

# Moving block and tapered block bootstrap for functional time series with an application to the $K$ -sample mean problem

DIMITRIOS PILAVAKIS, EFSTATHIOS PAPARODITIS and  
THEOFANIS SAPATINAS

*Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus.  
E-mail: stathisp@ucy.ac.cy*

We consider infinite-dimensional Hilbert space-valued random variables that are assumed to be temporal dependent in a broad sense. We prove a central limit theorem for the moving block bootstrap and for the tapered block bootstrap, and show that these block bootstrap procedures also provide consistent estimators of the long run covariance operator. Furthermore, we consider block bootstrap-based procedures for fully functional testing of the equality of mean functions between several independent functional time series. We establish validity of the block bootstrap methods in approximating the distribution of the statistic of interest under the null and show consistency of the block bootstrap-based tests under the alternative. The finite sample behaviour of the procedures is investigated by means of simulations. An application to a real-life dataset is also discussed.

*Keywords:* functional time series;  $K$ -sample mean problem; mean function; moving block bootstrap; spectral density operator; tapered block bootstrap

## 1. Introduction

In statistical analysis, conclusions are commonly derived based on information obtained from a random sample of observations. In an increasing number of fields, these observations are curves or images which are viewed as functions in appropriate spaces, since an observed intensity is available at each point on a line segment, a portion of a plane or a volume. Such observed curves or images are called ‘functional data’; see, for example, Ramsay and Dalzell [24], who also introduced the term ‘functional data analysis’ (FDA) which refers to statistical methods used for analysing this kind of data.

In this paper, we consider observations stemming from a stochastic process  $\mathbb{X} = (X_t, t \in \mathbb{Z})$  of Hilbert space-valued random variables which satisfies certain stationarity and weak dependence properties. Our goal is to infer properties of the stochastic process based on an observed stretch  $X_1, X_2, \dots, X_n$ , i.e., on a functional time series. In this context, we commonly need to calculate the distribution, or parameters related to the distribution, of some statistics of interest based on  $X_1, X_2, \dots, X_n$ . Since in a functional set-up such quantities typically depend in a complicated way on infinite-dimensional characteristics of the underlying stochastic process  $\mathbb{X}$ , their calculation is difficult in practice. As a result, resampling methods and, in particular, bootstrap methodologies are very useful.

For the case of independent and identically distributed (i.i.d.) Banach space-valued random variables, Giné and Zinn [6] proved the consistency of the standard i.i.d. bootstrap for the sample mean. For functional time series, Politis and Romano [22] established validity of the stationary bootstrap for the sample mean and for (bounded) Hilbert space-valued random variables satisfying certain mixing conditions. A functional sieve bootstrap procedure for functional time series has been proposed by Paparoditis [18]. Consistency of the non-overlapping block bootstrap for the sample mean and for near epoch dependent Hilbert space-valued random variables has been established by Dehling, Sharipov and Wendler [5]. However, up to date, consistency results are not available for the moving block bootstrap (MBB) or its improved versions, like the tapered block bootstrap (TBB), for functional time series. Notice that the MBB for real-valued time series was introduced by Künsch [12] and Liu and Singh [15]. The basic idea is to resample blocks of the time series and to joint them together in the order selected in order to form a new set of pseudo observations. This resampling scheme retains the dependence structure of the time series within each block and can be, therefore, used to approximate the distribution of a wide range of statistics. The TBB for real-valued time series was introduced by Paparoditis and Politis [19]. It uses a taper window to downweight the observations at the beginning and at the end of each resampled block and improves the bias properties of the MBB.

The aim of this paper is twofold. First, we prove consistency of the MBB and of the TBB for the sample mean function in the case of weakly dependent Hilbert space-valued random variables. Furthermore, we show that these bootstrap methods provide consistent estimators of the covariance operator of the sample mean function estimator, that is of the spectral density operator of the underlying stochastic process at frequency zero. We derive our theoretical results under quite general dependence assumptions on  $\mathbb{X}$ , that is, under  $L^2$ - $m$ -approximability assumptions, which are satisfied by a large class of commonly used functional time series models; see, for example, Hörmann and Kokoszka [7]. Second, we apply the above mentioned bootstrap procedures to the problem of fully functional testing of the equality of the mean functions between a number of independent functional time series. Testing the equality of mean functions for i.i.d. functional data has been extensively discussed in the literature; see, for example, Benko, Härdle and Kneip [1], Horváth and Kokoszka ([8], Chapter 5), Zhang [29] and Staicu, Lahiri and Carroll [26]. Bootstrap alternatives over asymptotic approximations have been proposed in the same context by Benko, Härdle and Kneip [1], Zhang, Peng and Zhang [28] and, more recently, by Paparoditis and Sapatinas [20]. Testing equality of mean functions for dependent functional data has also attracted some interest in the literature. Horváth, Kokoszka and Reeder [9] developed an asymptotic procedure for testing equality of two mean functions for functional time series. Since the limiting null distribution of a fully functional,  $L^2$ -type test statistic, depends on difficult to estimate process characteristics, tests are considered which are based on a finite number of projections. A projection-based test has also been considered by Horváth and Rice [10]. Although such tests lead to manageable limiting distributions, they have non-trivial power only for deviations from the null which are not orthogonal to the subspace generated by the particular projections considered.

In this paper, we show that the MBB and TBB procedures can be successfully applied to approximate the distribution under the null of such fully functional test statistics. This is achieved by designing the suggested block bootstrap procedures in such a way that the generated pseudo-observations satisfy the null hypothesis of interest. Notice that such block bootstrap-based testing methodologies are applicable to a broad range of possible test statistics. As an example,

we prove validity for the  $L^2$ -type test statistic recently proposed by Horváth, Kokoszka and Reeder [9].

The paper is organised as follows. In Section 2, the basic assumptions on the underlying stochastic process  $\mathbb{X}$  are stated and the MBB and TBB procedures for weakly dependent, Hilbert space-valued random variables, are described. Asymptotic validity of the block bootstrap procedures for estimating the distribution of the sample mean function is established and consistency of the long run covariance operator, that is, of the spectral density operator of the underlying stochastic process at frequency zero, is proven. Section 3 is devoted to the problem of testing equality of mean functions for several independent functional time series. Theoretical justifications of an appropriately modified version of the MBB and of the TBB procedure for approximating the null distribution of a fully functional test statistic is given and consistency under the alternative is shown. Numerical simulations and a real-life data example are presented and discussed in Section 4. Auxiliary results and proofs of the main results are deferred to the Appendix and to the supplementary material [21].

## 2. Block bootstrap procedures for functional time series

### 2.1. Preliminaries and assumptions

We consider a strictly stationary stochastic process  $\mathbb{X} = \{X_t, t \in \mathbb{Z}\}$ , where the random variables  $X_t$  are random functions  $X_t(\omega, \tau)$ ,  $\tau \in \mathcal{I}$ ,  $\omega \in \Omega$ ,  $t \in \mathbb{Z}$ , defined on a probability space  $(\Omega, A, P)$  and take values in the separable Hilbert-space of squared-integrable  $\mathbb{R}$ -valued functions on  $\mathcal{I}$ , denoted by  $L^2(\mathcal{I})$ . The expectation function of  $X_t$ ,  $\mathbb{E}X_t \in L^2(\mathcal{I})$ , is independent of  $t$ , and it is denoted by  $\mu$ . Throughout Section 2, we assume for simplicity that  $\mu = 0$ . We define  $\langle f, g \rangle = \int_{\mathcal{I}} f(\tau)g(\tau) d\tau$ ,  $\|f\|^2 = \langle f, f \rangle$  and the tensor product between  $f$  and  $g$  by  $f \otimes g(\cdot) = \langle f, \cdot \rangle g$ . For two Hilbert–Schmidt operators  $\Psi_1$  and  $\Psi_2$ , we denote by  $\langle \Psi_1, \Psi_2 \rangle_{\text{HS}} = \sum_{i=1}^{\infty} \langle \Psi_1(e_i), \Psi_2(e_i) \rangle$  the inner product which generates the Hilbert–Schmidt norm  $\|\Psi_1\|_{\text{HS}} = \sum_{i=1}^{\infty} \|\Psi_1(e_i)\|^2$ , for  $\{e_i, i = 1, 2, \dots\}$  an orthonormal basis of  $L^2(\mathcal{I})$ . Without loss of generality, we assume that  $\mathcal{I} = [0, 1]$  (the unit interval) and, for simplicity, integral signs without the limits of integration imply integration over the interval  $\mathcal{I}$ . We finally write  $L^2$  instead of  $L^2(\mathcal{I})$ .

To describe the dependent structure of the stochastic process  $\mathbb{X}$ , we use the notion of  $L^p$ - $m$ -approximability; see Hörmann and Kokoszka [7]. A stochastic process  $\mathbb{X} = \{X_t, t \in \mathbb{Z}\}$  with  $X_t$  taking values in  $L^2$ , is called  $L^2$ - $m$ -approximable if the following conditions are satisfied:

- (i)  $X_t$  admits the representation

$$X_t = f(\delta_t, \delta_{t-1}, \delta_{t-2}, \dots) \tag{1}$$

for some measurable function  $f : S^\infty \rightarrow L^2$ , where  $\{\delta_t, t \in \mathbb{Z}\}$  is a sequence of i.i.d. elements in  $L^2$ .

- (ii)  $\mathbb{E}\|X_0\|^2 < \infty$  and

$$\sum_{m \geq 1} \sqrt{\mathbb{E}\|X_t - X_{t,m}\|^2} < \infty, \tag{2}$$

where  $X_{t,m} = f(\delta_t, \delta_{t-1}, \dots, \delta_{t-m+1}, \delta_{t,t-m}^{(m)}, \delta_{t,t-m-1}^{(m)}, \dots)$  and, for each  $t$  and  $k$ ,  $\delta_{t,k}^{(m)}$  is an independent copy of  $\delta_t$ .

The intuition behind the above definition is that the function  $f$  in (1) should be such that the effect of the innovations  $\delta_i$  far back in the past becomes negligible, that is, these innovations can be replaced by other, independent, innovations. We somehow strengthen (2) to the following assumption.

**Assumption 1.**  $\mathbb{X}$  is  $L^2$ - $m$ -approximable and satisfies

$$\lim_{m \rightarrow \infty} m \sqrt{\mathbb{E} \|X_t - X_{t,m}\|^2} = 0.$$

Notice that the above assumption is satisfied by many linear and non-linear functional time series models considered in the literature; see, for example, Hörmann and Kokoszka [7].

### 2.2. The moving block bootstrap

The main idea of the MBB is to split the data into overlapping blocks of length  $b$  and to obtain the bootstrapped pseudo-time series by joining together the  $k$  independently and randomly selected blocks of observations in the order selected. Here,  $k$  is a positive integer satisfying  $b(k - 1) < n$  and  $bk \geq n$ . For simplicity of notation, we assume throughout the paper that  $n = kb$ . Since the dependence of the original time series is maintained within each block, it is expected that for weakly dependent time series, this bootstrap procedure will, asymptotically, correctly imitate the entire dependence structure of the underlying stochastic process if the block length  $b$  increases to infinity, at some appropriate rate, as the sample size  $n$  increases to infinity. Adapting this resampling idea to a functional time series  $\mathbf{X}_n = \{X_t, t = 1, 2, \dots, n\}$  stemming from a strictly stationary stochastic process  $\mathbb{X} = \{X_t, t \in \mathbb{Z}\}$  with  $X_t$  taking values in  $L^2$  and  $\mathbb{E}(X_t) = 0$ , leads to the following MBB algorithm.

- Step 1:* Let  $b = b(n)$ ,  $1 \leq b < n$ , be an integer. Denote by  $B_t = \{X_t, X_{t+1}, \dots, X_{t+b-1}\}$  the block of length  $b$  starting from observation  $X_t$ ,  $t = 1, 2, \dots, N$ , where  $N = n - b + 1$  is the number of such blocks available.
- Step 2:* Define i.i.d. integer-valued random variables  $I_1, I_2, \dots, I_k$  having a discrete uniform distribution assigning the probability  $1/N$  to each element of the set  $\{1, 2, \dots, N\}$ .
- Step 3:* Let  $B_i^* = B_{I_i}$ ,  $i = 1, 2, \dots, k$ , and denote by  $\{X_{(i-1)b+1}^*, X_{(i-1)b+2}^*, \dots, X_{ib}^*\}$  the elements of  $B_i^*$ . Join the  $k$  blocks in the order  $B_1^*, B_2^*, \dots, B_k^*$  together to obtain a new set of functional pseudo observations of length  $n$  denoted by  $X_1^*, X_2^*, \dots, X_n^*$ .

The above bootstrap algorithm can be potentially applied to approximate the distribution of some statistic  $T_n = T(X_1, X_2, \dots, X_n)$  of interest. For instance, let  $T_n = \bar{X}_n$  be the sample mean function of the observed stretch  $X_1, X_2, \dots, X_n$ , i.e.,  $\bar{X}_n = n^{-1} \sum_{t=1}^n X_t$ . We are interested in estimating the distribution of  $\sqrt{n}\bar{X}_n$ . For this, the bootstrap random variable  $\sqrt{n}(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*))$  is used, where  $\bar{X}_n^*$  is the mean function of the functional pseudo observations  $X_1^*, X_2^*, \dots, X_n^*$ , i.e.,

$\bar{X}_n^* = n^{-1} \sum_{t=1}^n X_t^*$  and  $\mathbb{E}^*(\bar{X}_n^*)$  is the (conditional on the observations  $\mathbf{X}_n$ ) expected value of  $\bar{X}_n^*$ . Straightforward calculations yield

$$\mathbb{E}^*(\bar{X}_n^*) = \frac{1}{N} \left[ \sum_{t=1}^n X_t - \sum_{t=1}^{b-1} (1 - t/b)(X_t + X_{n-t+1}) \right].$$

It is known that, under a variety of dependence assumptions on the underlying mean zero stochastic process  $\mathbb{X}$ , it holds true that  $\sqrt{n}\bar{X}_n \xrightarrow{d} \Gamma$  as  $n \rightarrow \infty$ , where  $\Gamma$  denotes a Gaussian process with mean zero and long run covariance operator  $2\pi\mathcal{F}_0$ . Furthermore,  $\|n\mathbb{E}(\bar{X}_n \otimes \bar{X}_n) - 2\pi\mathcal{F}_0\|_{\text{HS}} \rightarrow 0$  as  $n \rightarrow \infty$ . Here,  $\mathcal{F}_\omega = (2\pi)^{-1} \sum_{h \in \mathbb{Z}} C_h e^{-ih\omega}$ ,  $\omega \in \mathbb{R}$ , is the so-called spectral density operator of  $\mathbb{X}$  and  $C_h$  denotes the lag  $h$  autocovariance operator of  $\mathbb{X}$ , defined by  $C_h(\cdot) = \mathbb{E}\langle X_t, \cdot \rangle X_{t+h}$  for any  $h \in \mathbb{Z}$ ; see Panaretos and Tavakoli [16,17].

The following theorem establishes validity of the MBB procedure for approximating the distribution of  $\sqrt{n}\bar{X}_n$  and for providing a consistent estimator of the long run covariance operator  $2\pi\mathcal{F}_0$ .

**Theorem 2.1.** *Suppose that the mean zero stochastic process  $\mathbb{X} = (X_t, t \in \mathbb{Z})$  satisfies Assumption 1 and let  $X_1^*, X_2^*, \dots, X_n^*$  be a stretch of pseudo observations generated by the MBB procedure. Assume that the block size  $b = b(n)$  satisfies  $b^{-1} + bn^{-1/2} = o(1)$  as  $n \rightarrow \infty$ . Then, as  $n \rightarrow \infty$ ,*

(i)

$$d(\mathcal{L}(\sqrt{n}(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) \mid \mathbf{X}_n), \mathcal{L}(\sqrt{n}\bar{X}_n)) \rightarrow 0, \quad \text{in probability,}$$

where  $d$  is any metric metrizing weak convergence on  $L^2$  and  $\mathcal{L}(Z)$  denotes the law of the random element  $Z$ . Furthermore,

(ii)

$$\|n\mathbb{E}^*(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) \otimes (\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) - n\mathbb{E}(\bar{X}_n \otimes \bar{X}_n)\|_{\text{HS}} = o_P(1), \quad \text{in probability.}$$

### 2.3. The tapered block bootstrap

The TBB procedure is a modification of the block bootstrap procedure considered in Section 2.2 which is obtained by introducing a tapering of the random elements  $X_t$ . The tapering function down-weights the endpoints of each block  $B_t$ , towards zero, that is, towards the mean function of  $X_t$ . The pseudo observations are then obtained by choosing, with replacement,  $k$  appropriately scaled and tapered blocks of length  $b$  of centered observations and joining them together.

More precisely, the TBB procedure applied to the functional time series  $\mathbf{Y}_n = \{Y_t, t = 1, 2, \dots, n\}$  stemming from a strictly stationary,  $L^2$ -values, stochastic process  $\mathbb{Y} = (Y_t, t \in \mathbb{Z})$  with  $\mathbb{E}Y_t = 0$ , can be described as follows. Let  $X_1, X_2, \dots, X_n$  be the centered observations, that is,  $X_t = Y_t - \bar{Y}_n$ ,  $t = 1, 2, \dots, n$ , where  $\bar{Y}_n = n^{-1} \sum_{t=1}^n Y_t$ . Furthermore, let  $b = b(n)$ ,  $1 \leq b < n$ , be an integer and let  $w_n(\cdot)$ ,  $n = 1, 2, \dots$ , be a sequence of so-called data-tapering windows which satisfy the following assumption:

**Assumption 2.**  $w_n(\tau) \in [0, 1]$  and  $w_n(\tau) = 0$  for  $\tau \notin \{1, 2, \dots, n\}$ . Furthermore,

$$w_n(\tau) = w\left(\frac{\tau - 0.5}{n}\right), \quad (3)$$

where the function  $w : \mathbb{R} \rightarrow [0, 1]$  fulfills the conditions: (i)  $w(\tau) \in [0, 1]$  for all  $\tau \in \mathbb{R}$  with  $w(\tau) = 0$  if  $\tau \notin [0, 1]$ ; (ii)  $w(\tau) > 0$  for all  $\tau$  in a neighbourhood of  $1/2$ ; (iii)  $w(\tau)$  is symmetric around  $\tau = 0.5$ ; and (iv)  $w(\tau)$  is nondecreasing for all  $\tau \in [0, 1/2]$ .

Let

$$\tilde{B}_i = \left\{ w_b(1) \frac{b^{1/2}}{\|w_b\|_2} X_i, w_b(2) \frac{b^{1/2}}{\|w_b\|_2} X_{i+1}, \dots, w_b(b) \frac{b^{1/2}}{\|w_b\|_2} X_{i+b-1} \right\},$$

be a block of length  $b$  starting from  $X_t$ ,  $t = 1, 2, \dots, N$ , where each centered observation is multiplied by  $w_b(\cdot)$  and scaled by  $b^{1/2}/\|w_b\|_2$ ,  $\|w_b\|_2^2 = \sum_{i=1}^b w_b(i)$ . Let  $I_1, I_2, \dots, I_k$  be i.i.d. integers selected from a discrete uniform distribution which assigns probability  $1/N$  to each element of the set  $\{1, 2, \dots, N\}$ . Let  $B_i^* = \tilde{B}_{I_i}$ ,  $i = 1, 2, \dots, k$ , and denote the  $i$ -th block selected by  $\{X_{(i-1)b+1}^*, X_{(i-1)b+2}^*, \dots, X_{ib}^*\}$ . Join these blocks together in the order  $B_1^*, B_2^*, \dots, B_k^*$  to form the set of TBB pseudo observations  $X_1^*, X_2^*, \dots, X_n^*$ .

Notice that the ‘‘inflation’’ factor  $b^{1/2}/\|w_b\|_2$  is necessary to compensate for the decrease of the variance of the  $X_i^*$ 's effected by the shrinking caused by the window  $w_b$ ; see, also, Paparoditis and Politis [19]. Furthermore, the TBB procedure uses the centered time series  $X_1, X_2, \dots, X_n$  instead of the original time series  $Y_1, Y_2, \dots, Y_n$ , in order to shrink the end points of the blocks towards zero.

To estimate the distribution of  $\sqrt{n}\bar{Y}_n$  by means of the TBB procedure, the bootstrap random variable  $\sqrt{n}(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*))$  is used, where  $\bar{X}_n^* = n^{-1} \sum_{t=1}^n X_t^*$  and  $\mathbb{E}^*(\bar{X}_n^*)$  is the (conditional on the observations  $\mathbf{Y}_n$ ) expected value of  $\bar{X}_n^*$ . Straightforward calculations yield

$$\begin{aligned} \mathbb{E}^*(\bar{X}_n^*) &= \frac{1}{N} \frac{\|w_b\|_1}{\|w_b\|_2} \left[ \sum_{t=1}^n X_t - \sum_{t=1}^{b-1} \left( 1 - \frac{\sum_{s=1}^t w_b(s)}{\|w_b\|_1} \right) X_t \right. \\ &\quad \left. - \sum_{j=1}^{b-1} \left( 1 - \frac{\sum_{t=b-j+1}^b w_b(t)}{\|w_b\|_1} \right) X_{n-j+1} \right]. \end{aligned}$$

The following theorem establishes validity of the TBB procedure for approximating the distribution of  $\sqrt{n}\bar{Y}_n$  and for providing a consistent estimator of the long run covariance operator  $2\pi\mathcal{F}_0$ .

**Theorem 2.2.** *Suppose that the mean zero stochastic process  $\mathbb{Y}$  satisfies Assumption 1 and let  $w_n(\cdot)$ ,  $n = 1, 2, \dots$ , be a sequence of data-tapering windows satisfying Assumption 2. Furthermore, let  $X_t^*$ ,  $t = 1, 2, \dots, n$ , be a stretch of pseudo observations generated by the TBB procedure. Assume that the block size  $b = b(n)$  satisfies  $b^{-1} + bn^{-1/2} = o(1)$  as  $n \rightarrow \infty$ . Then,*

as  $n \rightarrow \infty$ ,

(i)

$$d(\mathcal{L}(\sqrt{n}(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \mid \mathbf{Y}_n), \mathcal{L}(\sqrt{n}\overline{Y}_n)) \rightarrow 0, \quad \text{in probability,}$$

where  $d$  is any metric metrizing weak convergence on  $L^2$ , and

(ii)

$$\|n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) - n\mathbb{E}(\overline{Y}_n \otimes \overline{Y}_n)\|_{\text{HS}} = o_P(1), \quad \text{in probability.}$$

**Remark 2.1.** The asymptotic validity of the MBB and TBB procedures established in Theorem 2.1 and Theorem 2.2, respectively, can be extended to cover also the case where maps  $\phi : L^2 \rightarrow D$  of the sample means  $\overline{X}_n$  (in the MBB case) and  $\overline{Y}_n$  (in the TBB case) are considered, when  $D$  is a normed space. For instance, such a result follows as an application of a version of the delta-method appropriate for the bootstrap and for maps  $\phi$  which are Hadamard differentiable at 0 tangentially to a subspace  $D_0$  of  $D$  (see Theorem 3.9.11 of van der Vaart and Wellner [27]). Extensions of such results to almost surely convergence and for other types of differentiable maps, like for instance Fréchet differentiable functionals (see Theorem 3.9.15 of van der Vaart and Wellner [27]) or quasi-Hadamard differentiable functionals (see Theorem 3.1 of Beutner and Zähle [2]), are more involved since they depend on the particular map  $\phi$  and the verification of some technical conditions.

### 3. Bootstrap-based testing of the equality of mean functions

Among different applications, the MBB and TBB procedures can be also used to perform a test of the equality of mean functions between several independent samples of a functional time series. In this case, both block bootstrap procedures have to be implemented in such a way that the pseudo observations  $X_1^*, X_2^*, \dots, X_n^*$  generated, satisfy the null hypothesis of interest.

#### 3.1. The set-up

Consider  $K$  independent functional time series  $\mathbf{X}_M = \{X_{i,t}, i = 1, 2, \dots, K, t = 1, 2, \dots, n_i\}$ , each one of which satisfies

$$X_{i,t} = \mu_i + \varepsilon_{i,t}, \quad t = 1, 2, \dots, n_i, \tag{4}$$

where, for each  $i \in \{1, 2, \dots, K\}$ ,  $\{\varepsilon_{i,t}, t \in \mathbb{Z}\}$  is a  $L^2$ - $m$ -approximable functional process and  $n_i$  denotes the length of the  $i$ -th time series. Let  $M = \sum_{i=1}^K n_i$  be the total number of observations and note that  $\mu_i(\tau)$ ,  $\tau \in \mathcal{T}$ , is the mean function of the  $i$ -th functional time series,  $i = 1, 2, \dots, K$ . The null hypothesis of interest is then,

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_K$$

and the alternative hypothesis

$$H_1 : \exists k, m \in \{1, 2, \dots, K\} \text{ with } k \neq m \text{ such that } \mu_k \neq \mu_m.$$

Notice that the above equality is in  $L^2$ , i.e.,  $\mu_k = \mu_m$  means that  $\|\mu_m - \mu_k\| = 0$  whereas  $\mu_k \neq \mu_m$  that  $\|\mu_m - \mu_k\| > 0$ .

### 3.2. Block bootstrap-based testing

The aim is to generate a set of functional pseudo observations  $\mathbf{X}_M^* = X_{i,t}^*$ ,  $i = 1, 2, \dots, K$ ,  $t = 1, 2, \dots, n_i$ , using either the MBB procedure or the TBB procedure in such a way that  $H_0$  is satisfied. These bootstrap pseudo-time series can then be used to estimate the distribution of some test statistic  $T_M$  of interest which is applied to test  $H_0$ . Toward this, the distribution of  $T_M^*$  is used as an estimator of the distribution of  $T_M$ , where  $T_M^*$  is the same statistical functional as  $T_M$  but calculated using the bootstrap functional pseudo-time series  $\mathbf{X}_M^*$ .

To implement the MBB procedure for testing the null hypothesis of interest, assume, without loss of generality, that the test statistic  $T_M$  rejects the null hypothesis when  $T_M > d_{M,\alpha}$ , where, for  $\alpha \in (0, 1)$ ,  $d_{M,\alpha}$  denotes the upper  $\alpha$ -percentage point of the distribution of  $T_M$  under  $H_0$ . The MBB-based testing procedure goes then as follows:

- Step 1:* Calculate the sample mean functions in each population and the pooled mean function, that is, calculate  $\bar{X}_{i,n_i} = (1/n_i) \sum_{t=1}^{n_i} X_{i,t}$ , for  $i = 1, 2, \dots, K$ , and  $\bar{X}_M = (1/M) \sum_{i=1}^K \sum_{t=1}^{n_i} X_{i,t}$ , and obtain the residual functions in each population, that is, calculate  $\hat{\varepsilon}_{i,t} = X_{i,t} - \bar{X}_{i,n_i}$ , for  $t = 1, 2, \dots, n_i$ ;  $i = 1, 2, \dots, K$ .
- Step 2:* For  $i = 1, 2, \dots, K$ , let  $b_i = b_i(n) \in \{1, 2, \dots, n - 1\}$  be the block size for the  $i$ -th functional time series and divide  $\{\hat{\varepsilon}_{i,t}, t = 1, 2, \dots, n_i\}$  into  $N_i = n_i - b_i + 1$  overlapping blocks of length  $b_i$ , say,  $B_{i,1}, B_{i,2}, \dots, B_{i,N_i}$ . Calculate the sample mean of the  $\xi$ -th observations of the blocks  $B_{i,1}, B_{i,2}, \dots, B_{i,N_i}$ , i.e.,  $\bar{\varepsilon}_{i,\xi} = (1/N_i) \sum_{t=1}^{N_i} \hat{\varepsilon}_{i,\xi+t-1}$ , for  $\xi = 1, 2, \dots, b_i$ .
- Step 3:* For simplicity assume that  $n_i = k_i b_i$  and for  $i = 1, 2, \dots, K$ , let  $q_1^i, q_2^i, \dots, q_{k_i}^i$  be i.i.d. integers selected from a discrete probability distribution which assigns the probability  $1/N_i$  to each element of the set  $\{1, 2, \dots, N_i\}$ . Generate bootstrap functional pseudo observations  $X_{i,t}^*$ ,  $t = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, K$ , as  $X_{i,t}^* = \bar{X}_M + \varepsilon_{i,t}^*$ , where

$$\varepsilon_{i,\xi+(s-1)b_i}^* = \hat{\varepsilon}_{i,q_s^i+\xi-1} - \bar{\varepsilon}_{i,\xi}, \quad s = 1, 2, \dots, k_i, \xi = 1, 2, \dots, b_i. \quad (5)$$

- Step 4:* Let  $T_M^*$  be the same statistic as  $T_M$  but calculated using the bootstrap functional pseudo-time series  $X_{i,t}^*$ ,  $t = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, K$ . Denote by  $D_{M,T}^*$  the distribution of  $T_M^*$  given  $\mathbf{X}_M$ . For  $\alpha \in (0, 1)$ , reject the null hypothesis  $H_0$  if  $T_M > d_{M,\alpha}^*$ , where  $d_{M,\alpha}^*$  denotes the upper  $\alpha$ -percentage point of the distribution of  $T_M^*$ , i.e.,  $\mathbb{P}(T_M^* > d_{M,\alpha}^*) = \alpha$ .

Note that the distribution  $D_{M,T}^*$  can be evaluated by Monte-Carlo.

To motivate the centering used in Step 3, denote, for  $i = 1, 2, \dots, K$ , by  $e_{i,t}^*$ ,  $t = 1, 2, \dots, n_i$ , the pseudo observations generated by applying the MBB procedure, described in Section 2.2,

directly to the residuals time series  $\hat{\varepsilon}_{i,t}$ ,  $t = 1, 2, \dots, n_i$ . Note that the  $e_{i,t}^*$ 's differ from the  $\varepsilon_{i,t}^*$ 's used in (5) by the fact that the later are obtained after centering. The sample mean  $\bar{\varepsilon}_{i,\xi}$ ,  $i = 1, 2, \dots, K$ ,  $\xi = 1, 2, \dots, b_i$ , calculated in Step 2, is the (conditional on  $\mathbf{X}_M$ ) expected value of the pseudo observations  $e_{i,t}^*$ ,  $t = 1, 2, \dots, n_i$ , where  $t = \xi \pmod{b_i}$ . Furthermore, for  $i = 1, 2, \dots, K$ , we generate the  $\varepsilon_{i,t}^*$ 's,  $t = 1, 2, \dots, n_i$ , by subtracting  $\bar{\varepsilon}_{i,\xi}$  from  $e_{i,sb+\xi}^*$ ,  $\xi = 1, 2, \dots, b$ ,  $s = 1, 2, \dots, k_i$ . This is done in order for the (conditional on  $\mathbf{X}_M$ ) expected value of  $\varepsilon_{i,t}^*$  to be zero. In this way, the generated set of pseudo time series  $X_{i,t}^*$ ,  $t = 1, 2, \dots, n_i$ ,  $i = 1, 2, \dots, K$ , satisfy the null hypothesis  $H_0$ . In particular, given  $\mathbf{X}_M = \{X_{i,t}, i = 1, 2, \dots, K, t = 1, 2, \dots, n_i\}$ , we have

$$\mathbb{E}^*(X_{i,\xi+(s-1)b_i}^*) = \bar{X}_M + \frac{1}{N_i} \sum_{t=1}^{N_i} [\hat{\varepsilon}_{i,t+\xi-1} - \bar{\varepsilon}_{i,\xi}] = \bar{X}_M,$$

for  $i = 1, 2, \dots, K$ ,  $\xi = 1, 2, \dots, b$  and  $s = 1, 2, \dots, k_i$ . That is, conditional on  $\mathbf{X}_M$ , the mean function of each functional pseudo-time series  $X_{i,1}^*, X_{i,2}^*, \dots, X_{i,n}^*$ ,  $i = 1, 2, \dots, K$ , is identical in each population and equal to the pooled sample mean function  $\bar{X}_M$ .

An algorithm based on the TBB procedure for testing the same pair of hypotheses can also be implemented by modifying appropriate the MBB-based testing algorithm. In particular, we replace Step 2 and Step 3 of this algorithm by the following steps:

*Step 2:* For  $i = 1, 2, \dots, K$ , let  $b_i = b_i(n) \in \{1, 2, \dots, n - 1\}$  be the block size for the  $i$ -th functional time series and  $N_i = n_i - b_i + 1$ . Let also  $\{\hat{\varepsilon}_{i,t}, t = 1, 2, \dots, n_i\}$  be the centered values of  $\{\hat{\varepsilon}_{i,t}, t = 1, 2, \dots, n_i\}$ , i.e.,  $\hat{\varepsilon}_{i,t} = \hat{\varepsilon}_{i,t} - \bar{\varepsilon}_i$ , where  $\bar{\varepsilon}_i = (1/n_i) \sum_{t=1}^{n_i} \hat{\varepsilon}_{i,t}$ . Also, let  $w_{n_i}(\cdot)$ ,  $n_i = 1, 2, \dots$ , be a sequence of data-tapering windows satisfying Assumption 2. Now, for  $t = 1, 2, \dots, N_i$ , let

$$\tilde{B}_{i,t} = \left\{ w_{b_i}(1) \frac{b_i^{1/2}}{\|w_{b_i}\|_2} \hat{\varepsilon}_{i,t}, w_{b_i}(2) \frac{b_i^{1/2}}{\|w_{b_i}\|_2} \hat{\varepsilon}_{i,t+1}, \dots, w_{b_i}(b_i) \frac{b_i^{1/2}}{\|w_{b_i}\|_2} \hat{\varepsilon}_{i,t+b_i-1} \right\},$$

$i = 1, 2, \dots, K,$

where  $\|w_{b_i}\|_2^2 = \sum_{i=1}^{b_i} w_{b_i}(i)$ . Here,  $\tilde{B}_{i,t}$  denotes the tapered block of  $\hat{\varepsilon}_{i,t}$ 's of length  $b_i$  starting from  $\hat{\varepsilon}_{i,t}$ . Furthermore, for  $i = 1, 2, \dots, K$ , calculate the sample mean of the  $\xi$ th observations of the blocks  $\tilde{B}_{i,1}, \tilde{B}_{i,2}, \dots, \tilde{B}_{i,N_i}$ , i.e.,

$$\bar{\varepsilon}_{i,\xi} = \frac{1}{N_i} \sum_{t=1}^{N_i} w_{b_i}(\xi) \frac{b_i^{1/2}}{\|w_{b_i}\|_2} \hat{\varepsilon}_{i,\xi+t-1}, \quad \xi = 1, 2, \dots, b_i.$$

*Step 3:* For  $i = 1, 2, \dots, K$ , let  $q_1^i, q_2^i, \dots, q_{k_i}^i$  be i.i.d. integers selected from a discrete probability distribution which assigns the probability  $1/N_i$  to each  $t \in \{1, 2, \dots, N_i\}$ . Generate bootstrap functional pseudo-observations  $X_{i,t}^+$ ,  $i = 1, 2, \dots, K$ ,  $t = 1, 2, \dots, n_i$

according to  $X_{i,t}^+ = \bar{X}_M + \epsilon_{i,t}^+$ , where

$$\epsilon_{i,\xi+(s-1)b_i}^+ = w_b(\xi) \frac{b_i^{1/2}}{\|w_{b_i}\|_2} \hat{\epsilon}_{i,k'_s+\xi-1} - \bar{\epsilon}_{i,\xi}, \quad s = 1, 2, \dots, k_i, \xi = 1, 2, \dots, b_i.$$

As in the case of the MBB-based testing, the generation of  $\epsilon_{i,t}^+, t = 1, 2, \dots, n_i, i = 1, 2, \dots, K$ , ensures that the functional pseudo-time series  $X_{i,t}^+, t = 1, 2, \dots, n_i, i = 1, 2, \dots, K$ , satisfy  $H_0$ , that is, given  $\mathbf{X}_M = \{X_{i,t}, i = 1, 2, \dots, K, t = 1, 2, \dots, n_i\}$ , we have that  $\mathbb{E}^+(X_{i,t}^+) = \bar{X}_M$ .

### 3.3. Bootstrap validity

Notice that, since the proposed block bootstrap-based methodologies are not designed having any particular test statistic in mind, they can be potentially applied to a wide range of test statistics. To prove validity of the proposed block bootstrap-based testing procedures, however, a particular test statistic has to be considered. For instance, one such test statistic is the fully functional test statistic proposed by Horváth, Kokoszka and Reeder [9] for the case of  $K = 2$  populations. Let  $X_{i,t}, i = 1, 2, t = 1, 2, \dots, n_i$ , be two independent samples of curves, satisfying model (4). For  $i \in \{1, 2\}$  and for  $(u, v) \in [0, 1]^2$ , denote by  $c_i(u, v)$  the kernels of the long run covariance operators  $2\pi \mathcal{F}_0^{(i)}$ , given by  $c_i(u, v) = \mathbb{E}[\epsilon_{i,0}(u)\epsilon_{i,0}(v)] + \sum_{j \geq 1} \mathbb{E}[\epsilon_{i,0}(u)\epsilon_{i,j}(v)] + \sum_{j \geq 1} \mathbb{E}[\epsilon_{i,0}(v)\epsilon_{i,j}(u)]$ . The test statistic considered in Horváth, Kokoszka and Reeder [9], evaluates then the  $L^2$ -distance of the two sample mean functions  $\bar{X}_{i,n_i} = n_i^{-1} \sum_{t=1}^{n_i} X_{i,t}, i = 1, 2$ , and it is given by

$$U_M = \frac{n_1 n_2}{M} \int (\bar{X}_{1,n_1}(\tau) - \bar{X}_{2,n_2}(\tau))^2 d\tau,$$

where  $M = n_1 + n_2$ . Horváth, Kokoszka and Reeder [9] proved that if  $\min\{n_1, n_2\} \rightarrow \infty$  and  $n_1/M \rightarrow \theta \in (0, 1)$  then, under  $H_0$ ,  $U_M$  converges weakly to  $\int \Gamma^2(\tau) d\tau$ , where  $\{\Gamma(\tau) : \tau \in \mathcal{I}\}$  is a Gaussian process satisfying  $\mathbb{E}(\Gamma(\tau)) = 0$  and  $\mathbb{E}(\Gamma(u)\Gamma(v)) = (1 - \theta)c_1(u, v) + \theta c_2(u, v)$ . Notice that calculation of critical values of the above test requires estimation of the distribution of  $\int \Gamma^2(\tau) d\tau$  which is a difficult task.

Although the test statistic  $U_M$  is quite appealing because it is fully functional, its limiting distribution is difficult to implement which demonstrates the importance of the bootstrap. To investigate the consistency properties of the bootstrap, we first establish a general result which allows for the consideration of different test statistics that can be expressed as functionals of the basic deviation process

$$\left\{ \sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1,n_1}(\tau) - \bar{X}_{2,n_2}(\tau)), \tau \in \mathcal{I} \right\}. \tag{6}$$

**Theorem 3.1.** *Let Assumption 1 be satisfied. Assume that  $\min\{n_1, n_2\} \rightarrow \infty, n_1/M \rightarrow \theta \in (0, 1)$  and that, for  $i \in \{1, 2\}$ , the block size  $b_i = b_i(n)$  fulfills  $b_i^{-1} + b_i n_i^{-1/2} = o(1)$ , as  $n_i \rightarrow \infty$ . Then, conditional on  $\mathbf{X}_M$ , as  $n_i \rightarrow \infty$ ,*

(i)  $\sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1,n_1}^* - \bar{X}_{2,n_2}^*) \Rightarrow \Gamma$ , in probability,

and, if additionally Assumption 2 is satisfied,

(ii)  $\sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1,n_1}^+ - \bar{X}_{2,n_2}^+) \Rightarrow \Gamma$ , in probability.

Here,  $\Rightarrow$  denotes weak convergence in  $L^2$ .

By Theorem 3.1 and the continuous mapping theorem, the suggested block bootstrap-based testing procedures can be successfully applied to consistently estimate the distribution of any test statistic of interest which is a continuous function of the basic deviation process (6). We elaborate on some examples. Below,  $P_{H_0}(Z \leq \cdot)$  denotes the distribution function of the random variable  $Z$  when  $H_0$  is true.

Consider for instance the test statistic  $U_M$ . Let

$$U_M^* = \frac{n_1 n_2}{M} \int (\bar{X}_{1,n_1}^*(\tau) - \bar{X}_{2,n_2}^*(\tau))^2 d\tau$$

and

$$U_M^+ = \frac{n_1 n_2}{M} \int (\bar{X}_{1,n_1}^+(\tau) - \bar{X}_{2,n_2}^+(\tau))^2 d\tau,$$

where  $\bar{X}_{i,n_i}^* = (1/n_i) \sum_{t=1}^{n_i} X_{i,t}^*$  and  $\bar{X}_{i,n_i}^+ = \frac{1}{n_i} \sum_{t=1}^{n_i} X_{i,t}^+$ ,  $i = 1, 2$ . We then have the following result.

**Corollary 3.1.** *Let the assumptions of Theorem 3.1 be satisfied. Then,*

- (i)  $\sup_{x \in \mathbb{R}} |P(U_M^* \leq x | \mathbf{X}_M) - P_{H_0}(U_M \leq x)| \rightarrow 0$ , in probability, and
- (ii)  $\sup_{x \in \mathbb{R}} |P(U_M^+ \leq x | \mathbf{X}_M) - P_{H_0}(U_M \leq x)| \rightarrow 0$ , in probability.

**Remark 3.1.** If the following type of one-sided tests  $H_0 : \mu_1 = \mu_2$  versus  $H_1 : \mu_1 > \mu_2$  or  $H'_1 : \mu_1 < \mu_2$  is of interest (where  $\mu_1 = \mu_2$  (resp  $\mu_1 > \mu_2$  or  $\mu_1 < \mu_2$ ) means  $\mu_1(\tau) = \mu_2(\tau)$  (resp  $\mu_1(\tau) > \mu_2(\tau)$  or  $\mu_1(\tau) < \mu_2(\tau)$ ) for all  $\tau \in \mathcal{I}$ ), then the following test statistic

$$\tilde{U}_M = \sqrt{\frac{n_1 n_2}{M}} \int (\bar{X}_{1,n_1}(\tau) - \bar{X}_{2,n_2}(\tau)) d\tau$$

can be used. In this case,  $H_0$  is rejected if  $\tilde{U}_M > \tilde{d}_{M,\alpha}$  or  $\tilde{U}_M < \tilde{d}_{M,\alpha}$ , respectively, with  $\tilde{d}_{M,\alpha}$  the upper  $\alpha$ -percentage point of the distribution of  $\tilde{U}_M$  under  $H_0$ . Consistent estimators of this distribution can be also obtained using the block bootstrap procedures discussed. In particular, the following results can be established:

- (i)  $\sup_{x \in \mathbb{R}} |P(\tilde{U}_M^* \leq x | \mathbf{X}_M) - P_{H_0}(\tilde{U}_M \leq x)| \rightarrow 0$ , in probability, and
- (ii)  $\sup_{x \in \mathbb{R}} |P(\tilde{U}_M^+ \leq x | \mathbf{X}_M) - P_{H_0}(\tilde{U}_M \leq x)| \rightarrow 0$ , in probability,

where

$$\tilde{U}_M^* = \sqrt{\frac{n_1 n_2}{M}} \int (\bar{X}_{1,n_1}^*(\tau) - \bar{X}_{2,n_2}^*(\tau)) d\tau$$

and

$$\tilde{U}_M^+ = \sqrt{\frac{n_1 n_2}{M}} \int (\bar{X}_{1,n_1}^+(\tau) - \bar{X}_{2,n_2}^+(\tau)) d\tau.$$

To elaborate, notice that using Theorem 1 of Horváth, Kokoszka and Reeder [9], we get, as  $n_1, n_2 \rightarrow \infty$ , that

$$\left( \frac{1}{\sqrt{n_1}} \sum_{j=1}^{n_1} (X_{1,j} - \mu_1), \frac{1}{\sqrt{n_2}} \sum_{j=1}^{n_2} (X_{2,j} - \mu_2) \right) \Rightarrow (\Gamma_1, \Gamma_2),$$

where  $\Gamma_1$  and  $\Gamma_2$  are two independent Gaussian random elements in  $L^2$  with mean zero and covariance operators  $C_1$  and  $C_2$  with kernels  $c_1(\cdot, \cdot)$  and  $c_2(\cdot, \cdot)$ , respectively. Under  $H_0$ , and for  $\tilde{\mu} = \mu_1 = \mu_2$  the common mean of the two populations, we have

$$\sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1,n_1}(\tau) - \bar{X}_{2,n_2}(\tau)) = \sqrt{\frac{n_2}{M}} \frac{1}{\sqrt{n_1}} \sum_{t=1}^{n_1} (X_{1,t} - \tilde{\mu}) - \sqrt{\frac{n_1}{M}} \frac{1}{\sqrt{n_2}} \sum_{t=1}^{n_2} (X_{2,t} - \tilde{\mu}),$$

which implies, for  $n_1, n_2 \rightarrow \infty$  and  $n_1/M \rightarrow \theta$ , that  $\tilde{U}_M \xrightarrow{d} \int \Gamma(\tau) d\tau$ , where  $\Gamma(\tau) = \sqrt{1-\theta}\Gamma_1(\tau) - \sqrt{\theta}\Gamma_2(\tau)$ ,  $\tau \in \mathcal{I}$ . Now, working along the same lines as in the proof of Theorem 3.1, it can be easily shown that  $\tilde{U}_M^*$  and  $\tilde{U}_M^+$  converges weakly to the same limit  $\int \Gamma(\tau) d\tau$ .

Another interesting class of test statistics for which Theorem 3.1 allows for a successful application of the suggested block bootstrap-based testing procedures, is the class of so-called projection-based tests. To elaborate, let  $\{\varphi_1, \varphi_2, \dots, \varphi_p\}$  be a set of  $p$  orthonormal functions in  $L_2$ . A common choice is to let  $\varphi_j$  be the orthonormalized eigenfunctions corresponding to the  $p$  largest eigenvalues of the covariance operator of the stochastic process  $\{\Gamma(\tau) = \sqrt{1-\theta}\Gamma_1(\tau) - \sqrt{\theta}\Gamma_2(\tau), \tau \in \mathcal{I}\}$ , which are assumed to be distinct and strictly positive. A test statistic  $S_{p,M}$  can then be considered which is defined as

$$S_{p,M} = \frac{n_1 n_2}{M} \sum_{k=1}^p \langle \bar{X}_{1,n_1} - \bar{X}_{2,n_2}, \hat{\varphi}_k \rangle^2,$$

and where  $\hat{\varphi}_k$  are estimators of  $\varphi_k$ ; see for instance Horváth, Kokoszka and Reeder [9] where studentized versions of  $\langle \bar{X}_{1,n_1} - \bar{X}_{2,n_2}, \hat{\varphi}_k \rangle$  have also been used.

The following result establishes consistency of the suggested block bootstrap methods also for this class of test statistics.

**Corollary 3.2.** *Let the assumptions of Theorem 3.1 be satisfied and assume that the  $p$  largest eigenvalues of the covariance operator of the stochastic process  $\{\Gamma(\tau) = \sqrt{1-\theta}\Gamma_1(\tau) - \sqrt{\theta}\Gamma_2(\tau), \tau \in \mathcal{I}\}$  are distinct and positive. Let  $\varphi_k, k = 1, 2, \dots, p$ , be the orthonormalized eigenfunctions corresponding to these eigenvalues and let  $\tilde{\varphi}_k$  and  $\hat{\varphi}_k$  be estimators of  $\varphi_k$  satisfying  $\max_{1 \leq k \leq p} \|\tilde{\varphi}_k - \tilde{c}_k \varphi_k\| \xrightarrow{P} 0$  and  $\max_{1 \leq k \leq p} \|\hat{\varphi}_k - \hat{c}_k \varphi_k\| \xrightarrow{P} 0$ , where  $\tilde{c}_k = \text{sign}(\langle \tilde{\varphi}_k, \varphi_k \rangle)$  and  $\hat{c}_k = \text{sign}(\langle \hat{\varphi}_k, \varphi_k \rangle)$ . Then,*

- (i)  $\sup_{x \in \mathbb{R}} |P(S_{p,M}^* \leq x | \mathbf{X}_M) - P_{H_0}(S_{p,M} \leq x)| \rightarrow 0$ , in probability, and
- (ii)  $\sup_{x \in \mathbb{R}} |P(S_{p,M}^+ \leq x | \mathbf{X}_M) - P_{H_0}(S_{p,M} \leq x)| \rightarrow 0$ , in probability,

where  $S_{p,M}^* = (n_1 n_2 / M) \sum_{k=1}^p \langle \bar{X}_{1,n_1}^* - \bar{X}_{2,n_2}^*, \tilde{\varphi}_k \rangle^2$  and  $S_{p,M}^+ = (n_1 n_2 / M) \sum_{k=1}^p \langle \bar{X}_{1,n_1}^+ - \bar{X}_{2,n_2}^+, \tilde{\varphi}_k \rangle^2$ .

**Remark 3.2.** In Corollary 3.2, we allow for  $\tilde{\varphi}_k$  to be a different estimator of  $\varphi_k$  than  $\hat{\varphi}_k$ , where the later is used in the test statistic  $S_{p,M}$ . For instance,  $\tilde{\varphi}_k$  could be the same estimator as  $\hat{\varphi}_k$  but based on the the bootstrap pseudo observations  $X_{i,t}^*$ ,  $i = 1, 2, \dots, k$  and  $t = 1, 2, \dots, n_i$ , respectively,  $X_{i,t}^+$ ,  $i = 1, 2, \dots, k$  and  $t = 1, 2, \dots, n_i$ . This will allow for the bootstrap statistics  $S_{p,M}^*$ , respectively  $S_{p,M}^+$ , to also imitate the effect of the estimation error of the unknown eigenfunctions  $\varphi_k$  on the distribution of  $S_{p,M}$ . Clearly, a simple and computationally easier alternative will be to set  $\tilde{\varphi}_k = \hat{\varphi}_k$ .

**Remark 3.3.** If the alternative hypothesis  $H_1$  is true, then under the same assumptions as in Theorem 4 of Horváth, Kokoszka and Reeder [9], we get that  $U_M \xrightarrow{P} \infty$ . Furthermore, under the same assumptions as in Theorem 6 of Horváth, Kokoszka and Reeder [9], we get that  $S_{p,M} \xrightarrow{P} \infty$  provided that  $\langle \mu_1 - \mu_2, \varphi_k \rangle \neq 0$  for at least one  $1 \leq k \leq p$ . Corollary 3.1 and Corollary 3.2 (together with Slutsky’s theorem) imply then, respectively, the consistency of the test  $U_M$  using the bootstrap critical values obtained from the distributions of  $U_M^*$  and  $U_M^+$ , and of the test  $S_{p,M}$  using the bootstrap critical values obtained from the distributions of  $S_{p,M}^*$  and  $S_{p,M}^+$ .

### 4. Numerical examples

We generated functional time series stemming from a first order functional autoregressive model (FAR(1))

$$\varepsilon_t(u) = \int \psi(u, v) \varepsilon_{t-1}(v) dv + B_t(u), \quad u \in [0, 1], \tag{7}$$

(see also Horváth, Kokoszka and Reeder [9]), and from a first order functional moving average model (FMA(1)),

$$\varepsilon_t(u) = \int \psi(u, v) B_{t-1}(v) dv + B_t(u), \quad u \in [0, 1]. \tag{8}$$

For both models, the kernel function  $\psi(\cdot, \cdot)$  is defined by

$$\psi(u, v) = \frac{e^{-(u^2+v^2)/2}}{4 \int e^{-t^2} dt}, \quad (u, v) \in [0, 1]^2, \tag{9}$$

and the  $B_t$ ’s are i.i.d. Brownian bridges. All curves were approximated using  $T = 21$  equidistant points  $\tau_1, \tau_2, \dots, \tau_{21}$  in the unit interval  $\mathcal{I}$  and transformed into functional objects using the

Fourier basis with 21 basis functions (see Section 3 of the supplementary material for additional simulations with a larger  $T$ ).

Implementation of the MBB and TBB procedures require the selection of the block size  $b$ . As it has been shown in Theorem 2.1 and Theorem 2.2,  $n\mathbb{E}^*[(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) \otimes (\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*))]$  is a consistent estimator of  $2\pi\mathcal{F}_0$ , with kernel

$$c_N(u, v) = \frac{1}{N} \sum_{i=1}^n X_i(u)X_i(v) + \sum_{h=1}^{b-1} \left(1 - \frac{h}{b}\right) \frac{1}{N} \sum_{i=1}^{n-h} [X_i(u)X_{i+h}(v) + X_{i+h}(u)X_i(v)] + o_p(1),$$

in the MBB case, and

$$\tilde{c}_N(u, v) = \frac{1}{N} \sum_{i=1}^n Y_i(u)Y_i(v) + \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \frac{1}{N} \sum_{i=1}^{n-h} [Y_i(u)Y_{i+h}(v) + Y_{i+h}(u)Y_i(v)] + o_p(1),$$

in the TBB case, with  $\mathcal{W}_h = \sum_{i=1}^{b-h} w_b(i)w_b(i+h)$ ,  $h = 0, 1, \dots, b-1$ . Now,  $c_N$  and  $\tilde{c}_N$  can be considered as lag-window estimators of the kernel  $c(u, v) = \sum_{i=-\infty}^{\infty} \mathbb{E}[X_0(u)X_i(v)]$ , using the Bartlett window with “truncation lag”  $b$  in the MBB case and using the same “truncation lag” with the window function  $W = \mathcal{W}_h/\|w_b\|$ , in the TBB case. The above observations suggest that the problem of choosing the block size  $b$  can be considered as a problem of choosing the truncation lag of a lag window estimator of the function  $c(u, v)$ . Choice of the truncation lag in the functional context has been recently discussed in Horváth, Rice and Whipple [11] and Rice and Shang [25]. Although different procedures to select the “truncation lag” have been proposed in the aforementioned papers, we found the simple rule of setting  $b_i = \lceil n_i^{1/3} \rceil$ , where  $\lceil x \rceil$  is the least integer greater than or equal to  $x$ , quite effective in our numerical examples. In the following, we denote by  $b^*$  this choice of  $b$ , which is used together with some other values of  $b_i$ . A simulation study has been first conducted in order to investigate the finite sample performance of the MBB and TBB procedures. For this, the problem of estimating the standard deviation function of the normalized sample mean  $\sqrt{n}\bar{X}_n(\tau)$ , i.e., of  $\sqrt{c(\tau, \tau)}$  for different values of  $\tau \in [0, 1]$  has been considered. The results obtained using both block bootstrap procedures have also been compared with those using the stationary bootstrap (SB). Realizations of length  $n = 100$  and  $n = 500$  from the functional time series models (7) and (8) have been used. The results obtained are presented and discussed in Section 2 of the supplementary material. Furthermore, Table 1 of the supplementary material presents results comparing the performance of projections-based tests when asymptotic and bootstrap approximations are used to obtain the critical values of the tests.

### 4.1. Testing equality of mean functions

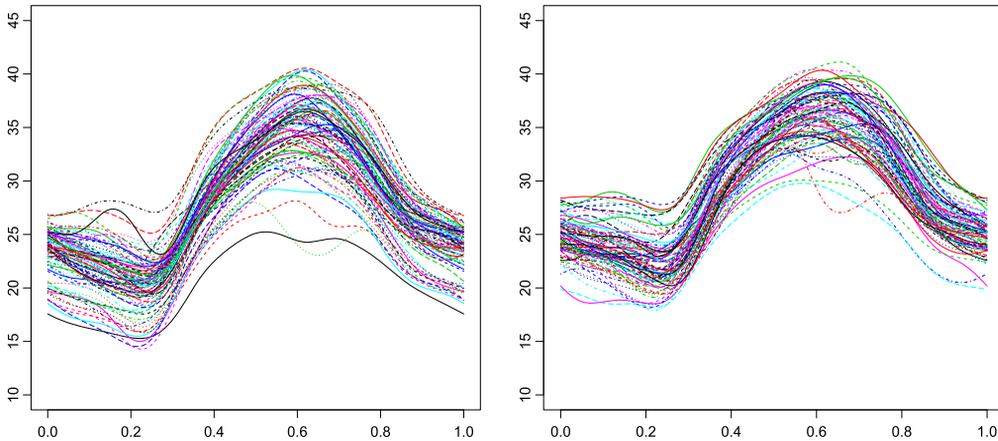
We investigate the size and power performance of the tests considered in Section 3.3. As can be seen in Section 2 of the supplementary material, the TBB estimators perform best in our simulations. For this reason, we concentrate in this section, on tests based on TBB critical values only. Two sample sizes  $n_1 = n_2 = 100$  and  $n_1 = n_2 = 200$  as well as a range of block sizes  $b =$

**Table 1.** Empirical size and power of the test based on TBB critical values and FAR(1) errors

$\gamma$	$n_1 = n_2 = 100$				$n_1 = n_2 = 200$			
	$b$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$b$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0	4	0.026	0.077	0.142	6	0.013	0.057	0.113
	6	0.015	0.061	0.112	8	0.010	0.052	0.115
	8	0.015	0.071	0.128	10	0.013	0.066	0.106
	$b^*$	0.027	0.074	0.143	$b^*$	0.013	0.057	0.113
0.2	4	0.048	0.135	0.206	6	0.058	0.160	0.237
	6	0.045	0.126	0.206	8	0.065	0.158	0.253
	8	0.034	0.118	0.185	10	0.070	0.162	0.247
	$b^*$	0.042	0.116	0.178	$b^*$	0.058	0.160	0.237
0.5	4	0.225	0.418	0.544	6	0.408	0.615	0.715
	6	0.200	0.374	0.499	8	0.411	0.632	0.759
	8	0.184	0.356	0.490	10	0.425	0.645	0.749
	$b^*$	0.218	0.424	0.532	$b^*$	0.408	0.615	0.715
0.8	4	0.584	0.772	0.853	6	0.864	0.966	0.980
	6	0.543	0.763	0.841	8	0.865	0.948	0.975
	8	0.529	0.739	0.831	10	0.843	0.948	0.976
	$b^*$	0.557	0.752	0.825	$b^*$	0.864	0.966	0.980
1	4	0.779	0.898	0.945	6	0.972	0.995	0.998
	6	0.746	0.891	0.941	8	0.975	0.994	0.999
	8	0.755	0.898	0.943	10	0.969	0.994	0.998
	$b^*$	0.769	0.901	0.945	$b^*$	0.972	0.995	0.998

$b_1 = b_2$ , are considered. The tests have been applied using three nominal levels, i.e.,  $\alpha = 0.01$ ,  $\alpha = 0.05$  and  $\alpha = 0.1$ . All bootstrap calculations are based on  $B = 1000$  bootstrap replicates and  $R = 1000$  model repetitions. To examine the empirical size and power behavior of the TBB-based test, the curves in the two samples were generated according to model (4) and with the errors  $\varepsilon_{i,t}$  following model (7), for  $i \in \{1, 2\}$ , with mean functions given by  $\mu_1(t) = 0$  and  $\mu_2(t) = \gamma t(1 - t)$  for the first and for the second population, respectively; see also Horváth, Kokoszka and Reeder [9]. The results obtained are shown in Table 1 for a range of values of  $\gamma$ . Notice that  $\gamma = 0$  corresponds to the null hypothesis.

As it is evident from this table, the TBB-based test statistic  $U_M^+$  has a good size behavior even in the case of  $n_1 = n_2 = 100$  observations while for  $n_1 = n_2 = 200$  observations the sizes of the TBB-based test are quite close to the nominal sizes for a range of block length values. It seems that the choice of the block size has a moderate effect on the power of the test. Furthermore, the power of the TBB-based test increases as the deviations from the null become larger (i.e., larger values of  $\gamma$ ) and/or as the sample size increases. Finally, using the suggested simple method to choose the block size  $b$ , the corresponding test has good size and power behavior in all cases.



**Figure 1.** Temperature curves: summer 2005 (left panel) and summer 2009 (right panel).

## 4.2. A real-life data example

We apply the TBB-based testing procedure to a data set consisting of the summer temperature measurements recorded in Nicosia, Cyprus, for the years 2005 and 2009. Our aim is to test whether there is a significant increase in the mean summer temperatures in 2009. The data consists of two samples of curves  $\{X_{i,t}(\tau), i = 1, 2, t = 1, 2, \dots, 92\}$ , where,  $X_{i,t}$  represents the temperature of day  $t$  of the summer 2005 for  $i = 1$  and of the summer 2009 for  $i = 2$ . More precisely,  $X_{i,1}$  represents the temperature of the 1st of June and  $X_{i,92}$  the temperature of the 31st of August. The temperature recordings have been taken in 15 minutes intervals, i.e., there are 96 temperature measurements for each day. These measurements have been transformed into functional objects using the Fourier basis with 21 basis functions. All curves are rescaled in order to be defined in the interval  $\mathcal{I}$ . Figure 1 shows the temperatures curves of the summer of 2005 and of 2009.

Since we are interested in checking whether there is an increase in the summer temperature in the year 2009 compared to 2005, the hypothesis of interest is  $H_0 : \mu_1(\tau) = \mu_2(\tau)$  versus  $H_1 : \mu_1(\tau) < \mu_2(\tau)$ , for all  $\tau \in \mathcal{I}$ . The  $p$ -values of the TBB-based test using the test statistic  $\tilde{U}_M$  are: 0.001 (for  $b = 4$ ), 0.003 (for  $b = 6$ ), 0.004 (for  $b = 8$ ) and 0.002 (for  $b = b^*$ ). These  $p$ -values have been obtained using  $B = 1000$  bootstrap replicates. As it is evident from these results, the  $p$ -values of the test statistic  $\tilde{U}_M$  are quite small leading to the rejection of  $H_0$  for all commonly used  $\alpha$ -levels.

## Appendix: Proofs

To prove Theorem 2.1 and Theorem 2.2, we first establish Lemma A.1 and Lemma A.2. The proofs of Lemma A.2 and Theorem 2.2 are given in Section 1 of the supplementary material.

Note also that, throughout the proofs, we use the fact that, by stationarity,  $\mathbb{E}\|X_{i,m} - X_i\| = \mathbb{E}\|X_{0,m} - X_0\|$  and  $\mathbb{E}\|X_{i,m}\| = \mathbb{E}\|X_i\| = \mathbb{E}\|X_0\|$  for all  $i \in \mathbb{Z}$ .

**Lemma A.1.** *Let  $g_b$  be a non-negative, continuous and bounded function defined on  $\mathbb{R}$ , satisfying  $g_b(0) = 1$ ,  $g_b(u) = g_b(-u)$ ,  $g_b(u) \leq 1$  for all  $u$ ,  $g_b(u) = 0$ , if  $|u| > c$ , for some  $c > 0$ . Suppose that  $(X_t, t \in \mathbb{Z})$  satisfies Assumption 1 and  $b = b(n)$  is a sequence of integers such that  $b^{-1} + bn^{-1/2} = o(1)$  as  $n \rightarrow \infty$ . Assume further that, for any fixed  $u$ ,  $g_b(u) \rightarrow 1$  as  $n \rightarrow \infty$ . Then, as  $n \rightarrow \infty$ ,*

$$\sum_{h=-b+1}^{b-1} g_b(h)\hat{\gamma}_h \xrightarrow{P} \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle),$$

where  $\hat{\gamma}_h = \frac{1}{n} \sum_{i=1}^{n-|h|} \langle X_i, X_{i+|h|} \rangle$  for  $-b + 1 \leq h \leq b - 1$ .

**Proof.** First, note by the independence of  $X_0$  and  $X_{i,i}$ , that  $\sum_{i=1}^{\infty} |\mathbb{E}\langle X_0, X_i \rangle| = \sum_{i=1}^{\infty} |\mathbb{E}\langle X_0, X_i - X_{i,i} \rangle| \leq \sum_{i=1}^{\infty} (\mathbb{E}\|X_0\|^2)^{1/2} (\mathbb{E}\|X_0 - X_{0,i}\|^2)^{1/2}$ , which implies by (2) that  $\sum_{i=-\infty}^{\infty} |\mathbb{E}\langle X_0, X_i \rangle| < \infty$ . Since  $n^{-1} \sum_{i=1}^n \langle X_i, X_i \rangle - \mathbb{E}\langle X_0, X_0 \rangle = o_P(1)$  as  $n \rightarrow \infty$ , it suffices to show that, as  $n \rightarrow \infty$ ,

$$\sum_{h=1}^{b-1} g_b(h) \frac{1}{n} \sum_{i=1}^{n-h} \langle X_i, X_{i+h} \rangle - \sum_{i \geq 1} \mathbb{E}\langle X_0, X_i \rangle = o_P(1). \tag{10}$$

Let  $c_{\infty}^+ = \sum_{i \geq 1} \mathbb{E}[\langle X_0, X_i \rangle]$ ,  $c_m^+ = \sum_{i \geq 1} \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle]$  and  $\hat{\gamma}_h^{(m)} = \frac{1}{n} \sum_{i=1}^{n-h} \langle X_{i,m}, X_{i+h,m} \rangle$ . Because

$$\left| \sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h - c_{\infty}^+ \right| \leq |c_m^+ - c_{\infty}^+| + \left| \sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)} - c_m^+ \right| + \left| \sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h - \sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)} \right|, \tag{11}$$

assertion (10) is proved by showing that there exists  $m_0 \in \mathbb{N}$  such that all three terms on the right-hand side of (11) can be made arbitrarily small in probability as  $n \rightarrow \infty$  for all  $m \geq m_0$ .

For the first term, we use the bound

$$\left| \sum_{i \geq 1} \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle] \right| \leq \left| \sum_{i=1}^m \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle] \right| + \left| \sum_{i=m+1}^{\infty} \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle] \right|, \tag{12}$$

and handle the first term on the right-hand side of (12) using  $\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle = \langle X_{0,m} - X_0, X_{i,m} \rangle + \langle X_0, X_{i,m} - X_i \rangle$ . Cauchy–Schwarz’s inequality and Assumption 1 yields that for

every  $\epsilon_1 > 0$ ,  $\exists m_1 \in \mathbb{N}$  such that

$$\begin{aligned} \left| \sum_{i=1}^m \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle] \right| &\leq 2 \sum_{i=1}^m (\mathbb{E} \|X_{0,m} - X_0\|^2 \mathbb{E} \|X_0\|^2)^{1/2} \\ &\leq 2(\mathbb{E} \|X_0\|^2)^{1/2} (m [\mathbb{E} \|X_{0,m} - X_0\|^2]^{1/2}) < \epsilon_1 \end{aligned}$$

for all  $m \geq m_1$ . For the second term of the right-hand side of (12), we get, using  $\langle X_0, X_i \rangle = \langle X_{i,i}, X_0 \rangle + \langle X_0, X_i - X_{i,i} \rangle$ , the fact that  $X_0$  and  $X_{i,i}$  as well as  $X_{0,m}$  and  $X_{i,m}$  are independent for  $i \geq m+1$  and Lemma 2.1 of Horváth and Kokoszka [8], that, for any  $\epsilon_2 > 0$ , there exists  $m_2 \in \mathbb{N}$  such that

$$\begin{aligned} \left| \sum_{i=m+1}^{\infty} \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle - \langle X_0, X_i \rangle] \right| &\leq \left| \sum_{i=m+1}^{\infty} \mathbb{E}[\langle X_{i,i}, X_0 \rangle] \right| + \left| \sum_{i=m+1}^{\infty} \mathbb{E}[\langle X_0, X_i - X_{i,i} \rangle] \right| \\ &\leq \sum_{i=m+1}^{\infty} (\mathbb{E} \|X_0\|^2 \mathbb{E} \|X_i - X_{i,i}\|^2)^{1/2} \\ &= (\mathbb{E} \|X_0\|^2)^{1/2} \sum_{i=m+1}^{\infty} (\mathbb{E} \|X_0 - X_{0,i}\|^2)^{1/2} < \epsilon_2 \end{aligned}$$

for all  $m \geq m_2$  because of (2). For the second term of (11), first note that, for every fixed  $m \geq 1$  and for any fixed  $h$ , we have that  $|\hat{\gamma}_h^{(m)} - \mathbb{E}[\langle X_{0,m}, X_{h,m} \rangle]| = o_p(1)$ . Furthermore, since  $\{X_{n,m}, n \in \mathbb{Z}\}$  is an  $m$ -dependent sequence,  $c_m^+ = \sum_{i=1}^m \mathbb{E}[\langle X_{0,m}, X_{i,m} \rangle]$ . Hence, the second term of the right-hand side of (11) is  $o_p(1)$ , if we show that  $|\sum_{h=m+1}^{b-1} g_b(h) \hat{\gamma}_h^{(m)}| = o_p(1)$ . We have  $\mathbb{E}[\sum_{h=m+1}^{b-1} g_b(h) \hat{\gamma}_h^{(m)}]^2 = n^{-2} \sum_{h_1=m+1}^{b-1} \sum_{h_2=m+1}^{b-1} \sum_{i_1=1}^{n-h_1} \sum_{i_2=1}^{n-h_2} g_b(h_1) g_b(h_2) \mathbb{E}(\langle X_{i_1,m}, X_{i_1+h_1,m} \rangle \langle X_{i_2,m}, X_{i_2+h_2,m} \rangle)$ . Since the sequence  $\{X_{i,m}, i \in \mathbb{Z}\}$  is  $m$ -dependent,  $X_{i,m}$  and  $X_{i+h,m}$  are independent for  $h \geq m+1$ , that is, using Lemma 2.1 of Horváth and Kokoszka [8] we have that  $\mathbb{E} \langle X_{i,m}, X_{i+h,m} \rangle = 0$  for the same  $h$ . Hence, the number of non-vanishing terms  $\mathbb{E}[\langle X_{i_1,m}, X_{i_1+h_1,m} \rangle \langle X_{i_2,m}, X_{i_2+h_2,m} \rangle]$  in the last equation above is of order  $O(nb)$  and, consequently,  $\mathbb{E}[\sum_{h=m+1}^{b-1} g_b(h) \hat{\gamma}_h^{(m)}]^2 = O(b/n) = o(1)$  from which the desired convergence follows by Markov's inequality. For the third term in (11), we show that, for  $m \geq m_0$ ,

$$\limsup_{n \rightarrow \infty} P \left( \left| \sum_{h=1}^{b-1} g_b(h) (\hat{\gamma}_h - \hat{\gamma}_h^{(m)}) \right| > \delta \right) = 0, \quad (13)$$

for all  $\delta > 0$ . From this, it suffices to show that, for  $m \geq m_0$ ,

$$\mathbb{E} \left| \sum_{h=1}^{b-1} g_b(h) (\hat{\gamma}_h - \hat{\gamma}_h^{(m)}) \right| = o(1). \quad (14)$$

Now, by the definitions of  $\hat{\gamma}_h$  and  $\hat{\gamma}_h^{(m)}$ , we have

$$\begin{aligned} \mathbb{E} \left| \sum_{h=1}^{b-1} g_b(h) (\hat{\gamma}_h - \hat{\gamma}_h^{(m)}) \right| &\leq \mathbb{E} \left| \frac{1}{n} \sum_{h=1}^m g_b(h) \sum_{i=1}^{n-h} (\langle X_i, X_{i+h} \rangle - \langle X_{i,m}, X_{i+h,m} \rangle) \right| \\ &\quad + \mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} (\langle X_i, X_{i+h} \rangle - \langle X_{i,m}, X_{i+h,m} \rangle) \right|. \end{aligned} \tag{15}$$

For the first term of the right-hand side of the above inequality, we use  $\langle X_i, X_{i+h} \rangle - \langle X_{i,m}, X_{i+h,m} \rangle = \langle X_i - X_{i,m}, X_{i+h} \rangle + \langle X_{i+h} - X_{i+h,m}, X_{i,m} \rangle$ , and we get, by to get, by Cauchy-Schwarz's inequality and simple algebra, that,

$$\begin{aligned} &\mathbb{E} \left| \frac{1}{n} \sum_{h=1}^m g_b(h) \sum_{i=1}^{n-h} (\langle X_i, X_{i+h} \rangle - \langle X_{i,m}, X_{i+h,m} \rangle) \right| \\ &\leq m [(\mathbb{E} \|X_0 - X_{0,m}\|^2 \mathbb{E} \|X_0\|^2)^{1/2} + (\mathbb{E} \|X_0 - X_{0,m}\|^2 \mathbb{E} \|X_{0,m}\|^2)^{1/2}]. \end{aligned}$$

Assumption 1 implies then that, for every  $\epsilon_3 > 0$ , there exists  $m_3 \in \mathbb{N}$  such that, for every  $m \geq m_3$ , the last quantity above is bounded by  $\epsilon_3$ . For the second term on the right-hand side of (15), we use the bound

$$\mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_i, X_{i+h} \rangle \right| + \mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_{i,m}, X_{i+h,m} \rangle \right|. \tag{16}$$

Note that the second summand of (16) is  $o(1)$ , while for the first term we use  $\langle X_i, X_{i+h} \rangle = \langle X_i, X_{i+h,h} \rangle + \langle X_i, X_{i+h} - X_{i+h,h} \rangle$  to get the bound

$$\mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_i, X_{i+h,h} \rangle \right| + \mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_i, X_{i+h} - X_{i+h,h} \rangle \right|. \tag{17}$$

For the last term of expression (17), we get, using (2), that for every  $\epsilon_4 > 0$ , there exists  $m_4 \in \mathbb{N}$  such that

$$\begin{aligned} \frac{1}{n} \sum_{h=m+1}^{b-1} \sum_{i=1}^{n-h} \mathbb{E} |\langle X_i, X_{i+h} - X_{i+h,h} \rangle| &\leq \sum_{h=m+1}^{b-1} \mathbb{E} |\langle X_0, X_h - X_{h,h} \rangle| \\ &\leq (\mathbb{E} \|X_0\|^2)^{1/2} \sum_{h=m+1}^{\infty} (\mathbb{E} \|X_0 - X_{0,h}\|^2)^{1/2} < \epsilon_4 \end{aligned}$$

for all  $m \geq m_4$ . Consider next the first term of (17). Because  $\langle X_i, X_{i+h,h} \rangle = \langle X_i - X_{i,h}, X_{i+h,h} \rangle + \langle X_{i,h}, X_{i+h,h} \rangle$ , we get for this term the bound

$$\mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_i - X_{i,h}, X_{i+h,h} \rangle \right| + \mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_{i,h}, X_{i+h,h} \rangle \right|. \quad (18)$$

The first term above is bounded by

$$\mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_i - X_{i,h}, X_{i+h,h} \rangle \right| \leq (\mathbb{E} \|X_0\|^2)^{1/2} \sum_{h=m+1}^{\infty} (\mathbb{E} \|X_0 - X_{0,h}\|^2)^{1/2}.$$

Thus, and by (2), for every  $\epsilon_5 > 0$ , there exists  $m_5 \in \mathbb{N}$  such that, for every  $m \geq m_5$ , this term is bounded by  $\epsilon_5$ . For the last term of (18), note that  $\{\langle X_{i,h}, X_{i+h,h} \rangle, i \in \mathbb{Z}\}$  is an  $2h$ -dependent stationary process, and since  $X_i$  and  $X_{i+h,h}$  are independent, that is,  $\mathbb{E} \langle X_i, X_{i+h,h} \rangle = 0$  for all  $i \in \mathbb{Z}$ ,  $\{\langle X_{i,h}, X_{i+h,h} \rangle, i \in \mathbb{Z}\}$  is then a mean zero  $2h$ -dependent stationary process which implies that  $n^{-1/2} \sum_{i=1}^n \langle X_{i,h}, X_{i+h,h} \rangle = O_P(1)$ . Using Portmanteau's theorem, and since the function  $f(x) = |x|$  is Lipschitz, we get that  $\mathbb{E} |n^{-1/2} \sum_{i=1}^n \langle X_{i,h}, X_{i+h,h} \rangle| = O(1)$ . Therefore,

$$\begin{aligned} \mathbb{E} \left| \frac{1}{n} \sum_{h=m+1}^{b-1} g_b(h) \sum_{i=1}^{n-h} \langle X_{i,h}, X_{i+h,h} \rangle \right| &\leq \frac{1}{\sqrt{n}} \sum_{h=m+1}^{b-1} \mathbb{E} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \langle X_{i,h}, X_{i+h,h} \rangle \right| \\ &= O(b/\sqrt{n}) = o(1), \end{aligned}$$

which concludes the proof of the lemma by choosing  $m_0 = \max\{m_1, m_2, m_3, m_4, m_5\}$ .  $\square$

**Lemma A.2.** *Suppose that  $(Y_t, t \in \mathbb{Z})$  satisfies Assumption 1 and that  $b = b(n)$  is a sequence of integers satisfying  $b^{-1} + bn^{-1/2} = o(1)$  as  $n \rightarrow \infty$ . Let  $w_n(\cdot), i = 1, 2, \dots$ , be a sequence of data-tapering windows satisfying Assumption 2. Then, as  $n \rightarrow \infty$ ,*

(i)

$$\sum_{|h| < b} \left( \frac{\mathcal{W}_{|h|}}{\|w_b\|_2^2} \right) \mathbb{E}[\langle Y_0, y \rangle \langle Y_h, y \rangle] \rightarrow \sum_{i=-\infty}^{\infty} \mathbb{E}[\langle Y_0, y \rangle \langle Y_i, y \rangle] \quad \text{for every } y \in L^2,$$

(ii)

$$\iint \{\tilde{c}_n(u, v) - c(u, v)\}^2 du dv = o_P(1),$$

where  $c(u, v) = \sum_{i=-\infty}^{\infty} \mathbb{E}[Y_0(u)Y_i(v)]$ ,  $\mathcal{W}_h = \sum_{i=1}^{b-h} w_b(i)w_b(i+h)$ ,  $h = 0, 1, \dots, b-1$  and

$$\tilde{c}_n(u, v) = \frac{1}{n} \sum_{i=1}^n Y_i(u)Y_i(v) + \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \frac{1}{n} \sum_{i=1}^{n-h} [Y_i(u)Y_{i+h}(v) + Y_{i+h}(u)Y_i(v)].$$

**Proof of Theorem 2.1.** By the triangle inequality and Theorem 1 of Horváth, Kokoszka and Reeder [9], the assertion of the theorem is established if we show that, as  $n \rightarrow \infty$ ,

$$\sqrt{n}(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) \Rightarrow \Gamma, \quad \text{in probability,} \tag{19}$$

where  $\Gamma$  is a Gaussian process in  $L^2$  with mean 0 and covariance operator  $C$  with kernel  $c(u, v) = \mathbb{E}(\Gamma(u)\Gamma(v))$  given for any  $u, v \in [0, 1]^2$  by

$$c(u, v) = \mathbb{E}[X_0(u)X_0(v)] + \sum_{i \geq 1} \mathbb{E}[X_0(u)X_i(v)] + \sum_{i \geq 1} \mathbb{E}[X_0(v)X_i(u)].$$

Using the notation  $S_n^* = \sqrt{n}(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*))$ , it follows from Proposition 7.4.2 of Laha and Rohatgi [13] that, to prove (19), it suffices to prove that,

- (L1)  $\langle S_n^*, y \rangle \xrightarrow{d} N(0, \sigma^2(y))$  for every  $y \in L^2$  where  $\sigma^2(y) = \langle C(y), y \rangle$ , and that
- (L2) the sequence  $\{S_n^*, n \in \mathbb{N}\}$  is tight.

Consider (L1). To establish the desired weak convergence, we prove that, as  $n \rightarrow \infty$ ,

$$\text{Var}^*(\langle S_n^*, y \rangle) \xrightarrow{P} \sigma^2(y) \tag{20}$$

and that

$$\frac{\langle S_n^*, y \rangle}{\sqrt{\text{Var}^*(\langle S_n^*, y \rangle)}} \xrightarrow{d} N(0, 1). \tag{21}$$

Consider (20) and notice that  $S_n^* = \frac{1}{\sqrt{k}} \sum_{i=1}^k [U_i^* - \mathbb{E}^*(U_i^*)]$ , where  $U_i^* = b^{-1/2}(X_{(i-1)b+1}^* + X_{(i-1)b+2}^* + \dots + X_{ib}^*)$ ,  $i = 1, 2, \dots, k$ . Due to the block bootstrap resampling scheme, the random variables  $U_i^*$ ,  $i = 1, 2, \dots, k$  are i.i.d. Thus, using  $\langle S_n^*, y \rangle = k^{-1/2} \sum_{i=1}^k [W_i^* - \mathbb{E}^*(W_i^*)]$ , where  $W_i^* = \langle U_i^*, y \rangle$ ,  $i = 1, 2, \dots, k$ , we have

$$\text{Var}^*(\langle S_n^*, y \rangle) = \mathbb{E}^*(W_1^*)^2 - (\mathbb{E}^*(W_1^*))^2. \tag{22}$$

Let  $\mu^* = \mathbb{E}^*(W_1^*)$  and  $U_i = b^{-1/2}(X_i + X_{i+1}, \dots + X_{i+b-1})$ ,  $i = 1, 2, \dots, N$ . We then have that

$$\mu^* = \frac{\sqrt{b}}{N} \left[ \sum_{i=1}^n \langle X_i, y \rangle - \sum_{j=1}^{b-1} \left(1 - \frac{j}{b}\right) [\langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle] \right]. \tag{23}$$

Therefore,  $\mathbb{E}^*(\mu^*) = 0$ . Using

$$\begin{aligned} & \left[ \sum_{i=1}^n \langle X_i, y \rangle - \sum_{j=1}^{b-1} \left(1 - \frac{j}{b}\right) (\langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle) \right]^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n \langle X_i, y \rangle \langle X_j, y \rangle - 2 \sum_{i=1}^n \sum_{j=1}^{b-1} \left(1 - \frac{j}{b}\right) \langle X_i, y \rangle [\langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle] \end{aligned}$$

$$+ \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{i}{b}\right) \left(1 - \frac{j}{b}\right) [\langle X_i, y \rangle + \langle X_{n-i+1}, y \rangle] [\langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle]$$

we get,

$$\mathbb{E}(\mu^*)^2 = \frac{b}{N^2} \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[\langle X_i, y \rangle \langle X_j, y \rangle] + O(b^2/n) = O(b^2/n), \tag{24}$$

where the last equality follows since, by Kronecker’s lemma,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}[\langle X_i, y \rangle \langle X_j, y \rangle] &= \sum_{|h|<n} \left(1 - \frac{|h|}{n}\right) \mathbb{E}[\langle X_0, y \rangle \langle X_h, y \rangle] \\ &\rightarrow \iint c(u, v) y(u) y(v) \, du \, dv \end{aligned} \tag{25}$$

as  $n \rightarrow \infty$ . Since  $\mathbb{E}^*(\mu^*) = 0$ , (24) implies that  $\mu^* = O_P(b/\sqrt{n})$ .

Consider next the first term of the right hand side of expression (22). For this term, we have

$$\begin{aligned} \mathbb{E}^*(W_1^*)^2 &= \frac{1}{N} \sum_{i=1}^N \langle U_i, y \rangle^2 \\ &= \frac{1}{N} \sum_{i=1}^n \langle X_i, y \rangle \langle X_i, y \rangle \\ &\quad + \sum_{h=1}^{b-1} \left(1 - \frac{h}{b}\right) \frac{1}{N} \sum_{i=1}^{n-h} [\langle X_i, y \rangle \langle X_{i+h}, y \rangle + \langle X_{i+h}, y \rangle \langle X_i, y \rangle] \\ &\quad - \frac{1}{N} \sum_{s=1}^{b-1} \left(1 - \frac{s}{b}\right) [\langle X_s, y \rangle \langle X_s, y \rangle + \langle X_{n-s+1}, y \rangle \langle X_{n-s+1}, y \rangle] \\ &\quad - \frac{1}{N} \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} \left(1 - \frac{j+t}{b}\right) [\langle X_j, y \rangle \langle X_{j+t}, y \rangle + \langle X_{n-j+1-t}, y \rangle \langle X_{n-j+1}, y \rangle \\ &\quad + \langle X_{j+t}, y \rangle \langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle \langle X_{n-j+1-t}, y \rangle]. \end{aligned} \tag{26}$$

Thus,

$$\begin{aligned} \mathbb{E}^*(W_1^*)^2 &= \frac{1}{N} \sum_{i=1}^n \langle X_i, y \rangle \langle X_i, y \rangle \\ &\quad + \sum_{h=1}^{b-1} \left(1 - \frac{h}{b}\right) \frac{1}{N} \sum_{i=1}^{n-h} [\langle X_i, y \rangle \langle X_{i+h}, y \rangle + \langle X_{i+h}, y \rangle \langle X_i, y \rangle] \end{aligned}$$

$$+ O_P(b/n) + O_P(b^2/n),$$

from which we get

$$\text{Var}^*(W_1^*) = \iint c_N(u, v)y(u)y(v) \, du \, dv + O_P(b^2/n), \tag{27}$$

where

$$c_N(u, v) = \frac{1}{N} \sum_{i=1}^n X_i(u)X_i(v) + \sum_{h=1}^{b-1} \left(1 - \frac{h}{b}\right) \frac{1}{N} \sum_{i=1}^{n-h} [X_i(u)X_{i+h}(v) + X_{i+h}(u)X_i(v)]. \tag{28}$$

By the ergodic theorem and equation (A.2) of Horváth, Kokoszka and Reeder [9], choosing the kernel  $K$  in their notation to be the kernel  $K(x) = (1 - |x|)\mathbb{1}_{[-1,1]}(x)$ , where  $\mathbb{1}_A(x)$  denotes the indicator function of  $A$ , it follows that

$$\iint [c_n(u, v) - c(u, v)]^2 \, du \, dv = o_P(1) \tag{29}$$

as  $n \rightarrow \infty$ , where  $c(u, v) = \sum_{i=-\infty}^{\infty} \mathbb{E}[X_0(u)X_i(v)]$  and  $c_n(u, v) = (N/n)c_N(u, v)$ . Using Cauchy–Schwarz’s inequality, we get that, as  $n \rightarrow \infty$ ,

$$\left| \iint (c_n(u, v) - c(u, v))y(u)y(v) \, du \, dv \right| \leq \left( \iint \{c_n(u, v) - c(u, v)\}^2 \, du \, dv \right)^{1/2} \|y\|^2 = o_P(1).$$

That is,

$$\iint c_n(u, v)y(u)y(v) \, du \, dv \xrightarrow{P} \iint c(u, v)y(u)y(v) \, du \, dv.$$

Since  $N/n \rightarrow 1$  as  $n \rightarrow \infty$ , we finally get from (27) that,

$$\text{Var}^*(S_n^*, y) = \text{Var}^*(W_1^*) \xrightarrow{P} \iint c(u, v)y(u)y(v) \, du \, dv = \sigma^2(y). \tag{30}$$

Consider next (21). Observe that  $W_i^* = \langle U_i^*, y \rangle, i = 1, 2, \dots, k$  are i.i.d. random variables and, therefore, it suffices to show that Lindeberg’s condition

$$\lim_{n \rightarrow \infty} \frac{1}{\tau_k^{*2}} \sum_{i=1}^k \mathbb{E}^*[(W_i^* - \mu^*)^2 \mathbb{1}(|W_i^* - \mu^*| > \varepsilon \tau_k^*)] = 0, \quad \text{for every } \varepsilon > 0, \tag{31}$$

is fulfilled, where  $\tau_k^{*2} = \sum_{i=1}^k \text{Var}^*(W_i^*) = k \text{Var}^*(W_1^*)$  and  $\mu^* = \mathbb{E}^*(W_i^*)$ . To establish (31), and because of (30), it suffices to show that, for any  $\delta > 0$  and as  $n \rightarrow \infty$ ,

$$P\left(\frac{1}{k} \sum_{i=1}^k \mathbb{E}^*[(W_i^* - \mu^*)^2 \mathbb{1}(|W_i^* - \mu^*| > \varepsilon \tau_k^*)] > \delta\right) \rightarrow 0. \tag{32}$$

Towards this, notice first that, for any two random variables  $X$  and  $Y$  and any  $\eta > 0$ , it yields that

$$\begin{aligned} & \mathbb{E}[|X + Y|^2 \mathbb{1}(|X + Y| > \eta)] \\ & \leq 4[\mathbb{E}|X|^2 \mathbb{1}(|X| > \eta/2) + \mathbb{E}|Y|^2 \mathbb{1}(|Y| > \eta/2)]; \end{aligned} \tag{33}$$

see Lahiri [14], p. 56. We then get by Markov’s inequality that

$$\begin{aligned} & P\left(\frac{1}{k} \sum_{t=1}^k \mathbb{E}^*[(W_t^* - \mu^*)^2 \mathbb{1}(|W_t^* - \mu^*| > \varepsilon \tau_k^*)] > \delta\right) \\ & \leq \delta^{-1} \mathbb{E}\{\mathbb{E}^*[(W_1^* - \mu^*)^2 \mathbb{1}(|W_1^* - \mu^*| > \varepsilon \tau_k^*)]\} \\ & = \delta^{-1} \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N (W_i - \mu^*)^2 \mathbb{1}(|W_i - \mu^*| > \varepsilon \tau_k^*)\right] \\ & = \delta^{-1} \mathbb{E}[(W_1 - \mu^*)^2 \mathbb{1}(|W_1 - \mu^*| > \varepsilon \tau_k^*)] \\ & \leq 4\delta^{-1} [\mathbb{E}W_1^2 \mathbb{1}(|W_1| > \varepsilon \tau_k^*/2) + \mathbb{E}(\mu^*)^2], \end{aligned} \tag{34}$$

where  $W_i = \langle U_i, y \rangle, i = 1, 2, \dots, N$ . Furthermore, we have

$$\mathbb{E}(W_1^2) = \mathbb{E}|\langle U_1, y \rangle|^2 = \sum_{|h|<b} \left(1 - \frac{|h|}{b}\right) \mathbb{E}[\langle X_0, y \rangle \langle X_h, y \rangle] \rightarrow \iint c(u, v) y(u) y(v) du dv,$$

as  $n \rightarrow \infty$ . Therefore, by the dominated convergence theorem,  $\lim_{n \rightarrow \infty} \mathbb{E}W_1^2 \mathbb{1}(|W_1| > \varepsilon \tau_k^*/2) = 0$ . Hence, using expression (24), we conclude that (34) converges to 0 as  $n \rightarrow \infty$ .

To establish (L2), it suffices, by Theorem 1.13 of Prokhorov [23] and Theorems 5.1 and 5.2 of Billingsley [3], to prove that  $\lim_{k \rightarrow \infty} \sup_{n \geq 1} \sum_{j=k}^{\infty} \mathbb{E}|\langle S_n^*, e_j \rangle|^2 = 0$ , where  $\{e_j, j \geq 1\}$  is a complete orthonormal basis of  $L^2$ . Using  $\mathbb{E}^*|\langle S_n^*, e_j \rangle|^2 = \text{Var}^*(\langle U_1^*, e_j \rangle)$  and Lemma 14 of Cerovecki and Hörmann [4], (L2) is satisfied if the following five conditions are fulfilled.

- (a)  $\text{Var}^*(\langle U_1^*, e_j \rangle) \geq 0 \forall j, n$ ;
- (b)  $\lim_{n \rightarrow \infty} \text{Var}^*(\langle U_1^*, e_j \rangle) = \Sigma_j$ , in probability;
- (c)  $\sum_{j \geq 1} \Sigma_j < \infty$ ;
- (d)  $\lim_{n \rightarrow \infty} \sum_{j \geq 1} \text{Var}^*(\langle U_1^*, e_j \rangle) = \sum_{j \geq 1} \Sigma_j$ , in probability;
- (e)  $\sum_{j \geq 1} \text{Var}^*(\langle U_1^*, e_j \rangle)$  is bounded for all  $n \geq 1$ , in probability.

Note that, by letting  $y = e_j$  in expression (30), property (b) follows with  $\Sigma_j = \iint c(u, v) e_j(u) \times e_j(v) du dv$ . To prove (c), notice that, by Proposition 6 of Horváth, Rice and Whipple [11], and since the stochastic process  $\{X_t, t \in \mathbb{Z}\}$  is  $L^2$ - $m$ -approximable, the covariance operator  $C$  with

kernel  $c(\cdot, \cdot)$  is trace class. Therefore,

$$\sum_{j \geq 1} \Sigma_j = \sum_{j \geq 1} \iint c(u, v) e_j(u) e_j(v) du dv = \sum_{j \geq 1} \lambda_j < \infty, \tag{35}$$

where  $\lambda_j, j \geq 1$  are the eigenvalues of  $C$ .

To establish (d), we get, using (23), that

$$\begin{aligned} \text{Var}^*(\langle U_1^*, e_j \rangle) &= \frac{1}{N} \sum_{i=1}^N \langle U_i, e_j \rangle^2 \\ &\quad - \left( \frac{\sqrt{b}}{N} \left[ \sum_{i=1}^n \langle X_i, e_j \rangle - \sum_{l=1}^{b-1} \left( 1 - \frac{l}{b} \right) [\langle X_l, e_j \rangle + \langle X_{n-l+1}, e_j \rangle] \right] \right)^2. \end{aligned} \tag{36}$$

By Parseval’s identity, we have,

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{1}{N} \sum_{i=1}^N |\langle U_i, e_j \rangle|^2 &= \frac{1}{N} \sum_{i=1}^N \|U_i\|^2 \\ &= \frac{1}{N} \sum_{i=1}^n \langle X_i, X_i \rangle + \sum_{h=1}^{b-1} \left( 1 - \frac{h}{b} \right) \frac{1}{N} \sum_{i=1}^{n-h} [\langle X_i, X_{i+h} \rangle + \langle X_{i+h}, X_i \rangle] \\ &\quad - \frac{1}{N} \sum_{s=1}^{b-1} \left( 1 - \frac{s}{b} \right) [\langle X_s, X_s \rangle + \langle X_{n-s+1}, X_{n-s+1} \rangle] \\ &\quad - \frac{1}{N} \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} \left( 1 - \frac{t+j}{b} \right) [\langle X_j, X_{j+t} \rangle + \langle X_{n-j+1-t}, X_{n-j+1} \rangle \\ &\quad + \langle X_{j+t}, X_j \rangle + \langle X_{n-j+1}, X_{n-j+1-t} \rangle]. \end{aligned}$$

Hence,

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{1}{N} \sum_{i=1}^N |\langle U_i, e_j \rangle|^2 &= \frac{1}{N} \sum_{i=1}^n \langle X_i, X_i \rangle \\ &\quad + \sum_{h=1}^{b-1} \left( 1 - \frac{h}{b} \right) \frac{1}{N} \sum_{i=1}^{n-h} [\langle X_i, X_{i+h} \rangle + \langle X_{i+h}, X_i \rangle] + O_P(b^2/n). \end{aligned} \tag{37}$$

Then, by letting  $g_b(h) = (1 - \frac{h}{b})$  in Lemma A.1, we get that, as  $n \rightarrow \infty$ ,

$$\sum_{j=1}^{\infty} \frac{1}{N} \sum_{i=1}^N \langle U_i, e_j \rangle^2 \xrightarrow{P} \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle). \tag{38}$$

For the second term of equation (36), we show that,

$$\sum_{j \geq 1} \left( \frac{\sqrt{b}}{N} \left[ \sum_{i=1}^n \langle X_i, e_j \rangle - \sum_{l=1}^{b-1} \left( 1 - \frac{l}{b} \right) [\langle X_l, e_j \rangle + \langle X_{n-l+1}, y \rangle] \right] \right)^2 = o_P(1), \tag{39}$$

as  $n \rightarrow \infty$ . Using  $\langle x, y \rangle = \sum_{j \geq 1} \langle x, e_j \rangle \langle y, e_j \rangle$ , we have

$$\begin{aligned} & \sum_{j \geq 1} \left[ \sum_{i=1}^n \langle X_i, e_j \rangle - \sum_{l=1}^{b-1} \left( 1 - \frac{l}{b} \right) (\langle X_l, e_j \rangle + \langle X_{n-l+1}, e_j \rangle) \right]^2 \\ &= \sum_{i=1}^n \sum_{l=1}^n \langle X_i, X_l \rangle - 2 \sum_{i=1}^n \sum_{l=1}^{b-1} \left( 1 - \frac{l}{b} \right) [\langle X_i, X_l \rangle + \langle X_i, X_{n-l+1} \rangle] \\ & \quad + \sum_{i=1}^{b-1} \sum_{l=1}^{b-1} \left( 1 - \frac{i}{b} \right) \left( 1 - \frac{l}{b} \right) [\langle X_i, X_l \rangle + \langle X_{n-i+1}, X_l \rangle \\ & \quad + \langle X_i, X_{n-l+1} \rangle + \langle X_{n-i+1}, X_{n-l+1} \rangle] \\ &= \sum_{i=1}^n \sum_{l=1}^n \langle X_i, X_l \rangle + O_P(nb) + O_P(b^2). \end{aligned}$$

Now note that, by the continuous mapping theorem and using Theorem 1 of Horváth, Kokoszka and Reeder [9], we get

$$\frac{1}{n} \sum_{i=1}^n \sum_{l=1}^n \langle X_i, X_l \rangle = \langle \sqrt{n \bar{X}_n}, \sqrt{n \bar{X}_n} \rangle = O_P(1). \tag{40}$$

Therefore,

$$\frac{b}{N^2} \left[ \sum_{i=1}^n \sum_{l=1}^n \langle X_i, X_l \rangle + O_P(nb) + O_P(b^2) \right] = O_P(b^2/n) = o_P(1),$$

which establishes (39). Hence, from (36), (38) and (39), we conclude that

$$\sum_{j \geq 1} \text{Var}^*(\langle U_1^*, e_j \rangle) \rightarrow \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle), \quad \text{in probability.} \tag{41}$$

Therefore, and by (35), property (d) is proved if we show that,

$$\sum_{j \geq 1} \lambda_j = \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle). \tag{42}$$

Using Mercer’s theorem, we have

$$\begin{aligned} \sum_{j \geq 1} \lambda_j &= \int c(u, u) \, du = \int \sum_{i=-\infty}^{\infty} \mathbb{E}[X_0(u)X_i(u)] \, du \\ &= \sum_{i=-\infty}^{\infty} \mathbb{E} \int [X_0(u)X_i(u)] \, du = \sum_{i=-\infty}^{\infty} \mathbb{E}\langle X_0, X_i \rangle. \end{aligned} \tag{43}$$

Notice that the above interchange of summation and integration is justified since, using Assumption 1, and the fact that  $X_0$  and  $X_{i,i}$  are independent for  $i \geq 1$ , we get

$$\begin{aligned} &\sum_{i=-\infty}^{\infty} \int |\mathbb{E}[X_0(u)X_i(u)]| \, du \\ &= \int |\mathbb{E}[X_0(u)X_0(u)]| \, du + 2 \sum_{i=1}^{\infty} \int |\mathbb{E}\{X_0(u)[X_i(u) - X_{i,i}(u)]\}| \, du \\ &\leq \int \mathbb{E}(X_0(u))^2 \, du + 2 \sum_{i=1}^{\infty} \left\{ \int \mathbb{E}[X_0(u)]^2 \, du \right\}^{1/2} \left\{ \int \mathbb{E}[X_i(u) - X_{i,i}(u)]^2 \, du \right\}^{1/2} \\ &\leq \mathbb{E}\|X_0\|^2 + 2(\mathbb{E}\|X_0\|^2)^{1/2} \sum_{i=1}^{\infty} (\mathbb{E}\|X_0 - X_{0,i}\|^2)^{1/2} < \infty. \end{aligned}$$

To prove (e), notice first that, by (36),  $\sum_{j=1}^{\infty} \text{Var}^*(\langle U_1^*, e_j \rangle) \leq \sum_{j=1}^{\infty} (1/N) \sum_{i=1}^N |\langle U_i, e_j \rangle|^2$  and, therefore, using (37), for any given  $n \geq 1$ ,  $\sum_{j=1}^{\infty} \text{Var}^*(\langle U_1^*, e_j \rangle)$  is bounded in probability. Furthermore, by (41), the sequence  $\{\sum_{j=1}^{\infty} \text{Var}^*(\langle U_1^*, e_j \rangle), n \geq 1\}$  converges in probability as  $n \rightarrow \infty$ .

Consider next assertion (ii) of the theorem. By the triangle inequality, it suffices to prove that as  $n \rightarrow \infty$ ,  $\|n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) - 2\pi\mathcal{F}_0\|_{\text{HS}} = o_P(1)$ . Now, recall that  $U_i^*$ ,  $i = 1, 2, \dots, n$ , are i.i.d., and note that

$$\begin{aligned} &n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) (y)(v) \\ &= \int \mathbb{E}^* [[U_1^*(u) - \mathbb{E}^*(U_1^*(u))][U_1^*(v) - \mathbb{E}^*(U_1^*(v))]] y(u) \, du, \end{aligned}$$

i.e.,  $n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*))$  is an integral operator with kernel

$$d(u, v) = \mathbb{E}^*[U_1^*(u)U_1^*(v)] - \mathbb{E}^*(U_1^*(u))\mathbb{E}^*(U_1^*(v)). \tag{44}$$

Now,

$$\begin{aligned}
 \mathbb{E}^*[U_1^*(u)U_1^*(v)] &= \frac{1}{N} \sum_{i=1}^n X_i(u)X_i(v) \\
 &+ \sum_{h=1}^{b-1} \left(1 - \frac{h}{b}\right) \frac{1}{N} \sum_{i=1}^{n-h} [X_i(u)X_{i+h}(v) + X_{i+h}(u)X_i(v)] \\
 &- \frac{1}{N} \sum_{s=1}^{b-1} \left(1 - \frac{s}{b}\right) [X_s(u)X_s(v) + X_{n-s+1}(u)X_{n-s+1}(v)] \\
 &- \frac{1}{N} \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} \left(1 - \frac{j+t}{b}\right) [X_j(u)X_{j+t}(v) \\
 &+ X_{n-j+1-t}(u)X_{n-j+1-t}(v) + X_{j+t}(u)X_j(v) \\
 &+ X_{n-j+1}(u)X_{n-j+1-t}(v)] \tag{45}
 \end{aligned}$$

and

$$\mathbb{E}^*(U_1^*(u)) = \frac{\sqrt{b}}{N} \left[ \sum_{i=1}^n X_i(u) - \sum_{j=1}^{b-1} \left(1 - \frac{j}{b}\right) (X_j(u) + X_{n-j+1}(u)) \right]. \tag{46}$$

Therefore,  $d(u, v) = c_N(u, v) + R(u, v)$ , where  $R(u, v)$  is defined as the difference of  $d(u, v)$  given in (44) and  $c_N(u, v)$  given in (28). Now, notice that  $2\pi \mathcal{F}_0(y)(v) = \int \sum_{h=-\infty}^{\infty} \mathbb{E}[X_0(u) \times X_h(v)]y(u) du$ , that is,  $2\pi \mathcal{F}_0$  is an integral operator with kernel  $c(u, v) = \sum_{h=-\infty}^{\infty} \mathbb{E}[X_0(u) \times X_h(v)]$ . Hence,

$$\begin{aligned}
 &\|n\mathbb{E}^*(\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) \otimes (\bar{X}_n^* - \mathbb{E}^*(\bar{X}_n^*)) - 2\pi \mathcal{F}_0\|_{\text{HS}} \\
 &= \iint [d(u, v) - c(u, v)]^2 du dv \\
 &\leq 2 \iint [c_N(u, v) - c(u, v)]^2 du dv + 2 \iint [R(u, v)]^2 du dv.
 \end{aligned}$$

Using (29) it suffices to prove that  $\iint [R(u, v)]^2 du dv = o_p(1)$ . To prove this, recall the inequality  $(\sum_{i=1}^L a_i)^2 \leq L \sum_{i=1}^L a_i^2$ , where  $L$  is a positive integer, and notice that, using (40),

$$\begin{aligned}
 &\frac{b^2}{N^4} \iint \left( \sum_{i=1}^n \sum_{j=1}^n X_i(u)X_j(v) \right)^2 du dv \\
 &= \frac{b^2}{N^2} \frac{1}{N} \sum_{i_1=1}^n \sum_{i_2=1}^n \int X_{i_1}(u)X_{i_2}(u) du \frac{1}{N} \sum_{j_1=1}^n \sum_{j_2=1}^n \int X_{j_1}(v)X_{j_2}(v) dv
 \end{aligned}$$

$$= \frac{b^2}{N^2} \left( \frac{1}{N} \sum_{i_1=1}^n \sum_{i_2=1}^n \langle X_{i_1}, X_{i_2} \rangle \right)^2 = O_P(b^2/N^2) = o_p(1). \tag{47}$$

Furthermore,

$$\begin{aligned} & \iint \left[ \frac{1}{N} \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} \left( 1 - \frac{j+t}{b} \right) X_j(u) X_{j+t}(v) \right]^2 du dv \\ & \leq \frac{1}{N^2} b^2 \iint \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} X_j^2(u) X_{j+t}^2(v) du dv \\ & = O_P(b^4/N^2) = o_p(1), \end{aligned} \tag{48}$$

where all other terms appearing in  $R(u, v)$  are handled similarly. This completes the proof of the theorem.  $\square$

**Proof of Theorem 3.1.** Consider assertion (i). For  $i = 1, 2$ , let  $\{e_{i,j}^*, j = 1, 2, \dots, n_i\}$  be the pseudo-observations generated by implementing the MBB procedure at  $\{\varepsilon_{i,j}, j = 1, 2, \dots, n_i\}$ . Using Theorem 2.1, it follows that, conditionally on  $\mathbf{X}_M$ , for  $i = 1, 2$ , and as  $n_1, n_2 \rightarrow \infty$ ,

$$\frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} (e_{i,j}^* - \mathbb{E}^*(e_{i,j}^*)) \Rightarrow \Gamma_i, \quad \text{in probability,}$$

where  $\Gamma_i$  is a Gaussian random element with mean zero and covariance operator  $C_i$  with kernel  $c_i(\cdot, \cdot)$ . Now, recall from Step 3 of the MBB-based testing algorithm that, for  $i = 1, 2$ , the pseudo-observations  $\varepsilon_{i,\xi+sb}^*(\tau), \xi = 1, 2, \dots, b, s = 0, 1, \dots, q_i, \tau \in \mathcal{I}$ , are generated by first applying the MBB procedure to  $\hat{\varepsilon}_{i,\xi+sb}(\tau), \xi = 1, 2, \dots, b, s = 0, 1, \dots, q_i, \tau \in \mathcal{I}$  and then  $\bar{\varepsilon}_{i,\xi}(\tau)$  is subtracted. Note further that  $\varepsilon_{i,j}(\tau) = \hat{\varepsilon}_{i,j}(\tau) + \bar{X}_{i,n_i} - \mu_i(\tau)$ . Thus,  $e_{i,\xi+sb}^*(\tau) = \varepsilon_{i,\xi+sb}^*(\tau) + \bar{\varepsilon}_{i,\xi}(\tau) + \bar{X}_{i,n_i}(\tau) - \mu_i(\tau)$  and, using expression (3.2), we get

$$\frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} (e_{i,j}^* - \mathbb{E}^*(e_{i,j}^*)) = \frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} (X_{i,j}^* - \mathbb{E}^*(X_{i,j}^*)) = \frac{1}{\sqrt{n_i}} \sum_{j=1}^{n_i} (X_{i,j}^* - \bar{X}_M).$$

Therefore, and conditionally on  $\mathbf{X}_M$ , as  $n_1, n_2 \rightarrow \infty$ ,

$$\left( \frac{1}{\sqrt{n_1}} \sum_{j=1}^{n_1} (X_{1,j}^* - \bar{X}_M), \frac{1}{\sqrt{n_2}} \sum_{j=1}^{n_2} (X_{2,j}^* - \bar{X}_M) \right) \Rightarrow (\Gamma_1, \Gamma_2), \quad \text{in probability,}$$

where  $\Gamma_1$  and  $\Gamma_2$  are two independent Gaussian random elements with mean zero and covariance operator  $C_1$  and  $C_2$  with kernel  $c_1(\cdot, \cdot)$  and  $c_2(\cdot, \cdot)$ , respectively. Since

$$\sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1,n_1}^* - \bar{X}_{2,n_2}^*) = \sqrt{\frac{n_2}{M}} \frac{1}{\sqrt{n_1}} \sum_{j=1}^{n_1} (X_{1,j}^* - \bar{X}_M) - \sqrt{\frac{n_1}{M}} \frac{1}{\sqrt{n_2}} \sum_{j=1}^{n_2} (X_{2,j}^* - \bar{X}_M),$$

and because  $n_1/M \rightarrow \theta$ , we get that, as  $n_1, n_2 \rightarrow \infty$ ,

$$\sqrt{\frac{n_1 n_2}{M}} (\bar{X}_{1, n_1}^* - \bar{X}_{2, n_2}^*) \Rightarrow \Gamma, \quad \text{in probability,}$$

where  $\Gamma = \sqrt{1-\theta}\Gamma_1 - \sqrt{\theta}\Gamma_2$ . The proof of assertion (ii) follows along the same lines using Theorem 2.2. This completes the proof of the theorem.  $\square$

## Acknowledgement

We would like to thank the two referees for their useful comments and suggestions which have improved the presentation of the paper.

## Supplementary Material

**Supplement to “Moving block and tapered block bootstrap for functional time series with an application to the  $K$ -sample mean problem”** (DOI: [10.3150/18-BEJ1099SUPP](https://doi.org/10.3150/18-BEJ1099SUPP); .pdf). The online supplement [21] contains the proofs of Lemma A.2 and Theorem 2.2 as well as some additional numerical results.

## References

- [1] Benko, M., Härdle, W. and Kneip, A. (2009). Common functional principal components. *Ann. Statist.* **37** 1–34. [MR2488343](#)
- [2] Beutner, E. and Zähle, H. (2016). Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electron. J. Stat.* **10** 1181–1222. [MR3499525](#)
- [3] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. [MR1700749](#)
- [4] Cerovecki, C. and Hörmann, S. (2017). On the CLT for discrete Fourier transforms of functional time series. *J. Multivariate Anal.* **154** 282–295. [MR3588570](#)
- [5] Dehling, H., Sharipov, O.Sh. and Wendler, M. (2015). Bootstrap for dependent Hilbert space-valued random variables with application to von Mises statistics. *J. Multivariate Anal.* **133** 200–215. [MR3282026](#)
- [6] Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869. [MR1055437](#)
- [7] Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. [MR2662361](#)
- [8] Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. *Springer Series in Statistics*. New York: Springer. [MR2920735](#)
- [9] 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](#)
- [10] Horváth, L. and Rice, G. (2015). Testing equality of means when the observations are from functional time series. *J. Time Series Anal.* **36** 84–108. [MR3300207](#)

- [11] Horváth, L., Rice, G. and Whipple, S. (2016). Adaptive bandwidth selection in the long run covariance estimator of functional time series. *Comput. Statist. Data Anal.* **100** 676–693. [MR3505826](#)
- [12] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241. [MR1015147](#)
- [13] Laha, R.G. and Rohatgi, V.K. (1979). *Probability Theory. Wiley Series in Probability and Mathematical Statistics*. New York–Chichester–Brisbane: Wiley. [MR0534143](#)
- [14] Lahiri, S.N. (2003). *Resampling Methods for Dependent Data. Springer Series in Statistics*. New York: Springer. [MR2001447](#)
- [15] Liu, R.Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap (East Lansing, MI, 1990)*. *Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.* 225–248. New York: Wiley. [MR1197787](#)
- [16] Panaretos, V.M. and Tavakoli, S. (2013). Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series. *Stochastic Process. Appl.* **123** 2779–2807. [MR3054545](#)
- [17] Panaretos, V.M. and Tavakoli, S. (2013). Fourier analysis of stationary time series in function space. *Ann. Statist.* **41** 568–603. [MR3099114](#)
- [18] Paparoditis, E. (2018). Sieve bootstrap for functional time series. *Ann. Statist.* **46** 3510–3538. [MR3852660](#)
- [19] Paparoditis, E. and Politis, D.N. (2001). Tapered block bootstrap. *Biometrika* **88** 1105–1119. [MR1872222](#)
- [20] Paparoditis, E. and Sapatinas, T. (2016). Bootstrap-based testing of equality of mean functions or equality of covariance operators for functional data. *Biometrika* **103** 727–733. [MR3551795](#)
- [21] Pilavakis, D., Paparoditis, E. and Sapatinas, T. (2019). Supplement to “Moving block and tapered block bootstrap for functional time series with an application to the  $K$ -sample mean problem.” DOI:10.3150/18-BEJ1099SUPP.
- [22] Politis, D.N. and Romano, J.P. (1994). Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap. *Statist. Sinica* **4** 461–476. [MR1309424](#)
- [23] Prokhorov, Yu.V. (1956). Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatn. Primen.* **1** 177–238. [MR0084896](#)
- [24] Ramsay, J.O. and Dalzell, C.J. (1991). Some tools for functional data analysis. *J. Roy. Statist. Soc. Ser. B* **53** 539–572. With discussion and a reply by the authors. [MR1125714](#)
- [25] Rice, G. and Shang, H.L. (2017). A plug-in bandwidth selection procedure for long-run covariance estimation with stationary functional time series. *J. Time Series Anal.* **38** 591–609. [MR3664648](#)
- [26] Staicu, A.-M., Lahiri, S.N. and Carroll, R.J. (2015). Significance tests for functional data with complex dependence structure. *J. Statist. Plann. Inference* **156** 1–13. [MR3267509](#)
- [27] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics. Springer Series in Statistics*. New York: Springer. [MR1385671](#)
- [28] Zhang, C., Peng, H. and Zhang, J.-T. (2010). Two samples tests for functional data. *Comm. Statist. Theory Methods* **39** 559–578. [MR2745305](#)
- [29] Zhang, J.-T. (2013). *Analysis of Variance for Functional Data*. Boca Raton, FL: CRC Press/CRC.

Received October 2017 and revised July 2018