Inference under functional proportional and common principal components models

Graciela Boente[1] , Daniela Rodriguez[1] and Mariela Sued[1]

[1] *Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina*

e–mail: gboente@dm.uba.ar drodrig@dm.uba.ar msued@dm.uba.ar

Abstract

In many situations, when dealing with several populations with different covariance operators, equality of the operators is assumed. Usually, if this assumption does not hold, one estimates the covariance operator of each group separately, which leads to a large number of parameters. As in the multivariate setting, this is not satisfactory since the covariance operators may exhibit some common structure. In this paper, we discuss the extension to the functional setting of common principal component model that has been widely studied when dealing with multivariate observations. Moreover, we also consider a proportional model in which the covariance operators are assumed to be equal up to a multiplicative constant. For both models, we present estimators of the unknown parameters and we obtain their asymptotic distribution. A test for equality against proportionality is also considered.

Some key words: Common principal components, Eigenfunctions, Functional data analysis, Hilbert-Schmidt operators, Kernel methods, Proportional model.

Corresponding Author
Graciela Boente
Moldes 1855, 3^o A
Buenos Aires, C1428CRA
Argentina
email: gboente@dm.uba.ar
FAX 54-11-45763375

Running Head: Functional common principal component model.

1 Introduction

Functional data analysis is an emerging field in statistics that has received considerable attention during the last decade due to its applications to many biological problems. It provides modern data analytical tools for data that are recoded as images or as a continuous phenomenon over a period of time. Because of the intrinsic nature of these data, they can be viewed as realizations of random functions $X_1(t), \ldots, X_n(t)$ often assumed to be in $L^2(\mathcal{I})$, with $\mathcal{I}$ a real interval or a finite dimensional Euclidean set. In this context, principal components analysis offers an effective way for dimension reduction and it has been extended from the traditional multivariate setting to accommodate functional data. In the functional data analysis literature, it is usually referred to as functional principal component analysis (FPCA). Since the pioneer work by Rao [16], further analysis on functional data has been developed, for instance, by Rice and Silverman [17] or Ramsay and Dalzell [13]. See also, Ramsay and Silverman [14], Ramsay and Silverman [15], Ferraty and Vieu [6]. In particular, functional principal component analysis was studied by Dauxois, Pousse and Romain [5], Besse and Ramsay [2], Pezzulli and Silverman [12], Silverman [18] and Cardot [4]. Several examples and applications can be found in these references.

Let us consider a random function $X(t)$ where $t \in \mathcal{I} = [0,1]$ with mean $\mu(t) = E(X(t))$ and covariance operator $\mathbf{\Gamma}$. Let $\gamma(s,t) = \text{COV}(X(s), X(t)),\ s,t \in \mathcal{I}$. Under general conditions, the covariance function may be expressed as

$$\gamma(s,t) = \sum_{i \geq 1} \lambda_i \phi_i(s) \phi_i(t)$$

where the λ_j are the ordered eigenvalues, $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ of the covariance operator and the functions ϕ_j the associated orthonormal eigenfunctions with the usual inner product in $L^2[0,1]$. Then, the spectral decomposition of the covariance operator, which is the analogous of a covariance matrix in a function space, allows to get a small dimension space which exhibits the main modes of variation of the data. Effectively, the well–known Karhunen–Loéve expansion allows to write the process as

$$X \ = \mu + \sum_{j=1}^{\infty} \beta_j \, \phi_j$$

where $\langle X - \mu, \phi_j \rangle = \beta_j$ are random scalar loadings such that $E(\beta_j) = 0$, $E(\beta_j^2) = \lambda_j$ and $E(\beta_j \, \beta_k) = 0$ for $j \neq k$. Note that the process can also be written as

$$X \ = \mu + \sum_{j=1}^{\infty} \lambda_j^{\frac{1}{2}} \, f_j \, \phi_j$$

with f_j random variables such that $E(f_j) = 0$, $E(f_j^2) = 1$, $E(f_j \ f_s) = 0$ for $j \neq s$. This representation provides a nice interpretation of the principal component analysis in the functional setting, since $\phi_1(t)$, $\phi_2(t), \ldots$ represent the major modes of variation of $X(t)$ over t.

In this paper, we go further since we generalize the previous ideas to the setting in which we are dealing with several populations. In many situations, we have independent observations $X_{i,1}(t), \cdots, X_{i,n_i}(t)$ from k independent samples of smooth random functions in $L^2[0,1]$ with mean μ_i and different covariance operators $\mathbf{\Gamma}_i$. However, as it is the case in the finite–dimensional setting, the covariance operators may exhibit some common structure and it is sensible to take it into account when estimating them. A simple generalization of equal covariance operators consists of assuming their proportionality, i.e., $\mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1$, for $1 \le i \le k$ and $\rho_1 = 1$.

The common principal components model, introduced by Flury [7] for $p-$th dimensional data, generalizes proportionality of the covariance matrices by allowing the matrices to have different eigenvalues but identical eigenvectors, that is, $\mathbf{\Sigma}_i = \boldsymbol{\beta} \mathbf{\Lambda}_i \boldsymbol{\beta}^{\mathrm{T}}$, $1 \le i \le k$, where the $\mathbf{\Lambda}_i$ are diagonal matrices and $\boldsymbol{\beta}$ is the orthogonal matrix of the common eigenvectors. This model can be viewed as a generalization of principal components to k groups, since the principal transformation is identical in all populations considered while the variances associated with them vary among groups. In biometric applications, principal components are frequently interpreted as independent factors determining the growth, size or shape of an organism. It seems therefore reasonable to consider a model in which the same factors arise in different, but related species. The common principal components model clearly serves this purpose.

A natural extension to the functional setting of the common principal components model introduced by Flury [7] is to assume that the covariance operators $\mathbf{\Gamma}_i$ have common eigenfunctions $\phi_j(t)$ but different eigenvalues λ_{ij}. In this sense, the processes $X_{i,1}(t)$, $1 \le i \le k$ can be written as

$$X_{i,1} \; = \mu_i + \sum_{j=1}^{\infty} \lambda_{ij}^{\frac{1}{2}} \, f_{ij} \, \phi_j$$

with $\lambda_{i1} \ge \lambda_{i2} \ge \ldots \ge 0$ and f_{ij} random variables such that $E(f_{ij}) = 0$, $E(f_{ij}^2) = 1$, $E(f_{ij} \; f_{is}) = 0$ for $j \ne s$ and so, the common eigenfunctions, as in the one–population setting, exhibit the same major modes of variation. We will denote this model the functional common principal component (FCPC) model. As in principal component analysis, the FCPC model could be used to reduce the dimensionality of the data, retaining as much as possible of the variability present in each of the populations. Besides, this model provides a framework for analyzing different population data that share their main modes of variation $\phi_1, \phi_2, \ldots$. It is worth noticing that when considering a functional proportional model, $X_{i,1}(t)$, $1 \le i \le k$ can be written as $X_{i,1} \; = \mu_i + \; \rho_i \sum_{j=1}^{\infty} \lambda_j^{\frac{1}{2}} \, f_{ij} \, \phi_j$, with $\rho_1 = 1$, $\lambda_1 \ge \lambda_2 \ge \ldots \ge 0$ and f_{ij} are random variables as described above. A similar problem was recently studied by Benko, Härdle and Kneip [1] who considered the case of $k = 2$ populations and provide tests for equality of means and equality of a fixed number of eigenfunctions.

The aim of this paper is to provide estimators of the common eigenfunctions under a FCPC model and to study their asymptotic behavior, as well as to consider estimators of the proportionality constants under a functional proportional model. In Section 2, we introduce the notation that will be used in the paper while in Section 3, we describe the estimators for the restricted models. Under a FCPC, two families of estimators for the common eigenfunctions are considered. Besides, the proportionality constant estimators defined under a functional proportional model allow to construct an asymptotic test to decide between equality against proportionality of the covariance operators which corresponds to the first two hierarchical levels considered in the finite–dimensional case, by Flury [9]. The asymptotic distribution of the given proposals is stated in Section 4. Proofs are given in the Appendix.

2 Notation and Preliminaries

Let $X_{i,1}(t), \cdots, X_{i,n_i}(t)$, $1 \leq i \leq k$, be independent observations from k independent samples of smooth random functions in $L^2(\mathcal{I})$, where $\mathcal{I} \subset I\!R$ is a finite interval, with mean μ_i. Without loss of generality, from now on, we will assume that $\mathcal{I} = [0,1]$. Denote by γ_i and $\mathbf{\Gamma}_i$ the covariance function and operator, respectively, related to each population. To be more precise, we are assuming that $\{X_{i,1}(t) : t \in \mathcal{I}\}$ are k stochastic processes defined in $(\Omega, \mathcal{A}, P)$ with continuous trajectories, mean μ_i and finite second moment, i.e., $E\left(X_{i,1}(t)\right) = \mu_i(t)$ and $E\left(X_{i,1}^2(t)\right) < \infty$ for $t \in \mathcal{I}$. Each covariance function $\gamma_i(t,s) = \text{COV}(X_{i,1}(s), X_{i,1}(t))$, $s,t \in \mathcal{I}$ has an associated linear operator $\mathbf{\Gamma}_i : L^2[0,1] \to L^2[0,1]$ defined as $\left(\mathbf{\Gamma}_i\, u\right)(t) = \int_0^1 \gamma_i(t,s)u(s)ds$, for all $u \in L^2[0,1]$. As in the case of one population, throughout this paper, we will assume that the covariance operators satisfy $\|\gamma_i\|^2 = \int_0^1 \int_0^1 \gamma_i^2(t,s)dtds < \infty$. The Cauchy-Schwartz inequality implies that $|\mathbf{\Gamma}_i u|^2 \leq \|\gamma_i\|^2 |u|^2$, where $|u|$ stands for the usual norm in the space $L^2[0,1]$.Therefore, $\mathbf{\Gamma}_i$ is a self–adjoint continuous linear operator. Moreover, $\mathbf{\Gamma}_i$ is a Hilbert-Schmidt operator. $\mathcal{F}$ will stand for the Hilbert space of such operators with inner product defined by $\langle \mathbf{\Gamma}_1, \mathbf{\Gamma}_2 \rangle_{\mathcal{F}} = \text{trace}(\mathbf{\Gamma}_1\mathbf{\Gamma}_2) = \sum_{j=1}^{\infty} \langle \mathbf{\Gamma}_1 u_j, \mathbf{\Gamma}_2 u_j \rangle$, where $\{u_j : j \geq 1\}$ is any orthonormal basis of $L^2[0,1]$ and $\langle u, v \rangle$ denotes the usual inner product in $L^2[0,1]$. Choosing a basis $\{\phi_{ij} : j \geq 1\}$ of eigenfunctions of $\mathbf{\Gamma}_i$ we have that $\|\mathbf{\Gamma}_i\|_{\mathcal{F}}^2 = \sum_{j=1}^{\infty} \lambda_{ij}^2 = \|\gamma_i\|^2 < \infty$, where $\{\lambda_{ij} : j \geq 1\}$ are the eigenvalues of $\mathbf{\Gamma}_i$. Note that under the FCPC model, the basis is the same for all populations.

As mentioned in the Introduction, when dealing with one population, non–smooth estimators of the eigenfunctions and eigenvalues of $\mathbf{\Gamma}$ were considered by Dauxois, Pousse and Romain [5], in a natural way through the empirical covariance operator. More precisely, the non–smooth estimators of the population functional principal component ϕ_k are the eigenfunction $\widehat{\phi}_k$ related to the $k-$th largest eigenvalue $\widehat{\lambda}_k$ of the random operator $\widehat{\mathbf{\Gamma}}_n$ where $\widehat{\mathbf{\Gamma}}_n$ is the linear operator related to the empirical covariance function $\widehat{\gamma}_n(t,s) = \sum_{j=1}^{n} \left(X_j(t) - \overline{X}(t)\right)\left(X_j(s) - \overline{X}(s)\right)/n$. Smooth versions of the previous estimates have been defined adding a penalty term or using a kernel approach. Smooth estimators of the covariance operators are useful when dealing with sparse data or when one wants to guarantee smoothness of the resulting common principal components. When dealing with one population, Ramsay and Silverman [15] argue for smoothness properties of the principal components as "for many data sets, PCA of functional data is more revealing if some type of smoothness is required to the principal components themselves". The same ideas apply when dealing with several populations sharing their eigenfunctions.

One way to perform smooth principal component analysis is through roughness penalties on the sample variance or on the L^2-norm, as defined by Rice and Silverman [17] and by Silverman [18], respectively, where consistency results were obtained. A different approach is a kernel–based one which corresponds to smooth the functional data and then perform PCA on the smoothed trajectories. In Boente and Fraiman [3] it is shown that the degree of regularity of kernel–based principal components is given by that of the kernel function used. See also Ramsay and Dalzell [13], Ramsay and Silverman [15] and Ferraty and Vieu [6]. Under a FCPC model, the kernel smoothing procedure becomes easier to implement and allows to derive the properties of the resulting estimators from those of the estimators of the covariance operator.

We will give two proposals to estimate the common eigenfunctions under a FCPC model. Both of them are based on estimators of the covariance operators. As mentioned above, for each population, one can consider either the non–smooth estimators studied in Dauxois, Pousse and Romain [5] or the kernel proposal studied in Boente and Fraiman [3], since under mild conditions they both have the same asymptotic distribution. For the sake of completeness, we briefly remind their definition in the actual setting.

The empirical covariance functions $\widehat{\gamma}_{i,\mathrm{R}}$ or the smoothed version of them $\widehat{\gamma}_{i,\mathrm{S}}(t,s)$ are defined as

$$\widehat{\gamma}_{i,\mathrm{R}}(s,t) = \frac{1}{n_i}\sum_{j=1}^{n_i}\left(X_{i,j}(s)-\overline{X}_i(s)\right)\left(X_{i,j}(t)-\overline{X}_i(t)\right) \tag{2.1}$$

$$\widehat{\gamma}_{i,\mathrm{S}}(s,t) = \frac{1}{n_i}\sum_{j=1}^{n_i}\left(X_{i,j,\,h}(s)-\overline{X}_{i,\,h}(s)\right)\left(X_{i,j,\,h}(t)-\overline{X}_{i,\,h}(t)\right)\,, \tag{2.2}$$

where $X_{i,j,\,h}(t)=\int K_h(t-s)X_{i,j}(s)ds$ are the smoothed trajectories and $K_h(.)=h^{-1}K(./h)$ is a nonnegative kernel function with smoothing parameter h, such that $\int K(u)du=1$ and $\int K^2(u)du<\infty$. The linear operators related to $\widehat{\gamma}_{i,\mathrm{R}}$ and $\widehat{\gamma}_{i,\mathrm{S}}$ will be denoted by $\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$ and by $\widehat{\mathbf{\Gamma}}_{i,\mathrm{S}}$, respectively. Methods for selecting the smoothing parameter h can be developed using cross–validation methods as it was described for penalizing methods in Section 7.5 in Ramsay and Silverman [15] but adapted to the problem of estimating the common directions, i.e., when considering the cross validation loss, the $i-$th sample should be centered with an estimator of μ_i.

Assume $n_i=\tau_i N$ with $0<\tau_i<1$ fixed numbers such that $\sum_{i=1}^k\tau_i=1$ and where $N=\sum_{i=1}^k n_i$ denotes the total number of observations in the sample. Define the weighted covariance function as $\gamma=\sum_{i=1}^k\tau_i\gamma_i$ and its related operator as $\mathbf{\Gamma}=\sum_{i=1}^k\tau_i\mathbf{\Gamma}_i$. Therefore, estimators of the weighted covariance function γ can be defined as $\widehat{\gamma}_{\mathrm{R}}=\sum_{i=1}^k\tau_i\widehat{\gamma}_{i,\mathrm{R}}$ and $\widehat{\mathbf{\Gamma}}_{\mathrm{R}}=\sum_{i=1}^k\tau_i\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$ or $\widehat{\gamma}_{\mathrm{S}}=\sum_{i=1}^k\tau_i\widehat{\gamma}_{i,\mathrm{S}}$ and $\widehat{\mathbf{\Gamma}}_{\mathrm{S}}=\sum_{i=1}^k\tau_i\widehat{\mathbf{\Gamma}}_{i,\mathrm{S}}$, the raw or smoothed estimators of γ and $\mathbf{\Gamma}$, respectively. It is worth noticing that our results do not make use of the explicit expression of the covariance operators, but they only require their consistency and asymptotic normality.

3 The proposals

3.1 Estimators of the common eigenfunctions and their size under a FCPC model

Let us assume that the FCPC model hold, i.e., that the covariance operators $\boldsymbol{\Gamma}_i$ have common eigenfunctions $\phi_j(t)$ but possible different eigenvalues λ_{ij} where λ_{ij} denotes the eigenvalue related to the eigenfunction ϕ_j, i.e., $\lambda_{ij} = \langle \phi_j, \boldsymbol{\Gamma}_i \phi_j \rangle$. Moreover, we will assume that the eigenvalues preserve the order among populations, i.e., throughout this paper we will assume that

A1. $\lambda_{i1} \geq \lambda_{i2} \geq \cdots \geq \lambda_{ip} \geq \lambda_{ip+1} \cdots$, for $1 \leq i \leq k$

A2. There exists ℓ such that for any $1 \leq j \leq \ell$, there exists $1 \leq i \leq k$ such that $\lambda_{ij} > \lambda_{i\,j+1}$.

Assumption **A2** is weaker than assuming that for any $j \geq 1$, there exists $1 \leq i \leq k$ such that $\lambda_{ij} > \lambda_{i\,j+1}$ since it allows for finite rank operators. Note that if $\sum_{i=1}^k \tau_i \lambda_{ij} > \sum_{i=1}^k \tau_i \lambda_{i\,j+1}$ for any $j \geq 1$, then **A2** is fulfilled for any value ℓ.

As mentioned in Section 2, we will assume that $n_i = \tau_i N$ with $0 < \tau_i < 1$ fixed numbers such that $\sum_{i=1}^k \tau_i = 1$ and $N = \sum_{i=1}^k n_i$.

The first proposal is based on the fact that under the FCPC model, the common eigenfunctions $\{\phi_j : j \geq 1\}$ are also a basis of eigenfunctions for the operator $\boldsymbol{\Gamma} = \sum_{i=1}^k \tau_i \boldsymbol{\Gamma}_i$, with eigenvalues given by

$$\nu_1 = \sum_{i=1}^k \tau_i \lambda_{i1} \geq \cdots \geq \nu_p = \sum_{i=1}^k \tau_i \lambda_{ip} \geq \nu_{p+1} = \sum_{i=1}^k \tau_i \lambda_{i\,p+1} \cdots .$$

Note that **A1** and **A2** entail that the first ℓ eigenfunctions will be related to the ℓ largest eigenvalues of the operator $\boldsymbol{\Gamma}$, having multiplicity one and being strictly positive. A first attempt to estimate the common eigenfunctions consists in considering the eigenfunctions $\widetilde{\phi}_j$ related to the largest eigenvalues $\widehat{\nu}_j$ of a consistent estimator $\widehat{\boldsymbol{\Gamma}}$ of $\boldsymbol{\Gamma}$, obtained as $\widehat{\boldsymbol{\Gamma}} = \sum_{i=1}^k \tau_i \widehat{\boldsymbol{\Gamma}}_i$ where $\widehat{\boldsymbol{\Gamma}}_i$ denotes any estimator of the $i-$th covariance operator. Examples of such estimators are, for instance, the empirical covariance functions or the smoothed version of them described in Section 2. The eigenvalue estimators can then be defined as $\widehat{\lambda}_{ij} = \langle \widetilde{\phi}_j, \widehat{\boldsymbol{\Gamma}}_i \widetilde{\phi}_j \rangle$.

The second proposal tries to improve the efficiency of the previous one for gaussian processes. To that purpose, we will have in mind that, in the finite–dimensional case, the maximum likelihood estimators of the common directions for normal data solve a system of equations involving both the eigenvalue and eigenvector estimators (see Flury, [7]). To be more precise, let $\mathbf{Y}_{i,1}, \ldots, \mathbf{Y}_{i,n_i}$, $1 \leq i \leq k$ be k independent samples of normally distributed random vectors in $I\!R^p$ with covariance matrices $\boldsymbol{\Sigma}_i$ satisfying a CPC model, i.e., such that $\boldsymbol{\Sigma}_i = \boldsymbol{\beta} \boldsymbol{\Lambda}_i \boldsymbol{\beta}^{\mathrm{T}}$, $1 \leq i \leq k$. Then, the maximum likelihood estimators, $\widehat{\boldsymbol{\beta}}$, of $\boldsymbol{\beta}$ solve the system of equations

$$\widehat{\boldsymbol{\beta}}_m^{\mathrm{T}} \left[\sum_{i=1}^{k} \tau_i \frac{\widehat{\lambda}_{im} - \widehat{\lambda}_{ij}}{\widehat{\lambda}_{im} \widehat{\lambda}_{ij}} \mathbf{S}_i \right] \widehat{\boldsymbol{\beta}}_j = 0 \qquad \text{for } m \neq j \tag{3.1}$$
$$\widehat{\boldsymbol{\beta}}_m^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_j = \delta_{mj} \ ,$$

where $\mathbf{S}_i = \sum_{j=1}^{n_i} \left(\mathbf{Y}_{i,j} - \overline{\mathbf{Y}}_{i,j} \right) \left(\mathbf{Y}_{i,j} - \overline{\mathbf{Y}}_{i,j} \right)^{\mathrm{T}} / n_i$ is the sample covariance matrix of the $i-$th population and $\widehat{\lambda}_{im} = \widehat{\boldsymbol{\beta}}_m^{\mathrm{T}} \mathbf{S}_i \widehat{\boldsymbol{\beta}}_m$.

Using consistent estimators of the eigenvalues, we generalize this system to the infinite–dimensional case. Effectively, let $\widehat{\lambda}_{ij}$ be initial estimators of the eigenvalues and $\widehat{\boldsymbol{\Gamma}}_i$ any consistent estimator of the covariance operator of the $i-$th population. Define for $j < \ell$ and $m < \ell$,

$$\widehat{\boldsymbol{\Gamma}}_{mj} = \sum_{i=1}^{k} \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im} \widehat{\lambda}_{ij}} \widehat{\boldsymbol{\Gamma}}_i \ , \tag{3.2}$$

which will be asymptotically well defined under **A2** if in addition $\lambda_{i\ell} > 0$ for $1 \leq i \leq k$. Let us consider the solution $\widehat{\phi}_j$ of the system of equations

$$\begin{cases} \delta_{mj} = & \langle \widehat{\phi}_m, \widehat{\phi}_j \rangle \\ 0 = & \langle \widehat{\phi}_m, \widehat{\boldsymbol{\Gamma}}_{mj} \widehat{\phi}_j \rangle \qquad 1 \leq j < m \ . \end{cases} \tag{3.3}$$

The FCPC method can be viewed as a simultaneous rotation of the axes yielding variables that are as uncorrelated as possible over the k groups. Moreover, as in the finite–dimensional setting (3.3) can be viewed as a generalized system of characteristic equations. If all the operators $\widehat{\boldsymbol{\Gamma}}_{mj}$ were identical to say $\widetilde{\boldsymbol{\Gamma}}$, then the characteristic eigenfunctions of $\widetilde{\boldsymbol{\Gamma}}$ will be a solution of (3.3). It is well known that for finite–dimensional normal populations with covariance matrices satisfying a CPC model, the solution of (3.1), being the maximum likelihood estimators, will provide efficient estimators of the common directions. This suggests that solving (3.3) will improve the asymptotic variance of the eigenfunctions of $\widehat{\boldsymbol{\Gamma}}$ for gaussian processes.

To summarize, the two proposals to estimate the common principal eigenfunctions and the eigenvalues of each population can be described as follows. Let $\widehat{\boldsymbol{\Gamma}}_i$ be a consistent estimator of the covariance operator of the $i-$th population

- *Proposal 1.* Define the pooled operator as $\widehat{\boldsymbol{\Gamma}} = \sum_{i=1}^{k} \tau_i \widehat{\boldsymbol{\Gamma}}_i$. Then, the estimators of the common eigenfunction ϕ_j can be defined as the eigenfunction $\widetilde{\phi}_j$ related to the $j-$th largest eigenvalue of $\widehat{\boldsymbol{\Gamma}}$. Besides, the estimator of the $j-$th eigenvalue of the $i-$th population, λ_{ij}, is defined as $\widetilde{\lambda}_{ij} = \langle \widetilde{\phi}_j, \widehat{\boldsymbol{\Gamma}}_i \widetilde{\phi}_j \rangle$.

- *Proposal 2.* Given initial consistent estimators of the eigenvalues $\widehat{\lambda}_{ij}$ define $\widehat{\boldsymbol{\Gamma}}_{mj}$ as in (3.2). Then, the solution $\{\widehat{\phi}_j\}_{j \geq 1}$ of the system (3.3) will provide estimators of the common eigenfunctions.

3.2 Computational methods for FCPC

To compute the family of estimators defined in (3.3), we can proceed as follows. Let $\{\alpha_s\}_{s\geq 1}$ be any orthonormal basis of $L^2[0,1]$, $p = p_N$ an increasing sequence of integers such that $p_N < N$ and define $Y_{i,j,s} = \langle \alpha_s, X_{i,j}\rangle$, for $1 \leq s \leq p$. When $\alpha_s = \phi_s$ the covariance matrices, $\mathbf{\Sigma}_i$, of $\mathbf{Y}_{i,1} = (Y_{i,1,1}, \ldots, Y_{i,1,p})^{\mathrm{T}}$ satisfy a CPC model since they are diagonal. However, since the eigenfunctions are our target, we have to consider a known orthonormal basis of $L^2[0,1]$, such as the Fourier basis. In this case, $\mathbf{\Sigma}_i$ can be approximated (through an order p truncation) by symmetric and non–negative definite commutable matrices with common eigenvectors $\boldsymbol{\beta}_j = (\langle \alpha_1, \phi_j\rangle, \ldots, \langle \alpha_p, \phi_j\rangle)^{\mathrm{T}}$. In order to obtain a solution $\widehat{\phi}_j$ of (3.3), we will use the solution $\widehat{\boldsymbol{\beta}}_j = (\widehat{\beta}_{j1}, \ldots, \widehat{\beta}_{jp})^{\mathrm{T}}$ of (3.1) where the matrices $\mathbf{S}_i$ are such that the $(m,s)-$component of $\mathbf{S}_i$ equals $\langle \alpha_m, \widehat{\mathbf{\Gamma}}_i \alpha_s\rangle$. Therefore, for $1 \leq j \leq p$, a solution $\widehat{\phi}_j$ of (3.3) that provides estimators of the common eigenfunctions can be computed as $\widehat{\phi}_j = \sum_{s=1}^{p} \widehat{\beta}_{js}\alpha_s$. It is worth noticing that this is equivalent to solving (3.3) with the truncated finite expansion (of order p) of $\widehat{\phi}_j = \sum_{s\geq 1}\langle \alpha_s, \widehat{\phi}_j\rangle \alpha_s$. Note that considering as $\mathbf{S}_i$ the sample covariance matrix of the $i-$th finite–dimensional observations $\{Y_{i,j}\}_{1\leq j\leq n_i}$ corresponds to the sample covariance operator $\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$. On the other hand, using as $\widehat{\mathbf{\Gamma}}_i$ the smoothed covariance estimators, $\widehat{\mathbf{\Gamma}}_{i,\mathrm{S}}$, to construct $\mathbf{S}_i$ is equivalent to define $\mathbf{Y}_{i,j}$ through the smoothed trajectories $X_{i,j,\,h}$ in the above description.

The approach of basis expansion of the functional data to obtain the principal components in a one–sample setting is discussed in Ramsay and Silverman [15] where they argue that the number p of basis functions depends on the sample size N and on the number of sampling points if the whole curve is not observed, on the level of smoothing imposed by using $p_N < N$ and on how efficient the basis reproduces the behavior of the data, among others. Moreover, they recommend to use a basis expansion of order p only to calculate more than a fairly small proportion of p eigenfunctions. A cross–validation method can be developed in our setting by choosing first the number k of principal components to be estimated. The procedure can be described as follows

- As in Ramsay and Silverman [15], we first center the data, i.e., we define $\widetilde{X}_{i,j} = X_{i,j} - \widehat{\mu}_i$ where $\widehat{\mu}_i$ denotes any estimator of the mean of the $i-$th population as the sample mean, for instance.
- For a fixed number of basis functions $k < p < N$, let $\widehat{\phi}_{m,p}^{(-(i,j))}$, $1 \leq m \leq k$, be the estimators of the common directions computed without the $j-$th observation of the $i-$th sample.
- Define $X_{i,j}^{\perp}(p) = \widetilde{X}_{i,j} - \pi\left(\widetilde{X}_{i,j}, \mathcal{H}_k^{(-(i,j))}(p)\right)$ where $\pi(X,\mathcal{H})$ denotes the projection of X over the closed subspace $\mathcal{H}$ and $\mathcal{H}_k^{(-(i,j))}(p)$ stands for the linear space spanned by $\widehat{\phi}_{1,p}^{(-(i,j))}, \ldots, \widehat{\phi}_{k,p}^{(-(i,j))}$. Note that in our situation, we have that $\langle \widehat{\phi}_s, \widehat{\phi}_j\rangle = \delta_{sj}$.
- Minimize the cross-validation scores $CV_k(p) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \|X_{i,j}^{\perp}(p)\|^2$.

If smoothed trajectories are considered, the minimization procedure involve both parameters p and h and it can be performed similarly to the proposals given by He, Müller and Wang [10] in functional canonical correlation analysis or by Kayano and Konishi [11] in functional principal component analysis. The advantage of this selection procedure is that is fully data driven. However, as it is well known in the one population setting, cross–validation may lead to unstable results and it is computationally expensive. Therefore, more research is needed in this direction in order to obtain more stable and faster data–driven procedures.

3.3 Estimators of the proportionality constants under a proportional model

Under a proportionality model $\mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1$, $2 \le i \le k$, $\rho_1 = 1$ and so, $\|\mathbf{\Gamma}_i\|_{\mathcal{F}}^2 = \sum_{j=1}^{\infty} \lambda_{ij}^2 = \rho_i^2 \|\mathbf{\Gamma}_1\|_{\mathcal{F}}^2$. Therefore, if $\|\mathbf{\Gamma}_1\|_{\mathcal{F}}^2 \neq 0$, we can define estimators of the proportionality constants ρ_i as

$$\widehat{\rho}_i^2 = \frac{\|\widehat{\mathbf{\Gamma}}_i\|_{\mathcal{F}}^2}{\|\widehat{\mathbf{\Gamma}}_1\|_{\mathcal{F}}^2} , \tag{3.4}$$

where $\widehat{\mathbf{\Gamma}}_i$, $1 \le i \le k$, are consistent estimators of the covariance operators, $\mathbf{\Gamma}_i$. Estimators $\{\widetilde{\phi}_j\}$ of the common eigenfunctions $\{\phi_j\}$ can be obtained as in Section 3.1 while, the eigenvalues of the first population, $\{\lambda_j\}$, can be estimated through $\widehat{\lambda}_j = \langle \widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_1 \widetilde{\phi}_j \rangle$.

Moreover, as in the finite–dimensional case, one can define a new family of estimators of λ_j. This family allows to construct estimators of the ratio λ_j/λ_1 more efficient than the previous one, for gaussian processes. Let $\widehat{\lambda}_{ij} = \langle \widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_i \widetilde{\phi}_j \rangle$ and let $\widehat{\rho}_i$ be defined as in (3.4), then, the eigenvalue estimators for the first population, $\widehat{\lambda}_j$, are defined as

$$\widehat{\lambda}_j = \sum_{i=1}^{k} \frac{\tau_i}{\widehat{\rho}_i} \widehat{\lambda}_{ij} \qquad 1 \le j \ . \tag{3.5}$$

4 Asymptotic properties of the eigenfunction and eigenvalue estimators under a FCPC model

4.1 Asymptotic distribution of Proposal 1

It is clear that consistency of each population covariance operator estimator ensures consistency of the pooled one. On the other hand, since the samples of the different populations are independent it is easy to derive the asymptotic distribution of the pooled operator estimator. We have thus the following result.

Lemma 4.1.1. *Let $\widehat{\mathbf{\Gamma}}_i$ be an estimator of the $i-th$ population covariance operator, $\widehat{\mathbf{\Gamma}} = \sum_{i=1}^{k} \tau_i \widehat{\mathbf{\Gamma}}_i$ be the pooled estimator defined in Proposal 1 and $\mathbf{\Gamma} = \sum_{i=1}^{k} \tau_i \mathbf{\Gamma}_i$.*

a) If $\widehat{\mathbf{\Gamma}}_i$ are consistent, i.e., if $\|\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i\|_{\mathcal{F}} \xrightarrow{a.s.} 0$ then $\|\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}\|_{\mathcal{F}} \xrightarrow{a.s.} 0$.

b) If $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ is zero mean gaussian random element of $\mathcal{F}$ with covariance operator $\mathbf{\Upsilon}_i$, then

$$\sqrt{N}\,(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}) \;=\; \sum_{i=1}^{k} \sqrt{\tau_i}\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \;\xrightarrow{\mathcal{D}}\; \mathbf{U}\,, \tag{4.1}$$

where $\mathbf{U} = \sum_{i=1}^{k} \sqrt{\tau_i}\mathbf{U}_i$ is a mean zero Gaussian process in $\mathcal{F}$ whose covariance operator is given by $\mathbf{\Upsilon} = \sum_{i=1}^{k} \tau_i \mathbf{\Upsilon}_i$.

Remark 4.1.1. As mentioned above, using the Central Limit Theorem in Hilbert spaces, Dauxois, Pousse and Romain [5] have shown that, when $E(\|X_{i,1}\|^4) < \infty$, for $1 \le i \le k$, if $\widehat{\mathbf{\Gamma}}_i$ are the raw empirical operators defined in (2.1), $\sqrt{n_i}\left(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i\right)$ converges in distribution to a zero mean gaussian random element, $\mathbf{U}_i$, of $\mathcal{F}$ with covariance operator $\mathbf{\Upsilon}_i$ given by

$$\mathbf{\Upsilon}_i \;=\; \sum_{m,r,o,p} s_{im}s_{ir}s_{io}s_{ip}E[f_{im}f_{ir}f_{io}f_{ip}]\,\phi_m\otimes\phi_r\tilde{\otimes}\phi_o\otimes\phi_p - \sum_{m,r}\lambda_{im}\lambda_{ir}\,\phi_m\otimes\phi_m\tilde{\otimes}\phi_r\otimes\phi_r \tag{4.2}$$

where $s_{im}^2 = \lambda_{im}$. On the other hand, as shown in Boente and Fraiman [3], if we choose $\widehat{\mathbf{\Gamma}}_i$ as the smoothed empirical operators defined in (2.2), then the same result holds under mild smoothness conditions on the covariance functions, if the bandwidth parameters for each population, h_{n_i}, satisfy that $n_i h_{n_i} \to 0$.

As proposed by Dauxois, Pousse and Romain [5], once we have consistent estimators of the covariance operator, a natural guess for estimating the eigenfunctions is to consider the corresponding eigenfunctions of the covariance estimators. Denote by $\widetilde{\phi}_j$ the eigenfunction related to the $j-$largest eigenvalue $\widehat{\nu}_j$ of $\widehat{\mathbf{\Gamma}}$. Then, from the results in Section 2.1 of Dauxois, Pousse and Romain [5], we get easily the following result.

Theorem 4.1.1. *Under the assumptions of Lemma 4.1.1b), for each eigenfunction ϕ_j of $\mathbf{\Gamma}$ related to the eigenvalue $\nu_j = \sum_{i=1}^{k} \tau_i\lambda_{ij}$ with multiplicity one, we have that*

$$\sqrt{N}(\widetilde{\phi}_j - \phi_j) \xrightarrow{\mathcal{D}} S_j\mathbf{U}\phi_j$$

where $\mathbf{U}$ is a mean zero gaussian process defined in (4.1) and S_j is the linear operator defined by

$$(S_j u)(t) \;=\; \sum_{m\neq j}\left\{\sum_{i=1}^{k}\tau_i(\lambda_{ij} - \lambda_{im})\right\}^{-1}\phi_m(t)\,\langle\phi_m, u\rangle\,.$$

Note that, from **A1** and **A2**, we get that for $j \le \ell$, ϕ_j is an eigenfunction related to an eigenvalue with multiplicity one and so, Theorem 4.1.1 can be used. In particular, we obtain the following result.

Corollary 4.1.1. *Let us assume that $\widehat{\mathbf{\Gamma}}_i$ is the raw empirical operator,$\widehat{\mathbf{\Gamma}}_{i,\mathrm{R}}$, defined in (2.1), that $E(\|X_{i,1}\|^4) < \infty$, for $1 \le i \le k$, and that **A1** and **A2** hold. For each eigenfunction ϕ_j of $\mathbf{\Gamma}$ related to the eigenvalue $\nu_j = \sum_{i=1}^{k} \tau_i\lambda_{ij}$ with multiplicity one, we have that*

a) $\sqrt{N}(\widetilde{\phi}_j - \phi_j, \phi_j) \xrightarrow{p} 0$

b) *For any* $j \neq m$ $\sqrt{N}\langle\widetilde{\phi}_j - \phi_j, \phi_m\rangle \;\rightarrow\; \mathcal{N}(0, \sigma^2_{mj})$ *with*

$$\sigma^2_{jm} \;=\; \left\{\sum_{i=1}^{k} \tau_i(\lambda_{ij} - \lambda_{im})\right\}^{-2} \sum_{i=1}^{k} \tau_i \lambda_{im}\lambda_{ij} E[f^2_{im} f^2_{ij}]$$

Moreover, if $X_{i,1}$ *are gaussian processes, for all* $1 \le i \le k$*, we get that*

$$\sigma^2_{jm} \;=\; \left\{\sum_{i=1}^{k} \tau_i(\lambda_{ij} - \lambda_{im})\right\}^{-2} \sum_{i=1}^{k} \tau_i \lambda_{im}\lambda_{ij}\;. \tag{4.3}$$

The following Theorem provides the asymptotic behavior of the eigenvalue estimators under mild conditions on the eigenfunction estimators.

Theorem 4.1.2. *Let* $\widehat{\mathbf{\Gamma}}_i$ *be an estimator of the covariance operator of the* $i-th$ *population such that* $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$*, where* $\mathbf{U}_i$ *is zero mean gaussian random element of* $\mathcal{F}$ *with covariance operator* $\mathbf{\Upsilon}_i$*. Let* $\widetilde{\phi}_j$ *be consistent estimators of the common eigenfunctions such that* $\sqrt{N}\left(\widetilde{\phi}_j - \phi_j\right) = O_p(1)$ *and define estimators of* λ_{ij} *as* $\widehat{\lambda}_{ij} = \langle\widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_i\widetilde{\phi}_j\rangle$*. For any fixed* m*, denote* $\widehat{\mathbf{\Lambda}}_i^{(m)} = \left\{\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right)\right\}_{1\le j\le m}$*. Then,*

a) *For each* $1 \le i \le k$*,* $\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right)$ *has the same asymptotic distribution as* $\sqrt{n_i}\left(\langle\phi_j, \widehat{\mathbf{\Gamma}}_i\phi_j\rangle - \lambda_{ij}\right)$.

b) *For any* m *fixed,* $\widehat{\mathbf{\Lambda}}_1^{(m)}, \ldots, \widehat{\mathbf{\Lambda}}_k^{(m)}$ *are asymptotically independent.*

c) *If, in addition, the covariance operator* $\mathbf{\Upsilon}_i$ *of* $\mathbf{U}_i$ *is given by (4.2), then* $\widehat{\mathbf{\Lambda}}_i^{(m)}$ *is jointly asymptotically normally distributed with zero mean and covariance matrix* $\mathbf{C}^{(i,m)}$ *such that* $\mathbf{C}^{(i,m)}_{jj} = \lambda^2_{ij}\left[E\left(f^4_{ij}\right) - 1\right]$ *and* $\mathbf{C}^{(i,m)}_{js} = \lambda_{ij}\lambda_{is}\left[E\left(f^2_{ij}f^2_{is}\right) - 1\right]$*, that is, the asymptotic variance of* $\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right)$ *is given by* $\lambda^2_{ij}\left[E\left(f^4_{ij}\right) - 1\right]$ *and the asymptotic correlations are given by*

$$\frac{E\left(f^2_{ij}f^2_{is}\right) - 1}{\left[E\left(f^4_{ij}\right) - 1\right]^{\frac{1}{2}}\left[E\left(f^4_{is}\right) - 1\right]^{\frac{1}{2}}}\;.$$

Moreover, in the normal case, we get that the components of $\widehat{\mathbf{\Lambda}}_i^{(m)}$ *are asymptotically independent with asymptotic variances* $2\lambda^2_{ij}$.

Remark 4.1.2. When all the populations have a gaussian distribution, Theorems 4.1.1 and 4.1.2 provide an expression for the asymptotic variance of the estimators that is related to that given in the finite–dimensional setting. A more general framework in which an analogous statement can be given, is when all the populations have the same distribution except for changes in the location parameter and the covariance operators. To be more precise, assume that the processes $X_{i,1}(t)$, $1 \le i \le k$ can be written as

$$X_{i,1} = \mu_i + \sum_{j=1}^{\infty} \lambda_{ij}^{\frac{1}{2}} f_{ij}\, \phi_j$$

with $\lambda_{i1} \ge \lambda_{i2} \ge \ldots \ge 0$ and f_{ij} random variables such that $E(f_{ij}) = 0$, $E(f_{ij}^2) = 1$, $E(f_{ij}\ f_{is}) = 0$ for $j \ne s$. Moreover, assume that for any j, f_{ij} and f_{1j} have the same distribution. In this case, the asymptotic variance of $\sqrt{N}\langle \widehat{\phi}_j - \phi_j, \phi_m \rangle$ reduces to

$$\sigma_{jm}^2 \ = \ E\left(f_{1m}^2 f_{1j}^2\right) \left\{\sum_{i=1}^{k} \tau_i(\lambda_{ij} - \lambda_{im})\right\}^{-2} \sum_{i=1}^{k} \tau_i \lambda_{im}\lambda_{ij} \ ,$$

that is, it is proportional to that obtained for gaussian processes.

Similarly, the asymptotic variance of $\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right)$ is given by $\lambda_{ij}^2 \left[E\left(f_{1j}^4\right) - 1\right]$ and the asymptotic correlations are the same for all populations.

It is worth noticing, that under the additional assumption of gaussian process $E(f_{ij}^4) = 3$ and so there is no need to estimate this quantity if we are seeking for confidence intervals or hypothesis testing. Under a more general framework, the expectation appearing in the previous expressions for the asymptotic variances, i.e., $E(f_{ij}^4)$ and $E\left(f_{1m}^2 f_{1j}^2\right)$ can easily be estimated using consistent estimators of the eigenvalues and eigenfunctions and the fact that $\langle X_{i,1}, \phi_j \rangle^2 = \lambda_{ij} f_{ij}^2$.

4.2 Asymptotic Properties of Proposal 2

In this section we will study the asymptotic behavior of the second proposal given in Section 3.1. We will show that a better efficiency is attained if $X_{i,1}(t)$ are gaussian random functions.

Denote $\mathbf{\Gamma}_{mj} = \sum_{i=1}^{k} \tau_i \left[(\lambda_{ij} - \lambda_{im}) / (\lambda_{im}\lambda_{ij})\right] \mathbf{\Gamma}_i$ and denote $\phi_j^\star$ any solution of

$$\begin{cases} \delta_{mj} = & \langle \phi_m^\star, \phi_j^\star \rangle \\ 0 = & \langle \phi_m^\star, \mathbf{\Gamma}_{mj} \phi_j^\star \rangle \qquad 1 \le j < m \ . \end{cases} \tag{4.4}$$

It is easy to see that if the covariance operators satisfy a FCPC model, then ϕ_j satisfies (4.4). The following result state the consistency of the estimators defined through (3.3).

Theorem 4.2.1. *Let $\widehat{\mathbf{\Gamma}}_i$ be consistent estimators of the covariance operator of the $i-th$ population, i.e., $\|\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i\|_{\mathcal{F}} \xrightarrow{a.s.} 0$. Moreover, assume that the FCPC model hold and let $\{\widehat{\lambda}_{ij}\}$ consistent estimators of the eigenvalues of the $i-th$ population $\{\lambda_{ij}\}$. Assume that for each j there exists $1 \le i_j \le k$ such that $\lambda_{i_j j} > 0$ and that the system (4.4) has a unique solution, then the solution $\widehat{\phi}_j$ of (3.3) provide consistent estimators of the common eigenfunctions ϕ_j.*

The following result states the asymptotic behavior of the coordinates $\{\langle\widehat{\phi}_j, \phi_s\rangle : s \geq 1\}$ of the common eigenfunctions estimators $\widehat{\phi}_j$ defined through Proposal 2 that will allow to establish the claimed improvement in efficiency for gaussian processes.

Theorem 4.2.2. *Let $\widehat{\mathbf{\Gamma}}_i$ be an estimator of the covariance operator of the $i-$th population such that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ is zero mean gaussian random element of $\mathcal{F}$ with covariance operator $\mathbf{\Upsilon}_i$ given by (4.2). Moreover, let $\widehat{\lambda}_{ij}$ be consistent estimators of the eigenvalues of the $i-$th population λ_{ij}. Assume **A1**, **A2** and that $\lambda_{i\ell} > 0$, for all $1 \leq i \leq k$. If the solution $\widehat{\phi}_j$ of (3.3) are consistent estimators of the common eigenfunctions ϕ_j such that either $\widehat{g}_j = \sqrt{N}\left(\widehat{\phi}_j - \phi_j\right) = O_p(1)$ or $N^{\frac{1}{4}}\left(\widehat{\phi}_j - \phi_j\right) = o_p(1)$ hold, then, for any $j \leq \ell$, $m \leq \ell$, $m \neq j$ we have that*

a) $\langle\widehat{g}_m, \phi_j\rangle$ has the same asymptotic distribution as $-\langle\widehat{g}_j, \phi_m\rangle$.

b) For $j < m$, $\langle\widehat{g}_j, \phi_m\rangle \xrightarrow{\mathcal{D}} \mathcal{N}(0, \theta^2_{jm})$, where

$$\theta^2_{jm} = \frac{\displaystyle\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}} E\left(f^2_{im} f^2_{ij}\right)}{\left\{\displaystyle\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}}\right\}^2} \tag{4.5}$$

Remark 4.2.1. Note that in the gaussian case, we get $E\left(f^2_{im} f^2_{ij}\right) = 1$ and so the asymptotic variance of coordinates of the common eigenfunction estimates, defined through Proposal 2, reduces to

$$\theta^2_{jm} = \left\{\sum_{i=1}^{k} \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}}\right\}^{-1}$$

On the other hand, the common eigenfunction estimates, defined through Proposal 1, have asymptotic variances σ^2_{jm} given by (4.3). Since $\theta^2_{jm} \leq \sigma^2_{jm}$, we obtain that the estimates of Proposal 2 are more efficient that those of Proposal 1 for gaussian processes.

Note that if we relax the gaussian distribution assumption by requiring, as in Remark 4.1.2, that f_{ij} and f_{1j} have the same distribution, for all j, then, the same conclusion holds.

4.3 Asymptotic Distribution of the proportionality constants

We will first state some results regarding the norm of a covariance operator estimator that will allow to derive easily the asymptotic behavior of the proportionality constant estimators defined in Section 3.3. Strong consistency follows easily from the continuity of the norm $\|\cdot\|_{\mathcal{F}}$ and the consistency of the covariance estimators of each population.

The following Theorem states the asymptotic distribution of the proportionality constants.

Theorem 4.3.1. *Let $X_{i,1}(t), \cdots, X_{i,n_i}(t)$ be independent observations from k independent samples of smooth random functions in $L^2[0,1]$ with gaussian distribution with mean μ_i and covariance operators $\mathbf{\Gamma}_i$ such that $\mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1$, $1 \le i \le k$, $\rho_1 = 1$. Let $\widehat{\mathbf{\Gamma}}_i$ be estimators of the covariance operators $\mathbf{\Gamma}_i$ such that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ are independent zero mean gaussian random elements of $\mathcal{F}$ with covariance operators $\mathbf{\Upsilon}_i$ given by (4.2). Let $\widehat{\rho}_i$ be defined as in (3.4). Then, if $\|\mathbf{\Gamma}_1\|_{\mathcal{F}} \neq 0$ and we denote by $\widehat{r}_i = \sqrt{N}\,(\widehat{\rho}_i - \rho_i)$, $\widehat{\mathbf{r}} = (\widehat{r}_2, \ldots, \widehat{r}_k)^{\mathrm{T}}$, we have that $\widehat{\mathbf{r}}$ is asymptotically normally distributed with zero mean and asymptotic variances given by*

$$\text{ASVAR}\,(\widehat{r}_i) \;=\; 2\,\rho_i^2\,(\tau_1 + \tau_i)\,\frac{\sum_{j\ge 1}\lambda_j^4}{\|\mathbf{\Gamma}_1\|_{\mathcal{F}}^4} \qquad 2 \le i \le k\,. \tag{4.6}$$

Moreover, if we denote by $\boldsymbol{\rho} = (\rho_2, \ldots, \rho_k)^{\mathrm{T}}$, we have that $\widehat{\mathbf{r}} \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}_{k-1}, \mathbf{B})$ where

$$\mathbf{B} = 2\frac{\sum_{j\ge 1}\lambda_j^4}{\|\mathbf{\Gamma}_1\|_{\mathcal{F}}^4}\left[\frac{1}{\tau_1}\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{T}} + \mathit{diag}\left(\frac{\rho_2^2}{\tau_2}, \ldots, \frac{\rho_k^2}{\tau_k}\right)\right]$$

It is worth noticing that $\sum_{j\ge 1}\lambda_j^4 = \|\mathbf{\Gamma}_1\,\mathbf{\Gamma}_1\|_{\mathcal{F}}^2$.

The following result gives the asymptotic distribution of the estimators of the ratio λ_j/λ_1.

Theorem 4.3.2. *Let $X_{i,1}(t), \cdots, X_{i,n_i}(t)$ be independent observations from k independent samples of smooth random functions in $L^2[0,1]$ with mean μ_i and covariance operators $\mathbf{\Gamma}_i$ such that $\mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1$, $1 \le i \le k$, $\rho_1 = 1$. Let $\widehat{\mathbf{\Gamma}}_i$ be estimators of the covariance operators $\mathbf{\Gamma}_i$ such that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ are independent zero mean gaussian random elements of $\mathcal{F}$ with covariance operators $\mathbf{\Upsilon}_i$ given by (4.2). Let $\widehat{\rho}_i$ consistent estimators of the proportionality constants ρ_i and $\widehat{\lambda}_j$ be defined as in (3.5) where $\widehat{\lambda}_{ij} = \langle \widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_i \widetilde{\phi}_j \rangle$ with $\sqrt{N}\left(\widetilde{\phi}_j - \phi_j\right) = O_p(1)$. Denote by $\widehat{\psi}_j = \sqrt{N}\left(\widehat{\lambda}_j/\widehat{\lambda}_1 - \lambda_j/\lambda_1\right)$ and for any fixed $p \ge 2$, let $\widehat{\boldsymbol{\psi}} = \left(\widehat{\psi}_2, \ldots, \widehat{\psi}_p\right)^{\mathrm{T}}$. Then, we have that $\widehat{\boldsymbol{\psi}}$ is asymptotically normally distributed with zero mean and asymptotic variances given by*

$$\begin{aligned}
\text{ASVAR}\left(\widehat{\psi}_j\right) &= \frac{\lambda_j^2}{\lambda_1^2}\sum_{i=1}^{k}\tau_i \text{VAR}\left(f_{ij}^2 - f_{i1}^2\right) \qquad\qquad 2 \le j \le p \\
\text{ASCOV}\left(\widehat{\psi}_j, \widehat{\psi}_m\right) &= \frac{\lambda_j\lambda_m}{\lambda_1^2}\sum_{i=1}^{k}\tau_i\left[E\left(f_{ij}^2 f_{im}^2\right) - E\left(f_{i1}^2 f_{im}^2\right) - E\left(f_{ij}^2 f_{i1}^2\right) + E\left(f_{i1}^4\right)\right]
\end{aligned}$$

for $2 \le j < m \le p$.

Remark 4.3.1 Note that if the process is gaussian, the asymptotic variance of $\widehat{\psi}_j = \sqrt{N}\left(\widehat{\lambda}_j/\widehat{\lambda}_1 - \lambda_j/\lambda_1\right)$ reduces to $\sigma_j^2 = 4\lambda_j^2/\lambda_1^2$ while the correlations are $1/2$ as in the finite–dimensional case. Moreover, these ratio estimators are more efficient than those obtained by considering as eigenvalue estimators $\langle \widetilde{\phi}_j, \widehat{\mathbf{\Gamma}}_1 \widetilde{\phi}_j \rangle$.

Theorem 4.3.1 can be used to test the hypothesis of equality of several covariance operators against proportionality. This corresponds to the two first levels of similarity considered in the finite–dimensional setting by Flury [9]. Effectively, assume that we want to test

$$H_0 : \mathbf{\Gamma}_1 = \mathbf{\Gamma}_2 = \ldots = \mathbf{\Gamma}_k \quad \text{against} \quad H_1 : \mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1 \,, 2 \le i \le k \ \text{ and } \exists\, i : \ \rho_i \neq 1 \,. \tag{4.7}$$

The estimators defined in Section 3.3 allow to construct a Wald statistic.

From now on, let $\boldsymbol{\gamma}_{\boldsymbol{\rho}} = \left(\frac{\tau_2}{\rho_2}, \ldots, \frac{\tau_k}{\rho_k}\right)^{\mathrm{T}}$ and denote

$$\mathbf{C}_{\boldsymbol{\rho}} = \left[\operatorname{diag}\left(\frac{\tau_2}{\rho_2^2}, \ldots, \frac{\tau_k}{\rho_k^2}\right) - \boldsymbol{\gamma}_{\boldsymbol{\rho}} \boldsymbol{\gamma}_{\boldsymbol{\rho}}^{\mathrm{T}} \right] .$$

The following result provides a test for (4.7).

Theorem 4.3.3. *Let $X_{i,1}(t), \cdots, X_{i,n_i}(t)$ be independent observations from k independent samples of smooth random functions in $L^2[0,1]$ with gaussian distribution with mean μ_i and covariance operators $\mathbf{\Gamma}_i$ such that $\mathbf{\Gamma}_i = \rho_i \mathbf{\Gamma}_1$, $1 \le i \le k$, $\rho_1 = 1$. Assume that we want to test (4.7) and that $\|\mathbf{\Gamma}_1\|_{\mathcal{F}} \neq 0$.*
Let $\widehat{\mathbf{\Gamma}}_i$ be estimators of the covariance operators $\mathbf{\Gamma}_i$ such that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, where $\mathbf{U}_i$ are independent zero mean gaussian random elements of $\mathcal{F}$ with covariance operators $\mathbf{\Upsilon}_i$ given by (4.2). Let $\widehat{\rho}_i$ be defined as in (3.4), $\widehat{\boldsymbol{\rho}} = (\widehat{\rho}_2, \ldots, \widehat{\rho}_k)^{\mathrm{T}}$ and $\mathcal{T}$ be defined as

$$\mathcal{T} = \sqrt{N}\, (\widehat{\boldsymbol{\rho}} - \mathbf{1}_{k-1})^{\mathrm{T}}\, \widehat{\mathbf{C}}\, (\widehat{\boldsymbol{\rho}} - \mathbf{1}_{k-1})\, \frac{\|\widehat{\mathbf{\Gamma}}_1\|_{\mathcal{F}}^4}{2\|\widehat{\mathbf{\Gamma}}_1 \widehat{\mathbf{\Gamma}}_1\|_{\mathcal{F}}^2} \,,$$

where $\widehat{\mathbf{C}} = \mathbf{C}_{\widehat{\boldsymbol{\rho}}}$ and $\mathbf{1}_{k-1}$ is the $k-th$ dimensional vector with all its components equal to 1. Then,

a) Under H_0, $\mathcal{T} \xrightarrow{\mathcal{D}} \chi_{k-1}^2$.

b) Under $H_{1,\mathbf{a}} : \mathbf{\Gamma}_i = \left(1 + a_i N^{-\frac{1}{2}}\right) \mathbf{\Gamma}_1$, we have that $\mathcal{T} \xrightarrow{\mathcal{D}} \chi_{k-1}^2(\theta)$ with

$$\theta = \mathbf{a}^{\mathrm{T}} \mathbf{C}_{\mathbf{1}_{k-1}} \mathbf{a}\, \frac{\|\mathbf{\Gamma}_1\|_{\mathcal{F}}^4}{2\|\mathbf{\Gamma}_1 \mathbf{\Gamma}_1\|_{\mathcal{F}}^2} = \left[\sum_{i=2}^{p} \tau_i\, a_i^2 - \left(\sum_{i=2}^{p} \tau_i\, a_i \right)^2 \right] \frac{\|\mathbf{\Gamma}_1\|_{\mathcal{F}}^4}{2\|\mathbf{\Gamma}_1 \mathbf{\Gamma}_1\|_{\mathcal{F}}^2} \,.$$

Therefore, a test, with asymptotic level α, rejects the null hypothesis when

$$\mathcal{T} > \chi_{k-1,1-\alpha}^2 \,,$$

with $P\left(\chi^2_{k-1} > \chi^2_{k-1,1-\alpha}\right) = \alpha$.

If the covariance operators are proportional the above testing procedure allows to decide if equality holds. If it does not, a modified discriminating rule using estimators of the proportional constants needs to be considered.

Acknowledgments

This research was partially supported by Grants X-018 from the Universidad de Buenos Aires, PID 5505 from CONICET and PICT 21407 from ANPCYT, Argentina.

Appendix

Proof of Corollary 4.1.1. The proof follows easily from the fact that the covariance operator of $V_j = S_j \mathbf{U}\phi_j$ is $\Sigma_{V_j}(s,t) = \sum_{r,m\neq j} c_m\, c_{mr}\, c_r\, \phi_m(s)\phi_r(t)$, where

$$c_m = \left\{\sum_{i=1}^{k} \tau_i \left(\lambda_{ij} - \lambda_{im}\right)\right\}^{-1} \quad \text{and} \quad c_{mr} = \sum_{i=1}^{k} \tau_i \lambda_{im}^{\frac{1}{2}} \lambda_{ir}^{\frac{1}{2}} \lambda_{ij}\, E\left(f_{im} f_{ir} f_{ij}^2\right) .\square$$

Proof of Theorem 4.1.2. The proof of b) follows easily from a). The proof of a) can be derived using analogous arguments to those considered in Flury [8] for the maximum likelihood estimates. Effectively, the consistency of $\widetilde{\phi}_j$ entails that $\sqrt{n_i}\langle\widetilde{\phi}_j, (\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i)\widetilde{\phi}_j\rangle$ and $\sqrt{n_i}\langle\phi_j, (\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i)\phi_j\rangle$ have the same the asymptotic distribution and so, the proof will be completed if we show that $\sqrt{n_i}\left[\langle\widetilde{\phi}_j, \mathbf{\Gamma}_i\widetilde{\phi}_j\rangle - \langle\phi_j, \mathbf{\Gamma}_i\phi_j\rangle\right] \xrightarrow{p} 0$.

Since $\langle\widetilde{\phi}_j, \widetilde{\phi}_j\rangle = 1$ and $\sqrt{n_i}(\widetilde{\phi}_j - \phi_j)$ is bounded in probability, using that $\langle\phi_j, \widetilde{\phi}_j - \phi_j\rangle = -(1/2)\langle\widetilde{\phi}_j - \phi_j, \widetilde{\phi}_j - \phi_j\rangle$, we get easily that $\sqrt{n_i}\langle\phi_j, \widetilde{\phi}_j - \phi_j\rangle \xrightarrow{p} 0$. On the other hand, we have that

$$\sqrt{n_i}\left(\langle\widetilde{\phi}_j, \mathbf{\Gamma}_i\widetilde{\phi}_j\rangle - \langle\phi_j, \mathbf{\Gamma}_i\phi_j\rangle\right) = U_{1n_i} + U_{2n_i} + U_{3n_i}$$

where

$$\begin{aligned} U_{1n_i} &= \sqrt{n_i}\langle\widetilde{\phi}_j - \phi_j, \mathbf{\Gamma}_i\left(\widetilde{\phi}_j - \phi_j\right)\rangle \\ U_{2n_i} &= \sqrt{n_i}\langle\widetilde{\phi}_j - \phi_j, \mathbf{\Gamma}_i\phi_j\rangle \\ U_{3n_i} &= \sqrt{n_i}\langle\phi_j, \mathbf{\Gamma}_i\left(\widetilde{\phi}_j - \phi_j\right)\rangle. \end{aligned}$$

Using that $\sqrt{N}\left(\widetilde{\phi}_j - \phi_j\right) = O_p(1)$, we obtain that $U_{1n_i} \xrightarrow{p} 0$. Besides, $U_{2n_i} + U_{3n_i} = 2\lambda_{ij}\sqrt{n_i}\langle\phi_j, \widetilde{\phi}_j - \phi_j\rangle \xrightarrow{p} 0$ concluding the proof of a).

c) As in a) we have that $\{\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right)\}_{1\leq j\leq p}$ has the same asymptotic distribution as $\left\{\langle\phi_j, \sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i)\phi_j\rangle\right\}_{1\leq j\leq p}$. Using that $\sqrt{n_i}(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i) \xrightarrow{\mathcal{D}} \mathbf{U}_i$, we get that $\widehat{\mathbf{\Lambda}}_i^{(p)} \xrightarrow{\mathcal{D}}$

$(\langle \phi_1, \mathbf{U}_i \phi_1 \rangle, \cdots, \langle \phi_p, \mathbf{U}_i \phi_p \rangle)$ which is a zero mean normally distributed random vector. The expression for its covariance matrix, follows easily using (4.2). □

Proof of Theorem 4.2.1. From the consistency of $\widehat{\mathbf{\Gamma}}_i$ and $\widehat{\lambda}_{ij}$, we get that $\widehat{\mathbf{\Gamma}}_{mj} \xrightarrow{a.s.} \mathbf{\Gamma}_{mj}$. Therefore, the solution $\{\widehat{\phi}_j\}_{j\geq 1}$ of the system (3.3) will converge to a solution $\phi_j^\star$ of the system (4.4). The assumed uniqueness, entails Fisher–consistency and thus, consistency. □

Proof of Theorem 4.2.2. Denote by $\widehat{\mathbf{Z}}_i = \sqrt{N}\left(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i\right)$. Then, we have that $\widehat{\mathbf{\Gamma}}_i = N^{-\frac{1}{2}}\widehat{\mathbf{Z}}_i + \mathbf{\Gamma}_i$ and $\widehat{\phi}_j = N^{-\frac{1}{2}}\widehat{g}_j + \phi_j$. Replacing in the first equation of (3.3), we get that $N^{1/2}\langle \widehat{\phi}_j - \phi_j, \widehat{\phi}_m - \phi_m \rangle + \langle \widehat{g}_j, \phi_m \rangle + \langle \widehat{g}_m, \phi_j \rangle = 0$, for all m, j. On the other hand, replacing in equation (3.3), we get that for $j \neq m$

$$\begin{aligned} \langle \widehat{g}_j, \phi_m \rangle + \langle \widehat{g}_m, \phi_j \rangle &= -\widehat{c}_{mj} \ , \\ \widehat{a}_{mj}\langle \widehat{g}_m, \phi_j \rangle + \widehat{b}_{mj}\langle \widehat{g}_j, \phi_m \rangle &= -\widetilde{u}_{mj} - \widehat{R}_{mj} \ , \end{aligned} \tag{A.1}$$

where

$$\begin{aligned} \widehat{c}_{mj} &= N^{1/2}\langle \widehat{\phi}_j - \phi_j, \widehat{\phi}_m - \phi_m \rangle \ , \quad \widehat{a}_{mj} = \sum_{i=1}^k \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}} \lambda_{ij} \\ \widehat{b}_{mj} &= \sum_{i=1}^k \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}} \lambda_{im} \ , \quad \widetilde{u}_{mj} = \sum_{i=1}^k \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}} \langle \widehat{\phi}_m, \widehat{\mathbf{Z}}_i \widehat{\phi}_j \rangle \quad \text{for } j < m \\ \widehat{R}_{mj} &= \sum_{i=1}^k \tau_i \frac{\widehat{\lambda}_{ij} - \widehat{\lambda}_{im}}{\widehat{\lambda}_{im}\widehat{\lambda}_{ij}} \langle \widehat{g}_m, \mathbf{\Gamma}_i(\widehat{\phi}_j - \phi_j) \rangle \ . \end{aligned}$$

Let us restrict the system of equations (A.1) only to those indexes with $1 \leq j < m \leq \ell$. Therefore, it can be written as the linear system $\widehat{\mathbf{B}}_\ell \ \widehat{\mathbf{G}} = \widehat{\mathbf{W}}$, where $\widehat{\mathbf{B}}_\ell$ is a matrix and $\widehat{\mathbf{G}} = \langle \widehat{g}_j, \phi_m \rangle_{1 \leq j \neq m \leq \ell}$.

Since $\widehat{g}_j = O_p(1)$ or $N^{\frac{1}{4}}\left(\widehat{\phi}_j - \phi_j\right) = o_p(1)$, we have that $\widehat{c}_{mj} \xrightarrow{p} 0$ and $\widehat{R}_{mj} \xrightarrow{p} 0$ which entails a). Moreover, the weak consistency of the eigenvalue estimators guarantees that $\widehat{a}_{mj} \xrightarrow{p} a_{mj}$ and $\widehat{b}_{mj} \xrightarrow{p} b_{mj}$, where

$$a_{mj} = \sum_{i=1}^k \tau_i \frac{\lambda_{ij} - \lambda_{im}}{\lambda_{im}} \qquad \text{and} \qquad b_{mj} = \sum_{i=1}^k \tau_i \frac{\lambda_{ij} - \lambda_{im}}{\lambda_{ij}}$$

and hence, convergence in probability of the matrix $\widehat{\mathbf{B}}_\ell$ to a matrix $\mathbf{B}_\ell$. Furthermore, the assumptions made on the eigenvalues λ_{ij} guarantee that $\mathbf{B}_\ell$ is non singular.

Using that $\widehat{\mathbf{Z}}_i \xrightarrow{\mathcal{D}} \mathbf{U}_i$ for each i, we get that $(\widetilde{u}_{mj})_{1 \leq j < m \leq \ell}$ is asymptotically normally distributed, i.e., $(\widetilde{u}_{mj})_{1 \leq j < m \leq \ell} \xrightarrow{\mathcal{D}} \mathbf{u}$ where $\mathbf{u}$ is a gaussian vector with zero mean, zero correlations and the variance of $u_{m,j}$ is given by

$$\sum_{i=1}^k \tau_i \frac{(\lambda_{im} - \lambda_{ij})^2}{\lambda_{im}\lambda_{ij}} E\left(f_{im}^2 f_{ij}^2\right) \ .$$

Therefore, $\widehat{\mathbf{W}} \xrightarrow{\mathcal{D}} \mathbf{W}$, a random vector with its first $s(s+1)$ rows equal to 0 and the last ones equal to $\mathbf{u}$ implying that $\widehat{\mathbf{G}} \xrightarrow{\mathcal{D}} \mathbf{B}_p^{-1}\mathbf{W}$ and so $(\widehat{g}_j, \phi_m) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \theta_{jm}^2)$, with θ_{jm}^2 defined in (4.5). □

In order to prove Theorem 4.3.1, we will need the following Lemma.

Lemma 1. *Let $X_1 \dots X_n$ be independent random elements $L^2([0,1])$ with covariance operator $\mathbf{\Gamma}$. Denote by λ_j the eigenvalues of $\mathbf{\Gamma}$. Let $\widehat{\mathbf{\Gamma}}$ be an estimator of the covariance operator $\mathbf{\Gamma}$ such that $\sqrt{n}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}) \xrightarrow{\mathcal{D}} \mathbf{U}$, where $\mathbf{U}$ is a zero mean gaussian random element of $\mathcal{F}$ with covariance operator $\mathbf{\Upsilon}$, given by*

$$\mathbf{\Upsilon} = \sum_{m,r,o,p} s_m s_r s_o s_p E\left(f_m f_r f_o f_p\right) \phi_m \otimes \phi_r \tilde{\otimes} \phi_o \otimes \phi_p - \sum_{m,r} \lambda_m \lambda_r \, \phi_m \otimes \phi_m \tilde{\otimes} \phi_r \otimes \phi_r \quad \text{(A.2)}$$

and $s_m^2 = \lambda_m$. Then, we have that $\sqrt{n}(\|\widehat{\mathbf{\Gamma}}\|_{\mathcal{F}}^2 - \|\mathbf{\Gamma}\|_{\mathcal{F}}^2) \xrightarrow{\mathcal{D}} 2(\mathbf{U}, \mathbf{\Gamma})_{\mathcal{F}} = 2Y$, where Y is a zero mean normal random variable with variance given by

$$\sigma_Y^2 = \sum_{m,p\geq 1} \lambda_m^2 \lambda_p^2 E\left(f_m^2 f_p^2\right) - \sum_{m,p\geq 1} \lambda_m^2 \lambda_p^2 \,.$$

Moreover, in the normal case, we get that $\sigma_Y^2 = 2 \sum_{j\geq 1} \lambda_j^4$.

Proof. Denote by $W_n = \sqrt{n}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma})$. Then, we have that $\sqrt{n}\left((\widehat{\mathbf{\Gamma}}, \widehat{\mathbf{\Gamma}})_{\mathcal{F}} - (\mathbf{\Gamma}, \mathbf{\Gamma})_{\mathcal{F}}\right) = 2(W_n, \mathbf{\Gamma})_{\mathcal{F}} + (W_n, \widehat{\mathbf{\Gamma}} - \mathbf{\Gamma})_{\mathcal{F}} = 2(W_n, \mathbf{\Gamma})_{\mathcal{F}} + o_p(1)$, since $W_n = O_p(1)$ and $\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma} = o_p(1)$. Finally, since $W_n \xrightarrow{\mathcal{D}} \mathbf{U}$ where $\mathbf{U}$ is zero mean gaussian random element of $\mathcal{F}$ with covariance operator $\mathbf{\Upsilon}$, given by (A.2), we get that $(W_n, \mathbf{\Gamma})_{\mathcal{F}} \xrightarrow{\mathcal{D}} (\mathbf{U}, \mathbf{\Gamma})_{\mathcal{F}}$ which is a mean zero normal random variable. From the explicit formula for $\mathbf{\Upsilon}$ given in (A.2), straightforward calculations lead to the expression for the variance of $(\mathbf{U}, \mathbf{\Gamma})_{\mathcal{F}}$. □

Proof of Theorem 4.3.1. Denote by $\widehat{T}_i = \sqrt{N}(\|\widehat{\mathbf{\Gamma}}_i\|_{\mathcal{F}}^2 - \|\mathbf{\Gamma}_i\|_{\mathcal{F}}^2)$. Then, using Lemma 1, we get that $\widehat{T}_i \xrightarrow{\mathcal{D}} T_i \sim \mathcal{N}(0, \sigma_i^2)$ where $\sigma_i^2 = 8\rho_i^4 \sum_{j\geq 1} \lambda_j^4 / \tau_i$ and $T_1, \dots, T_k$ are independent. It is easy to see that, for $2 \leq i \leq k$,

$$\widehat{r}_i = \frac{1}{2\,\rho_i\, \|\mathbf{\Gamma}_1\|_{\mathcal{F}}^2} \left(\widehat{T}_i - \rho_i^2 \widehat{T}_1\right) + o_p(1)$$

and so the asymptotic distribution of $\widehat{\mathbf{r}}$ is the distribution of $(Z_2, \dots, Z_k)^{\mathrm{T}}$ with $Z_i = \left(T_i - \rho_i^2 T_1\right) / \left(2\,\rho_i\, \|\mathbf{\Gamma}_1\|_{\mathcal{F}}^2\right)$. The proof follows now easily from the fact that T_i are independent and normally distributed. □

Proof of Theorem 4.3.2. The consistency of the eigenvalue estimators entails that it will be enough to obtain the asymptotic distribution of $\sqrt{N}\left(\widehat{\lambda}_j \lambda_1 - \lambda_j \widehat{\lambda}_1\right)$.

From the proof of Theorem 4.1.2, we have that $\sqrt{n_i}\left(\widehat{\lambda}_{ij} - \lambda_{ij}\right) = \widehat{t}_{ij} + r_{ij}$, where $r_{ij} \xrightarrow{p} 0$ and $\widehat{t}_{ij} = \langle \phi_j, \sqrt{n_i}\left(\widehat{\mathbf{\Gamma}}_i - \mathbf{\Gamma}_i\right) \phi_j \rangle$ is asymptotically normally distributed. Therefore, using that $\widehat{\rho}_i \xrightarrow{p} \rho_i$ and that $\widehat{t}_{ij}$ are bounded in probability, we have the following expansion

$$\widehat{\lambda}_j = \lambda_j \sum_{i=1}^{k} \frac{\tau_i \, \rho_i}{\widehat{\rho}_i} + \frac{1}{\sqrt{N}} \sum_{i=1}^{k} \frac{\sqrt{\tau_i}}{\rho_i} \widehat{t}_{ij} + \frac{1}{\sqrt{N}} R_j$$

with $R_j \xrightarrow{p} 0$. Hence,

$$\sqrt{N} \left(\widehat{\lambda}_j \lambda_1 - \lambda_j \widehat{\lambda}_1 \right) \quad = \quad \sum_{i=1}^{k} \frac{\sqrt{\tau_i}}{\rho_i} \left(\lambda_1 \widehat{t}_{ij} - \lambda_j \widehat{t}_{i1} \right) + o_p(1)$$

Now, the result follows easily using that $\widehat{\mathbf{t}}_i = \left(\widehat{t}_{i1}, \ldots, \widehat{t}_{ip} \right)$ are asymptotically independent and asymptotically normally distributed with variances and covariances given in Theorem 4.1.2 c). □

Proof of Theorem 4.3.3. The proof of a) follows immediately from Theorem 4.3.1. As in Theorem 4.3.1 denote by $\widehat{r}_i = \sqrt{N} \left(\widehat{\rho}_i - \mathbf{1}_{k-1} \right)$ and $\widehat{\mathbf{r}} = \left(\widehat{r}_2, \ldots, \widehat{r}_k \right)^{\mathrm{T}}$. The proof of b) follows immediately from the fact that, when $\boldsymbol{\Gamma}_i = \boldsymbol{\Gamma}_{i,N} = \left(1 + a_i N^{-\frac{1}{2}} \right) \boldsymbol{\Gamma}_1$, $\widehat{\mathbf{C}} \xrightarrow{p} \mathbf{C}_{\mathbf{1}_{k-1}}$ and $\widehat{\mathbf{r}} \xrightarrow{\mathcal{D}} N(\mathbf{a}, \mathbf{D})$ with $\mathbf{D} = 2 \, \| \boldsymbol{\Gamma}_1 \boldsymbol{\Gamma}_1 \|_{\mathcal{F}}^2 \left[\mathbf{1}_{k-1} \mathbf{1}_{k-1}^{\mathrm{T}} / \tau_1 + \operatorname{diag} \left(\tau_2^{-1}, \ldots, \tau_k^{-1} \right) \right] / \| \boldsymbol{\Gamma}_1 \|_{\mathcal{F}}^4$. □

References

[1] Benko, M.; Härdle, P. and Kneip, A. (2008). Common Functional Principal Components. *Ann. Statist.* to appear.

[2] Besse, P. and Ramsay, J. O. (1986). Principal component analysis of sampled functions. *Psychometrika*, **51**, 285-311.

[3] Boente, G. and Fraiman, R. (2000). Kernel-based functional principal components. *Statist. Probab. Lett.*, **48** , 335-345.

[4] Cardot, H. (2007). Conditional functional principal component analysis. *Scand. J. Statist.*. **34**, 317-335.

[5] Dauxois, J., Pousse, A. and Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *J. Multivar. Anal.*, **12**, 136-154.

[6] Ferraty, F. y Vieu, Ph. (2006). *Nonparametric Functional Data Analysis: Theory and Practice.* Springer Series in Statistcs, Springer, New York.

[7] Flury, B. K. (1984). Common principal components in k groups. *J. Am. Statist. Assoc.* **79,** 892-8.

[8] Flury, B. K. (1986). Asymptotic theory for common principal components. *Ann. Statist.* **14,** 418-30.

[9] Flury, B. (1988). *Common principal components and related multivariate models.* New York: John Wiley.

[10] He, G., Müller, H. G. and Wang, J. L. (2004). Methods of canonical analysis for functional data. *J. Statist. Plann. Inference*, **122**, 141-159.

[11] Kayano, M. and Konishi, S. (2008). Functional principal component analysis via regularized Gaussian basis expansions and its application to unbalanced data. In press in *J. Statist. Plann. Inference*.

[12] Pezzulli, S. D. and Silverman, B. W. (1993). Some properties of smoothed principal components analysis for functional data. *Comput. Statist. Data Anal.*, **8**, 1-16.

[13] Ramsay, J. O. and Dalzell, C. J. (1991). Some tools for functional data analysis (with discussion). *J. Roy. Statist. Soc. Ser. B*, **53**, 539-572.

[14] Ramsay, J. O. and Silverman, B. W. (2002). *Applied Functional Data Analysis* . Springer Series in Statistcs, Springer–Verlag

[15] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis, Second Edition*. Springer–Verlag.

[16] Rao, C.R. (1958). Some statistical methods for the comparison of growth curves. *Biometrics,* **14** 117.

[17] Rice, J. and Silverman, B. W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *J. Roy. Statist. Soc. Ser. B*, **53**, 233-243.

[18] Silverman, B. W. (1996). Smoothed functional principal components analysis by choice of norm. *Ann. Statist.*, **24**, 1-24.