Bernoulli **25**(4B), 2019, 3864–3882
https://doi.org/10.3150/19-BEJ1113

Principal components analysis of regularly varying functions

PIOTR KOKOSZKA[1], STILIAN STOEV[2] and QIAN XIONG[1]

[1]*Colorado State University. E-mail: Piotr.Kokoszka@colostate.edu*
[2]*University of Michigan.*

The paper is concerned with asymptotic properties of the principal components analysis of functional data. The currently available results assume the existence of the fourth moment. We develop analogous results in a setting which does not require this assumption. Instead, we assume that the observed functions are regularly varying. We derive the asymptotic distribution of the sample covariance operator and of the sample functional principal components. We obtain a number of results on the convergence of moments and almost sure convergence. We apply the new theory to establish the consistency of the regression operator in a functional linear model.

Keywords: functional data; principal components; regular variation

1. Introduction

A fundamental technique of functional data analysis is to replace infinite dimensional curves by coefficients of their projections onto suitable, fixed or data-driven, systems, see, for example, Bosq [2], Ramsay and Silverman [27], Horváth and Kokoszka [11], Hsing and Eubank [13]. A finite number of these coefficients encode the shape of the curves and are amenable to various statistical procedures. The best systems are those that lead to low dimensional representations, and so provide the most efficient dimension reduction. Of these, the functional principal components (FPCs) have been most extensively used, with hundreds of papers dedicated to the various aspects of their theory and applications.

If $X, X_1, X_2, \dots, X_N$ are mean zero i.i.d. functions in L^2 with $E\|X\|^2 < \infty$, then

$$X_n(t) = \sum_{j=1}^{\infty} \xi_{nj} v_j(t), \qquad E\xi_{nj}^2 = \lambda_j. \tag{1.1}$$

The FPCs v_j and the eigenvalues λ_j are, respectively, the eigenfunctions and the eigenvalues of the covariance operator $C : L^2 \to L^2$ defined by $C(x)(t) = \int \mathrm{Cov}(X(t), X(s))x(s)\,ds$. As such, the v_j are orthogonal. We assume they are normalized to unit norm. The v_j form an optimal orthonormal basis for dimension reduction measured by the L^2 norm, see, for example, Theorem 11.4.1 in Kokoszka and Reimherr [17].

The v_j and the λ_j are estimated by $\hat{v}_j$ and $\hat{\lambda}_j$ defined by

$$\int \hat{c}(t,s)\hat{v}_j(s)\,ds = \hat{\lambda}_j \hat{v}_j(t), \tag{1.2}$$

where

$$\hat{c}(t,s) = \frac{1}{N}\sum_{n=1}^{N} X_n(t)X_n(s). \tag{1.3}$$

Like the v_j, the $\hat{v}_j$ are defined only up to a sign. Thus, strictly speaking, in the formulas that follow, the $\hat{v}_j$ would need to be replaced with $\hat{c}_j\hat{v}_j$, where $\hat{c}_j = \operatorname{sign}\langle \hat{v}_j, v_j\rangle$. As is customary, to lighten the notation, we assume that the orientations of v_j and $\hat{v}_j$ match, that is, $\hat{c}_j = 1$.

Under the existence of the fourth moment,

$$E\|X\|^4 = \left\{\int X^2(t)\,dt\right\}^2 < \infty, \tag{1.4}$$

and assuming $\lambda_1 > \lambda_2 > \cdots$, it has been shown that for each $j \geq 1$,

$$\limsup_{N\to\infty} NE\|\hat{v}_j - v_j\|^2 < \infty, \qquad \limsup_{N\to\infty} NE(\hat{\lambda}_j - \lambda_j)^2 < \infty, \tag{1.5}$$

$$N^{1/2}(\hat{\lambda}_j - \lambda_j) \xrightarrow{d} N\big(0, \sigma_j^2\big), \tag{1.6}$$

$$N^{1/2}(\hat{v}_j - v_j) \xrightarrow{d} N(0, C_j), \tag{1.7}$$

for a suitably defined variance σ_j^2 and a covariance operator C_j. The above relations, especially (1.5), have been used to derive large sample justifications of inferential procedures based on the estimated FPCs $\hat{v}_j$. In most scenarios, one can show that replacing the $\hat{v}_j$ by the v_j and the $\hat{\lambda}_j$ by the λ_j is asymptotically negligible. Relations (1.5) were established by Dauxois et al. [3] and extended to weakly dependent functional time series by Hörmann and Kokoszka [10]. Relations (1.6) and (1.7) follow from the results of Kokoszka and Reimherr [16]. In case of continuous functions satisfying regularity conditions, they follow from the results of Hall and Hosseini-Nasab [9].

A crucial assumption for the relations (1.5)–(1.7) to hold is the existence of the fourth moment, i.e. (1.4), the i.i.d. assumption can be relaxed in many ways. Nothing is at present known about the asymptotic properties of the FPCs and their eigenvalues if (1.4) does not hold. Our objective is to explore what can be said about the asymptotic behavior of $\widehat{C}$, $\hat{v}_j$ and $\hat{\lambda}_j$ if (1.4) fails. We would thus like to consider the case of $E\|X_n\|^2 < \infty$ and $E\|X_n\|^4 = \infty$. Such an assumption is however too general. From mid 1980s to mid 1990s, similar questions were posed for scalar time series for which the fourth or even second moment does not exist. A number of results pertaining to the convergence of sample covariances and the periodogram have been derived under the assumption of regularly varying tails, for example, Davis and Resnick [4,5], Klüppelberg and Mikosch [15], Mikosch et al. [26], Kokoszka and Taqqu [18], Anderson and Meerschaert [1]; many others are summarized in the monograph of Embrechts et al. [8]. The assumption of regular variation is natural because non–normal stable limits can be derived by establishing a connection to random variables in a stable domain of attraction, which is characterized by regular variation. This is the approach we take. We assume that the functions X_n are regularly varying in the space L^2 with the index $\alpha \in (2,4)$, which implies $E\|X_n\|^2 < \infty$ and $E\|X_n\|^4 = \infty$. Suitable definitions and assumptions are presented in Section 2.

The paper is organized as follows. The remainder of the introduction provides a practical motivation for the theory we develop. It is not necessary to understand the contribution of the paper, but, we think, it gives a good feel for what is being studied. The formal exposition begins in Section 2, in which notation and assumptions are specified. Section 3 is dedicated to the convergence of the sample covariance operator (the integral operator with kernel (1.3)). These results are then used in Section 4 to derive various convergence results for the sample FPCs and their eigenvalues. Section 5 shows how the results derived in previous sections can be used in a context of a functional regression model. Its objective is to illustrate the applicability of our theory in a well known and extensively studied setting. It is hoped that it will motivate and guide applications to other problems of functional data analysis. All proofs which go beyond simple arguments are presented in Online material [19].

We conclude this introduction by presenting a specific data context. Denote by $P_i(t)$ the price of an asset at time t of trading day i. For the assets we consider in our illustration, t is time in minutes between 9:30 and and 16:00 EST (NYSE opening times) rescaled to the unit interval $(0, 1)$. The intraday return curve on day i is defined by $X_i(t) = \log P_i(t) - \log P_i(0)$. In practice, $P_i(0)$ is the price after the first minute of trading. The curves X_i show how the return accumulates over the trading day, see e.g. Lucca and Moench [22]; examples of are shown in Figure 1.

The first three sample FPCs, $\hat{v}_1$, $\hat{v}_2$, $\hat{v}_3$, are shown in Figure 2. They are computed, using (1.2), from minute-by-minute Walmart returns form July 05, 2006 to Dec 30, 2011, $N = 1378$ trading days. (This time interval is used for the other assets we consider.) The curves $\widehat{X}_i = \sum_{j=1}^{3} \hat{\xi}_{ij} \hat{v}_j$, with the scores $\hat{\xi}_{ij} = \int X_i(t)\hat{v}_j(t)\,dt$, visually approximate the curves X_i well. One can thus expect that the $\hat{v}_j$ (with properly adjusted sign) are good estimators of the population FPCs v_j in (1.1). Relations (1.5) and (1.7) show that this is indeed the case, *if* $E\|X\|^4 < \infty$. (The curves X_i can be assumed to form a stationary time series in L^2, see Horváth et al. [12].) We will now argue that the assumption of the finite fourth moment is not realistic, so, with the currently available theory, it is not clear if the $\hat{v}_j$ are good estimators of the v_j. If $E\|X\|^4 < \infty$, then $E\xi_{1j}^4 < \infty$ for every j. Figure 3 shows the Hill plots of the sample score $\hat{\xi}_{ij}$ for two stocks and for $j = 1, 2, 3$. Hill plots for other blue chip stocks look similar. These plots illustrate several properties. (1) It is reasonable to assume that the scores have Pareto tails. (2) The tail index α is smaller than

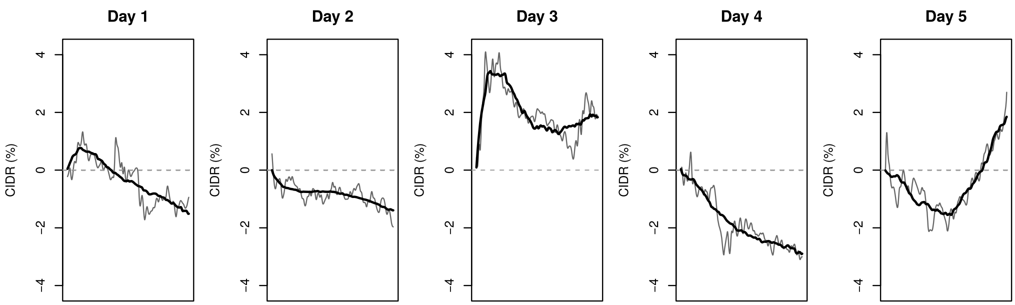

Figure 1. Five consecutive intraday return curves, Walmart stock. The raw returns are noisy grey lines. The smoother black lines are approximations $\widehat{X}_i(t) = \sum_{j=1}^{3} \hat{\xi}_{ij} \hat{v}_j$.

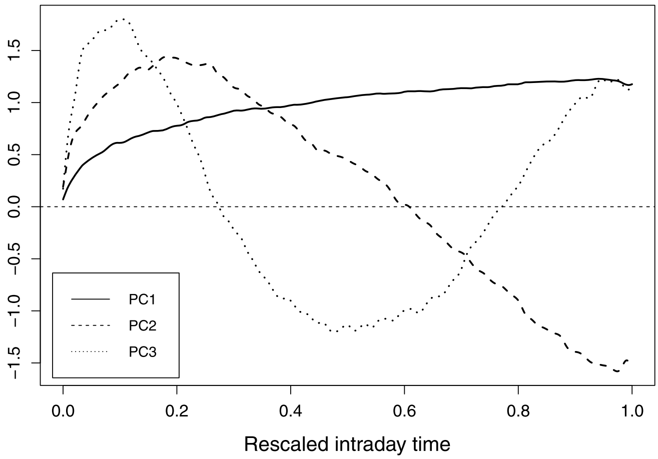

Figure 2. The first three sample FPCs of intraday returns on Walmart stock.

4, implying that the fourth moment does not exist. (3) It is reasonable to assume that the tail index does not depend on j and is between 2 and 4. With such a motivation, we are now able to formalize in the next section the setting of this paper.

2. Preliminaries

The functions X_n are assumed to be independent and identically distributed in L^2, with the same distribution as X, which is regularly varying with index $\alpha \in (2, 4)$. By $L^2 := L^2(\mathcal{T})$, we denote the usual separable Hilbert space of square integrable functions on some compact subset $\mathcal{T}$ of an Euclidean space. In a typical FDA framework, $\mathcal{T} = [0, 1]$, for example, Chapter 2 of Horváth and Kokoszka [11]. Regular variation in finite-dimensional spaces has been a topic of extensive research for decades, see, for example, Resnick [28,29] and Meerschaert and Scheffler [24]. We shall need the concept of regular variation of measures on *infinitely-dimensional* function spaces. To this end, we start by recalling some terminology and fundamental facts about regularly varying functions.

A measurable function $L : (0, \infty) \to \mathbb{R}$ is said to be slowly varying (at infinity) if, for all $\lambda > 0$,

$$\frac{L(\lambda u)}{L(u)} \to 1, \quad \text{as } u \to \infty.$$

Functions of the form $R(u) = u^\rho L(u)$ are said to be regularly varying with exponent $\rho \in \mathbb{R}$.

The notion of regular variation extends to measures and provides an elegant and powerful framework for establishing limit theorems. It was first introduced by Meerschaert [23] and has been since extended to Banach and even metric spaces using the notion of M_0 convergence (see,

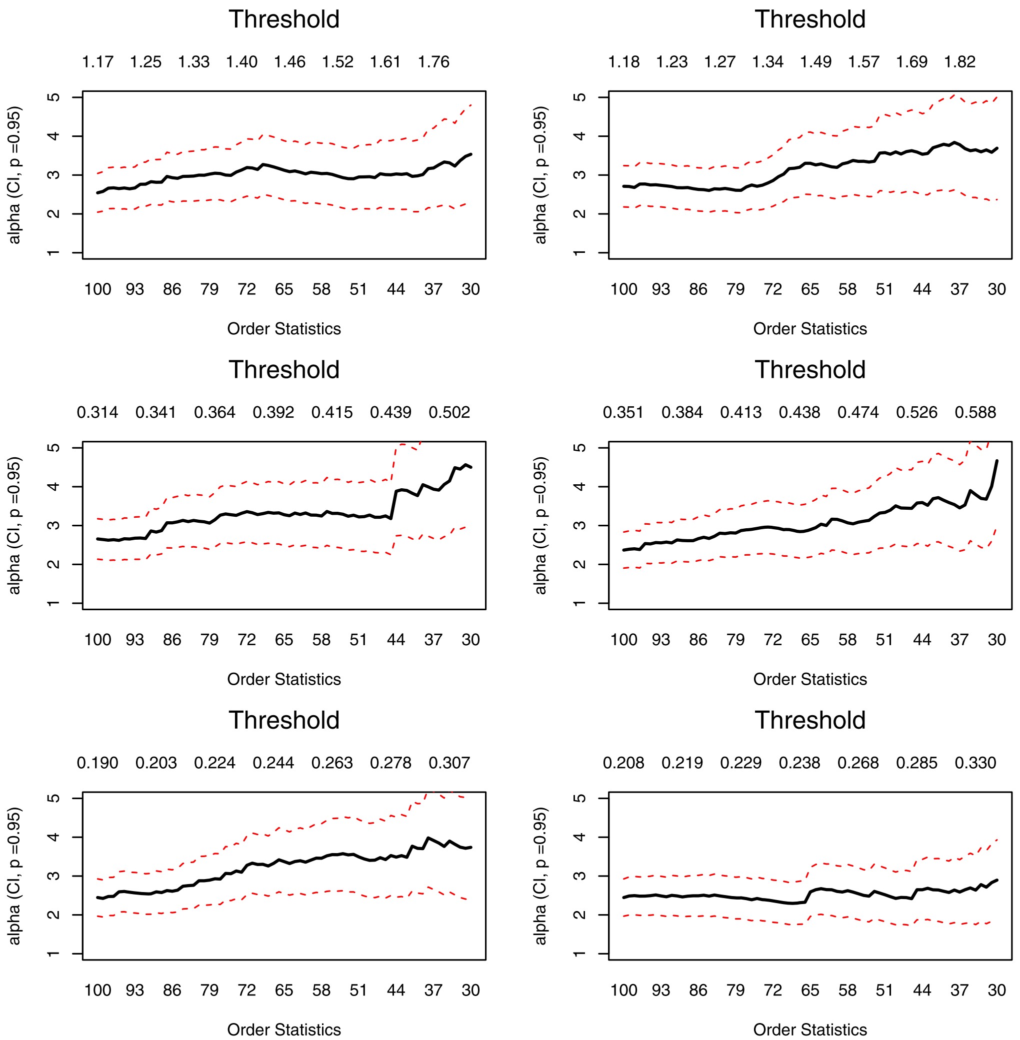

Figure 3. Hill plots (an estimate of α as a function of upper order statistics) for sample FPC scores for **Walmart** (left) and **IBM** (right). From top to bottom: levels $j = 1, 2, 3$.

e.g., Hult and Lindskog [14]). Even though we will work only with Hilbert spaces, we review the theory in a more general context.

Consider a separable Banach space $\mathbb{B}$ and let $B_\epsilon := \{z \in \mathbb{B} : \|z\| < \epsilon\}$ be the open ball of radius $\epsilon > 0$, centered at the origin. A Borel measure μ defined on $\mathbb{B}_0 := \mathbb{B} \setminus \{\mathbf{0}\}$ is said to be *boundedly finite* if $\mu(A) < \infty$, for all Borel sets that are bounded away from $\mathbf{0}$, that is, such that $A \cap B_\epsilon = \varnothing$, for some $\epsilon > 0$. Let $\mathbb{M}_0$ be the collection of all such measures. For $\mu_n, \mu \in \mathbb{M}_0$, we say that the μ_n converge to μ in the M_0 topology, if $\mu_n(A) \to \mu(A)$, for all bounded away from $\mathbf{0}$, μ-continuity Borel sets A, i.e., such that $\mu(\partial A) = 0$, where $\partial A := \overline{A} \setminus A^\circ$ denotes the boundary of A. The M_0 convergence can be metrized such that $\mathbb{M}_0$ becomes a complete separable metric space (Theorem 2.3 in Hult and Lindskog [14] and also Section 2.2 of Meiguet [25]). The following result is known, see, for example, Chapter 2 of Meiguet [25] and references therein.

Proposition 2.1. *Let X be a random element in a separable Banach space $\mathbb{B}$ and $\alpha > 0$. The following three statements are equivalent*:

(i) *For some slowly varying function L,*

$$P\big(\|X\| > u\big) = u^{-\alpha} L(u) \tag{2.1}$$

and

$$\frac{P(u^{-1}X \in \cdot)}{P(\|X\| > u)} \xrightarrow{M_0} \mu(\cdot), \quad u \to \infty, \tag{2.2}$$

where μ is a non-null measure on the Borel σ-field $\mathcal{B}(\mathbb{B}_0)$ of $\mathbb{B}_0 = \mathbb{B}\setminus\{\mathbf{0}\}$.

(ii) *There exists a probability measure Γ on the unit sphere $\mathbb{S}$ in $\mathbb{B}$ such that, for every $t > 0$,*

$$\frac{P(\|X\| > tu, X/\|X\| \in \cdot)}{P(\|X\| > u)} \xrightarrow{w} t^{-\alpha}\Gamma(\cdot), \quad u \to \infty.$$

(iii) *Relation* (2.1) *holds, and for the same spectral measure Γ in* (ii),

$$P\big(X/\|X\| \in \cdot \,|\, \|X\| > u\big) \xrightarrow{w} \Gamma(\cdot), \quad u \to \infty.$$

Definition 2.1. If any one of the equivalent conditions in Proposition 2.1 hold, we shall say that X is regularly varying with index α. The measures μ and Γ will be referred to as exponent and angular measures of X, respectively.

The measure Γ is sometimes called the spectral measure, but we will use the adjective "spectral" in the context of stable measures which appear in Section 3. It is important to distinguish the angular measure of a regularly varying random function and a spectral measure of a stable distribution, although they are related. We also note that we call α the tail index, and $-\alpha$ the tail exponent.

We will work under the following assumption.

Assumption 2.1. The random element X in the separable Hilbert space $H = L^2$ has mean zero and is regularly varying with index $\alpha \in (2, 4)$. The observations $X_1, X_2, \ldots$ are independent copies of X.

Assumption 2.1 is a coordinate free condition not related in any way to functional principal components. The next assumption relates the asymptotic behavior of the FPC scores to the assumed regular variation. It implies, in particular, that the expansion $X(t) = \sum_{j=1}^{\infty} \xi_j v_j(t)$ contains infinitely many terms, so that we study infinite dimensional objects. We will see in the proofs of Proposition 3.1 and Theorem 3.2 that under Assumption 2.1 the limit

$$Q_{nm} = \lim_{u\to\infty} \frac{P(\{\sum_{j=n}^{\infty} \xi_j^2\}^{1/2}\{\sum_{j=m}^{\infty} \xi_j^2\}^{1/2} > u)}{P(\sum_{j=1}^{\infty} \xi_j^2 > u)}$$

exists and is finite. We impose the following assumption related to condition (2.2).

Assumption 2.2. For every $n, m \geq 1$, $Q_{nm} > 0$.

Assumption 2.2 postulates, intuitively, that the tail sums $\sum_{j=n}^{\infty} \xi_j^2$ must have extreme probability tails comparable to that of $\|X\|^2$.

We now collect several useful facts that will be used in the following. The exponent measure μ satisfies

$$\mu(tA) = t^{-\alpha}\mu(A), \quad \forall t > 0, A \in \mathcal{B}(\mathbb{B}_0). \tag{2.3}$$

It admits the polar coordinate representation via the angular measure Γ. That is, if $x = r\theta$, where $r := \|x\|$ and $\theta = x/\|x\|$, for $x \neq \mathbf{0}$, we have

$$\mu(dx) = \alpha r^{-\alpha-1}\, dr\, \Gamma(d\theta). \tag{2.4}$$

This means that for every bounded measurable function f that vanishes on a neighborhood of $\mathbf{0}$, we have

$$\int_{\mathbb{B}} f(x)\mu(dx) = \int_{\mathbb{S}}\int_0^{\infty} f(r\theta)\alpha r^{-\alpha-1}\, dr\, \Gamma(d\theta).$$

There exists a sequence $\{a_N\}$ such that

$$NP(X \in a_N A) \to \mu(A), \tag{2.5}$$

for any set A in $\mathcal{B}(\mathbb{B}_0)$ with $\mu(\partial A) = 0$. One can take, for example,

$$a_N = N^{1/\alpha} L_0(N), \tag{2.6}$$

with a slowly varying function L_0 satisfying $L_0^{-\alpha}(N)L(N^{1/\alpha}L_0(N)) \to 1$.

We will work with Hilbert–Schmidt operators. A linear operator $\Psi : H \to H$ is Hilbert–Schmidt if $\sum_{j=1}^{\infty} \|\Psi(e_j)\|^2 < \infty$, where $\{e_j\}$ is any orthonormal basis of H. Every Hilbert–Schmidt operator is bounded. The space of Hilbert–Schmidt operators will be denoted by $\mathcal{S}$. It is itself a separable Hilbert space with the inner product

$$\langle \Psi_1, \Psi_2\rangle_{\mathcal{S}} = \sum_{j=1}^{\infty}\langle \Psi_1(e_j), \Psi_2(e_j)\rangle.$$

If Ψ is an integral operator defined by $\Psi(x)(t) = \int \psi(t,s)x(s)\, ds$, $x \in L^2$, then $\|\Psi\|_{\mathcal{S}}^2 = \iint \psi^2(t,s)\, dt\, ds$.

Relations (1.5) essentially follow from the bound

$$E\|\widehat{C}-C\|_{\mathcal{S}}^2 \le N^{-1}E\|X\|^4,$$

where the subscript $\mathcal{S}$ indicates the Hilbert–Schmidt norm. Under Assumption 2.1 such a bound is useless because, by (2.1), $E\|X\|^4 = \infty$. In fact, one can show that under Assumption 2.1, $E\|\widehat{C}\|_{\mathcal{S}}^2 = \infty$, so no other bound on $E\|\widehat{C}-C\|_{\mathcal{S}}^2$ can be expected. The following Proposition 2.2 implies however that under Assumption 2.1 the population covariance operator C is a Hilbert–Schmidt operator, and $\widehat{C} \in \mathcal{S}$ with probability 1. This means that the space $\mathcal{S}$ does provide a convenient framework.

Proposition 2.2. *Suppose X is a random element of L^2 with $E\|X\|^2 < \infty$ and $\widehat{C}$ is the sample covariance operator based on N i.i.d. copies of X. Then $C \in \mathcal{S}$ and $\widehat{C} \in \mathcal{S}$ with probability* 1.

Like all proofs, the proof of Proposition 2.2 is presented in the on-line material.

3. Limit distribution of $\widehat{C}$

We will show that $Nk_N^{-1}(\widehat{C}-C)$ converges to an $\alpha/2$-stable Hilbert–Schmidt operator, for an appropriately defined regularly varying sequence $\{k_N\}$. Unless stated otherwise, all limits in the following are taken as $N \to \infty$.

Observe that for any $x \in H$,

$$\begin{aligned}
Nk_N^{-1}(\widehat{C}-C)(x) &= Nk_N^{-1}\left(N^{-1}\sum_{n=1}^N \langle X_n, x\rangle X_n - E\big[\langle X_1, x\rangle X_1\big]\right)\\
&= k_N^{-1}\left(\sum_{n=1}^N \langle X_n, x\rangle X_n - NE\big[\langle X_1, x\rangle X_1\big]\right)\\
&= k_N^{-1}\left(\sum_{n=1}^N (X_n \otimes X_n)(x) - NE\big[(X_1 \otimes X_1)\big](x)\right), \qquad (3.1)
\end{aligned}$$

where $(X_n \otimes X_n)(x) = \langle X_n, x\rangle X_n$. Since the $X_n \otimes X_n$ are Hilbert–Schmidt operators, the last expression shows a connection between the asymptotic distribution of $\widehat{C}$ and convergence to a stable limit in the Hilbert space $\mathcal{S}$ of Hilbert–Schmidt operators. We therefore restate below, as Theorem 3.1, Theorem 4.11 of Kuelbs and Mandrekar [20] which provides conditions for the stable domain of attraction in a separable Hilbert space. The Hilbert space we will consider in the following will be $\mathcal{S}$ and the stability index will be $\alpha/2$, $\alpha \in (2,4)$. However, when stating the result of Kuelbs and Mandrekar, we will use a generic Hilbert space H and the generic stability index $p \in (0,2)$. Recall that for a stable random element $S \in H$ with index $p \in (0,2)$, there exists a *spectral* measure σ_S defined on the unit sphere $\mathbb{S}_H = \{z \in H : \|z\| = 1\}$, such that the characteristic functional of S is given by

$$E\exp\bigl\{i\langle x,S\rangle\bigr\}=\exp\biggl\{i\langle x,\beta_S\rangle-\int_{\mathbb{S}}\bigl|\langle x,s\rangle\bigr|^p\sigma_S(ds)+iC(p,x)\biggr\},\quad x\in H,\tag{3.2}$$

where

$$C(p,x)=\begin{cases}\tan\dfrac{\pi p}{2}\displaystyle\int_{\mathbb{S}}\langle x,s\rangle\bigl|\langle x,s\rangle\bigr|^{p-1}\sigma_S(ds) & \text{if } p\neq 1,\\ \dfrac{2}{\pi}\displaystyle\int_{\mathbb{S}}\langle x,s\rangle\log\bigl|\langle x,s\rangle\bigr|\sigma_S(ds) & \text{if } p=1.\end{cases}$$

We denote the above representation by $S\sim[p,\sigma_S,\beta_S]$. The p-stable random element S is necessarily regularly varying with index $p\in(0,2)$. In fact, its angular measure is precisely the normalized spectral measure appearing in (3.2), that is,

$$\Gamma_S(\cdot)=\frac{\sigma_S(\cdot)}{\sigma_S(\mathbb{S}_H)}.$$

Kuelbs and Mandrekar [20] derived sufficient and necessary conditions on the distribution of Z under which

$$b_N^{-1}\Biggl(\sum_{i=1}^N Z_i-\gamma_N\Biggr)\xrightarrow{d}S,\tag{3.3}$$

where the Z_i are i.i.d. copies of Z. They assume that the support of the distribution of S, equivalently of the distribution of Z, spans the whole Hilbert space H. In our context, we will need to work with Z whose distribution is not supported on the whole space. Denote by $L(Z)$ the smallest closed subspace which contains the support of the distribution of Z. Then $L(Z)$ is a Hilbert space itself with the inner product inherited from H. Denote by $\{e_j,\,j\in\mathbb{N}\}$ an orthonormal basis of $L(Z)$. We assume that this is an infinite basis because we consider infinite dimensional data. (The finite dimensional case has already been dealt with by Rvačeva [30].) Introduce the projections

$$\pi_m(z)=\sum_{j=m}^{\infty}\langle z,e_j\rangle e_j,\quad z\in H.$$

Theorem 3.1. *Let $Z_1,Z_2,\dots$ be i.i.d. random elements in a separable Hilbert space H with the same distribution as Z. Let $\{e_j,\,j\in\mathbb{N}\}$ be an orthonormal basis of $L(Z)$. There exist normalizing constants b_N and γ_N such that* (3.3) *holds* **if and only if**

$$\frac{P(\|\pi_m(Z)\|>tu)}{P(\|Z\|>u)}\to\frac{c_m}{c_1}t^{-p},\quad u\to\infty,\tag{3.4}$$

where for each $m\geq 1$, $c_m>0$, and $\lim_{m\to\infty}c_m=0$, and where

$$\frac{P(\|Z\|>u,Z/\|Z\|\in A)}{P(\|Z\|>u,Z/\|Z\|\in A^\star)}\to\frac{\sigma_S(A)}{\sigma_S(A^\star)},\quad u\to\infty,\tag{3.5}$$

for all continuity sets A, $A^\star\in\mathcal{B}(\mathbb{S}_H)$ with $\sigma_S(A^\star)>0$.

If (3.3) *holds, the sequence* b_N *must satisfy*

$$b_N \to \infty, \qquad \frac{b_N}{b_{N+1}} \to 1, \qquad N b_N^{-2} E\big(\|Z\|^2 I_{\{\|Z\| \le b_N\}}\big) \to \lambda_p \sigma_S(\mathbb{S}_H), \tag{3.6}$$

where

$$\lambda_p = \begin{cases} \dfrac{p(1-p)}{\Gamma(3-p)\cos(\pi p/2)}, & \text{if } p \neq 1 \\ 2/\pi, & \text{if } p = 1, \end{cases} \tag{3.7}$$

and $\Gamma(a) := \int_0^\infty e^{-x} x^{a-1}\,dx$, $a > 0$ *is the Euler gamma function. Furthermore, the* $\gamma_N \in H$ *may be chosen as*

$$\gamma_N = N E(Z I_{\{\|Z\| \le b_N\}}). \tag{3.8}$$

Remark 3.1. The origin of the constant λ_p appearing in (3.6) can be understood as follows. Consider the simple scalar case $H = \mathbb{R}$. Let Z be symmetric α-stable with $E[e^{iZx}] = e^{-c|x|^\alpha}$, $x \in \mathbb{R}$, where in this case, $c = \sigma(\mathbb{S}_H) \equiv \sigma(\{-1, 1\}) > 0$. Consider i.i.d. copies Z_i, $i = 1, 2, \ldots$ of Z and observe that by the p-stability property

$$\frac{1}{N^{1/\alpha}} \sum_{j=1}^N Z_j \stackrel{d}{=} Z \equiv S,$$

and hence (3.3) holds trivially with $b_N := N^{1/\alpha}$ and $\gamma_N := 0$.

On the other hand, by Proposition 1.2.15 on page 16 in Samorodnitsky and Taqqu [31], we have

$$P\big(|Z| > x\big) \sim \frac{c(1-p)}{\Gamma(2-p)\cos(\pi p/2)} x^{-p}, \quad \text{as } x \to \infty.$$

This along with an integration by parts and an application of Karamata's theorem yield $N b_N^{-2} E[Z^2 I_{\{|Z| \le b_N\}}] \to \lambda_p \sigma_S(\mathbb{S}_H)$, giving the constant in (3.6).

Proposition 3.1. *Conditions* (3.4) *and* (3.5) *in Theorem* 3.1 *hold if and only if* Z *is regularly varying in* H *with index* $p \in (0, 2)$ *and for each* $m \ge 1$, $\mu_Z(A_m) > 0$, *where*

$$A_m = \left\{ z \in H : \big\|\pi_m(z)\big\| = \left\| \sum_{j=m}^\infty \langle z, e_j \rangle e_j \right\| > 1 \right\}. \tag{3.9}$$

Our next objective is to show that if X is a regularly varying element of a separable Hilbert space H whose index is $\alpha > 0$, then the operator $Y = X \otimes X$ is regularly varying with index $\alpha/2$, in the space of Hilbert–Schmidt operators. If $y, z \in H$, then $y \otimes z$ is an element of $\mathcal{S}$ defined by $(y \otimes z)(x) = \langle y, x \rangle z$, $x \in H$. It is easy to check that $\|y \otimes z\|_{\mathcal{S}} = \|y\| \|z\|$. If $B_1, B_2 \subset H$, we denote by $B_1 \otimes B_2$ the subset of $\mathcal{S}$ defined as the set of operators of the form $x_1 \otimes x_2$, with

$x_1 \in B_1$, $x_2 \in B_2$. Denote by $\mathbb{S}_H$ the unit sphere in H centered at the origin, and by $\mathbb{S}_{\mathcal{S}}$ such a sphere in $\mathcal{S}$.

The next result is valid for all $\alpha > 0$.

Proposition 3.2. *Suppose X is a regularly varying element with index $\alpha > 0$ of a separable Hilbert space H. Then the operator $Y = X \otimes X$ is a regularly varying element with index $\alpha/2$ of the space $\mathcal{S}$ of Hilbert–Schmidt operators.*

Remark 3.2. The proof of Proposition 3.2 shows that the angular measure of $X \otimes X$ is supported on the *diagonal* $\{\Psi \in \mathbb{S}_{\mathcal{S}} : \Psi = x \otimes x \text{ for some } x \in \mathbb{S}_H\}$ and that $\Gamma_{X\otimes X}(B \otimes B) = \Gamma_X(B)$, $\forall B \subset \mathcal{B}(\mathbb{S}_H)$.

The next result specifies the limit distribution of the sums of the $X_i \otimes X_i$ based on the results derived so far.

Theorem 3.2. *Suppose Assumptions 2.1 and 2.2 hold. Then, there exist normalizing constants k_N and operators ψ_N such that*

$$k_N^{-1}\left(\sum_{i=1}^{N} X_i \otimes X_i - \psi_N\right) \xrightarrow{d} S, \tag{3.10}$$

where $S \in \mathcal{S}$ is a stable random operator, $S \sim [\alpha/2, \sigma_S, 0]$, where the spectral measure σ_S is defined on the unit sphere $\mathbb{S}_{\mathcal{S}} = \{y \in \mathcal{S} : \|y\|_{\mathcal{S}} = 1\}$. The normalizing constants may be chosen as follows

$$k_N = \left(\frac{\alpha}{4-\alpha}\right)^{2/\alpha} a_N^2, \qquad \psi_N = NE\big[(X \otimes X) I_{\{\|X\|^2 \le k_N\}}\big], \tag{3.11}$$

where a_N is defined by (2.6).

The final result of this section specifies the asymptotic distribution of $\widehat{C} - C$.

Theorem 3.3. *Suppose Assumptions 2.1 and 2.2 hold. Then,*

$$Nk_N^{-1}(\widehat{C} - C) \xrightarrow{d} S - \frac{\alpha}{\alpha-2}\int_{\mathbb{S}_H} (\theta \otimes \theta)\Gamma_X(d\theta), \tag{3.12}$$

where $S \in \mathcal{S}$ and $\{k_N\}$ are as in Theorem 3.2. ($k_N = N^{2/\alpha}L(N)$ for a slowly varying L.)

If the X_i are scalars, then the angular measure Γ_X is concentrated on $\mathbb{S}_H = \{-1, 1\}$, with $\Gamma_X(1) = p$, $\Gamma_X(-1) = 1 - p$, in the notation of Davis and Resnick [5]. Thus $\int_{\mathbb{S}_H} \theta^2 \Gamma_X(d\theta) = 1$, and we recover the centering $\alpha/(\alpha - 2)$ in Theorem 2.2 of Davis and Resnick [5]. Relation (3.12) explains the structure of this centering in a much more general context.

Theorem 3.3 readily leads to a strong law of large numbers which can be derived by an application of the following result, a consequence of Theorem 3.1 of de Acosta [6].

Theorem 3.4. *Suppose $Y_i, i \geq 1$, are i.i.d. mean zero elements of a separable Hilbert space with $E\|Y_i\|^\gamma < \infty$, for some $1 \leq \gamma < 2$. Then,*

$$\frac{1}{N^{1/\gamma}} \sum_{i=1}^{N} Y_i \xrightarrow{P} 0 \quad \textit{if and only if} \quad \frac{1}{N^{1/\gamma}} \sum_{i=1}^{N} Y_i \xrightarrow{a.s.} 0.$$

Set $Y_i = X_i \otimes X_i - E[X \otimes X]$. Then the Y_i are i.i.d. mean zero elements of $\mathcal{S}$ which, by Proposition 3.2, satisfy $E\|Y_i\|_{\mathcal{S}}^\gamma < \infty$, for any $\gamma \in (0, \alpha/2)$. Theorem 3.3 implies that for any $\gamma \in (0, \alpha/2)$, $N^{-1/\gamma} \sum_{i=1}^N Y_i \xrightarrow{P} 0$. Thus Theorem 3.4 leads to the following corollary.

Corollary 3.1. *Suppose Assumptions 2.1 and 2.2 hold. Then, for any $\gamma \in [1, \alpha/2)$, $N^{1-1/\gamma}\|\widehat{C} - C\|_{\mathcal{S}} \to 0$ with probability 1.*

4. Convergence of eigenfunctions and eigenvalues

We first formulate and prove a general result which allows us to derive the asymptotic distributions of the eigenfunctions and eigenvalues of an estimator of the covariance operator from the asymptotic distribution of the operator itself. The proof of this result is implicit in the proofs of the results of Section 2 of Kokoszka and Reimherr [16], which pertain to the asymptotic normality of the sample covariance operator if $E\|X\|^4 < \infty$. The result and the technique of proof are however more general, and can be used in different contexts, so we state and prove it in detail.

Assumption 4.1. Suppose C is the covariance operator of a random function X taking values in L^2 such that $E\|X\|^2 < \infty$. Suppose $\widehat{C}$ is an estimator of C which is a.s. symmetric, nonnegative–definite and Hilbert–Schmidt. Assume that for some random operator $Z \in \mathcal{S}$, and for some $r_N \to \infty$,

$$Z_N := r_N(\widehat{C} - C) \xrightarrow{d} Z.$$

In our setting, $Z \in \mathcal{S}$ is specified in (3.12), and $r_N = N^\beta L(N)$ for some $0 < \beta < 1/2$. More precisely,

$$r_N = N a_N^{-2}, \quad a_N = N^{1/\alpha} L_0(N), \alpha \in (2, 4).$$

We will work with the eigenfunctions and eigenvalues defined by

$$C(v_j) = \lambda_j v_j, \qquad \widehat{C}_j(\hat{v}_j) = \hat{\lambda}_j \hat{v}_j, \quad j \geq 1.$$

Assumption 4.1 implies that $\hat{\lambda}_j \geq 0$ and the $\hat{v}_j$ are orthogonal with probability 1. We assume that, like the v_j, the $\hat{v}_j$ have unit norms. To lighten the notation, we assume that $\operatorname{sign}\langle \hat{v}_j, v_j \rangle = 1$. This sign does not appear in any of our final results, it cancels in the proofs. We assume that both sets of eigenvalues are ordered in decreasing order. The next assumption is standard, it ensures that the population eigenspaces are one dimensional.

Assumption 4.2. $\lambda_1 > \lambda_2, \dots, > \lambda_p > \lambda_{p+1}$.

Set

$$T_j = \sum_{k \neq j} (\lambda_j - \lambda_k)^{-1} \langle Z, v_j \otimes v_k \rangle v_k.$$

Lemma I.2 in online material shows that the series defining T_j converges a.s. in L^2.

Theorem 4.1. *Suppose Assumptions 4.1 and 4.2 hold. Then,*

$$r_N\{\hat{v}_j - v_j, 1 \le j \le p\} \xrightarrow{d} \{T_j, 1 \le j \le p\}, \quad \text{in } (L^2)^p,$$

and

$$r_N\{\hat{\lambda}_j - \lambda_j, 1 \le j \le p\} \xrightarrow{d} \{\langle Z(v_j), v_j \rangle, 1 \le j \le p\}, \quad \text{in } \mathbb{R}^p.$$

If Z is an $(\alpha/2)$-stable random operator in $\mathcal{S}$, then the T_j are jointly $(\alpha/2)$-stable random functions in L^2, and $\langle Z(v_j), v_j \rangle$ are jointly $(\alpha/2)$-stable random variables. This follows directly from the definition of a stable distribution, for example, Section 6.2 of Linde [21]. Under Assumption 2.1, $r_N = N^{1-2/\alpha} L_0^{-2}(N)$. Theorem 4.1 thus leads to the following corollary.

Corollary 4.1. *Suppose Assumptions 2.1, 2.2 and 4.2 hold. Then,*

$$N^{1-2/\alpha} L_0^{-2}(N)\{\hat{v}_j - v_j, 1 \le j \le p\} \xrightarrow{d} \{T_j, 1 \le j \le p\}, \quad \text{in } (L^2)^p,$$

where the T_j are jointly $(\alpha/2)$-stable in L^2, and

$$N^{1-2/\alpha} L_0^{-2}(N)\{\hat{\lambda}_j - \lambda_j, 1 \le j \le p\} \xrightarrow{d} \{S_j, 1 \le j \le p\}, \quad \text{in } \mathbb{R}^p,$$

where the S_j are jointly $(\alpha/2)$-stable in $\mathbb{R}$.

Corollary 4.1 implies the rates in probability $\hat{v}_j - v_j = O_P(r_N^{-1})$ and $\hat{\lambda}_j - \lambda_j = O_P(r_N^{-1})$, with $r_N = N^{1-2/\alpha} L_0^{-2}(N)$. This means, that the distances between $\hat{v}_j$ and $\hat{\lambda}_j$ and the corresponding population parameters are approximately of the order $N^{2/\alpha-1}$, that is, are asymptotically larger that these distances in the case of $E\|X\|^4 < \infty$, which are of the order $N^{-1/2}$. Note that $2/\alpha - 1 \to -1/2$, as $\alpha \to 4$.

It is often useful to have some bounds on moments, analogous to relations (1.5). Since the tails of $\|T_j\|$ and $|S_j|$ behave like $t^{-\alpha/2}$, for example, Section 6.7 of Linde [21], $E\|T_j\|^\gamma < \infty$, $0 < \gamma < \alpha/2$, with an analogous relation for $|S_j|$. We can thus expect convergence of moments of order $\gamma \in (0, \alpha/2)$. The following theorem specifies the corresponding results.

Theorem 4.2. *If Assumptions 2.1 and 2.2 hold, then for each $\gamma \in (0, \alpha/2)$, there is a slowly varying function L_γ such that*

$$\limsup_{N \to \infty} N^{\gamma(1-2/\alpha)} L_\gamma(N) E\|\widehat{C} - C\|_{\mathcal{S}}^\gamma < \infty$$

and for $j \geq 1$,

$$\limsup_{N\to\infty} N^{\gamma(1-2/\alpha)} L_\gamma(N) E|\hat{\lambda}_j - \lambda_j|^\gamma < \infty.$$

If, in addition, Assumption 4.2 *holds, then for* $1 \leq j \leq p$,

$$\limsup_{N\to\infty} N^{\gamma(1-2/\alpha)} L_\gamma(N) E\|\hat{v}_j - v_j\|^\gamma < \infty.$$

Several cruder bounds can be derived from Theorem 4.2. In applications, it is often convenient to take $\gamma = 1$. Then $E\|\widehat{C} - C\|_{\mathcal{S}} \leq N^{2/\alpha-1} L_1(N)$. By Potter bounds, for example, Proposition 2.6 (ii) in Resnick [29], for any $\epsilon > 0$ there is a constant C_ϵ such that for $x > x_\epsilon$ $L_1(x) \leq C_\epsilon x^\epsilon$. For each $\alpha \in (2, 4)$, we can choose ϵ so small that $-\delta(\alpha) := 2/\alpha - 1 + \epsilon < 0$. This leads to the following corollary.

Corollary 4.2. *If Assumptions* 2.1 *and* 2.2 *hold, then for each* $\alpha \in (2, 4)$, *there are constant* C_α *and* $\delta(\alpha) > 0$ *such that*

$$E\|\widehat{C} - C\|_{\mathcal{S}} \leq C_\alpha N^{-\delta(\alpha)} \quad \text{and} \quad E\|\hat{\lambda}_j - \lambda_j\| \leq C_\alpha N^{-\delta(\alpha)}.$$

If, in addition, Assumption 4.2 *holds, then for* $1 \leq j \leq p$, $E\|\hat{v}_j - v_j\| \leq C_\alpha(j) N^{-\delta(\alpha)}$.

Corollary 4.2 implies that $E\|\widehat{C} - C\|_{\mathcal{S}}$, $E\|\hat{\lambda}_j - \lambda_j\|$ and $E\|\hat{v}_j - v_j\|$ tend to zero, for any $\alpha \in (2, 4)$.

5. An application: Functional linear regression

One of the most widely used tools of functional data analysis is the functional regression model, for example, Ramsay and Silverman [27], Horváth and Kokoszka [11], Kokoszka and Reimherr [17]. Suppose $X_1, X_2, \ldots, X_N$ are explanatory functions, $Y_1, Y_2, \ldots, Y_N$ are response functions, and assume that

$$Y_i(t) = \int_0^1 \psi(t, s) X_i(s)\, ds + \varepsilon_i(t), \quad 1 \leq i \leq N, \tag{5.1}$$

where $\psi(\cdot, \cdot)$ is the kernel of $\Psi \in \mathcal{S}$. The X_i are mean zero i.i.d. functions in $L^2 = L^2([0, 1])$, and so are the error functions ε_i. Consequently, the Y_i are i.i.d. in L^2. A question that has been investigated from many angles is how to consistently estimate the regression kernel $\psi(\cdot, \cdot)$. An estimator that has become popular following the work of Yao et al. [32] can be constructed as follows.

The population version of (5.1) is $Y(t) = \int \psi(t, s) X(s)\, ds + \varepsilon(t)$. Denote by v_i the FPCs of X and by u_j those of Y, so that

$$X(s) = \sum_{i=1}^\infty \xi_i v_i(s), \qquad Y(t) = \sum_{j=1}^\infty \zeta_j u_j(t).$$

If ε is independent of X, then, with $\lambda_\ell = E[\xi_\ell^2]$,

$$\psi(t,s) = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \frac{E[\xi_\ell \zeta_k]}{\lambda_\ell} u_k(t) v_\ell(s),$$

with the series converging in $L^2([0,1]\times[0,1])$, equivalently in $\mathcal{S}$, see Lemma 8.1 in Horváth and Kokoszka [11]. This motivates the estimator

$$\hat{\psi}_{KL}(t,s) = \sum_{k=1}^{K}\sum_{\ell=1}^{L} \frac{\hat{\sigma}_{\ell k}}{\hat{\lambda}_\ell} \hat{u}_k(t)\hat{v}_\ell(s),$$

where $\hat{u}_k$ are the eigenfunctions of $\widehat{C}_Y$ and $\hat{\sigma}_{\ell k}$ is an estimator of $E[\xi_\ell \zeta_k]$. Yao et al. [32] study the above estimator under the assumption that data are observed sparsely and with measurement errors. This requires two-stage smoothing, so their assumptions focus on conditions on the various smoothing parameters and the random mechanism that generates the sparse observations. Like in all work of this type, they assume that the underlying functions have finite fourth moments: $E\|X\|^4 < \infty$, $E\|\varepsilon\|^4 < \infty$, and so $E\|Y\|^4 < \infty$. Our objective is to show that if the X_i satisfy the assumptions of Section 2, then

$$\|\widehat{\Psi}_{KL} - \Psi\|_{\mathcal{L}} \overset{\text{a.s.}}{\to} 0, \tag{5.2}$$

as $N \to \infty$, and $K, L \to \infty$ at suitable rates determined by the rate of decay of the eigenvalues. The norm $\|\cdot\|_{\mathcal{L}}$ is the usual operator norm. The integral operators Ψ and $\widehat{\Psi}_{KL}$ are defined by their kernels $\psi(\cdot,\cdot)$ and $\hat{\psi}_{KL}(\cdot,\cdot)$, respectively. We focus on moment conditions, so we assume that the functions X_i, Y_i are fully observed, and use the estimator

$$\hat{\sigma}_{\ell k} = \frac{1}{N}\sum_{i=1}^{N} \hat{\xi}_{i\ell}\hat{\zeta}_{ik}, \quad \hat{\xi}_{i\ell} = \langle X_i, \hat{v}_\ell\rangle, \hat{\zeta}_{ik} = \langle Y_i, \hat{u}_k\rangle.$$

Since the regression operator Ψ is infinitely dimensional, we strengthen Assumption 4.2 to the following assumption.

Assumption 5.1. The eigenvalues $\lambda_i = E\xi_i^2$ and $\gamma_j = E\zeta_j^2$ satisfy

$$\lambda_1 > \lambda_2 > \cdots > 0, \qquad \gamma_1 > \gamma_2 > \cdots > 0.$$

Many issues related to the infinite dimension of the functional data in model (5.1) are already present when considering projections on the unobservable subspaces

$$\mathcal{V}_L = \text{span}\{v_1, v_2, \ldots, v_L\}, \qquad \mathcal{U}_K = \text{span}\{u_1, u_2, \ldots, u_K\}.$$

Therefore we first consider the convergence of the operator with the kernel

$$\psi_{KL}(t,s) = \sum_{k=1}^{K}\sum_{\ell=1}^{L} \frac{\sigma_{\ell k}}{\lambda_\ell} u_k(t) v_\ell(s).$$

Set $\sigma_{\ell k} = E[\xi_\ell \zeta_k]$ and observe that

$$\psi_{KL}(t,s) - \psi(t,s) = -\sum_{k>K \text{ or } \ell>L} \frac{\sigma_{\ell k}}{\lambda_\ell} u_k(t) v_\ell(s).$$

Therefore,

$$\|\Psi_{KL} - \Psi\|_{\mathcal{L}}^2 \le \|\Psi_{KL} - \Psi\|_{\mathcal{S}}^2 = \sum_{k>K \text{ or } \ell>L} \frac{\sigma_{\ell k}^2}{\lambda_\ell^2}. \tag{5.3}$$

The condition

$$\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \frac{\sigma_{\ell k}^2}{\lambda_\ell^2} < \infty, \tag{5.4}$$

which is Assumption (A1) of Yao et al. [32], implies that the remainder term is asymptotically negligible. It is instructive to rewrite condition (5.4) in a different form. Observe that

$$\sigma_{\ell k} = E\big[\xi_l \langle \Psi(X) + \varepsilon, u_k\rangle\big] = E\left[\xi_l \sum_{i=1}^{\infty} \xi_i \langle \Psi(v_i), u_k\rangle\right] = \lambda_\ell \langle \Psi(v_\ell), u_k\rangle. \tag{5.5}$$

Therefore,

$$\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \frac{\sigma_{\ell k}^2}{\lambda_\ell^2} = \sum_{\ell=1}^{\infty} \frac{1}{\lambda_\ell^2} \sum_{k=1}^{\infty} \lambda_\ell^2 \langle \Psi(v_\ell), u_k\rangle^2 = \sum_{\ell=1}^{\infty} \|\Psi(v_\ell)\|^2 = \|\Psi\|_{\mathcal{S}}^2. \tag{5.6}$$

We see that condition (5.4) simply means that Ψ is a Hilbert–Schmidt operator, and so it holds under our general assumptions on model (5.1).

The last assumption implicitly restricts the rates at which K and L tend to infinity with N. Under Assumption 5.1, the following quantities are well defined

$$\alpha_j = \min\{\lambda_j - \lambda_{j+1}, \lambda_{j-1} - \lambda_j\}, \quad j \ge 2, \alpha_1 = \lambda_1 - \lambda_2, \tag{5.7}$$

$$\beta_j = \min\{\gamma_j - \gamma_{j+1}, \gamma_{j-1} - \gamma_j\}, \quad j \ge 2, \beta_1 = \gamma_1 - \gamma_2. \tag{5.8}$$

Assumption 5.2. The truncation levels K and L tend to infinity with N in such a way that for some $\gamma \in (1, \alpha/2)$,

$$\limsup_{N\to\infty} \lambda_L^{-3/2} L^{1/2} N^{1/\gamma - 1} < \infty, \tag{5.9}$$

$$\limsup_{N\to\infty} \lambda_L^{-1} \left(\sum_{j=1}^{L} \alpha_j^{-1}\right) N^{1/\gamma - 1} < \infty, \tag{5.10}$$

$$\limsup_{N\to\infty} \lambda_L^{-1} K^{1/2} N^{1/\gamma - 1} < \infty, \tag{5.11}$$

$$\limsup_{N\to\infty} \lambda_L^{-1}\left\{\left(\sum_{k=1}^{K}\beta_k^{-1}\right)+\left(\sum_{k=1}^{K}\beta_k^{-2}\right)^{1/2}\right\}N^{1/\gamma-1}<\infty. \tag{5.12}$$

The conditions in Assumption 5.2 could be restated or unified; and could be replaced by slightly different conditions by modifying the technique of proof. The essence of this assumption is that K and L must tend to infinity sufficiently slowly, and the rate is influenced by index α; the closer α is to 4, the larger γ can be taken, so K and L can be larger.

Theorem 5.1. *Suppose model* (5.1) *holds with* $\Psi \in \mathcal{S}$, *the* X_i *and the* Y_i *satisfying Assumptions* 2.1 *and* 2.2, *and square integrable* ε_i, $E\|\varepsilon_i\|^2 < \infty$. *Then relation* (5.2) *holds under Assumptions* 5.1 *and* 5.2.

Acknowledgements

The authors have been supported by NSF grant "FRG: Collaborative Research: Extreme Value Theory for Spatially Indexed Functional Data" (1462067 CSU, 1462368 Michigan). We thank Professor Jan Rosiński for directing us to the work of de Acosta and Giné [7], and Mr. Ben Zheng for preparing the figures.

Supplementary Material

Supplement to "Principal components analysis of regularly varying functions" (DOI: 10.3150/19-BEJ1113SUPP; .pdf). Supplementary information

References

[1] Anderson, P.L. and Meerschaert, M.M. (1997). Periodic moving averages of random variables with regularly varying tails. *Ann. Statist.* **25** 771–785. MR1439323

[2] Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications. Lecture Notes in Statistics* **149**. New York: Springer. MR1783138

[3] Dauxois, J., Pousse, A. and Romain, Y. (1982). Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference. *J. Multivariate Anal.* **12** 136–154. MR0650934

[4] Davis, R. and Resnick, S. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13** 179–195. MR0770636

[5] Davis, R. and Resnick, S. (1986). Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14** 533–558. MR0840513

[6] de Acosta, A. (1981). Inequalities for B-valued random vectors with applications to the strong law of large numbers. *Ann. Probab.* **9** 157–161. MR0606806

[7] de Acosta, A. and Giné, E. (1979). Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. Verw. Gebiete* **48** 213–231. MR0534846

[8] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events. For Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613

[9] Hall, P. and Hosseini-Nasab, M. (2006). On properties of functional principal components analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 109–126. MR2212577

[10] Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. MR2662361

[11] Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. *Springer Series in Statistics*. New York: Springer. MR2920735

[12] Horváth, L., Kokoszka, P. and Rice, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82. MR3153649

[13] Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis*, *with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR3379106

[14] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910

[15] Klüppelberg, C. and Mikosch, T. (1994). Some limit theory for the self-normalised periodogram of stable processes. *Scand. J. Stat.* **21** 485–491. MR1310091

[16] Kokoszka, P. and Reimherr, M. (2013). Asymptotic normality of the principal components of functional time series. *Stochastic Process. Appl.* **123** 1546–1562. MR3027890

[17] Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. *Texts in Statistical Science Series*. Boca Raton, FL: CRC Press. MR3793167

[18] Kokoszka, P.S. and Taqqu, M.S. (1996). Parameter estimation for infinite variance fractional ARIMA. *Ann. Statist.* **24** 1880–1913. MR1421153

[19] Kokoszka, P., Stoev, S. and Xiong, Q. (2019). Supplement to "Principal components analysis of regularly varying functions." DOI:10.3150/19-BEJ1113SUPP.

[20] Kuelbs, J. and Mandrekar, V. (1974). Domains of attraction of stable measures on a Hilbert space. *Studia Math.* **50** 149–162. MR0345155

[21] Linde, W. (1986). *Probability in Banach Spaces – Stable and Infinitely Divisible Distributions*, 2nd ed. *A Wiley-Interscience Publication*. Chichester: Wiley. MR0874529

[22] Lucca, D.O. and Moench, E. (2015). The pre-FOMC announcement drift. *J. Finance* **70** 329–371.

[23] Meerschaert, M.M. (1984). *Multivariable Domains of Attraction and Regular Variation*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.) – University of Michigan. MR2633908

[24] Meerschaert, M.M. and Scheffler, H.-P. (2001). *Limit Distributions for Sums of Independent Random Vectors*: *Heavy Tails in Theory and Practice*. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1840531

[25] Meiguet, T. (2010). Heavy Tailed Functional Time series. Ph.D. Thesis, Universite catholigue de Louvain.

[26] Mikosch, T., Gadrich, T., Klüppelberg, C. and Adler, R.J. (1995). Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.* **23** 305–326. MR1331670

[27] Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

[28] Resnick, S.I. (1987). *Extreme Values*, *Regular Variation*, *and Point Processes*. *Applied Probability. A Series of the Applied Probability Trust* **4**. New York: Springer. MR0900810

[29] Resnick, S.I. (2007). *Heavy-Tail Phenomena*: *Probabilistic and Statistical Modeling*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424

[30] Rvačeva, E.L. (1962). On domains of attraction of multi-dimensional distributions. In *Select. Transl. Math. Statist. and Probability*, *Vol.* 2 183–205. Providence, R.I.: Amer. Math. Soc. MR0150795

[31] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes*: *Stochastic Models with Infinite Variance*. *Stochastic Modeling*. New York: CRC Press. MR1280932

[32] Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. MR2253106

Received June 2018 and revised December 2018