C. R. Math. Rep. Acad. Sci. Canada Vol. **46** (2) 2024, pp. 46–78

Homogeneity Tests for Several Distributions in Hilbert Space Based on Multiple Maximum Variance Discrepancy

Armando Sosthène Kali Balogoun and Guy Martial Nkiet

Presented by George A. Elliott, FRSC

Abstract. This paper deals with the problem of testing for the equality of k probability distributions on Hilbert spaces, with $k \geqslant 2$. We introduce a generalization of the maximum variance discrepancy called multiple maximum variance discrepancy. A consistent estimator of this measure is proposed as test statistic, and its asymptotic distribution under the null hypothesis is derived. Since this asymptotic distribution is that of an infinite sum of random variables, we then propose another test statistic obtained from an appropriate modification of the first one, and we get its asymptotic normality both under homogeneity hypothesis and under the alternative hypothesis, so introducing a faster test for homogeneity of distributions of random variables valued into a Hilbert space. A simulation study investigating the finite sample performances of the two introduced tests and comparing them to existing ones is provided.

Résumé. Cet article considère le problème de test d'égalité de k lois sur un espace de Hilbert, avec $k \geqslant 2$. Nous introduisons une généralisation de l'écart maximal de variance appelé écart maximal de variance multiple. Un estimateur convergent de cette mesure est proposé comme statistique de test, et sa loi asymptotique sous l'hypothèse nulle est déterminée. Puisque cette loi limite est celle d'une somme infinie de variables aléatoires, nous proposons ensuite une autre statistique de test obtenue à partir d'une modification appropriée de la première statistique, et nous obtenons sa normalité asymptotique aussi bien sous l'hypothèse nulle que sous l'hypothèse alternative, introduisant ainsi un test plus rapide d'homogénéité de lois de variables aléatoires à valeurs dans un espace de Hilbert. Une étude par simulation, permettant d'apprécier les performances des deux tests proposés et de les comparer à des tests existants, est fournie.

1. Introduction Thanks to recent advances in data storage and processing, functional data are frequently encountered in a wide range of scientific fields, including environmental science, finance, genetics, biology, geophysics, image and signal processing, to name just a few. Statistical modeling related to this kind of data leads to random variables valued into Hilbert spaces which is of growing

Received by the editors on April 29, 2024; revised May 21, 2024.
AMS Subject Classification: Primary: 62E20; secondary: 46E22.
Keywords: Asymptotic normality, kernel-based conditional dependence, reproducing kernel Hilbert space, functional data analysis.

interest since the last two decades. One important statistical inference problem for this kind of variables is that of testing homogeneity, i.e. the equality of the distributions of the variables that are considered. However, most of the works considering this problem in the context of functional data analysis propose methods which boil down to tests for the equality of certain characteristics of the distributions of the concerned functional variables such as means or covariance operators (e.g., [1,5,6,8,20]). On the other hand, the framework of kernel-based methods, that is methods based on kernel embeddings of probability measures in reproducing kernel Hilbert spaces (RKHS), allowed the emergence of powerful statistical inference methods for variables of the aforementioned type, including tests for homogeneity. More specifically, the maximum mean discrepancy (MMD) has been introduced in [12] for testing the equality of two distributions for which independent samples are observed. It has been extended later in [4] to the multiple case, that is the case of more than two distributions, through the introduction of the generalized maximum mean discrepancy (GMMD). Recently, a novel discrepancy measure, called maximum variance discrepancy (MVD), was introduced in [14] and used for testing the equality of two distributions, so leading to a new testing procedure which is shown from simulations to be more powerful than the one based on MMD. Therefore, there is an interest in extending this measure to the multiple case. This is what is done in this paper via the multiple maximum variance discrepancy (MMVD) that we introduce in Section 2. We then propose a consistent estimator of this measure as a statistic for testing homogeneity, and we derive its asymptotic distribution under the null hypothesis in Section 3. This distribution is that of an infinite sum of random variables and, therefore, can not be used for performing the test, so that one may be forced to resort to methods with high computational cost such as permutation or subsampling methods for achieving the test. Faced with a similar problem with the classical estimator of MMD, [15] adopted an approach, first introduced in [2], allowing to obtain asymptotic normality, both under the null hypothesis and under the alternative, for an estimator obtained from a modification of the first one. This approach was also used later in [3] for the case of GMMD. In this paper we adopt this approach for MMVD, so introducing in Section 4 another test statistic obtained from an appropriate modification of the first one. Asymptotic normality for this second statistic is the obtained both under the null hypothesis and the alternative one. This allows to propose another test for homogeneity of random variables valued into Hilbert spaces. Section 5 is devoted to the presentation of simulations made on synthetic functional data in order to appreciate the finite sample performances of the introduced tests, and to compare them to existing ones. All the proofs are postponed in Section 6.

2. The MMVD and Its Estimation

For $k \geqslant 2$ and $j \in \{1, \cdots, k\}$, we consider a random variable X_j valued into a separable Hilbert space $\mathcal{X}$ and denote by $\mathbb{P}_j$ its probability distribution. We want to define a measure of the discrepancy between $\mathbb{P}_1, \cdots, \mathbb{P}_k$. For doing that, as it is usual in kernel-based methods, we will embed these distributions into a RKHS, so that the discrepancy between them should be measured by that of the resulting embeddings. So, let $\mathcal{H}$ be this RKHS; it is a Hilbert space of functions from $\mathcal{X}$ to $\mathbb{R}$, endowed with a inner product $< \cdot, \cdot >_{\mathcal{H}}$ and associated to a kernel $K : \mathcal{X}^2 \to \mathbb{R}$, satisfying:

(i) for any $(x, y) \in \mathcal{X}^2, K(x, y) = K(y, x)$;
(ii) $\forall x \in \mathcal{X}$, $K(x, \cdot) \in \mathcal{H}$;
(iii) $\forall f \in \mathcal{H}$, $\forall x \in \mathcal{X}$, $f(x) =< K(x, \cdot), f >_{\mathcal{H}}$.

Assuming that K satisfies the following condition:

$$(\mathscr{A}_1): \ \|K\|_\infty := \sup_{(x,y)\in\mathcal{X}^2} K(x, y) < +\infty,$$

we consider, for any $j \in \{1, \cdots, k\}$, the mean and covariance operator embeddings defined as

$$m_j = \mathbb{E}\left(K\left(X_j, \cdot\right)\right) \quad \text{and} \quad V_j = \mathbb{E}\left[\left(K(X_j, \cdot) - m_j\right)^{\otimes^2}\right],$$

where $\otimes$ is the tensor product such that, for all $(a, b) \in \mathcal{H}^2$, $a \otimes b$ is the operator from $\mathcal{H}$ to itself which satisfies $(a \otimes b)(f) =< a, f >_{\mathcal{H}} b$, $\forall f \in \mathcal{H}$, and for any vector a we denote $a^{\otimes^2} = a \otimes a$. It is known (see, e.g., [9, 10]) that under $(\mathscr{A}_1)$, V_j is an Hilbert-Schmidt operator.

For the case of $k = 2$, the maximum variance discrepancy (MVD) was introduced in [14] in order to measure the discrepancy between $\mathbb{P}_1$ and $\mathbb{P}_2$. For dealing with the multiple case, we introduce below the multiple maximum variance discrepancy (MMVD) which is an extension of the MVD to the case of more than two populations.

DEFINITION 2.1. The multiple maximum variance discrepancy (MMVD), related to $\mathbb{P}_1, \cdots, \mathbb{P}_k$ and $\pi = (\pi_1, \cdots, \pi_k) \in (]0, 1[)^k$ with $\sum_{\ell=1}^k \pi_\ell = 1$, is the positive real number defined by

$$\mathrm{MMVD}^2(\mathbb{P}_1, \cdots, \mathbb{P}_k; \pi) = \sum_{j=1}^k \sum_{\substack{\ell=1\\ \ell\neq j}}^k \pi_\ell \parallel V_j - V_\ell \parallel_{\mathrm{HS}}^2,$$

where $\| \cdot \|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm of operators.

This MMVD, which definition recovers that of MVD given in [14], measures the discrepancy between the distributions $\mathbb{P}_1, \cdots, \mathbb{P}_k$. So, a consistent estimator of it can be used for testing for homogeneity, that is testing for the hypothesis

$$\mathscr{H}_0 : \mathbb{P}_1 = \mathbb{P}_2 = \cdots = \mathbb{P}_k \quad \text{against} \quad \mathscr{H}_1 : \exists (j, \ell) \in \{1, \cdots, k\}^2, \quad \mathbb{P}_j \neq \mathbb{P}_\ell.$$

More specifically, for $j \in \{1, \cdots, k\}$, letting $\{X_1^{(j)}, \cdots, X_{n_j}^{(j)}\}$ be a i.i.d. sample of X_j, we consider the empirical estimators of m_j and V_j defined, respectively, as

$$\widehat{m}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} K(X_i^{(j)}, \cdot) \quad \text{and} \quad \widehat{V}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} \Big(K(X_i^{(j)}, \cdot) - \widehat{m}_j\Big)^{\otimes^2}$$

which lead to the consistent estimator $\widehat{T}_n$ of $\mathrm{MMVD}^2(\mathbb{P}_1, \cdots, \mathbb{P}_k; \pi)$ given by

$$\widehat{T}_n \quad = \quad \sum_{j=1}^{k}\sum_{\substack{\ell=1 \\ \ell\neq j}}^{k} \pi_\ell \left\| \widehat{V}_j - \widehat{V}_\ell \right\|_{\mathrm{HS}}^2,$$

where $\pi_\ell = n_\ell/n$ with $n = \sum_{j=1}^k n_j$. This statistic can, therefore, be used as test statistic for testing for homogeneity. Note that a more detailed expression, that can be used for computing concretely $\widehat{T}_n$, is obtained from an expansion using properties of the tensor product $\otimes$, the reproducing property of $\mathcal{H}$ and the inner product $< \cdot, \cdot >_{\mathrm{HS}}$ defined by $< A, B >_{\mathrm{HS}}^2 = \mathrm{Tr}(AB^*)$ and which induces $\| \cdot \|_{\mathrm{HS}}$. Indeed, we have

(2.1)
$$\widehat{T}_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1 \\ \ell\neq j}}^{k} \pi_\ell \left\{ \left\|\widehat{V}_j\right\|_{\mathrm{HS}}^2 + \left\|\widehat{V}_\ell\right\|_{\mathrm{HS}}^2 - \frac{2}{n_j}\sum_{i=1}^{n_j} < \Big(K(X_i^{(j)}, \cdot) - \widehat{m}_j\Big)^{\otimes^2}, \widehat{V}_\ell >_{\mathrm{HS}} \right\},$$

and using the tensor product properties: $(a \otimes b)^* = b \otimes a$, $(a \otimes b)(c \otimes d) =< a, d >_{\mathcal{H}} c \otimes b$ and $\mathrm{Tr}(a \otimes b) =< a, b >_{\mathcal{H}}$ (see [7]), we get for any $(j, \ell) \in \{1, \cdots, k\}^2$ and any $(i, r) \in \{1, \cdots, n_j\} \times \{1, \cdots, n_\ell\}$:

$$\begin{aligned}
&< \Big(K(X_i^{(j)},\cdot)-\widehat{m}_j\Big)^{\otimes^2}, \Big(K(X_r^{(\ell)},\cdot)-\widehat{m}_\ell\Big)^{\otimes^2} >_{\mathrm{HS}}\\
&= \; < K(X_i^{(j)},\cdot)-\widehat{m}_j, K(X_r^{(\ell)},\cdot)-\widehat{m}_\ell >_{\mathcal{H}}^2\\
&= \; \Bigg(< K(X_i^{(j)},\cdot), K(X_r^{(\ell)},\cdot) >_{\mathcal{H}} - \frac{1}{n_\ell}\sum_{q=1}^{n_\ell} < K(X_i^{(j)},\cdot) K(X_q^{(\ell)},\cdot) >_{\mathcal{H}}\\
&\quad - \frac{1}{n_j}\sum_{p=1}^{n_j} < K(X_p^{(j)},\cdot), K(X_r^{(\ell)},\cdot) >_{\mathcal{H}}\\
&\quad + \frac{1}{n_j n_\ell}\sum_{p=1}^{n_j}\sum_{q=1}^{n_\ell} < K(X_p^{(j)},\cdot), K(X_q^{(\ell)},\cdot) >_{\mathcal{H}} \Bigg)^2 \qquad (2.2)\\
&= \; \Bigg(K(X_i^{(j)},X_r^{(\ell)}) - \frac{1}{n_\ell}\sum_{q=1}^{n_\ell} K(X_i^{(j)},X_q^{(\ell)}) - \frac{1}{n_j}\sum_{p=1}^{n_j} K(X_p^{(j)},X_r^{(\ell)})\\
&\quad + \frac{1}{n_j n_\ell}\sum_{p=1}^{n_j}\sum_{q=1}^{n_\ell} K(X_p^{(j)},X_q^{(\ell)}) \Bigg)^2\\
&= \; \mathcal{S}_{ir}^{j\ell},
\end{aligned}$$

where

$$\mathcal{S}_{ir}^{j\ell} = \left(\mathcal{K}_{ir}^{j\ell} - \frac{1}{n_\ell}\mathcal{K}_{i\bullet}^{j\ell} - \frac{1}{n_j}\mathcal{K}_{\bullet r}^{j\ell} + \frac{1}{n_j n_\ell}\mathcal{K}_{\bullet\bullet}^{j\ell}\right)^2 \qquad (2.3)$$

with

$$\mathcal{K}_{ir}^{j\ell} = K(X_i^{(j)},X_r^{(\ell)}),\ \mathcal{K}_{i\bullet}^{j\ell} = \sum_{q=1}^{n_\ell}\mathcal{K}_{iq}^{j\ell},\ \mathcal{K}_{\bullet r}^{j\ell} = \sum_{p=1}^{n_j}\mathcal{K}_{pr}^{j\ell} \ \text{ and } \ \mathcal{K}_{\bullet\bullet}^{j\ell} = \sum_{p=1}^{n_j}\sum_{q=1}^{n_\ell}\mathcal{K}_{pq}^{j\ell}.$$

Hence

$$\widehat{T}_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell\left\{\frac{1}{n_j^2}\sum_{i,r=1}^{n_j}\mathcal{S}_{ir}^{jj} + \frac{1}{n_\ell^2}\sum_{i,r=1}^{n_\ell}\mathcal{S}_{ir}^{\ell\ell} - \frac{2}{n_j n_\ell}\sum_{i=1}^{n_j}\sum_{r=1}^{n_\ell}\mathcal{S}_{ir}^{j\ell}\right\}. \qquad (2.4)$$

Example 2.1. This formula is easily applied to the case of functional variables. Indeed, if the $X_i^{(j)}$s are functional variables belonging, for instance, in $L^2([0,1])$ and observed on points $t_1,\dots,t_N$ of a fine grid in $[0,1]$ such that $t_1 = 0$ and $t_N = 1$, the statistic in (2.4) can easily be computed or approximated, depending on the used kernel. Its computation is obtained from that of $\mathcal{K}_{ir}^{j\ell}$. For example, if the Gaussian kernel is used, one has

$$\mathcal{K}_{ir}^{j\ell} = \exp\left(-\omega^2 \|X_i^{(j)} - X_r^{(\ell)}\|_{\mathcal{X}}^2\right) = \exp\left(-\omega^2 \int_0^1 \left(X_i^{(j)}(t) - X_r^{(\ell)}(t)\right)^2 dt\right),$$

where $\omega > 0$. This can be approximated by using trapezoidal rule, so leading to

$$\begin{aligned}\mathcal{K}_{ir}^{j\ell} \quad \simeq \quad & \exp\Bigg(-\omega^2 \sum_{m=1}^{N-1} \frac{t_{m+1} - t_m}{2} \left(\left(X_i^{(j)}(t_m) - X_r^{(\ell)}(t_m)\right)^2 \right. \\ & \left. + \left(X_i^{(j)}(t_{m+1}) - X_r^{(\ell)}(t_{m+1})\right)^2\right)\Bigg).\end{aligned}$$

3. Testing for Homogeneity

In this section, we consider the preceding estimator as test statistic and we derive its asymptotic distribution under $\mathscr{H}_0$. We assume that the observed i.i.d. samples satisfy the following conditions:

$(\mathscr{A}_2)$: $X_i^{(j)} \perp\!\!\!\perp X_r^{(\ell)}$ for $j \neq \ell$ and $(i, r) \in \{1, \cdots, n_j\} \times \{1, \cdots, n_\ell\}$, where $\perp\!\!\!\perp$ denotes stochastic independence;

$(\mathscr{A}_3)$: for $j \in \{1, \cdots, k\}$, there exists $\rho_j \in]0, 1[$ such that $\lim_{n_j \to +\infty} \left\{\sqrt{n}\,(\pi_j - \rho_j)\right\} = 0$, where $\pi_j = n_j/n$ with $n = \sum_{j=1}^k n_j$.

If $\mathscr{H}_0$ holds, then $m_1 = m_2 = \cdots = m_k =: m$ and $V_1 = V_2 = \cdots = V_k =: V$. Then, considering the kernel $\widetilde{K} : \mathcal{X}^2 \to \mathbb{R}$ defined by

$$\widetilde{K}(x, y) =< (K(x, \cdot) - m)^{\otimes 2} - V, (K(y, \cdot) - m)^{\otimes 2} - V >_{\text{HS}},$$

where $a^{\otimes 2} = a \otimes a$, and denoting by $\{\lambda_p\}_{p \geqslant 1}$ the nonincreasing sequence of eigenvalues of the integral operator associated with this kernel, that is the operator $S_{\widetilde{K}} : L^2(\mathbb{P}) \to L^2(\mathbb{P})$ defined by

$$S_{\widetilde{K}} g(x) = \int_{\mathcal{X}} \widetilde{K}(x, y) g(y) d\mathbb{P}(y), \quad \text{for} \quad g \in L^2(\mathbb{P}), \tag{3.1}$$

where $\mathbb{P} = \mathbb{P}_1 = \cdots = \mathbb{P}_k$, we have:

THEOREM 3.1. *We assume $(\mathscr{A}_1)$, $(\mathscr{A}_2)$ and $(\mathscr{A}_3)$. Then under $\mathscr{H}_0$, as $\min_{1 \leqslant j \leqslant k}(n_j) \to +\infty$,*

$$n\widehat{T}_n \xrightarrow{\mathscr{D}} \sum_{p=1}^{+\infty} \lambda_p \left\{(k-2)Z_p + \sum_{j=1}^k \left(\rho_j^{-1} Y_{j,p}^2 - 2 \sum_{\substack{\ell=1 \\ \ell \neq j}}^k \rho_\ell^{1/2} \rho_j^{-1/2} Y_{j,p} Y_{\ell,p}\right)\right\},$$

where $(Y_{j,p})_{p \geqslant 1,\, 1 \leqslant j \leqslant k}$ is a sequence of independent random variables having the standard normal distribution and $(Z_p)_{p \geqslant 1}$ is a sequence of independent random variables having the chi-squared distribution with k degrees of freedom.

Remark 3.1. This theorem generalizes Theorem 3.1 of [14]. Indeed, if $k = 2$ the previous limiting distribution is that of $\rho^{-1}(1-\rho)^{-1}\sum_{p=1}^{+\infty}\lambda_p\mathcal{W}_p$, where $\rho = \rho_1$ (hence, $\rho_2 = 1-\rho$), and

$$\begin{aligned}\mathcal{W}_p &= \rho(1-\rho)\left\{\frac{1}{\rho}Y_{1,p}^2 + \frac{1}{1-\rho}Y_{2,p}^2 - 2\left(\sqrt{\frac{1-\rho}{\rho}} + \sqrt{\frac{\rho}{1-\rho}}\right)Y_{1,p}Y_{2,p}\right\}\\ &= \rho(1-\rho)\left(\frac{1}{\sqrt{\rho}}Y_{1,p} - \frac{1}{\sqrt{1-\rho}}Y_{2,p}\right)^2\\ &\quad +2\rho(1-\rho)\left\{\frac{1}{\sqrt{\rho(1-\rho)}} - \sqrt{\frac{1-\rho}{\rho}} - \sqrt{\frac{\rho}{1-\rho}}\right\}Y_{1,p}Y_{2,p}\\ &= \mathcal{Z}_p^2\end{aligned}$$

with

$$\mathcal{Z}_p = \sqrt{\rho(1-\rho)}\left(\frac{1}{\sqrt{\rho}}Y_{1,p} - \frac{1}{\sqrt{1-\rho}}Y_{2,p}\right).$$

Since $Y_{1,p}$ and $Y_{2,p}$ are independent with standard normal distribution, it follows that $\mathcal{Z}_p$ has also the standard normal distribution. So, Theorem 3.1 of [14] is recovered.

Remark 3.2. As in [14], it is not easy to use the asymptotic distribution obtained in Theorem 3.1 for performing the test because it is an infinite sum and it is difficult to determine the weights contained in it. So, one may use the permutation method, or a subsampling method, for computing p-values from which the test can be achieved. Since such methods have high computational costs, it would be preferable to consider another test statistic for which the asymptotic null distribution can be used for determining a threshold to which this statistic can be compared in order to decide whether $\mathscr{H}_0$ is accepted or not.

4. **Modified Estimator and Asymptotic Normality** In this section, we introduce another estimator of MMVD, obtained from a modification of $\widehat{T}_n$ that consists in introducing a weight in the cross-product term of (2.1), and we use it as test statistic for testing for $\mathscr{H}_0$. Then, we get asymptotic normality for this statistic, what allows to easily perform the test.

For any $m \in \mathbb{N}^*$, letting $\{w_{l,n}(\gamma)\}_{1\leq l\leq m}$ be a triangular array of positive real numbers depending on a parameter $\gamma \in]0,1[$, we consider the estimator $\widehat{T}_{n,\gamma}$

given by:

$$\begin{aligned}\widehat{T}_{n,\gamma} &= \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell\left\{\left\|\widehat{V}_j\right\|_{\text{HS}}^2+\left\|\widehat{V}_\ell\right\|_{\text{HS}}^2-\frac{2}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)\right.\\ &\qquad \left.< \left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}_\ell >_{\text{HS}}\right\}\\ (4.1)\quad &= \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell\left\{\frac{1}{n_j^2}\sum_{i,r=1}^{n_j}\mathcal{S}_{ir}^{jj}+\frac{1}{n_\ell^2}\sum_{i,r=1}^{n_\ell}\mathcal{S}_{ir}^{\ell\ell}-\frac{2}{n_jn_\ell}\sum_{i=1}^{n_j}\sum_{r=1}^{n_\ell}w_{i,n_j}(\gamma)\,\mathcal{S}_{ir}^{j\ell}\right\},\end{aligned}$$

and we take it as test statistic. For obtaining its asymptotic normality, we suppose that the used sequence of weights satisfies the following conditions:

$(\mathscr{A}_4)$: There exists a strictly positive real number τ and an integer m_0 such that for all $m > m_0$:

$$m\left|\frac{1}{m}\sum_{i=1}^{m}w_{l,m}(\gamma)-1\right|\leq\tau.$$

$(\mathscr{A}_5)$: There exists $C > 0$ such that $\max\limits_{1\leq l\leq m} w_{l,m}(\gamma) \leqslant C$ for all $m \in \mathbb{N}^*$ and all $\gamma \in]0,1[$.

$(\mathscr{A}_6)$: For any $\gamma \in]0,1[$, $\lim\limits_{m\to+\infty} \frac{1}{m}\sum_{l=1}^{n} w_{l,m}^2(\gamma) = w^2(\gamma) > 1$.

A typical example of sequence satisfying the above assumption, given in [2], is $w_{l,m}(\gamma) = 1 + (-1)^l\,\gamma$. Other examples are $w_{l,m}(\gamma) = 1 + \sin(l\pi\gamma)$ and $w_{l,m}(\gamma) = 1 + \cos(l\pi\gamma)$.

Putting

$$V=\sum_{j=1}^{k}\rho_jV_j,\ \ \nu=\sum_{j=1}^{k}V_j,$$

and considering the maps $\mathcal{U}_j$ and $\mathcal{V}_j$ from $\mathcal{X}$ to $\mathbb{R}$ defined by:

$$\mathcal{U}_j(x) =< (K(x,\cdot)-m_j)^{\otimes^2}-V_j,(1-\rho_j+k\rho_j)V_j-\rho_j\nu >_{\text{HS}}$$

and

$$\mathcal{V}_j(x) =< (K(x,\cdot)-m_j)^{\otimes^2}-V_j,V-\rho_jV_j >_{\text{HS}},$$

we have:

THEOREM 4.1. *Assume conditions* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_6)$. *Then*

$$\sqrt{n}\Big(\widehat{T}_{n,\gamma} - MMVD^2(\mathbb{P}_1, \cdots, \mathbb{P}_k; \pi)\Big) \xrightarrow{\mathscr{D}} \mathcal{N}(0, \sigma_\gamma^2),$$

as $\min_{1\leqslant j\leqslant k}(n_j) \to +\infty$, *where* $\sigma_\gamma^2 = \sum_{j=1}^k 4\rho_j^{-1}\sigma_j^2(\gamma)$ *with:*

$$\sigma_j^2(\gamma) = Var\Big(\mathcal{U}_j(X_1^{(j)})\Big) + k^2(\gamma)\, Var\Big(\mathcal{V}_j(X_1^{(j)})\Big) - 2\,Cov\Big(\mathcal{U}_j(X_1^{(j)}), \mathcal{V}_j(X_1^{(j)})\Big).$$

This theorem gives asymptotic normality both under $\mathscr{H}_0$ and under $\mathscr{H}_1$. If $\mathscr{H}_0$ is true, then $\mathrm{MMVD}^2(\mathbb{P}_1, \cdots, \mathbb{P}_k; \pi) = 0$ and $\nu = kV$. Thus, for all $j \in \{1, \cdots, k\}$:

$$\mathcal{U}_j(x) = \mathcal{V}_j(x) = (1-\rho_j) < (K(x,\cdot) - \eta)^{\otimes^2} - V, V) >_{\mathrm{HS}},$$

where $V = V_1 = \cdots = V_k$ and $\eta = m_1 = \cdots = m_k$. Consequently,

$$\sqrt{n}\widehat{T}_{n,\gamma} \xrightarrow{\mathscr{D}} \mathcal{N}(0, \sigma_\gamma^2),$$

as $\min_{1\leqslant j\leqslant k}(n_j) \to +\infty$, where

$$\sigma_\gamma^2 = 4\left(w^2(\gamma) - 1\right)\theta^2 \sum_{j=1}^k \rho_j^{-1}(1-\rho_j)^2$$

with

$$\theta^2 = \mathrm{Var}\left(< \left(K(X_1^{(1)},\cdot) - \eta\right)^{\otimes^2} - V, V) >_{\mathrm{HS}}\right).$$

As σ_γ^2 is unknown we must replace it by a consistent estimator in order to perform the test. Putting

$$\widehat{V} = \sum_{j=1}^k \pi_j \widehat{V}_j,$$

$$\begin{aligned}
\widehat{\theta}_j^2 \quad &= \quad \frac{1}{n_j}\sum_{i=1}^{n_j} < \left(K(X_i^{(j)},\cdot) - \widehat{m}_j\right)^{\otimes^2}, \\
\widehat{V} \quad &> \quad {}^{2}_{\mathrm{HS}} - \left(\frac{1}{n_j}\sum_{i=1}^{n_j} < \left(K(X_i^{(j)},\cdot) - \widehat{m}_j\right)^{\otimes^2}, \widehat{V} >_{\mathrm{HS}}\right)^2,
\end{aligned}$$

$$\widehat{\theta}^2 = \sum_{j=1}^k \pi_j \widehat{\theta}_j^2 \quad \text{and} \quad \widehat{\sigma}_\gamma^2 = 4\left(w^2(\gamma) - 1\right)\widehat{\theta}^2 \sum_{j=1}^k \pi_j^{-1}(1-\pi_j)^2,$$

we have:

THEOREM 4.2. *Assume conditions* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_6)$. *Then, under* $\mathscr{H}_0$,

$$\sqrt{n}\frac{\widehat{T}_{n,\gamma}}{\widehat{\sigma}_\gamma} \xrightarrow{\mathscr{D}} \mathcal{N}(0,1),$$

as $\min_{1\leqslant j\leqslant k}(n_j) \to +\infty$.

This theorem allows to achieve the test in practice. More specifically, $\mathscr{H}_0$ is to be rejected when $\sqrt{n}\widehat{T}_{n,\gamma} > \widehat{\sigma}_\gamma\Phi^{-1}(1-\alpha)$, where α is the chosen significance level and Φ is the cumulative distribution function of the standard normal distribution.

Note that a more explicit expression, that can be used for concretly compute the variance estimate from the samples, can be obtained by using (2.2) and (2.3). Indeed, since

$$< \left(K(X_i^{(j)},\cdot) - \widehat{m}_j\right)^{\otimes^2}, \widehat{V} >_{\text{HS}} \quad = \quad \sum_{\ell=1}^{k}\frac{\pi_\ell}{n_\ell}\mathcal{S}_{i\bullet}^{j\ell}, \quad \text{where} \quad \mathcal{S}_{i\bullet}^{j\ell} = \sum_{r=1}^{n_\ell}\mathcal{S}_{ir}^{j\ell},$$

it follows:

$$\widehat{\theta}_j^2 = \frac{1}{n_j}\sum_{i=1}^{n_j}\left(\sum_{\ell=1}^{k}\frac{\pi_\ell}{n_\ell}\mathcal{S}_{i\bullet}^{j\ell}\right)^2 - \left(\frac{1}{n_j}\sum_{i=1}^{n_j}\sum_{\ell=1}^{k}\frac{\pi_\ell}{n_\ell}\mathcal{S}_{i\bullet}^{j\ell}\right)^2.$$

5. **Simulations** In this section, we present the results of simulations made on functional data in order to investigate the finite sample performances of the proposed tests and compare them to existing ones, namely the test of [4] based on GMMD and the FP test of [11]. We considered the case where $\mathcal{X} = L^2([0,1])$ with three populations ($k = 3$), and computed empirical sizes and powers through Monte Carlo simulations after generating functional data according to the three following models:

Model 1 : $X_j(t) = t(1-t) + \varepsilon_j(t)$, $j = 1,2,3$;

Model 2 : $X_1(t) = t(1-t)^5 + \varepsilon_1(t)$, $X_2(t) = t^2(1-t)^4 + \varepsilon_2(t)$, $X_3(t) = t^3(1-t)^3 + \nu_1(t)$;

Model 3 : $X_1(t) = t(1-t)^3 + \varepsilon_1(t)$, $X_2(t) = t(1-t)^3 - t + \nu_2(t)$, $X_3(t) = t(1-t)^3 + \varepsilon_2(t)$;

where $\varepsilon_1(t)$, $\varepsilon_2(t)$, $\varepsilon_3(t)$, $\nu_1(t)$ and $\nu_2(t)$ independent random variables such that $\varepsilon_j(t) \sim \mathcal{N}(0,t)$, $j = 1,2,3$, $\nu_1(t) \sim \mathcal{E}xp(t)$ and $\nu_2(t) \sim \mathcal{P}(t)$. Model 1, for which the hypothesis $\mathscr{H}_0$ holds, is used for computing empirical size. For the other two models $\mathscr{H}_0$ fails; so, they are used for computing empirical power.

We generated 2000 independent samples of each of the preceding processes in discretized versions on equispaced values $t_1, \cdots, t_{21}$ in the interval $[0,1]$, with $t_\ell = (\ell-1)/20$, $\ell = 1, \cdots, 21$. The sample sizes where $n_1 = n_2 = n_3 =$ 25, 50, 100, 200, 300. For performing our methods and that of [4], we used the Gaussian kernel $K(x,y) = \exp\left(-0.5\|x-y\|^2\right)$, and we computed the terms $\mathcal{K}_{ir}^{j\ell}$ as indicated in Example 2.1. Our first test and that of [4] were performed by using the permutation methods from which the corresponsing p-values were computed. Our secong test was performed with $\gamma = 0.41$. The nominal significance level was taken as $\alpha = 0.05$ for all tests. The obtained results are reported in Table 1 in which our tests based on (2.4) and (4.1) are denoted by M1a and M1b respectively, the one introduced in [4] is denoted by M2 and that of [11] is denoted by M3. They show that all the four methods are well calibrated for finite samples since the obtained empirical sizes are close to the nominal significance level, except when $n = 25$ for M2 and M3. Regarding the empirical power, it is seen that M1a and M1b performs very well, and give comparable results to M2. These three methods outperform M3 which gives very bad results for Model 3.

6. Proofs

6.1. Proof of Theorem 3.1 For $j \in \{1, \cdots, k\}$, putting

$$\widetilde{V}_j = \frac{1}{n_j}\sum_{i=1}^{n_j}\left(K(X_i^{(j)},\cdot) - m_j\right)^{\otimes 2},$$

we have $n\widehat{T}_n = A_n + B_n$, where

$$A_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\left(\frac{\pi_\ell}{\pi_j+\pi_\ell}\right)\left\{(n_j+n_\ell)\left(\left\|\widehat{V}_j-\widehat{V}_\ell\right\|^2-\left\|\widetilde{V}_j-\widetilde{V}_\ell\right\|^2\right)\right\},$$

$$B_n = n\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell\left\|\widetilde{V}_j-\widetilde{V}_\ell\right\|^2.$$

According to Lemma 1 of [15], $(n_j+n_\ell)\left(\left\|\widehat{V}_j-\widehat{V}_\ell\right\|^2-\left\|\widetilde{V}_j-\widetilde{V}_\ell\right\|^2\right) = o_P(1)$. Since $\lim_{n_j\to+\infty}\pi_j = \rho_j$ and $\lim_{n_l\to+\infty}\pi_\ell = \rho_\ell$, $\lim_{n_l,n_j\to+\infty}\frac{\pi_\ell}{\pi_j+\pi_\ell} = \frac{\rho_\ell}{\rho_j+\rho_\ell} = \frac{1}{1+\rho_j\rho_\ell^{-1}}$, we deduce that $A_n = o_P(1)$ and that $n\widehat{T}_n$ has the same limiting distribution

Table 1: Empirical sizes powers over 2000 replications for our tests based on (2.4) (denoted by M1a) and (4.1) (denoted by M1b), and the methods of [4] (denoted by M2) and [11] (denoted by M3), with nominal significance level $\alpha = 0.05$.

	$n_1 = n_2 = n_3$	M1a	M1b	M2	M3
Model 1	25	0.060	0.060	0.090	0.090
	50	0.050	0.060	0.060	0.070
	100	0.050	0.050	0.040	0.070
	200	0.040	0.050	0.040	0.060
	300	0.050	0.040	0.050	0.050
Model 2	25	0.960	0.910	0.940	0.780
	50	0.990	0.960	0.990	0.770
	100	1.000	0.990	1.000	0.890
	200	1.000	1.000	1.000	0.970
	300	1.000	1.000	1.000	0.990
Model 3	25	0.980	0.940	0.940	0.080
	50	0.990	0.980	0.990	0.080
	100	1.000	0.990	0.990	0.070
	200	1.000	1.000	1.000	0.060
	300	1.000	1.000	1.000	0.050

than B_n. It remains to determine this later. We have :

$$
\begin{aligned}
&\left\|\widetilde{V}_j-\widetilde{V}_\ell\right\|^2\\
= &\ \frac{1}{n_j^2}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j} < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes 2}-V, \left(K(X_r^{(j)},\cdot)-m_j\right)^{\otimes 2}-V >^2\\
&+\frac{1}{n_\ell^2}\sum_{i=1}^{n_\ell}\sum_{r=1}^{n_\ell} < \left(K(X_i^{(\ell)},\cdot)-m_\ell\right)^{\otimes 2}-V, \left(K(X_r^{(\ell)},\cdot)-m_\ell\right)^{\otimes 2}-V >^2\\
- &\ \frac{2}{n_jn_\ell}\sum_{i=1}^{n_j}\sum_{r=1}^{n_\ell} < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes 2}-V, \left(K(X_r^{(\ell)},\cdot)-m_\ell\right)^{\otimes 2}-V >^2\\
= &\ \frac{1}{n_j^2}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\widetilde{K}\left(X_i^{(j)},X_r^{(j)}\right)+\frac{1}{n_\ell^2}\sum_{i=1}^{n_\ell}\sum_{r=1}^{n_\ell}\widetilde{K}\left(X_i^{(\ell)},X_r^{(\ell)}\right)\\
- &\ \frac{2}{n_jn_\ell}\sum_{i=1}^{n_j}\sum_{r=1}^{n_\ell}\widetilde{K}\left(X_i^{(j)},X_r^{(\ell)}\right).
\end{aligned}
$$

Thus $B_n = C_n + D_n$, where

(6.1)
$$C_n=\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\left\{\frac{\pi_j^{-1}\pi_\ell}{n_j}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\widetilde{K}\left(X_i^{(j)},X_r^{(j)}\right)+\frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\sum_{r=1}^{n_\ell}\widetilde{K}\left(X_i^{(\ell)},X_r^{(\ell)}\right)\right\}$$

and

$$D_n=-2\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\frac{n\pi_\ell}{n_jn_\ell}\sum_{i=1}^{n_j}\sum_{r=1}^{n_\ell}\widetilde{K}\left(X_i^{(j)},X_r^{(\ell)}\right). \tag{6.2}$$

Since K is bounded, the integral operator $S_{\widetilde{K}}$ given in (3.1) is a Hilbert-Schmidt operator (see Theorem VI.22 in [18]). Therefore, from Mercer's theorem (see, e.g., [17]), we have:

$$\widetilde{K}(x,y)=\sum_{p=1}^{+\infty}\lambda_p e_p(x)e_p(y), \tag{6.3}$$

where $\{e_p\}_{p\geqslant 1}$ is an orthonormal basis of $L^2(\mathbb{P})$ such that e_p is an eigenvector of $S_{\widetilde{K}}$ associated to λ_p. Note that, from $\int_\chi \widetilde{K}(x,y)e_p(y)d\mathbb{P}(y)=\lambda_p e_p(x)$, we get:

$$\mathbb{E}\left(e_p(X_i^{(j)})\right)=\frac{1}{\lambda_p}\int_\chi \mathbb{E}\left(\widetilde{K}(X_i^{(j)},y)\right)e_p(y)d\mathbb{P}(y),$$

and since $\mathbb{E}\left(\widetilde{K}(X_i^{(j)},y)\right) =< \mathbb{E}\left(\left(K(X^{(\ell)},\cdot)-m_\ell\right)^{\otimes 2}-V\right),(K(y,\cdot)-m_\ell)^{\otimes 2}-V>= 0$, it follows that $\mathbb{E}\left(e_p(X_i^{(j)})\right)=0$. From (6.1) and (6.3),

$$\begin{aligned} C_n &= \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\left\{\frac{\pi_j^{-1}\pi_\ell}{n_j}\sum_{i=1}^{n_j}\sum_{r=1}^{n_j}\sum_{p=1}^{+\infty}\lambda_p e_p(X_i^{(j)})e_p(X_r^{(j)})\right\} \\ &\quad +\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\left\{\frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\sum_{r=1}^{n_\ell}\sum_{p=1}^{+\infty}\lambda_p e_p(X_i^{(\ell)})e_p(X_r^{(\ell)})\right\} \\ &= \sum_{p=1}^{+\infty}\lambda_p\left\{\sum_{j=1}^{k}\pi_j^{-1}\left(\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell\right)U_{n_j,p}^2\right\}+\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\sum_{p=1}^{+\infty}\lambda_p U_{n_\ell,p}^2, \end{aligned}$$

where $U_{n_j,p}=\frac{1}{\sqrt{n_j}}\sum_{i=1}^{n_j}e_p(X_i^{(j)})$. Since $\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_\ell = 1-\pi_j$, it follows

$$\begin{aligned} C_n &= \sum_{p=1}^{+\infty}\lambda_p\left\{\sum_{j=1}^{k}\pi_j^{-1}\left(1-\pi_j\right)U_{n_j,p}^2\right\}+\sum_{p=1}^{+\infty}\lambda_p\left\{\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}U_{n_\ell,p}^2\right\} \\ &= \sum_{p=1}^{+\infty}\lambda_p\left\{\sum_{j=1}^{k}\left(\pi_j^{-1}+k-2\right)U_{n_j,p}^2\right\}. \qquad (6.4) \end{aligned}$$

Further, from (6.2) and (6.3),

$$\begin{aligned} D_n &= -2\sum_{p=1}^{+\infty}\lambda_p\sum_{j=1}^{k}\frac{\pi_j^{-1/2}}{\sqrt{n_j}}\sum_{i=1}^{n_j}e_p(X_i^{(j)})\left(\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\frac{\pi_\ell^{1/2}}{\sqrt{n_\ell}}\sum_{r=1}^{n_\ell}e_p(X_r^{(\ell)})\right) \\ &= -2\sum_{p=1}^{+\infty}\lambda_p\sum_{j=1}^{k}\left\{\frac{\pi_j^{-1/2}}{\sqrt{n_j}}\sum_{i=1}^{n_j}e_p(X_i^{(j)})\right. \\ &\quad \left.\times\left(\sum_{\ell=1}^{k}\frac{\pi_\ell^{1/2}}{\sqrt{n_\ell}}\sum_{r=1}^{n_\ell}e_p(X_r^{(\ell)})-\frac{\pi_j^{1/2}}{\sqrt{n_j}}\sum_{r=1}^{n_j}e_p(X_r^{(j)})\right)\right\} \\ &= -2\sum_{p=1}^{+\infty}\lambda_p\left(\sum_{j=1}^{k}\pi_j^{-1/2}U_{n_j,p}\right)\left(\sum_{\ell=1}^{k}\pi_\ell^{1/2}U_{n_\ell,p}\right)+2\sum_{p=1}^{\infty}\lambda_p\sum_{j=1}^{k}U_{n_j,p}^2. \end{aligned}$$

Then, using the equalities $-2ab = (a-b)^2 - a^2 - b^2$ and $\pi_j^{-1/2} - \pi_j^{1/2} = \frac{1-\pi_j}{\sqrt{\pi_j}}$, it follows

$$\begin{aligned} D_n &= \sum_{p=1}^{+\infty} \lambda_p \left\{ 2\sum_{j=1}^{k} U_{n_j,p}^2 + \left(\sum_{j=1}^{k} \frac{1-\pi_j}{\sqrt{\pi_j}} U_{n_j,p} \right)^2 \right\} \\ &\quad - \sum_{p=1}^{+\infty} \lambda_p \left\{ \left(\sum_{j=1}^{k} \pi_j^{-1/2} U_{n_j,p} \right)^2 + \left(\sum_{j=1}^{k} \pi_j^{1/2} U_{n_j,p} \right)^2 \right\}. \end{aligned} \tag{6.5}$$

Using (6.4) and (6.5), we get

$$\begin{aligned} B_n &= \sum_{p=1}^{+\infty} \lambda_p \left\{ \sum_{j=1}^{k} \left(\pi_j^{-1} + k\right) U_{n_j,p}^2 + \left(\sum_{j=1}^{k} (\pi_j^{-1/2} - \pi_j^{1/2}) U_{n_j,p} \right)^2 \right\} \\ &\quad - \sum_{p=1}^{+\infty} \lambda_p \left\{ \left(\sum_{j=1}^{k} \pi_j^{-1/2} U_{n_j,p} \right)^2 + \left(\sum_{j=1}^{k} \pi_j^{1/2} U_{n_j,p} \right)^2 \right\} \\ &= \sum_{p=1}^{+\infty} \lambda_p \Phi_n(\mathcal{U}_{n,p}), \end{aligned}$$

where $\mathcal{U}_{n,p} = (U_{n_1,p}, \cdots, U_{n_k,p})$ and $\Phi_n : \mathbb{R}^p \to \mathbb{R}$ is the map defined by

$$\begin{aligned} \Phi_n(x) &= \sum_{j=1}^{k} \left(\pi_j^{-1} + k\right) x_j^2 + \left(\sum_{j=1}^{k} (\pi_j^{-1/2} - \pi_j^{1/2}) x_j \right)^2 \\ &\quad - \left(\sum_{j=1}^{k} \pi_j^{-1/2} x_j \right)^2 - \left(\sum_{j=1}^{k} \pi_j^{1/2} x_j \right)^2, \end{aligned}$$

where x_j is the j-th component of $x \in \mathbb{R}^p$. Since for any $(j,\ell) \in \{1, 2, \cdots, k\}^2$ with $j \neq \ell$, $U_{n_j,p}$ and $U_{n_\ell,p}$ are independent, we get by the central limit theorem, $U_{n_j,p} \xrightarrow{\mathscr{D}} Y_{j,p}$ as $n_j \to +\infty$ where $Y_{j,p} \sim \mathcal{N}(0,1)$, and $Y_{j,p}$ and $Y_{\ell,p}$ are independent if $j \neq \ell$. Consequently, $\mathcal{U}_{n,p} \xrightarrow{\mathscr{D}} \mathcal{U}_p := (Y_{1,p}, \cdots, Y_{k,p})$ as $\min_{1 \leqslant j \leqslant k}(n_j) \to +\infty$. So, considering the map $\Phi : \mathbb{R}^p \to \mathbb{R}$ defined by

$$\begin{aligned} \Phi(x) &= \sum_{j=1}^{k} \left(\rho_j^{-1} + k\right) x_j^2 + \left(\sum_{j=1}^{k} (\rho_j^{-1/2} - \rho_j^{1/2}) x_j \right)^2 \\ &\quad - \left(\sum_{j=1}^{k} \rho_j^{-1/2} x_j \right)^2 - \left(\sum_{j=1}^{k} \rho_j^{1/2} x_j \right)^2, \end{aligned}$$

we will show that

$$B_n \overset{\mathscr{D}}{\to} \sum_{p=1}^{+\infty} \lambda_p \, \Phi(\mathcal{U}_p) \quad \text{as} \quad \min_{1\leqslant j\leqslant k}(n_j) \to +\infty, \tag{6.6}$$

what leads to the required result since

$$\begin{aligned}\Phi(\mathcal{U}_p) &= \sum_{j=1}^{k} \left(\rho_j^{-1} + k - 2\right) Y_{j,p}^2 - 2\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k} \rho_\ell^{1/2}\rho_j^{-1/2} Y_{j,p}Y_{\ell,p} \\ &= (k-2)Z_p + \sum_{j=1}^{k}\left(\rho_j^{-1}Y_{j,p}^2 - 2\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell^{1/2}\rho_j^{-1/2}Y_{j,p}Y_{\ell,p}\right),\end{aligned}$$

where $Z_p = \sum_{j=1}^k Y_{j,p}^2 \sim \mathcal{X}_k^2$.

We denote by φ_U the characteristic function of the random variable U. The proof of (6.6) will be obtained from three steps.

First step: We put $S_n^{(q)} = \sum\limits_{p=1}^{q} \lambda_p \, \Phi(\mathcal{U}_{n,p})$ and we show that for all $\varepsilon > 0$ and all $t \in \mathbb{R}$, there exists $q \in \mathbb{N}^*$ such that $\left|\varphi_{B_n}(t) - \varphi_{S_n^{(q)}}(t)\right| < \frac{\varepsilon}{3}$ for n_j large enough $(j = 1, \cdots, k)$. Using the inequality $|e^{iz} - 1| \leqslant |z|$ for all $z \in \mathbb{R}$, we have for any $t \in \mathbb{R}$:

$$\begin{aligned}\left|\varphi_{B_n}(t) - \varphi_{S_n^{(q)}}(t)\right| &\leqslant \mathbb{E}\left(\left|e^{itB_n} - e^{itS_n^{(q)}}\right|\right) \leqslant \mathbb{E}\left(\left|e^{it\left(B_n - S_n^{(q)}\right)} - 1\right|\right) \\ &\leqslant |t|\,\mathbb{E}\left(\left|B_n - S_n^{(q)}\right|\right) \leqslant |t| \sum_{p=q+1}^{+\infty} \lambda_p \mathbb{E}\left(|\Phi_n(\mathcal{U}_{n,p})|\right).\end{aligned}$$

On the other hand, we have

$$\begin{aligned}\mathbb{E}\left(|\Phi_n(\mathcal{U}_{n,p})|\right) \leqslant\ & \sum_{j=1}^{k}\left(\pi_j^{-1} + k\right)\mathbb{E}\left(U_{n_j,p}^2\right) + \mathbb{E}\left(\left(\sum_{j=1}^{k}(\pi_i^{-1/2} - \pi_j^{1/2})U_{n_j,p}\right)^2\right) \\ & + \mathbb{E}\left(\left(\sum_{j=1}^{k}\pi_j^{-1/2}U_{n_j,p}\right)^2\right) + \mathbb{E}\left(\left(\sum_{j=1}^{k}\pi_j^{1/2}U_{n_j,p}\right)^2\right).\end{aligned}$$

Since $\mathbb{E}\left(e_p^2(X_i^{(j)})\right) = 1$ and $\mathbb{E}\left(e_p(X_i^{(j)})\, e_p(X_r^{(\ell)})\right) = \delta_{ir}\delta_{j\ell}$, it follows:

$$\mathbb{E}\left(U_{n_j,p}^2\right) = \frac{1}{n_j}\sum_{i=1}^{n_j}\mathbb{E}\left(e_p^2(X_i^{(j)})\right) + \frac{1}{n_j}\sum_{i=1}^{n_j}\sum_{\substack{r=1\\ r\neq i}}^{n_j}\mathbb{E}\left(e_p(X_i^{(j)})\, e_p(X_r^{(j)})\right) = 1$$

and for $j \neq \ell$,

$$\mathbb{E}\left(U_{n_j,p}U_{n_\ell,p}\right) = \frac{1}{\sqrt{n_j n_\ell}} \sum_{i=1}^{n_j} \sum_{r=1}^{n_\ell} \mathbb{E}\left(e_p(X_i^{(j)})\, e_p(X_r^{(\ell)})\right) = 0.$$

Therefore,

$$\begin{aligned} &\mathbb{E}\left(\left(\sum_{j=1}^{k}\left(\pi_j^{-1/2} - \pi_j^{1/2}\right) U_{n_j,p}\right)^2\right) \\ &= \sum_{j=1}^{k} \frac{(1-\pi_j)^2}{\pi_j}\mathbb{E}\left(U_{n_j,p}^2\right) + \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k} \frac{(1-\pi_j)(1-\pi_\ell)}{\sqrt{\pi_j\pi_\ell}}\mathbb{E}\left(U_{n_j,p}U_{n_\ell,p}\right) \\ &= \sum_{j=1}^{k} \frac{(1-\pi_j)^2}{\pi_j}, \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left(\left(\sum_{j=1}^{k}\pi_j^{-1/2}U_{n_j,p}\right)^2\right) \\ &= \sum_{j=1}^{k}\pi_j^{-1}\mathbb{E}\left(U_{n_j,p}^2\right) + \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_j^{-1}\pi_\ell^{-1}\mathbb{E}\left(U_{n_j,p}U_{n_\ell,p}\right) = \sum_{j=1}^{k}\pi_j^{-1}, \end{aligned}$$

$$\begin{aligned} &\mathbb{E}\left(\left(\sum_{j=1}^{k}\pi_j^{1/2}U_{n_j,p}\right)^2\right) \\ &= \sum_{j=1}^{k}\pi_j\mathbb{E}\left(U_{n_j,p}^2\right) + \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\pi_j^{1/2}\pi_\ell^{1/2}\mathbb{E}\left(U_{n_j,p}U_{n_\ell,p}\right) = 1, \end{aligned}$$

and, consequently,

$$\begin{aligned} \mathbb{E}\left(|\Phi_n(\mathcal{U}_{n,p})|\right) &\leqslant \sum_{j=1}^{k}\left(\pi_j^{-1} + k + \frac{(1-\pi_j)^2+1}{\pi_j} + \pi_j\right) \\ &= \sum_{j=1}^{k}\left(3\pi_j^{-1} + 2\pi_j + k - 2\right). \end{aligned}$$

Since $\lim_{n_j\to+\infty}\left(3\pi_j^{-1}+2\pi_j+k-2\right)=3\rho_j^{-1}+2\rho_j+k-2$, there exists $n_j^0\in\mathbb{N}^*$ such that, for any $n_j\geqslant n_j^0$, one has $3\pi_j^{-1}+2\pi_j+k-2\leqslant 3\rho_j^{-1}+2\rho_j+k-1$. Hence, for $n_1\geqslant n_1^0,\cdots,n_k\geqslant n_k^0$, $\mathbb{E}\left(|\Phi_n(\mathcal{U}_{n,p})|\right)\leqslant\sum_{j=1}^k\left(3\rho_j^{-1}+2\rho_j+k-1\right)$ and, consequently,

$$\left|\varphi_{B_n}(t)-\varphi_{S_n^{(q)}}(t)\right|\leqslant|t|\sum_{j=1}^k\left(3\rho_j^{-1}+2\rho_j+k-2\right)\sum_{p=q+1}^{+\infty}\lambda_p.$$

Since $\sum_{p=1}^{+\infty}\lambda_p<+\infty$, one has $\lim_{q\to+\infty}\left|\varphi_{B_n}(t)-\varphi_{S_n^{(q)}}(t)\right|=0$. So, there exists $q_0\in\mathbb{N}^*$ such that, for any $q\geqslant q_0$,

$$\left|\varphi_{B_n}(t)-\varphi_{S_n^{(q)}}(t)\right|<\frac{\varepsilon}{3}. \tag{6.7}$$

Second step: We consider $S_q=\sum_{p=1}^q\lambda_p\Phi(\mathcal{U}_p)$ and we show that we have $S_n^{(q)}\xrightarrow{\mathscr{D}}S_q$ as $\min\limits_{1\leqslant j\leqslant k}(n_j)\to+\infty$. It suffices to show that $S_n^{(q)}-S_q$ converges in probability to 0. Clearly,

$$\left|S_n^{(q)}-S_q\right|\leqslant\sum_{p=1}^q\lambda_p\Bigg(|\Phi_n(\mathcal{U}_{n,p})-\Phi(\mathcal{U}_{n,p})|+|\Phi(\mathcal{U}_{n,p})-\Phi(\mathcal{U}_p)|\Bigg). \tag{6.8}$$

Moreover, by using $a^2-b^2=(a-b)^2+2b(a-b)$, we get

$$\begin{aligned}
&|\Phi_n(\mathcal{U}_{n,p})-\Phi(\mathcal{U}_{n,p})|\\
&\quad\leqslant\Bigg\{\sum_{j=1}^k\left|\pi_j^{-1}-\rho_j^{-1}\right|+\left(\sum_{j=1}^k\left|\frac{1-\pi_j}{\sqrt{\pi_j}}-\frac{1-\rho_j}{\sqrt{\rho_j}}\right|\right)^2\\
&\quad+2\sum_{j=1}^k\sum_{\ell=1}^k\frac{1-\rho_\ell}{\sqrt{\rho_\ell}}\left|\frac{1-\pi_j}{\sqrt{\pi_j}}-\frac{1-\rho_j}{\sqrt{\rho_j}}\right|\\
&\quad+\left(\sum_{j=1}^k\left|\pi_j^{-1/2}-\rho_j^{-1/2}\right|\right)^2+2\sum_{j=1}^k\sum_{\ell=1}^k\rho_\ell^{-1/2}\left|\pi_j^{-1/2}-\rho_j^{-1/2}\right|\\
&\quad+\left(\sum_{j=1}^k\left|\pi_j^{1/2}-\rho_j^{1/2}\right|\right)^2+2\sum_{j=1}^k\sum_{\ell=1}^k\rho_\ell^{1/2}\left|\pi_j^{1/2}-\rho_j^{1/2}\right|\Bigg\}\|\mathcal{U}_{n,p}\|^2_{\mathbb{R}^p},
\end{aligned} \tag{6.9}$$

where $\|\cdot\|_{\mathbb{R}^p}$ denotes the Euclidean norm of $\mathbb{R}^p$. Since $\mathcal{U}_{n,p}$ converges in distribution to $\mathcal{U}_p$, we deduce from (6.8), (6.9) and of the continuity of Φ that $S_n^{(q)}-S_q$

converges in probability to 0 as $\min_{1\leqslant j\leqslant k}(n_j)\to+\infty$ and, therefore, that $S_n^{(q)}\stackrel{\mathscr{D}}{\to}S_q$ as $\min_{1\leqslant j\leqslant k}(n_j)\to+\infty$. Thus, there exists N_1 such that, for $\min_{1\leqslant j\leqslant k}(n_j)\geqslant N_1$, we have

$$\left|\varphi_{S_n^{(q)}}(t)-\varphi_{S_q}(t)\right|<\frac{\varepsilon}{3}. \tag{6.10}$$

Third step: We put $S=\sum_{p=1}^{+\infty}\lambda_p\Phi(\mathcal{U}_p)$ and we show that $S_q\stackrel{\mathscr{D}}{\to}S$ as $q\to+\infty$. We have:

$$\left|\varphi_{S_q}(t)-\varphi_S(t)\right| \leqslant \mathbb{E}\left(\left|e^{itS_q}-e^{itS}\right|\right)\leqslant|t|\,\mathbb{E}\left(|S_q-S|\right)\leqslant|t|\sum_{p=q+1}^{+\infty}\lambda_p\mathbb{E}\left(|\Phi(\mathcal{U}_p)|\right)$$

and

$$\begin{aligned}\mathbb{E}\left(|\Phi(\mathcal{U}_p)|\right) \leqslant\ & \sum_{j=1}^{k}\left(\rho_j^{-1}+k\right)\,\mathbb{E}\left(Y_{j,p}^2\right)+\mathbb{E}\left(\left(\sum_{j=1}^{k}\frac{1-\rho_j}{\sqrt{\rho_j}}Y_{j,p}\right)^2\right)\\ &+\mathbb{E}\left(\left(\sum_{j=1}^{k}\rho_j^{-1/2}Y_{j,p}\right)^2\right)+\mathbb{E}\left(\left(\sum_{j=1}^{k}\rho_j^{1/2}Y_{j,p}\right)^2\right).\end{aligned}$$

Since $\mathbb{E}\left(Y_{j,p}Y_{\ell,p}\right)=\delta_{j\ell}$, it follows

$$\begin{aligned}\mathbb{E}\left(\left(\sum_{j=1}^{k}\frac{1-\rho_j}{\sqrt{\rho_j}}Y_{j,p}\right)^2\right) &= \sum_{j=1}^{k}\frac{(1-\rho_j)^2}{\rho_j}\mathbb{E}\left(Y_{j,p}^2\right)\\ &\quad+\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\frac{(1-\rho_j)(1-\rho_\ell)}{\sqrt{\rho_j\rho_\ell}}\mathbb{E}\left(Y_{j,p}Y_{\ell,p}\right)\\ &= \sum_{j=1}^{k}\frac{(1-\rho_j)^2}{\rho_j},\end{aligned}$$

$$\begin{aligned}\mathbb{E}\left(\left(\sum_{j=1}^{k}\rho_j^{-1/2}Y_{j,p}\right)^2\right) &= \sum_{j=1}^{k}\rho_j^{-1}\mathbb{E}\left(Y_{j,p}^2\right)+\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_j^{-1/2}\rho_\ell^{-1/2}\mathbb{E}\left(Y_{j,p}Y_{\ell,p}\right)\\ &= \sum_{j=1}^{k}\rho_j^{-1},\end{aligned}$$

and

$$\mathbb{E}\left(\left(\sum_{j=1}^{k}\rho_j^{1/2}Y_{j,p}\right)^2\right) = \sum_{j=1}^{k}\rho_j\mathbb{E}\left(Y_{j,p}^2\right)+\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_j^{1/2}\rho_\ell^{1/2}\mathbb{E}\left(Y_{j,p}Y_{\ell,p}\right)=1.$$

Hence, $\mathbb{E}\left(|\Phi(\mathcal{U}_p)|\right)\leqslant\sum_{j=1}^{k}\left\{3\rho_j^{-1}+2\rho_j+k-2\right\}$ and

$$\left|\varphi_{S_q}(t)-\varphi_S(t)\right|\leqslant\sum_{j=1}^{k}\left\{3\rho_j^{-1}+2\rho_j+k-2\right\}\sum_{p=q+1}^{+\infty}\lambda_p.$$

Since $\sum_{p=1}^{+\infty}\lambda_p<+\infty$, we deduce that $\lim_{q\to+\infty}\left|\varphi_{S_q}(t)-\varphi_S(t)\right|=0$. So there exists $q_1\in\mathbb{N}^*$ such that, for any $q\geqslant q_1$,

$$\left|\varphi_{S_q}(t)-\varphi_S(t)\right|<\frac{\varepsilon}{3}. \tag{6.11}$$

Finally, putting $u=\max(q_0,q_1)$, $N_0=\max(n_1^0,\cdots,n_k^0)$ and using (6.7), (6.10) and (6.11), we deduce that if $\min_{1\leqslant j\leqslant k}(n_j)\geqslant\max(N_0,N_1)$ then

$$\begin{aligned}|\varphi_{B_n}(t)-\varphi_S(t)| \leqslant &\left|\varphi_{B_n}(t)-\varphi_{S_n^{(u)}}(t)\right|+\left|\varphi_{S_n^{(u)}}(t)-\varphi_{S_u}(t)\right|\\ &+|\varphi_{S_u}(t)-\varphi_S(t)|<\varepsilon.\end{aligned}$$

This shows that (6.6) holds.

6.2. Technical lemmas In this section, we give useful lemmas for proving Theorem 4.1. Let $\Delta_{j,\ell}=\|V_j-V_\ell\|_{\mathrm{HS}}^2$, $\widehat{\Delta}_{j,\ell}^{(n)}(\gamma)=\left\|\widehat{V}_j\right\|_{\mathrm{HS}}^2+\left\|\widehat{V}_\ell\right\|_{\mathrm{HS}}^2-\frac{2}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)$ $<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}_\ell>_{\mathrm{HS}}$ and $U_n=\sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left(\widehat{\Delta}_{j,\ell}^{(n)}(\gamma)-\Delta_{j,\ell}\right)$; then we have:

LEMMA 6.1. *Assume* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_4)$. *Then,*

$$\sqrt{n}\left(\widehat{T}_{n,\gamma}-MMVD^2(\mathbb{P}_1,\cdots,\mathbb{P}_k;\pi)\right)=U_n+o_P(1).$$

Proof. Clearly, we have the decomposition

$$\sqrt{n}\left(\widehat{T}_{n,\gamma}-\mathrm{MMVD}^2(\mathbb{P}_1,\cdots,\mathbb{P}_k;\pi)\right)=U_n+\epsilon_n,$$

where $\epsilon_n = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}(\pi_\ell - \rho_\ell)\,\widehat{\Delta}_{j,\ell}^{(n)}(\gamma)$, and it remains to prove that $\epsilon_n = o_P(1)$. Since $\epsilon_n = \epsilon_{1,n} - \epsilon_{2,n}$, where $\epsilon_{1,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}(\pi_\ell - \rho_\ell)\,\|\widehat{V}_j - \widehat{V}_\ell\|_{\mathrm{HS}}^2$ and

$$\epsilon_{2,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}(\pi_\ell - \rho_\ell)\left\{\frac{1}{n_j}\sum_{i=1}^{n_j}\big(w_{i,n_j}(\gamma)-1\big) < \Big(K(X_i^{(j)},\cdot)-\widehat{m}_j\Big)^{\otimes^2}, \widehat{V}_\ell >_{\mathrm{HS}}\right\},$$

it is enough to prove that $\epsilon_{1,n} = o_P(1)$ and $\epsilon_{2,n} = o_P(1)$. It is known from [13] (see the proof of Proposition 12) that $\left\|\widehat{V}_j - V_j\right\|_{\mathrm{HS}} = O_P(n_j^{-1/2})$. Then, using the inequality

$$\left\|\widehat{V}_j - \widehat{V}_\ell\right\|_{\mathrm{HS}}^2 \leqslant 4\left(\left\|\widehat{V}_j - V_j\right\|_{\mathrm{HS}}^2 + \|V_j - V_\ell\|_{\mathrm{HS}}^2 + \left\|V_\ell - \widehat{V}_\ell\right\|_{\mathrm{HS}}^2\right)$$

and Assumption $(\mathscr{A}_3)$, we deduce that each $\sqrt{n}(\pi_\ell - \rho_\ell)\,\|\widehat{V}_j - \widehat{V}_\ell\|_{\mathrm{HS}}^2$ converges in probability to 0 as $\min_{1\leqslant j\leqslant k}(n_j) \to +\infty$ and, consequently, that $\epsilon_{1,n} = o_P(1)$. Using the Cauchy-Schwartz inequality, the equality

$$\begin{aligned}
&\Big(K(X_i^{(j)},\cdot)-\widehat{m}_j\Big)^{\otimes^2}\\
&= \Big(K(X_i^{(j)},\cdot)-m_j\Big)^{\otimes^2} + \Big(K(X_i^{(j)},\cdot)-m_j\Big)\otimes(m_j-\widehat{m}_j)\\
&\quad + (m_j-\widehat{m}_j)\otimes\Big(K(X_i^{(j)},\cdot)-m_j\Big) + (\widehat{m}_j - m_j)^{\otimes^2},
\end{aligned} \tag{6.12}$$

the properties $(a\otimes b)^* = (b\otimes a)$, $\|a\otimes b\|_{\mathrm{HS}} = \|a\|_{\mathcal{H}}\|b\|_{\mathcal{H}}$, $\|T\|_{\mathrm{HS}} = \|T^*\|_{\mathrm{HS}}$ and the inequality $\|\widehat{V}_\ell\|_{\mathrm{HS}} \leqslant \|\widehat{V}_\ell - V_\ell\|_{\mathrm{HS}} + \|V_\ell\|_{\mathrm{HS}}$, we get

$$\left|\sqrt{n}(\pi_\ell - \rho_\ell)\left\{\frac{1}{n_j}\sum_{i=1}^{n_j}\big(w_{i,n_j}(\gamma)-1\big) < \Big(K(X_i^{(j)},\cdot)-\widehat{m}_j\Big)^{\otimes^2}, \widehat{V}_\ell >_{\mathrm{HS}}\right\}\right| \leqslant E_n + 2F_n + G_n,$$

where

$$\begin{aligned}
E_n &= \left|\sqrt{n}\,(\pi_\ell - \rho_\ell)\right|\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\big(w_{i,n_j}(\gamma)-1\big)\Big(K(X_i^{(j)},\cdot)-m_j\Big)^{\otimes^2}\right\|_{\mathrm{HS}}\\
&\quad\times\left\{\left\|\widehat{V}_\ell - V_\ell\right\|_{\mathrm{HS}} + \|V_\ell\|_{\mathrm{HS}}\right\},
\end{aligned}$$

$$F_n = \left|\sqrt{n}\,(\pi_\ell - \rho_\ell)\right| \, \|\widehat{m}_j - m_j\|_{\mathcal{H}} \left\| \frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right\|_{\mathcal{H}} \times\left\{ \left\|\widehat{V}_\ell - V_\ell\right\|_{\mathcal{H}} + \|V_\ell\|_{\text{HS}} \right\},$$

$$G_n = |\pi_\ell - \rho_\ell| \left|\frac{\sqrt{n}}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\right| \|\widehat{m}_j - m_j\|_{\mathcal{H}}^2 \left\{ \left\|\widehat{V}_\ell - V_\ell\right\|_{\mathcal{H}} + \|V_\ell\|_{\text{HS}} \right\},$$

Since

$$\begin{aligned}
&\left\|\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\text{HS}} \\
&= \left\|K(X_i^{(j)},\cdot)-m_j\right\|_{\mathcal{H}}^2 \leqslant 2\left\|K(X_i^{(j)},\cdot)\right\|_{\mathcal{H}}^2 + 2\,\|m_j\|_{\mathcal{H}}^2 \\
&= 2K(X_i^{(j)},X_i^{(j)}) + 2\,\|m_j\|_{\mathcal{H}}^2 \\
&\leqslant 2\|K\|_\infty + 2\,\|m_j\|_{\mathcal{H}}^2\,,
\end{aligned}$$

we deduce from Lemma 1 of [16] that

$$\left\| \frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\text{HS}} = o_P(1)$$

and

$$\left\| \frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right\|_{\mathcal{H}} = o_P(1).$$

These properties, together with $(\mathscr{A}_3)$ and the properties $\left\|\widehat{V}_\ell - V_\ell\right\|_{\text{HS}} = O_P(n_\ell^{-1/2})$ and $\|\widehat{m}_j - m_j\|_{\mathcal{H}} = O_P(n_j^{-1/2})$, obtained from the central limit theorem, allow to conclude that $E_n = o_P(1)$ and $F_n = o_P(1)$. Further, using $(\mathscr{A}_4)$ we get

$$\left|\frac{\sqrt{n}}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\right| = \sqrt{n}\left|\frac{1}{n_j}\sum_{i=1}^{n_j} w_{i,n_j}(\gamma)-1\right| \leqslant \frac{\sqrt{n}}{n_j}\tau = \frac{\pi_j^{-1}}{\sqrt{n}}\tau \tag{6.13}$$

for n_j large enough. Then, since $\lim_{n_j\to+\infty}\pi_j = \rho_j$ for any $j \in \{1,\cdots,k\}$, it follows that $G_n = o_P(1)$. So, we have proved that $\epsilon_{2,n} = o_P(1)$. □

Now, let us consider:

$$\Gamma_{1,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k} \rho_\ell \left\{ \left\|\widehat{V}_j - V_j\right\|_{\text{HS}}^2 + \left\|\widehat{V}_\ell - V_\ell\right\|_{\text{HS}}^2 \right\},$$

$$\Gamma_{2,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k} \rho_\ell < \widehat{V}_j - V_j, \widehat{V}_\ell - V_\ell >_{\text{HS}},$$

$$\Gamma_{3,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\frac{\rho_\ell}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)$$
$$\times\left\{<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}_\ell - V_\ell>_{\text{HS}} + <V_j,V_\ell>_{\text{HS}}\right\}.$$

Then, we have:

LEMMA 6.2. *Assume* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_4)$. *Then, for any* $l\in\{1,2,3\}$, *one has* $\Gamma_{l,n} = o_P(1)$.

Proof. Since $\left\|\widehat{V}_j - V_j\right\|_{\text{HS}} = O_P(n_j^{-1/2})$ for any $j\in\{1,,\cdots,k\}$, it can be deduced that $\sqrt{n}\left\|\widehat{V}_j - V_j\right\|^2_{\text{HS}} = O_P(\frac{\sqrt{n}}{n_j})$. Then, the property $\lim_{n_j\to+\infty}\frac{\sqrt{n}}{n_j}=0$ allows to conclude that $\left\|\widehat{V}_j - V_j\right\|_{\text{HS}} = o_P(1)$ for any $j\in\{1,,\cdots,k\}$, what yields $\Gamma_{1,n}=o_P(1)$. Using the Cauchy-Schwartz inequality, we obtain

$$|\Gamma_{2,n}| \leqslant \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left(\sqrt{n}\left\|\widehat{V}_j - V_j\right\|_{\text{HS}}\left\|\widehat{V}_\ell - V_\ell\right\|_{\text{HS}}\right).$$

Since

$$\sqrt{n}\left\|\widehat{V}_j - V_j\right\|_{\text{HS}}\left\|\widehat{V}_\ell - V_\ell\right\|_{\text{HS}} = O_P(\frac{\sqrt{n}}{\sqrt{n_j}\sqrt{n_\ell}}),$$

it follows from $\lim\limits_{n_j,n_\ell\to+\infty}\frac{\sqrt{n}}{\sqrt{n_j}\sqrt{n_\ell}}=0$ that

$$\sqrt{n}\left\|\widehat{V}_j - V_j\right\|_{\text{HS}}\left\|\widehat{V}_\ell - V_\ell\right\|_{\text{HS}} = o_P(1)$$

and, therefore, that $\Gamma_{2,n}=o_P(1)$. Further, using the Cauchy-Schwartz inequality, the equality (6.12) and the properties of tensor product and Hilbert-Schmidt norm used in the proof of Lemma 6.1, we get $|\Gamma_{3,n}|\leqslant H_n + 2I_n + J_n$, where

$$H_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\text{HS}}$$
$$\times\left\|\sqrt{n}\left(\widehat{V}_\ell - V_\ell\right)\right\|_{\text{HS}},$$

$$I_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right\|_{\mathcal{H}}$$
$$\times\left\|\widehat{m}-m_j\right\|_{\mathcal{H}}\left\|\sqrt{n}\left(\widehat{V}_\ell - V_\ell\right)\right\|_{\text{HS}},$$

$$J_n = \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left|\frac{\sqrt{n}}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\right| \times\left\{\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2\left\|\widehat{V}_\ell-V_\ell\right\|_{\mathrm{HS}}+\|V_j\|_{\mathrm{HS}}\|V_\ell\|_{\mathrm{HS}}\right\}.$$

From the proof of Lemma 6.1,

$$\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\mathrm{HS}}=o_P(1)$$

and

$$\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right\|_{\mathcal{H}}=o_P(1).$$

These properties, together with the equalities $\left\|\sqrt{n}\left(\widehat{V}_\ell-V_\ell\right)\right\|_{\mathrm{HS}}=O_P(\pi_\ell^{-1/2})$ and $\|\widehat{m}_j-m_j\|_{\mathcal{H}}=O_P(n_j^{-1/2})$ allow to conclude that $H_n=o_P(1)$ and $I_n=o_P(1)$. Further, using (6.13) and $\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2\left\|\widehat{V}_\ell-V_\ell\right\|_{\mathrm{HS}}=O_P(n_j^{-1}n_\ell^{-1/2})$, we obtain $J_n=o_P(1)$. □

Further, putting

$$\Gamma_{4,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\{\frac{2}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-V_j,V_j>_{\mathrm{HS}} +\frac{2}{n_\ell}\sum_{i=1}^{n_\ell}<\left(K(X_i^{(\ell)},\cdot)-\widehat{m}_\ell\right)^{\otimes^2}-V_\ell,V_\ell>_{\mathrm{HS}} -\frac{2}{n_\ell}\sum_{i=1}^{n_\ell}<\left(K(X_i^{(\ell)},\cdot)-\widehat{m}_\ell\right)^{\otimes^2}-V_\ell,V_j>_{\mathrm{HS}}\right\}$$

and

$$\Lambda_n = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\{\frac{2}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}-V_j,V_j>_{\mathrm{HS}} +\frac{2}{n_\ell}\sum_{i=1}^{n_\ell}<\left(K(X_i^{(\ell)},\cdot)-m_\ell\right)^{\otimes^2}-V_\ell,V_\ell>_{\mathrm{HS}} -\frac{2}{n_\ell}\sum_{i=1}^{n_\ell}<\left(K(X_i^{(\ell)},\cdot)-m_\ell\right)^{\otimes^2}-V_\ell,V_j>_{\mathrm{HS}}\right\},$$

we have:

LEMMA 6.3. *Assume* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_4)$. *Then,* $\Gamma_{4,n} = \Lambda_n + o_P(1)$.

Proof. Using (6.12), we get $\Gamma_{4,n} - \Lambda_n = 2\delta_n$, where

$$\begin{aligned}
\delta_n &= \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\Bigg[\frac{1}{n_j}\sum_{i=1}^{n_j}\Big\{ < \Big(K(X_i^{(j)},\cdot)-m_j\Big)\otimes(m_j-\widehat{m}_j)\,,V_j >_{\mathrm{HS}} \\
&\qquad + < (m_j-\widehat{m}_j)\otimes\Big(K(X_i^{(j)},\cdot)-m_j\Big)\,,V_j >_{\mathrm{HS}} \\
&\qquad + < (\widehat{m}_j-m_j)^{\otimes^2}\,,V_j >_{\mathrm{HS}}\Big\} \\
&\quad + \frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\Big\{ < \Big(K(X_i^{(\ell)},\cdot)-m_\ell\Big)\otimes(m_\ell-\widehat{m}_\ell)\,,V_\ell >_{\mathrm{HS}} \\
&\qquad + < (m_\ell-\widehat{m}_\ell)\otimes\Big(K(X_i^{(\ell)},\cdot)-m_\ell\Big)\,,V_\ell >_{\mathrm{HS}} \\
&\qquad + < (\widehat{m}_\ell-m_\ell)^{\otimes^2}\,,V_\ell >_{\mathrm{HS}}\Big\} \\
&\quad - \frac{1}{n_\ell}\sum_{i=1}^{n_\ell}\Big\{ < \Big(K(X_i^{(\ell)},\cdot)-m_\ell\Big)\otimes(m_\ell-\widehat{m}_\ell)\,,V_j >_{\mathrm{HS}} \\
&\qquad + < (m_\ell-\widehat{m}_\ell)\otimes\Big(K(X_i^{(\ell)},\cdot)-m_\ell\Big)\,,V_j >_{\mathrm{HS}} \\
&\qquad + < (\widehat{m}_\ell-m_\ell)^{\otimes^2}\,,V_j >_{\mathrm{HS}}\Big\}\Bigg] \\
&= \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\Big\{ < (\widehat{m}_\ell-m_\ell)^{\otimes^2}\,,V_j-V_\ell >_{\mathrm{HS}} - < (\widehat{m}_j-m_j)^{\otimes^2}\,,V_j >_{\mathrm{HS}}\Big\}.
\end{aligned}$$

Using the Cauchy-Schwartz inequality and a property of the Hilbert-Schmidt norm recalled in the proof of Lemma 6.1, we get

$$|\delta_n| \leqslant \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\Big\{\sqrt{n}\,\|\widehat{m}_\ell-m_\ell\|_{\mathcal{H}}^2\,\|V_j-V_\ell\|_{\mathrm{HS}} + \sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2\,\|V_j\|_{\mathrm{HS}}\Big\}.$$

Since $\sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2 = O_P(\sqrt{n}/n_j)$ and $\lim_{n_j\to+\infty}(\sqrt{n}/n_j) = 0$, we deduce from the preceding inequality that $\delta_n = o_P(1)$. □

Finally, putting

$$\Gamma_{5,n} = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\{\frac{2}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma) < \Big(K(X_i^{(j)},\cdot)-\widehat{m}_j\Big)^{\otimes^2} - V_j, V_\ell >_{\mathrm{HS}}\right\}$$

and

$$\Theta_n = \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\left\{\frac{2}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma) < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2} - V_j, V_\ell >_{\mathrm{HS}}\right\},$$

we have:

Lemma 6.4. *Assume* $(\mathscr{A}_1)$ *to* $(\mathscr{A}_4)$. *Then,* $\Gamma_{5,n} = \Theta_n + o_P(1)$.

Proof. Using (6.12), we get $\Gamma_{5,n} - \Theta_n = 2\zeta_n$, where

$$\begin{aligned}\zeta_n &= \sqrt{n}\sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\Bigg\{ < \left(\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right)\\ &\qquad\qquad \otimes\left(m_j-\widehat{m}_j\right), V_\ell >_{\mathrm{HS}}\\ &\quad + < (m_j-\widehat{m}_j)\otimes\left(\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right), V_\ell >_{\mathrm{HS}}\\ &\quad + < (\widehat{m}_j-m_j)^{\otimes^2}, V_\ell >_{\mathrm{HS}}\left(\frac{1}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)-2\right)\Bigg\}.\end{aligned}$$

Using the Cauchy-Schwartz inequality and a property of the Hilbert-Schmidt norm recalled in the proof of Lemma 6.1, we get

$$\begin{aligned}|\zeta_n| &\leqslant \sum_{j=1}^{k}\sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell\Bigg\{2\sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}\left\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\right\|_{\mathcal{H}}\\ &\qquad +\sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2\left(\left|\frac{1}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)-1\right|+1\right)\Bigg\}\|V_\ell\|_{\mathrm{HS}}.\end{aligned}$$

From the proof of Lemma 6.1, we have

$$\|\frac{1}{n_j}\sum_{i=1}^{n_j}\left(w_{i,n_j}(\gamma)-1\right)\left(K(X_i^{(j)},\cdot)-m_j\right)\|_{\mathcal{H}} = o_P(1).$$

Further, using (6.13), we see that $\left|\frac{1}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)-1\right|$ is bounded. These properties, together with the equalities

$$\sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}} = O_P(\pi_j^{-1/2})$$

and

$$\sqrt{n}\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2 = O_P(\sqrt{n}/n_j),$$

with $\lim_{n_j\to+\infty}\left(\pi_j^{-1}\right) = \rho_j^{-1}$ and $\lim_{n_j\to+\infty}\left(\sqrt{n}/n_j\right) = 0$, allow to conclude that $\zeta_n = o_P(1)$. □

6.3. Proof of Theorem 4.1 From Lemma 6.1 and Slutsky's theorem it is seen that the sequence $\sqrt{n}\left(\widehat{T}_{n,\gamma} - \mathrm{MMVD}^2(\mathbb{P}_1, \cdots, \mathbb{P}_k; \pi)\right)$ has the same limiting distribution than U_n; therefore, it remains to derive this latter. From easy calculations, we get $U_n = \Gamma_{1,n} - 2\Gamma_{2,n} - 2\Gamma_{3,n} + \Gamma_{4,n} - \Gamma_{5,n}$. Then, using Lemma 6.2, Lemma 6.3 and Lemma 6.4 we obtain $U_n = \Lambda_n - \Theta_n + o_P(1)$, from what it is deduced that the required limiting distribution is the same as that of $\Lambda_n - \Theta_n$. However, inverting sums in the second and third terms of Λ_n yields:

$$
\begin{aligned}
\Lambda_n &= \sum_{j=1}^{k} 2\pi_j^{-1/2} \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - m_j\right)^{\otimes^2} - V_j, \left(\sum_{\substack{\ell=1\\ \ell\neq j}}^{k} \rho_\ell\right) V_j >_{\mathrm{HS}} \\
&\quad + \sum_{\ell=1}^{k} 2\pi_\ell^{-1/2} \frac{1}{\sqrt{n_\ell}} \sum_{i=1}^{n_\ell} < \left(K(X_i^{(\ell)}, \cdot) - m_\ell\right)^{\otimes^2} - V_\ell, \left(\sum_{\substack{j=1\\ j\neq \ell}}^{k} 1\right) \rho_\ell V_\ell >_{\mathrm{HS}} \\
&\quad - \sum_{\ell=1}^{k} 2\pi_\ell^{-1/2} \frac{1}{\sqrt{n_\ell}} \sum_{i=1}^{n_\ell} < \left(K(X_i^{(\ell)}, \cdot) - m_\ell\right)^{\otimes^2} - V_\ell, \rho_\ell \sum_{\substack{j=1\\ j\neq \ell}}^{k} V_j >_{\mathrm{HS}} \\
&= \sum_{j=1}^{k} 2\pi_j^{-1/2} \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - m_j\right)^{\otimes^2} - V_j, (1-\rho_j)\, V_j >_{\mathrm{HS}} \\
&\quad + \sum_{\ell=1}^{k} 2\pi_\ell^{-1/2} \frac{1}{\sqrt{n_\ell}} \sum_{i=1}^{n_\ell} < \left(K(X_i^{(\ell)}, \cdot) - m_\ell\right)^{\otimes^2} - V_\ell, (k-1)\, \rho_\ell V_\ell >_{\mathrm{HS}} \\
&\quad - \sum_{\ell=1}^{k} 2\pi_\ell^{-1/2} \frac{1}{\sqrt{n_\ell}} \sum_{i=1}^{n_\ell} < \left(K(X_i^{(\ell)}, \cdot) - m_\ell\right)^{\otimes^2} - V_\ell, \rho_\ell\, (\nu - V_\ell) >_{\mathrm{HS}} \\
&= \sum_{j=1}^{k} 2\pi_j^{-1/2} \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - m_j\right)^{\otimes^2} - V_j, (1-\rho_j + k\rho_j)V_j - \rho_j \nu >_{\mathrm{HS}} \\
&= \sum_{j=1}^{k} 2\pi_j^{-1/2} \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} \mathcal{U}_j\left(X_i^{(j)}\right).
\end{aligned}
$$

On the other hand,

$$\begin{aligned}\Theta_n &= \sum_{j=1}^{k} 2\pi_j^{-1/2}\frac{1}{\sqrt{n_j}}\sum_{i=1}^{n_j} w_{i,n_j}(\gamma) < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2} - V_j, \sum_{\substack{\ell=1\\ \ell\neq j}}^{k}\rho_\ell V_\ell >_{\mathrm{HS}}\\ &= \sum_{j=1}^{k} 2\pi_j^{-1/2}\frac{1}{\sqrt{n_j}}\sum_{i=1}^{n_j} w_{i,n_j}(\gamma) < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2} - V_j, V-\rho_j V_j >_{\mathrm{HS}}\\ &= \sum_{j=1}^{k} 2\pi_j^{-1/2}\frac{1}{\sqrt{n_j}}\sum_{i=1}^{n_j} w_{i,n_j}(\gamma)\mathcal{V}_j\left(X_i^{(j)}\right).\end{aligned}$$

Hence

$$\Lambda_n - \Theta_n = \sum_{j=1}^{k} 2\pi_j^{-1/2}\mathcal{Y}_{n,j,\gamma}, \tag{6.14}$$

where $\mathcal{Y}_{n,j,\gamma} = \frac{1}{\sqrt{n_j}}\sum_{i=1}^{n_j}\left\{\mathcal{U}_j(X_i^{(j)}) - w_{i,n_j}(\gamma)\mathcal{V}_j(X_i^{(j)})\right\}$. By similar arguments as in the proof of Theorem 1 in [15] we obtain that, for any $\varepsilon > 0$,

$$\lim_{n_j\to+\infty}\left(s_{n,j,\gamma}^{-2}\sum_{i=1}^{n_j}\int_{\{x:|\mathcal{U}_j(x)-w_{i,n_j}(\gamma)\mathcal{V}_j(x)|>\varepsilon s_{n,j,\gamma}\}}\left(\mathcal{U}_j(x)-w_{i,n_j}(\gamma)\mathcal{V}_j(x)\right)^2 d\mathbb{P}_j(x)\right) = 0,$$

where $s_{n,j,\gamma}^2 = \sum_{i=1}^{n_j}\mathrm{Var}\left(\mathcal{U}_j(X_i^{(j)}) - w_{i,n_j}(\gamma)\mathcal{V}_j(X_i^{(j)})\right)$. Therefore, by Section 1.9.3 in [19] we get $\sqrt{n_j}s_{n,j,\gamma}^{-1}\mathcal{Y}_{n,j,\gamma}\xrightarrow{\mathscr{D}}\mathcal{N}(0,1)$ as $n_j\to+\infty$. However,

$$\begin{aligned}\left(\frac{s_{n,j,\gamma}}{\sqrt{n_j}}\right)^2 &= \mathrm{Var}\left(\mathcal{U}_j(X_1^{(j)})\right) + \left(\frac{1}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}^2(\gamma)\right)\mathrm{Var}\left(\mathcal{V}_j(X_1^{(j)})\right)\\ &\quad -2\left(\frac{1}{n_j}\sum_{i=1}^{n_j}w_{i,n_j}(\gamma)\right)\mathrm{Cov}\left(\mathcal{U}_j(X_1^{(j)}),\mathcal{V}_j(X_1^{(j)})\right);\end{aligned}$$

then, using $(\mathscr{A}_4)$ and $(\mathscr{A}_6)$, we get $\lim_{n_j\to+\infty}\left(n_j^{-1}s_{n,j,\gamma}^2\right) = \sigma_j^2(\gamma)$. Hence, $\mathcal{Y}_{n,j,\gamma}\xrightarrow{\mathscr{D}}\mathcal{N}\left(0,\sigma_j^2(\gamma)\right)$. Since $\mathcal{Y}_{n,j,\gamma}$ and $\mathcal{Y}_{n,\ell,\gamma}$ are independent when $j\neq\ell$, we deduce from (6.14) and from $\lim_{n_j\to+\infty}(\pi_j) = \rho_j$ that $\Lambda_n - \Theta_n\xrightarrow{\mathscr{D}}\mathcal{N}\left(0,\sigma_\gamma^2\right)$ as $\min_{1\leqslant j\leqslant}(n_j)\to+\infty$, where $\sigma_\gamma^2 = \sum_{j=1}^{s}4\rho_j^{-1}\sigma_j^2(\gamma)$.

6.4. Proof of Theorem 4.2 It is enough to prove that, for any $j \in \{1, \cdots, k\}$, the sequence $\widehat{\theta}_j^2$ converges in probability to θ^2 as $n_j \to +\infty$. Indeed, in this case $\widehat{\theta}^2$ converges in probability to $\theta^2 \sum_{j=1}^k \rho_j = \theta^2$ since $\lim_{n_j \to +\infty} \pi_j = \rho_j$, and thus $\widehat{\sigma}_\gamma^2$ converges in probability to σ_γ^2. We have $\widehat{\theta}_j^2 = \widehat{\theta}_{1,j}^2 - \widehat{\theta}_{2,j}^2$, where

$$\widehat{\theta}_{1,j}^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, \widehat{V} >_{\text{HS}}^2$$

and

$$\widehat{\theta}_{2,j} = \frac{1}{n_j} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, \widehat{V} >_{\text{HS}} .$$

Clearly,

$$\begin{aligned} \widehat{\theta}_{1,j}^2 &= \frac{1}{n_j} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, \widehat{V} - V >_{\text{HS}}^2 \\ &\quad + \frac{1}{n_j} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, V >_{\text{HS}}^2 \\ &\quad + \frac{2}{n_j} \sum_{i=1}^{n_j} < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, \widehat{V} - V >_{\text{HS}} \\ &\qquad \times < \left(K(X_i^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, V >_{\text{HS}}, \end{aligned} \tag{6.15}$$

and using the Cauchy-Schwartz inequality, a property of tensor product recalled in Lemma 6.1, the reproducing property of $\mathcal{H}$ and assumption $(\mathscr{A}_1)$, we get

$$\begin{aligned} &\frac{1}{n_j} \sum_{i=1}^{n_j} < \left(K(X_1^{(j)}, \cdot) - \widehat{m}_j \right)^{\otimes^2}, \widehat{V} - V >_{\text{HS}}^2 \\ &\leqslant \frac{1}{n_j} \sum_{i=1}^{n_j} \left\| K(X_i^{(j)}, \cdot) - \widehat{m}_j \right\|_{\mathcal{H}}^4 \left\| \widehat{V} - V \right\|_{\text{HS}}^2 \\ &\leqslant \frac{4}{n_j} \sum_{i=1}^{n_j} \left(\left\| K(X_i^{(j)}, \cdot) - m_j \right\|_{\mathcal{H}}^4 + \|\widehat{m}_j - m_j\|_{\mathcal{H}}^4 \right) \left\| \widehat{V} - V \right\|_{\text{HS}}^2 \\ &\leqslant \left(16 \|K\|_\infty^2 + 16 \|m_j\|_{\mathcal{H}}^4 + 4 \|\widehat{m}_j - m_j\|_{\mathcal{H}}^4 \right) \left\| \widehat{V} - V \right\|_{\text{HS}}^2 . \end{aligned} \tag{6.16}$$

Since $\left\| \widehat{V} - V \right\|_{\text{HS}} = \left\| \sum_{j=1}^k \pi_j \widehat{V}_j - \sum_{j=1}^k \rho_j V_j \right\|_{\text{HS}} \leqslant \sum_{j=1}^k \left(\pi_j \left\| \widehat{V}_j - V_j \right\|_{\text{HS}} + |\pi_j - \rho_j| \, \|V_j\|_{\text{HS}} \right)$, and $\lim_{n_j \to +\infty} \pi_j = \rho_j$ together with $\pi_j \left\| \widehat{V}_j - V_j \right\|_{\text{HS}} = O_P(\frac{\pi_j}{\sqrt{n_j}})$

and $\lim_{n_j\to+\infty}\frac{\pi_j}{\sqrt{n_j}}=0$, we deduce that $\| \widehat{V}-V \|_{\mathrm{HS}}=o_P(1)$. An use of this property together with $\|\widehat{m}_j-m_j\|_{\mathcal{H}}^4=O_P(n_j^{-2})$ in (6.16) yields

$$\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}-V>_{\mathrm{HS}}^2=o_P(1). \tag{6.17}$$

From similar reasoning than above, we obtain

$$\begin{aligned}&\left|\frac{2}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}-V>_{\mathrm{HS}}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}\right|\\
&\quad\leqslant\frac{4}{n_j}\sum_{i=1}^{n_j}\left(\left\|K(X_i^{(j)},\cdot)-m_j\right\|_{\mathcal{H}}^4+\|\widehat{m}_j-m_j\|_{\mathcal{H}}^4\right)\left\|\widehat{V}-V\right\|_{\mathrm{HS}}\|V\|_{\mathrm{HS}}\\
&\quad\leqslant\left(16\,\|K\|_\infty^2+16\,\|m_j\|_{\mathcal{H}}^4+4\,\|\widehat{m}_j-m_j\|_{\mathcal{H}}^4\right)\left\|\widehat{V}-V\right\|_{\mathrm{HS}}\|V\|_{\mathrm{HS}}\end{aligned}$$

from what we deduce that

$$\begin{aligned}&\frac{2}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}-V>_{\mathrm{HS}}\\
&\qquad\times<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}=o_P(1).\end{aligned}\tag{6.18}$$

Further,

$$\begin{aligned}&\left|\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2-\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2\right|\\
&=\left|\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}\right.\\
&\qquad\left.\times<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}+\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}\right|\\
&\leqslant\frac{1}{n_j}\sum_{i=1}^{n_j}\left\|\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\mathrm{HS}}\\
&\qquad\times\left\{\left\|K(X_i^{(j)},\cdot)-\widehat{m}_j\right\|_{\mathcal{H}}^2+\left\|K(X_i^{(j)},\cdot)-m_j\right\|_{\mathcal{H}}^2\right\}\|V\|_{\mathrm{HS}}^2\\
&\leqslant\frac{1}{n_j}\sum_{i=1}^{n_j}\left\|\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\mathrm{HS}}\\
&\qquad\times\left\{6\,\|K\|_\infty+2\,\|m_j\|_{\mathcal{H}}^2\right\}\|V\|_{\mathrm{HS}}^2,\end{aligned}\tag{6.19}$$

and using (6.12) we get

$$\begin{aligned}
&\frac{1}{n_j}\sum_{i=1}^{n_j}\left\|\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\mathrm{HS}}\\
&\quad\leqslant \frac{1}{n_j}\sum_{i=1}^{n_j}\left\{\left\|\left(K(X_i^{(j)},\cdot)-m_j\right)\otimes(m_j-\widehat{m}_j)\right\|_{\mathrm{HS}}\right.\\
&\quad\left.+\left\|\left(K(X_i^{(j)},\cdot)-m_j\right)\otimes(m_j-\widehat{m}_j)\right\|_{\mathrm{HS}}+\left\|(\widehat{m}_j-m_j)^{\otimes^2}\right\|_{\mathrm{HS}}\right\}\\
&\quad=\frac{1}{n_j}\sum_{i=1}^{n_j}\left\{2\left\|K(X_i^{(j)},\cdot)-m_j\right\|_{\mathcal{H}}\|\widehat{m}_j-m_j\|_{\mathcal{H}}+\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2\right\}\\
&\quad\leqslant 2\left(\|K\|_\infty^{1/2}+\|m_j\|_{\mathcal{H}}\right)\|\widehat{m}_j-m_j\|_{\mathcal{H}}+\|\widehat{m}_j-m_j\|_{\mathcal{H}}^2. \qquad (6.20)
\end{aligned}$$

This allows to obtain $\frac{1}{n_j}\sum_{i=1}^{n_j}\left\|\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}-\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}\right\|_{\mathrm{HS}}$ $=o_P(1)$ and, from (6.19), to conclude that

$$\begin{aligned}
&\left|\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2\right.\\
&\left.-\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2\right|=o_P(1). \qquad (6.21)
\end{aligned}$$

Therefore, from (6.15), (6.17), (6.18) and (6.21), it follows:

$$\widehat{\theta}_{1,j}^2=\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2+o_P(1).$$

Then using the law of large numbers and Slutsky's theorem, we deduce that $\widehat{\theta}_{1,j}^2$ converges in probability to $\mathbb{E}\left(<\left(K(X_1^{(j)},\cdot)-m_j\right)^{\otimes^2},V>_{\mathrm{HS}}^2\right)$ as $n_j\to+\infty$. On the other hand, using (6.17) we get

$$\begin{aligned}
\widehat{\theta}_{2,j} &= \frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},\widehat{V}-V>_{\mathrm{HS}}\\
&\quad+\frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}\\
&= \frac{1}{n_j}\sum_{i=1}^{n_j}<\left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2},V>_{\mathrm{HS}}+o_P(1).
\end{aligned}$$

Further, an use of the Cauchy-Schwartz inequality and (6.20) gives

$$\begin{aligned}&\frac{1}{n_j}\sum_{i=1}^{n_j} < \left(K(X_i^{(j)},\cdot)-\widehat{m}_j\right)^{\otimes^2}, V >_{\mathrm{HS}} \\ &= \frac{1}{n_j}\sum_{i=1}^{n_j} < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}, V >_{\mathrm{HS}} + o_P(1)\end{aligned}$$

and, therefore, $\widehat{\theta}_{2,j} = \frac{1}{n_j}\sum_{i=1}^{n_j} < \left(K(X_i^{(j)},\cdot)-m_j\right)^{\otimes^2}, V >_{\mathrm{HS}} + o_P(1)$. By the law of large numbers and Slutsky's theorem, $\widehat{\theta}_{2,j}$ converges in probability to $\mathbb{E}\left(< \left(K(X_1^{(j)},\cdot)-m_j\right)^{\otimes^2}, V >_{\mathrm{HS}}\right)$ as $n_j \to +\infty$. Finally, $\widehat{\theta}_j^2$ converges in probability, as $n_j \to +\infty$, to

$$\begin{aligned}\theta_j^2 &:= \mathbb{E}\left(< \left(K(X_1^{(j)},\cdot)-m_j\right)^{\otimes^2}, V >_{\mathrm{HS}}^2\right) \\ &\quad -\left(\mathbb{E}\left(< \left(K(X_1^{(j)},\cdot)-m_j\right)^{\otimes^2}, V >_{\mathrm{HS}}\right)\right)^2 \\ &= \mathrm{Var}\left(< \left(K(X_1^{(j)},\cdot)-m_j\right)^{\otimes^2} - V, V >_{\mathrm{HS}}\right)\end{aligned}$$

Since under $\mathscr{H}_0$ we have $\theta_j^2 = \theta^2$, we have shown that $\widehat{\theta}_j^2$ converges in probability, as $n_j \to +\infty$, to θ^2.

References

1. J.G. Aghoukeng Jiofack and G.M. Nkiet, Testing for equality of means of a Hilbert space valued random variable, C. R. Math. Acad. Sci. Paris **347** (2009), 1429-1433.

2. I.A. Ahmad, Modification of some goodness-of-fit statistics to yield asymptotic normal null distributions, Biometrika **80** (1993), 466–472.

3. A.S.K. Balogoun, G.M. Nkiet and C. Ogouyandjou, Asymptotic normality of a generalized maximum mean discrepancy estimator, Statist. Probab. Lett. **169** (2021), 108961.

4. A.S.K. Balogoun, G.M. Nkiet and C. Ogouyandjou, A k-sample test for functional data based on generalized maximum mean discrepancy, Lith. Math. J. **63** (2022), no 3, 489-496.

5. G. Boente, D. Rodriguez and M. Sued, Testing equality between several populations covariance operators, Ann. Inst. Statist. Math. **70** (2018), 917-950.

6. A. Cuevas, M. Febrero and R. Fraiman, An anova test for functional data, Comput. Statist. Data Anal. **47**(2004), 111-122.

7. J. Dauxois, Y. Romain and S. Viguier-Pla, Tensor products and statistics, Linear Algebra Appl. **210** (1994), 59-88.

8. S. Fremdt, J.G. Steinebach, L. Horváth and P. Kokoszka, Testing the equality of covariance operators in functional samples, Scand. J. Statist. **40**(2012), 138-152.

9. K. Fukumizu, F. Bach and A. Gretton, Statistical consistency of kernel canonical correlation analysis, J. Mach. Learn. Res. **8** (2007), 361-383.

10. K. Fukumizu, A. Gretton, X. Sun, B. Schölkopf, Kernel measures of conditional dependence, In J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, editors, Adv. Neural Inf. Process. Syst. **20** (2008), 489–496.

11. T. Górecki and Ł. Smaga, fdANOVA : an R software package for analysis of variance for univariate and multivariate functional data, Comput. Statist. **34** (2019), 571–597.

12. A. Gretton, K.M. Borgwardt, M.J. Rasch, B. Schölkopf and A.J. Smola, A kernel two-sample test, J. Mach. Learn. Res. **13** (2012), 723-773.

13. Z. Harchaoui, F. Bach and E. Moulines, Testing for homogeneity with kernel Fisher discriminant analysis, ArXiv: 0804.10226v1 (2007).

14. N. Makigusa, Two-sample test based on maximum variance discrepancy. Comm. Statist. Theory Methods **53** (2024), 5421-5438.

15. N. Makigusa and K. Naito, Asymptotic normality of a consistent estimator of maximum mean discrepancy in Hilbert space, Statist. Probab. Lett. **156** (2020), 108596.

16. T.K. Manfoumbi Djonguet, A. Mbina Mbina and G.M. Nkiet, Testing independence of functional variables by an Hilbert-Schmidt independence criterion estimator, Statist. Probab. Lett. **207** (2024), 110016.

17. H.Q. Minh, P. Niyogi and Y. Yao, Mercer's theorem, feature maps, and smoothing, In: Proc. of Computational Learning Theory (COLT), 2006.

18. M. Reed and B. Simon, Functional Analysis, Academic Press, New York, 1981.

19. R.J. Serfling, Approximation theorems of mathematical statistics, Wiley, New York, 1980.

20. C. Zhang, H. Peng and J.-T. Zhang, Two samples tests for functional data. Comm. Statist. Theory Methods **39** (2010), 559-578.

Département de Mathématiques, Faculté des Sciences et Techniques, Université Dan Dicko Dankoulodo de Maradi, BP 465 Maradi, Niger
e-mail: armando.balogoun@uddm.edu.ne

Département de Mathématiques et Informatique, Faculté des Sciences, Université des Sciences et Techniques de Masuku, BP 813 Franceville, Gabon
e-mail: guymartial.nkiet@univ-masuku.org