

# Functional Canonical Analysis for Square Integrable Stochastic Processes<sup>1</sup>

May 30, 2002 (Revision)

Guozhong He<sup>2</sup>, Hans-Georg Müller<sup>3</sup> and Jane-Ling Wang<sup>3</sup>

We study the extension of canonical correlation from pairs of random vectors to the case where a data sample consists of pairs of square integrable stochastic processes. Basic questions concerning the definition and existence of functional canonical correlation are addressed and sufficient criteria for the existence of functional canonical correlation are presented. Various properties of functional canonical analysis are discussed. We consider a canonical decomposition, in which the original processes are approximated by means of their canonical components.

*Key words:* Canonical correlation, canonical decomposition, covariance operator, functional data analysis, Hilbert-Schmidt operator, inverse problem

## 1. Introduction

Increasingly, data are collected in the form of random functions or curves. Such curve data may be generated by densely spaced repeated measurements, for example in longitudinal studies, or by automatic recordings of a quantity over time. This type of data is becoming more prevalent throughout the sciences and in financial markets, as automated on-line data

---

<sup>1</sup> Research supported by NSF Grants DMS 98-03627, DMS 99-71602, and NIH Grant 99-SC-NIH-1028.

<sup>2</sup> The California Diabetes Control Program, University of California, San Francisco/DHS, 601 North 7th Street, MS 675, Sacramento, CA 94234, USA.

<sup>3</sup> Department of Statistics, University of California, One Shields Avenue, Davis, CA 95616, USA.

collection facilities are becoming more ubiquitous. Functional data analysis (FDA) is concerned with data for which the  $i$ th observation consists of one or several infinite dimensional objects such as curves or surfaces (see the book by Ramsay and Silverman (1997) for an excellent overview). Typically, these objects are considered to be random elements of some functional space. Many research questions and statistical modeling issues are indeed best described in functional terms. This motivates the extension of classical concepts of multivariate data analysis (MDA), such as principal components analysis, canonical correlation analysis, and linear modeling, to the infinite-dimensional functional domain.

In this paper, we consider the situation where a data sample consists of pairs of observed functions. The study of the dependence between the two functions recorded for a sample of subjects is then often of interest. We thus aim at extending methods for analyzing the linear correlation between paired observations from the multivariate to the infinite-dimensional case.

Several approaches have been developed previously for extending multivariate canonical correlation (Hotelling, 1936) to the functional case. In early work on this problem, Hannan (1961) and Brillinger (1975) described canonical analysis for multivariate stationary time series. By invoking stationarity, the problem in this setting may be reduced to classical multivariate canonical analysis. A theoretical approach, based on angles between subspaces of functions was developed by Dauxois and Nkiet (1977). A sample version of smoothed functional canonical correlation was defined in Leurgans, Moyeed and Silverman (1993), who demonstrated the need for regularization in functional canonical correlation analysis. They implemented regularization via modified smoothing splines and demonstrated their technique with an application to the study of human gait movement data; compare also Olshen et al. (1989). For the related question of extending principal components from multivariate to functional data, we refer to Rice and Silverman (1991). A different and promising approach aiming at covariance rather than correlation for pairs of random curves was proposed in Service, Rice, and Chavez (1998). Regularization for canonical correlation amounts to restricting the dimension of the problem and can be achieved via a judicious

choice of the roughness penalty for smoothing splines as in Leurgans et al. (1993), or by alternative approaches that allow to avoid the inversion problem as in Service et al. (1998). A third approach that will be discussed below is to approximate processes by finite expansions, for example in terms of eigenfunctions, and to apply canonical correlation analysis to the resulting finite-dimensional principal components.

In this paper, we address two issues: First, the problem to define functional canonical correlation in infinite-dimensional space and to identify conditions under which it is well defined. Second, the representation of pairs of square integrable processes in terms of canonical basis functions.

The paper is organized as follows: In Section 2, we provide basic notation and introduce functional canonical correlation based on a classical definition. An alternative definition is the theorem of Section 3. The main results on existence can be found in Section 4, and a functional canonical representation is established in Section 5, including a discussion of special cases and examples. Proofs and auxiliary results are collected in Section 6, and some pertinent facts from functional analysis in an appendix.

## 2. Canonical Correlation for Random Vectors and Random Functions

Suppose we observe a sample of bivariate processes  $(X, Y)$ , where  $X \in L_2(T_1)$  and  $Y \in L_2(T_2)$  are jointly distributed  $L_2$ -processes,

$$\int E\|Z\|^2 = E[\langle Z Z \rangle] = E \int_T (Z(s))^2 ds < \infty, \quad \text{for } Z = X \text{ or } Z = Y.$$

Here,  $T_1$  and  $T_2$  are index sets (intervals or countable sets), and  $L_2(T_1), L_2(T_2)$  are two Hilbert spaces of square integrable functions on  $T_1$  and  $T_2$  with respect to measures  $\mu_1$  and  $\mu_2$ , (usually Lebesgue measure or counting measure), with scalar products  $\langle u, v \rangle = \int u(s)v(s)d\mu_i(s)$ , for  $i = 1, 2$ . Canonical correlation as defined for finite-dimensional vectors,  $X \in \mathbb{R}^{k_1}$ ,  $Y \in \mathbb{R}^{k_2}$ , and formally for stochastic processes  $X \in L_2(T_1)$ ,  $Y \in L_2(T_2)$  is characterized as follows: Let  $H_i = \mathbb{R}^{k_i}$  in the vector case,  $H_i = L_2(T_i)$  in the functional case. Then the first *canonical correlation*  $\rho_1$  and associated *weight functions or vectors*  $u_1$  and  $v_1$  are defined as

$$(1) \quad \rho_1 = \sup_{u \in H_1, v \in H_2} \text{Cov}(\langle u, X \rangle, \langle v, Y \rangle) = \text{Cov}(\langle u_1, X \rangle, \langle v_1, Y \rangle),$$

where  $u$  and  $v$  are subject to

$$(2) \quad \text{Var}(\langle u, X \rangle) = 1, \quad \text{Var}(\langle v, Y \rangle) = 1.$$

The  $k$ -th canonical correlation and weight functions  $\rho_k, u_k, v_k$  for  $(X, Y)$ , for  $k > 1$ , are defined as

$$(3) \quad \rho_k = \sup_{u \in H_1, v \in H_2} \text{Cov}(\langle u, X \rangle, \langle v, Y \rangle) = \text{Cov}(\langle u_k, X \rangle, \langle v_k, Y \rangle),$$

where  $u$  and  $v$  are subject to (2), and the  $k$ -th pair of canonical variables

$$(4) \quad (U_k, V_k) \text{ is uncorrelated with the } (k-1) \text{ pairs } \{(U_i, V_i), i = 1, \dots, k-1\},$$

where  $U_k = \langle u_k, X \rangle$  and  $V_k = \langle v_k, Y \rangle$ . We shall call  $(\rho_k, u_k, v_k, U_k, V_k)$  the  $k$ -th *canonical components*.

For any  $X$  and  $Y$ , we hold that  $X$  and  $Y$  are uncorrelated if all their canonical correlations are zero. This is equivalent to saying that:  $\text{Corr}(X, Y) = 0$  if and only if  $\rho_1 = 0$ , since  $\rho_1 \geq \rho_2 \geq \dots \geq 0$ .

Regarding the cases where  $X, Y$  are stochastic processes, we first consider the special case where processes  $X$  and  $Y$  can be represented by a finite number of orthonormal basis functions. For such finite-dimensional processes, functional canonical analysis is equivalent to the usual multivariate canonical analysis for the random coefficient vectors. To see this, let

$$X(t) = \mu_X(t) + \sum_{i=1}^{k_1} \xi_i \theta_i(t), \quad t \in T_1, \quad Y(t) = \mu_Y(t) + \sum_{i=1}^{k_2} \zeta_i \varphi_i(t), \quad t \in T_2,$$

where  $\{\theta_i\}$  and  $\{\varphi_i\}$  are the first  $k_1$ , respectively  $k_2$ , elements of orthonormal bases of  $L_2(T_1)$  and  $L_2(T_2)$ , respectively, and  $\{\xi_i\}$  and  $\{\zeta_i\}$  are random variables with zero means and finite variances. We adopt the notation

$$\begin{aligned} \boldsymbol{\theta}(t) &= (\theta_1(t), \dots, \theta_{k_1}(t))^T, \quad \boldsymbol{\varphi}(t) = (\varphi_1(t), \dots, \varphi_{k_2}(t))^T, \quad 1 < k_1, k_2 < \infty, \\ \boldsymbol{\xi} &= (\xi_1, \dots, \xi_{k_1})^T, \quad \boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_{k_2})^T, \end{aligned}$$

with

$$E[\boldsymbol{\xi}] = 0, \quad \text{Var}[\boldsymbol{\xi}] = R_{11}, \quad E[\boldsymbol{\zeta}] = 0, \quad \text{Var}[\boldsymbol{\zeta}] = R_{22}, \quad \text{Cov}[\boldsymbol{\xi}, \boldsymbol{\zeta}] = E[\boldsymbol{\xi} \boldsymbol{\zeta}^T] = R_{12}.$$

Without loss of generality we assume  $\mu_X(t) = 0, \mu_Y(t) = 0$ . Then processes  $X$  and  $Y$  can be

written in vector form as

$$(5) \quad X(t) = \xi^T \theta(t), \quad Y(t) = \zeta^T \varphi(t).$$

As demonstrated in the following theorem, canonical correlations for  $X$  and  $Y$  in this case are the same as the canonical correlations for the random vectors  $\xi$  and  $\zeta$ .

**Theorem 2.1** *The  $i$ -th canonical component of  $(\xi, \zeta)$ , defined by  $(\sigma_i, \mathbf{u}_i, \mathbf{v}_i)$ , is related to the  $i$ -th canonical component of  $(X(t), Y(t))$ ,  $(\rho_i, u_i(t), v_i(t))$ , through*

$$u_i(t) = \mathbf{u}_i^T \theta(t), \quad v_i(t) = \mathbf{v}_i^T \varphi(t), \quad \rho_i = \sigma_i.$$

Canonical correlation analysis for finite dimensional processes is therefore equivalent to multivariate canonical correlation. In this simple situation, functional canonical correlation analysis is then well defined. Difficulties arise however in the more realistic situation where processes  $X$  and  $Y$  are genuinely infinite-dimensional, for the following reasons: First, the definition in this case requires that there are countably many canonical correlations. This means that the cross-correlation operator  $R$  defined in (10) below must be compact. Secondly, the operator  $R$  involves inverse operators. Inversion of functional operators in functional space is delicate as a compact operator is not invertible in infinite-dimensional spaces. Thirdly, the canonical weight functions,  $u_i$  and  $v_i$ , may not be square integrable. We address these issues, which are genuine difficulties in Section 4. An alternative characterization of canonical correlation for the infinite-dimensional case that is well-known for the functional multivariate case and is useful for our investigation is studied in the next section.

### 3. Alternative Characterization of Canonical Correlation

Consider the case of random vectors, i.e.,  $H_1 = \mathbb{R}^{k_1}$ ,  $H_2 = \mathbb{R}^{k_2}$ , and  $r_{XX} = \text{Cov}(X) = E(X - EX)(X - EX)^T$ ,  $r_{YY} = \text{Cov}(Y)$  and  $r_{XY} = \text{Cov}(X, Y) = E(X - EX)(Y - EY)^T$ . The  $(k_1 \times k_1)$  respectively  $(k_2 \times k_2)$  covariance matrices  $r_{XX}$ ,  $r_{YY}$  are symmetric and nonnegative definite. It is then well known that

$$(6) \quad \rho_k = \sup_{u \in H_1, \langle u, r_{XX} u \rangle = 1, v \in H_2, \langle v, r_{YY} v \rangle \geq 1} \langle u, r_{XY} v \rangle = \langle u_k, r_{XY} v_k \rangle,$$

where for  $U_k = \langle u_k, X \rangle$ ,  $V_k = \langle v_k, Y \rangle$ , the pairs  $(U_k, V_k)$  are uncorrelated with  $(U_i, V_i)$  for

$i = 1, \dots, k-1$ ,  $k \leq \min(k_1, k_2)$ , and  $u_k, v_k$  are the  $k$ -th weight functions.

Extending this characterization of canonical correlation to the functional case, where  $H_1 = L_2(T_1)$ ,  $H_2 = L_2(T_2)$ , we define the covariance functions

$$\begin{aligned} r_{XX}(s, t) &= \text{Cov}[X(s), X(t)], \quad s, t \in T_1; \\ r_{YY}(s, t) &= \text{Cov}[Y(s), Y(t)], \quad s, t \in T_2; \\ r_{XY}(s, t) &= \text{Cov}[X(s), Y(t)], \quad s \in T_1, t \in T_2, \end{aligned}$$

and the covariance operator  $R_{XX}: L_2(T_1) \rightarrow L_2(T_1)$ ,

$$(7) \quad R_{XX}u(s) = \int_{T_1} r_{XX}(s, t)u(t)dt, \quad u \in L_2(T_1),$$

and analogously operators  $R_{YY}: L_2(T_2) \rightarrow L_2(T_2)$ , and  $R_{XY}: L_2(T_2) \rightarrow L_2(T_1)$ . Operators  $R_{XX}$  and  $R_{YY}$  are compact, self-adjoint, and nonnegative definite, and  $R_{XY}$  is compact.

Since

$$\begin{aligned} \text{Cov}(\langle u, X \rangle, \langle v, Y \rangle) &= E\{[\langle u, X \rangle - E(\langle u, X \rangle)][\langle v, Y \rangle - E(\langle v, Y \rangle)]\} \\ &= E\{[\langle u, X - E(X) \rangle][\langle v, Y - E(Y) \rangle]\} \\ &= \langle u, R_{XY}v \rangle, \end{aligned}$$

and

$$\text{Var}(\langle u, X \rangle) = \langle u, R_{XX}u \rangle, \quad \text{Var}(\langle v, Y \rangle) = \langle v, R_{YY}v \rangle,$$

a characterization of the  $k$ -th canonical correlation and weight functions analogous to (6) is given by:

$$(8) \quad \rho_k = \sup_{\substack{u \in L_2(T_1), \langle u, R_{XX}u \rangle = 1, \\ v \in L_2(T_2), \langle v, R_{YY}v \rangle = 1}} \langle u, R_{XY}v \rangle = \langle u_k, R_{XY}v_k \rangle,$$

where, in addition, for  $k > 1$ ,

$$(9) \quad (U_k, V_k) \text{ is uncorrelated with } (U_i, V_i) \text{ for } i = 1, \dots, k-1.$$

The canonical components are solely determined by the covariance functions of processes  $X$  and  $Y$ , and are not affected by their means. We therefore assume throughout the rest of the paper that the means vanish,

$$E[X(t)] = 0, \quad t \in T_1, \quad E[Y(s)] = 0, \quad s \in T_2.$$

#### 4. Existence of functional canonical correlations and functional canonical weight functions

Intuitively, maximizing the r.h.s. of (8), given the constraints (9), is equivalent to an eigenanalysis of the *cross-correlation operator*  $R$  of  $X$  and  $Y$ ,

$$(10) \quad R = R_{XX}^{-1/2} R_{XY} R_{YY}^{-1/2}.$$

That this in fact holds under certain assumptions on the processes is one of the basic results below (Theorem 4.8). Compactness of the cross-correlation operator  $R$  is therefore a natural condition in order to guarantee that functional canonical correlations exist and are interpretable. The basic problem is that, unlike the usual situation in the finite dimensional case, the square roots of covariance operators of  $L_2$ -processes are not invertible. In infinite-dimensional spaces canonical correlation corresponds to an inverse problem.

Our approach is to consider a subset of  $L_2$ , on which the inverse of a compact operator can be defined. Following Conway (1985), the range of  $R_{XX}^{-1/2}$ , given by

$$(11) \quad F_{XX} = \{R_{XX}^{1/2}h : h \in L_2(T_1)\},$$

is characterized by

$$F_{XX} = \{f \in L_2(T_1) : \sum_{i=1}^{\infty} \lambda_{Xi}^{-1} |\langle f, \theta_i \rangle|^2 < \infty, f \perp \ker(R_{XX})\},$$

where  $\{\lambda_{Xi}, \theta_i\}$  are the non-zero eigenvalues and eigenvectors of  $R_{XX}$ , and  $\ker(R_{XX}) = \{h \in L_2(T) : R_{XX}h = 0\}$ . Defining

$$(12) \quad F_{XX}^{-1} = \{h \in L_2(T_1) : h = \sum_{i=1}^{\infty} \lambda_{Xi}^{-1/2} \langle f, \theta_i \rangle \theta_i, f \in F_{XX}\},$$

we find that  $R_{XX}^{1/2}$  is a one-to-one mapping from the vector space  $F_{XX}^{-1} \subset L_2(T_1)$  onto the vector space  $F_{XX}$ . Thus restricting the domain of the operator  $R_{XX}^{1/2}$  to the subset  $F_{XX}^{-1}$ , we can define its inverse for  $f \in F_{XX}$  as

$$R_{XX}^{-1/2}f = \sum_{i=1}^{\infty} \lambda_{Xi}^{-1/2} \langle f, \theta_i \rangle \theta_i.$$

Then  $R_{XX}^{-1/2}$  satisfies the usual properties of an inverse in the sense that

$$R_{XX}^{1/2} R_{XX}^{-1/2} f = f, \text{ for all } f \in F_{XX}, \text{ and } R_{XX}^{-1/2} R_{XX}^{1/2} h = h, \text{ for all } h \in F_{XX}^{-1}.$$

Similarly, we define subspaces  $F_{YY}$ ,  $F_{YY}^{-1} \subset L_2(T_2)$  for  $R_{YY}^{1/2}$ , and define its inverse as

$$(13) \quad R_{YY}^{-1/2}f = \sum_{i=1}^{\infty} \lambda_{Yi}^{-1/2} \langle f, \varphi_i \rangle \varphi_i, \text{ for } f \in F_{YY}.$$

This process is reminiscent of finding a generalized inverse of a matrix.

Denote the adjoint operator of an operator  $A$  by  $A^*$ . The following condition will ensure that  $\text{dom } R = F_{YY}$  and  $\text{dom } R^* = F_{YY}$ , where  $\text{dom}$  refers to the domain on which these operators are defined.

**Condition 4.1** For  $F_{XX}$  and  $F_{YY}$  defined as above, let

$$R_{XY}R_{YY}^{-1/2}(F_{YY}) \subseteq F_{XX}; \quad R_{YX}R_{XX}^{-1/2}(F_{XX}) \subseteq F_{YY}.$$

In order to find sufficient assumptions on processes  $X$  and  $Y$  which imply that Condition 4.1 holds, we will use the Karhunen-Loève representations of square integrable processes given by

$$(14) \quad X(t) = \mu_X(t) + \sum_{i=1}^{\infty} \xi_i \theta_i(t), \quad t \in T_1, \quad \text{and} \quad Y(s) = \mu_Y(s) + \sum_{i=1}^{\infty} \zeta_i \phi_i(s), \quad s \in T_2.$$

Here, the  $\xi_i$  are uncorrelated r.v.'s with  $E\xi_i = 0$ ,  $E\xi_i^2 = \lambda_{Xi}$ , and the  $\zeta_i$  are uncorrelated r.v.'s with  $E\zeta_i = 0$ ,  $E\zeta_i^2 = \lambda_{Yi}$ , such that  $\sum_i \lambda_{Xi} < \infty$ ,  $\sum_i \lambda_{Yi} < \infty$ . The functions  $\{\theta_i, i = 1, 2, \dots\}$  and  $\{\phi_i, i = 1, 2, \dots\}$  are eigenfunctions of the covariance operators  $R_{XX}$  and  $R_{YY}$ , and as such are orthonormal.

**Proposition 4.2** Let  $X$  and  $Y$  be  $L_2$ -processes with Karhunen-Loève expansions as given by (14). Then Condition 4.1 holds if

$$(15) \quad \sum_{i,j=1}^{\infty} (E[\xi_i^2]E[\zeta_j^2])^{-1}(E[\xi_i\zeta_j])^2 < \infty, \quad \text{i.e.,} \quad \sum_{i,j=1}^{\infty} \text{Corr}^2(\xi_i, \zeta_j) < \infty.$$

We note that these are sufficient conditions, but are not necessary.

Let

$$R_0 = R^*R,$$

and  $\lambda_1 \geq \lambda_2 \geq \dots > 0$  be the (positive) eigenvalues of  $R_0$  with corresponding orthonormal eigenfunctions  $q_1, q_2, \dots$ , where  $q_i \in F_{YY}$ . Define

$$(16) \quad p_i = Rq_i / \sqrt{\lambda_i} \quad i = 1, 2, \dots$$

It is well known that canonical correlations and weight functions are obtained in the finite dimensional case by

$$(17) \quad \rho_i = \sqrt{\lambda_i}, \quad u_i = R_{XX}^{-1/2} p_i, \quad \text{and} \quad v_i = R_{YY}^{-1/2} q_i, \quad \text{for } i \geq 1.$$

However, in the infinite-dimensional case, the weight functions are not well defined in  $L_2$  whenever  $p_i \notin F_{XX}$  or  $q_i \notin F_{YY}$ . The following examples illustrate this problem. We use here and in the following the tensor notation, where an operator  $\theta \otimes \phi : H \rightarrow H$  is given by

$$(18) \quad (\theta \otimes \varphi)(h) = \langle h, \theta \rangle \varphi, \quad \text{for } h \in H.$$

**Example 4.3** Consider the case where  $X$  and  $Y$  have the same eigenfunctions in their Karhunen-Loève expansions, and suppose that coefficients of different indices are uncorrelated (this example was suggested by a referee). In this case,

$$R_{XX} = \sum_{i=1}^{\infty} \lambda_{Xi} \theta_i \otimes \theta_i, \quad R_{YY} = \sum_{i=1}^{\infty} \lambda_{Yi} \theta_i \otimes \theta_i, \quad R_{XY} = \sum_{i=1}^{\infty} E(\xi_i \zeta_i) \theta_i \otimes \theta_i,$$

and

$$R = R_{XX}^{-1/2} R_{XY} R_{YY}^{-1/2} = \sum_{i=1}^{\infty} \frac{E(\xi_i \zeta_i)}{(E\xi_i^2 E\zeta_i^2)^{1/2}} \theta_i \otimes \theta_i.$$

In order for  $R$  to be a well-defined Hilbert-Schmidt operator, we require

$$\sum_{i=1}^{\infty} \text{Corr}^2(\xi_i, \zeta_i) < \infty,$$

i.e., (15) is required already for the existence of the operator  $R$  and well-defined canonical correlations  $\rho_i = \text{Corr}(\xi_i, \zeta_i)$ . If (15) is not satisfied, then  $R$  is an unbounded operator and the canonical correlation are not all well-defined. If (15) is satisfied and  $R$  is a Hilbert-Schmidt operator. If in addition, (21) below is also satisfied, then the canonical weight functions will be  $u_i = \lambda_{Xi}^{-1/2} \theta_i$ ,  $v_i = \lambda_{Yi}^{-1/2} \theta_i$ . That these are well-defined will be ensured by Theorem 4.8.

**Example 4.4** Assume processes  $X$  and  $Y$  have Karhunen-Loève expansions (14) where the random variables  $\xi_i, \zeta_j$  satisfy

$$(19) \quad \lambda_{Xi} = E[\xi_i^2] = \frac{1}{i^2}, \quad \lambda_{Yj} = E[\zeta_j^2] = \frac{1}{j^2},$$

and let

$$(20) \quad E[\xi_i \zeta_j] = \frac{1}{(i+1)^2(j+1)^2}, \quad \text{for } i, j \geq 1.$$

We show in Section 6 that equations (19) and (20) can be satisfied by a pair of processes with appropriate operators  $R_{XX}$ ,  $R_{YY}$ , and  $R_{XY}$ . Then

$$\text{Corr}(\xi_i, \zeta_j) = \frac{E[\xi_i \zeta_j]}{\sqrt{E[\xi_i^2] E[\zeta_j^2]}} = \frac{ij}{(i+1)^2(j+1)^2}, \quad \text{for } i, j \geq 1,$$

with

$$\begin{aligned} \sum_{i,j=1}^{\infty} \text{Corr}^2(\zeta_i, \zeta_j) &= \sum_{i,j=1}^{\infty} \frac{i^2 j^2}{(i+1)^4 (j+1)^4} \\ &= \left( \sum_{i=1}^{\infty} \frac{i^2}{(i+1)^4} \right)^2 < \left( \sum_{i=2}^{\infty} \frac{1}{i^2} \right)^2 = \left( \frac{\pi^2}{6} - 1 \right)^2. \end{aligned}$$

Defining

$$c^2 = \sum_{i=1}^{\infty} \frac{i^2}{(i+1)^4}, \quad p = \frac{1}{c} \sum_{i=1}^{\infty} \frac{i}{(i+1)^2} \theta_i, \quad \text{and} \quad q = \frac{1}{c} \sum_{j=1}^{\infty} \frac{j}{(j+1)^2} \varphi_j,$$

and observing  $\|p\| = \|q\| = 1$ , and  $c^2 < 1$ ,

$$\begin{aligned} R &= \sum_{i,j=1}^{\infty} \text{Corr}(\zeta_i, \zeta_j) \theta_i \otimes \varphi_j = \sum_{i,j=1}^{\infty} \frac{ij}{(i+1)^2 (j+1)^2} \theta_i \otimes \varphi_j \\ &= \left( \sum_{i=1}^{\infty} \frac{i}{(i+1)^2} \theta_i \right) \otimes \left( \sum_{j=1}^{\infty} \frac{j}{(j+1)^2} \varphi_j \right) = c^2 p \otimes q. \end{aligned}$$

Since  $R^* R p = c^4 q \otimes q$ , we have

$$\lambda_1 = c^4, \quad \rho_1 = \sqrt{\lambda_1} = c^2, \quad p_1 = p, \quad \text{and} \quad q_1 = q.$$

However,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{\lambda_{X_i}} \langle p, \theta_i \rangle^2 = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{i^{-2}} \left( \frac{1}{c^2} \frac{i^2}{(i+1)^4} \right) = \infty,$$

which means that  $p_1 \notin F_{XX}$ , i.e.,  $u_1$  is not well defined within  $L_2$ .

The following condition is seen to guarantee the existence of well-defined weight functions.

**Condition 4.5** Let  $X$  and  $Y$  be  $L_2$ -processes which satisfy

$$(21) \quad (a) \quad \sum_{i,j=1}^{\infty} \frac{E^2[\zeta_i \zeta_j]}{\lambda_{X_i}^2 \lambda_{Y_j}} < \infty, \quad \text{and} \quad (b) \quad \sum_{i,j=1}^{\infty} \frac{E^2[\zeta_i \zeta_j]}{\lambda_{X_i} \lambda_{Y_j}^2} < \infty.$$

The following example demonstrates that Condition 4.5 is not satisfied by the processes  $X$  and  $Y$  of Example 4.3.

**Example 4.6** Let  $X$  and  $Y$  be the processes defined in Example 4.4. Then,

$$\frac{E^2[\xi_i \zeta_j]}{\lambda_{X_i}^2 \lambda_{Y_j}} = \frac{i^4 j^2}{(i+1)^4 (j+1)^4} > \frac{1}{j^2}, \text{ for } i, j \geq 1,$$

and

$$\frac{E^2[\xi_i \zeta_j]}{\lambda_{X_i} \lambda_{Y_j}^2} = \frac{i^2 j^4}{(i+1)^4 (j+1)^4} > \frac{1}{i^2}, \text{ for } i, j \geq 1,$$

so that Condition 4.5 is not satisfied for these processes.

**Example 4.7** Processes  $\tilde{X}$ ,  $\tilde{Y}$  which do satisfy Condition 4.5 are given via

$$\lambda_{\tilde{X}_i} = E[\tilde{\xi}_i^2] = \frac{1}{i^2}, \quad \lambda_{\tilde{Y}_j} = E[\tilde{\zeta}_j^2] = \frac{1}{j^2},$$

and

$$E[\tilde{\xi}_i \tilde{\zeta}_j] = \frac{1}{(i+1)^3 (j+1)^3}, \text{ for } i, j \geq 1.$$

One can show (Example 4.4), that such a pair of processes  $\tilde{X}$ ,  $\tilde{Y}$  exists. Then

$$\frac{E^2[\tilde{\xi}_i \tilde{\zeta}_j]}{\lambda_{\tilde{X}_i}^2 \lambda_{\tilde{Y}_j}} = \frac{i^4 j^2}{(i+1)^6 (j+1)^6} < \frac{1}{(i+1)^2 (j+1)^4},$$

and

$$\frac{E^2[\tilde{\xi}_i \tilde{\zeta}_j]}{\lambda_{\tilde{X}_i} \lambda_{\tilde{Y}_j}^2} = \frac{i^2 j^4}{(i+1)^6 (j+1)^6} < \frac{1}{(i+1)^4 (j+1)^2},$$

so that Condition 4.5 is satisfied.

Canonical correlations and weight functions exist under Condition 4.5, according to the following central result.

**Theorem 4.8** *Assume that  $L_2$ -processes  $X$  and  $Y$  satisfy Condition 4.5. Let  $(\lambda_i, q_i)$ ,  $i \geq 1$  be the  $i$ -th non-zero eigenvalue and eigenfunction of the operator  $R_0 = R^*R$ , where  $R = R_{XX}^{-1/2} R_{XY} R_{YY}^{-1/2}$  (see (10)). Defining  $p_i = Rq_i / \sqrt{\lambda_i}$ , the following holds:*

- (a)  $p_i \in F_{XX}$  and  $q_i \in F_{YY}$ ,  $i \geq 1$ ;
- (b)  $\rho_i = \sqrt{\lambda_i}$ ,  $u_i = R_{XX}^{-1/2} p_i$ , and  $v_i = R_{YY}^{-1/2} q_i$ ;
- (c)  $\text{Corr}(U_i, U_j) = \langle u_i, R_{XX} u_j \rangle = \langle p_i, p_j \rangle = \delta_{ij}$ ;
- (d)  $\text{Corr}(V_i, V_j) = \langle v_i, R_{YY} v_j \rangle = \langle q_i, q_j \rangle = \delta_{ij}$ ;

$$(e) \quad \text{Corr}(U_i, V_j) = \langle u_i, R_{XY} v_j \rangle = \langle p_i, R q_j \rangle = \rho_i \delta_{ij}.$$

According to Theorem 4.8, the usual properties of canonical correlations and canonical weights known from multivariate analysis (see (17)) extend to functional canonical analysis for  $L_2$ -processes, if Condition 4.5 is satisfied.

## 5. Canonical decomposition for pairs of random processes

Functional principal components analysis is based on the expansion of an  $L_2$ -process in terms of the eigenfunctions of its covariance operator, according to the Karhunen-Loève Theorem, extending principal components in multivariate analysis to function space. A similar extension of the finite-dimensional case to the case of  $L_2$ -processes was considered by Leurgans et al. (1993), Section 4.3, motivated by expanding  $(X, Y)$  as  $(\sum_i U_i u_i, \sum_i V_i v_i)$ , in terms of the canonical weight functions  $u_i, v_i$ . According to Theorem 5.1 below, an expansion of this type cannot be expected to converge in  $L_2$ .

An expansion of pairs of processes in terms of canonical weight functions is desirable, as it provides a natural approximation and description of pairs of square integrable processes. We discuss here the feasibility of this expansion for the functional case.

Since  $R_{XX}^{1/2} u_i = p_i$ ,  $R_{YY}^{1/2} v_i = q_i$ , we have

$$R_{XX} \sum U_i u_i = R_{XX} \sum \langle R_{XX}^{-1/2} p_i, X \rangle R_{XX}^{-1/2} p_i = R_{XX} \sum \langle p_i, R_{XX}^{-1/2} X \rangle p_i = X,$$

and  $R_{YY} \sum V_i v_i = Y$  for the finite dimensional case. This heuristic leads to the following extension for the infinite-dimensional case:

**Theorem 5.1 (Canonical Decomposition)** *Let  $\{(\rho_i, u_i, v_i, U_i, V_i), i \geq 1\}$  be the canonical components of  $X$  and  $Y$  with  $u_i \in L_2(T_1)$  and  $v_i \in L_2(T_2)$  for all  $i$ , and  $p_i, q_i$  defined as in (16). For  $k \geq 1$ , define projections  $P_k = P_{\text{span}\{p_1, \dots, p_k\}}$  and  $Q_k = P_{\text{span}\{q_1, \dots, q_k\}}$ , where  $\text{span}\{p_1, \dots, p_k\}$  and  $\text{span}\{q_1, \dots, q_k\}$  are the closed linear spaces generated by  $\{p_1, \dots, p_k\}$  and  $\{q_1, \dots, q_k\}$ . Define  $\bar{F}_{XX}, \bar{F}_{YY}$  to be the closures of subspaces  $F_{XX}, F_{YY}$  (see (11)) and  $P_{\bar{F}_{XX}}, P_{\bar{F}_{YY}}$  to be the projection operators to the subspaces. Then*

(a)  $X = X_{c,k} + X_{c,k}^\perp$ ,  $Y = Y_{c,k} + Y_{c,k}^\perp$  where

$$X_{c,k} = R_{XX}^{1/2} P_k R_{XX}^{-1/2} X = \sum_{i=1}^k U_i R_{XX} u_i, \quad \text{and} \quad X_{c,k}^\perp = (P_{\bar{F}_{XX}} - R_{XX}^{1/2} P_k R_{XX}^{-1/2}) X,$$

$$Y_{c,k} = R_{YY}^{1/2} Q_k R_{YY}^{-1/2} Y = \sum_{i=1}^k V_i R_{YY} v_i, \quad \text{and} \quad Y_{c,k}^\perp = (P_{\bar{F}_{YY}} - R_{YY}^{1/2} Q_k R_{YY}^{-1/2}) Y;$$

(b)  $(X_{c,k}, Y_{c,k})$  and  $(X, Y)$  share the same first  $k$  canonical components;

(c) The  $j$ -th canonical components of  $(X_{c,k}, Y_{c,k})$  are the same as the  $(j+k)$ -th canonical components of  $(X, Y)$ ,  $j \geq 1$ ;

(d)  $(X_{c,k}, Y_{c,k})$  and  $(X_{c,k}^\perp, Y_{c,k}^\perp)$  are uncorrelated, that is,

$$\text{Corr}(X_{c,k}, X_{c,k}^\perp) = 0, \text{Corr}(Y_{c,k}, Y_{c,k}^\perp) = 0, \text{Corr}(X_{c,k}, Y_{c,k}^\perp) = 0, \text{Corr}(Y_{c,k}, X_{c,k}^\perp) = 0;$$

(e) If  $k \rightarrow \infty$ , and  $P_\infty = P_{\text{span}\{p_i; i \geq 1\}}$ , then

$$X = X_{c,\infty} + X_{c,\infty}^\perp = \sum_{i=1}^{\infty} U_i R_{XX} u_i + (P_{\bar{F}_{XX}} - R_{XX}^{1/2} P_\infty R_{XX}^{-1/2}) X,$$

$$Y = Y_{c,\infty} + Y_{c,\infty}^\perp = \sum_{i=1}^{\infty} V_i R_{YY} v_i + (P_{\bar{F}_{YY}} - R_{YY}^{1/2} Q_\infty R_{YY}^{-1/2}) Y.$$

Further,  $(X_{c,\infty}, X_{c,\infty}^\perp)$  and  $(X, Y)$  share the same canonical components,  $\text{Corr}(X_{c,\infty}^\perp, Y_{c,\infty}^\perp) = 0$ , and  $(X_{c,\infty}, Y_{c,\infty})$  and  $(X_{c,\infty}^\perp, Y_{c,\infty}^\perp)$  are uncorrelated;

(f). If  $\text{span}\{p_i; i \geq 1\} = \bar{F}_{XX}$  and  $\text{span}\{q_i; i \geq 1\} = \bar{F}_{YY}$ , then  $X = X_{c,\infty}$  and  $Y = Y_{c,\infty}$ .

The proof is in Section 6.

The canonical decomposition introduced in Theorem 5.1 may be applied to derive estimation procedures for functional data analysis of pairs of curve data by approximating pairs of processes with a few significant canonical components, similar to approximating random vectors or random functions through a few principal components. The proposed canonical decomposition may also be useful in functional linear model settings. Such applications are described in He, Müller, and Wang (2001).

We conclude this section with two examples of applications of this canonical decomposition. The first example concerns the construction of bivariate processes with prescribed canonical components. Such constructions are useful for the study of functional canonical correlation and dependence between stochastic processes, and especially for Monte Carlo simulations.

**Example 5.2** Let  $\{p_i\} \subset L_2(T_1)$ ,  $\{q_i\} \subset L_2(T_2)$  be two given orthonormal systems. Assume  $\{U_i\}$  and  $\{V_i\}$  are two sets of independent random variables, such that

$$EU_i = 0, EU_i^2 = 1, EV_i = 0, \text{ and } EV_i^2 = 1, \text{ for } i \geq 1,$$

and

$$EU_i V_j = \rho_i \delta_{ij}, \text{ for } i, j \geq 1, \text{ where } \sum_i \rho_i^2 < \infty.$$

Further, let  $\{\lambda_{1i}\}$  and  $\{\lambda_{2i}\}$  be two positive and decreasing sequences such that  $\sum_i \lambda_{1i} < \infty$  and  $\sum_i \lambda_{2i} < \infty$ . We now obtain random processes  $X$  and  $Y$  which have the given

canonical components  $(\rho_i, U_i, V_i)$ , for  $i \geq 1$ , as well as covariance operators  $R_{11} = \sum_i \lambda_{1i} p_i \otimes p_i$  and  $R_{22} = \sum_i \lambda_{2i} q_i \otimes q_i$ , by setting

$$(23) \quad X = \sum_i U_i \sqrt{\lambda_{1i}} p_i, \quad Y = \sum_i V_i \sqrt{\lambda_{2i}} q_i.$$

In order to verify this, observe

$$r_{XX}(s, t) = E[X(s)X(t)] = \sum_{i,j} E[U_i U_j] \sqrt{\lambda_{1i} \lambda_{1j}} p_i(s) p_j(t) = \sum_i \lambda_{1i} p_i(s) p_i(t),$$

and conclude that  $R_{XX} = R_{11}$ . Similarly,  $R_{YY} = R_{22}$ . Furthermore,

$$r_{XY}(s, t) = E[X(s)Y(t)] = \sum_i \rho_i \sqrt{\lambda_{1i} \lambda_{2i}} p_i(s) q_i(t),$$

leading to

$$R_{XY} = R_{XX}^{1/2} [\sum_i \rho_i (p_i \otimes q_i)] R_{YY}^{1/2} = R_{XX}^{1/2} R R_{YY}^{1/2},$$

where

$$R = \sum_i \rho_i (p_i \otimes q_i).$$

This shows that the processes (23) have the required properties. Obviously,  $p_i \in H_1$  and  $q_i \in H_2$ ,  $i \geq 1$ , and thus the canonical weight functions are

$$u_i = R_1^{-1/2} p_i = p_i / \sqrt{\lambda_{1i}}, \quad v_i = R_2^{-1/2} q_i = q_i / \sqrt{\lambda_{2i}}.$$

Hence, comparing (23) and Theorem 4.1 (e), we have

$$X_{c, \infty} = X \quad \text{and} \quad Y_{c, \infty} = Y.$$

Finally, we note that expansion (23) coincides with Karhunen-Loève expansions (14) for processes  $X, Y$  with  $\xi_i = U_i \sqrt{\lambda_{1i}}$ , and  $\zeta_i = V_i \sqrt{\lambda_{2i}}$ , for the case where  $p_i, q_i$  are the eigenfunctions of  $X, Y$ , respectively.

Next, we consider processes with finite basis expansions. This illustrates the finite-dimensional special case of the general result.

**Example 5.3** (*Canonical Decomposition for Processes with Finite Base Functions*) Let  $X$  and  $Y$  be a pair of random processes which have a representation (14) with finitely many basis functions. Then the canonical decomposition for  $X$  and  $Y$  is equivalent to the canonical

decomposition for the corresponding vectors of random coefficients through

$$X(t) = X_{c,k}(t) + X_{c,k}^\perp(t) = \boldsymbol{\xi}_{c,k}^T \boldsymbol{\theta}(t) + \boldsymbol{\xi}_{c,k}^{\perp T} \boldsymbol{\theta}(t),$$

$$Y(t) = Y_{c,k}(t) + Y_{c,k}^\perp(t) = \boldsymbol{\zeta}_{c,k}^T \boldsymbol{\varphi}(t) + \boldsymbol{\zeta}_{c,k}^{\perp T} \boldsymbol{\varphi}(t),$$

for  $1 \leq k \leq \min(k_1, k_2)$ . Here

$$\boldsymbol{\xi} = \boldsymbol{\xi}_{c,k} + \boldsymbol{\xi}_{c,k}^\perp \quad \text{and} \quad \boldsymbol{\zeta} = \boldsymbol{\zeta}_{c,k} + \boldsymbol{\zeta}_{c,k}^\perp$$

are the canonical decompositions for  $\boldsymbol{\xi}$  and  $\boldsymbol{\zeta}$ . Letting  $R_{11} = \text{Cov}(\boldsymbol{\xi}_{c,k})$ ,  $R_{22} = \text{Cov}(\boldsymbol{\zeta}_{c,k})$  and denoting the canonical components of the random vectors  $(\boldsymbol{\xi}_{c,k}, \boldsymbol{\zeta}_{c,k})$  by  $(\rho_i, u_i, v_i)$ ,  $i = 1, \dots, k$ , the corresponding approximations for  $X, Y$  become

$$X_{c,k}(t) = \sum_{i=1}^k (\mathbf{u}_i^T \boldsymbol{\xi})(\mathbf{u}_i^T R_{11} \boldsymbol{\theta})(t), \quad \text{and} \quad Y_{c,k}(t) = \sum_{i=1}^k (\mathbf{v}_i^T \boldsymbol{\zeta})(\mathbf{v}_i^T R_{22} \boldsymbol{\varphi})(t).$$

## 6. Auxiliary results and proofs

We first establish several auxiliary results that will be needed for the proofs in this section. The definition and properties of Hilbert-Schmidt operators are included in the Appendix.

**Proposition 6.1** *If processes  $X$  and  $Y$  satisfy (15), then there exists a Hilbert-Schmidt operator  $A: L_2(T_2) \rightarrow L_2(T_1)$  such that  $A|_{F_{YY}} = R$  and  $A^*|_{F_{XX}} = R^*$ .*

**Proof of Proposition 6.1** Define an operator by the infinite matrix  $(r_{ij})_{i,j \geq 1}$  on the closed subspaces  $\text{span}\{\theta_i\}$  and  $\text{span}\{\varphi_i\}$ , with

$$(24) \quad r_{ij} = \frac{E[\xi_i \zeta_j]}{\sqrt{E[\xi_i^2]E[\zeta_j^2]}}, \quad \text{for } i, j \geq 1.$$

The inequality (15) implies that  $A$  is a Hilbert-Schmidt operator from the closed subspace  $\text{span}\{\theta_i : i \geq 1\} \subset L_2(T_1)$  into  $L_2(T_2)$ . We extend  $\text{dom}(A)$  to  $L_2(T_1)$  such that  $A|_{\text{span}\{\theta_i, i \geq 1\}^\perp} = 0$ , thus ensuring that the extended operator  $A$  is a Hilbert-Schmidt operator on  $L_2(T_1)$ . The proof for  $A^*$  is similar.  $\diamond$

### Proposition 6.2

(a) *Let  $X_n$  and  $Y_n$  be versions of processes  $X$  and  $Y$  which are truncated at the  $n$ -th component, that is*

$$(25) \quad X_n(s) = \sum_{i=1}^n \xi_i \theta_i(s), \quad s \in T_1, \quad \text{and} \quad Y_n(s) = \sum_{i=1}^n \zeta_i \phi_i(s), \quad s \in T_2,$$

and let  $R_n = (r_{ij})_{i,j=1,\dots,n}$ , be the cross-correlation matrix for  $X_n$  and  $Y_n$ . If  $\lim_{n \rightarrow \infty} \sum_{i,j=1}^n r_{ij}^2 < \infty$ , then  $R_n \rightarrow R$ , with respect to the linear operator norm, where  $R$  is a Hilbert-Schmidt operator.

(b) A process  $X$  has a finite basis expansion if  $X = X_n$  (25) for some  $n$ . If  $X$  or  $Y$  have finite basis expansions, then  $R$  is a Hilbert-Schmidt operator.

(c) If (15) holds, then  $R^*$ ,  $RR^*$ ,  $R^*R$  are Hilbert-Schmidt operators.

**Proof of Proposition 6.2** (a), (b), and (c) can be verified by checking inequality (15) of Proposition 4.2.  $\diamond$

**Proposition 6.3** Assume inequality (15) holds.

(a) Let  $\{(\lambda_i, q_i); i \geq 1\}$ , be the eigenvalues and eigenfunctions for the operator  $R_0 = R^*R$ . Then  $\{(\lambda_i, p_i); i \geq 1\}$ , defined by (16), is the set of eigenvalues and eigenfunctions for the operator  $RR^*$  and is an orthonormal system in  $\bar{F}_{XX}$ :

(b) Let  $\rho_i = \sqrt{\lambda_i}$ , then  $Rq_i = \rho_i p_i$  and  $R^*p_i = \rho_i q_i$ .

**Proof of Proposition 6.3**

$$(a) \quad \begin{aligned} RR^*(p_i) &= RR^*((1/\sqrt{\lambda_i})Rq_i) = (1/\sqrt{\lambda_i})R(R_0q_i) \\ &= (1/\sqrt{\lambda_i})R(\lambda_i q_i) = \lambda_i(1/\sqrt{\lambda_i})Rq_i = \lambda_i p_i. \end{aligned}$$

Therefore for any  $(\lambda_i, q_i)$ , there exists a corresponding eigenvalue and eigenfunction  $(\lambda_i, p_i)$  of  $RR^*$ , and vice versa. The  $\{p_i\}$  are orthonormal, as

$$\begin{aligned} \langle p_i, p_j \rangle &= \langle (1/\sqrt{\lambda_i})Rq_i, (1/\sqrt{\lambda_j})Rq_j \rangle = (1/\sqrt{\lambda_i \lambda_j}) \langle R^*Rq_i, q_j \rangle \\ &= (1/\sqrt{\lambda_i \lambda_j}) \langle \lambda_i q_i, q_j \rangle = (1/\sqrt{\lambda_i \lambda_j}) \lambda_i \delta_{ij} = \delta_{ij}. \end{aligned}$$

(b)  $R^*p_i = R^*(Rq_i/\rho_i) = (\rho_i^2/\rho_i)q_i = \rho_i q_i$  and  $Rq_i = R(R^*p_i/\rho_i) = \rho_i^2 p_i/\rho_i = \rho_i p_i$ .  $\diamond$

**Proposition 6.4** (Spectral decomposition of  $R$ ) Let  $R = R_{XX}^{-1/2} R_{XY} R_{YY}^{-1/2}$  be the correlation operators for  $X$  and  $Y$  which satisfy inequality (15), and let  $\rho_i$ ,  $p_i$ , and  $q_i$  be the  $i$ -th eigenvalue and eigenvector of  $R^*R$ , and  $RR^*$ , respectively. Then, the integral kernel of  $R$  (see (A1) in Appendix) is decomposed as

$$\text{KER}(R)(s, t) = \sum_{i=1}^{\infty} \rho_i p_i(s) q_i(t).$$

**Proof of Proposition 6.4** From Proposition 6.1,  $R = R_{XX}^{-1/2} R_{XY} R_{YY}^{-1/2} : F_{YY} \rightarrow F_{XX}$  is a

Hilbert-Schmidt operator, and therefore has an  $L_2$  integral kernel. Defining an integral operator  $A: F_{YY} \rightarrow F_{XX}$ , with integral kernel  $\text{KER}(A) = \sum_{i=1}^{\infty} \rho_i p_i(s) q_i(t)$ , we show that  $A = R$ .

First, for any  $v \in \text{span}\{q_i, i \geq 1\}^\perp$ ,

$$(Av)(s) = \int \text{KER}(A)(s,t) v(t) dt = \sum_{i=1}^{\infty} \rho_i p_i(s) \int q_i(t) v(t) dt = 0.$$

Furthermore,

$$\begin{aligned} \|Rv\|^2 &= \langle Rv, Rv \rangle = \langle v, R^* Rv \rangle \\ &= \langle v, R_0 v \rangle = \langle v, \sum_{i=1}^{\infty} \lambda_i P_i v \rangle, \end{aligned}$$

where  $R_0$  is a compact self-adjoint operator with spectral decomposition  $R_0 = \sum_i \lambda_i P_i$ . Here  $P_i$  is the projection operator from  $F_{YY}$  to the eigenspace defined by  $\text{span}\{\text{eigenfunctions corresponding to } \lambda_i\}$ . Thus,  $P_i v = 0$ ,  $i \geq 1$ . This implies  $Rv = 0$ . Next, for any  $j \geq 1$ ,

$$(Aq_j)(s) = \sum_{i=1}^{\infty} \rho_i p_i(s) \int q_i(t) q_j(t) d\mu_2 = \rho_j p_j(s) = (Rq_j)(s).$$

This shows that for  $v \in \text{span}\{q_i, i \geq 1\}$ , one has  $Av = Rv$ , which concludes the proof.  $\diamond$

**Proposition 6.5** *Assume  $L_2$ -processes  $X$  and  $Y$  satisfy Condition 4.5. Then,  $p_i \in F_{XX}$  and  $q_i \in F_{YY}$ , for  $i \geq 1$ .*

**Proof of Proposition 6.5** Write

$$R = \sum_{i,j=1}^{\infty} r_{ij} \theta_i \otimes \varphi_j,$$

where,  $r_{ij} = \sum_{i,j=1}^{\infty} E[\xi_i \zeta_j] (\lambda_{X_i} \lambda_{Y_j})^{-1/2}$ , for  $i, j \geq 1$ . Since Condition 4.5 implies (15),  $R$  is a Hilbert-Schmidt operator and can be written as

$$R = \sum_{i=1}^{\infty} \rho_i p_i \otimes q_i,$$

where  $(\rho_i^2, p_i)$  and  $(\rho_i^2, q_i)$  are the eigenvalues and eigenvectors of  $RR^*$  and  $R^*R$ , respectively (Conway, 1985). Then, for any fixed  $k \geq 1$ ,

$$Rq_k = \rho_k p_k = \sum_{i,j=1}^{\infty} r_{ij} \langle q_k, \varphi_j \rangle \theta_i.$$

From the definition of  $F_{XX}$ ,

$$p_k \in F_{XX} \text{ iff } \sum_{i=1}^{\infty} \lambda_{X_i}^{-1} \left( \sum_{j=1}^{\infty} r_{ij} \langle q_k, \varphi_j \rangle \right)^2 < \infty.$$

The right hand side is true because

$$\begin{aligned} \sum_{i=1}^{\infty} \lambda_{X_i}^{-1} \left( \sum_{j=1}^{\infty} r_{ij} \langle q_k, \varphi_j \rangle \right)^2 &= \sum_{i=1}^{\infty} \lambda_{X_i}^{-2} \left( \sum_{j=1}^{\infty} \lambda_{Y_j}^{-1/2} E[\xi_i \zeta_j] \langle q_k, \varphi_j \rangle \right)^2 \\ &\leq \sum_{i=1}^{\infty} \lambda_{X_i}^{-2} \left( \sum_{j=1}^{\infty} \lambda_{Y_j}^{-1} E^2[\xi_i \zeta_j] \sum_{j=1}^{\infty} \langle q_k, \varphi_j \rangle^2 \right) < \infty. \end{aligned}$$

The Cauchy inequality and the equality  $\sum_j \langle q_k, \phi_j \rangle^2 = 1$  are used in the derivations of the last two inequalities. By using Condition 4.5 (a), we show  $p_k \in F_{XX}$ . Similarly we prove  $q_k \in F_{YY}$  when Condition 4.5 (b) is satisfied.  $\diamond$

**Proof of Theorem 2.1** The covariance function for  $X$  is

$$r_{XX}(s, t) = E[X(s)X(t)] = \boldsymbol{\theta}^T(s)E[\boldsymbol{\zeta}\boldsymbol{\zeta}^T]\boldsymbol{\theta}(t) = \boldsymbol{\theta}^T(s)R_{11}\boldsymbol{\theta}(t).$$

Similarly, we have  $r_{YY}(s, t) = \boldsymbol{\varphi}^T(s)R_{22}\boldsymbol{\varphi}(t)$ , and  $r_{XY}(s, t) = \boldsymbol{\theta}^T(s)R_{12}\boldsymbol{\varphi}(t)$ . Because the covariance matrices  $R_{11}$  and  $R_{22}$  for the random coefficient vectors are finite dimensional, we may assume, without loss of generality, that they are full rank matrices. Any given  $u \in L^2(T_1)$  can be written as the sum of two components, with  $u_0 \in \text{span}\{\theta_1, \dots, \theta_{k_1}\}$ , and  $u_1 \in \text{span}\{\theta_1, \dots, \theta_{k_1}\}^\perp$ , such that

$$u(t) = u_0(t) + u_1(t) = \mathbf{u}^T \boldsymbol{\theta}(t) + u_1(t), \quad \text{where } \mathbf{u} \in R^{k_1}.$$

Then,

$$\begin{aligned} \langle u, X \rangle &= \langle u_0 + u_1, \boldsymbol{\zeta}^T \boldsymbol{\theta} \rangle = \langle \mathbf{u}^T \boldsymbol{\theta}, \boldsymbol{\zeta}^T \boldsymbol{\theta} \rangle \\ &= \mathbf{u}^T \langle \boldsymbol{\theta}, \boldsymbol{\theta}^T \rangle \boldsymbol{\zeta} = \mathbf{u}^T \mathbf{I}_{k_1} \boldsymbol{\zeta} = \mathbf{u}^T \boldsymbol{\zeta}, \end{aligned}$$

where  $\mathbf{I}_{k_1}$  is the  $k_1 \times k_1$  identity matrix. Similarly for any given  $v \in L^2(T_2)$ ,  $\langle v, Y \rangle = \mathbf{v}^T \boldsymbol{\zeta}$ , where  $\mathbf{v} \in R^{k_2}$ . Hence,

$$E[\langle u, X \rangle] = \mathbf{u}^T E[\boldsymbol{\zeta}] = 0,$$

$$E[\langle v, Y \rangle] = \mathbf{v}^T E[\boldsymbol{\zeta}] = 0,$$

$$\text{Var}(\langle u, X \rangle) = E[\langle u, X \rangle^2] = \mathbf{u}^T E[\boldsymbol{\zeta}\boldsymbol{\zeta}^T] \mathbf{u} = \mathbf{u}^T R_{11} \mathbf{u},$$

$$\text{Var}(\langle v, Y \rangle) = E[\langle v, Y \rangle^2] = \mathbf{v}^T E[\boldsymbol{\zeta}\boldsymbol{\zeta}^T] \mathbf{v} = \mathbf{v}^T R_{22} \mathbf{v},$$

$$\text{Cov}[\langle u, X \rangle, \langle v, Y \rangle] = E[\langle u, X \rangle \langle v, Y \rangle] = \mathbf{u}^T E[\boldsymbol{\zeta}\boldsymbol{\zeta}^T] \mathbf{v} = \mathbf{u}^T R_{12} \mathbf{v}.$$

Consider the first canonical correlation for  $X$  and  $Y$ ,

$$\rho_1 = \sup_{u \in L_2(T_1), v \in L_2(T_2)} \text{Cov}(\langle u, X \rangle, \langle v, Y \rangle) =: \text{Cov}(\langle u_1, X \rangle, \langle v_1, Y \rangle),$$

subject to

$$\text{Var}(\langle u, X \rangle) = 1, \quad \text{and} \quad \text{Var}(\langle v, Y \rangle) = 1.$$

This is seen to be equivalent to

$$\rho_1 = \sup_{\mathbf{u} \in R^{k_1}, \mathbf{v} \in R^{k_2}} \mathbf{u}^T R_{12} \mathbf{v} =: \mathbf{u}_1^T R_{12} \mathbf{v}_1,$$

subject to

$$\mathbf{u}^T R_{11} \mathbf{u} = 1, \quad \mathbf{v}^T R_{22} \mathbf{v} = 1.$$

This is exactly the definition of the first canonical correlation between the random vectors  $\xi$  and  $\zeta$ . On the other hand, if we start with the first canonical components for  $\xi$  and  $\zeta$ , given by  $(\rho_1, \mathbf{u}_1, \mathbf{v}_1)$ , we obtain the first canonical components for  $X(t)$  and  $Y(t)$  by

$$(\rho_1, u_1(t), v_1(t)) = (\rho_1, \mathbf{u}_1^T \boldsymbol{\theta}(t), \mathbf{v}_1^T \boldsymbol{\varphi}(t)).$$

Similarly, we can extend this to the second canonical component, and so on.  $\diamond$

**Proof of Proposition 4.2** To prove the first inclusion, i.e.,  $R_{XY} R_{YY}^{-1/2}(\mathbb{F}_{YY}) \subseteq \mathbb{F}_{XX}$ , we note that  $E[\xi_i^2] = \lambda_{Xi}$ , and  $E[\zeta_j^2] = \lambda_{Yj}$ . From (12) and (14),

$$R_{YY}^{-1/2} \mathbf{v} = \sum_{i=1}^{\infty} \lambda_{Yi}^{-1/2} \langle \mathbf{v}, \boldsymbol{\varphi}_i \rangle \boldsymbol{\varphi}_i, \quad \text{for } \mathbf{v} \in \mathbb{F}_{YY}, \text{ and}$$

$$R_{XY} R_{YY}^{-1/2} \mathbf{u} = \sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} \lambda_{Yj}^{-1/2} E[\xi_i \zeta_j] \langle \mathbf{v}, \boldsymbol{\varphi}_j \rangle \right) \boldsymbol{\theta}_i, \quad \text{for } \mathbf{v} \in \mathbb{F}_{YY}.$$

Thus if (15) holds, then

$$\begin{aligned} \sum_i \lambda_{Xi}^{-1} \left( \sum_j \lambda_{Yj}^{-1/2} E[\xi_i \zeta_j] \langle \mathbf{u}, \boldsymbol{\varphi}_j \rangle \right)^2 &\leq \sum_i \lambda_{Xi}^{-1} \left( \sum_j \lambda_{Yj}^{-1} (E[\xi_i \zeta_j])^2 \langle \mathbf{u}, \boldsymbol{\varphi}_j \rangle^2 \right) \\ &\leq \sum_{ij} \lambda_{Xi}^{-1} \lambda_{Yj}^{-1} (E[\xi_i \zeta_j])^2 \sum_j \langle \mathbf{u}, \boldsymbol{\varphi}_j \rangle^2 \\ &= \sum_{ij} \text{Corr}^2(\xi_i, \zeta_j) \|\mathbf{u}\|^2 < \infty. \end{aligned}$$

The second inclusion follows similarly.  $\diamond$

**Proof for Example 4.3** Set

$$a_i^2 = \frac{1}{i^2}, \quad c_i = \frac{1}{(i+1)^2}, \quad \text{for } i = 1, 2, \dots,$$

and define the operators

$$R_1 = \sum_{i=1}^{\infty} a_i^2 \boldsymbol{\theta}_i \otimes \boldsymbol{\theta}_i, \quad R_2 = \sum_{i=1}^{\infty} a_i^2 \boldsymbol{\varphi}_i \otimes \boldsymbol{\varphi}_i, \quad \text{and } R_3 = \sum_{i,j=1}^{\infty} c_i c_j \boldsymbol{\theta}_i \otimes \boldsymbol{\varphi}_j.$$

The property that  $\sum_i a_i^2 < \infty$  implies that  $R_1$  and  $R_2$  are self-adjoint and positive definite operators, and the property that  $\sum_{ij} c_i^2 c_j^2 < \infty$  implies that  $R_3$  is a Hilbert-Schmidt operator. By using (19), we show that  $R_1 = R_{XX}$ ,  $R_2 = R_{YY}$ . It remains to show that  $R_3$  can be interpreted as a cross-covariance operator  $R_{XY}$ . For any  $n \geq 1$ , define

$A_k = \text{Diag}(a_1, \dots, a_k)$ ,  $B_{nk} = (c_i c_j)_{i=1, \dots, n, j=1, \dots, k} = \mathbf{C}_n \mathbf{C}_k^T$ , for  $k = 1, \dots, n$ ,  
 where  $\mathbf{C}_l = (c_1, \dots, c_l)^T$ , and

$$M_{nk} = \begin{pmatrix} A_n & B_{nk} \\ B_{nk}^T & A_k \end{pmatrix} = \begin{pmatrix} A_n & \mathbf{C}_n \mathbf{C}_k^T \\ \mathbf{C}_k \mathbf{C}_n^T & A_k \end{pmatrix}, \text{ for } k = 1, \dots, n.$$

Then

$$M_{nn} = \begin{pmatrix} A_n & \mathbf{C}_n \mathbf{C}_n^T \\ \mathbf{C}_n \mathbf{C}_n^T & A_n \end{pmatrix} \rightarrow M_\infty = \begin{pmatrix} R_1 & R_3 \\ R_3^* & R_2 \end{pmatrix}, \text{ as } n \rightarrow \infty.$$

We want to show that, for  $n \geq 1$ ,  $M_{nn}$  is positive definite, so that  $M_\infty$  is also positive definite, and therefore defines a covariance operator for the pair of  $L_2$ -processes  $X$  and  $Y$  such that  $R_3 = R_{XY}$ . It will suffice to show that  $\text{Det}(M_{nk}) > 0$ , for  $k = 1, \dots, n$ . As for any  $n \geq 1$ ,

$$\sum_{i=1}^n \frac{c_i^2}{a_i^2} = \sum_{i=1}^n \frac{i^2}{(i+1)^4} < \sum_{i=1}^{\infty} \frac{1}{(i+1)^2} = \frac{\pi^2}{6} - 1 < 1,$$

we find indeed

$$\begin{aligned} \text{Det}(M_{nk}) &= \text{Det}(A_n) \text{Det}(A_k - B_{nk}^T A_n^{-1} B_{nk}) \\ &= \left( \prod_{i=1}^n a_i^2 \right) \text{Det}(A_k - \mathbf{C}_k \mathbf{C}_n^T A_n^{-1} \mathbf{C}_n \mathbf{C}_k^T) \\ &= \left( \prod_{i=1}^n a_i^2 \right) \text{Det}(A_k - \left( \sum_{i=1}^n \frac{c_i^2}{a_i^2} \right) \mathbf{C}_k \mathbf{C}_k^T) \\ &= \left( \prod_{i=1}^n a_i^2 \right) \text{Det}(A_k) \text{Det}\left(1 - \left( \sum_{i=1}^n \frac{c_i^2}{a_i^2} \right) \mathbf{C}_k^T A_k^{-1} \mathbf{C}_k\right) \\ &= \left( \prod_{i=1}^n a_i^2 \right) \left( \prod_{i=1}^k a_i^2 \right) \left[1 - \left( \sum_{i=1}^n \frac{c_i^2}{a_i^2} \right) \left( \sum_{i=1}^k \frac{c_i^2}{a_i^2} \right)\right] > 0. \end{aligned}$$

◇

**Proof of Theorem 4.8** Part (a) follows from Proposition 6.5. The proof for part (b) - (e) can be found in He, Müller, and Wang (2000). ◇

For the proof of Theorem 5.1, we need the following additional auxiliary results.

**Lemma 6.6** *Assume Condition 4.5 is satisfied. Then, for  $k \geq 1$ ,*

(a)  $\langle u, X_{c,k} \rangle = \langle u, X \rangle$  and  $\langle u, X_{c,k}^\perp \rangle = 0$ , if  $\mathbf{R}_{XX}^{1/2} u \in \text{span}\{p_i, i = 1, \dots, k\}$ ;

- (b)  $\langle v, Y_{c,k} \rangle = \langle v, Y \rangle$  and  $\langle v, Y_{c,k}^\perp \rangle = 0$ , if  $R_{YY}^{1/2}v \in \text{span}\{q_i, i = 1, \dots, k\}$ ;
- (c)  $\langle u, X_{c,k} \rangle = 0$  and  $\langle u, X_{c,k}^\perp \rangle = \langle u, X \rangle$ , if  $R_{XX}^{1/2}u \in \text{span}\{p_i, i = 1, \dots, k\}^\perp$ ;
- (d)  $\langle v, Y_{c,k} \rangle = 0$  and  $\langle v, Y_{c,k}^\perp \rangle = \langle v, Y \rangle$ , if  $R_{YY}^{1/2}v \in \text{span}\{q_i, i = 1, \dots, k\}^\perp$ .

*Proof.* We only prove (a) and (c).

(a).  $R_{XX}^{1/2}u \in \text{span}\{p_i, i = 1, \dots, k\}$  is equivalent to  $u \in \text{span}\{u_i, i = 1, \dots, k\}$ , and, for  $i \leq k$ ,

$$\begin{aligned} \langle u_i, X_{c,k} \rangle &= \langle R_{XX}^{-1/2} p_i, \sum_{j=1}^k U_j R_{XX}^{1/2} p_j \rangle = \sum_{j=1}^k U_j \langle p_i, p_j \rangle = U_i = \langle u_i, X \rangle, \\ \langle u_i, X_{c,k}^\perp \rangle &= \langle u_i, X - X_{c,k} \rangle = \langle u_i, X \rangle - \langle u_i, X_{c,k} \rangle = 0. \end{aligned}$$

(c). Let  $R_{XX}^{1/2}u \in \text{span}\{p_1, \dots, p_k\}^\perp$ . Then,

$$\begin{aligned} \langle u, X_{c,k} \rangle &= \langle u, \sum_{j=1}^k U_j R_{XX}^{1/2} u_j \rangle = \sum_{j=1}^k U_j \langle R_{XX}^{1/2} u, R_{XX}^{1/2} u_j \rangle = \sum_{j=1}^k U_j \langle R_{XX}^{1/2} u, p_j \rangle = 0, \\ \langle u_i, X_{c,k}^\perp \rangle &= \langle u_i, X - X_{c,k} \rangle = \langle u_i, X \rangle - \langle u_i, X_{c,k} \rangle = \langle u_i, X \rangle. \end{aligned} \quad \diamond$$

**Lemma 6.7** For  $1 \leq i \leq k$ ,  $u \in F_{XX}^{-1}$ ,  $v \in F_{YY}^{-1}$ ,

$$\begin{aligned} R_{XX}^{1/2}u \in \text{span}\{p_1, \dots, p_k\}^\perp &\text{ iff } \langle u, X \rangle \text{ and } \langle u_i, X \rangle \text{ are uncorrelated; and,} \\ R_{YY}^{1/2}v \in \text{span}\{q_1, \dots, q_k\}^\perp &\text{ iff } \langle v, Y \rangle \text{ and } \langle v_i, Y \rangle \text{ are uncorrelated.} \end{aligned}$$

*Proof.* Consequence of

$$\begin{aligned} \text{Corr}(\langle u, X \rangle, \langle u_i, X \rangle) &= E[\langle u, X \rangle \langle u_i, X \rangle] = \langle u, R_{XX} u_i \rangle \\ &= \langle R_{XX}^{1/2} u, R_{XX}^{1/2} u_i \rangle = \langle R_{XX}^{1/2} u, p_i \rangle. \end{aligned} \quad \diamond$$

### Proof of Theorem 5.1

(a) Note that for  $i \geq 1$ ,  $f \in F_{XX}$ ,

$$\begin{aligned} (p_i \otimes p_i) R_{XX}^{-1/2}(f) &= \langle p_i, R_{XX}^{-1/2}(f) \rangle p_i = \langle R_{XX}^{-1/2}(p_i), f \rangle p_i \\ &= R_{XX}^{-1/2}(p_i) \otimes p_i(f) = u_i \otimes p_i(f), \end{aligned}$$

that is,  $(p_i \otimes p_i) R_{XX}^{-1/2} = u_i \otimes p_i$  on  $F_{XX}$ . Hence,

$$P_k R_{XX}^{-1/2} = \sum_{i=1}^k (p_i \otimes p_i) R_{XX}^{-1/2} = \sum_{i=1}^k u_i \otimes p_i \text{ on } F_{XX}.$$

Since  $u_i \in L_2(T_1)$ ,  $i \geq 1$ , the operator  $\sum_{i=1}^k u_i \otimes p_i$  has the  $L_2$  kernel  $\sum_{i=1}^k p_i(s) u_i(t)$ , and therefore is a Hilbert-Schmidt operator on  $L_2(T_1)$ . As

$$\sup_{f \in F_{XX}} \|P_k R_{XX}^{-1/2} f\| = \sup_{f \in F_{XX}} \left\| \sum_{i=1}^k (u_i \otimes p_i) f \right\| \leq \left\| \sum_{i=1}^k u_i \otimes p_i \right\| \|f\|,$$

$P_k R_{XX}^{-1/2}$  can be extended to a bounded operator on  $F_{XX}$ , and further to a Hilbert-Schmidt operator on  $L_2(T_1)$ . Now

$$P_k R_{XX}^{-1/2} X = \sum_{i=1}^k \langle u_i, X \rangle p_i = \sum_{i=1}^k U_i R_{XX}^{1/2} u_i,$$

and this implies

$$X_{c,k} = R_{XX}^{1/2} P_k R_{XX}^{-1/2} X = \sum_{i=1}^k U_i R_{XX} u_i, \quad X_{c,k}^\perp = X - X_{c,k}.$$

(b) For  $\tilde{u} \in \bar{F}_{XX}$ , write  $\tilde{u} = P_k R_{XX}^{1/2} \tilde{u} + (P_{\bar{F}_{XX}} - P_k R_{XX}^{1/2}) \tilde{u} = \tilde{u}_1 + \tilde{u}_2$ , with  $\tilde{u}_1 \in \text{span}\{p_1, \dots, p_k\}$ ,  $\tilde{u}_2 \in \text{span}\{p_1, \dots, p_k\}^\perp$ . Similarly, for  $\tilde{v} \in \bar{F}_{YY}$ , write  $\tilde{v} = \tilde{v}_1 + \tilde{v}_2$ . From Lemma 6.6,

$$\langle \tilde{u}, X_{c,k} \rangle = \langle \tilde{u}_1, X_{c,k} \rangle = \langle \tilde{u}_1, X \rangle, \quad \text{and} \quad \langle \tilde{v}, Y_{c,k} \rangle = \langle \tilde{v}_1, Y_{c,k} \rangle = \langle \tilde{v}_1, Y \rangle,$$

and

$$\begin{aligned} & \sup_{\tilde{u} \in L_2(T_1), \tilde{v} \in L_2(T_2)} \text{Cov}(\langle \tilde{u}, X_{c,k} \rangle, \langle \tilde{v}, Y_{c,k} \rangle) \\ (26) \quad &= \sup_{\tilde{u}_1 \in \text{span}\{u_1, \dots, u_k\}, \tilde{v}_1 \in \text{span}\{v_1, \dots, v_k\}} \text{Cov}(\langle \tilde{u}_1, X \rangle, \langle \tilde{v}_1, Y \rangle) \\ &\leq \sup_{\tilde{u} \in L_2(T_1), \tilde{v} \in L_2(T_2)} \text{Cov}(\langle \tilde{u}, X \rangle, \langle \tilde{v}, Y \rangle). \end{aligned}$$

For  $i = 1, \dots, k$ ,

$$\langle u_i, X_{c,k} \rangle = \sum_{j=1}^k U_j \langle u_i, R_{XX} u_j \rangle = \sum_{j=1}^k U_j \delta_{ij} = U_i, \quad \langle v_i, Y_{c,k} \rangle = V_i,$$

and

$$\text{Cov}(\langle u_i, X_{c,k} \rangle, \langle v_i, Y_{c,k} \rangle) = E[U_i V_i] = \rho_i = \sup_{\tilde{u} \in L_2(T_1), \tilde{v} \in L_2(T_2)} \text{Cov}(\langle \tilde{u}, X \rangle, \langle \tilde{v}, Y \rangle).$$

This implies

$$\text{Var}(\langle u_i, X_{c,k} \rangle) = 1 \quad \text{and} \quad \text{Var}(\langle v_i, Y_{c,k} \rangle) = 1,$$

and for  $i > 1, j = 1, \dots, i-1$ ,

$$\text{Cov}(\langle u_i, X_{c,k} \rangle, \langle u_j, X_{c,k} \rangle) = 0 \quad \text{and} \quad \text{Cov}(\langle v_i, Y_{c,k} \rangle, \langle v_j, Y_{c,k} \rangle) = 0.$$

Hence,  $(\rho_i, u_i, v_i)$ ,  $i = 1, \dots, k$ , are the  $i$ -th canonical correlation and weight functions for  $X_{c,k}$  and  $Y_{c,k}$ , and are identical to those for  $X$  and  $Y$ .

(c) From Lemma 6.6, for  $\tilde{u} \in \bar{F}_{XX}$ , and  $\tilde{v} \in \bar{F}_{YY}$ . Let  $\tilde{u} = \tilde{u}_1 + \tilde{u}_2$  and  $\tilde{v} = \tilde{v}_1 + \tilde{v}_2$  as in the above. Then,

$$\langle \tilde{u}, X_{c,k}^\perp \rangle = \langle \tilde{u}_2, X_{c,k}^\perp \rangle = \langle \tilde{u}_2, X \rangle, \quad \text{which is uncorrelated with } U_i, \quad i = 1, \dots, k,$$

$$\langle \tilde{v}, Y_{c,k}^\perp \rangle = \langle \tilde{v}_2, Y_{c,k}^\perp \rangle = \langle \tilde{v}_2, Y \rangle, \quad \text{which is uncorrelated with } V_i, \quad i = 1, \dots, k.$$

From constraint (4) and the facts that  $\tilde{u}_2$  is uncorrelated with  $\text{span}\{u_1, \dots, u_k\}$ , and  $\tilde{v}_2$  is uncorrelated with  $\text{span}\{v_1, \dots, v_k\}$ , we have

Can Corr  $(X_{c,k}^\perp, Y_{c,k}^\perp)_i \leq \text{Can Corr}(X, Y)_{i+k}$  for  $i \geq 1$ .

On the other hand, we have that  $\{(u_{i+k}, v_{i+k}) : i \geq 1\}$  satisfy constraints (2) and (4) for  $(X_{c,k}, Y_{c,k})$ , so that

$$\begin{aligned} \text{Can Corr}(X_{c,k}^\perp, Y_{c,k}^\perp)_i &= \text{Cov}(\langle u_{i+k}, X_{c,k}^\perp \rangle, \langle v_{i+k}, Y_{c,k}^\perp \rangle) \\ &= \rho_{i+k} = \text{Can Corr}(X, Y)_{i+k}. \end{aligned}$$

Using a similar argument as for the proof of (b), we have that  $\{(\rho_{i+k}, u_{i+k}, v_{i+k}) : i \geq 1\}$ , the  $(i+k)$ -th canonical correlation and weight functions for  $X$  and  $Y$ , are the  $i$ -th canonical correlation and weight functions for  $X_{c,k}$  and  $Y_{c,k}$ .

(d) We only provide the proof of the third equality, since the proofs of the other equalities are similar. For  $\tilde{u} \in \bar{F}_{XX}$  and  $\tilde{v} \in \bar{F}_{YY}$ , let  $\tilde{u} = \tilde{u}_1 + \tilde{u}_2$  and  $\tilde{v} = \tilde{v}_1 + \tilde{v}_2$  as in the proof for (b). Then,

$$\langle \tilde{u}, X_{c,k} \rangle = \langle \tilde{u}_1, X \rangle, \quad \langle \tilde{v}, Y_{c,k}^\perp \rangle = \langle \tilde{v}_2, Y \rangle,$$

and

$$\text{Cov}(\langle \tilde{u}, X_{c,k} \rangle, \langle \tilde{v}, Y_{c,k}^\perp \rangle) = E[\langle \tilde{u}_1, X \rangle \langle \tilde{v}_2, Y \rangle] = \langle \tilde{p}_1, R\tilde{q}_2 \rangle,$$

where  $\tilde{p}_1 = R_{XX}^{1/2} \tilde{u}_1 \in \text{span}\{p_1, \dots, p_k\}$  and  $\tilde{q}_2 = (I_{\bar{F}_{YY}} - R_{YY}^{1/2}) \tilde{v}_2 \in \text{span}\{q_1, \dots, q_k\}^\perp$ . From Proposition 6.3 (b),  $R\tilde{q}_2 \in \text{span}\{p_1, \dots, p_k\}^\perp$ , and  $\langle \tilde{p}_1, R\tilde{q}_2 \rangle = 0$ .

(e) From (a),

$$X_{c,k} = \sum_{i=1}^k U_i R_{XX}^{1/2} p_i, \quad \text{for } k \geq 1.$$

Now the fact that

$$\sum_{i=1}^k E \|U_i R_{XX}^{1/2} p_i\|^2 = \sum_{i=1}^k E[U_i^2] \|R_{XX}^{1/2} p_i\|^2 = \sum_{i=1}^k \langle p_i, R_{XX}^{1/2} p_i \rangle,$$

and

$$\langle p_i, R_{XX} p_i \rangle = \sum_{j=1}^{\infty} \lambda_{X_j} \langle p_i, \theta_j \rangle^2$$

imply for all  $k \geq 1$ ,

$$\begin{aligned} \sum_{i=1}^k E \|U_i R_{XX}^{1/2} p_i\|^2 &= \sum_{i=1}^k \sum_{j=1}^{\infty} \lambda_{X_j} \langle p_i, \theta_j \rangle^2 = \sum_{j=1}^{\infty} \lambda_{X_j} \left( \sum_{i=1}^k \langle p_i, \theta_j \rangle^2 \right) \\ &= \sum_{j=1}^{\infty} \lambda_{X_j} \|P_k \theta_j\|^2 \leq \sum_{j=1}^{\infty} \lambda_{X_j} \|\theta_j\|^2 = \sum_{j=1}^{\infty} \lambda_{X_j} < \infty. \end{aligned}$$

Hence,

$$X_{c,k} = R_{XX}^{1/2} P_k R_{XX}^{-1/2} \rightarrow R_{XX}^{1/2} P_\infty R_{XX}^{-1/2} X = X_{c,\infty} = \sum_{i=1}^{\infty} U_i R_{XX} u_i, \text{ as } k \rightarrow \infty,$$

where the convergence is in the mean squared norm  $E\|\cdot\|_{L_2}^2$ . This implies (e).

(f) If  $\text{span}\{p_i : i \geq 1\} = \bar{F}_{XX}$ , then  $P_\infty = P_{\bar{F}_{XX}}$ , and  $X_{c,\infty}^\perp = 0$ .  $\diamond$

## Appendix

We compile here some facts from functional analysis, which are listed without proof. Details can be found in texts such as Conway (1985) or Dunford and Schwartz (1988).

We use the tensor product notation (18) throughout. Let  $H_1, H_2$  be Hilbert spaces. We denote the set of bounded linear operators  $A : H_1 \rightarrow H_2$  by  $B(H_1, H_2)$ .

A linear operator is *compact* if the set  $\{Ah \mid h \in H_1, \|h\| = 1\}$  has a compact closure in  $H_2$ . The *kernel* of a linear operator  $A$  is defined as  $\ker(A) = \{h \in H_1 \mid Ah = 0\}$ . A compact self-adjoint operator can be expressed as

$$A = \sum_i \mu_i e_i \otimes e_i,$$

for a sequence of  $\{\mu_i\}$  of real numbers and an orthonormal basis  $\{e_i\}$  of  $\ker(A)^\perp$ . For an operator  $A : L_2 \rightarrow L_2$ , and a subspace  $H \subset L_2$ , denote by  $A|_H$  the restriction to subspace  $H$ .

For  $k \in L_2(T_1 \times T_2)$ , the integral operator  $A : L_2(T_1) \rightarrow L_2(T_2)$  defined by

$$(Af)(t) = \int_{T_1} k(s, t) f(s) ds, \quad f \in L^2(T_1),$$

is a compact operator. We define the integral kernel of  $A$  as  $\text{KER}(A)(s, t) = k(s, t)$ .

The covariance operator  $R_{XX}$  (7) is a compact self-adjoint nonnegative operator, since  $\text{KER}(R_{XX})(s, t) = \text{KER}(R_{XX})(t, s)$ .

A bounded linear operator is a *Hilbert-Schmidt operator* if there exists an orthonormal basis  $\{e_i\}$  of  $H_1$  such that  $\sum_i \|Ae_i\|^2 < \infty$ . Properties of Hilbert-Schmidt operators are as follows:

If  $M_1$  and  $M_2$  are measurable subsets of  $\mathbb{R}^p$  and  $\mathbb{R}^q$  respectively, the operator  $A \in B(L_2(M_1), L_2(M_2))$  is a Hilbert-Schmidt operator if and only if there exists a  $k \in L_2(M_1 \times M_2)$  such that

$$(A1) \quad (Af)(s) = \int_{M_1} k(s, t) f(t) dt, \quad \text{for } f \in L_2(M_1),$$

i.e.,  $A$  is an integral operator. Every Hilbert-Schmidt operator is compact. An operator  $A \in B(H_1, H_2)$  is a Hilbert-Schmidt operator if and only if the adjoint operator  $A^*$  of  $A$  is a

Hilbert-Schmidt operator.

Let  $\{e_i : i \geq 1\}$  and  $\{e'_i : i \geq 1\}$  be orthonormal bases for Hilbert spaces  $H_1$  and  $H_2$ , respectively, and let  $(a_{ij})_{i,j \geq 1}$  be an (infinite) matrix with  $a_{ij} \in \mathbb{R}$ . We define an operator  $A_{HS} : H_1 \rightarrow H_2$  by

$$A_{HS}h = \sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} a_{ij} \langle e_j, h \rangle \right) e_i,$$

for  $h \in \{u \in H_1 : \lim_{i \rightarrow \infty} \sum_{j=1}^{\infty} a_{ij} \langle e_j, u \rangle \text{ exists, for } j \geq 1, \text{ and } \sum_{i=1}^{\infty} \left| \sum_{j=1}^{\infty} a_{ij} \langle e_j, u \rangle \right|^2 < \infty\}$ .

This is a Hilbert-Schmidt operator, if  $\sum_{i,j=1}^{\infty} |a_{ij}|^2 = C^2 < \infty$ .

### Acknowledgments

We wish to thank a referee for constructive comments and helpful suggestions.

### References

- Ash, R. B. and Gardner, M. F. (1975), *Topics in Stochastic Processes*, New York: Academic Press.
- Brillinger, D. R. (1975), *Time Series. Data Analysis and Theory*, Holt, Rinehart and Winston, New York.
- Conway, J. B. (1985), *A Course in Functional Analysis*, New York: Springer-Verlag.
- Dauxois, J. and Nkiet, G.M. (1997) Canonical analysis of two Euclidean subspaces and its applications. *Linear Algebra and its Appl.* 264, 355-388.
- Dunford, N. and Schwartz, J. T. (1988). *Linear Operators*, Parts I and II. New York: Wiley.
- Hannan, E.J. (1961). The general theory of canonical correlation and its relation to functional analysis. *J. Australian Math. Soc.* 2, 229-242.
- He, G., Müller, H-G, and Wang, J-L (2000), Extending correlation and regression from multivariate to functional data. *Asymptotics in Statistics and Probability*, ed. M. Puri, pp. 197-210.
- He, G., Müller, H-G, and Wang, J-L (2002), Functional Linear Modelling and Canonical Regression Analysis. Manuscript.

- Hotelling, H. (1936). Relations between two sets of variates. *Biometrika* **28**, 321-377.
- Leurgans, S.E., Moyeed, R.A. and Silverman, B.W. (1993), Canonical correlation analysis when the data are curves, *J. R. Statistics, Soc.* **B 55**, 725-740.
- Olshen, R.A., Biden, E.N., Wyatt, M.P. and Sutherland, D.H. (1989), Gait analysis and the bootstrap, *Ann. Stat.*, 1989, **17**, 1419-1440.
- Ramsay, J.O. and Silverman, B.W. (1997), *Functional Data Analysis*. New York: Springer.
- Service, S.K., Rice, J.A. and Chavez, F.P. (1998). Relationship between physical and biological variables during the upwelling period in Monterey Bay, CA. *Deep-Sea Research II*, 1669-1685.
- Rice, J.A. and Silverman, B.W. (1991), Estimating the mean and covariance structure nonparametrically when the data are curves, *J. Royal Statist. Soc.* **B 53**, 233-243.