



HAL
open science

Asymptotic theory of multiple-set linear canonical analysis

Guy Martial Nkiet

► **To cite this version:**

| Guy Martial Nkiet. Asymptotic theory of multiple-set linear canonical analysis. 2017. <hal-01511413>

HAL Id: hal-01511413

<https://hal.science/hal-01511413v1>

Preprint submitted on 20 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Asymptotic theory of multiple-set linear canonical analysis

Guy Martial Nkiet

Received: date / Revised: date

Abstract This paper deals with asymptotics for multiple-set linear canonical analysis (MSLCA). A definition of this analysis, that adapts the classical one to the context of Euclidean random variables, is given and properties of the related canonical coefficients are derived. Then, estimators of the MSLCA's elements, based on empirical covariance operators, are proposed and asymptotics for these estimators are obtained. More precisely, we prove their consistency and we obtain asymptotic normality for the estimator of the operator that gives MSLCA, and also for the estimator of the vector of canonical coefficients. These results are then used to obtain a test for mutual non-correlation between the involved Euclidean random variables.

Keywords Multiple set canonical analysis · asymptotic study · non-correlation tests

1 Introduction

Multiple-set linear canonical analysis (MSLCA), also known as generalized canonical correlation analysis, has been extensively discussed in the literature, see Kettenring (1971), Gifi (1991), Gardner et al. (2006), Takane et al. (2008), Tenenhaus and Tenenhaus (2011), as well as the further references contained therein. It is a statistical method that generalizes linear canonical analysis (LCA) to the case where more than two sets of variables are considered, which is of a real interest since in applied statistical studies it is common to collect data from the observation of several sets of variables on a given population. However, although this interest, several aspects under which LCA has been studied have never been addressed to MSLCA. For example, asymptotic theory for LCA and related applications have been tackled by several authors

Guy Martial Nkiet

Université des Sciences et Techniques de Masuku, Département de Mathématiques et Informatique, BP 943 Franceville, Gabon. E-mail: gnkiet@hotmail.com

(e.g., Muirhead and Waternaux (1980), Anderson (1999), Pousse (1992), Fine (2000), Dauxois et al. (2004)). It would be natural to wonder how the obtained results extend to the case of MSLCA but, to the best of our knowledge, such an approach has never been tackled.

In this paper, we introduce an asymptotic theory for MSLCA. For doing that, we first define in Section 2 the notion of MSLCA for Euclidean random variables, that is random variables valued into Euclidean vector spaces. This analysis is defined from a maximization problem under specified constraints, and shown to be obtained from spectral analysis of a suitable operator. Properties of the related eigenvalues, called canonical coefficients, are then given. In Section 3, we tackle the problem of estimating MSLCA. More precisely, estimators based on empirical covariance operators are introduced. Then, consistency of the obtained estimators is proved. Further, we derive the asymptotic distribution of the used estimator of the aforementioned operator, and also that of the estimator of the vector of canonical coefficients in the general case as well as in the case of elliptical distribution. Section 4 is devoted to the introduction of a test for mutual non-correlation between the random variables involved in MSLCA. The results obtained for asymptotic theory of MSLCA are then used in order to derive the asymptotic distribution of the used test statistic under null hypothesis.

2 Multiple-set canonical linear analysis of Euclidean random variables

For an integer $K \geq 2$, let us consider random variables X_1, \dots, X_K defined on a probability space (Ω, \mathcal{A}, P) and valued into Euclidean vector spaces $\mathcal{X}_1, \dots, \mathcal{X}_K$ respectively. Denoting by \mathbb{E} the mathematical expectation related to P , we assume that, for any $k \in \{1, \dots, K\}$, we have $\mathbb{E}(\|X_k\|_k^2) < +\infty$ where $\|\cdot\|_k$ denotes the norm induced by the inner product $\langle \cdot, \cdot \rangle_k$ of \mathcal{X}_k , and, without loss of generality, that $\mathbb{E}(X_k) = 0$. Each vector α in the vector space $\mathcal{X} := \mathcal{X}_1 \times \dots \times \mathcal{X}_K$ will be written as

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_K \end{pmatrix},$$

and we recall that \mathcal{X} is an Euclidean vector space equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ defined by:

$$\forall \alpha \in \mathcal{X}, \forall \beta \in \mathcal{X}, \langle \alpha, \beta \rangle_{\mathcal{X}} = \sum_{k=1}^K \langle \alpha_k, \beta_k \rangle_k.$$

We denote by $\|\cdot\|_{\mathcal{X}}$ the norm induced by this inner product. Considering the \mathcal{X} -valued random variable

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix},$$

we can give the following definition which adapts the classical definition of multiple-set canonical analysis (e.g., Gifi (1991), Gardner et al. (2006), Takane et al. (2008)) to the context of Euclidean random variables.

Definition 2.1. *The multiple-set linear canonical analysis (MSLCA) of X is the search of a sequence $(\alpha^{(j)})_{1 \leq j \leq q}$ of vectors of E , where $q = \dim(\mathcal{X})$, satisfying:*

$$\alpha^{(j)} = \arg \max_{\alpha \in C_j} \mathbb{E} \left(\langle \alpha, X \rangle_{\mathcal{X}}^2 \right), \quad (1)$$

where

$$C_1 = \left\{ \alpha \in \mathcal{X} / \sum_{k=1}^K \text{var}(\langle \alpha_k, X_k \rangle_k) = 1 \right\}, \quad (2)$$

and, for $j \geq 2$:

$$C_j = \left\{ \alpha \in C_1 / \sum_{k=1}^K \text{cov} \left(\langle \alpha_k^{(r)}, X_k \rangle_k, \langle \alpha_k, X_k \rangle_k \right) = 0, \forall r \in \{1, \dots, j-1\} \right\}. \quad (3)$$

Remark 2.1

1) The constraints sets given in (2) and (3) can be expressed by using covariance operators defined for $(k, \ell) \in \{1, \dots, K\}^2$ by:

$$V_{k\ell} = \mathbb{E} (X_\ell \otimes X_k) = V_{\ell k}^* \text{ and } V_k := V_{kk},$$

where \otimes denotes the tensor product such that, for any (x, y) , $x \otimes y$ is the linear map : $h \mapsto \langle x, h \rangle y$, and T^* denotes the adjoint of T . Indeed, it is easily seen that, for $(\alpha, \beta) \in \mathcal{X}^2$,

$$\text{var}(\langle \alpha_k, X_k \rangle_k) = \mathbb{E}(\langle \alpha_k, X_k \rangle_k^2) = \mathbb{E}(\langle \alpha_k, (X_k \otimes X_k)(\alpha_k) \rangle_k) = \langle \alpha_k, V_k \alpha_k \rangle_k,$$

and

$$\begin{aligned} \text{cov}(\langle \alpha_k, X_k \rangle_k, \langle \beta_\ell, X_\ell \rangle_\ell) &= \mathbb{E}(\langle \alpha_k, X_k \rangle_k \langle \beta_\ell, X_\ell \rangle_\ell) = \mathbb{E}(\langle \alpha_k, (X_\ell \otimes X_k)(\beta_\ell) \rangle_k) \\ &= \langle \alpha_k, V_{k\ell} \beta_\ell \rangle_k. \end{aligned}$$

Therefore,

$$C_1 = \left\{ \alpha \in \mathcal{X} / \sum_{k=1}^K \langle \alpha_k, V_k \alpha_k \rangle_k = 1 \right\}, \quad (4)$$

and

$$C_j = \left\{ \alpha \in C_1 / \sum_{k=1}^K \langle \alpha_k^{(r)}, V_k \alpha_k \rangle_k = 0, \forall r \in \{1, \dots, j-1\} \right\}. \quad (5)$$

2) For any $\alpha \in C_1$, one has:

$$\begin{aligned} \mathbb{E}(\langle \alpha, X \rangle_{\mathcal{X}}^2) &= \mathbb{E} \left(\left(\sum_{k=1}^K \langle \alpha_k, X_k \rangle_k \right)^2 \right) = \sum_{k=1}^K \sum_{\ell=1}^K \mathbb{E}(\langle \alpha_k, X_k \rangle_k \langle \alpha_\ell, X_\ell \rangle_\ell) \\ &= \sum_{k=1}^K \mathbb{E}(\langle \alpha_k, X_k \rangle_k^2) + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \mathbb{E}(\langle \alpha_k, X_k \rangle_k \langle \alpha_\ell, X_\ell \rangle_\ell) \\ &= \sum_{k=1}^K \text{var}(\langle \alpha_k, X_k \rangle_k) + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \alpha_k, V_{k\ell} \alpha_\ell \rangle_k \\ &= 1 + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \alpha_k, V_{k\ell} \alpha_\ell \rangle_k = 1 + \varphi(\alpha), \end{aligned}$$

where

$$\varphi(\alpha) = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \alpha_k, V_{k\ell} \alpha_\ell \rangle_k. \quad (6)$$

Then, the MSLCA of X is obtained by minimizing $\varphi(\alpha)$ under the constraints expressed in (4) and (5).

For $k \in \{1, \dots, K\}$, the covariance operator V_k is a self-adjoint non-negative operator. From now on, we assume that it is invertible. Let τ_k be the canonical projection defined as

$$\tau_k : \alpha \in \mathcal{X} \mapsto \alpha_k \in \mathcal{X}_k;$$

its adjoint τ_k^* of τ_k is the map given by:

$$\tau_k^* : t \in \mathcal{X}_k \mapsto \underbrace{(0, \dots, 0, t, 0, \dots, 0)}_{k-1 \text{ times}}^T \in \mathcal{X},$$

where we denote by a^T the transposed of a . Now, let us consider the operators of $\mathcal{L}(\mathcal{X})$ given by:

$$\Phi = \sum_{k=1}^K \tau_k^* V_k \tau_k \quad \text{and} \quad \Psi = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* V_{k\ell} \tau_\ell.$$

From the fact that $\tau_k \tau_\ell^* = \delta_{k\ell} I_k$, where $\delta_{k\ell}$ is the usual Kronecker symbol and I_k is the identity operator of \mathcal{X}_k , it is easily seen that Φ is also an invertible self-adjoint and non-negative operator, with $\Phi^{-1} = \sum_{k=1}^K \tau_k^* V_k^{-1} \tau_k$ and $\Phi^{-1/2} =$

$\sum_{k=1}^K \tau_k^* V_k^{-1/2} \tau_k$. The following theorem shows how to obtain a MSLCA of X . It just repeats a known result (e.g., Gifi (1991), Takane et al. (2008)) within the framework used for this paper.

Theorem 2.1. *Letting $\{\beta^{(1)}, \dots, \beta^{(q)}\}$ be an orthonormal basis of \mathcal{X} such that $\beta^{(j)}$ is an eigenvector of the operator $T = \Phi^{-1/2} \Psi \Phi^{-1/2}$ associated with the j -th largest eigenvalue ρ_j of T . Then, the sequence $(\alpha^{(j)})_{1 \leq j \leq q}$ given by:*

$$\alpha^{(j)} = \Phi^{-1/2} \beta^{(j)} = \left(V_1^{-1/2} \beta_1^{(j)}, \dots, V_K^{-1/2} \beta_K^{(j)} \right),$$

consists of solutions of (1) under the constraints (2) and (3), and we have: $\rho_j = \langle \beta^{(j)}, T \beta^{(j)} \rangle_E = \varphi(\alpha^{(j)})$.

Proof. Putting $\beta_k = V_k^{1/2} \alpha_k$ and $\beta^{(r)} = V_k^{1/2} \alpha_k^{(r)}$, we have:

$$\begin{aligned} \varphi(\alpha) &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle V_k^{-1/2} \beta_k, V_{k\ell} V_\ell^{-1/2} \beta_\ell \rangle_k \\ &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \beta_k, V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \beta_\ell \rangle_k =: \psi(\beta), \end{aligned} \quad (7)$$

where

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_K \end{pmatrix} \in \mathcal{X}.$$

Since $V_k = V_k^{1/2} V_k^{1/2}$, having $\alpha \in C_j$ is equivalent to having $\beta \in C'_j$, where:

$$C'_1 = \left\{ \beta \in \mathcal{X} / \sum_{k=1}^K \|\beta_k\|_k^2 = 1 \right\} = \{ \beta \in \mathcal{X} / \|\beta\|_{\mathcal{X}}^2 = 1 \}, \quad (8)$$

and for $j \geq 2$:

$$\begin{aligned} C'_j &= \left\{ \beta \in C_1 / \sum_{k=1}^K \langle \beta_k^{(r)}, \beta_k \rangle_k = 0, \forall r \in \{1, \dots, j-1\} \right\} \\ &= \left\{ \beta \in C_1 / \langle \beta^{(r)}, \beta \rangle_{\mathcal{X}} = 0, \forall r \in \{1, \dots, j-1\} \right\}. \end{aligned} \quad (9)$$

Further, for any $\beta \in \mathcal{X}$:

$$\begin{aligned} \Psi \Phi^{-1/2} \beta &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{j=1}^K \tau_k^* V_{k\ell} \tau_\ell \tau_j^* V_j^{-1/2} \tau_j \beta = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{j=1}^K \delta_{\ell j} \tau_k^* V_{k\ell} V_j^{-1/2} \tau_j \beta \\ &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* V_{k\ell} V_\ell^{-1/2} \tau_\ell \beta, \end{aligned}$$

and

$$\begin{aligned}
\Phi^{-1/2}\Psi\Phi^{-1/2}\beta &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{j=1}^K \tau_j^* V_j^{-1/2} \tau_j \tau_k^* V_{k\ell} V_\ell^{-1/2} \tau_\ell \beta \\
&= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{j=1}^K \delta_{jk} \tau_j^* V_j^{-1/2} V_{k\ell} V_\ell^{-1/2} \tau_\ell \beta \\
&= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \tau_\ell \beta \\
&= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \beta_\ell.
\end{aligned}$$

Thus,

$$\begin{aligned}
\langle \beta, \Phi^{-1/2}\Psi\Phi^{-1/2}\beta \rangle_{\mathcal{X}} &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \beta, \tau_k^* V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \beta_\ell \rangle_{\mathcal{X}} \\
&= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \tau_k \beta, V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \beta_\ell \rangle_k \\
&= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \langle \beta_k, V_k^{-1/2} V_{k\ell} V_\ell^{-1/2} \beta_\ell \rangle_k = \psi(\beta),
\end{aligned}$$

where ψ is defined in (7). Then, the MSLCA optimization problem reduces to the maximization of $\langle \beta, \Phi^{-1/2}\Psi\Phi^{-1/2}\beta \rangle_{\mathcal{X}}$ under the constraints (8) and (9). Since $T = \Phi^{-1/2}\Psi\Phi^{-1/2}$ is a self-adjoint operator, this is a well known maximization problem for which a solution is obtained from the spectral analysis of T as stated in the theorem. \square

Definition 2.2. The ρ_j 's are termed the canonical coefficients. The $\alpha^{(j)}$'s are termed vectors of canonical directions.

The following theorem gives some properties of the canonical coefficients.

Theorem 2.2.

- (i) $\forall j \in \{1, \dots, q\}$, $-1 \leq \rho_j \leq K(K-1)$.
- (ii) $\forall j \in \{1, \dots, q\}$, $\rho_j = 0 \Leftrightarrow \forall (k, \ell) \in \{1, \dots, K\}^2$, $k \neq \ell$, $V_{k\ell} = 0$.

Proof.

(i) First, using (6), we have for any $j \in \{1, \dots, q\}$,

$$\rho_j = \varphi(\alpha^{(j)}) = \mathbb{E} \left(\langle \alpha^{(j)}, X \rangle_x^2 \right) - 1 \geq -1.$$

On the other hand, we have:

$$\begin{aligned} \rho_j = \varphi(\alpha^{(j)}) &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \mathbb{E} \left(\langle \alpha_k^{(j)}, X_k \rangle_k \langle \alpha_\ell^{(j)}, X_\ell \rangle_\ell \right) \\ &\leq \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sqrt{\mathbb{E} \left(\langle \alpha_k^{(j)}, X_k \rangle_k^2 \right)} \sqrt{\mathbb{E} \left(\langle \alpha_\ell^{(j)}, X_\ell \rangle_\ell^2 \right)}. \end{aligned}$$

Since, for any $k \in \{1, \dots, K\}$, one has:

$$\mathbb{E} \left(\langle \alpha_k^{(j)}, X_k \rangle_k^2 \right) = \text{var} \left(\langle \alpha_k^{(j)}, X_k \rangle_k \right) \leq \sum_{\ell=1}^K \text{var} \left(\langle \alpha_\ell^{(j)}, X_\ell \rangle_\ell \right) = 1,$$

it follows that:

$$\rho_j \leq \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K 1 = K(K-1).$$

(ii) Since the ρ_j 's are the eigenvalues of T , we have:

$$\forall j \in \{1, \dots, q\}, \rho_j = 0 \Leftrightarrow T = 0 \Leftrightarrow \Psi = 0 \Leftrightarrow \forall (k, \ell) \in \{1, \dots, K\}^2, k \neq \ell, V_{k\ell} = 0.$$

□

Remark 2.2.

1) When $K = 2$, one has $\Phi = \tau_1^* V_1 \tau_1 + \tau_2^* V_2 \tau_2$ and $\Psi = \tau_1^* V_{12} \tau_2 + \tau_2^* V_{21} \tau_1$. Then it is easy to check that $T = \tau_1^* S \tau_2 + \tau_2^* S^* \tau_1$, where $S = V_1^{-1/2} V_{12} V_2^{-1/2}$. Let x be an eigenvector of T associated with an eigenvalue $\rho \neq 0$. We have $Tx = \rho x$, that is equivalent to having:

$$\tau_1^* (S \tau_2 x - \rho \tau_1 x) = -\tau_2^* (S^* \tau_1 x - \rho \tau_2 x).$$

This implies:

$$\begin{cases} S \tau_2 x = \rho \tau_1 x \\ S^* \tau_1 x = \rho \tau_2 x \end{cases}$$

and, putting $x_1 = \tau_1 x$ and $x_2 = \tau_2 x$, we obtain

$$x_2 = \rho^{-1} S^* x_1 \quad \text{and} \quad R x_1 = \rho^2 x_1, \tag{10}$$

where

$$R = S S^* = V_1^{-1/2} V_{12} V_2^{-1} V_{21} V_2^{-1/2}.$$

Conversely, if (10) holds then, putting $x = \tau_1^* x_1 + \tau_2^* x_2$, we have:

$$\begin{aligned} Tx &= \tau_1^* S \tau_2 x + \tau_2^* S^* \tau_1 x = \tau_1^* S x_2 + \tau_2^* S^* x_1 = \rho^{-1} \tau_1^* S S^* x_1 + \rho \tau_2^* x_2 \\ &= \rho^{-1} \tau_1^* R x_1 + \rho \tau_2^* x_2 = \rho (\tau_1^* x_1 + \tau_2^* x_2) = \rho x. \end{aligned}$$

Moreover, since

$$\|x_2\|_2 = \rho^{-1} \|S^* x_1\|_1 = \rho^{-1} \sqrt{\langle S^* x_1, S^* x_1 \rangle_2} = \rho^{-1} \sqrt{\langle S S^* x_1, x_1 \rangle_1} = \|x_1\|_1$$

and

$$\|x\|_{\mathcal{X}}^2 = \|x_1\|_1^2 + \|x_2\|_2^2$$

it follows that

$$\|x_1\|_1 = \|x_2\|_2 = \frac{1}{\sqrt{2}} \|x\|_{\mathcal{X}}.$$

2) The preceding remark shows the equivalence between MSLCA and linear canonical analysis (LCA) when $K = 2$. Recall that LCA of X_1 and X_2 is obtained from the spectral analysis of R (see, e.g., Dauxois and Pouse (1975), Pouse (1992), Fine (2000)). More precisely, $\{\beta^{(j)}, \rho_j\}_{1 \leq j \leq q}$ is defined as in Theorem 2.1 if, and only if, $\{u_1^{(j)}, u_2^{(j)}, \rho_j^2\}_{1 \leq j \leq q}$, where $u_\ell^{(j)} = \frac{1}{\sqrt{2}} \tau_\ell \beta^{(j)}$ ($\ell \in \{1, 2\}$), is a LCA of X_1 and X_2 .

3 Estimation and asymptotic theory

In this section, we deal with estimation of MSLCA. For $k = 1, \dots, K$, let $\{X_k^{(i)}\}_{1 \leq i \leq n}$ be an i.i.d. sample of X_k . We use empirical covariance operators for defining estimators of MSLCA elements. Then, consistency and asymptotic normality are obtained for the resulting estimators of the vectors of canonical directions and the canonical coefficients.

3.1 Estimation and almost sure convergence

For $(k, \ell) \in \{1, \dots, K\}^2$, let us consider the sample means and covariance operators:

$$\bar{X}_{k \cdot n} = \frac{1}{n} \sum_{i=1}^n X_k^{(i)}, \quad \widehat{V}_{k\ell \cdot n} = \frac{1}{n} \sum_{i=1}^n (X_\ell^{(i)} - \bar{X}_{\ell \cdot n}) \otimes (X_k^{(i)} - \bar{X}_{k \cdot n}), \quad \widehat{V}_{k \cdot n} := \widehat{V}_{kk \cdot n},$$

and the random operators valued into $\mathcal{L}(\mathcal{X})$ defined as

$$\widehat{\Phi}_n = \sum_{k=1}^K \tau_k^* \widehat{V}_{k \cdot n} \tau_k \quad \text{and} \quad \widehat{\Psi}_n = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* \widehat{V}_{k\ell \cdot n} \tau_\ell.$$

Then, we estimate T by

$$\widehat{T}_n = \widehat{\Phi}_n^{-1/2} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2}.$$

Considering the eigenvalues $\widehat{\rho}_{1 \cdot n} \geq \widehat{\rho}_{2 \cdot n} \cdots \geq \widehat{\rho}_{q \cdot n}$ of \widehat{T}_n , and $\{\widehat{\beta}_n^{(1)}, \dots, \widehat{\beta}_n^{(q)}\}$ an orthonormal basis of \mathcal{X} such that $\widehat{\beta}_n^{(j)}$ is an eigenvector of \widehat{T}_n associated with $\widehat{\rho}_{j \cdot n}$. Then, we estimate ρ_j by $\widehat{\rho}_{j \cdot n}$, and $\beta^{(j)}$ by $\widehat{\beta}_n^{(j)}$. The following theorem establishes strong consistency for these estimators.

Theorem 3.1. For any integer $j \in \{1, \dots, q\}$:

(i) $\widehat{\rho}_{j \cdot n}$ converge almost surely, as $n \rightarrow +\infty$, to ρ_j .

(ii) $\text{sign}(\langle \widehat{\beta}_n^{(j)}, \beta^{(j)} \rangle_{\mathcal{X}}) \widehat{\beta}_n^{(j)}$ converges almost surely, as $n \rightarrow +\infty$, to $\beta^{(j)}$ in \mathcal{X} .

Proof. From obvious applications of the strong law of large numbers, it is easily seen that \widehat{T}_n converges almost surely in $\mathcal{L}(\mathcal{X})$, as $n \rightarrow +\infty$ to T . Then using Lemma 1 in Ferré and Yao (2003), we obtain the inequality $|\widehat{\rho}_{j \cdot n} - \rho_j| \leq \|\widehat{T}_n - T\|$ from what (i) is deduced. Clearly, each $\beta^{(j)} \otimes \beta^{(j)}$ and is a projector onto an eigenspace. Therefore, using Proposition 3 in Dossou-Gbete and Pousse (1991), we deduce that $\widehat{\beta}_n^{(j)} \otimes \widehat{\beta}_n^{(j)}$ converges almost surely in $\mathcal{L}(\mathcal{X})$ to $\beta^{(j)} \otimes \beta^{(j)}$, as $n \rightarrow +\infty$. Using again Lemma 1 in Ferré and Yao (2003), we obtain the inequality

$$\left\| \text{sign}(\langle \widehat{\beta}_n^{(j)}, \beta^{(j)} \rangle_{\mathcal{X}}) \widehat{\beta}_n^{(j)} - \beta^{(j)} \right\|_{\mathcal{X}} \leq 2\sqrt{2} \left\| \widehat{\beta}_n^{(j)} \otimes \widehat{\beta}_n^{(j)} - \beta^{(j)} \otimes \beta^{(j)} \right\|$$

from what we deduce (ii). \square

3.2 Asymptotic distribution

In this section, we assume that, for $k \in \{1, \dots, K\}$, we have $\mathbb{E}(\|X_k\|_k^4) < +\infty$ and $V_k = I_k$, where I_k denotes the identity operator of \mathcal{X}_k . We first derive an asymptotic distribution for \widehat{T}_n , then we obtain these of the canonical coefficients.

Theorem 3.2. $\sqrt{n}(\widehat{T}_n - T)$ converges in distribution, as $n \rightarrow +\infty$, to a random variable U having a normal distribution in $\mathcal{L}(\mathcal{X})$, with mean 0 and covariance operator Γ equal to that of the random operator:

$$Z = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K -\frac{1}{2} (\tau_k^*(X_k \otimes X_k) V_{k\ell} \tau_\ell + \tau_\ell^* V_{\ell k} (X_k \otimes X_k) \tau_k) + \tau_k^*(X_\ell \otimes X_k) \tau_\ell.$$

Proof. Under the above assumptions,

$$\Phi = \sum_{k=1}^K \tau_k^* V_k \tau_k = \sum_{k=1}^K \tau_k^* \tau_k = I_{\mathcal{X}},$$

where $I_{\mathcal{X}}$ is the identity operator of \mathcal{X} , and

$$\begin{aligned} \sqrt{n}(\widehat{T}_n - T) &= \sqrt{n}(\widehat{\Phi}_n^{-1/2}\widehat{\Psi}_n\widehat{\Phi}_n^{-1/2} - \Psi) \\ &= \sqrt{n}(\widehat{\Phi}_n^{-1/2} - I_{\mathcal{X}})\widehat{\Psi}_n\widehat{\Phi}_n^{-1/2} + \sqrt{n}(\widehat{\Psi}_n - \Psi)\widehat{\Phi}_n^{-1/2} + \Psi(\sqrt{n}(\widehat{\Phi}_n^{-1/2} - I_{\mathcal{X}})) \\ &= -\widehat{\Phi}_n^{-1}(\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}))(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1}\widehat{\Psi}_n\widehat{\Phi}_n^{-1/2} + \sqrt{n}(\widehat{\Psi}_n - \Psi)\widehat{\Phi}_n^{-1/2} \\ &\quad - \Psi\widehat{\Phi}_n^{-1}(\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}))(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1}. \end{aligned} \quad (11)$$

Clearly,

$$V_{kl} = \mathbb{E}(\tau_{\ell}(X) \otimes \tau_k(X)) = \tau_k V \tau_{\ell}^*, \quad (12)$$

where $V = \mathbb{E}(X \otimes X)$. Moreover, putting

$$X^{(i)} = \begin{pmatrix} X_1^{(i)} \\ \vdots \\ X_K^{(i)} \end{pmatrix},$$

we have

$$\begin{aligned} \widehat{V}_{kl,n} &= \frac{1}{n} \sum_{i=1}^n X_{\ell}^{(i)} \otimes X_k^{(i)} - \overline{X}_{\ell,n} \otimes \overline{X}_{k,n} \\ &= \frac{1}{n} \sum_{i=1}^n \tau_{\ell}(X^{(i)}) \otimes \tau_k(X^{(i)}) - \tau_{\ell}(\overline{X}_n) \otimes \tau_k(\overline{X}_n) \\ &= \tau_k \widehat{V}_n \tau_{\ell}^*, \end{aligned} \quad (13)$$

where $\overline{X}_n = n^{-1} \sum_{i=1}^n X^{(i)}$ and

$$\widehat{V}_n = \frac{1}{n} \sum_{i=1}^n X^{(i)} \otimes X^{(i)} - \overline{X}_n \otimes \overline{X}_n. \quad (14)$$

Therefore, using (12) and (13), we obtain

$$\sqrt{n}(\widehat{\Psi}_n - \Psi) = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* \tau_k \widehat{H}_n \tau_{\ell}^* \tau_{\ell} = f(\widehat{H}_n), \quad (15)$$

where $\widehat{H}_n = \sqrt{n}(\widehat{V}_n - V)$ and f is the operator defined as

$$f : A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* \tau_k A \tau_{\ell}^* \tau_{\ell} \in \mathcal{L}(\mathcal{X}).$$

Further, since $I_{\mathcal{X}} = \sum_{k=1}^K \tau_k^* \tau_k$, we obtain

$$\sqrt{n}(\widehat{\Phi}_n - I_{\mathcal{X}}) = \sum_{k=1}^K \tau_k^* \tau_k \widehat{H}_n \tau_k^* \tau_k = g(\widehat{H}_n), \quad (16)$$

where g is the operator $g : A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=1}^K \tau_k^* \tau_k A \tau_k^* \tau_k \in \mathcal{L}(\mathcal{X})$. Then, using (11), (15) and (16), we obtain $\sqrt{n}(\widehat{T}_n - T) = \widehat{\varphi}_n(\widehat{H}_n)$, where $\widehat{\varphi}_n$ is the random operator from $\mathcal{L}(\mathcal{X})$ to itself defined by

$$\widehat{\varphi}_n(A) = -(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} g(A) \widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2} + f(A) \widehat{\Phi}_n^{-1/2} - \Psi(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} g(A) \widehat{\Phi}_n^{-1}.$$

Considering the operator

$$\varphi : A \in \mathcal{L}(\mathcal{X}) \mapsto -\frac{1}{2}g(A)\Psi + f(A) - \frac{1}{2}\Psi g(A) \in \mathcal{L}(\mathcal{X}),$$

and denoting by $\|\cdot\|_{\infty}$ (resp. $\|\cdot\|_{\infty\infty}$) the norm of $\mathcal{L}(\mathcal{X})$ (resp. $\mathcal{L}(\mathcal{L}(\mathcal{X}))$) defined by $\|A\|_{\infty} = \sup_{x \in \mathcal{X} - \{0\}} \|Ax\|_{\mathcal{X}} / \|x\|_{\mathcal{X}}$ (resp. $\|h\|_{\infty\infty} = \sup_{B \in \mathcal{L}(\mathcal{X}) - \{0\}} \|h(B)\|_{\infty} / \|B\|_{\infty}$) for any A (resp. h) in $\mathcal{L}(\mathcal{X})$ (resp. $\mathcal{L}(\mathcal{L}(\mathcal{X}))$), we have

$$\begin{aligned} \|\widehat{\varphi}_n(\widehat{H}_n) - \varphi(\widehat{H}_n)\|_{\infty} &= \left\| - \left((\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}} \right) g(\widehat{H}_n) \widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2} \right. \\ &\quad - \frac{1}{2}g(\widehat{H}_n) \left(\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2} - \Psi \right) + f(\widehat{H}_n) \left(\widehat{\Phi}_n^{-1/2} - I_{\mathcal{X}} \right) \\ &\quad \left. - \Psi \left((\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}} \right) g(\widehat{H}_n) \widehat{\Phi}_n^{-1} - \frac{1}{2}\Psi(\widehat{\Phi}_n^{-1} - I_{\mathcal{X}}) \right\|_{\infty} \\ &\leq \|(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}}\|_{\infty} \|g(\widehat{H}_n)\|_{\infty} \|\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2}\|_{\infty} \\ &\quad + \frac{1}{2}\|g(\widehat{H}_n)\|_{\infty} \|\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2} - \Psi\|_{\infty} + \|f(\widehat{H}_n)\|_{\infty} \|\widehat{\Phi}_n^{-1/2} - I_{\mathcal{X}}\|_{\infty} \\ &\quad + \|\Psi\|_{\infty} \|(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}}\|_{\infty} \|g(\widehat{H}_n)\|_{\infty} \|\widehat{\Phi}_n^{-1}\|_{\infty} \\ &\quad + \frac{1}{2}\|\Psi\|_{\infty} \|g(\widehat{H}_n)\|_{\infty} \|\widehat{\Phi}_n^{-1} - I_{\mathcal{X}}\|_{\infty} \\ &\leq \left(\|(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}}\|_{\infty} \|g\|_{\infty\infty} \|\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2}\|_{\infty} \right. \\ &\quad + \frac{1}{2}\|g\|_{\infty\infty} \|\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2} - \Psi\|_{\infty} + \|f\|_{\infty\infty} \|\widehat{\Phi}_n^{-1/2} - I_{\mathcal{X}}\|_{\infty} \\ &\quad + \|\Psi\|_{\infty} \|(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1} - \frac{1}{2}I_{\mathcal{X}}\|_{\infty} \|g\|_{\infty\infty} \|\widehat{\Phi}_n^{-1}\|_{\infty} \\ &\quad \left. + \frac{1}{2}\|\Psi\|_{\infty} \|g\|_{\infty\infty} \|\widehat{\Phi}_n^{-1} - I_{\mathcal{X}}\|_{\infty} \right) \|\widehat{H}_n\|_{\infty}. \end{aligned} \tag{17}$$

Using the strong law of large numbers, it is easy to verify that, for any $(k, \ell) \in \{1, \dots, K\}^2$ with $k \neq \ell$, $\widehat{V}_{k\ell \cdot n}$ (resp. $\widehat{V}_{k \cdot n}$) converge almost surely to $V_{k\ell}$ (resp. \widehat{V}_k), as $n \rightarrow +\infty$. Consequently, $\widehat{\Phi}_n$ (resp. $\widehat{\Psi}_n$) converge almost surely to $\Phi = I_{\mathcal{X}}$ (resp. Ψ), as $n \rightarrow +\infty$. This implies the almost sure convergence of $(\widehat{\Phi}_n^{-1/2} + I_{\mathcal{X}})^{-1}$ (resp. $\widehat{\Phi}_n^{-1} \widehat{\Psi}_n \widehat{\Phi}_n^{-1/2}$; resp. $\widehat{\Phi}_n^{-1}$; resp. $\widehat{\Phi}_n^{-1/2}$) to $\frac{1}{2}I_{\mathcal{X}}$ (resp. Ψ ; resp. $I_{\mathcal{X}}$; resp. $I_{\mathcal{X}}$), as $n \rightarrow +\infty$. Furthermore, denoting by $\|\cdot\|$ the norm of $\mathcal{L}(\mathcal{X})$ defined by $\|A\| = \sqrt{\text{tr}(AA^*)}$ and using the properties $(a \otimes b)(c \otimes d) =$

$a, d > c \otimes b$ and $\text{tr}(a \otimes b) = \langle a, b \rangle$ of the tensor product (see Dauxois et al. (1994)), we have:

$$\begin{aligned} \mathbb{E}(\|X \otimes X\|^2) &= \mathbb{E}(\text{tr}((X \otimes X)(X \otimes X))) = \mathbb{E}(\|X\|_{\mathcal{X}}^4) = \mathbb{E}\left(\left(\sum_{k=1}^K \|X_k\|_k^2\right)^2\right) \\ &= \sum_{k=1}^K \mathbb{E}(\|X_k\|_k^4) + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \mathbb{E}(\|X_k\|_k^2 \|X_\ell\|_\ell^2) \\ &\leq \sum_{k=1}^K \mathbb{E}(\|X_k\|_k^4) + \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sqrt{\mathbb{E}(\|X_k\|_k^4)} \sqrt{\mathbb{E}(\|X_\ell\|_\ell^4)} < +\infty. \end{aligned}$$

Then, the central limit theorem can be used. It gives the convergence in distribution, as $n \rightarrow +\infty$, of $\sqrt{n}(n^{-1} \sum_{i=1}^n X^{(i)} \otimes X^{(i)} - V)$ to a random variable H having the normal distribution in $\mathcal{L}(\mathcal{X})$ with mean equal to 0 and a covariance operator equal to that of $X \otimes X$. Since, by the central limit theorem again, $\sqrt{n}\bar{X}_n$ converges in distribution, as $n \rightarrow +\infty$, to a random variable having a normal distribution in \mathcal{X} with mean equal to 0 and a covariance operator equal to V , we deduce from the equality $\sqrt{n}(\bar{X}_n \otimes \bar{X}_n) = n^{-1/2}(\sqrt{n}\bar{X}_n) \otimes (\sqrt{n}\bar{X}_n)$ that $\sqrt{n}(\bar{X}_n \otimes \bar{X}_n)$ converges in probability to 0, as $n \rightarrow +\infty$. Therefore, from (14) and Slutsky's theorem, we deduce that \hat{H}_n converges in distribution, as $n \rightarrow +\infty$ to H . Then, from (17), we conclude that $\hat{\varphi}_n(\hat{H}_n) - \varphi(\hat{H}_n)$ converges in probability to 0, as $n \rightarrow +\infty$. Then, using again Slutsky's theorem, we deduce that $\hat{\varphi}_n(\hat{H}_n)$ and $\varphi(\hat{H}_n)$ both converge in distribution to the same distribution. Since φ is a linear map (and is, therefore, continuous), this distribution just is that of the random variable $U = \varphi(H)$, that is the normal distribution in $\mathcal{L}(\mathcal{X})$ with mean 0 and covariance operator equal to that of $Z = \varphi(X \otimes X)$. Clearly,

$$g(X \otimes X) = \sum_{k=1}^K \tau_k^* \tau_k(X \otimes X) \tau_k^* \tau_k = \sum_{k=1}^K \tau_k^* ((\tau_k(X)) \otimes (\tau_k(X))) \tau_k = \sum_{k=1}^K \tau_k^*(X_k \otimes X_k) \tau_k,$$

and

$$f(X \otimes X) = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* \tau_k(X \otimes X) \tau_\ell^* \tau_\ell = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^*(X_\ell \otimes X_k) \tau_\ell.$$

Then, since $\tau_k \tau_j^* = \delta_{kj} I_k$, it follows

$$g(X \otimes X) \Psi = \sum_{k=1}^K \sum_{j=1}^K \sum_{\substack{\ell=1 \\ \ell \neq j}}^K \tau_k^*(X_k \otimes X_k) \tau_k \tau_j^* V_{j\ell} \tau_\ell = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^*(X_k \otimes X_k) V_{k\ell} \tau_\ell$$

and

$$\begin{aligned}\Psi g(X \otimes X) &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{j=1}^K \tau_k^* V_{k\ell} \tau_\ell \tau_j^* (X_j \otimes X_j) \tau_j = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_k^* V_{k\ell} (X_\ell \otimes X_\ell) \tau_\ell \\ &= \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \tau_\ell^* V_{\ell k} (X_k \otimes X_k) \tau_k.\end{aligned}$$

Thus,

$$Z = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K -\frac{1}{2} (\tau_k^* (X_k \otimes X_k) V_{k\ell} \tau_\ell + \tau_\ell^* V_{\ell k} (X_k \otimes X_k) \tau_k) + \tau_k^* (X_\ell \otimes X_k) \tau_\ell.$$

□

Using the preceding theorem and results in Eaton and Tyler (1991,1994), we can now give asymptotic distributions for the canonical coefficients. We denote by $(\rho'_j)_{1 \leq j \leq r}$ (with $r \in \mathbb{N}^*$) the sequence of distinct eigenvalues of T in decreasing order, that is $\rho'_1 > \dots > \rho'_r$. Putting $m_0 = 0$, denoting by m_j the multiplicity of ρ'_j and putting $\nu_j = \sum_{k=0}^{j-1} m_k$ for any $j \in \{1, \dots, r\}$, it is clear that for any $i \in \{\nu_{j-1} + 1, \dots, \nu_j\}$ one has $\rho_i = \rho'_j$. Further, considering the eigenspace $E_j = \ker(T - \rho'_j I)$, we have the following decomposition in orthogonal direct sum: $\mathcal{X} = E_1 \oplus \dots \oplus E_r$. We denote by Π_j the orthogonal projector from \mathcal{X} onto E_j , and by Δ the continuous map which associates to each self-adjoint operator A the vector $\Delta(A)$ of its eigenvalues in nonincreasing order. For $j \in \{1, \dots, r\}$, we consider m_j -dimensional vector given by $v_j = \rho'_j \mathbb{J}_{m_j}$, where \mathbb{J}_q denotes the q -dimensional vector with elements all equal to 1, and the \mathbb{R}^{m_j} -valued random vector:

$$\hat{v}_j^n = \begin{pmatrix} \hat{\rho}_{\nu_{j-1}+1 \cdot n} \\ \vdots \\ \hat{\rho}_{\nu_j \cdot n} \end{pmatrix}.$$

Then, putting

$$\hat{\Lambda}_n = \begin{pmatrix} \hat{v}_1^n \\ \vdots \\ \hat{v}_r^n \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} v_1 \\ \vdots \\ v_r \end{pmatrix},$$

we have:

Theorem 3.3. $\sqrt{n}(\hat{\Lambda}_n - \Lambda)$ converges in distribution, as $n \rightarrow +\infty$, to the \mathbb{R}^p -valued random vector

$$\zeta = \begin{pmatrix} \Delta(\Pi_1 W \Pi_1) \\ \vdots \\ \Delta(\Pi_r W \Pi_r) \end{pmatrix}, \quad (18)$$

where W is a random variable having a normal distribution in $\mathcal{L}(\mathcal{X})$, with mean 0 and covariance operator Θ given by:

$$\Theta = \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) (e_m \otimes e_r) \tilde{\otimes} (e_s \otimes e_t)$$

with

$$C(m, r, s, t) = \sum_{k=1}^K \sum_{j=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{\substack{q=1 \\ q \neq j}}^K \left(\gamma_{k\ell j q}^{m, r, s, t} + \gamma_{k\ell j q}^{m, r, t, s} + \gamma_{k\ell j q}^{r, m, s, t} + \gamma_{k\ell j q}^{r, m, t, s} \right. \\ \left. - \theta_{k\ell j q}^{m, r, s, t} - \theta_{k\ell j q}^{r, m, s, t} - \theta_{k\ell j q}^{s, t, m, r} - \theta_{k\ell j q}^{t, s, m, r} + \lambda_{k\ell j q}^{m, r, s, t} \right),$$

$$\gamma_{k\ell j q}^{a, b, c, d} = \frac{1}{4} \mathbb{E} \left(\langle X_k, \tau_k \beta^{(a)} \rangle_k \langle X_k, V_{k\ell} \tau_\ell \beta^{(b)} \rangle_k \langle X_j, \tau_j \beta^{(c)} \rangle_j \langle X_j, V_{jq} \tau_q \beta^{(d)} \rangle_j \right),$$

$$\theta_{k\ell j q}^{a, b, c, d} = \frac{1}{2} \mathbb{E} \left(\langle X_k, \tau_k \beta^{(a)} \rangle_k \langle X_k, V_{k\ell} \tau_\ell \beta^{(b)} \rangle_k \langle X_j, \tau_j \beta^{(c)} \rangle_j \langle X_q, \tau_q \beta^{(d)} \rangle_q \right)$$

and

$$\gamma_{k\ell j q}^{a, b, c, d} = \mathbb{E} \left(\langle X_k, \tau_k \beta^{(a)} \rangle_k \langle X_\ell, \tau_\ell \beta^{(b)} \rangle_\ell \langle X_j, \tau_j \beta^{(c)} \rangle_j \langle X_q, \tau_q \beta^{(d)} \rangle_q \right).$$

Proof. Since $\Delta(\widehat{T}_n) = \widehat{\Lambda}_n$ and $\Delta(T) = \Lambda$, we deduce from Theorem 3.2 and the Theorem 2.1 of Eaton and Tyler (1994) that $\sqrt{n}(\widehat{\Lambda}_n - \Lambda)$ converges in distribution, as $n \rightarrow +\infty$, to the random variable given in (18) with $W = P^*UP$, where $P = \sum_{\ell=1}^p e_\ell \otimes \beta^{(\ell)}$. Clearly, W has a normal distribution with mean 0 and covariance operator Θ equal to that of P^*ZP . In order to give an explicit expression of Θ , let us first note that:

$$P^*ZP = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K -\frac{1}{2} (P^* \tau_k^* (X_k \otimes X_k) V_{k\ell} \tau_\ell P + P^* \tau_\ell^* V_{\ell k} (X_k \otimes X_k) \tau_k P) \\ + P^* \tau_k^* (X_\ell \otimes X_k) \tau_\ell P \\ = \sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K -\frac{1}{2} ((P^* \tau_\ell^* V_{\ell k} X_k) \otimes (P^* \tau_k^* X_k) + (P^* \tau_k^* X_k) \otimes (P^* \tau_\ell^* V_{\ell k} X_k)) \\ + (P^* \tau_\ell^* X_\ell) \otimes (P^* \tau_k^* X_k).$$

Since

$$P^* \tau_\ell^* V_{\ell k} X_k = \left(\sum_{m=1}^p \beta^{(m)} \otimes e_m \right) \tau_\ell^* V_{\ell k} X_k = \sum_{m=1}^p \langle \beta^{(m)}, \tau_\ell^* V_{\ell k} X_k \rangle_{\mathcal{X}} e_m \\ = \sum_{m=1}^p \langle \tau_\ell \beta^{(m)}, V_{\ell k} X_k \rangle_\ell e_m$$

and, similarly, $P^* \tau_k^* X_k = \sum_{m=1}^p \langle \tau_k \beta^{(m)}, X_k \rangle_k e_m$, it follows:

$$\begin{aligned} P^* Z P &= \sum_{m=1}^p \sum_{r=1}^p \left[\sum_{k=1}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K -\frac{1}{2} (\langle \tau_\ell \beta^{(m)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k \right. \\ &\quad + \langle \tau_\ell \beta^{(r)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(m)}, X_k \rangle_k \\ &\quad \left. + \langle \tau_\ell \beta^{(m)}, X_\ell \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k \right] e_m \otimes e_r. \end{aligned}$$

From:

$$\begin{aligned} \mathbb{E} \left(\langle \tau_\ell \beta^{(m)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k \right) &= \mathbb{E} \left(\langle (X_k \otimes X_k) (\tau_k \beta^{(r)}), V_{k\ell} \tau_\ell \beta^{(m)} \rangle_k \right) \\ &= \mathbb{E} \left(\langle X_k \otimes X_k \rangle (\tau_k \beta^{(r)}), V_{k\ell} \tau_\ell \beta^{(m)} \right) \\ &= \langle V_k \tau_k \beta^{(r)}, V_{k\ell} \tau_\ell \beta^{(m)} \rangle_k \\ &= \langle \tau_k \beta^{(r)}, V_{k\ell} \tau_\ell \beta^{(m)} \rangle_k, \end{aligned}$$

$$\mathbb{E} \left(\langle \tau_\ell \beta^{(r)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(m)}, X_k \rangle_k \right) = \langle \tau_k \beta^{(m)}, V_{k\ell} \tau_\ell \beta^{(r)} \rangle_k$$

and

$$\begin{aligned} \mathbb{E} \left(\langle \tau_\ell \beta^{(m)}, X_\ell \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k \right) &= \mathbb{E} \left(\langle (X_\ell \otimes X_k) (\tau_k \beta^{(r)}), \tau_\ell \beta^{(m)} \rangle_k \right) \\ &= \langle \mathbb{E} \left(X_\ell \otimes X_k \right) (\tau_k \beta^{(r)}), \tau_\ell \beta^{(m)} \rangle_k \\ &= \langle V_{k\ell} \tau_\ell \beta^{(m)}, \tau_k \beta^{(r)} \rangle_k, \end{aligned}$$

we deduce that $\mathbb{E}(P^* Z P) = 0$. Thus,

$$\Theta = \mathbb{E} \left((P^* Z P) \tilde{\otimes} (P^* Z P) \right) = \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) (e_m \otimes e_r) \tilde{\otimes} (e_s \otimes e_t),$$

where

$$C(m, r, s, t) = \sum_{k=1}^K \sum_{\substack{j=1 \\ \ell \neq k}}^K \sum_{\substack{\ell=1 \\ \ell \neq k}}^K \sum_{\substack{q=1 \\ q \neq j}}^K \mathbb{E} \left(Y_{k\ell}^{m,r} Y_{jq}^{s,q} \right)$$

with

$$\begin{aligned} Y_{k\ell}^{m,r} &= -\frac{1}{2} \left(\langle \tau_\ell \beta^{(m)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k + \langle \tau_\ell \beta^{(r)}, V_{\ell k} X_k \rangle_{\ell} \langle \tau_k \beta^{(m)}, X_k \rangle_k \right) \\ &\quad + \langle \tau_\ell \beta^{(m)}, X_\ell \rangle_{\ell} \langle \tau_k \beta^{(r)}, X_k \rangle_k. \end{aligned}$$

Further calculations give

$$\begin{aligned} \mathbb{E} \left(Y_{k\ell}^{m,r} Y_{jq}^{s,q} \right) &= \gamma_{k\ell j q}^{m,r,s,t} + \gamma_{k\ell j q}^{m,r,t,s} + \gamma_{k\ell j q}^{r,m,s,t} + \gamma_{k\ell j q}^{r,m,t,s} \\ &\quad - \theta_{k\ell j q}^{m,r,s,t} - \theta_{k\ell j q}^{r,m,s,t} - \theta_{k\ell j q}^{s,t,m,r} - \theta_{k\ell j q}^{t,s,m,r} + \lambda_{k\ell j q}^{m,r,s,t}. \end{aligned}$$

□

When T has simple eigenvalues, that is $\rho_1 > \rho_2 > \dots > \rho_q$, the preceding theorem has a simpler statement. We have:

Corollary 3.1. *When the eigenvalues of T are simple, $\sqrt{n}(\widehat{\Lambda}_n - \Lambda)$ converges in distribution, as $n \rightarrow +\infty$, to a random variable having a normal distribution in \mathbb{R}^p with mean 0 and covariance matrix $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq p}$ with:*

$$\sigma_{ij} = \sum_{1 \leq m, r, s, t \leq p} \beta_m^{(i)} \beta_r^{(i)} \beta_s^{(j)} \beta_t^{(j)} C(m, r, s, t).$$

Proof. In this case, $m_1 = \dots = m_p = 1$ and, for any $j \in \{1, \dots, p\}$, $\Pi_j = \beta^{(j)} \otimes \beta^{(j)}$. Thus

$$\begin{aligned} \Pi_j W \Pi_j &= ((\beta^{(j)} \otimes \beta^{(j)}) W (\beta^{(j)} \otimes \beta^{(j)})) = (\beta^{(j)} \otimes \beta^{(j)}) (\beta^{(j)} \otimes (W \beta^{(j)})) \\ &= \langle \beta^{(j)}, W \beta^{(j)} \rangle_{\mathcal{X}} \beta^{(j)} \otimes \beta^{(j)}, \end{aligned}$$

and, therefore, $\Delta(\Pi_j W \Pi_j) = \langle \beta^{(j)}, W \beta^{(j)} \rangle_{\mathcal{X}}$. Then, ζ is a linear function of W and, consequently, it has a normal distribution with mean 0 and covariance matrix $\Sigma = (\sigma_{ij})_{1 \leq i, j \leq p}$ with $\sigma_{ij} = \mathbb{E}(\langle \beta^{(i)}, W \beta^{(i)} \rangle_{\mathcal{X}} \langle \beta^{(j)}, W \beta^{(j)} \rangle_{\mathcal{X}})$. Denoting by $\langle \cdot, \cdot \rangle$ the inner product of operators defined by $\langle A, B \rangle = \text{tr}(AB^*)$, we have:

$$\langle W, \beta^{(j)} \otimes \beta^{(j)} \rangle = \text{tr}(W(\beta^{(j)} \otimes \beta^{(j)})) = \text{tr}(\beta^{(j)} \otimes (W \beta^{(j)})) = \langle \beta^{(j)}, W \beta^{(j)} \rangle_{\mathcal{X}},$$

it follows that

$$\begin{aligned} \sigma_{ij} &= \mathbb{E}(\langle \beta^{(i)}, W \beta^{(i)} \rangle_{\mathcal{X}} \langle \beta^{(j)}, W \beta^{(j)} \rangle_{\mathcal{X}}) \\ &= \mathbb{E}(\langle W, \beta^{(i)} \otimes \beta^{(i)} \rangle \langle W, \beta^{(j)} \otimes \beta^{(j)} \rangle) \\ &= \mathbb{E}(\langle (W \tilde{\otimes} W)(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} \rangle) \\ &= \langle \mathbb{E}(W \tilde{\otimes} W)(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} \rangle \\ &= \langle \Theta(\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} \rangle \\ &= \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) \langle (e_m \otimes e_r) \tilde{\otimes} (e_s \otimes e_t) (\beta^{(i)} \otimes \beta^{(i)}), \beta^{(j)} \otimes \beta^{(j)} \rangle \\ &= \sum_{1 \leq m, r, s, t \leq p} C(m, r, s, t) \langle e_m \otimes e_r, \beta^{(i)} \otimes \beta^{(i)} \rangle \langle e_s \otimes e_t, \beta^{(j)} \otimes \beta^{(j)} \rangle. \end{aligned}$$

Then, the required result is obtained from

$$\begin{aligned} \langle e_m \otimes e_r, \beta^{(i)} \otimes \beta^{(i)} \rangle &= \text{tr}((e_m \otimes e_r)(\beta^{(i)} \otimes \beta^{(i)})) \\ &= \text{tr}(\langle e_m, \beta^{(i)} \rangle_{\mathcal{X}} \beta^{(i)} \otimes e_r) \\ &= \langle e_m, \beta^{(i)} \rangle_{\mathcal{X}} \langle e_r, \beta^{(i)} \rangle_{\mathcal{X}} \\ &= \beta_m^{(i)} \beta_r^{(i)} \end{aligned}$$

and $\langle e_s \otimes e_t, \beta^{(j)} \otimes \beta^{(j)} \rangle = \beta_s^{(j)} \beta_t^{(j)}$. \square

4 Testing for mutual non-correlation

In this section, we consider the problem of testing for mutual non-correlation between X_1, X_2, \dots, X_K , that is testing for non-correlation between any pair (X_k, X_ℓ) for $(k, \ell) \in \{1, \dots, K\}^2$. The null hypothesis is

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

and the alternative is given by:

$$\mathcal{H}_1 : \exists (k, \ell) \in \{1, \dots, K\}^2, k \neq \ell, V_{k\ell} \neq 0,$$

that is testing for non-correlation between any pair (X_k, X_ℓ) of variables. We will first introduce a test statistic for this problem, then we will derive its asymptotic distribution under the null hypothesis, in the general case and in case X has an elliptical distribution, by using the results of the preceding section.

4.1 A test statistic

For $(k, \ell) \in \{1, \dots, K\}^2$, denoting by $\pi_{k\ell}$ the operator

$$\pi_{k\ell} : A \in \mathcal{L}(\mathcal{X}) \mapsto \tau_k A \tau_\ell^* \in \mathcal{L}(\mathcal{X}_\ell, \mathcal{X}_k),$$

we have for $k \neq \ell$:

$$\begin{aligned} \pi_{k\ell}(T) &= \pi_{k\ell}(\Phi^{-1/2} \Psi \Phi^{-1/2}) = \sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K \tau_k \tau_i^* V_i^{-1/2} V_{ij} V_j^{-1/2} \tau_j \tau_\ell^* \\ &= \sum_{i=1}^K \sum_{\substack{j=1 \\ j \neq i}}^K \delta_{ki} \delta_{j\ell} V_i^{-1/2} V_{ij} V_j^{-1/2} = V_k^{-1/2} V_{k\ell} V_\ell^{-1/2}. \end{aligned}$$

Therefore, \mathcal{H}_0 is equivalent to having $\sum_{k=2}^K \sum_{\ell=1}^{k-1} \text{tr}(\pi_{k\ell}(T) \pi_{k\ell}(T)^*) = 0$. This leads us to take as test statistic the random variable \widehat{S}_n given by:

$$\widehat{S}_n = \sum_{k=2}^K \sum_{\ell=1}^{k-1} \text{tr}(\pi_{k\ell}(\widehat{T}_n) \pi_{k\ell}(\widehat{T}_n)^*).$$

4.2 Asymptotic distribution under null hypothesis

For $k \in \{1, \dots, K\}$, we denote by $\{e_j^{(k)}; j = 1, \dots, p_k\}$ an orthonormal basis of \mathcal{X}_k , where $p_k = \dim(\mathcal{X}_k)$, and we consider the matrix:

$$\Gamma = \begin{pmatrix} \Gamma_{21,21} & \Gamma_{21,31} & \Gamma_{21,32} & \cdots & \Gamma_{21,KK-1} \\ \Gamma_{31,21} & \Gamma_{31,31} & \Gamma_{31,32} & \cdots & \Gamma_{31,KK-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \Gamma_{KK-1,21} & \Gamma_{KK-1,31} & \Gamma_{KK-1,32} & \cdots & \Gamma_{KK-1,KK-1} \end{pmatrix}, \quad (19)$$

where $\Gamma_{kl,rs}$ is the $p_k p_\ell \times p_r p_s$ matrix given by

$$\Gamma_{kl,rs} = \begin{pmatrix} \gamma_{1111}^{kl,rs} & \gamma_{1121}^{kl,rs} & \cdots & \gamma_{11p_r-1}^{kl,rs} & \cdots & \gamma_{111p_s}^{kl,rs} & \gamma_{112p_s}^{kl,rs} & \cdots & \gamma_{11p_r p_s}^{kl,rs} \\ \gamma_{2111}^{kl,rs} & \gamma_{2121}^{kl,rs} & \cdots & \gamma_{21p_r-1}^{kl,rs} & \cdots & \gamma_{211p_s}^{kl,rs} & \gamma_{212p_s}^{kl,rs} & \cdots & \gamma_{21p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{p_k 111}^{kl,rs} & \gamma_{p_k 121}^{kl,rs} & \cdots & \gamma_{p_k 1p_r-1}^{kl,rs} & \cdots & \gamma_{p_k 11p_s}^{kl,rs} & \gamma_{p_k 12p_s}^{kl,rs} & \cdots & \gamma_{p_k 1p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{1p_\ell 11}^{kl,rs} & \gamma_{1p_\ell 21}^{kl,rs} & \cdots & \gamma_{1p_\ell p_r-1}^{kl,rs} & \cdots & \gamma_{1p_\ell 1p_s}^{kl,rs} & \gamma_{1p_\ell 2p_s}^{kl,rs} & \cdots & \gamma_{1p_\ell p_r p_s}^{kl,rs} \\ \gamma_{2p_\ell 11}^{kl,rs} & \gamma_{2p_\ell 21}^{kl,rs} & \cdots & \gamma_{2p_\ell p_r-1}^{kl,rs} & \cdots & \gamma_{2p_\ell 1p_s}^{kl,rs} & \gamma_{2p_\ell 2p_s}^{kl,rs} & \cdots & \gamma_{2p_\ell p_r p_s}^{kl,rs} \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \gamma_{p_k p_\ell 11}^{kl,rs} & \gamma_{p_k p_\ell 21}^{kl,rs} & \cdots & \gamma_{p_k p_\ell p_r-1}^{kl,rs} & \cdots & \gamma_{p_k p_\ell 1p_s}^{kl,rs} & \gamma_{p_k p_\ell 2p_s}^{kl,rs} & \cdots & \gamma_{p_k p_\ell p_r p_s}^{kl,rs} \end{pmatrix}, \quad (20)$$

with

$$\gamma_{ijpq}^{kl,rs} = \mathbb{E} \left(\langle X_k, e_i^{(k)} \rangle_k \langle X_r, e_p^{(r)} \rangle_r \langle X_\ell, e_j^{(\ell)} \rangle_\ell \langle X_s, e_q^{(s)} \rangle_s \right).$$

Then, putting $d = \sum_{k=1}^K \sum_{\ell=1}^{k-1} p_k p_\ell$, we have:

Theorem 4.1. Under \mathcal{H}_0 , the sequence $n\widehat{S}_n$ converges in distribution, as $n \rightarrow +\infty$, to $\mathcal{Q} = \mathbb{W}^T \mathbb{W}$, where \mathbb{W} is a random variable having a centered normal distribution in \mathbb{R}^d with covariance matrix Γ .

Proof. First, note that under \mathcal{H}_0 we have $\Psi = 0$ and, therefore, $T = 0$. Then, the asymptotic distribution of \widehat{S}_n under \mathcal{H}_0 can be obtained Theorem 3.2 which shows that $\sqrt{n}\widehat{T}_n$ converges in distribution, as $n \rightarrow +\infty$, to a r.v. U having a normal distribution in $\mathcal{L}(\mathcal{X})$ with mean 0 and covariance operator equal to that of Z . Then, since the map $A \in \mathcal{L}(\mathcal{X}) \mapsto \sum_{k=2}^K \sum_{\ell=1}^{k-1} \text{tr}(\pi_{k\ell}(A)\pi_{k\ell}(A)^*) \in \mathbb{R}$ is continuous, we deduce that $n\widehat{S}_n$ converges in distribution, as $n \rightarrow +\infty$, to $\mathcal{Q} = \sum_{k=2}^K \sum_{\ell=1}^{k-1} \text{tr}(W_{k\ell}W_{k\ell}^*)$, where $W_{k\ell} = \pi_{k\ell}(U)$. We can write $W_{k\ell} = \sum_{i=1}^{p_k} \sum_{j=1}^{p_\ell} W_{k\ell}^{ij} e_j^{(\ell)} \otimes e_i^{(k)}$ where $W_{k\ell}^{ij} = \langle W_{k\ell}, e_j^{(\ell)} \otimes e_i^{(k)} \rangle$, and, using the equalities $(u \otimes v)(w \otimes z) = \langle u, z \rangle w \otimes v$ and $\text{tr}(u \otimes v) = \langle u, v \rangle$ (cf. Dauxois et al. (1994)), we have:

$$\text{tr}(W_{k\ell}W_{k\ell}^*) = \sum_{i=1}^{p_k} \sum_{j=1}^{p_\ell} \sum_{r=1}^{p_k} \sum_{s=1}^{p_\ell} \delta_{js} \delta_{ir} W_{k\ell}^{ij} W_{k\ell}^{rs} = \sum_{i=1}^{p_k} \sum_{j=1}^{p_\ell} \left(W_{k\ell}^{ij} \right)^2 = \mathbb{W}_{k,\ell}^T \mathbb{W}_{k,\ell}.$$

where δ denotes the usual Kronecker symbol and

$$\mathbb{W}_{k,\ell} = \left(W_{k\ell}^{11}, W_{k\ell}^{21}, \dots, W_{k\ell}^{p_k 1}, \dots, W_{k\ell}^{1p_\ell}, W_{k\ell}^{2p_\ell}, \dots, W_{k\ell}^{p_k p_\ell} \right)^T.$$

Hence, $\mathcal{Q} = \mathbb{W}^T \mathbb{W}$, where $\mathbb{W} = \left(\mathbb{W}_{2,1}^T, \mathbb{W}_{3,1}^T, \mathbb{W}_{3,2}^T, \dots, \mathbb{W}_{K,1}^T, \dots, \mathbb{W}_{K,K-1}^T \right)^T$. Since \mathbb{W} is a linear function of U , it is also normally distributed with mean 0

and covariance matrix Γ defined in (19) and (20) with $\gamma_{ijpq}^{k\ell,rs} = \mathbb{E} \left(W_{k\ell}^{ij} W_{rs}^{pq} \right)$. Further,

$$\begin{aligned} \gamma_{ijpq}^{k\ell,rs} &= \mathbb{E} \left(\langle W_{k\ell}, e_j^{(\ell)} \otimes e_i^{(k)} \rangle \langle W_{rs}, e_q^{(s)} \otimes e_p^{(r)} \rangle \right) \\ &= \mathbb{E} \left(\langle (W_{k\ell} \tilde{\otimes} W_{rs})(e_j^{(\ell)} \otimes e_i^{(k)}), e_q^{(s)} \otimes e_p^{(r)} \rangle \right) \\ &= \langle \mathbb{E}(W_{k\ell} \tilde{\otimes} W_{rs})(e_j^{(\ell)} \otimes e_i^{(k)}), e_q^{(s)} \otimes e_p^{(r)} \rangle, \end{aligned}$$

where $\tilde{\otimes}$ denotes the tensor product related to the inner product of operators $\langle A, B \rangle = \text{tr}(AB^*)$. Recalling that U has the same covariance operator than $\varphi(X \otimes X)$, that is $\mathbb{E}(U \tilde{\otimes} U) = \mathbb{E}((\varphi(X \otimes X)) \tilde{\otimes} (\varphi(X \otimes X)))$, we obtain

$$\begin{aligned} \mathbb{E}(W_{k\ell} \tilde{\otimes} W_{rs}) &= \mathbb{E}((\pi_{k\ell}(U)) \tilde{\otimes} (\pi_{rs}(U))) = \pi_{rs} \mathbb{E}(U \tilde{\otimes} U) \pi_{k\ell}^* \\ &= \pi_{rs} \mathbb{E}((\varphi(X \otimes X)) \tilde{\otimes} (\varphi(X \otimes X))) \pi_{k\ell}^* \\ &= \mathbb{E}((\pi_{k\ell}(\varphi(X \otimes X))) \tilde{\otimes} (\pi_{rs}(\varphi(X \otimes X)))) \\ &= \mathbb{E}((X_\ell \otimes X_k) \tilde{\otimes} (X_s \otimes X_r)). \end{aligned}$$

Thus, $\mathbb{E}(W_{k\ell} W_{rs}^T) = \mathbb{E} \left(\langle X_\ell \otimes X_k, e_j^{(\ell)} \otimes e_i^{(k)} \rangle \langle X_s \otimes X_r, e_q^{(s)} \otimes e_p^{(r)} \rangle \right)$

$$\begin{aligned} \gamma_{ijpq}^{k\ell,rs} &= \mathbb{E} \left(\langle X_\ell \otimes X_k, e_j^{(\ell)} \otimes e_i^{(k)} \rangle \langle X_s \otimes X_r, e_q^{(s)} \otimes e_p^{(r)} \rangle \right) \\ &= \mathbb{E} \left(\langle X_k, e_i^{(k)} \rangle \langle X_r, e_p^{(r)} \rangle \langle X_\ell, e_j^{(\ell)} \rangle \langle X_s, e_q^{(s)} \rangle \right). \end{aligned}$$

□

4.3 Case of elliptical distribution

Recall that a random variable valued into an Euclidean space E is said to have an elliptical distribution $\mathcal{E}(\mu, \Sigma, h)$ if its characteristic function is of the form:

$$\exp(i \langle t, \mu \rangle_E) h(\langle t, \Sigma t \rangle_E).$$

In this subsection, we assume that X has an elliptical distribution $\mathcal{E}(0, V, h)$. Then, the previous theorem has a more explicit formulation stated below.

Theorem 4.2. *When X has the elliptical distribution $\mathcal{E}(0, V, h)$ then, under \mathcal{H}_0 , the sequence $n\hat{S}_n$ converges in distribution, as $n \rightarrow +\infty$, to $4h''(0)\chi_d^2$. *Proof.* The \mathbb{R}^4 -valued random vector*

$$S = \begin{pmatrix} \langle X_k, e_i^{(k)} \rangle >_k \\ \langle X_r, e_p^{(r)} \rangle >_r \\ \langle X_\ell, e_j^{(\ell)} \rangle >_\ell \\ \langle X_s, e_q^{(s)} \rangle >_s \end{pmatrix}$$

can obviously be written as $S = L(X)$, where L is a linear map from \mathcal{X} to \mathbb{R}^4 . Then, S has the elliptical distribution $E(0, LVL^*, h)$ and, therefore,

$$\begin{aligned} \gamma_{ijpq}^{k\ell,rs} &= 4h''(0) \left[\mathbb{E} \left(\langle X_k, e_i^{(k)} \rangle_k \langle X_r, e_p^{(r)} \rangle_r \right) \mathbb{E} \left(\langle X_\ell, e_j^{(\ell)} \rangle_\ell \langle X_s, e_q^{(s)} \rangle_s \right) \right. \\ &\quad + \mathbb{E} \left(\langle X_k, e_i^{(k)} \rangle_k \langle X_\ell, e_j^{(\ell)} \rangle_\ell \right) \mathbb{E} \left(\langle X_r, e_p^{(r)} \rangle_r \langle X_s, e_q^{(s)} \rangle_s \right) \\ &\quad \left. + \mathbb{E} \left(\langle X_k, e_i^{(k)} \rangle_k \langle X_s, e_q^{(s)} \rangle_s \right) \mathbb{E} \left(\langle X_r, e_p^{(r)} \rangle_r \langle X_\ell, e_j^{(\ell)} \rangle_\ell \right) \right] \\ &= 4h''(0) \left[\langle V_{kr}e_p^{(r)}, e_i^{(k)} \rangle_k \langle V_{\ell s}e_q^{(s)}, e_j^{(\ell)} \rangle_\ell \right] \\ &\quad + \langle V_{k\ell}e_j^{(\ell)}, e_i^{(k)} \rangle_k \langle V_{rs}e_q^{(s)}, e_p^{(r)} \rangle_p \\ &\quad + \langle V_{ks}e_q^{(s)}, e_i^{(k)} \rangle_k \langle V_{\ell r}e_p^{(r)}, e_j^{(\ell)} \rangle_\ell \Big]. \end{aligned}$$

Since, under \mathcal{H}_0 , we have for any $(m, l) \in \{1, \dots, K\}^2$, $V_{ml} = \delta_{ml}V_l = \delta_{ml}I_l$, where δ denotes the usual Kronecker symbol, we finally obtain:

$$\gamma_{ijpq}^{k\ell,rs} = 4h''(0) (\delta_{kr}\delta_{\ell s}\delta_{ip}\delta_{jq} + \delta_{k\ell}\delta_{rs}\delta_{ij}\delta_{pq} + \delta_{ks}\delta_{\ell r}\delta_{iq}\delta_{jp}). \quad (21)$$

If $k = r$, $\ell = s$ with $k \in \{2, \dots, K\}$ and $\ell \in \{1, \dots, k-1\}$, since $\ell \neq k$ and, equivalently, $s \neq r$, it follows that

$$\gamma_{ijpq}^{k\ell,rs} = 4h''(0)\delta_{ip}\delta_{jq} = \begin{cases} 4h''(0) & \text{if } i = j = p = q \\ 0 & \text{otherwise} \end{cases}.$$

Hence,

$$\Gamma_{k\ell, k\ell} = 4h''(0) \mathbb{I}_{p_k p_\ell}, \quad (22)$$

where \mathbb{I}_m denotes the m -dimensional identity matrix. If $(k, \ell) \neq (r, s)$ with $(k, r) \in \{2, \dots, K\}^2$ and $(\ell, s) \in \{1, \dots, k-1\} \times \{1, \dots, r-1\}$, suppose that $\gamma_{ijpq}^{k\ell,rs} \neq 0$. Then, since $\ell \neq k$, $s \neq r$ and $(k, \ell) \neq (r, s)$, we deduce from (21) that we necessarily have $s = k$ and $\ell = r$. Thus $k < r$ and, since $r = \ell$, we obtain the inequality $k < \ell$. That is not possible since $\ell \in \{1, \dots, k-1\}$. So, we deduce that $\gamma_{ijpq}^{k\ell,rs} = 0$. Using (19), (20) and (22) we can conclude that $\Gamma = 4h''(0) \mathbb{I}_d$. Then, we deduce from Theorem 4.1 that $\mathcal{Q} = 4h''(0)\mathcal{Q}'$, where \mathcal{Q}' is a random variable with distribution equal to χ_d^2 . \square

References

1. T.W. Anderson, Asymptotic theory for canonical analysis, J. Multivariate. Anal., 70 (1999), 1-29.
2. J. Dauxois, G.M. Nkiet, Y. Romain, Linear relative canonical analysis of euclidean random variables, asymptotic study and some applications, Ann. Inst. Statist. Math., 56 (2004) 279-304.
3. J. Dauxois, A. Pousse, Une extension de l'analyse canonique. Quelques applications, Ann. Inst. H. Poincaré, 11 (1975) 355-379.
4. J. Dauxois, Y. Romain, S. Viguier, Tensor product and statistics, Linear Algebra Appl., 270 (1994) 59-88.

5. S. Dossou-Gbete, A. Pousse, Asymptotic study of eigenelements of a sequence of random selfadjoint operators, *Statistics*, 22 (1991) 479-491.
6. M.L. Eaton, D. Tyler, On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix, *Ann. Statist.*, 19 (1991) 260-271.
7. M.L. Eaton, D. Tyler, The asymptotic distribution of singular values with applications to canonical correlations and correspondence analysis, *J. Multivariate. Anal.*, 70 (1994), 238-264.
8. L. Ferré, A.F. Yao, Functional sliced inverse regression analysis, *Statistics*, 37 (2003) 475-488.
9. J. Fine, Etude asymptotique de l'analyse canonique. *Pub. Inst. Statist. Univ. Paris*, 44 (2000) 21-72.
10. S. Gardner, J.C. Gower, N.J. Le Roux, A synthesis of canonical variate analysis, generalised canonical correlation and Procrustes analysis, *Comput. Statist. Data Anal.*, 50 (2006) 107-134.
11. A. Gifi, *Nonlinear Multivariate Analysis*, Wiley, Chichester, 1991.
12. J.R. Kettenring, Canonical analysis of several sets of variables, *Biometrika*, 58 (1971) 433-451.
13. R.J. Muirhead, C.M. Waternaux, Asymptotic distributions in canonical analysis and other multivariate procedures for nonnormal populations, *Biometrika*, 67 (1980) 31-43.
14. A. Pousse, Etudes asymptotiques, *in* "Modèles pour l'analyse des données multidimensionnelles", J.J. Droesbeke, B. Fichet and P. Tassi, eds), *Economica*, Paris, 1992.
15. Y. Takane, H. Hwang, H. Abdi, Regularized multiple-set canonical correlation analysis, *Psychometrika*, 73 (2008), 753-775.
16. A. Tenenhaus, M. Tenenhaus, Regularized generalized canonical correlation analysis, *Psychometrika*, 76 (2011), 257-284.