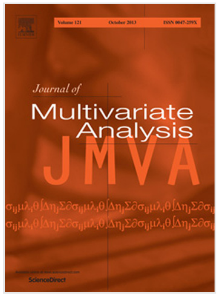
# Accepted Manuscript




Some asymptotic theory for Silverman's smoothed functional principal components in an abstract Hilbert space

Gamage Pemantha Lakraj, Frits Ruymgaart

PII: S0047-259X(16)30254-8
DOI: http://dx.doi.org/10.1016/j.jmva.2016.12.004
Reference: YJMVA 4198

To appear in: *Journal of Multivariate Analysis*

Received date: 2 December 2015

# Some asymptotic theory for Silverman's smoothed functional principal components in an abstract Hilbert space

Gamage Pemantha Lakraj[a,*], Frits Ruymgaart[a]

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

**Abstract**

Unlike classical principal component analysis (PCA) for multivariate data, one needs to smooth or regularize when estimating functional principal components. Silverman's method for smoothed functional principal components has nice theoretical and practical properties. Some theoretical properties of Silverman's method were obtained using tools in the $L^2$ and the Sobolev spaces. This paper proposes an approach, in a general manner, to study the asymptotic properties of Silverman's method in an abstract Hilbert space. This is achieved by exploiting the perturbation results of the eigenvalues and the corresponding eigenvectors of a covariance operator. Consistency and asymptotic distributions of the estimators are derived under mild conditions. First we restrict our attention to the first smoothed functional principal component and then extend the same method for the first $K$ smoothed functional principal components.

*Keywords:* Functional PCA, Smoothing, Hilbert space, Spectrum, Perturbation theory

*2010 MSC:* 00-01, 99-00

*Corresponding author

*Email addresses:* pemantha82@gmail.com (Gamage Pemantha Lakraj), h.ruymgaart@ttu.edu (Frits Ruymgaart)

## 1. Introduction

There are many situations in applied sciences where collected data are originally functions, curves or surfaces rather than vectors. Functional Data Analysis (FDA) is a major branch of research in statistics which deals with data of a continuous nature. In the recent past, a wide scope of contributions related to FDA has been made. The book by Horváth and Kokoszka [18] discusses theoretical and practical aspects related to principal components analysis, time series, change point detection, and spatial statistics in FDA. The essential mathematical concepts and results relevant to the theoretical development of FDA are explored in Hsing and Eubank [19]. The recent reviews by Cuevas [8] and Goia and Vieu [15] survey the main ideas and current literature on regression, classification, bootstrap methods, dimension reduction, semi-parametric modeling, and nonparametric techniques in FDA.

Since functional data are infinite-dimensional, we need to extract important information in order to get a thorough understanding of the structure of the data. Functional principal component analysis (FPCA) provides a finite basis system to represent infinite-dimensional functional data with high accuracy. Many important features and the dominant modes of variations of data are captured by these functional principal components. Early work on FPCA can be found in Deville [10], Dauxois et al. [9], Besse and Ramsay [4], Ramsay and Dalzell [30], Castro et al. [7], and Rice and Silverman [33].

There is a rich collection of literature related to recent work on FPCA. Chapter 8 of Ramsay and Silverman [31] gives a comprehensive discussion of FPCA in the context of the basis representation of the functional data. The properties of FPCA are explained through stochastic expansion and related results in [16]. Kokoszka and Reimherr [24] establish the asymptotic normality of the sample principal components of functional time series data. There are different ways of obtaining estimates in FPCA. Using penalized spline regression, Yao and Lee [36] propose an iterative estimation method for performing FPCA. Ocaña et al. [27] establish a procedure to formulate an algorithm to compute estimates of FPCA under general settings. An approach for robust estimators of functional principal components is given in [3].

FPCA has been applied to many theoretical and practical problems. A direct application of FPCA is functional principal component regression; see, e.g., Cardot et al. [6]. In the context of longitudinal data analysis, a random function usually represents an individual or an item observed at a small number of random points. Hall et al. [17] discuss the application of FPCA to longitudinal data. There are many other practical applications of FPCA in a variety of fields such as the analysis of income density curves [23], spectroscopy data [37], and financial time series data [22].

Since non-smooth functional principal components are too rough for accurate interpretations and advanced analysis, we need to smooth or regularize when estimating them. Many approaches have been proposed to estimate smoothed functional principal components. In one approach, data are smoothed first and FPCA is performed on the smoothed data. Kernel smoothed FPCA is based on this approach, and the asymptotic properties of these principal components are discussed in [5]. Rice and Silverman [33] propose another approach to smoothed FPCA where the variance of principal components is penalized based on a roughness penalty. Rather than penalizing the variance, Silverman [35] incorporates the roughness penalty into the orthonormality constraint in performing smoothed FPCA. An alternative approach to the estimation of FPCA using penalized rank 1 approximation to the data matrix is proposed in [20]. Two different versions of smoothed FPCA based on penalized splines with B-splines are discussed in [1].

There are many important applications of smoothed FPCA. The penalized-components functional version of principal component regression and partial least squares are introduced in [32] based on smoothed FPCA. Luo et al. [26] apply smoothed FPCA for testing association of the entire allelic spectrum of genetic variation. Proximity measures between functional mathematical objects are crucial in semi-parametric and nonparametric FDA. In some situations, semi-metric spaces are better adapted than metric spaces for FDA. As motivated in Section 3.4 of Ferraty and Vieu [12], we can use FPCA as a tool to build a class of semi-metrics. Section 4.3 of [2] exploits semi-functional partial linear modeling, involving functional principal component semi-metrics, to forecast electricity consumption data.

The method for smoothed FPCA in Silverman [35] is an important approach in many ways; see Qui and Zhao [29] for detailed discussion. This method can be studied comprehensively using operator theory in Hilbert spaces. Qui and Zhao [29] discuss some theoretical properties of Silverman's method using tools in $L^2$ and Sobolev spaces. However, we can generalize Silverman's method to an abstract Hilbert space and use perturbation theory to study its theoretical properties in a more general manner.

In this paper, we propose a new approach to study the asymptotic properties of Silverman's smoothed functional principal components in an abstract separable Hilbert space. Our arguments are related to those in Dauxois et al. [9] and involve both Cauchy contours and resolvents. We obtain asymptotic properties using results on the perturbed eigenvalues and eigenvectors of a sample smoothed covariance operator. Consistency and asymptotic distributions of the estimators are derived under mild conditions. For the sake of simplicity of presentation, first we restrict our attention to the first smoothed functional principal component and then extend the same method to the first $K$ principal components.

The paper is set out as follows. In Section 2, we give notations, definitions, assumptions, and the detailed background. Section 3 is devoted to define Silverman's method in an abstract separable Hilbert space and to review some properties. Our main results concerning the asymptotic properties of smoothed functional principal components are given in Section 4. Outlines of the proofs of the lemmas and the theorems in Section 4 are given in Section 5.

## 2. Theoretical framework

We present notations, definitions, assumptions, and theoretical background that are used throughout. Let $\mathbb{H}$ stand for an infinite-dimensional separable Hilbert space over the real numbers with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. The class of all bounded linear operators mapping $\mathbb{H}$ into itself is a Banach space, denoted by $\mathcal{L}$, with operator norm $\|\cdot\|_{\mathcal{L}}$. The subspace of $\mathcal{L}$ which contains Hilbert–Schmidt operators is denoted by $\mathcal{L}_{HS}$, with norm $\|\cdot\|_{HS}$. A simple example of an operator in $\mathcal{L}_{HS}$ is the operator $(a \otimes b)$, with $a, b \in \mathbb{H}$, defined by its action:

$$\forall_{x \in \mathbb{H}} \quad (a \otimes b)x = \langle x, b \rangle a.$$

A family of orthogonal projections $\{E(t)_{t \in \mathbb{R}}\}$ on Hilbert space $\mathbb{H}$ is called a resolution of identity supported by the compact interval $[m, M]$ if

1. $\mathrm{Im}E(s) \subset \mathrm{Im}E(t)$ whenever $s \leq t$,
2. $\mathrm{Im}E(s) = \cap\{\mathrm{Im}E(t) : t > s\}$,
3. $E(t) = 0$ if $t < m$,
4. $E(t) = I$ if $t > M$,

where $\mathrm{Im}E(t)$ represents the image of $E(t)$; see Section V.3 in [14] for details. Let $A \in \mathcal{L}$ be a Hermitian operator. For $t \in \mathbb{R}$, let $E(t)$ be the orthogonal projection of $\mathbb{H}$ onto the spectral subspace of $A$ associated with $(-\infty, t]$. Then, according to Theorem 3.2 in [14], $\{E(t)\}_{t \in \mathbb{R}}$ is a resolution of identity supported by the interval [m(A),M(A)], where

$$m(A) = \inf_{\|f\|=1} \langle Af, f \rangle, \quad M(A) = \sup_{\|f\|=1} \langle Af, f \rangle.$$

Furthermore, as stated in Theorem 2.1 of [14], $\sigma(A) \subset [m(A), M(A)]$, where $\sigma(A)$ is the spectrum of the operator $A$.

Let $U : \mathbb{H} \to \mathbb{H}$ be a bounded Hermitian linear operator, and let $\{E(\lambda)\}_{\lambda \in \mathbb{R}}$ be the resolution of identity for $U$. Then according to the spectral theorem of Hermitian operators,

$$f(U) = \int_{\sigma(U)} f(\lambda)\, dE(\lambda) \tag{2.1}$$

for all continuous functions on $\sigma(U)$; see Section V.4 of [14] and Theorem 9′ of Chapter 31 in [25] for details. Take $f(\lambda) = \lambda$ for all $\lambda \in \sigma(U)$; we then have the following Stieltjes integral representation for $U$:

$$U = \int_{\sigma(U)} \lambda \, dE(\lambda). \tag{2.2}$$

Let $T$ be a bounded linear operator in $\mathbb{H}$ with its spectrum $\sigma(T)$, and let $\Omega$ be a bounded open region in $\mathbb{C}$ with smooth boundary $\Gamma = \partial\Omega$, such that

$$\sigma(T) \subset \Omega, \quad \delta_\Gamma = \text{distance}(\Gamma, \sigma(T)) > 0.$$

The resolvent of the operator $T$, viz.

$$R(z) = (zI - T)^{-1}, \quad z \in \rho(T)$$

is bounded and analytic on the resolvent set $\rho(T) = \{\sigma(T)\}^c$. Let $D \supset \bar{\Omega}$ be an open neighborhood of $\bar{\Omega}$ and $\varphi : D \to \mathbb{C}$ be an analytic function. Then the following operator is well defined (see Section 2 in Gilliam et al. [13]):

$$\varphi(T) = \frac{1}{2\pi i} \oint_\Gamma \varphi(z) R(z) \, dz. \tag{2.3}$$

The mapping $\varphi \mapsto \varphi(T)$ establishes an algebra homomorphism. In particular, it follows that for any $\psi$ which is analytic on $D$,

$$\varphi(T)\psi(T) = (\varphi\psi)(T). \tag{2.4}$$

Furthermore, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathbb{H}$ equipped with the Borel $\sigma$-algebra $\mathcal{B}_\mathbb{H}$. A random variable in $\mathbb{H}$ is a mapping $X : \Omega \to \mathbb{H}$ such that $X$ is $(\mathcal{F}, \mathcal{B}_\mathbb{H})-$measurable. Assuming that $\mathrm{E}\|X\|^2 < \infty$, we can define the mean $\mu \in \mathbb{H}$ and the covariance operator $\Sigma : \mathbb{H} \to \mathbb{H}$ by the requirements

$$\forall_{a,b\in\mathbb{H}} \quad E\langle a, X\rangle = \langle a, \mu\rangle, \quad E\langle a, X - \mu\rangle\langle X - \mu, b\rangle = \langle a, \Sigma b\rangle,$$

respectively. Given a random sample $X_1, \ldots, X_n$ of independent copies of $X$, the sample mean and the sample standard deviation are defined by

$$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, \quad \hat{\Sigma} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}) \otimes (X_i - \bar{X}),$$

respectively.

We make the following main assumptions throughout the paper:

**Assumption 1.** We assume that the covariance operator $\Sigma$ is strictly positive with simple eigenvalues $\lambda_k > 0$, $k = 1, 2, \ldots$

**Remark 1.** Since the covariance operator $\Sigma$ is Hermitian and compact, we can arrange the eigenvalues of $\Sigma$ as a decreasing sequence, $\lambda_1 > \lambda_2 > \cdots$, which converges to zero. The spectrum of $\Sigma$ is $\sigma(\Sigma) = \{0, \lambda_1, \ldots\}$, which is a compact subset in $\mathbb{R}$. If $\{e_1, e_2, \ldots\}$ is a complete orthonormal basis of eigenvectors corresponding to eigenvalues $\{\lambda_1, \lambda_2, \ldots\}$, then the operator $\Sigma$ can be written as

$$\Sigma = \sum_{k=1}^{\infty} \lambda_k E_k,$$

where $E_k = e_k \otimes e_k$ is the orthogonal projection onto the one-dimensional subspace spanned by $e_k$.

**Assumption 2.** We assume that any random element $X \in \mathbb{H}$ is mean centered and

$$E\left\|X\right\|^2 < M < \infty, \quad E\left\|X\right\|^4 < \infty.$$

## 3. Silverman's method in an abstract Hilbert space

According to Silverman's method in [35], the first smoothed functional principal component is the maximizer of

$$\frac{\mathrm{var}\langle f, X\rangle}{\langle f, f\rangle + \alpha\langle D^{1/2}f, D^{1/2}f\rangle}$$

with respect to $f$ which is an element in a sufficiently smoothed function space, and $D$ is a linear differential operator. The roughness of the function $f$ is measured by $\langle D^{1/2}f, D^{1/2}f\rangle$, and the tuning parameter $\alpha > 0$ controls it. It is known from operator theory that the inverse of a linear differential operator is an integral operator; see for instance Chapter 10 in [21]. By combining this property of linear differential operators with the idea in Silverman's method, we can define smoothed functional principal components in a more general manner.

Here, we define smoothed functional principal components in an abstract Hilbert space. The operators involved are also abstract, but with properties that are similar to those in Silverman [35]. Let $S \in \mathcal{L}$ be a strictly positive Hermitian operator. It is known that the range $\mathcal{D}$ of $S$ is dense in $\mathbb{H}$. The Hermitian operator $S \in \mathcal{L}$ has a compact spectrum $\sigma(S)$ contained in an interval $[0, \sigma]$, $0 < \sigma < \infty$, with the resolution of identity $\{E(\lambda), 0 < \lambda \leq \sigma\}$. Then, according to (2.2), the operator $S$ can be represented as

$$S = \int_{\sigma(S)} \lambda \, dE(\lambda).$$

Since $S$ is strictly positive, the inverse of this operator $R : \mathcal{D} \to \mathbb{H}$ exists, and (2.1) yields

$$R = S^{-1} = \int_{\sigma(S)} \frac{1}{\lambda} \, dE(\lambda). \tag{3.1}$$

Furthermore, let us consider a Hermitian operator $(I + \alpha R)^{1/2}$, where $\alpha \geq 0$. Since eigenvalues are strictly positive, the operator $(I + \alpha R)^{1/2}$ is one-to-one and its inverse

$$T_\alpha = (I + \alpha R)^{-1/2} = \int_{\sigma(S)} \left( \frac{\lambda}{\lambda + \alpha} \right)^{1/2} dE(\lambda) \tag{3.2}$$

is bounded and Hermitian on $\mathbb{H}$. It is clear from (3.2) that $\|T_\alpha\|_{\mathcal{L}} \leq 1$. Using the operator $T_\alpha$, let us define an inner product as follows, for all $f, g \in \mathcal{D}$:

$$\begin{aligned} \langle f, g \rangle_\alpha &= \langle f, g \rangle + \alpha \langle R^{1/2} f, R^{1/2} g \rangle \\ &= \langle f, g \rangle + \alpha \langle f, Rg \rangle \\ &= \langle (I + \alpha R)^{1/2} f, (I + \alpha R)^{1/2} g \rangle \\ &= \langle T_\alpha^{-1} f, T_\alpha^{-1} g \rangle. \end{aligned}$$

This is akin to weighted Sobolev inner product and norms if $\mathcal{D}$ is an appropriate function space; see Zhikov [38].

With the above settings, now we can define Silverman's smoothed functional principal components, in a more general manner, in the following way:

$$\gamma_{\alpha,1} = \underset{f \in \mathcal{D},\ f \neq 0}{\arg\max} \frac{\langle f, \Sigma f \rangle}{\|f\|_\alpha^2} = \underset{f \in \mathcal{D},\ f \neq 0}{\arg\max} \frac{\langle f, T_\alpha^2 \Sigma f \rangle_\alpha}{\|f\|_\alpha^2},$$

and if $i > 1$, then

$$\gamma_{\alpha,i} = \underset{f \in \mathcal{D},\ f \neq 0}{\arg\max} \frac{\langle f, \Sigma f\rangle}{\|f\|_\alpha^2} = \underset{f \in \mathcal{D},\ f \neq 0}{\arg\max} \frac{\langle f, T_\alpha^2 \Sigma f\rangle_\alpha}{\|f\|_\alpha^2} \tag{3.3}$$

such that $\langle \gamma_{\alpha,i}, \gamma_{\alpha,j}\rangle_\alpha = 0$ for all $j < i$.

In this definition, the orthonormal constraints are imposed by the inner product $\langle \cdot, \cdot\rangle_\alpha$. Hence, these smoothed principal components are orthonormal with respect to $\langle \cdot, \cdot\rangle_\alpha$ but not $\langle \cdot, \cdot\rangle$. Further, Silverman's smoothed functional principal components can be considered as the eigenvectors of the operator $T_\alpha^2 \Sigma$.

Since the range of the operator $(I + \alpha R)^{1/2}$ is $\mathbb{H}$, we can find $g \in \mathbb{H}$ such that $(I + \alpha R)^{1/2} f = g$. Then it follows from definition (3.3) that

$$\begin{aligned}
\lambda_{\alpha,1} &= \max_{f \in \mathcal{D},\ f \neq 0} \frac{\langle f, \Sigma f\rangle}{\|f\|_\alpha^2} = \max_{f \in \mathcal{D},\ f \neq 0} \frac{\langle f, \Sigma f\rangle}{\left\|(I + \alpha R)^{1/2} f\right\|^2} \\
&= \max_{g \in \mathbb{H},\ g \neq 0} \frac{\langle g, (I + \alpha R)^{-1/2}\, \Sigma\, (I + \alpha R)^{-1/2}\, g\rangle}{\|g\|^2} \\
&= \max_{g \in \mathbb{H},\ g \neq 0} \frac{\langle g, T_\alpha \Sigma T_\alpha g\rangle}{\|g\|^2}. \qquad (3.4)
\end{aligned}$$

Note that $T_\alpha \Sigma T_\alpha$ is the covariance operator of the random variable $T_\alpha X$. Let $e_{\alpha,1}$ be an eigenvector corresponding to the largest eigenvalue $\lambda_{\alpha,1}$ of $T_\alpha \Sigma T_\alpha$. Then, as described in Proposition 5.4 of Ocaña et al. [28], Silverman's smoothed first functional principal component is

$$\gamma_{\alpha,1} = T_\alpha e_{\alpha,1},$$

which is an eigenvector corresponding to the largest eigenvalue $\lambda_{\alpha,1}$ of $T_\alpha^2 \Sigma$. Similarly, we can find higher order principal components.

The population covariance operator $\Sigma$ is usually unknown to us. Hence, we have to use the eigenvectors of a sample covariance operator $\hat{\Sigma}$ to estimate smoothed functional principal components. Using $\hat{\Sigma}$ we estimate $\lambda_{\alpha,1}$ and $e_{\alpha,1}$ by

$$\hat{\lambda}_{\alpha,1} = \text{largest eigenvalue of } T_\alpha \hat{\Sigma} T_\alpha$$

and

$$\hat{e}_{\alpha,1} = \underset{g \in \mathbb{H},\ g \neq 0}{\arg\max} \frac{\langle g, T_\alpha \hat{\Sigma} T_\alpha g\rangle}{\|g\|^2},$$

respectively. Hence, Silverman's smoothed first sample principal component is

$$\hat{\gamma}_{\alpha,1} = T_\alpha \hat{e}_{\alpha,1}, \tag{3.5}$$

which is an eigenvector of $T_\alpha^2 \hat{\Sigma}$ corresponding to the eigenvalue $\hat{\lambda}_{\alpha,1}$.

## 4. Asymptotic theory

In this section, we discuss some asymptotic properties of Silverman's smoothed functional principal component and the corresponding eigenvalues. With a view towards (3.4), let us write

$$\Sigma_\alpha = T_\alpha \Sigma T_\alpha, \quad \Pi_\alpha = \Sigma_\alpha - \Sigma.$$

The perturbation $\Pi_\alpha$ of $\Sigma$ is Hermitian and bounded. In order to study the asymptotic behavior of $\Pi_\alpha$, we consider the following inequality:

$$\begin{aligned} \|\Sigma_\alpha - \Sigma\|_{\mathcal{L}} &= \|T_\alpha \Sigma T_\alpha - \Sigma\|_{\mathcal{L}} \\ &\leq \|T_\alpha \Sigma T_\alpha - \Sigma T_\alpha\|_{\mathcal{L}} + \|\Sigma T_\alpha - \Sigma\|_{\mathcal{L}}. \end{aligned} \tag{4.1}$$

The first part of this decomposition is treated as

$$\begin{aligned} \|T_\alpha \Sigma T_\alpha - \Sigma T_\alpha\|_{\mathcal{L}} &= \|E\left(T_\alpha X \otimes T_\alpha X\right) - E\left(X \otimes T_\alpha X\right)\|_{\mathcal{L}} \\ &\leq E\left\|\left(T_\alpha X - X\right) \otimes T_\alpha X\right\|_{\mathcal{L}} \\ &\leq E\left\{\|(T_\alpha - I)X\| \, \|X\|\right\}, \end{aligned} \tag{4.2}$$

and the second part is treated as

$$\begin{aligned} \|\Sigma T_\alpha - \Sigma\|_{\mathcal{L}} &= \|E(X \otimes T_\alpha X) - E(X \otimes X)\|_{\mathcal{L}} \\ &\leq E\left\|X \otimes \left(T_\alpha X - X\right)\right\|_{\mathcal{L}} \\ &\leq E\left\{\|(T_\alpha - I)X\| \, \|X\|\right\}. \end{aligned} \tag{4.3}$$

Hence, combining (4.1), (4.2), and (4.3), we obtain

$$\|\Sigma_\alpha - \Sigma\|_{\mathcal{L}} \leq 2E\left\{\|(T_\alpha - I)X\| \, \|X\|\right\}.$$

The following lemma reveals the asymptotic behavior of $E\left\|(T_\alpha - I)X\right\|^2$.

**Lemma 1.** *Under Assumption 2 and Definition* (3.2)*, we have*

$$E\,\|(I-T_\alpha)X\|^2 = \mathcal{O}(\alpha^2) \quad \textit{as } \alpha \to 0.$$

Since $E\,\|X\|^2 < \infty$, Lemma 1 and the Cauchy–Schwarz inequality yield the following result:

$$\|\Sigma_\alpha - \Sigma\|_{\mathcal{L}} = \mathcal{O}(\alpha) \quad \text{as } \alpha \to 0. \tag{4.4}$$

Similarly, let us write

$$\hat{\Sigma}_\alpha = T_\alpha \hat{\Sigma} T_\alpha, \quad \hat{\Pi}_\alpha = \hat{\Sigma}_\alpha - \Sigma,$$

where $\hat{\Sigma}_\alpha$ the sample analogue to $\Sigma_\alpha$, and $\hat{\Pi}_\alpha$ is random. Under the condition $E\,\|X\|^4 < \infty$, Dauxois et al. [9] showed that there exists a Gaussian random element $\mathbb{G}$ with zero mean in $\mathcal{L}_{HS}$ (and hence in $\mathcal{L}$ by the Continuous Mapping Theorem), such that, as $n \to \infty$,

$$\sqrt{n}\,(\hat{\Sigma} - \Sigma) \rightsquigarrow \mathbb{G}, \tag{4.5}$$

where $\rightsquigarrow$ denotes convergence in distribution. Using (4.4) and (4.5), we see that, as $\alpha \to 0$ and $n \to \infty$,

$$\begin{aligned}
\|\hat{\Sigma}_\alpha - \Sigma\|_{\mathcal{L}} &= \|\Sigma_\alpha - \Sigma + \hat{\Sigma}_\alpha - \Sigma_\alpha\|_{\mathcal{L}} \\
&\le \|\Sigma - \Sigma_\alpha\|_{\mathcal{L}} + \|T_\alpha\|_{\mathcal{L}}\,\|\hat{\Sigma} - \Sigma\|_{\mathcal{L}}\,\|T_\alpha\|_{\mathcal{L}} \\
&= \mathcal{O}(\alpha) + \mathcal{O}_p\left(1/\sqrt{n}\right). \qquad (4.6)
\end{aligned}$$

This means that the random perturbation $\hat{\Pi}_\alpha$ of $\Sigma$ is small for a large sample size $n$ and a small value of smoothing parameter $\alpha$.

For the asymptotic expansion of eigenvalues and its corresponding eigenvectors of $\hat{\Sigma}_\alpha$, we need to consider the convergence rate of $\hat{\Pi}_\alpha$. Therefore, let us take

$$\alpha = \alpha(n) = o(n^{-1}) \tag{4.7}$$

and then (4.6) yields

$$\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}} = \mathcal{O}_p\left(\frac{1}{\sqrt{n}}\right). \tag{4.8}$$

Moreover, we have the following lemma, which is useful in discussing the asymptotic distributions of the smoothed functional principal components.

**Lemma 2.** *Under Condition* (4.7),

$$\|\sqrt{n}(\hat{\Sigma}_{\alpha(n)} - \hat{\Sigma})\|_{\mathcal{L}} = o_p(1).$$

Now observe that

$$\begin{aligned} \|T_\alpha^2\hat{\Sigma} - \Sigma\|_{HS} &= \|T_\alpha^2\Sigma - \Sigma + T_\alpha^2\hat{\Sigma} - T_\alpha^2\Sigma\|_{HS} \\ &\leq \|T_\alpha^2\Sigma - \Sigma\|_{HS} + \|T_\alpha^2\|_{\mathcal{L}}\|\hat{\Sigma} - \Sigma\|_{HS}. \end{aligned}$$

Using Lemma 1, we can show that $\|T_\alpha^2\Sigma - \Sigma\|_{HS} = \mathcal{O}(\alpha)$ as $\alpha \to 0$. Dauxois et al. [9] showed, by exploiting the Strong Law of Large Numbers in Hilbert spaces, that $\|\hat{\Sigma} - \Sigma\|_{HS}$ coverages a.s. to 0 as $n \to \infty$. Hence, under Condition (4.7), $T_{\alpha(n)}^2\hat{\Sigma}$ coverages a.s. to $\Sigma$ in $\mathcal{L}_{HS}$ as $n \to \infty$; let $\Omega_1$ be the convergence set.

We know that Silverman's sample smoothed functional principal components are the eigenvectors of the compact operator $T_\alpha^2\hat{\Sigma}$. For $i \in I$, where $I$ is either a beginning section of $\mathbb{N}$ or $\mathbb{N}$ itself, let $\hat{\gamma}_{\alpha,i}$ be the normalized eigenvector corresponding to the eigenvalue $\hat{\lambda}_{\alpha,i}$ of the operator $T_\alpha^2\hat{\Sigma}$. Then, under Assumptions 1–2 and Condition (4.7), we can deduce that $\hat{\gamma}_{\alpha(n),i}$ converges to $e_i$ on $\Omega_1$ as $n \to \infty$, and $\hat{\lambda}_{\alpha(n),i}$ converges $\lambda_i$ on $\Omega_1$ as $n \to \infty$; see p. 1091 in [11] and the proofs of Propositions 3 and 4 in [9] for details.

### *4.1. The spectrum of a sample smoothed covariance operator*

As shown in (3.5), estimators of Silverman's smoothed functional principal components are closely related to the eigenvectors of $\hat{\Sigma}_{\alpha(n)}$. Thus, it is necessary to study the spectral properties of a sample smoothed covariance operator $\hat{\Sigma}_{\alpha(n)}$ in order to get the asymptotic properties of estimators of Silverman's smoothed functional principal components.

We can use the resolvent $R(z)$, the spectrum of the covariance operator $\Sigma$, and the random perturbation $\hat{\Pi}_{\alpha(n)}$ to obtain the spectrum of $\hat{\Sigma}_{\alpha(n)}$. The operator $R(z) = (zI - \Sigma)^{-1}$ is the resolvent of the covariance operator $\Sigma$ and it is well defined and bounded for all $z \in \rho(\Sigma) = \{\sigma(\Sigma)\}^c = \{0, \lambda_1, \ldots\}^c$. For all $z \in \rho(\Sigma)$, it is known that

$$\|R(z)\|_{\mathcal{L}} = \frac{1}{\delta(z)}, \tag{4.9}$$

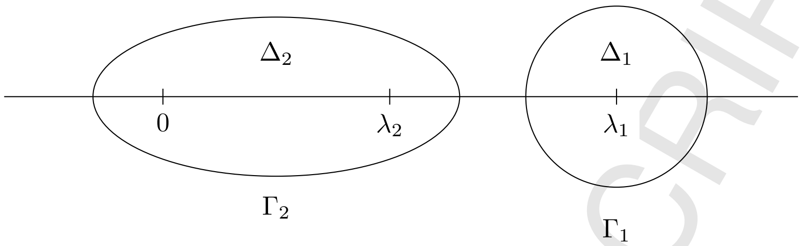



Figure 1: Contours which separate largest eigenvalue $\lambda_1$ of the covariance operator $\Sigma$ from the other eigenvalues.

where $\delta(z) = \text{distance}(z, \sigma(\Sigma))$.

Now we select two contours $\Gamma_1$ and $\Gamma_2$, as described below, such that the largest eigenvalue $\lambda_1$ of $\Sigma$ is in $\Gamma_1$ and the other eigenvalues are in $\Gamma_2$. Let $\Delta_1 \in \mathbb{C}$ be an open disc centered at eigenvalue $\lambda_1$ with radius $0 < \tau < (\lambda_1 - \lambda_2)/2$, and $\Gamma_1$ denote its smooth boundary. Also consider the smooth contour $\Gamma_2$ at distance $\tau$ from $[0, \lambda_2]$ and denote its interior by $\Delta_2$; see Figure 1. Let us write $\Gamma = \Gamma_1 \cup \Gamma_2$ and $\Delta = \Delta_1 \cup \Delta_2$ such that $\sigma(\Sigma) \subset \Delta$.

The asymptotic expansion for the perturbed eigenvalues and eigenvectors are only valid for sufficiently small perturbations. Therefore, let us introduce a subset

$$\Omega_n = \{\omega \in \Omega : \|\hat{\Pi}_{\alpha(n)}(\omega)\|_{\mathcal{L}} \leq n^{-1/3}\}, \tag{4.10}$$

of the sample space $\Omega$, and let $\Pr(\Omega_n)$ be the probability of the subset $\Omega_n$. Then we have

$$\Pr(\Omega_n) \to 1 \quad \text{as } n \to \infty, \tag{4.11}$$

under the condition $\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}} = \mathcal{O}_p\left(1/\sqrt{n}\right)$.

With the above contours fixed, it is clear from (4.9) and (4.10) that there exists $n_0 \in \mathbb{N}$ such that

$$\forall_{n \geq n_0} \quad \forall_{\omega \in \Omega_n} \quad \sup_{z \in \Gamma} \|\hat{\Pi}_{\alpha(n)}(\omega) R(z)\|_{\mathcal{L}} < \frac{1}{2}. \tag{4.12}$$

Henceforth we will tacitly assume that $n$ is at least $n_0$. Let $\hat{R}_{\alpha(n)}(z)$ be the resolvent of $\hat{\Sigma}_{\alpha(n)}$, whenever defined. Then for all $\omega \in \Omega_n$ the expansion

$$\begin{aligned} \hat{R}_{\alpha(n)}(z) &= (zI - \hat{\Sigma}_{\alpha(n)})^{-1} = \{zI - (\hat{\Pi}_{\alpha(n)} + \Sigma)\}^{-1} \\ &= (R^{-1}(z) - \hat{\Pi}_{\alpha(n)})^{-1} = R(z)\{I - \hat{\Pi}_{\alpha(n)}R(z)\}^{-1} \quad (4.13) \end{aligned}$$

is well defined for all $z \notin \Delta$. In particular, the above expansion is defined for $z \in \Gamma = \Gamma_1 \cup \Gamma_2$, and hence the spectrum of $\hat{\Sigma}_{\alpha(n)}$ satisfies

$$\sigma(\hat{\Sigma}_{\alpha(n)}) \subset \Delta_1 \cup \Delta_2 = \Delta.$$

Therefore, all the eigenvalues of $\hat{\Sigma}_{\alpha(n)}$ are in $\Delta$ for a sufficiently large $n$.

From (2.3), it is well known that

$$1_{\bar{\Delta}_1}(\Sigma) = \frac{1}{2\pi i}\oint_\Gamma 1_{\bar{\Delta}_1}(z)R(z)\,dz = E_1 = e_1 \otimes e_1,$$

where $E_1$ is the eigenprojection corresponding to the eigenvalue $\lambda_1$ of covariance operator $\Sigma$. Similarly, let us define

$$\hat{E}_{\alpha(n),1} = 1_{\bar{\Delta}_1}(\hat{\Sigma}_{\alpha(n)}) = \frac{1}{2\pi i}\oint_\Gamma 1_{\bar{\Delta}_1}(z)\hat{R}_{\alpha(n)}(z)\,dz.$$

This operator is Hermitian and (2.4) yields

$$\begin{aligned} \hat{E}_{\alpha(n),1}\hat{E}_{\alpha(n),1} &= 1_{\bar{\Delta}_1}(\hat{\Sigma}_{\alpha(n)})1_{\bar{\Delta}_1}(\hat{\Sigma}_{\alpha(n)}) = 1_{\bar{\Delta}_1}.1_{\bar{\Delta}_1}(\hat{\Sigma}_{\alpha(n)}) \\ &= \frac{1}{2\pi i}\oint_\Gamma 1_{\bar{\Delta}_1}(z)1_{\bar{\Delta}_1}(z)\hat{R}_{\alpha(n)}(z)\,dz = \hat{E}_{\alpha(n),1}. \end{aligned}$$

The operator $\hat{E}_{\alpha(n),1}$ is Hermitian and idempotent, and hence it is an orthogonal projection operator.

In addition, it can be shown that $\|E_1 - \hat{E}_{\alpha(n),1}\|_{\mathcal{L}} < 1$ for sufficiently small $\hat{\Pi}_{\alpha(n)}$; see, e.g., pp. 372–373 in [34]. Thus the operator $\hat{E}_{\alpha(n),1}$ is one-dimensional, and we can find an element $\hat{e}_{\alpha(n),1} \in \mathbb{H}$ with $\|\hat{e}_{\alpha(n),1}\| = 1$ such that

$$\hat{E}_{\alpha(n),1} = \hat{e}_{\alpha(n),1} \otimes \hat{e}_{\alpha(n),1}.$$

Using Corollary 3.2 in Gilliam et al. [13], it can be shown that

$$\hat{\Sigma}_{\alpha(n)}\hat{e}_{\alpha(n),1} = \langle \hat{\Sigma}_{\alpha(n)}\hat{e}_{\alpha(n),1}, \hat{e}_{\alpha(n),1}\rangle \hat{e}_{\alpha(n),1} = \hat{\lambda}_{\alpha(n),1}\hat{e}_{\alpha(n),1}.$$

Thus, $\hat{e}_{\alpha(n),1}$ is an eigenvector, and $\hat{E}_{\alpha(n),1}$ is the one-dimensional eigenprojection corresponding to the largest eigenvalue $\hat{\lambda}_{\alpha(n),1}$ of $\hat{\Sigma}_{\alpha(n)}$. Hence, for a sufficiently small perturbation $\hat{\Pi}_{\alpha(n)}$, the only point of $\sigma(\hat{\Sigma}_{\alpha(n)})$ inside the contour $\Gamma_1$ is $\hat{\lambda}_{\alpha(n),1}$.

*4.2. Asymptotic results for the smoothed first functional principal component*

The perturbation $\hat{\Pi}_{\alpha(n)}$ satisfies (4.12) on $\Omega_n$, and hence it follows from Corollary 3.3 in Gilliam et al. [13] that, on $\Omega_n$,

$$\hat{e}_{\alpha(n),1} = e_1 + Q_1 \hat{\Pi}_{\alpha(n)} e_1 + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}}^2), \tag{4.14}$$

where

$$Q_1 = \sum_{j=2}^{\infty} \frac{1}{\lambda_1 - \lambda_j} E_j.$$

Furthermore, Corollary 3.4 in Gilliam et al. [13] yields, on $\Omega_n$,

$$\hat{\lambda}_{\alpha(n),1} = \lambda_1 + \langle \hat{\Pi}_{\alpha(n)} e_1, e_1 \rangle + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}}^2). \tag{4.15}$$

Now we have the following main result.

**Theorem 1.** *Under Assumptions 1–2 and Condition* (4.7), $\sqrt{n}\left(\hat{e}_{\alpha(n),1} - e_1\right) \rightsquigarrow Q_1 \mathbb{G} e_1$ *in* $\mathbb{H}$ *and* $\sqrt{n}\,(\hat{\lambda}_{\alpha(n),1} - \lambda_1) \rightsquigarrow \langle \mathbb{G} e_1, e_1 \rangle$ *in* $\mathbb{R}$*, where* $\rightsquigarrow$ *denotes convergence in distribution.*

The following corollary gives the asymptotic distribution of Silverman's smoothed first sample principal component $\hat{\gamma}_{\alpha(n),1}$.

**Corollary 1.** *Under Assumptions 1–2 and Condition* (4.7), $\sqrt{n}\,(\hat{\gamma}_{\alpha(n),1} - e_1) \rightsquigarrow Q_1 \mathbb{G} e_1$ *in* $\mathbb{H}$.

So far we have focused on the first smoothed functional principal component and the corresponding eigenvalue. However, under some mild conditions, we can easily derive multivariate asymptotic distributions for the first $K$ smoothed principal components and the corresponding eigenvalues.

*4.3. Joint asymptotic distributions*

For fixed positive integer $K$, let $\lambda_1 > \cdots > \lambda_K > 0$ be the first $K$ eigenvalues of $\Sigma$. Then, we can find $\epsilon > 0$ such that $K+1$ intervals $[\lambda_1 - \epsilon, \lambda_1 + \epsilon], \ldots, [0, \lambda_{K+1} + \epsilon]$ are disjoint. For each $k \in \{1, \ldots, K\}$, let $\Gamma_k$ be the smooth boundary of the circle $C_k$ with center $\lambda_k$ and radius $\epsilon$ and $\Delta_k$ be the interior of the circle $C_k$. Also consider the smooth contour $\Gamma_{K+1}$ at distance $0 < \tau < \epsilon$ from $[0, \lambda_{K+1}]$ and denote its interior by $\Delta_{K+1}$. We take $\Gamma = \bigcup_{k=1}^{K+1} \Gamma_k$ and $\Delta = \bigcup_{k=1}^{K+1} \Delta_k$ such that $\sigma(\Sigma) \subset \Delta$. It is clear that the only point of the spectrum of the covariance operator $\Sigma$ in $\Delta_k$ is $\lambda_k$ for each $k \in \{1, \ldots, K\}$.

As shown in (4.13), the resolvent operator $\hat{R}_{\alpha(n)}(z)$ is well defined for all $\omega \in \Omega_n$ and $z \notin \Delta = \bigcup_{k=1}^{K+1} \Delta_k$. In particular, $\hat{R}_{\alpha(n)}(z)$ is well defined for all $z \in \Gamma = \bigcup_{k=1}^{K+1} \Gamma_k$. Thus, for a sufficiently small perturbation $\hat{\Pi}_{\alpha(n)}$,

$$\sigma(\hat{\Sigma}_{\alpha(n)}) \subset \bigcup_{k=1}^{K+1} \Delta_k = \Delta.$$

Then, similarly to (4.14) and (4.15), we can obtain the asymptotic expansions of the eigenvalue $\hat{\lambda}_{\alpha(n),k}$ and the corresponding eigenvector $\hat{e}_{\alpha(n),k}$ of the sample smooth covariance operator $\hat{\Sigma}_{\alpha(n)}$, viz.

$$\hat{\lambda}_{\alpha(n),k} = \lambda_k + \langle \hat{\Pi}_{\alpha(n)} e_k, e_k \rangle + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}}^2) \ \text{ on } \Omega_n, \ 1 \le k \le K \tag{4.16}$$

and

$$\hat{e}_{\alpha(n),k} = e_k + Q_k \hat{\Pi}_{\alpha(n)} e_k + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}}^2) \ \text{ on } \Omega_n, \ 1 \le k \le K, \tag{4.17}$$

where

$$Q_k = \sum_{j \neq k}^{\infty} \frac{1}{\lambda_k - \lambda_j} E_j. \tag{4.18}$$

Then, by exploiting (4.16) and (4.17), we obtain the following result.

**Proposition 1.** *Under Assumptions 1–2 and Condition* (4.7)

$$\{\sqrt{n}(\hat{e}_{\alpha(n),k} - e_k), 1 \le k \le K\} \rightsquigarrow (Q_1 \mathbb{G} e_1, \ldots, Q_K \mathbb{G} e_K) \ \textit{in} \ \mathbb{H}^K,$$

$$\{\sqrt{n}\left(\lambda_{\alpha(n),k} - \lambda_k\right), 1 \le k \le K\} \rightsquigarrow (\langle \mathbb{G} e_1, e_1 \rangle, \ldots, \langle \mathbb{G} e_K, e_K \rangle) \ \textit{in} \ \mathbb{R}^K,$$

*where* $\rightsquigarrow$ *denotes convergence in distribution.*

Similarly, we have the following result for Silverman's smoothed functional principal components.

**Corollary 2.** *Under Assumptions 1–2 and Condition* (4.7),

$$\{\sqrt{n}(\hat{\gamma}_{\alpha(n),k} - e_k), 1 \leq k \leq K\} \rightsquigarrow (Q_1 \mathbb{G} e_1, \ldots, Q_K \mathbb{G} e_K) \text{ in } \mathbb{H}^K.$$

## 5. Proofs

In this section, we provide outlines of the proofs of the lemmas and the theorems stated in Section 4.

**Proof of Lemma 1.** Exploiting (3.1) and (3.2), we obtain

$$(I - T_\alpha) = I - (I + \alpha R)^{-1/2} = \int_0^\sigma \left\{ 1 - \left( \frac{\lambda}{\lambda + \alpha} \right)^{1/2} \right\} dE(\lambda).$$

Let $Y$ be a random variable in $\mathcal{D} \subset \mathbb{H}$. It follows from (3.1) that

$$\begin{aligned} \|(I - T_\alpha)Y\|^2 &= \int_0^\sigma \left\{ 1 - \left( \frac{\lambda}{\lambda + \alpha} \right)^{1/2} \right\}^2 d\|E(\lambda)Y\|^2 \\ &\leq \int_0^\sigma \left( \frac{\alpha}{\lambda + \alpha} \right)^2 d\|E(\lambda)Y\|^2 \\ &\leq \alpha^2 \|RY\|^2 . \end{aligned} \tag{5.1}$$

Let $X$ be a random variable in $\mathbb{H}$. Since $\mathcal{D}$ is dense in $\mathbb{H}$, for all $\epsilon > 0$, there exists a random element $Y_\epsilon \in \mathcal{D}$ such that

$$\|X - Y_\epsilon\| \leq \frac{\epsilon}{\sqrt{8}}. \tag{5.2}$$

Then, observe that

$$\begin{aligned} \|(I - T_\alpha)X\|^2 &= \|(I - T_\alpha)(X - Y_\epsilon) + (I - T_\alpha)Y_\epsilon\|^2 \\ &\leq 2\|(I - T_\alpha)(X - Y_\epsilon)\|^2 + 2\|(I - T_\alpha)Y_\epsilon\|^2 \\ &\leq 2\|I - T_\alpha\|_{\mathcal{L}}^2 \|X - Y_\epsilon\|^2 + 2\|(I - T_\alpha)Y_\epsilon\|^2 . \end{aligned}$$

Since $\|T_\alpha\| \leq 1$, (5.1) and (5.2) yield $E\|(I - T_\alpha)X\|^2 \leq \epsilon^2 + 2\alpha^2 E\|RY_\epsilon\|^2$.

The random variable $RY_\epsilon$ is in $\mathbb{H}$, and hence it follows from Assumption 2 that $E\|RY_\epsilon\| < M < \infty$. Since $\epsilon > 0$ is an arbitrary value, we conclude that $E\|(I - T_\alpha)X\|^2 = \mathcal{O}(\alpha^2)$ as $\alpha \to 0$. □

**Proof of Lemma 2.** Suppose $X_1, \dots, X_n$ are independent copies of a random element $X \in \mathbb{H}$. Let us take $\|\sqrt{n}(\hat{\Sigma}_\alpha - \hat{\Sigma})\|_{\mathcal{L}} = \|\sqrt{n}\hat{Z}_\alpha\|_{\mathcal{L}}$. Then using Markov's inequality and the triangle inequality, we find that, for all $\epsilon > 0$,

$$\begin{aligned}
\Pr\left(\left\|\sqrt{n}Z_\alpha\right\|_{\mathcal{L}} > \epsilon\right) &= \Pr\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(T_\alpha X_i \otimes T_\alpha X_i - X_i \otimes X_i\right)\right\|_{\mathcal{L}} > \epsilon\right) \\
&\leq \frac{1}{\sqrt{n}\epsilon}\sum_{i=1}^{n} E\left\|\left(T_\alpha X_i \otimes T_\alpha X_i - X_i \otimes X_i\right)\right\|_{\mathcal{L}}.
\end{aligned}$$

Since $X_1, \dots, X_n$ are independent and $\|T_\alpha\|_{\mathcal{L}} \leq 1$,

$$\begin{aligned}
\Pr\left(\|\sqrt{n}Z_\alpha\|_{\mathcal{L}} > \epsilon\right) &\leq \frac{\sqrt{n}}{\epsilon} E\left\|\left(T_\alpha X \otimes T_\alpha X - X \otimes X\right)\right\|_{\mathcal{L}} \\
&= \frac{\sqrt{n}}{\epsilon} E\left\|T_\alpha X \otimes (T_\alpha X - X) + (T_\alpha X - X) \otimes X\right\|_{\mathcal{L}} \\
&\leq \frac{\sqrt{n}}{\epsilon} E\left(\|T_\alpha X\|\,\|T_\alpha X - X\| + \|T_\alpha X - X\|\,\|X\|\right) \\
&\leq \frac{2\sqrt{n}}{\epsilon} E\left\{\|X\|\,\|(T_\alpha - I)\,X\|\right\}.
\end{aligned}$$

Since $E\|X\|^2 < \infty$, it follows from Lemma 1 and the Cauchy–Schwarz inequality that

$$\Pr\left(\|\sqrt{n}Z_\alpha\|_{\mathcal{L}} > \epsilon\right) \leq \frac{2\sqrt{n}\mathcal{O}(\alpha)}{\epsilon}.$$

Since we take $\alpha = \alpha(n) = o(n^{-1})$ as $n \to \infty$,

$$\Pr(\|\sqrt{n}(\hat{\Sigma}_{\alpha(n)} - \hat{\Sigma})\|_{\mathcal{L}} > \epsilon) = \Pr\left(\|\sqrt{n}Z_{\alpha(n)}\|_{\mathcal{L}} > \epsilon\right) = o\left(1/\sqrt{n}\right).$$

This yields

$$\|\sqrt{n}(\hat{\Sigma}_{\alpha(n)} - \hat{\Sigma})\|_{\mathcal{L}} = o_p(1) \text{ as } n \to \infty.$$

□

**Proof of Theorem 1.** The combination of (4.8), (4.14), and (4.11) yields

$$\begin{aligned}
\sqrt{n}\left(\hat{e}_{\alpha(n),1} - e_1\right) &= \sqrt{n}\left\{(\hat{e}_{\alpha(n),1} - e_1)\mathbf{1}_{\Omega_n} + (\hat{e}_{\alpha(n),1} - e_1)\mathbf{1}_{\Omega_n^c}\right\} \\
&= \sqrt{n}\left\{Q_1\hat{\Pi}_{\alpha(n)}e_1 + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|^2)\right\}\mathbf{1}_{\Omega_n} + \\
&\quad \sqrt{n}\left(\hat{e}_{\alpha(n),1} - e_1\right)\mathbf{1}_{\Omega_n^c} \\
&= \sqrt{n}\,Q_1\hat{\Pi}_{\alpha(n)}e_1 + o_p(1) \quad \text{as } n \to \infty.
\end{aligned}$$

Now observe that

$$\sqrt{n}\,(\hat{\Pi}_{\alpha(n)}) = \sqrt{n}\,(\hat{\Sigma}_{\alpha(n)} - \Sigma) = \sqrt{n}\,(\hat{\Sigma} - \Sigma) + \sqrt{n}\,(\hat{\Sigma}_{\alpha(n)} - \hat{\Sigma}). \tag{5.3}$$

Thus, (4.5) and Lemma 2 give the first result.

Similarly, we can obtain that

$$\sqrt{n}\,(\hat{\lambda}_{\alpha(n)} - \lambda) = \sqrt{n}\,\langle \hat{\Pi}_\alpha(n) e_1, e_1\rangle + o_p(1) \quad \text{as} \quad n \to \infty,$$

by exploiting (4.8), (4.14), and Lemma 2. Hence, the second result follows from (4.5), and (5.3). □

**Proof of Corollary 1.** Silverman's smoothed first sample principal component $\hat{\gamma}_{\alpha(n),1}$ is the eigenvector corresponding to the largest eigenvalue $\hat{\lambda}_{\alpha(n),1}$ of $T_\alpha^2\hat{\Sigma}$. Now let us take

$$\hat{\Sigma}_\alpha = T_\alpha^2 \Sigma, \quad \hat{\Pi}_\alpha = \hat{\Sigma}_\alpha - \Sigma.$$

Under the condition $\alpha = \alpha(n) = o(n^{-1})$, it follows from Lemma 1 that

$$\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}} = \|T^2_{\alpha(n)}\hat{\Sigma} - \Sigma\|_{\mathcal{L}} = \mathcal{O}_p\left(1/\sqrt{n}\right) \quad \text{as } n \to \infty. \tag{5.4}$$

We select a subset $\Omega_n$ of the sample space $\Omega$, such that

$$\Omega_n = \{\omega : \|\hat{\Pi}_{\alpha(n)}(\omega)\|_{\mathcal{L}} = \|T^2_{\alpha(n)}\hat{\Sigma} - \Sigma\|_{\mathcal{L}} \le n^{-1/3}\},$$

and it follows from (5.4) and (4.11) that $\Pr(\Omega_n) \to 1$, as $n \to \infty$. On this subset $\Omega_n$, the perturbation $\hat{\Pi}_{\alpha(n)} = (T^2_{\alpha(n)}\hat{\Sigma} - \Sigma)$ satisfies (4.12). Thus, using Corollary 3.3 and Remark 3.6 in Gilliam et al. [13], we obtain that

$$\hat{\gamma}_{\alpha(n),1} = e_1 + Q_1\hat{\Pi}_{\alpha(n)}e_1 + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|^2_{\mathcal{L}}) \text{ on } \Omega_n. \tag{5.5}$$

Similar to Lemma 2, it can be shown that

$$\|\sqrt{n}\,(\hat{\Sigma}_{\alpha(n)} - \hat{\Sigma})\|_{\mathcal{L}} = \|\sqrt{n}\,(T^2_{\alpha(n)}\hat{\Sigma} - \hat{\Sigma})\|_{\mathcal{L}} = o_p(1) \text{ as } n \to \infty. \tag{5.6}$$

By exploiting (5.4), (5.5), and (5.6), we complete this proof similarly to the proof of the first result of Theorem 1 . □

**Proof of Proposition 1.** By exploiting (4.16), (4.17), and Condition (4.7), we can obtain the following results similarly to the proof of Theorem 1. For all $k \in \{1, \ldots, K\}$, and as $n \to \infty$,

$$\sqrt{n}\,(\hat{\lambda}_{\alpha(n),k} - \lambda_k) = \sqrt{n}\,\langle \hat{\Pi}_{\alpha(n)} e_k, e_k \rangle + o_p(1) \tag{5.7}$$

and

$$\sqrt{n}\,\big(\hat{e}_{\alpha(n),k} - e_k\big) = \sqrt{n}\,Q_k \hat{\Pi}_{\alpha(n)} e_k + o_p(1). \tag{5.8}$$

Now define linear maps $\phi_k : \mathcal{L} \to \mathbb{H}$ such that for any $T \in \mathcal{L}$ and all $k \in \{1, \ldots, K\}$,

$$\phi_k T = Q_k T e_k, \tag{5.9}$$

where $\mathcal{L}$ is the Banach space of all bounded operators which map $\mathbb{H}$ into itself. It is clear that the $\phi_k$'s are bounded, and hence they are continuous. Then, define a linear map $\Phi_K : \mathcal{L} \to \mathbb{H}^K$ such that $\Phi_K = (\phi_1, \ldots, \phi_K)$, where $\mathbb{H}^K$ is the product space of $K$ copies of $\mathbb{H}$. Since the $\phi_k$'s are continuous, $\Phi_K$ is also a continuous map. We know from (5.3), (4.5), and Lemma 2 that $\sqrt{n}\,(\hat{\Pi}_{\alpha(n)})$ converges in distribution to a Gaussian random element $\mathbb{G}$ with mean zero. Hence, according to the Continuous Mapping Theorem,

$$\Phi_K(\sqrt{n}\,\hat{\Pi}_{\alpha(n)}) \rightsquigarrow \Phi_K \mathbb{G} = (Q_1 \mathbb{G} e_1, \ldots, Q_K \mathbb{G} e_K). \tag{5.10}$$

By (5.8) and (5.10),

$$\{\sqrt{n}\,(\hat{e}_{\alpha(n),k} - e_k) : 1 \le k \le K\} \rightsquigarrow (Q_1 \mathbb{G} e_1, \ldots, Q_K \mathbb{G} e_K) \text{ in } \mathbb{H}^K.$$

Now consider linear maps $\psi_k : \mathcal{L} \to \mathbb{R}$, $1 \le k \le K$ such that for any $T \in \mathcal{L}$, $\psi_k T = \langle T e_k, e_k \rangle$. Then define $\Psi_K = (\psi_1, \ldots, \psi_K)$ which is a linear map from $\mathcal{L}$ to $\mathbb{R}^K$. Since the $\psi_k$'s are continuous maps, $\Psi_k$ is also continuous. Thus, the following result follows from the Continuous Mapping Theorem:

$$\Psi_K(\sqrt{n}\,\Pi_{\alpha(n)}) \rightsquigarrow \Psi_K \mathbb{G} = (\langle \mathbb{G} e_1, e_1 \rangle, \ldots, \langle \mathbb{G} e_K, e_K \rangle). \tag{5.11}$$

Then, it follows from (5.7) and (5.11) that

$$\{\sqrt{n}\,\big(\lambda_{\alpha(n),k} - \lambda_k\big) : 1 \le k \le K\} \rightsquigarrow (\langle \mathbb{G} e_1, e_1 \rangle, \ldots, \langle \mathbb{G} e_K, e_K \rangle) \text{ in } \mathbb{R}^K.$$

$\square$

**Proof of Corollary 2.** We can get the following result similarly to (5.5): for all $k \in \{1, \ldots, K\}$,

$$\hat{\gamma}_{\alpha(n),k} = e_k + Q_k \hat{\Pi}_{\alpha(n)} e_k + \mathcal{O}(\|\hat{\Pi}_{\alpha(n)}\|_{\mathcal{L}}^2) \text{ on } \Omega_n.$$

Then (5.4) and (4.11) yield that, for all $k \in \{1, \ldots, K\}$, and as $n \to \infty$,

$$\sqrt{n}\left(\hat{\gamma}_{\alpha(n),k} - e_k\right) = \sqrt{n}\, Q_k \hat{\Pi}_{\alpha(n)} e_k + o_p(1). \tag{5.12}$$

By exploiting (5.9), (5.10), and (5.12), we complete this proof similarly to the proof of the first result of Proposition 1. □

## References


[1] A. Aguilera, M. Aguilera-Morillo, Penalized pca approaches for B-spline expansions of smooth functional data, Applied Mathematics and Computation 219 (2013) 7805–7819.

[2] G. Aneiros-Pérez, P. Vieu, Nonparametric time series prediction: A semi-functional partial linear modeling, Journal of Multivariate Analysis 99 (2008) 834–857.

[3] J.L. Bali, G. Boente, D.E. Tyler, J.L. Wang, Robust functional principal components: A projection-pursuit approach, The Annals of Statistics 39 (2011) 2852–2882.

[4] P. Besse, J.O. Ramsay, Principal components analysis of sampled functions, Psychometrika 51 (1986) 285–311.

[5] G. Boente, R. Fraiman, Kernel-based functional principal components, Statistics & Probability Letters 48 (2000) 335–345.

[6] H. Cardot, F. Ferraty, P. Sarda, Functional linear model, Statistics & Probability Letters 45 (1999) 11–22.

[7] P. Castro, W. Lawton, E. Sylvestre, Principal modes of variation for processes with continuous sample curves, Technometrics 28 (1986) 329–337.

[8] A. Cuevas, A partial overview of the theory of statistics with functional data, Journal of Statistical Planning and Inference 147 (2014) 1–23.

[9] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference, Journal of Multivariate Analysis 12 (1) (1982) 136–154.

[10] J.-C. Deville, Méthodes statistiques et numériques de l'analyse harmonique, Ann. INSEE 15 (1974) 3–101.

[11] N. Dunford, J.T. Schwartz, Linear Operators. Part 2: Spectral Theory. Self Adjoint Operators in Hilbert Space, Interscience Publishers, 1963.

[12] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: tTheory and Practice, Springer, Berlin, 2006.

[13] D.S. Gilliam, T. Hohage, X. Ji, F. Ruymgaart, The Fréchet derivative of an analytic function of a bounded operator with some applications, International Journal of Mathematics and Mathematical Sciences 2009.

[14] I. Gohberg, S. Goldberg, M. Kaashoek, Classes of Linear Operators, vol. 1 of Operator Theory: Advances and Applications, Birkhäuser Basel, 2013.

[15] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics, Journal of Multivariate Analysis 146 (2016) 1–6.

[16] P. Hall, M. Hosseini-Nasab, On properties of functional principal components analysis, Journal of the Royal Statistical Society: Series B 68 (2006) 109–126.

[17] P. Hall, H.-G. Müller, J.L. Wang, Properties of principal component methods for functional and longitudinal data analysis, The Annals of Statistics 34 (2006) 1493–1517.

[18] L. Horváth, P. Kokoszka, Inference for Functional Data With Applications, Springer, Berlin, 2012.

[19] T. Hsing, R. Eubank, Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators, Wiley, New York, 2015.

[20] J. Z. Huang, H. Shen, A. Buja, Functional principal components analysis via penalized rank one approximation, Electronic Journal of Statistics 2 (2008) 678–695.

[21] J. Hunter, B. Nachtergaele, Applied Analysis, World Scientific, 2001.

[22] S. Ingrassia, G.D. Costanzo, Functional principal component analysis of financial time series, in: New Developments in Classification and Data Analysis, Springer, 2005, pp. 351–358.

[23] A. Kneip, K.J. Utikal, Inference for density families using functional principal component analysis, Journal of the American Statistical Association 96 (2001) 519–542.

[24] P. Kokoszka, M. Reimherr, Asymptotic normality of the principal components of functional time series, Stochastic Processes and their Applications 123 (2013) 1546–1562.

[25] P. Lax, Functional Analysis, Pure and Applied Mathematics, Wiley, New York, 2002.

[26] L. Luo, Y. Zhu, M. Xiong, Smoothed functional principal component analysis for testing association of the entire allelic spectrum of genetic variation, European Journal of Human Genetics 21 (2013) 217–224.

[27] F.A. Ocaña, A.M. Aguilera, M. Escabias, Computational considerations in functional principal component analysis, Computational Statistics 22 (2007) 449–465.

[28] F.A. Ocaña, A.M. Aguilera, M.J. Valderrama, Functional principal components analysis by choice of norm, Journal of Multivariate Analysis 71 (1999) 262–276.

[29] X. Qi, H. Zhao, Some theoretical properties of Silverman's method for smoothed functional principal component analysis, Journal of Multivariate Analysis 102 (2011) 741–767.

[30] J.O. Ramsay, C. Dalzell, Some tools for functional data analysis, Journal of the Royal Statistical Society. Series B (1991) 539–572.

[31] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, 2nd ed., Springer, Berlin, 2005.

[32] P.T. Reiss, R.T. Ogden, Functional principal component regression and functional partial least squares, Journal of the American Statistical Association 102 (2007) 984–996.

[33] J.A. Rice, B.W. Silverman, Estimating the mean and covariance structure nonparametrically when the data are curves, Journal of the Royal Statistical Society: Series B 53 (1991) 233–243.

[34] F. Riesz, B.S. Nagy, Functional Analysis. (Translated from the 2nd French edition by LF Boron), 1955.

[35] B.W. Silverman, Smoothed functional principal components analysis by choice of norm, The Annals of Statistics 24 (1996) 1–24.

[36] F. Yao, T. Lee, Penalized spline models for functional principal component analysis, Journal of the Royal Statistical Society: Series B 68 (2006) 3–25.

[37] F. Yao, H.-G. Müller, Functional quadratic regression, Biometrika 97 (2010) 49–64.

[38] V.V. Zhikov, Weighted Sobolev spaces, Sbornik: Mathematics 189 (1998) 1139–1170.