A median test for functional data

Zaineb SMIDA [*], Lionel CUCALA [†], Ali GANNOUN [‡], and Ghislain DURIF [§]

Institut Montpelliérain Alexander Grothendieck, CNRS, Université de Montpellier, France.

Abstract

The median test has been proven to be more powerful than the Student t-test and the Wilcoxon-Mann-Whitney test in heavy-tailed cases for univariate data. The multivariate extension of the median test, for multidimensional data, was demonstrated to be more efficient than the Hotelling T^2 and the Wilcoxon-Mann-Whitney tests for high dimensions and in very heavy-tailed cases. On the basis of these postulates, in this paper, we construct a median type test based on spatial ranks for functional data, i.e in infinite dimensional space, and we obtain asymptotic results. Then, we compare the proposed functional median test with numerous competing tests using simulated and real functional data : as in the univariate and multivariate cases, the proposed test is more adapted to heavy-tailed distributions.

Keywords: Functional Data, Gateaux Derivative, Heavy-Tailed Distributions, Two-Sample Location Test, Separable Hilbert Space, Smooth Banach Space.

AMS Subject Classifications: 60F17, 62G05, 62G10, 62G20.

1 Introduction

Statistical hypothesis testing plays an essential role in statistics (Lehmann, 1986; Lehmann and Romano, 2005). In nonparametric statistics, tests of hypotheses are known as nonparametric or distribution-free tests. It is not necessary to assume hypotheses on the shape of the distribution and estimate its parameters. These tests can be used to verify that two or more datasets come from identical populations.
Here, we will focus on this type of tests to solve the two-sample location problem which

[*]Email adress: zaineb.smida@umontpellier.fr (Corresponding author)
[†]Email adress: lionel.cucala@umontpellier.fr
[‡]Email adress: ali.gannoun@umontpellier.fr
[§]Email adress: ghislain.durif@umontpellier.fr

received a considerable attention in the past. More specifically, we consider the known two-sample problem with independent observations

$$\begin{aligned} X_1, \dots, X_m &\sim F \\ Y_1, \dots, Y_n &\sim G, \end{aligned}$$

where F and G are continuous distributions functions. Then, we only focus on the situation where the distribution function G is considered as a shifted version of F, i.e. $G(\cdot) = F(\cdot - \Delta)$. In this case, the null hypothesis of equality of F and G can be expressed as $H_0 : \Delta = 0$ against the alternative one which can be $\Delta \neq 0$.

For univariate data, Wilcoxon (1945) and Mann and Whitney (1947) proposed nonparametric tests based on ranks. Each of them defined their own test statistic which leads to the same test named Wilcoxon-Mann-Whitney. This test is more powerful than the Student's t-test for various non-Gaussian distributions (Blair and Higgins, 1980) and it is also asymptotically optimum in case of a logistic type density (Hàjek et al., 1999). Another test of hypothesis of the location problem is assigned to Mood (1950) and it is called the median test. Another version of this test based on ranks was presented in Van der Vaart (1998). This version of test is an asymptotically optimum in the case of a double exponential distribution (Capéraà and Cutsem, 1988; Hàjek et al., 1999). In fact, it is based on a statistic which counts the number of individuals from the second sample exceeding the median of the pooled sample unlike the Wilcoxon-Mann-Whitney test statistic which uses the sum of ranks of the second sample in the pooled sample. Nowadays, the median test is not often used because it is less powerful than the Wilcoxon-Mann-Whitney test when applied to Gaussian distributions (Mood, 1954). However, this test is more efficient, when dealing with symmetrical distributions with heavy-tails, than the Wilcoxon-Mann-Whitney one (Capéraà and Cutsem, 1988).

For multivariate data, several versions of the Hotelling, Wilcoxon-Mann-Whitney and median tests have been studied. See, for example, Puri and Sen (1971), Chakraborty and Chaudhuri (1999), Oja and Randles (2004), Oja (2010). The extension of univariate two-sample Mood test is called the sign test and it has the best efficiency in very heavy-tailed cases and for high dimensions : it also outperforms the Hotelling test in heavy-tailed cases (Oja and Randles , 2004).

Currently, the development of the sensoring and computing tools allows us to work with huge datasets. So, we have more and more access to data of functional type, for example the functional chemometric data and the electricity consumption of different regions (Ramsay and Silverman, 2005; Ferraty and Vieu, 2006; Kokoszka and Reimherr, 2017). These kinds of data are not real random variables or vectors but they are a collection of random elements like curves, surfaces, images, etc, and each sample variable is usually considered as a function. The main particularity of such data is the infinite dimension of the data space such as Banach and Hilbert spaces. Appropriate statistical tools are necessary to handle these types of data. Definitions of functional spaces and main convergence results such as central limit theorem are presented in Araujo and Giné (1980) and Bosq (2000). Definitions of various parameters as centrality and dispersion ones and their specific estimations are introduced in several papers such as Cuevas (2014), Chakraborty and Chaudhuri (2014b) and Goia and

Vieu (2016). Recent advances on various aspects of functional data analysis are presented in the latest publications of Aneiros et al. (2019) and Aneiros et al. (2022) which can be considered as reviews of the literature on these topics. Ever since it was popularized by Ferraty and Vieu (2006), nonparametric functional data analysis has become an active field of research (Geenens, 2015; Ling and Vieu, 2018; Chowdhury and Chaudhuri, 2020), mainly for modelling and regressing these specific kind of data.

Two-sample tests of hypotheses for functional data have also been proposed by several authors using either parametric or nonparametric techniques, such as in the univariate and multivariate settings. In the parametric case, to decide whether two samples of curves are issued from the same distribution, Horváth et al. (2013) proposed two test statistics for testing the equality of mean functions. These two test statistics are based on the orthogonal projections on the space generated by the eigenfunctions of an L^2-consistent estimator. This estimator is obtained from the asymptotic covariance operator of the difference between the two-sample mean functions. Among these two tests, one is the same as the Hotelling statistic in finite dimensional space. Still using a parametric approach, Cuevas et al. (2004) introduced an analog of the classical one-way analysis of variance (ANOVA) problem for functional data.

In a nonparametric setting, Chakraborty and Chaudhuri (2015) (see, also Chakraborty and Chaudhuri, 2014a) proposed a Wilcoxon-Mann-Whitney test based on spatial ranks. Their statistic is an extension of the one defined for example in Hàjek et al. (1999) and Van der Vaart (1998) for real valued random variables.

Our goal here is to construct an extension of the median test for processes valued in infinite dimensional spaces. The rest of this paper is organised as follows : in Section 2, we propose a median test statistic based on spatial ranks in Banach space and its extensions in separable Hilbert space. We also introduce a modified and more simple version of this statistic. Then, we study the asymptotic behavior of the latter one under the null hypothesis. We implement the test using its asymptotic distribution on the one hand and using the random permutation method on the other hand. Then, we derive the asymptotic distribution under some shrinking location shifts models and we describe how asymptotic power can be evaluated. To illustrate our theoretical results, we compare in Section 3 the performance of the proposed test with various other tests, either parametric or nonparametric, using simulated and real functional datasets. We conclude with a discussion of the methods and results.

2 The construction of the test

2.1 The introduction of the median statistics

First we recall what median tests look like in the univariate case. Let X and Y be two $\mathbb{R}$-valued random variables. We consider $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ two independent random samples of X and Y with distribution functions F and F_θ respectively, such that $\forall x \in \mathbb{R}; F_\theta(x) = F(x - \theta)$. The constant θ is called *the translation parameter*.

The median test statistic based on ranks (Capéraà and Cutsem, 1988) for testing the hypothesis

$$H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta \neq 0$$

is defined as

$$T_{\mathrm{Mo}} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{R_i > 0\}},$$

where $R_i = 1 + (\sum_{j=1}^{m} \mathbb{1}_{\{Y_i > X_j\}} + \sum_{k=1}^{n} \mathbb{1}_{\{Y_i > Y_k\}} - \frac{N+1}{2})$ is the centered rank of Y_i when $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ are ordered together in the same sample of size $N = n + m$.
This test is based on the number of observations of $Y_1, \dots, Y_n$ that is strictly greater than the global median of the N observations.
We consider also the following test statistic :

$$T'_{\mathrm{Mo}} = \frac{1}{n} \sum_{i=1}^{n} \mathrm{sign}(R_i), \tag{1}$$

where the sign function is $x \longmapsto \mathrm{sign}(x) = \begin{cases} \frac{x}{|x|} & \text{if} \;\; x \neq 0 \\ 0 & \text{if} \;\; x = 0 \end{cases}$.
These two test statistics are equivalent and related to the same test since $T'_{\mathrm{Mo}} = 2\mathrm{T}_{\mathrm{Mo}} - 1$.
To make things easier afterwards, we introduce a test statistic which counts the number of the observations $Y_1, \dots, Y_n$ that are greater than the median of the observations $X_1, \dots, X_m$ instead of the global median. In other words, in the univariate case it is equal to

$$T_{\mathrm{MED}} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{\hat{F}_m(Y_i) > \frac{1}{2}\}},$$

where $\hat{F}_m(x) = 1/m \sum_{j=1}^{m} \mathbb{1}_{\{X_j \leq x\}}$ is the empirical distribution function of $X_1, \dots, X_m$. This statistic is inspired from the work of Koul and Staudte (1972).
Our goal here is to construct an extension of T_{MED} in infinite dimensional space.

2.1.1 T_{MED} in the functional case

We shall now consider X and Y two independant random elements in a Banach space χ. We denote by χ^* its dual space, i.e., the space of the linear continuous functions on χ with values in $\mathbb{R}$, and χ^{**} its bidual space, i.e., the space of the linear continuous functions on χ^* with values in $\mathbb{R}$. Now, we suppose that :

- The space χ is smooth, i.e., the norm function $\|.\|_\chi$ is Gateaux differentiable at each $x \neq 0, x \in \chi$. We denote by $\mathbf{SGN}_x \in \chi^*$ its Gateaux derivative. This sign function is defined, for all $h \in \chi$, as

$$\mathbf{SGN}_x(h) = \begin{cases} \lim_{t \to 0} \frac{\|x+th\|_\chi - \|x\|_\chi}{t} & \text{if} \;\; x \neq 0 \\ 0 & \text{if} \;\; x = 0 \end{cases} .$$

- The space χ^* is smooth, i.e, the norm function $\|.\|_{\chi^*}$ is Gateaux differentiable at each $y \neq 0$, $y \in \chi^*$. We denote by $\mathbf{SGN}^*_y \in \chi^{**}$ its Gateaux derivative. This sign function is defined, for all $H \in \chi^*$, as

$$\mathbf{SGN}^*_y(H) = \begin{cases} \lim_{t \to 0} \frac{\|y+tH\|_{\chi^*} - \|y\|_{\chi^*}}{t} & \text{if} \quad y \neq 0 \\ 0 & \text{if} \quad y = 0 \end{cases}.$$

Also, we consider $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ independent random samples of X and Y from two probability measures P and Q on χ. We suppose that P and Q differ by a shift $\Delta \in \chi$. Then, for testing

$$H_0 : \Delta = 0 \quad \text{against} \quad H_1 : \Delta \neq 0,$$

the statistic T_{MED} becomes

$$\begin{aligned} \text{MED} &= \frac{1}{n} \sum_{i=1}^{n} \mathbf{SGN}^*_{\left\{\frac{1}{m}\sum\limits_{j=1}^{m} \mathbf{SGN}_{\{Y_i - X_j\}}\right\}} \\ &= \frac{1}{n} \sum_{i=1}^{n} \phi(F_m(Y_i)), \end{aligned} \tag{2}$$

where

* $\phi : u \mapsto \phi(u) = \mathbf{SGN}^*_u$ is a map on $\chi^* \backslash \{0\}$ (see, e.g., Corollary 4.2.12 in Borwein and Vanderwerff, 2010) that is continuous if the norm on χ^* is twice Gateaux differentiable. This sign function was used to proof asymptotic properties of the spatial depth in infinite dimensional spaces. See Theorem 4.1 in Chakraborty and Chaudhuri (2014b) for more details.

* $F_m : y \mapsto F_m(y) = \frac{1}{m} \sum\limits_{j=1}^{m} \mathbf{SGN}_{\{y - X_j\}}$ is the empirical spatial distribution associated to the iid observations $X_1, \ldots, X_m$ (Chakraborty and Chaudhuri, 2014b). This empirical spatial distribution has been used to develop the Wilcoxon-Mann-Whitney type test for two-sample problems in infinite dimensional spaces (Chakraborty and Chaudhuri, 2015). Remark that, in the univariate case, the empirical spatial distribution is equal to $2\hat{F}_m(y) - 1$ where $\hat{F}_m$ is the empirical distribution function of $X_1, \ldots, X_m$.

* We will also denote by $F(y)$ the spatial distribution of X at $y \in \chi$ which is equal to $\mathbb{E}\left[\mathbf{SGN}_{\{y-X\}}\right]$. For more details, see Chakraborty and Chaudhuri (2014b). Note that, in the univariate case, the spatial distribution is equal to $2\tilde{F}(y) - 1$ where $\tilde{F}$ is the distribution function of X.

As stated by Chakraborty and Chaudhuri (2014a), when the space χ is assumed to be an Hilbert one, sign functions become simpler so that $\mathbf{SGN}_x = \frac{x}{\|x\|_\chi}$ and $\mathbf{SGN}^*_y = \frac{y}{\|y\|_{\chi^*}}$. Thus, we might rewrite the statistic MED as follows :

$$\mathrm{MED} = \frac{1}{n}\sum_{i=1}^{n} \frac{\displaystyle\sum_{j=1}^{m} \frac{Y_i - X_j}{\|Y_i - X_j\|_\chi}}{\left\| \displaystyle\sum_{j=1}^{m} \frac{Y_i - X_j}{\|Y_i - X_j\|_\chi} \right\|_\chi}.$$

2.1.2 T'_{Mo} in the functional case

An extension of T'_{Mo} defined as (1) in the functional case can be written as

$$\mathrm{Mo} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{SGN}^*_{\left\{\sum_{k=1}^{n} \mathbf{SGN}_{Y_i - Y_k} + \sum_{j=1}^{m} \mathbf{SGN}_{Y_i - X_j}\right\}}.$$

When χ is assumed to be an Hilbert space, the statistic Mo becomes

$$\mathrm{Mo} = \frac{1}{n}\sum_{i=1}^{n} \frac{\displaystyle\sum_{k=1,k\neq i}^{n} \frac{Y_i - Y_k}{\|Y_i - Y_k\|_\chi} + \sum_{j=1}^{m} \frac{Y_i - X_j}{\|Y_i - X_j\|_\chi}}{\left\| \displaystyle\sum_{k=1,k\neq i}^{n} \frac{Y_i - Y_k}{\|Y_i - Y_k\|_\chi} + \sum_{j=1}^{m} \frac{Y_i - X_j}{\|Y_i - X_j\|_\chi} \right\|_\chi}.$$

2.2 Asymptotic distribution

In this subsection, we study the asymptotic normality of MED.
First, we introduce the following notations which will be used later :

- We denote by $\mathbf{G} := G(\mathbf{m}, \mathbf{C})$ the distribution of a Gaussian random element (say $\mathbf{G}$) in a separable Banach space χ with mean $\mathbf{m} \in \chi$ and covariance $\mathbf{C}$, where $\mathbf{C} : \chi^* \times \chi^* \to \mathbb{R}$ is a symmetric nonnegative definite continuous bilinear function. For all $\mathbf{l} \in \chi^*$, $\mathbf{l}(\mathbf{G})$ follows an univariate Gaussian distribution with mean $\mathbf{l}(\mathbf{m})$ and variance $\mathbf{C}(\mathbf{l}, \mathbf{l})$.

- For all $x, y \in \chi$, define

$$F_X(y) = \mathbb{E}\big[\mathbf{SGN}_{\{y-X\}}\big|y\big] , \tag{3}$$

$$F_Y(x) = \mathbb{E}\big[\mathbf{SGN}_{\{Y-x\}}\big|x\big] . \tag{4}$$

  These two functions are used to prove the theorem of the asymptotic normality of the Wilcoxon-Mann-Whitney test statistic under finite and shrinking locations shifts (Chakraborty and Chaudhuri, 2015). Moreover, we denote

$$\mu = \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)\}}\right]$$

  and

$$\tilde{\mu} = \mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X)\}}\right] .$$

  Remark that, under H_0, we have $\mu = \tilde{\mu}$.

- Let $\Gamma_1, \Gamma_2 : \chi^{***} \times \chi^{***} \to \mathbb{R}$ be the symmetric positive definite continuous bilinear operators defined as :

$$\Gamma_1(f, g) = \mathbb{E}\left[f\left(\mathbf{SGN}^*_{\{F_X(Y)\}}\right) g\left(\mathbf{SGN}^*_{\{F_X(Y)\}}\right)\right] - f(\mu)g(\mu) \tag{5}$$

and

$$\Gamma_2(f, g) = \mathbb{E}\left[f\left(\mathbf{SGN}^*_{\{F_Y(X)\}}\right) g\left(\mathbf{SGN}^*_{\{F_Y(X)\}}\right)\right] - f(\tilde{\mu})g(\tilde{\mu}), \tag{6}$$

where $f, g \in \chi^{***}$.

For our next theorem, we shall also consider the following assumptions and definition :

Assumption 1. *We assume that the norm in χ^* is twice Gateaux differentiable at every $x \neq 0$.*

From Assumption 1 (see also, e.g., Chapter 4, Section 6 in Borwein and Vanderwerff, 2010), let $\mathbf{J_x} : \chi^* \to \chi^{**}$ denote, when it exists, the Hessian of the function $g : x \mapsto \mathbb{E}\Big[\|F_X(Y) + x\|_{\chi^*}\Big|X_1, \dots, X_m\Big]$, $x \in \chi^*$. In particular, if we assume that χ is an Hilbert space, then χ^* is also an Hilbert one. Since the norms in Hilbert spaces are twice Gateaux differentiable (p. 6 in Chakraborty and Chaudhuri, 2014a), and if $Z = F_X(Y)$, the derivative of the map g is defined as :

$$\nabla_x g = \mathbb{E}\big[\mathbf{SGN}^*_{\{Z+x\}}\big|X_1, \dots, X_m\big] = \mathbb{E}\left[\frac{Z + x}{\|Z + x\|_{\chi^*}}\middle|X_1, \dots, X_m\right]$$

and its Hessian is:

$$\begin{array}{ccccccc} \mathbf{J}_x & : & \chi^* & \to & \chi^{**} & & \\ & & h & \mapsto & \mathbf{J}_x(h) & : & \chi^* \to \mathbb{R} \\ & & & & & & v \mapsto \{\mathbf{J}_x(h)\}(v) := \langle \mathbf{J}_x(h), v\rangle. \end{array}$$

Then, we have

$$\mathbf{J}_x = \mathbb{E}\left[\frac{1}{\|Z + x\|_{\chi^*}}\left(\mathbf{I}_{\chi^*} - \frac{(Z + x) \otimes (Z + x)}{\|Z + x\|^2_{\chi^*}}\right)\middle|X_1, \dots, X_m\right],$$

where $\mathbf{I}_{\chi^*}$ is the identity operator in χ^* and $u \otimes v(h) = \langle u, h\rangle.v$ for all $h, v \in \chi^*$. Thus,

$$\mathbf{J}_x(h) = \mathbb{E}\left[\frac{h}{\|Z + x\|_{\chi^*}} - \frac{\langle Z + x, h\rangle\,(Z + x)}{\|Z + x\|^3_{\chi^*}}\middle|X_1, \dots, X_m\right],$$

for all $h \in \chi^*$. More explicitly, $\mathbf{J}_x$ is given by

$$\{\mathbf{J}_x(h)\}(v) = \langle \mathbf{J}_x(h), v\rangle = \mathbb{E}\left[\frac{1}{\|Z + x\|_{\chi^*}}\left(\langle h, v\rangle - \frac{\langle Z + x, h\rangle\langle Z + x, v\rangle}{\|Z + x\|^2_{\chi^*}}\right)\middle|X_1, \dots, X_m\right],$$

for all $h, v \in \chi^*$.

Assumption 2. *The Hessian operator $\boldsymbol{J}_x$ defined as above exists for all $x \in \chi^*$ and there is a constant $c > 0$ such that*

$$\|\boldsymbol{J}_0\| \leq c.$$

Definition 1. *(Banach space of type 2)*
A Banach space χ is said to be of type 2 if there is a constant $b > 0$ such that for any $n \geq 1$ and independent zero mean random elements $V_1, \ldots, V_n$ in χ satisfying $\mathbb{E}(\|V_i\|^2) < \infty$, for all $i = 1, \ldots, n$, we have

$$\mathbb{E}(\|\sum_{i=1}^{n} V_i\|^2) \leq b \sum_{i=1}^{n} \mathbb{E}(\|V_i\|^2).$$

We will focus on this particular functional space since, as said in Chakraborty and Chaudhuri (2014a), type 2 Banach spaces are the only Banach spaces where the central limit theorem holds for every sequence of independent and identically distributed random elements, whose squared norms have finite expectations (see also Theorem 2.8, p. 53, in Bosq, 2000). Remark that the Hilbert and the L^p spaces with $p \in [2, \infty)$ are Banach of type 2 spaces (see for more details p. 51 in Bosq, 2000 or p. 159 in Araujo and Giné, 1980).

Then, using the previous notations and definition, the asymptotic normality of MED is given by the following theorem.

Theorem 1. *(Asymptotic Gaussianity of MED)*
Let $N = m + n$ and $m/N \to \lambda \in (0, 1)$ as $m, n \to \infty$. Assume that the bidual χ^{**} space is a separable and type 2 Banach space. Then, under assumptions (1) and (2), for any two probability measures P and Q on χ,

$$(mn/N)^{1/2}(MED - \mu) \quad \textit{converges weakly to} \quad G\left(0, \lambda\Gamma_1 + (1-\lambda)\Gamma_2\right)$$

as $m, n \to \infty$.

The proof of Theorem 1 is available in appendix A.1.

Remark. *For easier understanding, we develop μ in the univariate case. As defined before,*

$$\mu = \mathbb{E}\left[\boldsymbol{SGN}^*_{\{F_X(Y)\}}\right] = \mathbb{E}\left[\boldsymbol{SGN}^*_{\left\{\mathbb{E}\left[\boldsymbol{SGN}_{\{Y-X\}}|Y\right]\right\}}\right].$$

Thus,

$$\begin{aligned} F_X(Y) &= \mathbb{E}\big[\mathbf{SGN}_{\{Y-X\}}\big|Y\big] \\ &= \mathbb{E}\big[\mathbb{1}_{\{Y>X\}}\big|Y\big] - \mathbb{E}\big[\mathbb{1}_{\{Y<X\}}\big|Y\big] \\ &= 2\mathbb{E}\big[\mathbb{1}_{\{Y>X\}}\big|Y\big] - 1 \\ &= 2\tilde{F}_X(Y) - 1, \end{aligned}$$

where $\tilde{F}_X(\cdot)$ is the conditional distribution function of X given Y i.e. the projection of $\mathbb{1}_{\{X<\cdot\}}$ onto the subspace spanned by Y. Then, we obtain

$$\begin{aligned}\mu &= \mathbb{E}\left[sign(2\tilde{F}_X(Y)-1)\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\left\{2\tilde{F}_X(Y)-1>0\right\}}\right] - \mathbb{E}\left[\mathbb{1}_{\left\{2\tilde{F}_X(Y)-1<0\right\}}\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\left\{F_X(Y)>\frac{1}{2}\right\}}\right] - \mathbb{E}\left[\mathbb{1}_{\left\{F_X(Y)<\frac{1}{2}\right\}}\right]\\ &= \mathbb{E}\left[\mathbb{1}_{\left\{Y>\tilde{F}_X^{-1}\left(\frac{1}{2}\right)\right\}}\right] - \mathbb{E}\left[\mathbb{1}_{\{Y<\tilde{F}_X^{-1}\left(\frac{1}{2}\right)\}}\right]\\ &= 1-2\mathbb{E}\left[\mathbb{1}_{\{Y<\tilde{F}_X^{-1}\left(\frac{1}{2}\right)\}}\right]\\ &= 1-2\,G\left(\tilde{F}_X^{-1}\left(1/2\right)\right),\end{aligned}$$

where G is the distribution function of Y. Under H_0, we have $\mu = 0$.

2.3 Computing the significance

In this subsection, we propose two methods to compute the significance of the MED test statistic: the first one is based on Theorem 1 defined as above and the second one is based on Monte-Carlo simulations (Dwass, 1957). A comparison between these procedures will be presented in Section 3.

2.3.1 Using the asymptotic distribution

The significance of the test based on the MED statistic can rely on the asymptotic distribution of the statistic exhibited by Theorem 1. Since $\mu = 0$ under H_0, we shall reject the null hypothesis if $\left\|(mn/N)^{1/2}\mathrm{MED}\right\| > q_\alpha$ where q_α denotes the $(1-\alpha)$ quantile of the limiting distribution $\|G\left(0, \lambda\Gamma_1 + (1-\lambda)\Gamma_2\right)\|$ and α is the asymptotic size of the test. In order to compute this quantile, we need to derive the covariance operators Γ_1 and Γ_2 and the norm of the asymptotic distribution given by Theorem 1. We describe this estimation step when χ is a separable Hilbert space with norm $\|\cdot\|_\chi$.
Let $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ be two χ-valued samples. Thus, we proceed as follows.

- **Estimating the asymptotic covariance operator of the Gaussian process :**
  Since, the operator Γ_1 given by (5) is equal to

  $$\Gamma_1 = \mathbb{E}\left(\mathbf{SGN}^*_{\{F_X(Y)\}} \otimes \mathbf{SGN}^*_{\{F_X(Y)\}}\right) - \mu\otimes\mu,$$

  its empirical estimator is

  $$\hat{\Gamma}_1 = \frac{1}{n-1}\sum_{i=1}^{n}\left[\left(\frac{\hat{F}_X(Y_i)}{\|\hat{F}_X(Y_i)\|_\chi} - \hat{\mu}\right)\otimes\left(\frac{\hat{F}_X(Y_i)}{\|\hat{F}_X(Y_i)\|_\chi} - \hat{\mu}\right)\right],$$

where for all $i = 1, \ldots, n$, $\hat{F}_X(Y_i) = \frac{1}{m} \sum_{j=1}^{m} \frac{Y_i - X_j}{\|Y_i - X_j\|_{\mathcal{X}}}$ and $\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \frac{\hat{F}_X(Y_i)}{\|\hat{F}_X(Y_i)\|_{\mathcal{X}}}$.

In the usual case where the sample functions are observed in k equispaced points on $[0, 1]$, this covariance operator turns into a $k \times k$ covariance matrix. In addition, Γ_2 which is given by (6) can be estimated by $\hat{\Gamma}_2$, whose formulation is derived similarly to $\hat{\Gamma}_1$. Consequently, the asymptotic covariance operator of MED, defined as $\lambda\Gamma_1 + (1 - \lambda)\Gamma_2$, is estimated by

$$\hat{\lambda}\hat{\Gamma}_1 + (1 - \hat{\lambda})\hat{\Gamma}_2,$$

where $\hat{\lambda} = \frac{m}{N}$ and $N = m + n$.

- **Estimating the asymptotic distribution of the test statistic :**

  As stated in Horváth et al. (2013) and Chakraborty and Chaudhuri (2015), Theorem 1 implies that

  $$\left\| \left(\frac{mn}{N}\right)^{1/2} \mathrm{MED} \right\|_{\mathcal{X}}^2 = \sum_{k=1}^{\infty} \tau_k N_k,$$

  where for $k \geq 1$, the N_k's are independent chi-squared random variables, each with one degree of freedom, and the τ_k's are the eigenvalues of $\lambda\Gamma_1 + (1 - \lambda)\Gamma_2$. This agrees with the Karhunen-Loève expansion (for more details, see Theorem 2.1 of Horváth and Kokoszka, 2012 or Theorem IV.2.4 and Proposition 1.9 of Vakhania et al., 1987). Replacing the covariance operator $\lambda\Gamma_1 + (1-\lambda)\Gamma_2$ by the estimator we described earlier, we obtain a finite number of eigenvalues and, following Chakraborty and Chaudhuri (2015), we only retain the positive eigenvalues. The asymptotic distribution of the test statistic is thus approximated by the distribution of a finite sum of independent chi-squared random variables which can be easily generated.

2.3.2 Using random permutations

The method based on the asymptotic distribution suffers two limitations: the need to estimate the covariance operators Γ_1 and Γ_2 and the distance to the asymptotic distribution when m and n are quite small, which is often the case when comparing two samples of functions. Consequently, we may consider a test procedure based on Monte-Carlo simulations allowing to give an approximation of the null distribution (Dwass, 1957). The procedure is based on the following steps :

1. Let $X_{\text{obs}} = (X_1, \ldots, X_m)$ and $Y_{\text{obs}} = (Y_1, \ldots, Y_n)$. Among the $m + n$ observations of $(X_{\text{obs}}, Y_{\text{obs}})$, m of them are randomly chosen to create X_{perm} and the n others to create Y_{perm}.

2. The simulated median statistic S_{perm} is then computed using X_{perm} and Y_{perm} instead of X and Y.

3. Based on n_{perm} random permutations, the *p-value* of the median statistic is given by

$$p_{value} = \frac{1 + \sum_{l=1}^{n_{\text{perm}}} \mathbb{1}_{\left\{S_{\text{perm}}^{(l)} > S\right\}}}{n_{\text{perm}} + 1},$$

   where S is the value of the statistic computed using the observed data and $S_{perm}^{(l)}$ the value computed using the l^{th} random permutation.

4. We will reject H_0 when the *p-value* is below the level of the test.

2.4 Asymptotic power under shrinking location shifts

In this section, we give the asymptotic distribution of the test statistic MED under appropriate sequences of shrinking location shifts. In order to do that, we suppose that Y is distributed as $X + \Delta_N$, where

$$\Delta_N = \delta \left(\frac{mn}{N}\right)^{-1/2} \tag{7}$$

for some nonzero fixed $\delta \in \chi$ and $N = m + n \geq 1$ is the total size of the two samples. As said in Chakraborty and Chaudhuri (2015), these alternative hypotheses choice has been proved to be contiguous to the null one and it is useful to find nondegenerate asymptotic distributions of the test statistics under these alternatives (see Oja, 1999).
In order to derive the distribution of the median statistic under these alternative hypotheses, we need two more assumptions.

Assumption 3. *We assume that the norm in χ is twice Gateaux differentiable at every $y \neq 0$. In addition, we suppose that the Hessian of the map $y \longmapsto \mathbb{E}[\|y + Y - Z\|_\chi | Y]$, at $y \in \chi$, denoted by $\tilde{\boldsymbol{J}}_y : \chi \to \chi^*$, exists where Z is an independent copy of Y.*

Assumption 4. *Since we have assumed that the norm in χ^* is twice Gateaux differentiable at every $x \neq 0$ (Assumption 1), we suppose here that the Hessian of the function $u \mapsto \mathbb{E}\left[\left\|u + \mathbb{E}\big[\boldsymbol{SGN}_{\{Y-Z\}}\big|Y\big]\right\|_{\chi^*}\right]$, denoted by $\boldsymbol{H}_u : \chi^* \to \chi^{**}$, exists where Z is an independent copy of Y.*

Finally, we obtain the following theorem.

Theorem 2. *Let $N = m + n$ and $m/N \to \lambda \in (0, 1)$ as $m, n \to \infty$. Suppose that χ^{**} is a separable and type 2 Banach space. Assume that the distributions of X and Y are nonatomic. Then, under assumptions (1), (2), (3) and (4) and under the sequence of shrinking location shifts defined as (7),*

$$(mn/N)^{1/2} MED \ \text{ converges weakly to } \ G\left(\boldsymbol{H}_0(\tilde{\boldsymbol{J}}_0(\delta)), \lambda\Gamma_1 + (1-\lambda)\Gamma_2\right)$$

as $m, n \to \infty$.

The proof of Theorem 2 is available in appendix A.1.

2.4.1 Computing the power

For evaluating the asymptotic power of the test based on MED statistic we will use Theorem 2. According to the latter, we shall first estimate the asymptotic mean and then follow the procedure described in subsection 2.3.1. In order to do that, we shall first estimate the operators $\mathbf{H}_0$ and $\tilde{\mathbf{J}}_0$ to derive an estimator of the asymptotic mean of MED under the sequence of shrinking location shifts defined as (7). We describe this estimation step when χ is a separable Hilbert space with norm $\|\cdot\|_\chi$.
Let $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ be two χ-valued samples. Thus, we have

$$\tilde{\mathbf{J}}_0 = \mathbb{E}\left[\frac{1}{\|Y - X\|_\chi}\mathbf{I}_\chi - \frac{(Y-X)\otimes(Y-X)}{\|Y-X\|_\chi^3}\middle| Y\right],$$

where $\mathbf{I}_\chi$ stands for the identity operator on χ. Thus, an estimator of $\tilde{\mathbf{J}}_0$ is given by

$$\hat{\tilde{\mathbf{J}}}_{0,i} = \frac{1}{m}\sum_{j=1}^{m}\left[\frac{1}{\|Y_i - X_j\|_\chi}\mathbf{I}_\chi - \frac{(Y_i - X_j)\otimes(Y_i - X_j)}{\|Y_i - X_j\|_\chi^3}\right],$$

for all $i = 1, \dots, n$. Furthermore, since

$$\mathbf{H}_0 = \mathbb{E}\left[\frac{1}{\|F_X(Y)\|_\chi}\mathbf{I}_\chi - \frac{F_X(Y)\otimes F_X(Y)}{\|F_X(Y)\|_\chi^3}\right],$$

it can be estimated by

$$\hat{\mathbf{H}}_0 = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{\|\hat{F}_X(Y_i)\|_\chi}\mathbf{I}_\chi - \frac{\hat{F}_X(Y_i)\otimes\hat{F}_X(Y_i)}{\|\hat{F}_X(Y_i)\|_\chi^3}\right].$$

Consequently, the asymptotic mean of MED under the sequence of shrinking location shifts $\mathbf{H}_0(\tilde{\mathbf{J}}_0(\delta))$ can be estimated by

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\hat{\tilde{J}}_{0,i}(\delta)}{\|\hat{F}_X(Y_i)\|_\chi} - \frac{\left\langle \hat{F}_X(Y_i), \hat{\tilde{J}}_{0,i}(\delta)\right\rangle \hat{F}_X(Y_i)}{\|\hat{F}_X(Y_i)\|_\chi^3}\right].$$

A comparison with the test statistic introduced by Chakraborty and Chaudhuri (2015) is given in subsection 3.1.2. In addition, we give more details about the estimation of the covariance operators and the asymptotic mean under the sequence of shrinking location shifts of the Wilcoxon-Mann-Whitney test statistic introduced by Chakraborty and Chaudhuri (2015) in appendix A.2 .

3 Applications

3.1 A simulation study

In this subsection, we aim to compare the power of the two median statistics introduced in the previous section with those of the tests available in Chakraborty and Chaudhuri (2015),

Cuevas et al. (2004) and Horváth et al. (2013).
We set the separable Hilbert space $\chi = L^2[0,1]$. The test introduced by Chakraborty and Chaudhuri (2015) is based on the Wilcoxon-Mann-Whitney statistic, which is defined as a U-statistic (Borovskikh, 1996) like in the univariate case and can be rewrited in χ as follows

$$\mathrm{WMW} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\mathbf{SGN}_{Y_i - X_j} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{Y_i - X_j}{\|Y_i - X_j\|_\chi}.$$

The test statistic introduced by Cuevas et al. (2004) is defined by

$$\mathrm{CFF} = m\|\bar{X} - \bar{Y}\|_\chi^2,$$

where $\bar{X}$ and $\bar{Y}$ are the empirical means of the X_j's and Y_i's respectively for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$ and μ_g is the empirical mean of the pooled sample of the X_j's and Y_i's. Horváth et al. (2013) introduced the test statistics defined as follows

$$\mathrm{HKR1} = \frac{mn}{N}\sum_{l=1}^{p}\frac{\langle \bar{X} - \bar{Y}, \hat{\varphi}_l\rangle}{\hat{\lambda}_l},$$

and

$$\mathrm{HKR2} = \frac{mn}{N}\sum_{l=1}^{p}\langle \bar{X} - \bar{Y}, \hat{\varphi}_l\rangle.$$

Here, $\langle ., . \rangle$ denotes the inner product in χ, the $\hat{\lambda}_l$'s are the eigenvalues of the empirical pooled covariance of the X_j's and the Y_i's sorted in decreasing order of magnitude and the $\hat{\varphi}_l$'s are the corresponding empirical eigenfunctions. In this subsection, we used the usual empirical pooled covariance for the two tests HKR1 and HKR2 given by Horváth et al. (2013) and the numbers of projection directions p are chosen using the cumulative variance method that they provide. Now, let us consider the decomposition

$$X = \sum_{k=1}^{\infty} Z_k e_k,$$

where for all $k \geqslant 0$, $e_k = \sqrt{2}sin(t/\sigma_k)$ is an orthonormal basis of χ, $\sigma_k = ((k-0.5)\pi)^{-1}$ and the Z_k's are independent random variables which correspond to the projection of X on the Karhunen-Loève basis (Karhunen, 1947; Lévy and Loève, 1948). We have considered four scenarios:

(i) A standard Brownian motion (sBm), i.e. Z_k/σ_k follows a $\mathcal{N}(0,1)$ distribution.

(ii) A centered t process on $[0,1]$ with 5 degrees of freedom, i.e. $Z_k/\sigma_k \backsim t(5)$.

(iii) A Cauchy distribution with parameters 0 and 1, i.e. $Z_k/\sigma_k \backsim \mathcal{C}(0,1)$.

(iv) A double exponential distribution with parameters 0 and 1, i.e. $Z_k/\sigma_k \backsim \mathcal{D}exp(0,1)$.

The scenarios (i) and (ii) are studied in Chakraborty and Chaudhuri (2015) and we have chosen the scenarios (iii) and (iv) to study the performance of the differents tests using more heavy-tailed distributions.

3.1.1 Finite-size powers

Assume that Y is distributed as $X + \Delta$ and under the alternative hypotheses $H_1 : \Delta \neq 0$. Three choices are considered, namely : $\Delta_1(t) = c$, $\Delta_2(t) = ct$ and $\Delta_3(t) = ct(1-t)$ where $c > 0$ for all $t \in [0, 1]$. Figure 1 shows examples of simulated data.

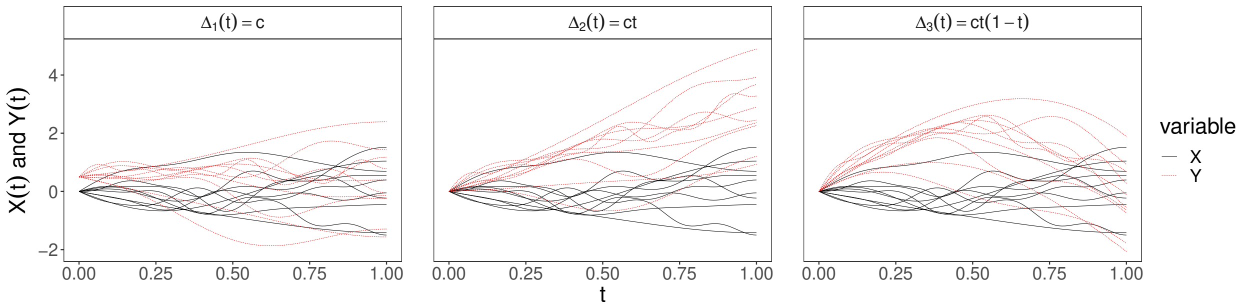

Figure 1: Examples of generated data using the scenario (i) with $c = 0.5$ (left panel), $c = 3$ (middle panel) and $c = 8$ (right panel). In black: 10 samples of X. In red: 10 samples of Y.

Before computing the powers of the tests previously introduced, we have derived the size of each of these tests, i.e. the probability for rejecting the null hypothesis when it is true. We have chosen $n_{sim} = 1000$ random simulations of (X, Y). For each simulated dataset, all test statistics and their critical values are derived in the same way as described in subsection 2.3 using the asymptotic and the permutation methods:

- To apply the asymptotic method: critical values of the test statistic MED are derived as described above. Similarly, those of WMW, HKR1 and CFF are calculated using their associated asymptotic theorems described in Chakraborty and Chaudhuri (2015), Cuevas et al. (2004) and Horváth et al. (2013) respectively.

- To apply the permutation procedure: we have used $n_{perm} = 999$ random permutations. The hypothesis H_0 is rejected if $p_{value} < \alpha$, where α is the significance level which is chosen equal to 0.05.

Then, using different sample sizes, we obtain the results gathered in Figure 2.

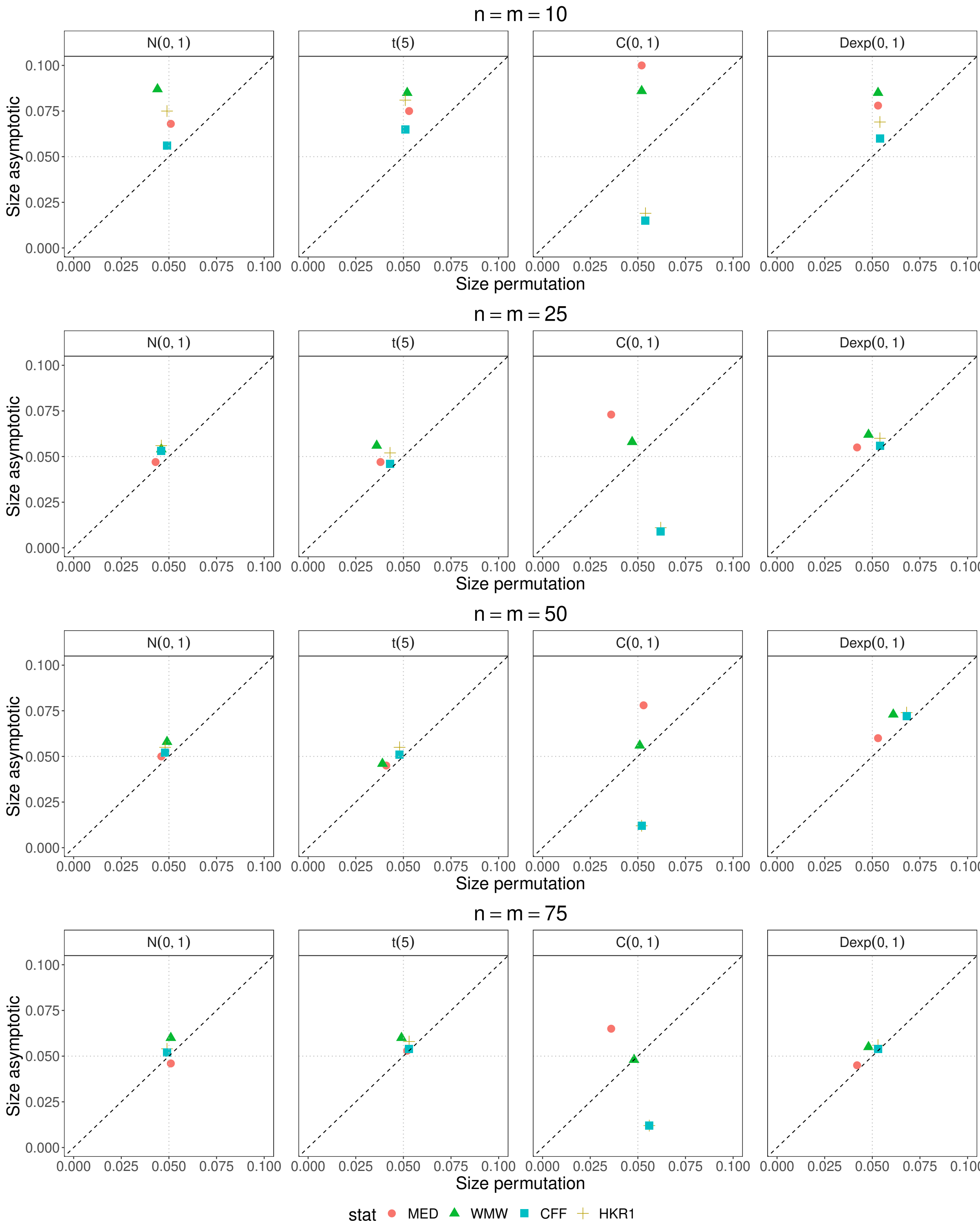

Figure 2: Sizes computed with the permutation method versus corresponding sizes computed with the asymptotic method for the different test statitics, MED, WMW, HKR1 and CFF, using the different scenarios (i), (ii), (iii) and (iv).

We see in Figure 2 that, using the asymptotic method and when the sample sizes are small ($n = m = 10$), the sizes of all the tests are different from the 5% nominal level whatever the distribution. However, the ones obtained using the permutation method are close to 5%: this seems logical since the asymptotic method is more adapted for large sample sizes. Moreover, using Cauchy distribution, we remark that the sizes of test statistics HKR1 and CFF are different from the nominal level whatever the sample size when using the asymptotic method: this may result from difficulties in estimating the covariance operators of these tests when the distribution has heavy tails.

Because of these results, we decided to first focus on simulations with limited sample sizes, using the permutation method to derive the corresponding statistical power of the different tests. The asymptotic method will be used in the next subsection to derive asymptotic powers. We set $m = n = 10$ and each sample curve is observed at 100 equidistant points on $[0, 1]$. Figure 3 presents the corresponding power results.

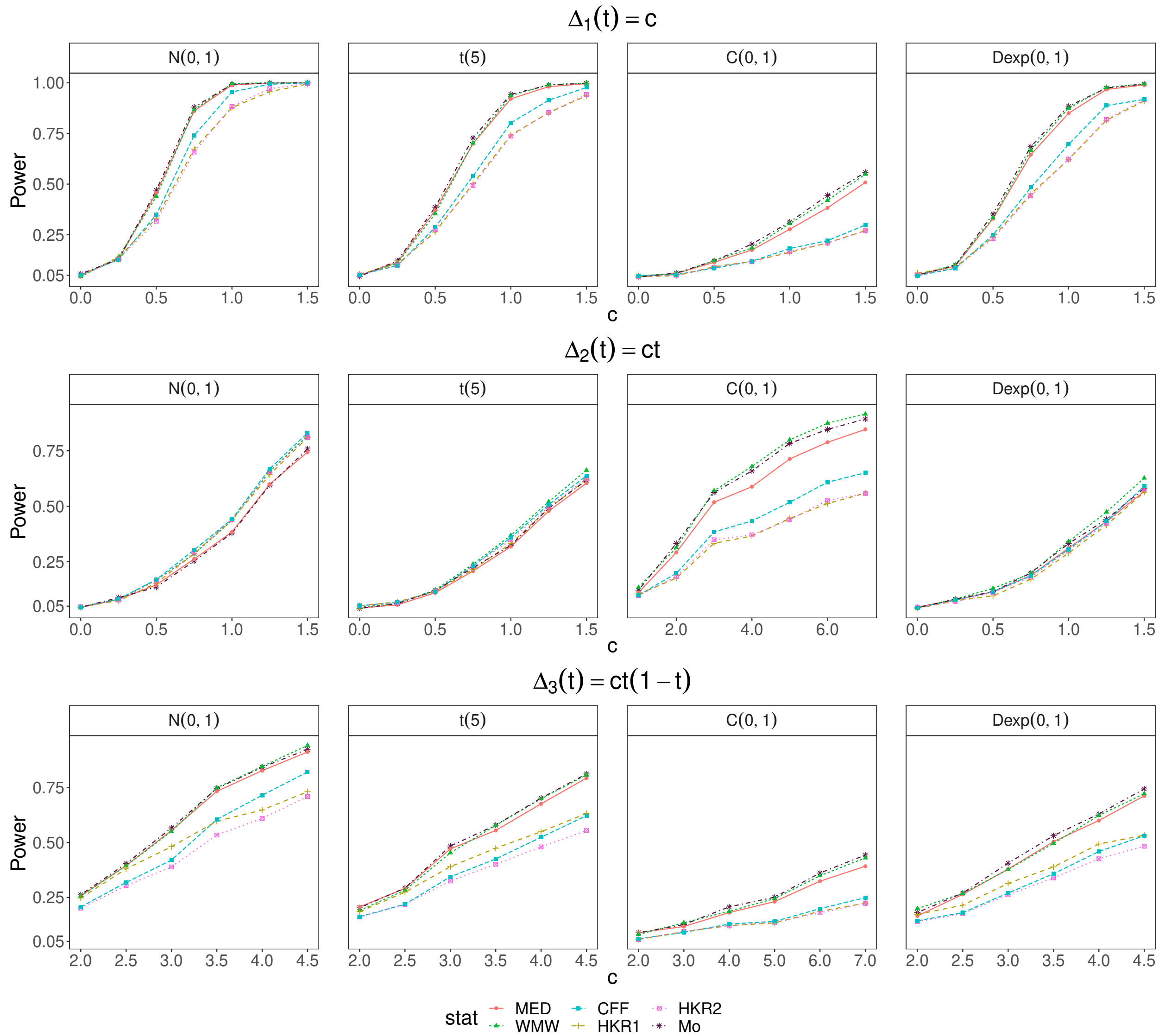

Figure 3: Values of the statistical power for the tests using MED, Mo, WMW, CFF, HKR1 and HKR2 statistics depending on the values of c, when $\Delta_1(t) = c$, $\Delta_2(t) = ct$ and $\Delta_3(t) = ct(1-t)$ using the different scenarios (i), (ii), (iii) and (iv), $n_{\text{perm}} = 999$, $n_{\text{sim}} = 1000$ and $n = m = 10$.

From Figure 3, we can say that :

- With $\Delta_1(t)$, the tests based on Mo, WMW and MED have similar powers under all distributions except the Cauchy one where the test based on MED is less powerful for large values of c. We can also see that the tests using the CFF, HKR1 and HKR2 statistics are less powerful than the ones based on MED, Mo and WMW for large values of c in all the settings we considered.

- With $\Delta_2(t)$ and under all the distributions except the Cauchy one, all the tests have similar powers for small values of c. With the $\mathcal{N}(0,1)$ distribution and for large values of c, the parametric test based on CFF outperforms all the other tests. However, using the Student process, the test based on WMW is more powerful than others. We notice that, using the heavy-tailed distributions $\mathcal{C}(0,1)$ and $\mathcal{D}exp(0,1)$, the proposed tests based on Mo and MED outperform the parametric tests CFF, HKR1 and HKR2.

- With $\Delta_3(t)$, the test using Mo statistic outperforms the test based on WMW against heavy-tailed distributions $t(5)$, $\mathcal{C}(0,1)$ and $\mathcal{D}exp(0,1)$ for large values of c and it has a similar power against sBm distribution. We remark also that the power of the test based on MED is very similar to the power of the tests based on WMW and Mo for small values of the shift c and for all the distributions. However, for large values of c, it is similar to the tests based on WMW and Mo for all the distributions except the $\mathcal{C}(0,1)$ one. We can also see that using the four distributions, the nonparametric tests based on spatial ranks using MED, Mo and WMW statistics performs better than the parametric mean-based ones using CFF, HKR1 and HKR2 statistics.

3.1.2 Asymptotic powers

In this subsection, we compare the asymptotic powers of the tests based on MED and WMW since their asymptotic distributions are known under the sequence of shrinking location shifts (Theorem 2 introduced above and Theorem 2 of Chakraborty and Chaudhuri, 2015).
To do so, recall that Y is distributed as $X+\Delta_N$, where Δ_N is given by (7). We have considered three choices of $\delta \in \chi$, namely $\delta_1(t) = c$, $\delta_2(t) = ct$ and $\delta_3(t) = ct(1-t)$, where $t \in [0,1]$ and $c > 0$. The ranges of δ_2 and δ_3 being smaller than the range of δ_1, it is combined with larger values of c. For evaluating the asymptotic powers of these tests, we have used 1000 sample functions from different distributions (i), (ii), (iii) and (iv). The asymptotic covariance of the tests based on MED and WMW are estimated respectively as described in subsection 2.3.1 and in appendix A.2. The estimators of the asymptotic means are respectively derived as explained in subsection 2.4.1 and in appendix A.2. Consequently, the asymptotic powers of the tests based on MED and WMW are computed from the Gaussian distributions given respectively by our Theorem 2 and Theorem 2 of Chakraborty and Chaudhuri (2015) with the appropriate estimated parameters. Figure 4 shows the corresponding results.

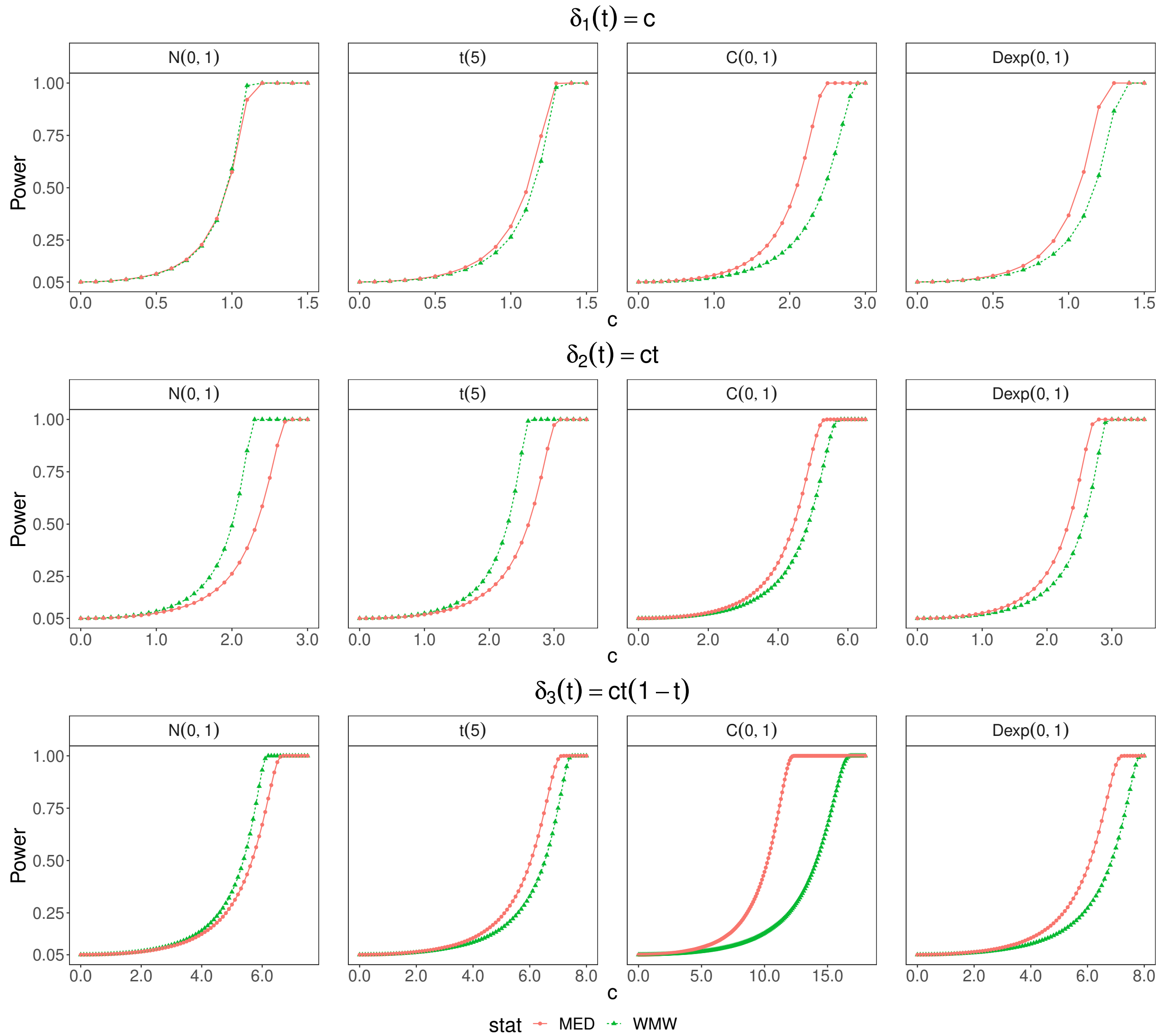

Figure 4: Values of the statistical power for the tests using MED and WMW statistics depending on the values of c, when $\delta_1(t) = c$, $\delta_2(t) = ct$ and $\delta_3(t) = ct(1-t)$ and $n = m = 1000$ using the different scenarios (i), (ii), (iii) and (iv).

From Figure 4, we see that both tests achieve the 5% nominal level asymptotically using different location shift models and different types of distributions and its asymptotic powers are equal when $c = 0$. Moreover,

- Under $\delta_1(t)$, our test based on MED and the one based on WMW have similar asymptotic powers in the case of the Gaussian distribution. However, our test outperforms the WMW one with $t(5)$ distribution and even more powerful with $\mathcal{C}(0,1)$ and $\mathcal{D}exp(0,1)$ which are more heavy-tailed.

- Under $\delta_2(t)$, with $\mathcal{C}(0,1)$ and $\mathcal{D}exp(0,1)$ distributions, our test is more powerful compared to the WMW one. However, with $\mathcal{N}(0,1)$ and $t(5)$ distributions, our test based on MED is less powerful using large values of c.

- Under $\delta_3(t)$, the asymptotic powers curves of our test and the WMW one are close with the $\mathcal{N}(0,1)$ distribution for small values of c. In all the other situations, our test outperforms the WMW one and especially with $\mathcal{C}(0,1)$ distribution where we can see a large difference between both curves.

3.2 An application to real data

In this subsection, we compare the two median tests based on MED and Mo with those based on WMW, CFF, HKR1 and HKR2 which are presented in the previous subsection using two datasets already analysed by Chakraborty and Chaudhuri (2014a) (for more details, see Ramsay and Silverman, 2005 and Ferraty and Vieu, 2006). In both datasets, each observation is an element in the separable Hilbert space $\chi = L^2[a,b]$.

3.2.1 Coffee data

This dataset can be downloaded from `http://www.cs.ucr.edu/~eamonn/time_series_data/`. It contains the spectroscopy values for 14 samples for two different types of coffee beans (Arabica and Robusta) recorded at 286 wavelengths (see Figure 5).

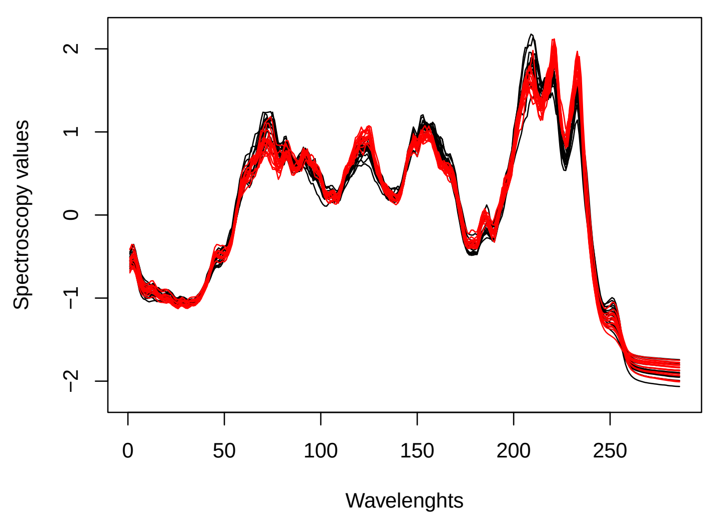

Figure 5: In red : Spectroscopy curves of Robusta beans. In black : Spectroscopy curves of Arabica beans.

Based on 999 random permutations, the p-values obtained on this dataset are equal to 0.001 except the one based on HKR2 which equals 0.003. The p-values obtained using the asymptotic method are also zero up to two decimal places. This leads us to reject the null hypothesis whatever the significance method. From Figure 5, the spectroscopy curves of the two coffee types are clearly different since the maximum values are not observed in the same wavelengths for Arabica and Robusta. Remark that the results given by Chakraborty and Chaudhuri (2014a) on the same dataset are completely different since their asymptotic p-values of the tests based on WMW (0.072), CFF (0.169), HKR1 (0.273) and HKR2 (0.273) fail to reject H_0.

3.2.2 Berkeley growth data

This dataset is available in the R package fda and contains the heights of 39 boys and 54 girls measured at 31 time points from age 1 to 18 (see Figure 6).

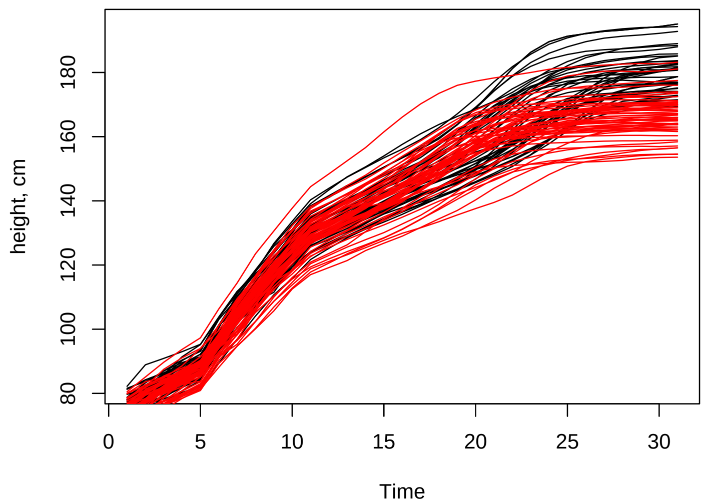

Figure 6: In red : Heights of the 54 girls. In black : Heights of the 39 boys.

All the p-values based on 999 random permutations and the asymptotic method are 0 up to two decimal places and are similar to the ones given by Chakraborty and Chaudhuri (2014a). These p-values exhibit a strong difference between the two distributions. So, we decided to evaluate the proportion of rejection of the null hypothesis to compare the behaviour of the different statistics when the level α is equal to 0.05. Such as done by Chakraborty and Chaudhuri (2014a), we have chosen randomly 20% subsamples of the 2 classes of the complete dataset and this subsampling was repeated 100 times. Results are given in Table 1.

As seen in Table 1, for this type of data, our tests based on Mo and MED statistics have the highest rate of rejection of the null hypothesis, close to the one obtained using

Statistic	MED	Mo	WMW	CFF	HKR1	HKR2
Proportion of rejection	1	1	0.99	0.87	0.28	0.25

Table 1: The proportions of rejection of the null hypothesis of the different test statistics MED, Mo, WMW, CFF, HKR1 and HKR2.

the statistic WMW of Chakraborty and Chaudhuri (2015) and much larger than the tests proposed by Cuevas et al. (2004) namely CFF and Horváth et al. (2013) namely HKR1 and HKR2.

4 Discussion

Nowadays and with the development of modern technology, scientists often observe functional datasets instead of multivariate ones. As a consequence, there is a need for testing procedures adapted to these infinite dimensional data. In this paper, we have proposed an extension of an existing nonparametric test based on ranks in the infinite dimensional spaces to compare two datasets (samples of curves). This is a median test in the functional case, similar to the rank-based test proposed by Capéraà and Cutsem (1988) and Van der Vaart (1998) in the univariate case. It can be noted that we introduce the notion of ranking functional elements through a sign function. The median test is one way to use this sign function for ranking functional elements but other possibilities have been investigated such as functional depth (Chakraborty and Chaudhuri, 2014b; Gijbels and Nagy, 2017; Estévez-Pérez and Vieu, 2021): these could lead to other types of nonparametric tests for comparing samples of functions that we shall investigate in a future work.
First, we proposed two median statistics in a Banach space then their equivalent in a particular case which is an Hilbert space. Second, we derived the asymptotic Gaussianity of one of the proposed median statistics under the null hypothesis and local alternatives proposed in section 2. Remark that, our median statistic not being a U-statistic (see Borovskikh, 1996) such as the Wilcoxon-Mann-Whitney one but based on two sign functions instead of one, this increases the complexity of the proofs. Computing significance and powers is possible using two different procedures, either based on the asymptotic distribution or on random permutations.
The application to simulated and real data shows that the median tests have good performance compared to the Wilcoxon-Mann-Whitney test proposed by Chakraborty and Chaudhuri (2015), the ANOVA test based on CFF and the mean-based tests introduced by Horváth et al. (2013). Moreover, when the distribution of the processes is heavy-tailed, the median test is as powerful as the WMW test for moderate sample size and asymptotically more powerful than any other test, either parametric or nonparametric.
We are willing to propose a R package in which these different parametric and nonparametric tests would be available and we are working on it. A perspective would be to develop the tests introduced in this paper and the Wilcoxon-Mann-Whitney one to compare more than two datasets in infinite dimensional space. To do so, we may follow the strategy used by Oja

(2010) in the multivariate case: this would become a multiple location test.
Recently, Smida et al. (2022) have proposed a nonparametric spatial scan statistic for detecting spatial clusters using functional data. This scan statistic was constructed using the Wilcoxon-Mann-Whitney two-sample test for functional data of Chakraborty and Chaudhuri (2015) and it is implemented in a R package named HDSpatialScan (Frévent et al., 2021). Another perspective would be to develop a new nonparametric scan statistic in the functional case using the median statistics introduced in this paper and to compare it with the nonparametric one based on the Wilcoxon-Mann-Whitney statistic.

A Appendix

A.1 Proof of theorems

Proof of theorem 1. The median statistic (2) is

$$\begin{aligned} \mathrm{MED} &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{SGN}^{*}_{\left\{\frac{1}{m}\sum\limits_{j=1}^{m}\mathbf{SGN}_{\{Y_i-X_j\}}\right\}} \\ &= \frac{1}{n}\sum_{i=1}^{n}\phi(F_m(Y_i)). \end{aligned}$$

We remark that the random elements $\mathbf{SGN}_{\{Y_i-X_j\}}$ are not independent for all $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,m\}$. To ensure the independence, our strategy is to use the conditional expectations $F_X(y)$ and $F_Y(x)$ defined respectively by (3) and (4) for all $x, y \in \chi$.
Thus, we can use the following decomposition :

$$\mathrm{MED} - \mathbb{E}[\mathrm{MED}|X_j; j=1,\dots,m] = L_n^{'} + L_m^{''} + K_{m,n}^{'},$$

where

$$L_n^{'} = \frac{1}{n}\sum_{i=1}^{n}\left[\phi(F_X(Y_i)) - \mathbb{E}(\phi(F_X(Y_i)))\right], \tag{8}$$

$$L_m^{''} = \frac{1}{m}\sum_{j=1}^{m}\left[\phi(F_Y(X_j)) - \mathbb{E}(\phi(F_Y(X_j)))\right] \tag{9}$$

and

$$K_{m,n}^{'} = \mathrm{MED} - \mathbb{E}[\mathrm{MED}|X_j; j=1,\dots,m] - L_n^{'} - L_m^{''}. \tag{10}$$

Remark that the same decomposition is used in subsection 14.1.1 of Van der Vaart (1998) for $\mathbb{R}$-valued random variables. Furthermore, consider

$$K_{m,n}^{''} = \mathbb{E}[\mathrm{MED}|X_j; j=1,\dots,m] - \mathbb{E}[\phi(F_X(Y))]. \tag{11}$$

Hence, we can write the following decomposition :

$$\mathrm{MED} - \mu = L_n^{\prime} + L_m^{''} + R_{m,n}^{\prime}, \tag{12}$$

where

$$R'_{m,n} = K^{'}_{m,n} + K^{''}_{m,n}.$$

The proof of the asymptotic distribution of MED can be split into four steps :

- **Step 1:** Show that L'_n converges in *law* to a Gaussian element.
- **Step 2:** Show that $L^{''}_m$ converges in *law* to a Gaussian element.
- **Step 3:** Show that $R'_{m,n}$ converges in *probability* to 0.
- **Step 4:** Conclude the asymptotic normality of MED.

Step 1 : Asymptotic behavior of L'_n

Let

$$L^{'}_n = \frac{1}{n}\sum_{i=1}^{n}\left[\phi(F_X(Y_i)) - \mathbb{E}(\phi(F_X(Y_i)))\right].$$

We want to prove here the asymptotic Gaussianity of $L^{'}_n$. For this purpose, for all $i = 1, \ldots, n$, let us write the sequence

$$\begin{aligned}\psi_n(Y_i) &= n^{-1/2}\left[\phi(F_X(Y_i)) - \mathbb{E}(\phi(F_X(Y_i)))\right] \\ &= n^{-1/2}\left[\mathbf{SGN}^*_{\{\mathbb{E}[\mathbf{SGN}_{\{Y_i-X\}}|Y_i]\}} - \mathbb{E}\left[\mathbf{SGN}^*_{\{\mathbb{E}[\mathbf{SGN}_{\{Y_i-X\}}|Y_i]\}}\right]\right].\end{aligned}$$

Note that $\mathbb{E}\left[\psi_n(Y_i)\right] = 0$. In order to show the asymptotic Gaussianity of $\sum_{i=1}^{n}\psi_n(Y_i)$, we check that the triangular array $\{\psi_n(Y_1), \ldots, \psi_n(Y_n)\}_{n=1}^{\infty}$ of rowwise independent and identically distributed random elements satisfies the three conditions of Corollary 7.8 in Araujo and Giné (1980).

- **Condition 1 :** Let us show that

  $$\forall\epsilon > 0, \lim_{n\to+\infty}\sum_{i=1}^{n}\mathbb{P}\left(\left\|\psi_n(Y_i)\right\|_{\chi^{**}} > \epsilon\right) = 0.$$

  Using the Bienaymé-Tchebychev inequality, we obtain : for any $\epsilon > 0$,

  $$\begin{aligned}\sum_{i=1}^{n}\mathbb{P}\left(\left\|\psi_n(Y_i)\right\|_{\chi^{**}} > \epsilon\right) &\leq \sum_{i=1}^{n}\frac{\mathbb{E}\left[\left\|\mathbf{SGN}^*_{\{F_X(Y_i)\}} - \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right]\right\|^3_{\chi^{**}}\right]}{\epsilon^3 n^{3/2}} \\ &\leq \frac{8}{\epsilon^3 n^{1/2}} \underset{n\to+\infty}{\longrightarrow} 0.\end{aligned}$$

  For the last inequality, we have used the fact that $\left\|\mathbf{SGN}^*_{\{x\}}\right\|_{\chi^{**}} \leq 1$ for all $x \in \chi^*$. Thus, the first condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

- **Condition 2 :** Let us show that

$$\forall f \in \chi^{***}, \lim_{n\to+\infty} \sum_{i=1}^{n} \mathbb{E}\left[f^2\left(\psi_n(Y_i) - \mathbb{E}(\psi_n(Y_i))\right] = \Gamma_1(f,f) < \infty.$$

Let us fix $f \in \chi^{***}$. Since f is linear, we may write

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\left[f^2\left[\psi_n(Y_i) - \mathbb{E}(\psi_n(Y_i))\right]\right] &= \sum_{i=1}^{n} \mathbb{E}\left[f^2\left[\psi_n(Y_i)\right]\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[f^2\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}} - \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right]\right)\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\left(f\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right) - f\left(\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right]\right)\right)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\left(f\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right) - \mathbb{E}\left[f\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right)\right]\right)^2\right].
\end{aligned}$$

We consider now, for all $i = 1, \ldots, n$,

$$W_i := f\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right) = f\left(\mathbf{SGN}^*_{\{\mathbb{E}\left[\mathbf{SGN}_{\{Y_i - X\}}\middle|Y_i\right]\}}\right).$$

Hence, we get

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{E}\left[f^2\left[\psi_n(Y_i)\right]\right] &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\left[W_i - \mathbb{E}\left(W_i\right)\right]^2\right] \\
&= \mathbb{E}\left[\left[W_1 - \mathbb{E}\left(W_1\right)\right]^2\right] \\
&= \mathbb{E}\left[W_1^2\right] - \mathbb{E}^2\left[W_1\right] \\
&= \Gamma_1(f,f) < \infty,
\end{aligned}$$

where Γ_1 is defined as (5). Thus, the second condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

- **Condition 3 :** Let us show that

$$\lim_{k\to+\infty} \overline{\lim_{n\to+\infty}} \sum_{i=1}^{n} \mathbb{E}\left[d^2\left(\psi_n(Y_i) - \mathbb{E}\left[\psi_n(Y_i)\right], \mathcal{F}_k\right)\right] = 0,$$

where $\{\mathcal{F}_k\}_{k\geq 1}$ is a sequence of finite dimensional subspaces of χ^{**} such that $\mathcal{F}_k \subseteq \mathcal{F}_{k+1}$ for all $k \geq 1$ and the closure of $\cup_{k=1}^{\infty}\mathcal{F}_k$ is equal to χ^{**}. This sequence exists because of the separability of χ^{**}. Also, for any $x \in \chi^{**}$ and any $k \geq 1$, we define $d(x,\mathcal{F}_k) = \inf\{\|x-y\|_{\chi^{**}} : y \in \mathcal{F}_k\}$. It is easy to prove that for all $k \geq 1$, the map $x \longmapsto d(x,\mathcal{F}_k)$

is continuous and bounded on any closed ball in χ^{**}.
First, since $\mathbb{E}\left[\psi_n(Y_i)\right]=0$, we have

$$\begin{aligned}\sum_{i=1}^{n} & \mathbb{E}\left[d^2\left(\psi_n(Y_i),\mathcal{F}_k\right)\right] \\ = & \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y_i)\}}\right],\mathcal{F}_k\right)\right] \\ = & \mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_X(Y_1)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y_1)\}}\right],\mathcal{F}_k\right)\right].\end{aligned}$$

From the choice of the sequence $\{\mathcal{F}_k\}_{k\leq 1}$, we obtain $d(x,\mathcal{F}_k)\longrightarrow 0$ as $k\to\infty$ for any $x\in\chi^{**}$. Thus,

$$\lim_{k\to+\infty}\overline{\lim_{n\to+\infty}}\sum_{i=1}^{n}\mathbb{E}\left[d^2\left(\psi_n(Y_i),\mathcal{F}_k\right)\right]=0$$

and the third condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

Therefore, using Corollary 7.8 in Araujo and Giné (1980), $\sum_{i=1}^{n}\psi_n(Y_i)$ converges weakly to a centered Gaussian random element in χ^{**} as $m,n\to\infty$. Moreover, the asymptotic covariance is Γ_1 which was obtained while checking the second condition presented as above. Finally,

$$\sqrt{n}L'_n=n^{-1/2}\sum_{i=1}^{n}\left[\phi(F_X(Y_i))-\mathbb{E}\left[\phi(F_X(Y_i))\right]\right]\xrightarrow{\mathcal{L}}\mathbf{G}(0,\Gamma_1) \tag{13}$$

weakly as $m,n\to\infty$.

Step 2 : Asymptotic behavior of L''_m

Let

$$L''_m=\frac{1}{m}\sum_{j=1}^{m}\left[\phi(F_Y(X_j))-\mathbb{E}(\phi(F_Y(X_j)))\right]. \tag{14}$$

Similarly to the previous step, our goal in this part is to prove the asymptotic Gaussianity of L''_m. To do that, for all $j=1,\ldots,m$, we consider

$$\begin{aligned}\tilde{\psi}_m(X_j) &= m^{-1/2}\left[\phi(F_Y(X_j))-\mathbb{E}(\phi(F_Y(X_j)))\right] \\ &= m^{-1/2}\left[\mathbf{SGN}^*_{\left\{\mathbb{E}\left[\mathbf{SGN}_{\{Y-X_j\}}\middle|X_j\right]\right\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\left\{\mathbb{E}\left[\mathbf{SGN}_{\{Y-X_j\}}\middle|X_j\right]\right\}}\right]\right].\end{aligned}$$

To show the asymptotic Gaussianity of $\sum_{j=1}^{m}\tilde{\psi}_m(X_j)$, we will use the same procedure as in step 1, replacing the array $\{\psi_n(Y_1),\ldots,\psi_n(Y_n)\}_{n=1}^{\infty}$ by $\{\tilde{\psi}_m(X_1),\ldots,\tilde{\psi}_m(X_m)\}_{m=1}^{\infty}$. Note that $\mathbb{E}\left[\tilde{\psi}_m(X_j)\right]=0$. Now,we shall check that the triangular array $\{\tilde{\psi}_m(X_1),\ldots,\tilde{\psi}_m(X_m)\}_{m=1}^{\infty}$ of rowwise independent and identically distributed random elements also satisfies the three conditions of Corollary 7.8 in Araujo and Giné (1980).

- **Condition 1 :** Let us show that

$$\forall \epsilon > 0, \lim_{m \to +\infty} \sum_{j=1}^{m} \mathbb{P}\left(\left\| \tilde{\psi}_m(X_j) \right\|_{\chi^{**}} > \epsilon \right) = 0.$$

Observe that for any $\epsilon > 0$,

$$\begin{aligned} \sum_{j=1}^{m} \mathbb{P}\left(\left\| \tilde{\psi}_m(X_j) \right\|_{\chi^{**}} > \epsilon \right) &\leq \sum_{j=1}^{m} \frac{\mathbb{E}\left[\left\| \mathbf{SGN}^*_{\{F_Y(X_j)\}} - \mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right] \right\|^3_{\chi^{**}} \right]}{\epsilon^3 m^{3/2}} \\ &\leq \frac{8}{\epsilon^3 m^{1/2}} \underset{m \to +\infty}{\longrightarrow} 0. \end{aligned}$$

For the last inequality, we have used the fact that $\left\| \mathbf{SGN}^*_{\{x\}} \right\|_{\chi^{**}} \leq 1$ for all $x \in \chi^*$. Thus, the first condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

- **Condition 2 :** Let us show that

$$\forall f \in \chi^{***}, \lim_{m \to +\infty} \sum_{j=1}^{m} \mathbb{E}\left[f^2 \left(\tilde{\psi}_m(X_j) - \mathbb{E}(\tilde{\psi}_m(X_j)) \right) \right] = \Gamma_2(f, f) < \infty.$$

Le us fix $f \in \chi^{***}$. Since f is linear and $\mathbb{E}(\tilde{\psi}_m(X_j)) = 0$, for all $j = 1, \ldots, m$, we can write

$$\begin{aligned} \sum_{j=1}^{m} \mathbb{E}\left[f^2 \left[\tilde{\psi}_m(X_j) \right] \right] &= \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}\left[f^2 \left(\mathbf{SGN}^*_{\{F_Y(X_j)\}} - \mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right] \right) \right] \\ &= \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}\left[\left(f\left(\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right) - f\left(\mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right] \right) \right)^2 \right] \\ &= \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}\left[\left(f\left(\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right) - \mathbb{E}\left[f\left(\mathbf{SGN}^*_{\{F_Y(X_j)\}} \right) \right] \right)^2 \right] \\ &= \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}\left[\left[v_j - \mathbb{E}(v_j) \right]^2 \right] \\ &= \mathbb{E}\left[\left[v_1 - \mathbb{E}(v_1) \right]^2 \right] \\ &= \mathbb{E}\left[v_1^2 \right] - \mathbb{E}^2\left[v_1 \right] \\ &= \Gamma_2(f, f) < \infty, \end{aligned}$$

where $v_j := \mathbf{SGN}^*_{\{F_Y(X_j)\}} = \mathbf{SGN}^*_{\{\mathbb{E}\left[\mathbf{SGN}_{\{Y - X_j\}} \middle| X_j \right]\}}$, for all $j = 1, \ldots, m$ and Γ_2 is defined as (6). Thus, the second condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

- **Condition 3 :** Let us show that

$$\lim_{k\to+\infty}\overline{\lim_{m\to+\infty}}\sum_{j=1}^{m}\mathbb{E}\left[d^2\left(\tilde{\psi}_m(X_j)-\mathbb{E}\left[\tilde{\psi}_m(X_j)\right],\mathcal{F}_k\right)\right]=0,$$

where $\{\mathcal{F}_k\}_{k\geq 1}$ is a sequence of finite dimensional subspaces of χ^{**} such that $\mathcal{F}_k\subseteq\mathcal{F}_{k+1}$ for all $k\geq 1$ and the closure of $\cup_{k=1}^{\infty}\mathcal{F}_k$ is equal to χ^{**}. This sequence exists because of the separability of χ^{**}. Also, for any $x\in\chi^{**}$ and any $k\geq 1$, we define $d(x,\mathcal{F}_k)=\inf\{\|x-y\|_{\chi^{**}}: y\in\mathcal{F}_k\}$. It is easy to prove that for all $k\geq 1$, the map $x\longmapsto d(x,\mathcal{F}_k)$ is continuous and bounded on any closed ball in χ^{**}.

We have $\mathbb{E}\left[\tilde{\psi}_m(X_j)\right]=0$, for all $j=1,\ldots,m$. Hence,

$$\begin{aligned}\sum_{j=1}^{m}\mathbb{E}\left[d^2\left(\tilde{\psi}_m(X_j),\mathcal{F}_k\right)\right] &= \frac{1}{m}\sum_{j=1}^{m}\mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_Y(X_j)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X_j)\}}\right],\mathcal{F}_k\right)\right]\\ &= \mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_Y(X_1)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_Y(X_1)\}}\right],\mathcal{F}_k\right)\right]\end{aligned}$$

So, we have

$$\lim_{k\to+\infty}\overline{\lim_{m\to+\infty}}\sum_{j=1}^{m}\mathbb{E}\left[d^2\left(\tilde{\psi}_m(X_j),\mathcal{F}_k\right)\right]=0.$$

The last equality is derived from the choice of the $\mathcal{F}_k$'s which implies that $d(x,\mathcal{F}_k)\to 0$ as $k\to\infty$ for all $x\in\chi^{**}$. Thus, the third condition of the Corollary 7.8 in Araujo and Giné (1980) holds.

Consequently, $\sum_{j=1}^{m}\tilde{\psi}_m(X_j)$ converges weakly to a centered Gaussian element in χ^{**} as $m,n\to\infty$. Moreover, its covariance asymptotic covariance is Γ_2 which was obtained while checking the second condition as above. Finally,

$$\sqrt{m}L''_m=m^{-1/2}\sum_{j=1}^{m}\left[\phi(F_Y(X_j))-\mathbb{E}\left[\phi(F_Y(X_j))\right]\right]\xrightarrow{\mathcal{L}}\mathbf{G}(0,\Gamma_2),\qquad(15)$$

weakly as $m,n\to\infty$.

Step 3 : Asymptotic behavior of $R'_{m,n}$

Since

$$R'_{m,n}=K'_{m,n}+K''_{m,n},$$

the convergence of $K^{'}_{m,n}$ and $K^{''}_{m,n}$ to $\mathbf{0}$ ensures that $R^{'}_{m,n}$ converges to $\mathbf{0}$. We have

$$\begin{aligned} K^{'}_{m,n} &= \text{MED} - \mathbb{E}[\text{MED}|X_j; j=1,\ldots,m] - L^{'}_n - L^{''}_m \\ &= \frac{1}{n}\sum_{i=1}^{n}[\phi(F_m(Y_i)) - \mathbb{E}[\phi(F_m(Y_i))|X_1,\ldots,X_m] - \phi(F_X(Y_i)) + \mathbb{E}(\phi(F_X(Y_i)))] \\ &- \frac{1}{m}\sum_{j=1}^{m}[\phi(F_Y(X_j)) - \mathbb{E}(\phi(F_Y(X_j)))] \\ &= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} h(X_j, Y_i), \end{aligned}$$

where $h(X_j, Y_i) = \phi(F_m(Y_i)) - \mathbb{E}[\phi(F_m(Y_i))|X_1,\ldots,X_m] - \phi(F_X(Y_i)) + \mathbb{E}(\phi(F_X(Y_i))) - \phi(F_Y(X_j)) + \mathbb{E}(\phi(F_Y(X_j)))$. Hence,

$$K^{'}_{m,n} = \frac{1}{n}\sum_{i=1}^{n} U(Y_i),$$

where $U(Y_i) = \frac{1}{m}\sum\limits_{j=1}^{m} h(X_j, Y_i)$ for all $i = 1,\ldots,n$. Since the X_j's and the Y_i's are independent, conditionally to $X_j, j = 1,\ldots,m$, the $U(Y_i)$'s for $i = 1,\ldots,n$ are independent and zero mean random elements. Using the fact that χ^{**} is a Banach space of type 2 (see, definition (1)), there is $b > 0$ such that

$$\mathbb{E}\left[\left\|K^{'}_{m,n}\right\|^2_{\chi^{**}}\middle|X_j, j=1,\ldots,m\right] \le \frac{b}{n^2m^2}\sum_{i=1}^{n}\mathbb{E}\left[\|U(Y_i)\|^2_{\chi^{**}}\middle|X_j, j=1,\ldots,m\right]. \tag{16}$$

Computing the expectations of both sides of (16) leads to

$$\mathbb{E}\left[\left\|K^{'}_{m,n}\right\|^2_{\chi^{**}}\right] \le \frac{b}{nm^2}\mathbb{E}\left[\left\|\sum_{j=1}^{m} h(X_j, Y_1)\right\|^2_{\chi^{**}}\right].$$

Now, we are willing to find an upper bound for $\mathbb{E}\left[\left\|\sum_{j=1}^{m} h(X_j, Y_1)\right\|^2_{\chi^{**}}\right]$. Since the X_j's and the Y_i's are independent, conditionally to Y_1, the $h(X_j, Y_1)$'s for $j = 1,\ldots,m$, are independent and zero mean random elements. Consequently, using the definition of Banach space of type 2 (see, definition (1)), there is $b > 0$ such that

$$\begin{aligned}
\mathbb{E}\left[\left\|\sum_{j=1}^{m} h(X_j, Y_1)\right\|_{\chi^{**}}^2\right] &= \mathbb{E}\left[\mathbb{E}\left[\left\|\sum_{j=1}^{m} h(X_j, Y_1)\right\|_{\chi^{**}}^2 \middle| Y_1\right]\right] \\
&\leq b\sum_{j=1}^{m} \mathbb{E}\left[\mathbb{E}\left[\|h(X_j, Y_1)\|_{\chi^{**}}^2 \middle| Y_1\right]\right] \\
&= bm\mathbb{E}\left[\|h(X_1, Y_1)\|_{\chi^{**}}^2\right].
\end{aligned}$$

Consequently, using the the fact that $\left\|\mathbf{SGN}^*_{\{x\}}\right\|_{\chi^{**}} \leq 1$ for all $x \in \chi^*$, we have

$$\mathbb{E}\left[\left\|K'_{m,n}\right\|_{\chi^{**}}^2\right] \leq \frac{36b^2}{mn}. \tag{17}$$

Now, we want to find an upper bound for $K''_{m,n}$.

Using Assumption (1), the map $g : x \mapsto \mathbb{E}\left[\|F_X(Y) + x\|_{\chi^*} \middle| X_1, \ldots, X_m\right]$, for all $x \in \chi^*$, is twice Gateaux differentiable and since $\mathbf{J}_x$ exists (see Assumption (2)), then, for all $h \in \chi^*$,

$$\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)+x+th\}} \middle| X_1, \ldots, X_m\right] = \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)+x\}} \middle| X_1, \ldots, X_m\right] + t\mathbf{J}_x(h) + \mathbf{R}(t),$$

where $\|\mathbf{R}(t)\|_{\chi^{**}} / t \to 0$ when $t \to 0$.

Consequently, when $x = 0$, $t = \frac{1}{m}$ and $h = \sum_{j=1}^{m} \left(\mathbf{SGN}_{\{Y_i - X_j\}} - \mathbb{E}\left[\mathbf{SGN}_{\{Y-X\}} \middle| Y\right]\right)$, we obtain $th = F_m(Y_i) - F_X(Y)$ and

$$\begin{aligned}
\mathbb{E}\left[\mathbf{SGN}^*_{\{F_m(Y_i)\}} \middle| X_1, \ldots, X_m\right] &= \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)\}} \middle| X_1, \ldots, X_m\right] \\
&+ \frac{1}{m}\mathbf{J}_0\left(\sum_{j=1}^{m} \left(\mathbf{SGN}_{\{Y_i - X_j\}} - \mathbb{E}\left[\mathbf{SGN}_{\{Y-X\}} \middle| Y\right]\right)\right) + \mathbf{R}\left(\frac{1}{m}\right).
\end{aligned}$$

Using the linearity of J_0, for all $i = 1, \ldots, n$, we have

$$\mathbb{E}\left[\mathbf{SGN}^*_{\{F_m(Y_i)\}} \middle| X_1, \ldots, X_m\right] = \mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)\}} \middle| X_1, \ldots, X_m\right] + \mathbf{J}_0\left(F_m(Y_i) - F_X(Y)\right) + \mathbf{R}\left(\frac{1}{m}\right).$$

Hence,

$$\begin{aligned}
K''_{m,n} &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[\phi(F_m(Y_i))|X_1, \ldots, X_m] - \mathbb{E}\left[\phi(F_X(Y))\right] \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} \left[\mathbf{J}_0\left(\mathbf{SGN}_{\{Y_i - X_j\}}\right) - \mathbf{J}_0\left(\mathbb{E}\left[\mathbf{SGN}_{\{Y-X\}} \middle| Y\right]\right)\right] + \mathbf{R}\left(\frac{1}{m}\right) \\
&= \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} \tilde{h}(Y_i, X_j) + \mathbf{R}\left(\frac{1}{m}\right),
\end{aligned}$$

where $\tilde{h}(Y_i, X_j) = \mathbf{J}_0\left(\mathbf{SGN}_{\{Y_i - X_j\}}\right) - \mathbf{J}_0\left(\mathbb{E}\big[\mathbf{SGN}_{\{Y-X\}}\big|Y\big]\right)$. Thus, we consider

$$K''_{m,n} = \frac{1}{n}\sum_{i=1}^{n}\tilde{\phi}(Y_i) + \mathbf{R}\left(\frac{1}{m}\right),$$

where $\tilde{\phi}(Y_i) = \frac{1}{m}\sum_{j=1}^{m}\tilde{h}(Y_i, X_j)$. From the definition of the operator J_0 and once again from the definition of type 2 Banach spaces, the independence between the samples of X and Y and the Y_i's being identically distributed, we get

$$\begin{aligned}
\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{\phi}(Y_i)\right\|^2_{\chi^{**}}\right] &= \mathbb{E}\left[\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{\phi}(Y_i)\right\|^2_{\chi^{**}}\middle| X_j, j=1,\ldots,m\right]\right] \\
&\leq \frac{b}{n^2}\sum_{i=1}^{n}\mathbb{E}\left[\mathbb{E}\left[\left\|\tilde{\phi}(Y_i)\right\|^2_{\chi^{**}}\middle| X_j, j=1,\ldots,m\right]\right] \\
&= \frac{b}{n}\mathbb{E}\left[\left\|\tilde{\phi}(Y_1)\right\|^2_{\chi^{**}}\right].
\end{aligned}$$

Conditionally to Y_1, using the definition of type 2 Banach spaces and the X_j's being identically distributed, we have

$$\begin{aligned}
\mathbb{E}\left[\left\|\tilde{\phi}(Y_1)\right\|^2_{\chi^{**}}\right] &= \mathbb{E}\left[\left\|\frac{1}{m}\sum_{j=1}^{m}\tilde{h}(Y_1, X_j)\right\|^2_{\chi^{**}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left\|\frac{1}{m}\sum_{j=1}^{m}\tilde{h}(Y_1, X_j)\right\|^2_{\chi^{**}}\middle| Y_1\right]\right] \\
&\leq \frac{b}{m}\mathbb{E}\left[\left\|\tilde{h}(Y_1, X_1)\right\|^2_{\chi^{**}}\right] \\
&= \frac{b}{m}\mathbb{E}\left[\left\|\mathbf{J}_0\left(\mathbf{SGN}_{\{Y_1 - X_1\}}\right) - \mathbf{J}_0\left(\mathbb{E}\big[\mathbf{SGN}_{\{Y-X\}}\big|Y\big]\right)\right\|^2_{\chi^{**}}\right] \\
&\leq \frac{4bc^2}{m}. \qquad (18)
\end{aligned}$$

The inequality (18) comes from $\mathbf{J}_0$ being a linear continuous map, the assumption (2) and the fact that $\left\|\mathbf{SGN}_{\{x\}}\right\|_{\chi^*} \leq 1$, for all $x \in \chi$. Thus, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left\|K''_{m,n}\right\|^2_{\chi^{**}}\right] &\leq 2\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{\phi}(Y_i)\right\|^2_{\chi^{**}}\right] + 2\mathbb{E}\left[\left\|\mathbf{R}\left(\frac{1}{m}\right)\right\|^2_{\chi^{**}}\right] \\
&\leq \frac{8b^2c^2}{mn} + 2\left\|\mathbf{R}\left(\frac{1}{m}\right)\right\|^2_{\chi^{**}}. \qquad (19)
\end{aligned}$$

Therefore, combining (17) and (19), we get

$$\begin{aligned}\mathbb{E}\left[\|R_{m,n}^{'}\|_{\chi^{**}}^2\right] &\leq 2\mathbb{E}\left[\|K_{m,n}^{'}\|_{\chi^{**}}^2\right] + 2\mathbb{E}\left[\|K_{m,n}^{''}\|_{\chi^{**}}^2\right] \\ &\leq \frac{72b^2}{mn} + \frac{16b^2c^2}{mn} + 4\left\|\mathbf{R}\left(\frac{1}{m}\right)\right\|_{\chi^{**}}^2\end{aligned}$$

Finally,

$$\mathbb{E}\left[\left\|\left(\frac{mn}{N}\right)^{1/2} R_{m,n}^{'}\right\|_{\chi^{**}}^2\right] \leq \frac{72b^2}{N} + \frac{16b^2c^2}{N} + \frac{4n}{mN}\left(\frac{\left\|\mathbf{R}\left(\frac{1}{m}\right)\right\|_{\chi^{**}}}{\frac{1}{m}}\right)^2.$$

Since $\frac{m}{N} \to \lambda \in (0,1)$ as $m, n \to \infty$, we obtain

$$\mathbb{E}\left[\left\|\left(\frac{mn}{N}\right)^{1/2} R_{m,n}^{'}\right\|_{\chi^{**}}^2\right] \underset{m,n\to\infty}{\longrightarrow} 0. \tag{20}$$

Step 4 : Asymptotic behavior of MED

Let's take the equation (12) again :

$$\mathrm{MED} - \mathbb{E}\left[\phi(F_X(Y))\right] = L_n^{'} + L_m^{''} + R_{m,n}^{'}$$

$\Leftrightarrow$

$$\sqrt{\frac{mn}{N}}\left[\mathrm{MED} - \mathbb{E}\left[\phi(F_X(Y))\right]\right] = \sqrt{\frac{mn}{N}} L_n^{'} + \sqrt{\frac{mn}{N}} L_m^{''} + \sqrt{\frac{mn}{N}} R_{m,n}^{'}.$$

From the convergence results (20), (13) and (15) achieved in steps 1, 2 and 3, we get

$$\sqrt{\frac{mn}{N}} L_n^{'} \textit{ converges weakly to } \mathbf{G}(0, \lambda\Gamma_1), \tag{21}$$

$$\sqrt{\frac{mn}{N}} L_m^{''} \textit{ converges weakly to } \mathbf{G}\left(0, (1-\lambda)\Gamma_2\right) \tag{22}$$

and

$$\sqrt{\frac{mn}{N}} R_{m,n}^{'} \textit{ converges in probability to } \mathbf{0} \tag{23}$$

as $m, n \to \infty$. Hence, using Slutsky lemma and the independence of $L_n^{'}$ and $L_m^{''}$, we obtain

$$\sqrt{\frac{mn}{N}}\left[\mathrm{MED} - \mu\right] \xrightarrow{\mathcal{L}} \mathbf{G}(0, \lambda\Gamma_1 + (1-\lambda)\Gamma_2), \tag{24}$$

weakly as $m, n \to \infty$. This completes the proof of the theorem 1. $\square$

Proof of theorem 2. Define

$$\begin{aligned}\rho(\Delta_N) &= \mathbb{E}\left(\mathbf{SGN}^*_{F_X(Y)}\right)\\ &= \mathbb{E}\left(\mathbf{SGN}^*_{\mathbb{E}[\mathbf{SGN}_{Y-X}|Y]}\right)\\ &= \mathbb{E}\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y-Z+\Delta_N}|Y\right]}\right),\end{aligned}$$

where $\Delta_N = \delta(\frac{mn}{N})^{-1/2}$ is the shrinking location shift defined by (7), for some fixed nonzero δ in χ and Z is an independent copy of Y.
As for proving Theorem 1, we consider the following decomposition :

$$\mathrm{MED} - \rho(\Delta_N) = \tilde{L}'_n + \tilde{L}''_m + S'_{m,n}, \tag{25}$$

where

$$\tilde{L}'_n = \frac{1}{n}\sum_{i=1}^{n}\left[\phi(F_X(Y_i)) - \rho(\Delta_N)\right], \tag{26}$$

$$\tilde{L}''_m = \frac{1}{m}\sum_{j=1}^{m}\left[\phi(F_Y(X_j)) - \tilde{\rho}(\Delta_N)\right], \tag{27}$$

$$S'_{m,n} = \mathrm{MED} - \rho(\Delta_N) - \tilde{L}'_n - \tilde{L}''_m, \tag{28}$$

and

$$\tilde{\rho}(\Delta_N) = \mathbb{E}\left(\mathbf{SGN}^*_{F_Y(X)}\right) = \mathbb{E}\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{W-X+\Delta_N}|X\right]}\right), \tag{29}$$

where W is an independent copy of X.

To find the asymptotic distribution of MED under the sequence of shrinking location shifts, we should prove the following steps :

- **Step 1:** Show that $\tilde{L}'_n$ converges in *law* to a Gaussian element.
- **Step 2:** Show that $\tilde{L}''_m$ converges in *law* to a Gaussian element.
- **Step 3:** Show that $S'_{m,n}$ converges in *probability* to 0.
- **Step 4:** Study the asymptotic behavior of $\rho(\Delta_N)$.
- **Step 5:** Conclude the asymptotic normality of MED.

Step 1 : Asymptotic behavior of $\tilde{L}'_n$

We have

$$\tilde{L}'_n = \frac{1}{n}\sum_{i=1}^{n}\left[\phi(F_X(Y_i)) - \rho(\Delta_N)\right].$$

Let us write,

$$\begin{aligned}\psi'_n(Y_i) &= n^{-1/2}\left[\phi(F_X(Y_i)) - \rho(\Delta_N)\right] \\ &= n^{-1/2}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_i-X}|Y_i\right]} - \mathbb{E}\left(\mathbf{SGN}^*_{F_X(Y)}\right)\right].\end{aligned}$$

Note that $\mathbb{E}(\psi'_n(Y_i)) = 0$, for all $i \in \{1,\ldots,n\}$. Then, we aim to verify the three conditions of Corollary 7.8 in Araujo and Giné (1980).

- **Condition** 1 **:** We get for any $\epsilon > 0$,

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{P}\left(\|\psi'_n(Y_i)\| > \epsilon\right) &\leq \sum_{i=1}^{n}\frac{\mathbb{E}\left(\left\|\mathbf{SGN}^*_{F_X(Y_i)} - \rho(\Delta_N)\right\|^3\right)}{\epsilon^3 n^{3/2}} \\ &= \frac{8}{\epsilon^3 n^{1/2}} \underset{n\to+\infty}{\longrightarrow} 0.\end{aligned}$$

- **Condition** 2 **:** Show that

$$\forall f \in \chi^{***}, \lim_{n\to+\infty}\sum_{i=1}^{n}\mathbb{E}\left[f^2\left(\psi'_n(Y_i) - \mathbb{E}(\psi'_n(Y_i))\right]\right. < \infty.$$

Using the linearity of f, we observe that

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{E}\left[f^2(\psi'_n(Y_i)\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[f^2\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_i-X}|Y_i\right]} - \mathbb{E}\left(\mathbf{SGN}^*_{\mathbb{E}[\mathbf{SGN}_{Y-X}|Y]}\right)\right)\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left((\tilde{W}_i - \mathbb{E}(\tilde{W}_i))^2\right),\end{aligned}$$

where $\tilde{W}_i = f(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_i-X}|Y_i\right]}), i = 1,\ldots,n$. Then, the Y_i's being identically distributed, we obtain

$$\sum_{i=1}^{n}\mathbb{E}\left[f^2(\psi'_n(Y_i)\right] = \mathbb{E}\left((\tilde{W}_1 - \mathbb{E}(\tilde{W}_1))^2\right).$$

Note that,

$$\tilde{W}_1 = f\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_1-X}|Y_1\right]}\right) = f\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_1-Z+\Delta_N}|Y_1\right]}\right),$$

where Z is an independent copy of Y_1. Moreover, assumption (3) implies that map $x \longmapsto \mathrm{SGN}_x$ is Fréchet continuous on $\chi\backslash\{0\}$: this follows from Theorem 4.6.15(a) and Proposition 4.6.16 of Borwein and Vanderwerff (2010). Then, we have

$$\mathbb{E}[\mathbf{SGN}_{Y_1-Z+\Delta_N}|Y_1] \longrightarrow \mathbb{E}[\mathbf{SGN}_{Y_1-Z}|Y_1]\,,$$

as $m,n \longrightarrow \infty$. Likewise, since the norm in χ^* is twice Gateaux differentiable (see assumption (1)), the map $x \longmapsto \mathbf{SGN}^*_x$ is continuous on $\chi^*\backslash\{0\}$ so that

$$\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_1-Z+\Delta_N}\middle|Y_1\right]} \longrightarrow \mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{Y_1-Z}\middle|Y_1\right]}$$

as $m,n \longrightarrow \infty$. Consequently,

$$\mathbb{E}\left((\tilde{W}_1-\mathbb{E}(\tilde{W}_1))^2\right) \underset{m,n\to+\infty}{\longrightarrow} \Gamma_1(f,f).$$

- **Condition** 3 **:** Let us show that

$$\lim_{k\to+\infty}\overline{\lim_{n\to+\infty}}\sum_{i=1}^{n}\mathbb{E}\left[d^2\left(\psi_n'(Y_i)-\mathbb{E}\left[\psi_n'(Y_i)\right],\mathcal{F}_k\right)\right]=0,$$

where $\{\mathcal{F}_k\}_{k\geq 1}$ is a sequence of finite dimensional subspaces of χ^{**} such that $\mathcal{F}_k \subseteq \mathcal{F}_{k+1}$ for all $k \geq 1$ and the closure of $\cup_{k=1}^{\infty}\mathcal{F}_k$ is equal to χ^{**}. This sequence exists because of the separability of χ^{**}. Also, for any $x \in \chi^{**}$ and any $k \geq 1$, we define $d(x,\mathcal{F}_k) = \inf\{\|x-y\|_{\chi^{**}} : y \in \mathcal{F}_k\}$. It is easy to prove that for all $k \geq 1$, the map $x \longmapsto d(x,\mathcal{F}_k)$ is continuous and bounded on any closed ball in χ^{**}.

For all $k \geq 1$, we have

$$\begin{aligned}
\sum_{i=1}^{n}\mathbb{E}\left[d^2(\psi_n'(Y_i),\mathcal{F}_k)\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_X(Y_i)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)\}}\right],\mathcal{F}_k\right)\right]\\
&= \mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\{F_X(Y_1)\}}-\mathbb{E}\left[\mathbf{SGN}^*_{\{F_X(Y)\}}\right],\mathcal{F}_k\right)\right]\\
&= \mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathrm{SGN}_{Y_1-Z+\Delta_N}\middle|Y_1\right]}-\mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathrm{SGN}_{Y-Z+\Delta_N}\middle|Y\right]}\right],\mathcal{F}_k\right)\right].
\end{aligned}$$

Since, for all $k \geq 1$, the map $x \longmapsto d(x,\mathcal{F}_k)$ is continuous, we obtain

$$\sum_{i=1}^{n}\mathbb{E}\left[d^2(\psi_n'(Y_i),\mathcal{F}_k)\right] \longrightarrow \mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathrm{SGN}_{Y_1-Z}\middle|Y_1\right]}-\mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}[\mathrm{SGN}_{Y-Z}|Y]}\right],\mathcal{F}_k\right)\right]$$

as $m,n \to \infty$. Moreover, from the choice of the $\mathcal{F}_k$'s, we obtain

$$\lim_{k\to+\infty}\mathbb{E}\left[d^2\left(\mathbf{SGN}^*_{\mathbb{E}\left[\mathrm{SGN}_{Y_1-Z}\middle|Y_1\right]}-\mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}[\mathrm{SGN}_{Y-Z}|Y]}\right],\mathcal{F}_k\right)\right]=0.$$

This completes the checking of the third condition of Corollary 7.8 in Araujo and Giné (1980).

Thus, since the three conditions are verified, we can say that $\sum_{i=1}^{n} \psi_n^{'}$ converges *weakly* to a centered Gaussian random element in χ^{**} as $m, n \to \infty$. Consequently, we have

$$\sqrt{n}\tilde{L}_n^{'} = n^{-1/2} \sum_{i=1}^{n} [\phi(F_X(Y_i)) - \rho(\Delta_N)] \xrightarrow{\mathcal{L}} \mathbf{G}(0, \Gamma_1) \tag{30}$$

weakly as $m, n \to \infty$.

Step 2: Asymptotic behavior of $\tilde{L}_m^{''}$

Using similar arguments as the ones used in the previous step (Asymptotic behavior of $\tilde{L}_n^{'}$), we obtain

$$\sqrt{m}\tilde{L}_m^{''} = m^{-1/2} \sum_{j=1}^{m} [\phi(F_Y(X_j)) - \tilde{\rho}(\Delta_N)] \xrightarrow{\mathcal{L}} \mathbf{G}(0, \Gamma_2) \tag{31}$$

weakly as $m, n \to \infty$.

Step 3: Asymptotic behavior of $S_{m,n}^{'}$

The asymptotic behavior of $S_{m,n}^{'}$ is obtained in the same way as the asymptotic behavior of $R_{m,n}^{'}$. Arguing as in the step 3 of the proof of Theorem 1, we show that

$$\mathbb{E}\left[\left\| \left(\frac{mn}{N}\right)^{1/2} S_{m,n}^{'} \right\|_{\chi^{**}}^{2} \right] \underset{m,n\to\infty}{\longrightarrow} 0. \tag{32}$$

Thus, $(mn)^{1/2}S_{m,n}^{'}$ converges in *probability* to 0 as $m, n \to \infty$ under the sequence of shrinking location shifts.

Step 4: Asymptotic behavior of $\rho(\Delta_N)$

As defined previously

$$\rho(\Delta_N) = \mathbb{E}\left(\mathbf{SGN}^{*}_{F_X(Y)}\right) \quad = \quad \mathbb{E}\left(\mathbf{SGN}^{*}_{\mathbb{E}\left[\mathbf{SGN}_{Y-Z+\Delta_N}|Y\right]}\right),$$

where $\Delta_N = \delta(mn/N)^{-1/2}$ for nonzero fixed $\delta \in \chi$ and Z is an independent copy of Y. As mentioned in assumption (3) in section 2.4, the norm in χ is assumed to be twice Gateaux differentiable at every $x \neq 0, x \in \chi$ and using the Hessian operator $\tilde{\mathbf{J}}_x : \chi \to \chi^{*}$ of the function $x \longmapsto \mathbb{E}[\|x+Y-Z\|_\chi|Y]$, we obtain

$$\begin{aligned}
\mathbb{E}\big[\mathbf{SGN}_{\{Y-Z+\Delta_N\}}\big|Y\big] &= \mathbb{E}\Big[\mathbf{SGN}_{\{Y-Z+\delta(\frac{mn}{N})^{-1/2}\}}\Big|Y\Big] \\
&= \mathbb{E}\big[\mathbf{SGN}_{\{Y-Z\}}\big|Y\big] + \left(\frac{mn}{N}\right)^{-1/2}\tilde{\mathbf{J}}_0(\delta) + \tilde{\mathbf{R}}\left(\frac{mn}{N}\right)^{-1/2} \\
&= \mathbb{E}\big[\mathbf{SGN}_{\{Y-Z\}}\big|Y\big] + \left(\frac{mn}{N}\right)^{-1/2}\Big(\tilde{\mathbf{J}}_0(\delta) + \tilde{\mathbf{R}}(1)\Big).
\end{aligned}$$

Consequently, we get

$$\begin{aligned}
\mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{\{Y-Z+\Delta_N\}}\middle|Y\right]}\right] &= \mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{\{Y-Z\}}\middle|Y\right]+\left(\frac{mn}{N}\right)^{-1/2}\left(\tilde{\mathbf{J}}_0(\delta)+\tilde{\mathbf{R}}(1)\right)}\right] \\
&= \mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{\{Y-Z\}}\middle|Y\right]+t'h'}\right],
\end{aligned}$$

where $t' = \left(\frac{mn}{N}\right)^{-1/2}$ and $h' = \Big(\tilde{\mathbf{J}}_0(\delta) + \tilde{\mathbf{R}}(1)\Big)$. Next, since the norm in χ^* is assumed to be twice Gateaux differentiable at every $y \neq 0, y \in \chi^*$ (see assumption (4) in subsection 2.4) and using the Hessian $\mathbf{H}_u : u : \chi^* \to \chi^{**}$ of the function $u \longmapsto \mathbb{E}\left[\|u + \mathbb{E}[\mathrm{SGN}_{Y-Z}|Y]\,\|_{\chi^*}\right]$, we have

$$\begin{aligned}
\mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{\{Y-Z\}}\middle|Y\right]+t'h'}\right] &= \mathbb{E}\left[\mathbf{SGN}^*_{\mathbb{E}\left[\mathbf{SGN}_{\{Y-Z\}}\middle|Y\right]}\right] + \left(\frac{mn}{N}\right)^{-1/2}\mathbf{H}_0(h') + \tilde{\tilde{\mathbf{R}}}\left(\frac{mn}{N}\right)^{-1/2} \\
&= \left(\frac{mn}{N}\right)^{-1/2}\mathbf{H}_0(\Big(\tilde{\mathbf{J}}_0(\delta) + \tilde{\mathbf{R}}(1)\Big)) + \tilde{\tilde{\mathbf{R}}}\left(\frac{mn}{N}\right)^{-1/2}.
\end{aligned}$$

The last equality follows from Y and Z being identically distributed ($\mu = 0$, under H_0. For the univariate case , see Remark 2.2).
Hence, we have

$$\left(\frac{mn}{N}\right)^{1/2}\rho(\Delta_N) = \mathbf{H}_0(\Big(\tilde{\mathbf{J}}_0(\delta) + \tilde{\mathbf{R}}(1)\Big)) + \tilde{\tilde{\mathbf{R}}}'(1).$$

Consequently, we get

$$\left(\frac{mn}{N}\right)^{1/2}\rho(\Delta_N) \longrightarrow \mathbf{H}_0(\Big(\tilde{\mathbf{J}}_0(\delta)\Big) \tag{33}$$

as $m, n \to \infty$.

Step 5: Asymptotic behavior of MED

Let's take again the decomposition (25) :

$$\mathrm{MED} - \rho(\Delta_N) = \tilde{L}'_n + \tilde{L}''_m + S'_{m,n}$$

$\Leftrightarrow$

$$\sqrt{\frac{mn}{N}}\left[\mathrm{MED} - \rho(\Delta_N)\right] = \sqrt{\frac{mn}{N}}\tilde{L}'_n + \sqrt{\frac{mn}{N}}\tilde{L}''_m + \sqrt{\frac{mn}{N}}S'_{m,n}.$$

Using the convergence results (32), (30) and (31) shown in steps 1, 2, and 3, we get

$$\sqrt{\frac{mn}{N}}\tilde{L}'_n \textit{ converges weakly to } \mathbf{G}(0, \lambda\Gamma_1), \tag{34}$$

$$\sqrt{\frac{mn}{N}}\tilde{L}''_m \textit{ converges weakly to } \mathbf{G}\left(0, (1-\lambda)\Gamma_2\right) \tag{35}$$

and

$$\sqrt{\frac{mn}{N}}S'_{m,n} \textit{ converges in probability to } \mathbf{0} \tag{36}$$

as $m, n \to \infty$ under the sequence of shrinking location shifts.
Moreover, using the convergence result (33) shown in the step 4 and the independence of $\tilde{L}'_n$ and $\tilde{L}''_m$, we obtain

$$\sqrt{\frac{mn}{N}}\mathrm{MED} \xrightarrow{\mathcal{L}} G\left(\mathbf{H}_0(\tilde{\mathbf{J}}_0(\delta)), \lambda\Gamma_1 + (1-\lambda)\Gamma_2\right)$$

weakly as $m, n \to \infty$ under the sequence of shrinking location shifts. This completes the proof of the Theorem 2. □

A.2 Estimation of the WMW covariance operator and its asymptotic mean under the sequence of shrinking location shifts

Here, we give some supplementary details about Chakraborty and Chaudhuri (2015). Consider χ is a Hilbert separable space. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be two samples taking values in χ.

- **Covariance operator of WMW:**

  The covariance of the WMW test statistic introduced in subsection 3.1 is equal to $(1-\lambda)\Pi_1 + \lambda\Pi_2$, where

$$\begin{aligned}\Pi_1 &= \mathbb{E}\left[\mathbb{E}[\mathbf{SGN}_{Y-X}|X] \otimes \mathbb{E}[\mathbf{SGN}_{Y-X}|X]\right] - \upsilon \otimes \upsilon \\ &= \mathbb{E}\left[F_Y(X) \otimes F_Y(X)\right] - \upsilon \otimes \upsilon\end{aligned}$$

  and

$$\begin{aligned}\Pi_2 &= \mathbb{E}\left[\mathbb{E}[\mathbf{SGN}_{Y-X}|Y] \otimes \mathbb{E}[\mathbf{SGN}_{Y-X}|Y]\right] - \upsilon \otimes \upsilon \\ &= \mathbb{E}\left[F_X(Y) \otimes F_X(Y)\right] - \upsilon \otimes \upsilon,\end{aligned}$$

  where $\upsilon = \mathbb{E}(\mathbf{SGN}_{Y-X})$, F_X and F_Y are defined respectively by (3) and (4) and the map $x \otimes x : \chi \longrightarrow \chi$ is defined as $\langle x \otimes x(f), g\rangle = \langle x, f\rangle\langle x, g\rangle$, for all $f, g \in \chi$.

Thus, the operator Π_1 is estimated by

$$\hat{\Pi}_1 \quad = \quad \frac{1}{m-1}\sum_{j=1}^{m}\left[(\hat{F}_Y(X_j)-\hat{\upsilon})\otimes(\hat{F}_Y(X_j)-\hat{\upsilon})\right]$$

where $\hat{\upsilon} = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{Y_i - X_j}{\|Y_i - X_j\|_\chi}$ and $\hat{F}_Y(X_j) = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i - X_j}{\|Y_i - X_j\|_\chi}$, for all $j = 1, \ldots, m$. Similarly, We estimate Π_2 by $\hat{\Pi}_2$.

- **Asymptotic mean of WMW:**

  According to Chakraborty and Chaudhuri (2015), the asymptotic mean of the WMW statistic under the sequence of shrinking location shifts is given by

  $$\mathbf{M}_0(\delta) = \mathbb{E}\left[\frac{\delta}{\|Y-X\|_\chi} - \frac{\langle Y-X, \delta\rangle (Y-X)}{\|Y-X\|_\chi^3}\right],$$

  for fixed nonzero $\delta \in \chi$ and where $\mathbf{M}_x$ is the Hessian of the function $x \longmapsto \mathbb{E}\left[\|x+Y-X\|_\chi\right]$, $x \in \chi$. Consequently, this latter can be estimated by

  $$\frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\frac{\delta}{\|Y_i - X_j\|_\chi} - \frac{\langle Y_i - X_j, \delta\rangle (Y_i - X_j)}{\|Y_i - X_j\|_\chi^3}\right].$$

References

Aneiros, G., Cao, R., Fraiman, R., Genest, C. and Vieu, P. (2019). Recent advances in functional data analysis and high-dimensional statistics. *Journal of Multivariate Analysis.* **170**, 3–9.

Aneiros, G., Horová, I., Hušková, M. and Vieu, P. (2022). On functional data analysis and related topics. *Journal of Multivariate Analysis.* **189**, 104861.

Araujo, A. and Giné, E. (1980). *The central limit theorem for real and Banach valued random variables.* John Wiley & Sons.

Blair, R. C. and Higgins, J. J. (1980). A Comparison of the Power of Wilcoxon's Rank-Sum Statistic to that of Student'st Statistic Under Various Nonnormal Distributions. *Journal of Educational Statistics.* **5**, 309–335.

Borovskikh, Y. V. (1996). *U-statistics in Banach spaces.* VSP International Science Publishers.

Borwein, J. M. and Vanderwerff, J. D. (2010). *Convex functions: Constructions, characterizations and counterexamples.* Cambridge University Press, Cambridge.

Bosq, D. (2000). *Linear processes in function spaces.* Lectures notes in statistics, Springer Verlag.

Capéraà, P. and Cutsem, B.V. (1988). *Méthodes et modèles en statistiques non paramétrique. Exposé fondamental.* Presses de l'université Laval.

Chakraborty, A. and Chaudhuri, P. (2014a). A Wilcoxon-Mann-Whitney type test for infinite dimensional data. *arXiv:1403.0201v1.*

Chakraborty, A. and Chaudhuri, P. (2014b). The spatial distribution in infinite dimensional spaces and related quantiles and depths. *The Annals of Statistics.* **42**, 1203–1231.

Chakraborty, A. and Chaudhuri, P. (2015). A Wilcoxon-Mann-Whitney type test for infinite-dimensional data. *Biometrika.* **102**, 239–246.

Chakraborty, B. and Chaudhuri, P. (1999). On affine invariant sign and rank tests in one and two-sample multivariate problems. *Multivariate analysis, design of experiments and survey sampling.* **159**, 499–522.

Chowdhury, J., Chaudhuri, P. (2020). Convergence rates for kernel regression in infinite-dimensional spaces. *Annals of the Institute of Statistical Mathematics*. **72**. 471–509.

Clifford, B. R. and Higgins, J.J. (1980). A comparison of the power of Wilcoxon's rank-sum statistic to that of Student's t statistic under various nonnormal distributions. *Journal of Educational Statistics.* **5**, 309–335.

Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *Journal of Statistical Planning and Inference.* **147**, 1–23.

Cuevas, A., Febrero, M. and Fraiman, R. (2004). An anova test for functional data. *Computational Statistics & Data Analysis.* **47**, 111—122.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *The Annals of Mathematical Statistics.* **28**, 181–187.

Estévez-Pérez, G. and Vieu, P. (2021). A new way for ranking functional data with applications in diagnostic test. *Computational Statistics.* **36**, 127–154.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis (Theory and practice).* Springer-Verlag, New York.

Frévent, C., Ahmed, MS., Soula, J., Smida, Z., Cucala, L., Dabo-Niang, S. and Genin, M. (2021). HDSpatialScan: Multivariate and Functional Spatial Scan Statistics. *https://CRAN.R-project.org/package=HDSpatialScan.*

Geenens, G. (2015). Moments, errors, asymptotic normality and large deviation principle in nonparametric functional regression. *Statistics & Probability Letters.* **107**, 369–377.

Gijbels, I. and Nagy, S. (2017). On a general definition of depth for functional data. *Statistical Science.* **32**, 630–639

Goia, A. and Vieu, P. (2016). An introduction to recent advances in high/infinite dimensional statistics. *Journal of Multivariate Analysis.* **146**, 1–6

Hàjek, J., Ȉidák, Z. and Sen, K. (1999). *Theory of Rank Tests (Second edition)*. Academic Press, United States of America.

Hoffmann-Jørgensen, J. and Pisier, G. (1976). The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probability.* **4**, 587–599.

Horváth, L., Kokoszka, P., and Reeder, R. (2013). Estimation of the mean of function time series and a two-sample problem. *Journal of the Royal Statistical Society. Series B.* **75**, 103–122.

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications.* Springer New York.

Karhunen. K. (1947). Uber lineare methoden in der wahrscheinlichkeitsrechnung. *Annales Academiae Scientarum Fennicae.* **37**, 3–79.

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis.* CRC Press.

Koul, H.L. and Staudte, R. G. (1972). Weak Convergence of Weighted Empirical Cumulatives Based on Ranks. *Ann. Math. Statist.* **43**, 832–841.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses (Second edition).* Springer-Verlag, New York.

Lehmann, E. L and Romano, J.P. (2005). *Testing Statistical Hypotheses (Third edition).* Springer-Verlag, New York.

Lévy, P. and Loève. M. (1948). *Processus stochastiques et mouvement brownien.* Gauthier-Villars, Paris.

Ling, N. and Vieu, P. (2018). Nonparametric modelling for functional data: selected survey and tracks for future. *Statistics.* **52**, 934–949.

Mann, H. B., Whitney D.R. (1947). On a test of whether one of two random variables is stochastically larger than the order. *Ann. Math. Statist.* **18**, 50–60.

Mood, A. M. (1950). *Introduction to the Theory of Statistics.* McGraw-Hi1l series in probability and statistics, New York.

Mood, A. M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. *Ann. Math. Statist* **25**, 514–522.

Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates: A review. *Scandinavian Journal of Statistics.* **26**, 319-–43.

Oja, H. (2010). *Multivariate Nonparametric Methods with R.* Springer, New York.

Oja, R. and Randles, H. R. (2004). Multivariate nonparametric tests. *Statistical Science.* **19**, 598–605.

Puri, M. L. and Sen, P. K. (1971). *Nonparametric Methods in Multivariate Analysis.* John Wiley & Sons, Inc, New York-London-Syndney.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis (Second edition).* Springer-Verlag New York.

Smida, Z., Cucala, L. Gannoun, A. and Durif, G. (2022). A Wilcoxon-Mann-Whitney spatial scan statistic for functional data. *Computational Statistics & Data Analysis.* **167**, 107378.

Vakhania, N., Tarieladze, V.I., Chobanyan, S.A. (1987). *Probability Distributions on Banach Spaces.* Dordrecht: Reidel. Translated from Russian.

Van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press.

Wilcoxon, F. (1945). Individual comparaisons by ranking methods. *Biometrics.*, **1**, 80–83.

Zapała, A. M. (2000). Jensen's Inequality for Conditional Expectations in Banach Spaces. *Real analysis exchange.* **26**, 541-552.