Journal of Multivariate Analysis 193 (2023) 105123

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Principal component analysis of infinite variance functional data

Piotr Kokoszka [a,*,1], Rafał Kulik [b,1]

[a] Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, United States
[b] University of Ottawa, Canada

ARTICLE INFO

Article history:
Received 20 February 2022
Received in revised form 13 October 2022
Accepted 13 October 2022
Available online 23 October 2022

AMS 2020 Subject Classification:
primary 62E20
secondary 60E07

Keywords:
Asymptotic theory
Functional data
Infinite variance
Principal components
Regular variation

ABSTRACT

Principal Components Analysis is a widely used approach of multivariate analysis. Over the past 30 years, it has gained renewed attention in the context of functional data, chiefly as a commonly used tool for dimension reduction or feature extraction. Consequently, a large body of statistical theory has been developed to justify its application in various contexts. This theory focuses on the convergence of sample PCs to their population counterparts in a multitude of statistical models and under diverse data collection assumptions. What such results have in common is the assumption that the population covariance operator exists.This paper is concerned with multivariate and functional data that have infinite variance and, consequently, for which the population covariance operator is not defined. However, the sample covariance operator and its eigenfunctions are always defined. It has been unknown what the asymptotic behavior of these important statistics is in the context of infinite variance multivariate or functional data. We derive suitable large sample theory. In particular, we specify normalizing sequences and conditions for suitably defined consistency. We study multivariate models in which explicit limits can be derived. These examples show that definitions, results and intuition developed for multivariate and functional data with finite variance need not apply in the setting we consider.

1. Introduction

Over the last two decades, Functional Data Analysis (FDA) has grown into a substantial field of statistics that has found applications in a number of disciplines and stimulated the development of novel statistical approaches and underlying theory. The field is now very rich and multifaceted, and it is at this point not possible to discuss most representative publications or comment on their impact or importance. We merely note the monographs and textbooks of Bosq [4], Ramsay and Silverman [25], Ferraty and Vieu [9], Ramsay et al. [24], Shi and Choi [29], Horváth and Kokoszka [12], Hsing and Eubank [13] and Kokoszka and Reimherr [16]. Many review papers focusing on specific aspects or applications of FDA are available, including many published by this journal, e.g., Goia and Vieu [11], Aneiros et al. [1] and Aneiros et al. [2].

As an outgrowth of its well-known multivariate counterpart, Functional Principal Component Analysis (FPCA) has been an important tool of FDA since the early days of the field and remains so. To explain the contribution of this paper, we

* Corresponding author.
E-mail address: Piotr.Kokoszka@colostate.edu (P. Kokoszka).

1 Both authors contributed equally.

https://doi.org/10.1016/j.jmva.2022.105123

begin by presenting the basics of the FPCA. Denote by $L^2 = L^2(\mathcal{U})$ the space of square integrable functions on a domain $\mathcal{U}$ with the inner product $\langle f, g\rangle = \int_{\mathcal{U}} f(u)g(u)du$. If $\mathcal{U}$ is a Polish space (complete and separable metric space), then L^2 is a separable Hilbert space. Suppose we observe functions $X_1, X_2, \ldots, X_N \in L^2$. Consider the sample covariance operator

$$\widehat{C}(x) = \frac{1}{N}\sum_{n=1}^{N}\left\langle X_n - \bar{X}_N, x\right\rangle (X_n - \bar{X}_N), \quad x \in L^2(\mathcal{U}).$$

Statistical software, see e.g., Ramsay et al. [24], can compute its orthonormal eigenfunction $\hat{v}_j$ and positive eigenvalues $\hat{\lambda}_j$ that satisfy

$$\widehat{C}(\hat{v}_j) = \hat{\lambda}_j\hat{v}_j, \quad j \in \{1, 2, \ldots, N\}, \tag{1}$$

and the sample scores $\hat{\xi}_{nj} = \left\langle X_n - \bar{X}_N, \hat{v}_j\right\rangle$. The $\hat{v}_j$ are the sample Functional Principal Components (FPCs). Clearly, no assumptions are needed to compute the sample quantities $\widehat{C}, \hat{\lambda}_j, \hat{v}_j$. To establish their convergence to population quantities, one must impose assumptions on the observations. If the X_i are independent with the same distribution as X, and $E\|X\|^4 < \infty$, the products in the definition of $\widehat{C}$ have finite second moment, so the CLT in a separable Hilbert space implies that $\widehat{C}$ converges with standard rate to the (population) covariance operator

$$C(x) = E\left[\langle X - EX, x\rangle (X - EX)\right], \quad x \in L^2(\mathcal{U}). \tag{2}$$

A more subtle question is if the $\hat{\lambda}_j$ and the $\hat{v}_j$ converge to the λ_j and v_j defined by $C(v_j) = \lambda_j v_j,\ j \geq 1$. This problem was solved by Dauxois et al. [5] whose results are reported e.g. in Bosq [4] and Horváth and Kokoszka [12] in greater generality. In particular, it is known that $\|\hat{v}_j - v_j\| \stackrel{P}{\to} 0$ and $\hat{\lambda}_j \stackrel{P}{\to} \lambda_j$, and asymptotic normality holds.

If one drops the assumption $E\|X\|^4 < \infty$, the standard Hilbert space CLT cannot be applied to the operator $\widehat{C}$. Even in the case of partial sums of scalar observations, to obtain convergence to a nondegenerate limit, one must assume that the observations have regularly varying tails. A basically complete theory is given in Gnedenko and Kolmogorov [10], which is summarized on a few pages in Section 2.2 of Embrechts et al. [8]. In the case of functional observations, it must be assumed that the X_i are regularly varying in L^2. Postponing the definitions and details to Section 2, if the index of regular variation α satisfies $\alpha \in (2, 4)$, then $E\|X\|^2 < \infty$, so the covariance operator C given by (2) is still well-defined. Asymptotic theory for this case is worked out in Kokoszka et al. [17]. One can still conclude that $\|\widehat{C} - C\| \stackrel{P}{\to} 0$, $\|\hat{v}_j - v_j\| \stackrel{P}{\to} 0$ and $\hat{\lambda}_j \stackrel{P}{\to} \lambda_j$, and specify asymptotics distributions, which are no longer normal.

If $\alpha < 2$, one cannot define the covariance operator C, because then $E\|X\|^2 = \infty$. Consequently one cannot consider its eigenfunctions v_j and eigenvalues λ_j. However, one can always compute the sample covariance operator $\widehat{C}$ and the FPCs $\hat{v}_j$. Some questions are then: Does $\widehat{C}$ and do the $\hat{v}_j$ converge to any limits? Since C and the v_j do not exist, what might those limits be? Can one establish convergence after a suitable normalization? What are the limits then? What assumptions must be imposed on the functional data to ensure their existence? This paper is concerned with providing precise answer to such questions. We provide exact formulas for various normalizing and centering sequences and show how they impact the limits. A data example motivating our theory is provided in the online material. This paper is concerned with the fundamental theory of PCA for infinite variance data. It is hoped that the understanding it provides will lead to the development of effective tools to handle such multivariate and functional data. An objective of this paper is thus to advance theory underlying the approximation

$$X_n(u) \approx \sum_{j=1}^{Q}\hat{\xi}_{nj}\hat{v}_j(u) \tag{3}$$

that is commonly used in FDA for dimension reduction or feature extraction. We want to understand the large sample behavior of the FPCs $\hat{v}_j$ and their scores $\hat{\xi}_{nj}$.

There is profound work on regularly varying functional data viewed as elements of abstract spaces. Such work focuses primary on aspects relevant to Extreme Value Theory, including the polar decomposition and the extremal index. Without attempting to give a full review, we note that following the work of Basrak and Segers [3], who studied the polar decomposition of a regularly varying multivariate time series, Meinguet and Segers [22] provided a detailed study of regularly varying time series in Banach spaces. Segers et al. [28] extended their results in two aspects: regular variation of the time series treated as a single random element in a sequence space and the polar decomposition in star-shaped metric spaces. It may be hoped that such general results combined with the results of this paper will motivate the development of useful statistical models for regularly varying functional data.

The remainder of the paper is organized as follows. In Section 2, we conveniently organize known results and prove corollaries that are needed in subsequent sections. Section 3 considers general, infinitely dimensional functional data. In Section 4, we focus on multivariate models in which asymptotic quantities can be computed explicitly. The results of Section 4 emphasize important differences between the cases of finite- and infinite variance multivariate or functional observations. Section 5 contains infinitely dimensional results that extend some results of Section 4. Longer and more technical proofs are collected in Sections 6 and 7. Online material contains a data example illustrating our theory.

2. Preliminary results

We begin with Proposition 1 and Remark 1 that follow from known results. We spell out relationships between various normalizing sequences that play a crucial role in the following. We first recall some definitions related to M_0 convergence, see Hult and Lindskog [14], Meinguet [21] and Lindskog et al. [20]. Suppose $\mathbb{X}$ a Polish space (a separable and complete metric space) and $\mathbb{B}$ is a separable Banach space. Denote by $M(\cdot)$ the set of positive measures on the Borel sets of a specified space. Denote further,

$$M_b(\mathbb{X}) = \{\mu \in M(\mathbb{X}) : \mu(\mathbb{X}) < \infty\},$$
$$M_0(\mathbb{B}) = \{\mu \in M(\mathbb{B} \setminus \{0\}) : \forall\, r > 0,\, \mu(\{x : \|x\| > r\}) < \infty\}.$$

Recall that the weak convergence of $\mu_N \in M_b(\mathbb{X})$ to $\mu \in M_b(\mathbb{X})$ (written as $\mu_N \xrightarrow{w} \mu$) means that for each bounded and continuous function $f : \mathbb{X} \to \mathbb{R}$, $\int f d\mu_N \to \int f d\mu$. Denote by $\mathcal{C}_{bd}(\mathbb{B})$ the set of bounded continuous functions on $\mathbb{B}$ that vanish on an open disk containing zero. We say that $\mu_N \in M_0(\mathbb{B})$ converges in M_0 to $\mu \in M_0(\mathbb{B})$ (written as $\mu_N \xrightarrow{M_0} \mu$) if

$$\int_{\mathbb{B}} f(z)\mu_N(dz) \to \int_{\mathbb{B}} f(z)\mu(dz), \quad \forall\, f \in \mathcal{C}_{bd}(\mathbb{B}).$$

Proposition 1. *Let X be a random element in a separable Banach space $\mathbb{B}$ and $\alpha > 0$. The following statements are equivalent:*

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

$$P(||X|| > u) = u^{-\alpha} L(u) \tag{4}$$

and

$$\frac{P(u^{-1}X \in \ \cdot\ )}{P(||X|| > u)} \xrightarrow{M_0} \mu(\cdot), \quad u \to \infty, \tag{5}$$

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

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

$$\frac{P(||X|| > tu, X/||X|| \in \cdot)}{P(||X|| > u)} \xrightarrow{w} t^{-\alpha}\Gamma(\cdot), \quad u \to \infty. \tag{6}$$

(iii) *Relation (4) holds, and for the same angular measure Γ as in (ii),*

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

(iv) *There is a sequence $\tilde{a}_N \to \infty$ such that with the same μ as in (i),*

$$\mu_N(\cdot) := NP\left(\tilde{a}_N^{-1}X \in \ \cdot\ \right) \xrightarrow{M_0} \mu(\cdot), \quad \text{as } N \to \infty. \tag{7}$$

(v) *There is a sequence $a_N \to \infty$ such that for the same angular measure Γ as in (ii),*

$$NP\left(\|X\| > ta_N, \frac{X}{\|X\|} \in \ \cdot\right) \xrightarrow{w} t^{-\alpha}\Gamma(\ \cdot\ ), \quad \text{as } N \to \infty.$$

Remark 1. Several points are in place here.

- If any of the conditions of Proposition 1 holds, we will write $X \in \mathrm{RV}_{-\alpha}(\mathbb{B})$.
- The angular measure Γ is related to the exponent measure μ via

$$\mu(dx) = c\,\alpha r^{-\alpha-1} dr\,\Gamma(d\theta), \quad r = \|x\|,\ \theta = \frac{x}{\|x\|}, \tag{8}$$

  where $c = \mu(\{x : \|x\| > 1\})$.
- Since Γ is the probability measure, condition (v) gives

$$NP\left(\|X\| > a_N\right) \to 1. \tag{9}$$

  We will refer to a_N as the quantile sequence of X (formally speaking, the quantile sequence of $\|X\|$).
- The sequences a_N and $\tilde{a}_N$ are regularly varying with index $1/\alpha$. They coincide if and only if μ is normalized in such the way that $\mu(\{x : \|x\| > 1\}) = 1$.
- In what follows, we consider random elements in a separable Hilbert space and use the norm $\|x\|^2 = \langle x, x\rangle$.

Suppose $\mathbb{H}$ is a separable Hilbert space and set $\mathbb{S}_{\mathbb{H}} = \{x \in \mathbb{H} :\ \|x\| = 1\}$.

Definition 1. A random element S in $\mathbb{H}$ is stable with index $p \in (0,2)$ if there is a finite measure σ_S on $\mathcal{B}(\mathbb{S}_{\mathbb{H}})$ and $\beta \in \mathbb{H}$ such that

$$E\left[\exp\{i\langle x, S\rangle\}\right] = \exp\left\{i\langle x, \beta\rangle - \int_{\mathbb{S}_{\mathbb{H}}} |\langle x, s\rangle|^p \sigma_S(ds) + iC(p,x)\right\},$$

where

$$C(p,x) = \begin{cases} \tan\frac{\pi p}{2} \int_{\mathbb{S}_{\mathbb{H}}} \langle x, s\rangle |\langle x, s\rangle|^{p-1}\sigma_S(ds), & p \neq 1, \\ \frac{2}{\pi}\int_{\mathbb{S}_{\mathbb{H}}} \langle x, s\rangle \log|\langle x, s\rangle|\sigma_S(ds), & p = 1. \end{cases}$$

For an orthonormal basis $\{\eta_k, k \geq 1\}$ in $\mathbb{H}$, define the projections

$$\pi_m(x) = \sum_{k=m}^{\infty} \langle x, \eta_k\rangle \eta_k$$

and consider the following condition.

Condition 1. The random element $Z \in \mathbb{H}$ satisfies the following conditions for $p \in (0,2)$:

$$\lim_{u\to\infty} \frac{P\left(\|Z\| > u,\ \|Z\|^{-1}Z \in A\right)}{P\left(\|Z\| > u,\ \|Z\|^{-1}Z \in A^*\right)} = \frac{\Gamma_Z'(A)}{\Gamma_Z'(A^*)}, \tag{10}$$

where Γ_Z' is a finite measure on $\mathbb{S}_{\mathbb{H}}$ and A, A^*, $\Gamma_Z'(A^*) > 0$, are continuity sets of Γ_Z', and

$$\lim_{u\to\infty} \frac{P(\|Z\| > u)}{P(\|\pi_m(Z)\| > tu)} = \frac{c_1 t^p}{c_m}, \tag{11}$$

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

Note that if $Z \in \mathrm{RV}_{-p}(\mathbb{H})$, then (10) holds and the angular measure Γ_Z in (6) and Γ_Z' in (10) are related by

$$\Gamma_Z = \Gamma_Z'/\Gamma_Z'(\mathbb{S}_{\mathbb{H}}) \tag{12}$$

because Γ_Z is a probability measure. Condition (11) ensures that the random objects we study are truly infinitely dimensional. (We study the finite dimensional case, which is instructive, in Section 4.) The following proposition, proven in Kokoszka et al. [17], establishes a connection between Condition 1 and regular variation.

Proposition 2. *Condition 1 holds if and only if $Z \in \mathrm{RV}_{-p}(\mathbb{H})$ with $p \in (0,2)$ and*

$$\forall\, m \geq 1, \quad \mu_Z(A_m) > 0, \tag{13}$$

where

$$A_m = \left\{ z \in \mathbb{H} : ||\pi_m(z)|| = \left\| \sum_{j=m}^{\infty} \langle z, \eta_j\rangle \eta_j \right\| > 1 \right\},$$

and where μ_Z is the exponent measure in (5).

Theorem 4.11 of Kuelbs and Mandrekar [18] directly implies the following theorem.

Theorem 1. *Let $Z_1, Z_2, \ldots$ be i.i.d. random elements in a separable Hilbert space $\mathbb{H}$ with the same distribution as Z. Suppose S is stable according to Definition 1. Then, there exist normalizing constants b_N and γ_N such that*

$$b_N^{-1}\left(\sum_{n=1}^{N} Z_n - \gamma_N\right) \xrightarrow{d} S, \tag{14}$$

if and only if Condition 1 holds with $\Gamma_Z' = \sigma_S$.

Remark 2 follows from the examination of the proof of Theorem 4.11 of Kuelbs and Mandrekar [18].

Remark 2. The sequence b_N must satisfy

$$b_N \to \infty, \quad \frac{b_N}{b_{N+1}} \to 1, \quad N b_N^{-2} E\left[||Z||^2 I_{\{||Z|| \leq b_N\}}\right] \to \lambda_p \sigma_S(\mathbb{S}_H), \tag{15}$$

where

$$\lambda_p = \begin{cases} \frac{p(1-p)}{\Gamma(3-p)\cos(\pi p/2)} & , \ p \neq 1 \\ 2/\pi & , \ p = 1, \end{cases} \tag{16}$$

and $\Gamma(a) := \int_0^\infty e^{-x} x^{a-1} dx$, $a > 0$, is Euler's gamma function.

Remark 3. If Condition 1 holds, then for arbitrary $p \in (0, 2)$ the $\gamma_N \in \mathbb{H}$ may be chosen as

$$\gamma_N = NE\left[ZI_{\{||Z|| \le b_N\}}\right]. \tag{17}$$

This choice does not yield $\beta = 0$ in Definition 1. If $p \in (0, 1)$, then γ_N can be chosen as $\gamma_N \equiv 0$, while if $p \in (1, 2)$, we can choose $\gamma_N = E[Z]$. These choices yield $\beta = 0$. Our theory will require $p \in (0, 1)$. Therefore, we set

$$\gamma_N = 0, \quad \text{if } \ p \in (0, 1). \tag{18}$$

The choice of the centering sequence does not follow from Kuelbs and Mandrekar [18], however, its form is well-known in the scalar case. See Theorem 8.3.1 in Kulik and Soulier [19], which in turn is based on an analogous result in Davis and Hsing [6].

Next, we want to replace the sequence b_N appearing in (14) with the quantile sequence a_N in the definition of the regular variation. We formulate the precise result we need as the following corollary to Theorem 1. We emphasize the assumption $p \in (0, 1)$.

Corollary 1. *Let $Z_1, Z_2, \ldots$ be i.i.d. random elements in a separable Hilbert space $\mathbb{H}$ with the same distribution as Z that satisfies Condition 1 with $p \in (0, 1)$ and has the angular measure Γ_Z ($Z \in \mathrm{RV}_{-p}(\mathbb{H})$ by Proposition 2). Then*

$$a_{N,Z}^{-1} \sum_{n=1}^{N} Z_n \xrightarrow{d} S_\infty^*, \tag{19}$$

where $a_{N,Z}$ is the quantile sequence of Z and S_∞^* is a p-stable random element with the characteristic functional

$$Ee^{i\langle x, S_\infty^* \rangle} = \exp\left\{-s_p \int_{\mathbb{S}_\mathbb{H}} \left[|\langle x, s\rangle|^p - i \tan\frac{\pi p}{2} \langle x, s\rangle \,|\langle x, s\rangle|^{p-1}\right] \Gamma_Z(ds)\right\}, \tag{20}$$

and where

$$s_p = \Gamma(1-p)\cos(\pi p/2). \tag{21}$$

Proof. By Theorem 1 and (18), we only need to find the relationship between b_N in (15) and the quantile sequence $a_{N,Z}$. From (15) we have

$$\lim_{N\to\infty} \frac{E\left[||Z||^2 I_{\{||Z|| \le b_N\}}\right]}{b_N^2 P(||Z|| > a_{N,Z})} = \lambda_p \sigma_S(\mathbb{S}_H).$$

On the other hand, Proposition 1.4.6 in Kulik and Soulier [19] gives

$$\lim_{N\to\infty} \frac{E\left[||Z||^2 I_{\{||Z|| \le b_N\}}\right]}{b_N^2 P(||Z|| > b_N)} = \frac{p}{2-p}.$$

Set $a_{N,Z} = \kappa b_N$. The above asymptotics give

$$\kappa = \left(\frac{2-p}{p}\lambda_p\right)^{1/p} \sigma_S^{1/p}(\mathbb{S}_H) = s_p^{-1/p}\sigma_S^{1/p}(\mathbb{S}_H) \tag{22}$$

with $s_p = \Gamma(1-p)\cos(\pi p/2)$.

Theorem 1 with $\gamma_N = 0$ gives

$$a_{N,Z}^{-1} \sum_{n=1}^{N} Z_n \xrightarrow{d} \kappa^{-1} S,$$

with S given in (14), with the location parameter $\beta = 0$. The characteristic function is

$$\begin{aligned}
&E\left[\exp\left\{i\left\langle x, \kappa^{-1}S\right\rangle\right\}\right] = E\left[\exp\left\{i\left\langle \kappa^{-1}x, S\right\rangle\right\}\right] \\
&= \exp\left\{-\kappa^{-p} \int_{\mathbb{S}_\mathbb{H}} |\langle x, s\rangle|^p \sigma_S(ds) + i \tan\frac{\pi p}{2}\kappa^{-p} \int_{\mathbb{S}_\mathbb{H}} \langle x, s\rangle\, |\langle x, s\rangle|^{p-1} \sigma_S(ds)\right\} \\
&= \exp\left\{-s_p \int_{\mathbb{S}_\mathbb{H}} |\langle x, s\rangle|^p \frac{\sigma_S(ds)}{\sigma_S(\mathbb{S}_\mathbb{H})} + i \tan\frac{\pi p}{2} s_p \int_{\mathbb{S}_\mathbb{H}} \langle x, s\rangle\, |\langle x, s\rangle|^{p-1} \frac{\sigma_S(ds)}{\sigma_S(\mathbb{S}_\mathbb{H})}\right\}.
\end{aligned}$$

Again, Theorem 1 gives that $\sigma_S = \Gamma_Z'$, with Γ_Z' from Condition 1. The result follows from (12).

For a given Hilbert space $\mathbb{H}$ let $\mathcal{S} = \mathcal{S}_{\mathbb{H}}$ be the space of Hilbert–Schmidt operators. If $x, y \in \mathbb{H}$, then $x \otimes y \in \mathcal{S}$ is defined by $(x \otimes y)(z) = \langle x, z\rangle\, y$, $z \in \mathbb{H}$. For $\Psi, \Phi \in \mathcal{S}$, we define

$$\langle \Psi, \Phi\rangle_{\mathcal{S}} = \sum_{j=1}^{\infty} \left\langle \Psi(e_j), \Phi(e_j)\right\rangle,$$

where $\{e_j, j \geq 1\}$ is an orthonormal basis in $\mathbb{H}$.

We want to link the regular variation of X to that of $X \otimes X$. We first state a simple fact that will allow us to deal with sample covariance operators even if the corresponding population covariance operator does not exist. We will use it with $\mathbb{H} = L^2(\mathcal{U})$.

Fact 1. If X is a random element of $\mathbb{H}$, then $X \otimes X \in \mathcal{S}$ with probability one and

$$\|X \otimes X\|_{\mathcal{S}} = \|X\|^2 \quad a.s.$$

Proof. For any orthonormal basis $\{e_j, j \geq 1\}$ in $\mathbb{H}$,

$$\|X \otimes X\|_{\mathcal{S}}^2 = \sum_{j=1}^{\infty} \| \left\langle X, e_j\right\rangle X\|^2 = \|X\|^2 \sum_{j=1}^{\infty} | \left\langle X, e_j\right\rangle |^2 = \|X\|^4.$$

Since $X \in \mathbb{H}$ a.s., it follows that $\|X\| < \infty$ a.s. and so $\|X \otimes X\|_{\mathcal{S}} < \infty$ a.s., and the norm identity follows.

The next result is taken from Kokoszka et al. [17].

Proposition 3. *If $X \in \mathrm{RV}_{-\alpha}(\mathbb{H})$, $\alpha > 0$, then $X \otimes X \in \mathrm{RV}_{-\alpha/2}(\mathcal{S}_{\mathbb{H}})$.*

3. General convergence results

In this section, we derive limiting behavior of the sample covariance operator and establish consistence of the sample FPCs. Proofs that are not given in this section, are given in Section 6, so as not to interrupt the narrative.

3.1. Asymptotic behavior of the sample covariance operator

On reflection, the following fact is elementary, but it emphasizes that no moment conditions are needed to define the *sample* covariance operator and that it has desirable properties with probability one.

Fact 2. Suppose $X_1, X_2, \ldots, X_N$ are random elements of $L^2(\mathcal{U})$. Then the sample covariance operator defined by

$$\widehat{C}(x) = \frac{1}{N} \sum_{n=1}^{N} \left\langle X_n - \bar{X}_N, x\right\rangle (X_n - \bar{X}_N), \quad x \in L^2(\mathcal{U}), \tag{23}$$

is a random element of $\mathcal{S}$ that is a.s. symmetric, $\langle \widehat{C}(x), y\rangle = \langle x, \widehat{C}(y)\rangle$, and nonnegative, $\langle \widehat{C}(x), x\rangle \geq 0$.

Proof. Since $L^2(\mathcal{U})$ is a vector space, $X_n - \bar{X}_N \in L^2(\mathcal{U})$. By Fact 1, $(X_n - \bar{X}_N) \otimes (X_n - \bar{X}_N)$ is a random element of $\mathcal{S}$. Since $\mathcal{S}$ is a vector space, $\widehat{C} \in \mathcal{S}$. The claims that $\widehat{C}$ is a.s. symmetric and nonnegative follow from its definition.

Next, we formulate the main assumption of this paper.

Assumption 1. The random function X is regularly varying in $L^2 = L^2(\mathcal{U})$ with tail index $\alpha \in (0, 2)$ and the angular measure Γ_X. The random functions $X_1, X_2, \ldots, X_N$ are i.i.d. copies of X.

In relation to (11) in Condition 1, we introduce the following assumption.

Assumption 2. For an orthonormal basis $e_j, j \geq 1$, in L^2, set $\pi_J(x) = \sum_{j \geq J} \left\langle x, e_j\right\rangle e_j$. We assume that for any $I, J \geq 1$, there are constants $c(I, J) > 0$ such that

$$\lim_{u\to\infty} \frac{P\left(\|\pi_I(X)\|\ \|\pi_J(X)\| > u\right)}{P\left(\|X\|^2 > u\right)} \to \frac{c(I, J)}{c(1, 1)},$$

and $c(I, J) \to 0$, as $I, J \to \infty$.

In conjunction with Assumption 1, Assumption 2 means that the distribution of X is regularly varying on the whole space L^2, not a finite dimensional subspace, but the projections $\pi_J(X)$ become asymptotically negligible relative to the distribution of X as $J \to \infty$.

We can now state the most general result of our paper. Recall that the constant s_p is defined by (21).

Theorem 2. *Suppose Assumptions 1 and 2 hold. Denote by a_N the quantile sequence of X defined by (9). Then*

$$Na_N^{-2}\widehat{C} \xrightarrow{d} C_\infty,$$

where C_∞ is an $(\alpha/2)$-stable Hilbert–Schmidt operator with the characteristic functional

$$E\left[\exp\left\{i\left\langle\phi, C_\infty\right\rangle_{\mathcal{S}}\right\}\right] = \exp\left\{-s_{\alpha/2}\int_{\mathbb{S}_{L^2}} \mid \langle\phi, x\otimes x\rangle_{\mathcal{S}}\mid^{\alpha/2}\Gamma_X(dx) + iC(\alpha,\phi)\right\},$$

with the scalar $C(\alpha,\phi)$ given by

$$C(\alpha,\phi) = s_{\alpha/2}\tan\left(\frac{\pi\alpha}{4}\right)\int_{\mathbb{S}_{L^2}} \langle\phi, x\otimes x\rangle_{\mathcal{S}} \mid \langle\phi, x\otimes x\rangle_{\mathcal{S}}\mid^{\alpha/2-1}\Gamma_X(dx).$$

Remark 4. The sequence a_N in Theorem 2 has the representation $a_N = N^{1/\alpha}L_a(N)$ with a slowly varying function L_a. Ignoring the slowly varying function, one can say that, roughly, $\widehat{C} \sim C_\infty N^{2/\alpha-1}$. Since $\alpha \in (0,2)$, $N^{2/\alpha-1} \to \infty$. Thus, $\widehat{C}$ does not converge to any finite limit and $\|\widehat{C}\|_{\mathcal{S}} \xrightarrow{P} \infty$.

Remark 5. The proof of Theorem 2 shows, c.f. Lemma 4, that the non-centered covariance operator

$$C_N = \frac{1}{N}\sum_{n=1}^{N} X_n \otimes X_n \tag{24}$$

has the same asymptotic distribution as the usual covariance operator $\widehat{C}$. This is the effect of very heavy tails, $\alpha/2 < 1$, that suppress averaging. This effect is fairly well-known in the case of scalar observations, see e.g. Section 4 of Davis and Resnick [7]. In the case of scalar observations with finite variance, the asymptotic distributions of $N^{-1}\sum_{i=1}^{N}(X_i - \bar{X}_N)^2$ and $N^{-1}\sum_{i=1}^{N} X_i^2$ differ by $\sqrt{N}(E[X_1])^2$.

Remark 6. The proof of Theorem 2 critically relies on Corollary 1, which in turn relies on the centering (18). The universal centering (17) in Theorem 1 leads to Theorem 3. We state it below in order to provide a precise results that shows a constant, deterministic shift in the limit.

Theorem 3. *Suppose the conditions of Theorem 2 hold and γ_N is given by (17) with $Z = X \otimes X$. Then*

$$a_N^{-2}\left(N\widehat{C} - \gamma_N\right) \xrightarrow{d} C_\infty - \Psi_X(\alpha),$$

where the deterministic operator $\Psi_X(\alpha)$ is given by

$$\Psi_X(\alpha) := \kappa^{\alpha/2-1}\frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}}(x\otimes x)\Gamma_X(dx),$$

with κ given in (22) with $p = \alpha/2$.

Remark 7. The above result involves the constant κ. It can be eliminated from the limit by replacing b_N in the definition of γ_N with the quantile sequence $a_{N,Z}$. From Lemma 7 we can conclude that

$$a_N^{-2}\left(N\widehat{C} - \tilde{\gamma}_N\right) \xrightarrow{d} C_\infty - \frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}}(x\otimes x)\Gamma_X(dx),$$

with

$$\tilde{\gamma}_N = NE\left[ZI_{\{||Z||_{\mathcal{S}}\le a_{N,Z}\}}\right].$$

To deal with the convergence of the FPCs, we will need almost sure convergence of $Na_N^{-2}\widehat{C}$ to C_∞. The following theorem is a direct consequence of Theorem 13 in Section IV.3 of Pollard [23] because every point of a separable metric space is completely regular in the sense of Definition 6 of Pollard [23]. The measurability conditions does not come into play because we work with Borel σ-algebras and Borel-measurable functions.

Theorem 4. *Suppose $\mathcal{X}$ is a separable metric space and μ_N, μ probability measures on $\mathcal{X}$. If the μ_N converge weakly to μ, then there are random elements X_N and X with distributions, respectively, μ_N and μ such that $X_N \to X$ almost surely.*

Using Theorem 4, we can prove the following result.

Lemma 1. *There are versions $\widetilde{C}_N$ and $\widetilde{C}_\infty$ of $\widehat{C}$ and C_∞ such that*

$$Na_N^{-1}\widetilde{C}_N \overset{a.s.}{\to} \widetilde{C}_\infty.$$

Moreover, $\widetilde{C}_N$ and $\widetilde{C}_\infty$ are symmetric and nonnegative with probability 1.

Proof. The almost sure convergence follows directly from Theorem 4. Since $L^2(\mathcal{U})$ is separable, the following sets are Borel subsets of $\mathcal{S}$:

$$B_{\rm sym} = \{\Phi \in \mathcal{S} : \langle \Phi(x), y\rangle = \langle x, \Phi(y)\rangle \ \text{ for all } x, y \in L^2\},$$

$$B_{\rm nn} = \{\Phi \in \mathcal{S} : \langle \Phi(x), x\rangle \geq 0 \text{ for all } x \in L^2\}.$$

By Fact 2, $\widehat{C}$ is symmetric and nonnegative almost surely, that is

$$P(\widehat{C} \in B_{\rm sym}) = 1, \quad P(\widehat{C} \in B_{\rm nn}) = 1.$$

Since $\widetilde{C}_N$ and $\widehat{C}$ have the same distribution,

$$P(\widetilde{C}_N \in B_{\rm sym}) = 1, \quad P(\widetilde{C}_N \in B_{\rm nn}) = 1.$$

Hence, each $\widetilde{C}_N$ is symmetric and nonnegative almost surely.

The map $\mathcal{S} \ni \Phi \mapsto \langle \Phi(x), x\rangle$ is continuous, so

$$\left\langle Na_N^{-1}\widetilde{C}_N(x), x\right\rangle \overset{a.s.}{\to} \left\langle \widetilde{C}_\infty(x), x\right\rangle \geq 0.$$

Similarly, we conclude that

$$\left\langle \widetilde{C}_\infty(x), y\right\rangle = \left\langle \widetilde{C}_\infty(y), x\right\rangle \quad a.s.$$

The above two relations are at this point established for fixed x and y, i.e. they hold on events Ω_x and $\Omega_{x,y}$ of probability 1. By the separability of L^2, we can chose a dense subset $\{x_i, i \geq 1\}$ such that $\langle \widetilde{C}_\infty(\omega)(x_i), x_i\rangle \geq 0$ for $\omega \in \Omega_c$ with $P(\Omega_c) = 1$. Since the map $L^2 \ni x \mapsto \langle \Phi(x), x\rangle$ is continuous,- we conclude that for $\omega \in \Omega_c$, $\langle \widetilde{C}_\infty(\omega)(x), x\rangle \geq 0$ for each $x \in L^2$. A similar argument shows that there is a probability 1 event on which the symmetry of $\widetilde{C}_\infty$ holds for all x, y.

We will use the following corollary.

Corollary 2. *The limiting operator C_∞ is symmetric and nonnegative almost surely.*

Proof. Since $\widetilde{C}_\infty$ and C_∞ have the same distribution, as in the proof of Lemma 1, $P(C_\infty \in B_{\rm sym}) = P(\widetilde{C}_\infty \in B_{\rm sym}) = 1$ and $P(C_\infty \in B_{\rm nn}) = P(\widetilde{C}_\infty \in B_{\rm nn}) = 1$.

3.2. Convergence of sample FPCs

We now turn to the study of the asymptotic behavior of the eigenfunctions and the eigenvalues of $\widehat{C}$, i.e. the sample functional principal components and their analysis of variance. If $E\|X\|^2 < \infty$, the population covariance operator C exists and one generally imposes the assumption $\lambda_1 > \lambda_2 > \cdots > \lambda_p > 0$ on its eigenvalues to ensure that the estimation targets are uniquely defined. The asymptotic distribution of the estimated eigenfunctions, the estimated FPCs $\hat{v}_j$ can then be formulated using the deterministic eigenfunctions v_j of C and their eigenvalues λ_j that are fixed numbers, see Kokoszka and Reimherr [15] and Kokoszka et al. [17]. In the setting of this paper, the population covariance operator does not exist and the objects discussed above are not defined.

Recall that if Ψ is a symmetric, nonnegative Hilbert–Schmidt operator, then it admits the spectral decomposition

$$\Psi(x) = \sum_{j=1}^{\infty} \lambda_j \left\langle x, e_j\right\rangle e_j, \ \ x \in L^2, \tag{25}$$

with $\lambda_j \geq 0$ and an orthonormal basis $\{e_j, j \geq 1\}$ depending on Ψ. Due to Fact 2, $\widehat{C}$ is a.s. symmetric and nonnegative. Hence, it admits decomposition (25). Since $\widehat{C}(x) \in \operatorname{span}(X_1, \ldots, X_N)$, there are at most N nonzero $\hat{\lambda}_j$ in the spectral representation of $\widehat{C}$, i.e. with probability 1,

$$\widehat{C}(x) = \sum_{j=1}^{N} \hat{\lambda}_j \left\langle \hat{v}_j, x\right\rangle \hat{v}_j, \quad x \in L^2. \tag{26}$$

The sample eigenvalues $\hat{\lambda}_j$ and eigenfunctions $\hat{v}_j$ are well-defined random objects.

If $\alpha > 2$, the covariance operator C defined by (2) has the spectral decomposition

$$C(x) = \sum_{j=1}^{\infty} \lambda_j \langle v_j, x \rangle v_j, \tag{27}$$

where λ_j's are its (deterministic) eigenvalues and v_j are (deterministic) functions, and X can be expressed as $X = \sum_{j=1}^{\infty} \xi_j v_j$ with $\xi_j = \langle X, v_j \rangle$. One can show that $q_N(\hat{\lambda}_j - \lambda_j)$ and $q_N(\hat{v}_j - v_j)$ converge in distribution at some specific rate $q_N \to \infty$, see Kokoszka and Reimherr [15] and Kokoszka et al. [17].

In case of $\alpha < 2$, the limit C_∞ in Theorem 2 is random. However, thanks to Corollary 2, C_∞ is both symmetric and nonnegative a.s., and hence it admits a spectral representation:

$$C_\infty(x) = \sum_{j=1}^{\infty} \Lambda_j \langle V_j, x \rangle V_j, \tag{28}$$

where the V_js are random functions, orthonormal with probability 1, and Λ_j are random variables, nonnegative with probability 1. The following two results (Proposition 4 and Theorem 5) clarify the relationship between the estimated quantities in (26) and the asymptotic quantities in (28).

Set $r_N := N a_N^{-2}$. By Remark 4, $r_N \to 0$. In the case of $\alpha > 2$, analogous normalizing constants tend to infinity. By Lemma 1, there are representations $\widetilde{C}_N$ and $\widetilde{C}_\infty$, respectively, of $\widehat{C}$ and C_∞ such that $r_N \widetilde{C}_N \stackrel{a.s.}{\to} \widetilde{C}_\infty$. In what follows, we drop the $\widetilde{\ }$ in the notation for the versions, unless there is a possibility for a confusion.

Our first result shows that *rescaled* eigenvalues $\hat{\lambda}_j$ in (26) approximate the Λ_js in (28).

Proposition 4. *Consider the $\hat{\lambda}_j$ in (26) and the Λ_js in (28). Under the conditions of Theorem 2, $r_N \hat{\lambda}_j \stackrel{d}{\to} \Lambda_j$. (Recall, $r_N := N a_N^{-2}$.)*

Proof. By Lemma 1, we can assume that $\widehat{C}$ and C_∞ are defined on the same probability space and $r_N \widehat{C} \to C_\infty$ almost surely, i.e.

$$\| r_N \widehat{C} - C_\infty \|_{\mathcal{S}} \to 0, \quad a.s. \quad (N \to \infty). \tag{29}$$

The sample covariance operator has the spectral decomposition (26). The limiting operator C_∞ has the spectral decomposition (28). Both are a.s. Hilbert–Schmidt and hence compact. Therefore, by Lemma 2.2 in Horváth and Kokoszka [12],

$$| r_N \hat{\lambda}_j - \Lambda_j | \le \| r_N \widehat{C} - C_\infty \|_{\mathcal{S}} \ \ a.s. \tag{30}$$

By (29), on the common probability space, $r_N \hat{\lambda}_j \to \Lambda_j$ a.s., implying the claim.

The next theorem shows that the sample FPCs $\hat{v}_j$ in (26) converge to the eigenfunctions V_j in (28) under a suitable condition that separates the eigenvalues to ensure that the limiting eigenspaces are a.s. one-dimensional. Our condition (31) is similar in spirit to the assumption $\lambda_1 > \lambda_2 > \cdots > 0$ used when the population covariance operator exists.

Theorem 5. *Suppose the assumptions of Theorem 2 hold, and j is such that*

$$\inf_{k \neq j} (\Lambda_j - \Lambda_k)^2 > 0 \ \ a.s. \tag{31}$$

Then, $\hat{v}_j \stackrel{d}{\to} \operatorname{sign}(\langle \hat{v}_j, V_j \rangle) V_j$.

The $\hat{v}_j$ and the V_j in Theorem 5 are arranged by decreasing eigenvalues for each outcome for which the convergence (29) holds. Detailed proof is given in the proof of Theorem 5.

In Section 4, we illustrate the results of this section in some special cases. We will see that even in the finite dimensional case, results widely used in FDA no longer hold. Consequently, outputs of standard FDA procedure must be interpreted with care.

4. Multivariate observations

The purpose of this section is to describe the structure of the limit C_∞ as well as the form of its eigenvalues and eigenfunctions in commonly encountered multivariate settings. This section contains discussion and informative arguments. More involved proofs are provided in Section 7. Throughout this section, we work under the following assumption.

Assumption 3. Suppose $X = \sum_{j=1}^{d} \xi_j v_j$, where the vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)$ is regularly varying with index $\alpha \in (0, 2)$ and the angular measure $\Gamma_{\boldsymbol{\xi}}$. The deterministic functions v_j satisfy $\|v_j\| = 1$, $\langle v_i, v_j \rangle = 0$ for $i \neq j$.

Assumption 3 is particularly relevant in simulation studies that, by necessity can use only finite sums. For example, the v_j's are often taken to be the eigenfunctions in the Karhunen–Loéve expansion of the Brownian motion or the Brownian bridge, i.e., respectively, $v_j(u) = \sqrt{2}\sin((j-1/2)\pi u)$, $v_j(u) = \sqrt{2}\sin(\pi ju)$, $u \in [0, 1]$. The results of this section pertain, in particular, to the relationship between the known and deterministic v_j, which could be viewed as population FPCs, and their estimators, if the ξ_j have infinite variance.

Lemma 2. *Under Assumption 3, $X \in \mathrm{RV}_{-\alpha}(\mathbb{H})$, where $\mathbb{H}$ is the subspace of L^2 space spanned by $\{v_1, \dots, v_d\}$.*

We reformulate Theorem 2 in the finite dimensional case. Instead of Assumptions 1 and 2, we impose Assumption 3. Note that Assumption 3 excludes Assumption 2 and as such the finite dimensional case does not follow from Theorem 2.

Let $|\cdot|$ be a norm on $\mathbb{R}^k$, $k = d^2$. For the vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$, define a vector Y in $\mathbb{R}^k$ by $Y = g(\boldsymbol{\xi})$ with $g(x_1, \dots, x_d) = (x_i x_j, i, j \in \{1, \dots, d\})$. Let d_N be a sequence defined by

$$NP\left(|Y| > d_N^2\right) = 1. \tag{32}$$

Theorem 6. *Suppose Assumption 3 holds. Then*

$$Nd_N^{-2}\widehat{C} \xrightarrow{d} C_\infty := \sum_{i,j=1}^{d} \Lambda_{i,j}(v_i \otimes v_j),$$

where the random vector $\tilde{\boldsymbol{\Lambda}} = (\Lambda_{i,j}, i, j \in \{1, \dots, d\})$ in $\mathbb{R}^k$, $k = d^2$, has the characteristic function

$$E\left[\exp\left\{i\langle\theta, \tilde{\boldsymbol{\Lambda}}\rangle\right\}\right] = \exp\left\{-s_p \int_{\mathbb{S}_{\mathbb{R}^d}} \left|\sum_{j=1}^{d} \theta_{i,j}x_i x_j\right|^p \left(1 - i\,\mathrm{sign}\left(\sum_{j=1}^{d} \theta_{i,j}x_i x_j\right)\tan\left(\frac{\pi p}{2}\right)\right)\Gamma_{\xi}(dx)\right\} \tag{33}$$

with $\theta = (\theta_{i,j}, i, j \in \{1, \dots, d\})$.

In what follows, we investigate special cases of Theorem 6.

4.1. One-dimensional case

We begin the illustration of Theorem 6 with the simplest possible case of $X = \xi v$, $v \in L^2$, $\|v\| = 1$ and ξ is a regularly varying random variable with index $\alpha \in (0, 2)$. That is, for all $t > 0$,

$$\lim_{u\to\infty} \frac{P(|\xi| > tu)}{P(|\xi| > u)} = t^{-\alpha}\ , \quad \lim_{u\to\infty} \frac{P(\xi > u)}{P(|\xi| > u)} = q_+ = 1 - \lim_{u\to\infty} \frac{P(\xi < -u)}{P(|\xi| > u)}. \tag{34}$$

The angular measure of ξ (in the sense of Proposition 1) is $\Gamma_\xi = q_+\delta_1 + q_-\delta_{-1}$, $q_+ + q_- = 1$. Note that X is regularly varying as well. Indeed, we have $\|X\| = |\xi|$ and

$$\frac{X}{\|X\|} = \mathrm{sign}(\xi)v\ .$$

For any $A \subseteq \mathbb{S}_{L^2}$, denote $-A = \{-a : a \in A\}$. Regular variation of ξ implies

$$\frac{P(\|X\| > tu, X/\|X\| \in A)}{P(\|X\| > u)} = \frac{P(|\xi| > tu, \mathrm{sign}(\xi)v \in A)}{P(|\xi| > u)} \to t^{-\alpha}\left\{q_+\delta_v(A) + q_-\delta_v(-A)\right\}, \quad u \to \infty.$$

Thus, X is regularly varying and its angular measure (again in the sense of Proposition 1) is $\Gamma_X = q_+\delta_v + q_-\delta_{-v}$. Hence, Assumption 3 holds. Then $Z = X \otimes X = \xi^2(v \otimes v)$ is also regularly varying and its angular measure is $\Gamma_Z = \delta_{v\otimes v}$. What is the limit C_∞ in this case? We will argue that

$$C_\infty = \Lambda(v \otimes v), \quad \Lambda \sim S_p(s_p^{1/p}, 1, 0). \tag{35}$$

We will approach it from two directions. We first present a direct argument. Suppose $X_n = \xi_n v$, $n \geq 1$, where the ξ_n are i.i.d. with the same distribution as ξ in (34). Then

$$C_N = N^{-1}\sum_{n=1}^{N} X_n \otimes X_n = (v \otimes v)N^{-1}\sum_{n=1}^{N} \xi_n^2.$$

Let c_N be defined by $NP(|\xi| > c_N) = 1$. By Lemma 9, $c_N^{-2}\sum_{n=1}^{N}\xi_n^2 \xrightarrow{d} S_p(s_p^{1/p}, 1, 0)$. Recall that a_N is chosen as $NP(\|X\| > a_N) = 1$. Since $\|X\| = |\xi|$, the c_N in Lemma 9 and the a_N in Proposition 1 coincide. This directly verifies that $Na_N^{-2}C_N$ converges to C_∞ in (35).

Now we show how this result follows from Theorem 6. We also make a link to Theorem 2, even though the latter theorem is not applicable here (Assumption 2 does not hold). For $\phi \in \mathcal{S}$ and C_∞ in (35),

$$Ee^{i\langle \phi, C_\infty \rangle} = \exp\left\{-s_p| \langle \phi, v \otimes v \rangle_{\mathcal{S}} |^p \left(1 - i\mathrm{sign}(\langle \phi, v \otimes v \rangle_{\mathcal{S}}) \tan\left(\frac{\pi p}{2}\right)\right)\right\}.$$

Note that with $p = \alpha/2$,

$$\int_{\mathbb{S}_{L^2}} | \langle \phi, x \otimes x \rangle_{\mathcal{S}} |^p \Gamma_X(dx) = (q_+ + q_-)| \langle \phi, v \otimes v \rangle_{\mathcal{S}} |^p = | \langle \phi, v \otimes v \rangle_{\mathcal{S}} |^p$$

and

$$\begin{aligned}\int_{\mathbb{S}_{L^2}} \langle \phi, x \otimes x \rangle_{\mathcal{S}} | \langle \phi, x \otimes x \rangle_{\mathcal{S}} |^{p-1} \Gamma_X(dx) = q_+ \langle \phi, v \otimes v \rangle_{\mathcal{S}} | \langle \phi, v \otimes v \rangle_{\mathcal{S}} |^{p-1} \\ + q_- \langle \phi, (-v) \otimes (-v) \rangle_{\mathcal{S}} | \langle \phi, (-v) \otimes (-v) \rangle_{\mathcal{S}} |^{p-1} \\ = \langle \phi, v \otimes v \rangle_{\mathcal{S}} | \langle \phi, v \otimes v \rangle_{\mathcal{S}} |^{p-1}.\end{aligned}$$

Therefore, the characteristic functional in Theorem 6 (and Theorem 2) coincides with the characteristic functional of C_∞ in (35). We summarize the discussion above as the following fact. We note that the limiting eigenvalue is random, while the limiting eigenfunction is deterministic.

Fact 3. If $X = \xi v$, where ξ satisfies (34) and v is a unit length element of L^2, then

$$Na_N^{-2} C_N \xrightarrow{d} C_\infty = S_p(s_p^{1/p}, 1, 0)(v \otimes v) \ \ (p = \alpha/2)$$

with a_N defined by $NP(|\xi| > a_N) = 1$.

4.2. Extremal independence

The next result identifies the operator C_∞ in the case of extremally independent components. It follows from Theorem 6, by considering the specific form of the angular measure as well as the relation between the sequence d_N in (32) and the quantile sequence a_N of $\|X\|$ defined by

$$NP(\|X\| > a_N) = 1.$$

Proposition 5. *Suppose Assumption 3 holds with the angular measure Γ_ξ concentrated at the points $e_j, j \in \{1, 2, \dots, d\}$, where the e_j are the standard coordinate vectors in $\mathbb{R}^d$. Let $\mathbb{H}$ be spanned by $\{v_1, \dots, v_d\}$. Then*

$$Na_N^{-2} C_N \xrightarrow{d} C_\infty = \sum_{j=1}^d \Lambda_j (v_j \otimes v_j),$$

where the Λ_j are independent p-stable random variables with the characteristic function

$$E\left[\exp\left\{i\theta \Lambda_j\right\}\right] = \exp\left\{-s_p \sigma_j^p |\theta|^p \left(1 - i\mathrm{sign}\,(\theta) \tan\left(\frac{\pi p}{2}\right)\right)\right\}$$

with $\sigma_j^p = \Gamma_\xi(e_j)$.

Remark 8. The characteristic function of $\mathbf{\Lambda} = (\Lambda_1, \dots, \Lambda_d)$ can be written as

$$E\left[\exp\left\{i\langle \theta, \mathbf{\Lambda} \rangle\right\}\right] = \exp\left\{-s_p \int_{\mathbb{S}_{\mathbb{R}^d}} \left(\sum_{j=1}^d \theta_j s_j^2\right)^p \left(1 - i\mathrm{sign}\left(\sum_{j=1}^d \theta_j s_j^2\right) \tan\left(\frac{\pi p}{2}\right)\right) \Gamma_\xi(ds)\right\}. \quad (36)$$

Remark 9. We note that C_∞ has the representation (28) with $V_j = v_j$, yielding $C_\infty(v_j) = \Lambda_j v_j$. The eigenvalues are random, but the eigenfunctions are deterministic. Notice also that since the Λ_j are independent,

$$\min_{k \in \{1,2,\dots,d\} \setminus \{j\}} (\Lambda_j - \Lambda_k)^2 > 0 \quad a.s.$$

Therefore, condition (31) holds, and the same proof as in the infinite dimensional case shows that $\hat{v}_j \xrightarrow{P} \mathrm{sign}(\langle \hat{v}_j, v_j \rangle) v_j$ for each $j \in \{1, 2, \dots, d\}$.

Remark 10. If the covariance operator C defined by (2) exists, then its largest eigenvalue and the corresponding eigenfunction (the first FPC) satisfy, respectively, $\lambda_1 = \max_{x\in\mathbb{S}_{L^2}} \langle C(x), x\rangle$ and $v_1 = \arg\max_{x\in\mathbb{S}_{L^2}} \langle C(x), x\rangle$. In the context of this paper, we can analogously define

$$\lambda_{1,\infty} = \max_{x\in\mathbb{S}_{L^2}} \langle C_\infty(x), x\rangle\,, \quad v_{1,\infty} = \arg\max_{x\in\mathbb{S}_{L^2}} \langle C_\infty(x), x\rangle\,. \tag{37}$$

Since all Λ_l's are positive, we have for any unit length $x = \sum_l \beta_l v_l \in L^2$,

$$\langle C_\infty(x), x\rangle = \sum_{j=1}^d \Lambda_j \left\langle v_j, \sum_l \beta_l v_l \right\rangle^2 = \sum_{j=1}^d \Lambda_j \beta_j^2 \le \Lambda_{\max} \sum_{j=1}^d \beta_j^2 = \Lambda_{\max}.$$

Therefore, $\lambda_{1,\infty} = \Lambda_{\max}$, where $\Lambda_{\max} = \max\{\Lambda_1, \Lambda_2, \ldots, \Lambda_d\}$. Also, we immediately obtain that $v_{1,\infty} = v_J$, where J is an integer-valued random variable defined by $J = \arg\max\{\Lambda_j, j \in \{1,\ldots,d\}$. We note that defining the first FPC via (37) is not the same as defining it as the first eigenfunction of the limit C_∞. Observe that $v_{1,\infty}$ is random because it can be any of the d deterministic functions used to define X in Assumption 3. A take away is that for infinite variance functional data these two definitions of FPCs are no longer equivalent.

Remark 11. The assumption of Proposition 5 holds when $\xi_1,\ldots,\xi_d$ are independent, all regularly varying with index $-\alpha$. If the ξ_j have the same distribution, then $\Gamma_\xi(e_j) = 1/d$ and hence the limiting random variables Λ_j are $S_p((s_p/d)^{1/p}, 1, 0)$. On the other hand, if $\xi_j \overset{d}{=} \gamma_j\tilde{\xi}_j$, $\gamma_j \neq 0$ and $\tilde{\xi}_j$, $j \in \{1,\ldots,d\}$, are i.i.d., then $\Gamma_\xi(e_j) = |\gamma_j|^\alpha / \sum_{i=1}^d |\gamma_i|^\alpha$. These quantities appear in Theorem 7 that extends Proposition 5 to the case $d = \infty$.

Remark 12. Let $I \subset \{1,\ldots,d\}$. It is possible that some of the random variables $\xi_j, j \in I$, have asymptotically smaller tails than $\xi_j, j \notin I$. Then $\Gamma_\xi(e_j) = 0$ for $j \in I$ and the sum in the statement of Proposition 5 reduces to the sum over $j \notin I$.

Remark 13. The assumption of Proposition 5 also holds when $\xi_j = A_jV_j$, where V_j are independent and regularly varying with index $-\alpha$, while $A_1,\ldots,A_d$ independent of $V_1,\ldots,V_d$ such that $E\left[|A_j|^{\alpha+\epsilon}\right] < \infty$ for some $\epsilon > 0$. Note that $A_1,\ldots,A_d$ do not need to be independent between themselves.

As noted in Remark 11, under suitable summability conditions, Proposition 5 can be extended to an infinite deterministic basis $v_1, v_2, \ldots$. The extension is presented in Section 5.

4.3. Extremal linear dependence

Now, we consider a special case of extremal dependence. Assume that ξ_0 is a regularly varying random variable with index $-\alpha$, $\alpha \in (0,2)$, and $\boldsymbol{\zeta} = (\zeta_1,\ldots,\zeta_d)$ is a random vector, independent of ξ_0, such that $E\left[|\boldsymbol{\zeta}|^{\alpha+\epsilon}\right] < \infty$ for the Euclidean norm $|\cdot|$ on $\mathbb{R}^d$ and $\epsilon > 0$. Define $\xi_i = \xi_0 + \zeta_i$, $i \in \{1,\ldots,d\}$. Then the random vector $\boldsymbol{\xi} = (\xi_1,\ldots,\xi_d)$ fulfills Assumption 3 and the angular measure $\Gamma_{\boldsymbol{\xi}}$ is concentrated on $(1,\ldots,1)/|(1,\ldots,1)| =: (x_0,\ldots,x_0)$. The characteristic exponent in (33) becomes

$$-s_p x_0^\alpha |\sum_{j=1}^d \theta_{i,j}|^{\alpha/2} \left(1 - i\mathrm{sign}\left(\sum_{j=1}^d \theta_{i,j}\right) \tan\left(\frac{\pi p}{2}\right)\right).$$

We note that this is the characteristic exponent of a random vector $(\Lambda,\ldots,\Lambda)$ in $\mathbb{R}^k$, $k = d^2$, where $\Lambda \sim S_p(s_p^{1/p}, 1, 0)$. Therefore, C_∞ in Theorem 6 has the representation

$$C_\infty = \Lambda \sum_{i,j=1}^d (v_i \otimes v_j).$$

Remark 14. Similarly to Remark 10, we want to determine $\lambda_{1,\infty}$ and $v_{1,\infty}$ given by (37). For $\Lambda_{i,j}$, $i,j \in \{1,\ldots,d\}$ in Theorem 6, let $\boldsymbol{\Sigma} = (\Lambda_{i,j})_{i,j=1}^d$ be a symmetric random matrix. For $x = \sum_{l=1}^d \beta_l v_l$, $\sum_{j=1}^d \beta_j^2 = 1$, we have

$$\langle C_\infty(x), x\rangle = \boldsymbol{\beta}^\top \boldsymbol{\Sigma} \boldsymbol{\beta} =: Q(\boldsymbol{\beta}).$$

The maximal value of the quadratic form $Q(\boldsymbol{\beta})$, subject to the constraints $\sum_{j=1}^d \beta_j^2 = 1$, equals $\Sigma_{\max}$, where $\Sigma_{\max}$ is the largest eigenvalue of the random matrix $\boldsymbol{\Sigma}$. Therefore, $\lambda_{1,\max} = \Sigma_{\max}$

In the current situation, $\boldsymbol{\Sigma} = \Lambda\mathbf{1}$, where $\mathbf{1}$ is $d\times d$ matrix of ones. The matrix has two eigenvalues: $d\Lambda$ (with multiplicity 1) and 0 (with multiplicity d). Hence, the eigenfunction associated to $d\Lambda$ is $\beta(v_1 + \cdots + v_d)$ with $\beta > 0$. Therefore, the eigenfunctions are deterministic, but they do not agree with v_j's in Assumption 3.

In the Online Material, we discuss and example of functional data that may exhibit extremal dependence.

5. Expansion using an infinite deterministic basis

In this section, we assume that $X = \sum_{j=1}^{\infty} \xi_j v_j$, where the v_j are deterministic functions satisfying $\|v_j\| = 1$, $\langle v_i, v_j \rangle = 0$ for $i \neq j$. In contrast to Assumption 3, we need to make specific assumptions on ξ_js in order to (a) ensure summability of the series; (b) guarantee regular variation.

Assumption 4. Suppose $X = \sum_{j=1}^{\infty} \xi_j v_j$, where ξ_j, $j \geq 1$, are independent random variables. Furthermore, there exists a regularly varying random variable ξ_0 with index $\alpha \in (0, 2)$ and a sequence of real numbers $\gamma_j \neq 0$ such that

$$\xi_j \overset{d}{=} \gamma_j \xi_0, \quad j \geq 1, \tag{38}$$

and there exists $\epsilon > 0$ such that

$$\sum_{j=1}^{\infty} |\gamma_j|^{(\alpha-\epsilon)\wedge 1} < \infty. \tag{39}$$

Remark 15. If $\alpha \in (1, 2)$ and $E[\xi_j] \neq 0$, then the summability of the series $\sum_{j=1}^{\infty} |\gamma_j|$ is necessary for X to be well-defined. If $E[\xi_j] = 0$, then it is enough to assume $\sum_{j=1}^{\infty} |\gamma_j|^{\alpha-\epsilon} < \infty$, see Corollary 4.2.1 in [19].

Proposition 6 is basically a special case of Proposition 7.1 in Meinguet and Segers [22]. A detailed proof is not presented due to a page limit, but is available upon request.

Proposition 6. *Under Assumption 4, $X \in \mathrm{RV}_{-\alpha}(\mathbb{H})$, where $\mathbb{H}$ is the subspace of L^2 space spanned by $\{v_1, v_2, \ldots, \}$.*

To establish the convergence of the sample covariance operator, we must strengthen Assumption 4 to the following assumption.

Assumption 5. Assume that $\xi_{0,j} \overset{d}{=} \gamma_j \tilde{\xi}_{0,j}$, where $\tilde{\xi}_{0,j}, j \geq 1$ are i.i.d regularly varying with index $\alpha \in (0, 2)$ and the same distribution as $\tilde{\xi}_{0,0}$. There exists $q \in (\alpha/2, \alpha \wedge 1)$ such that

$$\sum_{i=1}^{\infty} |\gamma_i|^q < \infty. \tag{40}$$

The sequences $\{\xi_{n,j}, j \geq 1\}$, $n \geq 1$, are i.i.d. copies of $\{\xi_{0,j}, j \geq 1\}$.

The main result of this section is an extension of Proposition 5 to the infinite-dimensional case. The long proof cannot be presented due to a page limit, but is available upon request.

Theorem 7. *Suppose Assumption 5 holds. Let $\mathbb{H}$ be spanned by $\{v_1, v_2, \ldots, \}$. Then*

$$N a_N^{-2} C_N \overset{d}{\to} C_\infty = \sum_{j=1}^{\infty} \Lambda_j (v_j \otimes v_j),$$

where the Λ_j are independent p-stable random variables with the characteristic function

$$E\left[\exp\left\{i\theta\Lambda_j\right\}\right] = \exp\left\{-s_p \sigma_j^p |\theta|^p \left(1 - i\,\mathrm{sign}\,(\theta) \tan\left(\frac{\pi p}{2}\right)\right)\right\} \tag{41}$$

with $\sigma_j^p = |\gamma_j|^\alpha / \sum_{i=1}^{\infty} |\gamma_i|^\alpha$.

The scaling parameters σ_j have interpretation as the values of the angular measure, see Remark 11.

6. Proofs of the results of Section 3

Proof of Theorem 2. Lemmas 3–4 lead to the asymptotic distribution of the operators C_N defined by (24). Lemma 5 shows that $\widehat{C}$ has the same asymptotic distribution.

Lemma 3. *Under the assumptions of Theorem 2, $Z = X \otimes X$ satisfies Condition 1 with $p = \alpha/2 \in (0, 1)$.*

Proof. By Assumption 1 and Proposition 3, $Z \in \mathrm{RV}_{-p}(\mathcal{S})$. Condition (10) thus follows from part (ii) of Proposition 1. To establish (11), recall that if $\left\{e_j, j \geq 1\right\}$ is an orthonormal basis in L^2, then $\left\{e_i \otimes e_j, i, j \geq 1\right\}$ is an orthonormal basis in $\mathcal{S}$. Identifying m with the pair (I, J), it is easy to check that $\|\pi_m(z)\|_{\mathcal{S}} = \|\pi_I(x)\| \|\pi_J(x)\|$. Consequently,

$$\frac{P(\|Z\|_{\mathcal{S}} > u)}{P(\|\pi_m(Z)\|_{\mathcal{S}} > tu)} = \frac{P(\|Z\|_{\mathcal{S}} > u)}{P(\|Z\|_{\mathcal{S}} > tu)} \frac{P(\|Z\|_{\mathcal{S}} > tu)}{P(\|\pi_m(Z)\|_{\mathcal{S}} > tu)} \to t^p \frac{c(1,1)}{c(I,J)}.$$

Lemma 4. *Suppose the assumptions of Theorem 2 hold and recall that the non-centered sample covariance operators C_N are defined by (24). Denote by a_N the quantile sequence of X defined by (9). Then*

$$Na_N^{-2}C_N \stackrel{d}{\to} C_\infty,$$

where C_∞ is specified in Theorem 2.

Proof. We want to apply Corollary 1 with $Z_n = X_n \otimes X_n$ and $\mathbb{H} = \mathcal{S}$. Lemma 3 shows that Assumptions 1 and 2 imply Condition 1 with $p = \alpha/2 \in (0, 1)$. Let $a_{N,Z}$ be the quantile sequence of $\|Z\| = \|X \otimes X\|$: $NP(\|Z\|_{\mathcal{S}} > a_{N,Z}) = 1$. Since $\|X \otimes X\|_{\mathcal{S}} = \|X\|^2$, we immediately get $a_{N,Z} = (a_N)^2$. By (19), $Na_N^{-2}C_N \stackrel{d}{\to} S^*_\infty$.

It remains to identify S^*_∞ in Corollary 1 with C_∞ in Theorem 2. By Remark 3.2. in Kokoszka et al. [17], if $Z = X \otimes X$, then the angular measure Γ_Z is concentrated on the diagonal

$$\mathbb{D}_{\mathcal{S}} = \left\{\Psi \in \mathbb{S}_{\mathcal{S}} : \psi = x \otimes x,\ x \in L^2\right\}$$

and $\Gamma_Z(B \otimes B) = \Gamma_X(B)$ for $B \in L^2$. Therefore, for $\phi \in \mathcal{S}$ and any function f such that $s \mapsto f(\langle \phi, s\rangle_{\mathcal{S}})$ is integrable over $\mathcal{S}$,

$$\int_{\mathbb{S}_{\mathcal{S}}} f(\langle \phi, s\rangle_{\mathcal{S}})\Gamma_Z(ds) = \int_{\mathbb{S}_{L^2}} f(\langle \phi, x \otimes x\rangle_{\mathcal{S}})\Gamma_X(dx).$$

This shows that the characteristic functionals of S^*_∞ and C_∞ are the same.

Lemma 5. *If Assumption 1 holds with $\alpha \in (0, 2)$, then*

$$Na_N^{-2}||C_N - \widehat{C}||_{\mathcal{S}} \stackrel{P}{\to} 0.$$

Proof. The result basically follows from the Marcińkiewicz–Zygmund law of large numbers (MZLLN), with some issues related to the fact that this law must be applied to random elements of a Hilbert space rather than to scalars. Let us recall the MZLLN: Suppose Y_n are i.i.d. (scalar) random variables and $r \in (0, 2)$. Then

$$N^{-1/r}\left[\sum_{n=1}^N Y_n - bN\right] \stackrel{a.s.}{\to} 0$$

if and only if $E|Y_1|^r < \infty$, where $b = 0$ if $r \in (0, 1)$ and $b = EY_1$ if $r \in [1, 2)$.

In our context, for any $x \in \mathbb{H}$, $C_N(x) - \widehat{C}(x) = \langle \bar{X}_N, x\rangle \bar{X}_N$, so

$$||C_N - \widehat{C}||_{\mathcal{S}} = ||\bar{X}_N||^2,$$

so we must show that $\sqrt{N}a_N^{-1}\bar{X}_N \stackrel{P}{\to} 0$. Note that

$$\sqrt{N}a_N^{-1}||\bar{X}_N|| \le N^{-1/2}a_N^{-1}\sum_{n=1}^N ||X_n||.$$

Consider $r \in (0, \min(\alpha, 1))$ to be specified later. Since $E||X_n||^r < \infty$, by the MZLLN, $N^{-1/r}\sum_{n=1}^N ||X_n|| \stackrel{a.s.}{\to} 0$, so we must ensure that $N^{-1/2}a_N^{-1}N^{1/r} \to 0$. By Remark 4, $a_N^{-1} < N^{-1/\alpha+\delta}$ for arbitrarily small $\delta >$ and sufficiently large N. Therefore, it is enough to ensure that

$$\frac{1}{r} < \frac{1}{2} + \frac{1}{\alpha}.$$

For $\alpha \in (0, 1)$, the above condition can be met by choosing r slightly smaller than α, for $\alpha \in [1, 2)$, slightly smaller than 1 because the RHS is greater than 1 due to $\alpha < 2$.

Proof of Theorem 3. Recall that a_N is the quantile sequence of $\|X\|$ and $a_{N,Z}$ is the quantile sequence of $\|Z\| = \|X \otimes X\|$. By Theorem 2, it is enough to show that

$$a_N^{-2}\gamma_N = Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\|_{\mathcal{S}} \le b_N\}}\right] \to \Psi_X(\alpha) \tag{42}$$

because $a_{N,Z} = a_N^2$. The convergence is not trivial because $E\left[\|Z\|\right] = \infty$, so a more subtle argument utilizing the regular variation of X (and so of Z) must be used. It suffices to establish the convergence in (42) in the weak topology of $\mathcal{S}$ because this is enough to ensure the convergence of characteristic functionals. If one can also prove tightness, this will imply the convergence in distribution. Since the limit is a constant in $\mathcal{S}$, the convergence is in probability, so the claim of Theorem 3 will follow.

We thus begin with establishing the convergence (42), up to a scaling constant that will be worked out later. Recall that if $(\mathbb{X}, \mathcal{A}, \mu)$ is a measure space and $\mathbb{B}$ is a separable Banach space, then a measurable function $f : \mathbb{X} \to \mathbb{B}$ is Bochner integrable if there is a sequence of simple functions $s_n : \mathbb{X} \to \mathbb{B}$ such that $\int_{\mathbb{X}} \|f - s_n\| d\mu \to 0$. The function f is Bochner integrable if and only of $\int_{\mathbb{X}} \|f\| d\mu < \infty$. We will apply the following general Lemma with $\mathbb{H} = \mathcal{S}$. In the proofs of this section $c = \mu(\{z : \|z\| > 1\})$.

Lemma 6. *If $Z \in \mathrm{RV}_{-p}(\mathbb{H})$, $p \in (0, 1)$, then*

$$M_N := N\tilde{a}_N^{-1} E\left[ZI_{\{\|Z\| \le \tilde{a}_N\}}\right] \to \int_{\mathbb{B}_1(\mathbb{H})} z\mu_Z(dz) =: M, \tag{43}$$

where $\mathbb{B}_r(\mathbb{H}) = \{z \in \mathbb{H} : \|z\| \le r\}$, $r > 0$, and $\tilde{a}_N$ is the normalizing sequence such that

$$\mu_N(\cdot) = NP(\tilde{a}_N^{-1} Z \in \cdot) \xrightarrow{M_0} \mu_Z(\cdot). \tag{44}$$

The convergence in (43) is in the weak topology of $\mathbb{H}$, i.e. for each $y \in \mathbb{H}$, $\langle y, M_N \rangle \to \langle y, M \rangle$.

Proof. Let Γ_Z be the angular measure of Z and write $\mathbb{B}_r = \mathbb{B}_r(\mathbb{H})$. We first verify that the limit in (43) exists in the sense of Bochner. This follows immediately from (8):

$$\int_{\mathbb{B}_1(\mathbb{H})} \|z\| \mu_Z(dz) = c \int_{\mathbb{S}_{\mathbb{H}}} \int_0^1 rpr^{-p-1} dr \; \Gamma(d\theta) = \frac{cp}{1-p} \Gamma_Z(\mathbb{S}_{\mathbb{H}}) = \frac{cp}{1-p},$$

since Γ_Z is the probability measure on $\mathbb{S}_{\mathbb{H}}$. The weak convergence (43) is equivalent to

$$\forall\, y \in \mathbb{H} \quad J_N := \int_{\mathbb{B}_1} \langle y, z \rangle \, \mu_N(dz) \to \int_{\mathbb{B}_1} \langle y, z \rangle \, \mu_Z(dz) =: J.$$

Fix $y \in \mathbb{H}$. Following Definition 2.5.4 in Meinguet [21], we use the superscript $^{(r)}$ to denote the restriction on a measure in M_0 to $\{x \in \mathbb{H} : \|x\| > r\}$. We approximate the integrals J_N and J, respectively, by

$$J_N^{(r)} = \int_{\mathbb{B}_1} \langle y, z \rangle \, \mu_N^{(r)}(dz), \quad J^{(r)} = \int_{\mathbb{B}_1} \langle y, z \rangle \, \mu_Z^{(r)}(dz)$$

and use the inequality

$$|J - J_N| \le |J - J^{(r)}| + |J^{(r)} - J_N^{(r)}| + |J_N^{(r)} - J_N|.$$

Fix $\epsilon > 0$. We first verify that for sufficiently small r, $|J - J^{(r)}| < \epsilon/3$. This follows from the bounds

$$|J - J^{(r)}| \le \|y\| \int_{\mathbb{B}_r} \|z\| \mu_Z(dz) = c\|y\| \Gamma_Z(\mathbb{S}) p \int_0^r u^{-p} du = cr^{1-p} \|y\| \frac{p}{1-p}.$$

Next, we establish a similar bound on $|J_N^{(r)} - J_N|$. Observe that $J_N^{(r)} - J_N = \left\langle y, N\tilde{a}_N^{-1} E\left[ZI_{\{\|z\| \le r\tilde{a}_N\}}\right]\right\rangle$. Hence $|J_N^{(r)} - J_N| \le \|y\| N\tilde{a}_N^{-1} E\left[\|Z\| I_{\{\|z\| \le r\tilde{a}_N\}}\right]$. Using relation (1.4.5a) in Kulik and Soulier [19] with $\beta = 1, \alpha = p, x = \tilde{a}_N, t = r$, we see that, as $N \to \infty$,

$$\begin{aligned} N\tilde{a}_N^{-1} E\left[\|Z\| I_{\{\|Z\| \le r\tilde{a}_N\}}\right] &\sim N\tilde{a}_N^{-1} \frac{p}{1-p} r^{1-p} \tilde{a}_N P(\|Z\| > a_N) \\ &= r^{1-p} \frac{p}{1-p} NP(\|Z\| > \tilde{a}_N). \end{aligned}$$

By part (iv) of Proposition 1, we therefore obtain

$$\limsup_{N\to\infty} |J_N^{(r)} - J_N| \le cr^{1-p} \|y\| \frac{p}{1-p}.$$

Now choose r so small that

$$\limsup_{N\to\infty} \left[|J^{(r)} - J| + |J_N^{(r)} - J_N| \right] < \frac{2\epsilon}{3}$$

and such that $\mu_N^{(r)}$ converges weakly to $\mu^{(r)}$, as $N \to \infty$. The latter requirement can be satisfied by Theorem 2.5.6 in Meinguet [21]. Since the integration in the definition of $J_N^{(r)}$ extends only over the unit ball, the uniform integrability condition is automatically met, and so $J_N^{(r)} \to J^{(r)}$. We conclude that for sufficiently large N, $|J - J_N| < \epsilon$.

In the next lemma, we specialize the limit in (43) to the case $Z = X \otimes X$. Note also that different scaling sequences are used.

Lemma 7. *Suppose the assumptions of Theorem 2 hold and $Z = X \otimes X$. Then*

$$Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\|_{\mathcal{S}} \le a_{N,Z}\}}\right] \to \frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}} (x \otimes x)\varGamma_X(dx), \tag{45}$$

$$Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\|_{\mathcal{S}} \le b_N\}}\right] \to \kappa^{\alpha/2-1}\frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}} (x \otimes x)\varGamma_X(dx), \tag{46}$$

in the weak topology of the Hilbert space $\mathcal{S}$, where κ is given in (22).

Proof. Recall that the Hilbert space here is $\mathbb{H} = \mathcal{S}$. Set $\mathbb{B}_r = \mathbb{B}_r(\mathcal{S})$, $\|\cdot\| = \|\cdot\|_{\mathcal{S}}$. Recall also that a_N and $a_{N,Z}$ are the quantile sequences of X and $Z = X \otimes X$, respectively. By Remark 1,

$$NP(a_{n,Z}^{-1}Z \in \cdot) \xrightarrow{M_0} \mu_Z(\cdot)$$

and $c = \mu_Z(z : \|z\| > 1) = 1$. Hence, (43) reads, by (8),

$$Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\| \le a_{N,Z}\}}\right] \to \int_{\mathbb{B}_1} z\mu_Z(dz) = c\int_0^1\int_{\mathbb{S}_{\mathcal{S}}} r\theta\; pr^{-p-1}dr\,\varGamma_Z(d\theta) = c\frac{p}{1-p}\int_{\mathbb{S}_{\mathcal{S}}}\theta\,\varGamma_Z(d\theta).$$

As noted in the proof of Lemma 4, if $Z = X \otimes X$, then $\varGamma_Z$ is concentrated on the diagonal $\mathbb{D}_{\mathcal{S}}$ and $\varGamma_Z(B \otimes B) = \varGamma_X(B)$ for $B \in L^2$. Therefore, $\varGamma_Z(\mathbb{S}_{\mathcal{S}}) = \varGamma_X(\mathbb{S}_{L^2})$ and $\int_{\mathbb{S}_{\mathcal{S}}}\theta\,\varGamma_Z(d\theta) = \int_{\mathbb{S}_{L^2}}(x \otimes x)\varGamma_X(dx)$, showing that M in (43) can be expressed as

$$M = \frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}} (x \otimes x)\varGamma_X(dx).$$

At the same time, with $\tilde{a}_N = b_N$, (43) and the above computation gives

$$Nb_N^{-1}E\left[ZI_{\{\|Z\| \le b_N\}}\right] \to \int_{\mathbb{B}_1} z\mu_Z'(dz) = c'\frac{p}{1-p}\int_{\mathbb{S}_{\mathcal{S}}}\theta\,\varGamma_Z(d\theta) = c'\frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}}(x \otimes x)\varGamma_X(dx),$$

where

$$NP(b_N^{-1}Z \in \cdot) \xrightarrow{M_0} \mu_Z'(\cdot)$$

and $c' = \mu_Z'(z : \|z\| > 1)$. From the proof of Corollary 1 we know that $a_{N,Z} = \kappa b_N$, hence $c' = \kappa^p$. Indeed:

$$1 \sim NP(\|Z\| > a_{N,Z}) = NP(\|Z\| > \kappa b_N) \sim \kappa^{-p}NP(\|Z\| > b_N) \sim \kappa^{-p}c'.$$

Thus,

$$Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\| \le b_N\}}\right] = Nb_N^{-1}E\left[ZI_{\{\|Z\| \le b_N\}}\right]\frac{b_N}{a_{N,Z}} \to \frac{c'}{\kappa}\frac{\alpha}{2-\alpha}\int_{\mathbb{S}_{L^2}}(x \otimes x)\varGamma_X(dx),$$

Lemma 8. *Under the assumptions of Theorem 2, the sequence of Dirac measures at the points $z_N := Na_{N,Z}^{-1}E\left[ZI_{\{\|Z\|_{\mathcal{S}} \le b_N\}}\right]$ is tight.*

Proof. Set again $\|\cdot\| = \|\cdot\|_{\mathcal{S}}$. A sequence of Dirac measures in a metric space is tight if and only if the points of the sequence form a relatively compact subset. In a separable Hilbert space, a subset is relatively compact if and only if it is bounded and has equi-small tails with respect to any (one) orthonormal system $\{\eta_k\}$, i.e.

$$\forall\, \epsilon > 0\;\; \exists\, m \ge 1\;\; \forall\, N \ge 1\;\; \sum_{k \ge m}\langle z_N, \eta_k\rangle^2 < \epsilon. \tag{47}$$

Since $a_{N,Z} = \kappa b_N$, it follows from relation (1.4.5a) in Kulik and Soulier [19] that

$$\|z_N\| \le Na_{N,Z}^{-1}E\left[\|Z\|I_{\{\|Z\| \le b_N\}}\right] \sim \kappa^p\frac{p}{1-p},$$

so the sequence z_N is bounded in $\mathcal{S}$. To verify (47), notice that it can be written as

$$\lim_{m\to\infty}\limsup_{N\to\infty}\|\pi_m(z_N)\| = 0$$

and observe that

$$\pi_m(z_N) = Na_{N,Z}^{-1}E\left[\pi_m(Z)I_{\{\|Z\| \le b_N\}}\right].$$

Therefore, identifying m in (11) with the pair (I, J) in Assumption 2, we get by relation (1.4.5a) in Kulik and Soulier [19],

$$\begin{aligned}\|\pi_m(z_N)\| &\le N a_{N,Z}^{-1} E\left[\|\pi_m(Z)\| I_{\{\|Z\|\le b_N\}}\right]\\ &\le N a_{N,Z}^{-1} E\left[\|\pi_m(Z)\| I_{\{\|\pi_m(Z)\|\le b_N\}}\right]\\ &\sim \frac{p}{1-p} N a_{N,Z}^{-1} b_N P\left(\|\pi_m(Z)\| > b_N\right)\\ &\sim \frac{p}{1-p}\kappa^{-1} NP\left(\|\pi_m(Z)\| > a_{N,Z}\right)\kappa^{p}.\end{aligned}$$

Next, we use the factorization

$$NP\left(\|\pi_m(Z)\| > a_{N,Z}\right) = \frac{P\left(\|\pi_m(Z)\| > a_{N,Z}\right)}{P(\|Z\| > a_{N,Z})} NP(\|Z\| > a_{N,Z}).$$

By (11),

$$\limsup_{N\to\infty} \|\pi_m(z_N)\| \le \frac{c_m}{c_1}\frac{p}{1-p}\kappa^{p-1}.$$

The claim thus follows because $c_m \to 0$ by Assumption 2.

Proof of Theorem 5. Set $c_j = \mathrm{sign}(\langle \hat{v}_j, V_j\rangle)$ and

$$\alpha_j = \inf_{k\ne j} |\Lambda_j - \Lambda_k|.$$

By Assumption (31), $\alpha_j > 0$ a.s., so the claim will follow once we have shown that (for the versions in Lemma 1)

$$||\hat{v}_j - c_j V_j|| \le \frac{2\sqrt{2}}{\alpha_j} ||r_N \widehat{C} - C_\infty||_{\mathcal{S}}. \tag{48}$$

(Recall that by Lemma 1 $\|r_N\widehat{C} - C_\infty\|_{\mathcal{S}} \overset{a.s.}{\to} 0$.) Set

$$D_j = ||C_\infty(\hat{v}_j) - \Lambda_j \hat{v}_j||, \quad S_j = \sum_{k\ne j} \langle \hat{v}_j, V_k\rangle^2.$$

To prove (48), it is enough to show that

$$||\hat{v}_j - c_j V_j||^2 \le 2S_j; \tag{49}$$

$$\alpha_j^2 S_j \le D_j^2; \tag{50}$$

$$D_j \le 2||r_N\widehat{C} - C_\infty||_{\mathcal{S}}. \tag{51}$$

We begin with the verification of (49). Since the V_k form an orthonormal basis in L^2 almost surely,

$$||\hat{v}_j - c_j V_j||^2 = \sum_{k=1}^{\infty} \left(\langle \hat{v}_j, V_k\rangle - c_j\langle V_j, V_k\rangle\right)^2 = \left(\langle \hat{v}_j, V_j\rangle - c_j\right)^2 + S_j. \tag{52}$$

Since $c_j^2 = 1$ and $c_j\langle \hat{v}_j, V_j\rangle = |\langle \hat{v}_j, V_j\rangle|$, $\left(\langle \hat{v}_j, V_j\rangle - c_j\right)^2 = (1 - |\langle \hat{v}_j, V_j\rangle|)^2$. Using

$$(1 - |\langle \hat{v}_j, V_j\rangle|)^2 = \sum_{k=1}^{\infty} \langle \hat{v}_j, V_k\rangle^2 - 2|\langle \hat{v}_j, V_j\rangle| + \langle \hat{v}_j, V_j\rangle^2,$$

we obtain

$$\left(\langle \hat{v}_j, V_j\rangle - c_j\right)^2 = S_j + 2\left(\langle \hat{v}_j, V_j\rangle^2 - |\langle \hat{v}_j, V_j\rangle|\right) \le S_j.$$

Combining the last bound with (52), we obtain (49).

We now turn to (50). Since the V_k form a basis in L^2 almost surely and since C_∞ is symmetric almost surely,

$$D_j^2 = \sum_{k=1}^{\infty}\left(\langle C_\infty(\hat{v}_j), V_k\rangle - \Lambda_j\langle \hat{v}_j, V_k\rangle\right)^2 = \sum_{k=1}^{\infty}\left(\langle \hat{v}_j, \Lambda_k V_k\rangle - \Lambda_j\langle \hat{v}_j, V_k\rangle\right)^2 = \sum_{k\ne j}(\Lambda_k - \Lambda_j)^2\langle \hat{v}_j, V_k\rangle^2 \ge \alpha_j^2 S_j.$$

It remains to verify (51). This follows from the decomposition

$$C_\infty(\hat{v}_j) - \Lambda_j\hat{v}_j = \left(C_\infty - r_N\widehat{C}\right)(\hat{v}_j) + \left(r_N\widehat{C} - \Lambda_j\right)(\hat{v}_j)$$

Note that $\left(r_n\widehat{C} - \Lambda_j\right)(\hat{v}_j) = \left(r_N\hat{\lambda}_j - \Lambda_j\right)(\hat{v}_j)$ and by (30), $|r_N\hat{\lambda}_j - \Lambda_j| \le ||r_N\widehat{C} - C_\infty||_{\mathcal{S}}$.

7. Proofs of the results of Section 4

Proof of Lemma 2. Since $\mathbb{H} = L^2$ is spanned by $\{v_1, \ldots, v_d\}$, any $A \subseteq \mathbb{S}_{L^2}$ has those the form $A = A_B = \{v = \sum_{j=1}^d \beta_j v_j : \|v\| = 1, (\beta_1, \ldots, \beta_d) \in B\}$ for some Borel set B of $\mathbb{R}^d$. Since $\|v\|^2 = \sum_{j=1}^d \beta_j^2$, $B \subseteq \mathbb{S}_{\mathbb{R}^d}$, where the latter unit sphere stems from the Euclidean norm $\|\cdot\|_d$. We note that $\sum_{j=1}^d \xi_j v_j \in A_B$ if and only if $(\xi_1, \ldots, \xi_d)/\|(\xi_1, \ldots, \xi_d)\|_d \in B$. Since $\|v_j\| = 1$, we have $\|X\|^2 = \sum_{j=1}^d \xi_j^2 = \|(\xi_1, \ldots, \xi_d)\|_d^2$. Hence, the assumed regular variation of $\boldsymbol{\xi} := (\xi_1, \ldots, \xi_d)$ gives

$$\frac{P(\|X\| > tu, X/\|X\| \in A_B)}{P(\|X\| > u)} = \frac{P(\|\boldsymbol{\xi}\| > tu, \boldsymbol{\xi}/\|\boldsymbol{\xi}\| \in B)}{P(\|\boldsymbol{\xi}\| > u)} \to t^{-\alpha} \Gamma_{\boldsymbol{\xi}}(B).$$

The regular variation of X follows by identifying $\Gamma_X(A_B) = \Gamma_{\boldsymbol{\xi}}(B)$.

7.1. Stable convergence of random vectors

The results of this section are needed to prove Theorem 6 and other results of Section 4.

Recall that a random variable Λ is p-stable, $p \in (0, 2) \setminus \{1\}$, if its characteristic function is given by

$$E\left[\exp\{i\theta\Lambda\}\right] = \exp\left\{ i\beta\theta - \tilde{\sigma}^p |\theta|^p \left(1 - i\gamma \operatorname{sign}(\theta) \tan\left(\frac{\pi p}{2}\right)\right)\right\}, \quad \theta \in \mathbb{R}, \tag{53}$$

with $\beta \in \mathbb{R}$, $\tilde{\sigma} > 0$, $\gamma \in [-1, 1]$. We write $\Lambda \sim S_p(\tilde{\sigma}, \gamma, \beta)$. A random vector $\boldsymbol{\Lambda} = (\Lambda_1, \ldots, \Lambda_q)$ is p-stable ($p \neq 1$) if

$$E \exp\{i\langle\theta, \boldsymbol{\Lambda}\rangle\} = \exp\left\{ i\langle\beta, \theta\rangle - \int_{\mathbb{S}_{\mathbb{R}^q}} |\langle\theta, s\rangle|^p \left(1 - i \operatorname{sign}(\langle\theta, s\rangle) \tan\left(\frac{\pi p}{2}\right)\right) \sigma_{\boldsymbol{\Lambda}}(ds) \right\},$$

where $\theta \in \mathbb{R}^q$, $\beta \in \mathbb{R}^q$, and $\sigma_{\boldsymbol{\Lambda}}$ is a finite measure on the unit sphere in $\mathbb{R}^q$. (Note that we use $\langle\cdot,\cdot\rangle$ for both inner product in L^2 and $\mathbb{R}^q$). We write $\boldsymbol{\Lambda} \sim S_p(\sigma_{\boldsymbol{\Lambda}}, \beta)$, see Samorodnitsky and Taqqu [27], Definition 2.3.2. In the scalar case, the spectral measure of $\boldsymbol{\Lambda} = \Lambda$ is concentrated on $\{-1, 1\}$ and (53) is obtained by setting $\tilde{\sigma} = (\sigma_\Lambda(\{1\}) + \sigma_\Lambda(\{-1\}))^{1/p}$, while $\beta = (\sigma_\Lambda(\{1\}) - \sigma_\Lambda(\{-1\}))/(\sigma_\Lambda(\{1\}) + \sigma_\Lambda(\{-1\}))$.

For our considerations, we need a well-known lemma.

Lemma 9. *Assume that ξ_n, $n \geq 1$, are independent, identically distributed and regularly varying random variables with index $\alpha \in (0, 2)$ and let c_N be defined by $NP(|\xi_1| > c_N) = 1$. Then*

$$c_N^{-2} \sum_{n=1}^N \xi_n^2 \xrightarrow{d} \sum_{n=1}^\infty P_n \sim S_p(s_p^{1/p}, 1, 0), \quad p = \frac{\alpha}{2},$$

where $\{P_j, j \geq 1\}$ is the sequence of points of a Poisson process with the mean measure ν_p, $\nu_p(dx) = px^{-p-1}dxI_{\{x>0\}}$ and s_p is given by (21).

Lemma 9 is basically a special case of Theorem 8.3.1 in Kulik and Soulier [19]. Since $p = \alpha/2 \in (0, 1)$, the series $\sum_{n=1}^\infty P_n$ is summable and defines a p-stable random variable $S_p(s_p^{1/p}, 1, 0)$, see Theorem 1.4.5 in Samorodnitsky and Taqqu [27] and Section 8.2.1 in Kulik and Soulier [19].

Let now $|\cdot|$ be a norm on $\mathbb{R}^d$. Let $\tilde{d}_N$ be defined by $NP(|\boldsymbol{\xi}| > \tilde{d}_N) = 1$. By Assumption 3, $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_d)$ is regularly varying, and hence Proposition 1(ii) gives

$$NP\left(|\boldsymbol{\xi}| > t\tilde{d}_N, \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \in \cdot\right) \xrightarrow{w} t^{-\alpha}\Gamma_{\boldsymbol{\xi}}(\,\cdot\,), \quad \text{as } N \to \infty, \tag{54}$$

where $\Gamma_{\boldsymbol{\xi}}$ is a probability measure on the $|\cdot|$-unit sphere. Let $\mu_{\boldsymbol{\xi}}$ be the associated exponent measure as in (7). That is

$$NP\left(\tilde{d}_N^{-1}\boldsymbol{\xi} \in \cdot\,\right) \xrightarrow{M_0} \mu_{\boldsymbol{\xi}}(\cdot).$$

The choice of $\tilde{d}_N$ gives $\mu_{\boldsymbol{\xi}}(\{x \in \mathbb{R}^d : |x| > 1\}) = 1$.

For $g : \mathbb{R}^d \to \mathbb{R}^k$, $k = d^2$, defined by $g(x_1, \ldots, x_d) = (x_i x_j, i, j \in \{1, \ldots, d\})$, set $Y = g(\xi_1, \ldots, \xi_d)$. According to Proposition 2.1.12 in Kulik and Soulier (2020), if the measure $\mu_Y := \mu_{\boldsymbol{\xi}} \circ g^{-1}$ is not identically equal to zero on $\mathbb{R}^k$, then Y is regularly varying with index $p = \alpha/2$ and the exponent measure $\mu_Y = \mu_{\boldsymbol{\xi}} \circ g^{-1}$. The corresponding scaling sequence is $\tilde{d}_N^2$:

$$NP\left(\tilde{d}_N^{-2} Y \in \cdot\,\right) \xrightarrow{M_0} \mu_Y(\cdot).$$

With a norm $|\cdot|$ on $\mathbb{R}^k$, we have

$$NP\left(|Y| > \tilde{d}_N^2\right) \to \mu_Y(\{u \in \mathbb{R}^k : |u| > 1\}) = \mu_{\boldsymbol{\xi}}(\{x \in \mathbb{R}^d : |g(x)| > 1\}).$$

Note that the latter expression may be different from 1, unless $g:\mathbb{R}\to\mathbb{R}_+$, $g(x)=x^2$. Set $c_p=(\mu_\xi(\{x\in\mathbb{R}^d:|g(x)|>1\}))^{1/(2p)}$ and $d_N=\tilde{d}_N c_p$. With this choice

$$NP\left(|Y|>d_N^2\right)\to 1 \tag{55}$$

and

$$NP\left(d_N^{-2}Y\in\ \cdot\ \right)\xrightarrow{M_0} c_p^{-2p}\mu_Y(\cdot)=\frac{\mu_Y(\cdot)}{\mu_Y(\{u\in\mathbb{R}^k:|u|>1\})}. \tag{56}$$

(Note that in the scalar case of $d=1$ we have $\tilde{d}_N=d_N$.) With the aforementioned norm $|\cdot|$ on $\mathbb{R}^k$, consider the polar coordinate transformation $T:\mathbb{R}^k\to\mathbb{R}_+\times\mathbb{S}^{k-1}$: $T(y)=(|y|,y/|y|)$. Using Eq. (2.2.4) in Kulik and Soulier (2020) we have $c_p^{-2p}\mu_Y\circ T^{-1}=\nu_p\otimes\Gamma_Y$, where ν_p is defined by $\nu_p(dx)=px^{-p-1}dxI_{\{x>0\}}$ and Γ_Y is the angular measure of $Y=g(\xi_1,\ldots,\xi_d)$. With this background, we can formulate the following Lemma.

Lemma 10. *Assume that $\boldsymbol{\xi}=(\xi_1,\ldots,\xi_d)$ is regularly varying with index $\alpha\in(0,2)$. Set $Y=g(\xi_1,\ldots,\xi_d)=(\xi_i\xi_j,\ i,j\in\{1,\ldots,d\})$ and choose d_N as in (55). Assume that $(\xi_{n,1},\ldots,\xi_{n,d})$, $n\in\{1,\ldots,N\}$, are independent, identically distributed copies of $\boldsymbol{\xi}$. Set $p=\alpha/2$. Then*

$$\left(d_N^{-2}\sum_{n=1}^{N}\xi_{n,i}\xi_{n,j},\ i,j=1,\ldots,d\right)\xrightarrow{d}\sum_{n=1}^{\infty}P_nW_n,$$

where $\{W_j,j\geq 1\}$ is a sequence of i.i.d. random vectors on $\mathbb{S}^{k-1}$ with the distribution Γ_Y, independent from the sequence $\{P_j,j\geq 1\}$ of the points of a Poisson process with the mean measure ν_p.

The $k=d^2$-dimensional limiting random vector $\widetilde{\boldsymbol{\Lambda}}=\sum_{n=1}^{\infty}P_nW_n$ is p-stable with the characteristic function

$$E\left[\exp\left\{i\left\langle\theta,\widetilde{\boldsymbol{\Lambda}}\right\rangle\right\}\right]=\exp\left\{-s_p\int_{\mathbb{S}_{\mathbb{R}^k}}|\langle\theta,s\rangle|^p\left(1-i\operatorname{sign}(\langle\theta,s\rangle)\tan\left(\frac{\pi p}{2}\right)\right)\Gamma_Y(ds)\right\}. \tag{57}$$

Proof. The proof is relatively classical, we sketch it for completeness without dealing with some particular technical details. Let Y_n, $n\in\{1,\ldots,N\}$ be independent copies of Y. Define

$$M_N=\sum_{n=1}^{N}\delta_{d_N^{-2}Y_n}$$

Then (56) and the classical result on Poisson convergence (see Theorem 5.3 in [26]) give $M_N\Rightarrow M$, where M is Poisson random measure with mean measure $c_p^{-2p}\mu_Y$ and $\Rightarrow$ denotes weak convergence in the space of Radon measures.

Since $c_p^{-2p}\mu_Y\circ T^{-1}=\nu_p\otimes\Gamma_Y$, Example 7.1.15 in Kulik and Soulier (2020) gives the following representation of M:

$$M:=\sum_{n=1}^{\infty}\delta_{J_n^{-1/p}W_n},$$

where $\{W_j,j\geq 1\}$ is a sequence of i.i.d. random vectors on $\mathbb{S}^{k-1}$ with distribution Γ_Y, independent from the sequence $\{J_j,j\geq 1\}$ of points of a unit rate homogeneous Poisson process. Equivalently,

$$M=\sum_{n=1}^{\infty}\delta_{P_nW_n},$$

where $\{P_j,j\geq 1\}$ are the points of a Poisson process with the mean measure ν_p.

Fix $\epsilon>0$. The summation functional

$$\sum_j\delta_{x_j}\to\sum_j x_j,\quad x_j\in\mathbb{R}^k,$$

is continuous on $\{x\in\mathbb{R}^k:|x|>\epsilon\}$; see page 215 in Resnick [26]. This implies

$$V_{n,\epsilon}:=d_N^{-2}\sum_{n=1}^{N}Y_nI_{\{|Y_n|>d_N^2\epsilon\}}\xrightarrow{d}\sum_{n=1}^{\infty}P_nW_nI_{\{|P_nW_n|>\epsilon\}}.$$

Since $p\in(0,1)$, the series $\sum_{n=1}^{\infty}P_n$ is summable (see Section 8.2.1 in Kulik and Soulier [19]). Moreover, since $|W_n|=1$, we can let $\epsilon\to 0$ on the right hand side of the above expression to get the limit in the form $\sum_{n=1}^{\infty}P_nW_n$.

To conclude, it suffices to show that $d_N^{-2}\sum_{n=1}^{N}Y_nI_{\{|Y_n|\leq d_N^2\epsilon\}}$ converges to zero in probability as first $n\to\infty$ and then $\epsilon\to 0$. With the norm $|\cdot|$ on $\mathbb{R}^k$ we have (see Proposition 1.4.6 in Kulik and Soulier [19])

$$\lim_{N\to\infty}Nd_N^{-2}E\left(|Y_n|I_{\{|Y_n|\leq d_N^2\epsilon\}}\right)=\epsilon^{1-p}\frac{p}{1-p}.$$

In conclusion,

$$d_N^{-2}\sum_{n=1}^{N} Y_n \xrightarrow{d} \sum_{n=1}^{\infty} P_n W_n.$$

Proof of Theorem 6. We use the finite-dimensional asymptotics from Lemma 10 along with the continuous mapping. As in Lemma 10, let d_N be defined by $NP\left(|Y| > d_N^2\right) = 1$, $Y = g(\xi_1,\dots,\xi_d)$, $g(x_1,\dots,x_d) = (x_i x_j, i,j \in \{1,\dots,d\})$. By Lemma 10 and the continuous mapping theorem,

$$Nd_N^{-2}C_N = \sum_{i,j=1}^{d}\left(d_N^{-2}\sum_{n=1}^{N}\xi_{n,i}\xi_{n,j}\right)(v_i\otimes v_j) \xrightarrow{d} \sum_{i,j=1}^{d}\Lambda_{i,j}(v_i\otimes v_j), \tag{58}$$

where a random vector $\tilde{\mathbf{\Lambda}} = (\Lambda_{i,j}, i,j\in\{1,\dots,d\})$ in $\mathbb{R}^k$ has the characteristic function given in (57). Let $G : \mathbb{S}_{\mathbb{R}^d} \to \mathbb{S}_{\mathbb{R}^k}$ be defined by

$$G(x_1,\dots,x_d) = \frac{(x_i x_j, i,j\in\{1,\dots,d\})}{|(x_i x_j, i,j\in\{1,\dots,d\})|},$$

with the norm $|\cdot|$ on $\mathbb{R}^k$. Then the angular measures of $\boldsymbol{\xi} = (\xi_1,\dots,\xi_d)$ and Y fulfill $\Gamma_{\boldsymbol{\xi}}\circ G^{-1} = \Gamma_Y$ and

$$\int_{\mathbb{S}_{\mathbb{R}^k}} |\langle\theta,s\rangle|^p\,\Gamma_Y(ds) = \int_{\mathbb{S}_{\mathbb{R}^d}} |\langle\theta,G(x)\rangle|^p\,\Gamma_{\boldsymbol{\xi}}(dx) = \int_{\mathbb{S}_{\mathbb{R}^d}} \left|\sum_{i,j=1}^{d}\theta_{i,j}x_i x_j\right|^p \Gamma_{\boldsymbol{\xi}}(dx), \tag{59}$$

where we used a representation $\theta = (\theta_{i,j}, i,j\in\{1,\dots,d\})$. The corresponding expression holds for the second integral in the characteristic functional in (57). Therefore, the characteristic function of $\tilde{\mathbf{\Lambda}}$ agrees with the one in (33).

Proof of Proposition 5. The angular measure Γ_ξ is concentrated on the intersection of the unit ball and the axes, the only possibility for the product $x_i x_j$ to be nonzero, is if $i=j$. Consequently, the joint characteristic function in (33) becomes

$$\exp\left\{-s_p\int_{\mathbb{S}_{\mathbb{R}^d}}\left|\sum_{j=1}^{d}\theta_{j,j}x_j^2\right|^p\left(1-i\operatorname{sign}(\langle\theta,s\rangle)\tan\left(\frac{\pi p}{2}\right)\right)\Gamma_\xi(dx)\right\}.$$

Since the angular measure has non-zero mass on e_j's only, the characteristic function reduces to

$$\prod_{j=1}^{d}\exp\left\{-s_p\sigma_j^p|\theta_{j,j}|^p\left(1-i\operatorname{sign}\left(\theta_{j,j}\right)\tan\left(\frac{\pi p}{2}\right)\right)\right\}$$

with $\sigma_j^p = \Gamma_\xi(e_j)$.

Now, we need to find the relationship between the scaling sequence d_N and the quantile sequence a_N: $NP(\|X\| > a_N) = 1$. We have

$$NP(\|X\| > a_N) = NP\left(\sum_{j=1}^{d}\xi_j^2 > a_N^2\right) = 1.$$

On the other hand, choose the norm on $\mathbb{R}^k$ ($k=d^2$) to be the Euclidean norm. Hence,

$$NP(|Y| > d_N^2) = NP\left(\sum_{i,j=1}^{d}\xi_j\xi_j > d_N^2\right) = 1.$$

Since the angular measure Γ_ξ is concentrated on e_j's, we have

$$\lim_{N\to\infty} NP(|\xi_i\xi_j| > d_N^2) = 0\,,\quad i\neq j;$$

see the discussion leading to Corollary 2.1.20 in Kulik and Soulier [19]. Now, if V is a regularly varying random variable, and U is such that $P(|U|>u)/P(|V|>u)\to 0$ as $u\to\infty$, then $P(U+V>u)/P(V>u)\to 1$ as $u\to\infty$. Therefore, the sequences a_N and d_N are asymptotically equivalent: $\lim_{N\to\infty} a_N/d_N = 1$.

Hence,

$$Na_N^{-2}C_N \xrightarrow{d} \sum_{j=1}^{d}\Lambda_j(v_j\otimes v_j), \tag{60}$$

where $\Lambda_j = \Lambda_{j,j}$ are independent p-stable random variables as in the statement of the theorem.

Remark 16. We illustrate that statement of Proposition 5 is in agreement with Theorem 2 even though the latter is not applicable in the current, finite dimensional situation. Recall the characteristic integrals in Theorem 2:

$$\int_{\mathbb{S}_{L^2}} |\langle \phi, x\otimes x\rangle_{\mathcal{S}}|^p \Gamma_X(dx)$$

and

$$\int_{\mathbb{S}_{L^2}} \langle \phi, x\otimes x\rangle_{\mathcal{S}} \, |\langle \phi, x\otimes x\rangle_{\mathcal{S}}|^{p-1} \Gamma_X(dx), \tag{61}$$

where Γ_X is the angular measures of X. Consider the map $H : \mathbb{S}_{\mathbb{R}^d} \to \mathbb{S}_{L^2}$ defined by $H(s_1,\ldots,s_d) = \sum_{j=1}^d s_j v_j$. Then $\Gamma_\xi \circ H^{-1} = \Gamma_X$ and the first characteristic integral becomes

$$\begin{aligned}\int_{\mathbb{S}_{L^2}} |\langle \phi, x\otimes x\rangle_{\mathcal{S}}|^p \Gamma_X(dx) &= \int_{\mathbb{S}_{\mathbb{R}^d}} \left|\left\langle \phi, \left[\sum_{i=1}^d s_i v_i\right] \otimes \left[\sum_{j=1}^d s_j v_j\right]\right\rangle_{\mathcal{S}}\right|^p \Gamma_\xi(ds)\\ &= \int_{\mathbb{S}_{\mathbb{R}^d}} \left|\sum_{i,j=1}^d s_i s_j \left\langle \phi, v_i\otimes v_j\right\rangle_{\mathcal{S}}\right|^p \Gamma_\xi(ds)\\ &= \int_{\mathbb{S}_{\mathbb{R}^d}} \left|\sum_{i,j=1}^d s_i s_j \left\langle \phi(v_i), v_j\right\rangle\right|^p \Gamma_\xi(ds)\,. \end{aligned}\tag{62}$$

The corresponding relationship holds for the second characteristic integral.

We apply Corollary 1. Using the property of the angular measure, we continue with (62) to get

$$\int_{\mathbb{S}_{L^2}} |\langle \phi, x\otimes x\rangle_{\mathcal{S}}|^p \Gamma_X(dx) = \int_{\mathbb{S}_{\mathbb{R}^d}} |\sum_{i,j=1}^d s_i s_j \left\langle \phi(v_i), v_j\right\rangle|^p \Gamma_\xi(ds) = \int_{\mathbb{S}_{\mathbb{R}^d}} |\sum_{j=1}^d s_j^2 \left\langle \phi(v_j), v_j\right\rangle|^p \Gamma_\xi(ds)\,.$$

with the corresponding expression for the second characteristic integral. Thus, the characteristic exponent becomes

$$-s_p \int_{\mathbb{S}_{\mathbb{R}^d}} |\sum_{j=1}^d s_j^2 \left\langle \phi(v_j), v_j\right\rangle|^p \Gamma_\xi(ds) \tag{63}$$

$$+ s_p i \tan\frac{\pi p}{2} \int_{\mathbb{S}_{\mathbb{R}^d}} \left(\sum_{j=1}^d s_j^2 \left\langle \phi(v_j), v_j\right\rangle\right) |\sum_{j=1}^d s_j^2 \left\langle \phi(v_j), v_j\right\rangle|^{p-1} \Gamma_\xi(ds). \tag{64}$$

Let now Λ_j be independent p-stable random variables as in the statement of Proposition. Set $\boldsymbol{\Lambda} = (\Lambda_1,\ldots,\Lambda_d)$. Let $A_\infty = \sum_{j=1}^d \Lambda_j (v_j\otimes v_j)$. Then

$$Ee^{i\langle \phi, A_\infty\rangle_{\mathcal{S}}} = Ee^{i\sum_{j=1}^d \Lambda_j \langle \phi, v_j\otimes v_j\rangle_{\mathcal{S}}} =: Ee^{i\sum_{j=1}^d \theta_j \Lambda_j}$$

with $\theta_j = \left\langle \phi, v_j\otimes v_j\right\rangle_{\mathcal{S}} = \left\langle \phi(v_j), v_j\right\rangle$. Using (36) it is immediate to see that the characteristic function of A_∞ coincides with the one induced by (63)–(64).

Acknowledgments

We thank Dr. Mihyun Kim for performing the numerical analysis reported in the Online Material. Piotr Kokoszka was partially supported by the United States NSF grants DMS–1914882 and DMS– 2123761. Rafał Kulik was supported by Canada NSERC grant.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2022.105123.

References

[1] G. Aneiros, R. Cao, R. Fraiman, C. Genest, P. Vieu, Recent advances in functional data analysis and high-dimensional statistics, J. Multivariate Anal. 170 (2019) 3–9.
[2] G. Aneiros, I. Horová, M. Husková, P. Vieu, On functional data analysis and related topics, J. Multivariate Anal. 189 (2022).
[3] B. Basrak, J. Segers, Regularly varying multivariate time series, Stochastic Process. Appl. 119 (2009) 1055–1080.

[4] D. Bosq, Linear Processes in Function Spaces, Springer, 2000.
[5] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for principal component analysis of a vector random function, J. Multivariate Anal. 12 (1982) 136–154.
[6] R. Davis, T. Hsing, Point process and partial sum convergence for weakly dependent random variables with infinite variance, Ann. Probab. 23 (1995) 879–917.
[7] R.A. Davis, S.I. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, Ann. Probab. 13 (1985) 179–195.
[8] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, Springer, Berlin, 1997.
[9] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: Theory and Practice, Springer, 2006.
[10] B.V. Gnedenko, A.N. Kolmogorov, Limit Distributions for Sums of Independent Random Variables, Addison-Wesley, 1954.
[11] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics, J. Multivariate Anal. 146 (2016) 1–6.
[12] L. Horváth, P. Kokoszka, Inference for Functional Data with Applications, Springer, 2012.
[13] T. Hsing, R. Eubank, Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators, Wiley, 2015.
[14] H. Hult, F. Lindskog, Regular variation for measures on metric spaces, Publ. L'Inst. Math. Nouvelle Série 80 (94) (2006) 121–140.
[15] P. Kokoszka, M. Reimherr, Asymptotic normality of the principal components of functional time series, Stochastic Process. Appl. 123 (2013) 1546–1562.
[16] P. Kokoszka, M. Reimherr, Introduction to Functional Data Analysis, CRC Press, 2017.
[17] P. Kokoszka, S. Stoev, Q. Xiong, Principal components analysis of regularly varying functions, Bernoulli 25 (2019) 3864–3882.
[18] J. Kuelbs, V. Mandrekar, Domains of attraction of stable measures on a Hilbert space, Studia Math. 50 (1974) 149–162.
[19] R. Kulik, P. Soulier, Heavy-Tailed Time Series, Springer, 2020.
[20] F. Lindskog, S. Resnick, J. Roy, Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps, Probab. Surv. 11 (2014) 270–314.
[21] T. Meinguet, Heavy Tailed Functional Time Series (Ph.D. thesis), Universite catholique de Louvain, 2010.
[22] T. Meinguet, J. Segers, Regularly Varying Time Series in Banach Spaces, Technical Report, UCLouvain, 2010, arXiv:1001.3262.
[23] D. Pollard, Convergence of Stochastic Processes, Springer, 1984.
[24] J. Ramsay, G. Hooker, S. Graves, Functional Data Analysis with R and MATLAB, Springer, 2009.
[25] J.O. Ramsay, B.W. Silverman, Functional Data Analysis, Springer, 2005.
[26] S.I. Resnick, Heavy–Tail Phenomena, Springer, 2007.
[27] G. Samorodnitsky, M.S. Taqqu, Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance, Chapman & Hall, 1994.
[28] J. Segers, Y. Zhao, T. Meinguet, Polar decomposition of regularly varying time series in star-shaped metric spaces, Extremes 20 (2017) 539–566.
[29] J.Q. Shi, T. Choi, Gaussian Process Regression Analysis for Functional Data, CRC Press, 2011.