Journal of Multivariate Analysis 102 (2011) 741–767

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Some theoretical properties of Silverman's method for Smoothed functional principal component analysis ☆

Xin Qi *, Hongyu Zhao

Department of Epidemiology and Public Health, Yale University, New Haven, CT 06520-8034, United States

ARTICLE INFO

Article history:
Received 25 March 2010
Available online 15 December 2010

AMS subject classifications:
primary 62G05
62H25
secondary 62G20
60F17

Keywords:
Functional PCA
Smoothing methods
Roughness penalty
Convergence rates
Functional central limit theorem
Asymptotic normality

ABSTRACT

Principal component analysis (PCA) is one of the key techniques in functional data analysis. One important feature of functional PCA is that there is a need for smoothing or regularizing of the estimated principal component curves. Silverman's method for smoothed functional principal component analysis is an important approach in a situation where the sample curves are fully observed due to its theoretical and practical advantages. However, lack of knowledge about the theoretical properties of this method makes it difficult to generalize it to the situation where the sample curves are only observed at discrete time points. In this paper, we first establish the existence of the solutions of the successive optimization problems in this method. We then provide upper bounds for the bias parts of the estimation errors for both eigenvalues and eigenfunctions. We also prove functional central limit theorems for the variation parts of the estimation errors. As a corollary, we give the convergence rates of the estimations for eigenvalues and eigenfunctions, where these rates depend on both the sample size and the smoothing parameters. Under some conditions on the convergence rates of the smoothing parameters, we can prove the asymptotic normalities of the estimations.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Principal component analysis (PCA) is one of the key techniques in multivariate analysis and functional data analysis. An important difference between classical PCA and functional PCA is that there is a need for smoothing or regularizing of the estimated principal component curves in functional PCA (see Chapter 9 in [12]). Many methods have been proposed to estimate the smoothed functional principal components when the sample curves are fully observed. A general overview of these methods and an extensive list of references can be found in [12]. The reader can find in Ferraty and Vieu [6] more discussions on theoretical aspects and nonparametric methods for functional data analysis. Functional PCA has many important applications. For example, functional principal component regression (see for instance [2]) is a direct application of functional principal component analysis.

The approach proposed in Silverman [15] is an important method for smoothing functional PCA (see Chapter 9 in [12]) due to its theoretical and practical advantages. First, the weak assumptions underlying this method make it applicable to data from many fields. Silverman [15] did not make any assumptions on the mean curves and sample curves. Hence, in addition to data with smooth random curves, this method can be applied to analyze data where the sample curves can be unsmooth or even discontinuous, such as those encountered in financial engineering, survival analysis and other fields. For covariance functions, Silverman [15] only assumed that they have series expansions by their eigenfunctions without imposing a smoothing constraint. This is attractive because the covariance functions are continuous but unsmooth in many

☆ Supported in part by NIH grants R01 GM59507, a pilot project from the Yale Pepper Center, and NSF grant DMS 0714817.
* Corresponding author.
E-mail addresses: xin.qi@yale.edu (X. Qi), hongyu.zhao@yale.edu (H. Zhao).

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.12.001

important models such as stochastic differential equation models in financial engineering and counting process models in survival analysis. Second, Silverman's method controls the smoothness of eigenfunction curves by directly imposing roughness penalties on these functions instead of on sample curves or covariance functions. Furthermore, this approach changes the eigenvalue and eigenfunction problems in the usual L^2 space to problems in another Hilbert space, the Sobolev space (with a norm different from the usual norm in the Sobolev space). Therefore, many powerful tools from the theory of Hilbert space can be employed to study the properties of this method. Third, this approach incorporates the smoothing step into the step for computing eigenvalues and eigenfunctions. Therefore, this method is computationally efficient with the same computational load as the usual unsmoothed functional PCA. Fourth, the estimates produced by this method are invariant under scale transformations. As pointed out by Huang et al. [8], the invariance property under scale transformations should be a guiding principle in introducing roughness penalties to functional PCA.

Despite all these advantages, lack of knowledge about the theoretical properties of this method makes it difficult to generalize it to the situations where the sample curves are only observed at discrete time points. Silverman [15] only proved consistency of the estimations as the sample size goes to infinity and the smoothing parameter goes to zero. Even the existence of the solutions to the successive optimization problems in this method is not established. It is not clear how the estimation errors depend on the sample size and the smoothing parameter. Asymptotic normalities of the estimations also need to be proved. In this paper, we aim to solve these open problems. In Section 2, we give the detailed background, basic notations and our main assumptions. In Section 3, Silverman's method is introduced and the existence theorem for the successive optimization problems is proven. Our main results appear in Section 4. Section 5 contains detailed proofs of our theorems.

2. Notations and main assumptions

We introduce notations and definitions used throughout the paper. Let $\mathbb{N}$ denote the collection of all the positive integers. We consider a finite time interval $[a, b]$. In this paper, we will mainly consider functions in the following two spaces, the L^2 space

$$L^2([a, b]) = \left\{ f : f \text{ is a measurable function on } [a, b] \text{ and } \int_a^b |f(t)|^2 \mathrm{d}t < \infty \right\},$$

and the Sobolev space

$$W_2^2([a, b]) = \{ f : f, f' \text{ are absolutely continuous on } [a, b] \text{ and } f'' \in L^2([a, b]) \},$$

where f' and f'' denote the first and second derivatives of f, respectively. For any $f, g \in L^2([a, b])$, define the usual inner product

$$(f, g) = \int_a^b f(t)g(t)\mathrm{d}t,$$

with corresponding squared norm $\|f\|^2 = (f, f)$. Given a smoothing parameter $\alpha > 0$, for any $f, g \in W_2^2([a, b])$, define

$$[f, g] = \int_a^b f''(t)g''(t)\mathrm{d}t$$

and the inner product

$$(f, g)_\alpha = (f, g) + \alpha[f, g]$$

with corresponding squared norm $\|f\|_\alpha^2 = (f, f)_\alpha$. Note that if $\alpha = 0$, we return to the $L^2([a, b])$ space. For any bounded operator B from $L^2([a, b])$ to $L^2([a, b])$, define the norm

$$\|B\| = \sup\{\|Bf\| : f \in L^2([a, b]) \text{ and } \|f\| \le 1\}. \tag{2.1}$$

For any measurable function $A(s, t)$ on $[a, b] \times [a, b]$, if

$$\int_a^b \int_a^b A^2(s, t)\mathrm{d}s\mathrm{d}t < \infty,$$

then $f \to \int_a^b A(s, t)f(t)\mathrm{d}t$ defines a bounded operator from $L^2([a, b])$ to $L^2([a, b])$. To simplify the notation, we just use A to denote this operator, that is

$$Af(s) = \int_a^b A(s, t)f(t)\mathrm{d}t,$$

and we have

$$\|A\| \le \left(\int_a^b \int_a^b A^2(s, t)\mathrm{d}s\mathrm{d}t \right)^{\frac{1}{2}}.$$

Let $X(t), a \le t \le b$ be a measurable stochastic process on $[a, b]$. Under Assumption 1, $X(t) \in L^2([a, b])$ a.s. Let $\{X_1(t), X_2(t), \dots, X_n(t)\}$ be i.i.d. sample curves from the distribution of $X(t)$. Assume that $EX(t) = \nu(t)$. Define Γ to be the covariance function

$$\Gamma(s, t) = E[(X(s) - \nu(s))(X(t) - \nu(t))], \quad \forall s, t \in [a, b],$$

and $\hat{\Gamma}_n$ to be the sample covariance function

$$\hat{\Gamma}_n(s, t) = \frac{1}{n}\sum_{p=1}^{n}(X_p(s) - \bar{X}(s))(X_p(t) - \bar{X}(s)), \quad \forall s, t \in [a, b],$$

where $\bar{X}$ is the sample's mean curve

$$\bar{X}(t) = \frac{1}{n}(X_1(t) + \cdots + X_n(t)).$$

We will give our basic assumptions below. Silverman [15] made three assumptions in Section 5.2 in order to prove the consistency result. Our assumptions are stronger than those in [15].

Assumption 1.

$$E[\|X\|^4] = E\left[\left(\int_a^b |X(t)|^2 dt\right)^2\right] < \infty. \tag{2.2}$$

Remark. (1) This assumption is stronger than the first assumption in Section 5.2 of [15]. Under condition (2.2), the central limit theorem for sample covariance function holds (see Section 2 in [3] and Chapter 10 in [10]).

(2) Assumption 1 is satisfied by many stochastic processes used in applications. For example, if $X(t)$ is a bounded process, it is obvious that (2.2)is true. Gaussian processes are an important class of stochastic processes which are widely used in statistics and other areas. Suppose that $X(t)$ is a Gaussian process with mean zero. Then

$$\begin{aligned} E[\|X\|^4] &= E\left[\left(\int_a^b |X(t)|^2 dt\right)^2\right] = \int_a^b\int_a^b E[X(t)^2X(s)^2]dtds \\ &= \int_a^b\int_a^b [\Gamma(s, s)\Gamma(t, t) + 2\Gamma(s, t)^2]dsdt \le \int_a^b\int_a^b 3\Gamma(s, s)\Gamma(t, t)dsdt \\ &= 3\left[\int_a^b \Gamma(t, t)dt\right]^2. \end{aligned}$$

Hence if $\Gamma(t, t)$ is integrable in $[a, b]$, which is satisfied by Gaussian processes commonly encountered in applications, (2.2) is true. Now let us consider the standard Brownian motion, the most widely studied Gaussian process. For the standard Brownian motion, $\Gamma(t, t) = t$, hence Assumption 1 is satisfied. It is well-known that its sample paths are continuous and nowhere differentiable almost surely. For non-Gaussian processes, let us consider a Poisson process with rate 1 in $[0, 1]$. Its sample paths are step functions only taking integer values and hence discontinuous. It is easy to verify that Assumption 1 is satisfied by Poisson processes.

(3) Under condition (2.2), we have

$$\begin{aligned} \int_a^b\int_a^b \Gamma(s, t)^2 dsdt &= \int_a^b\int_a^b (E[(X(s) - \nu(s))(X(t) - \nu(t))])^2 dsdt \\ &= \int_a^b\int_a^b (EX(t)X(s) - \nu(s)\nu(t))^2 dsdt \\ &\le \int_a^b\int_a^b 2(EX(t)X(s))^2 + 2\nu(s)^2\nu(t)^2 dsdt \\ &\le \int_a^b\int_a^b 2EX^2(t)X^2(s) + 2\nu(s)^2\nu(t)^2 dsdt \le 4E[\|X\|^4] < \infty. \end{aligned}$$

Therefore, the operator Γ is a Hilbert–Schmidt operator, hence it is a compact operator (see Section XI.6 in [5] or Section 97 in [13]). It follows that the set of eigenvalues of this operator are bounded and at most countable with at most one limit point at 0. Because the covariance operator Γ is always nonnegative-definite, all the eigenvalues are nonnegative. Let $\lambda_1 \ge \lambda_2 \ge \cdots \ge 0$ be the collection of all eigenvalues and the corresponding eigenfunctions are $\gamma_1, \gamma_2, \dots$. Every eigenfunction has been scaled to have L^2-norm 1. The set of all the eigenfunctions forms an orthonormal basis of $L^2([a, b])$. Furthermore, we have decomposition

$$\Gamma(s,t) = \sum_{j=1}^{\infty} \lambda_j \gamma_j(s)\gamma_j(t), \tag{2.3}$$

the series on the right-hand side converges in the L^2 sense. If Γ is a continuous function, the series on the right-hand side absolutely and uniformly converges. Although Silverman [15] did not assume that Γ is square integrable, he assumed the decomposition form of (2.3).

(4) We have

$$\Gamma\gamma_j = \lambda_j\gamma_j, \quad \forall j = 1, 2, \ldots.$$

(5) By (2.2), $X(s)$ is square integrable a.s. Hence, the sample covariance functions $\hat{\Gamma}_n$ satisfies

$$\int_a^b \int_a^b \hat{\Gamma}_n(s,t)^2 \mathrm{d}s\mathrm{d}t < \infty$$

a.s. Then we have that the eigenvalues $\hat{\lambda}_1 \ge \hat{\lambda}_2 \ge \cdots \ge 0$ since the operator $\hat{\Gamma}_n$ is nonnegative-definite. The corresponding eigenfunctions $\hat{\gamma}_j, j \in \mathbb{N}$ satisfying

$$\hat{\Gamma}_n\hat{\gamma}_j = \hat{\lambda}_j\hat{\gamma}_j, \quad \forall j = 1, 2, \ldots.$$

Suppose that we are interested in estimating the first K eigenvalues and eigenfunctions of Γ.

Assumption 2. Any eigenvalue λ_j, $1 \le j \le K$ has multiplicity 1, so that

$$\lambda_1 > \lambda_2 > \cdots > \lambda_K > \lambda_{K+1}.$$

Remark. This assumption is just the third assumption in Section 5.2 of [15]. If an eigenvalue has multiplicity 1, then the corresponding eigenfunction is uniquely determined up to a sign. If the multiplicity is larger than 1, the eigenfunctions cannot be uniquely determined up to a sign.

Assumption 3. The eigenfunctions γ_j, $1 \le j \le K$ belong to $W_2^2([a, b])$.

Remark. (1) This assumption is the second assumption in Section 5.2 of [15] and is essential in our paper.

(2) If the covariance function Γ satisfies some smoothness conditions, then Assumption 3 is true. For example, suppose that $\Gamma(s, t)$, $\frac{\partial\Gamma(s,t)}{\partial s}$ and $\frac{\partial^2\Gamma(s,t)}{\partial s^2}$ are all continuous on $[a, b] \times [a, b]$ (hence they are bounded and square integrable), one can easily verify that

$$\lambda_k\gamma_k''(s) = \int_a^b \frac{\partial^2\Gamma(s,t)}{\partial s^2}\gamma_k(t)\mathrm{d}t \quad \forall 1 \le k \le K.$$

Hence, by Cauchy–Schwarz inequality and $\|\gamma_k\| = 1$, we have

$$\lambda_k^2 \int_a^b (\gamma_k''(s))^2 \mathrm{d}s \le \int_a^b \int_a^b \left(\frac{\partial^2\Gamma(s,t)}{\partial s^2}\right)^2 \mathrm{d}s\mathrm{d}t < \infty \quad \forall 1 \le k \le K.$$

(3) There are many important random processes whose covariance matrices are not smooth, but the eigenfunctions corresponding to nonzero eigenvalues belong to $W_2^2([a, b])$. The simplest examples are standard Brownian motion and the Poisson process with rate 1 in time interval $[0, 1]$. Their covariance functions are the same and equal to $\min(s, t)$, $0 \le s, t \le 1$ (see Page 89 in the book [7]). The eigenvalues and eigenfunctions are

$$\lambda_j = \left(\frac{2}{(2j-1)\pi}\right)^2, \qquad \gamma_j = \sqrt{2}\sin\left(\frac{(2j-1)\pi t}{2}\right), \quad j = 1, 2, \ldots. \tag{2.4}$$

The next example is the famous Black–Scholes Model in finance. Let S_t denote the price of a stock at time t. Then S_t satisfies the following SDE,

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma S_t \mathrm{d}W_t,$$

where μ is the instantaneous mean return, σ is the instantaneous return volatility and W_t is a Brownian motion. The covariance function of S_t is smooth except at the points on the diagonal line $\{(s, t) : s = t\}$. The same is true for the following example. Consider the counting processes model in survival analysis. Let N_t be the number of the occurrences of the event in $[0, t]$. Then N_t satisfies

$$\mathrm{d}N_t = \lambda(t)\mathrm{d}t + \mathrm{d}M_t,$$

where $\lambda(t)$ is a smooth intensity function and M_t is a martingale.

Silverman [15] introduced a "half-smoothing" operator which plays an important role in this paper. We give a strict definition of this operator here. We first define an unbounded operator L in $L^2([a, b])$. The domain of L is

$$\mathcal{D}(L) = \{f \in L^2([a, b]) : f, f' \text{ are absolutely continuous and } f'' \in L^2([a, b])\},$$

and for any $f \in \mathcal{D}(L)$,

$$Lf = f''.$$

Then L is a closed but unbounded operator and $\mathcal{D}(L)$ is dense in $L^2([a, b])$ (for the definition of closed operators, see Chapter VIII of [13] or Chapter 13 of [14]). Let L^* be the adjoint operator of L. By the theorem in Section 118 of [13] or Theorem 13.13 in [14], $(I + \alpha L^*L)^{-1}$ is a bounded, positive self-adjoint operator with norm less than or equal to 1, where $\alpha \geq 0$ is the smoothing parameter. Now it follows from Theorems 12.33 and 13.31 in [14] that $(I + \alpha L^*L)^{-1}$ has a unique positive and self-adjoint square root S_α with norm less than or equal to 1 which is the "half-smoothing" operator in [15]. Therefore,

$$S_\alpha^2 = (I + \alpha L^*L)^{-1}, \tag{2.5}$$

and by Theorem 13.11(b) in [14], the inverse S_α^{-1} exists and is self-adjoint because $(I + \alpha L^*L)^{-1}$ is invertible.

3. Silverman's approach to smoothed functional PCA

In this section, we always assume that the independent sample curves

$$\{X_1(t), X_2(t), \ldots, X_n(t) : a \leq t \leq b\}$$

are entirely observed. We first consider the usual population functional principal components. The first population functional principal component is defined as the linear functional $\ell_1(X)$ of X which maximizes

$$\mathrm{Var}(\ell(X))$$

over all nonzero linear functionals ℓ in $L^2([a, b])$ with the norm $\|\ell\|$=1. The second population functional principal component is defined as the linear functional $\ell_2(X)$ of X which maximizes

$$\mathrm{Var}(\ell(X))$$

over all linear functional ℓ with the norm $\|\ell\| = 1$ and uncorrelated with $\ell_1(X)$. Similarly, we can define all the other population functional principal components, $\ell_3(X), \ldots$. Because X takes values in $L^2([a, b])$ which is a real Hilbert space, by the Riesz representation theorem, for any bounded linear functional ℓ, there is a unique $\gamma \in L^2([a, b])$ such that for any $f \in L^2([a, b])$,

$$\ell(f) = (\gamma, f) \quad \text{and} \quad \|\ell\| = \|\gamma\|.$$

Hence there exists $\gamma_j \in L^2([a, b]), j \in \mathbb{N}$, with $\|\gamma_j\| = 1$, such that the population functional principal components $\ell_j(X) = (\gamma_j, X), j \in \mathbb{N}$. γ_j is called the j-th principal component weight function or j-th principal component curve. Because

$$\mathrm{Var}(\ell_j(X)) = \mathrm{Var}(\gamma_j, X) = (\gamma_j, \Gamma\gamma_j), \quad \forall j \in \mathbb{N},$$

γ_1 is the solution of the following optimization problem,

$$\max_{\|\gamma\|=1} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2}. \tag{3.1}$$

The maximum value of (3.1) is just the largest eigenvalue λ_1 of Γ and γ_1 is the corresponding eigenfunction (see Section 2, Chapter 3 in [16]). γ_2 is the solution of the optimization problem,

$$\max_{\|\gamma\|=1, (\gamma,\gamma_1)=0} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2}. \tag{3.2}$$

The maximum value of (3.2) is just the second eigenvalue λ_2 of Γ and γ_2 is the corresponding eigenfunction. Similarly, γ_j is the eigenfunction corresponding to the eigenvalue λ_j which is also the variance of the j-th principal component.

Because the covariance function Γ is usually unknown, we cannot obtain the population principal component weight functions directly. Hence, people use the sample covariance function $\hat{\Gamma}_n$ to estimate Γ and use the eigenvalues and eigenfunctions of $\hat{\Gamma}_n$ to estimate the eigenvalues and eigenfunctions of Γ. We call them non-smooth estimators. However, the non-smooth principal component curves can show substantial variability (see Chapter 9 in [12]). There is a need for smoothing of the estimated principal component weight functions.

Silverman [15] (see also Chapter 9 in [12]) proposed a method of incorporating smoothing by replacing the usual L^2 norm with a norm that takes the roughness of the functions into account. Let α be a nonnegative smoothing parameter. Define the estimators $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ of $\{(\lambda_j, \gamma_j) : j \in \mathbb{N}\}$ to be the solutions of the following successive optimization problems. First, $\hat{\gamma}_1^{[\alpha]}$ is the solution of the optimization problem

$$\max_{\|\gamma\|=1} \frac{(\gamma, \hat{\Gamma}_n \gamma)}{(\gamma, \gamma) + \alpha[\gamma, \gamma]} = \max_{\|\gamma\|=1} \frac{(\gamma, \hat{\Gamma}_n \gamma)}{\|\gamma\|_\alpha^2}. \tag{3.3}$$

Let $\hat{\lambda}_1^{[\alpha]}$ be the maximum value of (3.3). For any $k \in \mathbb{N}$, if we have obtained $\{\hat{\gamma}_j^{[\alpha]}, j = 1, 2, \ldots, k-1\}$ and $\{\hat{\lambda}_j^{[\alpha]}, j = 1, 2, \ldots, k-1\}$, $\hat{\gamma}_k^{[\alpha]}$ is the solution of the optimization problem

$$\max_{\substack{\|\gamma\|=1, (\gamma, \hat{\gamma}_j^{[\alpha]})_\alpha = 0, \\ j=1,\ldots,k-1}} \frac{(\gamma, \hat{\Gamma}_n \gamma)}{\|\gamma\|_\alpha^2}, \tag{3.4}$$

and $\hat{\lambda}_k^{[\alpha]}$ is the maximum value of (3.4). Note that $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ depends on both the sample size n and the smoothing parameter α.

First of all, we need to show that the solutions $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ of the successive optimization problems (3.3) and (3.4) exist.

Theorem 3.1. *Under Assumption 1, the solutions $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ of the successive optimization problems (3.3) and (3.4) exist for any $\alpha \geq 0$ almost surely. Moreover, we have, for any $\gamma \in W_2^2([a, b])$ and $j \in \mathbb{N}$,*

$$(\hat{\Gamma}_n \hat{\gamma}_j^{[\alpha]}, \gamma) = \hat{\lambda}_j^{[\alpha]} (\hat{\gamma}_j^{[\alpha]}, \gamma)_\alpha. \tag{3.5}$$

Similarly, define $\{(\lambda_j^{[\alpha]}, \gamma_j^{[\alpha]}) : j \in \mathbb{N}\}$ to be the solutions of the successive optimization problems (3.3) and (3.4) with $\hat{\Gamma}_n$ replaced by Γ. Similarly, we have the following equalities for Γ and $\{(\lambda_j^{[\alpha]}, \gamma_j^{[\alpha]}) : j \in \mathbb{N}\}$

$$(\Gamma \gamma_j^{[\alpha]}, \gamma) = \lambda_j^{[\alpha]} (\gamma_j^{[\alpha]}, \gamma)_\alpha, \quad \forall j \in \mathbb{N}, \gamma \in W_2^2([a, b]). \tag{3.6}$$

Note that

$$\gamma_j^{[0]} = \gamma_j, \qquad \lambda_j^{[0]} = \lambda_j, \qquad \hat{\gamma}_j^{[0]} = \hat{\gamma}_j, \quad \lambda_j^{[0]} = \hat{\lambda}_j, \quad \forall j \in \mathbb{N}.$$

Theorem 1 in [15] gives the consistency of the estimators

$$\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$$

as $\alpha \to 0$ and $n \to \infty$.

4. Asymptotic theory

Fix a positive integer K. We will assume throughout this section that we want to estimate the first K principal component curves. For any $1 \leq k \leq K$, define

$$L_k = \max_{1 \leq j \leq k} \sqrt{[\gamma_j, \gamma_j]}.$$

Then under Assumption 3, L_k is finite and is a measure of roughness of the first k eigenfunctions of Γ. For standard Brownian motion and the Poisson process with rate 1 (see remark (3) after Assumption 3),

$$L_k = \left(\frac{(2k-1)\pi}{2}\right)^2, \quad k = 1, 2, \ldots.$$

For any $1 \leq k \leq K$, we have decompositions

$$\hat{\lambda}_k^{[\alpha]} - \lambda_k = (\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}) + (\lambda_k^{[\alpha]} - \lambda_k), \tag{4.1}$$

$$\hat{\gamma}_k^{[\alpha]} - \gamma_k = (\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}) + (\gamma_k^{[\alpha]} - \gamma_k). \tag{4.2}$$

The last terms $\lambda_k^{[\alpha]} - \lambda_k$, $\gamma_k^{[\alpha]} - \gamma_k$ on the right-hand sides of both (4.1) and (4.2) are nonrandom. They are the "bias terms" due to the introduction of α. We will give the upper bounds for norms of these terms. The first terms on the right-hand sides of both (4.1) and (4.2) are the "variation terms" due to the randomness of the sample curves. We will prove a functional central limit theorem for these terms. In order to avoid any confusion it should be pointed out that (4.1) and (4.2) are not the bias-variance decompositions in the strict sense because $\lambda_k^{[\alpha]}$ and $\gamma_k^{[\alpha]}$ are not the expectations of $\hat{\lambda}_k^{[\alpha]}$ and $\hat{\gamma}_k^{[\alpha]}$ respectively. Since it is hard to express or characterize the exact expectations of $\hat{\lambda}_k^{[\alpha]}$ and $\hat{\gamma}_k^{[\alpha]}$, the asymptotic properties of the usual bias and variation terms in the strict sense may not be easily studied. Heuristic calculations of the usual bias and variation terms in the strict sense were performed in Section 6 of [15].

Note that even if the multiplicity of λ_k is one, we cannot uniquely determine γ_k because $-\gamma_k$ is also an eigenfunction. In the following theorem, by "Given γ_k", we mean that not only γ_k is an eigenfunction, but also the direction of γ_k is given.

Define

$$\alpha_0 = \min_{1\le k\le K}\left\{\min\left\{\frac{\sqrt{1+\frac{2k(\lambda_{k-1}-\lambda_k)^2}{(k-1)\lambda_k\|\Gamma\|}}-1}{2kL_k^2}, \frac{\lambda_k-\lambda_{k+1}}{(8\sqrt{k}+16k)L_k^2\lambda_k}, \frac{(\lambda_{k-1}-\lambda_k)\left\{1+\frac{2\|\Gamma\|}{\lambda_k-\lambda_{k+1}}\right\}^{-\frac{1}{2}}}{4\sqrt{2k(k-1)}L_k^2\lambda_k}\right\}\right\}. \tag{4.3}$$

Theorem 4.1. *Under Assumptions 1–3, for any $1 \le k \le K$ and $0 \le \alpha \le \alpha_0$,*

$$\begin{aligned} 0 &\le \lambda_k - \lambda_k^{[\alpha]} \\ &\le \sqrt{2}\sqrt{k}L_k^2\lambda_k\alpha\left(1+O\left(\frac{\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}\alpha+\frac{k(k-1)L_k^4\lambda_k^2\|\Gamma\|}{(\lambda_{k-1}-\lambda_k)^2(\lambda_k-\lambda_{k+1})}\alpha^2\right)\right). \end{aligned} \tag{4.4}$$

Given γ_k, $1 \le k \le K$, we can uniquely choose $\gamma_k^{[\alpha]}$ for each $\alpha \in [0, \alpha_0]$ such that $\gamma_k^{[\alpha]}$ is a continuous function of α and $(\gamma_k^{[\alpha]}, \gamma_k) > 0$ for all $0 \le \alpha \le \alpha_0$, and we have

$$\|\gamma_k^{[\alpha]}-\gamma_k\| \le \sqrt{\alpha}\sqrt{\frac{4\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}}+\alpha\sqrt{4k(k-1)L_k^4\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2\left\{1+\frac{2\|\Gamma\|}{\lambda_k-\lambda_{k+1}}\right\}}. \tag{4.5}$$

Remark. (1) If K is fixed or bounded, we have

$$0 \le \lambda_k - \lambda_k^{[\alpha]} \le \sqrt{2}\sqrt{k}L_k^2\lambda_k\alpha + o(\alpha),$$
$$\|\gamma_k^{[\alpha]}-\gamma_k\| \le \sqrt{\alpha}\sqrt{\frac{4\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}}+o(\sqrt{\alpha}).$$

Hence, the convergence rates for eigenvalues and eigenfunctions are different. Eigenvalues have faster convergence rates than eigenfunctions.

(2) As $K \to \infty$, we have $\alpha_0 \to 0$. If we choose α in such a way that $0 \le \alpha \le \alpha_0$ and the right-hand sides of (4.4) and (4.5) converge to zero, then $\lambda_k^{[\alpha]} \to \lambda_k$ and $\gamma_k^{[\alpha]} \to \gamma_k$ for all $1 \le k \le K$.

(3) The convergence rates for both eigenvalues and eigenfunctions depend on L_k. If the eigenfunctions are less smooth, that is, L_k is large, then the convergence is slow.

(4) (4.4) and (4.5) give the upper bounds. However, the lower bounds are 0 for any $k \in \mathbb{N}$. Here is a simple example. Without loss of generality, let $k = 2$. Suppose $[a, b] = [0, 2\pi]$,

$$\begin{aligned}\Gamma(s,t) = {}& \frac{2}{\pi}\cos(s)\cos(t)+\frac{1}{2\pi}+\frac{1}{2\pi}\sin(s)\sin(t)+\frac{1}{\pi}\sum_{m=2}^{\infty}\left(\frac{1}{2m}\right)^3\cos(ms)\cos(mt) \\ &+\frac{1}{\pi}\sum_{m=2}^{\infty}\left(\frac{1}{2m+1}\right)^3\sin(ms)\sin(mt).\end{aligned}$$

Note that the right-hand side in the above equality converges both uniformly and in $L^2([0, 2\pi] \times [0, 2\pi])$ to a strictly positive definite covariance functions. Its first eigenvalue and eigenfunction are 2 and $\frac{1}{\sqrt{\pi}}\cos(t)$, the second ones are 1 and $\frac{1}{\sqrt{2\pi}}$. It is interesting to note that the eigenfunctions of Γ are the same as the solutions of the successive optimization problems (3.3) and (3.4). The first maximum value of the successive optimization problems (3.3) and (3.4) is $\frac{2}{1+\alpha}$ and the second one is still 1. That is, in this case, we have $\lambda_2^{[\alpha]} = \lambda_2$ and $\gamma_2^{[\alpha]} = \gamma_2$ for any α, hence the lower bounds are zeros.

Define $C_{\mathbb{R}}[0, \alpha_0]$ to be the normed space of all continuous real functions in $[0, \alpha_0]$ equipped with norm $\sup_{0\le\alpha\le\alpha_0} |\cdot|$. Let $\prod_{1\le j\le K} C_{\mathbb{R}}[0, \alpha_0]$ denote the product space of K copies of $C_{\mathbb{R}}[0, \alpha_0]$. Define $C_{L^2([a,b])}[0, \alpha_0]$ to be the normed space of all continuous functions in $[0, \alpha_0]$ taking values in $L^2([a, b])$ equipped with norm $\sup_{0\le\alpha\le\alpha_0} \|\cdot\|$. Similarly, we define $\prod_{1\le j\le K} C_{L^2([a,b])}[0, \alpha_0]$.

For each $1 \le k \le K$ and each n, we will view $\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]})$ as a stochastic process with index $\alpha \in [0, \alpha_0]$ and values in $L^2[a, b]$ and view $\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]})$ as a stochastic process with index $\alpha \in [0, \alpha_0]$ and values in $\mathbb{R}$. However, in the following subset in the probability space,

$$\Omega_0 = \{\omega : \text{ there exists at least one } \alpha \in [0, \alpha_0] \text{ such that } \hat{\lambda}_1^{[\alpha]}, \ldots, \hat{\lambda}_K^{[\alpha]} \text{ are not mutually different } \}, \tag{4.6}$$

$\hat{\gamma}_1^{[\alpha]}, \ldots, \hat{\gamma}_K^{[\alpha]}$ are not uniquely determined up to signs. We will show that Ω_0 is measurable and its probability goes to zero as $n \to \infty$ in the proof of the following theorem. Hence, how to define $\hat{\gamma}_1^{[\alpha]}, \ldots, \hat{\gamma}_K^{[\alpha]}$ in Ω_0 does not affect our asymptotic results. In order to make the development of our theory easier, we will use the following definition

$$\text{in } \Omega_0, \text{ define } \hat{\gamma}_k^{[\alpha]} = 0, \quad 1 \le k \le K. \tag{4.7}$$

Theorem 4.2. *Under Assumptions 1–3 and the definition (4.7), we can properly choose $\hat{\gamma}_k^{[\alpha]}$ in Ω_0^c to make the sequence*

$$\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K, 0 \le \alpha \le \alpha_0\}_n \tag{4.8}$$

of stochastic processes is measurable and has sample paths in

$$\prod_{k=1}^{K} C_{L^2([a,b])}[0, \alpha_0]$$

a.s. Furthermore, the sequence converges in distribution to a Gaussian random element with values in $\prod_{k=1}^{K} C_{L^2([a,b])}[0, \alpha_0]$ and mean zero. Similarly, the sequence

$$\{\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}), 1 \le k \le K, 0 \le \alpha \le \alpha_0\}_n \tag{4.9}$$

of stochastic processes has sample paths in $\prod_{k=1}^{K} C_{\mathbb{R}}[0, \alpha_0]$ a.s. and converges in distribution to a Gaussian random element with values in $\prod_{k=1}^{K} C_{\mathbb{R}}[0, \alpha_0]$ and mean zero.

Remark. (1) Recall the definition of Gaussian random elements in a separable Banach space. Suppose that X is a random element with values in a Banach space B with mean zero. Then X is a Gaussian element if for any bounded linear functional f, $f(X)$ is a Gaussian random variable. If X is a Gaussian random element, we can define its covariance operator Q. Q is a bounded operator from the dual space B' to B such that for any $f, g \in B'$, $g(Qf) = E[f(X)g(X)]$. Note that the distribution of a Gaussian element with values in a Banach space and mean zero is determined by its covariance operator. For further properties of Gaussian random elements in Banach spaces, see [10].

(2) The covariance operators (4.8) and (4.9) can be characterized by the "half-smoothing" operator S_α defined in (2.5) and the limit distribution of $\sqrt{n}(\hat{\Gamma}_n - \Gamma)$. However, the characterization involves some technical definitions. The reader can find the characterization in the proof of this theorem.

(3) The measurabilities and a.s. continuities of the sample paths of the processes (4.8) and (4.9) are not obvious at all.

(4) The convergences of (4.8) and (4.9) are weak convergences of probability measures in spaces $\prod_{k=1}^{K} C_{\mathbb{R}}[0, \alpha_0]$ and $\prod_{k=1}^{K} C_{L^2([a,b])}[0, \alpha_0]$, which are stronger than the convergences of only the marginal distributions of (4.8) and (4.9).

Now from Theorems 4.1 and 4.2, we have the following corollaries.

Corollary 4.1. *Under Assumptions 1–3, for any $1 \le k \le K$ and $0 \le \alpha \le \alpha_0$,*

$$\begin{aligned} |\hat{\lambda}_k^{[\alpha]} - \lambda_k| &\le |\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}| + \sqrt{2}\sqrt{k}L_k^2\lambda_k\alpha + o(\alpha), \\ \|\hat{\gamma}_k^{[\alpha]} - \gamma_k\| &\le |\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}| + \sqrt{\alpha}\sqrt{\frac{4\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k - \lambda_{k+1}}} + o(\sqrt{\alpha}) \end{aligned} \tag{4.10}$$

where

$$\sup_{0 \le \alpha \le \alpha_0} |\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}| = O_p\left(\frac{1}{\sqrt{n}}\right), \qquad \sup_{0 \le \alpha \le \alpha_0} |\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}| = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Remark. From Corollary 4.1, it seems that smoothing (that is, $\alpha > 0$) is unnecessary since when $\alpha = 0$, we get the best order $\frac{1}{\sqrt{n}}$. We clarify this problem by the following remarks.

(1) Both Silverman [15] and this paper consider the ideal situation where every sample curve is observed at all points in $[a, b]$ without any noise or measurement error. Although in this situation the estimates are consistent when $\alpha = 0$, smoothing is advantageous.

– First, because the "bias terms" and the "variation terms" are not the bias and the variation in the strict sense, they are correlated. Since the upper bounds on the right-hand sides of (4.10) are the sums of the upper bounds for bias terms and variation terms, the upper bounds in (4.10) are actually for the cases in which bias terms and variation terms are positively correlated. They are the worst cases when we introduce smoothing. In some cases such as those in Section 6.3 of [15], the mean squared errors for some $\alpha > 0$ are less than those for $\alpha = 0$. For these cases, it is possible that bias terms and variation terms are negatively correlated and hence the estimate errors should be much less than the upper bounds in (4.10). Section 6.4 of [15] gave an optimal α with order $O\left(\frac{1}{n}\right)$ for estimates of eigenfunctions. By Corollary 4.1, if we choose the optimal α, we obtain the best asymptotic rates $O\left(\frac{1}{\sqrt{n}}\right)$. Even for the worst cases, if we take $\alpha = O\left(\frac{1}{n}\right)$, we can obtain the rate $O\left(\frac{1}{\sqrt{n}}\right)$.
– Second, from a practical viewpoint, it is desirable that the estimates of principal component curves can keep the main patterns of the true principal component curves. However, the sample curves of many stochastic process are non-smooth or even discontinuous, such as examples in remark (3) after Assumption 3. Hence, their sample covariance functions have many local variations and so do the eigenfunctions of those sample covariance functions. In these cases, the local variations can be removed by using an appropriate amount of smoothing, that is, choosing an appropriate positive α.

(2) In practice, people cannot observe the entire sample curves. The observations can only be made at discrete points often with noise or measurement error. The observation points could be dense or sparse. If the sample curves are smooth and the observation points are dense, we can obtain smoothed estimate of each sample function and perform the usual functional PCA. This method cannot be applied to other situations. However, Silverman's method can be generalized to all these situations (see [11]). In our generalization, smoothing is essential and the smoothing parameters must be positive. The theoretical results in this paper has been applied to prove the consistency results in [11].

If α goes to 0 fast enough as $n \to \infty$, we have the following asymptotic normalities.

Corollary 4.2. *Under Assumptions 1–3, for any sequence $\{\alpha_n, n \geq 1\}$ with $\alpha_n = o_p\left(\frac{1}{\sqrt{n}}\right)$, the joint distributions of*

$$\{\sqrt{n}(\hat{\lambda}_1^{[\alpha_n]} - \lambda_1), \sqrt{n}(\hat{\lambda}_2^{[\alpha_n]} - \lambda_2), \ldots, \sqrt{n}(\hat{\lambda}_K^{[\alpha_n]} - \lambda_K)\}$$

converge to the same Gaussian distribution with mean zero. For any sequence $\{\alpha_n, n \geq 1\}$ with $\alpha_n = o_p\left(\frac{1}{n}\right)$, the joint distributions of

$$\{\sqrt{n}(\hat{\gamma}_1^{[\alpha_n]} - \gamma_1), \sqrt{n}(\hat{\gamma}_2^{[\alpha_n]} - \gamma_2), \ldots, \sqrt{n}(\hat{\gamma}_K^{[\alpha_n]} - \gamma_K)\}$$

converge to the same Gaussian distribution with mean zero.

Remark. Dauxois et al. [3] gave the asymptotic normalities of the eigenvalues and eigenfunctions of $\hat{\Gamma}_n$ and characterized the covariance operators of the limit Gaussian random elements. These results are special cases of Corollary 4.2 with all α_n equal to zeros. Therefore, by Corollary 4.2, all the limit Gaussian distributions in Corollary 4.2 are the same those in [3].

5. Proofs

Proof of Theorem 3.1. By remark (3) after Assumption 1, $\|\hat{\Gamma}_n\| < \infty$ a.s. Fix a sample and $\alpha \geq 0$ such that $\|\hat{\Gamma}_n\| < \infty$. Consider the Hilbert space $W_2^2([a, b])$ equipped with the inner product $(\cdot, \cdot)_\alpha$. For any $f, g \in W_2^2([a, b])$, the functional $(f, \hat{\Gamma}_n g)$ define a bilinear form in $W_2^2([a, b])$ and

$$|(f, \hat{\Gamma}_n g)| \leq \|\hat{\Gamma}_n\| \, \|f\| \, \|g\| \leq \|\hat{\Gamma}_n\| \, \|f\|_\alpha \|g\|_\alpha.$$

Hence, there is a unique bounded operator R_α in $W_2^2([a, b])$, such that for any $f, g \in W_2^2([a, b])$,

$$(f, \hat{\Gamma}_n g) = (f, R_\alpha g)_\alpha,$$

(see Section 84 in [13]). It is easy to see that R_α is symmetric and nonnegative-definite. We want to show that R_α is a compact operator (note that a compact operator is called completely continuous operator in [13]). By Definition 4 in Section 85 of [13], we only need to show that for any bounded sequence $\{f_m \in W_2^2([a, b]), m \in \mathbb{N}\}$, one can select a subsequence $\{f_{m_k}\}$ such that

$$(f_{m_k} - f_{m_l}, R_\alpha(f_{m_k} - f_{m_l}))_\alpha = (f_{m_k} - f_{m_l}, \hat{\Gamma}_n(f_{m_k} - f_{m_l})) \to 0, \tag{5.1}$$

as $k, l \to \infty$. Because $\hat{\Gamma}_n$ is a compact operator in $L^2([a, b])$ (see remark (2) after Assumption 1) and $\{f_m\}$ is also a bounded sequence in $L^2([a, b])$, one can select a subsequence $\{f_{m_k}\}$ such that $\{\hat{\Gamma}_n f_{m_k}\}$ converges, then (5.1) is true for $\{f_{m_k}\}$. Hence R_α is a compact operator. It has eigenvalues and eigenfunctions $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ with $\hat{\lambda}_1^{[\alpha]} \geq \hat{\lambda}_2^{[\alpha]} \geq \cdots \geq 0$. They are the solutions of the successive optimization problems (3.3) and (3.4) (see Chapter 3 of [16]). Now for any $\gamma \in W_2^2([a,b])$ and any $j \in \mathbb{N}$, because

$$R_\alpha \hat{\gamma}_j^{[\alpha]} = \hat{\lambda}_j^{[\alpha]} \hat{\gamma}_j^{[\alpha]},$$

we have

$$(\hat{\Gamma}_n \hat{\gamma}_j^{[\alpha]}, \gamma) = (R_\alpha \hat{\gamma}_j^{[\alpha]}, \gamma)_\alpha = \hat{\lambda}_j^{[\alpha]} (\hat{\gamma}_j^{[\alpha]}, \gamma)_\alpha. \quad \square$$

Proof of Theorem 4.1. The proof of the existence and uniqueness of the choices of the signs of $\gamma_k^{[\alpha]}$, $1 \le k \le K$ making them continuous functions of α will be postponed to the proof of Theorem 4.2 because we need some technical lemmas in the proof of Theorem 4.2. We will assume that we can choose the signs of $\gamma_k^{[\alpha]}$, $1 \le k \le K$ such that they are continuous function of α for all $0 \le \alpha \le \alpha_0$ and $\gamma_k^{[0]} = \gamma_k$, $1 \le k \le K$.

For any $1 \le k \le K$, let P_k be the orthogonal projection operator in $L^2([a,b])$ onto the space spanned by $\{\gamma_1, \ldots, \gamma_k\}$ and I be the identity operator in $L^2([a,b])$. Then $(I - P_k)$ is the orthogonal projection operator onto the closed subspace spanned by $\{\gamma_j, j \ge (k+1)\}$. $\square$

Lemma 1. *For any $k \in \mathbb{N}$, and $\alpha_1 \ge \alpha_2 \ge 0$*

$$\lambda_k^{[\alpha_1]} \le \lambda_k^{[\alpha_2]}, \qquad \hat{\lambda}_k^{[\alpha_1]} \le \hat{\lambda}_k^{[\alpha_2]}, \quad \forall k \in \mathbb{N}.$$

Proof. It follows Theorem 8.1 in Chapter 3 of [16]. $\square$

Lemma 2. *For any $1 \le k \le K$ and $\alpha \ge 0$, we have*

$$\|P_{k-1}\gamma_k^{[\alpha]}\|^2 \le \alpha^2 \left(\frac{\lambda_k}{\lambda_{k-1} - \lambda_k}\right)^2 (k-1) L_k^2 [\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]. \tag{5.2}$$

Proof. For any $j < k$, by (3.6), we have

$$\begin{aligned} \lambda_j (\gamma_k^{[\alpha]}, \gamma_j) &= (\gamma_k^{[\alpha]}, \Gamma \gamma_j) = (\Gamma \gamma_k^{[\alpha]}, \gamma_j) = \lambda_k^{[\alpha]} (\gamma_k^{[\alpha]}, \gamma_j)_\alpha \\ &= \lambda_k^{[\alpha]} \{(\gamma_k^{[\alpha]}, \gamma_j) + \alpha [\gamma_k^{[\alpha]}, \gamma_j]\}. \end{aligned}$$

So

$$(\lambda_j - \lambda_k^{[\alpha]})(\gamma_k^{[\alpha]}, \gamma_j) = \lambda_k^{[\alpha]} \alpha [\gamma_k^{[\alpha]}, \gamma_j].$$

By Assumption 2 and Lemma 1, $\lambda_j > \lambda_k \ge \lambda_k^{[\alpha]}$. Therefore,

$$(\gamma_k^{[\alpha]}, \gamma_j) = \frac{\lambda_k^{[\alpha]}}{\lambda_j - \lambda_k^{[\alpha]}} \alpha [\gamma_k^{[\alpha]}, \gamma_j]$$

and we have

$$\begin{aligned} \|P_{k-1}\gamma_k^{[\alpha]}\|^2 &= \sum_{j=1}^{k-1} (\gamma_k^{[\alpha]}, \gamma_j)^2 = \sum_{j=1}^{k-1} \left(\frac{\lambda_k^{[\alpha]}}{\lambda_j - \lambda_k^{[\alpha]}}\right)^2 \alpha^2 [\gamma_k^{[\alpha]}, \gamma_j]^2 \\ &\le \alpha^2 \left(\frac{\lambda_k}{\lambda_{k-1} - \lambda_k}\right)^2 \sum_{j=1}^{k-1} [\gamma_k^{[\alpha]}, \gamma_j]^2 \\ &\le \alpha^2 \left(\frac{\lambda_k}{\lambda_{k-1} - \lambda_k}\right)^2 \sum_{j=1}^{k-1} [\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}][\gamma_j, \gamma_j] \\ &\le \alpha^2 \left(\frac{\lambda_k}{\lambda_{k-1} - \lambda_k}\right)^2 (k-1) L_k^2 [\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}], \end{aligned}$$

where the last inequality in the second line follows from the Cauchy–Schwarz inequality. $\square$

Lemma 3. *For any $1 \le k \le K$ and any*

$$0 \le \alpha < \frac{\sqrt{1 + \frac{4k(\lambda_{k-1} - \lambda_k)^2}{(k-1)\lambda_k \|\Gamma\|}} - 1}{2kL_k^2},$$

(if $k = 1$, the right-hand side is defined to be infinity), we have

$$[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \le \frac{kL_k^2}{1 - \alpha \frac{\lambda_k}{(\lambda_{k-1}-\lambda_k)^2}(k-1)L_k^2(1+\alpha kL_k^2)\|\Gamma\|}. \tag{5.3}$$

Furthermore, if

$$0 \le \alpha \le \frac{\sqrt{1+\frac{2k(\lambda_{k-1}-\lambda_k)^2}{(k-1)\lambda_k\|\Gamma\|}} - 1}{2kL_k^2},$$

(if $k = 1$, the right-hand side is defined to be infinity), we have

$$[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \le 2kL_k^2. \tag{5.4}$$

For any $\alpha \ge 0$, we have

$$0 \le \lambda_k - \lambda_k^{[\alpha]} \le \alpha kL_k^2\lambda_k. \tag{5.5}$$

Hence, as $\alpha \to 0$, $\lambda_k^{[\alpha]} \to \lambda_k$.

Proof. Let $\operatorname{span}(\gamma_1, \ldots, \gamma_k)$ denote the linear subspace spanned by

$$\{\gamma_1, \ldots, \gamma_k\}.$$

From Theorem 5.1 (Poincare's Principle) in Chapter 3 of [16], we have

$$\begin{aligned}
\min_{0\ne\gamma\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2+\alpha[\gamma,\gamma]} \le \lambda_k^{[\alpha]} &= \frac{(\gamma_k^{[\alpha]}, \Gamma\gamma_k^{[\alpha]})}{\|\gamma_k^{[\alpha]}\|^2 + \alpha[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]} \\
&= \frac{(P_{k-1}\gamma_k^{[\alpha]} + (I-P_{k-1})\gamma_k^{[\alpha]}, \Gamma(P_{k-1}\gamma_k^{[\alpha]} + (I-P_{k-1})\gamma_k^{[\alpha]}))}{\|\gamma_k^{[\alpha]}\|^2 + \alpha[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]} \\
&= \frac{(P_{k-1}\gamma_k^{[\alpha]}, \Gamma P_{k-1}\gamma_k^{[\alpha]}) + ((I-P_{k-1})\gamma_k^{[\alpha]}, \Gamma(I-P_{k-1})\gamma_k^{[\alpha]})}{\|\gamma_k^{[\alpha]}\|^2 + \alpha[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]} \\
&\le \frac{\|\Gamma\|\cdot\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + \lambda_k}{\|\gamma_k^{[\alpha]}\|^2 + \alpha[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]},
\end{aligned} \tag{5.6}$$

where the equality in the third line of (5.6) is true because $(I - P_{k-1})$ is the orthogonal projection operator onto the closed subspace spanned by $\{\gamma_j, j \ge k\}$ which is orthogonal to $\operatorname{span}(\gamma_1, \ldots, \gamma_{k-1})$, and both of them are invariant subspaces of Γ. The last inequality in (5.6) holds because the largest eigenvalue of Γ restricted to the closed subspace spanned by $\{\gamma_j, j \ge k\}$ is λ_k and the L^2 norm of $(I - P_{k-1})\gamma_k^{[\alpha]}$ is less than 1. On the other hand, we have

$$\begin{aligned}
\min_{0\ne\gamma\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2+\alpha[\gamma,\gamma]} &= \min_{0\ne\gamma\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2\left(1+\frac{\alpha[\gamma,\gamma]}{\|\gamma\|^2}\right)} \\
&\ge \min_{0\ne\gamma\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)} \frac{(\gamma, \Gamma\gamma)}{\|\gamma\|^2\left(1+\max_{0\ne\beta\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)}\frac{\alpha[\beta,\beta]}{\|\beta\|^2}\right)} \\
&= \frac{1}{\left(1+\max_{0\ne\beta\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)}\frac{\alpha[\beta,\beta]}{\|\beta\|^2}\right)} \min_{0\ne\gamma\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)} \frac{(\gamma,\Gamma\gamma)}{\|\gamma\|^2} \\
&= \frac{\lambda_k}{\left(1+\max_{0\ne\beta\in\operatorname{span}(\gamma_1,\ldots,\gamma_k)}\frac{\alpha[\beta,\beta]}{\|\beta\|^2}\right)} \ge \frac{\lambda_k}{(1+\alpha kL_k^2)}.
\end{aligned} \tag{5.7}$$

The equality in the last line follows from the fact that the smallest eigenvalue of Γ in $\operatorname{span}(\gamma_1, \ldots, \gamma_k)$ is λ_k. The last inequality holds because that, for any $\beta \in \operatorname{span}(\gamma_1, \ldots, \gamma_k)$, let $\beta = \sum_{i=1}^k c_i\gamma_i$, where $c_1, \ldots, c_k$ are some real numbers, then we have

$$\frac{[\beta,\beta]}{\|\beta\|^2}=\frac{\left[\sum_{i=1}^{k}c_i\gamma_i,\sum_{i=1}^{k}c_i\gamma_i\right]}{\sum_{i=1}^{k}c_i^2}=\frac{\sum_{i=1}^{k}c_i^2[\gamma_i,\gamma_i]+\sum_{j\neq l}c_jc_l[\gamma_j,\gamma_l]}{\sum_{i=1}^{k}c_i^2}$$

$$\leq\frac{\sum_{i=1}^{k}c_i^2(\sqrt{[\gamma_i,\gamma_i]})^2+\sum_{j\neq l}c_jc_l\sqrt{[\gamma_j,\gamma_j]}\sqrt{[\gamma_l,\gamma_l]}}{\sum_{i=1}^{k}c_i^2}$$

$$=\frac{\left(\sum_{i=1}^{k}c_i\sqrt{[\gamma_i,\gamma_i]}\right)^2}{\sum_{i=1}^{k}c_i^2}\leq\frac{\left(\sum_{i=1}^{k}c_i^2\right)\left(\sum_{i=1}^{k}[\gamma_i,\gamma_i]\right)}{\sum_{i=1}^{k}c_i^2}\leq\sum_{i=1}^{k}[\gamma_i,\gamma_i]\leq kL_k^2,$$

where the inequality in the second line is due to the Cauchy–Schwarz inequality. Now from (5.6), (5.7) and Lemma 1, we have

$$\frac{\lambda_k}{(1+\alpha kL_k^2)}\leq\lambda_k^{[\alpha]}\leq\lambda_k.$$

From these inequalities, it can be derived that

$$0\leq\lambda_k-\lambda_k^{[\alpha]}\leq\alpha kL_k^2\lambda_k.$$

Therefore, $\lambda_k^{[\alpha]}\to\lambda_k$ as $\alpha\to 0$.

Again by (5.6), (5.7), and note that $\|\gamma_k^{[\alpha]}\|=1$, we have

$$\frac{\lambda_k}{(1+\alpha kL_k^2)}\leq\frac{\|\Gamma\|\cdot\|P_{k-1}\gamma_k^{[\alpha]}\|^2+\lambda_k}{\|\gamma_k^{[\alpha]}\|^2+\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]}=\frac{\|\Gamma\|\cdot\|P_{k-1}\gamma_k^{[\alpha]}\|^2+\lambda_k}{1+\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]}.$$

Then

$$\lambda_k(1+\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}])\leq\|\Gamma\|\cdot\|P_{k-1}\gamma_k^{[\alpha]}\|^2(1+\alpha kL_k^2)+\lambda_k(1+\alpha kL_k^2),$$

hence,

$$\lambda_k\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]\leq\|\Gamma\|\cdot\|P_{k-1}\gamma_k^{[\alpha]}\|^2(1+\alpha kL_k^2)+\lambda_k\alpha kL_k^2.$$

Now by (5.2), we have

$$[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]\leq\alpha\frac{\lambda_k}{(\lambda_{k-1}-\lambda_k)^2}(k-1)L_k^2[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}](1+\alpha kL_k^2)\|\Gamma\|+kL_k^2.$$

After rearranging the terms, we then obtain

$$[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]\left\{1-\alpha\frac{\lambda_k}{(\lambda_{k-1}-\lambda_k)^2}(k-1)L_k^2(1+\alpha kL_k^2)\|\Gamma\|\right\}\leq kL_k^2.$$

When the expression in braces on the left of the above inequality is positive, which is equivalent to

$$\alpha<\frac{\sqrt{1+\frac{4k(\lambda_{k-1}-\lambda_k)^2}{(k-1)\lambda_k\|\Gamma\|}}-1}{2kL_k^2},$$

(if $k=1$, the right-hand side is defined to be infinity), we have

$$[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]\leq\frac{kL_k^2}{1-\alpha\frac{\lambda_k}{(\lambda_{k-1}-\lambda_k)^2}(k-1)L_k^2(1+\alpha kL_k^2)\|\Gamma\|}. \tag{5.8}$$

When

$$\alpha\leq\frac{\sqrt{1+\frac{2k(\lambda_{k-1}-\lambda_k)^2}{(k-1)\lambda_k\|\Gamma\|}}-1}{2kL_k^2},$$

(if $k = 1$, the right-hand side is defined to be infinity), it can be shown that

$$1 - \alpha \frac{\lambda_k}{(\lambda_{k-1} - \lambda_k)^2}(k-1)L_k^2(1 + \alpha k L_k^2)\|\Gamma\| \geq \frac{1}{2},$$

and then it follows from (5.8) that

$$[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \leq 2kL_k^2. \quad \square$$

Lemma 4. *For any $1 \leq k \leq K$ and any*

$$0 \leq \alpha \leq \frac{\lambda_k - \lambda_{k+1}}{2kL_k^2\lambda_k}, \tag{5.9}$$

we have

$$\|(I - P_k)\gamma_k^{[\alpha]}\|^2 \leq \frac{2}{\lambda_k - \lambda_{k+1}}\left[\|\Gamma\|\alpha^2\left(\frac{\lambda_k}{\lambda_{k-1} - \lambda_k}\right)^2 (k-1)L_k^2[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] + \lambda_k\alpha[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}]^{\frac{1}{2}}L_k\right]. \tag{5.10}$$

Proof. By the following orthogonal decomposition

$$\gamma_k^{[\alpha]} = P_{k-1}\gamma_k^{[\alpha]} + (\gamma_k^{[\alpha]}, \gamma_k)\gamma_k + (I - P_k)\gamma_k^{[\alpha]}, \tag{5.11}$$

we have

$$\begin{aligned}(\gamma_k^{[\alpha]}, \Gamma\gamma_k^{[\alpha]}) &= (P_{k-1}\gamma_k^{[\alpha]}, \Gamma P_{k-1}\gamma_k^{[\alpha]}) + (\gamma_k^{[\alpha]}, \gamma_k)^2(\gamma_k, \Gamma\gamma_k) + ((I - P_k)\gamma_k^{[\alpha]}, \Gamma(I - P_k)\gamma_k^{[\alpha]}) \\ &\leq \|\Gamma\|\,\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + \lambda_k(\gamma_k^{[\alpha]}, \gamma_k)^2 + \lambda_{k+1}\|(I - P_k)\gamma_k^{[\alpha]}\|^2,\end{aligned} \tag{5.12}$$

where the last inequality follows from the fact that $(I - P_k)\gamma_k^{[\alpha]}$ belongs to the closed subspace spanned by $\{\gamma_j, j \geq k+1\}$ in which the largest eigenvalue of Γ is λ_{k+1}. On the other hand, by (3.6), we have

$$\begin{aligned}(\gamma_k^{[\alpha]}, \Gamma\gamma_k^{[\alpha]}) &= \lambda_k^{[\alpha]}(\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]})_\alpha = \lambda_k^{[\alpha]}\|\gamma_k^{[\alpha]}\|^2 + \alpha\lambda_k^{[\alpha]}[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \\ &= \lambda_k^{[\alpha]}\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + \lambda_k^{[\alpha]}(\gamma_k^{[\alpha]}, \gamma_k)^2 + \lambda_k^{[\alpha]}\|(I - P_k)\gamma_k^{[\alpha]}\|^2 + \alpha\lambda_k^{[\alpha]}[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}].\end{aligned} \tag{5.13}$$

From (5.12) and (5.13),

$$\begin{aligned}&\lambda_k^{[\alpha]}\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + \lambda_k^{[\alpha]}(\gamma_k^{[\alpha]}, \gamma_k)^2 + \lambda_k^{[\alpha]}\|(I - P_k)\gamma_k^{[\alpha]}\|^2 + \alpha\lambda_k^{[\alpha]}[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \\ &\quad \leq \|\Gamma\|\,\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + \lambda_k(\gamma_k^{[\alpha]}, \gamma_k)^2 + \lambda_{k+1}\|(I - P_k)\gamma_k^{[\alpha]}\|^2,\end{aligned}$$

then

$$\begin{aligned}(\lambda_k^{[\alpha]} - \lambda_{k+1})\|(I - P_k)\gamma_k^{[\alpha]}\|^2 &\leq (\|\Gamma\| - \lambda_k^{[\alpha]})\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + (\lambda_k - \lambda_k^{[\alpha]})(\gamma_k^{[\alpha]}, \gamma_k)^2 - \alpha\lambda_k^{[\alpha]}[\gamma_k^{[\alpha]}, \gamma_k^{[\alpha]}] \\ &\leq (\|\Gamma\| - \lambda_k^{[\alpha]})\|P_{k-1}\gamma_k^{[\alpha]}\|^2 + (\lambda_k - \lambda_k^{[\alpha]})(\gamma_k^{[\alpha]}, \gamma_k)^2.\end{aligned} \tag{5.14}$$

It follows from (5.9) that $\alpha k L_k^2\lambda_k \leq \frac{1}{2}(\lambda_k - \lambda_{k+1})$. Then by (5.5), we have

$$\lambda_k - \lambda_k^{[\alpha]} \leq \frac{1}{2}(\lambda_k - \lambda_{k+1}),$$

hence,

$$\lambda_k^{[\alpha]} - \lambda_{k+1} \geq \frac{1}{2}(\lambda_k - \lambda_{k+1}). \tag{5.15}$$

Because

$$\begin{aligned}\lambda_k(\gamma_k^{[\alpha]}, \gamma_k) &= (\gamma_k^{[\alpha]}, \Gamma\gamma_k) = (\Gamma\gamma_k^{[\alpha]}, \gamma_k) \\ &= \lambda_k^{[\alpha]}(\gamma_k^{[\alpha]}, \gamma_k)_\alpha \\ &= \lambda_k^{[\alpha]}\{(\gamma_k^{[\alpha]}, \gamma_k) + \alpha[\gamma_k^{[\alpha]}, \gamma_k]\},\end{aligned}$$

we have

$$(\lambda_k - \lambda_k^{[\alpha]})(\gamma_k^{[\alpha]}, \gamma_k) = \lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]}, \gamma_k]. \tag{5.16}$$

From (5.14)–(5.16),

$$\begin{aligned}\frac{1}{2}(\lambda_k-\lambda_{k+1})\|(I-P_k)\gamma_k^{[\alpha]}\|^2 &\le \|\Gamma\|\,\|P_{k-1}\gamma_k^{[\alpha]}\|^2+\lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]},\gamma_k](\gamma_k^{[\alpha]},\gamma_k)\\ &\le \|\Gamma\|\,\|P_{k-1}\gamma_k^{[\alpha]}\|^2+\lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]},\gamma_k]\\ &\le \|\Gamma\|\,\|P_{k-1}\gamma_k^{[\alpha]}\|^2+\lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]^{\frac{1}{2}}[\gamma_k,\gamma_k]^{\frac{1}{2}}.\end{aligned}$$

Now by Lemma 2,

$$\frac{1}{2}(\lambda_k-\lambda_{k+1})\|(I-P_k)\gamma_k^{[\alpha]}\|^2\le\|\Gamma\|\alpha^2\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2(k-1)L_k^2[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]+\lambda_k\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]^{\frac{1}{2}}L_k. \quad \square$$

Now we can prove Theorem 4.1. It follows from the definition (4.3) of α_0 that all the conditions in Lemmas 3 and 4 are satisfied. From the orthogonal decomposition

$$\gamma_k^{[\alpha]}=P_{k-1}\gamma_k^{[\alpha]}+(\gamma_k^{[\alpha]},\gamma_k)\gamma_k+(I-P_k)\gamma_k^{[\alpha]},$$

we have

$$1=\|\gamma_k^{[\alpha]}\|^2=\|P_{k-1}\gamma_k^{[\alpha]}\|^2+(\gamma_k^{[\alpha]},\gamma_k)^2+\|(I-P_k)\gamma_k^{[\alpha]}\|^2.$$

Hence, it follows from Lemmas 2 and 4 and (5.4) in Lemma 3 that

$$\begin{aligned}(\gamma_k^{[\alpha]},\gamma_k)^2 &= 1-\|P_{k-1}\gamma_k^{[\alpha]}\|^2-\|(I-P_k)\gamma_k^{[\alpha]}\|^2\ge 1-\alpha^2\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2(k-1)L_k^2[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]\\ &\quad-\frac{2}{\lambda_k-\lambda_{k+1}}\left[\|\Gamma\|\alpha^2\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2(k-1)L_k^2[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]+\lambda_k\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]^{\frac{1}{2}}L_k\right]\\ &\ge 1-\frac{2\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}\alpha-2k(k-1)L_k^4\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2\left\{1+\frac{2\|\Gamma\|}{\lambda_k-\lambda_{k+1}}\right\}\alpha^2. \qquad (5.17)\end{aligned}$$

Define

$$a=2k(k-1)L_k^4\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2\left\{1+\frac{2\|\Gamma\|}{\lambda_k-\lambda_{k+1}}\right\},\qquad b=\frac{2\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}. \qquad (5.18)$$

By solving the following inequalities,

$$a\alpha^2+b\alpha\le\frac{1}{2},\quad \alpha\ge 0,$$

we obtain $0\le\alpha\le\frac{\sqrt{b^2+2a}-b}{2a}$. Since

$$\begin{aligned}\frac{\sqrt{b^2+2a}-b}{2a}&=\frac{1}{\sqrt{b^2+2a}+b}\ge\frac{1}{2\sqrt{b^2+2a}}\ge\frac{1}{2\sqrt{2\max\{b^2,2a\}}}\\ &\ge\frac{1}{2\sqrt{2}}\min\left\{\frac{1}{b},\frac{1}{\sqrt{2a}}\right\}.\end{aligned}$$

By the definition (4.3) of α_0 and (5.18), we have

$$\alpha_0\le\frac{1}{2\sqrt{2}}\min\left\{\frac{1}{b},\frac{1}{\sqrt{2a}}\right\}\le\frac{\sqrt{b^2+2a}-b}{2a}.$$

Hence, for any $0\le\alpha\le\alpha_0$, we have $a\alpha^2+b\alpha\le\frac{1}{2}$. Now it follows from (5.17) that, for any $0\le\alpha\le\alpha_0$,

$$(\gamma_k^{[\alpha]},\gamma_k)^2\ge 1-b\alpha-a\alpha^2=\frac{1}{2}+\left(\frac{1}{2}-b\alpha-a\alpha^2\right)\ge\frac{1}{2}. \qquad (5.19)$$

Because $\gamma_k^{[\alpha]}$ is a continuous function of α, $(\gamma_k^{[\alpha]},\gamma_k)$ is also a continuous function of α and $(\gamma_k^{[0]},\gamma_k)=(\gamma_k,\gamma_k)=1$. Hence, it follows from (5.19) that $(\gamma_k^{[\alpha]},\gamma_k)>0$ for all $0\le\alpha\le\alpha_0$.

From (5.16), (5.17) and (5.4), we have

$$\begin{aligned}(\lambda_k-\lambda_k^{[\alpha]}) &= \frac{\lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]},\gamma_k]}{(\gamma_k^{[\alpha]},\gamma_k)} \le \frac{\lambda_k^{[\alpha]}\alpha|[\gamma_k^{[\alpha]},\gamma_k]|}{\sqrt{(\gamma_k^{[\alpha]},\gamma_k)^2}} \le \frac{\lambda_k^{[\alpha]}\alpha[\gamma_k^{[\alpha]},\gamma_k^{[\alpha]}]^{\frac{1}{2}}[\gamma_k,\gamma_k]^{\frac{1}{2}}}{\sqrt{(\gamma_k^{[\alpha]},\gamma_k)^2}} \\ &= \sqrt{2}\sqrt{k}L_k^2\lambda_k\alpha\left(1+O\left(\frac{\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}\alpha+\frac{k(k-1)L_k^4\lambda_k^2\|\Gamma\|}{(\lambda_{k-1}-\lambda_k)^2(\lambda_k-\lambda_{k+1})}\alpha^2\right)\right).\end{aligned}$$

By (5.17) and $(\gamma_k^{[\alpha]},\gamma_k)>0$, we have

$$\begin{aligned}\|\gamma_k^{[\alpha]}-\gamma_k\|^2 &= 2(1-(\gamma_k^{[\alpha]},\gamma_k)) \le 2(1-(\gamma_k^{[\alpha]},\gamma_k))(1+(\gamma_k^{[\alpha]},\gamma_k)) \\ &= 2(1-(\gamma_k^{[\alpha]},\gamma_k)^2),\end{aligned}$$

and thus

$$\|\gamma_k^{[\alpha]}-\gamma_k\| \le \sqrt{\alpha}\sqrt{\frac{4\sqrt{2}\sqrt{k}L_k^2\lambda_k}{\lambda_k-\lambda_{k+1}}}+\alpha\sqrt{4k(k-1)L_k^4\left(\frac{\lambda_k}{\lambda_{k-1}-\lambda_k}\right)^2\left\{1+\frac{2\|\Gamma\|}{\lambda_k-\lambda_{k+1}}\right\}}. \quad \square$$

Proof of Theorem 4.2. We first study the properties of the "half-smoothing" operators S_α. At the end of Section 2, we know that S_α is a bounded linear operator from $L^2([a,b])$ to $L^2([a,b])$ with norm less than or equal to 1. Moreover, S_α is a one to one (injective) map. Hence, its inverse S_α^{-1} exists. When $\alpha=0$, S_0 is just the identity operator I in $L^2([a,b])$. The following lemma gives the reason why S_α is called "half-smoothing" operators. □

Lemma 5. *The range of S_α (or the domain of S_α^{-1}) is $W_2^2([a,b])$. Moreover, for any $f\in W_2^2([a,b])$,*

$$\|S_\alpha^{-1}f\|^2=\|f\|_\alpha^2. \tag{5.20}$$

Proof. If $\alpha=0$, the results are trivial. Hence, we assume that $\alpha>0$. Since the space $C^\infty[a,b]$ of smooth functions is dense in space

$$(W_2^2([a,b]),\|\cdot\|_\alpha),$$

for any $f\in W_2^2([a,b])$, there exists a sequence $\{f_m\in C^\infty[a,b], m\in\mathbb{N}\}$ such that $\|f_m-f\|_\alpha\to 0$. One can see that the domain of $S_\alpha^{-2}=I+\alpha L^*L$ contains $C^\infty[a,b]$, hence $C^\infty[a,b]$ is also in the domain of S_α^{-1}. Now we compute

$$\begin{aligned}\|S_\alpha^{-1}f_l-S_\alpha^{-1}f_m\|^2 &= (S_\alpha^{-1}f_l-S_\alpha^{-1}f_m, S_\alpha^{-1}f_l-S_\alpha^{-1}f_m) \\ &= (f_l-f_m, S_\alpha^{-2}(f_l-f_m)) = (f_l-f_m,(I+\alpha L^*L)(f_l-f_m)) \\ &= (f_l-f_m,f_l-f_m)+(f_l-f_m,\alpha L^*L(f_l-f_m)) \\ &= (f_l-f_m,f_l-f_m)+\alpha(L(f_l-f_m),L(f_l-f_m)) \\ &= (f_l-f_m,f_l-f_m)+\alpha[f_l-f_m,f_l-f_m]=\|f_l-f_m\|_\alpha\to 0,\end{aligned} \tag{5.21}$$

as $m,l\to\infty$. Hence, $\{S_\alpha^{-1}f_m, m\in\mathbb{N}\}$ is a Cauchy sequence in $L^2([a,b])$. It converges to some function, say g, in $L^2([a,b])$. Since S_α is a bounded operator, $f_m=S_\alpha S_\alpha^{-1}f_m$ converges to $S_\alpha g$ in L^2-norm. However, f_m converges to f in $\|\cdot\|_\alpha$ norm, it also converges in L^2-norm. Therefore, $S_\alpha g=f$, that is, f is in the range of S_α. Hence, $W_2^2([a,b])$ is in the range of S_α. Because for any $m\in\mathbb{N}$, from a similar calculation as in (5.21),

$$\|S_\alpha^{-1}f_m\|^2=\|f_m\|_\alpha^2,$$

and

$$\|S_\alpha^{-1}f_m-S_\alpha^{-1}f\|\to 0, \quad \|f_m-f\|_\alpha\to 0,$$

we have $\|S_\alpha^{-1}f\|^2=\|f\|_\alpha^2$.

Now we show that the range of S_α is equal to $W_2^2([a,b])$. Since we have shown that $W_2^2([a,b])$ is in the range of S_α and S_α is a one-to-one map, we only need to show that the range of $W_2^2([a,b])$ under S_α^{-1} is $L^2([a,b])$. By (5.20) and the completeness of $(W_2^2([a,b]),\|\cdot\|_\alpha)$, the range of $W_2^2([a,b])$ under S_α^{-1} is a closed subspace of $L^2([a,b])$. If the range of $W_2^2([a,b])$ under S_α^{-1} is not $L^2([a,b])$, then we can find $0\ne h\in L^2([a,b])$ such that

$$(h,S_\alpha^{-1}f)=0, \quad \forall f\in W_2^2([a,b]).$$

Since one can see that the domain of $S_\alpha^{-2}=I+\alpha L^*L$ is contained in $W_2^2([a,b])$, we have

$$(h,S_\alpha^{-1}f)=0, \quad \forall f\in \text{domain of } S_\alpha^{-2}.$$

Then

$$(h, S_\alpha^{-1} f) = (S_\alpha^{-1} S_\alpha h, S_\alpha^{-1} f) = (S_\alpha h, S_\alpha^{-2} f) = 0, \quad \forall f \in \text{ domain of } S_\alpha^{-2}.$$

However, because the range of S_α^{-2} is the whole $L^2([a, b])$, we have $S_\alpha h = 0$. Hence $h = 0$ since S_α is a one-to-one map. We get a contradiction. Therefore, the range of S_α is equal to $W_2^2([a, b])$. □

Lemma 6. $\{(\hat{\lambda}_j^{[\alpha]}, S_\alpha^{-1}\hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ *and* $\{(\lambda_j^{[\alpha]}, S_\alpha^{-1}\gamma_j^{[\alpha]}) : j \in \mathbb{N}\}$ *are eigenvalues and eigenfunctions of the compact operators* $S_\alpha \hat{\Gamma}_n S_\alpha$ *and* $S_\alpha \Gamma S_\alpha$ *in* $L^2([a, b])$ *respectively. Moreover, there are no other eigenvalues for* $S_\alpha \hat{\Gamma}_n S_\alpha$ *and* $S_\alpha \Gamma S_\alpha$.

Note that the L^2 norms of $S_\alpha^{-1}\hat{\gamma}_j^{[\alpha]}$ and $S_\alpha^{-1}\gamma_j^{[\alpha]}$ may not be 1.

Proof. If $\alpha = 0$, the results are trivial. Hence, we assume that $\alpha > 0$. Because $\{(\hat{\lambda}_j^{[\alpha]}, \hat{\gamma}_j^{[\alpha]}) : j \in \mathbb{N}\}$ are solutions of the successive optimization problems (3.3) and (3.4), then by Lemma 5,

$$\frac{(S_\alpha^{-1}\hat{\gamma}_1^{[\alpha]}, S_\alpha \hat{\Gamma}_n S_\alpha S_\alpha^{-1}\hat{\gamma}_1^{[\alpha]})}{\|S_\alpha^{-1}\hat{\gamma}_1^{[\alpha]}\|^2} = \frac{(\hat{\gamma}_1^{[\alpha]}, \hat{\Gamma}_n \hat{\gamma}_1^{[\alpha]})}{\|\hat{\gamma}_1^{[\alpha]}\|_\alpha^2} = \hat{\lambda}_1^{[\alpha]} = \max_{0 \neq \gamma \in W_2^2([a,b])} \frac{(\gamma, \hat{\Gamma}_n \gamma)}{\|\gamma\|_\alpha^2}$$

$$= \max_{0 \neq \gamma \in W_2^2([a,b])} \frac{(S_\alpha^{-1}\gamma, S_\alpha \hat{\Gamma}_n S_\alpha S_\alpha^{-1}\gamma)}{\|S_\alpha^{-1}\gamma\|^2} = \max_{0 \neq \beta \in L^2([a,b])} \frac{(\beta, S_\alpha \hat{\Gamma}_n S_\alpha \beta)}{\|\beta\|^2}.$$

Hence, $(\hat{\lambda}_1^{[\alpha]}, S_\alpha^{-1}\hat{\gamma}_1^{[\alpha]})$ are the first eigenvalue and the corresponding eigenfunction of $S_\alpha \hat{\Gamma}_n S_\alpha$. Similarly, we can prove the conclusions for other eigenvalues and eigenfunctions. □

Define

$$H = \text{the Banach space of all compact bounded operators from} L^2([a, b]) \text{to} L^2([a, b]) \text{with norm defined in (2.1).} \quad (5.22)$$

For the definition and properties of compact operators in Banach spaces, we refer reader to Chapter 21 in [9]. Define a sequence of stochastic processes

$$\{Z_n(\alpha) = \sqrt{n}(S_\alpha \hat{\Gamma}_n S_\alpha - S_\alpha \Gamma S_\alpha),\ n \in \mathbb{N},\ 0 \le \alpha \le \alpha_0\},$$

which is indexed by α and takes values in H because both $\hat{\Gamma}_n$ and Γ are compact operators and S_α is a bounded operator. Note that $Z_n(0) = \sqrt{n}(\hat{\Gamma}_n - \Gamma)$. We follow the notations in [3]. Let F denote the space of Hilbert–Schmidt operators from $L^2([a, b])$ to $L^2([a, b])$. Then F is a Hilbert space with a inner product denoted by $\langle \cdot, \cdot \rangle_F$. By Assumption 1,

$$E[\|X\|^4] < \infty.$$

Thus $\hat{\Gamma}_n, \Gamma \in F$. It follows from Proposition 5 in [3] that $\{Z_n(0), n \in \mathbb{N}\}$, regarded as a sequence of random elements with values in F, converges in distribution to the Gaussian random element in F with mean 0 and covariance operator Q, where

$$Q = E[(X \otimes X - \Gamma)\tilde{\otimes}(X \otimes X - \Gamma)] = E[(X \otimes X)\tilde{\otimes}(X \otimes X)] - \Gamma \tilde{\otimes} \Gamma. \quad (5.23)$$

$X \otimes X$ denotes the bounded operator from $L^2([a, b])$ to $L^2([a, b])$ with $(X \otimes X)(\gamma) = (\gamma, X)X$ for any $\gamma \in L^2([a, b])$. $\Gamma \tilde{\otimes} \Gamma$ denotes the bounded operator from F to F with $(\Gamma \tilde{\otimes} \Gamma)(\Lambda) = \langle \Lambda, \Gamma \rangle_F \Gamma$ for any $\Lambda \in F$. The other terms in (5.23) are defined similarly. Note that according to the definition (5.23), Q is an operator from F to F. However, because F is a Hilbert space, there is an isometry between F and its dual space F'. Hence, Q can be regarded as a bounded operator from F' to F and then it satisfies the definition of covariance operators in remark (1) after Theorem 4.2. However, in this paper, we will consider the space H of compact operators which is larger than the space F of Hilbert–Schmidt operators (every Hilbert–Schmidt operator is compact). In the proof of Proposition 6 in [3], the authors used the fact that if A is a Hilbert–Schmidt operator, then $(A - zI)^{-1}$ is also a Hilbert–Schmidt operator, where z is a complex which is not an eigenvalue of A and I is the identity operator. However, this is not true in general. But $(A - zI)^{-1}$ is a bounded operator. Because the norm (2.1) in H is smaller than the norm in F, the embedding map $i : F \hookrightarrow H$ (i maps any Hilbert–Schmidt operator to itself) is a bounded operator. Then we have

Lemma 7. $\{Z_n(0), n \in \mathbb{N}\}$, *regarded as a sequence of random elements with values in* H, *converges in distribution to a Gaussian random element in* H *with mean zero and covariance operator* iQi^*, *where* i^* *is the adjoint operator of* i *and* Q *is defined in* (5.23).

Proof. It follows immediately from the following lemma. □

Lemma 8. *Suppose that* $\{X_n, n \ge 1\}$ *is a sequence of random elements with values in a Banach space* B. *If* X_n *converges in distribution to a Gaussian random element* X *with mean zero and covariance operator* Λ. *Let* T *be a bounded operator (that is, a continuous linear function) from* B *to another Banach space* C. *Then* $T(X_n)$ *converges in distribution to* $T(X)$ *which is also a Gaussian random element with mean zero and covariance operator* $T \Lambda T^*$, *where* T^* *is the adjoint operator of* T.

Proof. Since T is a continuous map from B to C, by continuous mapping theorem, $T(X_n)$ converges in distribution to $T(X)$. Now we show that $T(X)$ is an Gaussian random element. For any bounded linear functional $f \in C'$, $f \circ T \in B'$. Hence, $f(T(X)) = f \circ T(X)$ is a Gaussian random variable since X is Gaussian. Thus $T(X)$ is Gaussian and obviously its mean is zero. In order to compute its covariance operator, we introduce the following notations. For any $x \in B$, $y \in C$ and $f \in B'$, $g \in C'$, define $\langle x, f\rangle_B = f(x)$, $\langle y, g\rangle_C = g(y)$. By the definition of covariance operators (see remark (1) after Theorem 4.2) and the definition of adjoint operators, for any $g, h \in C'$,

$$\begin{aligned} E[g(T(X))h(T(X))] &= E[(g \circ T(X))(h \circ T(X))] = E[(T^*(g)(X))(T^*(f)(X))] \\ &= \langle \Lambda(T^*(g)), T^*(f)\rangle_B = \langle \Lambda T^*(g), T^*(f)\rangle_B = \langle T\Lambda T^*(g), f\rangle_C. \end{aligned}$$

Therefore, the covariance operator of TX is $T\Lambda T^*$. □

Lemma 9. *For any finite* $0 \le \alpha_1 < \cdots < \alpha_k \le \alpha_0$*, the sequence*

$$\{(Z_n(\alpha_1), \ldots, Z_n(\alpha_k)), n \in \mathbb{N}\}$$

converges in distribution to a Gaussian random element with values in H^k *and mean zero, where* H^k *is the product space of* k *copies of* H.

Proof. This lemma follows from Lemma 8 and the fact that

$$(Z_n(\alpha_1), \ldots, Z_n(\alpha_k)) = (S_{\alpha_1} Z_n(0) S_{\alpha_1}, \ldots, S_{\alpha_k} Z_n(0) S_{\alpha_k})$$

is a continuous and linear function of $Z_n(0)$ since S_{α_i}, $i = 1, \ldots, k$ are bounded operators. □

Unfortunately, S_α is not continuous as $\alpha \to 0$ under the norm (2.1). For example, let

$$[a, b] = [0, 2\pi], \quad f_n(t) = \frac{1}{\sqrt{2\pi}} e^{\text{int}}.$$

By (5.20),

$$\|S_\alpha^{-1} f_n\|^2 = \|f_n\|_\alpha^2 = \|f_n\|^2 + \alpha[f_n, f_n] = 1 + \alpha n^4.$$

Define $g_n = \frac{1}{\sqrt{1+\alpha n^4}} S_\alpha^{-1} f_n$. Then $\|g_n\| = 1$ and

$$\begin{aligned} \|(S_\alpha - I)g_n\| &= \left\|\frac{1}{\sqrt{1+\alpha n^4}} f_n - g_n\right\| \\ &\ge \|g_n\| - \left\|\frac{1}{\sqrt{1+\alpha n^4}} f_n\right\| = 1 - \frac{1}{\sqrt{1+\alpha n^4}}. \end{aligned}$$

Therefore, $\|S_\alpha - I\| \ge 1$ for all α. Note that $S_0 = I$. However, we have the following results.

Lemma 10. *For any* $f \in L^2([a, b])$, $\alpha \to S_\alpha f$ *is a continuous map from* $[0, \alpha_0]$ *to* $L^2([a, b])$.

Proof. Let E be the resolution of the identity for the self-adjoint operator S_{α_0} (for reference, see Chapter 12 of [14]). Because S_{α_0} is a positive operator with $\|S_{\alpha_0}\| \le 1$, $E_{f,f}$ is a bounded positive Borel measure in $[0, 1]$. Fix $\alpha \in [0, \alpha_0]$.

$$\begin{aligned} S_\alpha &= (I + \alpha L^* L)^{-\frac{1}{2}} = \left(\left(1 - \frac{\alpha}{\alpha_0}\right) I + \frac{\alpha}{\alpha_0}(I + \alpha_0 L^* L)\right)^{-\frac{1}{2}} \\ &= \left(\left(1 - \frac{\alpha}{\alpha_0}\right) I + \frac{\alpha}{\alpha_0} S_{\alpha_0}^{-2}\right)^{-\frac{1}{2}} = S_{\alpha_0}\left(\frac{\alpha}{\alpha_0} + \left(1 - \frac{\alpha}{\alpha_0}\right) S_{\alpha_0}^2\right)^{-\frac{1}{2}}. \end{aligned}$$

Now define a family of continuous functions on $[0, 1]$,

$$\varphi_\alpha(x) = \begin{cases} \dfrac{x}{\sqrt{\frac{\alpha}{\alpha_0} + \left(1 - \frac{\alpha}{\alpha_0}\right) x^2}}, & 0 < \alpha \le \alpha_0 \\ 1 & \alpha = 0, \end{cases}$$

then $S_\alpha = \varphi_\alpha(S_{\alpha_0})$. Let $\alpha' \in [0, \alpha_0]$ and $\alpha' \to \alpha$. It follows from Theorems 12.21 and 12.23 in Chapter 12 of [14] that

$$\|(S_{\alpha'} - S_\alpha)f\|^2 = \int_0^1 (\varphi_{\alpha'}(x) - \varphi_\alpha(x))^2 \mathrm{d}E_{f,f}(x).$$

The integrand on the right-hand side is bounded. If $\alpha \ne 0$, the integrand converges to 0 at each point in $[0, 1]$ as $\alpha' \to \alpha$. By the bounded convergence theorem, $\|(S_{\alpha'} - S_\alpha)f\|^2 \to 0$. If $\alpha = 0$, the integrand converges to 0 at each point in $[0, 1]$

except 0. If we can show that the measure value $E_{f,f}(\{0\})$ of $E_{f,f}$ on the set $\{0\}$ is zero, then by the bounded convergence theorem, we still have $\|(S_{\alpha'} - S_\alpha)f\|^2 \to 0$. In fact, for any $g \in L^2([a, b])$,

$$(g, S_{\alpha_0}E(\{0\})f) = \int_{\{0\}} x\mathrm{d}E_{g,f}(x) = 0.$$

Hence, $S_{\alpha_0}E(\{0\})f = 0$. Because S_{α_0} is a one-to-one operator, $E(\{0\})f = 0$. Therefore,

$$E_{f,f}(\{0\}) = (f, E(\{0\})f) = 0. \quad \square$$

Lemma 11. *For any compact operator Λ in $L^2([a, b])$, $\alpha \to S_\alpha \Lambda S_\alpha$ is a continuous map from $[0, \alpha_0]$ to H.*

Proof. By Lemma 11 in Section XI.9 of [5], there exists a sequence Λ_m of bounded operators having finite-dimensional range, such that $\|\Lambda_m - \Lambda\| \to 0$. If we can show that for each $m, \alpha \to S_\alpha \Lambda_m S_\alpha$ is a continuous map, then since $\|S_\alpha \Lambda_m S_\alpha - S_\alpha \Lambda S_\alpha\| \le \|\Lambda_m - \Lambda\| \to 0$ uniformly, $\alpha \to S_\alpha \Lambda S_\alpha$ is continuous. Now fix m and $0 \le \alpha \le \alpha_0$. Let $\{e_1, \ldots, e_k\}$ be an orthonormal basis of the range of Λ_m and $\alpha' \to \alpha$. For any $f \in L^2([a, b])$ with $\|f\| \le 1$,

$$\begin{aligned}
\|S_{\alpha'} \Lambda_m S_{\alpha'} f - S_\alpha \Lambda_m S_\alpha f\| &= \|(S_{\alpha'} - S_\alpha)\Lambda_m S_{\alpha'} f + S_\alpha \Lambda_m (S_{\alpha'} - S_\alpha) f\| \\
&\le \|(S_{\alpha'} - S_\alpha)\Lambda_m S_{\alpha'} f\| + \|S_\alpha \Lambda_m (S_{\alpha'} - S_\alpha) f\| \\
&= \|(S_{\alpha'} - S_\alpha)\Lambda_m (S_{\alpha'} - S_\alpha) f + (S_{\alpha'} - S_\alpha)\Lambda_m S_\alpha f\| + \|S_\alpha \Lambda_m (S_{\alpha'} - S_\alpha) f\| \\
&\le \|(S_{\alpha'} - S_\alpha)\| \, \|\Lambda_m (S_{\alpha'} - S_\alpha) f\| + \|(S_{\alpha'} - S_\alpha)\Lambda_m S_\alpha f\| \\
&\quad + \|S_\alpha\| \, \|\Lambda_m (S_{\alpha'} - S_\alpha) f\| \le 3\|\Lambda_m (S_{\alpha'} - S_\alpha) f\| + \|(S_{\alpha'} - S_\alpha)\Lambda_m S_\alpha f\|.
\end{aligned}$$

Because

$$\Lambda_m S_\alpha f = \sum_{i=1}^{k} (\Lambda_m S_\alpha f, e_i) e_i$$

$$\begin{aligned}
\|(S_{\alpha'} - S_\alpha)\Lambda_m S_\alpha f\| &\le \sum_{i=1}^{k} |(\Lambda_m S_\alpha f, e_i)| \, \|(S_{\alpha'} - S_\alpha) e_i\| \\
&\le \sum_{i=1}^{k} \|\Lambda_m\| \, \|(S_{\alpha'} - S_\alpha) e_i\|
\end{aligned}$$

which converges to 0 uniformly for all $f \in L^2([a, b])$ with $\|f\| \le 1$ by Lemma 10. Now

$$\begin{aligned}
\|\Lambda_m (S_{\alpha'} - S_\alpha) f\|^2 &= \sum_{i=1}^{k} |(\Lambda_m (S_{\alpha'} - S_\alpha) f, e_i)|^2 \\
&= \sum_{i=1}^{k} |(f, (S_{\alpha'} - S_\alpha)\Lambda_m^* e_i)|^2 \le \sum_{i=1}^{k} \|(S_{\alpha'} - S_\alpha)\Lambda_m^* e_i\|^2
\end{aligned}$$

which converges to 0 uniformly for all $f \in L^2([a, b])$ with $\|f\| \le 1$ by Lemma 10, where Λ_m^* is the adjoint operator of Λ_m. Hence, $\|S_{\alpha'} \Lambda_m S_{\alpha'} - S_\alpha \Lambda_m S_\alpha\| \to 0$. $\square$

In the next lemma, we assume that all the eigenfunctions have norms 1.

Lemma 12. *Suppose that $\alpha \to \Lambda(\alpha)$ is a continuous map from $[0, \alpha_0]$ to the subspace of positive compact operators in $L^2([a, b])$ in H. Assume that the first K eigenvalues of $\Lambda(\alpha)$ for any $\alpha \in [0, \alpha_0]$ are positive and mutually different, and each of them has multiplicity 1. Then given the first K eigenfunctions $\{e_k^{[0]}, 1 \le k \le K\}$ of $\Lambda(0)$, there exist unique choices of the first k eigenfunctions $\{e_k^{[\alpha]}, 1 \le k \le K\}$ of $\Lambda(\alpha)$ for any $\alpha \in (0, \alpha_0]$ such that $\alpha \to e_k^{[\alpha]}$ is a continuous map from $[0, \alpha_0]$ to $L^2([a, b])$ for any $1 \le k \le K$.*

Note that for each $1 \le k \le K$ and $0 \le \alpha \le \alpha_0$, there exist two eigenfunctions with norm 1 of $\Lambda(\alpha)$ corresponding its k-th eigenvalues and any one of the two eigenfunctions is equal to the other one multiplied by -1.

Proof. Let $\mu_1^{[\alpha]} > \cdots > \mu_K^{[\alpha]} > 0$ be the first K eigenvalues of $\Lambda(\alpha)$. Let $E^k(\alpha)$ be the orthogonal projection onto the space spanned by the $e_k^{[\alpha]}$, $1 \le k \le K$, $0 \le \alpha \le \alpha_0$. Note $E^k(\alpha)$ does not depend on the sign of $e_k^{[\alpha]}$.

We first show that for any $1 \le k \le K$, $E^k(\alpha)$ is a continuous function of from $[0, \alpha_0]$ to H. For any fixed $\alpha \in [0, \alpha_0]$, we can find a small positive number ϵ_α, such that the $K + 1$ intervals

$$[\mu_1^{[\alpha]} - \epsilon_\alpha, \mu_1^{[\alpha]} + \epsilon_\alpha], [\mu_2^{[\alpha]} - \epsilon_\alpha, \mu_2^{[\alpha]} + \epsilon_\alpha], \ldots, [\mu_{K+1}^{[\alpha]} - \epsilon_\alpha, \mu_{K+1}^{[\alpha]} + \epsilon_\alpha]$$

are disjoint. Since $\Lambda(\alpha)$ is a continuous function, we can choose a neighborhood $\mathcal{M}_\alpha$ of α in $[0,\alpha_0]$, such that for any $\alpha' \in \mathcal{M}_\alpha$

$$\max_{1\le k\le K+1} |\mu_k^{[\alpha']} - \mu_k^{[\alpha]}| \le \|\Lambda(\alpha') - \Lambda(\alpha)\| \le \frac{\epsilon_\alpha}{4}$$

where the first inequality follows from Corollary 4 in Section XI.9 of [5]. Now we define K circles on the complex plane $\mathbb{C}$,

$$C_k = \text{the circle with center } \mu_k^{[\alpha]} \text{ and radius } \epsilon_\alpha, \quad 1 \le k \le K.$$

Then one can see that for any $\alpha' \in \mathcal{M}_\alpha$, the disk bounded by the circle C_k only contains the k-th eigenvalues $\mu_k^{[\alpha']}$ of $\Lambda(\alpha')$. Hence, we have (see Section VII.3 of [4] or Definition 10.26 in [14])

$$E^k(\alpha') = \frac{1}{2\pi \mathrm{i}} \int_{C_k} (zI - \Lambda(\alpha'))^{-1} \mathrm{d}z,$$

for any $\alpha' \in \mathcal{M}_\alpha$. Since $(zI - \Lambda(\alpha'))^{-1}$ is a continuous function of $z \in C_k$ and C_k is a compact set, we have

$$M = \sup_{z\in C_k} \|(zI - \Lambda(\alpha'))^{-1}\| < \infty. \tag{5.24}$$

Since $\Lambda(\alpha)$ is a continuous function of α, for any $0 < \delta < 1$, we can find a neighborhood $\mathcal{N}_\alpha$ of α such that

$$\|\Lambda(\alpha') - \Lambda(\alpha)\| \le \frac{\delta}{M}, \quad \forall \alpha' \in \mathcal{N}_\alpha. \tag{5.25}$$

Now for any $\alpha' \in \mathcal{M}_\alpha \bigcap \mathcal{N}_\alpha$,

$$\begin{aligned}
\|E^k(\alpha') - E^k(\alpha)\| &\le \frac{1}{2\pi} \int_{C_k} \|(zI - \Lambda(\alpha'))^{-1} - (zI - \Lambda(\alpha))^{-1}\| \mathrm{d}z \\
&= \frac{1}{2\pi} \int_{C_k} \|(zI - \Lambda(\alpha) - (\Lambda(\alpha') - \Lambda(\alpha)))^{-1} - (zI - \Lambda(\alpha))^{-1}\| \mathrm{d}z \\
&= \frac{1}{2\pi} \int_{C_k} \|(zI - \Lambda(\alpha))^{-1}(I - (\Lambda(\alpha') - \Lambda(\alpha))(zI - \Lambda(\alpha))^{-1})^{-1} - (zI - \Lambda(\alpha))^{-1}\| \mathrm{d}z \\
&\le \frac{1}{2\pi} \int_{C_k} \|(zI - \Lambda(\alpha))^{-1}\| \, \|(I - (\Lambda(\alpha') - \Lambda(\alpha)) \, (zI - \Lambda(\alpha))^{-1})^{-1} - I\| \mathrm{d}z \\
&\le \frac{M}{2\pi} \int_{C_k} \left\| \left(I + \sum_{k=1}^{\infty} [(\Lambda(\alpha') - \Lambda(\alpha))(zI - \Lambda(\alpha))^{-1}]^k \right) - I \right\| \mathrm{d}z \\
&\le \frac{M}{2\pi} \int_{C_k} \sum_{k=1}^{\infty} [\|\Lambda(\alpha') - \Lambda(\alpha)\| \, \|(zI - \Lambda(\alpha))^{-1}\|]^k \mathrm{d}z \\
&\le \frac{M}{2\pi} \int_{C_k} \mathrm{d}z \sum_{k=1}^{\infty} \left[\frac{\delta}{M} M\right]^k \quad \text{by (5.24) and (5.25)} \\
&= \frac{M}{2\pi} \int_{C_k} \mathrm{d}z \frac{\delta}{1-\delta}.
\end{aligned} \tag{5.26}$$

Since δ can be arbitrarily small, $E^k(\alpha)$ is continuous at α.

Now we show that for any given $\alpha \in [0,\alpha_0]$, and given $e_k^{[\alpha]}$, there exists a neighborhood $[\alpha^1, \alpha^2]$ of α such that for any $\alpha' \in [\alpha^1,\alpha^2]$, we can uniquely choose $e_k^{[\alpha']}$ such that $e_k^{[\alpha']}$ is continuous in this neighborhood. Because $E^k(\alpha')$ is a continuous function of α', $\|E^k(\alpha')e_k^{[\alpha]}\|$ is a continuous function of α' and its value is 1 at $\alpha' = \alpha$. Hence, we can find a neighborhood $[\alpha^1,\alpha^2]$ of α such that $\|E^k(\alpha')e_k^{[\alpha]}\| \ge \frac{1}{2}$ for $\alpha' \in [\alpha^1,\alpha^2]$. Then

$$e_k^{[\alpha']} = \frac{E^k(\alpha')e_k^{[\alpha]}}{\|E^k(\alpha')e_k^{[\alpha]}\|},$$

are eigenfunctions and continuous in $[\alpha^1,\alpha^2]$. Now we show the uniqueness. Suppose $\tilde{e}_k^{[\alpha]}$, $\alpha' \in [\alpha^1,\alpha^2]$ is another choice of the eigenfunctions such that it is continuous and $\tilde{e}_k^{[\alpha]} = e_k^{[\alpha]}$. If for some $\alpha'' \in [\alpha^1,\alpha^2]$, $e_k^{[\alpha'']} \ne \tilde{e}_k^{[\alpha'']}$, we have $e_k^{[\alpha'']} = -\tilde{e}_k^{[\alpha'']}$. Since both the inner products $(e_k^{[\alpha]}, e_k^{[\alpha']})$ and $(e_k^{[\alpha]}, \tilde{e}_k^{[\alpha']})$ are continuous functions for $\alpha' \in [\alpha^1,\alpha^2]$. By the choice of $[\alpha^1,\alpha^2]$, $|(e_k^{[\alpha]}, e_k^{[\alpha'']})| = |(e_k^{[\alpha]}, \tilde{e}_k^{[\alpha'']})| \ge \frac{1}{2}$. Because $(e_k^{[\alpha]}, e_k^{[\alpha'']}) = -(e_k^{[\alpha]}, \tilde{e}_k^{[\alpha'']})$, one of them must be negative. Without loss of generality, we assume that $(e_k^{[\alpha]}, e_k^{[\alpha'']}) < 0$. Since $(e_k^{[\alpha]}, e_k^{[\alpha]}) = 1 > 0$, it follows from the intermediate value theorem that there is at least one point α''' between α and α'' such that $(e_k^{[\alpha]}, e_k^{[\alpha''']}) = 0$. However, it is impossible because

$$|(e_k^{[\alpha]}, e_k^{[\alpha''']})| = \|E^k(\alpha''')e_k^{[\alpha]}\| \geq \frac{1}{2}.$$

Hence we have proved the uniqueness.

Fix $e_k^{[0]}$. Let the set

$$\mathcal{V} = \{\alpha \in [0, \alpha_0] : \text{ we can uniquely choose } e_k^{[\alpha']} \text{ for } \alpha' \in [0, \alpha_0] \text{ such that } e_k^{[\alpha']} \text{ is continuous in } [0, \alpha]\}.$$

By the arguments in the last paragraph, $\mathcal{V}$ is nonempty. Now we show that the set $\mathcal{V}$ is an open set. Suppose that α^* is any point in $\mathcal{V}$. It follows from the last paragraph that there exists a neighborhood $[\alpha^1, \alpha^2]$ of α^* such that given $\mathrm{e}^{[\alpha^*]}$, we can uniquely choose the sign of $\mathrm{e}^{[\alpha]}$ for any $\alpha \in [\alpha^1, \alpha^2]$ to make $\mathrm{e}^{[\alpha]}$, $\alpha \in [\alpha^1, \alpha^2]$ a continuous function. We show that $[\alpha^1, \alpha^2] \subset \mathcal{V}$. Let α^{**} be any point in $[\alpha^1, \alpha^2]$. It is easy to see that we can choose the signs of $\mathrm{e}^{[\alpha]}$ for all $\alpha \in [0, \alpha^{**}]$ such that $\mathrm{e}^{[\alpha]}$ is a continuous function of α in $[0, \alpha^{**}]$. We only need to show the uniqueness of $\mathrm{e}^{[\alpha]}$. The uniqueness is obvious if $\alpha^{**} \geq \alpha^*$ since $\alpha^* \in \mathcal{V}$. Hence we assume that $\alpha^{**} < \alpha^*$. We will proceed by contradiction. Assume that there are two different continuous functions $\tilde{\mathrm{e}}^{[\alpha]}$ and $\tilde{\tilde{\mathrm{e}}}^{[\alpha]}$, $0 \leq \alpha \leq \alpha^{**}$. By the definition of $[\alpha^1, \alpha^2]$, we can choose a continuous function $\hat{\mathrm{e}}^{[\alpha]}$, $\alpha^{**} \leq \alpha \leq \alpha^*$. Define

$$\mathrm{e}'^{[\alpha]} = \begin{cases} \tilde{\mathrm{e}}^{[\alpha]} & \text{if } 0 \leq \alpha \leq \alpha^{**} \\ \hat{\mathrm{e}}^{[\alpha]} & \text{if } \alpha^{**} \leq \alpha \leq \alpha^* \text{ and } \tilde{\mathrm{e}}^{[\alpha^{**}]} = \hat{\mathrm{e}}^{[\alpha^{**}]} \\ -\hat{\mathrm{e}}^{[\alpha]} & \text{if } \alpha^{**} \leq \alpha \leq \alpha^* \text{ and } \tilde{\mathrm{e}}^{[\alpha^{**}]} = -\hat{\mathrm{e}}^{[\alpha^{**}]} \end{cases},$$

and

$$\mathrm{e}''^{[\alpha]} = \begin{cases} \tilde{\tilde{\mathrm{e}}}^{[\alpha]} & \text{if } 0 \leq \alpha \leq \alpha^{**} \\ \hat{\mathrm{e}}^{[\alpha]} & \text{if } \alpha^{**} \leq \alpha \leq \alpha^* \text{ and } \tilde{\tilde{\mathrm{e}}}^{[\alpha^{**}]} = \hat{\mathrm{e}}^{[\alpha^{**}]} \\ -\hat{\mathrm{e}}^{[\alpha]} & \text{if } \alpha^{**} \leq \alpha \leq \alpha^* \text{ and } \tilde{\tilde{\mathrm{e}}}^{[\alpha^{**}]} = -\hat{\mathrm{e}}^{[\alpha^{**}]}. \end{cases}$$

Then $\tilde{\mathrm{e}}^{[\alpha]}$ and $\tilde{\tilde{\mathrm{e}}}^{[\alpha]}$ are two different continuous functions in $[0, \alpha^*]$, which contradicts $\alpha^* \in \mathcal{V}$. Hence, $\mathcal{V}$ is an open set.

Now if we can prove that $\mathcal{V}$ is also a closed set, we have $\mathcal{V} = [0, \alpha_0]$. Let $\alpha_m \in \mathcal{V}$ be a sequence of positive numbers converging to $\alpha \in [0, \alpha_0]$. If for some m, $\alpha_m \geq \alpha$, it is obvious that $\alpha \in \mathcal{V}$. Hence we assume that $\alpha_m < \alpha$ for all m. Then we can uniquely choose the signs of of $e_k^{[\alpha']}$ such that $e_k^{[\alpha']}$ is continuous in $[0, \alpha)$. Let $e_k^{[\alpha]}$ be one of the two eigenfunctions with norm 1. Because for any $\alpha' < \alpha$

$$|(e_k^{[\alpha]}, e_k^{[\alpha']})^2 - 1| = |(e_k^{[\alpha]}, E^k(\alpha')e_k^{[\alpha]}) - (e_k^{[\alpha]}, E^k(\alpha)e_k^{[\alpha]})| \leq \|E^k(\alpha') - E^k(\alpha)\|$$

goes to zero as $\alpha' \to \alpha$, $(e_k^{[\alpha]}, e_k^{[\alpha']})^2 \to 1$. Since $e_k^{[\alpha']}$ is continuous in $[0, \alpha)$. $(e_k^{[\alpha]}, e_k^{[\alpha']})$ converges either to 1 or -1. In the latter case, we change $e_k^{[\alpha]}$ to $-e_k^{[\alpha]}$. Hence, without loss of generality, we assume that $(e_k^{[\alpha]}, e_k^{[\alpha']}) \to 1$ as $\alpha' \to \alpha$. Now one can see that $e_k^{[\alpha']}$ is continuous on $[0, \alpha]$ and its uniqueness is obvious. Hence, $\alpha \in \mathcal{V}$. We have proven that $\mathcal{V}$ is a close set. $\square$

Define $C_H[0, \alpha_0]$ to be the space of all the continuous function from $[0, \alpha_0] \to H$ (see Chapter 3 of [1]). For any $\{\Lambda(\alpha) : 0 \leq \alpha \leq \alpha_0\} \in C_H[0, \alpha_0]$, define a norm

$$\|\Lambda\| = \sup_{0 \leq \alpha \leq \alpha_0} \|\Lambda(\alpha)\|. \tag{5.27}$$

Under the norm (5.27), $C_H[0, \alpha_0]$ is a Banach space. Recall the definition

$$\{Z_n(\alpha) = \sqrt{n}(S_\alpha \hat{\Gamma}_n S_\alpha - S_\alpha \Gamma S_\alpha),\ n \in \mathbb{N}, 0 \leq \alpha \leq \alpha_0\}.$$

By Lemma 11, we can regard the stochastic processes Z_n in $[0, \alpha]$ as random elements with values in $C_H[0, \alpha_0]$. Define a linear map $\Theta : H \to C_H[0, \alpha_0]$ such that for any compact operator $U \in H$,

$$\Theta(U) = \{S_\alpha U S_\alpha, 0 \leq \alpha \leq \alpha_0\}. \tag{5.28}$$

Lemma 13. *Θ is a bounded operator and the sequence $\{Z_n, n \in \mathbb{N}\}$ of stochastic processes with sample paths in $C_H[0, \alpha_0]$ converges in distribution to the Gaussian random element with mean zero and covariance operator $\Theta i Q i^* \Theta^*$.*

Proof. Since the norm of S_α is less than or equal to 1, for any $V \in H$,

$$\sup_{0 \leq \alpha \leq \alpha_0} \|S_\alpha U S_\alpha - S_\alpha V S_\alpha\| \leq \|U - V\|.$$

Hence, the map (5.28) is continuous and hence a bounded operator. Since $Z_n = \Theta(Z_n(0))$, the lemma follows from Lemmas 7 and 8. $\square$

Now for any $1 \le k \le K$, define

$$\hat{\eta}_k^{[\alpha]} = \frac{S_\alpha^{-1}\hat{\gamma}_k^{[\alpha]}}{\|S_\alpha^{-1}\hat{\gamma}_k^{[\alpha]}\|}, \qquad \eta_k^{[\alpha]} = \frac{S_\alpha^{-1}\gamma_k^{[\alpha]}}{\|S_\alpha^{-1}\gamma_k^{[\alpha]}\|}. \tag{5.29}$$

Note that by Lemma 6, $\hat{\eta}_k^{[\alpha]}$ and $\eta_k^{[\alpha]}$ are the eigenfunctions of $S_\alpha \hat{\Gamma} S_\alpha$ and $S_\alpha \Gamma S_\alpha$ with norms 1. By (5.29) and because $\|\hat{\gamma}_k^{[\alpha]}\| = 1$ and $\|\gamma_k^{[\alpha]}\| = 1$, we have

$$\|S_\alpha \hat{\eta}_k^{[\alpha]}\| = \frac{1}{\|S_\alpha^{-1}\hat{\gamma}_k^{[\alpha]}\|}, \qquad \|S_\alpha \eta_k^{[\alpha]}\| = \frac{1}{\|S_\alpha^{-1}\gamma_k^{[\alpha]}\|}, \tag{5.30}$$

and

$$\hat{\gamma}_k^{[\alpha]} = \frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \hat{\eta}_k^{[\alpha]}\|}, \qquad \gamma_k^{[\alpha]} = \frac{S_\alpha \eta_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|}. \tag{5.31}$$

Define $\tilde{\epsilon}_k = \frac{\lambda_k - \lambda_{k+1}}{4}$, $1 \le k \le K$, and $\epsilon_K = \min_{1 \le k \le K} \tilde{\epsilon}_k$. Then the $K+1$ intervals

$$[\lambda_1 - \tilde{\epsilon}_1, \lambda_1 + \tilde{\epsilon}_1], [\lambda_2 - \tilde{\epsilon}_2, \lambda_2 + \tilde{\epsilon}_1], \ldots, [\lambda_K - \tilde{\epsilon}_K, \lambda_K + \tilde{\epsilon}_{K-1}], [\lambda_{K+1} - \tilde{\epsilon}_K, \lambda_{K+1} + \tilde{\epsilon}_K], \tag{5.32}$$

are disjoint. By the definition (4.3) of α_0 and (5.5) in Lemma 3, for any $0 \le \alpha \le \alpha_0$ and $1 \le k \le K$,

$$0 \le \lambda_k - \lambda_k^{[\alpha]} \le \alpha k L_k^2 \lambda_k \le \alpha_0 k L_k^2 \lambda_k \le \frac{\lambda_k - \lambda_{k+1}}{16 k L_k^2 \lambda_k} k L_k^2 \lambda_k \le \frac{\tilde{\epsilon}_k}{4}. \tag{5.33}$$

Hence, $\lambda_1^{[\alpha]}, \ldots, \lambda_K^{[\alpha]}$ are different mutually for all $0 \le \alpha \le \alpha_0$. Now given γ_k, $1 \le k \le K$, by Lemmas 11 and 12, we can uniquely choose the first K eigenfunctions $\{\eta_k^{[\alpha]}, 1 \le k \le K\}$ of $S_\alpha \Gamma S_\alpha$ such that $\eta_k^{[0]} = \gamma_k$ and $\eta_k^{[\alpha]}$, $1 \le k \le K$, are continuous functions of α. We have proved the claims about the continuity of $\gamma_k^{[\alpha]}$, $1 \le k \le K$ at the beginning of the proof of Theorem 4.1.

Now we define K circles in the complex plane $\mathbb{C}$,

$$\begin{aligned} C_1 &= \text{the circle with center } \lambda_1 \text{ and radius } \tilde{\epsilon}_1, \quad 1 \le k \le K, \\ C_k &= \text{the circle with center } \lambda_k + \frac{\tilde{\epsilon}_{k-1} - \tilde{\epsilon}_k}{2} \text{ and radius } \frac{\tilde{\epsilon}_{k-1} + \tilde{\epsilon}_k}{2}, \quad 1 \le k \le K. \end{aligned} \tag{5.34}$$

Note that the K discs bounded by C_k, $1 \le k \le K$ are disjoint and the intersections between these discs and the real line in the complex plane are just the first K intervals in (5.32). Let $E^k(\alpha)$ be the orthogonal projection onto the space spanned by the $\eta_k^{[\alpha]}$, $1 \le k \le K$, $0 \le \alpha \le \alpha_0$. Now because it follows from (5.33) that for any $0 \le \alpha \le \alpha_0$, $1 \le k \le K$, the disk bounded by the circle C_k only contains the k-th eigenvalues $\lambda_k^{[\alpha]}$ of $S_\alpha \Gamma S_\alpha$, for any $0 \le \alpha \le \alpha_0$, $1 \le k \le K$, we have

$$E^k(\alpha) = \frac{1}{2\pi \mathrm{i}} \int_{C_k} (zI - S_\alpha \Gamma S_\alpha)^{-1} \mathrm{d}z. \tag{5.35}$$

By Lemma 11, $S_\alpha \Gamma S_\alpha$ is a continuous function of α. Hence, by a similar calculation as in (5.26), it can be shown that $E^k(\alpha)$ is a continuous function of α.

Recall that we define in (4.6)

$$\Omega_0 = \{\omega : \text{ there exists at least one } \alpha \in [0, \alpha_0] \text{ such that } \hat{\lambda}_1^{[\alpha]}, \ldots, \hat{\lambda}_K^{[\alpha]} \text{ are not mutually different}\}.$$

Lemma 14. *Ω_0 is a measurable set and $P(\Omega_0) \to 0$ as $n \to \infty$.*

Proof. Consider the subset

$$\begin{aligned} \mathcal{E} = \{B : \ & B \text{ is a positive compact operator, its first } K \text{ eigenvalues are mutually different and} \\ & \text{each of them has multiplicity } 1\}. \end{aligned}$$

$\mathcal{E}$ is an open subset of the space of all positive compact operators which is closed in H, hence it is measurable. Let $(\Omega, \mathcal{F})$ be the probability space and $([0, \alpha_0], \mathcal{B}[0, \alpha_0])$ be the Lebesgue space. Since $S_\alpha \hat{\Gamma} S_\alpha$ has continuous sample paths, it is jointly measurable in $(\Omega \times [0, \alpha_0], \mathcal{F} \times \mathcal{B}[0, \alpha_0])$. One can see that Ω_0^c is the projection of the set $\{(\omega, \alpha) : S_\alpha \hat{\Gamma}_n S_\alpha \in \mathcal{E}\}$ to Ω. Therefore, Ω_0^c is measurable, so is Ω_0. By (5.33) and the definition of ϵ_K (just above (5.32)), we have

$$\Omega_0 \subset \left\{ \sup_{0 \le \alpha \le \alpha_0} \max_{1 \le k \le K+1} |\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}| > \frac{\epsilon_K}{4} \right\}.$$

By Corollary 4 in Section XI.9 of [5],

$$\sup_{0\le\alpha\le\alpha_0}\max_{1\le k\le K+1}|\hat{\lambda}_k^{[\alpha]}-\lambda_k^{[\alpha]}|\le\sup_{0\le\alpha\le\alpha_0}\|S_\alpha\hat{\Gamma}_nS_\alpha-S_\alpha\Gamma S_\alpha\|\le\|\hat{\Gamma}_n-\Gamma\|. \tag{5.36}$$

Hence,

$$P(\Omega_0)\le P\left(\|\hat{\Gamma}_n-\Gamma\|>\frac{\epsilon_K}{4}\right)\to 0 \tag{5.37}$$

by the law of large numbers. □

For any $\omega\in\Omega_0$, define $\hat{E}_n^k(\alpha)$ to be zero. For any $\omega\notin\Omega_0$, define $\hat{E}_n^k(\alpha)$ to be the orthogonal projection onto the space spanned by the k-th eigenfunction $\hat{\eta}_k^{[\alpha]}$ of $S_\alpha\hat{\Gamma}_nS_\alpha$ (note that $\hat{E}_n^k(\alpha)$ does not depend on the sign of $\hat{\eta}_k^{[\alpha]}$). By the same argument as in the proof of Lemma 12, we can show that $\hat{E}_n^k(\alpha)$ is a continuous function of $S_\alpha\hat{\Gamma}_nS_\alpha$, so it is measurable and continuous in α. Now let $\{e_m, m\in\mathbb{N}\}$ be a set of complete orthonormal basis functions in $L^2([a,b])$, we choose

$$\hat{\eta}_k^{[0]}=\frac{\hat{E}_n^k(0)\gamma_k}{\|\hat{E}_n^k(0)\gamma_k\|}\chi_{\{\hat{E}_n^k(0)\gamma_k\neq0\}}+\sum_{m=1}^{\infty}\frac{\hat{E}_n^k(0)e_m}{\|\hat{E}_n^k(0)e_m\|}\chi_{\{\hat{E}_n^k(0)\gamma_k=0,\hat{E}_n^k(0)e_j=0,1\le j\le m-1,\hat{E}_n^k(0)e_m\neq0\}} \tag{5.38}$$

in Ω_0^c and 0 in Ω_0, where χ is the indicator function. Then $\hat{\eta}_k^{[0]}$ is measurable and

$$(\hat{\eta}_k^{[0]},\eta_k^{[0]})\ge 0. \tag{5.39}$$

Now by Lemmas 11 and 12 and the definition of Ω_0, for any $\omega\notin\Omega_0$, we can uniquely choose $\hat{\eta}_k^{[\alpha]}$, $1\le k\le K$, such that $\hat{\eta}_k^{[\alpha]}$, $1\le k\le K$ are continuous functions of α. $\hat{\eta}_k^{[\alpha]}$ is measurable by the following lemma. By (5.31), $\hat{\gamma}_k^{[\alpha]}$, $1\le k\le K$ are continuous and measurable with $(\hat{\gamma}_k^{[0]},\gamma_k)\ge 0$.

Lemma 15. *If for any $1\le k\le K$, $\hat{\eta}_k^{[\alpha]}$ is a measurable map to $C_{L^2[a,b]}[0,\alpha_0]$.*

Proof. In Ω_0^c, $\hat{E}_n^k(\alpha)\hat{\eta}_k^{[0]}$ is a continuous function of α. Since $\|\hat{E}_n^k(0)\hat{\eta}_k^{[0]}\|=1$, let $\hat{T}^{(1)}=\inf\{\alpha,\|\hat{E}_n^k(\alpha)\hat{\eta}_k^{[0]}\|\le\frac{1}{2}\}\bigwedge\alpha_0$ in Ω_0^c. In Ω_0, define $\hat{T}^{(1)}=0$. Then $\hat{T}^{(1)}$ is a nonnegative random variable. By Lemma 12, we have in Ω_0^c, if $\alpha\le\hat{T}^{(1)}$,

$$\hat{\eta}_k^{[\alpha]}=\frac{\hat{E}_n^k(\alpha)\hat{\eta}_k^{[0]}}{\|\hat{E}_n^k(\alpha)\hat{\eta}_k^{[0]}\|}.$$

Define a random element

$$\zeta_1=\frac{\hat{E}_n^k(\hat{T}^{(1)})\hat{\eta}_k^{[0]}}{\|\hat{E}_n^k(\hat{T}^{(1)})\hat{\eta}_k^{[0]}\|}$$

in Ω_0^c and 0 in Ω_0. Define a random variable $\hat{T}^{(2)}=\inf\{\alpha\ge\hat{T}^{(1)},\|\hat{E}_n^k(\alpha)\zeta_1\|\le\frac{1}{2}\}\bigwedge\alpha_0$ and a random element

$$\zeta_2=\frac{\hat{E}_n^k(\hat{T}^{(2)})\zeta_1}{\|\hat{E}_n^k(\hat{T}^{(1)})\zeta_1\|}$$

in Ω_0^c and 0 in Ω_0. Similarly, we can define $(\hat{T}^{(3)},\zeta_3),\ldots$. One can show that for any $\omega\in\Omega^c$, there are only finite $\hat{T}^{(m)}(\omega)<\alpha_0$, $m=0,1,2,\ldots$, where $\hat{T}^{(0)}(\omega)=0$. Hence in Ω^c, we have

$$\hat{\eta}_k^{[\alpha]}=\sum_{m=0}^{\infty}\frac{\hat{E}_n^k(\alpha)\zeta_m}{\|\hat{E}_n^k(\alpha)\zeta_m\|}\chi_{[\hat{T}^{(m)},\hat{T}^{(m+1)})},$$

where $\zeta_0=\hat{\eta}_k^{[0]}$ and χ is the indicator function. Hence, $\hat{\eta}_k^{[\alpha]}$ is measurable. □

By (5.33) and (5.36), in the event $\{\|\hat{\Gamma}_n-\Gamma\|\le\frac{\epsilon_K}{4}\}\subset\Omega_0^c$, for any $0\le\alpha\le\alpha_0$, $1\le k\le K$, the disk bounded by the circle C_k only contains the k-th eigenvalues for $S_\alpha\hat{\Gamma}S_\alpha$ and $S_\alpha\Gamma S_\alpha$. Hence, in the event $\{\|\hat{\Gamma}_n-\Gamma\|\le\frac{\epsilon_K}{4}\}$, for any $0\le\alpha\le\alpha_0$, $1\le k\le K$, we have

$$E^k(\alpha)=\frac{1}{2\pi\mathrm{i}}\int_{C_k}(zI-S_\alpha\Gamma S_\alpha)^{-1}\mathrm{d}z,\qquad \hat{E}_n^k(\alpha)=\frac{1}{2\pi\mathrm{i}}\int_{C_k}(zI-S_\alpha\hat{\Gamma}_nS_\alpha)^{-1}\mathrm{d}z. \tag{5.40}$$

The proofs of the following Lemmas 16 and 17 follow the ideas of Section 2 in [3]. Define linear maps $\phi_k: C_H[0,\alpha_0]\to C_H[0,\alpha_0]$, $1\le k\le K$ such that for any $\Lambda\in C_H[0,\alpha_0]$ and $0\le\alpha\le\alpha_0$,

$$(\phi_k(\Lambda))(\alpha)=\frac{1}{2\pi\mathrm{i}}\int_{C_k}[(zI-S_\alpha\Gamma S_\alpha)^{-1}\Lambda(\alpha)(zI-S_\alpha\Gamma S_\alpha)^{-1}]\mathrm{d}z, \tag{5.41}$$

where $(\phi_k(\Lambda))(\alpha)$ denotes the value of $\phi_k(\Lambda)$ at the point α. Then define $\Phi_K = (\phi_1, \phi_2, \ldots, \phi_K)$ which is a linear map from $C_H[0, \alpha_0]$ to $\prod_{k=1}^K C_H[0, \alpha_0]$. One can verify that ϕ_k's are continuous. Hence Φ_K is a bounded operator.

Lemma 16. *The sequence* $\{\sqrt{n}(\hat{E}_n^k - E^k), 1 \le k \le K\}_n$ *of stochastic processes has sample paths in* $\prod_{k=1}^K C_H[0, \alpha_0]$ *a.s. and converges in distribution to a Gaussian random element with mean zero and covariance operator* $\Phi_K \Theta i Q i^* \Theta^* \Phi_K^*$.

Proof. In the event $\{\|\hat{\Gamma}_n - \Gamma\| \le \frac{\epsilon_K}{4}\}$, for each $z \in C_k$,

$$\begin{aligned}(zI - S_\alpha \hat{\Gamma}_n S_\alpha)^{-1} &= ((zI - S_\alpha \Gamma S_\alpha) - (S_\alpha \hat{\Gamma} S_\alpha - S_\alpha \Gamma_n S_\alpha))^{-1} \\ &= (zI - S_\alpha \Gamma S_\alpha)^{-1}(I - (S_\alpha \hat{\Gamma} S_\alpha - S_\alpha \Gamma_n S_\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1})^{-1}. \qquad (5.42)\end{aligned}$$

If

$$\sup_{0 \le \alpha \le \alpha_0} \|S_\alpha \hat{\Gamma} S_\alpha - S_\alpha \Gamma_n S_\alpha\| = \|\hat{\Gamma}_n - \Gamma\| \le \frac{1}{2}\tilde{\epsilon},$$

where

$$\tilde{\epsilon}^{-1} = \max_{1 \le k \le K} \sup_{0 \le \alpha \le \alpha_0} \sup_{z \in C_k} \|(zI - S_\alpha \Gamma S_\alpha)^{-1}\| < \infty,$$

then by (5.42), we have an absolutely convergent series expansion

$$(zI - S_\alpha \hat{\Gamma}_n S_\alpha)^{-1} = (zI - S_\alpha \Gamma S_\alpha)^{-1} \sum_{m=0}^{\infty} ((S_\alpha \Gamma S_\alpha - S_\alpha \hat{\Gamma}_n S_\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1})^m.$$

Hence,

$$(zI - S_\alpha \hat{\Gamma}_n S_\alpha)^{-1} - (zI - S_\alpha \Gamma S_\alpha)^{-1} = (zI - S_\alpha \Gamma S_\alpha)^{-1}(S_\alpha \Gamma S_\alpha - S_\alpha \hat{\Gamma}_n S_\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1} + \hat{U}_\alpha^n(z) \qquad (5.43)$$

where

$$\hat{U}_\alpha^n(z) = (zI - S_\alpha \Gamma S_\alpha)^{-1} \sum_{m=2}^{\infty} ((S_\alpha \Gamma S_\alpha - S_\alpha \hat{\Gamma}_n S_\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1})^m.$$

Hence, in the event $\{\|\hat{\Gamma}_n - \Gamma\| \le \frac{1}{2}\tilde{\epsilon}\}$,

$$\|\hat{U}_\alpha^n(z)\| \le \frac{2}{\tilde{\epsilon}^3}\|\hat{\Gamma}_n - \Gamma\|^2. \qquad (5.44)$$

Now in the event $\left\{\|\hat{\Gamma}_n - \Gamma\| \le \min\left(\frac{1}{2}\tilde{\epsilon}, \frac{\epsilon_K}{4}\right)\right\}$, by (5.42) and (5.43),

$$\begin{aligned}\sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha)) &= \frac{\sqrt{n}}{2\pi \mathrm{i}} \int_{C_k} [(zI - S_\alpha \hat{\Gamma} S_\alpha)^{-1} \mathrm{d}z - (zI - S_\alpha \Gamma_n S_\alpha)^{-1}] \mathrm{d}z \\ &= \phi_k(Z_n) + \frac{1}{2\pi \mathrm{i}} \int_{C_k} \sqrt{n} \hat{U}_\alpha^n(z) \mathrm{d}z. \qquad (5.45)\end{aligned}$$

Now we have from (5.44) and (5.45), for any $\delta > 0$,

$$\begin{aligned}P(\|\sqrt{n}(\hat{E}_n^k - E_n^k) - \phi_k(Z_n)\| > \delta) &\le P\left(\|\hat{\Gamma}_n - \Gamma\| > \min\left(\frac{1}{2}\tilde{\epsilon}, \frac{\epsilon_K}{4}\right)\right) \\ &\quad + P\left(\|\sqrt{n}(\hat{E}_n^k - E_n^k) - \phi_k(Z_n)\| > \delta, \|\hat{\Gamma}_n - \Gamma\| \le \min\left(\frac{1}{2}\tilde{\epsilon}, \frac{\epsilon_K}{4}\right)\right) \\ &\le P\left(\|\hat{\Gamma}_n - \Gamma\| > \min\left(\frac{1}{2}\tilde{\epsilon}, \frac{\epsilon_K}{4}\right)\right) + P\left(\|\frac{\sqrt{n}}{2\pi \mathrm{i}} \int_{C_k} \hat{U}_n(z) \mathrm{d}z\| > \delta\right) \\ &\le P\left(\|\hat{\Gamma}_n - \Gamma\| > \min\left(\frac{1}{2}\tilde{\epsilon}, \frac{\epsilon_K}{4}\right)\right) + P(\sqrt{n}\|\hat{\Gamma}_n - \Gamma\|^2 > \pi \delta \tilde{\epsilon}^3) \to 0, \qquad (5.46)\end{aligned}$$

as $n \to 0$. By Lemmas 8 and 13, $\Phi_K(Z_n) = (\phi_1(Z_n), \phi_2(Z_n), \ldots, \phi_K(Z_n))$ converges in distribution to the Gaussian element with mean zero and covariance operator $\Phi_K \Theta i Q i^* \Theta^* \Phi_K^*$. Now by (5.46), $\{\sqrt{n}(\hat{E}_n^k - E_n^k), 1 \le k \le K\}_n$ converges in distribution to the same distribution. □

Define the linear maps $\psi_k : C_H[0, \alpha_0] \to C_{L^2([a,b])}[0, \alpha_0]$, $1 \le k \le K$ such that for any $\Lambda \in C_H[0, \alpha_0]$,

$$\psi_k(\Lambda) = \{(I - E^k(\alpha))\Lambda(\alpha)\eta_k^{[\alpha]}, 0 \le \alpha \le \alpha_0\}. \qquad (5.47)$$

Then we define a linear map $\Psi_K : \Pi_{k=1}^{K} C_H[0,\alpha_0] \to \Pi_{k=1}^{K} C_{L^2([a,b])}[0,\alpha_0]$ such that for any $(\Lambda_1,\ldots,\Lambda_K) \in \Pi_{k=1}^{K} C_H[0,\alpha_0]$,

$$\Psi_K(\Lambda_1,\ldots,\Lambda_K) = (\psi_1(\Lambda_1),\ldots,\psi_K(\Lambda_K)). \tag{5.48}$$

It is easy to see that Ψ_K is a bounded operator.

Lemma 17. *The sequence $\{\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}),\ 1 \le k \le K\}_n$ of stochastic processes has sample paths in $\prod_{k=1}^{K} C_{L^2([a,b])}[0,\alpha_0]$ a.s. and converges in distribution to a Gaussian random element with mean zero and covariance operator $\Psi_K \Phi_K \Theta i Q i^* \Theta^* \Phi_K^* \Psi_K^*$.*

Proof. By the definitions (5.29) of $\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}$, $1 \le k \le K$, $(\eta_k^{[\alpha]}, \eta_k^{[\alpha]})^2 = \|\eta_k^{[\alpha]}\|^2 = 1$. In Ω_0^c, we have

$$\begin{aligned}\sup_{0\le\alpha\le\alpha_0} \sqrt{n}|(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})^2 - 1| &= \sup_{0\le\alpha\le\alpha_0} \sqrt{n}|(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})^2 - (\eta_k^{[\alpha]}, \eta_k^{[\alpha]})^2| \\ &= \sup_{0\le\alpha\le\alpha_0} \sqrt{n}|((\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - ((\eta_k^{[\alpha]}, \eta_k^{[\alpha]})\eta_k^{[\alpha]}, \eta_k^{[\alpha]})| \\ &= \sup_{0\le\alpha\le\alpha_0} \sqrt{n}|(\eta_k^{[\alpha]}, (\hat{E}_n^k(\alpha) - E^k(\alpha))\eta_k^{[\alpha]})|.\end{aligned}$$

By (5.46), $\sqrt{n}(\hat{E}_n^k(\alpha) - E^k(\alpha))$ and $\phi_k(Z_n)$ have the same limit distribution. Because for any $\Lambda \in C_H[0,\alpha_0]$,

$$\begin{aligned}(\eta_k^{[\alpha]}, \phi_k(\Lambda)\eta_k^{[\alpha]}) &= \left(\eta_k^{[\alpha]}, \int_{C_k} [(zI - S_\alpha \Gamma S_\alpha)^{-1}\Lambda(\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1}]\mathrm{d}z\,\eta_k^{[\alpha]}\right) \\ &= \int_{C_k} (\eta_k^{[\alpha]}, (zI - S_\alpha \Gamma S_\alpha)^{-1}\Lambda(\alpha)(zI - S_\alpha \Gamma S_\alpha)^{-1}\eta_k^{[\alpha]})\mathrm{d}z \\ &= \int_{C_k} (z - \lambda_k^{[\alpha]})^{-2}\mathrm{d}z\,(\eta_k^{[\alpha]}, \Lambda(\alpha)\eta_k^{[\alpha]}) = 0\end{aligned} \tag{5.49}$$

where we use the facts that

$$(zI - S_\alpha \Gamma S_\alpha)^{-1}\eta_k^{[\alpha]} = (z - \lambda_k^{[\alpha]})^{-1}\eta_k^{[\alpha]}, \qquad \int_{C_k} (z - \lambda_k^{[\alpha]})^{-2}\mathrm{d}z = 0.$$

So we have

$$\sup_{0\le\alpha\le\alpha_0} \sqrt{n}|(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})^2 - 1| \to 0$$

in probability. By (5.39) and the continuities of $\hat{\eta}_k^{[\alpha]}$ and $\eta_k^{[\alpha]}$, we have

$$\sup_{0\le\alpha\le\alpha_0} \sqrt{n}|(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - 1| \to 0, \tag{5.50}$$

in probability. Now

$$\begin{aligned}\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}) &= \sqrt{n}E^k(\alpha)(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}) + \sqrt{n}(I - E^k(\alpha))(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}) \\ &= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - 1)\eta_k^{[\alpha]} + \sqrt{n}(I - E^k(\alpha))\hat{\eta}_k^{[\alpha]} \\ &= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - 1)\eta_k^{[\alpha]} + \sqrt{n}(I - E^k(\alpha))\frac{\hat{E}_n^k(\alpha)\eta_k^{[\alpha]}}{(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})} \\ &= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - 1)\eta_k^{[\alpha]} + \frac{1}{(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})}(I - E^k(\alpha))\sqrt{n}(\hat{E}_n^k(\alpha) - E^k(\alpha))\eta_k^{[\alpha]} \\ &= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) - 1)\eta_k^{[\alpha]} + \frac{1}{(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]})}\psi_k(\sqrt{n}(\hat{E}_n^k(\alpha) - E^k(\alpha))).\end{aligned} \tag{5.51}$$

By (5.50), the first term in the last line converges to 0 in probability and $(\hat{\eta}_k^{[\alpha]}, \eta_k^{[\alpha]}) \to 1$ in probability. Hence, $(\sqrt{n}(\hat{\eta}_1^{[\alpha]} - \eta_1^{[\alpha]}),\ldots,\sqrt{n}(\hat{\eta}_K^{[\alpha]} - \eta_K^{[\alpha]}))$ has the same limit distribution as $\Psi_K(\sqrt{n}(\hat{E}_n^k(\alpha) - E^k(\alpha)))$ which converges to a Gaussian random element with mean zero and covariance operator $\Psi_K \Phi_K \Theta i Q i^* \Theta^* \Phi_K^* \Psi_K^*$ by Lemmas 8 and 16. □

Define the linear maps $\mathfrak{U}_k : C_H[0,\alpha_0] \to C_{\mathbb{R}}[0,\alpha_0]$, $1 \le k \le K$, such that for any $\Lambda \in C_H[0,\alpha_0]$,

$$\mathfrak{U}_k(\Lambda) = \{(\eta_k^{[\alpha]}, E^k(\alpha)(\psi_k(\Lambda))(\alpha)) + (\eta_k^{[\alpha]}, \Lambda(\alpha)\eta_k^{[\alpha]}) + ((\psi_k(\Lambda))(\alpha), E^k(\alpha)\eta_k^{[\alpha]}),\ 0 \le \alpha \le \alpha_0\},$$

where ψ_k is defined in (5.47), $(\psi_k(\Lambda))(\alpha)$ denotes the value of $\psi_k(\Lambda)$ at α. Define a linear map $\mho_K : \prod_{k=1}^K C_H[0,\alpha_0] \to \prod_{k=1}^K C_{\mathbb{R}}[0,\alpha_0]$ such that for any $(\Lambda_1,\ldots,\Lambda_K)$,

$$\mho_K(\Lambda_1,\ldots,\Lambda_K) = (\mathfrak{U}_1(\Lambda_1),\ldots,\mathfrak{U}_K(\Lambda_K)). \tag{5.52}$$

It is easy to see that $\mho_K$ is a bounded operator.

Lemma 18. *The sequence $\{\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}), 1 \le k \le K\}_n$ of stochastic processes has sample paths in $\prod_{k=1}^K C_{\mathbb{R}}[0,\alpha_0]$ a.s. and converges in distribution to a Gaussian random element with zero and covariance operator $\mho_K \Phi_K \Theta iQi^* \Theta^* \Phi_K^* \mho_K^*$.*

Proof. The continuities of $\hat{\lambda}_k^{[\alpha]}$ and $\lambda_k^{[\alpha]}$ follow from Lemma 11 and the inequalities

$$|\hat{\lambda}_k^{[\alpha]} - \hat{\lambda}_k^{[\alpha']}| \le \|S_\alpha \hat{\Gamma} S_\alpha - S_{\alpha'}\hat{\Gamma} S_{\alpha'}\|, \qquad |\lambda_k^{[\alpha]} - \lambda_k^{[\alpha']}| \le \|S_\alpha \Gamma S_\alpha - S_{\alpha'}\Gamma S_{\alpha'}\|,$$

for any $0 \le \alpha, \alpha' \le \alpha_0$. In Ω_0^c,

$$\begin{aligned}
\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}) &= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \hat{E}_n^k(\alpha)\hat{\eta}_k^{[\alpha]}) - (\eta_k^{[\alpha]}, E^k(\alpha)\eta_k^{[\alpha]})) \\
&= \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \hat{E}_n^k(\alpha)\hat{\eta}_k^{[\alpha]}) - (\hat{\eta}_k^{[\alpha]}, \hat{E}_n^k(\alpha)\eta_k^{[\alpha]})) + \sqrt{n}((\hat{\eta}_k^{[\alpha]}, \hat{E}_n^k(\alpha)\eta_k^{[\alpha]}) - (\hat{\eta}_k^{[\alpha]}, E_n^k(\alpha)\eta_k^{[\alpha]})) \\
&\quad + \sqrt{n}((\hat{\eta}_k^{[\alpha]}, E_n^k(\alpha)\eta_k^{[\alpha]}) - (\eta_k^{[\alpha]}, E_n^k(\alpha)\eta_k^{[\alpha]})) \\
&= (\hat{\eta}_k^{[\alpha]}, \hat{E}_n^k(\alpha)\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]})) + (\hat{\eta}_k^{[\alpha]}, \sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha))\eta_k^{[\alpha]}) \\
&\quad + (\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}), E_n^k(\alpha)\eta_k^{[\alpha]}).
\end{aligned} \tag{5.53}$$

By Lemmas 16 and 17, $\hat{\eta}_k^{[\alpha]} \to \eta_k^{[\alpha]}$ and $\hat{E}_k^{[\alpha]} \to E_k^{[\alpha]}$ in probability. Hence, by (5.53), $\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]})$ has the same limit distribution as

$$(\eta_k^{[\alpha]}, E_n^k(\alpha)\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]})) + (\eta_k^{[\alpha]}, \sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha))\eta_k^{[\alpha]}) + (\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}), E_n^k(\alpha)\eta_k^{[\alpha]})$$

which, by (5.51), has the same distribution as

$$\begin{aligned}
&(\eta_k^{[\alpha]}, E_n^k(\alpha)\psi_k(\sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha)))) + (\eta_k^{[\alpha]}, \sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha))\eta_k^{[\alpha]}) \\
&\quad + (\psi_k(\sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha))), E_n^k(\alpha)\eta_k^{[\alpha]}) \\
&= \mathfrak{U}_k(\sqrt{n}(\hat{E}_n^k(\alpha) - E_n^k(\alpha))).
\end{aligned}$$

Hence, $\{\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}), 1 \le k \le K\}_n$ has the same limit distribution as $\mho_K(\sqrt{n}(\hat{E}_n^1(\alpha) - E_n^1(\alpha)),\ldots,\sqrt{n}(\hat{E}_n^K(\alpha) - E_n^K(\alpha)))$ which converges to a Gaussian random element with mean zero and covariance operator $\mho_K \Phi_K \Theta iQi^* \Theta^* \Phi_K^* \mho_K^*$ by Lemmas 8 and 16. $\square$

Define a linear map $\Im_K : \prod_{k=1}^K C_{L^2([a,b])}[0,\alpha_0] \to \prod_{k=1}^K C_{L^2([a,b])}[0,\alpha_0]$ such that for any $(\Lambda_1,\ldots,\Lambda_K) \in \prod_{k=1}^K C_{L^2([a,b])}[0,\alpha_0]$,

$$\Im_K(\Lambda_1,\ldots,\Lambda_K) = \left\{\left(\frac{1}{\|S_\alpha \eta_1^{[\alpha]}\|}S_\alpha \Lambda_1(\alpha),\ldots,\frac{1}{\|S_\alpha \eta_K^{[\alpha]}\|}S_\alpha \Lambda_K(\alpha)\right), 0 \le \alpha \le \alpha_0\right\}. \tag{5.54}$$

$\Im_K$ is a bounded operator.

Lemma 19. *The sequence $\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K\}_n$ of stochastic processes has sample paths in $\prod_{k=1}^K C_{L^2([a,b])}[0,\alpha_0]$ a.s. and converges in distribution to a Gaussian random element with mean zero and covariance operator $\Im_K \Psi_K \Phi_K \Theta iQi^* \Theta^* \Phi_K^* \Psi_K^* \Im_K^*$.*

Proof. By (5.31),

$$\hat{\gamma}_k^{[\alpha]} = \frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \hat{\eta}_k^{[\alpha]}\|}, \qquad \gamma_k^{[\alpha]} = \frac{S_\alpha \eta_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|}.$$

Therefore,

$$\begin{aligned}
\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}) &= \sqrt{n}\left(\frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \hat{\eta}_k^{[\alpha]}\|} - \frac{S_\alpha \eta_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|}\right) \\
&= \sqrt{n}\left(\frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \hat{\eta}_k^{[\alpha]}\|} - \frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|}\right) + \sqrt{n}\left(\frac{S_\alpha \hat{\eta}_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|} - \frac{S_\alpha \eta_k^{[\alpha]}}{\|S_\alpha \eta_k^{[\alpha]}\|}\right) \\
&= \sqrt{n}\left(\frac{1}{\|S_\alpha \hat{\eta}_k^{[\alpha]}\|} - \frac{1}{\|S_\alpha \eta_k^{[\alpha]}\|}\right)S_\alpha \hat{\eta}_k^{[\alpha]} + \frac{1}{\|S_\alpha \eta_k^{[\alpha]}\|}S_\alpha(\sqrt{n}(\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]})).
\end{aligned} \tag{5.55}$$

Because

$$\|S_\alpha \hat{\eta}_k^{[\alpha]} - S_\alpha \eta_k^{[\alpha]}\| \le \|\hat{\eta}_k^{[\alpha]} - \eta_k^{[\alpha]}\| \to 0$$

in probability, by the definition (5.54) of $\Im_K$, (5.55) and Lemma 17, $\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K\}_n$ has the same limit distribution as $\Im_K(\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K\}_n)$ which converges to a Gaussian random element with mean zero and covariance operator $\Im_K \Psi_K \Phi_K \Theta iQi^* \Theta^* \Phi_K^* \Psi_K^* \Im_K^*$. □

Proof of Corollary 4.1. By Lemmas 18 and 19, the stochastic processes $\{\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}), 1 \le k \le K\}_n$ and $\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K\}_n$ convergence in distribution, hence they are tight by Theorem 5.2 in [1] since $C_{\mathbb{R}}[0, \alpha_0]$ and $C_{L^2[a,b]}[0, \alpha_0]$ are both complete and separable. Therefore, for any $\epsilon > 0$, one can find a positive number M depending on ϵ such that

$$\sup_n P(\max_{1\le k\le K} \sup_{0\le\alpha\le\alpha_0} |\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]})| \ge M) \le \epsilon,$$

$$\sup_n P(\max_{1\le k\le K} \sup_{0\le\alpha\le\alpha_0} |\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]})| \ge M) \le \epsilon.$$

In other words,

$$\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]} = O_p\left(\frac{1}{\sqrt{n}}\right),$$

$$\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]} = O_p\left(\frac{1}{\sqrt{n}}\right)$$

uniformly in α, which combines Theorem 4.1 to get our corollary. □

Proof of Corollary 4.2. First, we have decompositions

$$\sqrt{n}(\hat{\lambda}_k^{[\alpha_n]} - \lambda_k) = \sqrt{n}(\hat{\lambda}_k^{[\alpha_n]} - \lambda_k^{[\alpha_n]}) + \sqrt{n}(\lambda_k^{[\alpha_n]} - \lambda_k),$$
$$\sqrt{n}(\hat{\gamma}_k^{[\alpha_n]} - \gamma_k) = \sqrt{n}(\hat{\gamma}_k^{[\alpha_n]} - \gamma_k^{[\alpha_n]}) + \sqrt{n}(\gamma_k^{[\alpha_n]} - \gamma_k).$$

Under the conditions on α_n for eigenvalues and eigenfunctions respectively, by Theorem 4.1, we have $\sqrt{n}(\lambda_k^{[\alpha_n]} - \lambda_k) \to 0$ and $\sqrt{n}(\gamma_k^{[\alpha_n]} - \gamma_k) \to 0$ respectively. Since $\{\sqrt{n}(\hat{\gamma}_k^{[\alpha]} - \gamma_k^{[\alpha]}), 1 \le k \le K, 0 \le \alpha \le \alpha_0\}_n$ and $\{\sqrt{n}(\hat{\lambda}_k^{[\alpha]} - \lambda_k^{[\alpha]}), 1 \le k \le K, 0 \le \alpha \le \alpha_0\}_n$ converge in distribution by Theorem 4.2, they are tight. Hence, the asymptotic normalities of $\sqrt{n}(\hat{\lambda}_k^{[\alpha_n]} - \lambda_k^{[\alpha_n]})$ and $\sqrt{n}(\hat{\gamma}_k^{[\alpha_n]} - \gamma_k^{[\alpha_n]})$ follow from Theorem 4.2 and the following lemma. Then the corollary follows at once. □

Lemma 20. *Suppose that F is a metric space with distance d. Let $C_F[0, \alpha_0]$ denote the continuous function on $[0, \alpha_0]$ taking values in F. Suppose we have a sequence $\{Y_n(\alpha), 0 \le \alpha \le \alpha_0, n \in \mathbb{N}\}$ of stochastic processes has sample paths in $C_F[0, \alpha_0]$. Assume that Y_n is tight and $Y_n(0)$ converges in distribution to a random element Y in F, then for any sequence α_n of positive numbers converging to 0, $Y_n(\alpha_n)$ also converges in distribution to Y.*

Proof. First, we show that for any $\epsilon > 0$ we can find $\delta > 0$ such that

$$\sup_n P(\sup_{0\le\alpha',\alpha''\le\delta} d(Y_n(\alpha'), Y_n(\alpha'')) > \epsilon) \le \epsilon.$$

Since Y_n is tight, we can find a compact subset Ξ of $C_F[0, \alpha_0]$ such that

$$\sup_n P(Y_n \notin \Xi) \le \epsilon.$$

We can find a finite number of $\Lambda_1, \dots \Lambda_m \in \Xi$ such that for any $\Lambda \in \Xi$, we can find i such that $\sup_{0\le\alpha\le\alpha_0} d(\Lambda_i(\alpha), \Lambda(\alpha)) \le \frac{\epsilon}{3}$. Furthermore, we can find $\delta > 0$ such that,

$$\max_{1\le i\le m} \sup_{0\le\alpha',\alpha''\le\delta} d(\Lambda_i(\alpha'), \Lambda_i(\alpha'')) \le \frac{\epsilon}{3}.$$

Now it is easy to see that for any $\Lambda \in \Xi$,

$$\sup_{0\le\alpha',\alpha''\le\delta} d(\Lambda(\alpha'), \Lambda(\alpha'')) \le \epsilon.$$

Hence,

$$\sup_n P(\sup_{0\le\alpha',\alpha''\le\delta} d(Y_n(\alpha'), Y_n(\alpha'')) > \epsilon) \le \sup_n P(Y_n \notin \Xi) \le \epsilon.$$

If $\alpha_n \leq \delta$, we have

$$P(d(Y_n(0), Y_n(\alpha_n)) > \epsilon) \leq \epsilon.$$

Since ϵ is arbitrary, $d(Y_n(0), Y_n(\alpha_n)) \to 0$ in probability. □

References

[1] P. Billingsley, Convergence of Probability Measures, 2nd ed., Wiley-Interscience, 1999.
[2] H. Cardot, F. Ferraty, P. Sarda, Functional linear model, Statist. Probab. Lett. 45 (1999) 11–22.
[3] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a random vector function: some application to statistical inference, J. Multivariate Anal. 12 (1982) 136–154.
[4] N. Dunford, J.T. Schwartz, Linear Operators, General Theory, Part 1, Wiley-Interscience, 1988.
[5] N. Dunford, J.T. Schwartz, Linear Operators, Spectral Theory, Self Adjoint Operators in Hilbert Space, Part 2, Wiley-Interscience, 1988.
[6] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: Theory and Practice, Springer, 2006.
[7] P. Glasserman, Monte Carlo Methods in Financial Engineering, 1st ed., Springer, 2003.
[8] J.Z. Huang, H. Shen, A. Buja, Functional principal components analysis via penalized rank one approximation, Electron. J. Stat. 2 (2008) 678–695.
[9] P.D. Lax, Functional Analysis, in: Pure and Applied Mathematics: A Wiley-Interscience Series of Texts, Monographs and Tracts, Wiley-Interscience, 2002.
[10] M. Ledoux, M. Talagrand, Probability in Banach Spaces: Isoperimetry and Processes, in: Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, Band 23. A Series of Modern Surveys in Mathematics, Springer, 2006.
[11] X. Qi, H. Zhao, Functional principal component analysis for discretely observed functional data (2010) (submitted for publication, under review).
[12] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, 2nd ed., Springer, New York, 2005.
[13] F. Riesz, B. Sz.-Nagy, Functional Analysis, Dover Publications, 1990.
[14] W. Rudin, Functional Analysis, 2nd ed., McGraw-Hill Science/Engineering/Math, 1991.
[15] B.W. Silverman, Smoothed functional principal components analysis by choice of norm, Ann. Statist. 24 (1996) 1–24.
[16] H.F. Weinberger, Variational Methods for Eigenvalue Approximation, 2nd ed., in: CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial Mathematics, 1987.