



# Testing marginal homogeneity in Hilbert spaces with applications to stock market returns

Marc Ditzhaus<sup>1</sup> · Daniel Gaigall<sup>2</sup>

Received: 4 May 2021 / Accepted: 24 January 2022 / Published online: 14 February 2022  
© The Author(s) 2022

## Abstract

This paper considers a paired data framework and discusses the question of marginal homogeneity of bivariate high-dimensional or functional data. The related testing problem can be endowed into a more general setting for paired random variables taking values in a general Hilbert space. To address this problem, a Cramér-von-Mises type test statistic is applied and a bootstrap procedure is suggested to obtain critical values and finally a consistent test. The desired properties of a bootstrap test can be derived that are asymptotic exactness under the null hypothesis and consistency under alternatives. Simulations show the quality of the test in the finite sample case. A possible application is the comparison of two possibly dependent stock market returns based on functional data. The approach is demonstrated based on historical data for different stock market indices.

**Keywords** Marginal homogeneity · Functional data · Bootstrap test ·  $U$ -statistic · Cramér-von-Mises test · Stock market return

**Mathematics Subject Classification** 62G10 · 62G09

## 1 Introduction

Due to the availability of high-frequency data, statistical observations can be described and modeled by random functions, the so-called stochastic processes. Since classical methods are designed for vector-valued observations rather than for stochastic pro-

---

✉ Marc Ditzhaus  
marc.ditzhaus@tu-dortmund.de

Daniel Gaigall  
daniel.gaigall@insurance.uni-hannover.de

<sup>1</sup> Department of Statistics, TU Dortmund University, 44221 Dortmund, Germany

<sup>2</sup> Institute of Actuarial and Financial Mathematics, Leibniz University Hannover, 30167 Hannover, Germany

cesses, they usually cannot be applied in this situation. The field of functional data tries to close this gap. One popular solution to tackle this problem is to project the random functions to the real line and then apply one of the classical methods. For example, Cuesta-Albertos et al. (2006, 2007) applied the Kolmogorov–Smirnov goodness-of-fit test to randomly projected square-integrable functions. Cuevas and Fraiman (2009) extended this idea to more general spaces. Ditzhaus and Gaigall (2018) did the same by discussing observations with values in a general Hilbert space. In this way, stochastic processes as well as high-dimensional data can be discussed simultaneously. More interaction between these two fields is desirable as stated by Goia and Vieu (2016) and Cuevas (2014) in the functional data community as well as by Ahmed (2017) from the high-dimensional side. Extending the idea of goodness-of-fit, Bugni et al. (2009) studied the testing problem of whether the underlying distribution belongs to a pre-specified parametric family. Hall and Van Keilegom (2002) discussed pre-processing the functional data, which is in practice usually just available at finitely many time points, in the context of two-sample testing. Recently, the two-sample testing problem for multivariate functional data was discussed by Jiang et al. (2017) based on empirical characteristic functions.

In this paper, we use also the projection idea mentioned, but we address a paired-sample testing problem allowing for dependence. This approach and our test are completely new. We suggest a procedure for testing marginal homogeneity in separable Hilbert spaces, where we consider not just a few random projections but all projections from a sufficient large projection space. The advantage of our approach is that no additional randomness influences the result of the test. With a view to the consistency of the testing procedure, we apply a test statistic of Cramér–von-Mises type. See Anderson and Darling (1952) and Rosenblatt (1952) for Cramér–von-Mises tests in the usual cases of real-valued random variables and random vectors with real components, and Gaigall (2019) for the application of a Cramér–von-Mises test for the null hypothesis of marginal homogeneity for bivariate distributions on the Cartesian square of the real line. However, the demand for consistency has a price: the distribution of the test statistic under the null hypothesis is unknown, and so related quantiles are not available in practice. In contrast with the unpaired setting, exchangeability is not given in general under the null hypothesis of marginal homogeneity. This is already known for bivariate distributions on the Cartesian square of the real line, see Gaigall (2019). For that reason, a permutation test, such as they are used by Hall and Tajvidi (2002) and Bugni and Horowitz (2018) for the unpaired two-sample setting under functional data, or in Bugni and Horowitz (2018) in the more general situation of one control group against several treatment groups, is no option in our situation. To solve this problem, we offer a bootstrap procedure to determine critical values. We can show that the bootstrap test keeps the nominal level asymptotically under the null hypothesis. Moreover, we prove the consistency of the bootstrap procedure under any alternative.

While also other applications are possible, for instance to high-dimensional data, we especially focus on functional data and stock market returns. In the field of empirical finance, statistical inference of stock market prices is a widespread topic. Here, usual statistical procedures are typically applied to stock market log returns. One popular question is that of a suitable distributional model which matched with the observations.

For example, Göncü et al. (2016) apply different goodness-of-fit tests to data sets consisting of daily stock market index returns for several emerging and developed markets and Gray and French (2008) consider the distribution of log stock index returns of the S&P 500 and deduce that the distributions do not follow a normal distribution but demonstrate a greater ability for other distributional models. This topic is also treated based on high-frequency data, see Malevergne et al. (2005), where 5 min returns of the Nasdaq Composite index and 1 min returns of the S&P 500 are considered. Besides the topic of model selection for stock market returns, another interesting topic is the comparison of different stock market returns as it is done in Midesia et al. (2016), where annual pooled data of 100 conventional and Islamic stock returns are analyzed. In this context, the dependence of the different stock prices is obvious and already detected, see Min (2015), where the major 8 companies of the Korean stock market are investigated, and should be taken into account.

The paper is structured as follows. We first introduce the model and our general null hypothesis of marginal homogeneity in the paired sample setting in Sect. 2. We introduce a Cramér–von-Mises type test for the aforementioned testing problem and derive its asymptotic behavior with the help of the theory of  $U$ -statistics. The resulting asymptotic law under the null hypothesis can be transferred to a bootstrap counterpart of the test statistic. In addition to these theoretical findings, we study the small sample performance of the two resampling tests in a numerical simulation study presented in Sect. 3. Finally, the application to stock market indices is outlined in Sect. 4. We demonstrate the application of the test to the historical values of the stock market indices Nikkei Stock Average from Japan, Dow Jones Industrial Average from the US, and Standard & Poor's 500 from the US. The test confirms the intuition that the indices of the same country are much more comparable than indices of different countries. Note that all proofs are conducted in the Appendix.

## 2 Testing marginal homogeneity in Hilbert spaces

Let  $H$  be a Hilbert space, i.e., a real inner product space, where the inner product is denoted by  $\langle \cdot, \cdot \rangle$ . We suppose that  $H$  is separable with countable orthonormal basis  $O = \{e_i; i \in I\}$ , where  $e_i$  is the  $i$ -th basis element and the index set  $I$  is given by the natural numbers  $I = \mathbb{N}$  or the subset  $I = \{1, \dots, |I|\} \subset \mathbb{N}$ . Now, let paired observations be given

$$X_j = (X_{j,1}, X_{j,2}), \quad j = 1, \dots, n,$$

that are random variables with values in  $H \times H$ . We suppose that  $X_1, \dots, X_n$  are independent and identically distributed, and we suppose that the distribution  $P^{X_1}$  of  $X_1$  is unknown. For technical reasons, we suppose that for all  $i \in I$  the joint distribution of  $\langle X_{1,1}, e_i \rangle$  and  $\langle X_{1,2}, e_i \rangle$  is absolutely continuous with density  $f_i$ , where the set  $\{f_i(r, s) > 0; (r, s) \in \mathbb{R}^2\}$  is open and convex, compare with Gaigall (2019). While we allow any dependence structure between  $X_{j,1}$  and  $X_{j,2}$ , we like to infer the null hypothesis of marginal homogeneity

$$\mathcal{H} : P^{X_{j,1}} = P^{X_{j,2}} \quad \text{versus} \quad \mathcal{K} : P^{X_{j,1}} \neq P^{X_{j,2}}.$$

Our main application is the functional data case, where our observations are measurable and square-integrable real-valued functions  $X_{j,k}(t)$ ,  $t \in [0, T]$ ,  $k = 1, 2$ ,  $j = 1, \dots, n$ , on the interval  $[0, T]$ , and we consider the specific space  $H = L^2[0, T]$  containing all measurable and square-integrable real-valued functions on the interval  $[0, T]$  of length  $T \in (0, \infty)$  and equipped with the usual inner product  $\langle f, g \rangle = \int_0^T f(x)g(x) dx$ ,  $f, g \in H$ . A corresponding orthonormal basis is given by normalized Legendre polynomials. Another possible application is the high-dimensional case. Introducing the dimension  $d = |I|$ , considering the index set  $I = \{1, \dots, d\} \subset \mathbb{N}$ , and setting  $\mathbb{R}^I = \{f; f : I \mapsto \mathbb{R}\}$ , we consider random vectors  $X_{1,k}$ ,  $k = 1, 2$ , seen as random variables with values in the separable Hilbert space  $H = \{x; x \in \mathbb{R}^I, \sum_{i \in I} x(i)^2 r_i < \infty\}$  with inner product  $\langle x, x' \rangle = \sum_{i \in I} x(i)x'(i)r_i$ ,  $(x, x') \in H^2$ , where the elements  $e_i = \delta_i/\sqrt{r_i}$ ,  $i \in I$ , define a countable orthonormal basis  $O = \{e_i; i \in I\}$  and  $\eta = \sum_{i \in I} r_i \delta_i$  is a measure on the index set  $I$  with  $r_i \in (0, \infty)$  for all  $i \in I$ . This point of view enables an extension to the infinite-dimensional case  $d \in \mathbb{N} \cup \{\infty\}$ , where  $I = \mathbb{N}$  in the infinite-dimensional case  $d = \infty$ , and  $X_{1,k} = (X_{1,k}(i))_{i \in I}$ ,  $k = 1, 2$ , are now random sequences.

As postulated in the introduction, we project first the processes  $X_{j,i}$  to the real line and then apply a Cramér–von-Mises type test. With a view to the consistency of the testing procedure, we choose a Cramér–von-Mises type test statistic. We note that a Kolmogorov–Smirnov type test statistic, or a test statistic obtained by adding a suitable weight function, can be applied analogously. For the investigation of asymptotic properties of our test, we use that the Cramér–von-Mises distance is connected to von Mises’ type functionals, also known as  $V$ -Statistics. For that reason, the asymptotic properties of a test based on another test statistic have to be treated separately. In our approach, projection is done via the inner product, i.e., we consider  $\langle X_{j,i}, x \rangle$  for  $x \in H$ . We consider all projections  $x$  from a sufficient large projection space  $h \subset H$ . In fact, as explained in Ditzhaus and Gaigall (2018), the distributions of  $X_{1,1}$  and  $X_{1,2}$  coincide if and only if  $\langle X_{1,1}, x \rangle$  and  $\langle X_{1,2}, x \rangle$  have the same distribution for all projections  $x \in h$ , where

$$h = \left\{ \sum_{j=1}^k m_j e_{i_j}; k \in I, i_1, \dots, i_k \in I, i_1 < \dots < i_k, \sum_{j=1}^k m_j^2 = 1 \right\}.$$

This motivates the following test statistic:

$$\text{CvM}_n = \int D_n(x) \mathcal{P}(dx), \tag{2.1}$$

where  $\mathcal{P}$  is a suitable probability measure on the projection space  $h$  and  $D_n(x)$  is the usual two-sample Cramér–von-Mises distance when applying the projection  $x \in h$ .

Let

$$F_{n,i}(x, r) = \frac{1}{n} \sum_{j=1}^n I_{\langle x, X_{j,i} \rangle \leq r}, \quad (x, r) \in H \times \mathbb{R}, \quad i = 1, 2,$$

be the empirical distribution function of the real-valued random variables  $\langle x, X_{1,i} \rangle, \dots, \langle x, X_{n,i} \rangle$ . Then, the related two-sample Cramér–von-Mises distance is given by

$$D_n(x) = n \int [F_{n,1}(x, r) - F_{n,2}(x, r)]^2 \bar{F}_n(x, dr),$$

where  $\bar{F}_n = (F_{n,1} + F_{n,2})/2$ . While more general measures  $\mathcal{P}$  may be considered (compare to Ditzhaus and Gaigall 2018), we focus here to the following specific proposal. It is based on two probability measures  $\nu_1$  and  $\nu_2$  on the index set  $I$  such that  $\nu_j(\{i\}) > 0$  for all  $i \in I$  and  $j = 1, 2$ . These can be chosen arbitrarily in advance. In the case of infinite-dimensional Hilbert space with orthonormal basis  $O = \{e_i; i \in \mathbb{N}\}$ , it is possible to choose Poisson distributions shifted by 1. Otherwise, for Hilbert spaces of finite dimension  $d < \infty$  with orthonormal basis  $O = \{e_i; i = 1, \dots, d\}$ , it is possible to choose the uniform distribution on  $\{1, \dots, d\}$  for  $\nu_1$  and  $\nu_2$ . Anyhow, given these two probability measures, we generate a realization of  $\mathcal{P}$  as follows:

- STEP 1. Generate a realization  $k \in I$  of the distribution  $\nu_1$ .
- STEP 2. Independently of Step 1, generate  $i_1, \dots, i_k \in I$  by  $k$ -times sampling without replacement from the distribution  $\nu_2$ .
- STEP 3. Independently of Steps 1 and 2, generate a realization  $(m_1, \dots, m_k)$  of the uniform distribution on the unit circle in  $\mathbb{R}^k$ .
- STEP 4. Set  $x = \sum_{j=1}^k m_j e_{i_j}$ .

This step-by-step procedure determines  $\mathcal{P}$  uniquely and is considered throughout the remaining paper:

**Assumption 1** Consider  $\mathcal{P}$  given by STEPs 1–4.

As it is stated above, the concrete value of the test statistic can be obtained by Monte Carlo simulation. A possible implementation of the procedure is analogous to the implementation discussed in Section 2.2 in Ditzhaus and Gaigall (2018). In particular, a finite number of realizations from the probability measure  $\mathcal{P}$  can be used for the approximation of the integral with respect to  $\mathcal{P}$  in applications. In this regard, the practical implementation of the procedure is related to the proposal considered in Cuesta-Albertos et al. (2007), where the authors treat a one-sample goodness-of-fit problem for functional data. In detail, the maximum over Kolmogorov–Smirnov type test statistics for a finite and fixed number projections is considered there. The result is a randomized test. In contrast, we treat a marginal homogeneity problem on the basis of a paired-sample and consider the mean over Cramér–von-Mises type test statistics, where our theoretical results, stated next, are available for the theoretical criterion, i.e., if the number of projections tends to infinity and the mean is transferred in the expectation. In fact, by increasing the number of random projections it is possible to reduce the randomness in the testing procedure and to eliminate the randomness in the limit.

## 2.1 Asymptotic theory of the test statistic

For our asymptotic approach, we let  $n \rightarrow \infty$ . It is well known that the Cramér–von-Mises distance  $D_n$  is connected to von Mises’ type functionals, also known as  $V$ -Statistics, which are closely related to  $U$ -Statistics. For a deeper introduction to these kinds of statistics, we refer the reader to Koroljuk and Borovskich (1994) and Serfling (2001). Our statistic  $\text{CvM}_n$  can also be rewritten into a certain  $V$ -Statistic, and thus, the same theory can be applied to obtain the following result.

**Theorem 1** *Let  $\tau_1, \tau_2, \dots$  be a sequence of independent, standard normal random variables. Under the null hypothesis  $\mathcal{H}$  and Assumption 1,*

$$\text{CvM}_n \xrightarrow{d} 3 \sum_{i=1}^{\infty} \lambda_i \tau_i^2 = Z, \quad (2.2)$$

where  $(\lambda_i)_{i \in \mathbb{N}}$  is a sequence of non-negative numbers with  $\sum_{i=1}^{\infty} \lambda_i < \infty$  and  $\lambda_i > 0$  for at least one  $i \in \mathbb{N}$  implying that the distribution function of  $Z$  is continuous and strictly increasing on the non-negative half-line.

In the proofs, more information about  $(\lambda_i)_{i \in \mathbb{N}}$  is provided. In short, they are eigenvalues of a Hilbert–Schmidt operator corresponding to the kernel function of our  $V$ -statistic.

**Theorem 2** *Under the alternative  $\mathcal{K}$  and Assumption 1, our statistic  $\text{CvM}_n$  diverges, i.e.,*

$$\text{CvM}_n \xrightarrow{p} \infty \text{ as } n \rightarrow \infty.$$

In general, the test statistic  $\text{CvM}_n$  is not distribution-free under the null hypothesis, i.e., the distribution depends on the unknown distribution of  $X_1$ . As it can be seen in the proofs, the same applies to  $Z$ . Given that  $\alpha \in (0, 1)$  is the significance level, neither a  $(1 - \alpha)$ -quantile  $c_{n,1-\alpha}$  of  $\text{CvM}_n$  nor the  $(1 - \alpha)$ -quantile  $c_{1-\alpha}$  of  $Z$  is available as critical value in applications. To resolve this problem, we propose the estimation of the quantiles via bootstrapping in the spirit of Efron (1979) and follow the idea in Gaigall (2019), where the usual two-sample Cramér–von-Mises distance is applied to bivariate random vectors with values in  $\mathbb{R}^2$ . Note that under the null hypothesis  $\mathcal{H}$  the expectations  $E[F_{n,1}(x, y)] = E[F_{n,2}(x, y)]$ ,  $(x, y) \in H \times \mathbb{R}$ , coincide, and thus, we can rewrite our test statistic into

$$\begin{aligned} \text{CvM}_n = n \int \int & \{F_{n,1}(x, y) - E[F_{n,1}(x, y)] \\ & + E[F_{n,2}(x, y)] - F_{n,2}(x, y)\}^2 \bar{F}_n(x, dy) \mathcal{P}(dx). \end{aligned}$$

Denote by  $X_{jn}^* = (X_{jn,1}^*, X_{jn,2}^*)$ ,  $j = 1, \dots, n$ , a bootstrap sample from the original observations  $X_j$ ,  $j = 1, \dots, n$ , obtained by  $n$ -times sampling with replacement. Let

$F_{n,i}^*$ ,  $\bar{F}_n^*$  be the bootstrap counterparts of  $F_{n,i}$  and  $\bar{F}_n$ . Clearly, the conditional expectation  $E[F_{n,i}^*(x, y)|(X_j)_{j=1,\dots,n}]$  given the data  $(X_j)_{j=1,\dots,n}$  equals  $F_{n,i}(x, y)$ . Consequently, the bootstrap counterpart of our test statistic is

$$\begin{aligned} \text{CvM}_n^* = n \int \int & \left( F_{n,1}^*(x, y) - F_{n,1}(x, y) \right. \\ & \left. + F_{n,2}(x, y) - F_{n,2}^*(x, y) \right)^2 \bar{F}_n^*(x, dy) \mathcal{P}(dx). \end{aligned}$$

Let  $c_{n,1-\alpha}^*$  be a  $(1-\alpha)$ -quantile of  $\text{CvM}_n^*$  given the original observations  $X_1, \dots, X_n$ . In applications, concrete values of  $c_{n,1-\alpha}^*$  are obtained by Monte Carlo simulation. In the proofs, we show that the bootstrap statistic mimics asymptotically the limiting null distribution under the null hypothesis  $\mathcal{H}$  implying that  $c_{n,1-\alpha}^*$  is an appropriate estimator for the unknown quantile  $c_{n,1-\alpha}$  or  $c_{1-\alpha}$ , while  $\text{CvM}_n^*$  and  $c_{n,1-\alpha}^*$  remain asymptotically finite under general alternatives. This results in an asymptotically exact and consistent bootstrap test  $\varphi_n^* = I_{\text{CvM}_n > c_{n,1-\alpha}^*}$ .

**Theorem 3** Suppose that Assumption 1 holds. Then, as  $n \rightarrow \infty$ , we have

$$E[\varphi_n^*] = P(\text{CvM}_n > c_{n,1-\alpha}^*) \rightarrow \alpha I_{\mathcal{H}} + I_{\mathcal{K}},$$

where  $I.$  denotes the indicator function.

### 3 Simulations

Remembering that our test is suitable for random variables  $X_{i,j}$ ,  $i = 1, 2, j = 1, \dots, n$ , with values in a general separate Hilbert space, we consider the separable Hilbert space  $H$  consisting of all measurable and square-integrable functions on the unit interval  $[0, 1]$ . This space is endowed with the usual inner product  $\langle \cdot, \cdot \rangle$  and the normalized Legendre polynomials build a corresponding orthonormal basis  $O = \{e_i; i \in I\}$ ,  $I = \mathbb{N}$ . We obtain our test statistic (2.1) by Monte Carlo simulation based on 500 replications following STEP 1–4 from Sect. 2. Thereby, we choose in STEP 1 and STEP 2 a standard Poisson distribution shifted by 1, i.e., the distribution of  $N + 1$  for  $N \sim \text{Pois}(1)$ . In our simulations, the stochastic processes  $X_{j,i} = (X_{j,i}(t); t \in [0, 1])$  have the form

$$X_{j,i}(t) = a_i B_{j,i}(t) + b_i t(t-1), \quad t \in [0, 1], \quad i = 1, 2, \quad j = 1, \dots, n$$

for parameters  $a_i \in \mathbb{R} \setminus \{0\}$  and  $b_i \in \mathbb{R}$  and independent bivariate Brownian bridges  $B_j = (B_{j,1}, B_{j,2})$  on  $[0, 1]$ ,  $j = 1, \dots, n$ , with covariance structure

$$\text{Cov}(B_{j,1}(s), B_{j,2}(t)) = r(\min(s, t) - st), \quad s, t \in [0, 1], \quad j = 1, \dots, n$$

for a dependence parameter  $r \in [0, 1]$ . Each simulation is based on 5000 simulation runs. To obtain the critical values in the bootstrap procedure, we use Monte Carlo

**Table 1** Empirical sizes

$X_{j,1}(t)$	$X_{j,2}(t)$	$r$	$\alpha = 5\%$		$\alpha = 10\%$	
			$n = 20$	$n = 50$	$n = 20$	$n = 50$
$B_{j,1}(t)$	$B_{j,2}(t)$	0	5.7	5.1	10.3	9.5
		0.25	4.8	4.8	9.8	8.7
		0.5	3.9	4.6	8.4	9.0
$1.5B_{j,1}(t)$	$1.5B_{j,2}(t)$	0	5.5	5.3	10.4	10.1
		0.25	5.0	4.6	10.2	8.9
		0.5	3.6	5.1	8.1	10.1
$2B_{j,1}(t)$	$2B_{j,2}(t)$	0	5.0	5.8	10.6	11.3
		0.25	4.9	4.7	9.3	9.4
		0.5	3.5	4.4	7.4	9.2
$2.5B_{j,1}(t)$	$2.5B_{j,2}(t)$	0	6.1	5.2	11.6	10.0
		0.25	4.7	4.7	9.6	10.0
		0.5	3.6	4.0	8.6	8.3
$B_{j,1}(t) + 0.5t(1-t)$	$B_{j,2}(t) + 0.5t(1-t)$	0	5.2	5.1	10.1	10.3
		0.25	5.1	5.4	10.2	10.3
		0.5	3.6	4.5	8.0	8.6
$B_{j,1}(t) + t(1-t)$	$B_{j,2}(t) + t(1-t)$	0	5.7	4.8	10.6	9.2
		0.25	4.6	4.7	9.6	9.9
		0.5	3.5	4.2	7.4	8.7
$B_{j,1}(t) + 1.5t(1-t)$	$B_{j,2}(t) + 1.5t(1-t)$	0	5.5	5.8	11.0	11.0
		0.25	4.4	5.3	9.4	10.4
		0.5	3.8	4.0	7.5	8.5
$B_{j,1}(t) + 2t(1-t)$	$B_{j,2}(t) + 2t(1-t)$	0	5.5	5.1	10.6	10.3
		0.25	4.7	4.8	9.0	9.7
		0.5	3.4	4.8	8.1	9.3

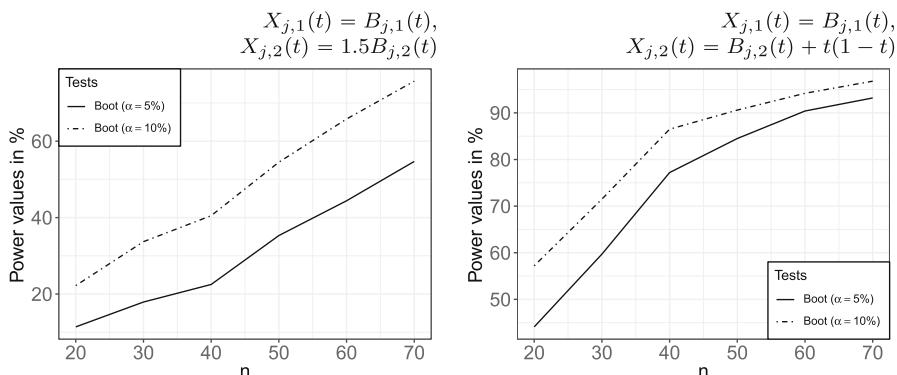
simulation based on 999 replications. Empirical sizes of the bootstrap test are displayed in Table 1. The simulations are conducted for parameters  $r \in \{0, 0.25, 0.5\}$ ,  $a_i \in \{1, 1.5, 2, 2.5\}$ , and  $b_i \in \{0, 0.5, 1, 1.5, 2\}$ , the sample sizes  $n \in \{20, 50\}$ , and significance levels  $\alpha \in \{5\%, 10\%\}$ . The empirical sizes are in almost all cases in a reasonable range around the nominal level  $\alpha$ . A systematic exception from this observations is the sizes of the bootstrap approach under the strong dependence setting ( $r = 0.5$ ) and the smaller sample size setup ( $n = 20$ ). In this case, the bootstrap decisions are rather conservative with corresponding empirical sizes from 3.4 to 3.9% with an average of 3.6% for  $\alpha = 5\%$  as well as values from 7.4 up to 8.6% and an average of 7.4% for  $\alpha = 10\%$ . However, increasing the sample sizes to  $n = 50$  improves the type-I error rate control; now we can observe in average a value of 4.5% for  $\alpha = 5\%$  and of 9.0% for  $\alpha = 10\%$ , respectively, under the strong dependence setting.

To complement this study for empirical sizes, we conduct additional simulations, now based on 1000 runs, under different alternative settings, where we vary the

**Table 2** Empirical power values for  $n = 20$ 

$X_{j,1}(t)$	$X_{j,2}(t)$	$r$	$\alpha = 5\%$	$\alpha = 10\%$
$B_{j,1}(t)$	$1.5B_{j,2}(t)$	0	10.7	19.9
		0.25	10.0	21.0
		0.5	12.0	25.2
$B_{j,1}(t)$	$2B_{j,2}(t)$	0	22.6	42.9
		0.25	36.3	58.7
		0.5	50.6	73.1
$B_{j,1}(t)$	$2.5B_{j,2}(t)$	0	47.2	74.3
		0.25	62.5	85.0
		0.5	84.7	94.3
$B_{j,1}(t)$	$B_{j,2}(t) + 0.5t(1-t)$	0	14.1	21.2
		0.25	13.0	21.6
		0.5	17.0	24.4
$B_{j,1}(t)$	$B_{j,2}(t) + t(1-t)$	0	35.2	46.9
		0.25	45.1	57.0
		0.5	52.8	66.4
$B_{j,1}(t)$	$B_{j,2}(t) + 1.5t(1-t)$	0	68.7	78.8
		0.25	78.7	86.4
		0.5	87.8	93.5
$B_{j,1}(t)$	$B_{j,2}(t) + 2t(1-t)$	0	92.6	96.0
		0.25	95.8	97.4
		0.5	98.9	99.4

parameters  $a_i \in \{1, 1.5, 2, 2.5\}$  and  $b_i \in \{0, 0.5, 1, 1.5, 2\}$  for the second component  $X_{j,2}$  while keeping them fixed as  $(a_i, b_i) = (1, 0)$  for the first one. In Table 2, the results for the small sample size setting ( $n = 20$ ) are displayed. Here, the power values increase for growing parameters  $a_i$  and  $b_i$ , respectively, as well as for growing dependence parameter  $r$ . It is seen that the power increases as the dependence parameter  $r$  increases in the most cases. A possible reason is that an increasing dependence parameter  $r$  reduces the differences of the first and the second components of the bivariate Brownian bridges  $B_j = (B_{j,1}, B_{j,2})$  on  $[0, 1]$ ,  $j = 1, \dots, n$ , such that the deterministic factors  $a_2 \in \{1, 1.5, 2, 2.5\}$  or the deterministic additive terms  $b_2 t(t-1)$ ,  $t \in [0, 1]$ ,  $b_2 \in \{0, 0.5, 1, 1.5, 2\}$ , in the second components  $X_{j,2}$ ,  $j = 1, \dots, n$ , cause clear differences in the components  $X_{j,1}$  and  $X_{j,2}$ ,  $j = 1, \dots, n$ , of the bivariate data. For two specific alternatives  $(a_i, b_i) = (1.5, 0), (1, 1)$  under moderate ( $r = 0.25$ ) dependence, the power behavior is, moreover, investigated under growing sample sizes  $n \in \{20, 30, \dots, 70\}$ , see Fig. 1. The resulting power curves have a clear positive slope indicating a reasonable power behavior for increasing sample sizes.



**Fig. 1** Power values for increasing sample sizes under moderate ( $r = 0.25$ ) dependence

#### 4 Applications to stock market returns

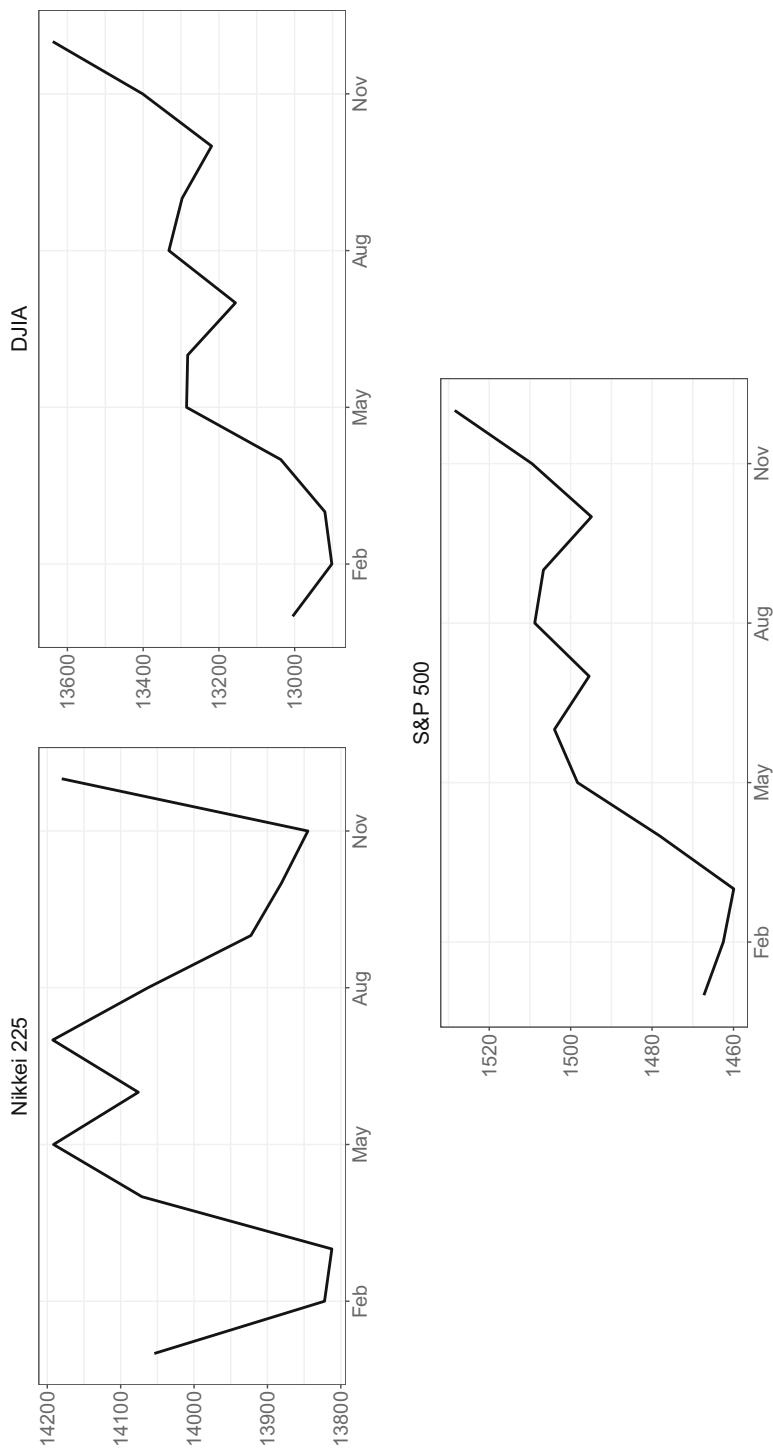
In a possible application, the observations are obtained from stock market returns. To be concrete, we consider two stock price processes and a time period  $[0, T_0]$ , where  $T_0 \in (0, \infty)$  is a time horizon. The stock price processes are denoted by

$$Y_i = (Y_i(t); t \in [0, T_0]), \quad i = 1, 2,$$

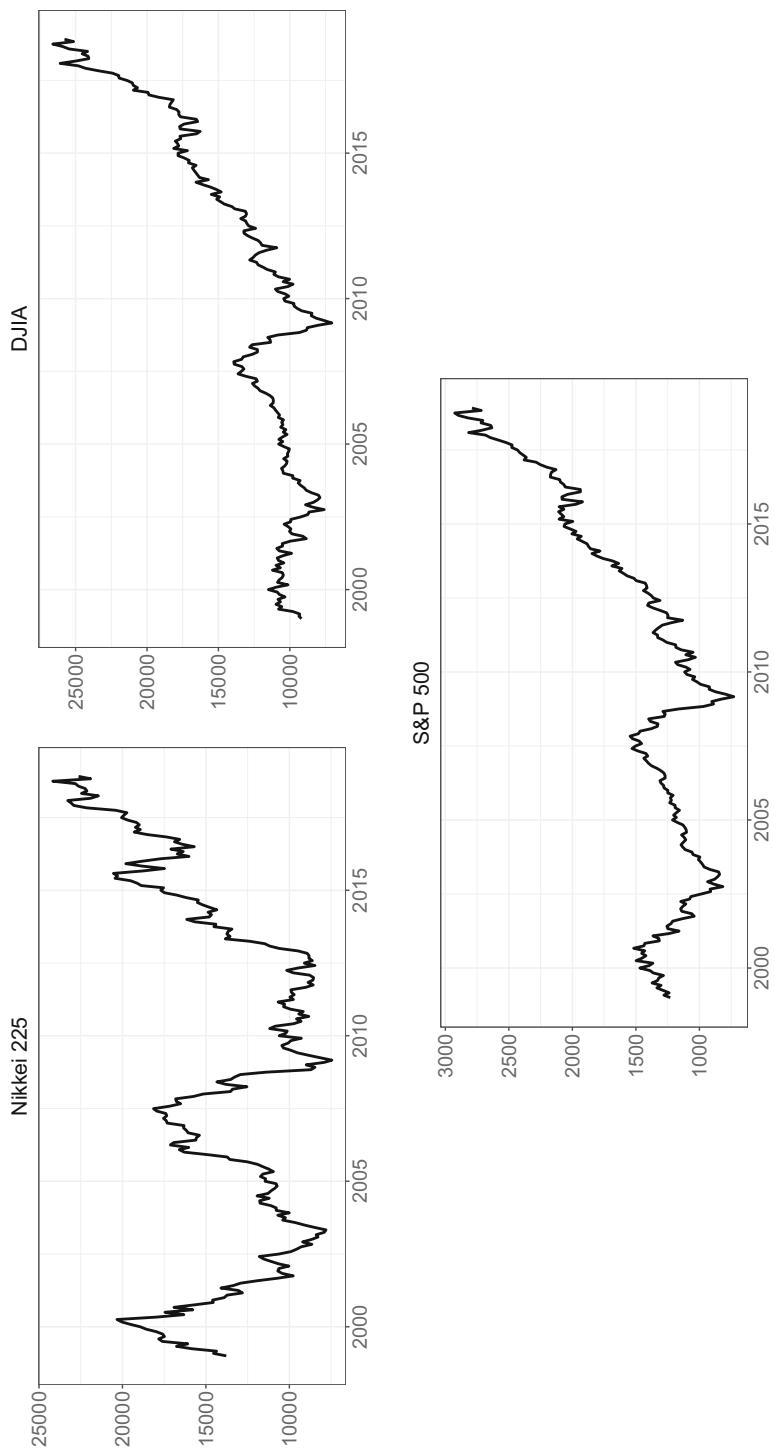
where  $Y_i$  is a random variable that takes values in the space of all measurable and square-integrable real-valued functions on  $[0, T_0]$ . Additional structural assumptions on the underlying stochastic process for the stock prices are required. In the classical models for stock prices, i.e., the exponential Lévy model, Black–Scholes model, and Merton model, the structural assumptions mentioned are independence and stationarity of the increments of a Lévy process. Seasonality effects, which are specific trends during certain periods of the stock price processes, can disturb the stationarity assumption. Figure 2 shows the mean monthly indices (open) of the Nikkei Stock Average (Nikkei 225), Dow Jones Industrial Average (DJIA), and Standard & Poor's 500 (S&P 500) from 01/01/1999 to 01/01/2019, where seasonality effects are clearly seen. We will tackle this problem by splitting up the time horizon into  $n$  equal-sized time intervals to obtain  $n$  observations for each stock price process. To be more specific, let  $T_0 = nT$  for  $T \in (0, \infty)$  then we consider the time periods  $[0, T], \dots, [(n-1)T, nT]$  and our observations are the log-returns during these periods

$$X_{j,i}(t) = \log \frac{Y_i(t + (j-1)T)}{Y_i((j-1)T)}, \quad t \in [0, T], \quad i = 1, 2, \quad j = 1, \dots, n,$$

which are themselves measurable and square-integrable real-valued functions on  $[0, T]$ . In particular, we have the specific space  $H = L^2[0, T]$  containing all measurable and square-integrable real-valued functions on the interval  $[0, T]$  of length  $T \in (0, \infty)$  and equipped with the usual inner product  $\langle f, g \rangle = \int_0^T f(x)g(x) dx$ ,  $f, g \in H$ . A corresponding orthonormal basis is given by normalized Legendre poly-



**Fig. 2** Mean monthly values from 01/01/1999 to 01/01/2000



**Fig. 3** Monthly values from 01/01/1999 to 01/01/2019

**Table 3** Empirical  $p$ -values of the test in percent

$X_{1,j}(t)$	$X_{2,j}(t)$	Monthly values	Weekly values	Daily values
DJIA	S&P 500	73.6	57.7	65.7
Nikkei 225	S&P 500	5.5	12.3	6.5
Nikkei 225	DJIA	5.6	11.8	10.9

nomials. The well-known models for stock prices mentioned imply an independent and identical distributed structure of the increments  $X_{j,i}$ . In models with time-dependent or stochastic volatility (volatility clustering for instance), this structure is may be violated. In fact, the theory of  $U$ -Statistics is well developed and covers also cases where the independent and identical distributed data structure is disturbed, see Chapters 2.3 and 2.4 of Lee (1990). We point out that our results are derived by application of the theory of  $U$ -Statistics in the independent and identical distributed data case, but it should be possible to extend and modify the approach in more general situations, however, under suitable regularity conditions.

In what follows, we demonstrate the application of our test to the values (open) of the stock market indices Nikkei Stock Average of Japan, Dow Jones Industrial Average of the US, and Standard & Poor's 500 of the US for the time period 01/01/1999 to 01/01/2019. For the demonstration of how the test works in applications, we consider different frequencies of the data that are monthly values, weekly values, and daily values, and a linear interpolation. The resulting time series (for the monthly values) are presented in Fig. 3 and can be seen as square-integrable functions on the interval  $[0, T_0]$  for  $T_0 = 20$  (years). To cover seasonality effects indicated in Fig. 2, we split the time horizon of 20 years into 20 subintervals each representing one year, i.e.,  $T = 1$  and  $n = 20$ . We apply our method to do pairwise comparisons of the indices, where the test statistic is again approximated by 500 random projections following STEP 1–4 and the shifted Poisson distribution is used in STEP 1 and STEP 2 as in Sect. 3. The resulting  $p$ -values for the bootstrap approach are displayed in Table 3 for 5000 resampling iterations, respectively. Since DJIA and S&P 500 reflect both the US market, it is not surprising that the test leads to a very high  $p$ -value and, thus, does not reject the null hypothesis. Comparisons of each of these US indices with the Japanese Nikkei 225 lead to  $p$ -values much lower, but even above the typical used 5%-benchmark. The results are in line with the first graphical impression, which we get in Fig. 3.

**Acknowledgements** The authors are grateful to the editor, the associate editor, and the two referees for their comments that substantially improved the paper's quality. Moreover, the authors gratefully acknowledge the computing time provided on the Linux HPC cluster at TU Dortmund (LiDO3), partially funded by the *Deutsche Forschungsgemeinschaft* as project 271512359.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## A Proofs

We prove Theorem 1 in a more general way. Instead, of CvM<sub>n</sub> we consider

$$S_n = n \int \int ([F_{1,n}(x, y) - F_1(x, y)] - [F_{2,n}(x, y) - F_2(x, y)])^2 \bar{F}_n(x, dy) \mathcal{P}(dx),$$

where

$$F_i(x, y) = P(\langle x, X_{1,i} \rangle \leq y), \quad (x, y) \in H \times \mathbb{R}, \quad i = 1, 2.$$

**Theorem 4** Suppose that Assumption 1 holds. Let  $\tau_1, \tau_2, \dots$  be a sequence of independent, standard normal random variables. Under the null hypothesis  $\mathcal{H}$  as well as under any alternative, we have

$$S_n \xrightarrow{d} 3 \sum_{i=1}^{\infty} \lambda_i \tau_i^2 = Z, \quad (\text{A.1})$$

where  $(\lambda_i)_{i \in \mathbb{N}}$  is a sequence of non-negative numbers with  $\sum_{i=1}^{\infty} \lambda_i < \infty$  and  $\lambda_i > 0$  for at least one  $i \in \mathbb{N}$  implying that the distribution function of  $Z$  is continuous and strictly increasing on the non-negative half-line.

**Proof** Let  $x_j = (x_{j,1}, x_{j,2}) \in H^2$ ,  $j \in \mathbb{N}$ . We introduce the asymmetric kernel  $f$  given, for every  $x_1, x_2, x_3 \in H^2$ , by

$$\begin{aligned} & f(x_1, x_2, x_3) \\ &= \frac{1}{2} \sum_{i=1}^2 \int [I_{\langle x_{1,1} - x_{3,i}, u \rangle \leq 0} - F_1(u, \langle x_{3,i}, u \rangle) - I_{\langle x_{1,2} - x_{3,i}, u \rangle \leq 0} + F_2(u, \langle x_{3,i}, u \rangle)] \\ &\quad \times [I_{\langle x_{2,1} - x_{3,i}, u \rangle \leq 0} - F_1(u, \langle x_{3,i}, u \rangle) - I_{\langle x_{2,2} - x_{3,i}, u \rangle \leq 0} + F_2(u, \langle x_{3,i}, u \rangle)] \mathcal{P}(du) \end{aligned} \quad (\text{A.2})$$

as well as its symmetric version  $\phi$  defined, for every  $x_1, x_2, x_3 \in H^2$ , by

$$\phi(x_1, x_2, x_3) = \frac{1}{3} \left( f(x_1, x_2, x_3) + f(x_2, x_3, x_1) + f(x_1, x_3, x_2) \right).$$

Clearly,

$$\begin{aligned} S_n &= \frac{1}{n^2} \sum_{i,j,k=1}^n f(X_i, X_j, X_k) = \frac{1}{n^2} \sum_{i,j,k=1}^n \phi(X_i, X_j, X_k) \\ &= \frac{1}{n^2} \sum_{i=1}^n \phi(X_i, X_i, X_i) + \frac{3}{n^2} \sum_{i,j=1; j \neq i}^n \phi(X_i, X_i, X_j) \\ &\quad + \frac{1}{n^2} \sum_{1 \leq i \neq j \neq k \leq n} \phi(X_i, X_j, X_k) \\ &=: U_{1,n} + U_{2,n} + U_{3,n}. \end{aligned} \tag{A.3}$$

By the classical strong law of large numbers as well as by its extension to  $U$ -statistics of higher degrees (e.g., Koroljuk and Borovskich 1994, Theorem 3.1.1)

$$\begin{aligned} nU_{1,n} &\xrightarrow{a.s.} E[\phi(X_1, X_1, X_1)] \in [-4, 4], \text{ and} \\ U_{2,n} &\xrightarrow{a.s.} 3 \cdot E[\phi(X_1, X_1, X_2)] \in [-4, 4]. \end{aligned}$$

Thus, the first summand in (A.3) vanishes and the second one converges to a constant, which we will specify later. For the third summand, we apply results about weak convergence of  $U$ -statistics. Following Koroljuk and Borovskich (1994), we first determine the rank of degeneracy of  $U_{3,n}$ . For that purpose, we introduce

$$g_1(x_1) = E[\phi(x_1, X_2, X_3)] \text{ and } g_2(x_1, x_2) = E[\phi(x_1, x_2, X_3)] - g_1(x_1) - g_1(x_2)$$

for all  $x_1, x_2 \in H^2$ . It is easy to check that for every  $x_1, x_2, x_3 \in H^2$  we have

$$E[f(X_1, x_2, x_3)] = E[f(x_1, X_2, x_3)] = 0. \tag{A.4}$$

Thus, we obtain for all  $x_1 \in H^2$

$$g_1(x_1) = E[\phi(x_1, X_2, X_3)] = E\{E[\phi(x_1, X_2, X_3)|X_3]\} = 0 = E[\phi(X_1, X_2, X_3)]. \tag{A.5}$$

The function  $(x_1, x_2) \mapsto g_2(x_1, x_2) = E[\phi(x_1, x_2, X_3)] = E[f(x_1, x_2, X_3)]/3$  is not constant equal to 0 with probability one, as explained at the end of the proof. Summing up, we can conclude that the rank of the  $U$ -statistic  $U_{3,n}$  is  $r = 2$  (Koroljuk and Borovskich 1994, p. 24). Consequently, applying Corollary 4.4.2 of Koroljuk and

Borovskich (1994) yields

$$U_{3,n} \xrightarrow{d} 3 \sum_{i=1}^{\infty} \lambda_i (\tau_i^2 - 1),$$

where  $\tau_1, \tau_2, \dots$  is a sequence of independent, standard normal random variables and  $(\lambda_i)_{i \in \mathbb{N}}$  is the collection of eigenvalues of the Hilbert–Schmidt operator  $h(\cdot) \mapsto E[\tilde{g}_2(X_1, X_2)h(X_2)|X_1]$  with respective eigenfunctions  $(\varphi_i)_{i \in \mathbb{N}}$ . By Lemma 1, see below,  $g_2$  is a degenerated and bounded Mercer kernel. In analogy to the argumentation of Leucht and Neumann (2013), see their proof of Theorem 2.1, an extension of Mercer’s Theorem Sun (2005) yields that for all  $x_1, x_2 \in H^2$

$$g_2(x_1, x_2) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(x) \varphi_i(y). \quad (\text{A.6})$$

The sum’s convergence does not only hold pointwisely but even absolutely and uniformly on every compact subset of the Cartesian square of the support of  $P^{X_1}$ . Moreover,

$$\lambda_i \geq 0 \text{ for all } i \in \mathbb{N} \text{ and } \sum_{i=1}^{\infty} \lambda_i = E(g_2(X_1, X_1)) = E(\phi(X_1, X_1, X_3)) < \infty. \quad (\text{A.7})$$

The latter is exactly the limit of  $U_{2,n}$ , which remained to be specified. Summing up, the stated result about the weak convergence of  $S_n$  follows when  $g_2$  is, as stated above, indeed not a constant function equal to 0. The latter as well as  $\lambda_i > 0$  for at least one  $i \in \mathbb{N}$  hold if  $\sum_{i=1}^{\infty} \lambda_i = E(g_2(X_1, X_1))$  is positive, which we discuss now. Note that  $E(g_2(X_1, X_1))$  is non-negative. Contrary to our claim, we assume now that  $E(g_2(X_1, X_1)) = 0$ . Then, either  $g_2$  is a constant function equal to 0, and thus, the rank of  $U_{3,n}$  is  $r = 3$  or the rank of the  $U_{3,n}$  is still  $r = 2$ . If  $r = 2$  is still true, then by our above argumentation it follows that  $\sum_{i=1}^{\infty} \lambda_i = 0$  implying  $\lambda_i = 0$  for all  $i \in \mathbb{N}$  and, consequently,  $S_n \xrightarrow{P} 0$ . If  $r = 3$ , then  $S_n \xrightarrow{P} 0$  follows from Theorem 4.4.2 of Koroljuk and Borovskich (1994). Thus, either way  $S_n \xrightarrow{P} 0$ . Our assumptions on the joint distribution of  $\langle X_{1,1}, e_i \rangle$  and  $\langle X_{1,2}, e_i \rangle$ ,  $i \in I$ , ensure that Assumption 1 and Assumption 2 in Gaigall (2019) are satisfied, and it follows from Theorem 1 and Theorem 3 in Gaigall (2019) that  $D_n(e_1) \xrightarrow{d} S$  as  $n \rightarrow \infty$ , where  $S$  is a real-valued random variable and not constantly zero. In all, from

$$S_n \geq D_n(e_1)v_1(\{1\}) \xrightarrow{d} S \cdot v_1(\{1\}) \text{ as } n \rightarrow \infty$$

together with  $v_1(\{1\}) > 0$  we obtain a contradiction. Thus,  $E(g_2(X_1, X_1)) > 0$  is true and, in particular,  $g_2$  is not a constant function equal to 0 and  $\lambda_i > 0$  for at least one  $i \in \mathbb{N}$ .  $\square$

**Lemma 1** Suppose that Assumption 1 holds. The function  $g_2$  is a degenerated and bounded Mercer kernel, i.e., it is continuous, symmetric and positive semidefinite.

**Proof** It is easy to see that  $g_2$  is bounded and symmetric. The degeneracy follows immediately from (A.4). For arbitrary  $k \in \mathbb{N}$  let  $c_1, \dots, c_k \in \mathbb{R}$ . Then,

$$\sum_{i,j=1}^k c_i c_j g_2(x_i, x_j) = \frac{1}{2} \sum_{\ell=1}^2 \mathbb{E} \left[ \int \left\{ \sum_{i=1}^k c_i \left[ I_{\langle x_i, x \rangle \leq \langle X_{3,\ell}, x \rangle} - F_1(x, \langle X_{3,\ell}, x \rangle) \right. \right. \right. \\ \left. \left. \left. - I_{\langle x_i, x \rangle \leq \langle X_{3,\ell}, x \rangle} + F_2(x, \langle X_{3,\ell}, x \rangle) \right] \right\}^2 \mathcal{P}(dx) \right] \geq 0.$$

Hence,  $g_2$  is positive semidefinite. For the continuity proof, let  $(x_{1n})_{n \in \mathbb{N}}$  and  $(x_{2n})_{n \in \mathbb{N}}$  be sequences in  $H^2$  such that  $\lim_{n \rightarrow \infty} x_{jn} = x_j \in H^2$ ,  $j = 1, 2$ . By Lemma 3.1 of Ditzhaus and Gaigall (2018)

$$I_{\langle x, x_{3,\ell} - y \rangle \neq 0} = 1 \quad (\text{A.8})$$

for  $\mathcal{P} \times P^{X_{1,\ell}}$ -almost all  $(x, x_{3,\ell})$  and every  $y \in H$ . This and the continuity of the inner product imply for  $\mathcal{P}$ -almost all  $x$  and every  $m, \ell \in \{1, 2\}$  that

$$\lim_{n \rightarrow \infty} I_{\langle x, x_{jn,m} \rangle \leq \langle x, X_{3,\ell} \rangle} = I_{\langle x, x_{j,m} \rangle \leq \langle x, X_{3,\ell} \rangle} \text{ with probability one.}$$

Consequently,  $g_2(x_{1n}, x_{2n})$  converges to  $g_2(x_1, x_2)$ .  $\square$

**Proof of Theorem 1** Since  $F_1 = F_2$  and, thus,  $S_n = \text{CvM}_n$  under the null hypotheses, the statement follows immediately from Theorem 4.  $\square$

**Proof of Theorem 2** First, observe that

$$n^{-1} \text{CvM}_n = n^{-1} S_n \\ + 2 \int \int [F_{n,1}(x, y) - F_1(x, y) - F_{n,2}(x, y) \\ + F_2(x, y)][F_1(x, y) - F_2(x, y)] \bar{F}_n(x, dy) \mathcal{P}(dx) \\ + \int \int [F_1(x, y) - F_2(x, y)]^2 \bar{F}_n(x, dy) \mathcal{P}(dx).$$

By Theorem 4,  $n^{-1} S_n$  converges in probability to 0. By the Cauchy–Schwarz inequality, the absolute value of the second summand is bounded from above by  $2\sqrt{n^{-1} S_n}$ . In particular, the second summand vanishes in probability as well. The third summand can be rewritten as

$$\frac{1}{2n} \sum_{j=1}^n \sum_{k=1}^2 \int [F_1(x, \langle x, X_{j,k} \rangle) - F_2(x, \langle x, X_{j,k} \rangle)]^2 \mathcal{P}(dx) = \frac{1}{2n} \sum_{j=1}^n g(X_j),$$

for an appropriate function  $g$ . By the strong law, this sum converges almost surely to

$$\begin{aligned} & \frac{1}{2} \sum_{k=1}^2 \int E\left[\left\{F_1(x, \langle x, X_{3,k} \rangle) - F_2(x, \langle x, X_{3,k} \rangle)\right\}^2\right] \mathcal{P}(dx) \\ &= \frac{1}{2} \sum_{k=1}^2 \int [F_1(x, y) - F_2(x, y)]^2 Q_k(dx, y), \end{aligned} \quad (\text{A.9})$$

where  $Q_k$  is the distribution introduced at the proof's end of Theorem 4. In analogy to the argumentation of Ditzhaus and Gaigall (2018) in the proof for their Theorem 3.2, we can conclude that each summand from (A.9) is strictly positive. Finally, we obtain

$$n^{-1} \text{CvM}_n \xrightarrow{p} \frac{1}{2} \sum_{k=1}^2 \int [F_1(x, y) - F_2(x, y)]^2 Q_k(dx, y) > 0.$$

□

**Proof of Theorem 3** From now on, we suppose that the data  $X_1, \dots, X_n$  are fixed and we operate on the conditional space. In particular, we can treat  $F_{n,i}$ ,  $i = 1, 2$ , as a non-random function, which converges, without loss of generality, pointwisely to  $F_i$ . First, we remark that the distribution of the bootstrap sample depends on the sample size. Moreover, the distribution of  $X_{in}^*$  converges weakly to the distribution of  $X_i$ . Thus, by Theorem 1.10.4 of Van der Vaart and Wellner (1996) we can assume without loss of generality that  $X_{in}^*$  converges to  $X'_i$  for all  $i \in \mathbb{N}$  with probability one, where  $X'_i$  has the same distribution as  $X_i$ , and that  $X'_r$  is independent from  $X_{1n}^*, \dots, X_{(r-1)n}^*, X_{(r+1)n}^*, \dots$  for all  $r \in \mathbb{N}$ . Now, define

$$\begin{aligned} f_n^*(x_1, x_2, x_3) \\ = \frac{1}{2} \sum_{k=1}^2 \int \left[ I_{\langle x, x_{1,1} - x_{3,k} \rangle \leq 0} - F_{n,1}(x, \langle x, x_{3,k} \rangle) - I_{\langle x, x_{1,2} - x_{3,k} \rangle \leq 0} + F_{n,2}(x, \langle x, x_{3,k} \rangle) \right] \\ \times \left[ I_{\langle x, x_{2,1} - x_{3,k} \rangle \leq 0} - F_{n,1}(x, \langle x, x_{3,k} \rangle) - I_{\langle x, x_{2,2} - x_{3,k} \rangle \leq 0} + F_{n,2}(x, \langle x, x_{3,k} \rangle) \right] \mathcal{P}(dx). \end{aligned}$$

Then, we have

$$\text{CvM}_n^* = \frac{1}{n^2} \sum_{i,j,k=1}^n f_n^*(X_{in}^*, X_{jn}^*, X_{kn}^*).$$

Define

$$S'_n = \frac{1}{n^2} \sum_{i,j,k=1}^n f(X'_i, X'_j, X'_k) \text{ and } \kappa_{n,i,j,k} = f_n^*(X_{in}^*, X_{jn}^*, X_{kn}^*) - f(X'_i, X'_j, X'_k),$$

where  $f$  is defined in (A.2). By Theorem 4,  $S'_n$  converges in distribution to  $Z$ . Combining this and

$$E[(\text{CvM}_n^* - S'_n)^2] = n^{-4} \sum_{i_1, \dots, i_6=1}^n E[\kappa_{n,i_1,i_2,i_3} \kappa_{n,i_4,i_5,i_6}] \rightarrow 0 \text{ as } n \rightarrow \infty \quad (\text{A.10})$$

under the null hypothesis, where the proof of (A.10) is given later, yields conditional convergence

$$\text{CvM}_n^* \xrightarrow{d} Z \text{ as } n \rightarrow \infty$$

given the observations  $X_1, \dots, X_n$  under  $\mathcal{H}$  for  $Z$  from Theorem 1. Consequently, we can deduce that  $c_{n,1-\alpha}^* \xrightarrow{P} c_{1-\alpha}$  under  $\mathcal{H}$  and, in particular, the statement under  $\mathcal{H}$  follows, compared to Lemma 1 of Janssen and Pauls (2003). For the statement under the alternative, it remains to show that  $(\text{CvM}_n^*)_{n \in \mathbb{N}}$  is a tight sequence of real-valued random variables, compared to Theorem 7 of Janssen and Pauls (2003), i.e., we have to show

$$\limsup_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P(|\text{CvM}_n^*| \geq K) = 0.$$

In contrast with (A.10), it remains now to show

$$\limsup_{n \rightarrow \infty} E[(\text{CvM}_n^* - S'_n)^2] \leq M < \infty. \quad (\text{A.11})$$

To sum up, we need to verify (A.10) for the statement under  $\mathcal{H}$  and (A.11) for the statement under  $\mathcal{K}$ . For this purpose, we divide the corresponding sum in (A.10) into the following six sums

$$I_{n,p} = n^{-4} \sum_{i_1, \dots, i_6=1}^n E[\kappa_{n,i_1,i_2,i_3} \kappa_{n,i_4,i_5,i_6}] I_{|\{i_1, \dots, i_6\}|=p}, \quad p = 1, \dots, 6.$$

First, we will prove that  $I_{n,p}$  converges to 0 for  $p \in \{1, 2, 3, 5, 6\}$  independently whether the null hypothesis or the alternative is true. In the end, we discuss  $I_{n,4}$  separately under the null hypothesis and the alternative. For all considerations below, remind that  $\kappa_{n,i,j,m}$  is uniformly bounded by 8. As a first consequence of this, we obtain that as  $n \rightarrow \infty$

$$I_{n,1} + I_{2,n} + I_{3,n} \leq \frac{8^2}{n^4} [n + (2^6 - 1)n(n-1) + (3^6 - 2^6)n(n-1)(n-2)] \rightarrow 0.$$

Let us have now a look on all summands with  $|\{i_1, \dots, i_6\}| = 5$ . Let  $r$  be the number that appears twice within the indices  $i_1, \dots, i_6$ . Observe that (A.4) also holds for the bootstrap sample, i.e.,

$$E[f_n^*(X_{1n}^*, x_2, x_3)] = E[f_n^*(x_1, X_{2n}^*, x_3)] = 0.$$

Combining this and (A.4) yields  $E[\kappa_{n,i,j,k} | X_r', X_r^*] = 0$  with probability one whenever  $|\{i, j, k\}| = 3$ . Consequently,

$$E[\kappa_{n,i_1,i_2,i_3} \kappa_{n,i_4,i_5,i_6}] = E\{E[\kappa_{n,i_1,i_2,i_3} | X_r', X_{r,*}] E[\kappa_{n,i_4,i_5,i_6} | X_r', X_{r,*}]\} = 0$$

Clearly, the same can be shown for the case  $|\{i_1, \dots, i_6\}| = 6$ . Hence,  $I_{n,5} + I_{n,6} = 0$ .

Now, we consider  $I_{n,4}$ . Due to the boundedness of  $\kappa_{n,i,j,m}$ , we always obtain

$$I_{n,4} \leq \frac{8^2}{n^4} (4^6 - 3^6) n(n-1)(n-2)(n-3) \leq 2^{20}.$$

From this and the previous considerations, we can conclude (A.11). Now, let us suppose that the null hypothesis is true. Due to symmetry  $\kappa_{n,i,j,k} = \kappa_{n,j,i,k}$  we get

$$\begin{aligned} I_{n,4} &\leq \frac{8}{n^4} (4^6 - 3^6) n(n-1)(n-2)(n-3) \\ &\quad \max \left\{ E[|\kappa_{n,1,1,2}|] + E[|\kappa_{n,1,2,2}|] + E[|\kappa_{n,1,3,2}|] \right\}. \end{aligned}$$

Consequently, it is sufficient for (A.10) to prove

$$\lim_{n \rightarrow \infty} \kappa_{n,1,j,2} = 0 \text{ in probability} \quad (\text{A.12})$$

for  $j = 1, 2, 3$ . From the continuity of the inner product, (A.8), the underlying independence and the convergence of  $X_{1n}^*$ ,  $X_{2n}^*$ ,  $X_{3n}^*$  we obtain that with probability one

$$\lim_{n \rightarrow \infty} I_{\langle x, X_{rn,k}^* \rangle \leq \langle x, X_{2n,l}^* \rangle} = I_{\langle x, X'_{r,k} \rangle \leq \langle x, X'_{2,l} \rangle} \text{ for } \mathcal{P}\text{-almost all } x, \quad (\text{A.13})$$

every  $r \in \{1, 3\}$  and  $k, l \in \{1, 2\}$ . Analogously, we have

$$\begin{aligned} F_{n,k}(x, \langle x, X_{2n,l}^* \rangle) &= E \left[ I_{\langle x, X_{1n,k}^* \rangle \leq \langle x, X_{2n,l}^* \rangle} | X_{2n,l}^* \right] \\ &\xrightarrow{a.s.} E \left[ I_{\langle x, X'_{1,k} \rangle \leq \langle x, X'_{2,l} \rangle} | X'_{2,l} \right] = F_k(x, \langle x, X'_{2,l} \rangle) \text{ as } n \rightarrow \infty. \end{aligned} \quad (\text{A.14})$$

Combining (A.13) and (A.14) shows (A.12) for the case  $j \in \{1, 3\}$ . The reason why we need to be more careful in the case  $j = 2$  is that, in general, (A.13) is false if  $r = j = 2$ . However, the integrals appearing in the limiting  $f(X'_1, X'_2, X'_2)$  vanish when they are restricted to the crucial (random) set  $A = \{x : \langle x, X'_{2,2} - X'_{2,1} \rangle = 0\}$ . To be more specific, since the null hypothesis is true and, hence,  $F_1 = F_2$ , we obtain

$$\begin{aligned} &\sum_{k=1}^2 \int_A \left[ I_{\langle x, X'_{1,1} - X'_{2,k} \rangle \leq 0} - F_1(x, \langle x, X'_{2,k} \rangle) - I_{\langle x, X'_{2,2} - X'_{2,k} \rangle \leq 0} + F_2(x, \langle x, X'_{2,k} \rangle) \right] \\ &\quad \times \left[ I_{\langle x, X'_{2,1} - X'_{2,k} \rangle \leq 0} - F_1(x, \langle x, X'_{2,k} \rangle) - I_{\langle x, X'_{2,2} - X'_{2,k} \rangle \leq 0} + F_2(x, \langle x, X'_{2,k} \rangle) \right] \\ &\quad \times \mathcal{P}(dx) \\ &= 2 \int_A \left[ I_{\langle x, X'_{1,1} - X'_{2,1} \rangle \leq 0} - F_1(x, \langle x, X'_{2,1} \rangle) - I_{\langle x, X'_{1,2} - X'_{2,1} \rangle \leq 0} + F_1(x, \langle x, X'_{2,1} \rangle) \right] \\ &\quad \times \left[ F_1(x, \langle x, X'_{2,1} \rangle) - F_1(x, \langle x, X'_{2,1} \rangle) \right] \mathcal{P}(dx) = 0. \end{aligned}$$

Thus, (A.12) follows again from (A.13), (A.14), the continuity of the inner product, and the convergence of  $X_{1n}^*$  and  $X_{2n}^*$ .  $\square$

## References

- Ahmed SE (2017) Big and complex data analysis. Contributions to statistics. Springer, Cham
- Anderson TW, Darling DA (1952) Asymptotic theory of certain “Goodness of Fit” criteria based on stochastic processes. Ann Math Stat 23:193–212
- Bugni FA, Horowitz JL (2018) Permutation tests for equality of distributions of functional data. ArXiv e-prints [arXiv:1803.00798](https://arxiv.org/abs/1803.00798)
- Bugni FA, Hall P, Horowitz JL, Neumann GR (2009) Goodness-of-fit tests for functional data. Econom J 12:S1–S18
- Cuesta-Albertos JA, Fraiman R, Ransford T (2006) Random projections and goodness-of-fit tests in infinite-dimensional spaces. Bull Braz Math Soc (N.S.) 37:477–501
- Cuesta-Albertos JA, del Barrio E, Fraiman R, Matrán C (2007) The random projection method in goodness of fit for functional data. Comput Stat Data Anal 51:4814–4831
- Cuevas A (2014) A partial overview of the theory of statistics with functional data. J Stat Plan Inference 147:1–23
- Cuevas A, Fraiman R (2009) On depth measures and dual statistics. A methodology for dealing with general data. J Multivar Anal 100:753–766
- Ditzhaus M, Gaigall D (2018) A consistent goodness-of-fit test for huge dimensional and functional data. J Nonparametric Stat 30:834–859
- Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7:1–26
- Gaigall D (2019) Testing marginal homogeneity of a continuous bivariate distribution with possibly incomplete paired data. Metrika 83:437–465
- Goia A, Vieu P (2016) An introduction to recent advances in high/infinite dimensional statistics. J Multivar Anal 146:1–6
- Göncü A, Oguz M, Karahan MO, Kuzubas TU (2016) A comparative goodness-of-fit analysis of distributions of some Lévy processes and Heston model to stock index returns. N Am J Econ Finance 36:69–83
- Gray BJ, French DW (2008) Empirical comparisons of distributional models for stock index returns. J Bus Finance Account 17:451–459
- Hall P, Tajvidi N (2002) Permutation tests for equality of distributions in high-dimensional settings. Biometrika 89:359–374
- Hall P, Van Keilegom I (2007) Two-sample tests in functional data analysis starting from discrete data. Stat Sin 17:1511–1531
- Janssen A, Pauls T (2003) How do bootstrap and permutation tests work? Ann Stat 31:768–806
- Jiang Q, Meintanis SG, Zhu L (2017) Two-sample tests for multivariate functional data. In: Aneiros G, Bongiorno E, Cao R, Vieu P (eds) Functional statistics and related fields. Contributions to statistics. Springer, Berlin
- Koroljuk VS, Borovskich XV (1994) Theory of  $U$ -statistics. Kluwer Academic Publishers Group, Dordrecht
- Lee AJ (1990)  $U$ -statistics: theory and practice. CRC Press, Boca Raton
- Leucht A, Neumann MH (2013) Dependent wild bootstrap for degenerate  $U$ - and  $V$ -statistics. J Multivar Anal 117:257–280
- Malevergne Y, Pisarenko V, Sornette D (2005) Empirical distributions of stock returns: between the stretched exponential and the power law? Quant Finance 5:379–401
- Mercer J (1909) Functions of positive and negative type and their connection with the theory of integral equations. Philos Trans R Soc A 209:415–446
- Midesia S, Basri H, Shabri M, Majid MSA (2016) The effects of asset management and profitability on stock returns: a comparative study between conventional and Islamic stock markets in Indonesia. Acad J Econ Stud 2:44–54
- Min S (2015) Goodness of fit and independence tests for major 8 companies of Korean stock market. Korean J Appl Stat 28:1245–1255
- Rosenblatt M (1952) Limit theorems associated with variants of the von Mises statistic. Ann Math Stat 23:617–623
- Serfling RS (2001) Approximation theorems of mathematical statistics. Wiley, New York
- Sun H (2005) Mercer theorem for RKHS on noncompact sets. J Complex 21:337–349
- Van der Vaart A, Wellner JA (1996) Weak convergence and empirical processes. With applications to statistics. Springer, New York

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.