)}{\lambda_{1j}}\right]$, using (A.20) and (A.21), we get that $\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(2)}_i, F\right) = \frac{2}{p}\rho_i\sum_{j=1}^{p}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}}; \sigma, G_0\right)$, for $i \neq 1$, and $\mathrm{PIF}_i\left(\mathbf{x}, \rho^{(2)}_\ell, F\right) = -\delta_{i1}\frac{2}{p}\rho_\ell\sum_{j=1}^{p}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_j}}; \sigma, G_0\right)$, for $\ell \neq i$. □

Proof of Theorem 2.4. The proof of the partial influnece function of the eigenvectors $\boldsymbol{\beta}_{\sigma,\rho,j}$ will be derived using analogous arguments as those considered in Boente, Pires and Rodrigues (2002). Let $\rho_{\ell,\epsilon,i} = \rho\left(F_{\epsilon,\mathbf{x},i}\right)$, $\boldsymbol{\beta}_{j,\epsilon,i} = \boldsymbol{\beta}_{\sigma,\rho,j}\left(F_{\epsilon,\mathbf{x},i}\right)$, $\lambda_{\ell j,\epsilon,i} = \sigma^2\left(F_\ell\left[\boldsymbol{\beta}_{j,\epsilon,i}\right]\right)$, for $\ell \neq i$ and $\lambda_{ij,\epsilon,i} = \sigma^2\left(F_{i,\epsilon,\mathbf{x}}\left[\boldsymbol{\beta}_{j,\epsilon,i}\right]\right)$, where we avoid the index ρ and σ for the sake of simplicity.

Now, $\boldsymbol{\beta}_{j,\epsilon,i}$ maximises $\varsigma_\rho\left(F_{\epsilon,\mathbf{x},i}\left[\mathbf{b}\right]\right)$ under the constraints $\boldsymbol{\beta}'_{j,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i} = 1$ and $\boldsymbol{\beta}'_{s,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i} = 0$ for $1 \leq s \leq j-1$. Therefore, $\boldsymbol{\beta}_{j,\epsilon,i}$ maximises

$$L(\mathbf{b},\gamma,\boldsymbol{\alpha}) = \frac{\tau_i}{\rho_{i,\epsilon,i}}\sigma^2\left(F_{i,\epsilon,\mathbf{x}}\left[\mathbf{b}\right]\right) + \sum_{i_0\neq i}\frac{\tau_{i_0}}{\rho_{i_0,\epsilon,i}}\sigma^2\left(F_{i_0}\left[\mathbf{b}\right]\right) - \gamma\left(\mathbf{b}'\mathbf{b}-1\right) - \sum_{s=1}^{j-1}\alpha_s\mathbf{b}'\boldsymbol{\beta}_{s,\epsilon,i}\ ,$$

and so it should satisfy

$$0 = \frac{\partial}{\partial\mathbf{b}}L(\mathbf{b},\gamma,\boldsymbol{\alpha})|_{\mathbf{b}=\boldsymbol{\beta}_{j,\epsilon,i}} = \psi(\epsilon) - 2\gamma\boldsymbol{\beta}_{j,\epsilon,i} - \sum_{s=1}^{j-1}\alpha_s\boldsymbol{\beta}_{s,\epsilon,i}\ , \tag{A.22}$$

with

$$\psi(\epsilon) \quad = \quad \frac{\tau_i}{\rho_{i,\epsilon,i}}\frac{\partial}{\partial\mathbf{b}}\sigma^2\left(F_{i,\epsilon,\mathbf{x}}\left[\mathbf{b}\right]\right)|_{\mathbf{b}=\boldsymbol{\beta}_{j,\epsilon,i}} + \sum_{i_0\neq i}\frac{\tau_{i_0}}{\rho_{i_0,\epsilon,i}}\frac{\partial}{\partial\mathbf{b}}\sigma^2\left(F_{i_0}\left[\mathbf{b}\right]\right)|_{\mathbf{b}=\boldsymbol{\beta}_{j,\epsilon,i}}\ . \tag{A.23}$$

Since $\boldsymbol{\beta}'_{j,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i} = 1$ and $\boldsymbol{\beta}'_{s,\epsilon,i}\boldsymbol{\beta}_{j,\epsilon,i} = 0$ for $1 \leq s \leq j-1$ we have that $\psi(\epsilon)'\boldsymbol{\beta}_{j,\epsilon,i} = 2\gamma$, $\psi(\epsilon)'\boldsymbol{\beta}_{s,\epsilon,i} = \alpha_s$, for $1 \leq s \leq j-1$. Using this in (A.22) and differentiating with respect to ϵ we obtain

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\psi(\epsilon)|_{\epsilon=0} \quad &= \quad \sum_{s=1}^{j}\Bigg[\left\{\psi(0)'\mathrm{PIF}_i\left(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,s},F\right)\right\}\boldsymbol{\beta}_s + \left\{\boldsymbol{\beta}'_s\frac{\partial}{\partial\epsilon}\psi(\epsilon)|_{\epsilon=0}\right\}\boldsymbol{\beta}_s + \\ &+ \quad \left\{\psi(0)'\boldsymbol{\beta}_s\right\}\mathrm{PIF}_i\left(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,s},F\right)\Bigg]\ .\end{aligned} \tag{A.24}$$

Since F_i is an elliptical distribution and $\sigma(G_0) = 1$, using the fact that $\sigma^2\left(F_i[\mathbf{b}]\right) = \mathbf{b}'\boldsymbol{\Sigma}_i\mathbf{b}$, we obtain $\psi(0) = 2\ \lambda_j\boldsymbol{\beta}_j$, which implies that $\psi(0)'\boldsymbol{\beta}_s = 0$ for $1 \leq s \leq j-1$. Write $\mathbf{P}_{j+1} = \mathbf{I}_p - \sum_{s=1}^{j}\boldsymbol{\beta}_s\boldsymbol{\beta}'_s$. Then (A.24) can be written as

$$\mathbf{P}_{j+1}\frac{\partial}{\partial\epsilon}\psi(\epsilon)|_{\epsilon=0} = 2\lambda_j\sum_{s=1}^{j}\boldsymbol{\beta}'_j\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,s},F)\boldsymbol{\beta}_s + 2\lambda_j\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,j},F)\ . \tag{A.25}$$

On the other hand, from (A.23) and since $\varsigma_\rho(F[\mathbf{b}]) = \mathbf{b}'\boldsymbol{\Sigma}_1\mathbf{b}$, we have that

$$\begin{aligned}\frac{\partial}{\partial\epsilon}\psi(\epsilon)\Big|_{\epsilon=0} \quad &= \quad -2\left[\sum_{\ell=1}^{k}\frac{\tau_\ell}{\rho_\ell(F)}\mathrm{PIF}_i(\mathbf{x},\rho_\ell,F)\right]\lambda_j\boldsymbol{\beta}_j + \\ &+ \quad 2\ \boldsymbol{\Sigma}\ \ \mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,j},F) + \tau_i\frac{\partial}{\partial\mathbf{b}}\mathrm{IF}\left(\mathbf{b}'\mathbf{x},\sigma^2,F_i[\mathbf{b}]\right)\Big|_{\mathbf{b}=\boldsymbol{\beta}_j}\ .\end{aligned} \tag{A.26}$$

Again from the equivariance of the scale estimator, we have that

$$\begin{aligned}\frac{\partial}{\partial\mathbf{b}}\mathrm{IF}\left(\mathbf{b}'\mathbf{x},\sigma^2,F_i[\mathbf{b}]\right)|_{\mathbf{b}=\boldsymbol{\beta}_j} \quad &= \quad 2\lambda_{ij}\boldsymbol{\beta}_j\mathrm{IF}\left(\frac{\boldsymbol{\beta}'_j\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}},\sigma^2,G_0\right) + \\ &+ \quad \lambda_{ij}\mathrm{DIF}\left(\frac{\boldsymbol{\beta}'_j\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}},\sigma^2,G_0\right)\left(\frac{\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}} - \frac{\boldsymbol{\beta}'_j\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}}\boldsymbol{\beta}_j\right)\ .\end{aligned} \tag{A.27}$$

From (A.25), (A.26) and (A.27) and using that $\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,j},F)'\boldsymbol{\beta}_j=0$, we obtain

$$\begin{aligned} 2\left(\mathbf{P}_{j+1}\boldsymbol{\Sigma}_1-\lambda_j\mathbf{I}_p\right)\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,j},F) &= 2\lambda_j\sum_{s=1}^{j-1}\boldsymbol{\beta}_j'\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,s},F)\boldsymbol{\beta}_s - \\ &- \tau_i\frac{\lambda_j^{\frac{1}{2}}}{\rho_i^{\frac{1}{2}}}\mathrm{DIF}\left(\frac{\boldsymbol{\beta}_j'\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}},\sigma^2,G_0\right)\mathbf{P}_{j+1}\mathbf{x}\,. \qquad \text{(A.28)}\end{aligned}$$

The matrix $\mathbf{P}_{j+1}\boldsymbol{\Sigma}_1-\lambda_j\mathbf{I}_p=\sum_{s=j+1}^{p}\lambda_s\boldsymbol{\beta}_s\boldsymbol{\beta}_s'-\lambda_j\mathbf{I}_p$ is a full rank matrix with inverse

$$\left(\mathbf{P}_{j+1}\boldsymbol{\Sigma}-\lambda_j\mathbf{I}_p\right)^{-1}=\sum_{s=j+1}^{p}\frac{1}{\lambda_s-\lambda_j}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'-\sum_{s=1}^{j}\frac{1}{\lambda_j}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'\,,$$

so that, $\left(\mathbf{P}_{j+1}\boldsymbol{\Sigma}_1-\lambda_j\mathbf{I}_p\right)^{-1}\boldsymbol{\beta}_s=-\frac{1}{\lambda_j}\boldsymbol{\beta}_s$ for $1\le s\le j-1$ and $\left(\mathbf{P}_{j+1}\boldsymbol{\Sigma}_1-\lambda_j\mathbf{I}_p\right)^{-1}\mathbf{P}_{j+1}=\sum_{s=j+1}^{p}\frac{1}{\lambda_s-\lambda_j}\boldsymbol{\beta}_s\boldsymbol{\beta}_s'$. Thus, from (A.28), and since $\lambda_{ij}=\rho_i\lambda_j$, after some calculations, we obtain for any $s\ge j+1$ that

$$\mathrm{PIF}_i(\mathbf{x},\boldsymbol{\beta}_{\sigma,\rho,j},F)'\boldsymbol{\beta}_s=\frac{1}{2(\lambda_j-\lambda_s)}\tau_i\frac{\lambda_j^{\frac{1}{2}}}{\rho_i^{\frac{1}{2}}}\mathrm{DIF}\left(\frac{\boldsymbol{\beta}_j'\mathbf{x}}{\lambda_{ij}^{\frac{1}{2}}},\sigma^2,G_0\right)\boldsymbol{\beta}_s'\mathbf{x}\,,$$

which implies (35), using the fact that $\mathrm{IF}\left(y,\sigma^2,G_0\right)=2\ \mathrm{IF}\left(y,\sigma,G_0\right)$.

Since $\lambda_{ij,\epsilon,i}=\sigma^2\left(F_{i,\epsilon,\mathbf{x}}\left[\boldsymbol{\beta}_{j,\epsilon,i}\right]\right)$ and, for $\ell\neq i$, $\lambda_{\ell j,\epsilon,i}=\sigma^2\left(F_\ell\left[\boldsymbol{\beta}_{j,\epsilon,i}\right]\right)=\boldsymbol{\beta}_{j,\epsilon,i}'\boldsymbol{\Sigma}_\ell\boldsymbol{\beta}_{j,\epsilon,i}$, the chain rule easily yields

$$\begin{aligned}\mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,\rho,\ell j},F) &= 0 \qquad \text{when } \ell\neq i\\ \mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,\rho,ij},F) &= 2\lambda_{ij}\ \mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}};\sigma,G_0\right)\,.\end{aligned}$$

Using Lemma A.1, the partial influence functions of the eigenvalues and for the proportionality constants follows now as in Theorem 2.3. □

Proof of Proposition 3.1. Using (11) and (28), we get that

$$\begin{aligned}\mathrm{ASVAR}\left(\widehat{\rho}_i\right) &= \frac{1}{\tau_1}E\left(\mathrm{PIF}_1(\mathbf{x},\rho_{\sigma,i},F)^2\right)+\frac{1}{\tau_i}\mathrm{VAR}_{F_i}\left[\frac{2\rho_i}{p}\sum_{j=1}^{p}\mathrm{IF}\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j}{\sqrt{\lambda_{ij}}},\sigma,G_0\right)\right]\\ &= \left(\frac{1}{\tau_1}+\frac{1}{\tau_i}\right)\frac{4\rho_i^2}{p^2}\mathrm{VAR}_G\left[\sum_{j=1}^{p}\mathrm{IF}\left(z_{1j},\sigma,G_0\right)\right]\,.\end{aligned}$$

When $G=N\left(\mathbf{0},\mathbf{I}_p\right)$, since z_{1j}, z_{1m} and z_{1r} are independent for $j\neq r$, $j\neq m$ and $m\neq r$ which implies that

$$\mathrm{ASVAR}\left(\widehat{\rho}_i\right)=\frac{4}{p}\rho_i^2\mathrm{ASVAR}\left(\sigma,G_0\right)\left(\frac{1}{\tau_1}+\frac{1}{\tau_i}\right)\,.$$

On the other hand, using (11) and (27), it can easily be seen

$$\mathrm{ASVAR}\left(\widehat{\lambda}_j\right) = \frac{1}{\tau_1}\mathrm{VAR}_{F_1}\left(\tau_1\mathrm{PIF}_1(\mathbf{x},\lambda_{\sigma,1j},F)+\lambda_j(1-\tau_1)A_1(\mathbf{x})\right)+\sum_{i\neq 1}\frac{\tau_i}{\rho_i^2}\mathrm{VAR}_{F_i}\left(\mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,ij},F)-\lambda_jA_i(\mathbf{x})\right)$$

Since,

$$\begin{aligned}
\mathrm{VAR}_{F_i}\left(\mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,ij},F)\right) &= 4\rho_i^2\lambda_j^2\mathrm{VAR}_G\left(\mathrm{IF}\left(z_{1j},\sigma,G_0\right)\right) = 4\rho_i^2\lambda_j^2\mathrm{ASVAR}\left(\sigma,G_0\right)\\
\mathrm{COV}_{F_i}\left(\mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,ij},F),\mathrm{PIF}_i(\mathbf{x},\lambda_{\sigma,is},F)\right) &= 4\rho_i^2\lambda_j\lambda_s\mathrm{COV}_G\left(\mathrm{IF}\left(z_{1j},\sigma,G_0\right),\mathrm{IF}\left(z_{1s},\sigma,G_0\right)\right)
\end{aligned}$$

$$\begin{aligned}
\mathrm{VAR}_{F_i}\left(A_i(\mathbf{x})\right) &= 4\frac{\rho_i^2}{p}\left[\mathrm{ASVAR}\left(\sigma,G_0\right)+(p-1)\mathrm{COV}_G\left(\mathrm{IF}\left(z_{11},\sigma,G_0\right),\mathrm{IF}\left(z_{12},\sigma,G_0\right)\right)\right]\\
\mathrm{COV}\left(\mathrm{PIF}_i\left(\mathbf{x},\lambda_{\sigma,ij},F\right),A_i(\mathbf{x})\right) &= 4\frac{\rho_i^2}{p}\lambda_j\left[\mathrm{ASVAR}\left(\sigma,G_0\right)+(p-1)\mathrm{COV}_G\left(\mathrm{IF}\left(z_{11},\sigma,G_0\right),\mathrm{IF}\left(z_{12},\sigma,G_0\right)\right)\right]\\
&= 4\frac{\rho_i^2}{p}\lambda_j\gamma
\end{aligned}$$

straightforward calculations led to

$$\mathrm{ASVAR}\left(\widehat{\lambda}_j\right) = 4\lambda_j^2\left\{\frac{1}{\tau_1}\frac{\gamma}{p}+\left(1-\frac{1}{p}\right)\kappa\right\} \qquad \mathrm{ASCOV}\left(\widehat{\lambda}_j,\widehat{\lambda}_s\right) = 4\lambda_j\lambda_s\left\{\frac{1}{\tau_1}\frac{\gamma}{p}-\frac{\kappa}{p}\right\} \qquad 1\le j\ne s\le p$$

$$\mathrm{ASVAR}\left(\widehat{\nu}_j\right) = 4\,\kappa\,\nu_j^2\left(\sum_{\ell=1}^{p}\nu_\ell^2-2\nu_j-+1\right) \qquad \mathrm{ASCOV}\left(\widehat{\nu}_j,\widehat{\nu}_s\right) = 4\,\kappa\,\nu_j\nu_s\left(\sum_{\ell=1}^{p}\nu_\ell^2-\nu_j-\nu_s\right) \qquad 1\le j\ \ne\ s\le p\,.$$

The proof of Proposition 3.2 follows analogously.

References

Boente, G. and Orellana, L. (2001). A Robust Approach to Common Principal Components. *Statistics in Genetics and in the Environmental Sciences.* L. T. Fernholz, S. Morgenthaler, and W. Stahel, editors, pp. 117–145. *Birkhäuser Verlag AG*, Basel, Switzerland.

Boente, G., Pires, A. and Rodrigues, I. (2002). Influence Functions and Outlier Detection under the Common Principal Components Model. *Biometrika*, **89**, 861-875.

Boente, G. and Orellana, L. (2004). Robust plug–in estimators in proportional scatter models. *Journal of Statistical Planning and Inference*, **122**, 95-110.

Croux, C. and Haesbroeck, G. (1999). Principal Component Analysis based on Robust estimators of the Covariance or Correlation Matrix: Influence Functions and Efficiencies. *Preprint 99/128.* Université Libre de Bruxelles.

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. (1996). A Fast Algorithm for Robust Principal Components based on Projection Pursuit. In *Compstat: Proceedings in Computational Statistics,* ed. A. Prat, Heidelberg: Physica–Verlag, 211–217.

Croux, C. and Ruiz–Gazen, A. (2000). High Breakdown Estimators for Principal Components: the Projection–Pursuit Approach Revisited. Working paper. Université Libre de Bruxelles.

Donoho, D. L. (1982). *Breakdown Properties of Multivariate Location Estimators.* Ph.D. qualifying paper, Harvard University.

Eriksen, P. S. (1987). Proportionality of Covariance Matrices, *Ann. Statist.*, **15**, 732–748.

Flury, B. (1984). Common Principal Components in Groups. *J. Amer. Statist. Assoc.***,** **79**, 892–898.

Flury, B. (1986). Proportionality of k Covariance Matrices. *Statist. Probab. Letters*, **4**, 29–33.

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

Flury, B. and Schmid, M. (1992). Quadratic Discriminant Functions with Constraints on the Covariance Matrices: Some Asymptotic Results. *J. Multivariate Analysis*, **40**, 244–261.

Flury, B., Schmid, M. and Narayanan, A. (1994). Error Rates in Quadratic Discrimination with Constraints on the Covariance Matrices. *J. Classification*, **11**, 101–120.

Guttman,I., Kim, D. Y. and Olkin,I. (1985). Statistical Inference for constants of proportionality. In *Multivariate Analysis–VI.* P. R. Krishnaiah, editor. pp. 257–280. *North–Holland*, New York.

Khatri, C. G. (1967). Some distribution problems connected with the characteristic roots of $\mathbf{S}_1\mathbf{S}_2^{-1}$. *Ann. Math. Statist.,* **38**, 944–948.

Kim, D. Y. (1971). *Statistical inference for constants of proportionality between covariance matrices.* Technical Report 59, Standford University, Department of Statistics, Stanford, CA.

Manly, B. F. and Rayner, J. C. W. (1987). The Comparison of Sample Covariance Matrices Using Likelihood Ratio Tests. *Biometrika*, **74**, 841–847.

Maronna, R. A. (1976). Robust M–Estimators of Multivariate Location and Scatter. *Ann. Statist.,* **4**, 51–67.

Novi Inverardi, P. L. and Flury, B. (1992). Robust Estimation of Common Principal Components. *Quaderni di Satistica e Matematica Applicata alle Scienze Economico–Social,* **14**, 49–79.

Owen, A. (1984). A neighborhood–based classifier for LANDSAT data. *Canad. J. of Statist.*, **12**, 191–200.

Pillai, K. C. S., Al–Ani, S. and Jouris, G. M. (1969). On the distribution of the ratios of the roots of a covariance matrix and Wilks' criterion for tests of three hypotheses. *Ann. Math. Statist.,* **40**, 2033–2040.

Pires, A. M. & Branco, J. (2002). Partial influence functions. *J. Mult. Anal.*, **83**, 451-468.

Rao, C. R. (1983). Likelihood tests for relationships between covariance matrices. In *Studies in Econometrics, Time Series and Multivariate Statistics.* S. Karlim, T. Ameniya, L. A. Goodman, editors. pp. 529–543. *Academic Press*, New York.

Stahel, W. (1981). *Robust estimation: Infinitesimal Optimality and Covariance Matrix Estimation.* Thesis (in German), ETH, Zurich.

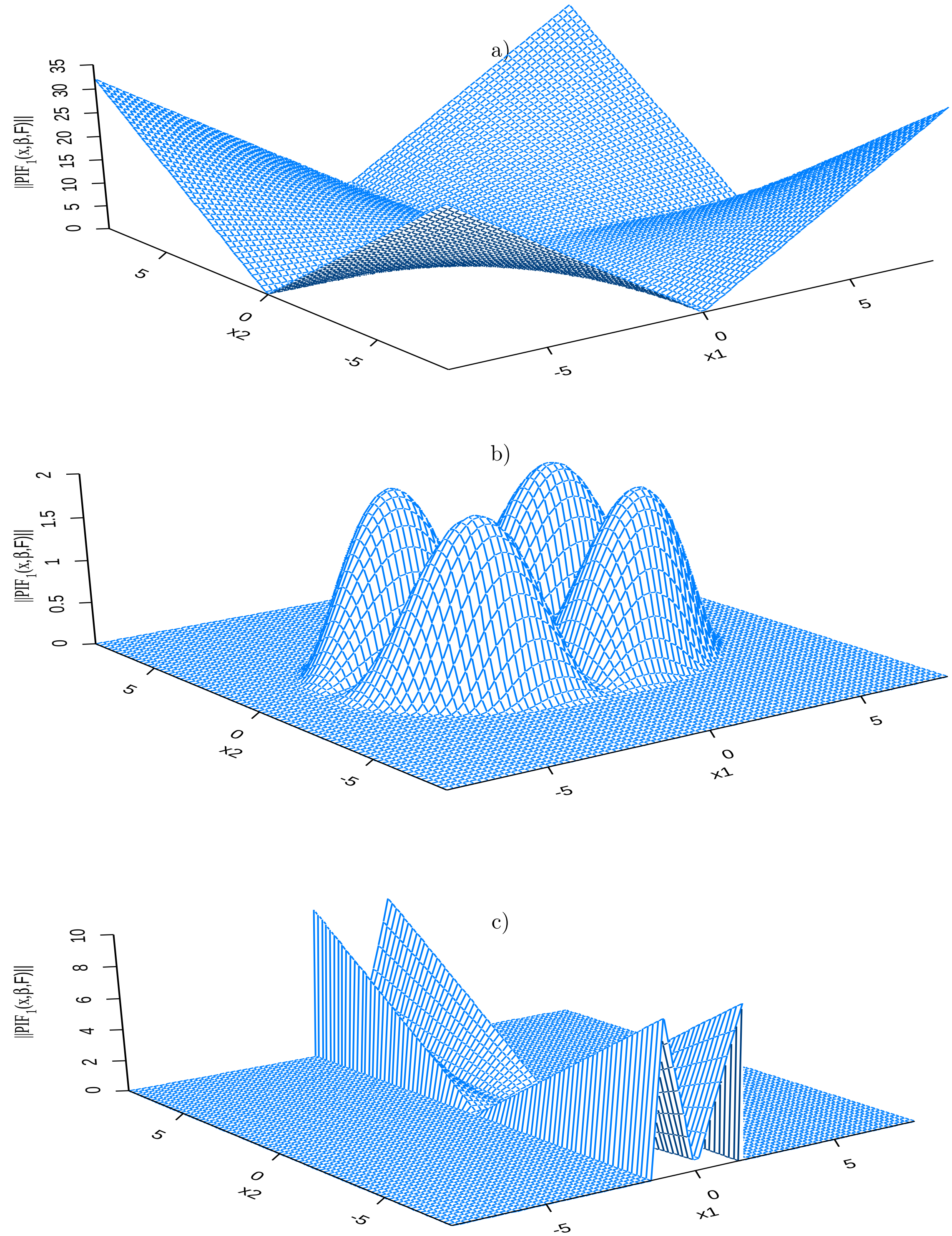

Figure 1: $\|\mathrm{PIF}_1(\mathbf{x}, \boldsymbol{\beta}_1, F)\|$ at $F = F_1 \times F_2$ with $F_1 = N(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N(\mathbf{0}, 4\mathrm{diag}(2,1))$. a) Maximum Likelihood Estimates b) Plug–in Estimates with an S–Scatter Matrix c) Projection Pursuit Estimates with an M–Scale Estimate

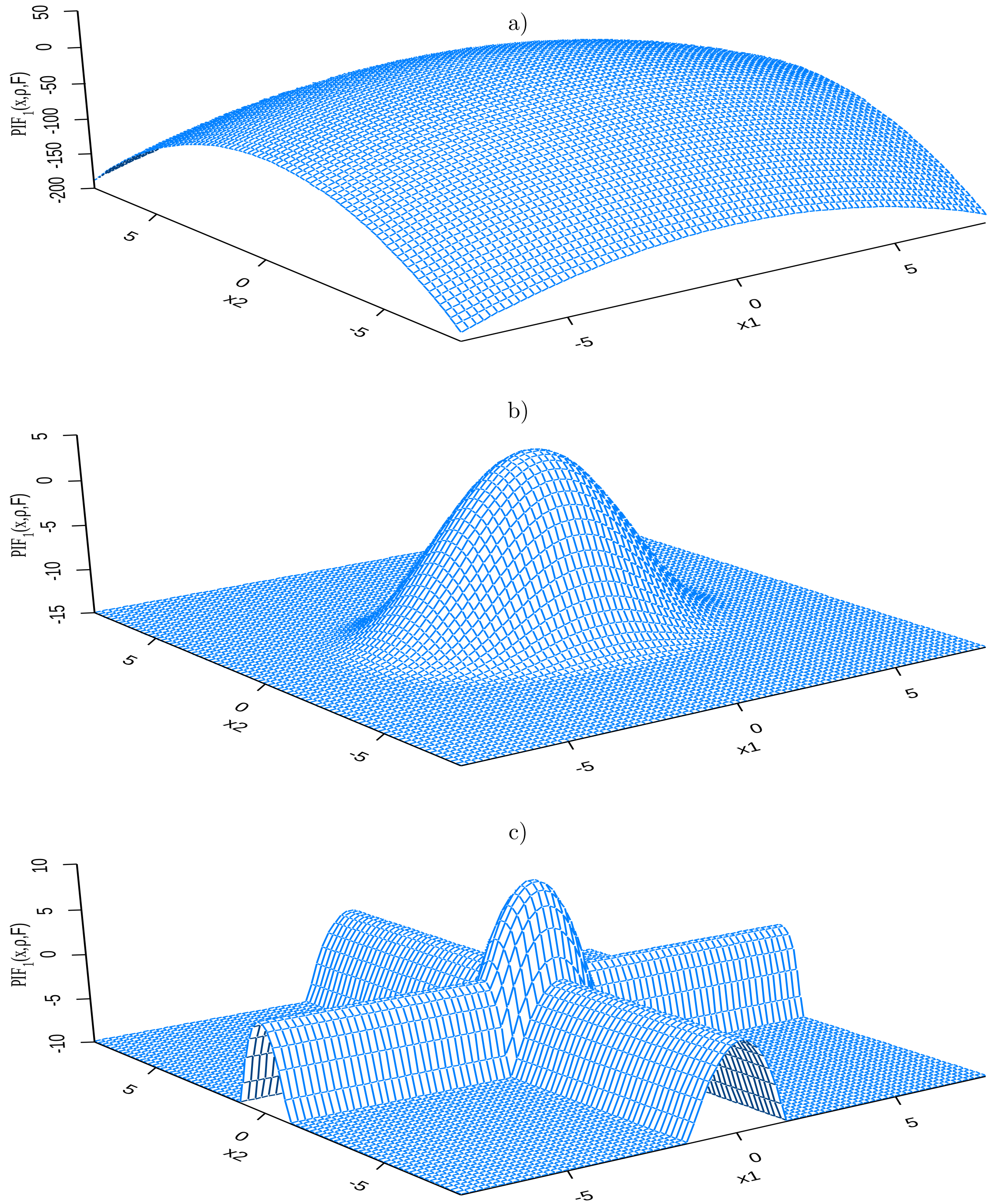

Figure 2: . $\mathrm{PIF}_1(\mathbf{x}, \rho, F)$ at $F = F_1 \times F_2$ with $F_1 = N(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N(\mathbf{0}, 4\mathrm{diag}(2,1))$ a) Maximum Likelihood Estimates b) Plug–in Estimates with an S–Scatter Matrix c) Projection Pursuit Estimates with an M–Scale Estimate

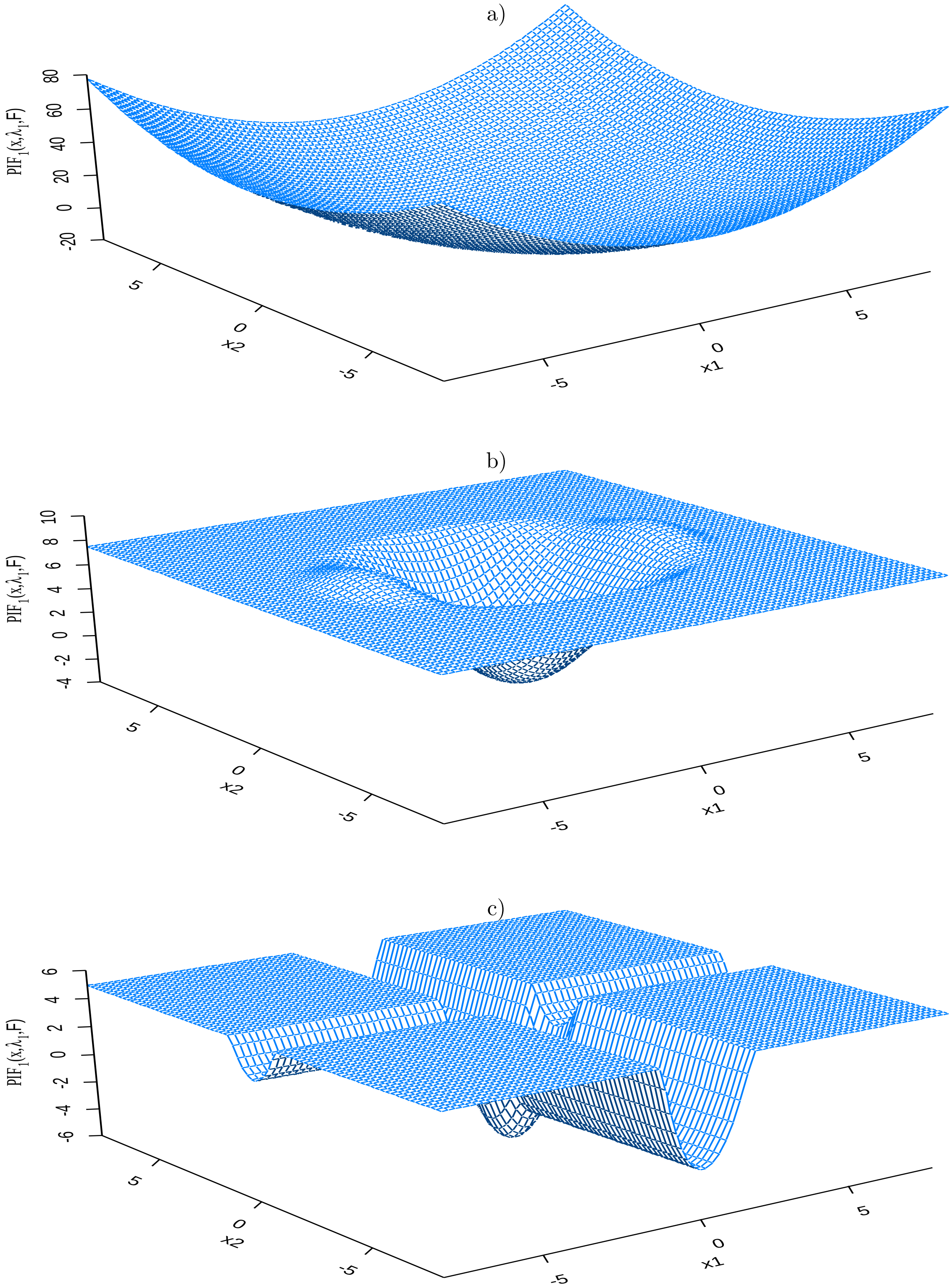

Figure 3: $\mathrm{PIF}_1(\mathbf{x}, \lambda_1, F)$ at $F = F_1 \times F_2$ with $F_1 = N\,(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N\,(\mathbf{0}, 4\mathrm{diag}(2,1))$. a) Maximum Likelihood Estimates b) Plug–in Estimates with an S–Scatter Matrix c) Projection Pursuit Estimates with an M–Scale Estimate

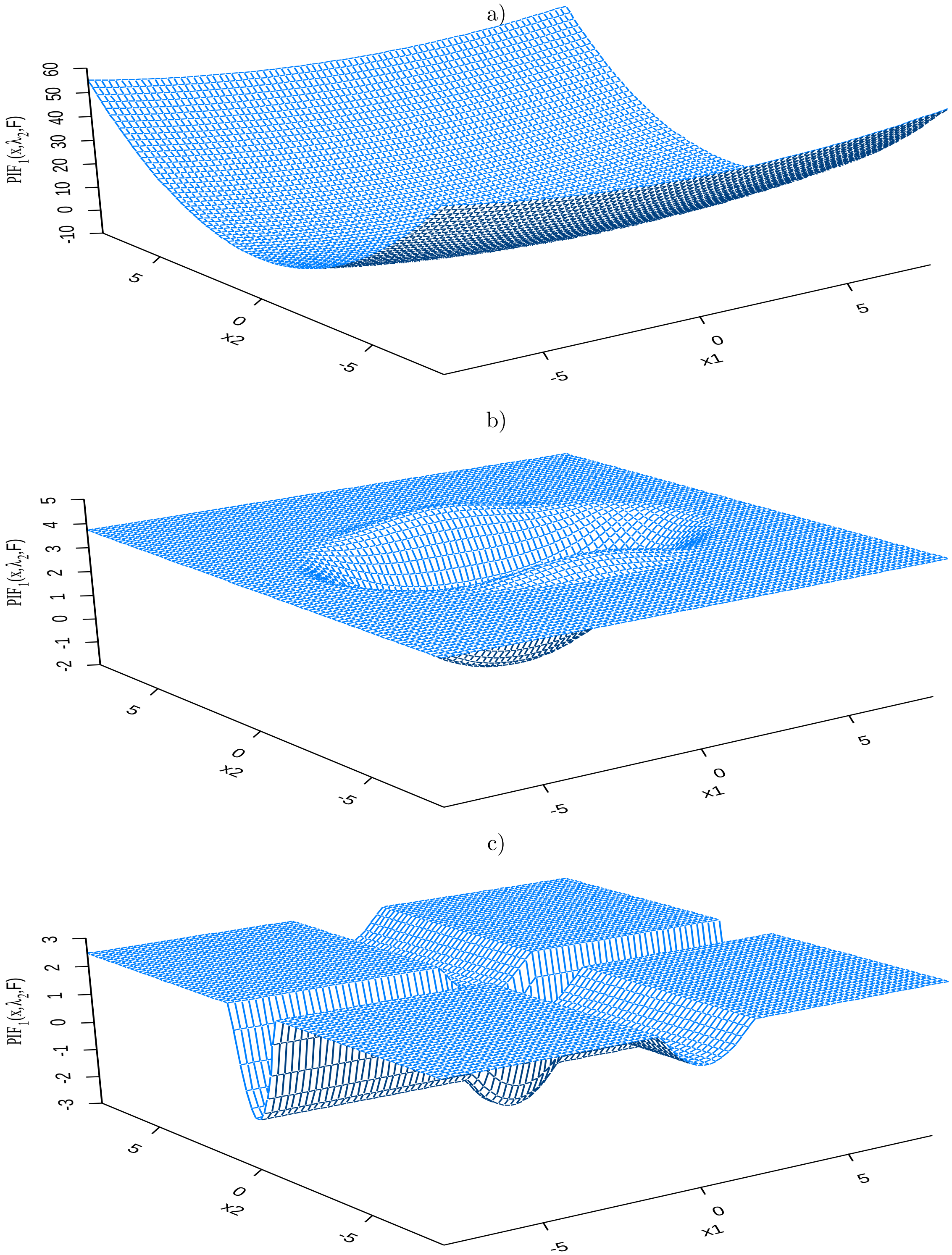

Figure 4: $\mathrm{PIF}_1(\mathbf{x}, \lambda_2, F)$ at $F = F_1 \times F_2$ with $F_1 = N(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N(\mathbf{0}, 4\mathrm{diag}(2,1))$. a) Maximum Likelihood Estimates b) Plug–in Estimates with an S–Scatter Matrix c) Projection Pursuit Estimates with an M–Scale Estimate

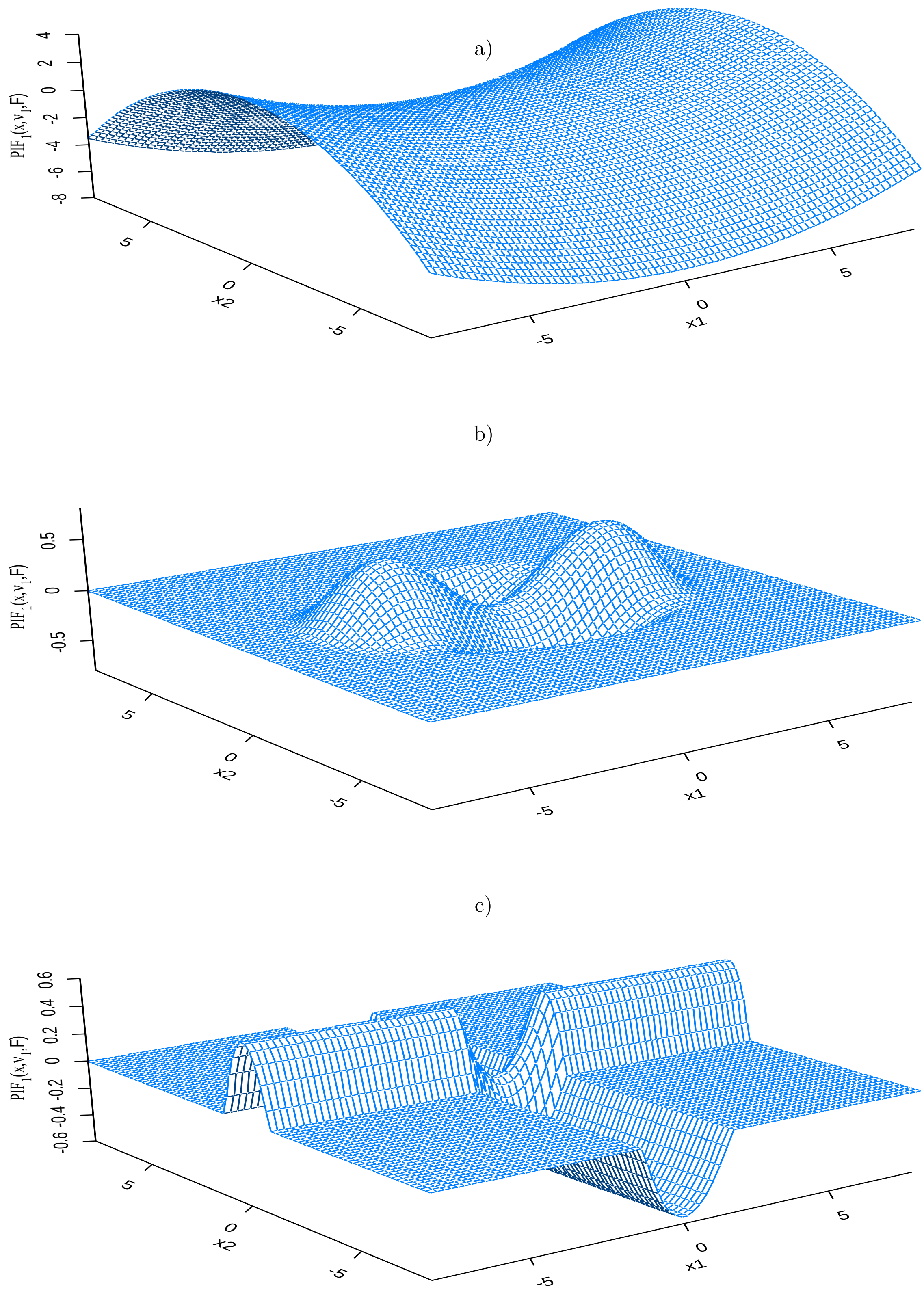

Figure 5: $\mathrm{PIF}_1(\mathbf{x}, \nu_1, F)$ at $F = F_1 \times F_2$ with $F_1 = N(\mathbf{0}, \mathrm{diag}(2,1))$ and $F_2 = N(\mathbf{0}, 4\mathrm{diag}(2,1))$. a) Maximum Likelihood Estimates b) Plug–in Estimates with an S–Scatter Matrix c) Projection Pursuit Estimates with an M–Scale Estimate

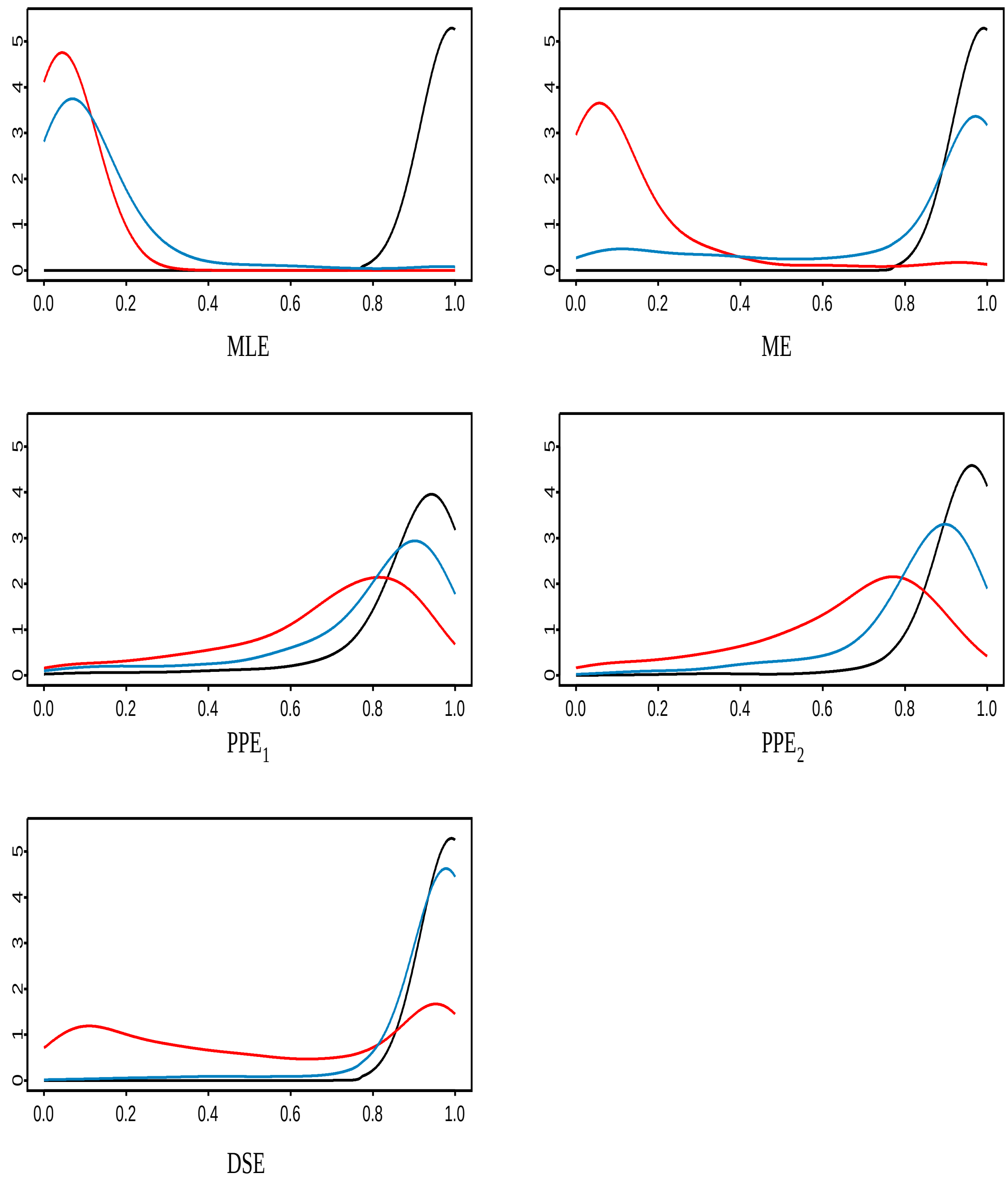

Figure 6: Density estimates of the cosinus of the angle between the estimated and the true direction related to the smallest eigenvalue. The densities of the estimates obtained under C_0, $C_{2,0.05}$ and $C_{2,0.1}$ are plotted in black, in light blue and in red, respectively.

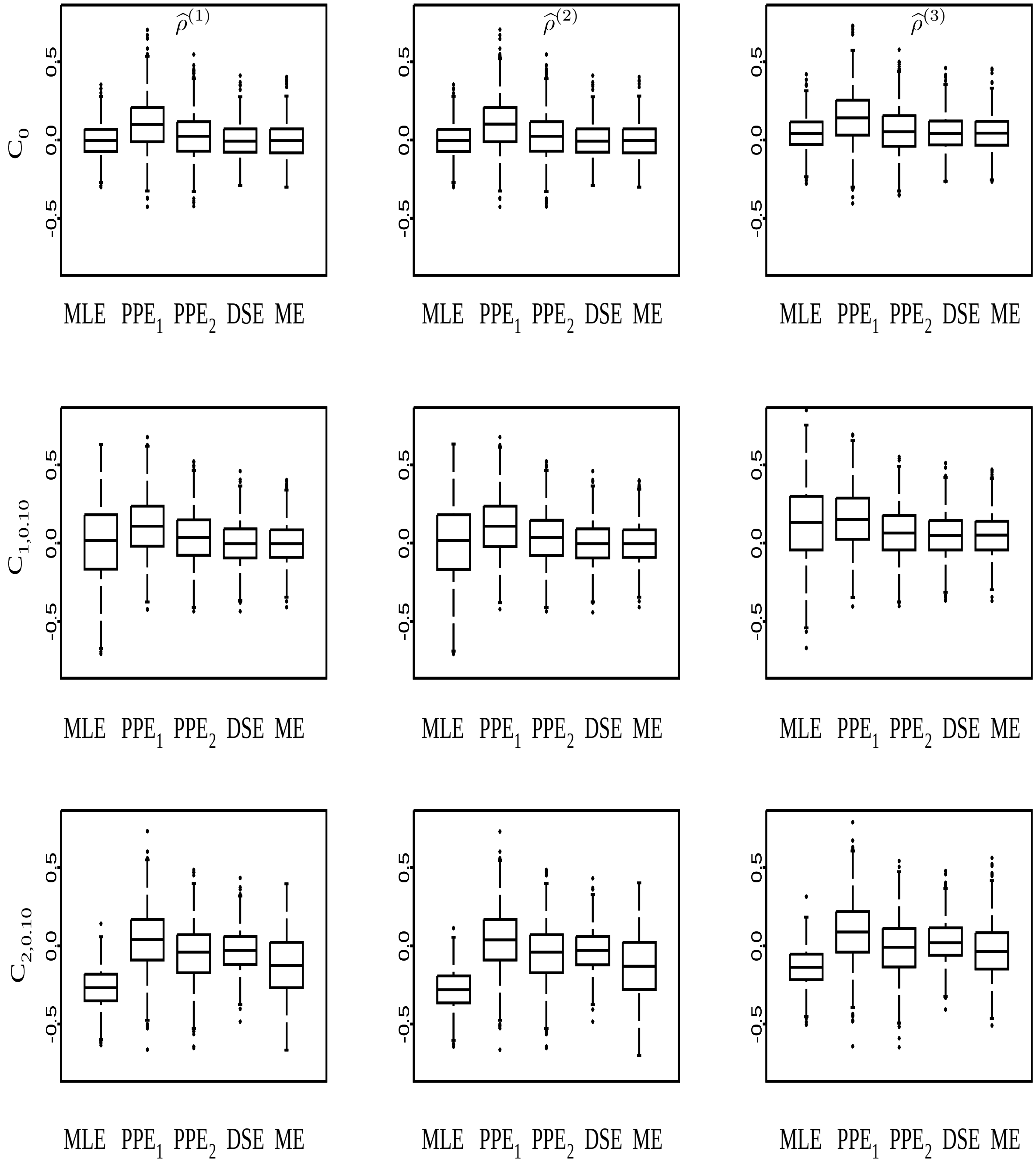

Figure 7: Boxplots of the centered logarithm of the estimates of the proportional constants, $\log(\widehat{\rho}) - \log(4)$

	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
λ	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
4	0.0414	0.2833	0.2046	0.0449	0.0431	0.1167	0.2832	0.1986	0.0493	0.0549
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
	0.3095	0.2932	0.2471	0.0479	0.0501	1.7748	0.3727	0.3421	0.0559	0.4399
	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
3	0.0700	0.5535	0.4069	0.0747	0.0728	0.2057	0.5140	0.3779	0.0809	0.0829
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
	1.3986	0.5893	0.4492	0.0816	0.0969	1.6549	0.7244	0.6642	0.1219	1.3372
	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
2	0.0358	0.3443	0.2060	0.0417	0.0376	0.1055	0.3286	0.2128	0.0435	0.0446
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
	1.7532	0.5070	0.3766	0.0721	0.1245	1.7628	0.7076	0.7313	0.9325	1.7135
	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
1	0.0117	0.1268	0.0650	0.0128	0.0124	0.0307	0.1246	0.0712	0.0135	0.0142
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
	1.8467	0.2740	0.2458	0.0326	0.0805	1.9208	0.5271	0.5820	0.9304	1.8625

Table 1: Median of the square distance between the estimated common principal directions and the true principal axes under a proportional model

Estimates $\widehat{\rho}^{(1)}$

	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
Mean	4.0146	4.5030	4.1569	4.0131	4.0139	4.1511	4.5219	4.1880	4.0179	4.0287
SD	0.4170	0.7845	0.6092	0.4441	0.4494	0.9917	0.8405	0.6975	0.5604	0.5555
MSE	0.1741	0.8687	0.3957	0.1974	0.2021	1.0063	0.9791	0.5219	0.3144	0.3094
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
Mean	3.2287	4.4318	4.0927	3.9617	3.9523	3.0868	4.2491	3.8789	3.9212	3.6070
SD	0.4656	0.7908	0.6511	0.4865	0.5761	0.3749	0.5250	0.8374	0.7065	0.7502
MSE	0.8123	0.8119	0.4326	0.2381	0.3342	0.9754	0.7634	0.5138	0.2819	0.7174

Estimates $\widehat{\rho}^{(2)}$

	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
Mean	4.0147	4.5016	4.1563	4.0132	4.0140	4.1517	4.5202	4.1876	4.0180	4.0288
SD	0.4171	0.7848	0.6092	0.4443	0.4496	0.9903	0.8405	0.6977	0.5608	0.5557
MSE	0.1742	0.8678	0.3956	0.1976	0.2023	1.0037	0.9772	0.5220	0.3149	0.3096
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
Mean	3.1998	4.4304	4.0922	3.9616	3.9494	3.0503	4.2479	3.8786	3.9206	3.5841
SD	0.4794	0.7906	0.6512	0.4868	0.5817	0.3814	0.8371	0.7065	0.5260	0.7744
MSE	0.8708	0.8106	0.4326	0.2384	0.3409	1.0483	0.7622	0.5139	0.2830	0.7728

Estimates $\widehat{\rho}^{(3)}$

	MLE	PPE_1	PPE_2	DSE	ME	MLE	PPE_1	PPE_2	DSE	ME
	$\boldsymbol{C_0}$					$\boldsymbol{C_{1,0.1}}$				
Mean	4.2009	4.7004	4.2909	4.2175	4.2074	4.6715	4.7287	4.3212	4.2391	4.2504
SD	0.4417	0.8300	0.6380	0.4719	0.4759	1.1294	0.8950	0.7254	0.6007	0.5945
MSE	0.2355	1.1800	0.4918	0.2700	0.2695	1.7268	1.3325	0.6294	0.4181	0.4161
	$\boldsymbol{C_{2,0.05}}$					$\boldsymbol{C_{2,0.1}}$				
Mean	3.6092	4.6361	4.2226	4.1665	4.1907	3.5129	4.4489	4.0107	4.1396	3.9566
SD	0.4729	0.8533	0.6818	0.5166	0.5988	0.4081	0.8931	0.7380	0.5557	0.7130
MSE	0.3765	1.1332	0.5145	0.2947	0.3950	0.4041	0.9994	0.5448	0.3283	0.5102

Table 2: Estimation of the proportionality constants