# Detecting relevant differences in the covariance operators of functional time series - a sup-norm approach

Holger Dette, Kevin Kokot

Ruhr-Universität Bochum

Fakultät für Mathematik

Bochum, Germany

e-mail: {holger.dette, kevin.kokot}@rub.de

June 10, 2020

**Abstract**

In this paper we propose statistical inference tools for the covariance operators of functional time series in the two sample and change point problem. In contrast to most of the literature the focus of our approach is not testing the null hypothesis of exact equality of the covariance operators. Instead we propose to formulate the null hypotheses in the form that "the distance between the operators is small", where we measure deviations by the sup-norm. We provide powerful bootstrap tests for these type of hypotheses, investigate their asymptotic properties and study their finite sample properties by means of a simulation study.

Keywords: covariance operator, functional time series, two sample problems, change point problems, CUSUM, relevant hypotheses, Banach spaces, bootstrap
AMS Subject Classification: 62G10, 62M10

## 1 Introduction

The field of functional data analysis has found considerable attention in the statistical literature as in many applications the observed data points exhibit certain degrees of dependence and smoothness and thus may naturally be regarded as discretized functions. Introductions to this topic can be found in the monographs of Bosq (2000), Ramsay and Silverman (2005), Ferraty and Vieu (2010), Horváth and Kokoszka (2012) and Hsing and Eubank (2015), among others.

Interest may, for example, be in comparing characteristic parameters of the random functions from two different samples (*two sample problem*) or in investigating whether a certain parameter of a functional time series remains stable over time (*change point problem*). In most cases the considered parameters (such as the mean) are functions themselves, which makes the analysis of this type of problems challenging. In the present paper we investigate the second-order properties of a stationary functional time series which are contained in its covariance operators and important for the understanding of the smoothness of the stochastic fluctuations of the data (Kraus and Panaretos, 2012). Most of the literature on this topic considers Hilbert space-valued random variables. The popularity of this approach is due to the fact that such a framework allows the development of dimension reduction techniques such as (functional) principal components. On the other hand dimension reduction may yield to a loss of information as data is projected on finite dimensional spaces, and several authors argue that it might be more reasonable to work in the space of functions directly (see, for example, Aue et al., 2018, for a recent reference).

In this paper we will develop methodology to compare the covariance operators of two functional time series and to detect changes in the covariance operator of a functional time series in the space of continuous functions defined on a compact interval. Thus - in contrast to most of the literature on this topic, which considers Hilbert space-valued objects - the random variables under consideration are (dependent) elements of a Banach space, and it is possible to compare the covariance operators in the sup-norm. Another important difference to the literature consists in the fact that the main focus of our approach is not on *classical* hypotheses of the form

$$H_0 : C_1 = C_2 \qquad \text{versus} \qquad H_1 : C_1 \neq C_2 \tag{1.1}$$

where $C_1$ and $C_2$ are either the covariance operators corresponding to the two samples or to the covariance operator before and after a change point. In contrast we consider *relevant* hypotheses of the form

$$H_0^\Delta : d(C_1, C_2) \leq \Delta \qquad \text{versus} \qquad H_1^\Delta : d(C_1, C_2) > \Delta \tag{1.2}$$

where $\Delta \geq 0$ is a given threshold and $d$ a suitable metric on the space of covariance operators (in our case the sup-norm). Note that hypotheses of the form (1.2) contain the classical hypotheses in (1.1) as a special case for the choice $\Delta = 0$, but we argue that the case $\Delta > 0$ is at least of equal interest. In fact, in many applications it is obvious that $C_1$ and $C_2$ can not exactly coincide but the deviation might be small. In such cases testing for exact equality may be questionable and it might be more reasonable to test for a relevant or significant deviation between the two covariance operators.

In the case of testing classical hypotheses the metric does not matter because under the null hypothesis the distance between $C_1$ and $C_2$ vanishes in any metric. However, this is not the case for relevant hypotheses of the form (1.2). In the present context two covariance operators with rather different shapes may still have a small $L^2$-distance, which makes an appropriate interpretation of the threshold $\Delta$ for practitioners difficult. As an alternative we propose to consider the maximum deviation between the covariance operators as metric in the hypotheses (1.2). On the one hand this metric makes the interpretation of the threshold $\Delta$ more easy. On the other hand it leads to a Banach space-based framework where no dimension reduction techniques are available and the development and theoretical justification of statistical methods are more challenging.

In Section 2 we review some basic properties of random variables in the space of continuous functions. In particular we define moments of order two through injective tensor products. We also state a central limit theorem for a stationary Banach-space valued process, which will be the basis for all theoretical arguments given in this paper. In Section 3 we develop statistical methods for the comparison of covariance operators in the two sample problem. In particular a test is proposed for the null hypothesis of no relevant difference between the covariance operators from two independent samples. As a special (and substantially simpler case) we also construct a new test for the classical hypotheses (1.1) with a simple structure and nice statistical properties. Section 4 is devoted to the change point problem, where methodology is developed to detect changes in the covariance operator of a functional time series. In all cases we make use of a multiplier bootstrap procedure to obtain critical values for the proposed tests. The theoretical justification of all methods is given in Section 6, while Section 5 contains a detailed simulation study to investigate the finite sample properties of the proposed tests. Although classical hypotheses are not the main focus of our work we also compare the new tests for the classical hypotheses with some of the currently available methodology and demonstrate that they provide powerful alternatives to the procedures, which have been proposed in the literature so far.

## 1.1 Related literature

There exists a considerable amount of literature considering the comparison of covariance operators in the two sample problem, where random functions in the Hilbert space of square-integrable functions and the classical null hypothesis of equal covariance operators are investigated. Panaretos et al. (2010) consider independent Gaussian data and describe an application to DNA minicircle data. Fremdt et al. (2013) extend the theoretical findings of these authors to a more general model such that non-Gaussian curves are also covered. In both references, functional principal components (FPCs) are used for dimension reduction. Kraus and Panaretos (2012) introduce the notion of a dispersion operator and propose a robust test, which is based on a truncated version of the Hilbert-Schmidt norm of a score operator defined via the dispersion operator. Zhang and Shao (2015) propose a pivotal test procedure based on FPCs and self-normalization and also provide inference tools for the eigensystem of the covariance operators.

Several authors argue that dimension reduction may yield to a loss of information and propose alternative procedures for the comparison of covariance operators in the two sample problem. Pigoli et al. (2014) discuss different distance measures between covariance operators and develop a permutation test and Paparoditis and Sapatinas (2016) propose a bootstrap test for the (classical) null hypothesis of equality of $K$ covariance operators. Cabassi et al. (2017) suggest to combine all pairwise comparisons between samples of independent data into a global test for this problem, where the Hilbert-Schmidt norm between the square roots of the covariance operators is used as a measure of deviation. Boente et al. (2018) provide a theoretical framework which clarifies the ability of the test to detect local alternatives. Pilavakis et al. (2020) develop a fully functional test for the equality of auto-covariance operators of temporally dependent time series, which is based on a moving block bootstrap. For independent data the $K$-sample problem has also been considered by Guo et al. (2016) who propose to estimate the supremum value of the sum of the squared differences between the estimated individual covariance functions and the pooled sample covariance function.

So far, the change point problem for covariance operators has found less attention in the literature. Jarušková (2013) uses FPCs to develop a test for the existence of a change point, while Stoehr et al. (2019) use the circular block bootstrap to construct a change point test. In particular these authors develop a test based on dimension reduction and two procedures which take the full functional structure into account. A fully functional test has also been proposed by Sharipov and Wendler (2019), who use a non overlapping block bootstrap to obtain critical values. More recently, Aue et al. (2020) propose statistical tests for detecting a change in the spectrum and in the trace of the covariance operator, respectively.

All these references consider the problem of testing classical hypotheses of the form (1.1). Recently Dette et al. (2020b) propose a comparison of covariance operators in the two sample problem and in the context of change point analysis by testing relevant hypotheses of the form (1.2), where an $L^2$-distance is used as metric. However, in the context of testing relevant hypotheses the norm matters as two covariance operators might be close in one norm but not in another. In particular, relevant deviations between covariance operators in the sup-norm have - to our best knowledge - not been considered so far and requires a different methodology as the space under consideration is a Banach but not a Hilbert space. There does not exist so much literature on functional data analysis considering Banach spaces and exemplarily we mention the recent work of Dette et al. (2020a) who considered relevant hypotheses for the mean function and Liebl and Reimherr (2019), who developed confidence bands for functional parameters.

## 2 $C(T)$-valued random variables

In this paper we consider random variables taking values in the Banach space of real-valued and continuous functions defined on a compact set $T$ and denote this space by $C(T)$, which is equipped with the sup-norm $\|X\|_\infty = \max_{t\in T} |X(t)|$ for any $X \in C(T)$. The underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is assumed to be complete and measurability is always meant with respect to the Borel $\sigma$-field generated by the open sets relative to the respective sup-norm.

Following Chapter 11 in Janson and Kaijser (2015), we use injective tensor products to define moments of $C(T)$-valued random variables and note that $C(T)^{\check{\otimes}k} = C(T^k)$ isometrically with the natural identification (Theorem 11.6). The $k$th moment of a $C(T)$-valued random variable $X$ exists, whenever $\mathbb{E}\big[\|X\|_\infty^k\big] < \infty$ (Theorem 11.25) and is defined by the function in $C(T)^{\check{\otimes}k} = C(T^k)$, which maps $(t_1, \ldots, t_k) \in T^k$ to

$$\mathbb{E}X^{\check{\otimes}k}(t_1, \ldots, t_k) = \mathbb{E}\big[X(t_1)\cdots X(t_k)\big]$$

(Theorem 11.10). Throughout this paper, we write $X^{\check{\otimes}2} = X\check{\otimes}X$ for any $X \in C(T)$ and mean the function in $C(T^2)$ defined by $(s,t) \mapsto X(s)X(t)$. Consequently, the covariance operator of a $C(T)$-valued random variable is defined by

$$C(\cdot,\cdot) = \operatorname{Cov}(X(\cdot), X(\cdot)) = \mathbb{E}\big[(X-\mu)^{\check{\otimes}2}(\cdot,\cdot)\big] \in C(T^2)$$

where $\mu = \mathbb{E}[X] \in C(T)$ is the expectation of $X$.

Let $\rho$ denote a metric on $T$ such that $(T,\rho)$ is totally bounded, then the metric $\rho_{\max}$ on $T^2$ is defined through $\rho_{\max}((s,t),(s',t')) = \max\{\rho(s,s'), \rho(t,t')\}$ and the expression $D(\omega, \rho_{\max})$ denotes the packing number with respect to the metric $\rho_{\max}$ on $T^2$ that is the maximal number of $\omega$-seperated points in $T^2$ (Van der Vaart and Wellner, 1996). Note that in this case $(T^2, \rho_{\max})$ is totally bounded as well.

In order to describe the dependence in the data we introduce the concept $\varphi$-mixing and denote by $\mathbb{P}(G|F)$ the conditional probability of $G$ given $F$. For two $\sigma$-fields $\mathcal{F}$ and $\mathcal{G}$ we define the coefficient

$$\phi(\mathcal{F},\mathcal{G}) = \sup\left\{|\mathbb{P}(G|F) - \mathbb{P}(G)| \colon F \in \mathcal{F},\ G \in \mathcal{G},\ \mathbb{P}(F) > 0\right\}. \tag{2.1}$$

For a given strictly stationary sequence $(\eta_j)_{j\in\mathbb{N}}$ of random variables in $C(T)$, denote by $\mathcal{F}_k^{k'}$ the $\sigma$-field generated by $(\eta_j \colon k \le j \le k')$. Then, the $k$th *$\varphi$-mixing coefficient* of $(\eta_j)_{j\in\mathbb{N}}$ is defined by

$$\varphi(k) = \sup_{k'\in\mathbb{N}} \phi(\mathcal{F}_1^{k'}, \mathcal{F}_{k'+k}^{\infty})$$

and the stationary time series $(\eta_j)_{j\in\mathbb{N}}$ is called *$\varphi$-mixing* whenever the sequence of mixing coefficients converges to zero as $k \to \infty$.

Given the preceding discussion, the analysis of the covariance operators of random variables in $C(T)$ can in some sense be regarded to the analysis of $C(T^2)$-valued random variables. More precisely, Theorem 11.7 in Janson and Kaijser (2015) implies that $C(T^2)$ is separable such that measurability issues are avoided, and Theorem 1.3 in Billingsley (1968) implies that any $C(T^2)$-valued random variable is tight. A random function $X$ in $C(T^2)$ is called Gaussian if and only if its finite dimensional vectors $(X(t_1), \ldots, X(t_k))$ follow a multivariate normal distribution for any $t_1, \ldots, t_k \in T^2$ and $k \in \mathbb{N}$.

**Assumption 2.1** *$(Z_j)_{j\in\mathbb{N}}$ is a sequence of $C(T)$-valued random variables such that*

$$Z_j = \mu + \eta_j\,, \quad j \in \mathbb{N}$$

*where $\mu \in C(T)$ denotes the expectation function and $(\eta_j)_{j\in\mathbb{N}}$ is a strictly stationary process.*

*(A1) The packing number $D(\omega, \rho_{\max})$ satisfies*

$$\int_0^\tau D(\omega, \rho_{\max})^{1/J}\, d\omega < \infty$$

*for some $\tau > 0$ and some even integer $J \geq 2$.*

*(A2) There is a constant $K$ such that*

$$\mathbb{E}[\|\eta_1\|_\infty^{4+\nu}] \leq K\,, \quad \mathbb{E}[\|\eta_1\|_\infty^{2J}] \leq \infty$$

*for some $\nu > 0$, where $J$ is the same integer as in (A1).*

*(A3) There exists a real-valued non-negative random variable $M$ and a constant $\tilde{K}$ such that, for any $j \in \mathbb{N}$, $\mathbb{E}\big[(\|\eta_j\|_\infty M)^J\big] < \tilde{K} < \infty$ and the inequality*

$$|\eta_j(t) - \eta_j(t')| \leq M\rho(t, t')$$

*holds almost surely for all $t, t' \in T$. The integer $J$ is the same as in (A1).*

*(A4) The process $(\eta_j)_{j\in\mathbb{N}}$ is $\varphi$-mixing with mixing coefficients satisfying, for some $\bar{\tau} \in (1/(2+2\nu), 1/2)$, the condition*

$$\sum_{k=1}^\infty k^{1/(1/2-\bar{\tau})}\varphi(k)^{1/2} < \infty\,, \quad \sum_{k=1}^\infty (k+1)^{J/2-1}\varphi(k)^{1/J} < \infty\,,$$

*where the constants $J$ and $\nu$ are the same as in (A1) and (A2), respectively.*

Note that Assumption 2.1 implies the existence of the covariance operator defined by

$$C(s,t) = \mathrm{Cov}(Z_j(s), Z_j(t)) = \mathbb{E}[(Z_j(s) - \mu(s))(Z_j(t) - \mu(t))]\,. \tag{2.2}$$

Condition (A4) on the summability of the mixing coefficients is satisfied if there exists an $a \in (0,1)$ such that $\varphi(k) \leq ca^k$ ($k \in \mathbb{N}$). For the formulation and a proof of a CLT of Banach space valued random variables we denote by the symbol "$\rightsquigarrow$" weak convergence in $(C(T))^k$ or $(C(T^2))^k$ and the symbol "$\xrightarrow{\mathcal{D}}$" denotes weak convergence in $\mathbb{R}^k$ for some $k \in \mathbb{N}$. The following result is proved in Section 6.

**Theorem 2.1** *Let $(Z_j)_{j\in\mathbb{N}}$ denote a stochastic process in $C(T)$ satisfying Assumption 2.1. Then,*

$$G_n = \frac{1}{\sqrt{n}}\sum_{j=1}^n ((Z_j - \bar{Z}_n)^{\breve{\otimes}2} - C) \rightsquigarrow Z$$

*in $C(T^2)$ as $n \to \infty$ where $\bar{Z}_n = 1/n \sum_{j=1}^n Z_j \in C(T)$, $C$ is defined by (2.2) and $Z$ is a centred Gaussian random variable with covariance operator*

$$\mathbb{C}((s,t),(s',t')) = \mathrm{Cov}(Z(s,t), Z(s',t')) = \sum_{i=-\infty}^{\infty} Cov\big(\eta_0^{\check{\otimes}2}(s,t), \eta_i^{\check{\otimes}2}(s',t')\big) . \tag{2.3}$$

In the remaining part of the paper, we consider the unit interval $T = [0,1]$ and, for a positive constant $\theta \in (0,1]$, the metric $\rho(s,t) = |s-t|^\theta$ on $[0,1]$. Consequently, on $T^2 = [0,1]^2$, we use the metric $\rho_{\max}((s,t),(s',t')) = \max\{\rho(s,s'), \rho(t,t')\}$ and the packing number of the square $[0,1]^2$ with respect to this metric satisfies $D(\omega, \rho_{\max}) \lesssim \lceil \omega^{-2/\theta} \rceil$ (to see this, consider the points $(k\omega^{1/\theta}, l\omega^{1/\theta})$ for $k, l = 0, \ldots, \lfloor \omega^{-1/\theta} \rfloor$). Therefore condition (A1) reduces to

$$\int_0^\tau D(\omega, \rho_{\max})^{1/J} \, d\omega \lesssim \int_0^\tau \lceil \omega^{-2/\theta} \rceil^{1/J} \, d\omega \lesssim \frac{\tau^{1-2/(J\theta)}}{1 - 2/(J\theta)} < \infty$$

and holds, whenever the even integer $J$ satisfies $J > 2/\theta$ and consequently, under this assumption, Hölder continuous processes satisfy (A1). Because the paths of the Brownian Motion $\{W(t)\}_{t\in[0,1]}$ are Hölder continuous of order $\theta$ for any $\theta \in (0, 1/2)$ and the random variable $\|W\|_\infty$ has moments of all order Assumption 2.1 is satisfied for the Brownian motion (we can use $J = 6$ in (A4) for this case). For general processes with less smoothness, that is a smaller constant $\theta$, we require a stronger summability assumption (A4) on the mixing coefficients and the existence of higher moments.

# 3 The two sample problem

Throughout this section, we consider two independent samples $(X_j : j = 1, \ldots, m)$ and $(Y_j : j = 1, \ldots, n)$ drawn from independent strictly stationary sequences $(X_j)_{j\in\mathbb{N}}$ and $(Y_j)_{j\in\mathbb{N}}$ in $C([0,1])$ with representations

$$X_j = \mu_1 + \eta_{1,j} , \quad Y_j = \mu_2 + \eta_{2,j} , \tag{3.1}$$

where $\mu_1, \mu_2 \in C([0,1])$ and $(\eta_{1,j})_{j\in\mathbb{N}}$, $(\eta_{2,j})_{j\in\mathbb{N}}$ are centred $C([0,1])$-valued processes satisfying the following assumption.

**Assumption 3.1** *The processes $(\eta_{1,j})_{j\in\mathbb{N}}$, $(\eta_{2,j})_{j\in\mathbb{N}}$ are independent centred strictly stationary processes satisfying Assumption 2.1 with metric $\rho(s,t) = |t-s|^\theta$ for some $\theta > 0$ such that $J\theta > 2$.*

In the following let

$$C_1(s,t) = \mathbb{E}[\eta_{1,j}(s)\eta_{1,j}(t)] = \mathrm{Cov}(X_1(s), X_1(t)) ,$$
$$C_2(s,t) = \mathbb{E}[\eta_{2,j}(s)\eta_{2,j}(t)] = \mathrm{Cov}(Y_1(s), Y_1(t))$$

denote the covariance operator of the first and the second sample, respectively. We measure the difference between $C_1$ and $C_2$ by their maximal deviation

$$d_\infty = \|C_1 - C_2\|_\infty = \sup_{s,t\in[0,1]} |C_1(s,t) - C_2(s,t)| \,, \tag{3.2}$$

and are interested in testing if there exists a relevant difference between the covariance operators, that is

$$H_0^\Delta : d_\infty \ \le \Delta \qquad \text{versus} \qquad H_1^\Delta : d_\infty > \Delta \,, \tag{3.3}$$

where $\Delta \in \mathbb{R}$ is a pre-specified constant. Note that the classical hypotheses

$$H_0 : C_1 = C_2 \qquad \text{versus} \qquad H_1 : C_1 \neq C_2 \tag{3.4}$$

are obtained for the choice $\Delta = 0$.
We denote by $\tilde{X}_{m,j} = X_j - \bar{X}_m$, $\tilde{Y}_{n,i} = Y_i - \bar{Y}_n$ the centred random curves (here $\bar{X}_m$ and $\bar{Y}_n$ denote the mean in the first and second sample, respecively), and estimate the maximal deviation $d_\infty$ in (3.2) between the two covariance operators by

$$\hat{d}_\infty := \sup_{s,t\in[0,1]} \left| \frac{1}{m-1} \sum_{j=1}^{m} \tilde{X}_{m,j}^{\check{\otimes}2}(s,t) - \frac{1}{n-1} \sum_{j=1}^{n} \tilde{Y}_{n,j}^{\check{\otimes}2}(s,t) \right| . \tag{3.5}$$

Now a reasonable decision rule is to reject the null hypothesis in (3.3) or (3.4) for large values of $\hat{d}_\infty$. Our first result provides the asymptotic properties of the statistic $\hat{d}_\infty$.

**Proposition 3.1** *If $\mu_1, \mu_2 \in C([0,1])$ and $(\eta_{1,j})_{j\in\mathbb{N}}$, $(\eta_{2,j})_{j\in\mathbb{N}}$ are strictly stationary and centred $C([0,1])$-valued processes satisfying Assumption 3.1 and $\frac{m}{m+n} \longrightarrow \lambda \in (0,1)$ as $m,n \to \infty$, the following assertions hold true.*

*(1) If $d_\infty = 0$, then*

$$\sqrt{m+n}\,\hat{d}_\infty \xrightarrow{\mathcal{D}} T = \sup_{s,t\in[0,1]} |Z(s,t)| \,, \tag{3.6}$$

*where $Z$ is a Gaussian random element in $C([0,1]^2)$ with covariance operator*

$$\mathbb{C} = \frac{1}{\lambda}\, \mathbb{C}_1 - \frac{1}{1-\lambda}\, \mathbb{C}_2 \,, \tag{3.7}$$

*and $\mathbb{C}_1$ and $\mathbb{C}_2$ are the long-run covariance operators defined by*

$$\mathbb{C}_1((s,t),(s',t')) = \sum_{i=-\infty}^{\infty} Cov\big(\eta_{1,0}^{\check{\otimes}2}(s,t), \eta_{1,i}^{\check{\otimes}2}(s',t')\big) \,, \tag{3.8}$$

$$\mathbb{C}_2((s,t),(s',t')) = \sum_{i=-\infty}^{\infty} Cov\big(\eta_{2,0}^{\check{\otimes}2}(s,t), \eta_{2,i}^{\check{\otimes}2}(s',t')\big) \,. \tag{3.9}$$

*(2) If $d_\infty > 0$, we have*

$$\sqrt{m+n}\,(\hat{d}_\infty - d_\infty) \xrightarrow{\mathcal{D}} T(\mathcal{E}) = \max\Big\{ \sup_{(s,t)\in\mathcal{E}^+} Z(s,t),\ \sup_{(s,t)\in\mathcal{E}^-} -Z(s,t)\Big\}\,, \tag{3.10}$$

*where $Z$ is a Gaussian random element in $C([0,1]^2)$ with covariance operator defined by* (3.7) *and*

$$\mathcal{E}^{\pm} = \big\{(s,t)\in[0,1]^2 \colon C_1(s,t) - C_2(s,t) = \pm d_\infty\big\} \tag{3.11}$$

*are the extremal sets of the difference of the covariance operators $C_1, C_2$.*

If $u_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the distribution of the random variable $T$ defined in (3.6), a consistent and asymptotic level $\alpha$ tests for the classical hypotheses in (3.4) can be obtained by rejecting the null hypothesis, whenever

$$\hat{d}_\infty > \frac{u_{1-\alpha}}{\sqrt{m+n}}\,.$$

Similarly, the null hypothesis in (3.3) is rejected if

$$\hat{d}_\infty > \Delta + \frac{u_{1-\alpha,\mathcal{E}}}{\sqrt{m+n}}$$

where $u_{1-\alpha,\mathcal{E}}$ is the $\alpha$-quantile of the distribution of the random variable $T(\mathcal{E})$ defined in (3.10). However, the quantile $u_{1-\alpha}$ depends on the long-run covariance operators $\mathbb{C}_1$ and $\mathbb{C}_2$ which are difficult to estimate. For the problem of testing relevant hypotheses the situation is even more complicated as the quantile $u_{1-\alpha,\mathcal{E}}$ additionally depends on the unknown extremal sets defined in (3.11), which have to be estimated as well. To deal with these problems we propose a bootstrap approach, which is explained for the classical and relevant hypotheses separately.

## 3.1 Classical hypotheses

In order to avoid the problem of estimating the long-run covariance operators we propose a bootstrap procedure to mimic the covariance structure of the distribution of the process

$$\frac{1}{m-1}\sum_{j=1}^{m} \tilde{X}_{m,j}^{\check{\otimes}2} - \frac{1}{n-1}\sum_{j=1}^{n} \tilde{Y}_{n,j}^{\check{\otimes}2} - (C_1 - C_2)$$

by a multiplier bootstrap process (note that the second term vanishes in the case $d_\infty = 0$). To be precise, we denote by $(\xi_k^{(1)})_{k\in\mathbb{N}}, \ldots, (\xi_k^{(R)})_{k\in\mathbb{N}}$ and $(\zeta_k^{(1)})_{k\in\mathbb{N}}, \ldots, (\zeta_k^{(R)})_{k\in\mathbb{N}}$ independent sequences of independent standard normal distributed random variables and define the $C([0,1]^2)$-valued processes $\hat{B}_{m,n}^{(1)}, \ldots, \hat{B}_{m,n}^{(R)}$ by

$$\begin{aligned}\hat{B}_{m,n}^{(r)} = \sqrt{n+m}\Big\{ &\frac{1}{m}\sum_{k=1}^{m-l_1+1} \frac{1}{\sqrt{l_1}}\Big(\sum_{j=k}^{k+l_1-1} \tilde{X}_{m,j}^{\check{\otimes}2} - \frac{l_1}{m}\sum_{i=1}^{m} \tilde{X}_{m,i}^{\check{\otimes}2}\Big)\xi_k^{(r)} \\ &- \frac{1}{n}\sum_{k=1}^{n-l_2+1} \frac{1}{\sqrt{l_2}}\Big(\sum_{j=k}^{k+l_2-1} \tilde{Y}_{n,j}^{\check{\otimes}2} - \frac{l_2}{n}\sum_{i=1}^{n} \tilde{Y}_{n,i}^{\check{\otimes}2}\Big)\zeta_k^{(r)}\Big\} \quad (r = 1,\ldots,R)\,.\end{aligned} \tag{3.12}$$

The parameters $l_1, l_2 \in \mathbb{N}$ define the block length such that $l_1/m \to 0$ and $l_2/n \to 0$ as $l_1, l_2, m, n \to \infty$. Note that the dependence on $l_1$ and $l_2$ is not reflected in the notation of the bootstrap processes. With these notations we define the bootstrap statistics

$$T_{m,n}^{(r)} = \sup_{s,t \in [0,1]} |\hat{B}_{m,n}^{(r)}(s,t)| \quad (r = 1, \ldots, R)\,, \tag{3.13}$$

and denote by $T_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}$ the empirical $(1-\alpha)$-quantile of the bootstrap sample $T_{m,n}^{(1)}, \ldots, T_{m,n}^{(R)}$. Then, rejecting the classical null hypothesis of equal covariance operators whenever

$$\hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}} \tag{3.14}$$

defines a bootstrap test for the classical hypotheses in (3.4). The following result provides the statistical properties of this test.

**Theorem 3.1** *Suppose that the error processes $(\eta_{1,j})_{j\in\mathbb{N}}$ and $(\eta_{2,j})_{j\in\mathbb{N}}$ in the representation* (3.1) *satisfy Assumption 3.1. Let $\hat{B}_{m,n}^{(1)}, \ldots, \hat{B}_{m,n}^{(R)}$ denote the bootstrap processes defined by* (3.12) *such that $l_1 = m^{\beta_1}$, $l_2 = n^{\beta_2}$ with*

$$0 < \beta_i < \nu_i/(2+\nu_i)\;, \quad \bar{\tau}_i > (\beta_i(2+\nu_i)+1)/(2+2\nu_i)$$

*where $\bar{\tau}_i, \nu_i$ are given in Assumption 2.1, $i = 1, 2$.*
*Then, under the classical null hypothesis $H_0 : C_1 = C_2$ in* (3.4) *we have*

$$\lim_{m,n,R\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}}\Big) = \alpha\,. \tag{3.15}$$

*Under the alternative $H_1 : C_1 \neq C_2$ in* (3.4) *it follows for any $R \in \mathbb{N}$,*

$$\liminf_{m,n\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}}\Big) = 1\,. \tag{3.16}$$

## 3.2 Relevant hypotheses

For testing relevant hypotheses it is crucial to estimate the extremal sets in (3.11) properly. For this purpose we propose

$$\hat{\mathcal{E}}_{m,n}^{\pm} = \Big\{(s,t) \in [0,1]^2 \colon \pm\big(\hat{C}_1(s,t) - \hat{C}_2(s,t)\big) \geq \hat{d}_\infty - \frac{c_{m,n}}{\sqrt{m+n}}\Big\} \tag{3.17}$$

as estimators of the sets $\mathcal{E}^{\pm}$ where $(c_{m,n})_{m,n\in\mathbb{N}}$ is a sequence of positive constants satisfying $\lim_{m,n\to\infty} c_{m,n}/\log(m+n) = c$ for some $c > 0$. For the construction of a test of the relevant hypotheses in (3.3) we recall the definition of the bootstrap process in (3.12) and define the statistics

$$K_{m,n}^{(r)} = \max\Big\{\sup_{(s,t)\in\hat{\mathcal{E}}_{m,n}^{+}} \hat{B}_{m,n}^{(r)}(s,t),\; \sup_{(s,t)\in\hat{\mathcal{E}}_{m,n}^{-}} \big(-\hat{B}_{m,n}^{(r)}(s,t)\big)\Big\} \qquad (r = 1, \ldots, R) \tag{3.18}$$

which serves as the bootstrap analogue of the statistic $T(\mathcal{E})$ defined in (3.10). If $K_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}$ denotes the empirical $(1-\alpha)$-quantile of the bootstrap sample $K_{m,n}^{(1)}, \ldots, K_{m,n}^{(R)}$ we propose to reject the null hypothesis of no relevant difference in the covariance operators at level $\alpha$, whenever

$$\hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}} \,. \tag{3.19}$$

The final result of this section states that this test is consistent and has asymptotic level $\alpha$.

**Theorem 3.2** *Suppose that the assumptions of Theorem 3.1 are satisfied and that* $\Delta > 0$.

*(1) Under the null hypothesis* $H_0 : d_\infty \leq \Delta$ *of no relevant difference in the covariance operators, it follows*

$$\lim_{m,n,R\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}}\Big) = \alpha \,, \tag{3.20}$$

*if* $\Delta = d_\infty$ *and, for any* $R \in \mathbb{N}$,

$$\lim_{m,n\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}}\Big) = 0 \,,$$

*if* $0 < d_\infty < \Delta$.

*(2) Under the alternative* $H_1 : d_\infty > \Delta$ *of a relevant difference in the covariance operators it follows for any* $R \in \mathbb{N}$

$$\liminf_{m,n\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{m+n}}\Big) = 1 \,.$$

# 4 Detecting changes in the covariance operator

In this section we study the change point problem for the covariance operator of an array $(X_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ of $C([0,1])$-valued random variables. For the consideration of relevant changes we require a dependence concept for an array $(\tilde{\eta}_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ of random variables in $C(T)$ with strictly stationary rows. For this purpose we denote by $\mathcal{F}_{k,n}^{k'}$ the $\sigma$-field generated by $(\tilde{\eta}_{n,j} \colon k \leq j \leq k')$. The $k$th *$\varphi$-mixing coefficient* of the array $(\tilde{\eta}_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ is then defined by

$$\varphi(k) = \sup_{n\in\mathbb{N}} \sup_{k'\in\{1,\ldots,n-k\}} \phi(\mathcal{F}_{1,n}^{k'}, \mathcal{F}_{k'+k,n}^{n})$$

and $(\tilde{\eta}_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ is called *$\varphi$-mixing* whenever $\varphi(k) \to 0$ as $k \to \infty$. For our theoretical investigations we make the following assumption.

**Assumption 4.1** *For some $\vartheta \in (0, \frac{1}{2}]$ there exists a number $s^* \in [\vartheta, 1-\vartheta]$ such that the random variables $(X_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ are given by $X_{n,j} = \mu + \tilde{\eta}_{n,j}$, where $\mu = \mathbb{E}[X_{n,j}]$ denotes the common expectation function,*

$$\tilde{\eta}_{n,j} = \begin{cases} \eta_{1,j} & \text{if } j \in \{1, \ldots, \lfloor s^* n \rfloor\} \\ \eta_{2,j} & \text{if } j \in \{\lfloor s^* n \rfloor + 1, \ldots, n\} \end{cases} \tag{4.1}$$

*and $(\eta_{1,j})_{n\in\mathbb{N}}$, $(\eta_{2,j})_{n\in\mathbb{N}}$ are centred strictly stationary processes satisfying conditions (A1) - (A3) of Assumption 2.1 with metric $\rho(s,t) = |s-t|^\theta$ for some $\theta > 0$ such that $\theta J > 2$. Furthermore it is assumed that the array $(\tilde{\eta}_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ is $\varphi$-mixing with mixing coefficients satisfying condition (A4) of Assumption 2.1.*

We denote by $C_1$ and $C_2$ the covariance operator before and after the change point. Recalling the definition of $d_\infty$ in (3.2) the relevant and classical hypotheses are given by (3.3) and (3.4), respectively. For the construction of a test for these hypotheses we consider a sequential empirical process on $C([0,1]^3)$ defined by

$$\hat{\mathbb{U}}_n(s,t,u) = \frac{1}{n}\Big( \sum_{j=1}^{\lfloor sn \rfloor} \tilde{X}_{n,j}^{\check{\otimes} 2}(t,u) + n\Big(s - \frac{\lfloor sn \rfloor}{n}\Big) \tilde{X}_{n,\lfloor sn \rfloor+1}^{\check{\otimes} 2}(t,u) - s \sum_{j=1}^{n} \tilde{X}_{n,j}^{\check{\otimes} 2}(t,u) \Big) \tag{4.2}$$

where $\tilde{X}_{n,j} = X_{n,j} - \bar{X}_n$ $(j = 1, \ldots, n;\ n \in \mathbb{N})$ and note that it can be shown that

$$\mathbb{E}\big[\hat{\mathbb{U}}_n(s,t,u)\big] = \big(s \wedge s^* - ss^*\big)\big(C_1(t,u) - C_2(t,u)\big) + o_{\mathbb{P}}(1)\,.$$

Consequently, it is reasonable to consider the statistic

$$\hat{\mathbb{M}}_n = \sup_{s\in[0,1]} \sup_{t,u\in[0,1]} |\hat{\mathbb{U}}_n(s,t,u)| \tag{4.3}$$

as an estimate of

$$s^*(1-s^*)\, d_\infty = s^*(1-s^*)\, \|C_1 - C_2\|_\infty\,.$$

The following result makes these heuristic arguments precise.

**Proposition 4.1** *If Assumption 4.1 is satisfied, the following statements hold true.*

*(1) If $d_\infty = 0$, then*

$$\sqrt{n}\, \hat{\mathbb{M}}_n \xrightarrow{\mathcal{D}} \check{T} = \sup_{(s,t,u)\in[0,1]^3} |\mathbb{W}(s,t,u)| \tag{4.4}$$

*where $\mathbb{W}$ is a Gaussian random element in $C([0,1]^3)$ with covariance operator*

$$\begin{aligned} &Cov(\mathbb{W}(s,t,u), \mathbb{W}(s',t',u')) \\ &= \big\{(s \wedge s' \wedge s^*) + ss's^* - s(s' \wedge s^*) - s'(s \wedge s^*)\big\}\, \mathbb{C}_1((t,u),(t',u')) \\ &\quad + \big\{(s \wedge s' - s^*)_+ + ss'(1-s^*) - s(s'-s^*)_+ - s'(s-s^*)_+\big\}\, \mathbb{C}_2((t,u),(t',u')) \end{aligned} \tag{4.5}$$

*and the long-run covariance operators* $\mathbb{C}_1, \mathbb{C}_2$ *are defined by*

$$\mathbb{C}_l((s,t),(s',t')) = \sum_{i=-\infty}^{\infty} Cov\big(\eta_{l,0}^{\check{\otimes}2}(s,t), \eta_{l,i}^{\check{\otimes}2}(s',t')\big) \qquad (l=1,2)\,. \tag{4.6}$$

*(2) If* $d_\infty > 0$*, we have*

$$\sqrt{n}\big(\hat{\mathbb{M}}_n - s^*(1-s^*)d_\infty\big) \xrightarrow{\mathcal{D}} \tilde{D}(\mathcal{E}) = \max\Big\{ \sup_{(t,u)\in\mathcal{E}^+} \mathbb{W}(s^*,t,u), \sup_{(t,u)\in\mathcal{E}^-} -\mathbb{W}(s^*,t,u)\Big\}\,, \tag{4.7}$$

*where* $\mathbb{W}$ *is a Gaussian random element in* $C([0,1]^3)$ *with covariance operator defined by* (4.5) *and* $\mathcal{E}^{\pm}$ *are the extremal sets defined in* (3.11).

As in the two sample problem we can form decision rules, rejecting the null hypothesis (classical or relevant) for large values of $\hat{\mathbb{M}}_n$. Note that this requires estimation of the long-run covariance operators and (in the case of relevant hypotheses) the estimation of the change point and the extremal sets. For the construction of an explicit test (based on a multiplier bootstrap) we investigate again classical and relevant hypotheses separately.

## 4.1 Classical hypotheses

Most of the literature on change point analysis of covariance operators investigates the classical hypotheses of the form (3.4), where $C_1$ and $C_2$ denote the covariance operator before and after the change point (see Jarušková, 2013; Sharipov and Wendler, 2019; Stoehr et al., 2019). In order to obtain critical values for a test for a structural break in the covariance operators we consider a $C([0,1]^3)$-valued bootstrap process defined by

$$\begin{aligned}\hat{B}_n^{(r)}(s,t,u) =& \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor sn\rfloor} \frac{1}{\sqrt{l}} \Big( \sum_{j=k}^{k+l-1} \hat{Y}_{n,j}(t,u) - \frac{l}{n}\sum_{j=1}^{n} \hat{Y}_{n,j}(t,u)\Big)\xi_k^{(r)} \\ &+ \sqrt{n}\Big(s - \frac{\lfloor sn\rfloor}{n}\Big)\frac{1}{\sqrt{l}}\Big( \sum_{j=\lfloor sn\rfloor+1}^{\lfloor sn\rfloor+l} \hat{Y}_{n,j}(t,u) - \frac{l}{n}\sum_{j=1}^{n} \hat{Y}_{n,j}(t,u)\Big)\xi_{\lfloor sn\rfloor+1}^{(r)}\,,\end{aligned} \tag{4.8}$$

if $\lfloor sn\rfloor \leq n-l$, where $(\xi_k^{(1)})_{k\in\mathbb{N}}, \dots, (\xi_k^{(R)})_{k\in\mathbb{N}}$ denote independent sequences of independent Gaussian random variables with mean 0 and variance 1 and

$$\hat{Y}_{n,j} = \tilde{X}_{n,j}^{\check{\otimes}2}(t,u) - (\hat{C}_2 - \hat{C}_1)\mathbb{1}\{j > \lfloor \hat{s}n\rfloor\} \qquad (j=1,\dots,n)\,.$$

The expressions

$$\hat{C}_1 = \frac{1}{\lfloor \hat{s}n\rfloor}\sum_{j=1}^{\lfloor \hat{s}n\rfloor} \tilde{X}_{n,j}^{\check{\otimes}2}(t,u) \qquad \text{and} \qquad \hat{C}_2 = \frac{1}{\lfloor (1-\hat{s})n\rfloor}\sum_{j=\lfloor \hat{s}n\rfloor+1}^{n} \tilde{X}_{n,j}^{\check{\otimes}2}(t,u)$$

are estimators of the covariance operator before and after the change point and

$$\hat{s} = \max\left\{\vartheta,\, \min\left\{\frac{1}{n} \arg\max_{1 \le k < n} \big\|\hat{\mathbb{U}}_n(k/n,\cdot,\cdot)\big\|_\infty,\, 1-\vartheta\right\}\right\} \tag{4.9}$$

is an estimator of the unknown change location $s^*$ (note that $s^* \in (\vartheta, 1-\vartheta)$ by assumption). In (4.8) the parameter $l \in \mathbb{N}$ denotes the block length satisfying $l/n \to 0$ as $l, n \to \infty$ and for any $t, u \in [0,1]$ and any $s \in [0,1]$ such that $\lfloor sn \rfloor > n - l$ we define

$$\hat{B}_n^{(r)}((n-l)/n, t, u) = \hat{B}_n^{(r)}(s, t, u)\,.$$

Finally, a bootstrap process is defined by

$$\hat{\mathbb{W}}_n^{(r)}(s,t,u) = \hat{B}_n^{(r)}(s,t,u) - s\hat{B}_n^{(r)}(1,t,u) \qquad (r = 1, \ldots, R) \tag{4.10}$$

and we consider the bootstrap statistic

$$\check{T}_n^{(r)} = \sup_{s,t,u \in [0,1]} \big|\hat{\mathbb{W}}_n^{(r)}(s,t,u)\big| \qquad (r = 1, \ldots, R)\,. \tag{4.11}$$

If $\check{T}_n^{\{\lfloor R(1-\alpha) \rfloor\}}$ denotes the empirical $(1-\alpha)$-quantile of the bootstrap sample $\check{T}_n^{(1)}, \check{T}_n^{(2)}, \ldots, \check{T}_n^{(R)}$, the classical null hypothesis (3.4) of no change in the covariance operators is rejected, whenever

$$\hat{\mathbb{M}}_n > \frac{\check{T}_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}}\,. \tag{4.12}$$

**Theorem 4.1** *Assume that the array $(X_{n,j}\colon n \in \mathbb{N}, j = 1, \ldots, n)$ satisfies Assumption 4.1. Further assume that $l = n^\beta$ for some constant $\beta \in (1/5, 2/7)$ such that the constant $\nu$ in (A2) satisfies $\nu \ge 4$ and*

$$(\beta(2+\nu)+1)/(2+2\nu) < \bar{\tau} < 1/2$$

*where $\bar{\tau}$ is defined in (A4).*
*Then, under the classical null hypothesis $H_0 : C_1 = C_2$, we have*

$$\lim_{n,R \to \infty} \mathbb{P}\bigg(\hat{\mathbb{M}}_n > \frac{\check{T}_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}}\bigg) = \alpha\,.$$

*Under the alternative $H_1 : C_1 \neq C_2$ we have, for any $R \in \mathbb{N}$,*

$$\liminf_{n \to \infty} \mathbb{P}\bigg(\hat{\mathbb{M}}_n > \frac{\check{T}_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}}\bigg) = 1\,.$$

## 4.2 Relevant hypotheses

Testing for a relevant change in the covariance operators as formulated in (3.3) is more complicated. In particular because - as indicated in Proposition 4.1 - it additionally requires the estimation of the extremal sets. To be precise we recall the definition of $\hat{\mathbb{M}}_n$ in (4.3) and define

$$\hat{d}_\infty = \frac{\hat{\mathbb{M}}_n}{\hat{s}(1-\hat{s})} \tag{4.13}$$

as an estimator of the maximal deviation of the covariance operator before and after the change point, and use

$$\hat{\mathcal{E}}_n^{\pm} = \Big\{(t,u) \in [0,1]^2 \colon \pm\big(\hat{C}_1(t,u) - \hat{C}_2(t,u)\big) \geq \hat{d}_\infty - \frac{c_n}{\sqrt{n}}\Big\}, \tag{4.14}$$

as the estimator of the extremal sets, where $(c_n)_{n\in\mathbb{N}}$ is a sequence of positive constants such that $\lim_{n\to\infty} c_n/\log(n) = c > 0$. In order to obtain a test for the relevant hypotheses in (3.3) define, for $r = 1, \ldots, R$, the bootstrap statistics

$$\check{K}_n^{(r)} = \frac{1}{\hat{s}(1-\hat{s})} \max\Big\{\sup_{(t,u)\in\hat{\mathcal{E}}_n^+} \hat{W}_n^{(r)}(\hat{s},t,u), \sup_{(t,u)\in\hat{\mathcal{E}}_n^-} \big(-\hat{W}_n^{(r)}(\hat{s},t,u)\big)\Big\}. \tag{4.15}$$

Then the null hypothesis of no relevant change in the covariance operators is rejected at level $\alpha$, whenever

$$\hat{d}_\infty > \Delta + \frac{\check{K}_n^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{n}}, \tag{4.16}$$

where $\check{K}_n^{\{\lfloor R(1-\alpha)\rfloor\}}$ is the empirical $(1-\alpha)$-quantile of the bootstrap sample $\check{K}_n^{(1)}, \check{K}_n^{(2)}, \ldots, \check{K}_n^{(R)}$. The following result shows that the bootstrap test for the relevant hypotheses is consistent and has asymptotic level $\alpha$.

**Theorem 4.2** *Let the assumption of Theorem 4.1 be satisfied and furthermore assume that the random variable $M$ in (A3) is bounded.*

*(1) Under the null hypothesis $H_0 : d_\infty \leq \Delta$ of no relevant difference in the covariance operators, we have*

$$\lim_{n,R\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{\check{K}_n^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{n}}\Big) = \alpha,$$

*if $\Delta = d_\infty$ and, for any $R \in \mathbb{N}$,*

$$\lim_{n\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{\check{K}_n^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{n}}\Big) = 0,$$

*if $0 < d_\infty < \Delta$.*

*(2) Under the alternative $H_1 : d_\infty > \Delta$ of a relevant difference in the covariance operators, we have for any $R \in \mathbb{N}$,*

$$\liminf_{n\to\infty} \mathbb{P}\Big(\hat{d}_\infty > \Delta + \frac{\check{K}_n^{\{\lfloor R(1-\alpha)\rfloor\}}}{\sqrt{n}}\Big) = 1\,.$$

# 5 Finite sample properties

## 5.1 Simulation study

In this section we study the finite sample properties of the test procedures developed in this paper and we also compare it with some competing procedures from the literature, which can be used under similar assumptions as considered here. The empirical rejection probabilities of the different tests have been calculated by 1000 simulation runs and 200 bootstrap statistics are used for the calculation of the bootstrap quantiles in each run.

### 5.1.1 Two sample problem

**Classical hypotheses:** In the following we investigate the finite sample properties of the test (3.14) for the classical null hypothesis of equal covariance operators in (3.4). For the sake of comparison, we use the same scenarios as considered in Paparoditis and Sapatinas (2016) who developed a bootstrap test for the hypotheses (3.4). Paparoditis and Sapatinas (2016) also applied the FPC test developed by Fremdt et al. (2013) to these scenarios, such that a comparison with the method developed by these authors is also possible. To be precise, curves are generated according to the model

$$\begin{aligned} X_i(t) &= \sum_{k=1}^{10} \big\{2^{1/2}k^{-1/2}\sin(\pi kt)V_{i,k} + k^{-1/2}\cos(2\pi kt)W_{i,k}\big\} \\ Y_j(t) &= c\sum_{k=1}^{10} \big\{2^{1/2}k^{-1/2}\sin(\pi kt)\tilde{V}_{j,k} + k^{-1/2}\cos(2\pi kt)\tilde{W}_{j,k}\big\} \end{aligned} \tag{5.1}$$

($i = 1, \ldots, m,\ j = 1, \ldots, n$), where the random variables $V_{i,k}, W_{i,k}, \tilde{V}_{j,k}, \tilde{W}_{j,k}$ are independent and $t_5$-distributed. The constant $c$ determines if the null hypothesis ($c = 1$) holds or not ($c \neq 1$). In order to obtain functional data objects, the curves are evaluated at 500 equidistant points in $[0, 1]$ and then the Fourier basis consisting of 49 basis functions is used to transform these function values into a functional data object (using the function "Data2fd" from the "fda" R-package). In Table 1 we display empirical rejection probabilities for two different sample sizes and different choices of $c$. Paparoditis and Sapatinas (2016) state that the procedure proposed by Fremdt et al. (2013) achieves the best results if two FPCs are used to represent the data, and therefore, the results of this procedure were obtained for this case.

|  | $c = 1$ |  |  | $c = 1.2$ |  |  |
|---|---|---|---|---|---|---|
| $n, m$ | 1% | 5% | 10% | 1% | 5% | 10% |
| 25 | 0.9<br>(0, 0.3) | 4.2<br>(0.6, 2.5) | 11.8<br>(2.2, 8.2) | 3.0<br>(0, 0.5) | 13.4<br>(1.6, 5.0) | 24.7<br>(3.9, 14.7) |
| 50 | 0.8<br>(0, 0.6) | 3.6<br>(1.6, 3.2) | 8.6<br>(4.1, 7.6) | 6.6<br>(0.3, 0.8) | 22.4<br>(2.6, 9.8) | 35.0<br>(7.2, 23.9) |
|  | $c = 1.4$ |  |  | $c = 1.6$ |  |  |
| $n, m$ | 1% | 5% | 10% | 1% | 5% | 10% |
| 25 | 10.3<br>(0, 1.6) | 32.3<br>(1.1, 16.8) | 51.0<br>(5.2, 36.8) | 22.4<br>(0, 4.7) | 54.7<br>(1.0, 33.8) | 73.8<br>(9.5, 61.2) |
| 50 | 27.8<br>(0.2, 12.8) | 58.9<br>(6.5, 46.1) | 75.1<br>(22.1, 67.6) | 55.0<br>(1.4, 37.0) | 83.3<br>(28.5, 79.6) | 91.8<br>(55.9, 90.3) |
|  | $c = 1.8$ |  |  | $c = 2$ |  |  |
| $n, m$ | 1% | 5% | 10% | 1% | 5% | 10% |
| 25 | 34.9<br>(0, 10.4) | 72.2<br>(3.6, 55.7) | 87.4<br>(23.0, 82.3) | 45.3<br>(0, 17.7) | 81.9<br>(7.0, 66.6) | 93.3<br>(50.5, 89.2) |
| 50 | 73.0<br>(6.6, 61.2) | 93.4<br>(57.4, 91.5) | 97.5<br>(82.1, 96.6) | 83.0<br>(24.5, 74.2) | 96.4<br>(83.6, 93.7) | 98.6<br>(95.7, 97.7) |

Table 1: *Rejection probabilities of the test* (3.14) *for the classical hypotheses* (3.4)*. The case* $c = 1$ *corresponds to the null hypothesis. The numbers in the brackets display the empirical rejection probabilities of the tests proposed by Fremdt et al. (2013) and Paparoditis and Sapatinas (2016), respectively.*

We observe that under the null, i.e. $c = 1$, the nominal level is well approximated by the test (3.14) and the alternatives are detected with reasonable probability. Moreover, in all considered scenarios under the alternative, the new procedure achieves a better power than the tests of Paparoditis and Sapatinas (2016) and Fremdt et al. (2013).

**Relevant hypotheses:** We now investigate the finite sample properties of the decision rule (3.19) for testing relevant hypotheses of the form (3.3) in the two sample problem. For this purpose we define different processes including independent random functions, functional moving average processes and non-Gaussian random curves.
For the data generation, we proceed similarly as in Sections 6.3 and 6.4 of Aue et al. (2015). We consider 21 $B$-spline basis functions $\nu_1, \dots, \nu_{21}$ and restrict to functions in the linear space $\mathbb{H} = \operatorname{span}\{\nu_1, \dots, \nu_{21}\}$. Then, for a sample of size $m \in \mathbb{N}$, random functions $\varepsilon_1, \dots, \varepsilon_m \in \mathbb{H} \subset C([0,1])$ are defined by

$$\varepsilon_j = \sum_{i=1}^{21} N_{i,j} \nu_i \,, \qquad j = 1, \dots, m \,, \tag{5.2}$$

where $N_{1,j}, N_{2,j}, \dots, N_{21,j}$ are independent normally distributed random variables with expectation zero and variance $\mathrm{Var}(N_{i,j}) = \sigma_i^2 = 1/i^2$. Independent and identically distributed Gaussian random functions are then obtained by

$$X_i = \varepsilon_i \qquad (i = 1, \dots, m)\,, \tag{5.3}$$

and we call $\{X_i\}_{i=1}^m$ fIID process. In order to obtain independent non-Gaussian curves, we replace the normally distributed random coefficients in (5.2) by independent $t_5$-distributed random variables, that is $N_{i,j} \sim t_5 \sqrt{3/(5i^2)}$. Then, the variances of the coefficients are the same as for the fIID processes and the corresponding setting is called the non-Gaussian process.
Using the processes in (5.2), fMA(2) processes can be defined by

$$X_i = \varepsilon_i + \kappa_1\, \varepsilon_{i-1} + \kappa_2\, \varepsilon_{i-2} \qquad (i = 1, \dots, m) \tag{5.4}$$

where $\kappa_1, \kappa_2 \in \mathbb{R}$ are parameters defining the dependency (for initialization define $\varepsilon_{-1}, \varepsilon_0$ as independent copies of $\varepsilon_1$). In the simulations, we set $\kappa_1 = 0.7, \kappa_2 = 0$ to obtain an fMA(1) processes and $\kappa_1 = 0.5, \kappa_2 = 0.3$ for an fMA(2) processes.
In order to test for a relevant difference in the covariance operators of two populations, we generate an independent second sample, $\tilde{Y}_1, \dots, \tilde{Y}_n$, in the same way and multiply it by a constant $a$ such that $Y_i = a\,\tilde{Y}_i$ $(i = 1, \dots, n)$. Consequently,

$$|C_1(s,t) - C_2(s,t)| = |C_1(s,t)(a^2 - 1)| \tag{5.5}$$

where $C_1, C_2$ are the covariance operators of $X_1$ and $Y_1$, respectively.
In the case of fIID and non-Gaussian processes defined by (5.3), the maximum of the covariance operator is given by

$$\max_{s,t \in [0,1]} \mathrm{Cov}(X_1(s), X_1(t)) = \max_{s,t \in [0,1]} \sum_{i=1}^{D} \nu_i(s)\nu_i(t)/i^2 = 1$$

which is attained at the point $(s,t) = (0,0)$. Consequently, we obtain for the sup-norm

$$\|C_1 - C_2\|_\infty = |a^2 - 1|$$

in both cases and the extremal sets are defined by $\mathcal{E}^+ = \{(0,0)\}, \mathcal{E}^- = \emptyset$. For fMA(2) processes of the form (5.4), we obtain

$$\|C_1 - C_2\|_\infty = |a^2 - 1|\,(1 + \kappa_1^2 + \kappa_2^2)\,.$$

In Table 2 we display empirical rejection probabilities for the hypotheses in (3.3) for the different types of processes and different choices of the sample sizes. In each case, we use $a = \sqrt{2}$ and define $\Delta$ such that $\Delta = |a^2 - 1|$ in the fIID and non-Gaussian setting and $\Delta = |a^2 - 1|\,(1 + \kappa_1^2 + \kappa_2^2)$ in the fMA(1) and fMA(2) setting. Throughout this section we call this situation the boundary of the hypotheses (3.3). For the estimation of the extremal sets, we use $c_{m,n} = 0.1 \log(n + m)$

| | fIID | | | non-Gaussian | | | fMA(1) | | | fMA(2) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $m, n$ | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| 50, 50 | 1.0 | 4.6 | 8.7 | 0 | 1.7 | 5.4 | 1.2 | 5.2 | 10.5 | 1.6 | 5.1 | 10.1 |
| 50, 100 | 0.9 | 4.7 | 10.1 | 0.6 | 3.9 | 9.8 | 2.1 | 7.6 | 13.5 | 2.4 | 7.6 | 11.9 |
| 100, 100 | 0.9 | 3.9 | 9.1 | 0.5 | 3.1 | 9.0 | 1.4 | 5.7 | 10.8 | 1.0 | 4.2 | 10.8 |

Table 2: *Simulated level of the test* (3.19) *for a relevant difference in the covariance operators at the boundary of the hypotheses in* (3.3)*, that is* $\|C_1 - C_2\|_\infty = \Delta$*.*

in (3.17) and the block lengths in the bootstrap process (3.12) are chosen as $l_1 = l_2 = 1$ in the fIID cases, as $l_1 = l_2 = 2$ in the fMA(1) and as $l_1 = l_2 = 3$ in the fMA(2) case. We observe a reasonable approximation of the nominal level of the test at the boundary of the hypotheses in all cases under consideration. The nominal level in the interior of the hypotheses, that is $\|C_1 - C_2\|_\infty < \Delta$ is usually much smaller (these results are not displayed).

Next we study the properties of the test (3.19) under the alternative in (3.3). As before two independent identically distributed samples are generated where the second sample is multiplied by a factor $a$. The threshold $\Delta$ is fixed and then empirical rejection probabilities are simulated for different choices of the constant $a$, such that the properties stated in Theorem 3.2 can be visualized. The results are displayed in Figure 1 for fMA(1) processes (with $\kappa_1 = 0.7, \kappa_2 = 0$) and non-Gaussian random curves. The threshold in (3.3) is set to $\Delta = 1 + \kappa_1^2$ and $\Delta = 1$, respectively. As illustrated before, the nominal level is reasonably well approximated in both cases and with increasing factor $a$, the empirical rejection probability also increases. It can be observed that the empirical rejection probability increases slightly faster in the fMA(1) case. An explanation of this behaviour consists in the fact that for the same factor $a$, the true maximal difference of the covariance operators is greater in the fMA(1) than in the non-Gaussian case.

### 5.1.2 Change point problem

**Classical hypotheses:** We begin with a comparison of the test (4.12) for the classical hypotheses (3.4) with two procedures which were recently proposed by Sharipov and Wendler (2019) and are based on the sup and $L^2$-norm of the CUSUM statistic. Following these authors we generate data from the model

$$X_{n,i}(t) = \begin{cases} \varepsilon_{X,i}(t)\,, & i < k^* = \lfloor s^* n \rfloor + 1 \\ \varepsilon_{X,i}(t)(1 + d_1 + d_2(1 + \sin(2\pi t)))\,, & i \geq k^* \end{cases} \tag{5.6}$$

where $\varepsilon_{X,1}, \ldots, \varepsilon_{X,n}$ are independent standard Brownian motions. A sample size of $n = 100$ is considered and the true change point is defined by $k^* = 51$. The empirical rejection probabilities of the three tests are displayed in Table 3. The level ($d_1 = d_2 = 0$) is approximated very well by all procedures under consideration. Moreover, the test (4.12) proposed in this paper is at least competitive in all cases under consideration. In the case $d_1 = 0.4, d_2 = 0$ the procedures of

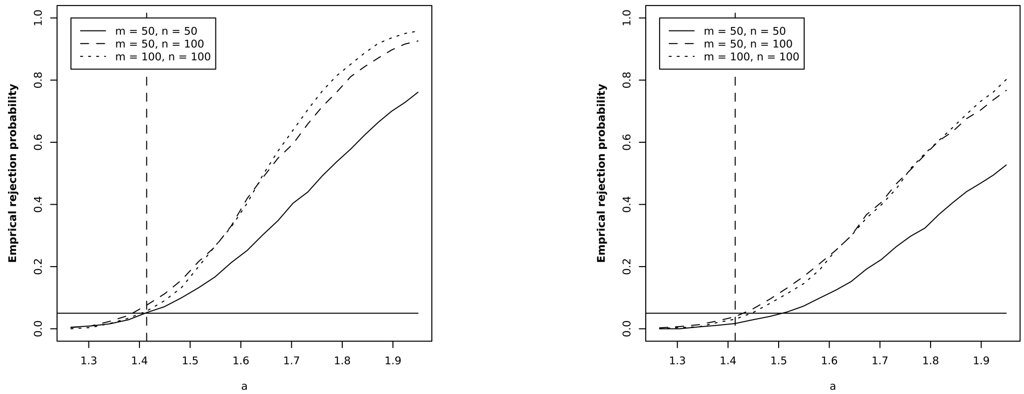


Figure 1: *Simulated rejection probabilities of the test* (3.19) *for a non-relevant difference in the covariance operators. Left panel fMA(1) with threshold* $\Delta = 1 + 0.7^2$*. Right panel non-Gaussian curves with threshold* $\Delta = 1$*. The second sample is multiplied by a for* $a = \sqrt{1.6}, \sqrt{1.7}, \ldots, \sqrt{3.8}$*, and the vertical lines represent the boundary of the null hypotheses (i.e.* $a = \sqrt{2}$*).*

Sharipov and Wendler (2019) perform slightly better but whenever $d_2 > 0$, the new procedure shows the best performance.

Next we provide a comparison with the procedure proposed by Stoehr et al. (2019). Following these authors, we simulate fAR(1) data where the errors (similar as in (5.2)) are defined by

$$e_j = \sum_{i=1}^{55} N_{i,j} \tilde{\nu}_i \,, \qquad j = 1, \ldots, n \,,$$

$\tilde{\nu}_1, \ldots, \tilde{\nu}_{55}$ denote the Fourier basis and the random coefficients $N_{1,j}, N_{2,j}, \ldots, N_{55,j}$ are independent normally distributed with expectation zero and variance $\mathrm{Var}(N_{i,j}) = \sigma_i^2$ ($i = 1, \ldots, 55$; $j = 1, \ldots, n$). The fAR(1) data are then defined by

$$X_{n,j} = \Psi(X_{n,j-1}) + e_j \,, \qquad j = 1, \ldots, n \,, \tag{5.7}$$

where the linear operator $\Psi$ is represented by a $55 \times 55$ matrix that is applied to the vector of the coefficients in the basis representation. Here the matrix with 0.4 on the diagonal and 0.1 on the superdiagonal and subdiagonal is chosen, such that the generated fAR(1) time series is stationary. For the alternative a change is inserted in the first $m$ leading eigendirections for $m = 2, 6, 25$ by adding an additional normally distributed noise term with variance $\sigma_\epsilon^2/m$ for the observations $X_{n,j}$ for $j > \lfloor 0.5n \rfloor$. The following three settings are considered:

Setting 1: $\sigma_i = 1$ for $i = 1, \ldots, 8$ and $\sigma_i = 0$ for $i = 9, \ldots, 55$, $\sigma_\epsilon = 1.5$

Setting 2: $\sigma_i = 3^{-i}$ for $i = 1, \ldots, 55$, $\sigma_\epsilon = 0.3$

Setting 3: $\sigma_i = i^{-1}$ for $i = 1, \ldots, 55$, $\sigma_\epsilon = 1$ .

| $d_1, d_2$ | 1% | 5% | 10% | $d_1, d_2$ | 1% | 5% | 10% |
|---|---|---|---|---|---|---|---|
| 0, 0 | 1.3<br>(0.4, 0.6) | 5.0<br>(4.4, 4.7) | 9.9<br>(10.0, 10.8) | 0.4, 0 | 19.3<br>(16.1, 19.8) | 44.5<br>(46.8, 50.1) | 61.0<br>(63.2, 65.4) |
| 0.8, 0 | 60.0<br>(56.0, 58.8) | 88.4<br>(88.0, 88.4) | 95.2<br>(96.0, 95.5) | 0, 0.4 | 22.4<br>(9.8, 12.5) | 48.9<br>(33.0, 38.6) | 65.3<br>(49.4, 55.4) |
| 0, 0.8 | 69.8<br>(45.8, 50.3) | 93.6<br>(82.9, 85.8) | 98.0<br>(93.8, 94.6) | 0.4, 0.4 | 63.4<br>(44.2, 49.1) | 89.3<br>(81.1, 82.2) | 95.8<br>(91.6, 92.1) |

Table 3: *Empirical rejection probabilities of the bootstrap test* (4.12) *for the classical hypotheses* (3.4) *of a structural break in the covariance operator. The numbers in the brackets display the empirical rejection probabilities of the test proposed in Sharipov and Wendler (2019) based on the supremum type integral type CUSUM statistic (for* $p = 3$*).*

| $m$ | Setting 1 | Setting 2 | Setting 3 |
|---|---|---|---|
| 0 | 4.7 (3.1) | 8.1 (5.0) | 3.9 (4.6) |
| 2 | 37.2 (22.8) | 92.5 (50.5) | 86.2 (30.4) |
| 6 | 81.1 (20.4) | 99.9 (98.9) | 99.9 (94.8) |
| 25 | 100 (29.0) | 100 (92.3) | 100 (97.3) |

Table 4: *Empirical rejection probabilities (at level* 5%*) of the bootstrap test* (4.12) *and the dimension reduction approach proposed in Stoehr et al. (2019) (numbers in the brackets).*

The empirical rejection probabilities of the test (4.12) with block length $l = 6$ and the test based on dimension reduction developed in Stoehr et al. (2019) are displayed in Table 4. We observe that in all cases under consideration the procedure proposed here yields an improvement with respect to the power. Note that Stoehr et al. (2019) also considered test procedures based on fully functional and weighted functional statistics. As these methods considerably overestimate the test level (see Figure 2 in the same reference), these procedures are not included in the comparison.

**Relevant hypotheses:** We conclude this section investigating the finite sample properties of the test defined by (4.16) for the hypotheses (3.3) of a relevant change in the covariance operator. For this purpose we consider similar scenarios as in Section 5.1.1. In all cases, the location of the change is set to $s^* = 0.5$ and the observations after the change point are multiplied by a constant $a$ such that (5.5) holds. For the estimation of the extremal sets, the parameter in (4.14) is set as $c_n = 0.1 \log(n)$.
In Table 5 empirical rejection probabilities are displayed for different processes at the boundary of the null hypothesis i.e. the observations after the change point are multiplied by $a = 2$ and the threshold $\Delta$ is defined in each case such that $\|C_1 - C_2\|_\infty = \Delta$. For fIID and non-Gaussian data the block length in (4.8) is set to $l = 1$ and the threshold is given by $\Delta = a^2 - 1$. The fMA(1) and fMA(2) data are defined by (5.4) with $\kappa_1 = 0.7, \kappa_2 = 0$ and $\kappa_1 = 0.5, \kappa_2 = 0.3$,

| | fIID | | | non-Gaussian | | | fMA(1) | | | fMA(2) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| 100 | 1.1 | 3.8 | 9.5 | 0 | 0.8 | 5.3 | 0.8 | 4.9 | 13.7 | 1.3 | 6.0 | 11.7 |
| 200 | 0.7 | 4.6 | 10.1 | 0.3 | 3.1 | 8.4 | 1.3 | 4.9 | 9.8 | 0.7 | 4.9 | 10.5 |

Table 5: *Simulated level of the test* (4.16) *for the hypotheses defined by* (3.3) *at the boundary of the hypotheses, that is* $\|C_1 - C_2\|_\infty = \Delta$.

respectively. The threshold parameter is set to $\Delta = (a^2 - 1)(1 + \kappa_1^2 + \kappa_2^2)$ in both cases and the block length in (4.8) is set to $l = 2$ and $l = 3$, respectively.

We observe that the nominal level is reasonably well approximated in most cases under consideration especially for the sample size $n = 200$. Only in the non-Gaussian case, the nominal level is underestimated for the sample size $N = 100$, but the approximation improves considerably for the sample size $N = 200$.

In Table 6, we show the empirical rejection probabilities of the test (4.16) and also for the test developed in Dette et al. (2020b) for scenarios in the interior of the null hypothesis of no relevant change point as well as under the alternative. We consider independent identically distributed Gaussian (fIID) and fMA(2) data and multiply the observations after the change point by different values $a = 1.8, 1.9, 2, 2.2, 2.4, 2.6$. In the fIID case the threshold parameter is given by $\Delta = 3$ and in the fMA(2) case it is $\Delta = 3 \cdot (1 + \kappa_1^2 + \kappa_2^2)$ (where still $\kappa_1 = 0.5, \kappa_2 = 0.3$). Consequently, the case $a = 2$ always corresponds to the boundary of the null hypothesis, and the cases $a < 2$ and $a > 2$ represent the interior of the null hypothesis and alternative. Since the procedure developed by Dette et al. (2020b) is based on a different metric, the threshold parameter $\Delta$ in the relevant hypotheses (1.2) is set to

$$\Delta = \int_{[0,1]} \int_{[0,1]} \{(1 - 2^2) C_1(s,t)\}^2 ds dt$$

for this test procedure. Consequently the boundary of the null hypothesis of no relevant change in the covariance operators (w.r.t. the corresponding metric) is also obtained for the factor $a = 2$ for both data models.

We mention again that the nominal level at the boundary of the hypotheses is reasonably well approximated by the test (4.16) while the test procedure developed in Dette et al. (2020b) is more conservative. In the interior of the null hypothesis ($a < 2$) the rejection probabilities of both tests are strictly smaller than the nominal level. This property is desirable as it means that the probability of a type I error is small in situations with a large deviation from the alternative. On the other hand, under the alternative the new test (4.16) has substantially more power than the test developed in Dette et al. (2020b).

| | fIID | | fMA(2) | |
|---|---|---|---|---|
| $a$ | $n = 100$ | $n = 200$ | $n = 100$ | $n = 200$ |
| 1.8 | 0.3 (0.4) | 0 (0) | 1.3 (0.1) | 1.0 (0.1) |
| 1.9 | 1.8 (0.9) | 0.1 (0.5) | 3.4 (0.4) | 1.4 (0.4) |
| 2.0 | 3.8 (2.3) | 4.6 (3.2) | 6.0 (1.0) | 4.9 (1.4) |
| 2.2 | 21.5 (9.8) | 33.6 (25.4) | 19.4 (6.2) | 27.2 (11.1) |
| 2.4 | 47.0 (23.3) | 74.9 (51.2) | 40.0 (15.6) | 65.9 (31.2) |
| 2.6 | 73.0 (37.9) | 96.0 (70.6) | 63.3 (26.5) | 88.0 (49.7) |

Table 6: *Simulated rejection probabilities of the test* (4.16) *for the hypotheses* (3.3) *of a relevant change in the covariance operator considering fIID and fMA(2) data (level* 5%*). The cases* $a < 2, a = 2$ *and* $a > 2$ *correspond to the interior, boundary of the null hypothesis and to the alternative. The numbers in brackets represent the empirical rejection probabilities of the procedure developed in Dette et al. (2020b).*

## 5.2 Data Example

Similar as Fremdt et al. (2013) and Paparoditis and Sapatinas (2016) we consider egg-laying curves of medflies (Mediterranean fruit flies, Ceratitis capitata). The original data consists of the number of eggs which were laid on each day during the lifetime of 1000 female medflies and a detailed description of the experiment can be found in Carey et al. (1998). Only medflies which lived at least 34 days are considered and split into two samples, the medflies which lived at most 43 days and those which lived at least 44 days. A Fourier basis consisting of 49 basis functions is used to transform the discrete observations to functional data $(X_j : j = 1, \ldots, 256)$ and $(Y_j : j = 1, \ldots, 278)$. The expressions $X_i(t)$ and $Y_j(t)$ denote the number of eggs which were laid on day $\lfloor 30t \rfloor$ by the $i$th short-lived and the $j$th long-lived medfly relative to the total number of eggs laid in the whole lifetime of the $i$th short-lived and the $j$th long-lived medfly, respectively ($t \in [0, 1], i = 1, \ldots, 256, j = 1, \ldots, 278$). First, the test (3.14) is used to study the classical hypotheses in (3.4). The window parameters in (3.12) are set to $l_1 = l_2 = 1$ since the egg-laying curves corresponding to the different medflies can be regarded as independent. For the calculation of critical values, 200 bootstrap samples are generated. The classical null hypothesis of equal covariance operators is then rejected at level 5% and can not be rejected at level 1%. The outcome when using the procedure developed in Fremdt et al. (2013) depends on the choice of the number of considered functional principal components $p$ and the procedure developed in Paparoditis and Sapatinas (2016) yields a $p$-value of 0.3% (see Table 3 in Paparoditis and Sapatinas (2016)). In Table 7 the empirical rejection probabilities of the test (3.19) are displayed for the relevant hypotheses in (3.3) for different choices of the threshold parameter $\Delta$. It can be seen that even for $\Delta = 0.0003$ i.e. when a maximal deviation of only 0.0003 is tolerated, the null hypothesis of no relevant difference between the covariance operators can not be rejected at all considered test levels. For $\Delta = 0.0002$ the null can be rejected at level 10% and for $\Delta = 0.0001$ also at level 5%. Although the classical null hypothesis of equal covariance operators is rejected at level 5%, these results may raise the question if the detected difference is really of practical relevance.

| $\Delta$ | 1% | 5% | 10% |
|---|---|---|---|
| 0.0001 | FALSE | TRUE | TRUE |
| 0.0002 | FALSE | FALSE | TRUE |
| 0.0003 | FALSE | FALSE | FALSE |

Table 7: *Summary of the outcome of the test* (3.19) *for the relevant hypotheses* (3.3) *for different values of* $\Delta$ *for the relative egg-laying curves of medflies. The label TRUE means that the null hypothesis is rejected and the label FALSE means that the null hypothesis is not rejected.*

# 6 Appendix: Proofs of main results

## 6.1 Proof of Theorem 2.1

We apply the central limit theorem as formulated in Theorem 2.1 in Dette et al. (2020a) to the sequence of $C(T^2)$-valued random variables $((Z_j-\mu)^{\check{\otimes}2})_{j\in\mathbb{N}} = (\eta_j^{\check{\otimes}2})_{j\in\mathbb{N}}$.
It can be easily seen that conditions (A1), (A2) and (A4) in this reference are satisfied. In order to see that the remaining condition (A3) also holds, we use the triangle inequality and Assumption 2.1 of the present work to obtain, for any $j\in\mathbb{N}$ and $s,t,s',t'\in T$,

$$\begin{aligned}|\eta_j(s)\eta_j(t)-\eta_j(s')\eta_j(t')| &\le |\eta_j(s)(\eta_j(t)-\eta_j(t'))| + |\eta_j(t')(\eta_j(s)-\eta_j(s'))| \\ &\le \|\eta_j\|_\infty \big(|\eta_j(t)-\eta_j(t')| + |\eta_j(s)-\eta_j(s')|\big) \\ &\le \|\eta_j\|_\infty\, M\, \big(\rho(t,t')+\rho(s,s')\big) \\ &\lesssim \|\eta_j\|_\infty\, M\, \rho_{\max}\big((t,s),(t',s')\big)\end{aligned}$$

where $\mathbb{E}\big[(\|\eta_j\|_\infty\, M)^J\big] \le \tilde{K} < \infty$ by (A3). Now observe that

$$\frac{1}{\sqrt{n}}\sum_{j=1}^n (Z_j-\bar{Z}_n)^{\check{\otimes}2} = \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j^{\check{\otimes}2} - \frac{1}{\sqrt{n}}\Big(\frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j\Big)^{\check{\otimes}2} = \frac{1}{\sqrt{n}}\sum_{j=1}^n \eta_j^{\check{\otimes}2} + o_{\mathbb{P}}(1)$$

which yields the claim since Theorem 2.1 in Dette et al. (2020a) can be applied to the sequence $(\eta_j^{\check{\otimes}2})_{j\in\mathbb{N}}$ as shown above.

## 6.2 Proof of Proposition 3.1

As the samples are independent, it directly follows from Theorem 2.1 that

$$
\begin{aligned}
&\sqrt{m+n}\Big(\frac{1}{m}\sum_{j=1}^{m}(\tilde{X}_{m,j}^{\check{\otimes}2}-C_1),\ \frac{1}{n}\sum_{j=1}^{n}(\tilde{Y}_{n,j}^{\check{\otimes}2}-C_2)\Big)\\
&\qquad=\sqrt{m+n}\Big(\frac{1}{m}\sum_{j=1}^{m}(\eta_{1,j}^{\check{\otimes}2}-C_1),\ \frac{1}{n}\sum_{j=1}^{n}(\eta_{2,j}^{\check{\otimes}2}-C_2)\Big)+o_{\mathbb{P}}(1)\rightsquigarrow\Big(\frac{1}{\sqrt{\lambda}}\ Z_1,\frac{1}{\sqrt{1-\lambda}}\ Z_2\Big)
\end{aligned}
$$

in $C([0,1]^2)^2$ as $m,n\to\infty$, where $Z_1$ and $Z_2$ are independent, centred Gaussian processes defined by their long-run covariance operators (3.8) and (3.9). By the continuous mapping theorem it follows that

$$
Z_{m,n}=\sqrt{m+n}\left(\frac{1}{m}\sum_{j=1}^{m}\tilde{X}_{m,j}^{\check{\otimes}2}-\frac{1}{n}\sum_{j=1}^{n}\tilde{Y}_{n,j}^{\check{\otimes}2}-(C_1-C_2)\right)\rightsquigarrow Z \tag{6.1}
$$

in $C([0,1]^2)$ as $m,n\to\infty$, where $Z$ is again a centred Gaussian process with covariance operator (3.7).

If $d_\infty=0$, the convergence in (6.1) together with the continuous mapping yield (3.6). If $d_\infty>0$, the asymptotic distribution of $\hat{d}_\infty$ can be deduced from Theorem B.1 in the online supplement of Dette et al. (2020a) or alternatively from the results in Cárcamo et al. (2020).

## 6.3 Proof of Theorem 3.1 and 3.2

**Proof of Theorem 3.1.** Using similar arguments as in the proof of Theorem 2.1, it follows that the process $\hat{B}_{m,n}^{(r)}$ in (3.12) admits the stochastic expansion

$$
\begin{aligned}
\hat{B}_{m,n}^{(r)}=\sqrt{n+m}\Bigg\{&\frac{1}{m}\sum_{k=1}^{m-l_1+1}\frac{1}{\sqrt{l_1}}\Bigg(\sum_{j=k}^{k+l_1-1}\eta_{1,j}^{\check{\otimes}2}-\frac{l_1}{m}\sum_{i=1}^{m}\eta_{1,j}^{\check{\otimes}2}\Bigg)\xi_k^{(r)}\\
&-\frac{1}{n}\sum_{k=1}^{n-l_2+1}\frac{1}{\sqrt{l_2}}\Bigg(\sum_{j=k}^{k+l_2-1}\eta_{2,j}^{\check{\otimes}2}-\frac{l_2}{n}\sum_{i=1}^{n}\eta_{2,j}^{\check{\otimes}2}\Bigg)\zeta_k^{(r)}\Bigg\}+o_{\mathbb{P}}(1)\,,
\end{aligned}
$$

and the sequences $(\eta_{1,j}^{\check{\otimes}2})_{j\in\mathbb{N}}$ and $(\eta_{2,j}^{\check{\otimes}2})_{j\in\mathbb{N}}$ satisfy Assumption 2.1 in Dette et al. (2020a). Thus, similar arguments as in the proof of Theorem 3.3 in the same reference yield

$$
\big(Z_{m,n},\hat{B}_{m,n}^{(1)},\dots,\hat{B}_{m,n}^{(R)}\big)\rightsquigarrow(Z,Z^{(1)},\dots,Z^{(R)}) \tag{6.2}
$$

in $C([0,1]^2)^{R+1}$ as $m,n\to\infty$ where the process $Z_{m,n}$ is defined in (6.1) and the random functions $Z^{(1)},\dots,Z^{(R)}$ are independent copies of $Z$ which is also defined in (6.1).

If $d_\infty=0$, the continuous mapping theorem implies

$$
\big(\sqrt{m+n}\,\hat{d}_\infty,\ T_{m,n}^{(1)},\dots,T_{m,n}^{(R)}\big)\xrightarrow{\mathcal{D}}(T,\ T^{(1)},\dots,T^{(R)}) \tag{6.3}
$$

in $\mathbb{R}^{R+1}$ as $m, n \to \infty$ where the statistic $\hat{d}_\infty$ is defined by (3.5), the bootstrap statistics $T^{(1)}_{m,n}, \ldots, T^{(R)}_{m,n}$ are defined by (3.13) and the random variables $T^{(1)}, \ldots, T^{(R)}$ are independent copies of $T$ which is defined by (3.6). Now, Lemma 4.2 in Bücher and Kojadinovic (2019) directly implies (3.15), that is

$$\lim_{m,n,R\to\infty} \mathbb{P}\bigg(\hat{d}_\infty > \frac{T^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}}{\sqrt{m+n}}\bigg) = \alpha\,.$$

For the application of this result, it is required that the distribution of the random variable $T$ has a continuous distribution function, which follows from Gaenssler et al. (2007). In order to show the consistency of test (3.14) in the case $d_\infty > 0$, write

$$\mathbb{P}\bigg(\hat{d}_\infty > \frac{T^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}}{\sqrt{m+n}}\bigg) = \mathbb{P}\big(\sqrt{m+n}\,(\hat{d}_\infty - d_\infty) + \sqrt{m+n}\,d_\infty > T^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}\big)$$

and note that, given (6.3) and (3.10), the assertion in (3.16) follows by simple arguments.

**Proof of Theorem 3.2.** First note that the same arguments as in the proof of Theorem 3.6 in Dette et al. (2020a) show that the estimators of the extremal sets defined by (3.17) are consistent that is

$$d_H(\hat{\mathcal{E}}^{\pm}_{m,n}, \mathcal{E}^{\pm}) \xrightarrow[m,n\to\infty]{\mathbb{P}} 0\,,$$

where $d_H$ denotes the Hausdorff distance. Thus, given the convergence in (6.2), the arguments in the proof of Theorem 3.7 in the same reference yield

$$\big(\sqrt{n+m}\ (\hat{d}_\infty - d_\infty),\ K^{(1)}_{m,n}, \ldots, K^{(R)}_{m,n}\big) \xrightarrow{\mathcal{D}} (T(\mathcal{E}),\ T^{(1)}(\mathcal{E}), \ldots, T^{(R)}(\mathcal{E})) \tag{6.4}$$

in $\mathbb{R}^{R+1}$ as $m, n \to \infty$ where the statistic $\hat{d}_\infty$ is defined by (3.5), the bootstrap statistics $K^{(1)}_{m,n}, \ldots, K^{(R)}_{m,n}$ are defined by (3.18) and the random variables $T^{(1)}(\mathcal{E}), \ldots, \mathcal{T}^{(\mathcal{R})}(\mathcal{E})$ are independent copies of $T(\mathcal{E})$ which is defined by (3.10). Note that this convergence holds true under the null and the alternative hypothesis.
If $\Delta = d_\infty$, Lemma 4.2 in Bücher and Kojadinovic (2019) directly implies (3.20) and again the results in Gaenssler et al. (2007) ensure that the limit $T(\mathcal{E})$ has a continuous distribution function.
If $\Delta \neq d_\infty$, write

$$\mathbb{P}\bigg(\hat{d}_\infty > \Delta + \frac{K^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}}{\sqrt{n+m}}\bigg) = \mathbb{P}\big(\sqrt{m+n}\,(\hat{d}_\infty - d_\infty) + \sqrt{m+n}\,(d_\infty - \Delta) > K^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}\big)\,.$$

Then it follows from (6.4) and simple arguments that, for any $R \in \mathbb{N}$,

$$\lim_{m,n\to\infty} \mathbb{P}\bigg(\hat{d}_\infty > \Delta + \frac{K^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}}{\sqrt{n+m}}\bigg) = 0 \quad \text{and} \quad \liminf_{m,n\to\infty} \mathbb{P}\bigg(\hat{d}_\infty > \Delta + \frac{K^{\{\lfloor R(1-\alpha)\rfloor\}}_{m,n}}{\sqrt{n+m}}\bigg) = 1$$

if $\Delta > d_\infty$ and $\Delta < d_\infty$, respectively. This proves the remaining assertions of Theorem 3.2.

## 6.4 Proof of Proposition 4.1

Let $C_{n,j}$ denote the covariance operator of $X_{n,j}$ defined by $C_{n,j}(s,t) = \mathrm{Cov}(X_{n,j}(s), X_{n,j}(t))$ and consider the sequential process

$$
\begin{aligned}
\hat{\mathbb{V}}_n(s) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} (\tilde{X}_{n,j}^{\check{\otimes}2} - C_{n,j}) + \sqrt{n}\Big(s - \frac{\lfloor sn \rfloor}{n}\Big)\big(\tilde{X}_{n,\lfloor sn \rfloor+1}^{\check{\otimes}2} - C_{n,\lfloor sn \rfloor+1}\big) \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} (\tilde{\eta}_{n,j}^{\check{\otimes}2} - C_{n,j}) + \sqrt{n}\Big(s - \frac{\lfloor sn \rfloor}{n}\Big)\big(\tilde{\eta}_{n,\lfloor sn \rfloor+1}^{\check{\otimes}2} - C_{n,\lfloor sn \rfloor+1}\big) + o_{\mathbb{P}}(1)
\end{aligned}
$$

which is an element of $C([0,1], C([0,1]^2))$. Note that $\{\hat{\mathbb{V}}_n(s)\}_{s\in[0,1]}$ can equivalently be regarded as an element of $C([0,1]^3)$ and we have the representation

$$
\hat{\mathbb{V}}_n = \tilde{\mathbb{V}}_{1,n} + \tilde{\mathbb{V}}_{2,n}\,, \tag{6.5}
$$

where the processes $\tilde{\mathbb{V}}_{1,n}, \tilde{\mathbb{V}}_{2,n} \in C([0,1]^3)$ are defined by

$$
\begin{aligned}
\tilde{\mathbb{V}}_{1,n}(s,t,u) &= \hat{\mathbb{V}}_{1,n}(s,t,u)\mathbb{1}\{\lfloor sn \rfloor < \lfloor s^*n \rfloor\} + \hat{\mathbb{V}}_{1,n}(\lfloor s^*n \rfloor/n, t, u)\mathbb{1}\{\lfloor sn \rfloor \geq \lfloor s^*n \rfloor\} \\
\tilde{\mathbb{V}}_{2,n}(s,t,u) &= (\hat{\mathbb{V}}_{2,n}(s,t,u) - \hat{\mathbb{V}}_{2,n}(\lfloor s^*n \rfloor/n, t, u))\mathbb{1}\{\lfloor sn \rfloor \geq \lfloor s^*n \rfloor\}
\end{aligned}
$$

$(s,t,u \in [0,1])$ and

$$
\hat{\mathbb{V}}_{l,n}(s) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} (\eta_{l,j}^{\check{\otimes}2} - C_l) + \sqrt{n}\Big(s - \frac{\lfloor sn \rfloor}{n}\Big)\big(\eta_{l,\lfloor sn \rfloor+1}^{\check{\otimes}2} - C_l\big) \qquad (l = 1,2)\,.
$$

Recall the definition of the array $(\tilde{\eta}_{n,j} \colon n \in \mathbb{N}, j = 1, \ldots, n)$ in (4.1). By Theorem 2.2 in Dette et al. (2020a) it follows that

$$
\hat{\mathbb{V}}_{l,n} \rightsquigarrow \mathbb{V}_l \qquad (l = 1,2)
$$

in $C([0,1]^3)$, where $\mathbb{V}_l$ is a centred Gaussian measure on $C([0,1]^3)$ characterized by the covariance operator

$$
\mathrm{Cov}\big(\mathbb{V}_l(s,t,u), \mathbb{V}_l(s',t',u')\big) = (s \wedge s')\, \mathbb{C}_l((t,u),(t',u')), \qquad l = 1,2
$$

and the long-run covariance operator $\mathbb{C}_l$ is defined in (4.6). From the continuous mapping theorem we obtain

$$
\tilde{\mathbb{V}}_{l,n} \rightsquigarrow \tilde{\mathbb{V}}_l \qquad (l = 1,2) \tag{6.6}
$$

in $C([0,1]^3)$, where $\tilde{\mathbb{V}}_1, \tilde{\mathbb{V}}_2$ are centred Gaussian measures on $C([0,1]^3)$ characterized by

$$
\tilde{\mathbb{V}}_1(s,t,u) = \mathbb{V}_1(s \wedge s^*, t, u)\,, \quad \tilde{\mathbb{V}}_2(s,t,u) = (\mathbb{V}_2(s,t,u) - \mathbb{V}_2(s^*,t,u))\mathbb{1}\{s \geq s^*\}
$$

with covariance operators

$$\operatorname{Cov}\big(\tilde{\mathbb{V}}_1(s,t,u),\tilde{\mathbb{V}}_1(s',t',u')\big) = (s\wedge s'\wedge s^*)\,\mathbb{C}_1((t,u),(t',u'))$$
$$\operatorname{Cov}\big(\tilde{\mathbb{V}}_2(s,t,u),\tilde{\mathbb{V}}_2(s',t',u')\big) = (s\wedge s'-s^*)_+\,\mathbb{C}_2((t,u),(t',u'))\,.$$

In the following we will show the weak convergence

$$\hat{\mathbb{V}}_n \rightsquigarrow \mathbb{V} \tag{6.7}$$

in $C([0,1]^3)$ as $n\to\infty$, where $\mathbb{V}\in C([0,1]^3)$ is a centred Gaussian random variable characterized by its covariance operator

$$\operatorname{Cov}(\mathbb{V}(s,t,u),\mathbb{V}(s',t',u')) = (s\wedge s'\wedge s^*)\,\mathbb{C}_1((t,u),(t',u')) + (s\wedge s'-s^*)_+\,\mathbb{C}_2((t,u),(t',u'))$$

and the long-run covariance operators $\mathbb{C}_1,\mathbb{C}_2$ are defined by (4.6). The convergence in (6.6) implies that the processes $\tilde{\mathbb{V}}_{1,n},\tilde{\mathbb{V}}_{2,n}$ are asymptotically tight and the representation in (6.5) yields that $\hat{\mathbb{V}}_n$ is asymptotically tight as well (see Section 1.5 in Van der Vaart and Wellner, 1996). In order to prove the convergence in (6.7) it consequently remains to show the convergence of the finite dimensional distributions. For this, we utilize the Crámer-Wold device and show that

$$\tilde{Z}_n = \sum_{j=1}^{q} c_j\hat{\mathbb{V}}_n(s_j,t_j,u_j) = \sum_{j=1}^{q} c_j\big\{\tilde{\mathbb{V}}_{1,n}(s_j,t_j,u_j)+\tilde{\mathbb{V}}_{2,n}(s_j,t_j,u_j)\big\}$$
$$\xrightarrow{\mathcal{D}} \tilde{Z} = \sum_{j=1}^{q} c_j\mathbb{V}(s_j,t_j,u_j)$$

for any $(s_1,t_1,u_1),\ldots,(s_q,t_q,u_q)\in[0,1]^3$, $c_1,\ldots,c_q\in\mathbb{R}$ and $q\in\mathbb{N}$. Asymptotic normality of $\tilde{Z}_n$ can be proved by the same arguments as in the proof of Theorem 2.1 in Dette et al. (2020a) and it remains to show that the variance of the random variable $\tilde{Z}_n$ converges to the variance of $\tilde{Z}$. Using (3.17) in Dehling and Philipp (2002) and assumptions (A2) and (A4) we obtain for any $(s,t,u),(s',t',u')\in[0,1]^3$

$$\begin{aligned}
&\operatorname{Cov}(\tilde{\mathbb{V}}_{1,n}(s,t,u),\tilde{\mathbb{V}}_{2,n}(s',t',u'))\\
&= \frac{1}{n}\sum_{j=1}^{\lfloor (s\wedge s^*)n\rfloor}\ \sum_{i=\lfloor s^*n\rfloor+1}^{\lfloor s'n\rfloor} \operatorname{Cov}(\tilde{\eta}_{n,j}^{\check{\otimes}2}(t,u),\tilde{\eta}_{n,i}^{\check{\otimes}2}(t',u')) + o(1)\\
&\lesssim \frac{1}{n}\sum_{j=1}^{\lfloor (s\wedge s^*)n\rfloor}\ \sum_{i=\lfloor s^*n\rfloor+1}^{\lfloor s'n\rfloor} \|\tilde{\eta}_{n,j}^{\check{\otimes}2}(t,u)\|_2\,\|\tilde{\eta}_{n,i}^{\check{\otimes}2}(t',u')\|_2\,\varphi(i-j)^{1/2} + o(1)\\
&\lesssim \frac{1}{n}\sum_{j=1}^{\lfloor (s\wedge s^*)n\rfloor}\ \sum_{i=\lfloor s^*n\rfloor+1}^{\lfloor s'n\rfloor} \varphi(i-j)^{1/2} + o(1)\\
&\lesssim \frac{1}{n}\sum_{i=1}^{\lfloor s'n\rfloor-1} i\varphi(i)^{1/2} + o(1) \underset{n\to\infty}{\longrightarrow} 0\,,
\end{aligned} \tag{6.8}$$

where the symbol "$\lesssim$" means less or equal up to a constant independent of $n$, and $\|X\|_2 = \mathbb{E}[X^2]^{1/2}$ denotes the $L^2$-norm of a real valued random variable $X$ (also note that we implicitly assume $\sum_{i=j}^k a_i = 0$ if $k < j$). Furthermore, assuming without loss of generality that $s \le s'$, we have

$$
\begin{aligned}
&\operatorname{Cov}(\tilde{\mathbb{V}}_{1,n}(s,t,u), \tilde{\mathbb{V}}_{1,n}(s',t',u')) \\
&= \frac{1}{n} \sum_{j=1}^{\lfloor (s\wedge s^*)n \rfloor} \sum_{i=1}^{\lfloor (s'\wedge s^*)n \rfloor} \operatorname{Cov}(\eta_{1,j}^{\check{\otimes}2}(t,u), \eta_{1,i}^{\check{\otimes}2}(t',u')) + o(1) \\
&= \frac{1}{n} \sum_{j=1}^{\lfloor (s\wedge s^*)n \rfloor} \Big( \sum_{i=1}^{\lfloor (s\wedge s^*)n \rfloor} + \sum_{i=\lfloor (s\wedge s^*)n \rfloor+1}^{\lfloor (s'\wedge s^*)n \rfloor} \Big) \operatorname{Cov}(\eta_{1,j}^{\check{\otimes}2}(t,u), \eta_{1,i}^{\check{\otimes}2}(t',u')) + o(1) \\
&= \frac{1}{n} \sum_{j=1}^{\lfloor (s\wedge s^*)n \rfloor} \sum_{i=1}^{\lfloor (s\wedge s^*)n \rfloor} \operatorname{Cov}(\eta_{1,j}^{\check{\otimes}2}(t,u), \eta_{1,i}^{\check{\otimes}2}(t',u')) + o(1)\,,
\end{aligned}
$$

where the last equality follows by the same arguments as used in (6.8). For the remaining expression we use the dominated convergence theorem to obtain

$$
\begin{aligned}
&\frac{1}{n} \sum_{j=1}^{\lfloor (s\wedge s^*)n \rfloor} \sum_{i=1}^{\lfloor (s\wedge s^*)n \rfloor} \operatorname{Cov}(\eta_{1,j}^{\check{\otimes}2}(t,u), \eta_{1,i}^{\check{\otimes}2}(t',u')) \\
&= \sum_{i=-(\lfloor (s\wedge s^*)n \rfloor - 1)}^{\lfloor (s\wedge s^*)n \rfloor - 1} \frac{\lfloor (s\wedge s^*)n \rfloor - |i|}{n} \operatorname{Cov}(\eta_{1,0}^{\check{\otimes}2}(t,u), \eta_{1,i}^{\check{\otimes}2}(t',u')) \underset{n\to\infty}{\longrightarrow} (s\wedge s^*)\, \mathbb{C}_1((t,u),(t',u'))
\end{aligned}
$$

which means that for any $(s,t,u), (s',t',u') \in [0,1]^3$

$$
\operatorname{Cov}(\tilde{\mathbb{V}}_{1,n}(s,t,u), \tilde{\mathbb{V}}_{1,n}(s',t',u')) \underset{n\to\infty}{\longrightarrow} (s\wedge s'\wedge s^*)\, \mathbb{C}_1((t,u),(t',u'))\,.
$$

By similar arguments we obtain

$$
\operatorname{Cov}(\tilde{\mathbb{V}}_{2,n}(s,t,u), \tilde{\mathbb{V}}_{2,n}(s',t',u')) \underset{n\to\infty}{\longrightarrow} (s\wedge s' - s^*)_+\, \mathbb{C}_2((t,u),(t',u'))
$$

and therefore we have

$$
\begin{aligned}
\operatorname{Var}(\tilde{Z}_n) &= \sum_{j=1}^q \sum_{j'=1}^q c_j c_{j'} \operatorname{Cov}(\hat{\mathbb{V}}_n(s_j,t_j,u_j), \hat{\mathbb{V}}_n(s_{j'},t_{j'},u_{j'})) \\
&= \sum_{j=1}^q \sum_{j'=1}^q c_j c_{j'} \big\{ \operatorname{Cov}(\tilde{\mathbb{V}}_{1,n}(s_j,t_j,u_j), \tilde{\mathbb{V}}_{1,n}(s_{j'},t_{j'},u_{j'})) \\
&\qquad\qquad + \operatorname{Cov}(\tilde{\mathbb{V}}_{2,n}(s_j,t_j,u_j), \tilde{\mathbb{V}}_{2,n}(s_{j'},t_{j'},u_{j'})) \big\} + o(1) \\
&\underset{n\to\infty}{\longrightarrow} \sum_{j=1}^q \sum_{j'=1}^q c_j c_{j'} \operatorname{Cov}(\mathbb{V}(s_j,t_j,u_j), \mathbb{V}(s_{j'},t_{j'},u_{j'})) = \operatorname{Var}(\tilde{Z})
\end{aligned}
$$

which finally proves (6.7).
Next we define the $C([0,1]^3)$-valued process

$$\hat{\mathbb{W}}_n(s,t,u) = \hat{\mathbb{V}}_n(s,t,u) - s\hat{\mathbb{V}}_n(1,t,u)\,, \qquad s,t,u \in [0,1]\,, \tag{6.9}$$

then the convergence in (6.7) and the continuous mapping theorem yield

$$\hat{\mathbb{W}}_n \rightsquigarrow \mathbb{W} \tag{6.10}$$

in $C([0,1]^3)$, where $\mathbb{W}$ is centred Gaussian defined by $\mathbb{W}(s,t,u) = \mathbb{V}(s,t,u) - s\mathbb{V}(1,t,u)$ with covariance operator given by (4.5). Finally, recall the definition of the process $(\hat{\mathbb{U}}_n : n \in \mathbb{N})$ in (4.2) and note that, in contrast to $\hat{\mathbb{W}}_n$, this process is not centred. Consequently, if $d_\infty = 0$, we have $\sqrt{n}\,\mathbb{U}_n = \hat{\mathbb{W}}_n$ and the convergence in (6.10) and the continuous mapping theorem directly yield (4.4).
If $d_\infty > 0$ assertion (4.7) is a consequence of the weak convergence in (6.10) and Theorem B.1 in the online supplement of Dette et al. (2020a) and also of the results in Cárcamo et al. (2020).

## 6.5 Proof of Theorem 4.1 and 4.2

**Proof of Theorem 4.1.** It can be shown that the bootstrap processes in (4.8) can be written

$$\begin{aligned}\hat{B}_n^{(r)}(s,t,u) =& \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor sn \rfloor} \frac{1}{\sqrt{l}} \Big( \sum_{j=k}^{k+l-1} \tilde{Y}_{n,j}(t,u) - \frac{l}{n} \sum_{j=1}^{n} \tilde{Y}_{n,j}(t,u) \Big) \xi_k^{(r)} \\ &+ \sqrt{n} \Big( s - \frac{\lfloor sn \rfloor}{n} \Big) \frac{1}{\sqrt{l}} \Big( \sum_{j=\lfloor sn \rfloor+1}^{\lfloor sn \rfloor+l} \tilde{Y}_{n,j}(t,u) - \frac{l}{n} \sum_{j=1}^{n} \tilde{Y}_{n,j}(t,u) \Big) \xi_{\lfloor sn \rfloor+1}^{(r)} + o_{\mathbb{P}}(1)\end{aligned}$$

for $r = 1, \dots, R$ where

$$\tilde{Y}_{n,j} = \tilde{\eta}_{n,j}^{\check{\otimes}2}(t,u) - (\hat{C}_2 - \hat{C}_1)\mathbb{1}\{j > \lfloor \hat{s}n \rfloor\}$$

for $j = 1, \dots, n$ ($n \in \mathbb{N}$) and the array $(\tilde{\eta}_{n,j}^{\check{\otimes}2} : n \in \mathbb{N},\ j = 1, \dots, n)$ satisfies (A1), (A3) and (A4) of Assumption 2.1 in Dette et al. (2020a). The convergence in (6.10) and similar arguments as in the proof of Theorem 4.3 in the same reference then imply

$$\big(\hat{\mathbb{W}}_n, \hat{\mathbb{W}}_n^{(1)}, \dots, \hat{\mathbb{W}}_n^{(R)}\big) \rightsquigarrow (\mathbb{W}, \mathbb{W}^{(1)}, \dots, \mathbb{W}^{(R)}) \tag{6.11}$$

in $C([0,1]^3)^{R+1}$ as $n \to \infty$ where the process $\hat{\mathbb{W}}_n$ is defined by (6.9), the bootstrap counterparts $\hat{\mathbb{W}}_n^{(1)}, \dots, \hat{\mathbb{W}}_n^{(R)}$ are defined by (4.10) and the random variables $\mathbb{W}^{(1)}, \dots, \mathbb{W}^{(R)}$ are independent copies of $\mathbb{W}$ which is defined by its covariance operator (4.5).
If $d_\infty = 0$, the continuous mapping theorem directly implies

$$\big(\hat{\mathbb{M}}_n, \check{T}_n^{(1)}, \dots, \check{T}_n^{(R)}\big) \xrightarrow{\mathcal{D}} (\check{T}, \check{T}^{(1)}, \dots, \check{T}^{(R)})$$

in $\mathbb{R}^{R+1}$ as $n \to \infty$ where the statistic $\hat{\mathbb{M}}_n$ is defined by (4.3), the bootstrap statistics $\check{T}_n^{(1)}, \ldots, \check{T}_n^{(R)}$ are defined by (4.11) and the random variables $\check{T}^{(1)}, \ldots, \check{T}^{(R)}$ are independent copies of the random variable $\check{T}$ defined by (4.4). Now the same arguments as in the discussion starting from equation (6.3) imply the assertions made in Theorem 4.1.

**Proof of Theorem 4.2.** We first mention that it follows by similar arguments as given in the proof of Theorem 4.2 in Dette et al. (2020a) that the estimator of the unknown change location defined by (4.9) satisfies

$$|\hat{s} - s^*| = O_{\mathbb{P}}(n^{-1})$$

whenever $d_\infty > 0$. Whenever $d_\infty = 0$, suppose that the estimate $\hat{s}$ converges weakly to a $[\vartheta, 1-\vartheta]$-valued random variable which is denoted by $s_{\max}$. Then, if $d_\infty > 0$, the convergence in (4.7) and Slutsky's theorem yield

$$\sqrt{n}\big(\hat{d}_\infty - d_\infty\big) \xrightarrow{\mathcal{D}} D(\mathcal{E}) = \tilde{D}(\mathcal{E})/[s^*(1-s^*)]\,, \tag{6.12}$$

where $\tilde{D}(\mathcal{E})$ is the same as in (4.7) and the statistic $\hat{d}_\infty$ is defined by (4.13).
The same arguments as in the proof of Theorem 3.6 in Dette et al. (2020a) again yield that the estimators of the extremal sets defined by (4.14) are consistent. The convergence in (6.11) and similar arguments as in the proof of Theorem 4.4 in the same reference then yield

$$\big(\sqrt{n}\ (\hat{d}_\infty - d_\infty),\ \check{K}_n^{(1)}, \ldots, \check{K}_n^{(R)}\big) \xrightarrow{\mathcal{D}} (D(\mathcal{E}),\ D^{(1)}(\mathcal{E}), \ldots, D^{(R)}(\mathcal{E})) \tag{6.13}$$

in $\mathbb{R}^{R+1}$ as $n \to \infty$ where the bootstrap statistics $\check{K}_n^{(1)}, \ldots, \check{K}_n^{(R)}$ are defined by (4.15) and the random variables $D^{(1)}(\mathcal{E}), \ldots, D^{(R)}(\mathcal{E})$ are independent copies of $D(\mathcal{E})$ which is defined by (6.12). The convergence in the preceding equation holds true under the null and the alternative hypothesis and now the same arguments as in the discussion starting from equation (6.4) imply the assertions made in Theorem 4.2.

**Acknowledgements** This research was partially supported by the Collaborative Research Center 'Statistical modeling of nonlinear dynamic processes' (*Sonderforschungsbereich 823, Teilprojekt A1, C1*) and the Research Training Group 'High-dimensional phenomena in probability - fluctuations and discontinuity' (*RTG 2131*). The authors are grateful to Christina Stoehr for sending us the results of Stoehr et al. (2019) and to Martina Stein, who typed parts of this manuscript with considerable technical expertise.

# References

Aue, A., Dubart Norinho, D., and Hörmann, S. (2015). On the prediction of stationary functional time series. *Journal of the American Statistical Association*, 110:378–392.

Aue, A., Rice, G., and Sönmez, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(3):509–529.

Aue, A., Rice, G., and Sönmez, O. (2020). Structural break analysis for spectrum and trace of covariance operators. *Environmetrics*, 31(1):e2617. e2617 env.2617.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Boente, G., Rodriguez, D., and Sued, M. (2018). Testing equality between several populations covariance operators. *Annals of the Institute of Statistical Mathematics*, 70(4):919–950.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications*. Lecture Notes in Statistics. Springer New York.

Bücher, A. and Kojadinovic, I. (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *Journal of Theoretical Probability*, 32:1145–1165.

Cabassi, A., Pigoli, D., Secchi, P., and Carter, P. A. (2017). Permutation tests for the equality of covariance operators of functional data with applications to evolutionary biology. *Electronic Journal of Statistics*, 11(2):3815–3840.

Cárcamo, J., Rodríguez, L.-A., and Cuevas, A. (2020). Directional differentiability for supremum-type functionals: statistical applications. *Bernoulli, to appear; ArXiv e-print 1902.01136*.

Carey, J. R., Liedo, P., Müller, H.-G., Wang, J.-L., and Chiou, J.-M. (1998). Relationship of age patterns of fecundity to mortality, longevity, and lifetime reproduction in a large cohort of mediterranean fruit fly females. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 53:B245–51.

Dehling, H. and Philipp, W. (2002). *Empirical Process Techniques for Dependent Data*, pages 3–113. Birkhäuser Boston, Boston, MA.

Dette, H., Kokot, K., and Aue, A. (2020a). Functional data analysis in the banach space of continuous functions. *Annals of Statistics*, 48(2):1168–1192.

Dette, H., Kokot, K., and Volgushev, S. (2020b). Testing relevant hypotheses in functional time series via self-normalization. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, n/a(n/a).

Ferraty, F. and Vieu, P. (2010). *Nonparametric Functional Data Analysis*. Springer-Verlag, New York.

Fremdt, S., Steinebach, J. G., Horváth, L., and Kokoszka, P. (2013). Testing the equality of covariance operators in functional samples. *Scandinavian Journal of Statistics*, 40(1):138–152.

Gaenssler, P., Molnár, P., and Rost, D. (2007). On continuity and strict increase of the cdf for the sup-functional of a gaussian process with applications to statistics. *Results in Mathematics*, 51(1):51–60.

Guo, J., Zhou, B., and Zhang, J.-T. (2016). A supremum-norm based test for the equality of several covariance functions. *Computational Statistics & Data Analysis*, 124.

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

Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to linear Operators*. Wiley, New York.

Janson, S. and Kaijser, S. (2015). Higher moments of Banach space valued random variables. *Memoirs of the American Mathematical Society*, 238.

Jarušková, D. (2013). Testing for a change in covariance operator. *Journal of Statistical Planning and Inference*, 143(9):1500 – 1511.

Kraus, D. and Panaretos, V. M. (2012). Dispersion operators and resistant second-order functional data analysis. *Biometrika*, 99(4):813–832.

Liebl, D. and Reimherr, M. (2019). Fast and fair simultaneous confidence bands for functional parameters. *arXiv:1910.00131.*

Panaretos, V. M., Kraus, D., and Maddocks, J. H. (2010). Second-order comparison of Gaussian random functions and the geometry of DNA minicircles. *Journal of the American Statistical Association*, 105(490):670–682.

Paparoditis, E. and Sapatinas, T. (2016). Bootstrap-based testing of equality of mean functions or equality of covariance operators for functional data. *Biometrika*, 103(3):727–733.

Pigoli, D., Aston, J. A. D., Dryden, I. L., and Secchi, P. (2014). Distances and inference for covariance operators. *Biometrika*, 101(2):409–422.

Pilavakis, D., Paparoditis, E., and Sapatinas, T. (2020). Testing equality of autocovariance operators for functional time series. *Journal of Time Series Analysis*, n/a(n/a).

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

Sharipov, O. S. and Wendler, M. (2019). Bootstrapping covariance operators of functional time series. *ArXiv e-print 1904.06721v2*.

Stoehr, C., Aston, J. A. D., and Kirch, C. (2019). Detecting changes in the covariance structure of functional time series with application to fMRI data. *ArXiv e-print 1903.00288*.

Van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications in Statistics*. Springer, New York.

Zhang, X. and Shao, X. (2015). Two sample inference for the second-order property of temporally dependent functional data. *Bernoulli*, 21(2):909–929.