

# FUNCTIONAL DATA ANALYSIS IN THE BANACH SPACE OF CONTINUOUS FUNCTIONS

BY HOLGER DETTE<sup>1,\*</sup>, KEVIN KOKOT<sup>1,\*\*</sup> AND ALEXANDER AUE<sup>2</sup>

<sup>1</sup>*Fakultät für Mathematik, Ruhr-Universität Bochum, \*holger.dette@ruhr-uni-bochum.de; \*\*kevin.kokot@ruhr-uni-bochum.de*

<sup>2</sup>*Department of Statistics, University of California, aae@ucdavis.edu*

Functional data analysis is typically conducted within the  $L^2$ -Hilbert space framework. There is by now a fully developed statistical toolbox allowing for the principled application of the functional data machinery to real-world problems, often based on dimension reduction techniques such as functional principal component analysis. At the same time, there have recently been a number of publications that sidestep dimension reduction steps and focus on a fully functional  $L^2$ -methodology. This paper goes one step further and develops data analysis methodology for functional time series in the space of all continuous functions. The work is motivated by the fact that objects with rather different shapes may still have a small  $L^2$ -distance and are therefore identified as similar when using a  $L^2$ -metric. However, in applications it is often desirable to use metrics reflecting the visualization of the curves in the statistical analysis. The methodological contributions are focused on developing two-sample and change-point tests as well as confidence bands, as these procedures appear to be conducive to the proposed setting. Particular interest is put on relevant differences; that is, on not trying to test for exact equality, but rather for prespecified deviations under the null hypothesis.

The procedures are justified through large-sample theory. To ensure practicability, nonstandard bootstrap procedures are developed and investigated addressing particular features that arise in the problem of testing relevant hypotheses. The finite sample properties are explored through a simulation study and an application to annual temperature profiles.

**1. Introduction.** Due to the recent dramatic evolution in advanced data collection technologies, the development of statistical methodology for the analysis of functional data sampled over time and/or space has become an active field of research. Most of the literature has dealt with developing Hilbert space-based methodology for which there exists by now a fully fledged theory. The interested reader is referred to the various monographs of Bosq [11], Ferraty and Vieu [23], Horváth and Kokoszka [28] and Ramsay and Silverman [40] for up-to-date accounts. However, the integral role of smoothness has been discussed at length in Ramsay and Silverman [40] and virtually all functions fit in practice are at least continuous. In such cases, dimension reduction techniques can incur a loss of information and fully functional methods can prove advantageous. More recently, Aue et al. [6], Bucchia and Wendler [13] and Horváth et al. [30] discussed fully functional methodology in a Hilbert space framework.

Since all functions utilized for practical purposes are at least continuous, and often smoother than that, it might be more natural to develop methodology for functional data in the space of continuous functions. This is the approach pursued in the present paper, in particular in two-sample and change-point problems for Banach space-valued time series

---

Received October 2017; revised February 2019.

*MSC2010 subject classifications.* 62G10, 62G15, 62M10.

*Key words and phrases.* Banach spaces, functional data analysis, time series, relevant hypotheses, two-sample tests, change-point tests, bootstrap.

satisfying mixing conditions. While it might thus be reasonable to build statistical analysis adopting this point of view, there are certain difficulties associated with it. Giving up on the theoretically convenient Hilbert space setting means that substantially more effort has to be put into the derivation of theoretical results, especially if one is interested in the incorporation of dependent functional observations. Section 2 of this paper gives an introduction to Banach space methodology and states some basic results, in particular an invariance principle for a sequential process in the space of continuous functions.

The theoretical contributions will be utilized for the development of relevant two-sample and change-point tests in Sections 3 and 4, respectively. Here, the usefulness of the proposed approach becomes more apparent as differences between two smooth curves are hard to detect in practice. Additionally, small discrepancies might perhaps not even be of importance in many applied situations. Therefore, the “relevant” setting is adopted that is not trying to test for exact equality under the null hypothesis, but allows for prespecified deviations from an assumed null function. For example, if  $C([0, 1])$ , the space of continuous functions on the compact interval  $[0, 1]$ , is equipped with the sup-norm  $\|f\| = \sup_{t \in [0, 1]} |f(t)|$ , and  $\mu_1$  and  $\mu_2$  are the mean functions corresponding to two samples, interest is in hypotheses of the form

$$(1.1) \quad H_0: \|\mu_1 - \mu_2\| \leq \Delta \quad \text{and} \quad H_1: \|\mu_1 - \mu_2\| > \Delta,$$

where  $\Delta \geq 0$  denotes a prespecified constant. The classical case of testing perfect equality, obtained by the choice  $\Delta = 0$ , is therefore a special case of (1.1). However, in applications it might be reasonable to think about this choice carefully and to define precisely the size of change which one is really interested in. In particular, testing relevant hypotheses avoids the consistency problem as mentioned in Berkson [9], that is, any consistent test will detect any arbitrary small change in the mean functions if the sample size is sufficiently large. One may also view this perspective as a particular form of a bias-variance trade-off. The problem of testing for a relevant difference between two (one-dimensional) means and other (finite-dimensional) parameters has been discussed by numerous authors in biostatistics (see Wellek [44] for a recent review), but to the best of our knowledge these testing problems have not been considered in the context of functional data. It turns out that from a mathematical point of view the problem of testing relevant (i.e.,  $\Delta > 0$ ) hypotheses is substantially more difficult than the classical problem (i.e.,  $\Delta = 0$ ). In particular, it is not possible to work with stationarity under the null hypothesis, making the derivation of a limit distribution of a corresponding test statistic or the construction of a bootstrap procedure substantially more difficult.

Section 3 develops corresponding two-sample tests for the Banach space  $C([0, 1])$ . Section 4 extends these results to the change-point setting (see Aue and Horváth [4] for a recent review of change-point methodology for time series). Here, one has to deal with the additional complexity of locating the unknown time of change. Several new results for change-point analysis of functional data in  $C([0, 1])$  are put forward. A specific challenge here is the fact that the asymptotic null distribution of test statistics for hypotheses of the type (1.1) depends on the set of extremal points of the unknown difference  $\mu_1 - \mu_2$ , and is therefore not distribution-free. Most notable for both the two-sample and the change-point problem is the construction of nonstandard bootstrap tests for relevant hypotheses to solve this problem. The bootstrap is theoretically validated and then used to determine cut-off values for the proposed procedures.

Another area of application that lends itself naturally to Banach space methodology is that of constructing confidence bands for the mean function of a collection of potentially temporally dependent, continuous functions. There has been recent work by Choi and Reimherr [17] on this topic in a Hilbert space framework for functional parameters of independent functions based on geometric considerations. Here, results for confidence bands for the mean

difference in a two-sample framework are added in Section 3.1.2. Natural modifications allow for the inclusion of the one-sample case. One of the main differences between the two approaches is that the proposed bands hold pointwise, while those constructed from Hilbert space theory are valid only in an  $L^2$ -sense. This property is appealing for practitioners, because two mean curves can have a rather different shape, yet the  $L^2$ -norm of their difference might be very small.

The finite-sample properties of the relevant two-sample and change-point tests and, in particular, the performance of the bootstrap procedures are evaluated with the help of a Monte Carlo simulation study in Section A of the Supplementary Material [21]. A number of scenarios are investigated, with the outcomes showing that the proposed methodology performs reasonably well. Furthermore, in Section 5 an application to a prototypical data example is given, namely two-sample and change-point tests for annual temperature profiles recorded at measuring stations in Australia.

The outline of the rest of this paper is as follows. Section 2 introduces the basic notions of the proposed Banach space methodology and gives some preliminary results. Section 3 discusses the two-sample problem and Section 4 is concerned with change-point analysis. Empirical aspects are highlighted in Section 5 and in Section A of the Supplementary Material. Proofs of the main results can also be found in see Section B of the Supplementary Material.

**2.  $C(T)$ -Valued random variables.** In this section, some basic facts are provided about central limit theorems and invariance principles for  $C(T)$ -valued random variables, where  $C(T)$  is the set of continuous functions from  $T$  into the real line  $\mathbb{R}$ . In what follows, unless otherwise mentioned,  $C(T)$  will be equipped with the sup norm  $\|\cdot\|$ , defined by  $\|f\| = \sup_{t \in T} |f(t)|$ , thus making  $(C(T), \|\cdot\|)$  a Banach space. The natural Borel  $\sigma$ -field  $\mathcal{B}(T)$  over  $C(T)$  is then generated by the open sets relative to the sup norm  $\|\cdot\|$ . Measurability of random variables on  $(\Omega, \mathcal{A}, P)$  taking values in  $C(T)$  is understood to be with respect to  $\mathcal{B}(T)$ . The underlying probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  is assumed complete. It is further assumed that there is a metric  $\rho$  on  $T$  such that  $(T, \rho)$  is totally bounded. The fact that  $T$  is metrizable implies that  $C(T)$  is separable and measurability issues are avoided (see Theorem 11.7 in Janson and Kaijser [32]). Moreover, any random variable  $X$  in  $C(T)$  is tight (see Theorem 1.3 in Billingsley [10]).

Let  $X$  be a random variable on  $(\Omega, \mathcal{A}, P)$  taking values in  $C(T)$ . There are different ways to formally introduce expectations and higher-order moments of Banach space-valued random variables (see Janson and Kaijser [32]). The expectation  $\mathbb{E}[X]$  of a random variable  $X$  in  $C(T)$  exists as an element of  $C(T)$  whenever  $\mathbb{E}[\|X\|] < \infty$ . The  $k$ th moment exists whenever  $\mathbb{E}[\|X\|^k] = \mathbb{E}[\sup_{t \in T} |X(t)|^k] < \infty$ . As pointed out in Chapter 11 of Janson and Kaijser [32],  $k$ th order moments may be computed through pointwise evaluation as  $\mathbb{E}[X(t_1) \cdots X(t_k)]$ . The case  $k = 2$  is important as it allows for the computation of covariance kernels in a pointwise fashion.

A sequence of random variables  $(X_n : n \in \mathbb{N})$  converges in distribution or weakly to a random variable  $X$  in  $C(T)$ , whenever it is asymptotically tight and its finite-dimensional distributions converge weakly to the finite-dimensional distributions of  $X$ , that is,

$$(X_n(t_1), \dots, X_n(t_k)) \Rightarrow (X(t_1), \dots, X(t_k))$$

for any  $t_1, \dots, t_k \in T$  and any  $k \in \mathbb{N}$ , where the symbol “ $\Rightarrow$ ” indicates convergence in distribution in  $\mathbb{R}^k$ .

A centered random variable  $X$  in  $C(T)$  is said to be Gaussian if its finite-dimensional distributions are multivariate normal, that is, for any  $t_1, \dots, t_k$ ,  $(X(t_1), \dots, X(t_k)) \sim \mathcal{N}_k(0, \Sigma)$ , where the  $(i, j)$ th entry of the covariance matrix  $\Sigma$  is given by  $\mathbb{E}[X(t_i)X(t_j)]$ ,  $i, j = 1, \dots, k$ .

The distribution of  $X$  is hence completely characterized by its covariance function  $k(t, t') = \mathbb{E}[X(t)X(t')]$ ; see Chapter 2 of Billingsley [10].

In general Banach spaces, deriving conditions under which the central limit theorem (CLT) holds is a difficult task, significantly more complex than the counterpart for real-valued random variables. In Banach spaces, finiteness of second moments of the underlying random variables does not provide a necessary and sufficient condition. Elaborate theory has been developed to resolve the issue, resulting in notions of Banach spaces of type 2 and cotype 2 (see the book by Ledoux and Talagrand [35] for an overview). However, the Banach space of continuous functions on a compact interval does not possess the requisite type and cotype properties and further assumptions are needed in order to obtain the CLT, especially to incorporate time series of continuous functions into the framework. To model the dependence of the observations, the notion of  $\varphi$ -mixing sequences  $(\eta_j : j \in \mathbb{N})$  of  $C(T)$ -valued random variables is introduced; see, for example, Bradley [12]. First, for any two  $\sigma$ -fields  $\mathcal{F}$  and  $\mathcal{G}$ , define

$$\phi(\mathcal{F}, \mathcal{G}) = \sup\{|\mathbb{P}(G|F) - \mathbb{P}(G)| : F \in \mathcal{F}, G \in \mathcal{G}, \mathbb{P}(F) > 0\},$$

where  $\mathbb{P}(G|F)$  denotes the conditional probability of  $G$  given  $F$ . Next, denote by  $\mathcal{F}_k^{k'}$  the  $\sigma$ -field generated by  $(\eta_j : k \leq j \leq k')$ . Then define the  $\varphi$ -mixing coefficient as

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

and call the sequence  $(\eta_j : j \in \mathbb{N})$   $\varphi$ -mixing whenever  $\lim_{k \rightarrow \infty} \varphi(k) = 0$ .

In order to obtain a CLT as well as an invariance principle for sequences of  $\varphi$ -mixing random elements in  $C(T)$ , the following conditions are imposed.

ASSUMPTION 2.1.  $(X_{n,j} : n \in \mathbb{N}, j = 1, \dots, n)$  is an array of  $C(T)$ -valued random variables where, for any  $j = 1, \dots, n$  and  $n \in \mathbb{N}$ ,

$$(2.1) \quad X_{n,j} = \eta_j + \mu_{n,j}$$

with expectations  $\mathbb{E}[X_{n,j}] = \mu_{n,j}$  and error process  $(\eta_j : j \in \mathbb{N}) \subset C(T)$ . Furthermore, the following conditions are assumed to hold:

(A1) There is a constant  $K$  such that, for all  $j \in \mathbb{N}$ ,

$$\mathbb{E}[\|\eta_j\|^{2+\nu}] \leq K, \quad \mathbb{E}[\|\eta_j\|^J] < \infty$$

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

(A2) The error process  $(\eta_j : j \in \mathbb{N})$  is stationary.

(A3) There exists a real-valued nonnegative random variable  $M$  with  $\mathbb{E}[M^J] < \infty$ , such that, for any  $n \in \mathbb{N}$  and  $j = 1, \dots, n$ , the inequality

$$|X_{n,j}(t) - X_{n,j}(t')| \leq M\rho(t, t')$$

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

(A4)  $(\eta_j : j \in \mathbb{N})$  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  $\nu$  and  $J$  are the same as in (A1).

Note that these assumptions can be formulated for sequences of random variables  $(X_n : n \in \mathbb{N})$  in  $C(T)$  in a similar way. Since triangular arrays satisfying Assumption 2.1 only differ in their means from row to row, it follows that the covariance structure is the same in each row, that is,

$$\text{Cov}(X_{n,j}(t), X_{n,j'}(t')) = \text{Cov}(\eta_j(t), \eta_{j'}(t')) = \gamma(j - j', t, t')$$

for all  $n \in \mathbb{N}$  and  $j, j' = 1, \dots, n$  (note that  $\gamma(-j, t, t') = \gamma(j, t', t)$ ). Assumption 2.1 implies the following CLT which is proved in Section B.2 of the Supplementary Material. Throughout this paper, the symbol  $\rightsquigarrow$  denotes weak convergence in  $(C(T))^k$  or  $(C([0, 1] \times T))^k$  for some  $k \in \mathbb{N}$  and  $D(\omega, \rho)$  is the packing number with respect to the metric  $\rho$  that is the maximal number of  $\omega$ -separated points in  $T$  (see van der Vaart and Wellner [43]).

**THEOREM 2.1.** *Let  $(X_{n,j} : n \in \mathbb{N}, j = 1, \dots, n)$  denote a triangular array of random variables in  $C(T)$  with expectations  $E[X_{n,j}] = \mu_{n,j}$  such that Assumption 2.1 is satisfied and  $\int_0^\tau D(\omega, \rho)^{1/J} d\omega < \infty$  for some  $\tau > 0$ . Then*

$$G_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_{n,j} - \mu_{n,j}) \rightsquigarrow Z$$

in  $C(T)$ , where  $Z$  is a centered Gaussian random variable with covariance function

$$(2.2) \quad C(s, t) = \text{Cov}(Z(s), Z(t)) = \sum_{i=-\infty}^{\infty} \gamma(i, s, t).$$

**REMARK 2.1.**

(a) Condition (A4) in Assumption 2.1 is satisfied for  $\varphi$ -mixing processes with exponentially decreasing mixing coefficients, that is,  $\varphi(k) \leq ca^k$  ( $k \in \mathbb{N}$ ) for some  $a \in (0, 1)$ .

(b) In the following sections, we focus on the interval  $T = [0, 1]$  equipped with the metric  $\rho(s, t) = |s - t|^\theta$  for a positive constant  $\theta \in (0, 1]$ . In this case, the packing number satisfies  $D(\omega, \rho) \lesssim \lceil \tau^{-1/\theta} \rceil$ , which implies

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

whenever the even integer  $J$  satisfies  $J > 1/\theta$ . Consequently, Theorem 2.1 can be applied to Hölder continuous processes under this assumption. For example, the paths of the Brownian motion on the interval  $[0, 1]$  are Hölder continuous of order  $\theta$  for any  $\theta \in (0, 1/2)$ , and in this case we have to assume  $J \geq 4$  in Assumption 2.1. In general, less smoothness requires a stronger summability assumption on the mixing coefficients. For processes with Hölder continuous paths with  $\theta > 1/2$ , thus including Lipschitz continuity,  $J = 2$  is sufficient to obtain the CLT in Theorem 2.1.

Next, we will verify a weak invariance principle for the process  $(\hat{V}_n : n \in \mathbb{N})$  given by

$$(2.3) \quad \begin{aligned} \hat{V}_n(s) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor sn \rfloor} (X_{n,j} - \mu_{n,j}) \\ &\quad + \sqrt{n} \left( s - \frac{\lfloor sn \rfloor}{n} \right) (X_{n, \lfloor sn \rfloor + 1} - \mu_{n, \lfloor sn \rfloor + 1}), \end{aligned}$$

useful for the change-point analysis proposed in Section 4. Note that the process  $(\hat{V}_n(s) : s \in [0, 1])$  is an element of the Banach space  $C([0, 1], C(T)) = \{\phi : [0, 1] \rightarrow C(T) \mid \phi \text{ is continuous}\}$ , where the norm on this space is given by

$$(2.4) \quad \sup_{s \in [0, 1]} \sup_{t \in T} |\phi(s, t)| = \|\phi\|_{C([0, 1] \times T)}.$$

Note also that for any  $s \in [0, 1]$  the quantity  $\phi(s)$  is an element of  $C(T)$  (i.e., a real-valued continuous function with domain  $T$ ). Denote by  $\phi(s, t)$  the value of  $\phi(s)$  at the point  $t \in T$ . Moreover, each element of  $C([0, 1], C(T))$  can equivalently be regarded as an element of  $C([0, 1] \times T)$ . Here and throughout this paper, the notation  $\|\cdot\|$  is used to denote any of the arising sup-norms as the corresponding space can be identified from the context. We also make frequently use of the notation  $s \wedge s' = \min\{s, s'\}$ . The proof of the following result is postponed to Section B.2 of the Supplementary Material.

**THEOREM 2.2.** *Let the assumptions of Theorem 2.1 be satisfied. Then the weak invariance principle holds, that is,*

$$(2.5) \quad \hat{V}_n \rightsquigarrow \mathbb{V}$$

in  $C([0, 1] \times T)$ , where  $\mathbb{V}$  is a centered Gaussian measure on  $C([0, 1] \times T)$  characterized by

$$(2.6) \quad \text{Cov}(\mathbb{V}(s, t), \mathbb{V}(s', t')) = (s \wedge s')C(t, t'),$$

and the long-run covariance function  $C$  is given in (2.2).

**REMARK 2.2.** It was pointed out by the referees that it might be of interest to investigate if similar statements hold for other dependency concepts. There is a large amount of literature discussing a CLT (often as a consequence of strong approximations) for Banach space valued random variables under mixing conditions as stated in Theorem 2.1. The discussion of the sufficient conditions for such statements is very sophisticated and we refer exemplarily to the work of Kuelbs and Philipp [34], Dehling and Philipp [20] and Dehling [19]. For example, it can be shown using the results in Dehling [19], page 400, that under Assumption 2.1 a central limit theorem is valid for absolute regular sequences if the corresponding mixing coefficients  $\beta(k)$  satisfy similar conditions as considered here. Therefore, Theorem 2.1 also holds under this concept of dependency. As this statement is used intensively for the asymptotic analysis in the following sections the statistical methodology for the two sample case as studied in Section 3 can also be developed for absolute regular sequences.

Another dependency concept is  $L^p$ - $m$ -approximability, which is frequently used for Hilbert space valued time series (see Hörmann and Kokoszka [27], Berkes et al. [8]). One can define a similar concept for the situation of  $C(T)$ -valued time series considered here where we essentially require that the error process in model (2.1) can be approximated by an  $m$ -dependent process. More precisely, this means that it admits a representation of the form  $\eta_j = f(\varepsilon_j, \varepsilon_{j-1}, \varepsilon_{j-2}, \dots)$  with a sequence  $(\varepsilon_n : n \in \mathbb{N})$  of random variables and there exists, for each  $j$ , an independent copy  $(\varepsilon_n^{(j)} : n \in \mathbb{N})$  of  $(\varepsilon_n : n \in \mathbb{N})$  such that the random variables  $\eta_{j,m} = f(\varepsilon_j, \dots, \varepsilon_{j-m+1}, \varepsilon_{j-m}^{(j)}, \varepsilon_{j-m-1}^{(j)}, \dots)$  satisfy

$$\sum_{m=1}^{\infty} \mathbb{E}[\|\eta_m - \eta_{m,m}\|^2]^{1/2} < \infty$$

(note that  $\|\cdot\|$  is the sup-norm on  $C(T)$ ). Replacing in Assumption 2.1 the condition (A4) by

$$\sum_{m=1}^{\infty} m^{1/(1/2-\bar{\tau})} \mathbb{E}[\|\eta_m - \eta_{m,m}\|^2]^{1/2} < \infty,$$

$$\sum_{m=1}^{\infty} (m+1)^{J/2-1} \mathbb{E}[\|\eta_m - \eta_{m,m}\|^J]^{1/J} < \infty,$$

a CLT can be proved for Banach space-valued time series of the form (2.1) with an  $m$ -approximable error process in  $C(T)$ . These results can then be used to develop similar statistical methodology as in Section 3 for  $m$ -approximable  $C([0, 1])$ -valued time series. Moreover, it can be shown by similar arguments as given in Example 2.1 in Hörmann and Kokoszka [27] that this dependency concept includes fAR(1) processes.

On the other hand, the step from a CLT to an invariance principle as stated in Theorem 2.2 is more complicated and has not been studied so intensively in the literature (see Kuelbs [33] or Garling [25] for some early references for independent sequences). For the proof of Theorem 2.2 in the Supplementary Material, we use results of Samur [41, 42], which require the assumption of  $\varphi$ -mixing sequences. An extension of these results to other mixing concepts might be possible, but is beyond the scope of the present paper.

**3. The two-sample problem.** From now on, consider the case  $T = [0, 1]$ , as this is the canonical choice for functional data analysis. The corresponding metric is given by  $\rho(s, t) = |s - t|^\theta$  (for some  $\theta \in (0, 1]$ ). Two-sample problems have a long history in statistics and the corresponding tests are among the most applied statistical procedures. For the functional setting, there have been a number of contributions as well. Two are worth mentioning in the present context. Hall and van Keilegom [26] studied the effect of smoothing when converting discrete observations into functional data. Horváth et al. [29] introduced two-sample tests for  $L^p$ - $m$  approximable functional time series based on Hilbert-space theory. In the following, a two-sample test is proposed in the Banach-space framework of Section 2. To this end, consider two independent samples  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_n$  of  $C([0, 1])$ -valued random variables. Under (A2) in Assumption 2.1, expectation functions and covariance kernels exist and are denoted by  $\mu_1 = \mathbb{E}[X_1]$  and  $\mu_2 = \mathbb{E}[Y_1]$ , and  $k_1(t, t') = \text{Cov}(X_1(t), X_1(t'))$  and  $k_2(t, t') = \text{Cov}(Y_1(t), Y_1(t'))$ , respectively. Interest is then in the size of the maximal deviation

$$d_\infty = \|\mu_1 - \mu_2\| = \sup_{t \in [0, 1]} |\mu_1(t) - \mu_2(t)|$$

between the two mean curves, that is, in testing the hypotheses of a relevant difference

$$(3.1) \quad H_0: d_\infty \leq \Delta \quad \text{versus} \quad H_1: d_\infty > \Delta,$$

where  $\Delta \geq 0$  is a prespecified constant determined by the user of the test. Note again that the “classical” two-sample problem  $H_0: \mu_1 = \mu_2$  versus  $H_0: \mu_1 \neq \mu_2$ —which, to the best of our knowledge, has not been investigated for  $C([0, 1])$ -valued data yet—is contained in this setup as the special case  $\Delta = 0$ . Observe also that tests for relevant differences between two finite-dimensional parameters corresponding to different populations have been considered mainly in the biostatistical literature, for example, in Wellek [44]. It is assumed throughout this section that the samples are balanced in the sense that

$$(3.2) \quad \frac{m}{n+m} \rightarrow \lambda \in (0, 1)$$

as  $m, n \rightarrow \infty$ . Additionally, let  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_n$  be sampled from independent time series  $(X_j: j \in \mathbb{N})$  and  $(Y_j: j \in \mathbb{N})$  that satisfy Assumption 2.1 with  $J > 1/\theta$ . Under

these conditions, both functional time series satisfy the CLT and it then follows from Theorem 2.1 that

$$(3.3) \quad \begin{aligned} &\sqrt{n+m} \left( \frac{1}{m} \sum_{j=1}^m (X_j - \mu_1), \frac{1}{n} \sum_{j=1}^n (Y_j - \mu_2) \right) \\ &\rightsquigarrow \left( \frac{1}{\sqrt{\lambda}} Z_1, \frac{1}{\sqrt{1-\lambda}} Z_2 \right), \end{aligned}$$

where  $Z_1$  and  $Z_2$  are independent, centered Gaussian processes possessing covariance functions

$$(3.4) \quad C_1(t, t') = \sum_{j=-\infty}^{\infty} \gamma_1(j, t, t') \quad \text{and} \quad C_2(t, t') = \sum_{j=-\infty}^{\infty} \gamma_2(j, t, t'),$$

respectively. Here,  $\gamma_1$  and  $\gamma_2$ , correspond to the respective sequences  $(X_j : j \in \mathbb{N})$  and  $(Y_j : j \in \mathbb{N})$  and are defined in the discussion after Assumption 2.1. Now, the weak convergence in (3.3) and the independence of the samples imply immediately that

$$(3.5) \quad Z_{m,n} = \sqrt{n+m} \left( \frac{1}{m} \sum_{j=1}^m X_j - \frac{1}{n} \sum_{j=1}^n Y_j - (\mu_1 - \mu_2) \right) \rightsquigarrow Z$$

in  $C([0, 1])$  as  $m, n \rightarrow \infty$ , where  $Z = Z_1/\sqrt{\lambda} + Z_2/\sqrt{1-\lambda}$  is a centered Gaussian process with covariance function

$$(3.6) \quad C(t, t') = \text{Cov}(Z(t), Z(t')) = \frac{1}{\lambda} C_1(t, t') + \frac{1}{1-\lambda} C_2(t, t').$$

Under the convergence in (3.5), the statistic

$$(3.7) \quad \hat{d}_\infty = \left\| \frac{1}{m} \sum_{j=1}^m X_j - \frac{1}{n} \sum_{j=1}^n Y_j \right\|$$

is a reasonable estimator of the maximal deviation  $d_\infty = \|\mu_1 - \mu_2\|$ , and the null hypothesis in (3.1) is rejected for large values of  $\hat{d}_\infty$ . In order to develop a test with a prespecified asymptotic level, the limit distribution of  $\hat{d}_\infty$  is determined in the following. For this purpose, let

$$(3.8) \quad \mathcal{E}^\pm = \{t \in [0, 1] : \mu_1(t) - \mu_2(t) = \pm d_\infty\}$$

if  $d_\infty > 0$ , and define  $\mathcal{E}^+ = \mathcal{E}^- = [0, 1]$  if  $d_\infty = 0$ . Finally, denote by  $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^-$  the set of extremal points of the difference  $\mu_1 - \mu_2$  of the two mean functions. The first main result establishes the asymptotic distribution of the statistic  $\hat{d}_\infty$ .

**THEOREM 3.1.** *If  $X_1, \dots, X_m$  and  $Y_1, \dots, Y_n$  are sampled from independent time series  $(X_j : j \in \mathbb{N})$  and  $(Y_j : j \in \mathbb{N})$  in  $C([0, 1])$ , each satisfying Assumption 2.1 with metric  $\rho(s, t) = |s - t|^\theta, \theta \in (0, 1], J\theta > 1$ , then*

$$(3.9) \quad T_{m,n} = \sqrt{n+m}(\hat{d}_\infty - d_\infty) \xrightarrow{\mathcal{D}} T(\mathcal{E}),$$

where

$$(3.10) \quad T(\mathcal{E}) = \max \left\{ \sup_{t \in \mathcal{E}^+} Z(t), \sup_{t \in \mathcal{E}^-} -Z(t) \right\},$$

and the centered Gaussian process  $Z$  is given by (3.6) and the sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$  are defined in (3.8).



It should be emphasized that the limit distribution depends in a complicated way on the set  $\mathcal{E}$  of extremal points of the difference  $\mu_1 - \mu_2$  and is therefore not distribution-free, even in the case of i.i.d. data. In particular, there can be two sets of processes with corresponding mean functions  $\mu_1, \mu_2$  and  $\tilde{\mu}_1, \tilde{\mu}_2$  such that  $\|\mu_1 - \mu_2\| = \|\tilde{\mu}_1 - \tilde{\mu}_2\|$ . However, the respective limit distributions in Theorem 3.1 will be entirely different if the corresponding sets of extremal points  $\mathcal{E}$  and  $\tilde{\mathcal{E}}$  do not coincide. The proof of Theorem 3.1 is given in Section B.3 of the Supplementary Material. In the case  $d_\infty = 0$ ,  $\mathcal{E}^+ = \mathcal{E}^- = [0, 1]$  and it follows for the random variable  $T$  ( $[0, 1]$ ) in Theorem 3.1 that

$$(3.11) \quad T = \max_{t \in [0,1]} |Z(t)|.$$

Here, the result is a simple consequence of the weak convergence (3.5) of the process  $Z_{m,n}$  (see Theorem 2.1) and the continuous mapping theorem.

However, Theorem 3.1 provides also the distributional properties of the statistic  $\hat{d}_\infty$  in the case  $d_\infty > 0$ . This is required for testing the hypotheses of a relevant difference between the two mean functions (i.e., the hypotheses in (3.1) with  $\Delta > 0$ ), which is of primary interest here. In this case, the weak convergence of an appropriately standardized version of  $\hat{d}_\infty$  does *not* follow from the weak convergence (3.5), as the process inside the supremum in (3.7) is not centered. In fact, additional complexity enters in the proofs because even under the null hypothesis observations cannot be easily centered. For details, refer to Section B.3 of the Supplementary Material.

### 3.1. Asymptotic inference.

3.1.1. *Testing the classical hypothesis*  $H_0: \mu_1 \equiv \mu_2$ . Theorem 3.1 also provides the asymptotic distributions of the test statistic  $\hat{d}_\infty$  in the case of two identical mean functions, that is, if  $\mu_1 \equiv \mu_2$ . This is the situation investigated in Hall and Van Keilegom [26] and Horváth et al. [29] in Hilbert-space settings. Here, it corresponds to the special case  $\Delta = 0$ , and thus  $d_\infty = 0$ ,  $\mathcal{E}^\pm = [0, 1]$ . Consequently,

$$T_{m,n} \xrightarrow{\mathcal{D}} T \quad (m, n \rightarrow \infty),$$

where the random variable  $T$  is defined in (3.11). An asymptotic level  $\alpha$  test for the classical hypotheses

$$(3.12) \quad H_0: \mu_1 = \mu_2 \quad \text{versus} \quad H_1: \mu_1 \neq \mu_2$$

may hence be obtained by rejecting  $H_0$  whenever

$$(3.13) \quad \hat{d}_\infty > \frac{u_{1-\alpha}}{\sqrt{n+m}},$$

where  $u_{1-\alpha}$  is the  $(1 - \alpha)$ -quantile of the distribution of the random variable  $T$  defined in (3.11). Note that this quantile only depends on the long-run covariance operator, which has to be estimated in applications (see, e.g., Horváth et al. [29] for such an estimator). Using Theorem 3.1, it is easy to see that the test defined by (3.13) is consistent and has asymptotic level  $\alpha$ .

3.1.2. *Confidence bands.* The methodology developed so far can easily be applied to the construction of simultaneous asymptotic confidence bands for the difference of the mean functions. There is a rich literature on confidence bands for functional data in Hilbert spaces. The available work includes Degras [18], who dealt with confidence bands for nonparametric regression with functional data; Cao et al. [16], who studied simultaneous confidence bands for the mean of dense functional data based on polynomial spline estimators; Cao [15], who

developed simultaneous confidence bands for derivatives of functional data when multiple realizations are at hand for each function, exploiting within-curve correlation; and Zheng et al. [45] who treated the sparse case. Most recently, Choi and Reimherr [17] extracted geometric features akin to Mahalanobis distances to build confidence bands for functional parameters.

The results presented here are the first of their kind relating to Banach space-valued functional data. The first theorem uses the limit distribution obtained in Theorem 3.1 to construct asymptotic simultaneous confidence bands for the two-sample case. A corresponding bootstrap analog will be developed in the next section. Confidence bands for the one-sample case can be constructed in a similar fashion using standard arguments and the corresponding results are consequently omitted.

**THEOREM 3.2.** *Let the assumptions of Theorem 3.1 be satisfied and, for  $\alpha \in (0, 1)$ , denote by  $u_{1-\alpha}$  the  $(1 - \alpha)$ -quantile of the random variable  $T$  defined in (3.11) and define the functions*

$$\mu_{m,n}^{\pm}(t) = \frac{1}{m} \sum_{j=1}^m X_j - \frac{1}{n} \sum_{j=1}^n Y_j \pm \frac{u_{1-\alpha}}{\sqrt{n+m}}.$$

Then the set

$$C_{\alpha,m,n} = \{ \mu \in C([0, 1]) : \mu_{m,n}^-(t) \leq \mu(t) \leq \mu_{m,n}^+(t) \text{ for all } t \in [0, 1] \}$$

defines a simultaneous asymptotic  $(1 - \alpha)$  confidence band for  $\mu_1 - \mu_2$ , that is,

$$\lim_{m,n \rightarrow \infty} \mathbb{P}(\mu_1 - \mu_2 \in C_{\alpha,m,n}) = 1 - \alpha.$$

Note that, unlike their Hilbert-space counterparts, the simultaneous confidence bands given in Theorem 3.2 (and their bootstrap analogs in Section 3.2.1) hold for all  $t \in [0, 1]$  and not only almost everywhere, making the proposed bands more easily interpretable and perhaps more useful for applications.

**REMARK 3.1.** As pointed out by a referee there might be situations, where pointwise (instead of uniform) confidence bands are of interest. These bands can easily be derived from the theory developed so far. For example, for a fixed  $t_0 \in [0, 1]$  it follows from (3.5) and the continuous mapping theorem that  $Z_{m,n}(t_0)$  is asymptotically normal distributed with mean 0 and variance  $\sigma^2 = C(t_0, t_0)$ , where  $C$  is defined in (3.6). Therefore, if  $z_{\beta}$  is the  $\beta$ -quantile of the  $\mathcal{N}(0, 1)$  distribution, and  $\hat{\sigma}^2$  is an estimator of the long-run variance, an asymptotic confidence interval for the difference  $\mu_1(t_0) - \mu_2(t_0)$  is given by  $[d_{m,n}^-, d_{m,n}^+]$ , where

$$d_{m,n}^{\pm} = \frac{1}{m} \sum_{j=1}^m X_j(t_0) - \frac{1}{n} \sum_{j=1}^n Y_j(t_0) \pm z_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{n+m}}.$$

**3.1.3. Testing for a relevant difference.** Recall the definition of the random variable  $T(\mathcal{E})$  in Theorem 3.1, then the null hypothesis of no relevant difference in (3.1) is rejected at level  $\alpha$ , whenever the inequality

$$(3.14) \quad \hat{d}_{\infty} > \Delta + \frac{u_{1-\alpha,\mathcal{E}}}{\sqrt{n+m}}$$

holds, where  $u_{\alpha,\mathcal{E}}$  denotes the  $\alpha$ -quantile of the distribution of  $T(\mathcal{E})$  ( $\alpha \in (0, 1)$ ). A conservative test avoiding the use of quantiles depending on the set of extremal points  $\mathcal{E}$  can be obtained observing the inequality

$$(3.15) \quad T(\mathcal{E}) \leq T,$$

where the random variable  $T$  is defined in (3.11). If  $u_\alpha$  denotes the  $\alpha$ -quantile of the distribution of  $T$ , then (3.15) implies

$$u_{\alpha, \mathcal{E}} \leq u_\alpha$$

and a conservative asymptotic level  $\alpha$  test is given by rejecting the null hypothesis in (3.1), whenever the inequality

$$(3.16) \quad \hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{n+m}}$$

holds. The properties of the tests (3.14) and (3.16) depend on the size of the distance  $d_\infty$  and will be explained below. In particular, observe the following properties for the test (3.16):

(a) If  $d_\infty < \Delta$ , Slutsky’s theorem yields that

$$\begin{aligned} & \lim_{n,m \rightarrow \infty} \mathbb{P}\left(\hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{n+m}}\right) \\ &= \lim_{n,m \rightarrow \infty} \mathbb{P}\left(\sqrt{n+m}(\hat{d}_\infty - d_\infty) > \sqrt{n+m}(\Delta - d_\infty) + u_{1-\alpha}\right) = 0. \end{aligned}$$

(b) If  $d_\infty = \Delta$ , we have

$$(3.17) \quad \begin{aligned} & \limsup_{n,m \rightarrow \infty} \mathbb{P}\left(\hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{n+m}}\right) \\ &= \limsup_{n,m \rightarrow \infty} \mathbb{P}\left(\sqrt{n+m}(\hat{d}_\infty - d_\infty) > \sqrt{n+m}(\Delta - d_\infty) + u_{1-\alpha}\right) \\ &\leq \lim_{n,m \rightarrow \infty} \mathbb{P}\left(\sqrt{n+m}(\hat{d}_\infty - d_\infty) > u_{1-\alpha, \mathcal{E}}\right) = \alpha. \end{aligned}$$

(c) If  $d_\infty > \Delta$ , the same calculation as in (a) implies

$$\lim_{n,m \rightarrow \infty} \mathbb{P}\left(\hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{n+m}}\right) = 1,$$

proving that the test defined in (3.16) is consistent.

(d) If the mean functions  $\mu_1$  and  $\mu_2$  define a boundary point of the hypotheses, that is,  $d_\infty = \Delta$  and either  $\mathcal{E}^+ = [0, 1]$  or  $\mathcal{E}^- = [0, 1]$ , then

$$T(\mathcal{E}) = \max_{t \in [0,1]} Z(t) \quad \text{or} \quad T(\mathcal{E}) = \max_{t \in [0,1]} -Z(t),$$

and consequently

$$\lim_{n,m \rightarrow \infty} \mathbb{P}\left(\hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{m+n}}\right) = \frac{\alpha}{2}.$$

Using similar arguments, it can be shown that the test (3.14) satisfies

$$(3.18) \quad \lim_{n,m \rightarrow \infty} \mathbb{P}\left(\hat{d}_\infty > \Delta + \frac{u_{1-\alpha, \mathcal{E}}}{\sqrt{n+m}}\right) = \begin{cases} 0 & \text{if } d_\infty < \Delta, \\ \alpha & \text{if } d_\infty = \Delta, \\ 1 & \text{if } d_\infty \geq \Delta. \end{cases}$$

Summarizing, the tests for the hypothesis (3.1) of no relevant difference between the two mean functions defined in (3.14) and (3.16) have asymptotic level at most  $\alpha$  and are consistent. However, the discussion given above also shows that the test (3.16) is conservative, even when  $\mathcal{E} = [0, 1]$ .

The quantile  $u_{1-\alpha, \mathcal{E}}$  of the test (3.14) can be estimated as follows. Let  $\hat{C}_{n,m}$  denote an estimator of the long-run covariance kernel defined in (3.6) (see, e.g., Horváth et al. [29]) and

let  $\hat{\mathcal{E}}_{m,n}^+$  and  $\hat{\mathcal{E}}_{m,n}^-$  denote consistent estimates of the extremal sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$ , respectively (see, e.g., the definition (3.25) and Theorem 3.6 in the following section). Now let  $\hat{u}_{1-\alpha,m,n}$  denote the  $(1 - \alpha)$ -quantile of the distribution of the statistic

$$\max \left\{ \sup_{t \in \hat{\mathcal{E}}_{m,n}^+} \hat{\mathbb{G}}_{n,m}(t), \sup_{t \in \hat{\mathcal{E}}_{m,n}^-} -\hat{\mathbb{G}}_{n,m}(t) \right\},$$

where  $\hat{\mathbb{G}}_{n,m}$  is a centered Gaussian process with covariance kernel  $\hat{C}_n$ , then it can be shown that the test, which rejects the null hypothesis, whenever

$$\hat{d}_\infty > \Delta + \hat{u}_{1-\alpha,m,n} / \sqrt{n+m}$$

has asymptotic level  $\alpha$  and is consistent. However, it turns out that the finite sample properties of this test are very sensitive with respect to the estimate  $\hat{C}_{n,m}$  of the long-run covariance operator and for this reason we discuss a bootstrap approach in the following section.

3.2. *Bootstrap.* In order to use the tests (3.13), (3.14) and (3.16) for classical and relevant hypotheses, the quantiles of the distribution of the random variables  $T(\mathcal{E})$  and  $T$  defined in (3.10) and (3.11) need to be estimated, which depend on certain features of the data generating process. The law  $T(\mathcal{E})$  involves the unknown set of extremal points  $\mathcal{E}$  of the differences of the mean functions. Moreover, the distributions of  $T(\mathcal{E})$  and  $T$  depend on the long-run covariance function (3.6). There are methods available in the literature to consistently estimate the covariance function (see, e.g., Horváth et al. [29]). In practice, however, it is difficult to reliably approximate the infinite sums in (3.6) and, therefore, an easily implementable bootstrap procedure is proposed in the following.

It turns out that a different and nonstandard bootstrap procedure will be required for testing relevant hypotheses than for classical hypotheses (and the construction of confidence bands) as in this case the null distribution depends on the set of extremal points  $\mathcal{E}$ . The corresponding resampling procedure requires a substantially more sophisticated analysis. Therefore, the analysis of bootstrap tests for the classical hypothesis and bootstrap confidence intervals is given first and discussion of bootstrap tests for relevant hypotheses is deferred to Section 3.2.2.

3.2.1. *Bootstrap confidence intervals and tests for the classical hypothesis  $H_0 : \mu_1 = \mu_2$ .* Following Bücher and Kojadinovic [14], the use of a multiplier block bootstrap is proposed, noting in passing that other resampling concepts such as the stationary bootstrap (see, e.g., Politis and Romano [37]) or the tapered bootstrap (see, e.g., Paparoditis and Politis [36]) can be adjusted as well to address the problem of testing for hypotheses with respect to the sup-norm.

To be precise, let

$$(\xi_k^{(1)} : k \in \mathbb{N}), \dots, (\xi_k^{(R)} : k \in \mathbb{N})$$

and

$$(\zeta_k^{(1)} : k \in \mathbb{N}), \dots, (\zeta_k^{(R)} : k \in \mathbb{N})$$

denote independent sequences of independent random variables with mean 0 and variance 1, and define the  $C([0, 1])$ -valued processes  $\hat{B}_{m,n}^{(1)}, \dots, \hat{B}_{m,n}^{(R)}$  through

$$\begin{aligned} \hat{B}_{m,n}^{(r)}(t) = \sqrt{n+m} & \left\{ \frac{1}{m} \sum_{k=1}^{m-l_1+1} \frac{1}{\sqrt{l_1}} \left( \sum_{j=k}^{k+l_1-1} X_j(t) - \frac{l_1}{m} \sum_{j=1}^m X_j(t) \right) \xi_k^{(r)} \right. \\ (3.19) \quad & \left. - \frac{1}{n} \sum_{k=1}^{n-l_2+1} \frac{1}{\sqrt{l_2}} \left( \sum_{j=k}^{k+l_2-1} Y_j(t) - \frac{l_2}{n} \sum_{j=1}^n Y_j(t) \right) \zeta_k^{(r)} \right\} \quad (r = 1, \dots, R) \end{aligned}$$

for  $t \in [0, 1]$ , where  $l_1, l_2 \in \mathbb{N}$  denote window sizes such that  $l_1/m \rightarrow 0$  and  $l_2/n \rightarrow 0$  as  $l_1, l_2, m, n \rightarrow \infty$ . Note that the dependence on  $l_1$  and  $l_2$  is not reflected in our notation.

The intuition of this definition is as follows. We center every (local) mean  $\bar{X}_k^{(\ell)} = \ell^{-1} \sum_{j=k}^{k+\ell-1} X_j$  using the global mean  $\bar{X}_m = m^{-1} \sum_{j=1}^m X_j$ . Then we do the same for the second sample ( $Y_j: j = 1, \dots, n$ ). Consequently, the inflated differences

$$D_{k,X} = \sqrt{l_1}(\bar{X}_k^{(l_1)} - \bar{X}_m) \quad \text{and} \quad D_{k,Y} = \sqrt{l_2}(\bar{Y}_k^{(l_2)} - \bar{Y}_n)$$

have approximately the same mean function (namely 0) and, as  $l_1, l_2 \rightarrow \infty$ , their long-run covariances are given by the kernels  $C_1$  and  $C_2$  in (3.4), respectively. Now the multiplication with the independent random variables  $\xi_k^{(r)}$  and  $\zeta_k^{(r)}$  yields an analog of the process  $Z_{m,n}$  in (3.5), where the summands are conditionally independent, approximately centered and have asymptotically correct long-run covariance. The following result makes these heuristic arguments precise and is a fundamental tool for the theoretical investigations of all bootstrap procedures proposed in this paper. It is proved in Section B.3 of the Supplementary Material.

**THEOREM 3.3.** *Suppose that  $(X_j: j \in \mathbb{N})$  and  $(Y_j: j \in \mathbb{N})$  satisfy Assumption 2.1 with metric  $\rho(s, t) = |s - t|^\theta$ ,  $\theta \in (0, 1]$ ,  $J\theta > 1$  and let  $\hat{B}_{m,n}^{(1)}, \dots, \hat{B}_{m,n}^{(R)}$  denote the bootstrap processes defined by (3.19) such that  $l_1 = m^{\beta_1}, l_2 = n^{\beta_2}$  with*

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

and  $\nu_i$  given in Assumption 2.1,  $i = 1, 2$ . Moreover, assume for the multipliers in (3.19)  $\mathbb{E}|\xi_1^{(r)}|^J < \infty$  and  $\mathbb{E}|\zeta_1^{(r)}|^J < \infty$ . Then

$$(3.20) \quad (Z_{m,n}, \hat{B}_{m,n}^{(1)}, \dots, \hat{B}_{m,n}^{(R)}) \rightsquigarrow (Z, Z^{(1)}, \dots, Z^{(R)})$$

in  $C([0, 1])^{R+1}$  as  $m, n \rightarrow \infty$ , where  $Z_{m,n}$  is defined in (3.5) and  $Z^{(1)}, \dots, Z^{(R)}$  are independent copies of the centered Gaussian process  $Z$  defined by (3.6).

Note that Theorem 3.3 holds under the null hypothesis and alternative. It leads to the following results regarding confidence bands and tests for the classical hypothesis (3.12) based on the multiplier bootstrap. To this end, observe that for the statistics

$$T_{m,n}^{(r)} = \|\hat{B}_{m,n}^{(r)}\|, \quad r = 1, \dots, R,$$

the continuous mapping theorem yields

$$(3.21) \quad (\sqrt{n + md_\infty} \hat{T}_\infty, T_{m,n}^{(1)}, \dots, T_{m,n}^{(R)}) \Rightarrow (T, T^{(1)}, \dots, T^{(R)}),$$

where the random variables  $T^{(1)}, \dots, T^{(R)}$  are independent copies of the statistic  $T$  defined in (3.11). Now, if  $T_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}$  is the empirical  $(1 - \alpha)$ -quantile of the bootstrap sample  $T_{m,n}^{(1)}, \dots, T_{m,n}^{(R)}$ , the following results are obtained.

**THEOREM 3.4.** *Let the assumptions of Theorem 3.3 be satisfied and define the functions*

$$\mu_{m,n}^{R,\pm}(t) = \frac{1}{m} \sum_{j=1}^m X_j - \frac{1}{n} \sum_{j=1}^n Y_j \pm \frac{T_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n + m}}.$$

Then

$$\hat{C}_{\alpha,m,n}^R = \{\mu \in C([0, 1]): \mu_{m,n}^{R,-}(t) \leq \mu(t) \leq \mu_{m,n}^{R,+}(t) \text{ for all } t \in [0, 1]\}$$

defines a simultaneous asymptotic  $(1 - \alpha)$  confidence band for  $\mu_1 - \mu_2$ , that is,

$$\lim_{R \rightarrow \infty} \liminf_{m,n \rightarrow \infty} \mathbb{P}(\mu_1 - \mu_2 \in \hat{C}_{\alpha,m,n}^R) \geq 1 - \alpha.$$

Pointwise bootstrap confidence intervals can be constructed similarly as indicated in Remark 3.1 using the weak convergence in equation (3.20) and the continuous mapping theorem. The details are omitted for the sake of brevity.

This section is concluded with a corresponding statement regarding the bootstrap test for the classical hypotheses in (3.12), which rejects the null hypothesis whenever

$$(3.22) \quad \hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n+m}},$$

where the statistic  $\hat{d}_\infty$  is defined in (3.7).

**THEOREM 3.5.** *Let the assumptions of Theorem 3.3 be satisfied, then the test (3.22) has asymptotic level  $\alpha$  and is consistent for the hypotheses (3.12). More precisely, under the null hypothesis of no difference in the mean functions,*

$$(3.23) \quad \lim_{R \rightarrow \infty} \limsup_{m,n \rightarrow \infty} \mathbb{P} \left( \hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n+m}} \right) = \alpha,$$

and, under the alternative, for any  $R \in \mathbb{N}$ ,

$$(3.24) \quad \liminf_{m,n \rightarrow \infty} \mathbb{P} \left( \hat{d}_\infty > \frac{T_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n+m}} \right) = 1.$$

**3.2.2. Testing for relevant differences in the mean functions.** The problem of constructing an appropriate bootstrap test for the hypotheses of no relevant difference in the mean functions is substantially more complicated. The reason for these difficulties consists in the fact that in the case of relevant hypotheses the limit distribution of the corresponding test statistic is complicated. In contrast to the problem of testing the classical hypotheses (3.12), where it is sufficient to mimic the distribution of the statistic  $T$  in (3.11) (corresponding to the case  $\mu_1 \equiv \mu_2$ ) one requires the distribution of the statistic  $T(\mathcal{E})$ , which depends in a sophisticated way on the set of extreme points of the (unknown) difference  $\mu_1 - \mu_2$ . Under the null hypothesis  $\|\mu_1 - \mu_2\| \leq \Delta$ , these sets can be very different, ranging from a singleton to the full interval  $[0, 1]$ . As a consequence, the construction of a valid bootstrap procedure requires appropriate consistent estimates of the sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$  introduced in Theorem 3.1.

For this purpose, recall the definition of the Hausdorff distance between two sets  $A, B \subset \mathbb{R}$ , given by

$$d_H(A, B) = \max \left\{ \sup_{x \in A} \inf_{y \in B} |x - y|, \sup_{y \in B} \inf_{x \in A} |x - y| \right\}$$

and denote by  $K([0, 1])$  the set of all compact subsets of the interval  $[0, 1]$ . First, define estimates of the extremal sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$  by

$$(3.25) \quad \hat{\mathcal{E}}_{m,n}^\pm = \left\{ t \in [0, 1] : \pm (\hat{\mu}_1(t) - \hat{\mu}_2(t)) \geq \hat{d}_\infty - \frac{c_{m,n}}{\sqrt{m+n}} \right\},$$

where  $\lim_{m,n \rightarrow \infty} c_{m,n} / \log(m+n) = c$  for some  $c > 0$ . Our first result shows that the estimated sets  $\hat{\mathcal{E}}_{m,n}^+$  and  $\hat{\mathcal{E}}_{m,n}^-$  are consistent for  $\mathcal{E}^+$  and  $\mathcal{E}^-$ , respectively.

**THEOREM 3.6.** *Let the assumptions of Theorem 3.3 be satisfied, then*

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

where the sets  $\hat{\mathcal{E}}_{m,n}^\pm$  are defined by (3.25).

The main implication of Theorem 3.6 consists in the fact that the random variable

$$\max_{t \in \hat{\mathcal{E}}_{m,n}^+} \hat{B}_{m,n}(t)$$

converges weakly to the random variable  $\max_{t \in \mathcal{E}^+} Z(t)$ . Note that  $\hat{B}_{m,n} \rightsquigarrow Z$  by Theorem 3.3 and that  $d_H(\hat{\mathcal{E}}_{m,n}^+, \mathcal{E}^+) \rightarrow 0$  in probability by the previous theorem, but the combination of both statements is more delicate and requires a continuity argument which is given in Section B.3 of the Supplementary Material, where the following result is proved.

**THEOREM 3.7.** *Let the assumptions of Theorem 3.3 be satisfied and define, for  $r = 1, \dots, R$ ,*

$$(3.26) \quad K_{m,n}^{(r)} = \max \left\{ \max_{t \in \hat{\mathcal{E}}_{m,n}^+} \hat{B}_{m,n}^{(r)}(t), \max_{t \in \hat{\mathcal{E}}_{m,n}^-} (-\hat{B}_{m,n}^{(r)}(t)) \right\}.$$

Then

$$(3.27) \quad \begin{aligned} &(\sqrt{n+m}(\hat{d}_\infty - d_\infty), K_{m,n}^{(1)}, \dots, K_{m,n}^{(R)}) \\ &\Rightarrow (T(\mathcal{E}), T^{(1)}(\mathcal{E}), \dots, T^{(R)}(\mathcal{E})), \end{aligned}$$

in  $\mathbb{R}^{R+1}$ , where  $d_\infty = \|\mu_1 - \mu_2\|$ , the statistic  $\hat{d}_\infty$  is defined in (3.7) and the variables  $T^{(1)}(\mathcal{E}), \dots, T^{(R)}(\mathcal{E})$  are independent copies of  $T(\mathcal{E})$  defined in Theorem 3.1.

Theorem 3.7 leads to a simple bootstrap test for the hypothesis of no relevant change. To be precise, let  $K_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}$  denote the empirical  $(1 - \alpha)$ -quantile of the bootstrap sample  $K_{m,n}^{(1)}, \dots, K_{m,n}^{(R)}$ , then the null hypothesis of no relevant change is rejected at level  $\alpha$ , whenever

$$(3.28) \quad \hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n+m}}.$$

The final result of this section shows that the test (3.28) is consistent and has asymptotic level  $\alpha$ . The proof is obtained by similar arguments as given in the proof of Theorem 3.5, which are omitted for the sake of brevity.

**THEOREM 3.8.** *Let the assumptions of Theorem 3.3 be satisfied. Then, under the null hypothesis of no relevant difference in the mean functions,*

$$\lim_{R \rightarrow \infty} \limsup_{m,n \rightarrow \infty} \mathbb{P} \left( \hat{d}_\infty > \Delta + \frac{K_{m,n}^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n+m}} \right) \leq \alpha,$$

and, under the alternative of a relevant difference in the mean functions, for any  $R \in \mathbb{N}$ ,

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

**4. Change-point analysis.** Change-point problems arise naturally in a number of applications (e.g., in quality control, economics and finance; see Aue and Horváth [4] for a recent review). In the functional framework, applications have centered around environmental and climate observations (see Aue et al. [5, 6]) and intra-day finance data (see Horváth et al. [28]). One of the first contributions in the area are Berkes et al. [7] and Aue et al. [3] who developed change-point analysis in a Hilbert space setting for independent data. Generalizations to time series of functional data in Hilbert spaces are due to Aston and Kirch [1, 2]. For Banach

spaces, to the best of our knowledge, the only contributions to change-point analysis available in the literature are due to Račkauskas and Suquet [38, 39], who have provided theoretical work analyzing epidemic alternatives for independent functions based on Hölder norms and dyadic interval decompositions. This section details new results on change-point analysis for  $C([0, 1])$ -valued functional data. The work is the first to systematically exploit a time series structure of the functions as laid out in Section 2.

4.1. *Asymptotic inference.* More specifically, the problem of testing for a (potentially relevant) change-point is considered for triangular arrays  $(X_{n,j} : n \in \mathbb{N}, j = 1, \dots, n)$  of  $C([0, 1])$ -valued random variables satisfying Assumption 2.1 with corresponding metric given by  $\rho(s, t) = |s - t|^\theta$  (for some  $\theta \in (0, 1]$ ). Denote by  $\mu_{n,j} = \mathbb{E}[X_{n,j}] \in C([0, 1])$  the expectation of  $X_{n,j}$  and assume (as in the discussion after Assumption 2.1) that

$$\gamma(j - j', t, t') = \text{Cov}(X_{n,j}(t), X_{n,j'}(t'))$$

is the covariance kernel common to all random functions in the sample. Parametrize with  $s^* \in (\vartheta, 1 - \vartheta)$ , where  $\vartheta \in (0, 1)$  is a constant, the location of the change-point, so that the mean functions satisfy

$$(4.1) \quad \mu_1 = \mu_{n,1} = \dots = \mu_{n,\lfloor ns^* \rfloor} \quad \text{and} \quad \mu_2 = \mu_{n,\lfloor ns^* \rfloor + 1} = \dots = \mu_{n,n}.$$

Then, for any  $n \in \mathbb{N}$ , both  $X_{n,1}, \dots, X_{n,\lfloor ns^* \rfloor}$  and  $X_{n,\lfloor ns^* \rfloor + 1}, \dots, X_{n,n}$  consist of (asymptotically) identically distributed but potentially dependent random functions. Let again

$$d_\infty = \|\mu_1 - \mu_2\|$$

denote the maximal deviation between the mean functions before and after the change-point. Interest is then in testing the hypotheses of a relevant change, that is,

$$(4.2) \quad H_0: d_\infty \leq \Delta \quad \text{versus} \quad H_1: d_\infty > \Delta,$$

where  $\Delta \geq 0$  is a prespecified constant. The relevant change-point test setting may be viewed in the context of a bias-variance trade-off. In the time series setting, one is often interested in accurate predictions of future realizations. However, if the stretch of observed functions suffers from a structural break, then only those functions sampled after the change-point should be included in the prediction algorithm because these typically require stationarity. This reduction of observations, however, inevitably leads to an increased variability that may be partially offset with a bias incurred through the relevant approach: if the maximal discrepancy  $d_\infty$  in the mean functions remains below a suitably chosen threshold  $\Delta$ , then the mean-squared prediction error obtained from predicting with the whole sample might be smaller than the one obtained from using only the noncontaminated post-change sample. In applications to financial data, the size of the allowable bias could also be dictated by regulations imposed on, say, investment strategies (Dette and Wied [22] specifically mention Value at Risk as one such example).

Recall the definition of the sequential empirical process in (2.3), where the argument  $s \in [0, 1]$  of this process is used to search over all potential change locations. Note that  $(\hat{V}_n(s, t) : (s, t) \in [0, 1]^2)$  can be regarded as an element of the Banach space  $C([0, 1]^2)$  (see the discussion before Theorem 2.2). Define the  $C([0, 1]^2)$ -valued process

$$(4.3) \quad \hat{W}_n(s, t) = \hat{V}_n(s, t) - s\hat{V}_n(1, t), \quad s, t \in [0, 1],$$

then, under the assumptions in Theorem 2.1, Theorem 2.2 and the continuous mapping theorem show that

$$(4.4) \quad \hat{W}_n \rightsquigarrow W$$



in  $C([0, 1]^2)$ , where

$$\mathbb{W}(s, t) = \mathbb{V}(s, t) - s\mathbb{V}(1, t).$$

In particular,  $\mathbb{W}$  is a centered Gaussian measure on  $C([0, 1]^2)$  defined by

$$(4.5) \quad \text{Cov}(\mathbb{W}(s, t), \mathbb{W}(s', t')) = (s \wedge s' - ss')C(t, t').$$

In order to define a test for the hypothesis of a relevant change-point defined by (4.2) consider the sequential empirical process  $(\hat{U}_n : n \in \mathbb{N})$  on  $C([0, 1]^2)$  given by

$$(4.6) \quad \hat{U}_n(s, t) = \frac{1}{n} \left( \sum_{j=1}^{\lfloor sn \rfloor} X_{n,j}(t) + n \left( s - \frac{\lfloor sn \rfloor}{n} \right) X_{n, \lfloor sn \rfloor + 1}(t) - s \sum_{j=1}^n X_{n,j}(t) \right).$$

Evaluating its expected value shows that, in contrast to  $\hat{W}_n$ , the process  $\hat{U}_n$  is typically not centered and the equality

$$\sqrt{n}\hat{U}_n = \hat{W}_n$$

holds only in the case  $\mu_1 = \mu_2$ . A straightforward calculation shows that

$$\mathbb{E}[\hat{U}_n(s, t)] = (s \wedge s^* - ss^*)(\mu_1(t) - \mu_2(t)) + o_{\mathbb{P}}(1)$$

uniformly in  $(s, t) \in [0, 1]^2$ . As the function  $s \mapsto s \wedge s^* - ss^*$  attains its maximum in the interval  $[0, 1]$  at the point  $s^*$ , the statistic

$$(4.7) \quad \hat{M}_n = \sup_{s \in [0, 1]} \sup_{t \in [0, 1]} |\hat{U}_n(s, t)|$$

is a reasonable estimate of

$$s^*(1 - s^*)d_\infty = s^*(1 - s^*)\|\mu_1 - \mu_2\|.$$

It is therefore proposed to reject the null hypothesis in (4.2) for large values of the statistic  $\hat{M}_n$ . The following result specifies the asymptotic distribution of  $\hat{M}_n$ .

**THEOREM 4.1.** *Assume  $d_\infty > 0$ ,  $s^* \in (0, 1)$  and let  $(X_{n,j} : n \in \mathbb{N}, j = 1, \dots, n)$  be an array of  $C([0, 1])$ -valued random variables satisfying Assumption 2.1 with  $\rho(s, t) = |s - t|^\theta$ ,  $\theta \in (0, 1]$  and  $J\theta > 1$ . Then*

$$(4.8) \quad \mathbb{D}_n = \sqrt{n}(\hat{M}_n - s^*(1 - s^*)d_\infty) \xrightarrow{\mathcal{D}} D(\mathcal{E}) = \max \left\{ \sup_{t \in \mathcal{E}^+} \mathbb{W}(s^*, t), \sup_{t \in \mathcal{E}^-} -\mathbb{W}(s^*, t) \right\},$$

where the statistic  $\hat{M}_n$  is defined in (4.7),  $\mathbb{W}$  is the centered Gaussian measure on  $C([0, 1]^2)$  characterized by (4.5),  $\mathcal{E} = \mathcal{E}^+ \cup \mathcal{E}^-$  and the sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$  are defined in (3.8).

The proof of Theorem 4.1 is given in Section B.4 of the Supplementary Material. The limit distribution of  $\mathbb{D}_n$  is rather complicated and depends on the set  $\mathcal{E}$  which might be different for functions  $\mu_1 - \mu_2$  with the same sup-norm  $d_\infty$  but different corresponding set  $\mathcal{E}$ . It is also worthwhile to mention that the condition  $d_\infty > 0$  is essential in Theorem 4.1. In the remaining case  $d_\infty = 0$ , the weak convergence of  $\hat{M}_n$  simply follows from  $\sqrt{n}\hat{U}_n = \hat{W}_n$ , (4.4) and the continuous mapping theorem, that is,

$$(4.9) \quad \sqrt{n}\hat{M}_n \xrightarrow{\mathcal{D}} \check{T} = \sup_{(s,t) \in [0, 1]^2} |\mathbb{W}(s, t)|$$

whenever  $d_\infty = 0$ .

If  $d_\infty > 0$ , the true location of the change-point  $s^*$  is unknown and, therefore, has to be estimated from the available data. The next theorem, which is proved in Section B.4 of the Supplementary Material, proposes one such estimator and specifies its large-sample behavior in form of a rate of convergence.

**THEOREM 4.2.** *Assume  $d_\infty > 0$ ,  $s^* \in (0, 1)$  and let  $(X_{n,j} : n \in \mathbb{N}, j = 1, \dots, n)$  be an array of  $C([0, 1])$ -valued random variables satisfying Assumption 2.1, where the random variable  $M$  in Assumption (A3) is bounded and  $\rho(s, t) = |s - t|^\theta$  with  $\theta \in (0, 1]$ ,  $J\theta > 1$ . Then the estimator*

$$(4.10) \quad \tilde{s} = \frac{1}{n} \arg \max_{1 \leq k < n} \|\hat{U}_n(k/n, \cdot)\|$$

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

Recall that the possible range of change locations is restricted to the open interval  $(\vartheta, 1 - \vartheta)$  and define the modified change-point estimator

$$(4.11) \quad \hat{s} = \max\{\vartheta, \min\{\tilde{s}, 1 - \vartheta\}\},$$

where  $\tilde{s}$  is given by (4.10). Since  $|\hat{s} - s^*| \leq |\tilde{s} - s^*|$ , it follows that

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

if  $d_\infty > 0$ , and, if  $d_\infty = 0$  suppose that  $\hat{s}$  converges weakly to a  $[\vartheta, 1 - \vartheta]$ -valued random variable  $s_{\max}$ .

**COROLLARY 4.1.** *Let the assumptions of Theorem 4.2 be satisfied and define*

$$(4.12) \quad \hat{d}_\infty = \frac{\hat{M}_n}{\hat{s}(1 - \hat{s})}$$

as an estimator of  $d_\infty$ . Then

$$\sqrt{n}(\hat{d}_\infty - d_\infty) \Rightarrow T(\mathcal{E}) = D(\mathcal{E})/[s^*(1 - s^*)],$$

where  $D(\mathcal{E})$  is defined in (4.8).

**REMARK 4.1.** A consistent level  $\alpha$  test for the hypotheses (4.2) is constructed along the lines of the two-sample case discussed in Section 3.

(a) Consider first the case  $\Delta > 0$ , that is, a relevant hypothesis. If  $d_\infty > 0$ , implying the existence of a change-point  $s^* \in (0, 1)$ , then the inequality

$$(4.13) \quad T(\mathcal{E}) \leq T = \frac{1}{s^*(1 - s^*)} \sup_{t \in [0, 1]} |\mathbb{W}(s^*, t)|$$

holds. If  $u_{\alpha, \mathcal{E}}$  denotes the quantile of  $T(\mathcal{E})$ , then

$$u_{\alpha, \mathcal{E}} \leq u_\alpha$$

for all  $\alpha \in (0, 1)$ . Consequently, similar arguments as given in Section 3.1.3 show that the test which rejects the null hypothesis of no relevant change if

$$(4.14) \quad \hat{d}_\infty > \Delta + \frac{u_{1-\alpha}}{\sqrt{n}}$$

is consistent and has asymptotic level  $\alpha$ . Note that an estimator of the long-run covariance function is needed in order to obtain the  $\alpha$ -quantile  $u_\alpha$  of the distribution of  $T$ . Moreover, the test (4.14) is conservative, even when the set  $\mathcal{E}$  of extremal points of the unknown difference  $\mu_1 - \mu_2$  is the whole interval  $[0, 1]$  (in this case the level is in fact  $\alpha/2$  instead of  $\alpha$ —see the discussion at the end of Section 3.1.3).

(b) In the case of testing, the classical hypotheses

$$H_0: \mu_1 = \mu_2 \quad \text{versus} \quad H_1: \mu_1 \neq \mu_2,$$

that is,  $\Delta = 0$ , the test described in (4.14) needs to be slightly altered. The asymptotic distribution of  $\hat{M}_n$  under  $H_0$  can be obtained from (4.9) and now it can be seen that rejecting  $H_0$  whenever

$$\hat{d}_\infty > \frac{\check{u}_{1-\alpha}}{\sqrt{n}},$$

where  $\check{u}_{1-\alpha}$  denotes the  $(1 - \alpha)$ -quantile of the distribution of the random variable  $\check{T}$  defined by (4.9), yields a consistent asymptotic level  $\alpha$  test. Again an estimator of the long-run covariance function is required to simulate the quantile  $\check{u}_{1-\alpha}$  from the corresponding Gaussian process.

4.2. *Bootstrap.* In order to avoid the difficulties mentioned in the previous remark, a bootstrap procedure is developed and its consistency is shown. To be precise, denote by

$$\hat{\mu}_1 = \frac{1}{[\hat{sn}]} \sum_{j=1}^{[\hat{sn}]} X_{n,j} \quad \text{and} \quad \hat{\mu}_2 = \frac{1}{[(1 - \hat{s})n]} \sum_{j=[\hat{sn}]+1}^n X_{n,j}$$

estimators for the expectation before and after the change-point. Let  $(\xi_k^{(1)}: k \in \mathbb{N}), \dots, (\xi_k^{(R)}: k \in \mathbb{N})$  denote  $R$  independent sequences of independent sub-Gaussian random variables with mean 0 and variance 1, and consider the  $C([0, 1]^2)$ -valued processes  $\hat{B}_n^{(1)}, \dots, \hat{B}_n^{(R)}$  defined by

$$\begin{aligned} \hat{B}_n^{(r)}(s, t) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{[sn]} \frac{1}{\sqrt{l}} \left( \sum_{j=k}^{k+l-1} \hat{Y}_{n,j}(t) - \frac{l}{n} \sum_{j=1}^n \hat{Y}_{n,j}(t) \right) \xi_k^{(r)} \\ (4.15) \quad &+ \sqrt{n} \left( s - \frac{[sn]}{n} \right) \frac{1}{\sqrt{l}} \left( \sum_{j=[sn]+1}^{[sn]+l} \hat{Y}_{n,j}(t) - \frac{l}{n} \sum_{j=1}^n \hat{Y}_{n,j}(t) \right) \xi_{[sn]+1}^{(r)}, \end{aligned}$$

where  $l \in \mathbb{N}$  is a bandwidth parameter satisfying  $l/n \rightarrow 0$  as  $l, n \rightarrow \infty$  and

$$\hat{Y}_{n,j} = X_{n,j} - (\hat{\mu}_2 - \hat{\mu}_1) \mathbb{1}\{j > [sn]\}$$

for  $j = 1, \dots, n$  ( $n \in \mathbb{N}$ ). Note that it is implicitly assumed that

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

for any  $t \in [0, 1]$  and any  $s \in [0, 1]$  such that  $[sn] > n - l$ . Next, define

$$\hat{W}_n^{(r)}(s, t) = \hat{B}_n^{(r)}(s, t) - s \hat{B}_n^{(r)}(1, t); \quad r = 1, \dots, R.$$

**THEOREM 4.3.** *Let  $\hat{B}_n^{(1)}, \dots, \hat{B}_n^{(R)}$  denote the bootstrap processes defined by (4.15), where  $l = n^\beta$  for some  $\beta \in (1/5, 2/7)$  and assume that the underlying array  $(X_{n,j}: j = 1, \dots, n; n \in \mathbb{N})$  satisfies Assumption 2.1 with metric  $\rho(s, t) = |s - t|^\theta$ ,  $\theta \in (0, 1]$ ,  $J\theta > 1$  in (A3) and  $\nu \geq 2$  in (A1) and*

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

in (A4). Moreover, assume additionally for the multipliers in (4.15)  $\mathbb{E}|\xi_1^{(r)}|^J < \infty$  Then

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

in  $C([0, 1]^2)^{R+1}$ , where  $\hat{W}_n$  and  $\mathbb{W}$  are defined in (4.3) and (4.5), respectively, and  $\mathbb{W}^{(1)}, \dots, \mathbb{W}^{(R)}$  are independent copies of  $\mathbb{W}$ .

The proof of Theorem 4.3 is provided in Section B.4 of the Supplementary Material.

We now consider a resampling procedure for the classical hypotheses, that is,  $\Delta = 0$  in (4.2). For that purpose, define, for  $r = 1, \dots, R$ ,

$$(4.16) \quad \check{T}_n^{(r)} = \max\{|\widehat{\mathbb{W}}_n^{(r)}(s, t)| : s, t \in [0, 1]\}.$$

Then, by the continuous mapping theorem,

$$(\sqrt{n}\widehat{\mathbb{M}}_n, \check{T}_n^{(1)}, \dots, \check{T}_n^{(R)}) \Rightarrow (\check{T}, \check{T}^{(1)}, \dots, \check{T}^{(R)})$$

in  $\mathbb{R}^{R+1}$ , where  $\check{T}^{(1)}, \dots, \check{T}^{(R)}$  are independent copies of the random variable  $\check{T}$  defined in (4.9). If  $\check{T}_n^{\{[R(1-\alpha)]\}}$  is the empirical  $(1 - \alpha)$ -quantile of the bootstrap sample  $\check{T}_n^{(1)}, \check{T}_n^{(2)}, \dots, \check{T}_n^{(R)}$ , the classical null hypothesis  $H_0 : \mu_1 = \mu_2$  of no change-point is rejected, whenever

$$(4.17) \quad \widehat{\mathbb{M}}_n > \frac{\check{T}_n^{\{[R(1-\alpha)]\}}}{\sqrt{n}}.$$

It follows by similar arguments, as given in Section B.3 of the Supplementary Material, that this test is consistent and has asymptotic level  $\alpha$  in the sense of Theorem 3.5, that is,

$$\begin{aligned} \lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_{H_0} \left( \widehat{\mathbb{M}}_n > \frac{\check{T}_n^{\{[R(1-\alpha)]\}}}{\sqrt{n}} \right) &= \alpha, \\ \lim_{n \rightarrow \infty} \inf \mathbb{P}_{H_1} \left( \widehat{\mathbb{M}}_n > \frac{\check{T}_n^{\{[R(1-\alpha)]\}}}{\sqrt{n}} \right) &= 1, \end{aligned}$$

for any  $R \in \mathbb{N}$ . The details are omitted for the sake of brevity.

We now continue developing bootstrap methodology for the problem of testing for a relevant change-point, that is,  $\Delta > 0$  in (4.9). It turns out that the theoretical analysis is substantially more complicated as the null hypothesis defines a set in  $C([0, 1])$ . Similar as in (3.25), the estimates of the extremal sets  $\mathcal{E}^+$  and  $\mathcal{E}^-$  are defined by

$$(4.18) \quad \hat{\mathcal{E}}_n^\pm = \left\{ t \in [0, 1] : \pm (\hat{\mu}_1(t) - \hat{\mu}_2(t)) \geq \hat{d}_\infty - \frac{c_n}{\sqrt{n}} \right\},$$

where  $\lim_{n \rightarrow \infty} c_n / \log(n) = c > 0$  and  $\hat{d}_\infty$  is given in (4.12). Consider bootstrap analogs

$$(4.19) \quad T_n^{(r)} = \frac{1}{\hat{s}(1 - \hat{s})} \max \left\{ \sup_{t \in \hat{\mathcal{E}}_n^+} \widehat{W}_n^{(r)}(\hat{s}, t), \sup_{t \in \hat{\mathcal{E}}_n^-} (-\widehat{W}_n^{(r)}(\hat{s}, t)) \right\}$$

( $r = 1, \dots, R$ ) of the statistic

$$\sqrt{n}(\hat{d}_\infty - d_\infty)$$

in Corollary 4.1, where  $d_\infty = \|\mu_1 - \mu_2\|$ .

**THEOREM 4.4.** *Let the assumptions of Theorem 4.3 be satisfied, then if  $d_\infty > 0$ ,*

$$(\sqrt{n}(\hat{d}_\infty - d_\infty), T_n^{(1)}, \dots, T_n^{(R)}) \Rightarrow (T(\mathcal{E}), T^{(1)}, \dots, T^{(R)})$$

in  $\mathbb{R}^{R+1}$ , where  $T^{(1)}, \dots, T^{(R)}$  are independent copies of the random variable  $T(\mathcal{E})$  defined in Corollary 4.1.

A test for the hypothesis of a relevant change-point in time series of continuous functions is now obtained by rejecting the null hypothesis in (4.2), whenever

$$(4.20) \quad \hat{d}_\infty > \Delta + \frac{T_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}},$$

where  $T_n^{\{\lfloor R(1-\alpha) \rfloor\}}$  is the empirical  $(1 - \alpha)$ -quantile of the bootstrap sample  $T_n^{(1)}, T_n^{(2)}, \dots, T_n^{(R)}$ . It follows by similar arguments, as given in Section B.3 of the Supplementary Material, that this test is consistent and has asymptotic level  $\alpha$  in the sense of Theorem 3.8, that is,

$$\lim_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}_{H_0} \left( \hat{d}_\infty > \Delta + \frac{T_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}} \right) \leq \alpha$$

and

$$\liminf_{n \rightarrow \infty} \mathbb{P}_{H_1} \left( \hat{d}_\infty > \Delta + \frac{T_n^{\{\lfloor R(1-\alpha) \rfloor\}}}{\sqrt{n}} \right) = 1,$$

for any  $R \in \mathbb{N}$ . The details are omitted for the sake of brevity.

**5. Data example.** A detailed simulation study investigating the finite sample properties of the new methodology is given in the Supplementary Material. In this section, we illustrate the new approach in two applications to annual temperature profiles are reported in this section. Data of this kind were recently used in Fremdt et al. [24] in support of methodology designed to choose the dimension of the projection space obtained with fPCA. For all examples, functions were generated from daily values through representation in a Fourier basis consisting of 49 basis functions, where reasonable deviations from this preset do not qualitatively change the outcome of the analyses to follow.

*5.1. Two-sample tests.* For the two-sample testing problem, annual temperature profiles were obtained from daily temperatures recorded at measuring stations in Cape Otway (1865–2011), a location close to the southernmost point of Australia, and Sydney (1859–2011), a city on the eastern coast of Australia. This led to  $m = 147$ , respectively,  $n = 153$  functions for the two samples. Differences in the temperature profiles are expected due to different climate conditions, so the focus of the relevant tests is on working out how big the discrepancy might be.

To set up the test for the hypotheses (3.1), the statistic in (3.7) was computed, resulting in the value  $\hat{d}_\infty = 5.73$ . To see whether this is significant, the proposed bootstrap methodology was applied. To estimate the extremal sets in (3.8), the estimators in (3.25) were utilized with  $c_{m,n} = 0.1 \log(m + n) = 0.570$  and as bandwidth parameters  $l_1 = l_2 = 1$  were used. The resulting bootstrap quantiles are reported in the second row of Table 1. Also reported in this table are the results of the bootstrap procedure in (3.28) for various levels  $\alpha$  and relevance  $\Delta$ . Note that the maximum difference in mean the functions is achieved at  $t = 0.99$ , toward the end of December, and consequently during the Australian summer. The results show that there is strong evidence in the data to support the hypothesis that the maximal difference is at least  $\Delta = 5.4$ , but that there is no evidence that the maximal difference is even larger than  $\Delta = 5.6$ . Several intermediate values of  $\Delta$  led to weaker support of the alternative. The left panel of Figure 1 displays the difference in mean functions graphically.

TABLE 1

Summary of the bootstrap two sample procedure for relevant hypotheses with varying  $\Delta$  for the annual temperature curves. The label TRUE refers to a rejection of the null, the label FALSE to a failure to reject the null

$\Delta$	99%	97.5%	95%	90%
$q$	5.138	4.201	3.757	3.009
5.4	TRUE	TRUE	TRUE	TRUE
5.45	FALSE	TRUE	TRUE	TRUE
5.5	FALSE	FALSE	TRUE	TRUE
5.55	FALSE	FALSE	TRUE	TRUE
5.6	FALSE	FALSE	FALSE	FALSE

5.2. *Change-point tests.* Following Fremdt et al. [24], annual temperature curves were obtained from daily minimum temperatures recorded in Melbourne, Australia. This led to 156 annual temperature profiles ranging from 1856 to 2011 to which the change-point test for the relevant hypotheses in (4.2) was applied based on the rejection decision in (4.14). To compute the test statistic  $\hat{d}_\infty$  in (4.12), note that the estimated change-point in (4.11) was  $\hat{s} = 0.62$  (corresponding to the year 1952). This gives  $\hat{d}_\infty = 1.765$ . To see whether this value leads to a rejection of the null, the multiplier bootstrap procedure was utilized with bandwidth parameter  $l = 1$ , leading to the rejection rule in (4.20). In order to apply this procedure, first the extremal sets  $\hat{\mathcal{E}}^+$  and  $\hat{\mathcal{E}}^-$  in (4.18) were selected, choosing  $c_n = 0.1 \log n = 0.504$ . This yielded the bootstrap quantiles reported in the second row of Table 2.

Several values for  $\Delta$ , determining which deviations are to be considered relevant, were then examined. The results of the bootstrap testing procedure are summarized in Table 2. It can be seen that the null hypothesis of no relevant change was rejected at all considered levels for the smaller choice  $\Delta = 1.2$ . On the other extreme, for  $\Delta = 1.4$ , the test never rejected. For the intermediate values  $\Delta = 1.25, 1.3, 1.35$ , the null was rejected at the 2.5%, 5% and 10% level, at the 5% and 10% level, and at the 10% level, respectively. Estimating the mean functions before and after  $\hat{s}$  (1962) shows that the maximum difference of the mean functions is approximately 1.765, lending further credibility to the conducted analyses. The right panel of Figure 1 displays both mean functions for illustration. It can be seen that the mean difference is maximal during the Australian summer (in February), indicating that the mean functions of minimum temperature profiles have been most drastically changed during the hottest part of the year. The results here are in agreement with the findings put forward in Hughes et al. [31], who reported that average temperatures in Antarctica have risen due to increases in minimum temperatures.

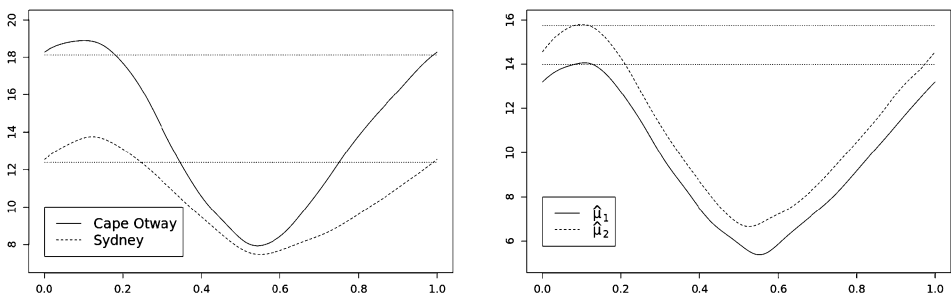


FIG. 1. Mean functions for the Australian temperature data. Left panel: Estimated mean functions of the Cape Otway and Sydney series for the two-sample case. Right panel: Estimated mean functions before and after the estimated change-point for the Melbourne temperature series.

TABLE 2

Summary of the bootstrap change-point procedure for relevant hypotheses with varying  $\Delta$  for the annual temperature curves. The label TRUE refers to a rejection of the null, the label FALSE to a failure to reject the null

$\Delta$	99%	97.5%	95%	90%
$q$	6.632	6.278	5.603	4.697
1.2	TRUE	TRUE	TRUE	TRUE
1.25	FALSE	TRUE	TRUE	TRUE
1.3	FALSE	FALSE	TRUE	TRUE
1.35	FALSE	FALSE	FALSE	TRUE
1.4	FALSE	FALSE	FALSE	FALSE

In summary, the results in this section highlight that there is strong evidence in the data for an increase in the mean function of Melbourne annual temperature profiles, with the maximum difference between “before” and “after” mean functions being at least 1.25 degrees centigrade. There is weak evidence that this difference is at least 1.35 degrees centigrade, but there is no support for the relevant hypothesis that it is even larger than that.

**Acknowledgments.** The authors thank Martina Stein, who typed parts of this manuscript with considerable technical expertise, Axel Bücher, Herold Dehling, Stanislav Volgushev, Joshua Gösmann and Daniel Meißner for very helpful discussions. The authors are also grateful to three referees for their constructive comments on an earlier version of this paper. Part of the research was done while A. Aue was visiting Ruhr-Universität Bochum as a Simons Visiting Professor of the Mathematical Research Institute Oberwolfach. Parts of this paper have been written during a visit of A. Aue and H. Dette at the Isaac Newton Institute, Cambridge, and these authors would like to thank the institute for its hospitality.

This work was supported in part by NSF Grants DMS-1305858 and DMS-1407530, and 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*) of the German Research Foundation.

## SUPPLEMENTARY MATERIAL

**Monte Carlo study and proofs** (DOI: [10.1214/19-AOS1842SUPPA](https://doi.org/10.1214/19-AOS1842SUPPA); .pdf). This supplement contains a detailed simulation study investigating the finite sample properties of the methodology introduced in this paper. Furthermore, the proofs of the theorems can be found here.

**R code** (DOI: [10.1214/19-AOS1842SUPPB](https://doi.org/10.1214/19-AOS1842SUPPB); .zip). This supplement contains all the R-scripts which were used for the simulations in this paper.

## REFERENCES

- [1] ASTON, J. A. D. and KIRCH, C. (2012). Detecting and estimating changes in dependent functional data. *J. Multivariate Anal.* **109** 204–220. MR2922864 <https://doi.org/10.1016/j.jmva.2012.03.006>
- [2] ASTON, J. A. D. and KIRCH, C. (2012). Evaluating stationarity via change-point alternatives with applications to fMRI data. *Ann. Appl. Stat.* **6** 1906–1948. MR3058688 <https://doi.org/10.1214/12-AOAS565>
- [3] AUE, A., GABRYS, R., HORVÁTH, L. and KOKOSZKA, P. (2009). Estimation of a change-point in the mean function of functional data. *J. Multivariate Anal.* **100** 2254–2269. MR2560367 <https://doi.org/10.1016/j.jmva.2009.04.001>
- [4] AUE, A. and HORVÁTH, L. (2013). Structural breaks in time series. *J. Time Series Anal.* **34** 1–16. MR3008012 <https://doi.org/10.1111/j.1467-9892.2012.00819.x>

- [5] AUE, A., NORINHO, D. D. and HÖRMANN, S. (2015). On the prediction of stationary functional time series. *J. Amer. Statist. Assoc.* **110** 378–392. MR3338510 <https://doi.org/10.1080/01621459.2014.909317>
- [6] AUE, A., RICE, G. and SÖNMEZ, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 509–529. MR3798876 <https://doi.org/10.1111/rssb.12257>
- [7] BERKES, I., GABRYS, R., HORVÁTH, L. and KOKOSZKA, P. (2009). Detecting changes in the mean of functional observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 927–946. MR2750251 <https://doi.org/10.1111/j.1467-9868.2009.00713.x>
- [8] BERKES, I., HORVÁTH, L. and RICE, G. (2013). Weak invariance principles for sums of dependent random functions. *Stochastic Process. Appl.* **123** 385–403. MR3003356 <https://doi.org/10.1016/j.spa.2012.10.003>
- [9] BERKSON, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *J. Amer. Statist. Assoc.* **33** 526–536.
- [10] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York. MR0233396
- [11] BOSQ, D. (2000). *Linear Processes in Function Spaces: Theory and Applications. Lecture Notes in Statistics* **149**. Springer, New York. MR1783138 <https://doi.org/10.1007/978-1-4612-1154-9>
- [12] BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** 107–144. MR2178042 <https://doi.org/10.1214/154957805100000104>
- [13] BUCCHIA, B. and WENDLER, M. (2017). Change-point detection and bootstrap for Hilbert space valued random fields. *J. Multivariate Anal.* **155** 344–368. MR3607900 <https://doi.org/10.1016/j.jmva.2017.01.007>
- [14] BÜCHER, A. and KOJADINOVIC, I. (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli* **22** 927–968. MR3449804 <https://doi.org/10.3150/14-BEJ682>
- [15] CAO, G. (2014). Simultaneous confidence bands for derivatives of dependent functional data. *Electron. J. Stat.* **8** 2639–2663. MR3292952 <https://doi.org/10.1214/14-EJS967>
- [16] CAO, G., YANG, L. and TODEM, D. (2012). Simultaneous inference for the mean function based on dense functional data. *J. Nonparametr. Stat.* **24** 359–377. MR2921141 <https://doi.org/10.1080/10485252.2011.638071>
- [17] CHOI, H. and REIMHERR, M. (2018). A geometric approach to confidence regions and bands for functional parameters. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 239–260. MR3744720 <https://doi.org/10.1111/rssb.12239>
- [18] DEGRAS, D. A. (2011). Simultaneous confidence bands for nonparametric regression with functional data. *Statist. Sinica* **21** 1735–1765. MR2895997 <https://doi.org/10.5705/ss.2009.207>
- [19] DEHLING, H. (1983). Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. Wahrsch. Verw. Gebiete* **63** 393–432. MR0705631 <https://doi.org/10.1007/BF00542537>
- [20] DEHLING, H. and PHILIPP, W. (1982). Almost sure invariance principles for weakly dependent vector-valued random variables. *Ann. Probab.* **10** 689–701. MR0659538
- [21] DETTE, H., KOKOT, K. and AUE, A. (2020). Supplement to “Functional data analysis in the Banach space of continuous functions.” <https://doi.org/10.1214/19-AOS1842SUPPA>, <https://doi.org/10.1214/19-AOS1842SUPPB>.
- [22] DETTE, H. and WIED, D. (2016). Detecting relevant changes in time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 371–394. MR3454201 <https://doi.org/10.1111/rssb.12121>
- [23] FERRATY, F. and VIEU, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics*. Springer, New York. MR2229687
- [24] FREMDT, S., HORVÁTH, L., KOKOSZKA, P. and STEINEBACH, J. G. (2014). Functional data analysis with increasing number of projections. *J. Multivariate Anal.* **124** 313–332. MR3147328 <https://doi.org/10.1016/j.jmva.2013.11.009>
- [25] GARLING, D. J. H. (1976). Functional central limit theorems in Banach spaces. *Ann. Probab.* **4** 600–611. MR0423449 <https://doi.org/10.1214/aop/1176996030>
- [26] HALL, P. and VAN KEILEGOM, I. (2007). Two-sample tests in functional data analysis starting from discrete data. *Statist. Sinica* **17** 1511–1531. MR2413533
- [27] HÖRMANN, S. and KOKOSZKA, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. MR2662361 <https://doi.org/10.1214/09-AOS768>
- [28] HORVÁTH, L. and KOKOSZKA, P. (2012). *Inference for Functional Data with Applications. Springer Series in Statistics*. Springer, New York. MR2920735 <https://doi.org/10.1007/978-1-4614-3655-3>
- [29] HORVÁTH, L., KOKOSZKA, P. and REEDER, R. (2013). Estimation of the mean of functional time series and a two-sample problem. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 103–122. MR3008273 <https://doi.org/10.1111/j.1467-9868.2012.01032.x>



- [30] HORVÁTH, L., KOKOSZKA, P. and RICE, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82. MR3153649 <https://doi.org/10.1016/j.jeconom.2013.11.002>
- [31] HUGHES, G. L., SUBBA RAO, S. and SUBBA RAO, T. (2007). Statistical analysis and time-series models for minimum/maximum temperatures in the Antarctic Peninsula. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463** 241–259.
- [32] JANSON, S. and KAIJSER, S. (2015). Higher moments of Banach space valued random variables. *Mem. Amer. Math. Soc.* **238** vii+110. MR3402381 <https://doi.org/10.1090/memo/1127>
- [33] KUELBS, J. (1973). The invariance principle for Banach space valued random variables. *J. Multivariate Anal.* **3** 161–172. MR0328986 [https://doi.org/10.1016/0047-259X\(73\)90020-1](https://doi.org/10.1016/0047-259X(73)90020-1)
- [34] KUELBS, J. and PHILIPP, W. (1980). Almost sure invariance principles for partial sums of mixing  $B$ -valued random variables. *Ann. Probab.* **8** 1003–1036. MR0602377
- [35] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes. Ergebnisse der Mathematik und Ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **23**. Springer, Berlin. MR1102015 <https://doi.org/10.1007/978-3-642-20212-4>
- [36] PAPARODITIS, E. and POLITIS, D. N. (2001). Tapered block bootstrap. *Biometrika* **88** 1105–1119. MR1872222 <https://doi.org/10.1093/biomet/88.4.1105>
- [37] POLITIS, D. N. and ROMANO, J. P. (1994). The stationary bootstrap. *J. Amer. Statist. Assoc.* **89** 1303–1313. MR1310224
- [38] RAČKAUSKAS, A. and SUQUET, C. (2004). Hölder norm test statistics for epidemic change. *J. Statist. Plann. Inference* **126** 495–520. MR2088755 <https://doi.org/10.1016/j.jspi.2003.09.004>
- [39] RAČKAUSKAS, A. and SUQUET, C. (2006). Testing epidemic changes of infinite dimensional parameters. *Stat. Inference Stoch. Process.* **9** 111–134. MR2249179 <https://doi.org/10.1007/s11203-005-0728-5>
- [40] RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2168993
- [41] SAMUR, J. D. (1984). Convergence of sums of mixing triangular arrays of random vectors with stationary rows. *Ann. Probab.* **12** 390–426. MR0735845
- [42] SAMUR, J. D. (1987). On the invariance principle for stationary  $\phi$ -mixing triangular arrays with infinitely divisible limits. *Probab. Theory Related Fields* **75** 245–259. MR0885465 <https://doi.org/10.1007/BF00354036>
- [43] VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Series in Statistics*. Springer, New York. MR1385671 <https://doi.org/10.1007/978-1-4757-2545-2>
- [44] WELLEK, S. (2010). *Testing Statistical Hypotheses of Equivalence and Noninferiority*, 2nd ed. CRC Press, Boca Raton, FL. MR2676002 <https://doi.org/10.1201/EBK1439808184>
- [45] ZHENG, S., YANG, L. and HÄRDLE, W. K. (2014). A smooth simultaneous confidence corridor for the mean of sparse functional data. *J. Amer. Statist. Assoc.* **109** 661–673. MR3223741 <https://doi.org/10.1080/01621459.2013.866899>