Available online at www.sciencedirect.com

ScienceDirect

Journal of Multivariate Analysis 99 (2008) 1083–1104

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Some properties of canonical correlations and variates in infinite dimensions

J. Cupidon[a], R. Eubank[b], D. Gilliam[a,*,1], F. Ruymgaart[a,2]

[a]*Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409, USA*

[b]*Mathematics and Statistics, Arizona State University, Tempe, AZ, 85278-1804, USA*

Received 14 July 2007

Available online 10 August 2007

Abstract

In this paper the notion of functional canonical correlation as a maximum of correlations of linear functionals is explored. It is shown that the population functional canonical correlation is in general well defined, but that it is a supremum rather than a maximum, so that a pair of canonical variates may not exist in the spaces considered. Also the relation with the maximum eigenvalue of an associated pair of operators and the corresponding eigenvectors is not in general valid. When the inverses of the operators involved are regularized, however, all of the above properties are restored. Relations between the actual population quantities and their regularized versions are also established. The sample functional canonical correlations can be regularized in a similar way, and consistency is shown at a fixed level of the regularization parameter.

AMS 2000 subject classification: primary 62H20; 62M99; secondary 32L15

Keywords: Functional population canonical correlation; Regularization of operators

1. Introduction

By restriction to a suitable subspace it may always be assumed without loss of generality that a covariance matrix or operator is injective. In the Euclidean case injectivity is equivalent to the requirement that the range of the covariance matrix is the entire space, and under this condition

* Corresponding author.

E-mail address: gilliam@math.ttu.edu (D. Gilliam).

[1] Research supported in part by AFOSR Grant #-FA9550-07-1-0214.

[2] Research supported by NSF Grant DMS-0605167.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2007.07.007

both the population canonical correlation and its sample analogue (provided only that the sample size is at least as big as the dimension) are well defined.

In Hilbert spaces—tacitly assumed to be of infinite dimension throughout the paper—the situation is different. The definition of canonical correlation can still be given as in the Euclidean case provided that "maximum'' is replaced with "supremum''. It will be shown (Section 3) that in general the value of this supremum is not attained for a pair of maximizers in the subspaces under consideration; i.e., a pair of canonical variates does not in general exist. This is a significant difference with the Euclidean case where the maximum is always assumed and yields a corresponding pair of canonical variables. Also the interpretation of the canonical correlation as the maximum eigenvalue of an associated operator breaks down.

The mathematical reason for this is that the covariance operator is compact and therefore its inverse unbounded and only densely defined. Because the range of the operator is still dense there does not exist a nonzero vector perpendicular to the range, and the definition of canonical correlation as a supremum of correlations still makes sense. Other properties, however, are lost in the infinite dimensional case.

Much can be salvaged, however, if regularized versions of inverses of covariance operators (or their square roots) are employed. It has been observed in the literature [8] that regularization is a necessity when the sample canonical correlation is to be considered. Indeed, a sample covariance operator always has a range of finite dimension equal to the sample size at most and hence cannot even be injective. Without some modifications or regularization the sample canonical correlation cannot be properly defined. However, it will be argued below that regularization is also expedient when dealing with the population; this regularization yields what will briefly be referred to as a regularized canonical correlation and regularized canonical variables.

In fact it will be shown in Section 2 that

(1) the regularized canonical correlations converge to the actual population canonical correlation, as the regularization parameter tends to zero;
(2) in case the actual population canonical correlation happens to be a global maximum, in some strong sense, attained at a unique pair of canonical variates, then the regularized canonical variates converge to this pair when the regularization parameter tends to zero;
(3) for a given value of the regularization parameter there exists a pair of associated positive compact Hermitian operators with the same largest eigenvalue, such that the regularized canonical correlation equals this eigenvalue;
(4) regularized canonical variates are given by a corresponding pair of eigenvectors.

Section 3 is devoted to correlating two parts of standard Brownian motion on disjoint time intervals, for several dependence structures. On the one hand this class of examples is of generic nature, and on the other hand canonical correlations and variates can be actually computed numerically because the basic ingredients are explicitly known. In passing we need to derive the Karhunen–Loève expansion for an arbitrary compact subinterval of the positive half-line, which does not follow trivially from the standard expansion on the unit interval if the left-hand endpoint is strictly positive. The class contains a number of illuminating examples that corroborate or complement the general theory. An example of the latter category is the counterexample, already mentioned above, where the canonical correlation is not assumed for a pair of canonical variates. Also two rather arbitrary dependent, dynamical processes are considered in Section 4.

In Section 5 we will briefly consider the sample canonical correlation and variates. As has already been observed, regularization is now necessary and the sample analogues share the four properties mentioned above for the population. It is natural to estimate a regularized population canonical correlation or variate by its sample analogue with the same regularization parameter. Given a value of the regularization parameter we conjecture that asymptotic normality, and hence consistency of the sample canonical correlation and variates can be obtained exploiting properties 3 and 4 above in conjunction with perturbation theory for functions of compact operators. For the Euclidean case see Ruymgaart and Yang [12]. This is beyond the scope of the present paper, but a simple direct proof of the consistency of the regularized sample canonical correlation is included.

The use of regularization for canonical correlation analysis dates back to Vinod [14]. In a functional data analysis context regularization techniques have been employed, for instance, by Silverman [13], Bosq [2] and Mas [10], and Leurgans et al. [8] for functional principal component analysis, functional times series analysis, and functional canonical correlation analysis respectively. An interesting practical example of two dependent random processes is, for instance, the angles made by hip and knee during the gait cycle, studied by Leurgans et al. [8]. While there are some questions concerning the necessity of regularization in functional principal components, it has already been observed above that Leurgans et al. [8] state that regularization is "absolutely essential'' for functional empirical canonical correlation, and that in its absence "every possible function can arise as a canonical variate with perfect canonical correlation''.

The standard method for adaptive selection of regularization parameters with functional data is the leave-one-out cross-validation method pioneered by Rice and Silverman [11] for functional principal components. A parallel of that approach has been used by Leurgans et al. [8] for functional canonical correlation. We will not explore that issue in this paper.

He et al. [6] addressed the problem of canonical correlation analysis for processes admitting a Karhunen–Loève type representation. In order to avoid difficulties with inversion of certain covariance operators, they restrict attention to subspaces where these inverses exist. Our use of regularization allows for the treatment of processes that cannot be handled by these authors. A possible relation with alternating projections and the alternating conditional expectation algorithm might be explored [1].

It should be noted that throughout this introduction we have for convenience talked somewhat loosely about the canonical correlation, where more precisely we should have discussed the squared principal canonical correlation. In the sections that follow we will adhere to this more precise formulation and use the abbreviation (R) SPCC for the (regularized) squared principal canonical correlation.

2. Definitions and basic properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathbb{H}$ an infinite dimensional separable Hilbert space over the real numbers with inner product $\langle \cdot, \cdot \rangle$, norm $\| \cdot \|$, and σ-field $\mathcal{B}_{\mathbb{H}}$ of Borel sets, and $X : \Omega \to \mathbb{H}$ a random element in $\mathbb{H}$, i.e., an $(\mathcal{F}, \mathcal{B}_{\mathbb{H}})$-measurable mapping. Assuming that $\mathbb{E}\|X\|^2 < \infty$ the mean $\mu = \mathbb{E}X$ is uniquely determined by the relations

$$\mathbb{E}\langle f, X \rangle = \langle f, \mu \rangle \quad \forall f \in \mathbb{H}, \tag{2.1}$$

and its covariance operator S by

$$\mathbb{E}\langle f, X - \mu \rangle \langle X - \mu, g \rangle = \langle f, Sg \rangle \quad \forall f, g \in \mathbb{H}; \tag{2.2}$$

see Laha and Rohatgi [7].

Such a covariance operator is nonnegative Hermitian and has a finite trace $\mathbb{E}\|X\|^2$, so that it is also compact. If 0 is an eigenvalue we can restrict S to the orthocomplement of its null space. We will therefore assume without loss of generality that S is injective and hence strictly positive.

Suppose that $\mathbb{H}_1$ and $\mathbb{H}_2$ are two subspaces of $\mathbb{H}$ such that

$$\mathbb{H} = \mathbb{H}_1 \oplus \mathbb{H}_2, \quad \mathbb{H}_1 \perp \mathbb{H}_2. \tag{2.3}$$

Denote the orthogonal projection onto $\mathbb{H}_j$ by P_j, let $X_j = P_j X$, and S_{ij} the restriction of S to $\mathbb{H}_j$ and $\mathbb{H}_i$, i.e.,

$$S_{ij} = P_i S P_j, \quad i, j = 1, 2. \tag{2.4}$$

Because the P_i are bounded and S is compact, each operator S_{ij} is still compact Debnath and Mikusiński [4]. The S_{jj} are in addition strictly positive Hermitian and have spectral representations

$$S_{jj} = \sum_{k=1}^{\infty} \lambda_{jk} \varphi_{jk} \otimes \varphi_{jk}, \quad j = 1, 2, \tag{2.5}$$

where the eigenvalues, repeated according to their finite multiplicities, satisfy $\lambda_{jk} > 0$ and $\lambda_{j1} \geqslant \lambda_{j2} \geqslant \cdots \downarrow 0$, and where the φ_{jk} form an orthonormal basis of corresponding eigenvectors for the space $\mathbb{H}_j$. Let us also note that

$$S_{12}^* = (P_1 S P_2)^* = P_2 S P_1 = S_{21}. \tag{2.6}$$

For brevity let us write

$$\mathbb{H}_1^0 = \mathbb{H}_1 \backslash \{0\}, \quad \mathbb{H}_2^0 = \mathbb{H}_2 \backslash \{0\}. \tag{2.7}$$

Definition 2.1. The squared principal canonical correlation (SPCC) is defined as

$$\rho^2 = \sup_{\substack{f \in \mathbb{H}_1^0 \\ g \in \mathbb{H}_2^0}} \rho_{f,g}^2 = \sup_{\substack{f \in \mathbb{H}_1^0 \\ g \in \mathbb{H}_2^0}} \frac{\{\mathbb{E}\langle f, X_1 - \mu_1 \rangle \langle X_2 - \mu_2, g \rangle\}^2}{\mathbb{E}\langle X_1 - \mu_1, f \rangle^2 \mathbb{E}\langle X_2 - \mu_2, g \rangle^2}. \tag{2.8}$$

If this supremum is assumed for certain $\overset{\vee}{f} \in \mathbb{H}_1^0$, $\overset{\vee}{g} \in \mathbb{H}_2^0$, the random variables $\langle X_1, \overset{\vee}{f} \rangle$ and $\langle X_2, \overset{\vee}{g} \rangle$ are called a pair of corresponding canonical variates.

Theorem 2.1. *Assuming that S is injective, the SPCC is well defined and we have*

$$0 \leqslant \rho^2 \leqslant 1, \tag{2.9}$$

$$\rho_{f,g}^2 = \frac{\langle f, S_{12} g \rangle^2}{\langle f, S_{11} f \rangle \langle g, S_{22} g \rangle}, \quad f \in \mathbb{H}_1^0, \quad g \in \mathbb{H}_2^0. \tag{2.10}$$

Proof. Because S_{11} and S_{22} are strictly positive it follows that $\langle f, S_{11} f \rangle > 0$ for all $f \in \mathbb{H}_1^0$ and $\langle g, S_{22} g \rangle > 0$ for all $g \in \mathbb{H}_2^0$. Furthermore, we have by (2.2) and (2.4)

$$\begin{aligned} &\mathbb{E}\langle f, X_1 - \mu_1 \rangle \langle X_2 - \mu_2, g \rangle \\ &\quad = \mathbb{E}\langle f, P_1(X - \mu) \rangle \langle P_2(X - \mu), g \rangle \\ &\quad = \mathbb{E}\langle P_1 f, (X - \mu) \rangle \langle (X - \mu), P_2 g \rangle \\ &\quad = \langle P_1 f, S P_2 g \rangle = \langle f, S_{12} g \rangle. \end{aligned} \tag{2.11}$$

It can be similarly shown that $\mathbb{E}\langle X_1 - \mu_1, f\rangle^2 = \langle f, S_{11} f\rangle$ and $\mathbb{E}\langle X_2 - \mu_2, g\rangle^2 = \langle g, S_{22} g\rangle$. This proves the equality in (2.10) and that $\rho^2_{f,g}$ and hence ρ^2 are well defined. By Schwarz's inequality we see that $0 \leqslant \rho^2_{f,g} \leqslant 1$ and this entails (2.9). □

Remark 2.1. The expression on the right in (2.10) is symmetric in f and g since by (2.6) we have $\langle f, S_{12} g\rangle = \langle S_{21} f, g\rangle$. Because $f \in \mathbb{H}_1^0$ and $g \in \mathbb{H}_2^0$, we could actually simply write $\langle f, Sg\rangle$ rather than $\langle f, S_{12} g\rangle$. We prefer the latter, however, for clarity; a similar remark holds true for the denominator in (2.10).

Remark 2.2. In the Euclidean case ρ^2 can be interpreted as the largest eigenvalue of

$$S_{11}^{-1/2} S_{12} S_{22}^{-1} S_{21} S_{11}^{-1/2}. \tag{2.12}$$

Such a representation turns out to be helpful to derive asymptotic normality of the sample SPCC [12]. In the present infinite dimensional situation the inverse S^{-1} of S is an unbounded operator, and so are S_{11}^{-1} and S_{22}^{-1}. Hence the above mentioned product of operators will be unbounded, so that this interpretation can no longer be valid.

Remark 2.3. It will be shown by means of an example in Section 3 that canonical variates as defined in Definition 2.1 will not in general exist.

Because of the deficiencies signaled in Remarks 2.2 and 2.3, regularization will be applied to better deal with the population SPCC, just as regularization is needed when dealing with the sample SPCC [8]. Let us replace S with $\alpha I + S$ for $\alpha > 0$, where I is the identity operator on $\mathbb{H}$. Although the latter operator is still strictly positive Hermitian with pure point spectrum, it is no longer compact and consequently cannot be a covariance operator. This is, however, of no real importance in the sequel as we will take the quotient on the right in (2.10) rather than the expression on the right in (2.8) as the starting point for the definition of regularized SPCC. Let us subsequently replace S_{ij} with $P_i(\alpha I + S)P_j = \alpha P_i P_j + S_{ij}$. Because $P_i P_j = 0$ for $i \neq j$, and $P_j P_j = I_j$ (essentially the identity operator on $\mathbb{H}_j$) we have

$$P_i(\alpha I + S)P_j = S_{ij} \quad \text{for } i \neq j, \tag{2.13}$$

$$P_j(\alpha I + S)P_j = \alpha I_j + S_{jj}, \tag{2.14}$$

for all $\alpha > 0$.

Definition 2.2. The regularized squared principal canonical correlation (RSPCC) is defined as

$$\rho^2(\alpha) = \sup_{\substack{f \in \mathbb{H}_1^0 \\ g \in \mathbb{H}_2^0}} \rho^2_{f,g}(\alpha) = \sup_{\substack{f \in \mathbb{H}_1^0 \\ g \in \mathbb{H}_2^0}} \frac{\langle f, S_{12} g\rangle^2}{\langle f, (\alpha I_1 + S_{11}) f\rangle \langle g, (\alpha I_2 + S_{22}) g\rangle}. \tag{2.15}$$

If this supremum is assumed for certain $\check{f}_\alpha \in \mathbb{H}_1^0$, $\check{g}_\alpha \in \mathbb{H}_2^0$, the random variables $\langle X, \check{f}_\alpha\rangle$ and $\langle X, \check{g}_\alpha\rangle$ are called a pair of corresponding regularized canonical variates.

Theorem 2.2. *For injective S we have*

$$0 \leqslant \rho^2_{f,g}(\alpha) \uparrow \rho^2_{f,g} \leqslant 1 \quad \text{as } \alpha \downarrow 0, \tag{2.16}$$

$$0 \leqslant \rho^2(\alpha) \uparrow \rho^2 \leqslant 1 \quad \text{as } \alpha \downarrow 0. \tag{2.17}$$

Proof. The result in (2.16) is immediate from the definition of $\rho^2_{f,g}(\alpha)$ in (2.15) and from the fact that $\langle f, (\alpha I_1 + S_{11}) f\rangle = \langle f, S_{11} f\rangle + \alpha\|f\|^2 \downarrow \langle f, S_{11} f\rangle$, as $\alpha \downarrow 0$, and from a similar fact for g. To verify (2.17) choose an arbitrary $\varepsilon > 0$. There exist $f_\varepsilon, g_\varepsilon$ such that $\rho^2_{f_\varepsilon,g_\varepsilon} > \rho^2 - 1/2\varepsilon$. Because of (2.16) there exists $\alpha_\varepsilon > 0$ such that $\rho^2_{f_\varepsilon,g_\varepsilon}(\alpha) > \rho^2_{f_\varepsilon,g_\varepsilon} - 1/2\varepsilon > \rho^2 - \varepsilon$ for all $0 < \alpha < \alpha_e$. $\square$

The function $(f, g) \mapsto \rho^2_{f,g}$ is easily seen to be continuous. We will say that this function has a strong global maximum at $\overset{\vee}{f} \in \mathbb{H}^0_1, \overset{\vee}{g} \in \mathbb{H}^0_2$ if for each $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that

$$\begin{cases} 0 \leqslant \rho^2_{\overset{\vee}{f},\overset{\vee}{g}} - \rho^2_{f,g} < \varepsilon \quad \text{implies} \\ \|\overset{\vee}{f} - f\| + \|\overset{\vee}{g} - g\| < \delta(\varepsilon) \quad \text{with } \delta(\varepsilon) \downarrow 0, \text{ as } \varepsilon \downarrow 0. \end{cases} \tag{2.18}$$

It should be noted, however, that according to Theorem 3.1 a maximum does not in general exist.

Theorem 2.3. *If $(f, g) \mapsto \rho^2_{f,g}$ has a strong global maximum at $\overset{\vee}{f} \in \mathbb{H}^0_1, \overset{\vee}{g} \in \mathbb{H}^0_2$, we have*

$$\|\overset{\vee}{f}_\alpha - \overset{\vee}{f}\| + \|\overset{\vee}{g}_\alpha - \overset{\vee}{g}\| \to 0 \quad \textit{as } \alpha \downarrow 0. \tag{2.19}$$

Proof. Choose an arbitrary $\eta > 0$, and $\varepsilon = \varepsilon(\eta)$ in (2.18) sufficiently small to ensure that $\delta(\varepsilon) < \eta$. We have $0 \leqslant \rho^2_{\overset{\vee}{f},\overset{\vee}{g}} - \rho^2_{\overset{\vee}{f}_\alpha,\overset{\vee}{g}_\alpha} = \rho^2_{\overset{\vee}{f},\overset{\vee}{g}} - \rho^2_{\overset{\vee}{f}_\alpha,\overset{\vee}{g}_\alpha}(\alpha) + \rho^2_{\overset{\vee}{f}_\alpha,\overset{\vee}{g}_\alpha}(\alpha) - \rho^2_{\overset{\vee}{f}_\alpha,\overset{\vee}{g}_\alpha}$. It is immediate from (2.16) that there exists $\alpha(\varepsilon) > 0$ such that $0 \leqslant \rho^2_{\overset{\vee}{f},\overset{\vee}{g}} - \rho^2_{\overset{\vee}{f}_\alpha,\overset{\vee}{g}_\alpha} = \rho^2 - \rho^2(\alpha) < \varepsilon$ for all $0 < \alpha < \alpha(\varepsilon)$. Apparently the condition in (2.18) is satisfied and therefore we may conclude that $\|\overset{\vee}{f} - \overset{\vee}{f}_\alpha\| + \|\overset{\vee}{g} - \overset{\vee}{g}_\alpha\| < \delta(\varepsilon) < \eta$ for $0 < \alpha < \alpha(\varepsilon)$. Since η was arbitrary (2.19) follows. $\square$

In order to prepare for the next result, let us recall the following well-known fact (see for instance, [5]). Let $T : \mathbb{H} \to \mathbb{H}$ be a Hermitian operator with pure point spectrum. Then we have

$$\max_{f \neq 0} \frac{\langle f, Tf\rangle}{\langle f, f\rangle} = \max_{\|f\|=1} \langle f, Tf\rangle = \text{largest eigenvalue of } T = \langle \overset{\vee}{f}, T\overset{\vee}{f}\rangle, \tag{2.20}$$

where $\overset{\vee}{f}$ is any eigenvector of unit length corresponding to this largest eigenvalue. The maximizer is unique (apart from the sign) if the eigenspace is one-dimensional.

Let us now introduce the operators

$$R_\alpha = (\alpha I_1 + S_{11})^{-1/2} S_{12} (\alpha I_2 + S_{22})^{-1} S_{21} (\alpha I_1 + S_{11})^{-1/2}, \tag{2.21}$$

$$Q_\alpha = (\alpha I_2 + S_{22})^{-1/2} S_{21} (\alpha I_1 + S_{11})^{-1} S_{12} (\alpha I_2 + S_{22})^{-1/2}. \tag{2.22}$$

For each $\alpha > 0$ these operators are compact because S_{12} and S_{21} are compact and the other factors are bounded (see the remark below (2.4)); they are also strictly positive Hermitian.

According to (2.20) there exist maximizers $\overset{\vee}{f}_\alpha \in \mathbb{H}_1$, $\overset{\vee}{g}_\alpha \in \mathbb{H}_2$, of unit length, such that

$$\max_{\substack{f\in\mathbb{H}_1\\ \|f\|=1}} \langle f, R_\alpha f\rangle = \langle \overset{\vee}{f}_\alpha, R_\alpha \overset{\vee}{f}_\alpha\rangle, \tag{2.23}$$

$$\max_{\substack{g\in\mathbb{H}_2\\ \|g\|=1}} \langle g, Q_\alpha g\rangle = \langle \overset{\vee}{g}_\alpha, Q_\alpha \overset{\vee}{g}_\alpha\rangle. \tag{2.24}$$

It should also be noted that for each $\alpha > 0$, $p > 0$

$$\text{range of } (\alpha I_j + S_{jj})^p = \mathbb{H}_j, \quad j = 1, 2, \tag{2.25}$$

and that its inverse is bounded and has spectral representation (cf. (2.5))

$$(\alpha I_j + S_{jj})^{-p} = \sum_{k=1}^{\infty} \frac{1}{(\alpha + \lambda_{jk})^p} \varphi_{jk} \otimes \varphi_{jk}. \tag{2.26}$$

Theorem 2.4. *For each $\alpha > 0$ we have*

$$\begin{aligned}\rho^2(\alpha) &= \text{largest eigenvalue of } R_\alpha = \langle \overset{\vee}{f}_\alpha, R_\alpha \overset{\vee}{f}_\alpha\rangle\\ &= \text{largest eigenvalue of } Q_\alpha = \langle \overset{\vee}{g}_\alpha, Q_\alpha \overset{\vee}{g}_\alpha\rangle.\end{aligned} \tag{2.27}$$

Hence the supremum defining $\rho^2(\alpha)$ is in fact a maximum, and the corresponding pair of regularized canonical variates is

$$\langle X_1, \overset{\vee}{f}_\alpha\rangle, \ \langle X_2, \overset{\vee}{g}_\alpha\rangle. \tag{2.28}$$

These canonical variates are essentially unique if the eigenspaces corresponding to the maximal eigenvalue are one-dimensional.

Proof. To find $\rho^2(\alpha)$ in (2.15) let us first fix $f \in \mathbb{H}_1^0$ and set $v = (\alpha I_2 + S_{22})^{1/2} g$. According to (2.25) v runs through all of $\mathbb{H}_2^0$ when g does. Therefore we may write

$$\begin{aligned}\sup_{g\in\mathbb{H}_2^0} \frac{\langle f, S_{12} g\rangle^2}{\langle g, (\alpha I_2 + S_{22}) g\rangle} &= \sup_{\|v\|=1} \frac{\langle f, S_{12}(\alpha I_2 + S_{22})^{-1/2} v\rangle^2}{\langle v, v\rangle}\\ &= \sup_{\|v\|=1} \langle (\alpha I_2 + S_{22})^{-1/2} S_{21} f, v\rangle.\end{aligned} \tag{2.29}$$

According to the Schwarz inequality this supremum is a maximum attained for

$$\overset{\vee}{v} = \frac{1}{\|(\alpha I_2 + S_{22})^{-1/2} S_{21} f\|} (\alpha I_2 + S_{22})^{-1/2} S_{21} f. \tag{2.30}$$

Substituting this $\overset{\vee}{v}$ back into (2.29) it follows that the supremum in that expression is a maximum equal to $\|(\alpha I_2 + S_{22})^{-1/2} S_{21} f\|^2$.

Let us next set $u = (\alpha I_1 + S_{11})^{1/2} f$ where, again by (2.25), u runs through all of $\mathbb{H}_1^0$ when f does. The supremum in (2.15) of actual interest now equals

$$\begin{aligned}\rho^2(\alpha) &= \sup_{f \in \mathbb{H}_1^0} \frac{\|(\alpha I_2 + S_{22})^{-1/2} S_{21} f\|^2}{\langle f, (\alpha I_1 + S_{11}) f\rangle} \\ &= \sup_{f \in \mathbb{H}_1^0} \frac{\langle f, S_{12}(\alpha I_2 + S_{22})^{-1} S_{21} f\rangle}{\langle f, (\alpha I_1 + S_{11}) f\rangle} \\ &= \sup_{\|u\|=1} \langle u, R_\alpha u\rangle. \end{aligned} \tag{2.31}$$

The supremum is in fact a maximum and assumed at any $\overset{\vee}{f}_\alpha$ as defined in (2.23). A similar reasoning can be pursued for Q_α. □

Remark 2.4. It has already been observed that for $\alpha = 0$ the value ρ^2 of the SPCC may not be assumed. We will now show that the method of proof of Theorem 2.4 breaks down in this case. Let us return, for instance, to the first substitution $v = (\alpha I_2 + S_{22})^{1/2} g$. This substitution worked conveniently because $(\alpha I_2 + S_{22})^{1/2}$ is bijective. This property no longer holds true if $\alpha = 0$. The operator $S_{22}^{1/2}$ is still injective, but it is not surjective. The same problem arises with $S_{11}^{1/2}$. In fact, it is not hard to see that

$$\text{range of } S_{jj}^{1/2} = \left\{ u \in \mathbb{H}_j : \sum_{k=1}^{\infty} \frac{\langle u, \varphi_{jk}\rangle^2}{\lambda_{jk}} < \infty \right\}, \tag{2.32}$$

using the notation in (2.5), and we no longer have an unrestricted supremum. When regularization is introduced the pattern of the proof becomes similar to that in the Euclidean case; see, for instance, Mardia et al. [9].

3. Correlating Brownian motions

3.1. Introduction

The above general theory for SPCC's applies more explicitly to pairs of stochastic processes that admit Karhunen–Loève expansions. Since we intend to provide concrete examples with numerical calculations we will restrict ourselves to standard Brownian motion processes on two disjoint subintervals of [0, 1]. It turns out that the Karhunen–Loève expansion of standard Brownian motion on an interval $[a, b]$, with $0 \leqslant a < b < \infty$, does not follow trivially from the expansion on [0, 1] if $a > 0$. Therefore in passing we will obtain the expansion on an arbitrary interval $[a, b]$ by first solving the ensuing Sturm–Liouville problem on that interval (Section 3.2). Then, in Section 3.3, some Karhunen–Loève expansions on selected subintervals will be given and several dependence structures will be specified. Explicit calculations of RSPCC's and their corresponding pairs of regularized canonical variates will be done in Section 3.4, where also an example is given of a SPCC whose value is not assumed in the Hilbert spaces considered.

3.2. The Sturm–Liouville problem on an arbitrary compact subinterval of the positive half-line

Suppose that $0 \leqslant a < b < \infty$, and let $\{W(t), t \geqslant 0\}$ be standard Brownian motion , meaning that $W(0) = 0$, $W(t) \overset{\mathrm{d}}{=} N(0, t)$ for each $t > 0$, and $\mathbb{E}W(s)W(t) = s \wedge t$ for $s, t \geqslant 0$. We will restrict this Brownian motion to $[a, b]$. As usual $S(s, t) = s \wedge t$, $(s, t) \in [a, b] \times [a, b]$, is considered as the kernel of an integral operator, also denoted by S without confusion, mapping $L^2(a, b)$ into itself according to

$$(Sf)(s) = \int_a^b S(s, t) f(t)\, dt, \quad f \in L^2(a, b). \tag{3.1}$$

This operator is compact, Hermitian and strictly positive. Finding its eigenvalues and eigenfunctions leads to the usual Sturm–Liouville problem, but with different boundary conditions.

Assume that for some $\lambda = \mu^{-2} > 0$ we have

$$(S\varphi)(s) = \int_a^s t\varphi(t)\, dt + s \int_s^b \varphi(t)\, dt = \lambda\varphi(s), \quad a \leqslant s \leqslant b. \tag{3.2}$$

Differentiating twice yields

$$\varphi''(s) + \mu^2 \varphi(s) = 0, \quad a \leqslant s \leqslant b. \tag{3.3}$$

We have the boundary conditions

$$\varphi(a) - a\varphi'(a) = 0, \quad \varphi'(b) = 0. \tag{3.4}$$

The general solution of the system (3.3) and (3.4) is of the form

$$\varphi(s) = \alpha \sin(\mu s) + \beta \cos(\mu s), \quad a \leqslant s \leqslant b. \tag{3.5}$$

Exploiting the boundary conditions we arrive at a system of equations after some trigonometric manipulations. For this system to have a nontrivial solution we must distinguish the cases $a = 0$ and $a > 0$.

If $a = 0$ the condition for a nontrivial solution is

$$\cos(\mu b) = 0. \tag{3.6}$$

Solutions of (3.6) are given by $\mu_m = (m - 1/2)\,\pi/b$ or by the eigenvalues

$$\lambda_m = \left\{ \frac{b}{\left(m - \frac{1}{2}\right)\pi} \right\}^2, \quad m \in \mathbb{N}. \tag{3.7}$$

The corresponding eigenfunctions are

$$\varphi_m(s) = c_m \cos(\mu_m(b - s)), \quad 0 \leqslant s \leqslant b, \ m \in \mathbb{N}, \tag{3.8}$$

where c_m is a normalization constant given by

$$c_m = \left[\int_0^b \cos^2(\mu_m(b - s))\, ds \right]^{-1/2}. \tag{3.9}$$

From this it follows easily that in the usual situation where $a = 0$ and $b = 1$ we obtain

$$\lambda_m = \left\{ \frac{1}{\left(m - \frac{1}{2}\right)\pi} \right\}^2, \quad \varphi_m(s) = \sqrt{2} \sin\left(\left(m - \frac{1}{2}\right)\pi s\right), \quad 0 \leqslant s \leqslant 1, \ m \in \mathbb{N}. \tag{3.10}$$

For $a > 0$, however, the eigenvalues can only be implicitly determined. In this case the condition for a nontrivial solution is

$$\tan(\mu(b-a)) = \frac{1}{a\mu}. \tag{3.11}$$

Setting $\xi = \mu(b-a)$, we need to solve

$$\tan(\xi) = \left(\frac{b-a}{a}\right)\frac{1}{\xi}, \quad \xi > 0. \tag{3.12}$$

It is not hard to see that there are infinitely many solutions

$$\xi_1 < \xi_2 < \cdots \uparrow \infty \quad \text{such that } \frac{\xi_m}{(m-1)\pi} \to 1 \text{ as } m \to \infty. \tag{3.13}$$

In terms of the eigenvalues $\lambda_1 > \lambda_2 > \cdots$ this means that

$$\lambda_m \sim \left\{\frac{b-a}{(m-1)\pi}\right\}^2 \quad \text{as } m \to \infty, \tag{3.14}$$

and for the corresponding eigenfunctions we have

$$\varphi_m(s) = c_m \cos\left(\xi_m \frac{b-s}{b-a}\right), \quad 0 < a \leqslant s \leqslant b, \tag{3.15}$$

where the c_m are normalization constants given by

$$c_m = \left[\int_a^b \cos^2\left(\xi_m \frac{b-s}{b-a}\right) ds\right]^{-1/2}. \tag{3.16}$$

It can be shown by direct calculation that

$$c_m \sim \sqrt{\frac{2}{b-a}} \quad \text{as } m \to \infty, \tag{3.17}$$

in either case, i.e., for each $0 \leqslant a < b < \infty$.

3.3. Standard Brownian motion on selected pairs of intervals

For the most part we will restrict to the specific pair of subintervals $[a, b] = [0, 1/2]$ and $[a, b] = [1/2, 1]$. The choice of $1/2$ as the right-hand end point of the first interval turns out to be convenient in certain calculations. In order to remain in keeping with the set-up in Section 2 we will choose $\mathbb{H} = L^2(0, 1)$ and consider $L^2(0, 1/2)$ and $L^2(1/2, 1)$ as subspaces $\mathbb{H}_1$ and $\mathbb{H}_2$ of $L^2(0, 1)$.

Choosing $a = 0$ and $b = 1/2$, the λ_m in (3.7) and the φ_m in (3.8) reduce to

$$\lambda_m = \lambda_{1m} = \left\{\frac{1}{(2m-1)\,\pi}\right\}^2, \quad m \in \mathbb{N}, \tag{3.18}$$

$$\varphi_m(s) = \varphi_{1m}(s) = 2\,\mathbf{1}_{[0,1/2]}(s)\sin((2m-1)\pi s), \quad 0 \leqslant s \leqslant 1, \quad m \in \mathbb{N}, \tag{3.19}$$

since the c_m all reduce to 2 in this case.

Let $U_{11}, U_{12}, \ldots$ be independent random variables on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$U_{1m} \overset{\mathrm{d}}{=} N(0, \lambda_{1m}), \quad m \in \mathbb{N}. \tag{3.20}$$

The random element $X_1 : \Omega \to \mathbb{H}_1$ is now defined to be given by

$$X_1(s) = \sum_{m=1}^{\infty} U_{1m}\varphi_{1m}(s), \quad 0 \leqslant s \leqslant 1. \tag{3.21}$$

It is clear that

$$X_1 \overset{\mathrm{d}}{=} \text{standard Brownian motion on } [0, 1/2]. \tag{3.22}$$

When we choose $a = 1/2$ and $b = 1$, the λ_m in (3.14) and the φ_m in (3.15) reduce to

$$\lambda_m = \lambda_{2m} \sim \left\{ \frac{1}{2\,(m-1)\,\pi} \right\}^2 \quad \text{as } m \to \infty, \tag{3.23}$$

$$\varphi_m(s) = \varphi_{2m}(s) = c_{2m}\mathbf{1}_{[1/2,1]}(s)\cos(2\xi_m(1-s)), \quad 0 \leqslant s \leqslant 1, \quad m \in \mathbb{N}, \tag{3.24}$$

where the ξ_m in (3.13) and the $c_m = c_{2m}$ in (3.16) satisfy

$$\xi_m \sim (m-1)\pi, \quad c_m \sim 2 \ \text{ as } m \to \infty. \tag{3.25}$$

Consider a sequence $U_{21}, U_{22}, \ldots$ of independent random variables on $L^2(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$U_{2m} \overset{\mathrm{d}}{=} N(0, \lambda_{2m}), \quad m \in \mathbb{N}. \tag{3.26}$$

and let us define the random element $X_2 : \Omega \to \mathbb{H}_2$ by the expansion

$$X_2(s) = \sum_{m=1}^{\infty} U_{2m}\varphi_{2m}(s), \quad 0 \leqslant s \leqslant 1. \tag{3.27}$$

For this process we have

$$X_2 \overset{\mathrm{d}}{=} \text{ standard Brownian motion on } [1/2, 1]. \tag{3.28}$$

At this point we will introduce one general and two special dependence structures between the random variables in (3.20) and (3.26) by specifying

$$\gamma_{km} = \mathbb{E}\, U_{1k}U_{2m}. \tag{3.29}$$

By the Schwarz inequality we see that these numbers must always satisfy

$$\left|\gamma_{km}\right| \leqslant \sqrt{\lambda_{1k}\lambda_{2m}}. \tag{3.30}$$

We will see now that they must in fact satisfy a much more stringent condition. The independence of the $U_{j1}, U_{j2}, \ldots$ entails that the standard random variables $\widetilde{U}_{jm} = U_{jm}/\sqrt{\lambda_{jm}}$ form an orthonormal system in the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Let us write inner product and norm in this space as $\langle \cdot, \cdot \rangle_{\mathbb{P}}$ and $\| \cdot \|_{\mathbb{P}}$. It follows from Bessel's inequality that $\lambda_{1k} = \|U_{1k}\|_{\mathbb{P}}^2 \geqslant \sum_{m=1}^{\infty} \langle U_{1k}, \widetilde{U}_{2m} \rangle_{\mathbb{P}}^2 = \sum_{m=1}^{\infty} \gamma_{km}^2/\lambda_{2m}$. By symmetry this yields the pair of inequalities

$$\sum_{m=1}^{\infty} \frac{\gamma_{km}^2}{\lambda_{1k}\lambda_{2m}} \leqslant 1 \ \ \forall k, \quad \sum_{k=1}^{\infty} \frac{\gamma_{km}^2}{\lambda_{1k}\lambda_{2m}} \leqslant 1 \ \ \forall m. \tag{3.31}$$

3.4. Some special cases and numerical approximations

Case 1: Let us first consider the general situation for arbitrary γ_{km}. The operators S_{ij} are given by (see (2.4) and (2.5))

$$S_{jj} = \sum_{m=1}^{\infty} \lambda_{jm} \varphi_{jm} \otimes \varphi_{jm}, \quad j = 1, 2, \tag{3.32}$$

$$S_{12} = \mathbb{E} X_1 \otimes X_2 = \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \gamma_{km} \varphi_{1k} \otimes \varphi_{2m}, \tag{3.33}$$

$$S_{21} = \mathbb{E} X_2 \otimes X_1 = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \gamma_{km} \varphi_{2m} \otimes \varphi_{1k}. \tag{3.34}$$

To compute $\rho^2(\alpha)$ we will exploit Theorem 2.4 and first determine the action of the operator R_α in (2.19) under the present circumstances. For this purpose choose an arbitrary $f = \sum_{k=1}^{\infty} x_k \varphi_{1k} \in \mathbb{H}_1$, where $x_k = \langle f, \varphi_{1k} \rangle$, and note that we have successively

$$(\alpha I_1 + S_{11})^{-1/2} f = \sum_{k=1}^{\infty} \frac{x_k}{\sqrt{\alpha + \lambda_{1k}}} \varphi_{1k} = f_{(1)}, \tag{3.35}$$

$$\begin{aligned} S_{21} f_{(1)} &= \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \gamma_{km} \langle f_{(1)}, \varphi_{1k} \rangle \varphi_{2m} \\ &= \sum_{m=1}^{\infty} \left\{ \sum_{k=1}^{\infty} \frac{\gamma_{km} x_k}{\sqrt{\alpha + \lambda_{1k}}} \right\} \varphi_{2m} = f_{(2)}, \end{aligned} \tag{3.36}$$

$$\begin{aligned} (\alpha I_2 + S_{22})^{-1} f_{(2)} &= \sum_{m=1}^{\infty} \frac{\langle f_{(2)}, \varphi_{2m} \rangle}{\alpha + \lambda_{2m}} \varphi_{2m} \\ &= \sum_{m=1}^{\infty} \frac{1}{\alpha + \lambda_{2m}} \left\{ \sum_{k=1}^{\infty} \frac{\gamma_{km} x_k}{\sqrt{\alpha + \lambda_{1k}}} \right\} \varphi_{2m} = f_{(3)}, \end{aligned} \tag{3.37}$$

$$\begin{aligned} S_{12} f_{(3)} &= \sum_{k=1}^{\infty} \left\{ \sum_{m=1}^{\infty} \gamma_{km} \langle f_{(3)}, \varphi_{2m} \rangle \right\} \varphi_{1k} \\ &= \sum_{k=1}^{\infty} \left[\sum_{m=1}^{\infty} \frac{\gamma_{km}}{\alpha + \lambda_{2m}} \left\{ \sum_{j=1}^{\infty} \frac{\gamma_{jm} x_j}{\sqrt{\alpha + \lambda_{1j}}} \right\} \right] \varphi_{1k} = f_{(4)}, \end{aligned} \tag{3.38}$$

and finally

$$\begin{aligned} (\alpha I_1 + S_{11})^{-1/2} f_{(4)} &= \sum_{k=1}^{\infty} \frac{\langle f_{(4)}, \varphi_{1k} \rangle}{\sqrt{\alpha + \lambda_{1k}}} \varphi_{1k} \\ &= \sum_{k=1}^{\infty} \frac{1}{\sqrt{\alpha + \lambda_{1k}}} \left[\sum_{m=1}^{\infty} \frac{\gamma_{km}}{\alpha + \lambda_{2m}} \left\{ \sum_{j=1}^{\infty} \frac{\gamma_{jm} x_j}{\sqrt{\alpha + \lambda_{1j}}} \right\} \right] \varphi_{1k} \\ &= R_\alpha f. \end{aligned} \tag{3.39}$$

Choosing $\varphi_{11}, \varphi_{12}, \ldots$ as an orthonormal basis of $L^2(0, 1/2)$, the operator $R_\alpha : L^2(0, 1/2) \to L^2(0, 1/2)$ corresponds to an operator $\mathcal{R}_\alpha : \ell^2 \to \ell^2$ with elements

$$\mathcal{R}_\alpha(k, j) = \frac{1}{\sqrt{\alpha + \lambda_{1k}}} \left\{ \sum_{m=1}^{\infty} \frac{\gamma_{km}\gamma_{jm}}{\alpha + \lambda_{2m}} \right\} \frac{1}{\sqrt{\alpha + \lambda_{1j}}} \ . \tag{3.40}$$

The operator $\mathcal{R}_\alpha$ has the same properties as R_α and is strictly positive compact Hermitian. According to Theorem 2.4 we have

$$\rho^2(\alpha) = \max_{\substack{x \in \ell^2 \\ \|x\|=1}} \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} x_k \mathcal{R}_\alpha(k, j) x_j. \tag{3.41}$$

A maximizer $\overset{\vee}{x}_\alpha$ yields the canonical variate

$$\langle X_1, \overset{\vee}{f}_\alpha \rangle \quad \text{where } \overset{\vee}{f}_\alpha = \sum_{k=1}^{\infty} \overset{\vee}{x}_{\alpha,k} \varphi_{1k}. \tag{3.42}$$

Similarly $Q_\alpha : L^2(1/2, 1) \to L^2(1/2, 1)$, see (2.20), corresponds to an operator $\mathcal{Q}_\alpha : \ell^2 \to \ell^2$ with elements

$$\mathcal{Q}_\alpha(k, j) = \frac{1}{\sqrt{\alpha + \lambda_{2k}}} \left\{ \sum_{m=1}^{\infty} \frac{\gamma_{mk}\gamma_{mj}}{\alpha + \lambda_{1m}} \right\} \frac{1}{\sqrt{\alpha + \lambda_{2j}}}. \tag{3.43}$$

If the corresponding quadratic form is maximized by a vector $\overset{\vee}{y}_\alpha \in \ell^2$ of unit length we obtain a second canonical variate

$$\langle X_2, \overset{\vee}{g}_\alpha \rangle \quad \text{where } \overset{\vee}{g}_\alpha = \sum_{k=1}^{\infty} \overset{\vee}{y}_{\alpha,k} \varphi_{2k}. \tag{3.44}$$

Of course application of Theorem 2.2 yields that the value ρ^2 of the SPCC also can be obtained as

$$\lim_{\alpha \downarrow 0} \rho^2(\alpha), \tag{3.45}$$

with $\rho^2(\alpha)$ as in (3.41).

For numerical calculations it remains to make an explicit choice for the γ_{km} in (3.29). We will not do that here, but rather in the more general situation of Section 4.

Case 2: Here we will consider the special instance of Case 1 where

$$U_{jk} = \langle W, \varphi_{jk} \rangle, \quad j = 1, 2, \ \ k \in \mathbb{N}, \tag{3.46}$$

and where W is standard Brownian motion on the positive half-line. In this situation we have

$$\begin{aligned} \gamma_{km} &= \mathbb{E}\langle W, \varphi_{1k} \rangle \langle W, \varphi_{2m} \rangle \\ &= \int_{s=0}^{1/2} \int_{t=1/2}^{1} \mathbb{E} W(s) \varphi_{1,k}(s) W(t) \varphi_{2m}(t)\, ds\, dt \\ &= \int_{s=0}^{1/2} \int_{t=1/2}^{1} \varphi_{1k}(s)(s \wedge t) \varphi_{2m}(t)\, ds\, dt \end{aligned}$$

$$
\begin{aligned}
&= \int_{s=0}^{1/2} s\varphi_{1k}(s)\,ds \cdot \int_{t=1/2}^{1} \varphi_{2m}(t)\,dt \\
&= \langle E, \varphi_{1k}\rangle\langle 1, \varphi_{2m}\rangle,
\end{aligned} \tag{3.47}
$$

where the functions E and 1 are given by

$$
E(s) = s \cdot \mathbf{1}_{[0,1/2]}(s), \quad 1(s) = \mathbf{1}_{[1/2,1]}(s), \quad 0 \leqslant s \leqslant 1. \tag{3.48}
$$

Apart from (3.22) and (3.28) we now have the even stronger property

$$
(X_1, X_2) \overset{\mathrm{d}}{=} \text{standard Brownian motion on } [0, 1], \tag{3.49}
$$

due to the fact that all the U_{jk} are constructed from one single standard Brownian motion. This property suggests that the SPCC might be 1. The variates $X_1(1/2)$ and $X_2(1/2)$ suggest themselves as potential candidates for a pair of canonical variates, but they cannot be expressed as inner products: in fact delta-functions would be needed for such representations.

Let us choose, however, the functions

$$
f_\varepsilon(s) = \frac{1}{\sqrt{\varepsilon}}\mathbf{1}_{[1/2-\varepsilon,1/2]}(s), \quad g_\varepsilon(s) = \frac{1}{\sqrt{\varepsilon}}\mathbf{1}_{[1/2,1/2+\varepsilon]}(s), \quad 0 \leqslant s \leqslant 1, \tag{3.50}
$$

with unit norm. We have

$$
\begin{aligned}
\rho^2_{f_\varepsilon, g_\varepsilon} &= \frac{\left[\int_{s=1/2-\varepsilon}^{1/2} s \left\{\int_{s=1/2}^{1/2+\varepsilon} dt\right\} ds\right]^2}{\left\{\int_{s=1/2-\varepsilon}^{1/2}\int_{u=1/2-\varepsilon}^{1/2} (s \wedge u)\,ds\,du\right\}\left\{\int_{t=1/2}^{1/2+\varepsilon}\int_{v=1/2}^{1/2+\varepsilon} (t \wedge v)\,dt\,dv\right\}} \\
&= \frac{(1-\varepsilon)^2}{(1-4\varepsilon/3)(1+2\varepsilon/3)} \to 1 \quad \text{as } \varepsilon \downarrow 0.
\end{aligned} \tag{3.51}
$$

This means that we must have

$$
\rho^2 = 1, \tag{3.52}
$$

indeed. The situation indicates that this value may not be assumed for a pair of functions in $L^2(0, 1/2)$ and $L^2(1/2, 1)$. We will now give a formal proof of this conjecture.

Theorem 3.1. *The value of the SPCC is not in general assumed for a pair of canonical variates. A case in point is the present situation.*

Proof. Due to the special choice of the γ_{km} in (3.47) it follows from (3.33) that

$$
\begin{aligned}
\langle f, S_{12} g\rangle &= \left\langle \sum_{\ell=1}^{\infty} \langle f, \varphi_{1\ell}\rangle\varphi_{1\ell}, \sum_{k=1}^{\infty}\left\{\sum_{m=1}^{\infty} \gamma_{km}\langle g, \varphi_{2m}\rangle\right\}\varphi_{1k}\right\rangle \\
&= \left(\sum_{k=1}^{\infty} \langle E, \varphi_{1k}\rangle\langle f, \varphi_{1k}\rangle\right)\left(\sum_{m=1}^{\infty} \langle 1, \varphi_{2m}\rangle\langle g, \varphi_{2m}\rangle\right) \\
&= \langle f, E\rangle\langle g, 1\rangle.
\end{aligned} \tag{3.53}
$$

(This can also be seen from (3.49) which implies that $\mathbb{E}\langle f, X_1\rangle\langle X_2, g\rangle = \mathbb{E}\langle f, W\rangle\langle W, g\rangle$.) Hence the supremum in Definition 2.1 factorizes in this case, and we have

$$
\rho^2 = \sup_{f \in \mathbb{H}_1^0} \frac{\langle f, E\rangle^2}{\langle f, S_{11} f\rangle} \cdot \sup_{g \in \mathbb{H}_2^0} \frac{\langle g, 1\rangle^2}{\langle g, S_{22} g\rangle}. \tag{3.54}
$$

It will now be shown that the first supremum on the right in (3.54) cannot be attained for an element in $\mathbb{H}_1^0$.

Let us write $f \in \mathbb{H}_1^0$ as $f = \sum_{k=1}^{\infty} \langle f, \varphi_{1k} \rangle \varphi_{1k} = \sum_{k=1}^{\infty} x_k \varphi_{1k}$, and note that (cf. (3.18), (3.19))

$$\begin{aligned} \langle E, \varphi_{1k} \rangle &= 2 \int_0^{1/2} s \sin((2k-1)\pi s)\, ds \\ &= (-1)^{k+1} 2 \{(2k-1)\pi\}^{-2} = (-1)^{k+1} 2\lambda_{1k}. \end{aligned} \tag{3.55}$$

It follows that

$$\sup_{f \in \mathbb{H}_1^0} \frac{\langle f, E \rangle^2}{\langle f, S_{11} f \rangle} = \sup_{f \in \mathbb{H}_1^0} \frac{\left(\sum_{k=1}^{\infty} x_k \langle E, \varphi_{1k} \rangle\right)^2}{\sum_{k=1}^{\infty} x_k^2 \lambda_{1k}} = 4 \sup_{\substack{f \in \mathbb{H}_1^0 \\ \|f\|=1}} \frac{\left(\sum_{k=1}^{\infty} x_k \lambda_{1k}\right)^2}{\sum_{k=1}^{\infty} x_k^2 \lambda_{1k}}, \tag{3.56}$$

where we may assume that

$$0 \leqslant x_k \leqslant 1 \quad \text{for all } k, \ \sum_{k=1}^{\infty} x_k^2 = 1. \tag{3.57}$$

Suppose that $\check{f} = \sum_{k=1}^{\infty} \check{x}_k \varphi_{1k}$ realizes this supremum (with the $\check{x}_k$ satisfying (3.57)). Let us set $0 < \Lambda_1 = \sum_{k=1}^{\infty} \lambda_{1k} < \infty$, so that the numbers

$$p_k = \frac{\lambda_{1k}}{\Lambda_1}, \quad k \in \mathbb{N}, \tag{3.58}$$

define the density of a discrete probability distribution on $\mathbb{N}$. The supremum may then be written as

$$4 \frac{\left(\sum_{k=1}^{\infty} \check{x}_k \lambda_{1k}\right)^2}{\sum_{k=1}^{\infty} \check{x}_k^2 \lambda_{1k}} = 4\Lambda_1 \frac{\left(\sum_{k=1}^{\infty} \check{x}_k p_k\right)^2}{\sum_{k=1}^{\infty} \check{x}_k^2 p_k}. \tag{3.59}$$

According to the Schwarz (or Jensen's) inequality for "integrals'' with respect to this discrete probability distribution we have

$$\left(\sum_{k=1}^{\infty} \check{x}_k p_k\right)^2 \leqslant \sum_{k=1}^{\infty} \check{x}_k^2 p_k, \tag{3.60}$$

with equality if and only if for some $\xi \in \mathbb{R}$

$$\check{x}_k = \xi \quad \text{for all } k. \tag{3.61}$$

An $\check{f}$ with such coordinates is either the zero function (if $\xi = 0$), and therefore not of norm 1, or not in $\mathbb{H}_1^0$ at all (if $\xi \neq 0$). Consequently we must have a strict inequality which means that we have

$$\sup_{f \in \mathbb{H}_1^0} \frac{\langle f, E \rangle^2}{\langle f, S_{11} f \rangle} < 4 \sum_{k=1}^{\infty} \lambda_{1k}, \tag{3.62}$$

if a maximizer would exist.

On the other hand we may argue that the left hand side of (3.62) equals

$$\lim_{\alpha\downarrow 0}\max_{f\in\mathbb{H}_1^0}\frac{\langle f, E\rangle^2}{\langle f, (\alpha I_1+S_{11})f\rangle}. \tag{3.63}$$

Put $(\alpha I_1+S_{11})^{1/2}f=u$. Then the maximum in (3.63) equals

$$\max_{\substack{u\in\mathbb{H}_1^0\\ \|u\|=1}}\langle u, (\alpha I_1+S_{11})^{-1/2}E\rangle^2. \tag{3.64}$$

According to the Schwarz inequality this maximum is assumed for $\overset{\vee}{u}_\alpha=(\alpha I_1+S_{11})^{-1/2}E/\|(\alpha I_1+S_{11})^{-1/2}E\|$, and the value of the maximum equals

$$\begin{aligned}\|(\alpha I_1+S_{11})^{-1/2}E\|^2&=\sum_{k=1}^{\infty}\frac{\langle E, \varphi_{1k}\rangle^2}{\alpha+\lambda_{1k}}\\&=4\sum_{k=1}^{\infty}\frac{\lambda_{1k}^2}{\alpha+\lambda_{1k}}\to 4\sum_{k=1}^{\infty}\lambda_{1k}\quad\text{as }\alpha\downarrow 0.\end{aligned} \tag{3.65}$$

Hence the value of the supremum on the left in (3.62) must be equal to $4\sum_{k=1}^{\infty}\lambda_{1k}$. This yields a contradiction and hence a maximizer $\overset{\vee}{f}$ cannot exist. □

Remark 3.1. Maintaining the model assumption (3.46) we may restrict standard Brownian motion to subintervals $[0, a]$ and $[b, 1]$ with arbitrary $0<a\leqslant b<1$. It seems a fair guess that in this more general case

$$\rho^2=\mathrm{Corr}(W(a), W(b))=\frac{a}{b}, \tag{3.66}$$

and that again this value of the SPCC is not assumed for a pair of canonical variates in $L^2(0, a)$ and $L^2(b, 1)$. Note that now ρ^2 can be any number in $(0, 1]$.

Returning to the special intervals $[0, 1/2]$ and $[1/2, 1]$ again, it is of some interest to compute $\rho^2(\alpha)$ and a corresponding pair of canonical variates, for $\alpha>0$. Because $\sum_{m=1}^{\infty}\gamma_{km}\gamma_{jm}/(\alpha+\lambda_{2m})=\sum_{m=1}^{\infty}\langle E, \varphi_{1k}\rangle\langle 1, \varphi_{2m}\rangle\langle E, \varphi_{1j}\rangle\langle 1, \varphi_{2m}\rangle/(\alpha+\lambda_{2m})=\Lambda_2(\alpha)\langle E, \varphi_{1k}\rangle\langle E, \varphi_{1j}\rangle$, where

$$\Lambda_2(\alpha)=\sum_{m=1}^{\infty}\frac{\langle 1, \varphi_{1m}\rangle^2}{\alpha+\lambda_{2m}}=\|(\alpha I_2+S_{22})^{-1/2}1\|^2, \tag{3.67}$$

the operator $\mathcal{R}_\alpha$ in (3.40) reduces to one with elements

$$\mathcal{R}_\alpha(k, j)=\Lambda_2(\alpha)\frac{\langle E, \varphi_{1k}\rangle\langle E, \varphi_{1j}\rangle}{\sqrt{(\alpha+\lambda_{1k})(\alpha+\lambda_{1j})}}. \tag{3.68}$$

Similarly $\mathcal{Q}_\alpha$ reduces to an operator with elements

$$\mathcal{Q}_\alpha(k, j)=\Lambda_1(\alpha)\frac{\langle 1, \varphi_{2k}\rangle\langle 1, \varphi_{2j}\rangle}{\sqrt{(\alpha+\lambda_{2k})(\alpha+\lambda_{2j})}}, \tag{3.69}$$

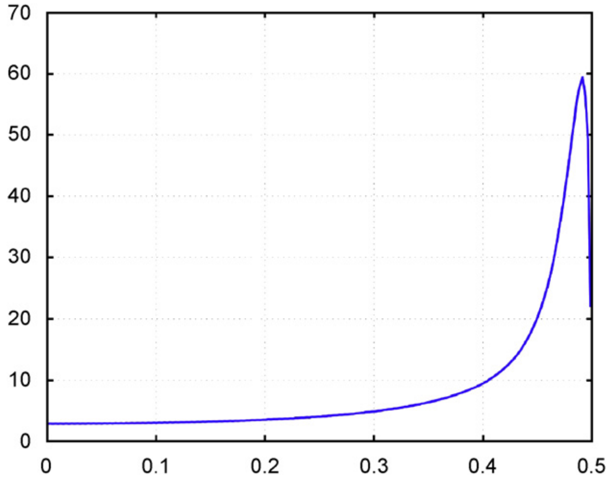

Fig. 1. Plot of $\overset{\vee}{f}$ for $\alpha = .0001$.

see (3.43), where

$$\Lambda_1(\alpha) = \sum_{m=1}^{\infty} \frac{\langle E, \varphi_{1m} \rangle^2}{\alpha + \lambda_{1m}} = \|(\alpha I_1 + S_{11})^{-1/2} E\|^2. \tag{3.70}$$

Thus we obtain from (3.41)

$$\rho^2(\alpha) = \Lambda_2(\alpha) \max_{\substack{x \in \ell^2 \\ \|x\|=1}} \left\{ \sum_{k=1}^{\infty} x_k \frac{\langle E, \varphi_{1k} \rangle}{\sqrt{\alpha + \lambda_{1k}}} \right\}^2. \tag{3.71}$$

The usual application of Schwarz's inequality yields the maximizer $\overset{\vee}{x}$ with the coordinates $\overset{\vee}{x}_k = \langle E, \varphi_{1k} \rangle / (\sqrt{\Lambda_1(\alpha)} \sqrt{\alpha + \lambda_{1k}})$. It follows that

$$\rho^2(\alpha) = \Lambda_1(\alpha) \Lambda_2(\alpha), \tag{3.72}$$

and that the canonical variate corresponding to $\overset{\vee}{x}$ is determined by the function

$$\overset{\vee}{f}_\alpha = \frac{1}{\sqrt{\Lambda_1(\alpha)}} \sum_{k=1}^{\infty} \frac{\langle E, \varphi_{1k} \rangle}{\sqrt{\alpha + \lambda_{1k}}} \varphi_{1k}. \tag{3.73}$$

Note the symmetry in $\Lambda_1(\alpha)$ and $\Lambda_2(\alpha)$ in the expression for $\rho^2(\alpha)$, although this expression is obtained from $\mathcal{R}_\alpha$. Also recall (3.55) for explicit evaluation of $\overset{\vee}{f}_\alpha$. Similarly it can be shown that the second canonical variate is determined by the function

$$\overset{\vee}{g}_\alpha = \frac{1}{\sqrt{\Lambda_2(\alpha)}} \sum_{k=1}^{\infty} \frac{\langle 1, \varphi_{2k} \rangle}{\sqrt{\alpha + \lambda_{2k}}} \varphi_{2k}. \tag{3.74}$$

The numerical calculations show that, for small $\alpha > 0$, the value of $\rho^2(\alpha)$ is close to $\rho^2 = 1$ (see Fig. 3). They also corroborate the conjecture that in a limiting sense the canonical variates are

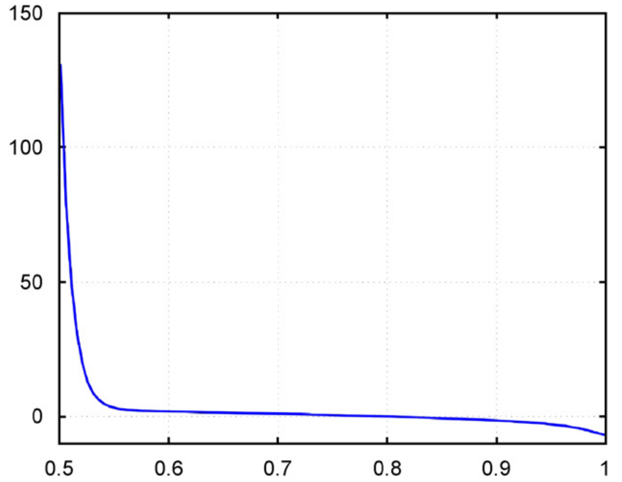

Fig. 2. Plot of $\overset{\vee}{g}$ for $\alpha = .0001$.

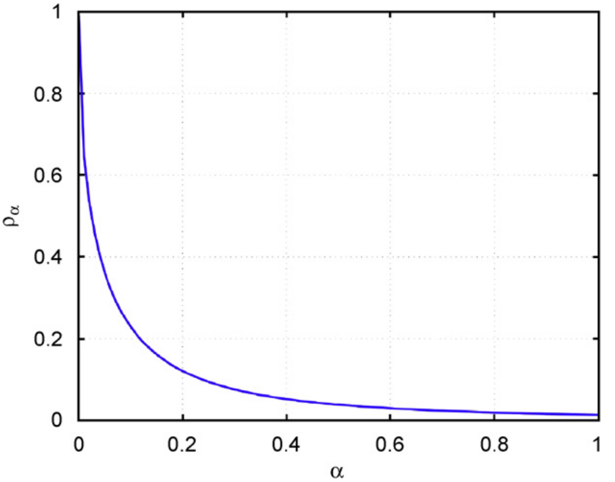

Fig. 3. Plot of $\rho^2(\alpha)$ for $0 \leqslant \alpha \leqslant 1$.

given by delta functions. Indeed, $\overset{\vee}{f}_\alpha$ is close to a delta-function with mass 1 concentrated near $1/2$ and to the left of that point (see Fig. 1), and $\overset{\vee}{g}_\alpha$ concentrates mass 1 near $1/2$ but to the right of that point (see Fig. 2).

Case 3: This case is hardly of any real interest and is merely included to establish another extreme. Suppose that $\{U_{11}, U_{12}, \ldots\} \perp\!\!\!\perp \{U_{21}, U_{22}, \ldots\}$. This entails that $\gamma_{km} = 0$ for all k and m and hence

$$X_1 \perp\!\!\!\perp X_2. \tag{3.75}$$

This naturally leads to $\rho^2 = 0 = \rho^2(\alpha)$ for all $\alpha > 0$. Also note that $\mathcal{R}_\alpha$ and $\mathcal{Q}_\alpha$ are the zero operator in this case.

4. Two arbitrary dynamical processes

Let $\{W(t), 0 \leqslant t \leqslant 1\}$ denote standard Brownian motion. In Sections 3.3 and 3.4 we were concerned with two processes X_1 and X_2 in orthogonal subspaces satisfying $X_1 \stackrel{\mathrm{d}}{=} W\mathbf{1}_{[0,1/2]}$, $X_2 \stackrel{\mathrm{d}}{=} W\mathbf{1}_{[1/2,1]}$, and such that X_1 and X_2 were dependent. In this subsection we will consider two arbitrary dynamical processes, with time intervals of equal length, for which a spectrum of dependencies will be specified. In order to deal with this situation as a special case of the general formulation in Sections 1 and 2, we consider the first process on [0, 1/2], and the second on [1/2, 1]. We restrict ourselves to processes with a Karhunen–Loève expansion.

Let $\widetilde{\varphi}_{1m}, \widetilde{\varphi}_{2m}$ ($m \in \mathbb{N}$) be two sequences of functions defined on [0, 1], such that

$$\varphi_{1m} = \mathbf{1}_{[0,1/2]}\widetilde{\varphi}_{1m}, \quad \varphi_{2m} = \mathbf{1}_{[1/2,1]}\widetilde{\varphi}_{2m}, \quad m \in \mathbb{N}, \tag{4.1}$$

are orthonormal bases in $L^2(0, 1/2)$ and $L^2(1/2, 1)$ respectively. Let

$$\lambda_{j1} > \lambda_{j2} > \cdots \downarrow 0, \quad j = 1, 2, \tag{4.2}$$

such that $\sum_{k=1}^{\infty} \lambda_{jk} < \infty$ ($j = 1, 2$), and suppose that

$$\begin{cases} Z_{jm} (j = 1, 2; m \in \mathbb{N}) \text{ are uncorrelated random variables} \\ \text{with mean 0 and variance 1.} \end{cases} \tag{4.3}$$

Choose $a_m, b_m \in \mathbb{R}$ such that

$$a_m^2 + b_m^2 = 1, \quad m \in \mathbb{N}. \tag{4.4}$$

We are now in a position to define the dynamical processes.

With the φ_{jm} as in (4.1), let us set

$$X_j = \sum_{m=1}^{\infty} U_{jm}\varphi_{jm}, \quad j = 1, 2, \tag{4.5}$$

where the random variables U_{jm} are given by

$$U_{1m} = \sqrt{\lambda_{1m}}Z_{1m}, \quad U_{2m} = \sqrt{\lambda_{2m}}(a_m Z_{1m} + b_m Z_{2m}). \tag{4.6}$$

Note that the U_{1m} are uncorrelated with variance λ_{1m} and, because of (4.4), the U_{2m} are uncorrelated with variance λ_{2m}. The "combined" process is

$$X = X_1 + X_2. \tag{4.7}$$

If P_1 and P_2 are projections of $\mathbb{H} = L^2(0, 1)$ on, respectively, $\mathbb{H}_1 = L^2(0, 1/2)$ and $\mathbb{H}_2 = L^2(1/2, 1)$, we indeed have

$$X_j = P_j X, \quad j = 1, 2. \tag{4.8}$$

A calculation completely analogous to the one in Section 3.4, Case 1, shows that in this case the canonical correlation $\rho^2(\alpha)$ between the processes is given by (3.41). Since

$$\gamma_{km} = \mathbb{E}U_{1k}U_{2m} = \sqrt{\lambda_{1k}\lambda_{2m}}a_m\delta_{km}, \tag{4.9}$$

see also (3.29), it follows that $\rho^2(\alpha)$ is the largest eigenvalue of the matrix $\mathcal{R}_\alpha$ with elements (cf. (3.40))

$$\mathcal{R}_\alpha(k, j) = \frac{1}{\sqrt{\alpha + \lambda_{1k}}} \left\{ \sum_{m=1}^{\infty} \frac{\gamma_{km}\gamma_{jm}}{\alpha + \lambda_{2m}} \right\} \frac{1}{\sqrt{\alpha + \lambda_{1j}}} = \begin{cases} \dfrac{a_k^2 \lambda_{1k}\lambda_{2k}}{(\alpha + \lambda_{1k})(\alpha + \lambda_{2k})}, & k = j, \\ 0, & k \neq j. \end{cases} \tag{4.10}$$

Apparently $\mathcal{R}_\alpha$ is a diagonal matrix. If we assume

$$1 \geqslant a_1^2 \geqslant a_2^2 \geqslant \cdots, \tag{4.11}$$

the largest eigenvalue of this matrix equals

$$\rho^2(\alpha) = \frac{a_1^2 \lambda_{11}\lambda_{21}}{(\alpha + \lambda_{11})(\alpha + \lambda_{21})}. \tag{4.12}$$

This example provides a whole range of possible values for the canonical correlations, depending on the choice of a_1^2. If $a_1^2 = 0$, we have $a_m^2 = 0$ for all m and obviously $X_1 \perp\!\!\!\perp X_2$ with $\rho^2(\alpha) = 0$. If a_1^2 is close to 1, and α is close to 0 we obtain $\rho^2(\alpha)$ close to 1.

5. The sample RSPCC

For the SPCC ρ^2 to be well defined we have required S to be injective. For all $\alpha > 0$ the RSPCC $\rho^2(\alpha)$, however, remains well defined even if S is not injective. If $X_1, \ldots, X_n$ is a random sample of independent copies of X, the usual estimator of S is

$$\widehat{S} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}) \otimes (X_i - \overline{X}) \quad \text{where } \overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i. \tag{5.1}$$

Because the range of $\widehat{S}$ has at most dimension n, this operator can never be injective, not even when S is. Therefore regularization is indispensable. Let us define

$$\widehat{S}_{ij} = P_i \widehat{S} P_j, \quad i, j = 1, 2. \tag{5.2}$$

Definition 5.1. The sample RSPCC is defined as

$$\widehat{\rho}^2(\alpha) = \max_{\substack{f \in \mathbb{H}_1^0 \\ g \in \mathbb{H}_2^0}} \frac{\langle f, \widehat{S}_{12} g\rangle^2}{\langle f, (\alpha I_1 + \widehat{S}_{11}) f\rangle \langle g, (\alpha I_2 + \widehat{S}_{22}) g\rangle}. \tag{5.3}$$

A corresponding pair of unit norm maximizers $\widehat{f}_\alpha \in \mathbb{H}_1^0$ and $\widehat{g}_\alpha \in \mathbb{H}_2^0$ determine the corresponding sample canonical variates.

Let us define the operators $\widehat{R}_\alpha$ and $\widehat{Q}_\alpha$ as in (2.21) and (2.22) but with $\widehat{S}_{ij}$ substituted for S_{ij}. Similarly let $\widehat{f}_\alpha$ and $\widehat{g}_\alpha$ be the maximizers of the quadratic forms in (2.23) and (2.24), respectively, when R_α and Q_α are replaced with $\widehat{R}_\alpha$ and $\widehat{Q}_\alpha$. The following is immediate from Theorem 2.3.

Theorem 5.1. *For each $\alpha > 0$ we have*

$$\begin{aligned}\widehat{\rho}^2(\alpha) &= \textit{largest eigenvalue of } \widehat{R}_\alpha = \langle \widehat{f}_\alpha, \widehat{R}_\alpha \widehat{f}_\alpha \rangle \\ &= \textit{largest eigenvalue of } \widehat{Q}_\alpha = \langle \widehat{g}_\alpha, \widehat{Q}_\alpha \widehat{g}_\alpha \rangle.\end{aligned} \tag{5.4}$$

The canonical variates are essentially unique if the eigenspaces corresponding to the maximal eigenvalue are one-dimensional.

This interpretation of the sample RSPCC and its canonical variates might enable us to determine the asymptotic normality of the correlation and possibly of its variates as well. A proof could be based on that in Ruymgaart and Yang [12] for matrices. It would require perturbation theory for functions of compact operators. We hope to deal with this topic, which is of a rather different nature, in a forthcoming paper. Asymptotic normality of the sample RSPCC would, of course, entail its consistency. Let us conclude this paper with a simple, direct proof of the consistency, for which perturbation theory is not required. See also Leurgans et al. [8] for their consistency result.

Theorem 5.2. *Suppose that $\mathbb{E}\|X\|^4 < \infty$. For fixed $\alpha > 0$ we have*

$$\widehat{\rho}^2(\alpha) \xrightarrow{\mathrm{p}} \rho^2(\alpha) \quad \textit{as } n \to \infty. \tag{5.5}$$

Proof. Under the present assumption it is known [3] that

$$\sqrt{n}(\widehat{S} - S) \xrightarrow{\mathrm{d}} \mathcal{G} \quad \text{as } n \to \infty \text{ in } \mathcal{L}(\mathrm{HS}), \tag{5.6}$$

where $\mathcal{L}(\mathrm{HS})$ is the Hilbert space of all Hilbert–Schmidt operators mapping $\mathbb{H}$ into itself, equipped with the inner product $\langle \cdot, \cdot \rangle_{\mathrm{HS}}$. Since $\|T\| \leqslant \|T\|_{\mathrm{HS}}$, for $T \in \mathcal{L}(\mathrm{HS})$, where $\|T\|$ is the ordinary operator norm, it follows via the continuous mapping theorem that

$$\|\widehat{S}_{ij} - S_{ij}\| \xrightarrow{\mathrm{p}} 0 \quad \text{as } n \to \infty, \;\; i, j = 1, 2. \tag{5.7}$$

For brevity let us write

$$\widehat{a}(f, g) = \langle f, \widehat{S}_{12} g \rangle^2, \quad a(f, g) = \langle f, S_{12} g \rangle^2, \tag{5.8}$$

$$\widehat{b}(f) = \langle f, (\alpha I_1 + \widehat{S}_{11}) f \rangle, \quad b(f) = \langle f, (\alpha I_1 + S_{11}) f \rangle, \tag{5.9}$$

$$\widehat{c}(g) = \langle g, (\alpha I_2 + \widehat{S}_{22}) g \rangle, \quad c(g) = \langle g, (\alpha I_2 + S_{22}) g \rangle. \tag{5.10}$$

We will also write capitals to indicate maxima or suprema over unit balls like, for instance

$$\max_{\substack{f \in \mathbb{H}_1^0, g \in \mathbb{H}_2^0 \\ \|f\|=1, \|g\|=1}} = \mathop{\mathrm{MAX}}_{f,g}. \tag{5.11}$$

With these notational conventions we have

$$\begin{aligned}|\widehat{\rho}^2(\alpha) - \rho^2(\alpha)| &= \left| \mathop{\mathrm{MAX}}_{f,g} \frac{\widehat{a}(f,g)}{\widehat{b}(f)\widehat{c}(g)} - \mathop{\mathrm{MAX}}_{f,g} \frac{a(f,g)}{b(f)c(g)} \right| \\ &\leqslant \mathop{\mathrm{SUP}}_{f,g} \left| \frac{\widehat{a}(f,g)b(f)c(g) - a(f,g)\widehat{b}(f)\widehat{c}(g)}{\widehat{b}(f)\widehat{c}(g)b(f)c(g)} \right| \\ &\leqslant \frac{1}{\alpha^4} \mathop{\mathrm{SUP}}_{f,g} \left| \widehat{a}(f,g)b(f)c(g) - a(f,g)\widehat{b}(f)\widehat{c}(g) \right|,\end{aligned} \tag{5.12}$$

because, for instance

$$\operatorname*{MIN}_{f} \widehat{b}(f) = \text{smallest eigenvalue of } (\alpha I_1 + \widehat{S}_{11}) = \alpha. \tag{5.13}$$

Next observe that

$$\begin{aligned}
&\operatorname*{SUP}_{f,g} \left|\widehat{a}(f,g)b(f)c(g) - a(f,g)\widehat{b}(f)\widehat{c}(g)\right| \\
&\quad \leqslant \operatorname*{SUP}_{f,g} |\widehat{a}(f,g) - a(f,g)|\, b(f)c(g) \\
&\qquad + \operatorname*{SUP}_{f,g} \left|\widehat{b}(f) - b(f)\right| |a(f,g)|c(g) \\
&\qquad + \operatorname*{SUP}_{f,g} |\widehat{c}(g) - c(g)|\, |a(f,g)|\widehat{b}(f).
\end{aligned} \tag{5.14}$$

The first term in this upper bound is bounded above by

$$\begin{aligned}
&\operatorname*{SUP}_{f,g} \left\{\left|\langle f, (\widehat{S}_{12} - S_{12})g\rangle\langle f, (\widehat{S}_{12} + S_{12})g\rangle\right| b(f)c(g)\right\} \\
&\quad \leqslant \operatorname*{SUP}_{f,g} \Big\{\|f\|^4\|g\|^4\|\widehat{S}_{12} - S_{12}\|_{\mathcal{L}} \\
&\qquad \times \|\widehat{S}_{12} + S_{12}\|_{\mathcal{L}}\|\alpha I_1 + S_{11}\|_{\mathcal{L}}\|\alpha I_2 + S_{22}\|_{\mathcal{L}}\Big\} \\
&\quad = \|\widehat{S}_{12} - S_{12}\|_{\mathcal{L}}\|\widehat{S}_{12} + S_{12}\|_{\mathcal{L}}\|\alpha I_1 + S_{11}\|_{\mathcal{L}}\|\alpha I_2 + S_{22}\|_{\mathcal{L}} \\
&\quad \xrightarrow{\mathrm{p}} 0 \quad \text{as } n \to \infty,
\end{aligned} \tag{5.15}$$

because of (5.7). The other two terms may be dealt with similarly. □

Acknowledgment

The authors are grateful to a referee for useful comments.

References

[1] P. Bickel, Personal communication, 2006.
[2] D. Bosq, Linear Processes in Function Spaces. Lecture Notes in Statistics, Springer, New York, 2000.
[3] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference, J. Multivar. Anal. 12 (1982) 136–154.
[4] L. Debnath, P. Mikusiński, Introduction to Hilbert Spaces with Applications, Academic Press, New York, 1999.
[5] B. Friedman, Principles and Techniques of Applied Mathematics, Dover, New York, 1990.
[6] G. He, H.-G. Müller, J.-L. Wang, Functional canonical analysis for square integrable stochastic processes, J. Multivariate Anal. 85 (2002) 54–77.
[7] R.G. Laha, V.K. Rohatgi, Probability Theory, Wiley, New York, 1979.
[8] S.E. Leurgans, R.A. Moyeed, B.W. Silverman, Canonical correlation analysis when the data are curves, J. Roy. Statist. Soc. B 55 (1993) 725–740.
[9] K.V. Mardia, J.T. Kent, J.M. Bibby, Multivariate Analysis, Academic Press, New York, 1979.
[10] A. Mas, Estimation d'opérateurs de corrélation de processus fonctionnels: lois limites, tests, déviations modérées, Thèse de Doctorat, Université Paris VI, 2000.
[11] J. Rice, B.W. Silverman, Estimating the mean and covariance structure nonparametrically when the data are curves, J. Roy. Statist. Soc. B 53 (1991) 233–243.
[12] F.H. Ruymgaart, S. Yang, Some applications of Watson's perturbation approach to random matrices, J. Multivariate Anal. 60 (1997) 48–60.
[13] B.W. Silverman, Smoothed functional principal component analysis by choice of norm, Ann. Statist. 24 (1996) 1–24.
[14] H.D. Vinod, Canonical ridge and the econometrics of joint production, J. Econometr. 4 (1976) 147–166.