

The spatial sign covariance operator: Asymptotic results and applications

Graciela Boente*, Daniela Rodriguez, Mariela Sued

Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina

ARTICLE INFO

Article history:

Received 3 November 2017

Available online 9 October 2018

Keywords:

Asymptotic distribution

Fisher-consistency

Functional data

Spatial sign covariance operator

Spherical principal components

AMS subject classifications:

62E20

62G35

ABSTRACT

Due to increased recording capability, functional data analysis has become an important research topic. For functional data, the study of outlier detection and/or the development of robust statistical procedures started only recently. One robust alternative to the sample covariance operator is the sample spatial sign covariance operator. In this paper, we study the asymptotic behavior of the sample spatial sign covariance operator centered at an estimated location. Among possible applications of our results, we derive the asymptotic distribution of the principal directions obtained from the sample spatial sign covariance operator and we develop a testing procedure to detect differences between the scatter operators of two populations. The test performance is illustrated through a Monte Carlo study for small sample sizes.

© 2018 Elsevier Inc. All rights reserved.

1. Introduction

Functional data analysis is concerned with the analysis of samples of curves registered in a continuous time period. A more general and inclusive framework that can accommodate cases where the observations are images or surfaces is to consider realizations of a random element X on a Hilbert space. The area has attracted much interest in the statistical community and has undergone a major development since technological advances in data collection and storage require procedures specifically designed for dealing with such data. It has been extensively discussed that simplifying a functional model by discretizing the observations as sequences of numbers can often fail to capture some of its important characteristics, such as the smoothness and continuity of the underlying functions. Statistical methods to analyze such functional data may be found, e.g., in [14,15,22,23,31]. For a summary of recent advances in functional statistics, see [9,20].

In this setting, the analysis of the covariance operator arises in many applied contexts. In particular, functional principal component analysis is a common tool to explore the characteristics of the data within a space of reduced dimension, since it allows to obtain a finite-dimensional representation of the data keeping the most significant components of the process. This finite-dimensional representation has also been used to obtain regression estimators under either a functional regression or a semi-linear model; see, e.g., [21,26]. It can further be combined with other techniques to classify data. Moreover, it is a powerful tool to detect atypical observations or outliers in the data set, when combined with a robust estimation procedure.

As is well known, the principal directions are the eigenfunctions of the covariance operator. Using this property to provide robust estimators of the principal directions, Locantore et al. [27] define spherical principal components; see also [19]. Other procedures to estimate robustly the principal directions include projection-pursuit [2], the robust approach from [32], the M -type smoothing spline estimators proposed in [25] and S -estimators [6]. In particular, the spherical principal components

* Correspondence to: Departamento de Matemáticas, FCEyN, UBA, Ciudad Universitaria, Pabellón 1, Buenos Aires, C1428EHA, Argentina.

E-mail addresses: gboente@dm.uba.ar (G. Boente), drodri@dm.uba.ar (D. Rodriguez), msued@dm.uba.ar (M. Sued).

are the eigenfunctions of the spatial sign operator which correspond to the sample covariance operator of the centered curves projected on the unit sphere as defined in Section 2.

One key point when deriving detection rules is that the robust functional principal direction estimators are indeed estimating the target directions. In this sense, random elements having an elliptical distribution, as introduced in Bali and Boente [1], provide a framework to study the consistency of some of the above mentioned procedures. We recall their definition in Section 2.1. In particular, for elliptically distributed random elements, the projection-pursuit estimators and the spherical principal components are Fisher consistent; see [2,7]. Moreover, Boente et al. [7] proved that for a random element with elliptical distribution, the linear space spanned by the q eigenfunctions associated to the q larger eigenvalues of the spatial sign covariance operator provides the best q -dimensional approximation to the centered process, in the sense of having the stochastically smallest residual squared norm among all linear spaces of dimension q . This result does not require the existence of a second order moment for the process. To the best of our knowledge, the asymptotic distribution of the robust principal direction estimators mentioned above is unknown. One of the goals of this paper is to derive the asymptotic distribution of the spherical principal component estimators through that of the sample spatial sign covariance operator.

A more recent statistical problem is that of testing for equality or proportionality between the covariance operators of two populations. For instance, Ferraty et al. [16] considered tests for comparing groups of curves based on a comparison of their covariances. By the Karhunen–Loève expansion, this is equivalent to testing whether all the samples have the same set of functional principal components sharing also their size. When considering only two populations, Benko et al. [3], Fremdt et al. [18] and Panaretos et al. [29] used this characterization to develop test statistics. In particular, Benko et al. [3] proposed two-sample bootstrap tests for specific aspects of the spectrum of functional data, such as the equality of a subset of eigenfunctions. As for Fremdt et al. [18] and Panaretos et al. [29], they considered an approach based on the projection of the data over a suitably chosen finite-dimensional space, such as that defined by the functional principal components of each population. The results in [18] generalized those provided in [29] which assume that the processes have a Gaussian distribution. More recently, Pigoli et al. [30] developed a two-sample test for comparing covariance operators using different distances between covariance operators. Their procedure is based on a permutation test and assumes that the two samples have the same mean; otherwise, an approximate permutation test is considered after the processes are centered using their sample means.

Some authors have also considered robust proposals for this problem. Kraus and Panaretos [24] introduced a class of dispersion operators and proposed a procedure for testing for equality of dispersion operators among two populations. Recently, Boente et al. [4] extended the classical two-population problem, presenting a test for equality of covariance operators among k populations in which the asymptotic distribution of the sample covariance operator plays a crucial role in deriving the asymptotic distribution of the proposed statistic. It is well known that the presence of outliers in the sample might lead to invalid conclusions. This motivates the development of robust procedures to deal with problems of this sort.

In this paper, as an application of our results, we present a test of equality for the spatial sign covariance operators between two populations. The statistic mimics the one presented for the classical setting and, as in the classical setting, its asymptotic distribution depends on that of the empirical spatial sign covariance operator for each population. It is worth noting that for functional elliptical distributions, equality of spatial sign covariance operators guarantees that the considered populations have the same principal components. Unlike the classical case, where the estimation of the mean plays no role in the asymptotic distribution of the covariance operator estimator, the imputation of an estimated location when defining the spatial sign covariance estimator requires some special consideration. One of the goals of this paper is to present a detailed proof of the asymptotic distribution of the sample spatial sign covariance estimator, which extends to the functional setting the results given by Dürre et al. [13] in the finite-dimensional case.

The paper is organized as follows. In Section 2, we introduce the notation to be used in the paper as well as the spatial sign covariance operator with an unknown location, while Section 3 deals with its consistency and asymptotic normality. Section 4 considers the application of the obtained results to two situations: the study of the asymptotic distribution of the spherical principal component estimators and the proposal of a testing procedure to detect differences between the spatial sign covariance operators of two populations, whose performance is also numerically studied for small samples. Proofs are relegated to the [Appendix](#).

2. The spatial covariance operator

Let \mathcal{H} be a separable Hilbert space, such as $L^2(\mathcal{I})$ for some bounded interval \mathcal{I} , with inner product $\langle \cdot, \cdot \rangle$ and norm $\|u\| = \langle u, u \rangle^{1/2}$. The functional sign of $u \in \mathcal{H}$, is defined as $s(u) = u/\|u\|$ for $u \neq 0$ and $s(0) = 0$. Let X be a random element taking values in \mathcal{H} . For a given $v \in \mathcal{H}$, the spatial or sign covariance operator of X centered at v is defined by

$$\Gamma^s(v) = \mathbb{E} \{s(X - v) \otimes s(X - v)\},$$

where \otimes denotes the tensor product on \mathcal{H} , e.g., for $u_1, u_2 \in \mathcal{H}$, the operator $u_1 \otimes u_2 : \mathcal{H} \rightarrow \mathcal{H}$ is defined as $(u_1 \otimes u_2)w = \langle u_2, w \rangle u_1$. Note that $u_1 \otimes u_2$ is a compact operator that belongs to \mathcal{F} , the Hilbert space of Hilbert–Schmidt operators over \mathcal{H} . Recall that for $\gamma \in \mathcal{F}$, γ^* denotes the adjoint of the operator γ , while for $\gamma_1, \gamma_2 \in \mathcal{F}$, the inner product in \mathcal{F} is defined as $\langle \gamma_1, \gamma_2 \rangle_{\mathcal{F}} = \text{trace}(\gamma_1^* \gamma_2) = \sum_{\ell=1}^{\infty} \langle \gamma_1 u_{\ell}, \gamma_2 u_{\ell} \rangle$, and so the norm equals

$$\|\gamma\|_{\mathcal{F}} = \langle \gamma^*, \gamma \rangle_{\mathcal{F}}^{1/2} = \left(\sum_{\ell=1}^{\infty} \|\gamma u_{\ell}\|^2 \right)^{1/2},$$

with $\{u_1, u_2, \dots\}$ any orthonormal basis of \mathcal{H} . These definitions are independent of the choice of basis. Given independent random elements X_1, \dots, X_n , distributed as X , for each $v \in \mathcal{H}$ define the sample version of $\Gamma^s(v)$ as

$$\widehat{\Gamma}_n^s(v) = \frac{1}{n} \sum_{i=1}^n s(X_i - v) \otimes s(X_i - v).$$

The Law of Large Numbers in \mathcal{F} entails that, for any $v \in \mathcal{H}$, $\widehat{\Gamma}_n^s(v)$ converges almost surely to $\Gamma^s(v)$. Moreover, the asymptotic distribution can be obtained from the Central Limit Theorem in \mathcal{F} ; see, e.g., Dauxois et al. [11].

Typically, the spatial operator is centered using a given location parameter μ which, in most cases, can be written as a functional $\mu = \mu(P)$ of the process distribution P . In a robust context, several location functionals $\mu(P)$ have been considered. Among others, the functional geometric median or spatial median $\mu_{\text{GM}} = \mu_{\text{GM}}(P)$ of the process X defined as

$$\mu_{\text{GM}} = \underset{u \in \mathcal{H}}{\operatorname{argmin}} E(\|X - u\| - \|X\|), \quad (1)$$

is the usual choice to center the data when using the spatial operator. However, our results do not restrict to this particular location. To describe the overall shape of the process, Fraiman and Muñiz [17] define the α -trimmed mean which may also be considered to center the data. Furthermore, a location parameter $\mu(P)$ can be defined through a suitable depth notion by considering the associated notion of median as the deepest point; see [9,10,28]. Note that these procedures provide Fisher-consistent functionals when considering a symmetric process around $v \in \mathcal{H}$, meaning that $X - v$ and $v - X$ have the same distribution. Effectively, in such a case, the spatial median μ_{GM} , the trimmed mean from [17] and any median or trimmed median defined from a depth function are equal to v , which is the natural target location parameter.

From now on, regardless of the location functional μ used, the notation $\Gamma^s(\mu)$ will refer to the spatial operator centered at $\mu = \mu(P)$. Taking into account that in most situations, μ is unknown, an estimator of μ must be considered when estimating $\Gamma^s(\mu)$. More precisely, we will consider a consistent estimator $\widehat{\mu}_n$ of μ to ensure that $\widehat{\Gamma}_n^s(\widehat{\mu}_n)$ provides a valid estimator of $\Gamma^s(\mu)$. The asymptotic properties of $\widehat{\Gamma}_n^s(\widehat{\mu}_n)$ are presented in Section 3.

2.1. Some general comments

As mentioned in the Introduction, the sample spatial operator $\widehat{\Gamma}_n^s(\widehat{\mu}_n)$ has been used as an alternative to the sample covariance operator when considering robust estimation procedures. In particular, it has been considered when one suspects that the underlying distribution may not have finite moments. Elliptical random elements have been introduced in [1] and further studied in [7]. For completeness, we recall their definition.

Let X be a random element in a separable Hilbert space \mathcal{H} . We say that X has an elliptical distribution with parameters $\mu \in \mathcal{H}$ and $\Gamma : \mathcal{H} \rightarrow \mathcal{H}$, where Γ is a self-adjoint, positive semi-definite and compact operator, if and only if for any linear and bounded operator $A : \mathcal{H} \rightarrow \mathbb{R}^d$ we have that the vector AX has a d -variate elliptical distribution with location parameter $A\mu$, shape matrix $A\Gamma A^*$ and characteristic generator φ , i.e., $AX \sim \mathcal{E}_d(A\mu, A\Gamma A^*, \varphi)$, where $A^* : \mathbb{R}^d \rightarrow \mathcal{H}$ denotes the adjoint operator of A . We write $X \sim \mathcal{E}(\mu, \Gamma, \varphi)$ and μ and Γ are called the location and the scatter operator of X , respectively. Hence, elliptical families provide a more general setting than considering Gaussian random elements and the sign operator gives a useful tool to obtain Fisher-consistent estimators of the principal directions, i.e., estimators consistent to the eigenfunctions of the elliptical process scatter operator, even when second moments do not exist; see [7].

Note that, as mentioned in [7], $X \sim \mathcal{E}(\mu, \Gamma, \varphi)$ if and only if $\langle u, X \rangle \sim \mathcal{E}_1(\langle u, \mu \rangle, \langle u, \Gamma u \rangle, \varphi)$ for all $u \in \mathcal{H}$. It is well known that, when $E\|X\|^2 < \infty$, $E(X) = \mu$ and Γ is up to a constant the covariance operator of X . Furthermore, using that an elliptical process is symmetric around μ , we get that $\mu(P) = \mu$ for any location functional $\mu(P)$ which is Fisher-consistent for symmetric processes. In particular, the functional geometric median defined in (1) equals the location μ .

For elliptical random elements, two situations may arise: either the scatter operator Γ has a finite rank q or it does not have a finite rank. In the first case, the process X has a finite q -dimensional expansion, viz.

$$X = \mu + \sum_{j=1}^q \lambda_j^{1/2} \xi_j \phi_j, \quad (2)$$

where ϕ_j are the eigenfunctions of Γ related to the eigenvalues λ_j ordered in decreasing order and $\xi = (\xi_1, \dots, \xi_q)^T \sim \mathcal{E}_q(0, \mathbf{I}_q, \varphi)$, i.e., ξ has a spherical distribution. In this setting, the asymptotic behavior of $\widehat{\Gamma}_n^s(\widehat{\mu}_n)$ may be derived from the results given in [13], since the distribution of $\operatorname{diag}(\lambda_1^{1/2}, \dots, \lambda_q^{1/2})\xi$ is symmetric around 0 . In contrast, if Γ does not have a finite rank, Proposition 2.1 in [7] states that the process is a scale mixture of Gaussian distributions. More precisely, there exists a zero mean Gaussian random element Y and a random variable $V > 0$ independent of Y such that

$$X = \mu + VY. \quad (3)$$

Without loss of generality, throughout the paper, we will assume that Γ is the covariance operator of Y . The results given in Section 3 include this case but they also provide consistency and asymptotic normality results in a more general framework than elliptical families.

3. Asymptotic results

The following results establish the consistency and the asymptotic normality of the spatial sign covariance operator with an unknown location. The proofs are relegated to the [Appendix](#). From now on, the notation $u_n \xrightarrow{\text{as}} u$ in \mathcal{H} means that $\|u_n - u\| \xrightarrow{\text{as}} 0$, while for random operators $\Upsilon_n \in \mathcal{F}$, the convergence $\Upsilon_n \xrightarrow{\text{as}} \Upsilon$ in \mathcal{F} stands for $\|\Upsilon_n - \Upsilon\|_{\mathcal{F}} \xrightarrow{\text{as}} 0$.

Theorem 1. Let $\mu \in \mathcal{H}$ be a given location parameter of the process. Assume that $\hat{\mu}_n$ is an estimator of μ such that $\hat{\mu}_n \xrightarrow{\text{as}} \mu$ in \mathcal{H} as $n \rightarrow \infty$ and $E(\|X - \mu\|^{-1}) < \infty$. Then, we have $\hat{\Gamma}^s(\hat{\mu}_n) \xrightarrow{\text{as}} \Gamma^s(\mu)$ in \mathcal{F} as $n \rightarrow \infty$.

Remark 1. As we mentioned, several location parameters $\mu = \mu(P)$ can be considered. In particular, the spatial median μ_{GM} defined in (1) is the usual choice to center the data when the spatial covariance operator is considered. In the functional case, different methods have been proposed to provide estimators of μ_{GM} . On one hand, Gervini [19] shows that the sample spatial median, denoted $\hat{\mu}_{n,\text{GM}}$ and defined as the solution of the empirical version of (1), can be found solving a convex n -dimensional minimization problem. Furthermore, $\hat{\mu}_{n,\text{GM}}$ is strongly consistent with respect to the weak topology in \mathcal{H} , i.e., for any $u \in \mathcal{H}$, $\langle \hat{\mu}_{n,\text{GM}}, u \rangle$ converges almost surely to $\langle \mu_{\text{GM}}, u \rangle$. On the other hand, Cardot et al. [8] propose to estimate the spatial median through an algorithm that can be seen as a stochastic gradient algorithm. Theorem 3.1 in [8] shows that, under mild conditions, this estimator converges in norm to the geometric median almost surely. This result guarantees the existence of strong consistent estimators of the spatial median (with respect to the strong topology in \mathcal{H}) and hence, that of the estimators of the spatial sign covariance operator. Besides, when μ and $\hat{\mu}_n$ are the α -trimmed mean defined in Fraiman and Muñiz [17] and its corresponding estimators, Theorem 3.2 in [17] together with Theorem 1 entails that $\hat{\Gamma}^s(\hat{\mu}_n)$ is a consistent estimator of $\Gamma^s(\mu)$.

In order to study the asymptotic distribution of $\hat{\Gamma}^s(\hat{\mu}_n)$, let \mathcal{B} be the Banach space of linear and continuous operators from \mathcal{H} to \mathcal{F} , i.e., $\mathcal{B} = \{\mathcal{T} : \mathcal{H} \rightarrow \mathcal{F} : \mathcal{T} \text{ linear and continuous}\}$ and let $\|\mathcal{T}\|_{\mathcal{B}} = \sup_{\|u\| \leq 1} \|\mathcal{T}(u)\|_{\mathcal{F}}$. The following assumptions will be required.

(A.1) $\sqrt{n}(\hat{\mu}_n - \mu) = O_{\text{Pr}}(1)$, where $O_{\text{Pr}}(1)$ denotes a random element $v_n \in \mathcal{H}$ such that $\|v_n\|$ is bounded in probability.

(A.2) $E(\|X - \mu\|^{-3/2}) < \infty$.

Theorem 2. Under Assumptions (A.1) and (A.2), we have $\sqrt{n}\{\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}^s(\mu)\} = \sqrt{n}G_X(\hat{\mu}_n - \mu) + o_{\text{Pr}}(1)$, where $o_{\text{Pr}}(1)$ stands for a remainder term $R_n \in \mathcal{F}$ such that $\|R_n\|_{\mathcal{F}} \xrightarrow{P} 0$ as $n \rightarrow \infty$, $G_X = 2F_X - S_X \in \mathcal{B}$, with

$$F_X(u) = E \left\{ \frac{\langle X - \mu, u \rangle}{\|X - \mu\|^4} (X - \mu) \otimes (X - \mu) \right\} \quad \text{and} \quad S_X(u) = \left\{ u \otimes E \left(\frac{X - \mu}{\|X - \mu\|^2} \right) + E \left(\frac{X - \mu}{\|X - \mu\|^2} \right) \otimes u \right\}. \quad (4)$$

Remark 2. Assumption (A.1) is satisfied when μ is the spatial median μ_{GM} , taking $\hat{\mu}_n$ as the average of the stochastic gradient algorithm estimator, presented in [8], where the asymptotic distribution of this estimator is obtained; see their Theorem 3.4. Regarding the assumptions $E(\|X - \mu\|^{-1}) < \infty$ and (A.2), they require that the probability mass is not too strongly concentrated near μ , as noted in [13] in the multivariate case. In particular, assume that the process X has a finite Karhunen–Loève expansion, viz.

$$X = \mu + \sum_{k=1}^q y_k \phi_k,$$

where $\phi_k \in \mathcal{H}$ are orthonormal and y_k are random variables, then $E_\gamma = E(\|X - \mu\|^{-\gamma}) = E(\|\mathbf{y}\|^{-\gamma})$, with $\mathbf{y} = (y_1, \dots, y_q)^\top$. Hence, if $q \geq 2$, $E_\gamma < \infty$ for $\gamma \in \{1, 3/2\}$ when \mathbf{y} has a bounded density at 0 while a weaker requirement may be given when \mathbf{y} has an elliptical distribution; see Remark V in [13]. For properly infinite-dimensional processes, $E_\gamma < \infty$ if there exists an orthonormal basis $\{\psi_1, \psi_2, \dots\}$ in \mathcal{H} such that, for some $q \geq 2$, the random vector $\mathbf{y} = (\langle X, \psi_1 \rangle, \dots, \langle X, \psi_q \rangle)^\top$ is such that $E(\|\mathbf{y}\|^{-\gamma}) < \infty$. In particular, for elliptically distributed random elements, one may take as the basis the eigenfunctions of the scatter operator defining the distribution. When the scatter operator of the elliptical distribution does not have a finite rank, invoking (3), we get that $E_\gamma = EV^{-\gamma}E\|Y\|^{-\gamma}$, where Y is a Gaussian process with covariance operator Γ and $V > 0$ is a random variable independent of Y . Hence, $E_\gamma < \infty$ if and only if $EV^{-\gamma} < \infty$. In particular, when V is such that $k/V^2 \sim \chi_k^2$, which corresponds to the functional version of a multivariate \mathcal{T} -distribution with k degrees of freedom, we have $E_\gamma < \infty$.

Remark 3. Note that when $F_X \equiv 0$ and $E\{(X - \mu)\|X - \mu\|^{-2}\} = 0$, Theorem 2 provides an extension to the functional data setting of the result given in Theorem 2 of [13]. More precisely, in this case $\sqrt{n}\{\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}^s(\mu)\} = o_{\text{Pr}}(1)$, meaning that the asymptotic behavior of the spatial sign covariance operator is not affected by the imputation of a location estimator. In particular, if X has a symmetric distribution around μ and $\hat{\mu}_n$ stands for the estimator defined in [8], then $\sqrt{n}\{\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}^s(\mu)\} = o_{\text{Pr}}(1)$, so that the asymptotic distribution of $\hat{\Gamma}^s(\hat{\mu}_n)$ can be obtained from that of $\hat{\Gamma}^s(\mu)$ using the Central Limit Theorem. Note that this assertion holds for elliptical families.

Furthermore, if $X \sim \mathcal{E}(\mu, \Gamma, \varphi)$ and Γ does not have a finite rank, we get that $\Gamma^s(\mu) = E\{s(Y) \otimes s(Y)\}$, the sign operator of the Gaussian process Y given in (3). Besides, when the estimator $\hat{\mu}_n$ of μ is the functional median defined in [8], noticing that

$$\hat{\Gamma}_n^s(\mu) = \frac{1}{n} \sum_{i=1}^n s(Y_i) \otimes s(Y_i),$$

we get that $\sqrt{n} \{\hat{\Gamma}_n^s(\hat{\mu}_n) - \Gamma^s(\mu)\}$ converges in \mathcal{F} to a zero mean Gaussian element with covariance operator equal to the covariance operator of $s(Y) \otimes s(Y)$.

For multivariate data, Theorem 2 of [13] gives the asymptotic distribution of the spatial sign operator. **Corollary 1** extends this result to the functional setting. In the general situation in which one cannot guarantee that $F_X \equiv 0$ and $E\{(X - \mu)\|X - \mu\|^{-2}\} = 0$, a joint asymptotic distribution between the estimator of the given location parameter and $\hat{\Gamma}^s(\mu)$ is needed.

Corollary 1. Assume that (A.2) holds and that $(\sqrt{n}(\hat{\mu}_n - \mu), \sqrt{n}(\hat{\Gamma}^s(\mu) - \Gamma^s(\mu))) \rightsquigarrow Z$ as $n \rightarrow \infty$, where Z is a zero mean Gaussian random object in $\mathcal{H} \times \mathcal{F}$, with covariance operator $\Upsilon : \mathcal{H} \times \mathcal{F} \rightarrow \mathcal{H} \times \mathcal{F}$. Then, $\sqrt{n} \{\hat{\Gamma}^s(\hat{\mu}_n) - \Gamma^s(\mu)\}$ converges in \mathcal{F} to a zero mean Gaussian element with covariance operator given by $(G_X \Pi_{\mathcal{H}} + \Pi_{\mathcal{F}}) \Upsilon (G_X \Pi_{\mathcal{H}} + \Pi_{\mathcal{F}})^*$, where $\Pi_{\mathcal{H}}$ and $\Pi_{\mathcal{F}}$ are the projection operators from $\mathcal{H} \times \mathcal{F}$ to \mathcal{H} and \mathcal{F} , respectively. Moreover, $G_X^* = 2F_X^* - S_X^*$ with F_X^* and S_X^* the adjoint operators of F_X and S_X , respectively given by

$$S_X^*(\Upsilon) = \Upsilon \left(E \frac{X - \mu}{\|X - \mu\|^2} \right) + \Upsilon^* \left(E \frac{X - \mu}{\|X - \mu\|^2} \right), \quad F_X^*(\Upsilon) = E \left\{ \frac{\langle (X - \mu) \otimes (X - \mu), \Upsilon \rangle_{\mathcal{F}}}{\|X - \mu\|^4} (X - \mu) \right\}.$$

4. Applications

In this section, we consider two applications of the results obtained in Section 3. We first derive the asymptotic behavior of the principal direction estimators obtained as the eigenfunctions of $\hat{\Gamma}^s(\hat{\mu}_n)$ from **Corollary 1**. Our second application uses the asymptotic distribution of the sample spatial sign operator to present a procedure to test equality among sign covariance operators.

4.1. On the asymptotic behavior of the spherical principal direction estimators

Robust estimators of the principal directions for functional data have been extensively studied since the spherical principal components proposed in [27] and studied in [19]. As mentioned in the Introduction, Fisher-consistency of several proposals including the spherical principal directions has been studied in a more general framework than Gaussian random elements, without requiring finite moments, such as that given by elliptically distributed random elements.

When considering the spherical principal directions, two possible situations may arise: either (a) the distribution is concentrated on a finite-dimensional subspace or (b) the rank of $\Gamma^s(\mu)$ is infinite, where μ stands for a given location parameter of X which is typically the functional spatial median μ_{GM} . Gervini [19] showed that the spherical principal direction estimators are Fisher-consistent for the principal directions when the process admits a Karhunen–Loève expansion with only finitely many terms, while Boente et al. [7] proved that the spherical principal components are in fact Fisher-consistent for any elliptical distribution. More precisely, assume that either of the following conditions holds:

- (a) Representation (2) holds for X , where $\lambda_1 \geq \dots \geq \lambda_k > 0$, $\phi_k \in \mathcal{H}$ are orthonormal and ξ_k are random variables such that $(\xi_1, \dots, \xi_q)^T$ has symmetric and exchangeable marginals.
- (b) $X \sim \mathcal{E}(\mu, \Gamma, \varphi)$ and denote $\lambda_1 \geq \lambda_2 \geq \dots$ the eigenvalues of the scatter operator Γ with associated eigenfunctions ϕ_j .

Note that in case a), the scatter operator $\Gamma = \lambda_1 \phi_1 + \dots + \lambda_q \phi_q$ has finite rank. As shown in [7, 19], the eigenfunctions of $\Gamma^s(\mu)$ are those of Γ and in the same order. More precisely, if $\lambda_1^s \geq \lambda_2^s \geq \dots$ stand for the ordered eigenvalues of $\Gamma^s(\mu)$, under a) or b), we have that ϕ_k is the eigenfunction of $\Gamma^s(\mu)$ related to the eigenvalue λ_k^s , meaning that the spatial principal directions are Fisher-consistent. Moreover, we also have that $\lambda_j^s > \lambda_{j+1}^s$ if $\lambda_j > \lambda_{j+1}$.

Beyond Fisher-consistency, consistency and order of consistency are also desirable properties for any robust procedure. However, for most of the proposed methods, only consistency results were obtained. In this section, we derive the asymptotic distribution of the spherical principal direction estimators, which correspond to the eigenfunctions of the spatial sign operator estimator. In this sense, our result provides the first asymptotic normality result for robust principal direction estimators in a general setting.

Even though the asymptotic behavior of the eigenfunctions of $\hat{\Gamma}^s(\mu)$ can easily be obtained from the Central Limit Theorem and the results in [11], **Theorem 2** states that this asymptotic behavior may not be the same when the given location is unknown and estimated. However, it should be stressed that for elliptical distributed random elements or under the symmetry assumptions required in Gervini [19] to ensure Fisher consistency, we have that the asymptotic behavior of the eigenfunctions of $\hat{\Gamma}^s(\hat{\mu}_n)$ is that of the eigenfunctions of $\Gamma^s(\mu)$, since as mentioned in Remark 3, $\sqrt{n} \{\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}^s(\mu)\} = o_{\text{Pr}}(1)$.

Similar arguments to those considered in [11] and Corollary 1 allow to obtain the asymptotic distribution of the spatial principal direction estimators beyond elliptical families. To this end, denote $\lambda_1^s, \lambda_2^s, \dots$ the sequence of eigenvalues of $\Gamma^s(\mu)$ ordered in decreasing order and as $\phi_1^s, \phi_2^s, \dots$ their related eigenfunctions. Let $\hat{\phi}_1^s, \hat{\phi}_2^s, \dots$ be the eigenfunctions of $\hat{\Gamma}^s(\hat{\mu}_n)$ related to the ordered eigenvalues $\hat{\lambda}_1^s \geq \hat{\lambda}_2^s \geq \dots$. Recall that if the process has an elliptical distribution with scatter operator Γ , $\phi_j^s = \phi_j$ the j th eigenfunction of Γ .

Define $\Lambda_i = \{j \in \mathbb{N} : \lambda_j^s = \lambda_i^s\}$, $\Lambda = \{i \in \mathbb{N} : \text{card}(\Lambda_i) = 1\}$ and the projection operators

$$\Pi_i^s = \sum_{j \in \Lambda_i} \phi_j^s \otimes \phi_j^s \quad \text{and} \quad \hat{\Pi}_i^s = \sum_{j \in \Lambda_i} \hat{\phi}_j^s \otimes \hat{\phi}_j^s.$$

The following result is a direct consequence of Propositions 3, 4, 6, 10 in [11] and Corollary 1. Taking into account that the i th principal direction is defined up to a sign change when the eigenvalue λ_i^s has multiplicity 1, in the sequel, we choose the direction of the eigenfunction estimator so that $\langle \hat{\phi}_i^s, \phi_i^s \rangle > 0$.

Proposition 1. Assume that (A.2) holds and that $(\sqrt{n}(\hat{\mu}_n - \mu), \sqrt{n}\{\hat{\Gamma}^s(\mu) - \Gamma^s(\mu)\}) \rightsquigarrow Z$, where Z is a zero mean Gaussian random object in $\mathcal{H} \times \mathcal{F}$, with covariance operator $\Upsilon : \mathcal{H} \times \mathcal{F} \rightarrow \mathcal{H} \times \mathcal{F}$. Denote as $\Pi_{\mathcal{H}}$ and $\Pi_{\mathcal{F}}$ the projection operators from $\mathcal{H} \times \mathcal{F}$ to \mathcal{H} and \mathcal{F} , respectively and as U a zero mean Gaussian random object in \mathcal{F} with covariance operator $\Upsilon^s = (G_{\mathcal{H}}\Pi_{\mathcal{H}} + \Pi_{\mathcal{F}})\Upsilon(G_{\mathcal{H}}\Pi_{\mathcal{H}} + \Pi_{\mathcal{F}})^*$. The following statements then hold.

- (i) $\hat{\Pi}_i^s \xrightarrow{\text{as}} \Pi_i^s$ in \mathcal{F} as $n \rightarrow \infty$. Moreover, for any $i \in \Lambda$, $\hat{\phi}_i^s \xrightarrow{\text{as}} \phi_i^s$ in \mathcal{H} as $n \rightarrow \infty$.
- (ii) As $n \rightarrow \infty$, $\sqrt{n}(\hat{\Pi}_i^s - \Pi_i^s)$ converges in distribution to the zero mean Gaussian random element of \mathcal{F} given by $\Delta_i U \Pi_i^s + \Pi_i^s U \Delta_i$, where

$$\Delta_i = \sum_{\ell \in \Lambda - \Lambda_i} \frac{1}{\lambda_i^s - \lambda_\ell^s} \phi_\ell^s \otimes \phi_\ell^s.$$

Furthermore, when $i \in \Lambda$, we have that as $n \rightarrow \infty$, $\sqrt{n}(\hat{\phi}_i^s - \phi_i^s) \rightsquigarrow (\Delta_i U)(\phi_i)$, which is a zero mean Gaussian process in \mathcal{H} .

Note that when $i \in \Lambda$, $\Delta_i = \sum_{\ell \neq i} \{1/(\lambda_i^s - \lambda_\ell^s)\} \phi_\ell^s \otimes \phi_\ell^s$.

4.2. Tests for equality of the sign covariance operators

The asymptotic distribution of the spatial covariance operator given in Corollary 1 allows to construct a test of equality between spatial covariance operators between two different populations. More precisely, assume that for $i \in \{1, 2\}$, we have independent observations $X_{i,1}, \dots, X_{i,n_i}$ such that $X_{i,j} \sim X_i \sim P_i$, for all $j \in \{1, \dots, n_i\}$. Let $\mu_i = \mu(P_i)$ stand for a given location parameter of the i th population. For the sake of simplicity, let $\Gamma_i^s = E\{s(X_i - \mu_i) \otimes s(X_i - \mu_i)\}$ be the spatial sign covariance operator of the i th population. We are interested in testing the null hypothesis

$$\mathcal{H}_0 : \Gamma_1^s = \Gamma_2^s \quad \text{vs.} \quad \mathcal{H}_1 : \Gamma_1^s \neq \Gamma_2^s.$$

As in [4], we will reject the null hypothesis when the difference between the estimated spatial sign covariance operators is large. Namely, if for each $i \in \{1, 2\}$, $\hat{\Gamma}_i^s$ stands for a consistent estimator of Γ_i^s based on $X_{i,1}, \dots, X_{i,n_i}$, we define

$$T_n^s = n \|\hat{\Gamma}_2^s - \hat{\Gamma}_1^s\|_{\mathcal{F}}^2, \tag{5}$$

where $n = n_1 + n_2$. The asymptotic distribution of T_n^s can be obtained from the asymptotic distribution of $\sqrt{n}(\hat{\Gamma}_i^s - \Gamma_i^s)$, as stated in the following proposition, which can be considered as a robust version of Corollary 1 in [4]. Its proof can be obtained using Theorem 1 from the above-mentioned paper.

Proposition 2. For $i \in \{1, 2\}$, let $X_{i,1}, \dots, X_{i,n_i} \in \mathcal{H}$ be independent observations from two independent populations with given location parameters μ_i . Denote as $\Gamma_i^s = E\{s(X_i - \mu_i) \otimes s(X_i - \mu_i)\}$ the spatial sign covariance operator of the i th population. Assume that $n_i/n \rightarrow \tau_i$ with $\tau_i \in (0, 1)$ as $n = n_1 + n_2 \rightarrow \infty$. Let $\hat{\Gamma}_i^s$ be independent estimators of the i th population spatial sign covariance operator such that $\sqrt{n_i}(\hat{\Gamma}_i^s - \Gamma_i^s) \rightsquigarrow U_i$ as $n \rightarrow \infty$, with U_i a zero mean Gaussian random element with covariance operator Υ_i . Denote $\Upsilon_w : \mathcal{F} \rightarrow \mathcal{F}$ the linear operator defined as $\Upsilon_w = (1/\tau_1)\Upsilon_1 + (1/\tau_2)\Upsilon_2$ and let $\theta_1, \theta_2, \dots$ stand for the sequence of eigenvalues of Υ_w ordered in decreasing order. Then, as $n \rightarrow \infty$, we have

$$n \|\hat{\Gamma}_2^s - \Gamma_2^s - (\hat{\Gamma}_1^s - \Gamma_1^s)\|_{\mathcal{F}}^2 \rightsquigarrow \sum_{\ell \geq 1} \theta_\ell Z_\ell^2,$$

with $Z_\ell \sim \mathcal{N}(0, 1)$ independent. In particular, if $\mathcal{H}_0 : \Gamma_1^s = \Gamma_2^s$ holds, we have, as $n \rightarrow \infty$,

$$n \|\hat{\Gamma}_2^s - \hat{\Gamma}_1^s\|_{\mathcal{F}}^2 \rightsquigarrow \sum_{\ell \geq 1} \theta_\ell Z_\ell^2. \tag{6}$$

The asymptotic results obtained in Section 3, in particular Corollary 1 and Remark 3, invite to consider as estimators of the sign operator

$$\widehat{\Gamma}_i^s = \frac{1}{n_i} \sum_{j=1}^{n_i} s(X_{i,j} - \widehat{\mu}_{n_i}) \otimes s(X_{i,j} - \widehat{\mu}_{n_i}),$$

with $\widehat{\mu}_{n_i}$ any consistent estimators of the functional geometric median $\mu_i = \mu_{\text{GM}}(P_i)$ of the process X_i satisfying (A.1) such as the spatial median estimators given in [8]; see Remark 1. In such a case, as noted in [4], Eq. (6) motivates the use of bootstrap methods, to decide whether to reject the null hypotheses. Let q_1, q_2, \dots be a sequence of integers such that $q_n \rightarrow \infty$ as $n \rightarrow \infty$. Then, the bootstrap algorithm can be implemented as follows:

Step 1: For $i \in \{1, 2\}$ and given the sample $X_{i,1}, \dots, X_{i,n_i}$, let $\widehat{\gamma}_i$ be consistent estimators of γ_i . Define

$$\widehat{\gamma}_w = \widehat{\tau}_1^{-1} \widehat{\gamma}_1 + \widehat{\tau}_2^{-1} \widehat{\gamma}_2$$

with $\widehat{\tau}_i = n_i/(n_1 + n_2)$.

Step 2: Denote by $\widehat{\theta}_\ell$ the positive eigenvalues of $\widehat{\gamma}_w$, $\widehat{\theta}_1 \geq \widehat{\theta}_2 \geq \dots$.

Step 3: Generate $Z_1^*, \dots, Z_{q_n}^*$ iid such that $Z_i^* \sim \mathcal{N}(0, 1)$ and let $\mathcal{U}_n^* = \sum_{j=1}^{q_n} \widehat{\theta}_j Z_j^{*2}$.

Step 4: Repeat Step 3 N_B times, to get N_B values of \mathcal{U}_{nr}^* for all $r \in \{1, \dots, N_B\}$.

The $(1-\alpha)$ -quantile of the asymptotic null distribution of T_n^s can be approximated by the $(1-\alpha)$ -quantile of the empirical distribution of $\mathcal{U}_{n1}^*, \dots, \mathcal{U}_{nN_B}^*$. Besides, the p -value can be estimated by $\widehat{p} = s/N_B$, where s equals the number of \mathcal{U}_{nr}^* which are larger or equal than the observed value of the statistic T_n^s .

Remark 4. When the sequence q_1, q_2, \dots is such that $q_n/\sqrt{n} \rightarrow 0$ as $n \rightarrow \infty$, the validity of the bootstrap procedure can be derived from Theorem 3 in [4] if $n_i/n \rightarrow \tau_i$ with $\tau_i \in (0, 1)$ and the estimators of γ_w are such that $\sqrt{n} \|\widehat{\gamma}_w - \gamma_w\|_{\mathcal{F}} = O_{\text{Pr}}(1)$, ensuring that the asymptotic significance level of the test based on the bootstrap critical value is indeed α . For each $i \in \{1, 2\}$, the sample covariance operator of $Y_i = s(X_i - \mu_i) \otimes s(X_i - \mu_i)$ provides a consistent estimator $\widehat{\gamma}_i$ of γ_i .

One important point is how to choose the sequence q_1, q_2, \dots in practice. As in other settings, q_n can be chosen so that $\widehat{\theta}_1 + \dots + \widehat{\theta}_{q_n}$ represents a given percentage of the trace of $\widehat{\gamma}_w$. Another possibility is to select q_n by relating it to a given number M of eigenvalues of the pooled operator $\Gamma_{\text{POOL}}^s = \tau_1 \Gamma_1^s + \tau_2 \Gamma_2^s$, which is easily estimated using its empirical counterpart. This is the procedure used in the simulation study considered in Section 4.2.1, where more details are given. It is well known that the choice of data-driven selectors for smoothing parameters such as q_n is more difficult in testing than in estimation problems. However, the numerical results obtained in Section 4.2.1 show that the procedure is quite robust to the choice of q_n .

Remark 5. Proposition 2 ensures that, under mild assumptions, it is possible to provide a test to decide whether $\Gamma_1^s = \Gamma_2^s$. An important point to highlight is what this null hypothesis represents, e.g., in terms of the covariance operators of the two populations, when they exist. Let us consider the situation in which the two populations have an elliptical distribution, i.e., $X_i \sim \mathcal{E}(\mu_i, \Gamma_i, \phi_i)$ for each $i \in \{1, 2\}$. Recall that the eigenfunctions of Γ_i^s are those of Γ_i and in the same order, while the eigenvalues of the sign covariance operator Γ_i^s , denoted $\lambda_{i,\ell}^s$, are shrunk with respect to those of Γ_i (that are denoted as $\lambda_{i,\ell}$) as follows

$$\lambda_{i,\ell}^s = \lambda_{i,\ell} \text{E} \left(\xi_{i,\ell}^2 / \sum_{j=1}^{\infty} \lambda_{i,j} \xi_{i,j}^2 \right), \quad (7)$$

where $\xi_{i,j} = \lambda_{i,\ell}^{-1/2} \langle X_i - \mu_i, \phi_{i,j} \rangle$ with $\phi_{i,\ell}$ the eigenfunction of Γ_i .

Assume that $\varphi_1 = \varphi_2$, i.e., if the two populations have the same underlying distribution up to location and scatter. Note that if the scatter operators are proportional, i.e., if $\Gamma_2 = \rho \Gamma_1$ for some positive constant ρ , then $\Gamma_1^s = \Gamma_2^s$. Thus, when the two populations have the same elliptical distribution up to location and scatter, the test based on T_n^s provides a way of testing the proportionality of the scatter operators, even when second moments do not exist. Note that when second moments exist, the covariance operator of X_i is proportional to Γ_i , hence the statistic T_n^s allows to test proportionality between the two covariance operators. Note also that when $\Gamma_1^s = \Gamma_2^s$, both scatter operators have the same rank and share the same eigenfunctions. Also, if the scatter operators have finite rank, from Proposition 1 in [12], we get that $\Gamma_1^s = \Gamma_2^s$ if and only if $\Gamma_2 = \rho \Gamma_1$ for some positive constant ρ . Hence, for finite-rank scatter operators, testing proportionality of the scatter operators is equivalent to testing equality of the spatial sign operators.

4.2.1. Monte Carlo study

This section contains the results of a simulation study designed to illustrate the finite-sample performance of the testing procedure described in Section 4.2, under the null hypothesis and different alternatives, when atypical data are introduced in the samples. The numerical study also aims to compare the performance of the test based on the sign operator with that based on the sample covariance operator introduced in [4]. The R code for computing the test statistic is available at <http://mate.dm.uba.ar/~drodrig/programs/sign.rar>.

We have performed $N = 1000$ replications taking samples of size $n_1 = n_2 = 100$. For each $i \in \{1, 2\}$, the generated samples $X_{i,1}, \dots, X_{i,n_i}$ are such that $X_{i,j} \sim X_i \in L^2(0, 1)$. In all cases, each trajectory was observed at $m = 100$ equidistant points in the interval $[0, 1]$ and we performed $N_B = 5000$ bootstrap replications. To summarize the test performance, we computed the observed frequency of rejections over replications with nominal level $\alpha = 0.05$.

Simulation settings

The distributions of the two populations correspond, under the null hypothesis, to independent centered Brownian motion processes, henceforth denoted as $\mathcal{BM}(0, 1)$. Hence, both processes have the same spatial sign operators and also the same covariance operators. To check the test's power performance, we consider the same alternatives as in [4] and also Gaussian alternatives. More precisely, we generate independent observations $X_{1,1}, \dots, X_{1,n_1}$ with the same distribution as X_1 with $X_1 \sim \mathcal{BM}(0, 1)$. Likewise, $X_{2,1}, \dots, X_{2,n_2}$ are generated with the same distribution as X_2 , where the second population has a distribution according to either one of the following models:

Model 1: $X_2 \sim Y_1 + \delta_n Y_2^2$, where Y_1 and Y_2 are independent $\mathcal{BM}(0, 1)$ and $\delta_n = \Delta n^{-1/4}$ with $n = n_1 + n_2$ while Δ takes values from 0 to 8 in Step 1 and from 10 to 20 in Step 2. The case $\Delta = 0$ corresponds to the null hypothesis in which both processes have a Gaussian distribution.

Model 2: $X_2 \sim Y_1 + \delta_n Y_2$, where Y_1 and Y_2 are independent, $Y_1 \sim \mathcal{BM}(0, 1)$, Y_2 is a Gaussian process with covariance kernel $\text{cov}\{Y_2(t), Y_2(s)\} = \exp(-|s - t|/0.2)$ and $\delta_n = \Delta n^{-1/4}$ with $n = n_1 + n_2$ while $\Delta \in \{0, 0.5, 1, 1.5, 2, 2.5, 3, 4, 5\}$. In this case, both processes have a Gaussian distribution under the null and under the alternative, which implies that for each population, the spatial sign operator has the same eigenfunctions as the covariance operator. Moreover, the eigenvalues of the spatial operator and those of the covariance operator of the i th population are related through (7) with $\xi_{i,\ell} \sim \mathcal{N}(0, 1)$ independent of each other.

Denote as \mathcal{T}_1 and $\text{Bi}(1, p)$ the standard Cauchy or univariate Student t distribution with 1 degree of freedom and the Bernoulli distribution with success probability p , respectively. To analyze the behavior when atypical data are introduced in the sample, for each generated sample, we consider the following contamination. We first generate two independent samples $V_{i,1}, \dots, V_{i,n_i} \in \mathbb{R}$ such that $V_{i,j} \sim \mathcal{T}_1$ for each $i \in \{1, 2\}$. We then generate the contaminated samples, denoted $X_{i,j}^{(c)}$, as

$$X_{i,j}^{(c)} = (1 - B_j) X_{i,j} + B_j |V_{i,j}| X_{i,j},$$

where $B_j \sim \text{Bi}(1, 0.1)$ are independent and independent of $(X_{1,j}, V_{1,j}, X_{2,j}, V_{2,j})$. Note that under the null hypothesis, both populations have the same elliptical distribution since they can be written as $W_{i,j} X_{i,j}$ with $W_{i,j} = (1 - B_j) + B_j |V_{i,j}|$ a positive random variable independent of $X_{i,j}$ and $W_{1,j} \sim W_{2,j}$.

The test statistics

We computed two test statistics, the statistic based on the spatial sign operator defined above and the procedure defined in [4]. The test statistic given in the latter paper is defined as $T_n = n \|\hat{T}_1 - \hat{T}_2\|^2$, where

$$\hat{T}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{i,j} - \bar{X}_i) \otimes (X_{i,j} - \bar{X}_i)$$

is the sample covariance operator. This testing method is designed to test equality of the two populations covariance operators, which is fulfilled when $\Delta = 0$. In contrast, the statistic T_n^s defined in (5) is designed to test equality of the spatial operators, i.e., $\Gamma_1^s = \Gamma_2^s$. As mentioned in Remark 5, this null hypothesis is fulfilled when the scatter operators related to the elliptical distribution are proportional, which holds when $\Delta = 0$, both for clean and contaminated samples. When computing the spatial sign operators $\hat{\Gamma}_i^s$, we center the data with the functional median computed through the function `l1median` from the R package `pcaPP`.

The testing procedure requires bootstrap calibration. For that purpose, following the procedure described in [4], we project the centered data onto the M largest principal components of the pooled operators $\hat{\Gamma}_{\text{POOL}}$, where the pooled operator was adapted to the testing procedure used. More precisely,

$$\hat{\Gamma}_{\text{POOL}} = \frac{n_1 \hat{\Gamma}_1 + n_2 \hat{\Gamma}_2}{n}$$

when the test statistic is based on the sample covariance matrices, while

$$\hat{\Gamma}_{\text{POOL}}^s = \frac{n_1 \hat{\Gamma}_1^s + n_2 \hat{\Gamma}_2^s}{n}$$

when the test statistic corresponds to the sample sign operator. Usually the number M is chosen so as to explain a given percentage of the total operator trace which for the sign operator equals 1. If this percentage is large enough, the predicted data constructed using the first M spherical principal components give a good approximation of the original sample suggesting that the bootstrap procedure will provide adequate critical values. Once the number of principal directions is

Table 1

Percentage of the total variance explained by the first M principal components, under the null hypothesis, i.e., for $\Delta = 0$, when using the test $T_{B,M}$ or $T_{B,M}^S$.

	Clean samples				Contaminated samples			
	M				M			
	3	10	20	30	3	10	20	30
$T_{B,M}$	0.934	0.981	0.991	0.995	0.962	0.992	0.996	0.998
$T_{B,M}^S$	0.828	0.950	0.979	0.989	0.828	0.950	0.979	0.989

selected, we project the data over the linear space spanned by them and the pooled covariance operator, denoted \mathcal{V}_w in Proposition 2, is then estimated through a finite-dimensional matrix.

Note that, in this sense, the goal of this numerical study is twofold. It aims not only to study the performance of the sign operator testing procedure for clean and contaminated samples but also to analyze the effect that different values of q_n may have on its level and power. Note that, in our numerical study as in [4], q_n is related to the number M of principal directions through $q_n = M(M+1)/2$. Further note that the condition $q_n/\sqrt{n} \rightarrow 0$ entails that the number M of components should increase to infinity but at a rate slower than $n^{1/4}$. To analyze the effect of the finite-dimensional approximation, we choose different values for the number of principal directions, viz. $M \in \{3, 10, 20, 30\}$. With the selected number of principal directions, we explained more than 80% of the total variability; see Table 1.

When the populations have a Gaussian distribution, the asymptotic covariance operator of the sample covariance operator

$$\hat{\Gamma}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{i,j} - \bar{X}_i) \otimes (X_{i,j} - \bar{X}_i)$$

can be estimated using the eigenvalues and eigenfunctions of $\hat{\Gamma}_i$. Taking into account that, under the null hypothesis, the processes are Gaussian, we have also used this approximation when considering the sample covariance operator.

From now on we denote as $T_{B,M}$ and $T_{B,M}^S$, for $M \in \{3, 10, 20, 30\}$, the bootstrap calibration of the statistics T_n and T_n^S , respectively, computed using M principal components. Finally, $T_{B,G}$ stands for the bootstrap calibration of T_n computed using the Gaussian approximation.

Simulation results

For the alternatives given through Models 1 and 2, Figs. 1–2 plot the frequencies of rejection at the 5% level for the procedure based on the sample covariance operator (upper panels) and for the test based on the sample spatial sign operator (lower panels) when considering the uncontaminated and contaminated samples.

As noted in [4], when using the Gaussian approximation, the test based on the sample covariance operators shows an improvement in size for uncontaminated samples. However, when the data are contaminated, the level breaks-down and the test becomes uninformative.

Note that when projecting the data on the first M principal components, the empirical size of the tests based on the bootstrap calibration either using the sample covariance or the spatial sign operators is quite close to the nominal one, for uncontaminated samples. To analyze the significance of the empirical size, we check whether the empirical size is significantly different from the nominal level $\alpha = 0.05$ by testing $\mathcal{H}_{0,\pi} : \pi = \alpha$ with nominal level γ , where π stands for the value such that $\pi_n \xrightarrow{P} \pi$ with π_n the empirical size of the considered test. This null hypothesis is rejected at level γ vs. $\mathcal{H}_{1,\pi} : \pi \neq \alpha$ if $\pi_n \notin [a_1(\alpha), a_2(\alpha)]$, where $a_j(\alpha) = \alpha + (-1)^j z_{\gamma/2} \{\alpha(1-\alpha)/N\}^{1/2}$ for $j \in \{1, 2\}$. If $\mathcal{H}_{0,\pi} : \pi = \alpha = 0.05$ is not rejected, the testing procedure is considered accurate, while if $\pi_n < a_1(\alpha)$ the testing procedure is conservative and when $\pi_n > a_2(\alpha)$ the test is liberal. For clean samples, both procedures are accurate with significance level $\gamma = 0.01$. In contrast, when the samples are contaminated, the test based on the sample covariance operator becomes conservative with empirical size not exceeding 0.011 for any value of M , while that based on the sign operator preserves its empirical size. See Tables 2–5 in [5] for details.

Regarding the behavior under the alternative, for uncontaminated data, the procedure based on the spatial sign operator shows a loss of power with respect to the sample covariance operator when the alternatives follow Model 1. For the Gaussian alternatives, however, the sign test has a much better performance, reaching a higher power in particular when Δ varies between 1 and 3. For both models, the test $T_{B,M}^S$ is stable for the considered contaminations, while the procedure based on the sample covariance operator shows an important loss of power, since the frequency of rejection never exceeds 0.5 or 0.2 under Models 1 and 2, respectively, for any value of the selected number of principal directions M .

It is also worth noting that, for the models and contaminations considered, the testing procedure is quite stable, both in terms of level and power, to the choice of the dimension M of the finite-dimensional approximation and so to the choice of q_n .

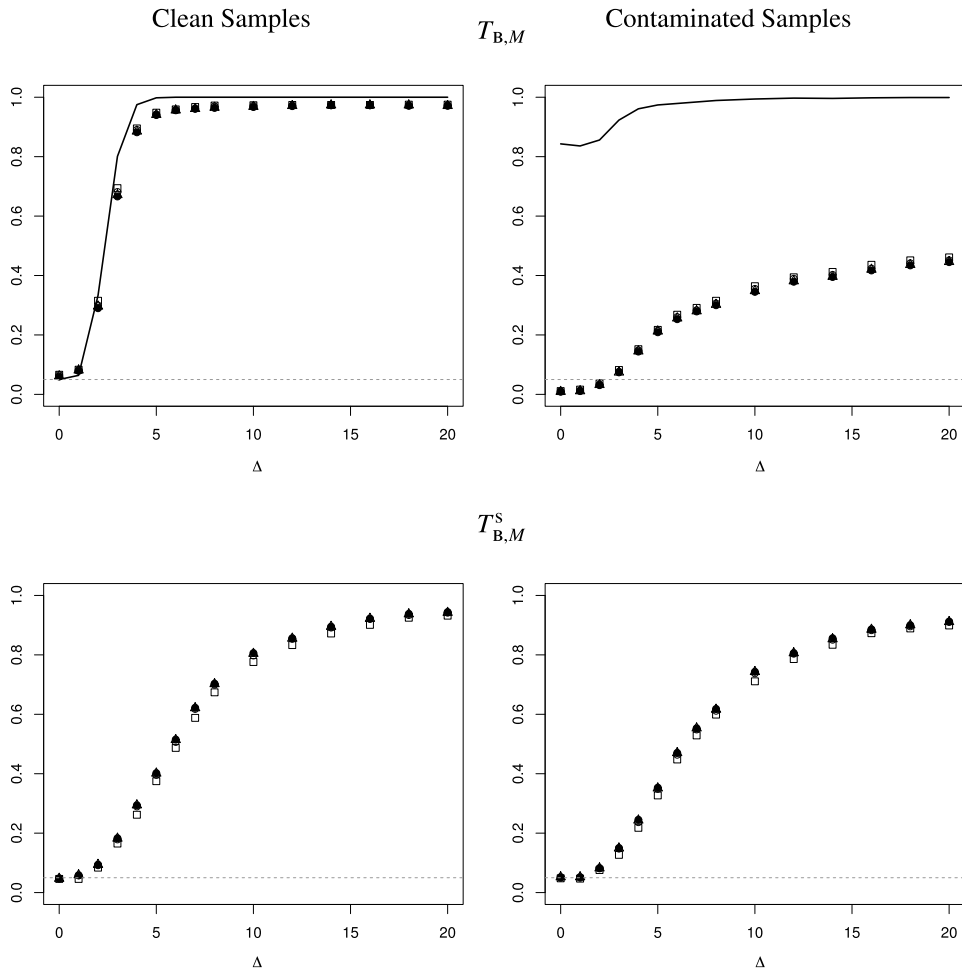


Fig. 1. Frequency of rejection, under Model 1, for the bootstrap tests $T_{B,M}$ based on the sample covariance operators (upper plots) and $T_{B,M}^S$ based on the spatial sign operator (lower plots). The solid lines correspond to $T_{B,G}$, i.e., when the eigenvalues θ_ℓ are estimated using that the processes are Gaussian. The square, circles, triangles and filled circles correspond to $M = 3, 10, 20$ and 30 , respectively.

Acknowledgments

The authors wish to thank two anonymous referees and the Editor-in-Chief, Christian Genest, for valuable comments which led to an improved version of the original paper. This research was partially supported by Grants PIP 112-201101-00742 from CONICET, PICT 2014-0351 and 201-0377 from ANPCYT, 20020130100279BA and 20020150200110BA from the Universidad de Buenos Aires at Argentina and the Spanish Project MTM2016-76969P from the Ministerio de Ciencia e Innovación at Spain.

Appendix. Proofs

Throughout this section, without loss of generality, we will assume that $\mu = 0$. Furthermore, we will denote as $\hat{\Gamma}_0^s = \hat{\Gamma}^s(0)$, $\Gamma_i(\hat{\mu}_n) = s(X_i - \hat{\mu}_n) \otimes s(X_i - \hat{\mu}_n)$ and $\Gamma_i = s(X_i) \otimes s(X_i)$.

Proof of Theorem 1. Note that the Strong Law of Large Numbers entails that it is enough to prove that $\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}_0^s \xrightarrow{as} 0$. Consider the random set $\mathcal{A}_n = \{x \in \mathcal{H} : \|x - \hat{\mu}_n\| \geq \|x\|/2\}$. We have

$$\|\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}_0^s\|_{\mathcal{F}} \leq \frac{1}{n} \sum_{X_i \in \mathcal{A}_n} \|\Gamma_i(\hat{\mu}_n) - \Gamma_i\|_{\mathcal{F}} + \frac{1}{n} \sum_{X_i \notin \mathcal{A}_n} \|\Gamma_i(\hat{\mu}_n) - \Gamma_i\|_{\mathcal{F}} = A_{n,1} + A_{n,2} \quad (\text{A.1})$$

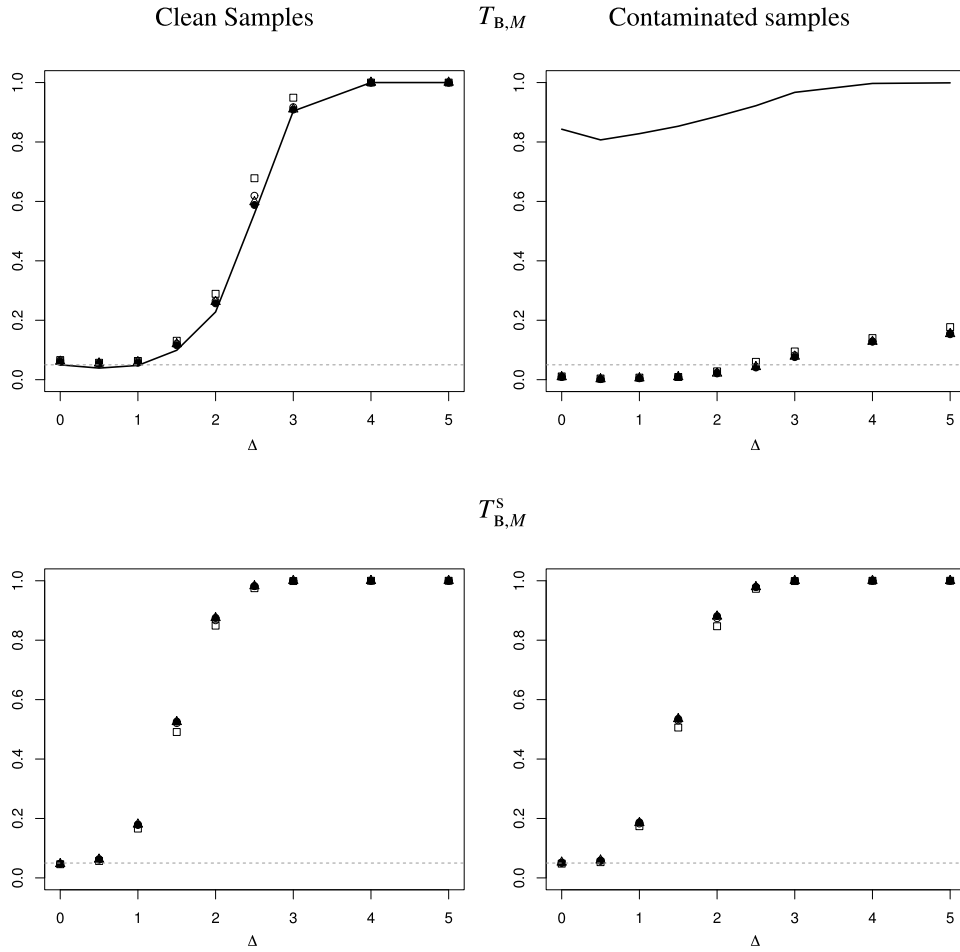


Fig. 2. Frequency of rejection for the bootstrap tests $T_{B,M}$ based on the sample covariance operators (upper plots) and $T_{B,M}^S$ based on the spatial sign operator (lower plots). The solid lines correspond to $T_{B,G}$, i.e., when the eigenvalues θ_ℓ are estimated using that the processes are Gaussian. The square, circles, triangles and filled circles correspond to $M = 3, 10, 20$ and 30 , respectively. The alternatives correspond to Model 2.

To show that $A_{n,1} \xrightarrow{\text{as}} 0$, note that straightforward calculations lead to the bound

$$\| \Gamma_i(\hat{\mu}_n) - \Gamma_i \|_{\mathcal{F}}^2 = \frac{2}{\|X_i\|^2 \|X_i - \hat{\mu}_n\|^2} (\|X_i\|^2 \|\hat{\mu}_n\|^2 - \langle \hat{\mu}_n, X_i \rangle^2) \leq \frac{4\|X_i\|^2 \|\hat{\mu}_n\|^2}{\|X_i\|^2 \|X_i - \hat{\mu}_n\|^2}.$$

Note that if $X_i \in \mathcal{A}_n$, we have $\| \Gamma_i(\hat{\mu}_n) - \Gamma_i \|_{\mathcal{F}}^2 \leq 16 \|\hat{\mu}_n\|^2 / \|X_i\|^2$, which implies that

$$\frac{1}{n} \sum_{X_i \in \mathcal{A}_n} \| \Gamma_i(\hat{\mu}_n) - \Gamma_i \|_{\mathcal{F}} \leq 4 \|\hat{\mu}_n\| \frac{1}{n} \sum_{i=1}^n \frac{1}{\|X_i\|}.$$

Therefore, using that $\hat{\mu}_n \xrightarrow{\text{as}} 0$, $E(\|X\|^{-1}) < \infty$ and the Strong Law of Large Numbers we conclude that $A_{n,1} \xrightarrow{\text{as}} 0$.

It remains to show that the second term $A_{n,2}$ in the right-hand side of (A.1) converges almost surely to zero. The fact that $\| \Gamma_i(\hat{\mu}_n) \|_{\mathcal{F}} = \| \Gamma_i \|_{\mathcal{F}} = 1$ implies that

$$A_{n,2} = \frac{1}{n} \sum_{X_i \notin \mathcal{A}_n} \| \Gamma_i(\hat{\mu}_n) - \Gamma_i \|_{\mathcal{F}} \leq \frac{2}{n} \sum_{i=1}^n Z_{n,i},$$

where $Z_{n,i} = \mathbf{1}_{\mathcal{A}_n^c}(X_i)$.

Note that the assumption $E(\|X\|^{-1}) < \infty$ implies that $\Pr(\|X\| = 0) = 0$. Hence, for any $\epsilon > 0$, let $\delta > 0$ be such that $\Pr(\|X\| \leq \delta) \leq \epsilon$ and denote $Z_{\delta,i} = \mathbf{1}_{B_\delta}(X_i)$, where $B_\delta = \{\|x\| \leq \delta\}$. Then,

$$\frac{1}{n} \sum_{i=1}^n Z_{n,i} \leq \frac{1}{n} \sum_{i=1}^n Z_{\delta,i} + \frac{1}{n} \sum_{i=1}^n (Z_{n,i} - Z_{\delta,i})_+ = B_{n,1} + B_{n,2},$$

where $a_+ = \max(a, 0)$. The Strong Law of Large Numbers entails that $B_{n,1} \xrightarrow{\text{as}} \Pr(\|X\| \leq \delta) \leq \epsilon$. To show that $B_{n,2} \xrightarrow{\text{as}} 0$, note that $\{\|\hat{\mu}_n\| \leq \delta/2\} \subset \{(Z_{n,i} - Z_{\delta,i})_+ = 0\}$. Hence, using that $\hat{\mu}_n \xrightarrow{\text{as}} 0$, we get that there exists a null probability set \mathcal{N} such that for $\omega \notin \mathcal{N}$, there exists n_0 such that, for every integer $n > n_0$, $\|\hat{\mu}_n\| \leq \delta/2$ implying that

$$B_{n,2} = \frac{1}{n} \sum_{i=1}^n (Z_{n,i} - Z_{\delta,i})_+ = 0$$

and concluding the proof. \square

Proof of Theorem 2. Note that $\Gamma_i(\hat{\mu}_n) - \Gamma_i$ can be written as follows

$$\begin{aligned} \Gamma_i(\hat{\mu}_n) - \Gamma_i &= \|X_i\|^{-2} \{\|X_i\|^2 \Gamma_i(\hat{\mu}_n) - X_i \otimes X_i\} \\ &= \|X_i\|^{-2} \{(\|X_i - \hat{\mu}_n\|^2 + \|\hat{\mu}_n\|^2 + 2\langle X_i - \hat{\mu}_n, \hat{\mu}_n \rangle) \Gamma_i(\hat{\mu}_n) - X_i \otimes X_i\} \\ &= \|X_i\|^{-2} \{\hat{\mu}_n \otimes \hat{\mu}_n - \hat{\mu}_n \otimes X_i - X_i \otimes \hat{\mu}_n + (2\langle X_i, \hat{\mu}_n \rangle - \|\hat{\mu}_n\|^2) \Gamma_i(\hat{\mu}_n)\}. \end{aligned} \quad (\text{A.2})$$

Therefore,

$$\sqrt{n} \{\hat{\Gamma}^s(\hat{\mu}_n) - \hat{\Gamma}_0^s\} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{\Gamma_i(\hat{\mu}_n) - \Gamma_i\} = S_{n,1} - S_{n,2} - S_{n,3} + 2S_{n,4} - S_{n,5},$$

where

$$\begin{aligned} S_{n,1} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\hat{\mu}_n \otimes \hat{\mu}_n}{\|X_i\|^2} = n (\hat{\mu}_n \otimes \hat{\mu}_n) \left(\frac{1}{n^{3/2}} \sum_{i=1}^n \frac{1}{\|X_i\|^2} \right), \\ S_{n,2} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\hat{\mu}_n \otimes X_i}{\|X_i\|^2} = \sqrt{n} \hat{\mu}_n \otimes \left(\frac{1}{n} \sum_{i=1}^n \frac{X_i}{\|X_i\|^2} \right), \\ S_{n,3} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i \otimes \hat{\mu}_n}{\|X_i\|^2} = \frac{1}{n} \sum_{i=1}^n \frac{X_i}{\|X_i\|^2} \otimes \sqrt{n} \hat{\mu}_n, \\ S_{n,4} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^2} \Gamma_i(\hat{\mu}_n), \\ S_{n,5} &= \|\hat{\mu}_n\|^2 \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\Gamma_i(\hat{\mu}_n)}{\|X_i\|^2} = n \|\hat{\mu}_n\|^2 \frac{1}{n^{3/2}} \sum_{i=1}^n \frac{\Gamma_i(\hat{\mu}_n)}{\|X_i\|^2}. \end{aligned}$$

Note that (A.2) entails that $EV_i^{2/3} < \infty$ where $V_i = 1/\|X_i\|^2$, so Marcinkiewicz's Strong Law of Large Numbers implies that

$$\frac{1}{n^{3/2}} \sum_{i=1}^n \frac{1}{\|X_i\|^2} \xrightarrow{\text{as}} 0. \quad (\text{A.3})$$

Hence, Assumptions (A.1) and (A.2) together with the Strong Law of Large Numbers and the fact that $\|\Gamma_i(\hat{\mu}_n)\|_{\mathcal{F}} = 1$ entail that $S_{n,j} \xrightarrow{p} 0$ for $j \in \{1, 5\}$.

The decomposition of $\Gamma_i(\hat{\mu}_n) - \Gamma_i$ obtained in (A.2) entails that $S_{n,4}$ can be written as $S_{n,4} = S_{n,41} + S_{n,42} - S_{n,43} - S_{n,44} + S_{n,45} - S_{n,46}$, where

$$\begin{aligned} S_{n,41} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} X_i \otimes X_i, & S_{n,42} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} \hat{\mu}_n \otimes \hat{\mu}_n, \\ S_{n,43} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} \hat{\mu}_n \otimes X_i, & S_{n,44} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} X_i \otimes \hat{\mu}_n, \\ S_{n,45} &= \frac{2}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} \langle X_i, \hat{\mu}_n \rangle \Gamma_i(\hat{\mu}_n), & S_{n,46} &= \|\hat{\mu}_n\|^2 \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\langle X_i, \hat{\mu}_n \rangle}{\|X_i\|^4} \Gamma_i(\hat{\mu}_n). \end{aligned}$$

Using again Marcinkiewicz's Strong Law of Large Numbers, we get that

$$\frac{1}{n^2} \sum_{i=1}^n \frac{1}{\|X_i\|^3} \xrightarrow{\text{as}} 0,$$

since $E\|X_i\|^{-3/2} < \infty$ by (A.2). Hence, (A.1) and $\|F_i(\hat{\mu}_n)\|_{\mathcal{F}} = 1$ entail that $S_{n4j} \xrightarrow{p} 0$, for $j \in \{2, 6\}$. Moreover, using (A.3), we obtain that $S_{n4j} \xrightarrow{p} 0$ for $j \in \{3, 4, 5\}$.

It remains to study the asymptotic behavior of $S_{n,2}$, $S_{n,3}$ and S_{n41} . We will show that

$$S_{n41} - \sqrt{n} F_X(\hat{\mu}_n - \mu) = o_{\text{Pr}}(1), \quad (\text{A.4})$$

$$S_{n,2} + S_{n,3} - \sqrt{n} S_X(\hat{\mu}_n - \mu) = o_{\text{Pr}}(1). \quad (\text{A.5})$$

To derive (A.4), denote as $W_i : \mathcal{H} \rightarrow \mathcal{F}$ the random objects in \mathcal{B} defined as $W_i(u) = (\langle X_i, u \rangle / \|X_i\|^4) X_i \otimes X_i$ for $u \in \mathcal{H}$. It is easy to see that $\|W_i\|_{\mathcal{B}} \leq \|X_i\|^{-1}$ and Assumption (A.2) guarantee that $E(\|X\|^{-1}) < \infty$, hence the Strong Law of Large Numbers on \mathcal{B} allows to conclude that

$$\frac{1}{n} \sum_{i=1}^n \frac{\langle X_i, \cdot \rangle}{\|X_i\|^4} X_i \otimes X_i \xrightarrow{\text{as}} F_X,$$

where F_X is defined in (4), which together with (A.1) concludes the proof of (A.4).

To obtain (A.5), note that the Strong Law of Large Numbers on \mathcal{H} and the fact that $E\|X_i\|^{-1} < \infty$ imply that

$$\frac{1}{n} \sum_{i=1}^n \frac{X_i}{\|X_i\|^2} \xrightarrow{\text{as}} E\left(\frac{X}{\|X\|^2}\right).$$

Thus, if we define a sequence $\mathcal{T}_1, \mathcal{T}_2, \dots$ of random objects in \mathcal{B} as

$$\mathcal{T}_n(u) = u \otimes \frac{1}{n} \sum_{i=1}^n \frac{X_i}{\|X_i\|^2} + \frac{1}{n} \sum_{i=1}^n \frac{X_i}{\|X_i\|^2} \otimes u$$

for any $u \in \mathcal{H}$, we obtain that $\mathcal{T}_n \xrightarrow{\text{as}} S_X$, where S_X is defined in (4). Hence, using (A.1), we obtain (A.4), concluding the proof. \square

Proof of Corollary 1. Note that from Theorem 2 we get that

$$\sqrt{n} \{\widehat{\Gamma}^s(\hat{\mu}_n) - \Gamma^s(\mu)\} = \sqrt{n} \{\widehat{\Gamma}^s(\mu) - \Gamma^s(\mu)\} + \sqrt{n} G_X(\hat{\mu}_n - \mu) + o_{\text{Pr}}(1).$$

Now, the results follow immediately defining, for any fixed $v \in \mathcal{H}$, the operators $R_v : \mathcal{H} \rightarrow \mathcal{F}$ and $L_v : \mathcal{H} \rightarrow \mathcal{F}$ as $R_v(u) = u \otimes v$ and $L_v(u) = v \otimes u$ and using that $R_v^*(\gamma) = \gamma(v)$ and $L_v^*(\gamma) = \gamma^*(v)$. \square

References

- [1] J. Bali, G. Boente, Principal points and elliptical distributions from the multivariate setting to the functional case, *Statist. Probab. Lett.* 79 (2009) 1858–1865.
- [2] J. Bali, G. Boente, D. Tyler, J. Wang, Robust functional principal components: A projection-pursuit approach, *Ann. Statist.* 39 (2011) 2852–2882.
- [3] M. Benko, W.K. Härdle, A. Kneip, Common functional principal components, *Ann. Statist.* 37 (2009) 1–34.
- [4] G. Boente, D. Rodriguez, M. Sued, Testing equality of several covariance operators, *Ann. Inst. Statist. Math.* 70 (2018) 19–950.
- [5] G. Boente, D. Rodriguez, M. Sued, The spatial sign covariance operator: Asymptotic results and applications, 2018. Available at <http://arxiv.org/abs/1804.04210>.
- [6] G. Boente, M. Salibián-Barrera, S-estimators for functional principal component analysis, *J. Amer. Statist. Assoc.* 110 (2015) 1100–1111.
- [7] G. Boente, M. Salibián-Barrera, D. Tyler, A characterization of elliptical distributions and some optimality properties of principal components for functional data, *J. Multivariate Anal.* 131 (2014) 254–264.
- [8] H. Cardot, P. Cénac, P. Zitt, Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm, *Bernoulli* 19 (2013) 18–431.
- [9] A. Cuevas, A partial overview of the theory of statistics with functional data, *J. Statist. Plann. Inference* 147 (2014) 1–23.
- [10] A. Cuevas, M. Febrero, R. Fraiman, Robust estimation and classification for functional data via projection-based depth notions, *Comput. Statist.* 22 (2007) 481–496.
- [11] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a vector random function: Some applications to statistical inference, *J. Multivariate Anal.* 12 (1982) 136–154.
- [12] A. Dürre, D. Tyler, D. Vogel, On the eigenvalues of the spatial sign covariance matrix in more than two dimensions, *Statist. Probab. Lett.* 111 (2016) 80–85.
- [13] A. Dürre, D. Vogel, D. Tyler, The spatial sign covariance matrix with unknown location, *J. Multivariate Anal.* 130 (2014) 107–117.
- [14] F. Ferraty, Y. Romain, *The Oxford Handbook of Functional Data Analysis*, Oxford University Press, 2010.
- [15] F. Ferraty, P. Vieu, *Nonparametric Functional Data Analysis: Theory and Practice*, Springer, New York, 2006.
- [16] F. Ferraty, P. Vieu, S. Viguier-Pla, Factor-based comparison of groups of curves, *Comput. Statist. Data Anal.* 51 (2007) 4903–4910.
- [17] R. Fraiman, G. Muñiz, Trimmed means for functional data, *Test* 10 (2001) 419–440.

- [18] S. Fremdt, J. Steinbach, L. Horváth, P. Kokoszka, Testing the equality of covariance operators in functional samples, *Scand. J. Stat.* 40 (2013) 138–152.
- [19] D. Gervini, Robust functional estimation using the median and spherical principal components, *Biometrika* 95 (2008) 587–600.
- [20] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics, *J. Multivariate Anal.* 146 (2016) 1–6.
- [21] P. Hall, J. Horowitz, Methodology and convergence rates for functional linear regression, *Ann. Statist.* 35 (2007) 70–91.
- [22] L. Horváth, P. Kokoszka, *Inference for Functional Data with Applications*, Springer, New York, 2012.
- [23] T. Hsing, R. Eubank, *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*, Wiley, New York, 2015.
- [24] D. Kraus, V. Panaretos, Dispersion operators and resistant second-order functional data analysis, *Biometrika* 99 (2012) 813–832.
- [25] S. Lee, H. Shin, N. Billor, M -type smoothing splines estimators for principal functions, *Comput. Statist. Data Anal.* 66 (2013) 89–100.
- [26] H. Lian, Functional partial linear model, *J. Nonparametr. Stat.* 23 (2011) 115–128.
- [27] N. Locantore, J. Marron, D. Simpson, N. Tripoli, J. Zhang, K. Cohen, Robust principal components for functional data, *Test* 8 (1999) 1–28.
- [28] S. López-Pintado, J. Romo, Depth-based inference for functional data, *Comput. Statist. Data Anal.* 51 (2007) 4957–4968.
- [29] V. Panaretos, D. Kraus, J. Maddocks, Second-order comparison of Gaussian random functions and the geometry of DNA minicircles, *J. Amer. Statist. Assoc.* 105 (2010) 670–682.
- [30] D. Pigoli, J. Aston, I. Dryden, P. Secchi, Distances and inference for covariance operators, *Biometrika* 101 (2014) 409–422.
- [31] J.O. Ramsay, B.W. Silverman, *Functional Data Analysis*, Springer, Berlin, 2005.
- [32] P. Sawant, N. Billor, H. Shin, Functional outlier detection with robust functional principal component analysis, *Comput. Statist.* 27 (2012) 83–102.