Journal of Artificial Intelligence Research 73 (2022) 1355-1384 Submitted 08/2021; published 04/2022

Marginal Distance and Hilbert-Schmidt Covariances-Based Independence Tests for Multivariate Functional Data

Mirosław Krzyśko MKRZYSKO@AMU.EDU.PL
Interfaculty Institute of Mathematics and Statistics
Calisia University-Kalisz
Nowy Świat 4, 62-800 Kalisz, Poland

Łukasz Smaga (corresponding author) LS@AMU.EDU.PL
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Uniwersytetu Poznańskiego 4, 61-614 Poznań, Poland

Piotr Kokoszka PIOTR.KOKOSZKA@COLOSTATE.EDU
Department of Statistics
Colorado State University
Fort Collins, CO 80523, USA

Abstract

We study the pairwise and mutual independence testing problem for multivariate functional data. Using a basis representation of functional data, we reduce this problem to testing the independence of multivariate data, which may be high-dimensional. For pairwise independence, we apply tests based on distance and Hilbert-Schmidt covariances as well as their marginal versions, which aggregate these covariances for coordinates of random processes. In the case of mutual independence, we study asymmetric and symmetric aggregating measures of pairwise dependence. A theoretical justification of the test procedures is established. In extensive simulation studies and examples based on a real economic data set, we investigate and compare the performance of the tests in terms of size control and power. An important finding is that tests based on distance and Hilbert-Schmidt covariances are usually more powerful than their marginal versions under linear dependence, while the reverse is true under non-linear dependence.

1. Introduction

One of the fundamental problems in statistics is testing of independence between a number of random variables of different types. There are many applications of this problem; for example, independent component analysis (Matteson & Tsay, 2017), graphical models (Gan et al., 2019; Li & McCormick, 2019), variable selection (Li et al., 2012; Shao & Zhang, 2014) and many practical problems (Hua & Ghosh, 2015; Kong et al., 2012).

Pairwise independence concerns the independence of only two variables, while mutual independence applies to two or more variables. Of course, the latter implies the former, but the reverse is generally not true. Naturally, pairwise independence is more often considered in the literature. For univariate data, tests based on Pearson's coefficient of correlation (Pearson, 1895) and its nonparametric counterparts, such as Kendall's tau (Kendall, 1938) and Spearman's rho (Spearman, 1904) coefficients, are the classical tests. A more recent test procedure is that of Bergsma and Dassios (2014), who extended Kendall's tau test. Since

for testing independence of functional data we use independence tests for multivariate data, we focus on this case. In the low-dimensional context there are plentiful methods (see, for example, Gretton et al., 2008; Székely et al., 2007; Taskinen et al., 2003, and the references therein). On the other hand, the high-dimensional scenario is more rarely considered in the literature (see e.g., Pan et al., 2014; Székely & Rizzo, 2013; Yang & Pan, 2015; Zhu et al., 2020). Mutual independence is also less frequently investigated. For this problem, Leung and Drton (2018), Yao et al. (2018) and Jin and Matteson (2018) proposed methods based on aggregating pairwise dependence measures.

In this paper, we extend the independence tests based on distance covariance and Hilbert-Schmidt covariance introduced by Székely et al. (2007) and Gretton et al. (2008) respectively, to multivariate functional data. Székely and Rizzo (2013) extended the distance covariance based test to the high-dimensional setting. These covariances are the dependence metrics that target linear dependence as well as non-linear and non-monotone dependence. They are applicable, in principle, applicable in the high-dimensional context. However, recently Zhu et al. (2020) showed that the tests based on the distance and Hilbert-Schmid covariances can capture only linear dependence in the high-dimensional scenario. Fortunately, they also proposed tests based on aggregation of marginal sample covariances which capture pairwise non-linear dependence. Thus, these tests overcome the drawback of the tests based on both covariances. More details about them are presented in Section 3.

1.1 Multivariate Functional Data

Multivariate data are often repeatedly observed at different time or space points. For simplicity, we will write *time points* or *design time points*, regardless of what they refer to. In classical analysis, the final data are called doubly multivariate data. However, great advances in computational and analytical techniques result in large numbers of design time points, which implies that the performance of methods for such data will decrease. Fortunately, such data can also be treated in a more convenient way in many applications. Namely, they can be seen as realizations of some random processes at given time points. Then a single observation is a vector of functions, curves or surfaces representing random processes. Such data are called functional data, and their analysis is referred to as the functional data analysis (FDA), which is a relatively new branch of statistics.

Functional data appear naturally in applications in fields such as life sciences, chemometrics, environmental science, economics and engineering. Some examples are as follows: temperature or precipitation in a given location over 10 years, share index variation over one hour, knee flexion angle over a complete gait cycle, socio-economic conditions in given countries over a few years (see Section 7). A good review of the main FDA methods – for example, classification, cluster analysis, hypothesis testing, principal component analysis, regression analysis – as well as their applications can be found in the following books: Ferraty and Vieu (2006), Horváth and Kokoszka (2012), Ramsay and Silverman (2002, 2005), and Zhang (2013).

Functional data and their analysis have several desirable properties that distinguish them from their classical counterparts. First of all, the curse of dimensionality is avoided in the functional data framework, since there may be any number of time points at which observations are measured, and this usually has only a minor effect on the results of analysis.

Secondly, the design time points do not have to be equally spaced in a given interval, which means that missing values for functional observations are not a problem in general. Finally, the values observed at different time points for a single subject may be dependent. These properties follow from the fact that in FDA a single observation is treated as a whole function or curve.

1.2 Testing Independence of Functional Data

In this paper, we consider the problem of independence testing for multivariate functional data. To deal effectively with this problem, we use a basis representation of functional data, which can be seen as a dimensionality reduction method of infinite-dimensional functional data; each multivariate functional observation can be represented by a random vector, perhaps of high dimension. To test independence of functional data, we apply independence tests for multivariate data obtained from the basis representation. The FDA methods based on the basis representation are simple, but also powerful (see Aguilera et al., 2021; Górecki et al., 2016, 2020; Horváth & Kokoszka, 2012; Krzyśko & Waszak, 2013; Lin et al., 2021; Ramsay & Silverman, 2005, among many other contributions).

The tests described above were first considered by Górecki et al. (2016, 2020), who used the tests based on distance covariance and Hilbert-Schmidt covariance. Górecki et al. (2020) also showed that functional versions of independence tests perform better than the direct application of multivariate methods to raw functional data, when the dependence is non-linear, while the raw and functional approaches give similar results in the case of linear dependence. We obtained similar results for raw and functional marginal covariances-based tests, which is presented in the Appendix A. Thus, we continue their work and extend their methods and results in several directions. First, they used biased estimators for the covariances, which has some effect on their performance. We extend these tests to those using unbiased estimators for the covariances, which results in better test procedures. Next, we also consider the t-tests based on both covariances, which for pairwise independence testing, have similar properties to permutation tests, but are much less computationally intensive. Nevertheless, the most important aspect of the present work is that we also apply the tests proposed by Zhu et al. (2020) based on aggregation of marginal sample covariances. Why is this so important? The reason is very simple. The drawback of the standard covariance-based tests seems to apply not only to multivariate data, but also to multivariate functional data. Fortunately, the tests based on aggregation of marginal sample covariances are still free of this drawback and can effectively detect non-linear dependence. We show this in intensive simulation studies, and illustrate it using an example with a pillar data set containing economic variables describing European countries in the period 2008-2015. Note that our simulations are much more elaborate than those of Górecki et al. (2016, 2020). The above results are true for pairwise independence, but we also show that most of them still hold for mutual independence, after applying the methods of Jin and Matteson (2018) to the functional data framework. We also establish a theoretical justification of all proposed tests.

To sum up, the extensions of the results of previous papers and the differences between them and the present work are as follows:

- Górecki et al. (2016, 2020): There were considered the permutation tests based on the distance and Hilbert-Schmidt covariances for pairwise independence only. In contrast, we: (1) consider the t-tests which seem to have the same good properties as the permutation tests, while being less time consuming; (2) investigate marginal covariance-based tests, which are better than classical covariances in identifying non-linear dependence; (3) show that the classical covariances may fail in detecting non-linear dependence of functional data; (4) extend the pairwise independence test procedures to mutual independence ones; (5) show the theoretical justification of applying tests for random vectors to functional data; (6) conduct much more extensive simulation studies, which consider simulation data generated for example based on: basis representation of functional data, different stochastic processes (Wiener process, Ornstein-Uhlenbeck process, the Brownian bridge process), dependent functional coordinates.

- Zhu et al. (2020): They considered the classical and marginal versions of the covariances in the independence testing problem for multivariate data. We extend their results from multivariate vector data to multivariate functional data, as follows: First, we reduce the dimension of the functional data using their basis representation which projects the functional observations onto multivariate ones. To such representation of functional data, we apply the covariance-based independence tests. Such a method is generally more powerful than simply applying multivariate methods to raw functional data, which we also show. Moreover, we present how to use their marginal covariances in testing mutual independence, instead of just pairwise independence. We also establish that the properties of covariance-based tests transfer to the functional data framework and mutual independence testing. These all are theoretically and empirically justified.

- Jin and Matteson (2018): They showed how to aggregate the distance covariance to test the mutual independence for multivariate data. We present that their methodology is also applicable for functional data. We also show that one can use the other pairwise independence measures instead of the distance covariance. Moreover, we establish that the mentioned above good or bad properties of the covariance-based tests transfer to the mutual independence testing.

The remainder of the paper is organized as follows: Section 2 describes functional data and the statistical hypotheses in a formal way, and presents the tests based on the basis representation of functional data. In Sections 3 and 5 respectively, the pairwise and mutual independence testing methods are described, while Section 4 contains the theoretical justification of the test procedures. Section 6 contains a description of the numerical experiments conducted and discussion of their results. In Section 7, illustrative real data examples are given, while Section 8 concludes the paper.

2. Functional Data and Statistical Hypotheses

From a theoretical standpoint, we suppose that the multivariate functional data are random processes in a certain Hilbert space. Let $L_2^p(I)$, where $I = [a, b]$, $a, b \in \mathbb{R}$ and $a < b$, be the Hilbert space of p-dimensional vectors of square integrable functions defined on

the interval I. Assume that $\mathbf{X}_1, \ldots, \mathbf{X}_k$ are k random processes belonging to the spaces $L_2^{p_1}(I_1), \ldots, L_2^{p_k}(I_k)$ respectively, where $I_i = [a_i, b_i]$, $a_i, b_i \in \mathbb{R}$ and $a_i < b_i$, $i = 1, \ldots, k$. We would like to test the independence of $\mathbf{X}_1, \ldots, \mathbf{X}_k$, i.e.,

$$H_0 : \mathbf{X}_1, \ldots, \mathbf{X}_k \text{ are mutually independent, } H_1 : \neg H_0. \tag{1}$$

To solve this problem, we equivalently express the null hypothesis in the finite dimensional space. For this purpose, we reduce the dimension of the functional data by using their basis representation. Note that we reduce the dimension of functional data treated as elements of an infinite-dimensional Hilbert space, not the dimensions $p_1, \ldots, p_k$ of multivariate processes $\mathbf{X}_1, \ldots, \mathbf{X}_k$.

For $i = 1, \ldots, k$, let $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip_i})^\top$ and let $\{\phi_{ijl}\}_{l=1}^{\infty}$ be a basis in $L_2^1(I_i)$, $j = 1, \ldots, p_i$. Then each component of process $\mathbf{X}_i$ can be represented as a linear combination of an infinite number of basis functions, i.e.,

$$X_{ij}(t_i) = \sum_{l=1}^{\infty} \alpha_{ijl}\phi_{ijl}(t_i), \tag{2}$$

where the coefficients α_{ijl} are random variables. Since the basis functions are fixed, the coefficients α_{ijl} are responsible for the randomness of the processes $\mathbf{X}_i$. Thus, the independence of the random processes $\mathbf{X}_1, \ldots, \mathbf{X}_k$ is equivalent to the independence of

$$((\alpha_{11l})_{l=1}^{\infty}, \ldots, (\alpha_{1p_1l})_{l=1}^{\infty}), \ldots, ((\alpha_{k1l})_{l=1}^{\infty}, \ldots, (\alpha_{kp_kl})_{l=1}^{\infty}).$$

This equivalence is independent of the basis since it is fixed. Similar arguments have been used in other statistical problems for functional data as, for example, the analysis of variance considered in Lin et al. (2021).

However, representation (2) cannot be applied in practice. For any practical analysis, it must be truncated to a finite sum:

$$X_{ij}(t_i) \approx \sum_{l=1}^{B_{ij}} \alpha_{ijl}\phi_{ijl}(t_i), \quad i = 1, \ldots, k,\ j = 1, \ldots, p_i,\ t_i \in I_i.$$

The quality of the above representation depends of the choice of the basis functions ϕ_{ijl} and their number B_{ij}. For specific functions, we want to select basis functions in such a way that a relatively small number of them is needed to achieve a good approximation. This is often possible for functions with some specific properties, like smoothness or periodicity, but in general, we may need large B_{ij}. This motivates the theoretical justification in Section 4. For estimation of the coefficients α_{ijl}, the least squares method and the roughness penalty approach are usually applied. The choice of basis may affect the test procedure proposed below due to the quality of approximation, which is a combined effect of the choice of the basis and the number of coefficients, B_{ij}, as reported in the simulation studies (see the last paragraph of Section 6.3).

We use throughout the following matrix notation, which is the basis representation of the process $\mathbf{X}_i$:

$$\mathbf{X}_i(t_i) \approx \mathbf{\Phi}_i(t_i)\boldsymbol{\alpha}_i, \tag{3}$$

where

$$\begin{aligned}\boldsymbol{\Phi}_i(t_i) &= \operatorname{diag}(\boldsymbol{\phi}_{i1}^\top(t_i), \ldots, \boldsymbol{\phi}_{ip_i}^\top(t_i)),\\ \boldsymbol{\phi}_{ij}(t_i) &= (\phi_{ij1}(t_i), \ldots, \phi_{ijB_{ij}}(t_i))^\top,\\ \boldsymbol{\alpha}_i &= (\alpha_{i11}, \ldots, \alpha_{i1B_{i1}}, \ldots, \alpha_{ip_i1}, \ldots, \alpha_{ip_iB_{ip_i}})^\top \in \mathbb{R}^{B_i},\end{aligned}$$

$B_i = B_{i1} + \cdots + B_{ip_i}$, $t_i \in I_i$, $i = 1, \ldots, k$ and $j = 1, \ldots, p_i$. Of course, the matrices $\boldsymbol{\Phi}_i(t_i)$ are fixed, while the vectors $\boldsymbol{\alpha}_i$ are random.

Taking account of the basis representation (3), we can conclude that the hypotheses in (1) corresponding to the independence and dependence of the random processes $\mathbf{X}_1, \ldots, \mathbf{X}_k$ can be practically verified by testing the following hypotheses:

$$H_0^v : \boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k \text{ are mutually independent, } H_1^v : \neg H_0^v \tag{4}$$

corresponding to the independence and dependence respectively of the random vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$. The testing problem (1) is not equivalent to the testing problem (4). One can only say that the rejection of H_0^v in (4) implies the rejection of H_0 in (1). This is common to all functional tests that use expansions (see, for example, Lin et al., 2021). Nevertheless, to verify (1), we can use methods for testing (4). This task should be approached carefully, since the dimensions $B_1, \ldots, B_k$ of the vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ can be quite large, often larger than the number of observations in a sample. For this reason, in the following sections, we consider methods which can be used in a broad range of cases, in particular in cases of high dimensions.

Of course, for estimation and inference, we need to have a sample. Assume that $\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in}$ are independent realizations of random processes $\mathbf{X}_i$ for $i = 1, \ldots, k$ and $n \in \mathbb{N}$, $n > 3$. Let $\mathbf{X}_{ij}(t_i) \approx \boldsymbol{\Phi}_i(t_i)\boldsymbol{\alpha}_{ij}$, $i = 1, \ldots, k$, $j = 1, \ldots, n$, $t_i \in I_i$ be the basis representations (3) of the observations and

$$\boldsymbol{\alpha}_{ij} = (\alpha_{i11j}, \ldots, \alpha_{i1B_{i1}j}, \ldots, \alpha_{ip_i1j}, \ldots, \alpha_{ip_iB_{ip_i}j})^\top \in \mathbb{R}^{B_i}.$$

Since our considerations are based throughout on the vectors of coefficients $\boldsymbol{\alpha}_{ij}$, we use the following notation:

- $\mathbf{A}_i = (\boldsymbol{\alpha}_{i1}, \ldots, \boldsymbol{\alpha}_{in})^\top$ – the $n \times B_i$ random sample matrix,
- $\mathcal{A}_{il} = (\alpha_{il1}^*, \ldots, \alpha_{iln}^*)^\top$ – the $n \times 1$ random component-wise sample, where $l = 1, \ldots, B_i$ and $\boldsymbol{\alpha}_{ij} = (\alpha_{i1j}^*, \ldots, \alpha_{iB_ij}^*)^\top$.

3. Testing Pairwise Independence

In this section, we consider eight test procedures for pairwise independence with $k = 2$ random processes $\mathbf{X}_1$ and $\mathbf{X}_2$. The tests are based on the distance and Hilbert-Schmidt covariances, as well as their marginal versions.

3.1 Distance Covariance and Its Marginal Version

To test the hypotheses given in (4), the distance covariance introduced by Székely et al. (2007) may be used. This is a measure of dependence between two random vectors

of arbitrary dimensions, defined as

$$dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \int_{\mathbb{R}^{B_1+B_2}} \frac{\|\varphi_{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2}(\mathbf{t},\mathbf{s}) - \varphi_{\boldsymbol{\alpha}_1}(\mathbf{t})\varphi_{\boldsymbol{\alpha}_2}(\mathbf{s})\|^2}{c_{B_1}c_{B_2}\|\mathbf{t}\|^{1+B_1}\|\mathbf{s}\|^{1+B_2}} d\mathbf{t}d\mathbf{s},$$

where $\varphi_{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2}$ is the joint characteristic function of $(\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top)^\top$, $\varphi_{\boldsymbol{\alpha}_1}$ and $\varphi_{\boldsymbol{\alpha}_2}$ are the characteristic functions of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ respectively, $\|\cdot\|$ is the complex Euclidean norm and $c_b = \pi^{(1+b)/2}/\Gamma((1+b)/2)$. The above definition supports the interpretation of distance covariance, while for estimation, the following alternative form is more useful (Székely & Rizzo, 2009). Let $(\boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2')$ and $(\boldsymbol{\alpha}_1'', \boldsymbol{\alpha}_2'')$ be independent copies of $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$. Then

$$\begin{aligned} dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) =& E\|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_1'\|\|\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_2'\| \\ &+ E\|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_1'\|E\|\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_2'\| \\ &- 2E\|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_1'\|\|\boldsymbol{\alpha}_2 - \boldsymbol{\alpha}_2''\|. \end{aligned}$$

Since $dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = 0$ is equivalent to independence of $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, a permutation test with test statistic being an estimator for $dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ is proposed for verifying the dependence between $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, including non-linear and non-monotonic dependences. (For an explanation of the permutation method, see Section 5.) This test for (1) was considered by Górecki et al. (2016) with a biased estimator for $dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$. Here we consider an unbiased estimator, described below, which results in a better test procedure (data not shown). The unbiased estimator for $dCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ is called the $\mathcal{U}$-centering-based unbiased sample distance covariance and was defined by Székely and Rizzo (2014) as follows:

$$dCov_n^2(\mathbf{A}_1, \mathbf{A}_2) = (\widetilde{\mathbf{A}} \cdot \widetilde{\mathbf{B}}) = \frac{1}{n(n-3)} \sum_{p \neq q} \widetilde{a}_{pq}\widetilde{b}_{pq},$$

where $\widetilde{\mathbf{A}} = (\widetilde{a}_{pq})_{p,q=1}^n$ and $\widetilde{\mathbf{B}} = (\widetilde{b}_{pq})_{p,q=1}^n$ are the $\mathcal{U}$-centered versions of $\mathbf{A} = (a_{pq})_{p,q=1}^n$ and $\mathbf{B} = (b_{pq})_{p,q=1}^n$ respectively, i.e.,

$$\widetilde{a}_{pq} = a_{pq} - \frac{1}{n-2}\sum_{s=1}^n a_{ps} - \frac{1}{n-2}\sum_{r=1}^n a_{rq} + \frac{1}{(n-1)(n-2)}\sum_{r=1}^n\sum_{s=1}^n a_{rs}$$

for $p \neq q$ and 0 otherwise, and $a_{pq} = \|\boldsymbol{\alpha}_{1p} - \boldsymbol{\alpha}_{1q}\|$, $b_{pq} = \|\boldsymbol{\alpha}_{2p} - \boldsymbol{\alpha}_{2q}\|$.

To test (4), the distance correlation-based t-test proposed by Székely and Rizzo (2013) can also be used. They showed that, under suitable assumptions, the following test statistic

$$T_{dCov}(\mathbf{A}_1, \mathbf{A}_2) = \frac{\sqrt{v-1}dCor_n^2(\mathbf{A}_1, \mathbf{A}_2)}{\sqrt{1 - (dCor_n^2(\mathbf{A}_1, \mathbf{A}_2))^2}}$$

has t_{v-1}-distribution with $v-1$ degrees of freedom asymptotically, where $v = n(n-3)/2$ and

$$dCor_n^2(\mathbf{A}_1, \mathbf{A}_2) = \frac{dCov_n^2(\mathbf{A}_1, \mathbf{A}_2)}{\sqrt{dCov_n^2(\mathbf{A}_1, \mathbf{A}_1)dCov_n^2(\mathbf{A}_2, \mathbf{A}_2)}}$$

is the sample distance correlation. Then the corresponding critical region and p-value are $\{T_{dCov}(\mathbf{A}_1, \mathbf{A}_2) \geq t_{v-1,\alpha}\}$ and $P(t_{v-1} \geq T_{dCov}(\mathbf{A}_1, \mathbf{A}_2))$ respectively, where $t_{v-1,\alpha}$ is the upper 100α percentile of the t_{v-1}-distribution.

The tests based on the distance covariance should be applicable to arbitrary positive integers B_1 and B_2, including high-dimensional scenarios. However, recently Zhu et al. (2020) showed that this may not be the case for the sample distance covariance-based tests. Namely, they proved that the sample distance covariance between two random vectors can be approximated by the sum of squared componentwise sample cross-covariances, which indicates that the test based on this estimator can only capture linear dependence in high dimension. Moreover, they also showed that the distance correlation-based t-test has trivial limiting power when the two random vectors are non-linearly dependent but uncorrelated componentwise. To overcome such problems, Zhu et al. (2020) proposed tests based on the following aggregation of marginal sample distance covariances:

$$mdCov_n^2(\mathbf{A}_1, \mathbf{A}_2) = \sqrt{\binom{n}{2}} \sum_{p=1}^{B_1} \sum_{q=1}^{B_2} dCov_n^2(\mathcal{A}_{1p}, \mathcal{A}_{2q}).$$

The first test is the permutation test with $mdCov_n^2(\mathbf{A}_1, \mathbf{A}_2)$ as the test statistic. The second test procedure is the t-test based on

$$T_{mdCov}(\mathbf{A}_1, \mathbf{A}_2) = \frac{\sqrt{v-1} mdCor_n^2(\mathbf{A}_1, \mathbf{A}_2)}{\sqrt{1-(mdCor_n^2(\mathbf{A}_1, \mathbf{A}_2))^2}}$$

having the same asymptotic distribution as $T_{dCov}(\mathbf{A}_1, \mathbf{A}_2)$, where $mdCor_n^2(\mathbf{A}_1, \mathbf{A}_2)$ is defined analogously to $dCor_n^2(\mathbf{A}_1, \mathbf{A}_2)$. In contrast to the sample covariance-based tests, which test independence by treating all components of a vector jointly as a whole, the marginal sample distance covariance-based tests capture pairwise low-dimensional non-linear dependence. This implies that the latter tests can detect non-linear dependence in high dimensions, while the former may not have this property. This was demonstrated in the simulation studies of Zhu et al. (2020) for vector data, and in those of Section 6 below for functional data. Similar findings apply to Hilbert-Schmidt covariance, which we discuss in the next section.

3.2 Hilbert-Schmidt Covariance and Its Marginal Version

Now we consider the Hilbert-Schmidt covariance proposed by Gretton et al. (2005, 2008), which is a generalization of the distance covariance. Hilbert-Schmidt covariance can be defined similarly to the distance covariance with kernel values instead of Euclidean distance, i.e.,

$$\begin{aligned} hCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) =& EK(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1') L(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2') \\ &+ EK(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1') EL(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2') \\ &- 2EK(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_1') L(\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2''), \end{aligned}$$

where K and L are selected kernels. When these kernels are universal (e.g., Gaussian and Laplacian kernels; see Section 6) and defined on compact domains, $hCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = 0$ if and only if $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ are independent (Gretton et al., 2005, Theorem 4). Thus, the permutation test based on a biased estimator for $hCov^2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ was considered by Górecki

et al. (2020) for testing (1). The natural unbiased sample Hilbert-Schmidt covariance is defined as follows:

$$hCov_n^2(\mathbf{A}_1, \mathbf{A}_2) = (\widetilde{\mathbf{K}} \cdot \widetilde{\mathbf{L}}) = \frac{1}{n(n-3)} \sum_{p \neq q} \widetilde{k}_{pq} \widetilde{l}_{pq},$$

where $\widetilde{\mathbf{K}} = (\widetilde{k}_{pq})_{p,q=1}^n$ and $\widetilde{\mathbf{L}} = (\widetilde{l}_{pq})_{p,q=1}^n$ are the $\mathcal{U}$-centered versions of $\mathbf{K} = (k_{pq})_{p,q=1}^n$ and $\mathbf{L} = (l_{pq})_{p,q=1}^n$ respectively, where $k_{pq} = l_{pq} = 0$ for $p = q$ and otherwise $k_{pq} = K(\boldsymbol{\alpha}_{1p}, \boldsymbol{\alpha}_{1q})$, $l_{pq} = L(\boldsymbol{\alpha}_{2p}, \boldsymbol{\alpha}_{2q})$.

Zhu et al. (2020) extended the $T_{dCov}(\mathbf{A}_1, \mathbf{A}_2)$ test to the t-test based on the Hilbert-Schmidt covariance with test statistic

$$T_{hCov}(\mathbf{A}_1, \mathbf{A}_2) = \frac{\sqrt{v-1} hCor_n^2(\mathbf{A}_1, \mathbf{A}_2)}{\sqrt{1 - (hCor_n^2(\mathbf{A}_1, \mathbf{A}_2))^2}}.$$

Moreover, they established that the $hCov_n^2(\mathbf{A}_1, \mathbf{A}_2)$ and $T_{hCov}(\mathbf{A}_1, \mathbf{A}_2)$ tests have the same drawbacks as the tests based on distance covariance (see Section 3.1). For this reason, Zhu et al. (2020) also proposed a permutation test based on the marginally aggregated sample Hilbert-Schmidt covariances

$$mhCov_n^2(\mathbf{A}_1, \mathbf{A}_2) = \sqrt{\binom{n}{2}} \sum_{p=1}^{B_1} \sum_{q=1}^{B_2} hCov_n^2(\mathcal{A}_{1p}, \mathcal{A}_{2q})$$

and the t-test based on

$$T_{mhCov}(\mathbf{A}_1, \mathbf{A}_2) = \frac{\sqrt{v-1} mhCor_n^2(\mathbf{A}_1, \mathbf{A}_2)}{\sqrt{1 - (mhCor_n^2(\mathbf{A}_1, \mathbf{A}_2))^2}}.$$

To sum up, for testing (1) with $k = 2$, we have eight tests:

- four permutation tests based on $dCov_n^2$, $mdCov_n^2$, $hCov_n^2$, $mhCov_n^2$,
- four t-tests based on T_{dCov}, T_{mdCov}, T_{hCov}, T_{mhCov}.

For simplicity, in the following, we will refer to the tests based on standard distance and Hilbert-Schmidt covariances as the joint distance and Hilbert-Schmidt covariances-based tests, or even more simply as joint covariances-based tests if applicable. Similarly, the tests based on marginally aggregated covariances will be refered to as marginal covariances-based tests.

Note that in this section, we applied these multivariate tests to vectors of coefficients of the basis representation. However, it is possible to define appropriate functional objects, for example characteristic functions or kernels, to obtain general functional distance covariance, functional Hilbert-Schmidt covariance, etc. For the former, this was done by Górecki et al. (2016, 2020), but for practical application, they finally also used the basis representation, which greatly simplifies work with general objects. Here we extend their practical solution. Moreover, in the next sections, we present a theoretical justification of the test procedures, and we consider the case $k \geq 2$, which is a further extension of the results of Górecki et al. (2016, 2020) and of Zhu et al. (2020), who investigated the case of two processes only.

4. Theoretical Justification of the Test Procedures

All procedures considered in this paper are theoretically justified under weak assumptions by the application of the general results established by Zhu et al. (2020), chiefly by the results in their Appendix A.2 dedicated to studentized test statistics in the High Dimension Medium Sample Size, HDMSS, framework and the unified approach they develop. The differences will be highlighted in the following, after suitable notation has been introduced. Basically, the setting we consider in this paper is more complex because we consider several samples and the objects in each sample are expansion coefficients. In the most general justification, we would thus need to work with four-dimensional tensors, and the idea might be lost in the manipulations of coefficients with four integer subscripts. To focus on the essence of the justification, we therefore first consider the case of $k = 2$ and change the notation slightly to make the exposition of the theory easier to follow. In Remark 1, we comment on the assumptions needed in the general case.

We thus consider two populations of functions

$$(X_1, \ldots, X_p)^\top, \quad (Y_1, \ldots, Y_q)^\top \tag{5}$$

that satisfy the following assumption

Assumption 1 *Each component X_i, $i = 1, \ldots, p$, is a square integrable (i.e., $E\|X_i\|^2 < \infty$) random function in $L_2(I_X)$ and each Y_j, $j = 1, \ldots, q$, is a random function in $L_2(I_Y)$ satisfying analogous assumptions.*

The common domain I_X could be replaced by different domain I_{X_i}, $i = 1, \ldots, p$, because we ultimately work with expansion coefficients. The same comment applies to the functions Y_i.

Let $\{\phi_l\}_{l=1}^\infty$ be a complete orthonormal system in $L_2(I_X)$, and $\{\psi_l\}_{l=1}^\infty$ such a system in $L_2(I_Y)$. Under Assumption 1, the following infinite expansions exist:

$$X_i = \sum_{l=1}^{\infty} \alpha_{il}\phi_l, i = 1, \ldots, p, \quad Y_j = \sum_{l=1}^{\infty} \beta_{il}\psi_l, j = 1, \ldots, q.$$

Suppose we observe iid realizations

$$(X_{t1}, \ldots, X_{tp})^\top, \quad (Y_{t1}, \ldots, Y_{tq})^\top, \quad t = 1, \ldots, n.$$

The tests are based on approximations

$$X_{ti} \approx \sum_{l=1}^{B_{X,i}} \alpha_{til}\phi_l, \quad Y_{tj} \approx \sum_{l=1}^{B_{Y,j}} \beta_{tjl}\psi_l, \quad t = 1, \ldots, n.$$

As we already noticed in Section 2, if the truncation levels $B_{X,i}$ and $B_{Y,j}$ are treated as fixed numbers, the testing problem (1) is not equivalent to the testing problem (4). However, asymptotic equivalence can be ensured if we let the truncation levels $B_{X,i}$ and $B_{Y,j}$ increase to infinity with the sample size n. To formulate a suitable assumption, we need to introduce

some notation. The functional object $(X_1, \ldots, X_p)^\top$ is approximately represented by the vector

$$\mathbf{A} = (\alpha_{11}, \ldots, \alpha_{1B_{X,1}}, \ldots, \alpha_{p1}, \ldots \alpha_{pB_{X,p}})^\top \tag{6}$$

whose length is

$$P = \sum_{i=1}^{p} B_{X,i}.$$

Similarly, the object $(Y_1, \ldots, Y_q)^\top$ is represented by the vector

$$\mathbf{B} = (\beta_{11}, \ldots, \beta_{1B_{Y,1}}, \ldots, \beta_{q1}, \ldots \beta_{qB_{Y,q}})^\top$$

whose length is

$$Q = \sum_{j=1}^{q} B_{Y,j}.$$

Assumption 2 *The sample size n tends to infinity and, as $n \to \infty$,*

$$\min(B_{X,1}, \ldots, B_{X,p}) \to \infty \quad \text{and} \quad \min(B_{Y,1}, \ldots, B_{Y,q}) \to \infty.$$

Assumption 2 clearly implies the HDMSS condition of Zhu et al. (2020), i.e., $n \wedge P \wedge Q \to \infty$.

Consider the $n \times P$ matrix

$$\mathbf{A}_n = \begin{pmatrix} \alpha_{111} & \cdots & \alpha_{11B_{X,1}} & \cdots & \alpha_{1p1} & \cdots & \alpha_{1pB_{X,p}} \\ \alpha_{211} & \cdots & \alpha_{21B_{X,1}} & \cdots & \alpha_{2p1} & \cdots & \alpha_{2pB_{X,p}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_{n11} & \cdots & \alpha_{n1B_{X,1}} & \cdots & \alpha_{np1} & \cdots & \alpha_{npB_{X,p}} \end{pmatrix},$$

and an analogously defined $n \times Q$ matrix $\mathbf{B}_n$. The difference between the functional setting of this paper and the setting of Zhu et al. (2020) is that we use the matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ of coefficients rather than the matrices of observations.

We now formulate our test procedures within the unified framework of Zhu et al. (2020) that covers all tests statistics we consider. Suppose $K(\cdot, \cdot)$ and $L(\cdot, \cdot)$ are *bounded* bivariate kernels. The Gaussian and Laplacian kernels studied in the next section are clearly bounded by 1. Since the kernels are bounded, the following random variables are well-defined. For $i = 1, \ldots, p$ and $l(i) = 1, \ldots, B_{X,i}$, set

$$\begin{aligned} K_{st}(i, l(i)) = {} & K(\alpha_{s,i,l(i)}, \alpha_{t,i,l(i)}) \\ & - E[K(\alpha_{s,i,l(i)}, \alpha_{t,i,l(i)}) | \alpha_{s,i,l(i)}] \\ & - E[K(\alpha_{s,i,l(i)}, \alpha_{t,i,l(i)}) | \alpha_{t,i,l(i)}] \\ & + E[K(\alpha_{s,i,l(i)}, \alpha_{t,i,l(i)})], \quad s, t = 1, \ldots, n, \end{aligned}$$

and define $L_{st}(j, l(j))$, $j = 1, \ldots, q$, $l(j) = 1, \ldots, B_{Y,j}$ analogously.

Next we construct P matrices

$$\mathbf{K}(i, l(i)) = (K_{st}(i, l(i)))_{s,t=1}^{n}, \ i = 1, \ldots, p, \ l(i) = 1, \ldots, B_{X,i}$$

and Q matrices

$$\mathbf{L}(j,l(j)) = (L_{st}(j,l(j)))_{s,t=1}^{n}\,,\ j=1,\ldots,q,\ l(j)=1,\ldots,B_{Y,j}.$$

Suppose $\mathbf{G}=(g_{st})_{s,t=1}^{n}$ and $\mathbf{H}=(h_{st})_{s,t=1}^{n}$ are $n\times n$ matrices. Set

$$\tilde{g}_{st} = g_{st} - \frac{1}{n-2}\sum_{v=1}^{n} g_{sv} - \frac{1}{n-2}\sum_{u=1}^{n} g_{ut} + \frac{1}{(n-1)(n-2)}\sum_{u,v=1}^{n} g_{uv},\quad \text{if } s\neq t,$$

and $\tilde{g}_{st}=0$ if $s=t$. Define the entries $\tilde{h}_{st}$ analogously, and consider

$$(\widetilde{\mathbf{G}}\cdot\widetilde{\mathbf{H}}) := \frac{1}{n(n-3)}\sum_{s,t=1}^{n}\tilde{g}_{st}\tilde{h}_{st}.$$

Now we can define the *unified covariance distance*

$$\mathrm{uCov}_n^2(\mathbf{A}_n,\mathbf{B}_n) = \frac{1}{\sqrt{PQ}}\sum_{i=1}^{p}\sum_{l(i)=1}^{B_{X,i}}\sum_{j=1}^{q}\sum_{l(j)=1}^{B_{Y,j}}(\widetilde{\mathbf{K}}(i,l(i))\cdot\widetilde{\mathbf{L}}(j,l(j))).$$

To ensure the asymptotic validity of the test procedures, we must reformulate Assumption D5 in Zhu et al. (2020) within our functional context. To do it, we need to introduce more notation. Consider the following quantities

$$U(\mathbf{A}_s,\mathbf{A}_t) := \frac{1}{\sqrt{P}}\sum_{i=1}^{p}\sum_{l(i)=1}^{B_{X,i}} K_{st}(i,l(i)),\quad V(\mathbf{B}_s,\mathbf{B}_t) := \frac{1}{\sqrt{Q}}\sum_{j=1}^{q}\sum_{l(j)=1}^{B_{Y,j}} L_{st}(j,l(j)),$$

where $\mathbf{A}_s$ and $\mathbf{A}_t$ are the sth and tth rows of the matrix $\mathbf{A}_n$ corresponding to the sth and tth observations, respectively, in the sample for $(X_1,\ldots,X_p)^\top$; $\mathbf{B}_s$ are $\mathbf{B}_t$ are defined analoguously.

Assumption 3 *Denoting with "prime", $'$, independent copies of the vector* $\mathbf{A}$ *given by* (6)*, we assume that*

$$\frac{1}{\sqrt{n}}\frac{E[U^4(\mathbf{A},\mathbf{A}')]}{(E[U^2(\mathbf{A},\mathbf{A}')])^2}\to 0 \tag{7}$$

and

$$\frac{E[U(\mathbf{A},\mathbf{A}')U(\mathbf{A}',\mathbf{A}'')U(\mathbf{A}'',\mathbf{A}''')U(\mathbf{A}''',\mathbf{A})]}{(E[U^2(\mathbf{A},\mathbf{A}')])^2}\to 0. \tag{8}$$

We impose analogous conditions on $V(\cdot,\cdot)$*.*

Just as in the case of directly observable vectors considered in Zhu et al. (2020), Assumption 3 does not have a clear, intuitive interpretation. It is needed to apply a martingale central limit theorem. Conditions (7) and (8) are abstract, technical assumptions needed to control the growth of the fourth moments of the partial sums in the definition of $U(A_s,A_s)$ relative to the second moments. To a rough approximation, one can say that the kurtosis of the partial sums must be controlled to ensure a normal limit.

To lighten the notation, set

$$\mathcal{U}(\mathbf{A}_n, \mathbf{B}_n) = \mathrm{uCov}_n^2(\mathbf{A}_n, \mathbf{B}_n), \quad \mathcal{C}(\mathbf{A}_n, \mathbf{B}_n) = \frac{\mathcal{U}(\mathbf{A}_n, \mathbf{B}_n)}{\sqrt{\mathcal{U}(\mathbf{A}_n, \mathbf{A}_n)\mathcal{U}(\mathbf{B}_n, \mathbf{B}_n)}}$$

and define the test statistic

$$T_n = \kappa_n \frac{\mathcal{C}(\mathbf{A}_n, \mathbf{B}_n)}{\sqrt{1 - \mathcal{C}^2(\mathbf{A}_n, \mathbf{B}_n)}}, \qquad \kappa_n = \sqrt{\frac{n(n-3)}{2} - 1}. \tag{9}$$

As in Section A.2 of Zhu et al. (2020), the following theorem provides an asymptotic justification for all tests considered in this paper. Its proof follows by direct verification that the assumptions we formulated imply the assumptions of Proposition A.2.1 established by Zhu et al. (2020). (Recall that we assume throughput that the realizations indexed by $t = 1, \ldots, n$ are iid.)

Theorem 1 *If Assumptions 1, 2 and 3 hold, and the functional objects $(X_1, \ldots, X_p)^\top$ and $(Y_1, \ldots, Y_q)^\top$ are independent, then $T_n \xrightarrow{d} N(0, 1)$.*

Remark 1 *Observe that in Assumptions 1, 2 and 3 conditions are imposed on each of the two groups separately. Under independence there is no connection between the two groups. In the case of k samples, not necessarily $k = 2$, Assumptions 1, 2 and 3 must be replaced by analogous assumptions specifying conditions on the functions, and objects derived from them, in each group separately.*

5. Testing Mutual Independence

In this section, we extend the results of Section 3 for pairwise processes ($k = 2$) to the case of $k \geq 2$ processes, i.e., mutual independence. For this purpose, we apply the asymmetric and symmetric measures of mutual dependence given by Jin and Matteson (2018), which capture mutual dependence by aggregating pairwise dependence. They also considered other methods, but these two had the best finite sample performance, hence we omit the others.

Let us introduce the following notation. The subset of the vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ to the right of $\boldsymbol{\alpha}_i$ is defined as

$$\boldsymbol{\alpha}_{i^+} = \left(\boldsymbol{\alpha}_{i+1}^\top, \ldots, \boldsymbol{\alpha}_k^\top\right)^\top \in \mathbb{R}^{B_{i+1}+\cdots+B_k}$$

for $i = 1, \ldots, k-1$. On the other hand, the subset of the vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ that excludes $\boldsymbol{\alpha}_i$ only is defined as:

$$\boldsymbol{\alpha}_{-i} = \left(\boldsymbol{\alpha}_1^\top, \ldots, \boldsymbol{\alpha}_{i-1}^\top, \boldsymbol{\alpha}_{i+1}^\top, \ldots, \boldsymbol{\alpha}_k^\top\right)^\top \in \mathbb{R}^{B_1+\cdots+B_{i-1}+B_{i+1}+\cdots+B_k}$$

for $i = 1, \ldots, k$. Using this notation, Jin and Matteson (2018) proposed the following asymmetric and symmetric measures of mutual dependence of vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$:

$$R(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k) = \frac{1}{k-1} \sum_{i=1}^{k-1} V^2(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_{i^+}),$$

$$S(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k) = \frac{1}{k} \sum_{i=1}^{k} V^2(\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_{-i}),$$

where V^2 is a dependence measure for two vectors. They also proved that if $V^2(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0$ is equivalent to independence of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, then (under mild conditions) $R(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k) = 0$ or $S(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k) = 0$ if and only if H_0^v in (4) holds. Moreover, under H_1^v in (4), $R(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)$ and $S(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)$ are strictly positive. Thus, we can reject the null hypothesis in (4) for large values of the following estimators for $R(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)$ and $S(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)$:

$$R_n(\mathbf{A}_1, \ldots, \mathbf{A}_k) = \frac{1}{k-1} \sum_{i=1}^{k-1} V_n^2(\mathbf{A}_i, \mathbf{A}_{i^+}),$$

$$S_n(\mathbf{A}_1, \ldots, \mathbf{A}_k) = \frac{1}{k} \sum_{i=1}^{k} V_n^2(\mathbf{A}_i, \mathbf{A}_{-i}),$$

where V_n^2 is an estimator for V^2, $\mathbf{A}_{i^+} = (\mathbf{A}_{i+1}, \ldots, \mathbf{A}_k)$ is the sample corresponding to $\boldsymbol{\alpha}_{i^+}$, and $\mathbf{A}_{-i} = (\mathbf{A}_1, \ldots, \mathbf{A}_{i-1}, \mathbf{A}_{i+1}, \ldots, \mathbf{A}_k)$ corresponds to $\boldsymbol{\alpha}_{-i}$. Since the asymptotic null distributions of R_n and S_n are complicated, the permutation method is used to approximate the null distributions of these test statistics. In this method, the permuted sample of the pooled sample

$$\boldsymbol{\alpha}_{11}, \ldots, \boldsymbol{\alpha}_{1n}, \boldsymbol{\alpha}_{21}, \ldots, \boldsymbol{\alpha}_{2n}, \ldots, \boldsymbol{\alpha}_{k1}, \ldots, \boldsymbol{\alpha}_{kn}$$

is as follows

$$\boldsymbol{\alpha}_{11}, \ldots, \boldsymbol{\alpha}_{1n}, \boldsymbol{\alpha}_{2\pi_1(1)}, \ldots, \boldsymbol{\alpha}_{2\pi_1(n)}, \ldots, \boldsymbol{\alpha}_{k\pi_{k-1}(1)}, \ldots, \boldsymbol{\alpha}_{k\pi_{k-1}(n)},$$

where the permutations $\pi_1, \ldots, \pi_{k-1}$ are uniformly chosen from the symmetric group $\mathcal{S}_n$, i.e., the set of all $n!$ permutations of $(1, \ldots, n)$. Then the p-value for the permutation test is equal to the proportion of times that the test statistic based on the permuted samples is greater than that based on the original sample. In general, it is not possible to use all permuted samples in practical implementation, and so a large number of them is applied.

As V_n^2, Jin and Matteson (2018) used the sample distance covariance $dCov_n^2$. However, the remaining test statistics of Section 3 can also be considered. Of course, $dCov_n^2$, T_{dCov}, $hCov_n^2$ and T_{hCov} seem to be preferred as they are consistent with the theoretical results of Jin and Matteson (2018), i.e., the quantities, which they estimate vanish if and only if the random vectors are independent. Nevertheless, as we will show in simulation studies in the next section, the R_n and S_n tests based on these four test statistics seem to have the same drawbacks as the joint covariances-based tests for two processes described in Section 3. Fortunately, we will also establish that the R_n and S_n tests using $mdCov_n^2$, T_{mdCov}, $mhCov_n^2$ and T_{mhCov} overcome these drawbacks, similarly as in the case $k = 2$.

6. Simulation Studies

In this section, we study the finite sample performance of the tests considered in Sections 3 and 5 in terms of size control and power. The empirical size of a test should be close to the significance level, which we set at $\alpha = 5\%$. Then the test maintain the type I error level accurately. On the other hand, the larger the power of the test, the better it is. The empirical sizes and powers of the tests were estimated as the proportions of rejections of the null hypothesis on the basis of 500 simulation replications, when the data were generated under the null and alternative hypothesis respectively. All numerical experiments in this paper were performed in the R program (R Core Team, 2020). The code is available from the authors upon request.

6.1 Test Procedures

For pairwise independence ($k = 2$), we considered 12 tests:

- four tests based on distance covariance: $dCov_n^2$, $mdCov_n^2$, T_{dCov}, T_{mdCov},
- eight tests based on Hilbert-Schmidt covariance: $hCov_n^2$, $mhCov_n^2$, T_{hCov}, T_{mhCov} with
  - Gaussian kernel $K(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/(2\gamma^2))$,
  - Laplacian kernel $K(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|/\gamma)$.

For mutual independence with $k = 3$, we had 24 tests, which were the above tests combined with the R and S methods described in Section 5. We used the same kernel (Gaussian or Laplacian) for all samples, but the bandwidth parameter γ was taken separately for each sample as the median distance between points in a sample, i.e., $\gamma_i = median\{\|\boldsymbol{\alpha}_{ip} - \boldsymbol{\alpha}_{iq}\| : p, q = 1, \ldots, n; p \neq q\}$ (Gretton et al., 2009). Following Zhu et al. (2020), we set 200 permuted samples to estimate the p-values of permutation tests based on the test statistics $dCov_n^2$, $mdCov_n^2$, $hCov_n^2$, $mhCov_n^2$, R_n and S_n. For $k = 2$, the T_{dCov}, T_{mdCov}, T_{hCov} and T_{mhCov} tests were implemented using the t-distribution approach.

6.2 Simulation Experiments

We generated functional data in the following three models. In Model 1 below, we used the B-spline basis as the basis representation of the functional data, since the Fourier basis was applied to generate simulation data. On the other hand, in Models 2 and 3, we considered both Fourier and B-spline bases. For simplicity, all numbers of basis functions B_{ij} were set equal to five. The coefficients of the basis representation were estimated by the least squares method.

Model 1 We considered pairwise independence ($k = 2$). We generated $n = 20$ observations with dimensions $p_* = p_1 = p_2 \in \{3, 6\}$ using the basis representation (3). Namely, the functional data were generated by their values in an equally spaced grid of 50 points $t_{11} = t_{21} = 0, \ldots, t_{1,50} = t_{2,50} = 1$ in $I_1 = I_2 = [0, 1]$ in the following way:

$$\begin{bmatrix} \mathbf{X}_{1j}(t_{1u}) \\ \mathbf{X}_{2j}(t_{2u}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_1(t_{1u}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi}_2(t_{2u}) \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{1j} \\ \boldsymbol{\alpha}_{2j} \end{bmatrix} + \boldsymbol{\varepsilon}_{ju},$$

where $j = 1, \ldots, n$, $u = 1, \ldots, 50$, the matrices $\boldsymbol{\Phi}_l$ were as in (3) and contained the Fourier basis functions only, $(\boldsymbol{\alpha}_{1j}^\top, \boldsymbol{\alpha}_{2j}^\top)^\top$ are $10p_*$-dimensional random vectors, and $\boldsymbol{\varepsilon}_{ju}^\top = (\varepsilon_{ju,1}, \ldots, \varepsilon_{ju,2p_*})$ were the measurement errors such that $\varepsilon_{ju,v} \sim N(0, 0.025a_{jv})$ were independent and a_{jv} was the range of the v-th row of the following matrix:

$$\begin{bmatrix} \boldsymbol{\Phi}_1(t_{11})\boldsymbol{\alpha}_{1j} & \ldots & \boldsymbol{\Phi}_1(t_{1,50})\boldsymbol{\alpha}_{1j} \\ \boldsymbol{\Phi}_2(t_{21})\boldsymbol{\alpha}_{2j} & \ldots & \boldsymbol{\Phi}_2(t_{2,50})\boldsymbol{\alpha}_{2j} \end{bmatrix}.$$

The random vectors $(\boldsymbol{\alpha}_{1j}^\top, \boldsymbol{\alpha}_{2j}^\top)^\top$ were generated in the following three settings, which were identical or similar to those considered by Zhu et al. (2020) in their experiments for multivariate data:

Setting 1 We generated i.i.d. samples $\boldsymbol{\alpha}_{1j}$ and $\boldsymbol{\alpha}_{2j}$ for $j = 1, \ldots, n$ under the null hypothesis in the following three cases:

(*i*) $\boldsymbol{\alpha}_{ij} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \mathbf{I}_{p_*})$ for $i = 1, 2$,

(*ii*) $\boldsymbol{\alpha}_{1j} \sim AR_{0.5}(1)$ and $\boldsymbol{\alpha}_{2j} \sim AR_{-0.5}(1)$, where $AR_\rho(1)$ denotes the Gaussian autoregressive model of order 1 with parameter ρ,

(*iii*) $\boldsymbol{\alpha}_{ij} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \boldsymbol{\Sigma}_{5p_*})$ for $i = 1, 2$, where $\boldsymbol{\Sigma}_a = (0.7^{|p-q|})_{p,q=1}^a$.

Setting 2 We generated i.i.d. samples $\boldsymbol{\alpha}_{1j} = (\alpha_{1j,1}, \ldots, \alpha_{1j,5p_*})^\top$ and $\boldsymbol{\alpha}_{2j} = (\alpha_{2j,1}, \ldots, \alpha_{2j,5p_*})^\top$ for $j = 1, \ldots, n$ under the alternative hypothesis in the following four cases:

(*i*) $\boldsymbol{\alpha}_{1j} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \mathbf{I}_{5p_*})$ and $\alpha_{2j,v} = \alpha_{1j,v}^2$ for $v = 1, \ldots, 5p_*$,

(*ii*) $\boldsymbol{\alpha}_{1j} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \boldsymbol{\Sigma}_{5p_*})$ and $\alpha_{2j,v} = \alpha_{1j,v}^2$ for $v = 1, \ldots, 5p_*$, $\boldsymbol{\Sigma}_a = (0.7^{|p-q|})_{p,q=1}^a$,

(*iii*) $\boldsymbol{\alpha}_{1j} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \mathbf{I}_{5p_*})$ and $\alpha_{2j,v} = \log|\alpha_{1j,v}|$ for $v = 1, \ldots, 5p_*$,

(*iv*) $\boldsymbol{\alpha}_{1j} \sim N_{5p_*}(\mathbf{0}_{5p_*}, \mathbf{I}_{5p_*})$ and $\alpha_{2j,v} = \alpha_{1j,v}/2 + Z_{j,v}$ for $v = 1, \ldots, 5p_*$, where $Z_{j,v} \sim N(0, 0.7)$ were independent.

Setting 3 We generated i.i.d. samples $\boldsymbol{\alpha}_{1j} = (\alpha_{1j,1}, \ldots, \alpha_{1j,5p_*})^\top$ and $\boldsymbol{\alpha}_{2j} = (\alpha_{2j,1}, \ldots, \alpha_{2j,5p_*})^\top$ for $j = 1, \ldots, n$ under the alternative hypothesis in the following four cases:

(*i*) $\alpha_{1j,1}, \ldots, \alpha_{1j,5p_*}$ were i.i.d. of uniform distribution $U(-1, 1)$ and $\alpha_{2j,v} = \alpha_{1j,v}^2$ for $v = 1, \ldots, 5p_*$,

(*ii*) $\alpha_{1j,1}, \ldots, \alpha_{1j,5p_*}$ were i.i.d. of uniform distribution $U(0, 1)$ and $\alpha_{2j,v} = 4\alpha_{1j,v}^3 - 3.6\alpha_{1j,v} + 0.8$ for $v = 1, \ldots, 5p_*$,

(*iii*) $\alpha_{1j,v} = \sin(z_{1j,v})$ and $\alpha_{2j,v} = \cos(z_{1j,v})$ for $v = 1, \ldots, 5p_*$, where $z_{1j,1}, \ldots, z_{1j,5p_*}$ were i.i.d. of uniform distribution $U(0, 2\pi)$,

(*iv*) $\alpha_{1j,1}, \ldots, \alpha_{1j,5p_*}$ were i.i.d. of uniform distribution $U(-1, 1)$ and $\alpha_{2j,v} = \alpha_{1j,v}/2 + Z_{j,v}$ for $v = 1, \ldots, 5p_*$, where $Z_{j,v} \sim N(0, 0.5)$ were independent.

Model 2 Here, we considered pairwise independence ($k = 2$) as in Model 1, but we generated $n = 20$ observations with dimensions $p_1 = p_2 = 3$ using well-known stochastic processes instead of direct use of the basis representation.

We considered the Wiener process, the Ornstein-Uhlenbeck process and the Brownian bridge separately. We set $I_1 = I_2 = [0, 1]$ and $m = 25$ as the number of equally spaced design time points in $[0, 1]$ at which the discrete functional data for the processes considered were observed. The Wiener process and the Ornstein-Uhlenbeck process were observed in points $t_1 = 0, \ldots, t_{m+1} = 1$, but the final observations used were those for the m points $t_2, \ldots, t_{m+1}$, because we removed the first zero value of these processes for t_1. The Brownian bridge was observed at the points $t_1 = 0, \ldots, t_{m+2} = 1$. However, the final observations used were those for the m points $t_2, \ldots, t_{m+1}$, as we removed the first and the last zero value of the Brownian bridge for t_1 and t_{m+2}.

Let Z denote a stochastic process chosen from the Wiener process, the Ornstein-Uhlenbeck process and the Brownian bridge. Then the coordinates of each observation in the first sample $\mathbf{X}_{11} = (X_{111}, X_{112}, X_{113})^\top, \ldots, \mathbf{X}_{1n} = (X_{1n1}, X_{1n2}, X_{1n3})^\top$ were independent realizations of the process Z. Of course, for each observation such realizations were generated independently. The second sample $\mathbf{X}_{21} = (X_{211}, X_{212}, X_{213})^\top, \ldots, \mathbf{X}_{2n} = (X_{2n1}, X_{2n2}, X_{2n3})^\top$ was generated in each of the following four cases ($j = 1, \ldots, n$; $v = 1, 2, 3$):

(i) The observations were obtained in the same way as for the first sample but independently;

(ii) $X_{2jv} = X_{1jv}^2$;

(iii) $X_{2jv} = \log |X_{1jv}|$;

(iv) $X_{2jv} = X_{1jv}/2 + Y_{jv}$, where Y_{jv} is the process of i.i.d. variables of distribution $N(0, 1.2)$.

In case (i) the null hypothesis held, while in the remaining cases the alternative hypothesis held.

In the above method of data generation, the variables of multivariate functional data were independent. Thus we call this the independent setting. Additionally, we considered the following dependent setting. Let $\mathbf{Y}_{1j} \sim N_9(\mathbf{0}_9, \mathbf{\Sigma})$, $j = 1, \ldots, n$ be independent random vectors, where $\mathbf{\Sigma} = \sigma\left((1-\rho)\mathbf{I}_9 + \rho\mathbf{1}_9\mathbf{1}_9^\top\right)$ with $\sigma = 0.01$ and $\rho = 0.1$. Then the dependent coordinates of the multivariate functional data of the first sample were generated as $\mathbf{X}_{1j}(t) + \mathbf{\Phi}(t)\mathbf{Y}_{1j}$, where $\mathbf{X}_{1j}$ were observations obtained in the independent setting, $t \in [0, 1]$, and the 3×9 matrix $\mathbf{\Phi}$ was as in Section 2 and contained three Fourier basis functions for each variable. In case (i), the second sample was generated in the same way.

Model 3 We considered mutual independence with $k = 3$ groups of multivariate functional data. We generated three samples with $n = 30$ observations each and equal dimensions $p_1 = p_2 = p_3 = 3$. The first and second (respectively the third) samples were obtained in the same way as the first (respectively the second) sample in Model 2. Naturally, the first two samples in this model were generated independently of each other. Similarly to Model 2, the null hypothesis held in case (i), while it did not hold in the other cases, since then the first and third processes were dependent.

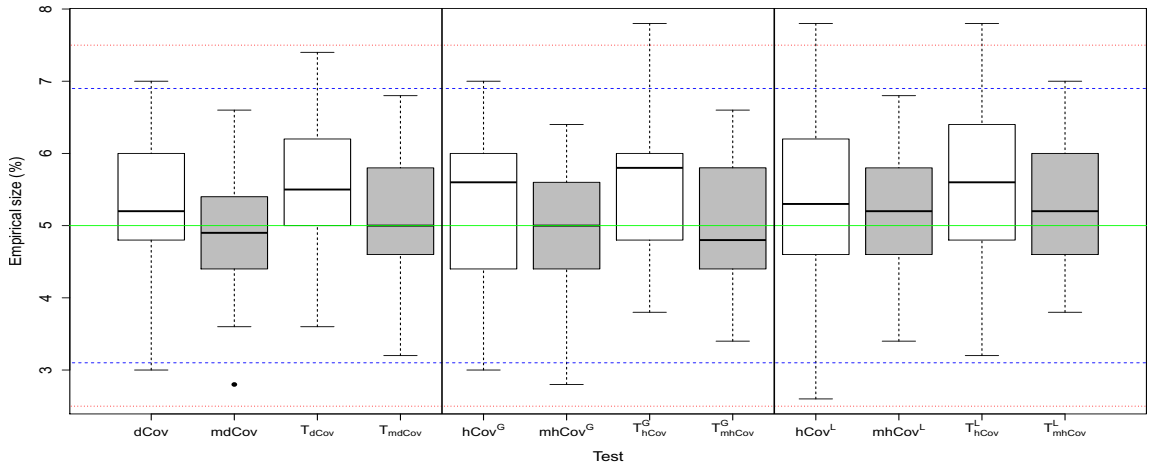

Figure 1: Box-and-whisker plots for empirical sizes (as percentages) of all tests obtained in Models 1–3. The solid, dashed and dotted horizontal lines represent the significance level 5% and the 95% and 99% binomial proportion confidence intervals $[3.1\%, 6.9\%]$ and $[2.5\%, 7.5\%]$ respectively. The two vertical lines divide the tests into three groups: on the left, tests based on distance covariance; in the middle, tests based on Hilbert-Schmidt covariance with Gaussian kernel ($hCov^G$, $mhCov^G$, T^G_{hCov}, T^G_{mhCov}); on the right, tests based on Hilbert-Schmidt covariance with Laplacian kernel ($hCov^L$, $mhCov^L$, T^L_{hCov}, T^L_{mhCov}).

6.3 Simulation Results

To save space, the simulation results for Models 1–3 are summarized by box-and-whisker and "line" plots in Figures 1–5 to illustrate the main findings of the simulation. The exact simulation results are available from the authors upon request.

In Figure 1, we can observe that all tests maintain the type I error level quite well. Their empirical sizes belong to the binomial proportion 95% and 99% confidence intervals $[3.1\%, 6.9\%]$ and $[2.5\%, 7.5\%]$ respectively (Duchesne & Francq, 2015) in almost all cases, as they should. There are only a few exceptions, mainly for the t-tests based on the joint distance and Hilbert-Schmidt covariances. This follows from the asymptotic character of these tests, i.e., they use a critical value based on the asymptotic distribution of test statistics. Moreover, this indicates that the t-tests are at least slightly more liberal than the permutation tests. It also seems that the tests based on marginally aggregated covariances control the type I error level at least slightly better then the tests based on joint distance and Hilbert-Schmidt covariances.

Figures 2 and 3 present a power comparison between joint and marginal versions of the distance and Hilbert-Schmidt covariances-based tests. We can observe that their performance depends on the character of dependence. For detecting non-linear dependence, the tests based on marginally aggregated covariances are better in terms of power than their joint counterparts. On the other hand, the reverse is usually true when the observations are linearly dependent. This is especially evident for Hilbert-Schmidt covariance-based tests,

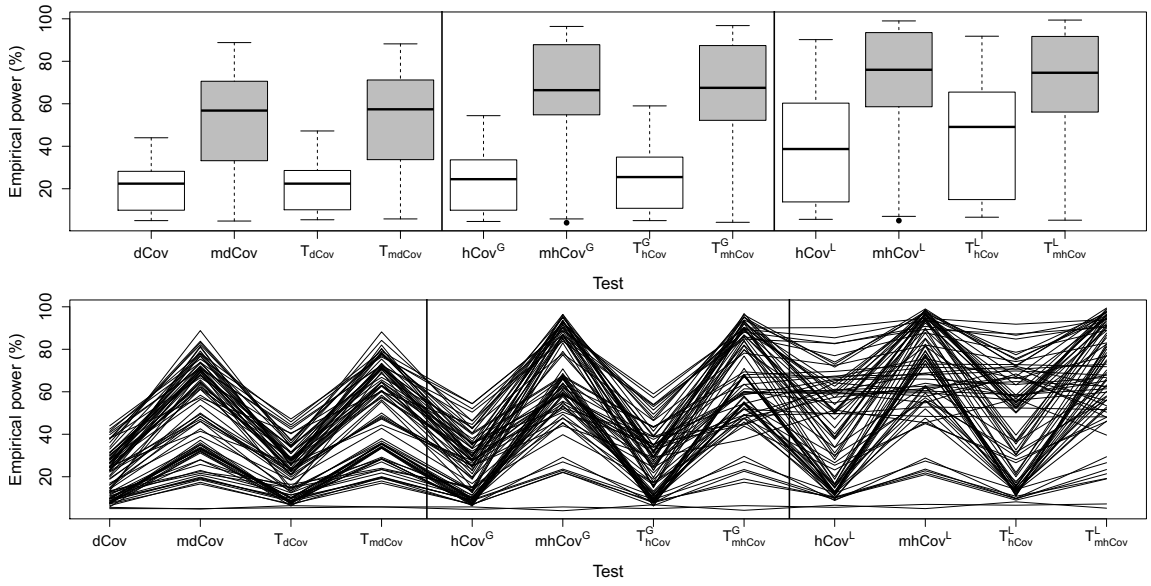

Figure 2: Box-and-whisker and "line" plots for empirical powers (as percentages) of all tests obtained in the case of non-linear dependence, i.e., Model 1, settings 2–3, cases (i)–(iii); Models 2–3 in cases (ii)–(iii). The two vertical lines divide the tests into three groups: on the left, tests based on distance covariance; in the middle, tests based on Hilbert-Schmidt covariance with Gaussian kernel ($hCov^G$, $mhCov^G$, T^G_{hCov}, T^G_{mhCov}); on the right, tests based on Hilbert-Schmidt covariance with Laplacian kernel ($hCov^L$, $mhCov^L$, T^L_{hCov}, T^L_{mhCov}).

while for distance covariance-based tests, this holds to a much lesser extent. These results are consistent with those of Zhu et al. (2020) for multivariate data. Namely, they indicate that the joint covariance-based tests can capture linear dependence of multivariate functional data very well, but are much less powerful, or even fail, when detection of non-linear dependence is required. This can be explained by the fact that these tests use the basis representation of functional data, which reduces them to possibly high-dimensional multivariate data. On the other hand, the marginal versions of the tests better capture non-linear dependence, but may be worse at detecting linear dependence.

Under non-linear dependence, the distance covariance-based tests are less powerful than the Hilbert-Schmidt covariance-based tests with Gaussian kernel, which is overcome by using these tests with the Laplacian kernel. For linear dependence, the reverse is true for marginally aggregated covariance-based tests, while the power of all joint covariance-based tests is much more stable. The permutation tests and the corresponding t-tests are comparable in terms of power, but the latter are often slightly more powerful than the former, which may be explained by the slightly liberal character of the t-tests.

In Figures 4-5, the results for $k = 3$ groups are presented. Thus, we compare the R_n and S_n methods considered in Section 5. These two methods seem to perform very similarly, but there are some differences, which we will now indicate. The S_n tests using the marginally

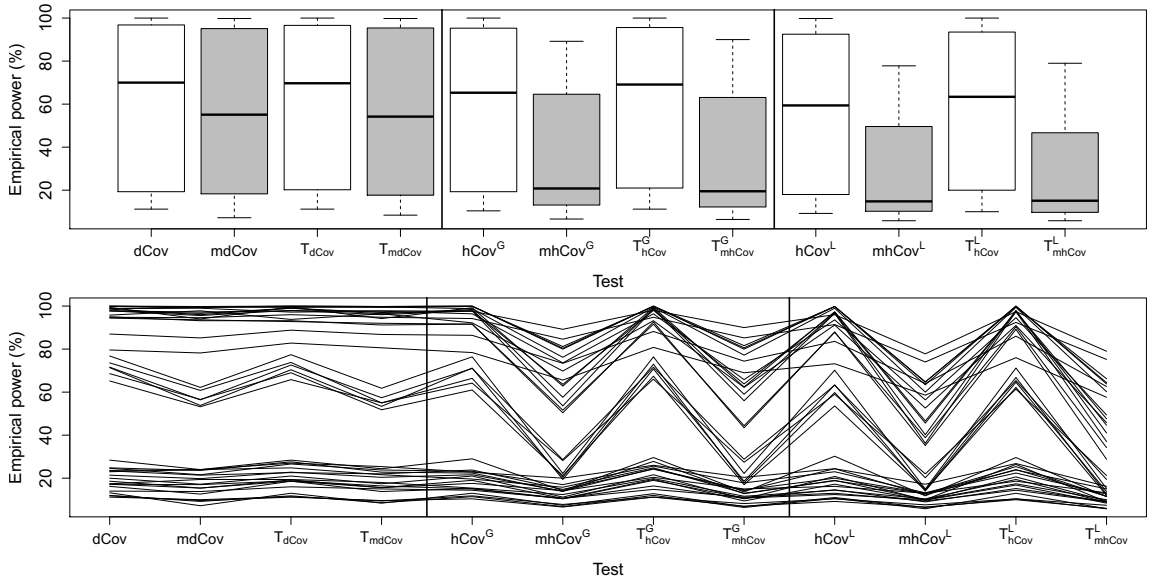

Figure 3: Same as Figure 2 but for linear dependence, i.e., case (iv) in Model 1 with settings 2–3 and in Models 2–3.

aggregated covariances are usually at least slightly more powerful then the R_n tests. Under non-linear dependence, the reverse holds for joint covariances, except for the $hCov_n^2$ tests. Nevertheless, it seems that both the R_n and S_n methods capture the mutual dependence by combining the pairwise dependence measure. We can also observe that the above findings about the behaviour of joint and marginal versions of tests for pairwise dependence also hold true for the R_n and S_n tests; in particular, this extends the results of Zhu et al. (2020). Moreover, the use of a pairwise dependence measure in the R_n and S_n methods which does not have to be equal to zero if and only if the processes are independent (e.g., $mdCov^2$) is also reasonable and results in powerful test procedures.

All of the above observations hold true for both Fourier- and B-spline-based tests. However, we can observe that the tests using a Fourier basis are more powerful than the B-spline-based tests (data not shown). This can perhaps be explained by the construction of the data (i.e., the Fourier basis is used in this construction) and the number of basis functions in the basis representation was set to five for both bases.

7. Real Data Example

For illustrative purposes, we re-analyse the pillar data set using the tests under consideration. This data set was investigated in Górecki et al. (2020) using functional canonical correlation analysis and and independent test based on Hilbert-Schmidt covariance with biased estimator.

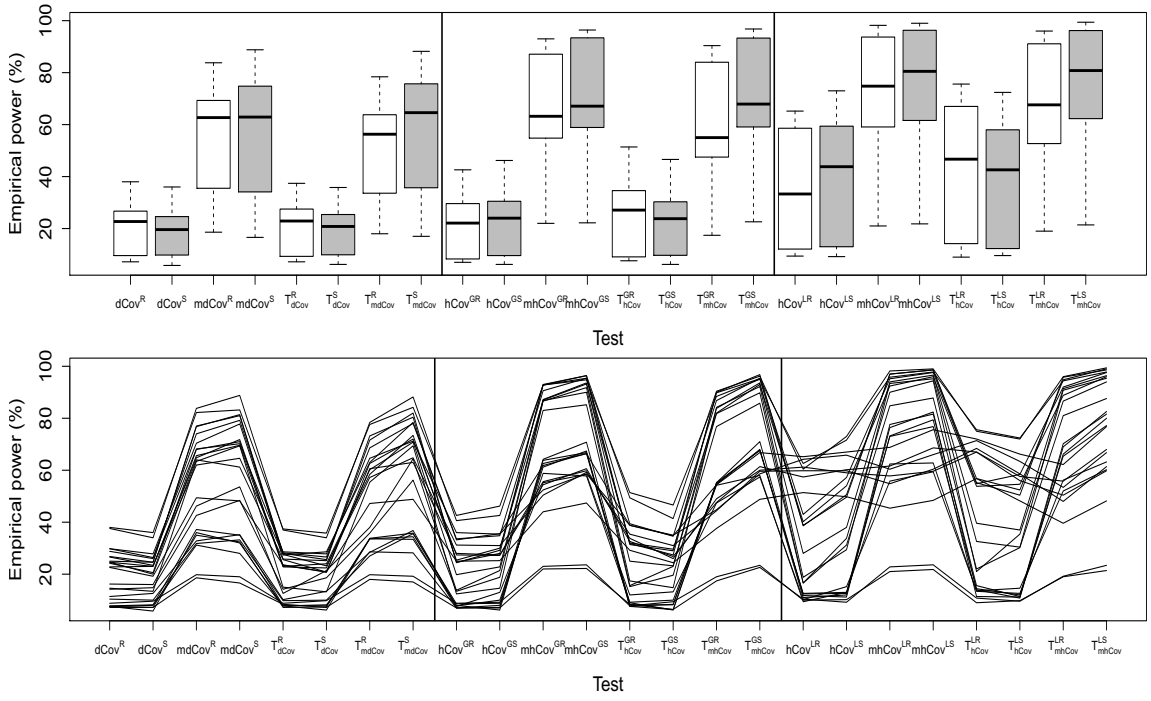

Figure 4: Box-and-whisker and "line" plots for empirical powers (as percentages) of R_n and S_n tests obtained in the case of non-linear dependence, i.e., Model 3 in cases (ii)–(iii). The two vertical lines divide the tests into three groups: on the left, tests based on distance covariance; in the middle, tests based on Hilbert-Schmidt covariance with Gaussian kernel ($hCov^{GR}$, $hCov^{GS}$, $mhCov^{GR}$, $mhCov^{GS}$, T^{GR}_{hCov}, T^{GS}_{hCov}, T^{GR}_{mhCov}, T^{GS}_{mhCov}); on the right, tests based on Hilbert-Schmidt covariance with Laplacian kernel ($hCov^{LR}$, $hCov^{LS}$, $mhCov^{LR}$, $mhCov^{LS}$, T^{LR}_{hCov}, T^{LS}_{hCov}, T^{LR}_{mhCov}, T^{LS}_{mhCov}).

7.1 Pillar Data Set

In the pillar data set, we have twelve pillars, which are groups of variables describing 38 European countries in the period 2008-2015. The data set was constructed based on the annual reports of the World Economic Forum (WEF) (`http://www.weforum.org`). The countries are listed in Table 3 of Górecki et al. (2020). The pillars are as follows: 1 – institutions (17), 2 – infrastructure (6), 3 – macroeconomic environment (2), 4 – health and primary education (7), 5 – higher education and training (6), 6 – goods market efficiency (10), 7 – labor market efficiency (6), 8 – financial market development (5), 9 – technological readiness (4), 10 – market size (4), 11 – business sophistication (9), 12 – innovation (5). The numbers in parentheses are the number of variables of each pillar, which are listed in Table 2 of Górecki et al. (2016). These variables describe various socio-economic conditions or spheres for individual states.

Note that the pillar data are doubly multivariate data, since all variables of all pillars are measured repeatedly in time, i.e., for 2008/2009, 2009/2010, ..., 2014/2015. This indicates

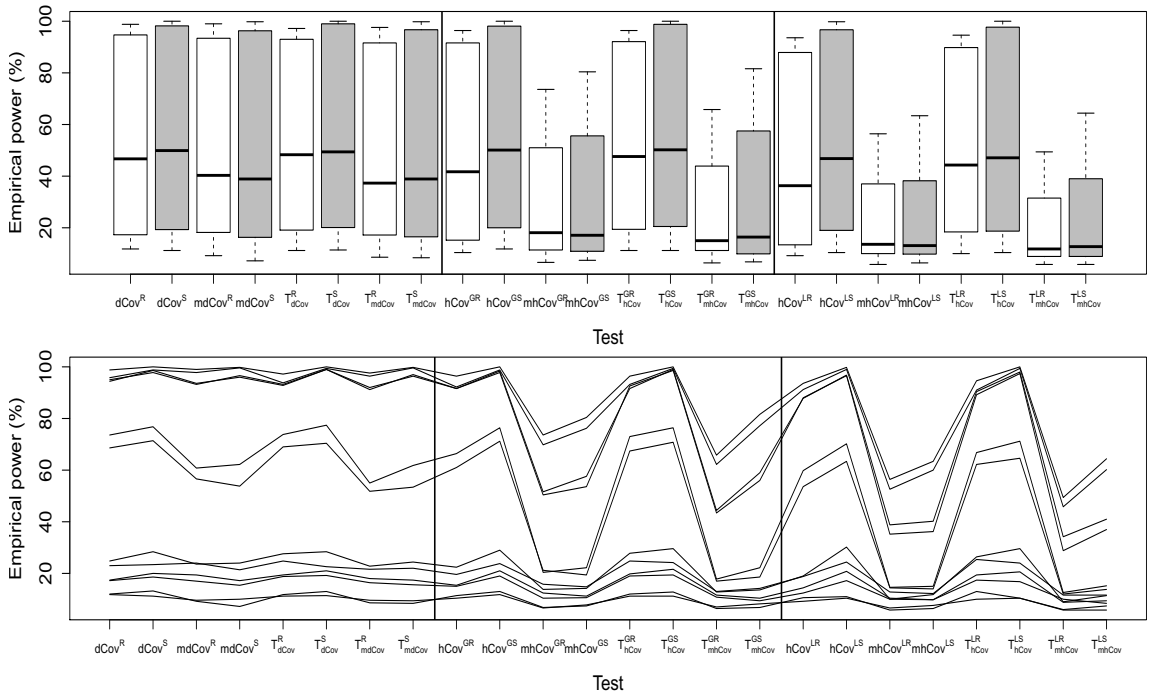

Figure 5: Same as Figure 4 but for linear dependence, i.e., Model 3 in case (iv).

that we can treat them as multivariate functional data measured at seven time points. Namely, each pillar is represented as a multivariate random process with the appropriate number of variables (e.g., pillar 1 has 17 functional coordinates), which are observed at $t_1 = 0.5, \ldots, t_7 = 6.5$, belonging to the interval $[0.5, 6.5]$. Thus, we have $n = 38$ functional observations characterized by twelve random processes $\mathbf{X}_i \in L_2^{p_i}([0.5, 6.5])$, $i = 1, \ldots, 12$, with $p_1 = 17, p_2 = 6, \ldots, p_{12} = 5$.

7.2 Results of Testing of Independence of Pillars

It is of interest to test the independence of particular pillars. For this purpose, the tests considered in Section 6 were applied. For the basis representation we used five Fourier basis functions, since the number of design time points is small. Nevertheless, for most pillars, the number of variables in the basis representation is quite large and even greater than the number of observations; for instance, for the first pillar, we have $17 \cdot 5 = 97$ variables. The p-values of the permutation tests were estimated based on 1000 permuted samples.

First, we consider pairwise independence ($k = 2$) between all pairs of pillars. The resulting p-values are presented in Tables 1-3. All tests detect the dependence of most pairs of pillars. For example, the p-values of all tests for pillars 1 and 4 are always equal to zero, indicating their dependence. However, there are some pairs of pillars for which the tests have very different p-values. This applies mainly to comparisons of pillars 2 and 10 with other pillars. These two pillars are dependent on each other, but the distance covariance-based tests indicate that each of them is independent of some other pillars. On the other hand, the Hilbert-Schmidt covariance-based tests give a lower number of such independent

$dCov_n^2$ (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
$mdCov_n^2$ (utm)	1		4.8	1.1	0.0	0.0	0.0	0.0	0.0	0.0	22.9	0.0	0.0
	2	5.4		1.8	17.2	2.8	3.4	5.3	22.5	18.4	0.0	0.6	0.7
	3	2.6	1.4		0.0	1.4	0.6	0.1	1.1	11.1	7.4	0.1	0.4
	4	0.0	35.9	0.0		0.0	0.0	0.0	0.0	0.0	65.6	0.0	0.0
	5	0.0	2.7	3.0	0.0		0.0	0.0	0.0	0.0	11.2	0.0	0.0
	6	0.0	3.2	3.0	0.0	0.0		0.0	0.0	0.0	10.2	0.0	0.0
	7	0.0	5.8	0.2	0.0	0.0	0.0		0.0	0.0	22.7	0.0	0.0
	8	0.0	18.4	12.5	0.0	0.0	0.0	0.0		0.0	37.3	0.0	0.0
	9	0.0	22.1	28.4	0.0	0.0	0.0	0.0	0.0		61.5	0.0	0.0
	10	24.9	0.0	5.4	74.2	8.0	7.6	23.2	29.3	59.7		3.0	4.2
	11	0.0	0.7	0.4	0.0	0.0	0.0	0.0	0.0	0.0	2.0		0.0
	12	0.0	0.9	2.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	
T_{dCov} (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
T_{mdCov} (utm)	1		3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	33.7	0.0	0.0
	2	4.2		0.0	26.0	0.6	0.6	3.2	25.9	25.2	0.0	0.0	0.0
	3	0.3	0.0		0.0	0.0	0.0	0.0	0.1	10.1	3.4	0.0	0.0
	4	0.0	47.6	0.0		0.0	0.0	0.0	0.0	0.0	68.5	0.0	0.0
	5	0.0	0.6	0.4	0.0		0.0	0.0	0.0	0.0	14.2	0.0	0.0
	6	0.0	0.4	0.7	0.0	0.0		0.0	0.0	0.0	11.0	0.0	0.0
	7	0.0	4.4	0.0	0.0	0.0	0.0		0.0	0.0	28.9	0.0	0.0
	8	0.0	22.6	14.7	0.0	0.0	0.0	0.0		0.0	47.1	0.0	0.0
	9	0.0	31.5	39.6	0.0	0.0	0.0	0.0	0.0		68.7	0.0	0.0
	10	36.6	0.0	1.7	74.0	9.2	6.3	28.7	38.4	67.9		0.4	2.0
	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3		0.0
	12	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	

Table 1: P-values (as percentages) of the $dCov_n^2$, $mdCov_n^2$, T_{dCov} and T_{mdCov} tests for pairwise independence for the pillar data set (ltm - lower triangular of matrix, utm - upper triangular of matrix)

instances. Finally, the $mhCov_n^2$ and T_{mhCov} tests, which seem to be the most powerful for the pillars data set (as explained below), suggest the independence of pillars 8 and 10 only (at significance level of 5%).

It seems that the above remarks are consistent with the simulation results for the case of non-linear dependence. We explain this as follows. First of all, the p-values of the tests based on marginally aggregated covariances are usually smaller than those of the corresponding joint covariances-based tests. Moreover, in some cases, the decisions suggested by the joint and marginal versions of the tests are different; for instance, the $mhCov_n^2$ and T_{mhCov} tests reject the independence of pillars 9 and 10, while the $hCov_n^2$ and T_{hCov} tests do not. All this confirms that the marginally aggregated covariances-based tests have better ability to detect non-linear dependence than the joint covariance tests. Furthermore, the p-values of the distance covariance-based tests are greater than those of the Hilbert-Schmidt covariance-based tests with Gaussian kernel, which are greater than the p-values of the latter tests using the Laplacian kernel. This is consistent with the increase in the power of these tests

$hCov_n^2$ (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
$mhCov_n^2$ (utm)	1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
	2	0.8		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
	3	0.4	0.5		0.0	0.1	0.0	0.0	0.0	0.0	0.3	0.0	0.0
	4	0.0	1.1	0.0		0.0	0.0	0.0	0.6	0.0	1.2	0.0	0.0
	5	0.0	0.2	1.2	0.0		0.0	0.0	0.0	0.0	0.5	0.0	0.0
	6	0.0	0.2	0.4	0.0	0.0		0.0	0.0	0.0	0.1	0.0	0.0
	7	0.0	0.0	0.5	0.0	0.0	0.0		0.0	0.0	0.3	0.0	0.0
	8	0.0	4.0	8.4	0.0	0.0	0.0	0.0		0.0	12.5	0.0	0.0
	9	0.0	10.4	17.1	0.0	0.0	0.0	0.0	0.0		1.1	0.0	0.0
	10	15.1	0.0	3.7	6.8	3.5	1.1	2.7	13.9	35.0		0.1	0.1
	11	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.2		0.0
	12	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	
T_{hCov} (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
T_{mhCov} (utm)	1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
	2	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.1	0.0		0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0
	5	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.1	0.0	0.0
	6	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0
	7	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.1	0.0	0.0
	8	0.0	2.4	6.1	0.0	0.0	0.0	0.0		0.0	9.8	0.0	0.0
	9	0.0	12.6	18.2	0.0	0.0	0.0	0.0	0.0		0.2	0.0	0.0
	10	13.3	0.0	2.1	5.7	0.8	0.1	0.5	16.1	42.6		0.0	0.0
	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0
	12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

Table 2: P-values (as percentages) of the $hCov_n^2$, $mhCov_n^2$, T_{hCov} and T_{mhCov} tests with Gaussian kernel for pairwise independence for the pillar data set (ltm - lower triangular of matrix, utm - upper triangular of matrix)

observed in simulations for non-linear dependence. Finally, the behaviour of permutation tests and corresponding t-tests is similar, but the latter usually have smaller p-values than the former. This can be explained by the slightly more powerful character of the t-tests in comparison with the permutation tests.

Now, we consider testing of mutual dependence with $k = 3$, for all triples of pillars. To save space, the p-values of all tests for all 220 triples of pillars are not shown, but available from the authors upon request. Unfortunately, they are not easy to follow, even though in 99 cases the p-values of all tests are equal to zero. Thus, we summarize the results in Table 4, where for each test, the proportions of rejections of the null hypothesis for all tests are presented. These indicate that most of the null hypotheses of independence are rejected, and some of the tests reject all of them. The triples of pillars which were found independent mainly contain the pillars 2 or 10, similarly as for pairwise independence. Moreover, these proportions, as well as closer analysis of the p-values, indicate that the comments made in the above paragraph also hold true in the mutual independence case. Finally, we can

$hCov_n^2$ (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
$mhCov_n^2$ (utm)	1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
	2	0.4		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
	3	0.2	0.0		0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.3	0.0		0.0	0.0	0.0	0.6	0.0	0.8	0.0	0.0
	5	0.0	0.3	0.8	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
	6	0.0	0.0	0.2	0.0	0.0		0.0	0.0	0.0	0.2	0.0	0.0
	7	0.0	0.0	0.6	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0
	8	0.0	2.6	4.3	0.0	0.0	0.0	0.0		0.0	10.0	0.0	0.0
	9	0.0	6.2	28.2	0.0	0.0	0.0	0.0	0.0		0.4	0.0	0.0
	10	7.1	0.0	3.0	4.0	2.2	0.2	2.1	14.3	31.7		0.0	0.2
	11	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.1		0.0
	12	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	
T_{hCov} (ltm)		1	2	3	4	5	6	7	8	9	10	11	12
T_{mhCov} (utm)	1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	2	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	4	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0
	5	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0
	6	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0
	7	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	0.0	0.0
	8	0.0	1.1	1.6	0.0	0.0	0.0	0.0		0.0	8.8	0.0	0.0
	9	0.0	5.4	33.5	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0
	10	3.2	0.0	1.2	1.4	0.3	0.0	0.6	16.2	39.1		0.0	0.0
	11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0
	12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	

Table 3: P-values (as percentages) of the $hCov_n^2$, $mhCov_n^2$, T_{hCov} and T_{mhCov} tests with Laplacian kernel for pairwise independence for the pillar data set (ltm - lower triangular of matrix, utm - upper triangular of matrix)

observe that the R_n and S_n methods perform fairly well. They also behave very similarly to each other. The only exceptions to this are the methods based on the T_{dCov} and T_{mdCov} tests, where the R_n tests reject the hypothesis of mutual independence more often than the S_n tests.

8. Conclusions

We have considered both pairwise and mutual independence testing problems for multivariate functional data. For these problems, we proposed tests based on the basis representation of functional data, which apply the independence tests for multivariate data obtained in this representation. We focused on tests based on the commonly used distance and Hilbert-Schmidt covariances, and their marginal versions as recently proposed in the literature. The theoretical justification of the test procedures was established. Moreover, in intensive simulation studies and real data examples, the obtained tests were compared in terms of maintenance of the type I error level and power. Most of them control the type I error

M	$dCov$	$mdCov$	T_{dCov}	T_{mdCov}
R	73.6	75.5	82.3	82.3
S	74.5	75.5	76.4	76.4
	Gaussian kernel			
M	$hCov$	$mhCov$	T_{hCov}	T_{mhCov}
R	93.6	100.0	95.9	100.0
S	92.3	100.0	93.2	100.0
	Laplacian kernel			
M	$hCov$	$mhCov$	T_{hCov}	T_{mhCov}
R	97.3	100.0	97.3	100.0
S	96.8	100.0	97.7	100.0

Table 4: Proportions (as percentages) of rejections of the null hypothesis in all tests for mutual independence with $k = 3$, for all triples of pillars

level very well, but the power comparison is more complex. The numerical experiments indicate that the drawbacks of the distance and Hilbert-Schmidt covariances-based tests, as well as those of their marginal versions, seem to hold also for functional data. In particular, we observed that the distance and Hilbert-Schmidt covariances-based tests are usually more (respectively less) powerful than their marginal versions under linear (respectively non-linear) dependence. Moreover, in general, the distance covariance-based tests are more powerful than the Hilbert-Schmidt covariance-based tests under linear dependence, while the reverse is true under non-linear dependence. These findings, among others, show that no one test is uniformly superior to the others. Thus, constructing a test procedure, which can successfully detect both linear and non-linear dependence in high-dimensional data or multivariate functional data seems to be an interesting direction for future research.

Acknowledgments

The authors are grateful to Mr. Changbo Zhu for sharing us the code for Zhu et al. (2020). We are also grateful to the anonymous referees for many valuable comments and constructive suggestions which resulted in the present version. A part of calculations for simulation study and real data example was made at the Poznań Supercomputing and Networking Center.

Appendix A. Comparison of Raw and Functional Marginal Approaches

As we noticed in Section 1.2, the functional versions of independence tests usually perform better than the direct application of multivariate methods to raw functional data. This was shown by Górecki et al. (2020) for the joint covariances-based tests. In this appendix, we briefly present that this seems to hold also for marginal covariances-based tests, what was suggested by one of the reviewers. Namely, we conducted the simulation studies in Model 1 with $p_* = 3$ for all marginal covariances-based tests in two versions: raw and functional.

The raw marginal covariances-based tests were constructed as follows: Let $\mathbf{D}_{ij}$, $i = 1, 2$, $j = 1, 2, 3$, be matrices of dimension 20×50 ($n = 20$, the number of design time points is equal to 50), which represent the raw functional observations of the three variables of the two random processes. For each process, these raw data were combined into the 20×150 matrix $\mathbf{D}_i = (\mathbf{D}_{i1}\ \mathbf{D}_{i2}\ \mathbf{D}_{i3})$, $i = 1, 2$. The matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ correspond to two samples of multivariate data. Then, to such data, the marginal covariances-based tests were applied.

As a result of these simulation studies, we obtained the empirical sizes and powers of the raw and functional marginal covariances-based tests. In terms of control of the type I error level, there are no significant differences and the raw and functional marginal covariances-based test procedures perform equally well. However, the power performance is more interesting. The empirical powers of all tests are summarized by box-and-whisker plots in Figure 6. We can easily observe that the findings are similar to those for the joint covariances-based tests as was studied by Górecki et al. (2020) (see Section 1.2). Namely, for the non-linear dependence, the functional marginal covariances-based tests usually outperform their raw versions. This is especially seen for the test procedures based on the Hilbert-Schmidt covariance. On the other hand, in the case of linear dependence, the raw and functional marginal covariances-based tests perform very similarly, but the former seems to be slightly more powerful than the latter. To sum up, the functional versions of the tests generally seem to be better than direct application of the multivariate methods to the raw functional data.

References

Aguilera, A. M., Acal, C., Aguilera-Morillo, M. C., Jiménez-Molinos, F., & Roldán, J. B. (2021). Homogeneity problem for basis expansion of functional data with applications to resistive memories. *Mathematics and Computers in Simulation*, *186*, 41–51.

Bergsma, W., & Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli*, *20*, 1006–1028.

Duchesne, P., & Francq, C. (2015). Multivariate hypothesis testing using generalized and {2}-inverses - with applications. *Statistics*, *49*, 475–496.

Ferraty, F., & Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*. Springer.

Gan, L., Narisetty, N. N., & Liang, F. (2019). Bayesian regularization for graphical models with unequal shrinkage. *Journal of the American Statistical Association*, *114*, 1218–1231.

Górecki, T., Krzyśko, M., Ratajczak, W., & Wołyński, W. (2016). An extension of the classical distance correlation coefficient for multivariate functional data with applications. *Statistics in Transition new series*, *17*, 449–466.

Górecki, T., Krzyśko, M., & Wołyński, W. (2020). Independence test and canonical correlation analysis based on the alignment between kernel matrices for multivariate functional data. *Artificial Intelligence Review*, *53*, 475–499.

Gretton, A., Bousquet, O., Smola, A., & Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *ALT, Springer-Verlag pages 63–77*.

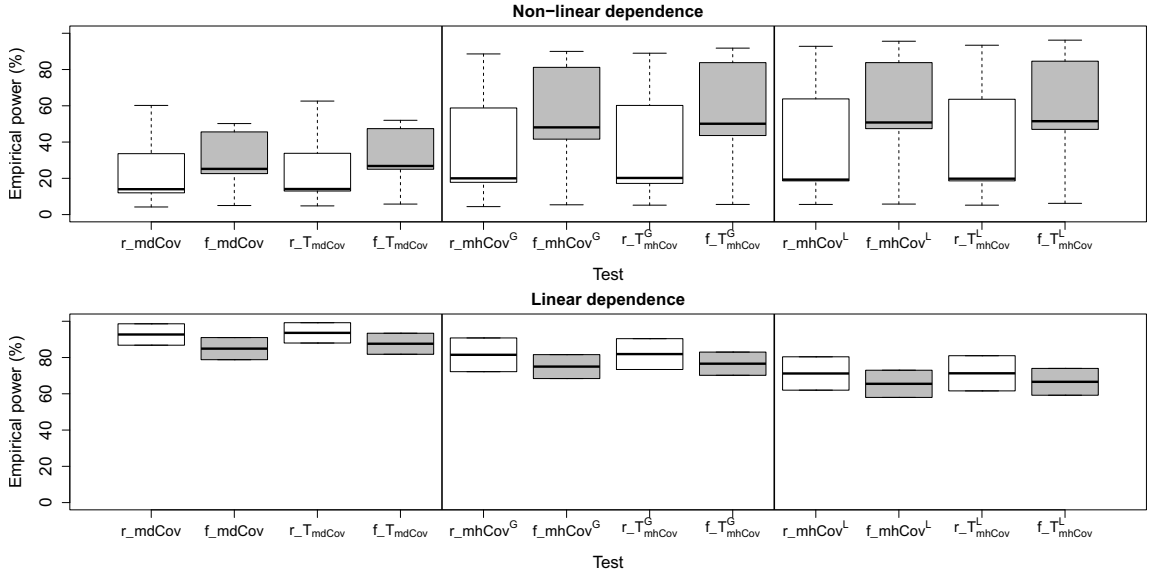

Figure 6: Box-and-whisker plots for empirical powers (przedostatnias percentages) of the raw and functional marginal covariances-based tests obtained in Model 1 with $p_* = 3$, denoted by prefix "r" and "f", respectively. The two vertical lines divide the tests into three groups: on the left, tests based on distance covariance; in the middle, tests based on Hilbert-Schmidt covariance with Gaussian kernel ($mhCov^G$, T^G_{mhCov}); on the right, tests based on Hilbert-Schmidt covariance with Laplacian kernel ($mhCov^L$, T^L_{mhCov}).

Gretton, A., Fukumizu, K., Harchaoui, Z., & Sriperumbudur, B. K. (2009). A fast, consistent kernel two-sample test. In *Advances in Neural Information Processing Systems 22, Curran Associates, Inc., pages 673–681.*

Gretton, A., Fukumizu, K., Teo, C. H., Song, L., Schölkopf, B., & Smola, A. (2008). A kernel statistical test of independence. In *Advances in Neural Information Processing Systems 20, Cambridge, MA. MIT Press, pages 585–592.*

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

Hua, W. Y., & Ghosh, D. (2015). Equivalence of kernel machine regression and kernel distance covariance for multidimensional phenotype association studies. *Biometrics*, *71*, 812–820.

Jin, Z., & Matteson, D. S. (2018). Generalizing distance covariance to measure and test multivariate mutual dependence via complete and incomplete V-statistics. *Journal of Multivariate Analysis*, *168*, 304–322.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, *30*, 81–93.

Kong, J., Klein, B. E., Klein, R., Lee, K. E., & Wahba, G. (2012). Using distance correlation and SS-ANOVA to assess associations of familial relationships, lifestyle factors,

diseases, and mortality. *Proceedings of the National Academy of Sciences*, *109*, 20352–20357.

Krzyśko, M., & Waszak, Ł. (2013). Canonical correlation analysis for functional data. *Biometrical Letters*, *50*, 95–105.

Leung, D., & Drton, M. (2018). Testing independence in high dimensions with sums of rank correlations. *The Annals of Statistics*, *46*, 280–307.

Li, R., Zhong, W., & Zhu, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association*, *107*, 1129–1139.

Li, Z., & McCormick, T. H. (2019). An expectation conditional maximization approach for Gaussian graphical models. *Journal of Computational and Graphical Statistics*, *28*, 767–777.

Lin, Z., Lopes, M. E., & Müller, H. G. (2021). High-dimensional MANOVA via bootstrapping and its application to functional and sparse count data. *Journal of the American Statistical Association*, 10.1080/01621459.2021.1920959.

Matteson, D. S., & Tsay, R. S. (2017). Independent component analysis via distance covariance. *Journal of the American Statistical Association*, *112*, 623–637.

Pan, G., Gao, J., & Yang, Y. (2014). Testing independence among a large number of high-dimensional random vectors. *Journal of the American Statistical Association*, *109*, 600–612.

Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*, 240–242.

R Core Team (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

Ramsay, J. O., & Silverman, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer-Verlag.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis, Second Edition*. Springer.

Shao, X., & Zhang, J. (2014). Martingale difference correlation and its use in high-dimensional variable screening. *Journal of the American Statistical Association*, *109*, 1302–1318.

Spearman, C. E. (1904). The proof and measurement of association between two things. *American Journal of Psychiatry*, *15*, 72–101.

Székely, G. J., & Rizzo, M. L. (2009). Brownian distance covariance. *The Annals of Applied Statistics*, *3*, 1236–1265.

Székely, G. J., & Rizzo, M. L. (2013). The distance correlation t-test of independence in high dimension. *Journal of Multivariate Analysis*, *117*, 193–213.

Székely, G. J., & Rizzo, M. L. (2014). Partial distance correlation with methods for dissimilarities. *The Annals of Statistics*, *42*, 2382–2412.

Székely, G. J., Rizzo, M. L., & Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, *35*, 2769–2794.

Taskinen, S., Kankainen, A., & Oja, H. (2003). Sign test of independence between two random vectors. *Statistics and Probability Letters*, *62*, 9–21.

Yang, Y., & Pan, G. (2015). Independence test for high dimensional data based on regularized canonical correlation coefficients. *The Annals of Statistics*, *43*, 467–500.

Yao, S., Zhang, X., & Shao, X. (2018). Testing mutual independence in high dimension via distance covariance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *80*, 455–480.

Zhang, J. T. (2013). *Analysis of Variance for Functional Data*. Chapman & Hall, London.

Zhu, C., Zhang, X., Yao, S., & Shao, X. (2020). Distance-based and RKHS-based dependence metrics in high dimension. *The Annals of Statistics*, *48*, 3366–3394.