Robust multiple-set linear canonical analysis based on minimum covariance determinant estimator

Ulrich Djemby Bivigou and Guy Martial Nkiet
Université des Sciences et Techniques de Masuku, Franceville, Gabon

ABSTRACT
In this paper, we introduce a robust version of multiple-set linear canonical analysis (MSLCA) by using the MCD estimator of the covariance operator of the involved random vector. The related influence functions are derived and are shown to be bounded. Asymptotic properties of the introduced robust MSLCA are obtained and allow us to propose a robust test for mutual non correlation. A simulation study, which shows that this test outperforms classical ones in the presence of disturbed data, is presented.

1. Introduction
Many multivariate statistical methods are based on empirical covariance operators. That is the case for multivariate regression, principal components analysis, factor analysis, linear discriminant analysis, linear canonical analysis, multiple-set linear canonical analysis, and so on. However, these empirical covariance operators are known to be extremely sensitive to outliers. That is an undesirable property that makes the preceding methods themselves sensitive to outliers. For overcoming this problem, robust alternatives have been proposed for these methods, mainly by replacing the aforementioned empirical covariance operators by robust estimators. In this vein, robust versions of several multivariate statistical methods have been introduced, especially for multiple regression (Rousseeuw et al. 2004), principal components analysis (Croux and Haesbroeck 2000; Cui, He, and Ng 2003), linear discriminant analysis (Croux, Filmozer, and Joossens 2008), linear canonical analysis (Croux and Dehon 2002), linear canonical analysis (Croux and Dehon 2002). Multiple-set linear canonical analysis (MSLCA) is an important multivariate statistical method that analyzes the relationship between more than two random vectors, so generalizing linear canonical analysis. It has been introduced for many years (e.g., Gifi 1990) and has been studied since then under different aspects (e.g., Takane, Hwang, and Abdi 2008; Tenenhaus and Tenenhaus 2011; Hwang et al. 2012). A formulation of MSLCA within the context of Euclidean random variables has been made recently (Nkiet 2017) and allowed to obtain an asymptotic theory for this analysis when it is estimated by using empirical covariance operators. To the best of our knowledge, such estimation of MSLCA is the one that have been tackled in the literature, despite the fact that it is non
robust because its depends on a non robust estimator of covariance operator. So, there is an interest in introducing a robust estimation of MSLCA as it was done for other multivariate statistical methods. This can be done by using a robust estimator of the covariance operator of the involved random vector instead of the empirical covariance operator. Among such robust estimators, the minimum covariance determinant (MCD) estimator has been extensively studied (Butler, Davies, and Jhun 1993; Croux and Haesbroeck 1999; Cator and Lopuhaä 2010, 2012), and it is known to have good robustness properties. Also, its asymptotic properties have been obtained (Butler, Davies, and Jhun 1993; Cator and Lopuhaä 2010, 2012) mainly under elliptical distribution.

In this paper, we propose a robust version of MSLCA based on MCD estimator of the covariance operator. We start by recalling, in Section 2, the notion of MSLCA for Euclidean random variables. In Section 3, we introduce a robust estimation of MSLCA (denoted by RMSLCA) by using the MCD estimator of the covariance operator on which this analysis is defined. Then we derive the influence function of the operator that determines RMSLCA and we prove that it is bounded. We also derive the influence functions of the canonical coefficients and the canonical directions. Section 4 is devoted to asymptotic properties of RMSLCA. We obtain limiting distributions that are then used in Section 5 for defining a robust test for mutual non correlation. The proofs of all theorems and propositions are postponed in the Appendix section.

2. Multiple-set linear canonical analysis

For an integer $K \geq 2$, we consider random variables $X_1, ..., X_K$ with values into Euclidean vector spaces $\mathcal{X}_1, ..., \mathcal{X}_K$ respectively. Denoting by $\langle \cdot, \cdot \rangle_k$ the inner product of $\mathcal{X}_k$, and by $\| \cdot \|_k$ the associated norm, we assume that the following condition holds:

$(A_1)$: for $k \in \{1, ..., K\}$, we have $E(X_k) = 0$ and $E(\|X_k\|^2_k) < +\infty$.

Furthermore, we consider the random vector $X = (X_1, ..., X_K)$ with values into the Euclidean vector space $\mathcal{X} := \mathcal{X}_1 \times \cdots \mathcal{X}_K$ with induced inner product denoted by $\langle \cdot, \cdot \rangle_{\mathcal{X}}$. We can give the following definition of multiple-set linear canonical analysis (see Nkiet 2017):

**Definition 1.** The multiple-set linear canonical analysis (MSLCA) of $X$ is the search of a sequence $(a^{(j)})_{1 \leq j \leq q}$ of vectors of $\mathcal{X}$, where $q = \dim(\mathcal{X})$, satisfying:

$$a^{(j)} = \arg \max_{a \in C_j} \mathbb{E}\left(\langle a, X \rangle^2_{\mathcal{X}}\right),$$

(1)

where $C_1 = \left\{ a \in \mathcal{X} / \sum_{k=1}^K \text{var}(\langle a_k, X_k \rangle_k) = 1 \right\}$ and for $j \geq 2$:

$$C_j = \left\{ a \in C_1 / \sum_{k=1}^K \text{cov}(\langle a^{(r)}_k, X_k \rangle_k, \langle a_k, X_k \rangle_k) = 0, \quad \forall r \in \{1, ..., j-1\} \right\}.$$

A solution of the above maximization problem is obtained from spectral analysis of an operator that will know be specified. For $(k, \ell) \in \{1, ..., K\}^2$, let us consider the covariance operators $V_{k\ell} = \mathbb{E}(X_\ell \otimes X_k) = V^*_{\ell k}$ and $V_k := V_{kk}$, where $\otimes$ denotes the tensor
product such that \( x \otimes y \) is the linear map: \( h \mapsto <x, h \rangle y \), and \( A^* \) denotes the adjoint of \( A \). Letting \( \tau_k \) be the canonical projection \( \tau_k : x \in \mathcal{X} \mapsto x_k \in \mathcal{X}_k \), we consider the operators defined as \( \Phi = \sum_{k=1}^{K} \tau_k^* V_k \tau_k \) and \( \Psi = \sum_{k=1}^{K} \sum_{\ell \neq k} \tau_k^* V_k \tau_\ell \). The covariance operator \( V_k \) is a self-adjoint and positive operator; we assume throughout this paper that it is invertible. Then, it is easy to check that \( \Phi \) is also self-adjoint positive and invertible operator, and we put \( T = \Phi^{-1/2} \Psi \Phi^{-1/2} \). The spectral analysis of this last operator gives a solution of the maximization problem specified in Definition 1. Indeed, if \( \{\beta^{(1)},...,\beta^{(q)}\} \) is an orthonormal basis of \( \mathcal{X} \) such that \( \beta^{(j)} \) is an eigenvector of \( T \) associated with the \( j \)-th largest eigenvalue \( \rho_j \), then we obtain a solution of Equation (1) by taking \( \alpha^{(j)} = T^{1/2} \beta^{(j)} \), and we have \( \rho_j = < \beta^{(j)}, T \beta^{(j)} >_\mathcal{X} \). The \( \rho_j \)'s are termed the canonical coefficients and the \( \alpha^{(j)} \)'s are termed the canonical directions. Note that \( T \) can be expressed as a function of the covariance operator \( V = \mathbb{E}(X \otimes X) \) of \( X \). Indeed, by using properties of tensor products (see Dauxois, Romain, and Viguier 1994), it is easy to check that \( V_{k \ell} = \tau_k V \tau_\ell^* \) and \( V_k = \tau_k V \tau_k^* \). Therefore, \( T = f(V)^{-1/2} g(V) f(V)^{-1/2} \), where

\[
f(A) = \sum_{k=1}^{K} \tau_k^* A \tau_k \quad \text{and} \quad g(A) = \sum_{k=1}^{K} \sum_{\ell \neq k} \tau_k^* A \tau_\ell.
\]

(2)

3. Robust estimation

In this section, we introduce a robust estimation of MSLCA by replacing \( V \) by a robust estimator. More precisely, we use the minimum covariance determinant (MCD) estimator of \( V \). We consider the following assumption:

\((\mathcal{A}_2)\): the distribution \( \mathbb{P}_X \) of \( X \) is an elliptical contoured distribution with density \( f_X(x) = (\det(V))^{-1/2} h(<x, V^{-1}x >_\mathcal{X}) \), where \( h : [0, +\infty[ \rightarrow [0, +\infty[ \) is a function having a strictly negative derivative \( h' \).

3.1. Estimation based on MCD estimator

Letting \( \gamma \) be a fixed real such that \( 0 < \gamma < 1 \), we consider a subsample \( S \subset \{X^{(1)},...,X^{(n)}\} \) of size \( h_n \geq \lfloor n \gamma \rfloor \), where \( X^{(i)} = (X_1^{(i)},...,X_K^{(i)}) \), and we define the empirical mean and covariance operator based on this subsample by:

\[
\hat{M}_n(S) = \frac{1}{h_n} \sum_{X^{(i)} \in \mathcal{S}} X^{(i)}
\]

and

\[
\hat{V}_n(S) = \frac{1}{h_n} \sum_{X^{(i)} \in \mathcal{S}} \left( X^{(i)} - \hat{M}_n(S) \right) \otimes \left( X^{(i)} - \hat{M}_n(S) \right).
\]

We denote by \( \hat{S}_n \) the subsample of \( \{X^{(1)},...,X^{(n)}\} \) which minimizes the determinant \( \det(\hat{V}_n(S)) \) of \( \hat{V}_n(S) \) over all subsamples of size \( h_n \). Then, the MCD estimators of the
mean and the covariance operator of $X$ are $\hat{M}_n(\hat{S}_n)$ and $\hat{V}_n(\hat{S}_n)$, respectively. It is well known that these estimators are robust and have high breakdown points (see, e.g., Rousseeuw et al. 2004). From them, we can introduce an estimator of MSLCA which is expected to be also robust. Indeed, putting $\tilde{V}_n := \hat{V}_n(\hat{S}_n)$, we consider the random operators $\Phi_n = \sum_{k=1}^{K} \tau_k \tilde{V}_{k,n} \tau_k$ and $\Psi_n = \sum_{k=1}^{K} \sum_{\ell \neq k} \tau_k \tilde{V}_{k,\ell,n} \tau_\ell$, where $\tilde{V}_{k,n} = \tau_k \tilde{V}_n \tau_k^*$ and $\tilde{V}_{k,\ell,n} = \tau_k \tilde{V}_n \tau_\ell^*$, and we estimate $T$ by

$$\tilde{T}_n = \Phi_n^{-1/2} \Psi_n \Phi_n^{-1/2}.$$  

Considering the eigenvalues $\tilde{\rho}_{1,n} \geq \tilde{\rho}_{2,n} \cdots \geq \tilde{\rho}_{q,n}$ of $\tilde{T}_n$, and $\{\tilde{\beta}_n^{(1)}, \ldots, \tilde{\beta}_n^{(q)}\}$ an orthonormal basis of $\mathcal{X}$ such that $\tilde{\beta}_n^{(j)}$ is an eigenvector of $\tilde{T}_n$ associated with $\tilde{\rho}_{j,n}$, we estimate $\rho_j$ by $\tilde{\rho}_{j,n}$, $\beta_j^{(j)}$ by $\tilde{\beta}_n^{(j)}$, and $\alpha^{(j)}$ by $\tilde{x}_n^{(j)} = \Phi_n^{-1/2} \tilde{\beta}_n^{(j)}$. This gives a robust MSLCA that we denote by RMSLCA.

### 3.2. Influence functions

For studying the effect of a small amount of contamination at a given point on MSLCA it is important, as usual in robustness literature (see Hamble et al. 1986), to use influence function. Recall that the influence function of a functional $S$ at $\mathbb{P}$ is defined as

$$\text{IF}(x; S, \mathbb{P}) = \lim_{\varepsilon \to 0} \frac{S((1-\varepsilon)\mathbb{P} + \varepsilon \delta_x) - S(\mathbb{P})}{\varepsilon},$$

where $\delta_x$ is the Dirac measure putting all its mass in $x$. In order to derive the influence functions related to the above estimator of MSLCA, we have to specify the functional that corresponds to it. For doing that, we will first recall the functional associated to the above MCD estimator of covariance operator. Let us consider the set $E_\gamma = \{x \in \mathcal{X}, x, V^{-1}x > \mathcal{X} \leq \gamma^2\}$, where $\gamma^2$ is determined by the equation

$$\frac{2\pi^{q/2}}{\Gamma(q/2)} \int_0^{\gamma^2} t^{q-1} h(t^2) \, dt = \gamma,$$

$\Gamma$ being the usual gamma function. The functional $\mathbb{V}_{1,\gamma}$ related to the aforementioned MCD estimator of $V$ is defined in Cator and Lopuhaä (2010) (see also Butler, Davies, and Jhun (1993) and Croux and Haesbroeck (1999)) by

$$\mathbb{V}_{1,\gamma}(\mathbb{P}) = \frac{1}{\gamma} \int_{E_\gamma} (x - M_{\mathbb{P}}(E_\gamma)) \otimes (x - M_{\mathbb{P}}(E_\gamma)) d\mathbb{P}(x),$$

where

$$M_{\mathbb{P}}(B) = \frac{1}{\gamma} \int_{B} x \, d\mathbb{P}(x).$$

It is known that $\mathbb{V}_{1,\gamma}(\mathbb{P}_x) = \sigma_{\gamma}^2 V$ where

$$\sigma_{\gamma}^2 = \frac{2\pi^{q/2}}{\gamma} \frac{\int_0^{\gamma^2} t^{q-1} h(t^2) \, dt}{\Gamma(q/2)}.$$  

$$\sigma_{\gamma}^2 = \frac{2\pi^{q/2}}{\gamma} \frac{\Gamma(q/2)}{\Gamma(q/2)} \int_0^{\gamma^2} t^{q-1} h(t^2) \, dt.$$  

(4)
Therefore, the functional \( T_{1,\gamma} \) related to \( T \) is defined as

\[
T_{1,\gamma}(\mathbb{P}) = f(\mathbb{V}_{1,\gamma}(\mathbb{P}))^{-1/2} g(\mathbb{V}_{1,\gamma}(\mathbb{P})) f(\mathbb{V}_{1,\gamma}(\mathbb{P}))^{-1/2}
\]

where \( f \) and \( g \) are given in Equation (2). Now, we can give the influence functions related to RMSLCA of \( \mathcal{X}_k \). We make the following assumption:

(\( \mathcal{A}_3 \)): For all \( k \in \{1, \ldots, K\} \), we have \( V_k = I_k \), where \( I_k \) denotes the identity operator of \( \mathcal{X}_k \).

First, putting

\[
\kappa_0 = \frac{\pi^{q/2}}{(q + 2)\Gamma(q/2 + 1)} \int_0^{r(\gamma)} t^{q+3} h'(t^2) \, dt,
\]

and \( T_\gamma = T_{1,\gamma}(\mathbb{P}_X) \), we have:

**Theorem 1.** We suppose that the assumptions (\( \mathcal{A}_1 \)) to (\( \mathcal{A}_3 \)) hold. Then

\[
\text{IF}(x; T_\gamma, \mathbb{P}_X) = -\frac{\sigma_\gamma^{-2}}{2\kappa_0} I_{E_\gamma}(x) \phi(x, V),
\]

where

\[
\phi(x, V) = \sum_{k=1}^K \sum_{\ell=1}^K \frac{1}{2} \frac{r_\gamma^k(x_k \otimes x_k) V_{k\ell} \tau_\ell}{\tau_\ell} - \frac{1}{2} \tau_\ell^k V_{\ell k}(x_k \otimes x_k) \tau_k + \tau_\ell^k(x_\ell \otimes x_k) \tau_\ell.
\]

The following proposition shows that RMSLCA is robust as the preceding influence function is bounded. We denote by \( \| \cdot \|_{\infty} \) the usual operators norm defined by \( \|A\|_{\infty} = \sup_{x \in \mathcal{X}, x \neq 0} (\|Ax\|_{\mathcal{Y}}/\|x\|_{\mathcal{X}}) \).

**Proposition 1.** We suppose that the assumptions (\( \mathcal{A}_1 \)) to (\( \mathcal{A}_3 \)) hold. Then,

\[
\sup_{x \in \mathcal{X}} \|\text{IF}(x; T_\gamma, \mathbb{P}_X)\|_{\infty} \leq \frac{\sigma_\gamma^{-2}}{2|\kappa_0|} K(K - 1)(\|V\|_{\infty} + 1)\|V^{1/2}\|_{\infty}^2 r^2(\gamma).
\]

Now, we give in the following theorem, the influence functions related to the canonical coefficients and the canonical directions obtained from RMSLCA. For \( j \in \{1, \ldots, q\} \), denoting by \( \mathbb{R}_{\gamma,j} \) (resp. \( \mathbb{B}_{\gamma,j} \); resp. \( \mathbb{A}_{\gamma,j} \)) the functional such that \( \mathbb{R}_{\gamma,j}(\mathbb{P}) \) is the \( j \)-th largest eigenvalue of \( T_{1,\gamma}(\mathbb{P}) \) (resp. the associated eigenvector; resp. \( \mathbb{A}_{\gamma,j}(\mathbb{P}) = f(\mathbb{V}_{1,\gamma}(\mathbb{P}))^{-1/2} \mathbb{B}_{\gamma,j}(\mathbb{P}) \)), we put \( \rho_{\gamma,j} = \mathbb{R}_{\gamma,j}(\mathbb{P}_X), \beta_{\gamma,j}^{(i)} = \mathbb{B}_{\gamma,j}(\mathbb{P}_X) \) and \( \alpha_{\gamma,j}^{(i)} = \mathbb{A}_{\gamma,j}(\mathbb{P}_X) \). Considering

\[
\nu_0 = \frac{2\pi^{q/2}}{\Gamma(q/2)} h(r(\gamma)^2) r(\gamma)^{q-1} \sigma_\gamma, \quad \nu_1 = \frac{r(\gamma)}{\sigma_\gamma}, \quad \nu_2 = \frac{2\nu_0\nu_2^3}{\sigma_\gamma (q + 2)} - \frac{2\gamma}{\sigma_\gamma},
\]

\[
\kappa_1 = -\frac{r(\gamma)^2}{q\gamma}, \quad \kappa_2 = \frac{\sigma_\gamma \nu_2 + 2\gamma}{q\gamma \sigma_\gamma \nu_2}, \quad \kappa_3 = \frac{2}{\sigma_\gamma \nu_2}, \quad \kappa_4 = \frac{r(\gamma)^2 - q\sigma_\gamma^2}{q}
\]

and putting
We suppose, in addition, that 

\[ \text{covariance operator equal to that of the random operator } I \]

where \( I \) denotes the identity operator of \( \mathcal{X} \), and 

\[ \rho_i = \frac{1}{\sqrt{n}} \sum_{k=1}^{K} \sum_{\ell=1}^{K} \rho_{ij} \left( < \beta^{(m)}_{k}, x_k > k < x_{\ell}, \beta^{(j)}_{\ell} > \ell \right) \]

\[ - \frac{1}{2} < \beta^{(m)}_{k}, x_k > k < x_k, V_{k\ell} \beta^{(j)}_{\ell} > k \]

\[ - \frac{1}{2} < x_k, V_{k\ell} \beta^{(m)}_{\ell} > \ell < x_k, \beta^{(j)}_{\ell} > k \]

\[ - \frac{1}{2} \left( \sum_{k=1}^{K} \tau^{(j)}_{k}(x_k \otimes x_k)\tau_{k} + < \beta^{(j)}_{k}, x_k >_{k}^{2} \right) - 2I \beta^{(j)} \]

where \( I \) denotes the identity operator of \( \mathcal{X} \), we have:

**Theorem 2.** We suppose that the assumptions \((\mathcal{A}_1)\) to \((\mathcal{A}_3)\) hold. Then, for any \( x \in \mathcal{X} \) and any \( j \in \{1, \ldots, q\} \), we have:

(i) \( \text{IF}(x; \rho_{1j}, \mathbb{P}_X) = - \frac{\sigma^2}{2\kappa_0} I_{E_1}(x) \sum_{k=1}^{K} \sum_{\ell=1}^{K} < \beta^{(j)}_{k}, x_k > k < x_{\ell}, \beta^{(j)}_{\ell} > \ell. \)

(ii) We suppose, in addition, that \( \rho_1 > \rho_2 > \cdots > \rho_q \). Then:

\[ \text{IF}(x; \beta^{(j)}_{j}, \mathbb{P}_X) = - \frac{\sigma^3}{2\kappa_0} I_{E_1}(x) \lambda_j(x, V) \]

\[ + \sigma^3 \left( \left( \frac{1}{2\kappa_0} - \kappa_1 - \kappa_2 \| V^{-1/2} x \|_{\mathcal{X}}^2 \right) I_{E_1}(x) - \kappa_4 \right) \beta^{(j)}. \]

**4. Asymptotic distributions**

In this section we deal with asymptotic expansion for RMSLCA. We first establish asymptotic normality for \( T_n \) and then we derive the asymptotic distribution of the canonical coefficients.

**Theorem 3.** Under the assumptions \((\mathcal{A}_1)\) to \((\mathcal{A}_3)\), \( \sqrt{n}(T_n - T) \) converges in distribution, as \( n \to +\infty \), to a random variable \( U_j \), having a normal distribution with mean 0 and covariance operator equal to that of the random operator

\[ Z_j = \sigma^2 \kappa_3 I_{E_1}(X) \sum_{k=1}^{K} \sum_{\ell=1}^{K} \left\{ - \frac{1}{2} \tau_{k}^{(j)}(X_k \otimes X_k) V_{k\ell} \tau_{\ell} + \tau_{k}^{(j)} V_{k\ell} (X_k \otimes X_k) \tau_{k} \right\} \]

\[ + \tau_{k}^{(j)}(X_k \otimes X_k) \tau_{\ell} \] \[ + \left( \sigma^2 - 1 \right) w\left( \| V^{-1/2} x \|_{\mathcal{X}}^2 \right) \sum_{k=1}^{K} \sum_{\ell=1}^{K} \tau_{k}^{(j)} V_{k\ell} \tau_{\ell}, \]

where \( w : [0, +\infty [ \to \mathbb{R} \) is the function defined by

\[ w(t) = I_{[0, r(\cdot)]}(t)(\kappa_1 + \kappa_2 t^2) + \kappa_4, \]

and \( \kappa_1, \kappa_2, \kappa_3 \) and \( \kappa_4 \) are given in Equation (6).

This theorem allows to obtain asymptotic distributions for the canonical coefficients. We have:
Corollary 1. We suppose that the assumptions (\(\mathcal{A}_1\)) to (\(\mathcal{A}_3\)) hold and that the canonical coefficients satisfy: \(\rho_1 > \rho_2 > \cdots > \rho_q\). Then, for any \(j \in \{1, \ldots, q\}\), \(\sqrt{n}(\hat{\rho}_{j,n} - \rho_j)\) converges in distribution, as \(n \to +\infty\), to the normal distribution \(N(0, \nu_j^2)\), where:

\[
\nu_j^2 = \text{Var} \left( \sum_{k=1}^K \sum_{\ell=1}^K \left\{ \sigma_{\gamma}^{-2} \kappa_3 I_{E_t}(X) < \beta^{(j)}_{k}, X_k > \kappa_2 \| X \|_{X}^{\frac{2}{\gamma}} \right. \right) - 1 \left( I_{E_t}(X) \left( \kappa_1 + \kappa_2 \| V^{-1/2}X \|_{X}^{2} \right) + \kappa_4 \right) < \beta^{(j)}_{\ell}, V_{k\ell} \beta^{(j)}_{\ell} > \kappa_3 \right\}.
\]

5. Robust test for mutual non correlation

In this section we consider the problem of testing for mutual non correlation between \(X_1, X_2, \ldots, X_K\). This is testing for the null hypothesis

\[\mathcal{H}_0 : \forall (k, \ell) \in \{1, \ldots, K\}^2, k \neq \ell, V_{k\ell} = 0\]

against the alternative

\[\mathcal{H}_1 : \exists (k, \ell) \in \{1, \ldots, K\}^2, k \neq \ell, V_{k\ell} \neq 0.\]

This testing problem was already considered in Nkiet (2017); a test statistic which depends on empirical covariance operator was then proposed and its asymptotic distribution under the null hypothesis was derived. Since the resulting testing method may be non robust, it is interesting to propose a new method that depends instead on a robust estimator of the covariance operator of \(X\). Here, we introduce a test statistic constructed similarly to the one of Nkiet (2017), but with the MCD estimator of the aforementioned covariance operator. It is then defined as

\[
\tilde{S}_n = \sum_{k=2}^K \sum_{\ell=1}^{k-1} \text{tr} \left( \pi_{k\ell} (\tilde{T}_n) \pi_{k\ell} (\tilde{T}_n)^* \right),
\]

where \(\tilde{T}_n\) is the estimator given in Equation (3) and \(\pi_{k\ell}\) is the operator: \(A \mapsto \tau_{k} A \tau_{\ell}^*\).

5.1. Asymptotic distribution under null hypothesis

Let us consider

\[
\tau = \frac{\sigma_{\gamma}^{-4} \kappa_3^2}{q(q+1)} \mathbb{E}(I_{E_t}(X) \| X \|_{X}^{4}).
\]

Then, we have:

**Theorem 4.** Suppose that the assumptions (\(\mathcal{A}_1\)) to (\(\mathcal{A}_3\)) hold. Then, under \(\mathcal{H}_0\), the sequence \(\tau^{-1} n\tilde{S}_n\) converges in distribution, as \(n \to +\infty\), to \(\chi_d^2\), where \(d = \sum_{k=1}^K \sum_{\ell=1}^{k-1} p_k p_\ell\) with \(p_k = \text{dim}(X_k)\).

**Remark 1.** For performing this test in practice one has to estimate the unknown parameter \(\tau\). This can be done by using an estimate \(\hat{k}_3\) of \(\kappa_3\) as in Equation (6) by replacing \(r(\gamma)\) (resp. \(\sigma_{\gamma}\)) by an estimate \(\hat{r}\) (resp. \(\hat{\sigma}\)), and by considering
\[
\hat{r} = \frac{\hat{\sigma}^{-4} K_3^2}{q(q+1)n} \sum_{i=1}^{n} \mathbf{1}_{\tilde{E}}(X^{(i)}) \|X^{(i)}\|_\chi^4,
\]
where \(\tilde{E} = \left\{ x \in \mathcal{X}, <x, \tilde{V}_n^{-1}x> \leq r^2 \right\}\). A consistent estimator \(\hat{r}\) of \(r(\gamma)\) is defined in Butler, Davies, and Jhun (1993).

### 5.2. Simulations

In order to observe the performance of the above introduced robust test, we made simulations within a framework corresponding to the case of \(K = 3, p_1 = 2, p_2 = 3\) and \(p_3 = 5\). We computed empirical powers over 1000 replications with nominal significance level \(\alpha = 0.05\) from the proposed method, denoted here by MCD, and two known methods. These known methods are the likelihood ratio test (LRT) defined in Anderson (1984), and the test introduced in Nkiet (2017) based on classical empirical covariance operators (EMP). For sample sizes \(n = 50, 75, 85\), we generated 1000 independent replicates of a sample of a 10-dimensional random vector \(X\) having a contaminated normal distribution given by \((1 - e/100) N(0, \Sigma) + (e/100) N(m_1, I_{10})\), where \(m_1 = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1)^T, I_k\) is the \(k \times k\) identity matrix, and \(\Sigma\) is a symmetric \(10 \times 10\) matrix given by

\[
\Sigma = \begin{pmatrix}
\Sigma_1 & \Sigma_{12} & \Sigma_{13} \\
\Sigma_{21} & \Sigma_2 & \Sigma_{23} \\
\Sigma_{31} & \Sigma_{32} & \Sigma_3
\end{pmatrix}
\]

with \(\Sigma_1 = I_2, \Sigma_2 = I_3, \Sigma_3 = I_5\). For \((k, \ell) \in \{1, 2, 3\}\) with \(k < \ell, \Sigma_{k\ell}\) is a \(p_k \times p_\ell\) matrix with general term \(\sigma_{ij}^{k\ell} = (\delta_{k\ell} - i - j) \times 0.001\), where \(\delta_{12} = 133, \delta_{13} = 130\) and \(\delta_{23} = 128\).
The percentage of contamination $e$ was taken from 0% to 50%. The obtained results are given in Figures 1–3. They show that the proposed robust method (MCD) outperforms the two others, with a gap which increases as $e$ increases, especially for $e > 25\%$. Furthermore, as $n$ increases LRT and EMP are getting closer to each other and still outperformed by MCD.

**Figure 2.** Empirical power versus percentage of contamination for MCD (red), LRT (blue), EMP (black) with sample size $n = 75$ and nominal significance level $\alpha = 0.05$.

**Figure 3.** Empirical power versus percentage of contamination for MCD (red), LRT (blue), EMP (black) with sample size $n = 85$ and nominal significance level $\alpha = 0.05$. 

The percentage of contamination $e$ was taken from 0% to 50%. The obtained results are given in Figures 1–3. They show that the proposed robust method (MCD) outperforms the two others, with a gap which increases as $e$ increases, especially for $e > 25\%$. Furthermore, as $n$ increases LRT and EMP are getting closer to each other and still outperformed by MCD.
6. Conclusion

In this paper we have introduced a robust version of multiple-set linear canonical analysis (RMSLCA), obtained by replacing the involved covariance operator by its MCD estimator, that allows to study the relationship between more than two random vectors while being less sensitive to outliers than the traditional version. The related influence functions have been derived and was shown to be bounded. The asymptotic distribution of the operator that gives RMSLCA is given as well as that of its eigenvalues. This permitted to propose a robust test for mutual non correlation which is shown, by a simulation study, to perform better than classical methods, especially when there are outliers. All this shows the interest of using, in applications, this robust method rather than the classic method since, in practice, it is common to have outliers in a dataset.

References

Anderson, T. W. 1984. *An introduction to multivariate statistical analysis*. New York: Wiley.

Butler, R. W., P. L. Davies, and M. Jhun. 1993. Asymptotics for the minimum covariance determinant estimator. *The Annals of Statistics* 21 (3):1385–400. doi:10.1214/aos/1176349264.

Cator, E. A., and H. P. Lopuhaä. 2010. Asymptotic expansion of the minimum covariance determinant estimators. *Journal of Multivariate Analysis* 101 (10):2372–88. doi:10.1016/j.jmva.2010.06.009.

Cator, E. A., and H. P. Lopuhaä. 2012. Central limit theorem and influence function of the MCD estimators at general multivariate distributions. *Bernoulli* 18 (2):520–51. doi:10.3150/11-BEJ353.

Croux, C., and C. Dehon. 2002. Analyse canonique basée sur des estimateurs robustes de la matrice des covariances. *Revue de statistique appliquée* 50:51–26.

Croux, C., P. Filmozer, and K. Joossens. 2008. Classification efficiencies for robust linear discriminant analysis. *Statistica Sinica* 18:603–18. doi:10.1093/biomet/87.3.603.

Cui, H., X. He, and K. W. Ng. 2003. Asymptotic distributions of principal components based on robust dispersions. *Biometrika* 90 (4):953–66. doi:10.1093/biomet/90.4.953.

Dauxois, J., Y. Romain, and S. Viguier. 1994. Tensor product and statistics. *Linear Algebra and Its Applications* 210:59–88. doi:10.1016/0024-3795(94)90466-9.

Dossou-Gbete, S., and A. Pousse. 1991. Asymptotic study of eigenelements of a sequence of random selfadjoint operators. *Statistics* 22 (3):479–91. doi:10.1080/02331889108802329.

Gifi, A. 1990. *Nonlinear multivariate analysis*. New York: Wiley.

Hample, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. 1986. *Robust statistics: The approach based on influence functions*. New York: Wiley.

Hwang, H., K. Jung, Y. Takane, and T. S. Woodward. 2012. Functional multiple-set canonical correlation analysis. *Psychometrika* 77 (1):48–64. doi:10.1007/s11336-011-9234-4.

Lopuhaä, H. P. 1997. Asymptotic expansion of S-estimators of location and covariance. *Statistica Neerlandica* 51 (2):220–37. doi:10.1111/1467-9574.00051.

Nkiet, G. M. 2017. Asymptotic theory of multiple-set linear canonical analysis. *Mathematical Methods of Statistics* 26 (3):196–211. doi:10.3103/S1066530717030036.

Rousseeuw, P. J., S. Van Aelst, K. Van Driessen, and J. A. Gulló. 2004. Robust multivariate regression. *Technometrics* 46 (3):293–305. doi:10.1198/004017004000000329.

Takane, Y., H. Hwang, and H. Abdi. 2008. Regularized multiple-set canonical correlation analysis. *Psychometrika* 73 (4):753–75. doi:10.1007/s11336-008-9065-0.
Tenenhaus, A., and M. Tenenhaus. 2011. Regularized generalized canonical correlation analysis. *Psychometrika* 76 (2):257–84. doi:10.1007/s11336-011-9206-8.

**Appendix**

**A.1. Proof of Theorem 1**

It is shown in Croux and Haesbroeck (1999) that under spherical distribution $\mathbb{P}^0_X$, one has the equality

$$\text{IF}(x; V_\gamma, \mathbb{P}_X) = -(2\kappa_0)^{-1} I_{\{||x||_X \leq r(\gamma)\}} (x \otimes x) + w(||x||_X) \, I,$$

where $w$ is the function defined in Equation (7). Then affine equivariant property implies that under elliptical model given in assumption ($\mathcal{A}_2$) we have:

$$\text{IF}(x; V_\gamma, \mathbb{P}_X) = V^{1/2} \left( \frac{-1}{2\kappa_0} I_{E}(x) \otimes (V^{-1/2}x) + w(||V^{-1/2}x||_X) \, I \right) V^{1/2}$$

where

$$A = \frac{-1}{2\kappa_0} I_{E}(x) \otimes x + w(||V^{-1/2}x||_X) \, V.$$ 

Putting $V_\gamma = V_{1,\gamma}(\mathbb{P}_X) = \sigma_\gamma^2 \, V$, we have $f(V_{1,\gamma}(\mathbb{P}_X)) = \sigma_\gamma^2 \, f(V) = \sigma_\gamma^2 \, I$. Thus

$$\mathcal{T}_\gamma(\mathbb{P}_{x,x}) - \mathcal{T}_\gamma(\mathbb{P}_X) = \sigma_\gamma^{-2} \left\{ A^{-1/2}_{x,\gamma} g(V_{x,\gamma}(\mathbb{P}_{x,x})) A^{-1/2}_{x,\gamma} - g(V_\gamma) \right\}$$

$$= \sigma_\gamma^{-2} \left\{ (A^{-1/2}_{x,\gamma} - I) g(V_{x,\gamma}(\mathbb{P}_{x,x})) A^{-1/2}_{x,\gamma} + g(V_{\gamma}(\mathbb{P}_{x,x}) - V_{\gamma}(\mathbb{P}_X)) A^{-1/2}_{x,\gamma} + g(V_\gamma) (A^{-1/2}_{x,\gamma} - I) \right\}$$

where $A_{x,\gamma} = \sigma_\gamma^{-2} g(V_{x,\gamma}(\mathbb{P}_{x,x}))$. Then, using the equality

$$A^{-1/2} - I = -A^{-1}(A - I)(A^{-1/2} + I)^{-1},$$

we obtain:

$$\mathcal{T}_\gamma(\mathbb{P}_{x,x}) - \mathcal{T}_\gamma(\mathbb{P}_X) = \sigma_\gamma^{-2} \left\{ -\sigma_\gamma^{-2} A^{-1}_{x,\gamma} f(V_{x,\gamma}(\mathbb{P}_{x,x}) - \sigma_\gamma^2 V)(A^{-1/2}_{x,\gamma} + I)^{-1} g(V_\gamma) A^{-1/2}_{x,\gamma}$$

$$+ g(V_{\gamma}(\mathbb{P}_{x,x}) - \sigma_\gamma^2 V) A^{-1/2}_{x,\gamma}$$

$$- \sigma_\gamma^{-2} g(V_{\gamma}(\mathbb{P}_{x,x}) - \sigma_\gamma^2 V)(A^{-1/2}_{x,\gamma} + I)^{-1} \right\}.$$ 

Then, since $\lim_{\epsilon \to 0} A_{x,\gamma} = I$, we obtain by using the continuity of the maps $A \mapsto A^{-1}$ and $A \mapsto A^{-1/2}$:

$$\text{IF}(x; T_\gamma, \mathbb{P}_X) = \sigma_\gamma^{-2} \left\{ -\sigma_\gamma^{-2} \frac{1}{2} f(\text{IF}(x; V_\gamma, \mathbb{P}_X)) g(V_\gamma)$$

$$+ g(\text{IF}(x; V_\gamma, \mathbb{P}_X)) - \sigma_\gamma^{-1} \frac{1}{2} g(V_\gamma) f(\text{IF}(x; V_\gamma, \mathbb{P}_X)) \right\}$$

$$= \sigma_\gamma^{-2} \left\{ -\frac{1}{2} f(\text{IF}(x; V_\gamma, \mathbb{P}_X)) g(V)$$

$$+ g(\text{IF}(x; V_\gamma, \mathbb{P}_X)) - \frac{1}{2} g(V) f(\text{IF}(x; V_\gamma, \mathbb{P}_X)) \right\}.$$
A.3. Proof of Theorem 2

(i). From Lemma 3 in Croux and Dehon (2002) we obtain

\[
\text{IF}(x; T_\gamma, P_X) = \langle \beta^{(j)}, \text{IF}(x; T_\gamma, P_X) \beta^{(j)} \rangle_{X'}
\]

\[
= -\frac{\sigma^{-2}}{2K_0} E_{i}(x) \langle \beta^{(j)}, \phi(x, V) \beta^{(j)} \rangle_{X'}
\]

\[
= -\frac{\sigma^{-2}}{2K_0} E_{i}(x) \sum_{k=1}^{K} \sum_{l \neq k} < \beta^{(j)}_k, x_k > x < x_l - V_{ik} x_k, \beta^{(j)}_l > .
\]

(ii). Since \( f(\sigma^{-2} V_{\gamma}) = f(V) = 1 \), we obtain by applying the second part of Lemma 3 in Croux and Dehon (2002):
IF \( x; \beta_{\gamma}^{(j)}; P_X \) = \( \sum_{m=1}^{q} \frac{1}{\rho_j - \rho_m} < \beta^{(m)} \), IF \( x; T_{\gamma}; P_X \) \( \beta^{(j)} > x \beta^{(m)} \) 

\( - \frac{1}{2} < \beta^{(j)} \), IF \( x; f \left( \sigma_{\gamma}^{-2} V_{\gamma} \right); P_X \) \( \beta^{(j)} > x \beta^{(j)} \)

Further, IF \( x; f \left( \sigma_{\gamma}^{-2} V_{\gamma} \right); P_X \) = \( \sigma_{\gamma}^{-2} f(\text{IF}(x; V_{\gamma}; P_X)) \) and from Equation (A1) it follows

IF \( x; f \left( \sigma_{\gamma}^{-2} V_{\gamma} \right); P_X \) = \( - \frac{\sigma_{\gamma}^{-2}}{2 \kappa_0} \mathbf{1}_{E_{\gamma}}(x) f(x \otimes x) + \sigma_{\gamma}^{-2} w \left( \| V^{-1/2} x \|_{\chi} \right) f(V) \)

\( = - \frac{\sigma_{\gamma}^{-2}}{2 \kappa_0} \mathbf{1}_{E_{\gamma}}(x) \left( f(x \otimes x) - I \right) \)

\( + \sigma_{\gamma}^{-2} \left\{ \mathbf{1}_{E_{\gamma}}(x) \left( - \frac{1}{2 \kappa_0} + \kappa_1 + \kappa_2 \| V^{-1/2} x \|_{\chi}^2 + \kappa_4 \right) I \right\} \beta^{(j)} \),

where \( w \) is the function defined in Equation (7). This equality together with Equation (A4) and Theorem 1 imply

IF \( x; \beta_{\gamma}^{(j)}; P_X \) = \( - \frac{\sigma_{\gamma}^{-2}}{2 \kappa_0} \mathbf{1}_{E_{\gamma}}(x) \sum_{m=1}^{q} \frac{1}{\rho_j - \rho_m} < \beta^{(m)} \), \( \phi(x, V) \beta^{(j)} > x \beta^{(m)} \)

\( + \frac{\sigma_{\gamma}^{-2}}{4 \kappa_0} \mathbf{1}_{E_{\gamma}}(x) \left( f(x \otimes x) - I \right) \beta^{(j)} > x \beta^{(j)} \)

\( + \sigma_{\gamma}^{-2} \left\{ \left( \frac{1}{4 \kappa_0} - \frac{\kappa_1}{2} - \frac{\kappa_2}{2} \| V^{-1/2} x \|_{\chi}^2 \right) \mathbf{1}_{E_{\gamma}}(x) - \frac{\kappa_4}{2} \right\} \beta^{(j)} \).

On the other hand, since

\( z^{(j)}_\gamma (P_X) = (f(V_{\gamma}(P_X)))^{-1/2} \beta^{(j)}_\gamma (P_X) = f(V_{\gamma})^{-1/2} \beta^{(j)}_\gamma (P_X) = \sigma_{\gamma}^{-1} \beta^{(j)}_\gamma (P_X) \),

it follows

\( z^{(j)}_\gamma (P_{\epsilon, x}) - z^{(j)}_\gamma (P_X) = f(V_{\gamma}(P_{\epsilon, x}))^{-1/2} \beta^{(j)}_\gamma (P_{\epsilon, x}) - \sigma_{\gamma}^{-1} \beta^{(j)}_\gamma (P_X) \)

\( = \sigma_{\gamma}^{-1} \left\{ A_{\epsilon, \gamma}^{-1/2} \beta^{(j)}_\gamma (P_{\epsilon, x}) - \beta^{(j)}_\gamma (P_X) \right\} \)

\( = \sigma_{\gamma}^{-1} \left\{ A_{\epsilon, \gamma}^{-1/2} \left( \beta^{(j)}_\gamma (P_{\epsilon, x}) - \beta^{(j)}_\gamma (P_X) \right) + \left( A_{\epsilon, \gamma}^{-1/2} - I \right) \beta^{(j)}_\gamma (P_X) \right\} \),

where \( A_{\epsilon, \gamma} = \sigma_{\gamma}^{-2} f(V_{\gamma}(P_{\epsilon, x})) \). Then using Equation (A2), we obtain:

\( z^{(j)}_\gamma (P_{\epsilon, x}) - z^{(j)}_\gamma (P_X) = \sigma_{\gamma}^{-1} \left\{ A_{\epsilon, \gamma}^{-1/2} \left( \beta^{(j)}_\gamma (P_{\epsilon, x}) - \beta^{(j)}_\gamma (P_X) \right) \right\}

\( - A_{\epsilon, \gamma}^{-1} \left( A_{\epsilon, \gamma}^{-1/2} - I \right)^{-1} \beta^{(j)}_\gamma (P_X) \)

\( = \sigma_{\gamma}^{-1} \left\{ A_{\epsilon, \gamma}^{-1/2} \left( \beta^{(j)}_\gamma (P_{\epsilon, x}) - \beta^{(j)}_\gamma \right) \right\}

\( - \sigma_{\gamma}^{-2} A_{\epsilon, \gamma}^{-1/2} f(V_{\gamma}(P_{\epsilon, x}) - V_{\gamma}(P_X)) \left( A_{\epsilon, \gamma}^{-1/2} + I \right)^{-1} \beta^{(j)}_\gamma (P_X) \).

From the continuity of the maps \( A \mapsto A^{-1}, A \mapsto A^{-1/2} \), and the equality \( \lim_{\epsilon \to 0} A_{\epsilon, \gamma} = I \), we deduce from Equation (A6) that
Using affine equivariant property, we deduce from Eq. (A.25) in Cator and Lopuhaä (2010) that:

$$\text{IF}(x; z_j^{(j)}, \mathbb{P}_X) = \lim_{\epsilon \to 0} \frac{\alpha_j^{(j)}(\mathbb{P}_{x+\epsilon}) - \alpha_j^{(j)}(\mathbb{P}_X)}{\epsilon}$$

$$= \frac{1}{\sigma_j^2} \left\{ \text{IF}(x; \beta_j^{(j)}, \mathbb{P}_X) - \frac{\sigma_j^{-2}}{2} f(\text{IF}(x; V_j, \mathbb{P}_X)) \beta_j^{(j)} \right\}.$$

Then, the required result is obtain by using Equations (A1) and (A5).

### A.4. Proof of Theorem 3

#### A.4.1. A preliminary lemma

The following lemma gives the asymptotic distribution of the random variable

$$\hat{H}_n = \sqrt{n} \left( \hat{V}_n - \sigma^2 V \right).$$

**Lemma 1.** We assume that assumptions (\(A_1\)) to (\(A_3\)) hold. Then, \(\hat{H}_n\) converges in distribution, as \(n \to +\infty\), to a random variable \(H_x\) having a normal distribution \(N(0, \Lambda)\), where \(\Lambda\) is the covariance operator of

$$Z = \kappa_3 \ 1_{E_i}(X) \ X \otimes X + w\left( \|V^{-1/2}X\|_X \right) \ V,$$

and \(w\) is the function given in Equation (7).

**Proof.** Using affine equivariant property, we deduce from Eq. (A.25) in Cator and Lopuhaä (2010) that:

$$\hat{H}_n = V^{1/2} \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \frac{v(\|V^{-1/2}X(i)\|_X)}{\|V^{-1/2}X(i)\|_X^2} (V^{-1/2}X(i)) \otimes (V^{-1/2}X(i)) \right. \right.$$ \n
$$\left. + w\left( \|V^{-1/2}X(i)\|_X \right) \ I + o_p(1) \right) V^{1/2}$$ \n
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( \frac{v(\|V^{-1/2}X(i)\|_X)}{\|V^{-1/2}X(i)\|_X^2} (X(i) \otimes X(i) + w\left( \|V^{-1/2}X(i)\|_X \right) \ V) + o_p(1) \right)$$ \n
$$= \hat{W}_n + o_p(1),$$

where \(\hat{W}_n = n^{-1/2} \sum_{i=1}^{n} Z_i\), with

$$Z_i = \frac{v(\|V^{-1/2}X(i)\|_X)}{\|V^{-1/2}X(i)\|_X^2} (X(i) \otimes X(i) + w\left( \|V^{-1/2}X(i)\|_X \right) \ V),$$

and \(v : [0 + \infty] \to [0 + \infty]\) is the function defined by \(v(t) = \kappa_3 \ 1_{[0, r(t)]}(t) \ t^2\). Slutsky’s theorem permits to conclude that \(\hat{H}_n\) has the same limiting distribution than \(\hat{W}_n\), which can be obtained by using central limit theorem. For doing that, we will first show that \(E(Z_i) = 0\). Putting \(Y^{(i)} = V^{-1/2}X(i)\), we have

$$E(Z_i) = V^{1/2} \mathbb{E} \left( \frac{v(\|Y^{(i)}\|_X)}{\|Y^{(i)}\|_X^2} \ Y^{(i)} \otimes Y^{(i)} + w\left( \|Y^{(i)}\|_X \right) \ I \right) V^{1/2},$$

(A8)

and since \(Y^{(i)}\) has a spherical distribution, we deduce from Cator and Lopuhaä (2010) (see p. 2387) that \(E(v(\|Y^{(i)}\|_X) + q w(\|Y^{(i)}\|_X)^2) = 0, E(v(\|Y^{(i)}\|_X)^2) < +\infty, \text{ and } E(w(\|Y^{(i)}\|_X)^2) < +\infty\). Hence \(E(w(\|Y^{(i)}\|_X) = -q^{-1}E(v(\|Y^{(i)}\|_X), \text{ and Equation (A8) becomes:}$$
\[
\mathbb{E}(Z_i) = V^{1/2} \left( \mathbb{E} \left( \frac{\nu(Y^{(i)}_{\|Y\|_X})}{\|Y^{(i)}\|_X^2} \cdot Y^{(i)} \otimes Y^{(i)} \right) - \frac{1}{q} \mathbb{E} \left( \nu(\|Y^{(i)}\|_X) \right)^{1/2} \right) V^{1/2}.
\]

From the proof of Theorem 4.2 in Cator and Lopuhaä (2010) (see p. 2386) we get
\[
\mathbb{E}(\kappa_3 \mathbf{1}_{[0,r]}(\|Y^{(i)}\|_X) \cdot Y^{(i)} \otimes Y^{(i)}) = \frac{1}{q} \mathbb{E}(\kappa_3 \mathbf{1}_{[0,r]}(\|Y^{(i)}\|_X) \cdot \|Y^{(i)}\|_X^2) I
\]

\[
= \frac{1}{q} \mathbb{E}(\nu(\|Y^{(i)}\|_X)) I.
\]

Hence, \( \mathbb{E}(Z_i) = 0 \). Now, using the central limit theorem we conclude that \( \hat{W}_n \) converges in distribution, as \( n \to +\infty \), to a normal distribution \( N(0, \Lambda) \), where \( \Lambda \) is the covariance operator of
\[
Z = \mathbb{E}(V^{-1/2}X) \cdot X \otimes X + w(\|V^{-1/2}X\|_X) \cdot V
\]

(A9)

\[\text{A.4.2. Proof of the theorem}\]

Arguing as in the proof of Theorem 3.2 in Nkiet (2017) (see p. 203), we have the equality \( \sqrt{n}( \hat{\nu}_n - T) = \bar{\phi}_n(H_n) \), where \( H_n \) is given in Equation (A7) and \( \bar{\phi}_n \) is the random operator defined by
\[
\bar{\phi}_n(A) = -\sigma_0^{-1} f(\hat{V}_n)^{-1} f(A)(\sigma_0 f(\hat{V}_n)^{-1/2} + I)^{-1} g(\hat{V}_n) f(\hat{V}_n)^{-1/2}
\]

\[+ \sigma_0^{-1} g(A) f(\hat{V}_n)^{-1/2} - g(V) f(\hat{V}_n)^{-1} f(A)(\sigma_0 f(\hat{V}_n)^{-1/2} + I)^{-1}.
\]

Considering the linear map \( \varphi_\gamma \) from \( \mathcal{X} \) to itself defined as
\[
\varphi_\gamma(A) = \sigma_\gamma^{-2} \left( -\frac{1}{2} f(A) g(V) + g(A) - \frac{1}{2} g(V) f(A) \right),
\]

and denoting by \( \|\cdot\|_\infty \) and \( \|\cdot\|_\infty \) the norms \( \|A\|_{\infty} = \sup_{x \in \mathcal{X}} \|Ax\|_{\mathcal{X}} \) and \( \|Q\|_{\infty} = \sup_{B \neq 0} \|Q(B)\|_{\infty} / \|B\|_{\infty} \), we have:
\[
\|\bar{\phi}_n(H_n) - \varphi_\gamma(H_n)\|_{\infty} \leq \|\bar{\phi}_n - \varphi_\gamma\|_{\infty} \|H_n\|_{\infty}
\]

(A10)

and
\[
\|\bar{\phi}_n - \varphi_\gamma\|_{\infty} \leq \left( \sigma_\gamma^{-1} \|f\|_{\infty} \|\sigma_0 f(\hat{V}_n)^{-1/2} + I\|^{-1} g(\hat{V}_n) f(\hat{V}_n)^{-1/2} \right)
\]

\[+ \sigma_\gamma^{-1} \|f\|_{\infty} \|g(V)\|_{\infty} + \sigma_\gamma^{-1} \|g\|_{\infty}
\]

\[+ \|f\|_{\infty} \|g(V)\|_{\infty} \left( \sigma_0 f(\hat{V}_n)^{-1/2} + I \right)^{-1} \]

\[\times \|f(\hat{V}_n)^{-1/2} - \sigma_\gamma^{-1} \|_{\infty}
\]

\[+ \left( \sigma_\gamma^{-3} \|f\|_{\infty} \|g(\hat{V}_n) f(\hat{V}_n)^{-1/2} \|_{\infty}
\]

\[+ \|f\|_{\infty} \|g(V)\|_{\infty} \|\sigma_0 f(\hat{V}_n)^{-1/2} + I\|^{-1} - \frac{1}{2} \|I\|_{\infty}
\]

\[+ \sigma_\gamma^{-3} \|f\|_{\infty} \|g\|_{\infty} \|f(\hat{V}_n)^{-1/2} \|_{\infty} \|\hat{V}_n - \sigma_\gamma^2 V\|_{\infty}.
\]

(A11)
Lemma 1 implies that \( \hat{V}_n \) converges in probability to \( \sigma^2 \gamma \gamma V \), as \( n \to +\infty \). Then, using the continuity of maps \( f, g, A \mapsto A^{-1} \) and \( A \mapsto A^{-1/2} \), we deduce that \( f(\hat{V}_n) \) (resp. \( f(\hat{V}_n)^{-1} \)); resp. \( f(\hat{V}_n)^{-1/2} \); resp. \( g(\hat{V}_n) \)) converges in probability, as \( n \to +\infty \), to \( \sigma^2 \gamma \gamma \) (resp. \( \sigma^2 \gamma^{-2} \gamma \gamma \)); resp. \( \sigma^2 \gamma^{-1} \gamma \gamma \); resp. \( \sigma^2 \gamma g(V) \)). Consequently, from Equations (A10) and (A11) we deduce that \( \hat{\phi}_n(\hat{H}_n) - \phi_\gamma(\hat{H}_n) \) converge in probability to 0 as \( n \to +\infty \). Slutsky’s theorem allows to conclude that \( \hat{\phi}_n(\hat{H}_n) \) and \( \phi_\gamma(\hat{H}_n) \) both converge to the same distribution, that is the distribution of \( \phi_\gamma(H) \). Since \( \phi_\gamma \) is linear this distribution is the normal distribution with mean equal to 0 and covariance operator equal to that of the random variable:

\[
Z_\gamma = \phi_\gamma(Z) = \kappa_3 \mathbf{1}_{E_\gamma}(X) \phi_\gamma(X \otimes X) + w(\|V^{-1/2}X\|_X) \phi_\gamma(V).
\]

Besides

\[
\phi_\gamma(X \otimes X) = \sigma^{-2}_\gamma \sum_{k=1}^{K} \sum_{\ell=1}^{K} \left\{ -\frac{1}{2} (\tau_k^* (X_k \otimes X_k) V_{k\ell} \tau_\ell + \tau_\ell^* V_{k\ell} (X_k \otimes X_k) \tau_k) + \tau_k^* (X_k \otimes X_k) \tau_\ell \right\},
\]

and from \( f(V) = \mathbb{I} \), it follows:

\[
\phi_\gamma(V) = \sigma^{-2}_\gamma (g(V) - g(V)) = \left( \sigma^{-2}_\gamma - 1 \right) g(V) = \left( \sigma^{-2}_\gamma - 1 \right) \sum_{k=1}^{K} \sum_{\ell=1}^{K} \tau_k^* V_{k\ell} \tau_\ell.
\]

Thus

\[
Z_\gamma = \sigma^{-2}_\gamma \sum_{k=1}^{K} \sum_{\ell=1}^{K} \left\{ -\frac{1}{2} (\tau_k^* (X_k \otimes X_k) V_{k\ell} \tau_\ell + \tau_\ell^* V_{k\ell} (X_k \otimes X_k) \tau_k) + \tau_k^* (X_k \otimes X_k) \tau_\ell \right\} + \left( \sigma^{-2}_\gamma - 1 \right) w(\|V^{-1/2}X\|_X) \sum_{k=1}^{K} \sum_{\ell=1}^{K} \tau_k^* V_{k\ell} \tau_\ell.
\]

**A.5. Proof of Corollary 1**

We deduce from Theorem 3 and Proposition 4 in Dossou-Gbete and Pousse (1991) that \( \sqrt{n}(\hat{\rho}_{j,n} - \rho_j) \) converges in distribution, as \( n \to +\infty \), to \( \Delta(\Pi_j U_j, \Pi_j) \), where \( \Pi_j \) is the orthogonal projector onto the eigenspace associated with \( \rho_j \) and \( \Delta \) is the continuous map which associates to each self-adjoint operator \( A \) the vector \( \Delta(A) \) of its eigenvalues in non increasing order. Clearly, \( \Pi_j = \beta^{(j)} \otimes \beta^{(j)} \) and, therefore, using the properties of tensor product (see Dauxois, Romain, and Viguier 1994):

\[
\Pi_j U_j, \Pi_j = (\beta^{(j)} \otimes \beta^{(j)}) U_j (\beta^{(j)} \otimes \beta^{(j)}) = (\beta^{(j)} \otimes \beta^{(j)}) (\beta^{(j)} \otimes (U_j \beta^{(j)}))
= < \beta^{(j)}, W \beta^{(j)} > X \beta^{(j)} \otimes \beta^{(j)}.
\]

Hence, \( \Delta(\Pi_j U_j, \Pi_j) = < \beta^{(j)}, U_j \beta^{(j)} > X \). Since the map \( A \mapsto < \beta^{(j)}, A \beta^{(j)} >_X \) is linear, \( \Delta(\Pi_j U_j, \Pi_j) \) has the normal distribution \( \mathcal{N}(0, \nu_j^2) \), where \( \nu_j^2 = \mathbb{E}(< \beta^{(j)}, U_j \beta^{(j)} >_X^2) \). However, \( U_j \) has the same covariance operator than \( Z_\gamma \). Thus \( \nu_j^2 = \mathbb{E}(< \beta^{(j)}, Z_\gamma \beta^{(j)} >_X^2) = \text{Var}(< \beta^{(j)}, Z_\gamma \beta^{(j)} >_X) \). Using the equalities
< β′(j), τk′(Xk ⊗ Xk)Vkτβ′(j)>X = < β′(j), Xk>_k < Xk, V_kβ′(j)>_k,
< β′(j), τk′V_k(Xk ⊗ Xk)τkβ′(j)>X = < β′(j), Xk>_k < Xk, V_kβ′(j)>_k,
< β′(j), τk′(Xk ⊗ Xk)τkβ′(j)>X = < β′(j), Xk>_k < β′(j), X_k>_k,

and < β′(j), τk′V_kτβ′(j)>X = < β′(j), V_kβ′(j)>_k, we get

< β′(j), Z_jβ′(j)>X = Σ_k Σ_i

+ (σ_−2 1(E) X (k_1 + k_2 ||V^{−1/2}X||_X^2) + k_4) < β′(j), V_kβ′(j)>_k.

A.6. Proof of Theorem 4

Under H_0 we have T = 0 and we deduce from Theorem 3 that √nT_n converges in distribution, as n → +∞, to U_j. Since the map A → Σ_k Σ_i tr(π_k(A)π_i(A)^*) is continuous, we deduce that nS_n converges in distribution, as n → +∞, to

Q_j = Σ_k tr(π_k(π_i(U_j)π_i(U_j)^*)).

By a similar reasoning to that of the proof of Theorem 4.1 in Nkiet (2017) we obtain that ..

where W_j is a random variable having normal distribution N(0, Θ) in R^4 with

Θ =

where Θ_k, rs is the p_k, p_r × p_p, matrix given by

and

i_{ijp} = < E(π_k(U_j) ⊗ π_r(U_j)) (C_j ⊗ C_r), C_i ⊗ C_p >

= < E(π_k(Z_j) ⊗ π_r(Z_j)) (C_j ⊗ C_r), C_i ⊗ C_p >
where $\otimes$ denotes the tensor product related to the inner product of operators given by $\langle A, B \rangle = \text{tr}(AB^\star)$ and $\{e_i^{(k)}\}_{1 \leq i \leq p_k}$ is an orthonormal basis of $\mathcal{X}_k$. Since under $\mathcal{H}_0, Z_\gamma = \sigma_{\gamma}^{-2} \kappa_3 1_{E_\gamma}(X) \sum_{k=1}^K \sum_{m=1}^K \tau_k \tau_j(X_m \otimes X_j) \tau_m \tau_j$, we obtain

$$
\pi_{kl}(Z_\gamma) = \sigma_{\gamma}^{-2} \kappa_3 1_{E_\gamma}(X) \sum_{j=1}^K \sum_{m=1}^K \tau_k \tau_j(X_m \otimes X_j) \tau_m \tau_j
$$

$$
= \sigma_{\gamma}^{-2} \kappa_3 1_{E_\gamma}(X) \sum_{j=1}^K \sum_{m=1}^K \delta_{kj} \delta_{tm}(X_m \otimes X_j)
$$

$$
= \sigma_{\gamma}^{-2} \kappa_3 1_{E_\gamma}(X)(X_t \otimes X_k).
$$

Hence

$$
\gamma_{ijpt}^{kl} = \sigma_{\gamma}^{-4} \kappa_3^2 \mathbb{E}\left(1_{E_\gamma}(X) < \left((X_t \otimes X_k) \otimes ((X_s \otimes X_r)) \right) \left(e_{ijpq}^{(k)} \otimes e_{ijpq}^{(r)} \right) > \right)
$$

$$
= \sigma_{\gamma}^{-4} \kappa_3^2 \mathbb{E}\left(1_{E_\gamma}(X) < X_t \otimes X_k, e_{ijpq}^{(k)} > X_s \otimes X_r, e_{ijpq}^{(r)} > \right)
$$

$$
= \sigma_{\gamma}^{-4} \kappa_3^2 \mathbb{E}\left(1_{E_\gamma}(X) < X_t, k \neq k, X_r, e_{ijpq}^{(r)} > r \right)
$$

Note that under $\mathcal{H}_0$ we have $V = \mathbb{I}$, then $X$ has a spherical distribution with density $f_X(x) = h(||x||_X^2)$. Therefore, if $(k, \ell) = (r, u)$ and $(i, j) = (p, t)$ with $\ell \neq k$ and $u \neq r$, then $\gamma_{ijpq}^{kl}$ equals an integral of the form

$$
\int z(||x||_X^2) x_a^2 x_b^2 dx,
$$

with $a \neq b$, where $z$ is a suitable function from $[0, +\infty[$ to itself. Then from Lemma 1 in Lopuhaä (1997), we deduce that

$$
\int z(||x||_X^2) x_a^2 x_b^2 dx = \frac{1}{q(q+1)} \int z(||x||_X^2) ||x||_X^4 dx
$$

and, therefore, that

$$
\gamma_{ijpq}^{kl} = \frac{\sigma_{\gamma}^{-4} \kappa_3^2}{q(q+1)} \mathbb{E}\left(1_{E_\gamma}(X)||X||_X^4 \right).
$$

Otherwise, if one of the conditions $(k, \ell) = (r, u), (i, j) = (p, t), \ell \neq k, u \neq r$ does not hold then $\gamma_{ijpq}^{kl}$ equals an integral of the form

$$
\int z(||x||_X^2) x_a x_b x_c x_d dx
$$

in which at least two of the indices $a, b, c, d$ are different. From elementary calculus obtained by changing to spherical coordinates we obtain that this integral equals 0 and, therefore, $\gamma_{ijpq}^{kl} = 0$. We deduce that

$$
\Theta = \frac{\sigma_{\gamma}^{-4} \kappa_3^2}{q(q+1)} \mathbb{E}\left(1_{E_\gamma}(X)||X||_X^4 \right) I_d
$$

where $I_d$ is the $d \times d$ identity matrix. Thus, $Q_{\gamma} = \frac{\sigma_{\gamma}^{-4} \kappa_3^2}{q(q+1)} \mathbb{E}\left(1_{E_\gamma}(X)||X||_X^4 \right) Q'$ where $Q'$ is a random variable with distribution equal to $\gamma_{ijpq}^2$. 