The Annals of Statistics
1998, Vol. 26, No. 4, 1254–1278

NONLINEAR CANONICAL ANALYSIS AND INDEPENDENCE TESTS

BY JACQUES DAUXOIS AND GUY MARTIAL NKIET

Université Paul Sabatier and Université de Masuku

Measures of association between two random variables (r.v.) that are symmetric nondecreasing functions of the canonical coefficients provided by the nonlinear canonical analysis of the r.v.'s are studied. These measures can be used to characterize independence. Their estimators are obtained by estimating a suitable approximation to nonlinear canonical analysis. When some conditions ae satisfied, asymptotic distributions of the estimators, both under the independence hypothesis and under dependence, are given. A class of tests of independence with asymptotic level of significance can be investigated.

1. Introduction. Several measures of association between two random variables (r.v.) based on the canonical analysis of these r.v.'s have been introduced in the literature, and a review of such measures can be found in Cramer and Nicewander (1979) and Lazraq and Cléroux (1988). A general enough study of these measures is proposed in Lin (1987) and, more recently, in Dauxois and Nkiet (1997), where a global study of the induced tests is developed. In these works the measures considered are constructed using linear canonical analysis (LCA) of the r.v.'s and, thus, characterize only lack of a linear relationship. It can be interesting to look for another class of measures of association which characterize independence.

The properties of nonlinear canonical analysis (NLCA) suggest that such a class can be derived from this analysis, but there is little work in this direction. The canonical coefficients from NLCA have already been used for testing independence, but uniquely when the related random variables are categorical [see, for instance, Tsai and Sen (1990)]. That is restrictive because one knows that, in this particular case, NLCA is a LCA of suitable random vectors, and thus it suffices to use classical measures of association. Consequently, it appears that the use of NLCA is more interesting when the related random variables are not categorical.

In this paper, we propose a class of measures of association between random variables with values in any measurable spaces, constructed using symmetric nondecreasing functions and the canonical coefficients derived from NLCA. The properties of these measures show that they are appropriate to characterize independence without any assumption on the distribution of

Received July 1995; revised March 1998.

AMS 1991 *subject classifications*. Primary 62H20; secondary 62H15, 62G10.

Key words and phrases. Measure of association, nonlinear canonical analysis, symmetric nondecreasing functions, approximating and estimating nonlinear canonical analysis, independence tests.

the random variables considered. Using results about estimation of NLCA [Dauxois and Pousse (1975), Lafaye de Micheaux (1978)] and asymptotic theory of canonical analysis [Dossou-Gbete and Pousse (1991), Pousse (1992)], a class of estimators is proposed for each measure, and the asymptotic properties of these estimators are given. This leads us to introduce a class of independence tests, and we show that this class includes tests based on the usual chi-squared index. A test of the introduced class is obtained from a few steps. First, the chosen infinite-dimensional measure of association (based on NLCA) is approximated by a finite-dimensional one (based on a suitable LCA); then the latter is estimated from an i.i.d. sample; then the test is based on a simple function of the estimate. In Section 6.4 simulations are furnished with the objective of evaluating performances of some tests of the introduced class and comparing them to classical ones. The results suggest that these tests are generally more powerful than classical tests and that the use of B-spline approximation in the estimation of NLCA induces more powerful tests than that based on the usual chi-squared index.

In order to emphasize the main points of the paper, certain proofs of lemmas and propositions are left to Section 7.

2. Nonlinear canonical analysis (NLCA) of random variables. We consider a probability space $(\Omega, \mathscr{A}, P)$ such that the Hilbert space $L^2(P)$ of random variables with finite second-order moment is separable.

Let X and Y be random variables defined on $(\Omega, \mathscr{A}, P)$, with values in measurable spaces $(\Omega_X, \mathscr{A}_X)$ and $(\Omega_Y, \mathscr{A}_Y)$, respectively, and with probability distribution measures denoted by P_X and P_Y. We denote by $L^2(P_X)$ the space of measurable real functions φ defined on Ω_X and such that $\mathbb{E}(\varphi^2(X)) < +\infty$, and by $L^2(P_Y)$ the space analogous to $L^2(P_X)$ with respect to Y.

Nonlinear canonical analysis (NLCA) is the search of two variables $f_1 = \varphi_1(X)\,[\varphi_1 \in L^2(P_X)]$ and $g_1 = \psi_1(Y)\,[\psi_1 \in L^2(P_Y)]$, such that the pair (f_1, g_1) maximizes $\langle f, g\rangle = \mathbb{E}(fg)$ under the constraints $\mathbb{E}(f^2) = \mathbb{E}(g^2) = 1$, with iterations under orthonormality constraints. This means that, for $i \geq 2$, one searches for two variables $f_i = \varphi_i(X)$ $[\varphi_i \in L^2(P_X)]$ and $g_i = \psi_i(Y)$ $[\psi_i \in L^2(P_Y)]$, such that the pair (f_i, g_i) is a solution for the above maximization problem with the additional constraints $\langle f, f_k\rangle = 0$, $\langle g, g_k\rangle = 0$, for all $k \in \{1, \dots, i-1\}$.

Considering the subspaces

$$H_X = \{\varphi(X);\, \varphi \in L^2(P_X)\} \quad \text{and} \quad H_Y = \{\psi(Y);\, \psi \in L^2(P_Y)\}$$

of $L^2(P)$, it is known [see Dauxois and Pousse (1975)] that the solution for the NLCA problem is obtained, for example, from spectral analysis of the (linear) self-adjoint operator $T = \mathbb{E}^X\mathbb{E}^Y_{|H_X}$, that is, the restriction of $\mathbb{E}^X\mathbb{E}^Y$ at H_X, where $\mathbb{E}^X$ and $\mathbb{E}^Y$ are the conditional expectations relative to X and Y, respectively. If T is a compact operator, NLCA exists and is characterized by a triple:

$$\{(\rho_i)_{i=0,\dots,N}, (\varphi_i(X))_{i=0,\dots,N_1}, (\psi_i(Y))_{i=0,\dots,N_2}\},$$

where N, N_1, N_2 are elements of $\mathbb{N} \cup \{+\infty\}$. In this triple, $(\rho_i^2)_{i=0,\dots,N}$ is the sequence of nonzero eigenvalues of T arranged in decreasing order and repeated according to multiplicity; the systems $(\varphi_i)_{i=0,\dots,N_1}$ and $(\psi_i)_{i=0,\dots,N_2}$ are orthonormal bases of $L^2(P_X)$ and $L^2(P_Y)$, respectively, satisfying

$$\forall\ i = 0,\dots,N, \qquad T(\varphi_i(X)) = \rho_i^2\varphi_i(X) \quad \text{and} \quad \psi_i(Y) = \rho_i^{-1}\mathbb{E}^Y(\varphi_i(X)).$$

The ρ_i's are termed the nonlinear canonical coefficients. Letting $\mathbb{1}_\Omega$ be the constant random variable with value equal to 1, one can see that $T(\mathbb{1}_\Omega) = \mathbb{1}_\Omega$. So the first canonical terms are $\rho_0 = 1$ and $\varphi_0(X) = \psi_0(Y) = \mathbb{1}_\Omega$ and are termed the trivial canonical terms.

As an important property of NLCA, one knows that X and Y are independent if, and only if, for each $i \neq 0$, one has $\rho_i = 0$. This shows the interest of constructing measures of association by the use of nonlinear canonical coefficients.

Remark 2.1. In fact, NLCA is a generalization of linear canonical analysis of random vectors $X = (X_1,\dots,X_p)'$ and $Y = (Y_1,\dots,Y_q)'$. Indeed, LCA of X and Y is the computation of the above variables f_i and g_i in the forms $f_i = \sum_{i=1}^p a_k^i X_k$ and $g_i = \sum_{i=1}^q b_k^i Y_k$. In LCA these variables are obtained from the spectral analysis of the finite rank operator $R = V_1^{-1/2}V_{12}V_2^{-1}V_{21}V_1^{-1/2}$, with $V_1 = \mathbb{E}(X \otimes X)$, $V_{12} = \mathbb{E}(Y \otimes X) = V_{21}^*$, $V_2 = \mathbb{E}(Y \otimes Y)$, where $\otimes$ denotes the tensor product between vectors (see Section 7.3). These operators are the classical ones with matricial expressions $\mathbb{E}(XX')$, $\mathbb{E}(YX')$ and $\mathbb{E}(YY')$, respectively; they are covariance operators when X and Y are centered random vectors. Throughout this paper we use operators and tensor products; for details about the matricial correspondences, one may refer to Dauxois, Romain and Viguier (1994).

Remark 2.2. Both LCA and NLCA are particular cases of canonical analysis of Hilbertian subspaces H_1 and H_2 of a real separable Hilbert space H with inner product $\langle \cdot,\cdot \rangle_H$. Canonical analysis of H_1 and H_2 is the computation of unit vectors $f_1 \in H_1$ and $g_1 \in H_2$ that maximize $\langle f_1, g_1 \rangle_H$ with iterations under orthonormality constraints. The solution for this problem is obtained from spectral analysis of several operators. One of them is the self-adjoint operator $T = \Pi_{H_1}\Pi_{H_2 | H_1}$, where Π_E stands for the orthogonal projector onto the closed subspace E of H. Clearly, one can also use the operator $\Pi_{H_1}\Pi_{H_2}$. When H is the space $L^2(P)$, LCA (resp. NLCA) is obtained by taking $H_1 = \text{span}(X_1,\dots,X_p)$ and $H_2 = \text{span}(Y_1,\dots,Y_q)$ (resp. $H_1 = H_X$ and $H_2 = H_Y$).

The trivial canonical terms $\rho_0 = 1$ and $\varphi_0(X) = \psi_0(Y) = \mathbb{1}_\Omega$ do not give any information about independence of X and Y. They can be avoided by restricting NLCA to centered variables. Let us consider the spaces

$$H'_X = \{f \in H_X \mid \mathbb{E}(f) = 0\},$$
$$H'_Y = \{g \in H_Y \mid \mathbb{E}(g) = 0\};$$

one can verify that, $\forall f \in H'_X$, $\mathbb{E}^X \mathbb{E}^Y_{|H_X}(f) = \Pi_{H'_X} \Pi_{H'_Y \mid H_X}(f)$. Canonical analysis of H'_X and H'_Y, which is called centered NLCA of X and Y, provide the same canonical terms as NLCA of X and Y except for the trivial terms. Since the spectral decomposition of T is

$$T = \sum_{i=0}^{N} \rho_i^2 \varphi_i(X) \otimes \varphi_i(X) = \mathbb{1}_\Omega \otimes \mathbb{1}_\Omega + \sum_{i=1}^{N} \rho_i^2 \varphi_i(X) \otimes \varphi_i(X),$$

it is clear that centered NLCA of X and Y is obtained from spectral analysis of $S = T - \mathbb{1}_\Omega \otimes \mathbb{1}_\Omega$.

REMARK 2.3. One can also obtain NLCA of X and Y from spectral analysis of $T' = \mathbb{E}^Y \mathbb{E}^X_{|H_Y}$. This operator has the same eigenvalues as T; for all $i = 0, \dots, N$, the variable $\psi_i(Y)$ is a unitary eigenvector of T' associated with ρ_i^2 and $\varphi_i(X) = \rho_i^{-1} \mathbb{E}^X(\psi_i(Y))$. This shows the symmetry of NLCA.

REMARK 2.4. Generally, one has $\lim_{i \to +\infty}(\rho_i) = 0$. So the decreasing complete sequence $\lambda = (\rho_i^2)_{i=0,\dots,N}$ of eigenvalues of T is an element of the space c_0 of numerical sequences $(u_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to +\infty} u_n = 0$. Nevertheless, if there exists a measurable function f such that $X = f(Y)$, then H_X is included in H_Y; thus, all the canonical coefficients are equal to 1 and the previous property does not hold. In this latter case, we set $\lambda = \mathbb{1}$, where $\mathbb{1}$ denotes the numerical sequence for which all the terms are equal to 1.

3. Measures of association based on NLCA. Several measures of association in the literature are constructed as functions of the canonical coefficients deriving from LCA. They have the form $\Phi(\rho_1^2, \dots, \rho_s^2)$, where Φ is a symmetric nondecreasing function [Jensen and Mayer (1977)] and the ρ_i's are the aforementioned canonical coefficients.

We want to define measures of association constructed analogously by the use of canonical coefficients deriving from centered NLCA. Since the sequence of these coefficients belongs to c_0, we first need to extend the definition of symmetric nondecreasing functions so that their definition domain could be subsets of c_0.

3.1. *Symmetric nondecreasing functions defined on subsets of c_0.*

DEFINITION 3.1. A symmetric nondecreasing function is a real function Φ defined on a subset $\mathscr{D}_\Phi$ of c_0 such that the following hold:

1. For each $x = (x_n)_{n \in \mathbb{N}} \in \mathscr{D}_\Phi$ and any permutation $\sigma: \mathbb{N} \to \mathbb{N}$, the sequence $x_\sigma = (x_{\sigma(n)})_{n \in \mathbb{N}}$ belongs to $\mathscr{D}_\Phi$ and $\Phi(x) = \Phi(x_\sigma)$.
2. For each $x = (x_n)_{n \in \mathbb{N}} \in \mathscr{D}_\Phi$ and each $y = (y_n)_{n \in \mathbb{N}} \in \mathscr{D}_\Phi$ such that $\forall n \in \mathbb{N}$, $|x_n| \le |y_n|$, one has $\Phi(x) \le \Phi(y)$.
3. There exists a nondecreasing real function f_Φ with definition domain $\{x \in \mathbb{R} \mid (x, 0, \dots) \in \mathscr{D}_\Phi\}$ containing 0, such that

$$\Phi(x, 0, \dots) = f_\Phi(|x|) \quad \text{and} \quad \big(f_\Phi(x) = 0 \Leftrightarrow x = 0\big).$$

Clearly, any symmetric nondecreasing function Φ is a nonnegative map and one has

$$\Phi(x) = 0 \quad \Leftrightarrow \quad \forall\, n \in \mathbb{N}, \qquad x_n = 0. \tag{3.1}$$

Although the sequence $\mathbb{1} = (1, 1, \dots)$ does not belong to c_0, we can extend the domain of some symmetric nondecreasing functions at this point. Indeed, let $\mathbb{1}^{(n)} = (1, \dots, 1, 0, \dots)$ be the sequence of c_0 having 1 at the n first places and 0 elsewhere. Suppose that, for each $n \in \mathbb{N}^*$, $\mathbb{1}^{(n)}$ belongs to $\mathscr{D}_\Phi$ and that the nondecreasing sequence $(\Phi(\mathbb{1}^{(n)}))_{n \in \mathbb{N}^*}$ is bounded, we can set $\Phi(\mathbb{1}) = \lim_{n \to +\infty} \Phi(\mathbb{1}^{(n)})$.

Example 3.1. The symmetric norming functions Φ_p [Gohberg and Krejn (1971)] defined by

$$\Phi_p(x) = \Big(\sum_{n \in \mathbb{N}} |x_n|^p \Big)^{1/p}, \qquad 1 \le p < +\infty,$$

and

$$\Phi_\infty(x) = \max_{n \in \mathbb{N}} |x_n|$$

are examples of symmetric nondecreasing functions. If p is finite, the natural domain of Φ_p is the space l^p of real sequences $x = (x_n)_{n \in \mathbb{N}}$ such that $\sum_{n \in \mathbb{N}} |x_n|^p < +\infty$, and if $p = +\infty$, it is the space c_0. The definition domain of Φ_∞ is extendible to $\mathbb{1}$ as previously indicated and one has $\Phi_\infty(\mathbb{1}) = 1$.

3.2. *Definition and properties of measures of association based on NLCA.* Here, we propose a class of measures of association constructed through the canonical coefficients of centered NLCA.

Let (X, Y) be a pair of random variables having a NLCA. We consider a symmetric nondecreasing function Φ such that $\Phi(\mathbb{1}) = 1$; we assume that the decreasing sequence λ of squares of coefficients derived from the centered NLCA of X and Y belongs to $\mathscr{D}_\Phi \cup \{\mathbb{1}\}$. Thus we can set $r_\Phi(X, Y) = \Phi(\lambda)$. So we define a measure of association r_Φ and we have the following proposition.

Proposition 3.1. *The following properties hold for* r_Φ:

1. $r_\Phi(X, Y) = r_\Phi(Y, X)$;
2. *we have* $r_\Phi(X, Y) = 0$ *if, and only if,* X *and* Y *are independent random variables*;
3. *if for all real measurable functions* φ *defined on* Ω_X *there exists a real measurable function* ψ *defined on* Ω_Y *such that, almost surely,* $\varphi(X) = \Psi(Y)$, *then we have* $r_\Phi(X, Y) = 1$;
4. *if* φ *and* ψ *are bijective bimeasurable functions defined on* Ω_X *and* Ω_Y, *respectively, we have* $r_\Phi(\varphi(X), \psi(Y)) = r_\Phi(X, Y)$;
5. *when* (X, Y) *has the bivariate standard normal distribution with correlation coefficient* ρ, *there exists a nondecreasing real function* h_Φ *such that* $r_\Phi(X, Y) = h_\Phi(\rho^2)$.

PROOF. Property 1 is an obvious consequence of the NLCA symmetry. Using (3.1), we obtain 2.

If for all real measurable functions φ defined on Ω_X there exists a real measurable function ψ defined on Ω_Y such that, almost surely, $\varphi(X) = \psi(Y)$, then H_X is included in H_Y. Since $\lambda = \mathbb{1}$, we have $r_\Phi(X, Y) = 1$.

Property 4 comes from the fact that, putting $U = \varphi(X)$ and $V = \Psi(Y)$ and considering H_U and H_V the analogues of H_X with respect to U and V, respectively, we have $H_U = H_X$ and $H_V = H_Y$.

When (X, Y) has the bivariate standard normal distribution with correlation coefficient ρ, then [see Dauxois and Pousse (1975)] the NLCA of X and Y is the triple

$$\left\{(\rho^i)_{i \in \mathbb{N}}, (\phi_i(X))_{i \in \mathbb{N}}, (\phi_i(Y))_{i \in \mathbb{N}}\right\},$$

where the ϕ_i's are the Hermite polynomials. Hence

$$r_\Phi(X, Y) = \Phi(\rho^2, \rho^4, \rho^6, \dots) = h_\Phi(\rho^2),$$

where $h_\Phi(x) = \Phi(|x|, |x|^2, |x|^3, \dots)$ defines a nondecreasing function from $[0, +\infty[$ to itself. $\square$

REMARK 3.1. The previous properties of $r_\Phi(X, Y)$ are very close to conditions proposed by Rényi (1959) for good measures of dependence. In particular, property 2 shows that the measures of association considered here are appropriate to evaluate independence without any assumption about the distribution of the pair (X, Y).

We now provide some examples of measures of association which can be constructed using centered NLCA.

EXAMPLE 3.2. Let Φ be the symmetric nondecreasing function (with definition domain l^p) given by $\Phi(x) = 1 - \exp(-\sum_{n=1}^{+\infty} |x_n|^p)$. One has $\Phi(\mathbb{1}) = 1$. It leads to the measure of association

$$r_{1,p}(X, Y) = 1 - \exp\left(-\sum_{n=1}^{+\infty} \rho_n^{2p}\right).$$

EXAMPLE 3.3. Putting

$$\Phi(x) = \sqrt{\frac{\sum_{n=1}^{+\infty} |x_n|^p}{1 + \sum_{n=1}^{+\infty} |x_n|^p}},$$

we obtain the measure of association

$$r_{2,p}(X, Y) = \sqrt{\frac{\sum_{n=1}^{+\infty} \rho_n^{2p}}{1 + \sum_{n=1}^{+\infty} \rho_n^{2p}}}.$$

EXAMPLE 3.4. The symmetric nondecreasing function $\Phi_\infty(x) = \sqrt{\max_{n\in\mathbb{N}^*}|x_n|}$ permits us to obtain, as a measure of association, the Rényi maximal coefficient. Indeed, one has $r_\infty(X,Y) = \Phi_\infty(\lambda) = \max_{n\in\mathbb{N}^*}|\rho_n| = \rho_1$.

EXAMPLE 3.5. If (X,Y) has the bivariate standard normal distribution with correlation coefficient ρ, then one has $\rho_n = \rho^n$, $\forall\ n \in \mathbb{N}$. So, if $|\rho| < 1$, we obtain

$$r_{1,p}(X,Y) = 1 - \exp\left(-\frac{\rho^{2p}}{1-\rho^{2p}}\right),$$

$$r_{2,p}(X,Y) = |\rho|^p \quad\text{and}\quad r_\infty(X,Y) = |\rho|.$$

4. Approximating measures of association. Although NLCA can be determined in the case of (X,Y) having a bivariate normal distribution and in other special cases [see Lancaster (1969), Dauxois and Pousse (1975) and Buja (1990)], it is generally impossible to determine and it is the same for measures of association using NLCA coefficients. Nevertheless, since there exist convergent approximations for NLCA, it is possible to deduce approximations for the measures of association.

4.1. *A general approach to approximating NLCA*. For all $n \in \mathbb{N}$, we denote by $\mathscr{V}_X^n$ (resp. $\mathscr{V}_Y^n$) the subspace spanned by a linearly independent system $(\varphi_i^n)_{1\le i\le p_n}$ [resp. $(\psi_i^n)_{1\le i\le q_n}$] in $L^2(P_X)$ [resp. $L^2(P_Y)$].

We assume that the sequences $(\mathscr{V}_X^n)_{n\in\mathbb{N}}$ and $(\mathscr{V}_Y^n)_{n\in\mathbb{N}}$ are nondecreasing, that is

$$\forall\ n \in \mathbb{N},\qquad \mathscr{V}_X^n \subset \mathscr{V}_X^{n+1},\qquad \mathscr{V}_Y^n \subset \mathscr{V}_Y^{n+1},$$

and that their unions $\bigcup_n \mathscr{V}_X^n$ and $\bigcup_n \mathscr{V}_Y^n$ are dense in $L^2(P_X)$ and $L^2(P_Y)$, respectively. Then [see Dauxois and Pousse (1977) and Lafaye de Micheaux (1978)] the canonical analysis of the subspaces (cf. Remark 2.2)

$$\mathscr{W}_X^n = \operatorname{span}\big(\varphi_1^n(X),\dots,\varphi_{p_n}^n(X)\big)$$

and

$$\mathscr{W}_Y^n = \operatorname{span}\big(\psi_1^n(Y),\dots,\psi_{q_n}^n(Y)\big)$$

converges, as $n \to +\infty$, to the NLCA of X and Y; that is, the sequence of operators $\Pi_{\mathscr{W}_X^n}\Pi_{\mathscr{W}_Y^n}$ converges uniformly to $\mathbb{E}^X\mathbb{E}^Y$. Clearly, this canonical analysis is the LCA of random vectors

$$\varphi^n(X) = \big(\varphi_1^n(X),\dots,\varphi_{p_n}^n(X)\big)'$$

and

$$\psi^n(Y) = \big(\psi_1^n(Y),\dots,\psi_{q_n}^n(Y)\big)'.$$

Considering the operators

$$V_1^n = \mathbb{E}(\varphi^n(X)\otimes\varphi^n(X)),\qquad V_2^n = \mathbb{E}(\psi^n(Y)\otimes\psi^n(Y))$$

and

$$V_{12}^n = \mathbb{E}(\psi^n(Y)\otimes\varphi^n(X)) = (V_{21}^n)^*,$$

this LCA is obtained from the spectral analysis of

$$R_n = (V_1^n)^{-1/2} V_{12}^n (V_2^n)^{-1} V_{21}^n (V_1^n)^{-1/2}.$$

Let us remark that since $\{\varphi_1^n(X), \dots, \varphi_{p_n}^n(X)\}$ and $\{\psi_1^n(Y), \dots, \psi_{q_n}^n(Y)\}$ are systems consisting of linearly independent random variables, V_1^n and V_2^n are invertible and R_n is well defined.

From now on we assume that there exist vectors $\alpha^n = (\alpha_1^n, \dots, \alpha_{p_n}^n)'$ and $\beta^n = (\beta_1^n, \dots, \beta_{q_n}^n)'$ such that

$$\begin{aligned} &\forall (x, y) \in \Omega_X \times \Omega_Y, \qquad \langle \alpha^n, \varphi^n(x) \rangle_{\mathbb{R}^{p_n}} = \sum_{i=1}^{p_n} \alpha_i^n \varphi_i^n(x) = 1, \\ &\qquad\qquad\qquad\qquad\qquad\quad \langle \beta^n, \psi^n(x) \rangle_{\mathbb{R}^{q_n}} = \sum_{i=1}^{q_n} \beta_i^n \psi_i^n(x) = 1, \end{aligned} \tag{4.1}$$

where, for all $p \in \mathbb{N}^*$, $\langle \cdot, \cdot \rangle_{\mathbb{R}^p}$ stands for the usual inner product of $\mathbb{R}^p$. This means that $\mathbb{1}_\Omega$ belongs to $\mathscr{W}_X^n \cap \mathscr{W}_Y^n$ and thus, the greatest eigenvalue of T_n is $(\rho_0^n)^2 = 1$, and it is associated with the canonical pair (f_0^n, g_0^n), where f_0^n and g_0^n are both equal to $\mathbb{1}_\Omega$. Hence, the vector $e_0^n = \alpha^n$ is an obvious eigenvector of R_n associated with $(\rho_0^n)^2 = 1$, and it is easily seen that it is a unit eigenvector. From the spectral decomposition $R_n = \sum_{i=0}^{p_n - 1} (\rho_i^n)^2 e_i^n \otimes e_i^n$, we deduce that, in order to approximate the centered NLCA, we may consider the spectral analysis of $S_n = R_n - \alpha^n \otimes \alpha^n$.

Finally, the centered NLCA of X and Y is approximated by a sequence of suitable LCA's. To define this sequence, for each $n \in \mathbb{N}$ we first choose functions φ_i^n, $1 \le i \le p_n$, and ψ_i^n, $1 \le i \le q_n$, belonging to $L^2(P_X)$ and $L^2(P_Y)$, respectively, such that the spaces $\mathscr{V}_X^n$ and $\mathscr{V}_Y^n$ spanned by those functions have the previously mentioned properties; then we consider the nontrivial terms of the LCA of the random vectors $\varphi^n(X)$ and $\psi^n(Y)$.

The functions φ_i^n and ψ_i^n can be obtained from nondecreasing sequences of partitions, say σ_X^n and σ_Y^n, of Ω_X and Ω_Y, respectively.

We now exhibit examples of such constructions. The first example comes from Dauxois and Pousse (1975) and the second from Lafaye de Micheaux (1978); the case of random vectors may be easily given.

Example 4.1 (Step functions). Let $\sigma_X^n = \{J_1^n, \dots, J_{p_n}^n\}$ and $\sigma_Y^n = \{K_1^n, \dots, K_{q_n}^n\}$ define nondecreasing sequences of partitions of Ω_X and Ω_Y, respectively. Putting $p_{i\cdot}^n = P(X \in J_i^n)$ and $p_{\cdot j}^n = P(Y \in K_j^n)$, let us consider the functions $\varphi_i^n = (p_{i\cdot}^n)^{-1/2} \mathbb{1}_{J_i^n}$ and $\psi_j^n = (p_{\cdot j}^n)^{-1/2} \mathbb{1}_{K_j^n}$, where, for all sets A, $\mathbb{1}_A$ stands for its characteristic function. The sequences $(\mathscr{V}_X^n)_{n \in \mathbb{N}}$ and $(\mathscr{V}_Y^n)_{n \in \mathbb{N}}$ are nondecreasing and their unions are dense in $L^2(P_X)$ and $L^2(P_Y)$, respectively. The operator R_n admits the matricial expression

$$\left(\frac{1}{\sqrt{p_{i\cdot}^n p_{j\cdot}^n}} \sum_{k=1}^{q_n} \frac{p_{ik}^n p_{jk}^n}{p_{\cdot k}^n} \right)_{1 \le i, j \le p_n}$$

with $p_{ij}^n = P(X \in J_i^n; Y \in K_j^n)$. Thus, the related NLCA approximation is the sequence of correspondence analyses of the contingency tables (p_{ij}^n). Since

(4.1) is satisfied with

$$\alpha^n = \left(\sqrt{p^n_{1\cdot}}, \dots, \sqrt{p^n_{p_n\cdot}}\right) \quad \text{and} \quad \beta^n = \left(\sqrt{p^n_{\cdot 1}}, \dots, \sqrt{p^n_{\cdot q_n}}\right),$$

the operator S_n admits the matricial expression

$$s^n_{ij} = \frac{1}{\sqrt{p^n_{i\cdot} p^n_{j\cdot}}} \sum_{k=1}^{q_n} \frac{p^n_{ik}\, p^n_{jk}}{p^n_{\cdot k}} - \sqrt{p^n_{i\cdot} p^n_{j\cdot}}\,;$$

by a simple calculation, we obtain

$$s^n_{ij} = \frac{1}{\sqrt{p^n_{i\cdot} p^n_{j\cdot}}} \sum_{k=1}^{q_n} \frac{\left(p^n_{ik} - p^n_{i\cdot} p^n_{\cdot k}\right)\left(p^n_{jk} - p^n_{j\cdot} p^n_{\cdot k}\right)}{p^n_{\cdot k}}. \tag{4.2}$$

EXAMPLE 4.2 (B-splines). Let X and Y be two real random variables. For all $n \in \mathbb{N}^*$, let us consider dyadic subdivision of the interval $[-n, n)$. We obtain a partition σ^n_X $(= \sigma^n_Y)$ of $\mathbb{R}$ consisting of the intervals $(-\infty, -n)$, $[n, +\infty)$ and all the dyadic intervals $[2^{-n}(k-1), 2^{-n}k)$ which subdivide $[-n, n)$. For $s \geq 1$, we consider the B-spline functions of order s computed on the above partition of $\mathbb{R}$. The subspaces $\mathscr{V}^n_X$ $(= \mathscr{V}^n_Y)$ spanned by these functions satisfy the required properties [see Lafaye de Micheaux (1978)]. So they can be used for approximating centered NLCA as indicated in this section.

4.2. *Approximation of measures of association*. Now, given a symmetric nondecreasing function Φ, we want to approximate the measure of association defined by $r_\Phi(X, Y) = \Phi(\lambda)$ by a suitable numerical sequence. This is possible by using the previous results on NLCA approximation.

When subspaces $\mathscr{V}^n_X$ and $\mathscr{V}^n_Y$ having the required properties have been chosen, we denote by λ_n the sequence (regarded as an element of $\mathbb{R}^{p_n}$) of eigenvalues of S_n arranged in decreasing order and repeated according to multiplicity.

Moreover, we consider the restriction of Φ at $\mathbb{R}^{p_n}$, that is, the real function defined as

$$\Phi_n(x_1, \dots, x_{p_n}) = \Phi(x_1, \dots, x_{p_n}, 0, \dots)$$

with definition domain

$$\mathscr{D}_{\Phi_n} = \left\{x \in \mathbb{R}^{p_n};\ (x_1, \dots, x_{p_n}, 0, \dots) \in \mathscr{D}_\Phi\right\}.$$

Assuming that each λ_n belongs to $\mathscr{D}_{\Phi_n}$, we put $r^n_\Phi(X, Y) = \Phi_n(\lambda_n)$. Then we have the following proposition.

PROPOSITION 4.1. *If Φ is continuous, the numerical sequence $(r^n_\Phi(X, Y))_{n \in \mathbb{N}^*}$ converges, as $n \to +\infty$, to $r_\Phi(X, Y)$.*

REMARK 4.1. Let us recall that c_0 is a Banach space with norm

$$\|x\|_\infty = \max_{n \in \mathbb{N}} |x_n|.$$

The previous continuity hypothesis of Φ is relative to this norm. This result shows that we obtain an approximation of the measure of association $r_\Phi(X, Y)$ by the sequence $r_\Phi^n(X, Y)$.

We now consider an example of such an approximation.

EXAMPLE 4.3. Considering $r(X, Y) = 1 - \exp(-\sum_{i=1}^{+\infty} (\rho_i^n)^2)$ (see Example 3.2), if the subspaces $\mathscr{V}_X^n$ and $\mathscr{V}_Y^n$ as in Example 4.1 are chosen, we have to approximate $r(X, Y)$ by the sequence $(r_n(X, Y))_{n \in \mathbb{N}^*}$ such that $r_n(X, Y) = 1 - \exp(-\sum_{n=1}^{p_n - 1} \rho_n^2) = 1 - \exp(-\mathrm{tr}(S_n))$. Using (4.2), we obtain

$$r_n(X, Y) = 1 - \exp\left(- \sum_{i=1}^{p_n} \sum_{j=1}^{p_n} \frac{\left(p_{ij}^n - p_{i\cdot}^n p_{\cdot j}^n\right)^2}{p_{i\cdot}^n p_{\cdot j}^n} \right);$$

so this measure is based on the usual chi-squared index.

5. Estimating a measure of association. For any fixed sufficiently large n, we consider an approximation of $r_\Phi(X, Y)$ provided by $r_\Phi^n(X, Y)$ obtained as specified in the previous section.

Let $\{(X_i, Y_i)\}_{1 \le i \le m}$ be an i.i.d. sample of size m, where each pair (X_i, Y_i) has the same distribution as (X, Y). The aim of this section is to show how to estimate the approximation of $r_\Phi(X, Y)$ and to underline the asymptotic properties of the related estimator.

5.1. *An estimator for* $r_\Phi^n(X, Y)$. A natural way to estimate $r_\Phi^n(X, Y)$ is to take an estimator $S_{n,m}$ of S_n and to consider the random variable $r_\Phi^{n,m}(X, Y) = \Phi_n(\lambda_{n,m})$, where $\lambda_{n,m} = (\lambda_i^{n,m})_{1 \le i \le p_n}$ is the sequence (regarded as an element of $\mathbb{R}^{p_n}$) of eigenvalues of $S_{n,m}$ arranged in decreasing order and repeated according to multiplicity.

Let us consider the following random operators:

$$V_1^{n,m} = \frac{1}{m} \sum_{i=1}^m \varphi^n(X_i) \otimes \varphi^n(X_i), \qquad V_2^{n,m} = \frac{1}{m} \sum_{i=1}^m \psi^n(Y_i) \otimes \psi^n(Y_i)$$

and

$$V_{12}^{n,m} = \frac{1}{m} \sum_{i=1}^m \psi^n(Y_i) \otimes \varphi^n(X_i) = (V_{21}^{n,m})^*.$$

We obtain an estimator of R_n by putting

$$R_{n,m} = (V_1^{n,m})^{-1/2} V_{12}^{n,m} (V_2^{n,m})^{-1} V_{21}^{n,m} (V_1^{n,m})^{-1/2}.$$

By the law of large numbers, $V_1^{n,m}$ (resp. $V_2^{n,m}$, $V_{12}^{n,m}$) converges almost surely, uniformly, as $m \to +\infty$, to V_1^n (resp. V_2^n, V_{12}^n). Since V_1^n and V_2^n are invertible, this shows that $R_{n,m}$ is well defined for a sufficiently large m and that $R_{n,m}$ converges almost surely, uniformly, as $m \to +\infty$, to R_n.

REMARK 5.1. In the literature there exist methods which permit us to estimate the canonical functions φ_1 and ψ_1 associated with the greatest nontrivial canonical coefficient ρ_1 [see, e.g., Breiman and Friedman (1985)]. The method introduced here may not be used for this estimation problem. Indeed, the fact that a sequence (A_n) of operators converges uniformly to an operator A does not generally ensure the convergence of the kth eigenvector of A_n to that of A [see Dunford and Schwartz (1963)]. So the canonical functions obtained from the spectral analysis $S_{n,m}$ are not convergent estimators of the φ_i's and ψ_i's.

For determining an estimator of S_n we must first look for a trivial eigenvector of $R_{n,m}$ associated with the eigenvalue 1.

LEMMA 5.1. *Almost surely*, $(V_1^{n,m})^{1/2}\alpha^n$ *is a unit eigenvector of* $R_{n,m}$ *associated with the eigenvalue* 1.

This lemma shows that

$$S_{n,m} = R_{n,m} - \left[(V_1^{n,m})^{1/2} \alpha^n \right] \otimes \left[(V_1^{n,m})^{1/2} \alpha^n \right]$$

may be an estimator of S_n. From the previous convergence results, it is a convergent estimator.

Putting $r_\Phi^{n,m}(X,Y) = \Phi_n(\lambda_{n,m})$, we obtain the following proposition.

PROPOSITION 5.1. *If* Φ *is a continuous symmetric nondecreasing function, then, almost surely,* $r_\Phi^{n,m}(X,Y)$ *converges as* $m \to +\infty$ *to* $r_\Phi^n(X,Y)$.

PROOF. Since $S_{n,m}$ converges almost surely and uniformly to S_n, there exists [see Dunford and Schwartz (1963)] a permutation σ of $\{1,\dots,p_n\}$ such that, for all $i \in \{1,\dots,p_n\}$, $\lambda_{\sigma(i)}^{n,m}$ converges almost surely to λ_i^n. Let us put $\lambda_\sigma^{n,m} = (\lambda_{\sigma(i)}^{n,m})_{1 \le i \le p_n}$. Then by the continuity of Φ_n, which comes from that of Φ, $\Phi_n(\lambda_\sigma^{n,m})$ converges almost surely as $m \to +\infty$ to $\Phi_n(\lambda_n)$. By the symmetry of Φ_n, one has $\Phi_n(\lambda_\sigma^{n,m}) = \Phi_n(\lambda_{n,m})$ and the result follows. □

REMARK 5.2. In practice, for computing an estimate of the measure of association when data $(x_k, y_k)_{1 \le k \le m}$ are observed, one must first transform these data by the chosen functions φ_i^n and ψ_i^n so as to obtain two matrices $(\varphi_i^n(x_k))_{k,i}$ and $(\psi_i^n(y_k))_{k,i}$. Then, after centering the columns of these matrices, one has to calculate the usual canonical analysis from them. So the required estimate is $\Phi_n(l)$, where l is the sequence of canonical correlation coefficients.

5.2. *Asymptotic distribution of the estimators*. Here we determine the asymptotic distribution of $r_\Phi^{n,m}(X,Y)$, both in the general case and when X and Y are independent random variables.

From now on we assume that the considered symmetric nondecreasing function Φ_n is twice differentiable in an open set $\mathscr{O}$ containing λ and each λ_n, and that there exists a real $M_{\Phi_n} > 0$ such that, for any $x \in \mathscr{O}$, the second-order differential $D^2\Phi_n(x)$ at x satisfies $\|D^2\Phi_n(x)\| \leq M_{\Phi_n}$ (of course, $\|\cdot\|$ is the norm of continuous bilinear forms).

Moreover, let $\lambda_1^{n*}, \dots, \lambda_{s_n}^{n*}$ be the sequence of eigenvalues of S_n arranged in strictly decreasing order ($\lambda_1^{n*} > \cdots > \lambda_{s_n}^{n*}$). For each $i = 1, \dots, s_n$ we set $\mathscr{I}_i^n = \{j;\ \lambda_j^n = \lambda_i^{n*}\}$; from the properties of Φ_n (see Section 7.4), $(\partial\Phi/\partial x_k)(\lambda_n)$ is constant for $k \in \mathscr{I}_i^n$. Let $K_{\Phi_n}^i(\lambda_n)$ be this constant. We assume that there exists $i \in \{1, \dots, s_n\}$ such that $K_{\Phi_n}^i(\lambda_n) \neq 0$. Then we have the following proposition.

PROPOSITION 5.2. *$\sqrt{m}\,(r_\Phi^{n,m}(X,Y) - r_\Phi^n(X,Y))$ converges in distribution, as $m \to +\infty$, to a centered Gaussian random variable with variance $\sigma_{\Phi_n}^2(\lambda_n) = \sum_{i=1}^{s_n}\sum_{j=1}^{s_n} K_{\Phi_n}^i(\lambda_n)K_{\Phi_n}^j(\lambda_n)\sigma_{ij}^n$, where σ_{ij}^n is a real number depending on the canonical terms provided by the spectral analysis of S_n.*

The entire expression of σ_{ij}^n can be found in (7.4).

If X and Y are independent random variables, then $\lambda_n = 0$ and $\sigma_{\Phi_n}^2(\lambda_n) = 0$ (see Section 7.6). So $\sqrt{m}\,(r_\Phi^{n,m}(X,Y) - r_\Phi^n(X,Y))$ converges in probability, as $m \to +\infty$, to 0. Nevertheless, assuming that the constant $K_{\Phi_n} = (\partial\Phi/\partial x_k)(0)$ is different from 0, we have the following proposition.

PROPOSITION 5.3. *When X and Y are independent, $mK_{\Phi_n}^{-1}r_\Phi^{n,m}(X,Y)$ converges in distribution, as $m \to +\infty$, to $\chi^2_{(p_n-1)(q_n-1)}$.*

REMARK 5.3. It is interesting to note that these results hold whatever type of approximation has been chosen for approximating NLCA. Thus, this lemma is a generalization of the properties of the chi-squared index estimator.

6. A class of independence tests.

These results permit us to construct a class of tests based on measures of association described above.

6.1. *Constructing the test*. The independence hypothesis of X and Y, which is denoted by H_0, is equivalent to the fact that $r_\Phi(X,Y) = \Phi(\lambda) = 0$, where Φ is a suitable symmetric nondecreasing function. Since we have, in practice, an approximation of the measure $r_\Phi(X,Y)$, we replace the latter hypothesis by $r_\Phi^n(X,Y) = \Phi_n(\lambda_n) = 0$ (this new hypothesis is denoted by $H_{0,\Phi}^n$), where n is assumed to be sufficiently large. So we can choose $r_\Phi^{n,m}(X,Y) = \Phi_n(\lambda_{n,m})$ as a test statistic. Of course, concretely, this test works with $\lambda_{n,m}$; so it will be a test with an asymptotic level of significance.

The limiting distributions of $r_\Phi^{n,m}(X,Y)$ under the null hypothesis and its alternative $r_\Phi^n(X,Y) \neq 0$ come from Propositions 5.2 and 5.3, respectively.

The approximated critical region (for a significance level α) is

$$C_{\Phi_n}^{n,m}(\alpha) = \left\{K_{\Phi_n}^{-1} r_{\Phi}^{n,m}(X,Y) > m^{-1} t_{\alpha}^{n}\right\},$$

where t_α^n is such that $P(\chi^2_{(p_n-1)(q_n-1)} > t_\alpha^n) = \alpha$.

6.2. *Convergence of the test*. The following proposition shows that the test is convergent.

PROPOSITION 6.1. *Under hypothesis* $H_{1,\Phi}^n$: $r_\Phi^n(X,Y) > 0$, *for each* $\alpha \in [0,1]$, *one has* $\lim_{m\to+\infty} P(C_{\Phi_n}^{n,m}(\alpha)) = 1$.

EXAMPLE 6.1. Let Φ be the symmetric nondecreasing function $\Phi(x) = 1 - \exp(-\sum_{n=1}^{+\infty} |x_n|)$. Considering the approximation of centered NLCA obtained from step functions (see Example 4.2), an estimator of the related measure of association is given by

$$r_\Phi^{n,m}(X,Y) = 1 - \exp\left(-\sum_{k=1}^{p_n} (\rho_k^{n,m})^2\right) = 1 - \exp(-\operatorname{tr}(S_{n,m})).$$

By a simple calculation, the expression of $\operatorname{tr}(S_{n,m})$ may be given and the previous relation becomes

$$r_\Phi^{n,m}(X,Y) = 1 - \exp\left(-\sum_{i=1}^{p_n}\sum_{j=1}^{q_n} \frac{\left(p_{ij}^{n,m} - p_{i\cdot}^{n,m} p_{\cdot j}^{n,m}\right)^2}{p_{i\cdot}^{n,m} p_{\cdot j}^{n,m}}\right),$$

where $p_{ij}^{n,m}$, $p_{i\cdot}^{n,m}$ and $p_{\cdot j}^{n,m}$ denote the usual estimators (frequencies) of the probabilities p_{ij}^n, $p_{i\cdot}^n$ and $p_{\cdot j}^n$, respectively, derived from a sample of size m.

So we obtain a test based on the chi-squared independence test statistic.

6.3. *Simulations*. In this section, we illustrate the previous procedure for testing independence by applying it to various data sets. In order to evaluate performance on finite samples, the procedure is applied to simulated data from bivariate random variables (X, Y) with known distributions.

The objective is to estimate the powers of some tests of our class and to compare these powers to those of classical independence tests. In all the examples, we considered the symmetric nondecreasing function

$$\Phi(x) = \frac{\sum_i |x_i|}{1 + \sum_i |x_i|}.$$

The estimates of the related measure of association were computed by using the NLCA approximations provided by step functions (see Example 4.1) and B-spline functions of orders 2 and 3. All these functions were obtained from a partition of $\mathbb{R}$ using dyadic intervals as in Example 4.2. The induced independence tests were to be compared in terms of power with the test based on the empirical correlation coefficient r and the nonparametric Spearman's tau and Kendall's rho tests. In all the examples we used level of significance

$\alpha = 0.05$. For computing the power estimates, in each of the following examples, 100 data sets were generated, each set consisting of a number of bivariate observations. For the NLCA approximations, we took $n = 1$ or $n = 2$. All the results are given in Table 1.

Our first example consists of m bivariate observations $(x_k, y_k)_{1 \le k \le m}$ generated from the pair (X, Y) such that X has the standard normal distribution and $Y = X^2$. One can see that, for this example, the tests of our class (step functions, spline 2 and spline 3) are more powerful than others. Although these results are not surprising because, for the considered model, X and Y are uncorrelated; the results illustrate that our tests are more appropriate for testing independence. Indeed, they report on independence without any assumption on the distribution of (X, Y) whereas the tests based on r, Spearman's tau and Kendall's rho detect only the lack of correlation, which is not generally equivalent to independence. Of course, this first example is a little extreme because Y depends on X through a function.

As an example where the dependence between X and Y is not functional but stochastic, we generated 500 bivariate observations from a pair (X, Y) having the uniform distribution on the unit disk $D = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 < 1\}$. In this example, X and Y are also uncorrelated. Even in this case, our tests lead to greater values for the power estimates than the classical tests.

The next issue we address is how do our tests perform, relative to the classical tests, when the data are generated from correlated random variables. We generated 500 observations from a bivariate random variable (X, Y) having the standard normal distribution with correlation coefficient $\rho = 0.8$. In this example, the classical tests are more powerful, the tests with B-spline approximations give a good level for the power estimates and the chi-squared type test (step functions) is less interesting. The superiority of the classical tests does not surprise because one knows that, in the bivariate normal case, the r-test is uniformly the most powerful. These last results lead us to research what happens in the nonnormal case.

TABLE 1

Power estimates for (X, Y) having several distributions, level $\alpha = 0.05$: I, $Y = X^2$ and $X \rightsquigarrow N(0, 1)$; II, uniform distribution on the unit disk; III, bivariate standard normal distribution with $\rho = 0.8$; IV, mixture $P_{\theta, 0.25}$ with $\theta = 0.50$; V, mixture $P_{\theta, 0.25}$ with $\theta = 0.75$

			Powers					
	$\boldsymbol{n}$	$\boldsymbol{m}$	**Step functions**	**Spline 2**	**Spline 3**	$\boldsymbol{r}$	**Rho**	**Tau**
I	1	200	1.00	1.00	1.00	0.43	0.10	0.23
	2	400	0.95	1.00	1.00	0.44	0.14	0.24
II	2	500	0.81	0.99	0.99	0.03	0.03	0.02
III	2	500	0.38	0.99	0.73	1.00	1.00	1.00
IV	2	500	0.41	1.00	1.00	0.51	0.78	0.81
V	2	500	1.00	1.00	1.00	0.39	0.33	0.44

For $\theta \in [0,1]$, let $P_{\theta,\rho}$ be the mixture

$$P_{\theta,\rho} = \theta Q_1 + (1-\theta) Q_{2,\rho},$$

where Q_1 is the distribution of the bivariate random variable (X, X^2) such that X has the standard normal distribution and $Q_{2,\rho}$ is the standard bivariate normal distribution with correlation coefficient ρ. Our last example consists of 500 bivariate observations generated from the distribution $P_{\theta,\rho}$ with $\rho = 0.25$. Clearly, a pair (X,Y) with such a distribution consists of correlated random variables X and Y and is nonnormal if $\theta \neq 0$. The results suggest that, for data coming from a nonnormal population, our tests perform better than the classical tests. They illustrate that, unlike ours, the r-test is not robust with respect to departures from the normal distribution and shows the value of our procedures as nonparametric tests for independence. Although the Spearman's rho and Kendall's tau are also nonparametric, it seems that they generally induce smaller values for the power estimates than ours.

Another fact which emerges from the table is that the tests with B-spline approximations of NLCA are generally more powerful than the test based on the chi-squared index (step functions). This raises the possibility to improve the power of the chi-squared test by using other NLCA approximations.

7. Proofs.

7.1. *Proof of Proposition* 4.1. The sequence of eigenvalues of $\Pi_{\mathscr{W}_X^n}\Pi_{\mathscr{W}_Y^n}$ arranged in decreasing order and repeated according to multiplicity is $\lambda'_n = (\lambda_1^n, \dots, \lambda_{p_n}^n, 0, \dots)$ and that of $\mathbb{E}^X\mathbb{E}^Y$ is $\lambda = (\rho_i^2)_{i=0,\dots,N}$. From Lemma XI.9.4 in Dunford and Schwartz (1963), it follows that,

$$\forall\ i = 0, \dots, N, \qquad |\lambda_i^n - \rho_i^2| \le \|\Pi_{\mathscr{W}_X^n}\Pi_{\mathscr{W}_Y^n} - \mathbb{E}^X\mathbb{E}^Y\|,$$

where $\|\cdot\|$ denotes the usual operator norm.

Thus

$$\|\lambda'_n - \lambda\|_\infty \le \|\Pi_{\mathscr{W}_X^n}\Pi_{\mathscr{W}_Y^n} - \mathbb{E}^X\mathbb{E}^Y\|.$$

Since $\Pi_{\mathscr{W}_X^n}\Pi_{\mathscr{W}_Y^n}$ converges uniformly to $\mathbb{E}^X\mathbb{E}^Y$, the latter inequality shows that $(\lambda'_n)_{n\in\mathbb{N}}$ converges in c_0 to λ. We deduce that, for all continuous nondecreasing functions Φ having the specified properties, one has $\lim_{n\to+\infty}\Phi(\lambda'_n) = \Phi(\lambda)$. $\square$

7.2. *Proof of Lemma* 5.1. Let us put $u_i^n = (\varphi_i^n(X_1), \dots, \varphi_i^n(X_m))'$ and $v_j^n = (\psi_j^n(Y_1), \dots, \psi_j^n(Y_m))'$. For all $\omega \in \Omega$, the linear map

$$L_n(\omega)\colon \alpha \in \mathbb{R}^{p_n} \mapsto \sum_{i=1}^{p_n} \alpha_i u_i^n(\omega) \in \mathbb{R}^m$$

admits the adjoint operator

$$L_n(\omega)^*\colon \gamma \in \mathbb{R}^m \mapsto \big(\dots, \langle u_i^n(\omega), \gamma\rangle_{\mathbb{R}^m}, \dots\big) \in \mathbb{R}^{p_n}.$$

Denoting by $M_n(\omega)$ the analogue of $L_n(\omega)$ with respect to the $v_j^n(\omega)$'s, $j = 1, \dots, q_n$, one can verify the following equalities:

$$mV_1^{n,m}(\omega) = L_n(\omega)^* L_n(\omega), \qquad mV_2^{n,m}(\omega) = M_n(\omega)^* M_n(\omega)$$

and

$$mV_{12}^{n,m}(\omega) = M_n(\omega)^* L_n(\omega).$$

Putting

$$A_n(\omega) = m^{-1}\big(V_1^{n,m}(\omega)\big)^{-1} L_n(\omega)^*$$

and

$$B_n(\omega) = m^{-1} M_n(\omega)\big(V_2^{n,m}(\omega)\big)^{-1} M_n(\omega)^*,$$

since (4.1) implies $L_n(\omega)\alpha^n = M_n(\omega)\beta^n$, we have

$$B_m(\omega) L_n(\omega)\alpha^n = B_m(\omega) M_n(\omega)\beta^n = M_n(\omega)\beta^n = L_n(\omega)\alpha^n;$$

thus

$$\begin{aligned}&\big(V_1^{n,m}(\omega)\big)^{-1} V_{12}^{n,m}(\omega)\big(V_2^{n,m}(\omega)\big)^{-1} V_{21}^{n,m}(\omega)\alpha^n\\ &\qquad = A_n(\omega) B_n(\omega) L_n(\omega)\alpha^n = A_n(\omega) L_n(\omega)\alpha^n = \alpha^n.\end{aligned}$$

Premultiplying this last equality by $(V_1^{n,m}(\omega))^{1/2}$, we obtain

$$R_{n,m}(\omega)\big(V_1^{n,m}(\omega)\big)^{1/2}\alpha^n = \big(V_1^{n,m}(\omega)\big)^{1/2}\alpha^n.$$

The result comes from the equalities

$$\Big\|\big(V_1^{n,m}(\omega)\big)^{1/2}\alpha^n\Big\|^2_{\mathbb{R}^{p_n}} = \langle V_1^{n,m}(\omega)\alpha^n, \alpha^n\rangle_{\mathbb{R}^{p_n}} = m^{-1}\|L_n(\omega)\alpha^n\|^2_{\mathbb{R}^{p_n}}$$

and the fact that, from (4.1), we have $\|L_n(\omega)\alpha^n\|^2_{\mathbb{R}^{p_n}} = m$. □

7.3. *Asymptotic distributions of* $S_{n,m}$. Here we determine asymptotic distributions of the random operator $S_{n,m}$ both in the general case and when X and Y are independent random variables.

In this section, when E and F are two Euclidean spaces $\mathscr{L}(E, F)$ denotes the space of linear maps from E to F. When E and F are identical, it is denoted by $\mathscr{L}(E)$. For a pair (u, v) in $E \times F$, the tensor product $u \otimes v$ is the element of $\mathscr{L}(E, F)$ defined as,

$$\forall\, h \in E, \qquad (u \otimes v)h = \langle u, h\rangle_E v,$$

where $\langle \cdot, \cdot \rangle_E$ denotes an inner product in E. The tensor product of operators is denoted by $\tilde{\otimes}$ and it is defined as previously, relative to the inner product $(A \mid B) = \mathrm{tr}(AB^*)$, where tr denotes the trace operator and B^* is the adjoint of B.

Without loss of generality, we can assume that the systems $(\varphi_i^n)_{1 \le i \le p_n}$ and $(\psi_i^n)_{1 \le i \le q_n}$ are orthonormal in $L^2(P_X)$ and $L^2(P_Y)$, respectively, and that $p_n \le q_n$. Hence, one has $V_1^n = I_{p_n}$, $V_2^n = I_{q_n}$ and $R_n = V_{12}^n V_{21}^n$, where, for all $p \in \mathbb{N}^*$, I_p stands for the identity of $\mathbb{R}^p$.

Denoting by r_n the range of R_n, we consider $(\lambda_i^n)_{0 \le i \le r_n - 1}$ the decreasing sequence of nonzero eigenvalues of R_n, $(e_i^n)_{0 \le i \le r_n - 1}$ a sequence of orthonormal eigenvectors of R_n such that e_i^n is associated with λ_i^n, and we put $h_i^n = (\rho_i^n)^{-1} V_{21}^n e_i^n$, where $\rho_i^n = \sqrt{\lambda_i^n}$.

Completing the above systems so as to obtain orthonormal bases $(e_i^n)_{0 \leq i \leq p_n - 1}$ and $(h_i^n)_{0 \leq i \leq q_n - 1}$ of $\mathbb{R}^{p_n}$ and $\mathbb{R}^{q_n}$ respectively, one has

$$\varphi^n(X) = \sum_{i=0}^{p_n - 1} f_i^n e_i^n \quad \text{and} \quad \psi^n(Y) = \sum_{i=0}^{q_n - 1} g_i^n h_i^n, \tag{7.1}$$

where $f_i^n = \langle \varphi^n(X), e_i^n \rangle_{\mathbb{R}^{p_n}}$ and $g_i^n = \langle \psi^n(X), h_i^n \rangle_{\mathbb{R}^{q_n}}$. These last variables are the canonical variables, so we have $f_0^n = g_0^n = \mathbb{1}_\Omega$ and, for $i > 0$ and $j > 0$,

$$\mathbb{E}(f_i^n) = \mathbb{E}(g_j^n) = 0, \qquad \mathbb{E}(f_i^n f_j^n) = \delta_{ij},$$
$$\mathbb{E}(g_i^n g_j^n) = \delta_{ij} \quad \text{and} \quad \mathbb{E}(f_i^n g_j^n) = \delta_{ij} \rho_i^n,$$

where δ_{ij} denotes the Kronecker symbol.

Considering the random variable

$$F_{ik}^n = \lambda_i^n (f_i^n f_k^n - \delta_{ik}) - 2\sqrt{\lambda_i^n}\left(f_k^n g_i^n - \sqrt{\lambda_i^n}\,\delta_{ik}\right) + \sqrt{\lambda_i^n \lambda_k^n}\,(g_i^n g_k^n - \delta_{ik}) + \delta_{i0}(f_k^n - \delta_{k0}),$$

we have the following lemma.

LEMMA 7.1. *The random operator $U_{n,m} = \sqrt{m}\,(S_{n,m} - S_n)$ converges in distribution, in $\mathscr{L}(\mathbb{R}^{p_n})$, as $m \to +\infty$, to a centered Gaussian random operator U_n with covariance operator*

$$K_n = \frac{1}{4} \sum_{0 \leq i, j \leq r_n - 1} \sum_{0 \leq k, l \leq p_n - 1} \mathbb{E}(F_{ik}^n F_{jl}^n) \tau_{ik}^n \,\tilde{\otimes}\, \tau_{jl}^n,$$

where $\tau_{ik}^n = e_i^n \otimes e_k^n + e_k^n \otimes e_i^n$.

PROOF. We can write $\sqrt{m}\,(S_{n,m} - S_n) = \sqrt{m}\,(R_{n,m} - R_n) + C_{n,m}$, with

$$\begin{aligned} C_{n,m} &= -\sqrt{m}\left(\left[(V_1^{n,m})^{1/2} \alpha^n\right] \otimes \left[(V_1^{n,m})^{1/2} \alpha^n\right] - \alpha^n \otimes \alpha^n\right) \\ &= -\sqrt{m}\left(\left[\left((V_1^{n,m})^{1/2} - I_{p_n}\right)\alpha^n\right] \otimes \left[\left((V_1^{n,m})^{1/2} - I_{p_n}\right)\alpha^n\right]\right. \\ &\qquad \left. + \left[\left((V_1^{n,m})^{1/2} - I_{p_n}\right)\alpha^n\right] \otimes \alpha^n + \alpha^n \otimes \left[\left((V_1^{n,m})^{1/2} - I_{p_n}\right)\alpha^n\right]\right). \end{aligned}$$

Putting

$$Z^n = (\varphi^n(X), \psi^n(Y)), \qquad Z_i^n = (\varphi^n(X_i), \psi^n(Y_i)),$$
$$V_n = \mathbb{E}(Z^n \otimes Z^n) \quad \text{and} \quad V_{n,m} = m^{-1} \sum_{i=1}^m Z_i^n \otimes Z_i^n,$$

we deduce from the central limit theorem that the random operator $H_{n,m} = \sqrt{m}\,(V_{n,m} - V_n)$ converges in distribution, in $\mathscr{L}(\mathbb{R}^{p_n + q_n})$, as $m \to +\infty$, to a centered Gaussian random operator H_n having the same covariance operator as $Z^n \otimes Z^n$.

For each element S of $\mathscr{L}(\mathbb{R}^{p_n+q_n})$ one can associate a matrix (relative to the canonical basis of $\mathbb{R}^{p_n+q_n}$) in the form

$$\begin{pmatrix} S_1 & S_{12} \\ S_{21} & S_2 \end{pmatrix},$$

and, identifying each operator with its matrix, let us consider the following operators:

$$u_1\colon S \in \mathscr{L}(\mathbb{R}^{p_n+q_n}) \mapsto S_1 \in \mathscr{L}(\mathbb{R}^{p_n}),$$
$$u_2\colon S \in \mathscr{L}(\mathbb{R}^{p_n+q_n}) \mapsto S_{12} \in \mathscr{L}(\mathbb{R}^{q_n}, \mathbb{R}^{p_n}),$$
$$u_3\colon S \in \mathscr{L}(\mathbb{R}^{p_n+q_n}) \mapsto S_{21} \in \mathscr{L}(\mathbb{R}^{p_n}, \mathbb{R}^{q_n}),$$
$$u_4\colon S \in \mathscr{L}(\mathbb{R}^{p_n+q_n}) \mapsto S_2 \in \mathscr{L}(\mathbb{R}^{q_n}).$$

From the equality

$$(V_1^{n,m})^{1/2} - I_{p_n} = \left(V_1^{n,m} - I_{p_n}\right)\left((V_1^{n,m})^{1/2} + I_{p_n}\right)^{-1}$$

we deduce $C_{n,m} = B_1^{n,m}(H_{n,m}) + B_2^{n,m}(H_{n,m})$, where $B_1^{n,m}$ and $B_2^{n,m}$ are the following random operators:

$$B_1^{n,m}(S) = -m^{-1/2}\left(\left[u_1(S)\left((V_1^{n,m})^{1/2} + I_{p_n}\right)^{-1}\alpha^n\right]\right.$$
$$\left.\otimes\left[u_1(S)\left((V_1^{n,m})^{1/2} + I_{p_n}\right)^{-1}\alpha^n\right]\right),$$

$$B_2^{n,m}(S) = -\left[u_1(S)\left((V_1^{n,m})^{1/2} + I_{p_n}\right)^{-1}\alpha^n\right] \otimes \alpha^n$$
$$+ \alpha^n \otimes \left[u_1(S)\left((V_1^{n,m})^{1/2} + I_{p_n}\right)^{-1}\alpha^n\right].$$

Almost surely, as $m \to +\infty$, $B_1^{n,m}$ converges uniformly to the null operator and $B_2^{n,m}$ converges uniformly to the operator defined as

$$B_2^n(S) = -2^{-1}\big((S_1\alpha^n) \otimes \alpha^n + \alpha^n \otimes (S_1\alpha^n)\big).$$

Moreover, one knows [cf. Pousse (1992)] that we have

$$\sqrt{m}\,(R_{n,m} - R_n) = \sum_{j=3}^{6} B_j^{n,m}(H_{n,m}),$$

where $B_j^{n,m}$, $3 \le j \le 6$, is a random operator which converges almost surely uniformly as $m \to +\infty$ to an operator B_j^n, with

$$B_3^n(S) = -\tfrac{1}{2}\big(u_1(S)R_n + R_n u_1(S)\big), \qquad B_4^n(S) = u_2(S)V_{21}^n,$$
$$B_5^n(S) = V_{12}^n u_3(S), \qquad B_6^n(S) = -V_{12}^n u_4(S) V_{21}^n.$$

Hence, we have the equality

$$U_{n,m} = \sqrt{m}\,(S_{n,m} - S_n) = \sum_{j=1}^{6} B_j^{n,m}(H_{n,m}),$$

from which we deduce that $U_{n,m}$ converges in distribution, as $m \to +\infty$, to the centered Gaussian random operator $U_n = \sum_{j=2}^{6} B_j^n(H_n)$. The covariance operator of U_n is that of $\sum_{j=2}^{6} B_j^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n))$. From relation (7.1) and the spectral decompositions

$$R_n = \sum_{i=0}^{+\infty} \lambda_i^n e_i^n \otimes e_i^n \quad \text{and} \quad V_{12}^n = \sum_{i=0}^{+\infty} \sqrt{\lambda_i^n}\, h_i^n \otimes e_i^n,$$

we obtain

$$\begin{aligned} &B_2^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) \\ &\qquad = -\tfrac{1}{2} \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} \delta_{i0}(f_k^n - \delta_{k0}) e_i^n \otimes e_k^n + e_k^n \otimes e_i^n, \\ &\qquad B_3^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) \\ &\qquad = -\tfrac{1}{2} \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} \lambda_i^n (f_i^n f_k^n - \delta_{ik}) e_i^n \otimes e_k^n + e_k^n \otimes e_i^n, \end{aligned}$$

and

$$B_4^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) = \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} \sqrt{\lambda_i^n}\left(f_k^n g_i^n - \sqrt{\lambda_i^n}\, \delta_{ik}\right) e_i^n \otimes e_k^n,$$

$$B_5^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) = \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} \sqrt{\lambda_i^n}\left(f_k^n g_i^n - \sqrt{\lambda_i^n}\, \delta_{ik}\right) e_k^n \otimes e_i^n,$$

$$B_6^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) = \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} \sqrt{\lambda_i^n \lambda_k^n}\,(g_i^n g_k^n - \delta_{ik}) e_i^n \otimes e_k^n.$$

From these equalities, we deduce

$$\sum_{j=2}^{6} B_j^n(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) = -\tfrac{1}{2} \sum_{i=0}^{r_{n-1}} \sum_{k=0}^{p_{n-1}} F_{ik}^n \tau_{ik}^n,$$

where $\tau_{ik}^n = e_i^n \otimes e_k^n + e_k^n \otimes e_i^n$. Thus, the covariance operator of U_n is as stated in the lemma. □

LEMMA 7.2. *When X and Y are independent random variables the following hold*:

1. *The random operator $\sqrt{m}\, S_{n,m}$ converges in probability in $\mathscr{L}(\mathbb{R}^{p_n})$, as $m \to +\infty$, to the null operator.*
2. *The random operator $mS_{n,m}$ converges in distribution, in $\mathscr{L}(\mathbb{R}^{p_n})$, as $m \to +\infty$, to a random operator $\mathscr{W}_n$ having the Wishart distribution $W_{p_n-1}(q_n - 1, I_{p_n-1})$, where I_{p_n-1} is the identity map of $\mathbb{R}^{p_n-1}$.*

PROOF. When X and Y are independent, we have

$$V_{12}^{n,m} = \mathbb{E}(\psi^n(Y) \otimes \varphi^n(X)) = \mathbb{E}(\psi^n(Y)) \otimes \mathbb{E}(\varphi^n(X)).$$

From (7.1), $\mathbb{E}(\varphi^n(X)) = e_0^n = \alpha^n$ and $\mathbb{E}(\psi^n(Y)) = h_0^n$; hence

$$\begin{aligned} S_n &= V_{12}^{n,m} V_{21}^{n,m} - \alpha^n \otimes \alpha^n \\ &= (\nu_0^n \otimes \xi_0^n)(\xi_0^n \otimes \nu_0^n) - \alpha^n \otimes \alpha^n = \xi_0^n \otimes \xi_0^n - \alpha^n \otimes \alpha^n = 0. \end{aligned}$$

Moreover, since $r_n = 1$ (thus, for $i \geq 1$, $\lambda_i^n = 0$), we deduce that, for all pairs (i, k), the random variable F_{ik}^n is almost surely null. So K_n is null and thus U_n is the constant random operator equal to 0. From Lemma 7.1 it follows that $\sqrt{m}\, S_{n,m}$ converges in distribution (thus, in probability) in $\mathscr{L}(\mathbb{R}^{p_n})$, as $m \to +\infty$, to the null operator.

Considering the operator u_2 introduced in the proof of Lemma 7.1, we have

$$\begin{aligned} mS_{n,m} &= m(V_1^{n,m})^{-1/2} V_{12}^{n,m} (V_2^{n,m})^{-1} V_{21}^{n,m} (V_1^{n,m})^{-1/2} \\ &= (V_1^{n,m})^{-1/2} u_2(H_{n,m})(V_2^{n,m})^{-1} u_2(H_{n,m})^* (V_1^{n,m})^{-1/2}. \end{aligned}$$

Hence, $mS_{n,m}$ converges in distribution, as $m \to +\infty$, to the random operator $\mathscr{W}_n = u_2(H_n)u_2(H_n)^*$. Since H_n is a centered Gaussian operator, so is $u_2(H_n)$. Its covariance operator is equal to that of $u_2(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n))$. Using (7.1) we obtain

$$u_2(Z^n \otimes Z^n - \mathbb{E}(Z^n \otimes Z^n)) = \sum_{i=1}^{p_n - 1} \sum_{j=1}^{q_n - 1} f_i^n g_j^n h_j^n \otimes e_i^n.$$

Thus we can write $u_2(H_n) = \sum_{i=1}^{p_n-1} \sum_{j=1}^{q_n-1} z_{ij}^n h_j^n \otimes e_i^n$, where z_{ij}^n is a real random variable which satisfies

$$\mathbb{E}\big(z_{ij}^n z_{kl}^n\big) = \mathbb{E}\big(f_i^n f_k^n g_j^n g_l^n\big) = \mathbb{E}(f_i^n f_k^n)\mathbb{E}\big(g_j^n g_l^n\big) = \delta_{ik}\delta_{jl}.$$

This last equality shows that the $q_n - 1$ column vectors of the matrix relative to $u_2(H_n)$ are independent centered Gaussian random vectors having a covariance operator equal to $I_{p_n - 1}$. Hence, $\mathscr{W}_n$ has the Wishart distribution $W_{p_n-1}(q_n - 1, I_{p_n-1})$. $\square$

7.4. *A property of symmetric nondecreasing functions*. In Section 5.2 we have stated that $(\partial\Phi_n/\partial x_k)(\lambda_n)$ is constant when k belongs to $\mathscr{I}_i^n$ and that $(\partial\Phi_n/\partial x_k)(0)$ is equal to a constant whatever is the value of k in $\{1, \dots, p_n\}$. Here we prove a property of symmetric nondecreasing functions from which these assertions are deduced.

Let Φ be a symmetric nondecreasing function defined on $\mathbb{R}^p$. For any vector $x = (x_1, \dots, x_p) \in \mathbb{R}^p$, the x_i's are valued in a set which will be enumerated in the form $\{x_1^*, \dots, x_r^*\}$, $r \leq p$, where $x_1^* > x_2^* > \cdots > x_r^*$. For $i = 1, \dots, r$, we set $E_i(x) = \{k;\ 1 \leq k \leq p,\ x_k = x_i^*\}$.

LEMMA 7.3. *If Φ admits first partial derivatives at $x \in \mathbb{R}^p$, then there exists a real $K_\Phi^i(x)$, $i = 1, \dots, r$, such that, for any $k \in E_i(x)$, one has $(\partial\Phi/\partial x_k)(x) = K_\Phi^i(x)$.*

PROOF. Let $\Phi_{k,x}$ be the kth partial map of Φ at x, that is, the map defined on a suitable subset A_x of $\mathbb{R}^p$ by $\Phi_{k,x}(t) = \Phi(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_n)$. Let s_i be the cardinality of $E_i(x)$; since Φ is a symmetric function, for $h \in \mathbb{R}$, $k \in E_i(x)$ and $j \in E_i(x)$ with $j \neq k$, one has

$$\begin{aligned}\Phi_{k,x}(x_k + h) &= \Phi\Big(\underbrace{x_1^*, \dots, x_1^*}_{s_1 \text{ terms}}, \dots, \underbrace{x_i^* + h, x_i^*, \dots, x_i^*}_{s_i \text{ terms}}, \dots, \underbrace{x_r^*, \dots, x_r^*}_{s_r \text{ terms}}\Big)\\ &= \Phi_{j,x}\big(x_j + h\big),\end{aligned}$$

from which it follows that

$$\begin{aligned}\frac{\partial \Phi}{\partial x_k}(x) &= \lim_{h \to 0} \frac{\Phi_{k,x}(x_k + h) - \Phi_{k,x}(x_k)}{h}\\ &= \lim_{h \to 0} \frac{\Phi_{j,x}\big(x_j + h\big) - \Phi_{j,x}(x_j)}{h} = \frac{\partial \Phi}{\partial x_j}(x). \qquad \square\end{aligned}$$

As a consequence of the previous lemma, we have that, if Φ admits first partial derivatives at the null vector of $\mathbb{R}^p$, since $E_1(0) = \{1, \dots, p\}$, $(\partial \Phi_n / \partial x_k)(0)$ is constant when k belongs to $\{1, \cdots, p\}$.

7.5. *Asymptotic distribution of $r_\Phi^{n,m}(X, Y)$ in the general case.* Here we consider notation and assumptions introduced in Section 5.2. Almost surely, we can write

$$\Phi_n(\lambda_{n,m}) - \Phi_n(\lambda_n) = D\Phi_n(\lambda_n)(\lambda_{n,m} - \lambda_n) + R_{\Phi_n}(\lambda_{n,m}, \lambda_n).$$

By Taylor's formula we have, almost surely,

$$|R_{\Phi_n}(\lambda_{n,m}, \lambda_n)| \leq \frac{M_{\Phi_n}}{2} \|\lambda_{n,m} - \lambda_n\|^2_{\mathbb{R}^{p_n}}, \tag{7.2}$$

where $\|\cdot\|_{\mathbb{R}^p}$ is the usual Euclidean norm of $\mathbb{R}^p$.

LEMMA 7.4. *The random variable $\sqrt{m}\, R_{\Phi_n}(\lambda_{n,m}, \lambda_n)$ converges in probability, as $m \to +\infty$, to* 0.

PROOF. Let us put $U_{n,m} = \sqrt{m}\,(S_{n,m} - S_n)$. From Lemma 7.1, we know that $U_{n,m}$ converges in distribution, as $m \to +\infty$, to a centered Gaussian random operator U_n. Considering the orthogonal projector $Q_i^{n,m}$ onto the eigenspace of $S_{n,m}$ associated with $\lambda_i^{n,m}$, and the random operator

$$\xi_i^{n,m}(U_{n,m}) = \sqrt{m}\,(Q_i^{n,m} S_{n,m} Q_i^{n,m} - \lambda_i^{n*} Q_i^{n,m}),$$

one knows that [see Dossou-Gbete and Pousse (1991)] that the sequence $(\sqrt{m}\,(\lambda_k^{n,m} - \lambda_i^{n*}))_{k \in \mathscr{I}_i^n}$ consists of all the eigenvalues of $\xi_i^{n,m}(U_{n,m})$. Hence

$$\begin{aligned}\|\sqrt{m}\,(\lambda_{n,m} - \lambda_n)\|^2_{\mathbb{R}^{p_n}} &= \sum_{i=1}^{p_n} \big(\sqrt{m}\,(\lambda_i^{n,m} - \lambda_i^n)\big)^2\\ &= \sum_{i=1}^{s_n} \sum_{k \in \mathscr{I}_i^n} \big(\sqrt{m}\,(\lambda_k^{n,m} - \lambda_i^{n*})\big)^2\\ &= \sum_{i=1}^{s_n} \operatorname{tr}(\xi_i^{n,m}(U_{n,m})\, \xi_i^{n,m}(U_{n,m})^*),\end{aligned}$$

where tr denotes the trace operator. Since $(\xi_i^{n,m}(U_{n,m}))_{1\leq i\leq s_n}$ converges in distribution, as $m\to+\infty$, to the centered Gaussian random variable $(\xi_i^n(U_n))_{1\leq i\leq s_n}$, where $\xi_i^n(U_n)=Q_i^nU_nQ_i^n$ and Q_i^n is the orthogonal projector onto the eigenspace of S_n associated with λ_i^{n*} [see Dossou-Gbete and Pousse (1991)], the random variable $\|\sqrt{m}\,(\lambda_{n,m}-\lambda_n)\|^2_{\mathbb{R}^{p_n}}$ converges in distribution to $\sum_{i=1}^{s_n}\operatorname{tr}(\xi_i^n(U_n)\xi_i^n(U_n)^*)$. From (7.2) it follows that

$$|\sqrt{m}\,R_{\Phi_n}(\lambda_{n,m},\lambda_n)|\leq\frac{M_{\Phi_n}}{2\sqrt{m}}\|\sqrt{m}\,(\lambda_{n,m}-\lambda_n)\|^2_{\mathbb{R}^{p_n}},$$

and the assertion follows. □

PROOF OF PROPOSITION 5.2. From the previous lemma, we deduce that the random variables $\sqrt{m}\,(\Phi_n(\lambda_{n,m})-\Phi_n(\lambda_n))$ and $\sqrt{m}\,D\Phi_n(\lambda_n)(\lambda_{n,m}-\lambda_n)$ have the same asymptotic distribution. Moreover, we have

$$\begin{aligned}\sqrt{m}\,D\Phi_n(\lambda_n)(\lambda_{n,m}-\lambda_n)&=\sum_{i=1}^{p_n}\frac{\partial\Phi_n}{\partial x_i}(\lambda_n)\sqrt{m}\,(\lambda_i^{n,m}-\lambda_i^n)\\&=\sum_{i=1}^{s_n}\sum_{k\in\mathscr{I}_i^n}\frac{\partial\Phi_n}{\partial x_k}(\lambda_n)\sqrt{m}\,(\lambda_k^{n,m}-\lambda_i^{n*})\\&=\sum_{i=1}^{s_n}K^i_{\Phi_n}(\lambda_n)\operatorname{tr}(\xi_i^{n,m}(U_{n,m})),\end{aligned}$$

this random variable converges in distribution, as $m\to+\infty$, to the random variable $\nu_n=\sum_{i=1}^{s_n}K^i_{\Phi_n}(\lambda_n)\operatorname{tr}(\xi_i^n(U_n))$. As $(\xi_i^n(U_n))_{1\leq i\leq s_n}$ is centered Gaussian, so is ν_n. The variance is

$$\sigma^2_{\Phi_n}(\lambda_n)=\sum_{i=1}^{s_n}\sum_{j=1}^{s_n}K^i_{\Phi_n}(\lambda_n)K^j_{\Phi_n}(\lambda_n)\mathbb{E}\Big(\operatorname{tr}(\xi_i^n(U_n))\operatorname{tr}\big(\xi_j^n(U_n)\big)\Big).\tag{7.3}$$

Let us introduce the usual inner product $(A\mid B)=\operatorname{tr}(AB^*)$ of rank finite operators. We can write

$$\begin{aligned}&\mathbb{E}\Big(\operatorname{tr}(\xi_i^n(U_n))\operatorname{tr}\big(\xi_j^n(U_n)\big)\Big)\\&\quad=\mathbb{E}\Big(\big(\xi_i^n(U_n)\,\tilde{\otimes}\,\xi_j^n(U_n)\big)(I_{p_n})\mid(I_{p_n})\Big)\\&\quad=\mathbb{E}\Big(\operatorname{tr}\Big(\big(\xi_i^n(U_n)\,\tilde{\otimes}\,\xi_j^n(U_n)\big)(I_{p_n})\Big)\Big)=\mathbb{E}\Big(\operatorname{tr}\Big(\big(\xi_j^n\big(U_n\,\tilde{\otimes}\,U_n\big)\xi_i^{n*}\big)(I_{p_n})\Big)\Big)\\&\quad=\mathbb{E}\Big(\big(\xi_j^nK_n\,\xi_i^{n*}\big)(I_{p_n})\Big)=\operatorname{tr}\Big(\big(\xi_j^n(K_n(Q_i^n))\big)\Big).\end{aligned}$$

From the expression of K_n and the equality $Q_i^n=\sum_{k\in\mathscr{I}_i^n}e_k^n\otimes e_k^n$ we deduce

$$K_n(Q_i^n)=2^{-1}\sum_{p\in\mathscr{I}_i^n}\sum_{0\leq k\leq r_n-1}\sum_{0\leq l\leq p_n-1}\mathbb{E}\big(F^n_{pp}F^n_{kl}\big)\tau^n_{kl};$$

hence

$$\xi_j^n(K_n(Q_i^n)) = Q_j^n(K_n(Q_i^n))Q_j^n = \frac{1}{2}\sum_{p\in\mathscr{I}_i^n}\sum_{(q,l)\in(\mathscr{I}_j^n)^2}\mathbb{E}(F_{pp}^nF_{ql}^n)\tau_{ql}^n,$$

and it follows that

$$\operatorname{tr}\Big(\big(\xi_j^n(K_n(Q_i^n))\big)\Big) = \sum_{p\in\mathscr{I}_i^n}\sum_{q\in\mathscr{I}_j^n}\mathbb{E}(F_{pp}^nF_{qq}^n).$$

Consequently, the variance is

$$\sigma_{\Phi_n}^2(\lambda_n) = \sum_{1\le i,j\le s_n} K_{\Phi_n}^i(\lambda_n)K_{\Phi_n}^j(\lambda_n)\sigma_{ij}^n,$$

with

$$\sigma_{ij}^n = \sum_{p\in\mathscr{I}_i^n}\sum_{q\in\mathscr{I}_j^n}\mathbb{E}(F_{pp}^nF_{qq}^n). \tag{7.4}$$ □

7.6. *Asymptotic distribution of $r_\Phi^{n,m}(X,Y)$ in case X and Y are independent*. When X and Y are independent, one has $\lambda_n = 0$ and since for all pairs (i,j) we have $F_{ij}^n = 0$ (almost surely) we deduce from (7.4) that $\sigma_{\Phi_n}^2(\lambda_n) = 0$. Otherwise, almost surely, we can set $\Phi_n(\lambda_{n,m}) = D\Phi_n(0)(\lambda_{n,m}) + R_{\Phi_n}(\lambda_{n,m},0)$, and we have

$$|R_{\Phi_n}(\lambda_{n,m},0)| \le \frac{M_{\Phi_n}}{2}\|\lambda_{n,m}\|_{\mathbb{R}^{p_n}}^2. \tag{7.5}$$

LEMMA 7.5. *The random variable $m\,R_{\Phi_n}(\lambda_{n,m},0)$ converges in probability, as $m\to+\infty$, to* 0.

PROOF. Let Δ be the continuous function which associates with each operator the decreasing sequence of its eigenvalues. Since $\sqrt{m}\,(S_{n,m})$ converges in distribution ($m\to+\infty$) to 0, using Theorem 5.1 in Billingsley (1968), we obtain the following convergence in distribution (thus, in probability):

$$\lim_{m\to+\infty}\|\sqrt{m}\,\lambda_{n,m}\|_{\mathbb{R}^{p_n}} = \lim_{m\to+\infty}\|\Delta\big(\sqrt{m}\,(S_{n,m})\big)\|_{\mathbb{R}^{p_n}} = 0.$$

In consequence, Lemma 7.5 is deduced from the fact that, using (7.5), we have almost surely

$$|mR_{\Phi_n}(\lambda_{n,m},0)| \le 2^{-1}M_{\Phi_n}\|\sqrt{m}\,\lambda_{n,m}\|_{\mathbb{R}^{p_n}}^2.$$ □

PROOF OF PROPOSITION 5.3. We deduce from the preceding lemma that $m\Phi_n(\lambda_{n,m})$ has the same limit distribution as $mD\Phi_n(0)(\lambda_{n,m})$. Since we have

$$mD\Phi_n(0)(\lambda_{n,m}) = \sum_{i=1}^{p_n}\frac{\partial\Phi_n}{\partial x_i}(0)\,m\lambda_i^{n,m} = K_{\Phi_n}\operatorname{tr}(mS_{n,m}),$$

the limit distribution of $mK_{\Phi_n}^{-1}\Phi_n(\lambda_{n,m})$ is that of $\mathrm{tr}(mS_{n,m})$. The latter is the distribution of $\mathrm{tr}(\mathscr{W}_n)$, that is, the chi-squared distribution with $(p_n - 1) \times (q_n - 1)$ degrees of freedom. $\square$

7.7. *Proof of Proposition* 6.1. Under $H_{1,\Phi}^n$, one has $|\Phi_n(\lambda_n)| > 0$. Let ε be a real number such that the inequality $0 < \varepsilon < |\Phi_n(\lambda_n)|$ holds. Then

$$P\big(\big||\Phi_n(\lambda_{n,m})| - |\Phi_n(\lambda_n)|\big| \le \varepsilon\big) \le P\big(|\Phi_n(\lambda_{n,m})| \ge |\Phi_n(\lambda_n)| - \varepsilon\big) \le 1. \tag{7.6}$$

Since $|\Phi_n(\lambda_{n,m})|$ converges in probability, as $m \to +\infty$, to $|\Phi_n(\lambda_n)|$, one has the convergence of $P(||\Phi_n(\lambda_{n,m})| - |\Phi_n(\lambda_n)|| \le \varepsilon)$ to 1. Then from (7.6) we deduce that

$$P\big(|K_{\Phi_n}^{-1}\Phi_n(\lambda_{n,m})| \ge |K_{\Phi_n}^{-1}|(|\Phi_n(\lambda_n)| - \varepsilon)\big)$$

converges, as $m \to +\infty$, to 1. Moreover, for a sufficiently large m, one has

$$m^{-1}t_\alpha^n < |K_{\Phi_n}^{-1}|(|\Phi_n(\lambda_n)| - \varepsilon).$$

It follows that, for some m_0, if $m > m_0$, then we have

$$\begin{aligned} P\big(|K_{\Phi_n}^{-1}\Phi_n(\lambda_{n,m})| &\ge |K_{\Phi_n}^{-1}|(|\Phi_n(\lambda_n)| - \varepsilon)\big) \\ &\le P\big(|K_{\Phi_n}^{-1}\Phi_n(\lambda_{n,m})| > m^{-1}t_\alpha^n\big) \le 1. \end{aligned}$$

From this inequality, we deduce the equality

$$\lim_{m \to +\infty} P\big(m|K_{\Phi_n}^{-1}\Phi_n(\lambda_{n,m})| > t_\alpha^n\big) = 1. \qquad \square$$

Acknowledgments. We are deeply grateful to Professors Henri Caussinus, Joe Wittaker and Jim Ramsay for their help in the preparation of this paper. Also, we thank the referees for their valuable comments.

REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Analysis*. Wiley, New York.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BLACHER, R. (1988). A new form for the chi-squared test of independence. *Statistics* **19** 519–536.

BREIMAN, L. and FRIEDMAN, J. H. (1985). Estimating optimal transformations for multiple regression and correlation. *J. Amer. Statist. Assoc.* **80** 580–598.

BUJA, A. (1990). Remarks on functional canonical variates, alternating least squares methods and ACE. *Ann. Statist.* **18** 1032–1069.

CRAMER, E. M. and NICEWANDER, W. A. (1979). Some symmetric invariant measures of multivariate association. *Psychometrika* **41** 347–352.

DAUXOIS, J. and NKIET, G. M. (1997). Testing for the lack of a linear relationship. *Statistics* **30** 1–23.

DAUXOIS, J. and POUSSE, A. (1975). Une extension de l'analyse canonique. Quelques applications. *Ann. Inst. H. Poincaré* **11** 355–379.

DAUXOIS, J. and POUSSE, A. (1977). Some convergence problems in factor analysis. In *Recent Developments in Statistics* (Barra et al., eds.). North-Holland, Amsterdam.

DAUXOIS, J., ROMAIN, Y. and VIGUIER, S. (1994). Tensor products and statistics. *Linear Algebra Appl.* **270** 59–88.

DOSSOU-GBETE, S. and POUSSE, A. (1991). Asymptotic study of eigenelements of a sequence of random selfadjoint operators. *Statistics* **22** 479–491.

DUNFORD, N. and SCHWARTZ, J. T. (1963). *Linear Operators* **2**. Wiley, New York.

EATON, M. L. (1983). *Multivariate Statistics. A Vector Approach*. Wiley, New York.

GOHBERG, I. C. and KREJN, M. G. (1971). *Introduction à la Théorie des Opérateurs Linéaires non Autoadjoints dans un Espace Hilbertien*. Dunod, Paris.

JENSEN, D. R. and MAYER, L. S. (1977). Some variational results and their applications in multiple inference. *Ann. Statist.* **5** 922–931.

LAFAYE DE MICHEAUX, D. (1978). Approximation d'analyses canoniques non linéaires et analyses factorielles privilégiantes. Thèse de Docteur Ingénieur, Univ. Nice.

LANCASTER, H. O. (1969). *The Chi-Squared Distributions*. Wiley, New York.

LAZRAQ, A. and CLÉROUX, R. (1988). Etude comparative de différentes mesures de liaison entre deux variables aléatoires. *Statist. Anal. Données B* **13** 39–58.

LIN, P.-E. (1987). Measures of association between vectors. *Comm. Statist. Theory Methods* **16** 321–338.

O'NEILL, M. E. (1978). Asymptotic distributions of the canonical correlation from contingency tables. *Austral. J. Statist.* **20** 75–82.

POUSSE, A. (1992). Etudes asymptotiques, In *Modèles pour l'Analyse des Données Multidimensionnelles* (Droesbeke et al., eds.). Economica, Paris.

RÉNYI, A. V. (1959). On measures of dependence. *Acta. Math. Acad. Sci. Hungar.* **10** 441–451.

SCHUMAKER, L. L. (1981). *Spline Functions. Basic Theory*. Wiley, New York.

TSAI, MING-TAN M. and SEN, P. K. (1990). Tests for independence in two-way contingency tables based on canonical analysis. *Calcutta Statist. Assoc. Bull.* **40** 109–123.

LABORATOIRE DE STATISTIQUE
ET PROBABILITÉS, UMR-CNRS C55830
UNIVERSITÉ PAUL SABATIER
118, ROUTE DE NARBONNE
TOULOUSE
FRANCE

DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ DES SCIENCES
ET TECHNIQUES DE MASUKU
BP 943, FRANCEVILLE
GABON