

Paired sample tests in infinite dimensional spaces

Anirvan Chakraborty and Probal Chaudhuri

Theoretical Statistics and Mathematics Unit,

Indian Statistical Institute

203, B. T. Road, Kolkata - 700108, INDIA.

emails: vanchak@gmail.com, probal@isical.ac.in

Abstract

The sign and the signed-rank tests for univariate data are perhaps the most popular non-parametric competitors of the t test for paired sample problems. These tests have been extended in various ways for multivariate data in finite dimensional spaces. These extensions include tests based on spatial signs and signed ranks, which have been studied extensively by Hannu Oja and his coauthors. They showed that these tests are asymptotically more powerful than Hotelling's T^2 test under several heavy tailed distributions. In this paper, we consider paired sample tests for data in infinite dimensional spaces based on notions of spatial sign and spatial signed rank in such spaces. We derive their asymptotic distributions under the null hypothesis and under sequences of shrinking location shift alternatives. We compare these tests with some mean based tests for infinite dimensional paired sample data. We show that for shrinking location shift alternatives, the proposed tests are asymptotically more powerful than the mean based tests for some heavy tailed distributions and even for some Gaussian distributions in infinite dimensional spaces. We also investigate the performance of different tests using some simulated data.

Keywords: contaminated data, Gâteaux derivative, paired sample problem, spatial sign, spatial signed rank, smooth Banach space, t process.

1 Introduction

For univariate data, two nonparametric competitors of the t test for one sample and paired sample problems are the sign test and the Wilcoxon signed-rank test. It is well known that these two tests enjoy certain robustness properties, and are more powerful than the t test when the underlying distribution has heavier tails than the Gaussian distribution. The sign and the signed-rank tests have been extended in several ways for data in \mathbb{R}^d . [Puri and Sen \(1971\)](#) considered extensions based on coordinatewise signs and ranks. [Randles \(1989\)](#) and [Peters and Randles \(1990\)](#) studied

extensions of the sign test and the signed-rank test, respectively, using the notion of interdirections. Chaudhuri and Sengupta (1993) proposed a class of multivariate extensions of the sign test based on data-driven transformations of the sample observations. Hallin and Paindaveine (2002) considered multivariate signed-rank type tests based on interdirections and the ranks of the sample observations computed using pseudo-Mahalanobis distances. Hettmansperger et al. (1994) and Hettmansperger et al. (1997) considered multivariate versions of sign and signed-rank tests using simplices (see also Oja (1999)). Möttönen and Oja (1995), Chakraborty et al. (1998) and Marden (1999) used spatial signs and signed ranks (see also Oja (2010)) to construct extensions of sign and signed-rank tests for multivariate data. Möttönen et al. (1997) and Chakraborty et al. (1998) studied the asymptotic efficiency of spatial sign and signed-rank tests relative to Hotelling's T^2 test. They showed that the tests based on spatial signs and signed ranks are asymptotically more powerful than the T^2 test under heavy tailed distributions like multivariate t distributions. Further, while Hotelling's T^2 test is optimal for multivariate Gaussian distributions, it was shown that as the data dimension increases, the performance of the spatial sign and the spatial signed-rank tests become closer to Hotelling's T^2 test under multivariate Gaussian distributions.

Nowadays, we often come across data, which are curves or functions observed over an interval, and are popularly known as functional data. Such data are very different from multivariate data in finite dimensional spaces because the data dimension is much larger than the sample size, and also due to the fact that different sample observations may be observed at different sets of points in the interval. However, this type of data can be conveniently handled by viewing them as a sample in some infinite dimensional space, e.g., the space of real-valued functions defined over an interval. Many of the above-mentioned tests cannot be used to analyze such data. This is because the definitions of some of them involve hyperplanes, simplices etc. constructed using the data, and thus these tests require the data dimension to be smaller than the sample size. Some of the other tests involve inverses of covariance matrices computed from the sample, and such empirical covariance matrices are singular, when the data dimension is larger than the sample size.

Many of the function spaces, where functional data lie, are infinite dimensional Banach spaces. In this paper, we investigate a sign and a signed-rank type test for paired sample problems based on notions of spatial sign and spatial signed rank in such spaces. We derive the asymptotic distributions of the proposed spatial sign and signed-rank statistics under the null hypothesis as well as under suitable sequences of shrinking location shift alternatives. We also compare the asymptotic powers of these tests with some of the paired-sample mean based tests for infinite dimensional data. It is found that these tests outperform the mean based competitors, when the underlying distribution has heavy tails and also for some Gaussian distributions of the data. Some simulation studies are carried out to demonstrate the performance of different tests.

2 Sign and Signed-rank type tests

The construction and the study of the paired sample sign and signed-rank type statistics for data in general Banach spaces (finite or infinite dimensional) will require several concepts and tools from functional analysis and probability theory in Banach spaces. Let \mathcal{X} be a Banach space with norm $\|\cdot\|$. Let us denote its dual space by \mathcal{X}^* , which is the Banach space of real-valued continuous linear functions defined over \mathcal{X} .

Definition 1. *The norm in \mathcal{X} is said to be Gâteaux differentiable at a nonzero $\mathbf{x} \in \mathcal{X}$ with derivative, say, $\mathbf{S}_{\mathbf{x}} \in \mathcal{X}^*$ if $\lim_{t \rightarrow 0} t^{-1}(\|\mathbf{x} + t\mathbf{h}\| - \|\mathbf{x}\|) = \mathbf{S}_{\mathbf{x}}(\mathbf{h})$ for all $\mathbf{h} \in \mathcal{X}$. A Banach space \mathcal{X} is said to be smooth if the norm in \mathcal{X} is Gâteaux differentiable at every nonzero $\mathbf{x} \in \mathcal{X}$.*

As a convention, we take $\mathbf{S}_{\mathbf{x}} = \mathbf{0}$ if $\mathbf{x} = \mathbf{0}$. Hilbert spaces and L_p spaces for $p \in (1, \infty)$ are smooth. For a Hilbert space \mathcal{X} , we have $\mathbf{S}_{\mathbf{x}} = \mathbf{x}/\|\mathbf{x}\|$. If \mathcal{X} is a L_p space for some $p \in (1, \infty)$, then $\mathbf{S}_{\mathbf{x}}(\mathbf{h}) = \int \text{sign}\{\mathbf{x}(\mathbf{s})\}|\mathbf{x}(\mathbf{s})|^{p-1}\mathbf{h}(\mathbf{s})d\mathbf{s}/\|\mathbf{x}\|^{p-1}$ for all $\mathbf{h} \in \mathcal{X}$.

Definition 2. *A random element \mathbf{X} in the Banach space \mathcal{X} is said to be Bochner integrable if there exists a sequence $\{\mathbf{X}_n\}_{n \geq 1}$ of simple functions in \mathcal{X} such that $\mathbf{X}_n \rightarrow \mathbf{X}$ almost surely and $E(\|\mathbf{X}_n - \mathbf{X}\|) \rightarrow 0$ as $n \rightarrow \infty$.*

It is known that the Bochner expectation of \mathbf{X} exists if $E(\|\mathbf{X}\|) < \infty$. We refer to Chapters 4 and 5 in [Borwein and Vanderwerff \(2010\)](#) and Sect. 2 in Chapter 3 of [Araujo and Giné \(1980\)](#) for more details. Henceforth, the expectation of any Banach space valued random element will be in the Bochner sense. The spatial sign of $\mathbf{x} \in \mathcal{X}$ is given by $\mathbf{S}_{\mathbf{x}}$, and its spatial rank with respect to the distribution of a random element $\mathbf{X} \in \mathcal{X}$ is defined as $\Psi_{\mathbf{x}} = E(\mathbf{S}_{\mathbf{x} - \mathbf{X}})$. If \mathcal{X} is a Hilbert space, then it follows that $\Psi_{\mathbf{x}} = E\{(\mathbf{x} - \mathbf{X})/\|\mathbf{x} - \mathbf{X}\|\}$. In particular, if $\mathcal{X} = \mathbb{R}^d$ equipped with the Euclidean norm, $\mathbf{S}_{\mathbf{x}}$ and $\Psi_{\mathbf{x}}$ reduce to the usual multivariate spatial sign and spatial rank in \mathbb{R}^d (see [Oja \(2010\)](#)).

Let $(\mathbf{X}_1, \mathbf{Y}_1), (\mathbf{X}_2, \mathbf{Y}_2) \dots, (\mathbf{X}_n, \mathbf{Y}_n)$ be i.i.d. paired observations, where the \mathbf{X}_i 's and the \mathbf{Y}_i 's take values in a smooth Banach space \mathcal{X} . Let $\mathbf{W}_i = \mathbf{Y}_i - \mathbf{X}_i$, $1 \leq i \leq n$, and define $\boldsymbol{\nu} = E(\mathbf{S}_{\mathbf{W}_1})$. A paired sample sign statistic using spatial signs for testing the hypothesis $H_0^{(1)} : \boldsymbol{\nu} = \mathbf{0}$ against $H_1^{(1)} : \boldsymbol{\nu} \neq \mathbf{0}$ is defined as

$$\mathbf{T}_S = n^{-1} \sum_{1 \leq i \leq n} \mathbf{S}_{\mathbf{W}_i}.$$

We reject $H_0^{(1)}$ when $\|\mathbf{T}_S\|$ is large. Next, define $\boldsymbol{\theta} = E(\mathbf{S}_{\mathbf{W}_1 + \mathbf{W}_2})$. A paired sample signed-rank statistic using spatial signed ranks for testing the hypothesis $H_0^{(2)} : \boldsymbol{\theta} = \mathbf{0}$ against $H_1^{(2)} : \boldsymbol{\theta} \neq \mathbf{0}$ is given by

$$\mathbf{T}_{SR} = 2\{n(n-1)\}^{-1} \sum_{1 \leq i < j \leq n} \mathbf{S}_{\mathbf{W}_i + \mathbf{W}_j}.$$

We reject $H_0^{(2)}$ for large values of $\|\mathbf{T}_{SR}\|$. Recently, a two sample Wilcoxon–Mann–Whitney type test based on spatial ranks in infinite dimensional spaces have been studied by [Chakraborty and Chaudhuri \(2014\)](#). Note that if $\mathcal{X} = \mathbb{R}$, $\mathbf{S}_x = \text{sign}(x)$. Thus, \mathbf{T}_S and \mathbf{T}_{SR} reduce to the univariate sign and signed-rank statistics if $\mathcal{X} = \mathbb{R}$. Moreover, if $\mathcal{X} = \mathbb{R}^d$, then \mathbf{T}_S and \mathbf{T}_{SR} are the spatial sign and signed-rank statistics for finite dimensional multivariate data studied by [Möttönen and Oja \(1995\)](#), [Möttönen et al. \(1997\)](#) and [Marden \(1999\)](#). Note that the hypothesis $H_0^{(1)} : \boldsymbol{\nu} = \mathbf{0}$ (respectively, $H_0^{(2)} : \boldsymbol{\theta} = \mathbf{0}$) is equivalent to the hypothesis that the spatial median of \mathbf{W}_1 (respectively, $\mathbf{W}_1 + \mathbf{W}_2$) is zero. Suppose that $\mathbf{Y}_1 - \mathbf{X}_1$ has a symmetric distribution about some $\boldsymbol{\eta} \in \mathcal{X}$, i.e., the distribution of $\mathbf{Y}_1 - \mathbf{X}_1 - \boldsymbol{\eta}$ and $\boldsymbol{\eta} - \mathbf{Y}_1 + \mathbf{X}_1$ are the same. Then, it follows that both of $H_0^{(1)}$ and $H_0^{(2)}$ become the hypothesis $\boldsymbol{\eta} = \mathbf{0}$. This holds, in particular, if \mathbf{X}_1 and \mathbf{Y}_1 are exchangeable, i.e., the distributions of $(\mathbf{X}_1, \mathbf{Y}_1)$ and $(\mathbf{Y}_1, \mathbf{X}_1)$ are the same. Further, if the distribution of $\mathbf{Y}_1 - \mathbf{X}_1$ is symmetric and its mean exists, then both of $H_0^{(1)}$ and $H_0^{(2)}$ are equivalent to the hypothesis $E(\mathbf{Y}_1 - \mathbf{X}_1) = \mathbf{0}$.

2.1 Asymptotic distributions and the implementation of the tests

In order to study the asymptotic distributions of \mathbf{T}_S and \mathbf{T}_{SR} , we introduce the following definitions.

Definition 3. A Banach space \mathcal{X} is said to be of type 2 if there exists a constant $b > 0$ such that for any $m \geq 1$ and independent zero mean random elements $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_m$ in \mathcal{X} with $E(\|\mathbf{U}_i\|^2) < \infty$, $1 \leq i \leq m$, we have $E(\|\sum_{i=1}^m \mathbf{U}_i\|^2) \leq b \sum_{i=1}^m E(\|\mathbf{U}_i\|^2)$.

Definition 4. A Banach space \mathcal{X} is said to be p -uniformly smooth for some $p \in (1, 2]$ if for every $q \geq 1$ there exists a constant $\alpha_q > 0$ such that for any zero mean martingale sequence $(\mathbf{M}_m, \mathcal{G}_m)_{m \geq 1}$ in \mathcal{X} , we have $E(\|\mathbf{M}_m\|^q) \leq \alpha_q \sum_{i=1}^m E(\|\mathbf{M}_i - \mathbf{M}_{i-1}\|^p)^{q/p}$. Here, the sequence $(\mathbf{M}_m)_{m \geq 1}$ is adapted to the filtration $(\mathcal{G}_m)_{m \geq 1}$.

Any 2-uniformly smooth Banach space is of type 2. Hilbert spaces are 2-uniformly smooth, and L_p spaces are \tilde{p} -uniformly smooth, where $\tilde{p} = \min(p, 2)$ for $p \in (1, \infty)$.

Definition 5. A continuous linear operator $\mathbf{C} : \mathcal{X}^* \rightarrow \mathcal{X}^*$ is said to be symmetric if $\mathbf{y}\{\mathbf{C}(\mathbf{x})\} = \mathbf{x}\{\mathbf{C}(\mathbf{y})\}$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}^*$. It is said to be positive if $\mathbf{x}\{\mathbf{C}(\mathbf{x})\} > 0$ for all $\mathbf{x} \in \mathcal{X}^*$.

Definition 6. A random element \mathbf{Z} in a separable Banach space \mathcal{X} is said to have a Gaussian distribution with mean $\mathbf{m} \in \mathcal{X}$ and covariance \mathbf{C} , which we denote by $G(\mathbf{m}, \mathbf{C})$, if for any $\mathbf{u} \in \mathcal{X}^*$, $\mathbf{u}(\mathbf{Z})$ has a Gaussian distribution on \mathbb{R} with mean $\mathbf{u}(\mathbf{m})$ and variance $\mathbf{u}\{\mathbf{C}(\mathbf{u})\}$. Here, $\mathbf{C} : \mathcal{X}^* \rightarrow \mathcal{X}^*$ is a symmetric positive continuous linear operator.

We refer to Sect. 7 of Chapter 3 in [Araujo and Giné \(1980\)](#), [Borovskikh \(1991\)](#) and Sect. 2.4 in Chapter IV of [Vakhania et al. \(1987\)](#) for further details. Define $\mathbf{\Pi}_1 : \mathcal{X}^{**} \rightarrow \mathcal{X}^*$ and $\mathbf{\Pi}_2 : \mathcal{X}^{**} \rightarrow \mathcal{X}^*$ as

$$\begin{aligned} \mathbf{\Pi}_1(\mathbf{f}) &= E[\mathbf{f}(\mathbf{S}\mathbf{w}_1)\mathbf{S}\mathbf{w}_1] - \{\mathbf{f}(\boldsymbol{\nu})\}\boldsymbol{\nu}, \quad \text{and} \\ \mathbf{\Pi}_2(\mathbf{f}) &= 4(E[\mathbf{f}\{E(\mathbf{S}\mathbf{w}_1 + \mathbf{w}_2 | \mathbf{W}_1)\}E(\mathbf{S}\mathbf{w}_1 + \mathbf{w}_2 | \mathbf{W}_1)] - \{\mathbf{f}(\boldsymbol{\theta})\}\boldsymbol{\theta}), \end{aligned}$$

where $\mathbf{f} \in \mathcal{X}^{**}$. So, $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ are symmetric positive continuous linear operators. For Banach space valued random elements \mathbf{U} and \mathbf{V} defined on the same probability space with \mathbf{U} having finite Bochner expectation, the conditional expectation of \mathbf{U} given \mathbf{V} exists and can be properly defined (see Vakhania et al. (1987, pp. 125 – 128)). Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq n$, be i.i.d. paired observations with the \mathbf{X}_i 's and the \mathbf{Y}_i 's taking values in a smooth Banach space \mathcal{X} . The next theorem gives the asymptotic distributions of \mathbf{T}_S and \mathbf{T}_{SR} .

Theorem 1. *Suppose that the dual space \mathcal{X}^* is a separable and type 2 Banach space. Then, for any probability measure P on \mathcal{X} , $n^{1/2}(\mathbf{T}_S - \boldsymbol{\nu})$ converges weakly to a Gaussian limit $G(\mathbf{0}, \mathbf{\Pi}_1)$ as $n \rightarrow \infty$. Further, if \mathcal{X}^* is p -uniformly smooth for some $p \in (4/3, 2]$, we have weak convergence of $n^{1/2}(\mathbf{T}_{SR} - \boldsymbol{\theta})$ to a Gaussian limit $G(\mathbf{0}, \mathbf{\Pi}_2)$ as $n \rightarrow \infty$.*

Proof. Using the central limit theorem for i.i.d. random elements in a separable and type 2 Banach space (see Araujo and Giné (1980, Thm. 7.5(i))), we get that $n^{1/2}(\mathbf{T}_S - \boldsymbol{\nu})$ converges weakly to $G(\mathbf{0}, \mathbf{\Pi}_1)$.

Note that $\mathbf{T}_{SR} - \boldsymbol{\theta}$ is a Banach space valued U -statistic with kernel $\mathbf{h}(\mathbf{w}_i, \mathbf{w}_j) = \mathbf{S}_{\mathbf{w}_i + \mathbf{w}_j} - \boldsymbol{\theta}$, which satisfies $E\{\mathbf{h}(\mathbf{W}_i, \mathbf{W}_j)\} = \mathbf{0}$. By the Hoeffding type decomposition for Banach space valued U -statistics (see Borovskikh (1991, p. 430)), we have

$$\mathbf{T}_{SR} - \boldsymbol{\theta} = \frac{2}{n} \sum_{i=1}^n [E\{\mathbf{S}_{\mathbf{w}_i + \mathbf{W}'} \mid \mathbf{W}_i\} - \boldsymbol{\theta}] + \mathbf{R}_n,$$

where \mathbf{W}' is an independent copy of \mathbf{W}_1 . So, $\mathbf{R}_n = 2[n(n-1)]^{-1} \sum_{1 \leq i < j \leq n} \tilde{\mathbf{h}}(\mathbf{W}_i, \mathbf{W}_j)$, where $\tilde{\mathbf{h}}(\mathbf{w}_i, \mathbf{w}_j) = \mathbf{h}(\mathbf{w}_i, \mathbf{w}_j) - E\{\mathbf{h}(\mathbf{W}_i, \mathbf{W}_j) \mid \mathbf{W}_i = \mathbf{w}_i\} - E\{\mathbf{h}(\mathbf{W}_i, \mathbf{W}_j) \mid \mathbf{W}_j = \mathbf{w}_j\}$. Note that $\|\tilde{\mathbf{h}}(\mathbf{w}_i, \mathbf{w}_j)\| \leq 4$. Using the boundedness of $\tilde{\mathbf{h}}(\cdot, \cdot)$ and Thm. 5.1 in Borovskikh (1991), it follows that for any $q \in [1, p]$,

$$E(\|n^{1/2}\mathbf{R}_n\|^q) \leq \tilde{\alpha}_q n^{2-(3q/2)} \quad (1)$$

for every $n \geq 2$ and a constant $\tilde{\alpha}_q$. Thus, if $p > 4/3$, $E(\|n^{1/2}\mathbf{R}_n\|^q)$ converges to zero as $n \rightarrow \infty$ for any $q \in (4/3, p]$. This implies that $n^{1/2}\mathbf{R}_n$ converges to zero *in probability* as $n \rightarrow \infty$.

Now, $n^{-1/2} \sum_{i=1}^n E\{\mathbf{h}(\mathbf{W}_i, \mathbf{W}') \mid \mathbf{W}_i\}$ converge weakly to $G(\mathbf{0}, \mathbf{\Pi}_2)$ as $n \rightarrow \infty$ by the central limit theorem for i.i.d. random elements in a separable type 2 Banach space (see Araujo and Giné (1980, Thm. 7.5(i))). This, together with the fact that $n^{1/2}\mathbf{R}_n$ converges to zero *in probability*, completes the proof. \square

Let $c_{1\alpha}$ and $c_{2\alpha}$ denote the $(1-\alpha)$ quantiles of the distributions of $\|G(\mathbf{0}, \mathbf{\Pi}_1)\|$ and $\|G(\mathbf{0}, \mathbf{\Pi}_2)\|$, respectively. The test based on \mathbf{T}_S rejects $H_0^{(1)}$ when $\|n^{1/2}\mathbf{T}_S\| > c_{1\alpha}$ and the test based on \mathbf{T}_{SR} rejects $H_0^{(2)}$ when $\|n^{1/2}\mathbf{T}_{SR}\| > c_{2\alpha}$.

Corollary 1. *The asymptotic sizes of these tests based on \mathbf{T}_S and \mathbf{T}_{SR} will be the same as their nominal level. Further, these tests are consistent whenever $\boldsymbol{\nu} \neq \mathbf{0}$ and $\boldsymbol{\theta} \neq \mathbf{0}$, respectively. So, if*

the distribution of $\mathbf{Y}_1 - \mathbf{X}_1$ is symmetric and its spatial median is non-zero, then these tests are consistent. In particular, these tests are consistent for location shift alternatives.

We next describe how to compute the critical value of the tests based on \mathbf{T}_S and \mathbf{T}_{SR} using their asymptotic distributions under the null hypotheses obtained in Thm. 1. Suppose that \mathcal{X} is a separable Hilbert space and \mathbf{Z} is a zero mean Gaussian random element in \mathcal{X} with covariance operator \mathbf{C} . Then, it can be shown using the spectral decomposition of the compact self-adjoint operator \mathbf{C} (see Thm. IV.2.4, and Thm. 1.3 and Cor. 2 in pp. 159-160 in Vakhania et al. (1987)) that the distribution of $\|\mathbf{Z}\|^2$ is a weighted sums of independent chi-square variables each with one degree of freedom, where the weights are the eigenvalues of the \mathbf{C} . Thus, if \mathcal{X} is a separable Hilbert space, the asymptotic distributions of $\|n^{1/2}\mathbf{T}_S\|^2$ and $\|n^{1/2}\mathbf{T}_{SR}\|^2$ are weighted sums of independent chi-square variables each with one degree of freedom, where the weights are the eigenvalues of $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$, respectively. The eigenvalues of $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ can be estimated by the eigenvalues of $\widehat{\mathbf{\Pi}}_1$ and $\widehat{\mathbf{\Pi}}_2$, which are defined as

$$\begin{aligned}\widehat{\mathbf{\Pi}}_1 &= \frac{1}{n-1} \left\{ \sum_{i=1}^n \left(\frac{\mathbf{W}_i}{\|\mathbf{W}_i\|} - \widehat{\boldsymbol{\nu}} \right) \otimes \left(\frac{\mathbf{W}_i}{\|\mathbf{W}_i\|} - \widehat{\boldsymbol{\nu}} \right) \right\}, \\ \widehat{\mathbf{\Pi}}_2 &= \frac{4}{n-1} \left\{ \sum_{i=1}^n \left(\frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mathbf{W}_i + \mathbf{W}_j}{\|\mathbf{W}_i + \mathbf{W}_j\|} - \widehat{\boldsymbol{\theta}} \right) \otimes \left(\frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n \frac{\mathbf{W}_i + \mathbf{W}_j}{\|\mathbf{W}_i + \mathbf{W}_j\|} - \widehat{\boldsymbol{\theta}} \right) \right\}.\end{aligned}$$

Here, $\mathbf{x} \otimes \mathbf{x} : \mathcal{X} \rightarrow \mathcal{X}$ is the tensor product in the Hilbert space \mathcal{X} , which is defined as $\langle (\mathbf{x} \otimes \mathbf{x})(\mathbf{f}), \mathbf{g} \rangle = \langle \mathbf{x}, \mathbf{f} \rangle \langle \mathbf{x}, \mathbf{g} \rangle$ for $\mathbf{f}, \mathbf{g}, \mathbf{x} \in \mathcal{X}$. Further, $\widehat{\boldsymbol{\nu}} = n^{-1} \sum_{i=1}^n \mathbf{W}_i / \|\mathbf{W}_i\|$ and $\widehat{\boldsymbol{\theta}} = 2\{n(n-1)\}^{-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^n (\mathbf{W}_i + \mathbf{W}_j) / \|\mathbf{W}_i + \mathbf{W}_j\|$. The critical values $c_{1\alpha}$ and $c_{2\alpha}$ can be obtained by simulating from the estimated asymptotic distributions of \mathbf{T}_S and \mathbf{T}_{SR} . On the other hand, if \mathcal{X} is a general Banach space satisfying the assumptions of Thm. 1, we no longer have the weighted chi-square representations for the asymptotic distributions of \mathbf{T}_S and \mathbf{T}_{SR} under the null hypotheses. However, we can estimate $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ by their empirical counterparts, which are defined in a similar way as the definitions above. We can simulate from the asymptotic Gaussian distributions with the estimated covariance operators to compute the critical values of the tests.

3 Asymptotic distributions under shrinking alternatives

Consider i.i.d. paired observations $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq n$, where the \mathbf{X}_i 's and the \mathbf{Y}_i 's take values in a smooth Banach space \mathcal{X} . In this section, we shall derive the asymptotic distribution of \mathbf{T}_S and \mathbf{T}_{SR} under sequences of shrinking location shift alternatives, where $\mathbf{W}_i = \mathbf{Y}_i - \mathbf{X}_i$ is symmetrically distributed about $n^{-1/2}\boldsymbol{\eta}$ for some fixed nonzero $\boldsymbol{\eta} \in \mathcal{X}$ and $1 \leq i \leq n$. For some of the finite dimensional multivariate extensions of the sign and the signed-rank tests, such alternative hypotheses have been shown to be contiguous to the null and yields nondegenerate

asymptotic distributions of the test statistics (see, e.g., [Randles \(1989\)](#); [Möttönen et al. \(1997\)](#); [Oja \(1999\)](#)).

Definition 7. *The norm in \mathcal{X} is said to be twice Gâteaux differentiable at $\mathbf{x} \neq \mathbf{0}$ with Hessian (or second order Gâteaux derivative) $\mathbf{H}_{\mathbf{x}}$, which is a continuous linear map from \mathcal{X} to \mathcal{X}^* , if $\lim_{t \rightarrow 0} t^{-1}(\mathbf{S}_{\mathbf{x}+t\mathbf{h}} - \mathbf{S}_{\mathbf{x}}) = \mathbf{H}_{\mathbf{x}}(\mathbf{h})$ for all $\mathbf{h} \in \mathcal{X}$. Here, the limit is assumed to exist in the norm topology of \mathcal{X}^* .*

Norms in Hilbert spaces and L_p spaces for $p \in [2, \infty)$ are twice Gâteaux differentiable. We refer to Chapters 4 and 5 in [Borwein and Vanderwerff \(2010\)](#) for further details. For the next theorem, let us assume that the norm in \mathcal{X} is twice Gâteaux differentiable at every $\mathbf{x} \neq \mathbf{0}$, and denote the Hessians of the functions $\mathbf{x} \mapsto E\{\|\mathbf{x} - \mathbf{W}_1\| - \|\mathbf{W}_1\|\}$ and $\mathbf{x} \mapsto E\{\|2\mathbf{x} - \mathbf{W}_1 - \mathbf{W}_2\| - \|\mathbf{W}_1 + \mathbf{W}_2\|\}$ by $\mathbf{J}_{\mathbf{x}}^{(1)}$ and $\mathbf{J}_{\mathbf{x}}^{(2)}$, respectively. The following theorem gives the asymptotic distributions of \mathbf{T}_S and \mathbf{T}_{SR} under the sequence of shrinking alternatives mentioned earlier.

Theorem 2. *Suppose that \mathcal{X}^* is a separable and type 2 Banach space. Assume that the distribution of \mathbf{W}_1 is nonatomic, and $\mathbf{J}_{\mathbf{0}}^{(1)}$ exists. Then, $n^{1/2}\mathbf{T}_S$ converges weakly to a Gaussian limit $G\{\mathbf{J}_{\mathbf{0}}^{(1)}(\boldsymbol{\eta}), \boldsymbol{\Pi}_1\}$ as $n \rightarrow \infty$. Further, if \mathcal{X}^* is a p -uniformly smooth Banach space for some $p \in (4/3, 2]$ and $\mathbf{J}_{\mathbf{0}}^{(2)}$ exists, we have weak convergence of $n^{1/2}\mathbf{T}_{SR}$ to a Gaussian limit $G\{\mathbf{J}_{\mathbf{0}}^{(1)}(\boldsymbol{\eta}), \boldsymbol{\Pi}_2\}$. Here, the expectations in the definitions of $\mathbf{J}_{\mathbf{0}}^{(1)}$ and $\mathbf{J}_{\mathbf{0}}^{(2)}$ are taken with respect to the symmetric distribution of \mathbf{W}_1 about zero under the null hypothesis.*

Proof. We first derive the asymptotic distribution of \mathbf{T}_{SR} . Let $\boldsymbol{\eta}_n = n^{-1/2}\boldsymbol{\eta}$. Applying the Hoeffding type decomposition for Banach space valued U -statistics as in the proof of Thm. 1, it follows that

$$\mathbf{T}_{SR} - \boldsymbol{\theta}(\boldsymbol{\eta}_n) = \frac{2}{n} \sum_{i=1}^n \{E(\mathbf{S}_{\mathbf{W}_i + \mathbf{W}'} \mid \mathbf{W}_i) - \boldsymbol{\theta}(\boldsymbol{\eta}_n)\} + \tilde{\mathbf{R}}_n, \quad (2)$$

where \mathbf{W}' is an independent copy of \mathbf{W} . Arguing as in the proof of Thm. 1, it can be shown that $E(\|n^{1/2}\tilde{\mathbf{R}}_n\|^q)$ satisfies the bound obtained in (1) for every $n \geq 2$ and any $q \in (4/3, p]$. Thus, $n^{1/2}\tilde{\mathbf{R}}_n \rightarrow 0$ in probability as $n \rightarrow \infty$ under the sequence of shrinking shifts.

By definition, $\boldsymbol{\theta}(\boldsymbol{\eta}_n) = E(\mathbf{S}_{\mathbf{W}_1^0 + \mathbf{W}_2^0 + 2\boldsymbol{\eta}_n}) = E(\mathbf{S}_{2\boldsymbol{\eta}_n - \mathbf{W}_1^0 - \mathbf{W}_2^0})$, where $\mathbf{W}_i^0 = \mathbf{W}_i - \boldsymbol{\eta}_n$ has the same distribution as that of \mathbf{W}_i under the null hypothesis for $1 \leq i \leq n$. So, $n^{1/2}\boldsymbol{\theta}(\boldsymbol{\eta}_n)$ converges to $\mathbf{J}_{\mathbf{0}}^{(2)}(\boldsymbol{\eta})$ as $n \rightarrow \infty$.

Define $\boldsymbol{\Phi}_n(\mathbf{W}_i) = 2n^{-1/2}\{E(\mathbf{S}_{\mathbf{W}_i + \mathbf{W}'} \mid \mathbf{W}_i) - \boldsymbol{\theta}(\boldsymbol{\eta}_n)\}$. So, $E\{\boldsymbol{\Phi}_n(\mathbf{W}_i)\} = \mathbf{0}$. To prove the asymptotic Gaussianity of $\sum_{i=1}^n \boldsymbol{\Phi}_n(\mathbf{W}_i)$, it is enough to show that the triangular array $\{\boldsymbol{\Phi}_n(\mathbf{W}_1), \boldsymbol{\Phi}_n(\mathbf{W}_2), \dots, \boldsymbol{\Phi}_n(\mathbf{W}_n)\}_{n=1}^{\infty}$ of rowwise i.i.d. random elements satisfy the conditions of Corollary 7.8 in [Araujo and Giné \(1980\)](#).

Observe that for any $\epsilon > 0$,

$$\sum_{i=1}^n P\{\|\boldsymbol{\Phi}_n(\mathbf{W}_i)\| > \epsilon\} \leq \frac{8}{(\epsilon n)^{3/2}} \sum_{i=1}^n E\{\|E(\mathbf{S}_{\mathbf{W}_i + \mathbf{W}'} \mid \mathbf{W}_i) - \boldsymbol{\theta}(\boldsymbol{\eta}_n)\|^3\} \leq \frac{64}{(\epsilon^3 n)^{1/2}}.$$

Thus, $\lim_{n \rightarrow \infty} \sum_{i=1}^n P\{\|\Phi_n(\mathbf{W}_i)\| > \epsilon\} = 0$ for every $\epsilon > 0$, which ensures that condition (1) of Corollary 7.8 in [Araujo and Giné \(1980\)](#) holds.

We next verify condition (2) of Corollary 7.8 in [Araujo and Giné \(1980\)](#). Let us fix $\mathbf{f} \in \mathcal{X}$. Since $\|\mathbf{S}_{\mathbf{x}}\| = 1$ for all $\mathbf{x} \neq \mathbf{0}$, we can choose $\delta = 1$ in that condition (2). Then, using the linearity of \mathbf{f} , we have

$$\sum_{i=1}^n E[\mathbf{f}^2\{\Phi_n(\mathbf{W}_i)\}] = 4n^{-1} \sum_{i=1}^n E[\{V_{n,i} - E(V_{n,i})\}^2], \quad (3)$$

where $V_{n,i} = \mathbf{f}\{E(\mathbf{S}_{\mathbf{w}_i + \mathbf{w}'} \mid \mathbf{W}_i)\}$. Since the \mathbf{W}_i 's are identically distributed, the right hand side of (3) equals $4E[\{V_{n,1} - E(V_{n,1})\}^2]$. Note that $V_{n,1} = \mathbf{f}\{E(\mathbf{S}_{\mathbf{w}_1 + \mathbf{w}_2^0 + \eta_n} \mid \mathbf{W}_1)\}$. Since the norm in \mathcal{X} is assumed to be twice Gâteaux differentiable, it follows from Thm. 4.6.15(a) and Proposition 4.6.16 in [Borwein and Vanderwerff \(2010\)](#) that the norm in \mathcal{X} is Fréchet differentiable. This in turn implies that the map $\mathbf{x} \mapsto \mathbf{S}_{\mathbf{x}}$ is continuous on $\mathcal{X} \setminus \{\mathbf{0}\}$ (see [Borwein and Vanderwerff \(2010, Corollary 4.2.12\)](#)). Using this fact, it follows from the dominated convergence theorem for Banach space valued random elements that

$$\begin{aligned} E(\mathbf{S}_{\mathbf{w}_1 + \mathbf{w}_2^0 + \eta_n} \mid \mathbf{W}_1 = \mathbf{w}_1) &= E(\mathbf{S}_{\mathbf{w}_1 + \mathbf{w}_2^0 + \eta_n}) \\ &\rightarrow E(\mathbf{S}_{\mathbf{w}_1 + \mathbf{w}_2^0}) = E(\mathbf{S}_{\mathbf{w}_1^0 + \mathbf{w}_2^0} \mid \mathbf{W}_1^0 = \mathbf{w}_1) \end{aligned} \quad (4)$$

as $n \rightarrow \infty$ for almost all values of \mathbf{w}_1 . Thus, $E(V_{n,1})$ converges to $E[\mathbf{f}\{E(\mathbf{S}_{\mathbf{w}_1^0 + \mathbf{w}_2^0} \mid \mathbf{W}_1^0)\}]$ as $n \rightarrow \infty$ by the usual dominated convergence theorem. Similarly, $E(V_{n,1}^2)$ converges to $E[\mathbf{f}^2\{E(\mathbf{S}_{\mathbf{w}_1^0 + \mathbf{w}_2^0} \mid \mathbf{W}_1^0)\}]$ as $n \rightarrow \infty$. So, $\sum_{i=1}^n E[\mathbf{f}^2\{\Phi_n(\mathbf{W}_i)\}] \rightarrow \mathbf{\Pi}_2(\mathbf{f}, \mathbf{f})$ as $n \rightarrow \infty$, where $\mathbf{\Pi}_2$ is as defined before Thm. 1 in Sect. 2 and the expectations in that definition are taken under the null hypothesis. This completes the verification of condition (2) of Corollary 7.8 in [Araujo and Giné \(1980\)](#).

Finally, for the verification of condition (3) of Corollary 7.8 in [Araujo and Giné \(1980\)](#), suppose that $\{\mathcal{F}_k\}_{k \geq 1}$ is a sequence of finite dimensional subspaces of \mathcal{X}^* such that $\mathcal{F}_k \subseteq \mathcal{F}_{k+1}$ for all $k \geq 1$, and the closure of $\bigcup_{k=1}^{\infty} \mathcal{F}_k$ is \mathcal{X}^* . Such a sequence of subspaces exists because of the separability of \mathcal{X}^* . For any $\mathbf{x} \in \mathcal{X}^*$ and any $k \geq 1$, we define $d(\mathbf{x}, \mathcal{F}_k) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in \mathcal{F}_k\}$. It is straightforward to verify that for every $k \geq 1$, the map $\mathbf{x} \mapsto d(\mathbf{x}, \mathcal{F}_k)$ is continuous and bounded on any closed ball in \mathcal{X}^* . Thus, using (4), it follows that $\boldsymbol{\theta}(\eta_n) \rightarrow 0$ as $n \rightarrow \infty$, and

$$\begin{aligned} \sum_{i=1}^n E[d^2\{\Phi_n(\mathbf{W}_i), \mathcal{F}_k\}] &= 4n^{-1} \sum_{i=1}^n E[d^2\{E(\mathbf{S}_{\mathbf{w}_i + \mathbf{w}'} + \eta_n \mid \mathbf{W}_i) - \boldsymbol{\theta}(\eta_n), \mathcal{F}_k\}] \\ &= 4E[d^2\{E(\mathbf{S}_{\mathbf{w}_1 + \mathbf{w}_2^0 + \eta_n} \mid \mathbf{W}_1) - \boldsymbol{\theta}(\eta_n), \mathcal{F}_k\}] \\ &\rightarrow 4E[d^2\{E(\mathbf{S}_{\mathbf{w}_1^0 + \mathbf{w}_2^0} \mid \mathbf{W}_1^0), \mathcal{F}_k\}] \end{aligned}$$

as $n \rightarrow \infty$. From the choice of the \mathcal{F}_k 's, it can be shown that $d(\mathbf{x}, \mathcal{F}_k) \rightarrow 0$ as $k \rightarrow \infty$ for all $\mathbf{x} \in \mathcal{X}^*$. So, we have $\lim_{k \rightarrow \infty} E[d^2\{E(\mathbf{S}_{\mathbf{w}_1^0 + \mathbf{w}_2^0} \mid \mathbf{W}_1^0), \mathcal{F}_k\}] = 0$, and this completes the verification of condition (3) of Corollary 7.8 in [Araujo and Giné \(1980\)](#).

Thus, $\sum_{i=1}^n \Phi_n(\mathbf{W}_i)$ converges *weakly* to a zero mean Gaussian distribution in \mathcal{X} as $n \rightarrow \infty$. Further, its asymptotic covariance is $\mathbf{\Pi}_2$, which was obtained while checking condition (2) of Corollary 7.8 in Araujo and Giné (1980). Thus, it follows from equation (2) at the beginning of the proof that

$$n^{1/2}\{\mathbf{T}_{SR} - \boldsymbol{\theta}(\boldsymbol{\eta}_n)\} \longrightarrow G(\mathbf{0}, \mathbf{\Pi}_2)$$

weakly as $n \rightarrow \infty$ under the sequence of shrinking shifts. The above weak convergence and the fact that $n^{1/2}\boldsymbol{\theta}(\boldsymbol{\eta}_n)$ converges to $\mathbf{J}_0^{(2)}(\boldsymbol{\eta})$ as $n \rightarrow \infty$ complete the derivation of the asymptotic distribution of \mathbf{T}_{SR} .

Since \mathbf{T}_S is a sum of independent random elements, its asymptotic distribution can be obtained by using arguments similar to those used to derive the asymptotic distribution of $\sum_{i=1}^n \Phi_n(\mathbf{W}_i)$ given above. It follows that $n^{1/2}\mathbf{T}_S$ has an asymptotic Gaussian distribution with mean $\mathbf{J}_0^{(1)}(\boldsymbol{\eta})$ and covariance $\mathbf{\Pi}_1$ as $n \rightarrow \infty$ under the sequence of shrinking alternatives considered in the theorem. \square

Let \mathcal{X} be a separable Hilbert space and $\mathbf{Y}_1 - \mathbf{X}_1 = \sum_{k=1}^{\infty} Z_k \phi_k$ for an orthonormal basis ϕ_1, ϕ_2, \dots of \mathcal{X} and the Z_k 's are real-valued random variables. Then, the expectations defining $\mathbf{J}_0^{(1)}$ and $\mathbf{J}_0^{(2)}$ are finite if any two dimensional marginal of (Z_1, Z_2, \dots) has a density that is bounded on bounded subsets of \mathbb{R}^2 . If \mathcal{X} is a L_p space for some $p \in (1, \infty)$, then $\mathbf{J}_0^{(1)}$ and $\mathbf{J}_0^{(2)}$ exist if $E(\|\mathbf{W}_1\|^{-1})$ and $E(\|\mathbf{W}_1 + \mathbf{W}_2\|^{-1})$ are finite.

We next compare the asymptotic powers of the tests based on \mathbf{T}_S and \mathbf{T}_{SR} with some paired sample mean based tests for functional data in $L_2[a, b]$, where $a, b \in \mathbb{R}$. Cuevas et al. (2004) studied a test for analysis of variance in $L_2[a, b]$, and the test statistic for the two sample problem is based on $\|\bar{\mathbf{X}} - \bar{\mathbf{Y}}\|^2$. We consider the natural paired sample analog of this statistic and define it as $T_1 = \|\bar{\mathbf{W}}\|^2$. Horváth et al. (2013) studied a couple of two sample tests in $L_2[a, b]$ based on the projections of the sample functions onto the subspace formed by finitely many eigenfunctions of the sample pooled covariance operator. We consider the paired sample analogs of these tests, and the corresponding test statistics are defined as $T_2 = \sum_{k=1}^L \langle \bar{\mathbf{W}}, \hat{\boldsymbol{\psi}}_k \rangle^2$ and $T_3 = \sum_{k=1}^L \hat{\lambda}_k^{-1} \langle \bar{\mathbf{W}}, \hat{\boldsymbol{\psi}}_k \rangle^2$. Here, the $\hat{\lambda}_k$'s denote the eigenvalues of the empirical covariance of the \mathbf{W}_i 's in descending order of magnitudes, and the $\hat{\boldsymbol{\psi}}_k$'s are the corresponding empirical eigenfunctions. Consider i.i.d. paired observations $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \leq i \leq n$, where the \mathbf{X}_i 's and the \mathbf{Y}_i 's take values in $L_2[a, b]$. The next theorem gives the asymptotic distributions of T_1, T_2 and T_3 under sequences of shrinking location shift alternatives, where $\mathbf{W}_i = \mathbf{Y}_i - \mathbf{X}_i$ is symmetrically distributed about $n^{-1/2}\boldsymbol{\eta}$ for some fixed nonzero $\boldsymbol{\eta} \in \mathcal{X}$ and $1 \leq i \leq n$.

Theorem 3. (a) If $E(\|\mathbf{W}_1\|^2) < \infty$, nT_1 converges weakly to $\sum_{k=1}^{\infty} \lambda_k \chi_{(1)}^2(\beta_k^2/\lambda_k)$ as $n \rightarrow \infty$. Here, the λ_k 's are the eigenvalues of the covariance operator $\boldsymbol{\Sigma}$ of \mathbf{W}_1 in decreasing order of magnitudes, the $\boldsymbol{\psi}_k$'s are the eigenfunctions corresponding to the λ_k 's, $\beta_k = \langle \boldsymbol{\eta}, \boldsymbol{\psi}_k \rangle$, and $\chi_{(1)}^2(\beta_k^2/\lambda_k)$ denotes the non-central chi-square variable with one degree of freedom and non-centrality parameter

β_k^2/λ_k for $k \geq 1$.

(b) Suppose that for some $L \geq 1$, we have $\lambda_1 > \dots > \lambda_L > \lambda_{L+1} > 0$. Assume that $E(\|\mathbf{W}_1\|^4) < \infty$. Then, nT_2 converges weakly to $\sum_{k=1}^L \lambda_k \chi_{(1)}^2(\beta_k^2/\lambda_k)$, and nT_3 converges weakly to $\sum_{k=1}^L \chi_{(1)}^2(\beta_k^2/\lambda_k)$ as $n \rightarrow \infty$.

Proof. (a) As in the proof of Theorem 2, we denote $\boldsymbol{\eta}_n = n^{-1/2}\boldsymbol{\eta}$, and $\mathbf{W}_i^0 = \mathbf{W}_i - \boldsymbol{\eta}_n$ has the same distribution as that of \mathbf{W}_i under the null hypothesis for $1 \leq i \leq n$. Now, by the central limit theorem for i.i.d. random elements in a separable Hilbert space (see Araujo and Giné (1980, Thm. 7.5(i))), it follows that $n^{1/2}\overline{\mathbf{W}}^0$ converges weakly to $G(\mathbf{0}, \boldsymbol{\Sigma})$ as $n \rightarrow \infty$. Thus, $n^{1/2}\overline{\mathbf{W}} = n^{1/2}(\overline{\mathbf{W}}^0 + \boldsymbol{\eta}_n)$, converges weakly to $G(\boldsymbol{\eta}, \boldsymbol{\Sigma})$ as $n \rightarrow \infty$. Now, the distribution of $\|G(\boldsymbol{\eta}, \boldsymbol{\Sigma})\|^2$ is the same as that of $\sum_{k=1}^{\infty} \lambda_k \chi_{(1)}^2(\beta_k^2/\lambda_k)$ using the spectral decomposition of the compact self-adjoint operator $\boldsymbol{\Sigma}$. This proves part (a) of the proposition.

(b) Let $\mathbf{V} = (\langle \overline{\mathbf{W}}, \boldsymbol{\psi}_1 \rangle, \langle \overline{\mathbf{W}}, \boldsymbol{\psi}_2 \rangle, \dots, \langle \overline{\mathbf{W}}, \boldsymbol{\psi}_L \rangle)$ and $\tilde{\boldsymbol{\beta}} = (\beta_1, \dots, \beta_L)$. It follows from the central limit theorem in \mathbb{R}^L that the distribution of $n^{1/2}(\mathbf{V} - n^{-1/2}\tilde{\boldsymbol{\beta}})$ converges weakly to the L -variate Gaussian distribution $N_L(\mathbf{0}, \boldsymbol{\Lambda}_L)$ as $n \rightarrow \infty$ under the sequence of shrinking shifts considered, where $\boldsymbol{\Lambda}_L$ is the diagonal matrix $\text{Diag}(\lambda_1, \dots, \lambda_L)$. Thus, under the given sequence of shifts, the distribution of $n^{1/2}\mathbf{V}$ converges weakly to the L -variate Gaussian distribution $N_L(\tilde{\boldsymbol{\beta}}, \boldsymbol{\Lambda}_L)$ distribution as $n \rightarrow \infty$.

From arguments similar to those in the proof of Thm. 5.3 in Horváth and Kokoszka (2012), and using the assumptions in the present theorem, we get

$$\max_{1 \leq k \leq L} n^{1/2} |\langle \overline{\mathbf{W}}, \hat{\boldsymbol{\psi}}_k - \hat{c}_k \boldsymbol{\psi}_k \rangle| = o_P(1) \quad (5)$$

as $n \rightarrow \infty$ under the sequence of shrinking shifts. Here $\hat{\boldsymbol{\psi}}_k$ is the empirical version of $\boldsymbol{\psi}_k$ and $\hat{c}_k = \text{sign}(\langle \hat{\boldsymbol{\psi}}_k, \boldsymbol{\psi}_k \rangle)$. In view of (5), the limiting distribution of $n \sum_{k=1}^L (\langle \overline{\mathbf{W}}, \hat{\boldsymbol{\psi}}_k \rangle)^2$ is the same as that of $n \sum_{k=1}^L (\langle \overline{\mathbf{W}}, \hat{c}_k \boldsymbol{\psi}_k \rangle)^2 = n \|\mathbf{V}\|^2$, and the latter converges weakly to $\|N_L(\tilde{\boldsymbol{\beta}}, \boldsymbol{\Lambda}_L)\|^2$ as $n \rightarrow \infty$. Thus, nT_2 converges weakly to $\sum_{k=1}^L \lambda_k \chi_{(1)}^2(\beta_k^2/\lambda_k)$ as $n \rightarrow \infty$ under the sequence of shrinking shifts considered.

It also follows using similar arguments as in the proof of Thm. 5.3 in Horváth and Kokoszka (2012) that under the assumptions of the present theorem, we have

$$\max_{1 \leq k \leq L} n^{1/2} \hat{\lambda}_k^{-1/2} |\langle \overline{\mathbf{W}}, \hat{\boldsymbol{\psi}}_k - \hat{c}_k \boldsymbol{\psi}_k \rangle| = o_P(1)$$

as $n \rightarrow \infty$ under the given sequence of shrinking shifts. Similar arguments as in the case of T_2 yield the asymptotic distribution of nT_3 , and this completes the proof. \square

3.1 Comparison of asymptotic powers of different tests

We now compare the asymptotic powers of the tests based on \mathbf{T}_S , \mathbf{T}_{SR} , T_1 , T_2 and T_3 under shrinking location shifts. For this we have taken the random element $\mathbf{W}_i \in L_2[0, 1]$ to have the same

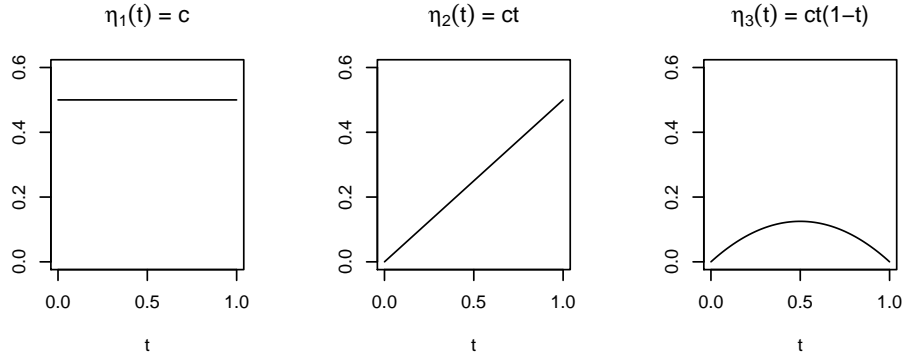


Figure 1: Plots of the shift alternatives

distribution as that of $\sum_{k=1}^{\infty} 2^{1/2} \{(k-0.5)\pi\}^{-1} Z_k \sin\{(k-0.5)\pi t\}$, where the Z_k 's are independent random variables for $k \geq 1$. We have considered two distributions of the Z_k 's, namely, the Z_k 's having $N(0, 1)$ distributions, and $Z_k = U_k(V/5)^{-1/2}$ with the U_k 's having $N(0, 1)$ distributions and V having a chi-square distribution with 5 degrees of freedom independent of the U_k 's for each $k \geq 1$. The latter choice is made to compare the performance of the tests based on \mathbf{T}_{SR} and \mathbf{T}_S with the mean based tests that use T_1 , T_2 and T_3 , when the underlying distribution is heavy tailed. The two distributions of \mathbf{W}_i considered here correspond to the Karhunen-Loève expansions of the standard Brownian motion and the t process (see Yu et al. (2007)) on $[0, 1]$ with 5 degrees of freedom having zero mean and covariance kernel $K(t, s) = \min(t, s)$, $t, s \in [0, 1]$, respectively. We call them the sBm and the $t(5)$ distributions, respectively. We have chosen five degrees of freedom for the t distribution so that the finiteness of the fourth moment required in Thm. 3 is satisfied. In a finite sample study in the next section, we will also consider t processes with one and three degrees of freedom, which violate some of the moment assumptions in Thm. 3. We have taken three choices of the location shift $\boldsymbol{\eta}$, namely, $\boldsymbol{\eta}_1(t) = c$, $\boldsymbol{\eta}_2(t) = ct$ and $\boldsymbol{\eta}_3(t) = ct(1-t)$, where $t \in [0, 1]$ and $c > 0$. The plots of the shifts are given in Fig. 1 below.

For evaluating the asymptotic powers of these tests, we have used Thm. 2 and Thm. 3. For each of the two distributions of \mathbf{W}_i , we have generated 5000 sample functions from it. The operators $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ are estimated as described in Sect. 2 using this sample, and the operator $\mathbf{\Sigma}$ is estimated by the sample covariance operator. The eigenvalues and the eigenfunctions of $\mathbf{\Pi}_1$, $\mathbf{\Pi}_2$ and $\mathbf{\Sigma}$ are then estimated by the eigenvalues and the eigenfunctions of the estimates of these operators. We have estimated $\mathbf{J}_0^{(1)}(\boldsymbol{\eta})$ and $\mathbf{J}_0^{(2)}(\boldsymbol{\eta})$ by their sample analogs. The asymptotic powers of the tests are then computed through 1000 Monte Carlo simulations from the asymptotic Gaussian distributions with the estimated parameters. We have used the cumulative variance method described in Horváth et al. (2013) to compute the number L associated with the tests based on T_2 and T_3 .

We now discuss the asymptotic powers of different tests under the sBm distribution. It is seen

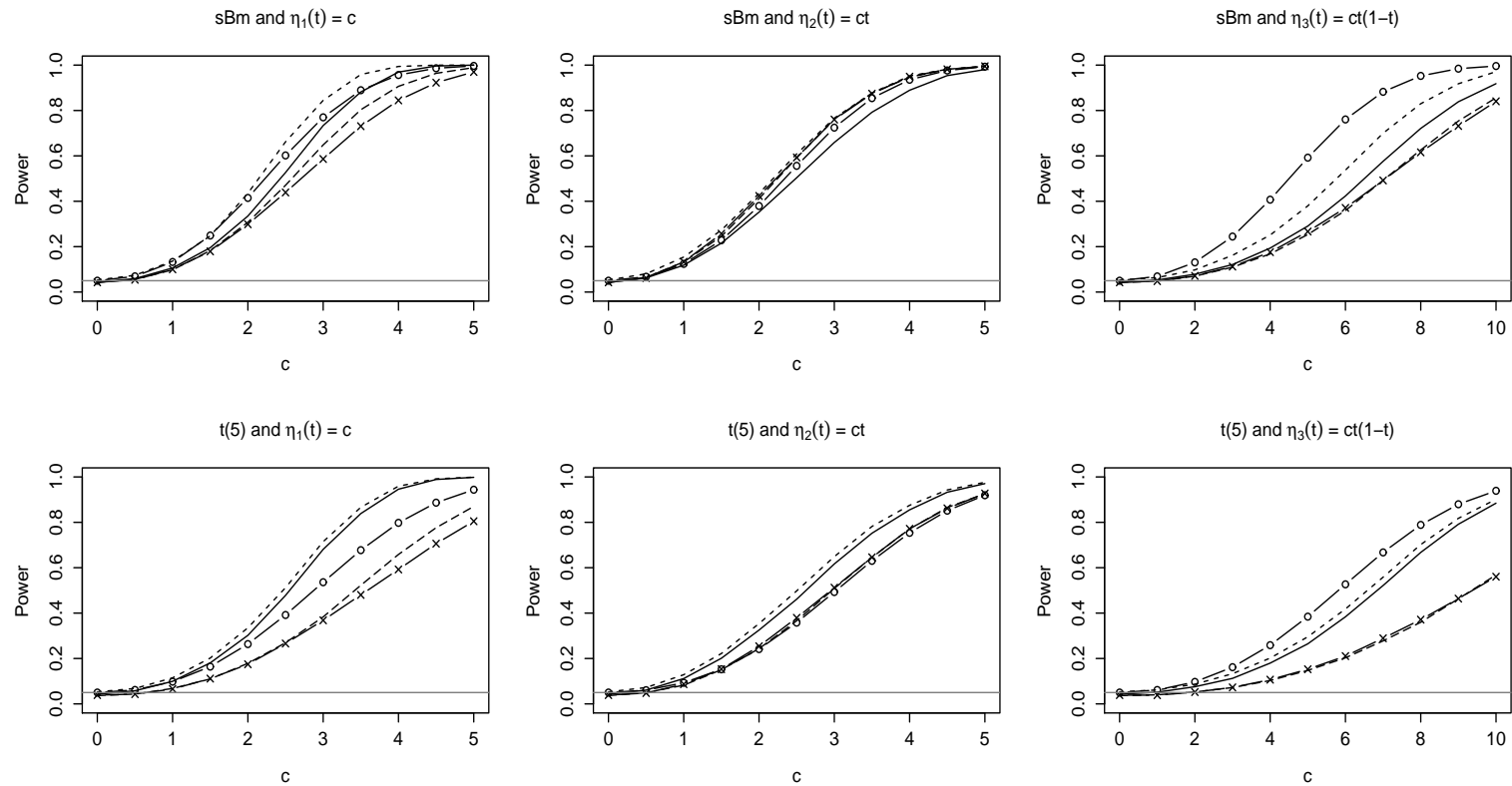


Figure 2: Plots of the asymptotic powers of the tests based on \mathbf{T}_S (—), \mathbf{T}_{SR} (---), T_1 (···), T_2 (- × -) and T_3 (- o -) for the sBm and the t(5) distributions under shrinking location shifts.

from Fig. 2 that the tests based on \mathbf{T}_S and \mathbf{T}_{SR} asymptotically outperform the tests based on T_1 and T_2 for the shifts $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_3(t)$. The test based on \mathbf{T}_{SR} is asymptotically more powerful than the test based on T_3 for all large values of c in the case of the shift $\boldsymbol{\eta}_1(t)$. However, for the shift $\boldsymbol{\eta}_3(t)$, the test based on T_3 is asymptotically more powerful than the test based on \mathbf{T}_{SR} . For the shift $\boldsymbol{\eta}_2(t)$ under the sBm distribution all the tests considered except the test based on \mathbf{T}_S have similar asymptotic powers, and the latter test is asymptotically less powerful than the other competitors under the sBm distribution.

We next consider the t(5) distribution. The tests based on \mathbf{T}_S and \mathbf{T}_{SR} asymptotically outperform all the competing tests for all the models considered, except the test based on T_3 for the shift $\boldsymbol{\eta}_3(t)$. The heavy tails of the t(5) distribution adversely affect the performance of the mean based tests, but the tests based on \mathbf{T}_S and \mathbf{T}_{SR} , which use spatial signs and spatial signed ranks, are less affected. For the shift $\boldsymbol{\eta}_3(t)$ under the t(5) distribution, although the test based on T_3 asymptotically outperforms the tests based on \mathbf{T}_S and \mathbf{T}_{SR} , its performance degrades significantly in comparison to its performance under the sBm distribution, which has lighter tails.

Between the two tests proposed in this paper, the test based on \mathbf{T}_{SR} is asymptotically more powerful than the test based on \mathbf{T}_S for both the distributions and the three shift alternatives considered in this paper. However, the powers of the two tests are quite close under the t(5) distribution.

4 Comparison of finite sample powers of different tests

We now carry out a comparative study of the finite sample empirical powers of the tests considered in the previous section for location shift alternatives. We consider the distribution of \mathbf{W}_i as in Sect. 3, i.e., \mathbf{W}_i has the same distribution as that of $\sum_{k=1}^{\infty} 2^{1/2} \{(k-0.5)\pi\}^{-1} Z_k \sin\{(k-0.5)\pi t\}$, where the Z_k 's are independent random variables for $k \geq 1$. We have considered four distributions for the \mathbf{W}_i 's, namely, the sBm and the t(5) distributions used in Sect. 3 as well as the t(1) and the t(3) distributions. For the t(1) (respectively, t(3)) distribution, $Z_k = U_k/(V/r)^{1/2}$, where the U_k 's are independent $N(0, 1)$ variables and V has a chi-square distribution with $r = 1$ (respectively, $r = 3$) degree of freedom independent of the U_k 's for each $k \geq 1$. The t(1) and the t(3) distributions are chosen to investigate the performance of the mean based tests when the moment conditions required by them fail to hold. We have chosen $n = 20$, and each sample curve is observed at 250 equispaced points in $[0, 1]$. We consider the same three location shifts that were considered in Sect. 3.1, namely, $\boldsymbol{\eta}_1(t) = c$, $\boldsymbol{\eta}_2(t) = ct$ and $\boldsymbol{\eta}_3(t) = ct(1-t)$ for $t \in [0, 1]$ and $c > 0$. All the sizes and the powers are evaluated by averaging over 1000 Monte-Carlo simulations. The estimated standard errors of the empirical sizes (respectively, powers) of different tests are of the order of 10^{-3} (respectively, 10^{-2} or less) for all the distributions considered. The empirical power curves of the tests are plotted in Fig. 3 and Fig. 4.

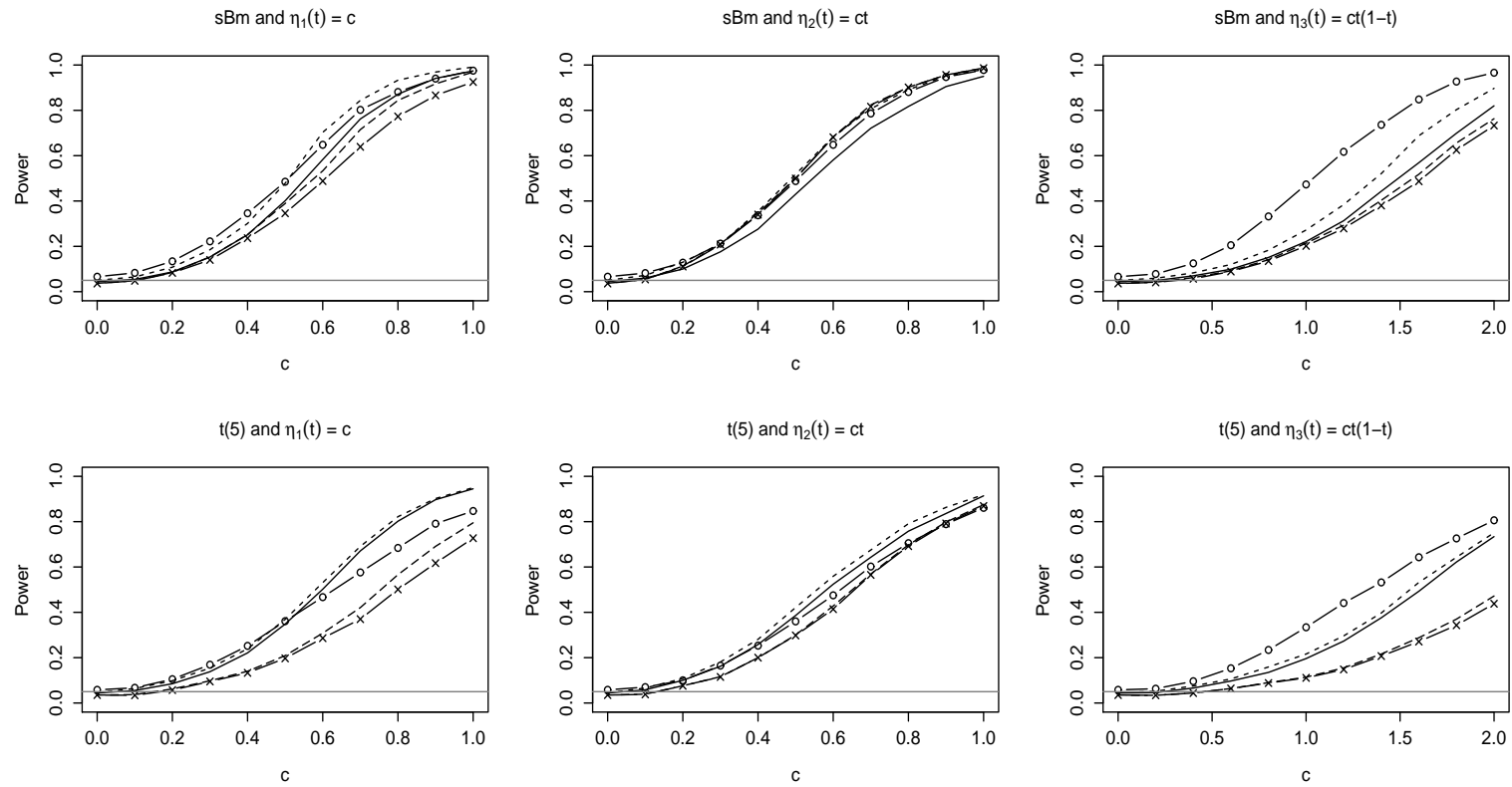


Figure 3: Plots of the empirical powers of the tests based on \mathbf{T}_S (—), \mathbf{T}_{SR} (- - -), T_1 (- · - ·), T_2 (- × -) and T_3 (- o -) for the sBm and the t(5) distributions.

For each of the tests, the difference between its observed size and the nominal 5% level is statistically insignificant under the sBm, the $t(3)$ and the $t(5)$ distributions. However, the sizes of the tests based on T_1 , T_2 and T_3 are 1%, 1% and 2.1% for the $t(1)$ distribution, all of which are significantly lower than the nominal level. On the other hand, the sizes of the tests based on spatial signs and signed ranks are not significantly different from the nominal level even under $t(1)$ distribution.

We first discuss the performance of different tests under the sBm distribution. The test based on \mathbf{T}_{SR} is significantly more powerful than the tests based on T_1 and T_2 for the shifts $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_3(t)$. The power of the test based on \mathbf{T}_S is not significantly different from that of the test based on T_1 for these two shifts although the former test has slightly more power than the latter test in our simulations. The test based on \mathbf{T}_S significantly outperforms the test based on T_2 for these two shifts. The test based on \mathbf{T}_{SR} is significantly more powerful than the test based on T_3 for large values of c in $\boldsymbol{\eta}_1(t)$. The test based on T_3 significantly outperforms the tests based on \mathbf{T}_S and \mathbf{T}_{SR} for the shift $\boldsymbol{\eta}_3(t)$. The power of the test based on \mathbf{T}_{SR} is not significantly different from the powers of the tests based on T_1 , T_2 and T_3 for the shift $\boldsymbol{\eta}_2(t)$, but the test based on \mathbf{T}_S is significantly less powerful than all its competitors for this shift under the sBm distribution.

We next consider the $t(5)$ distribution. The tests based on \mathbf{T}_S and \mathbf{T}_{SR} are significantly more powerful than the tests based on T_1 and T_2 for all the three shift alternatives considered. The tests based on \mathbf{T}_S and \mathbf{T}_{SR} are significantly more powerful than the test based on T_3 for large values of c in the shifts $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$. Like in the case of the sBm distribution, the test based on T_3 is significantly more powerful than the tests based on \mathbf{T}_S and \mathbf{T}_{SR} for the shift $\boldsymbol{\eta}_3(t)$. As in the asymptotic power study in Sect. 3, the performance of the mean based tests degrades significantly under the heavy tailed $t(5)$ distribution, while the tests based on spatial signs and signed ranks are less affected. The findings of this finite sample power study under the sBm and the $t(5)$ distributions are very similar to those of the asymptotic power study.

The tests based on T_1 , T_2 and T_3 have very low powers for all the shift alternatives considered under the $t(1)$ distribution and are significantly outperformed by the tests based on \mathbf{T}_S and \mathbf{T}_{SR} . The non-existence of moments of the $t(1)$ distribution severely affects the performance of the mean based tests, but the tests based on spatial signs and signed ranks are relatively less affected.

Under the $t(3)$ distribution, the tests based on \mathbf{T}_S and \mathbf{T}_{SR} significantly outperform the other competitors for the shifts $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$. The test based on T_3 has significantly more power than the tests based on \mathbf{T}_S and \mathbf{T}_{SR} only for $c \in (1, 1.5)$ in the shift $\boldsymbol{\eta}_3(t)$. However, the two latter tests significantly outperform the tests based on T_1 and T_2 for the shift $\boldsymbol{\eta}_3(t)$.

Among the two proposed tests, for all the shifts considered, the spatial sign test based on \mathbf{T}_S is significantly more powerful than the spatial signed-rank test based on \mathbf{T}_{SR} for the very heavy-tailed $t(1)$ distribution, while it is significantly less powerful under the light tailed sBm distribution. The powers of these two tests are not significantly different for the $t(3)$ and the $t(5)$ distributions.

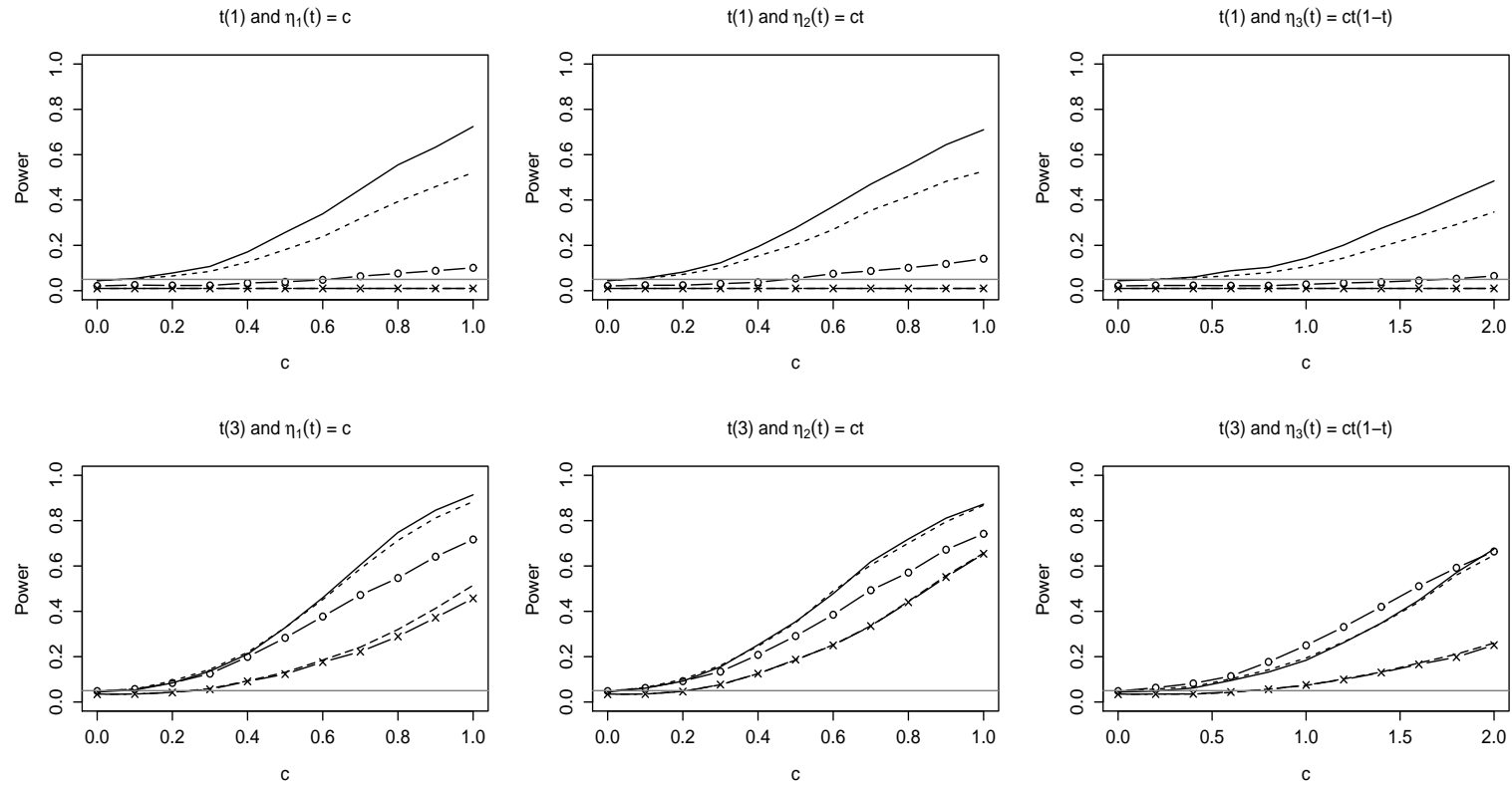


Figure 4: Plots of the asymptotic powers of the tests based on \mathbf{T}_S (—), \mathbf{T}_{SR} (---), T_1 (···), T_2 (- × -) and T_3 (- o -) for the $t(1)$ and the $t(3)$ distributions.

However, for the $t(3)$ distribution, the spatial sign test is slightly more powerful than the spatial signed-rank test for all the shifts considered, while the situation is reversed for the $t(5)$ distribution. These observations are similar to the relative performance of the spatial sign and signed-rank tests for finite dimensional data under multivariate t distributions (see Möttönen et al. (1997)).

4.1 Robustness study of different tests

It is known that the univariate sign and signed-rank tests are robust against the presence of outliers in the data, unlike some of the mean based tests like the t test. The influence functions of the sign and signed-rank tests are bounded, and as a result, the sizes and the powers of these tests, even under moderately high contamination proportions, are not much different from their powers in the uncontaminated case (see Hampel et al. (1986, Chapter 3)). The definition of the influence function for a test as discussed in Hampel et al. (1986, Chapter 3, p. 191) can be extended to the infinite dimensional setup by using the notion of a Gâteaux derivative. It can be shown that the influence functions of the tests based on \mathbf{T}_S and \mathbf{T}_{SR} are given by $[\mathbf{J}_0^{(1)}]^{-1}(\mathbf{S}_w)$ and $2[\mathbf{J}_0^{(2)}]^{-1}\{E(\mathbf{S}_{w+\mathbf{w}_1})\}$ provided the inverses exist, where $\mathbf{w} \in \mathcal{X}$, and $\mathbf{J}_0^{(1)}$ and $\mathbf{J}_0^{(2)}$ are defined before Thm. 2. When \mathcal{X} is a separable Hilbert space, both of these influence functions are bounded in norm under the assumptions analogous to those given in Proposition 2.1 in Cardot et al. (2013). So, it is expected that the tests based on \mathbf{T}_S and \mathbf{T}_{SR} will be robust in such cases.

We now conduct an empirical study to assess the robustness of spatial sign and signed-rank tests proposed in this paper in comparison with the mean based tests using T_1 , T_2 and T_3 . Let the distribution of $\mathbf{Y}_1 - \mathbf{X}_1 \in L_2[0, 1]$ be of the form $(1 - \epsilon)P + \epsilon Q$, where P is the sBm distribution considered earlier, Q is the Brownian motion with covariance kernel $K(t, s) = 16 \min(t, s)$, $t, s \in [0, 1]$, and assume that the contamination proportion ϵ takes values $1/20$, $3/20$ and $5/20$. So, even under contamination, *the null hypothesis remains unchanged*. As before, the sample size is chosen to be $n = 20$, and each sample curve is observed at 250 equispaced points in $[0, 1]$. For computing the powers of the tests in the presence of outliers in the data, we have considered the location shift $\Delta(t) = 0.8t$, $t \in [0, 1]$. This choice ensures that the powers of the tests, even under contamination, are not too close to the nominal 5% level nor too close to one, and thus a meaningful comparison between the tests can be made. The following table gives the sizes and the powers of different tests, which are evaluated by averaging over 1000 Monte-Carlo simulations, for various contamination models considered.

It is seen from Table 1 that except the tests based on T_1 and T_2 , the sizes of the other tests considered are close to the nominal 5% level in the contaminated as well as the uncontaminated situations. The sizes of the mean based tests that use T_1 and T_2 are much larger than the nominal level under contamination though the null hypothesis remains valid. This is probably because the asymptotic critical values of these two tests are under-estimated in the presence of contamination in

Table 1: Sizes and powers of some tests at nominal 5% level

ϵ	Size					Power		
	\mathbf{T}_S	\mathbf{T}_{SR}	T_1	T_2	T_3	\mathbf{T}_S	\mathbf{T}_{SR}	T_3
0	0.044	0.05	0.037	0.036	0.066	0.817	0.891	0.881
1/20	0.053	0.056	0.118	0.115	0.057	0.777	0.818	0.722
3/20	0.044	0.046	0.288	0.281	0.04	0.665	0.616	0.446
5/20	0.044	0.05	0.397	0.37	0.043	0.521	0.423	0.273

the data. Interestingly, the size of the mean based test using T_3 is unaffected under contamination. It seems that the standardization involved in the statistic T_3 keeps the size of the test under control even when there are outliers in the data.

For our power study in the presence of outliers in the data, we have excluded the tests based on T_1 and T_2 since they have very high sizes under contamination. The powers of all the tests decrease from their powers in the uncontaminated situation since the contamination increases the variability in the sample. The powers of the proposed spatial sign and signed-rank tests are significantly higher than the power of the mean based test using T_3 for all the contamination models considered. Recall that the test based on T_3 was significantly more powerful than the test based on \mathbf{T}_S , and its power was not significantly different from that of the test based on \mathbf{T}_{SR} for the same location shift in the uncontaminated case.

The powers of the tests based on spatial sign and signed rank are comparable when the contamination proportion is at most 3/20. However, when the proportion of contamination is 5/20, the spatial sign test becomes significantly more powerful than the spatial signed-rank test. This behaviour of the spatial sign and signed-rank tests is similar to that of the univariate sign and signed-rank tests, where it is known that the sign test is more robust than the signed-rank test for higher levels of contamination (see Hampel et al. (1986, Chapter 3)).

Concluding remarks

In this paper, we have studied a spatial sign test and a spatial signed-rank test for paired sample problems in infinite dimensional spaces. The tests are infinite dimensional extensions of the multivariate spatial sign and signed-rank tests considered earlier by Möttönen and Oja (1995), Möttönen et al. (1997) and Marden (1999) for finite dimensional data. We have shown that the asymptotic distributions of the proposed test statistics are Gaussian after appropriate centering and scaling. It is shown that both of the proposed tests are consistent for a class of alternatives

that includes the standard location shift alternatives. It is observed that under suitable sequences of shrinking location shift alternatives, the spatial sign and signed-rank tests are asymptotically more powerful than some of the mean based paired sample tests for infinite dimensional data when the underlying distribution has heavy tails. Further, even for some infinite dimensional Gaussian distributions, the spatial sign and signed-rank tests considered in this paper are more powerful than some of the mean based tests for paired sample problems. The asymptotic results are corroborated by finite sample simulation results.

The proposed spatial sign and signed-rank test statistics can be computed very easily, and the associated tests can be implemented using their asymptotic Gaussian distributions. The covariance operators of the asymptotic Gaussian distributions can be easily estimated as discussed in Sect. 2.1. For data in a Hilbert space, the implementations of the proposed tests become further simplified in view of the fact that in such a space, the squared norm of a Gaussian random element is distributed as a weighted sum of independent chi-square variables each with one degree of freedom.

The tests based on spatial signs and signed ranks studied in this paper do not require any moment assumption unlike some mean based tests for infinite dimensional data. These tests are also robust against contamination of the sample by outliers unlike the mean based tests considered in the paper. Between the two tests proposed in this paper, it is observed that the spatial sign test is better for some very heavy-tailed distributions, while the spatial signed-rank test outperforms it in some cases when the distribution has lighter tails. Further, the spatial sign test is more robust than the spatial signed-rank test, when there is a large amount of contamination in the sample.

Acknowledgement

Research of the first author is partially supported by the Shyama Prasad Mukherjee Fellowship of the Council of Scientific and Industrial Research, Government of India.

References

- Araujo, A., Giné, E. (1980) *The central limit theorem for real and Banach valued random variables*. John Wiley & Sons, New York-Chichester-Brisbane.
- Borovskikh, Yu. V.(1991) *UB-statistics*. *Teor. Veroyatnost. i Primenen.*, **36**, 417–433.
- Borwein, J. M., Vanderwerff, J. D. (2010) *Convex functions: constructions, characterizations and counterexamples*. Cambridge University Press, Cambridge.
- Cardot, H., Cénac, P., Zitt, P-A. (2013) Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm. *Bernoulli*, **19**, 18–43.

- Chakraborty, B., Chaudhuri, P., Oja, H. (1998) Operating transformation retransformation on spatial median and angle test. *Statist. Sinica*, **8**, 767–784.
- Chakraborty, A., Chaudhuri, P. (2014) A Wilcoxon-Mann-Whitney type test for infinite dimensional data. Technical Report. arXiv:1403.0201.
- Chaudhuri, P., Sengupta, D. (1993) Sign tests in multidimension: inference based on the geometry of the data cloud. *J. Amer. Statist. Assoc.*, **88**, 1363–1370.
- Cuevas, A., Febrero, M., Fraiman, R. (2004) An anova test for functional data. *Comput. Statist. Data Anal.*, **47**, 111–122.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel, W. A. (1986) *Robust statistics: The approach based on influence functions*. John Wiley & Sons, Inc., New York.
- Hallin, M., Paindaveine, D. (2002) Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks. *Ann. Statist.*, **30**, 1103–1133.
- Hettmansperger, T. P., Nyblom, J., Oja, H. (1994) Affine invariant multivariate one-sample sign tests. *J. Roy. Statist. Soc. Ser. B*, **56**, 221–234.
- Hettmansperger, T. P., Möttönen, J., Oja, H. (1997) Affine-invariant multivariate one-sample signed-rank tests. *J. Amer. Statist. Assoc.*, **92**, 1591–1600.
- Horváth, L., Kokoszka, P. (2012) *Inference for functional data with applications*. Springer, New York.
- Horváth, L., Kokoszka, P., Reeder, R. (2013) Estimation of the mean of functional time series and a two-sample problem. *J. R. Stat. Soc. Ser. B*, **75**, 103–122.
- Marden, J. I. (1999) Multivariate rank tests. In *Multivariate analysis, design of experiments, and survey sampling* (S. Ghosh ed.), 401–432. Dekker, New York.
- Möttönen, J., Oja, H. (1995) Multivariate spatial sign and rank methods. *J. Nonparametr. Statist.*, **5**, 201–213.
- Möttönen, J., Oja, H., Tienari, J. (1997) On the efficiency of multivariate spatial sign and rank tests. *Ann. Statist.*, **25**, 542–552.
- Oja, H. (1999) Affine invariant multivariate sign and rank tests and corresponding estimates: a review. *Scand. J. Statist.*, **26**, 319–343.
- Oja, H. (2010) *Multivariate nonparametric methods with R: An approach based on spatial signs and ranks*. Springer, New York.

- Peters, D., Randles, R. H. (1990) A multivariate signed-rank test for the one-sample location problem. *J. Amer. Statist. Assoc.*, **85**, 552–557.
- Puri, M. L., Sen, P. K. (1971) *Nonparametric methods in multivariate analysis*. John Wiley & Sons, Inc., New York-London-Sydney.
- Randles, R. H. (1989) A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.*, **84**, 1045–1050.
- Vakhania, N. N., Tarieladze, V. I., Chobanyan, S. A. (1987) *Probability Distributions on Banach Spaces (Translated from the Russian and with a preface by Wojbor A. Woyczynski)*. D. Reidel Publishing Co., Dordrecht.
- Yu, S., Tresp, V., Yu, K. (2007) Robust multi-task learning with t-processes. In *Proceedings of the 24th International Conference on Machine Learning (Oregon, 2007)*, 1103–1110. Omnipress, Corvallis, OR.