University of Cyprus

DEPARTMENT OF MATHEMATICS AND STATISTICS

BLOCK BOOTSTRAP METHODS FOR FUNCTIONAL TIME SERIES

DOCTOR OF PHILOSOPHY DISSERTATION

PILAVAKIS DIMITRIOS

2019

DEPARTMENT OF MATHEMATICS AND STATISTICS

BLOCK BOOTSTRAP METHODS FOR FUNCTIONAL TIME SERIES

PILAVAKIS DIMITRIOS

A Dissertation Submitted to the University of Cyprus in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

September 2019

VALIDATION PAGE

Doctoral Candidate: Pilavakis Dimitrios

Doctoral Thesis Title: Block Bootstrap Methods for Functional Time Series

The present Doctoral Dissertation was submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy at the ***Department of Mathematics and Statistics*** *and was approved on the 30th of September 2019 by the members of the* ***Examination Committee.***

Examination Committee:

Research Co-Supervisor: Paparoditis Efstathios, Professor,
Department of Mathematics and Statistics,
University of Cyprus

Research Co-Supervisor: Sapatinas Theofanis, Professor,
Department of Mathematics and Statistics,
University of Cyprus

Committee Member: Agapiou Sergios, Assistant Professor,
Department of Mathematics and Statistics,
University of Cyprus

Committee Member: Jentsch Carsten, Professor,
Department of Statistics,
Technische Universität Dortmund

Committee Member: Kreiss Jens-Peter, Professor,
Department of Mathematics,
Technische Universität Braunschweig

DECLARATION OF DOCTORAL CANDIDATE

The present doctoral dissertation was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy of the University of Cyprus. It is a product of original work of my own, unless otherwise mentioned through references, notes, or any other statements.

Pilavakis Dimitrios

Abstract

Functional data analysis is a fast-growing research area in statistics that deals with statistical analysis of infinite-dimensional (functional) data. It is therefore important to extend the theory and methods used for finite-dimensional to the setting of infinite-dimensional data. This dissertation applies bootstrap methods to functional data that are assumed to be weakly dependent in a broad sense and it is based on two main pillars.

The first pillar of this dissertation addresses the consistency of the moving block and of the tapered block bootstrap applied to functional time series. More precisely, central limit theorems for the moving block bootstrap and for the tapered block bootstrap for the sample mean are proved. It is also shown, that these block resampling procedures provide consistent estimators of the covariance operator of the sample mean function and therefore of the spectral density operator of the underlying functional process at frequency zero. A central limit theorem for the moving block bootstrap applied to the lag h sample covariance operator is also proved.

The second pillar deals with the application of bootstrap-based methodologies for testing hypotheses about the equality of certain characteristics of the distributions between several independent, populations in functional time series context. More precisely, algorithms based on both the moving block and the tapered block bootstrap procedures for the important problem of testing the equality of the mean functions of several populations are developed. A moving block bootstrap based procedure for testing the equality of the covariance operator between several independent functional populations is also considered. The basic idea behind these testing methodologies is to bootstrap the observed functional time series in such a way that the obtained functional pseudo-observations satisfy the null hypothesis of interest. Therefore, the suggested block bootstrap-based testing methodologies are applicable to a broad range of possible test statistics.

Validity of the proposed bootstrap methods in approximating the distribution of some fully functional test statistics under the null is established. In addition, the finite sample behaviour of the bootstrap procedures proposed is investigated by means of simulations. Simulations are also conducted to gauge the size and power properties of the suggested block bootstrap-based testing methodologies. Applications to real-life data sets are also examined.

Περιληψη

Η ανάλυση συναρτησιακών δεδομένων είναι ένας ταχέως αναπτυσσόμενος τομέας έρευνας της στατιστικής που ασχολείται με τη στατιστική ανάλυση απειροδιάστατων (συναρτησιακών) δεδομένων. Επομένως, είναι σημαντικό να επεκταθεί η θεωρία και οι μέθοδοι που χρησιμοποιούνται για τη στατιστική ανάλυση πεπερασμένης διάστασης δεδομένων και στην περίπτωση των δεδομένων άπειρης διάστασης. Η διατριβή αυτή εφαρμόζει τις μεθόδους bootstrap σε συναρτησιακά δεδομένα τα οποία υποθέτουμε ότι είναι ασθενώς εξαρτημένα με μια ευρεία έννοια και βασίζεται σε δύο κύριους πυλώνες.

Ο πρώτος πυλώνας αυτής της διατριβής ασχολείται με τη συνέπεια του moving block και του tapered block bootstrap όταν οι μέθοδοι αυτοί εφαρμόζονται σε συναρτησιακές χρονοσειρές. Πιο συγκεκριμένα, στην διατριβή αυτή, αποδεικνύονται κεντρικά οριακά θεωρήματα για το moving block και το tapered block bootstrap που αφορούν τον δειγματικό μέσο όρο. Επίσης, αποδεικνύεται ότι αυτές οι διαδικασίες αναδειγματοληψίας παρέχουν συνεπείς εκτιμήτριες του τελεστή συνδιακύμανσης της μέσης συνάρτησης του δείγματος, άρα και της φασματικής πυκνότητας της υπό εξέταση συναρτησιακής διαδικασίας, σε μηδενική συχνότητα. Επίσης, αποδεικνύεται ένα κεντρικό οριακό θεώρημα για το moving block bootstrap που αφορά τον δειγματικό τελεστή συνδιακύμανσης σε h χρονικές υστερήσεις.

Ο δεύτερος πυλώνας ασχολείται με την εφαρμογή μεθοδολογιών, που βασίζονται σε μεθόδους bootstrap, για τον έλεγχο υποθέσεων σχετικά με την ισότητα ορισμένων χαρακτηριστικών των κατανομών ανεξάρτητων πληθυσμών, χρησιμοποιώντας συναρτησιακές χρονοσειρές. Συγκεκριμένα, προτείνονται αλγόριθμοι που βασίζονται τόσο στο moving block όσο και στο tapered block bootstrap για το σημαντικό πρόβλημα του στατιστικού ελέγχου της ισότητας των μέσων συναρτήσεων πολλών ανεξάρτητων συναρτησιακών πληθυσμών. Προτείνεται, επίσης, μια διαδικασία που βασίζεται στο moving block bootstrap για τον έλεγχο της ισότητας του τελεστή συνδιακύμανσης μεταξύ διαφόρων ανεξάρτητων συναρτησιακών πληθυσμών. Η βασική ιδέα των προτεινόμενων διαδικασιών για τον στατι-

στικό έλεγχο υποθέσεων, είναι οι τεχνικές αναδειγματοληψίας να εφαρμοστούν με τέτοιο τρόπο, έτσι ώστε οι δημιουργηθείσες ψευδο-παρατηρήσεις να ικανοποιούν τη μηδενική υπόθεση του έλεγχου. Αυτό έχει ως αποτέλεσμα οι προτεινόμενοι αλγόριθμοι που βασίζονται στις μεθόδους block bootstrap να είναι εφαρμόσιμες σε ένα ευρύ φάσμα πιθανών ελεγχοσυναρτήσεων.

Όσον αφορά τις προτεινόμενες μεθόδους, αποδεικνύεται η εγκυρότητα τους προσεγγίζοντας την κατανομή μερικών ελεγχοσυναρτήσεων κάτω από τη μηδένικη υπόθεση. Επιπρόσθετα, διερευνάται μέσω προσομοιώσεων, η αποτελεσματικότητα των προτεινόμενων διαδικασιών, όταν εφαρμοστούν σε ένα πεπερασμένο δείγμα. Προσομοιώσεις διεξάγονται επίσης για να ελεγχθεί το επίπεδο σημαντικότητας και η ισχύς των προτεινόμενων μεθοδολογιών στατιστικού ελέγχου οι οποίες βασίζονται σε bootstrap διαδικασίες. Τέλος, εξετάζεται η αποτελεσματικότητα των προτεινόμενων διαδικασιών στατιστικού έλεγχου με την εφαρμογή τους σε πραγματικά δεδομένα.

Acknowledgments

Undertaking this PhD has been a truly unique and rewarding experience for me.

I would like to thank my supervisors Dr Paparoditis and Dr Sapatinas for their support and guidance throughout this journey. Without their supervision and constant feedback this PhD would not have been achievable. I am also thankful for the excellent example they provided as successful professors and mathematicians.

I would also like to thank the Committee members who gave detailed insightful comments and constructive feedback about my PhD research work.

I am grateful to my family and my parents who raised me with a love for science and mathematics and supported me in all my pursuits.

Last but not least, I would like to thank my wife Eleni, for her unconditional love, encouragement, enthusiasm and faithful support throughout my PhD studies.

To my wife, Eleni

Contents

List of Figures

List of Tables

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, we study phenomena that are continuous in time or space and therefore observations can be considered as curves or images. These observations 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' and the statistical methods used for analysing this kind of data are called 'functional data analysis' (FDA). FDA dealing with independent and identically distributed (i.i.d.) random variables has received considerable attention in the statistical literature during the last decades. However the i.i.d. assumption suffers in many cases, especially when the data are obtained sequentially over time, where there is a natural dependency in the functional sample. Such temporally dependent functional data are called 'functional time series'.

In functional time series analysis the aim is to infer properties of the functional stochastic process based on an observed stretch $X_1, X_2, \ldots, X_n$, i.e., on a functional time series. In this context, usually the distribution or parameters related to the distribution of some statistics of interest based on $X_1, X_2, \ldots, X_n$ needs to be estimated. Since in a functional set-up such quantities typically depend in a complicated way on difficult to estimate infinite-dimensional characteristics of the underlying stochastic process $\mathbb{X}$, their calculation is difficult in practice. As a result, in functional time series analysis, resampling methods and, in particular, bootstrap methodologies are very useful.

This work falls into this sub-field of functional data analysis that is, we focus on functional time series, stemming for a stochastic process $\mathbb{X} = (X_t,\, t \in \mathbb{Z})$ of Hilbert space-valued random variables which is L^p-m-approximable, a dependence assumption which is satisfied by large classes of commonly used functional time series models; see, e.g., Hörmann and Kokoszka (2010). This dissertation contributes to the current state of the art by establishing the validity of block bootstrap procedures in the functional time series context.

1.1 Thesis objectives

The contribution of this dissertation to the FDA is twofold. First, we prove consistency of the moving block bootstrap (MBB) and of the tapered block bootstrap (TBB) for the sample mean function in the case of weakly dependent (L^p-m-approximable), Hilbert space-valued random variables. Furthermore, we show that these bootstrap methods provide consistent estimators of the covariance operator of the mean function estimator, that is of the spectral density operator of the underlying functional stochastic process at frequency zero. We also prove a CLT for the MBB applied to approximate the distribution of the sample covariance operator. Second, we propose general bootstrap-based testing procedures for the important problem of comparing the mean functions or the covariance operators between several populations and which are applicable to a wide range of test statistics of interest. The basic idea of the suggested procedures is to generate the functional pseudo observations in such a way that the null hypothesis of interest is satisfied. For each hypothesis testing theoretical justification for approximating the null distribution of certain fully functional test statistics are given. Furthermore, simulations are carried out for each case to investigate the finite sample performance of the proposed algorithms.

1.2 Dissertation outline

The objectives previously described are unfold in this thesis in the form of five chapters. After this introductory chapter, Chapter 2 provides some concepts, tools and notations which are central in Functional Data Analysis. The chapter begins with a brief introduction to the data that motivate this research and its representation using basis functions. Then, an introduction to the theory of operators in Hilbert spaces is given. After that, by focusing to the Hilbert space of square integrable functions, some notations and properties of operators, is given. Later on, this chapter focuses on the properties of random samples in the space of square integrable functions and it gives the extension of the summary statistics to the functional framework. Moreover, the L^p-m-approximability, which is the weak dependence structure of the stochastic process considered in this dissertation, is presented. Finally, some basic definitions and results for the asymptotic behavior of the stochastic process considered are given.

In Chapter 3 a central limit theorem for the moving block bootstrap and for the tapered block bootstrap is proved. Also it is shown that these block bootstrap procedures provide consistent estimators of the spectral density operator of the underlying stochastic process at frequency zero. We conclude the chapter by addressing the important problem of comparing the mean functions between independent k-populations. Block bootstrap based procedures for testing the equality of mean functions between several independent functional time series are proposed. For these algorithms the generated pseudo observations satisfy the null hypothesis of interest therefore, the suggested methods can be applied to a broad range of test statistics of interest. Theoretical results that justify the validity of the suggested bootstrap-based procedures applied to test statistics considered in the literature are established. In Section 3.4, the finite sample performance of the MBB, of the TBB and of the stationary bootstrap (SB) is investigated by estimating the standard deviation function of the normalized sample mean function. Then, simulations are carried out to examine the finite sample size and power performance of the suggested tapered block bootstrap-based testing procedures. An application to a real-life data set is also discussed. Finally, auxiliary results and proofs of the main results are presented concluding this chapter.

Chapter 4 is devoted to the important problem of testing the equality of the lag-zero autocovariance operators of several independent functional time series. Firstly, the asymptotic validity of the MBB procedure applied to estimate the distribution of the lag-h sample autocovariance operator, for any (fixed) lag h, $h \in \mathbb{Z}$ is established. Then, a moving block bootstrap algorithm is proposed for testing the hypothesis of interest, which is based on bootstrapping the time series of tensor products, and generates pseudo random elements that satisfy the null hypothesis of interest. The finite sample size and power performance of the suggested moving block bootstrap-based testing procedure is illustrated through simulations and an application to a real-life data set is discussed. This chapter ends with some auxiliary results and the presentation of proofs of the main result obtained.

Finally, Chapter 5 provides some concluding remarks summarizing the contributions of this thesis and discussing some future developments.

2

Notation and Setup

Statistics is concerned with the analysis of data obtained from observations of random variables. The data that motivate this dissertation are observed in the form of curves, i.e. each observation is a real-valued function of the form $X_t(\tau)$, $\tau \in [a, b]$. More precisely 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 dependence properties. We suppose that 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 (Ω, A, P) and take values in the separable Hilbert-space of squared-integrable $\mathbb{R}$-valued functions on $\mathcal{I}$, denoted by $L^2(\mathcal{I})$. In this section, the notation used in the dissertation and the necessary background for supporting the main contributions of this thesis are introduced. More precisely, in Section 2.1 a brief introduction to Functional Time Series is given. Section 2.2 introduces some fundamental concepts of the theory of operators. Section 2.3 focuses on the space $L^2(\mathcal{I})$ of square integrable functions and describes some fundamental concepts. In Section 2.4 some basic properties of random samples in the space $L^2(\mathcal{I})$ and some useful results are given. The notion of weak dependence used in this dissertation is presented in Section 2.5. In Section 2.6, some basic definitions and results regarding the asymptotic theory of infinite dimensional spaces are given. We conclude this section, by giving some basic results regarding asymptotic behavior of the sample mean function and covariance operator of the functional time series considered in this dissertation.

2.1 Functional Time Series

Statistics is a branch of mathematics dealing with the collection, organization, analysis, interpretation, and presentation of observations taken on a sample with the aim of making inferences about the general population from which the sample is drawn. These data could appear in various forms. For instance, consider a data set obtained by recording the electricity consumption in a 15 minutes interval for 10 days, i.e., we have 96 electricity consumption measurements for each day. We may assume that the data set consists of 960 observations where x_t represents the t-th observation and it is a scalar quantity. Alternatively, x_t could be a vector of length 96 representing the observations corresponding to the t-th day. An alternative approach is to represent observations of electricity consumption as functions, i.e., $X_t(u)$ is a function representing the electricity consumption of the entire t-th day. For the latter case, to convert the discrete trajectories into functions we interpolate the data using a basis of $L^2(\mathcal{I})$, where $\mathcal{I}$ represents a 24-hour interval. For this conversion, by letting $X_t(u_j)$ be the j-th measurement of electricity consumption on day t, each vector $(X_t(u_1), X_t(u_2), \ldots, X_t(u_{96}))$ is approximated by an expansion of the form

$$X_t(u_j) \approx \sum_{k=1}^{K} c_{t,k}\phi_k(u_j)$$

where ϕ_k, $k = 1, 2, \ldots$ are basis functions in $L^2(\mathcal{I})$, for example the Fourier basis or the B-spline basis functions. Since, the basis functions are defined on the whole 24-hour interval we might express the functional data $X_t(u)$ as

$$X_t(u) \approx \sum_{k=1}^{K} c_{t,k}\phi_k(u).$$

In such cases the obtained set of functions $\{X_t,\, t = 1, 2, \ldots, n\}$ is called a functional time series.

Figure 2.1 illustrates the functional data approach stated above where the dotted vertical lines separate days. To convert the discrete data in functional form the Fourier basis with 49 basis functions has been used. As it can be seen, the curves $X_t(u)$ are obtained by splitting a continuous time record into daily curves.

Figure 2.2 demonstrates the dependence between the random elements X_t and X_s

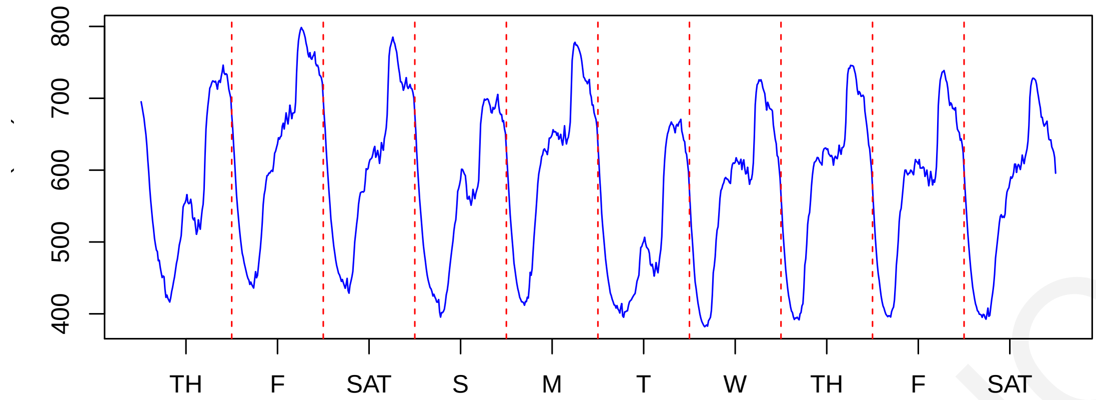

Figure 2.1: Electricity consumption in Cyprus from 1/1/2009 00:00 to 10/1/2009 23:45, recorded every fifteen minutes. The vertical lines separate days and t-th day's graph represents the observation $x_t(u)$.

for $s \neq t$ by showing the estimation of the correlation coefficient between the random variables $X_t(u)$ and $X_{t+h}(u)$ for different values of u and lag $h = 1, 2$. As evident the curves $X_t(u)$ are dependent.

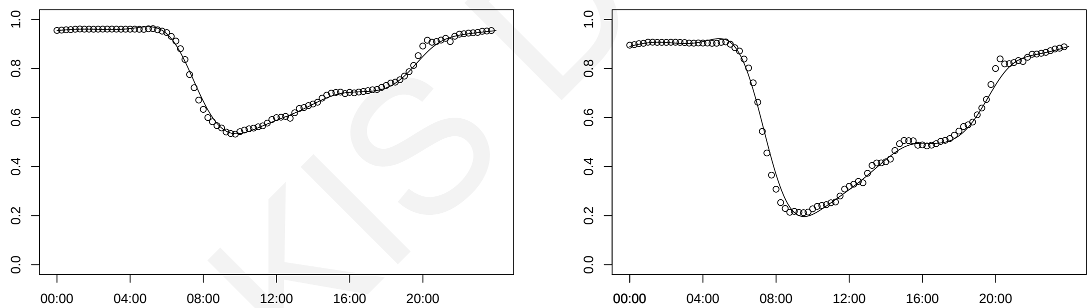

Figure 2.2: Estimated correlation values between the curves $X_t(u_j)$ and $X_{t+h}(u_j)$ for $j = 1, 2, \dots, 96$ for lag $h = 1$ (left) and $h = 2$ (right)

In the above example, each function was created from the same number of observations which were equally spaced. However, functional data can also arise in other cases. For example, when measurements on human subjects are made, it is often difficult to ensure that they are made at the same time in the life of the subject and there may be different numbers of measurements for different subjects. A typical example are height curves i.e. $X_t(u)$ is the height of subject t at time u after birth. (see, e.g., Tuddenham and Snyder, (1954).)

2.2 Compact Operators in Hilbert spaces

Operators are the basic mathematical tool to deal with functional data. In this section we will focus on compact operators and a brief description of their main theoretical properties is given.

We consider a separable Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ which generates the norm $\| \cdot \|$. A continuous and bounded linear operator $\Psi : \mathcal{H} \to \mathcal{H}$ is said to be compact if there exist two orthonormal bases $\{e_i, i = 1, 2, \ldots\}$ and $\{\epsilon_i, i = 1, 2, \ldots\}$ of $\mathcal{H}$ and a real sequence $\{\lambda_i, i = 1, 2, \ldots\}$ converging to zero as $i \to \infty$, such that

$$\Psi(x) = \sum_{i=1}^{\infty} \lambda_i \langle x, \epsilon_i \rangle e_i, \; x \in \mathcal{H}.$$

Note that, the λ_i may be assumed positive because one can replace e_i by $-e_i$. The above representation is called the singular value decomposition of Ψ. An operator having the above singular value decomposition is said to be a Hilbert-Schmidt operator if $\sum_{i=1}^{\infty} \lambda_i^2 < \infty$. We denote by $\mathcal{L}$ the space of Hilbert-Schmidt operators. For two Hilbert-Schmidt operators Ψ_1 and Ψ_2, we denote by

$$\langle \Psi_1, \Psi_2 \rangle_{HS} = \sum_{i=1}^{\infty} \langle \Psi_1(e_i), \Psi_2(e_i) \rangle$$

the inner product which generates the Hilbert-Schmidt norm $\|\Psi_1\|_{HS}^2 = \sum_{i=1}^{\infty} \|\Psi_1(e_i)\|^2$. Here $\{e_i, i = 1, 2, \ldots\}$ is an arbitrary orthonormal basis of $\mathcal{H}$. Note that the value of $\langle \Psi_1, \Psi_2 \rangle_{HS}$ is independent of the choice of the basis and that $\|\Psi\|_{HS}^2 = \sum_{i=1}^{\infty} \lambda_i^2$. We also define the tensor product $\Psi_1 \otimes \Psi_2 : \mathcal{L} \to \mathcal{L}$ between the operators Ψ_1 and Ψ_2 by $\Psi_1 \otimes \Psi_2(\cdot) = \langle \Psi_1, \cdot \rangle_{HS} \Psi_2$. Note that $\Psi_1 \otimes \Psi_2$ is an operator acting on the space of Hilbert-Schmidt operators.

Another important family of operators is the trace-class operators. A compact operator Ψ is said to be nuclear or trace-class if $\sum_{i=1}^{\infty} \lambda_i < \infty$. In this case, the trace of Ψ, is given by

$$tr(\Psi) = \sum_{i=1}^{\infty} \langle \Psi(e_i), e_i \rangle$$

where the sum converges absolutely and is independent of the choice of the orthonormal basis. It can be shown that if Ψ is trace-class $tr(\Psi) = \sum_{i=1}^{\infty} \lambda_i$.

2.3 THE HILBERT SPACE $L^2(\mathcal{I})$

In the following, we focus on the separable Hilbert space $L^2(\mathcal{I})$, that is, the set of all measurable real–valued functions f defined on $\mathcal{I}$ satisfying $\int_{\mathcal{I}} f^2(u)\,\mathrm{d}u < \infty$. The space $L^2(\mathcal{I})$ is a separable Hilbert space with the inner product

$$\langle f, g\rangle = \int_{\mathcal{I}} f(u)g(u)\,\mathrm{d}u$$

which generates the norm $\|f\|^2 = \langle f, f\rangle$. Notice that if $f, g \in L^2(\mathcal{I})$ the equality $f = g$ means $\|f - g\| = 0$ whereas $f \neq g$ that $\|f - g\| > 0$.

Let $\{e_i, i = 1, 2, \ldots\}$ be an orthonormal basis of $L^2(\mathcal{I})$. Then every $f \in L^2(\mathcal{I})$ can be written as

$$f = \sum_{i=1}^{\infty} \langle f, e_i\rangle e_i.$$

Therefore,

$$\langle f, g\rangle = \sum_{i=1}^{\infty} \langle f, e_i\rangle\langle g, e_i\rangle$$

and Parseval's equality

$$\|f\|^2 = \sum_{i=1}^{\infty} \langle f, e_i\rangle^2$$

follow. We define the tensor product $f \otimes g : L^2(\mathcal{I}) \to L^2(\mathcal{I})$ between f and g by $f \otimes g(\cdot) = \langle f, \cdot\rangle g$.

From now on, and without loss of generality, we assume that interval $\mathcal{I}$ is normalized to be a unit interval, i.e, $\mathcal{I} = [0, 1]$ and, for simplicity, integral signs without the limits of integration imply integration over the interval $\mathcal{I}$. We finally write L^2 instead of $L^2(\mathcal{I})$.

An important class of operators in L^2 are the integral operators defined by:

$$\Psi(x(u)) = \int \psi(u, v)x(v)\,\mathrm{d}v$$

where $\psi(u, v)$ is called the kernel of the operator Ψ. Such operators are Hilbert-Schmidt if and only if $\iint \psi^2(u, v)\,\mathrm{d}u\mathrm{d}v < \infty$. The integral operator Ψ is said to be symmetric if $\psi(u, v) = \psi(v, u)$ and positive-definite if for all square-integrable functions $f(u)$, $\iint \psi(u, v)f(u)f(v)\,\mathrm{d}u\mathrm{d}v \geq 0$. In this case $\psi(u, v)$ has the representation

$$\psi(u, v) = \sum_{i=1}^{\infty} \lambda_i e_i(u)e_i(v) \tag{2.3.1}$$

where $\{e_i, i = 1, 2, \ldots\}$ is an orthonormal basis of L^2 consisting of eigenfunctions of Ψ such that the corresponding sequence of eigenvalues $\{\lambda_i, i = 1, 2, \ldots\}$ is nonnegative, i.e., $\Psi(e_i) = \lambda_i e_i$. Representation 2.3.1 is known as Mercer's theorem and if the kernel ψ is continuous the convergence is absolute and with respect to the L^2 norm. From Mercer's theorem it follows directly that $\int \psi(u, u)\,\mathrm{d}u = \sum_{i=1}^{\infty} \lambda_i$.
If Ψ_1 and Ψ_2 are Hilbert-Schmidt integral operator with kernels $\psi_1(u, v)$ and $\psi_2(u, v)$, respectively, then $\langle \Psi_1, \Psi_2 \rangle_{HS} = \iint \psi_1(u, v)\psi_2(u, v)\,\mathrm{d}u\mathrm{d}v$ and $\|\Psi_1\|_{HS}^2 = \iint \psi_1^2(u, v)\,\mathrm{d}u\mathrm{d}v$.

2.4 RANDOM VARIABLE IN $L^2(\mathcal{I})$

The stationarity of a stochastic process $\mathbb{X} = \{X_t,\, t \in \mathbb{Z}\}$ is an indispensable property in the functional time series analysis.

Definition 2.4.1. A stochastic process $\mathbb{X} = \{X_t,\, t \in \mathbb{Z}\}$, is called strictly stationary if, for all finite sets of indices $Q \subseteq \mathbb{Z}$, the joint distribution of $(X_{t+q},\, q \in \mathbb{Q})$, does not depend on $t \in \mathbb{Z}$.

The samples $X_1, X_2, \ldots, X_n$ of curves that we consider in this dissertation and introduced in Section 2.1 are viewed as the outcomes of a strictly stationary stochastic process $\mathbb{X} = \{X_t,\, t \in \mathbb{Z}\}$, where the random variables X_t take values in L^2 and are dependent, in a broad sense which is made precise in Section 2.5.

If X_t is integrable, i.e., $\mathbb{E}\|X_t\| = \mathbb{E}\int X_t^2(u)\,\mathrm{d}u < \infty$, there is a unique function $\mu \in L^2(\mathcal{I})$ such that $\mathbb{E}\langle X, y\rangle = \langle \mu, y\rangle$, for all $y \in L^2(\mathcal{I})$. The function μ is called the expectation function of X_t, $\mathbb{E}X_t \in L^2(\mathcal{I})$ and is independent of t by the stationarity of $\mathbb{X}$. If in addition $\mathbb{E}\|X_t\|^2 < \infty$, then the covariance operator of $\mathbb{X}$ at lag $h \in \mathbb{Z}$ exists and is defined by

$$\mathcal{C}_h = \mathbb{E}(X_t - \mu) \otimes (X_{t+h} - \mu)$$

which is independent of t by the stationarity of $\mathbb{X}$. By the definition of the covariance operator $\mathcal{C}_h$, it follows that $\mathcal{C}_h$ is an integral operator with real valued kernel

$$c_h(u, v) = \mathbb{E}[(X_t(u) - \mu(u))(X_{t+h}(v) - \mu(v))].$$

Therefore,

$$\|\mathcal{C}_h\|_{HS}^2 = \iint c_h^2(u, v)\,\mathrm{d}u\mathrm{d}v.$$

As it can be seen, each value of $c_h(u,v)$ measures the joint variability of the functional variables X_t and X_{t+h} at points u and v respectively. If

$$\sum_{h\in\mathbb{Z}} \|\mathcal{C}_h\|_{HS} < \infty, \tag{2.4.1}$$

the series $\sum_{h\in\mathbb{Z}} c_h(u,v)e^{ih\omega}$, $\omega \in [-\pi,\pi]$ where i denotes the imaginary unit, converges, and the operator $\mathcal{F}_\omega$ whose kernel is

$$f_\omega(u,v) = (2\pi)^{-1} \sum_{h\in\mathbb{Z}} c_h(u,v)e^{-ih\omega}$$

is called the spectral density operator of $\mathbb{X}$ at frequency ω and is defined by

$$\mathcal{F}_\omega = (2\pi)^{-1} \sum_{h\in\mathbb{Z}} \mathcal{C}_h e^{-ih\omega},$$

see Panaretos and Tavakoli (2013). We then have $\mathcal{C}_h = \int_{-\pi}^{\pi} \mathcal{F}_\omega e^{ih\omega}\, \mathrm{d}\omega$.

Having an observed stretch $X_1, X_2, \ldots, X_n$, the mean function μ is estimated by the sample mean, $\overline{X}_n$, which is given by

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

The operator $\mathcal{C}_h$ is commonly estimated by the corresponding sample autocovariance operator, which is given by

$$\widehat{\mathcal{C}}_h = \begin{cases} n^{-1}\sum_{t=1}^{n-h}(X_t - \overline{X}_n)\otimes(X_{t+h}-\overline{X}_n), & \text{if } 0 \le h < n, \\ n^{-1}\sum_{t=1}^{n+h}(X_{t-h} - \overline{X}_n)\otimes(X_t-\overline{X}_n), & \text{if } -n < h < 0, \\ 0, & \text{otherwise,} \end{cases}$$

and the covariance kernel $c_h(u,v)$ estimated by its sample counterpart

$$\widehat{c}_h(u,v) = \begin{cases} n^{-1}\sum_{t=1}^{n-h}(X_t(u) - \overline{X}_n(u))(X_{t+h}(v)-\overline{X}_n(v)), & \text{if } 0 \le h < n, \\ n^{-1}\sum_{t=1}^{n+h}(X_{t-h}(u) - \overline{X}_n(u))(X_t(v)-\overline{X}_n(v)), & \text{if } -n < h < 0, \\ 0, & \text{otherwise.} \end{cases}$$

We conclude this section, by illustrating the notions introduced above using a real

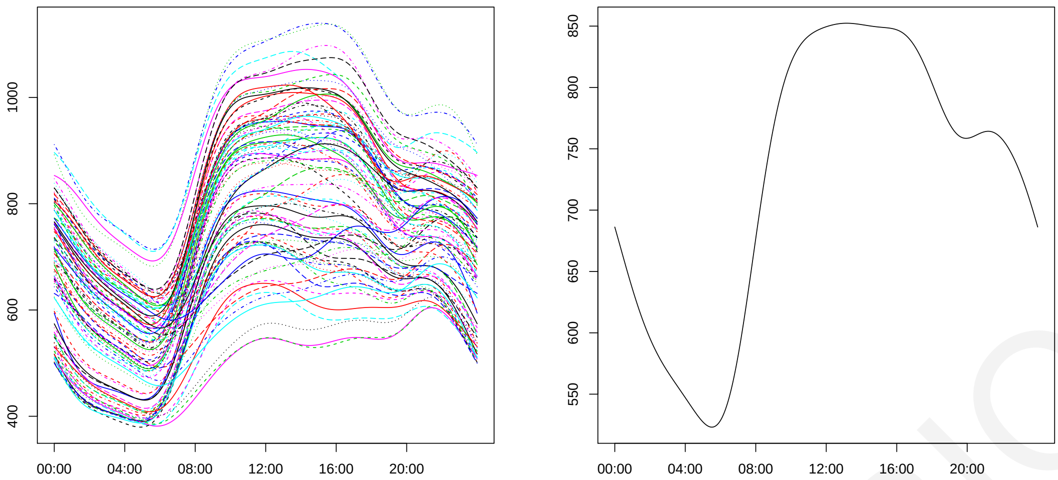

Figure 2.3: Electricity consumption curves (left) and their mean function estimation (right)

life example. The data set considered consists of 15-minutes measurements of the electricity consumption in Cyprus in Summer 2010, i.e., from 1 June 2010 through 31 August 2010. We use the $\mathbb{R}$ software with 49 Fourier basis functions to transform the raw discrete data to functional data as explained in Section 2.1. The resulting functional time series and the estimation of the mean function are displayed in Figure 2.3. Whereas Figure 2.4 illustrates the estimation of the covariance kernel at lag zero.

2.5 L^p-m-APPROXIMABLE

For the purpose of this dissertation, and in order to describe the dependence structure of the stochastic process $\mathbb{X}$, we use the notion of L^p-m-approximability; see Hörmann and Kokoszka (2010). A stochastic process $\mathbb{X} = \{X_t, t \in \mathbb{Z}\}$ with X_t taking values in L^2, is called L^p-m-approximable if the following conditions are satisfied:

(i) X_t admits the representation

$$X_t = f(\delta_t, \delta_{t-1}, \delta_{t-2}, \dots) \tag{2.5.1}$$

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

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

$$\sum_{m\geq 1} (\mathbb{E}\|X_t - X_{t,m}\|^p)^{1/p} < \infty, \tag{2.5.2}$$

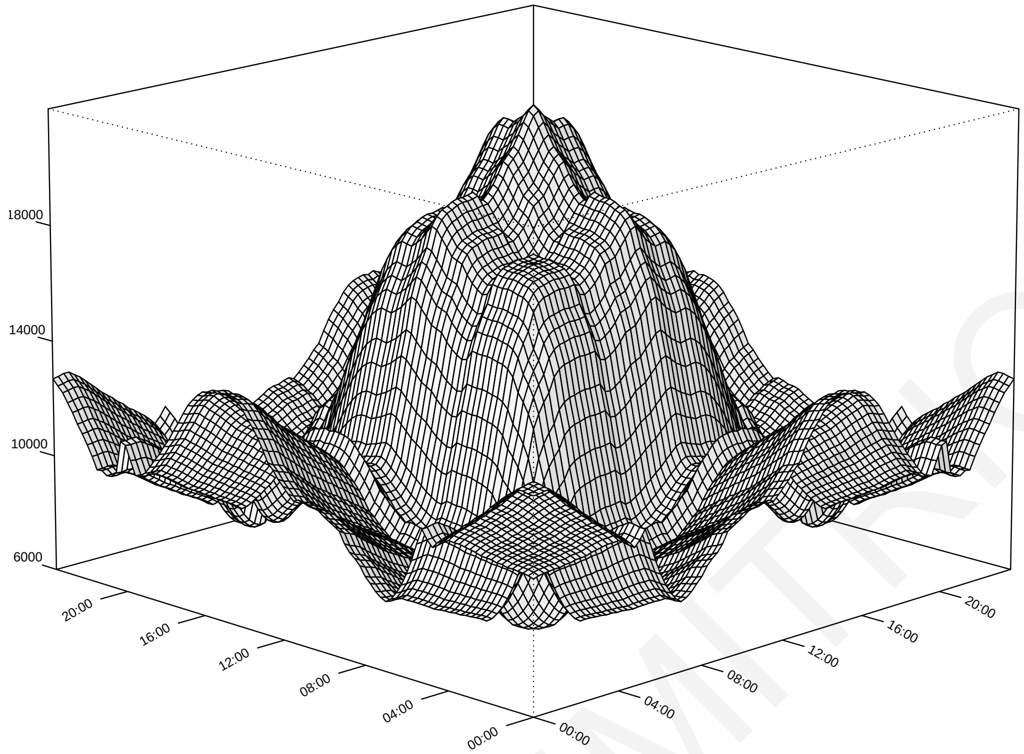

Figure 2.4: Estimated covariance kernel at lag zero of the electricity consumption data

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

The intuition behind the above definition is that the function f in (2.5.1) should be such that the effect of the innovations δ_i far back in the past becomes negligible, that is, these innovations can be replaced by other, independent, innovations.
By (2.5.1) the stochastic process $\mathbb{X} = (X_t; t \in \mathbb{Z})$ is strictly stationary. L^p-m-approximability implies that for each $m \geq 1$, the sequences $(X_{t,m}; t \in \mathbb{Z})$ are strictly stationary and m-dependent, and $X_{t,m}$ and X_t have the same distribution. Furthermore, from the above definition it is easily seen that $\mathbb{E}\|X_{t,m} - X_t\|^p = \mathbb{E}\|X_{0,m} - X_0\|^p$ and $\mathbb{E}\|X_{t,m}\|^p = \mathbb{E}\|X_t\|^p = \mathbb{E}\|X_0\|^p$ for $p \in \mathbb{N}$ and for all $t \in \mathbb{Z}$.

Kokoszka and Reimherr (2013) proved, that, L^4-m-approximability of $\mathbb{X}$ implies that the tensor product $\{X_t \otimes X_t,\ t \in \mathbb{Z}\}$ is L^2-m-approximable with $X_{t,m} \otimes X_{t,m}$ be the m-dependent approximation of $X_t \otimes X_t$.

Furthermore, Hörmann *et al.* (2015) proved that if $\mathbb{X}$ is L^2-m-approximable then (2.4.1) holds and the spectral density operator $\mathcal{F}_\omega$, $\omega \in [-\pi, \pi]$ is trace-class.

For the stochastic process $\mathbb{X}$ considered in this dissertation, we somehow strengthen condition (2.5.2) to the following condition.

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

$$\lim_{m\to\infty} m\left(\mathbb{E}\|X_t - X_{t,m}\|^p\right)^{1/p} = 0.$$

2.6 Central Limit Theorem and Related Results

When dealing with random variables, a pivotal concept in asymptotic derivation are Gaussian processes. In the finite dimensional case, a Gaussian distribution describes the distribution of random variables which are scalars or vectors (for multivariate distributions) and is fully defined by its mean value or its mean vector and its covariance value or its covariance matrix. Whereas, in the infinite dimensional case, a Gaussian process, defines a distribution over infinite dimensional variables, e.g., functions or operators, and it is fully specified by a mean and a covariance function.

Definition 2.6.1. An $\mathcal{H}$-valued random element $\mathcal{Z}$ is Gaussian on $\mathcal{H}$ if for all $h \in \mathcal{H}$ the real random variable $\langle \mathcal{Z}, h\rangle$ has a Gaussian distribution on $\mathbb{R}$.

One of the central topics in asymptotic theory is that of the weak convergence. To define this convergence in L^2 we will need the following definitions

Definition 2.6.2. Suppose $(X_n; n \in \mathbb{N})$ and X are random elements in L^2 with distributions P_{X_n} and P_X respectively. We say that P_{X_n} converges weakly to P_X if

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \quad \text{as } n \to \infty$$

for every bounded and continuous real function f on L^2.

Definition 2.6.3. Suppose $(X_n; n \in \mathbb{N})$ and X are random elements in L^2. We say that (X_n) converges in distribution to X, if the distribution P_{X_n} of X_n, converges weakly to the distribution P_X of X as $n \to \infty$. We denote this by $X_n \Rightarrow X$.

In infinite dimensional spaces, to prove weak convergence of a random sequence, a useful property is that of tightness.

Definition 2.6.4. A sequence of random variables $(X_n; n \in \mathbb{N})$ in L^2 is said to be tight if for every $\varepsilon > 0$ there exist a compact set $K_\varepsilon \subset L^2$ such that for all $n \in \mathbb{N}$:

$$P(X_n \in K_\varepsilon) > 1 - \varepsilon.$$

The following result, which is stated as Proposition 7.4.2 of Laha and Rohatgi (1979), gives a convenient criterion for weak convergence and tell us why tightness is an important property.

Theorem 2.6.1. *Let $(X_n; n \in \mathbb{N})$ be a stochastic process and X be a random element of L^2. Then $X_n \Rightarrow X$ as $n \to \infty$ if and only if*

(a) $\langle X_n, y \rangle \Rightarrow \langle X, y \rangle \qquad \forall\, y \in L^2$,

(b) the sequence $(X_n; n \in \mathbb{N})$ is tight.

Note that Condition (a) of the above theorem, is the weak convergence of real-valued random variables and can be proved by applying an appropriate central limit theorem. Concerning Condition (b) the following results give some sufficient conditions for tightness.

Lemma 2.6.1. *Let $\{W_{n,t},\ 1 \le t \le n, 1 \le n < \infty\}$ be a double array of random elements of a Hilbert space $\mathcal{H}$, strictly stationary for each n and with $\mathbb{E}W_{n,t} = 0$ and $\mathbb{E}\|W_{n,t}\|^2 < \infty$. If*

$$\lim_{n\to\infty} \mathbb{E} \left\| \sum_{t=1}^{n} W_{n,t} \right\|^2 \quad \text{exists and is finite}$$

then the sequence $\{W_{n,t}, 1 \le t \le n\}$ is tight.

The above lemma, is given in Remark 3.3 of Chen and White (1998). Another useful result for proving tightness, which is derived from Theorem 1.13 of Prokhorov (1956) is the following.

Theorem 2.6.2. *A zero mean sequence $\{W_{n,t},\ 1 \le t \le n, 1 \le n < \infty\}$ of square integrable elements on L^2 is tight if there exists a complete orthonormal system $\{e_j, j \ge 1\}$ in L^2 such that*

$$\lim_{k\to\infty} \sup_{n\ge 1} \sum_{j=k}^{\infty} \mathbb{E} \left| \left\langle \sum_{i=1}^{n} W_{n,t}, e_j \right\rangle \right|^2 = 0$$

A useful criterion for proving the above condition is given in Lemma 14 of Cerovecki and Hörmanm (2017) and is stated below.

Lemma 2.6.2. *Consider sequences $(p_j^n, j \ge 1)$, $n \ge 0$ satisfying the following properties:*

1. $p_j^n \ge 0$ *for all* j, n,

2. $\lim_{n\to\infty} p_j^n = p_j^0$,

3. $\sum_{j=1}^{n} p_j^0 = p < \infty$,

4. $\lim_{n\to\infty} \sum_{j=1}^{\infty} p_j^n = p$,

5. $\sum_{j=1}^{n} p_j^n < \infty$ *for all* $n \geq 1$.

Then

$$\lim_{k\to\infty} \sup_n \sum_{j>k} p_j^n = 0$$

We conclude this section with two central limit theorems for L^p-m-approximable stochastic processes. The first theorem concerns the sample mean and is stated as Theorem 1 of of Horváth *et al.* (2013), whereas the second concerns the covariance operator at lag 0 and is stated as Theorem 3 of Kokoszka and Reimherr (2013).

Theorem 2.6.3. *Suppose* $(X_t; t \in \mathbb{Z})$ *satisfies Assumption 1 with* $p = 2$*. Then*

$$\sqrt{n}(\overline{X}_n - \mu) \Rightarrow \mathcal{Z}_\mu$$

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

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

Theorem 2.6.4. *Suppose* $(X_t; t \in \mathbb{Z})$ *is an* L^4-m*-approximable stochastic process in* L^2*. Then*

$$\sqrt{n}(\widehat{\mathcal{C}}_0 - \mathcal{C}_0) \Rightarrow \mathcal{Z}_{\mathcal{C}_0}$$

where $\mathcal{Z}_{\mathcal{C}_0}$ *is a zero mean Gaussian Hilbert-Schmidt operator with covariance operator*

$$\begin{aligned} \Gamma = \mathbb{E}[((X_0 - \mu) \otimes (X_0 - \mu) - C_0) \otimes ((X_0 - \mu) \otimes (X_0 - \mu) - C_0)] \\ + 2\sum_{t=1}^{\infty} \mathbb{E}[((X_0 - \mu) \otimes (X_0 - \mu) - C_0) \otimes ((X_t - \mu) \otimes (X_t - \mu) - C_0)]. \end{aligned}$$

3

Moving Block and Tapered Block Bootstrap for Functional Time Series with an Application to the K-Sample Mean Problem

Abstract

We consider infinite-dimensional Hilbert space-valued random variables that are assumed to be weakly 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 data set is also discussed.

3.1 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, e.g., Ramsay and Dalzell (1991), 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 focus on functional time series, that is 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 $X_1, X_2, \ldots, X_n$, 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 $X_1, X_2, \ldots, X_n$. 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 (1990) proved the consistency of the standard i.i.d. bootstrap for the sample mean. For functional time series, Politis and Romano (1994) 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 (2017). 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 *et al.* (2015). 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 (1989) and Liu and Singh (1992). 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 (2001). 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 and therefore of the spectral density operator of the underlying stochastic process at frequency zero. We derive our theoretical results under quite general dependence assumptions on $\mathbb{X}$, i.e., under L^2-m-approximability assumptions, which are satisfied by a large class of commonly used functional time series models; see, e.g., Hörmann and Kokoszka (2010). 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, e.g., Benko *et al.* (2009), Hórvath and Kokoszka (2012, Chapter 5), Zhang (2013) and Staicu *et al.* (2015). Bootstrap alternatives over asymptotic approximations have been proposed in the same context by Benko *et al.* (2009), Zhang *et al.* (2010) and, more recently, by Paparoditis and Sapatinas (2016). Testing equality of mean functions for dependent functional data has also attracted some interest in the literature. Horváth *et al.* (2013) 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 (2015). 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 *et al.* (2013).

The paper is organised as follows. In Section 3.2, 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, i.e., of the spectral density operator of the underlying stochastic process at frequency zero, is proven. Section 3.3 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 Section 5.

3.2 Block Bootstrap Procedures for Functional Time Series

3.2.1 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 (Ω, 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 μ. Throughout Section 3.2, we assume for simplicity that $\mu = 0$. We define $\langle f, g \rangle = \int_{\mathcal{I}} f(\tau)g(\tau)\mathrm{d}\tau$, $\|f\|^2 = \langle f, f \rangle$ and the tensor product between f and g by $f \otimes g(\cdot) = \langle f, \cdot \rangle g$. For two Hilbert-Schmidt operators Ψ_1 and Ψ_2, we denote by $\langle \Psi_1, \Psi_2 \rangle_{HS} = \sum_{i=1}^{\infty} \langle \Psi_1(e_i), \Psi_2(e_i) \rangle$ the inner product which generates the Hilbert-Schmidt norm $\|\Psi_1\|_{HS} = \sum_{i=1}^{\infty} \|\Psi_1(e_i)\|^2$, for $\{e_i, i = 1, 2, \ldots\}$ 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 (2010). 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

$$X_t = f(\delta_t, \delta_{t-1}, \delta_{t-2}, \ldots) \tag{3.2.1}$$

for some measurable function $f : S^\infty \to L^2$, 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

$$\sum_{m\geq 1} \sqrt{\mathbb{E}\|X_t - X_{t,m}\|^2} < \infty, \tag{3.2.2}$$

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

The intuition behind the above definition is that the function f in (3.2.1) should be such that the effect of the innovations δ_i far back in the past becomes negligible, that is, these innovations can be replaced by other, independent, innovations. We somehow strengthen (3.2.2) to the following assumption.

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

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

Notice that the above assumption is satisfied by many linear and non-linear functional time series models cconsidered in the literature; see, e.g., Hörmann and Kokoszka (2010).

3.2.2 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 $\mathbf{X}_{\boldsymbol{n}} = \{X_t,\, t = 1, 2, \ldots, n\}$ 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 $B_t = \{X_t, X_{t+1}, \ldots, X_{t+b-1}\}$ the block of length b starting from observation X_t, $t = 1, 2, \ldots, N$, where $N = n - b + 1$ is the number of such blocks available.

Step 2 : Define i.i.d. integer-valued random variables $I_1, I_2, \ldots, I_k$ having a discrete uniform distribution assigning the probability $1/N$ to each element of the set $\{1, 2, \ldots, N\}$.

Step 3 : Let $B_i^* = B_{I_i}$, $i = 1, 2, \ldots, k$, and denote by $\{X^*_{(i-1)b+1}, X^*_{(i-1)b+2}, \ldots, X^*_{ib}\}$ the elements of B_i^*. Join the k blocks in the order $B_1^*, B_2^*, \ldots, B_k^*$ together to obtain a new set of functional pseudo observations of length n denoted by $X_1^*, X_2^*, \ldots, X_n^*$.

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

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

It is known that, under a variety of dependence assumptions on the underlying mean zero stochastic process $\mathbb{X}$, it holds true that $\sqrt{n}\overline{X}_n \xrightarrow{d} \Gamma$ as $n \to \infty$, where Γ denotes a Gaussian process with mean zero and long run covariance operator $2\pi\mathcal{F}_0$. Furthermore,

$\|n\mathbb{E}(\overline{X}_n \otimes \overline{X}_n) - 2\pi\mathcal{F}_0\|_{HS} \to 0$ as $n \to \infty$. 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 (2013a,b).

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

Theorem 3.2.1. *Suppose that the mean zero stochastic process* $\mathbb{X} = (X_t, t \in \mathbb{Z})$ *satisfies Assumption 2 and let* $X_1^*, X_2^*, \ldots, X_n^*$ *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* $n \to \infty$*. Then, as* $n \to \infty$,

(i) $$d(\mathcal{L}(\sqrt{n}\,(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \mid \mathbf{X_n}),\, \mathcal{L}(\sqrt{n}\,\overline{X}_n)) \to 0, \quad \textit{in probability,}$$

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

3.2.3 The Tapered Block Bootstrap

The TBB procedure is a modification of the block bootstrap procedure considered in Section 3.2.2 which is obtained by introducing a tapering of the random elements X_t. The tapering function down-weights the endpoints of each block B_i, towards zero, i.e., 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 $\mathbf{Y_n} = \{Y_t,\, t = 1, 2, \ldots, n\}$ stemming from a strictly stationary, L^2-valued, stochastic process $\mathbb{Y} = (Y_t,\, t \in \mathbb{Z})$, can be described as follows. Let $X_1, X_2, \ldots, X_n$ be the centered observations, i.e., $X_t = Y_t - \overline{Y}_n$, $t = 1, 2, \ldots, n$, where $\overline{Y}_n = n^{-1}\sum_{t=1}^n Y_t$. Furthermore, let $b = b(n)$, $1 \leq b < n$, be an integer and let $w_n(\cdot)$, $n = 1, 2, \ldots$, be a sequence of so-called data-tapering windows which satisfy the following assumption:

Assumption 3. $w_n(\tau) \in [0,1]$ and $w_n(\tau) = 0$ for $\tau \notin \{1, 2, \ldots, n\}$. Furthermore,

$$w_n(\tau) = w\left(\frac{\tau - 0.5}{n}\right), \tag{3.2.3}$$

where the function $w : \mathbb{R} \to [0,1]$ 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 τ 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

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

be a block of length b starting from X_t, $t = 1, 2, \ldots, N$, where each centered observation is multiplied by $w_b(\cdot)$ and scaled by $b^{1/2}/\|w_b\|_2$, where $\|w_b\|_2^2 = \sum_{i=1}^{b} w_b^2(i)$ and $\|w_b\|_1 = \sum_{i=1}^{b} w_b(t)$. Let $I_1, I_2, \ldots, I_k$ be i.i.d. integers selected from a discrete uniform distribution which assigns probability $1/N$ to each element of the set $\{1, 2, \ldots, N\}$. Let $B_i^* = \widetilde{B}_{I_i}$, $i = 1, 2, \ldots, k$, and denote the i-th block selected by $\{X^*_{(i-1)b+1}, X^*_{(i-1)b+2}, \ldots, X^*_{ib}\}$. Join these blocks together in the order $B_1^*, B_2^*, \ldots, B_k^*$ to form the set of TBB pseudo observations $X_1^*, X_2^*, \ldots, X_n^*$.

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 (2001). Furthermore, the TBB procedure uses the centered time series $X_1, X_2, \ldots, X_n$ instead of the original time series $Y_1, Y_2, \ldots, Y_n$, in order to shrink the end points of the blocks towards zero.

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

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

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

Theorem 3.2.2. *Suppose that the mean zero stochastic process $\mathbb{Y}$ satisfies Assumption 2 and let $w_n(\cdot)$, $n = 1, 2, \ldots,$ be a sequence of data-tapering windows satisfying Assumption 3. Furthermore, let $X_t^*, t = 1, 2, \ldots, n$, 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 $n \to \infty$. Then, as $n \to \infty$,*

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

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

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

Remark 3.2.1. The asymptotic validity of the MBB and TBB procedures established in Theorem 3.2.1 and Theorem 3.2.2, respectively, can be extended to cover also the case where maps $\phi : L^2 \to D$ 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 ϕ 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 (1996)). 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 de Vaart and Wellner (1996)) or quasi-Hadamard differentiable functionals (see Theorem 3.1 of Beutner and Zähe (2016)), are more involved since they depend on the particular map ϕ and the verification of some technical conditions.

3.3 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 $X_1^*, X_2^*, \ldots, X_n^*$ generated, satisfy the null hypothesis of interest.

3.3.1 The set-up

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

$$X_{i,t} = \mu_i + \varepsilon_{i,t},\ t = 1, 2, \ldots, n_i, \tag{3.3.1}$$

where, for each $i \in \{1, 2, \ldots, K\}$, $\{\varepsilon_{i,t}, t \in \mathbb{Z}\}$ 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 $\mu_i(\tau)$, $\tau \in \mathcal{I}$, is the mean function of the i-th functional time series, $i = 1, 2, \ldots, K$. The null hypothesis of interest is then,

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

and the alternative hypothesis

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

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 $\|\mu_m - \mu_k\| > 0$.

3.3.2 Block Bootstrap-based testing

The aim is to generate a set of functional pseudo observations $\mathbf{X^*_M} = X^*_{i,t}$, $i = 1, 2 \ldots, K$, $t = 1, 2, \ldots, n_i$, 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 α-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, i.e., calculate $\overline{X}_{i,n_i} = (1/n_i)\sum_{t=1}^{n_i} X_{i,t}$, for $i = 1, 2 \ldots, K$, and $\overline{X}_M = (1/M)\sum_{i=1}^{K}\sum_{t=1}^{n_i} X_{i,t}$, and obtain the residual functions in each population, i.e., calculate $\hat{\varepsilon}_{i,t} = X_{i,t} - \overline{X}_{i,n_i}$, for $t = 1, 2, \ldots, n_i$; $i = 1, 2 \ldots, K$.

Step 2 : For $i = 1, 2, \ldots, K$, let $b_i = b_i(n) \in \{1, 2, \ldots, n-1\}$ be the block size for the i-th functional time series and divide $\{\hat{\varepsilon}_{i,t},\ t = 1, 2, \ldots, n_i\}$ into $N_i = n_i - b_i + 1$ overlapping blocks of length b_i, say, $B_{i,1}, B_{i,2}, \ldots, B_{i,N_i}$. Calculate the sample mean of the ξ-th observations of the blocks $B_{i,1}, B_{i,2}, \ldots, B_{i,N_i}$, i.e., $\overline{\varepsilon}_{i,\xi} = (1/N_i)\sum_{t=1}^{N_i} \hat{\varepsilon}_{i,\xi+t-1}$, for $\xi = 1, 2, \ldots, b_i$.

Step 3 : For simplicity assume that $n_i = k_i b_i$ and for $i = 1, 2, \ldots, K$, let $q_1^i, q_2^i, \ldots, q_{k_i}^i$ be i.i.d. integers selected from a discrete probability distribution which assigns the probability $1/N_i$ to each element of the set $\{1, 2, \ldots, N_i\}$. Generate bootstrap functional pseudo observations $X_{i,t}^*$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$, as $X_{i,t}^* = \overline{X}_M + \varepsilon_{i,t}^*$, where

$$\varepsilon^*_{i,\xi+(s-1)b_i} = \hat{\varepsilon}_{i,q_s^i+\xi-1} - \overline{\varepsilon}_{i,\xi},\ s = 1, 2, \ldots, k_i,\ \xi = 1, 2, \ldots, b_i. \tag{3.3.2}$$

Step 4 : Let T_M^* be the same statistic as T_M but calculated using the bootstrap functional pseudo-time series $X_{i,t}^*$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$. 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 α-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 $i = 1, 2, \ldots, K$, by $e_{i,t}^*$, $t = 1, 2, \ldots, n_i$, the pseudo observations generated by applying the MBB procedure, described in Section 3.2.2, directly to the residuals time series $\hat{\varepsilon}_{i,t}$, $t = 1, 2, \ldots, n_i$. Note that the $e_{i,t}^*$'s differ from the $\varepsilon_{i,t}^*$'s used in (3.3.2) by the fact that the later are obtained after centering. The sample mean $\overline{\varepsilon}_{i,\xi}$, $i = 1, 2, \ldots, K$, $\xi = 1, 2, \ldots, b_i$, calculated in Step 2, is the (conditional on $\mathbf{X_M}$) expected value of the pseudo observations $e_{i,t}^*$, $t = 1, 2, \ldots, n_i$, where $t = \xi \pmod{b_i}$. Furthermore, for $i = 1, 2, \ldots, K$, we generate the $\varepsilon_{i,t}^*$'s, $t = 1, 2, \ldots, n_i$, by subtracting $\overline{\varepsilon}_{i,\xi}$ from $e_{i,sb+\xi}^*$, $\xi = 1, 2, \ldots, b$, $s = 1, 2, \ldots k_i$. 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 $X_{i,t}^*$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$, satisfy the null hypothesis H_0. In particular, given $\mathbf{X_M} = \{X_{i,t},\, i = 1, 2 \ldots, K,\, t = 1, 2, \ldots, n_i\}$, we have

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

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

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 $i = 1, 2, \ldots, K$, let $b_i = b_i(n) \in \{1, 2, \ldots, n-1\}$ be the block size for the i-th functional time series and $N_i = n_i - b_i + 1$. Let also $\{\hat{\epsilon}_{i,t},\, t = 1, 2, \ldots, n_i\}$ be the centered values of $\{\hat{\varepsilon}_{i,t},\, t = 1, 2, \ldots, n_i\}$, i.e., $\hat{\epsilon}_{i,t} = \hat{\varepsilon}_{i,t} - \overline{\varepsilon}_i$, where $\overline{\varepsilon}_i = (1/n_i)\sum_{t=1}^{n_i}\hat{\varepsilon}_{i,t}$. Also, let $w_{n_i}(\cdot)$, $n_i = 1, 2, \ldots$, be a sequence of data-tapering windows satisfying Assumption 3. Now, for $t = 1, 2, \ldots, N_i$, let

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

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

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

Step 3 : For $i = 1, 2, \ldots, K$, let $q_1^i, q_2^i, \ldots, q_{k_i}^i$ be i.i.d. integers selected from a discrete probability distribution which assigns the probability $1/N_i$ to each $t \in \{1, 2, \ldots, N_i\}$. For $i = 1, 2, \ldots, K$, $t = 1, 2, \ldots, n_i$ generate bootstrap functional pseudo-observations $X^+_{i,t}$, according to $X^+_{i,t} = \overline{X}_M + \epsilon^+_{i,t}$, where

$$\epsilon^+_{i,\xi+(s-1)b_i} = w_b(\xi)\frac{b_i^{1/2}}{\|w_{b_i}\|_2}\hat{\epsilon}_{i,k_s^i+\xi-1} - \bar{\epsilon}_{i,\xi},\; s = 1, 2, \ldots, k_i,\; \xi = 1, 2, \ldots, b_i.$$

As in the case of the MBB-based testing, the generation of $\epsilon^+_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$, ensures that the functional pseudo-time series $X^+_{i,t}$, $t = 1, 2, \ldots, n_i$, $i =$

$1, 2, \dots, K$, 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 $\mathbb{E}^+(X^+_{i,t}) = \overline{X}_M$.

3.3.3 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 *et al.* (2013) 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 (3.3.1). 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 $2\pi\mathcal{F}^{(i)}_0$, given by $c_i(u, v) = \mathbb{E}[\varepsilon_{i,0}(u)\varepsilon_{i,0}(v)] + \sum_{j\geq 1} \mathbb{E}[\varepsilon_{i,0}(u)\varepsilon_{i,j}(v)] + \sum_{j\geq 1} \mathbb{E}[\varepsilon_{i,0}(v)\varepsilon_{i,j}(u)]$. The test statistic considered in Horváth *et al.* (2013), evaluates then the L^2-distance of the two sample mean functions $\overline{X}_{i,n_i} = n_i^{-1} \sum_{t=1}^{n_i} X_{i,t}$, $i = 1, 2$, and it is given by

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

where $M = n_1 + n_2$. Horváth *et al.* (2013) proved that if $\min\{n_1, n_2\} \to \infty$ and $n_1/M \to \theta \in (0, 1)$ then, under H_0, U_M converges weakly to $\int \Gamma^2(\tau)\,\mathrm{d}\tau$, 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 $\int \Gamma^2(\tau)\,\mathrm{d}\tau$ 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

$$\Big\{\sqrt{\frac{n_1 n_2}{M}}\big(\overline{X}_{1,n_1}(\tau) - \overline{X}_{2,n_2}(\tau)\big), \tau \in \mathcal{I}\Big\}. \tag{3.3.3}$$

Theorem 3.3.1. *Let Assumption 2 be satisfied. Assume that* $\min\{n_1, n_2\} \to \infty$, $n_1/M \to \theta \in (0, 1)$ *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* $n_i \to \infty$. *Then, conditional on* $\mathbf{X_M}$, *as* $n_i \to \infty$,

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

and, if additionally Assumption 3 is satisfied,

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

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

By Theorem 3.3.1 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 (3.3.3). 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

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

and

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

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

Corollary 3.3.1. *Let the assumptions of Theorem 3.3.1 be satisfied. Then,*

(i) $\sup_{x\in\mathbb{R}}\big|P(U^*_M \leq x \mid \mathbf{X_M}) - P_{H_0}(U_M \leq x)\big| \to 0$, *in probability, and*

(ii) $\sup_{x\in\mathbb{R}}\big|P(U^+_M \leq x \mid \mathbf{X_M}) - P_{H_0}(U_M \leq x)\big| \to 0$, *in probability.*

Remark 3.3.1. 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

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

can be used. In this case, H_0 is rejected if $\widetilde{U}_M > \widetilde{d}_{M,\alpha}$ or $\widetilde{U}_M < \widetilde{d}_{M,\alpha}$, respectively, with $\widetilde{d}_{M,\alpha}$ the upper α-percentage point of the distribution of $\widetilde{U}_M$ 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) $\sup_{x\in\mathbb{R}}\big|P(\tilde{U}_M^* \leq x \mid \mathbf{X_M}) - P_{H_0}(\tilde{U}_M \leq x)\big| \to 0$, in probability, *and*

(ii) $\sup_{x\in\mathbb{R}}\big|P(\tilde{U}_M^+ \leq x \mid \mathbf{X_M}) - P_{H_0}(\tilde{U}_M \leq x)\big| \to 0$, in probability,

where

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

and

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

To elaborate, notice that using Theorem 1 of Horváth *et al.* (2013), we get, as $n_1, n_2 \to \infty$, 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 Γ_1 and Γ_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

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

which implies, for $n_1, n_2 \to \infty$ and $n_1/M \to \theta$, that $\widetilde{U}_M \xrightarrow{d} \int \Gamma(\tau)\, \mathrm{d}\tau$, 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 3.3.1, it can be easily shown that $\widetilde{U}_M^*$ and $\widetilde{U}_M^+$ converges weakly to the same limit $\int \Gamma(\tau)\, \mathrm{d}\tau$.

Another interesting class of test statistics for which Theorem 3.3.1 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 $\{\varphi_1, \varphi_2, \ldots, \varphi_p\}$ be a set of p orthonormal functions in L_2. A common choice is to let φ_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

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

and where $\widehat{\varphi}_k$ are estimators of φ_k; see for instance Horváth *et al.* (2013) where studentized versions of $\langle \overline{X}_{1,n_1} - \overline{X}_{2,n_2}, \widehat{\varphi}_k\rangle$ have also been used.

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

Corollary 3.3.2. *Let the assumptions of Theorem 3.3.1 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* φ_k*,* $k = 1, 2, \ldots, p$*, be the orthonormalized eigenfunctions corresponding to these eigenvalues and let* $\widetilde{\varphi}_k$ *and* $\widehat{\varphi}_k$ *be estimators of* φ_k *satisfying* $\max_{1\leq k\leq p} \|\widetilde{\varphi}_k - \widetilde{c}_k\varphi_k\| \xrightarrow{P} 0$ *and* $\max_{1\leq k\leq p} \|\widehat{\varphi}_k - \widehat{c}_k\varphi_k\| \xrightarrow{P} 0$*, where* $\widetilde{c}_k = sign\big(\langle\widetilde{\varphi}_k, \varphi_k\rangle\big)$ *and* $\widehat{c}_k = sign\big(\langle\widehat{\varphi}_k, \varphi_k\rangle\big)$*. Then,*

(i) $\sup_{x\in\mathbb{R}}\big|P(S^*_{p,M} \leq x \mid \mathbf{X_M}) - P_{H_0}(S_{p,M} \leq x)\big| \to 0,$ *in probability, and*

(ii) $\sup_{x\in\mathbb{R}}\big|P(S^+_{p,M} \leq x \mid \mathbf{X_M}) - P_{H_0}(S_{p,M} \leq x)\big| \to 0,$ *in probability,*

where $S^*_{p,M} = (n_1n_2/M)\sum_{k=1}^{p}\langle\overline{X}^*_{1,n_1} - \overline{X}^*_{2,n_2}, \widetilde{\varphi}_k\rangle^2$ *and* $S^+_{p,M} = (n_1n_2/M)\sum_{k=1}^{p}\langle\overline{X}^+_{1,n_1} - \overline{X}^+_{2,n_2}, \widetilde{\varphi}_k\rangle^2$.

Remark 3.3.2. In Corollary 3.3.2, we allow for $\widetilde{\varphi}_k$ to be a different estimator of φ_k than $\widehat{\varphi}_k$, where the later is used in the test statistic $S_{p,M}$. For instance, $\widetilde{\varphi}_k$ could be the same estimator as $\widehat{\varphi}_k$ but based on the the bootstrap pseudo observations $X^*_{i,t}$, $i = 1, 2, \ldots, k$ and $t = 1, 2, \ldots, n_i$, respectively, $X^+_{i,t}$, $i = 1, 2, \ldots, k$ and $t = 1, 2, \ldots, n_i$. 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 φ_k on the distribution of$S_{p,M}$. Clearly, a simple and computationally easier alternative will be to set $\widetilde{\varphi}_k = \widehat{\varphi}_k$.

Remark 3.3.3. If the alternative hypothesis H_1 is true, then under the same assumptions as in Theorem 4 of Horváth *et al.* (2013), we get that $U_M \xrightarrow{P} \infty$. Furthermore, under the same assumptions as in Theorem 6 of Horváth *et al.* (2013), we get that $S^*_{p,M} \xrightarrow{P} \infty$ and $S^+_{p,M} \xrightarrow{P} \infty$ provided that $\langle\mu_1 - \mu_2, \varphi_k\rangle \neq 0$ for at least one $1 \leq k \leq p$. Corollary 3.3.1 and Corollary 3.3.2 (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^+_{pM}.

3.4 Numerical Examples

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

$$\varepsilon_t(u) = \int \psi(u,v)\varepsilon_{t-1}(v)\,\mathrm{d}v + B_t(u), \quad u \in [0,1], \tag{3.4.1}$$

(see also Horváth *et al.* (2013)), and from a first order functional moving average model (FMA(1)),

$$\varepsilon_t(u) = \int \psi(u,v)B_{t-1}(v)\,\mathrm{d}v + B_t(u), \quad u \in [0,1]. \tag{3.4.2}$$

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

$$\psi(u,v) = \frac{\mathrm{e}^{-(u^2+v^2)/2}}{4\int \mathrm{e}^{-t^2}\mathrm{d}t}, \ (u,v) \in [0,1]^2, \tag{3.4.3}$$

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

Implementation of the MBB and TBB procedures require the selection of the block size b. As it has been shown in Theorem 3.2.1 and Theorem 3.2.2, $n\mathbb{E}^*[(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*))]$ is a consistent estimator of $2\pi\mathcal{F}_0$, with kernel

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

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 $\mathcal{W}_h = \sum_{i=1}^{b-h} w_b(i)w_b(i+h)$, $h = 0, 1, \ldots, b-1$. 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 *et al.* (2016) and Rice and Shang (2016). Although different procedures to select the "truncation lag" have been proposed in the aforementioned papers, we found the simple rule of setting $b_i = \left\lceil n_i^{1/3} \right\rceil$, 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.

For the TBB procedure, the following window has been used

$$w(\tau) = \begin{cases} \tau/0.43 & \text{if } \tau \in [0, 0.43] \\ 1 & \text{if } \tau \in [0.43, 1-0.43] \\ (1-\tau)/0.43 & \text{if } \tau \in [1-0.43, 1]. \end{cases}$$

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 $\sqrt{n}\overline{X}_n(\tau)$, the limit of which is $\sigma(\tau) = \sqrt{c(\tau,\tau)}$ for different values of $\tau \in [0,1]$ has been considered.In the following, we denote by $\hat{\sigma}(\tau)$ this estimator. 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 (3.4.1) and (3.4.2) have been used. The results obtained are presented and discussed in Section 3.4.1. Furthermore, Table 3.2 presents results comparing the performance of projections-based tests when asymptotic and bootstrap approximations are used to obtain the critical values of the tests.

3.4.1 Estimating the standard deviation of the mean function estimator

Realizations of length $n = 100$ and $n = 500$ from the functional time series models 3.3.1 with errors following either the FAR(1) model (3.4.1) or the FMA(1) model (3.4.2) have been generated and the standard deviation of the normalized sample mean $\sqrt{n}\overline{X}_n(\tau) = (1/\sqrt{n})\sum_{i=1}^{n} X_i(\tau)$ has been estimated, over a set of $\tau \in \mathcal{I}$, using the MBB, the TBB and the SB procedures. The exact standard deviation has been estimated using $100,000$ replications of the models considered. $R = 1000$ replications of each data generating process have been used where, for each replication, $B = 1000$ bootstrap pseudo-time series have been generated in order to evaluate the bootstrap estimators.

Since the results of both block bootstrap methods are, for small sample sizes, sensitive with respect to the choice of the block size b, we first present some simulations results which demonstrate the capabilities of these block bootstrap methods for functional time series. For this, we present, in some sense, the less biased results that can been obtained using the three different block bootstrap methods. That is, we present the results obtained when the block size b used has been selected as the one which minimizes the absolute averaged relative bias $T^{-1}\sum_{i=1}^{T}\left|\sigma^*_{j,b}(\tau_i)/\hat{\sigma}(\tau_i) - 1\right|$ for $j = 1, 2$. Here, $\sigma^*_{1,b}(\tau)$ and $\sigma^*_{2,b}(\tau)$ denote the MBB and TBB estimators of $\sigma(\tau)$, respectively, using the block size b. The same criterion has been used to choose the "best" probability p of the geometric distribution involved in the SB procedure i.e., the one which leads to the smallest overall in the sense described above. For the FAR(1) model and for $n = 100$, the block sizes selected using the described procedure were $b = 5$, $b = 8$ and $p = 0.25$ for the MBB, the TBB and the SB procedure, respectively. For $n = 500$, the corresponding values were $b = 10$, $b = 18$ and $p = 0.1$. For the FMA(1) model, for $n = 100$ and $n = 500$, we obtained the parameters: $b = 4$ and $b = 14$ for the MBB, $b = 6$ and $b = 10$ for the TBB, and $p = 0.5$ and $p = 0.125$ for the SB, respectively. The block bootstrap estimates of $\sigma(\tau)$ obtained using these block sizes for the FAR(1) model are presented in Figure 3.1 and for the FMA(1) model in Figure 3.2.

As it is seen from these figures, the TBB estimates perform best with the MBB estimates being better than the SB estimates. For both sample sizes considered, the block bootstrap estimators perform better in the case of the FMA(1) model than in the case of the FAR(1) model while for the FMA(1) model, the TBB estimates are quite good even for $n = 100$ observations. The results using all three bootstrap methods are

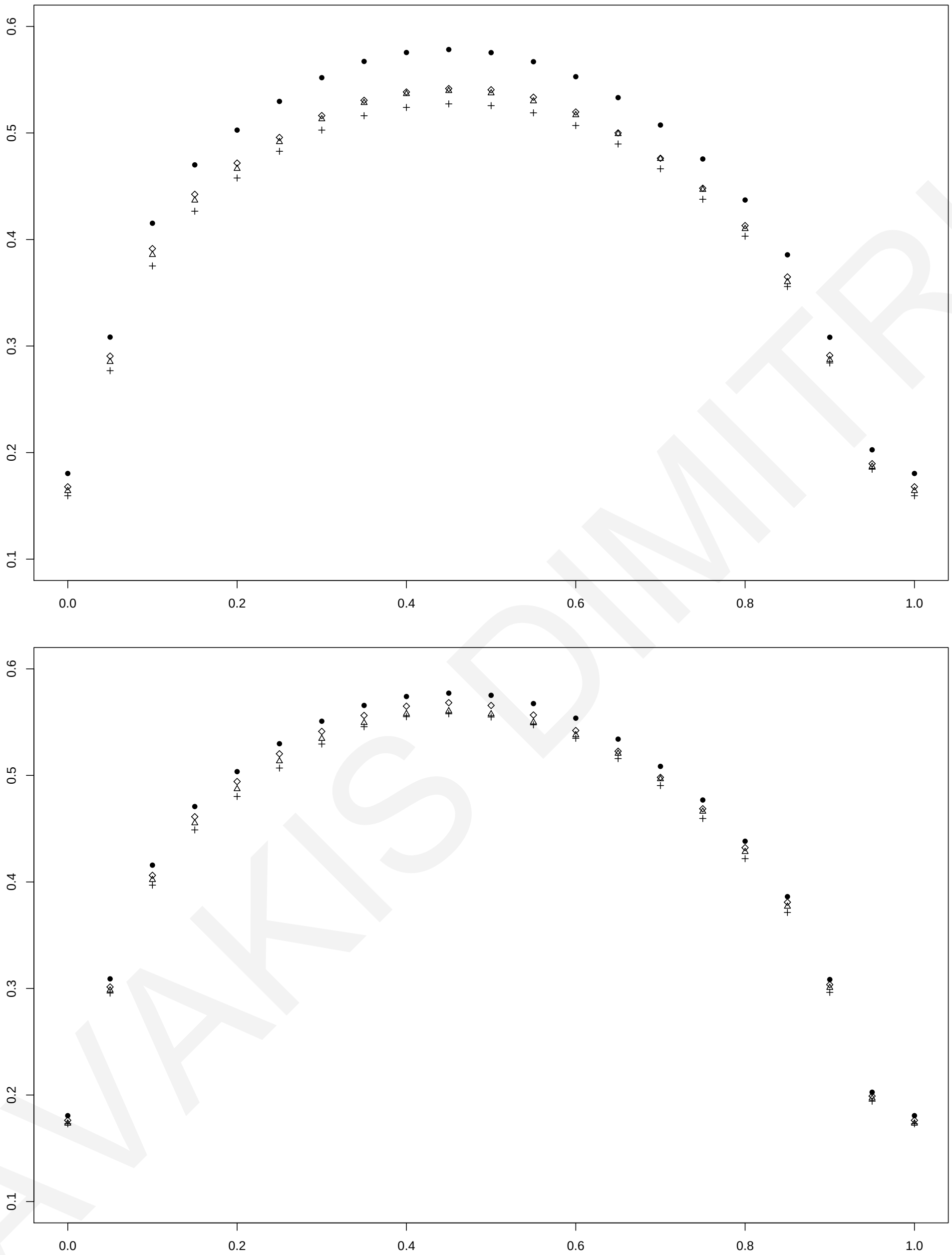

Figure 3.1: Comparison of different bootstrap estimates of the standard deviation $\sigma(\tau_i)$ of the normalized sample mean $\sqrt{n}\overline{X}_n(\tau_j)$ for FAR(1) time series and for a set of values $\tau_j \in [0, 1]$. The first figure refers to $n = 100$ and the second to $n = 500$. The estimated exact standard deviation is denoted by $\bullet$ while the mean estimates of the standard deviation of the TBB are denoted by "$\diamond$", of the MBB by "$\triangle$", and of the SB by "$+$".

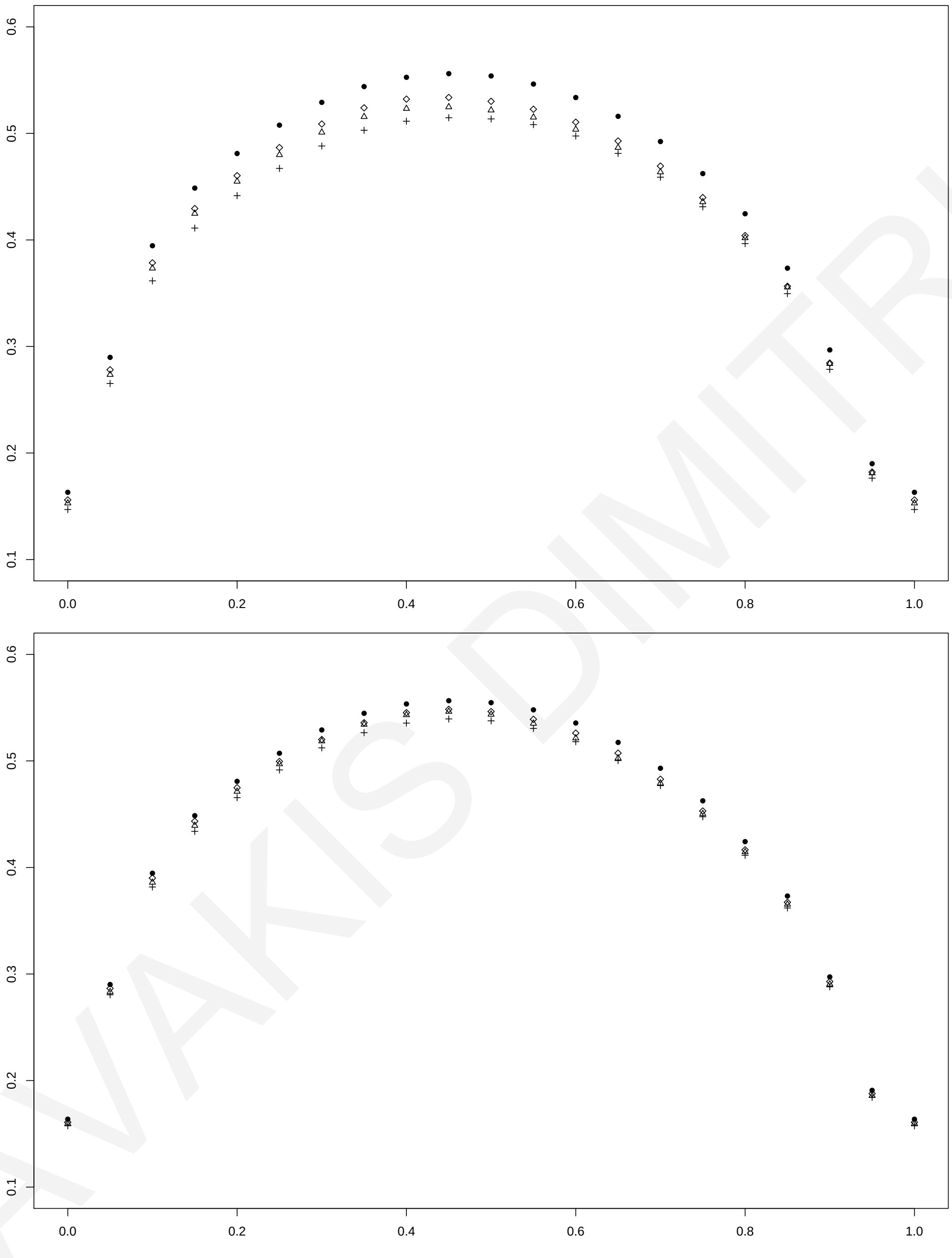

Figure 3.2: Comparison of different bootstrap estimates of the standard deviation $\sigma(\tau_i)$ of the normalized sample mean $\sqrt{n}\overline{X}_n(\tau_j)$ for FMA(1) time series and for a set of values $\tau_j \in [0, 1]$. The first figure refers to $n = 100$ and the second to $n = 500$. The estimated exact standard deviation is denoted by $\bullet$ while the mean estimates of the standard deviation of the TBB are denoted by "$\diamond$", of the MBB by "$\triangle$", and of the SB by "$+$".

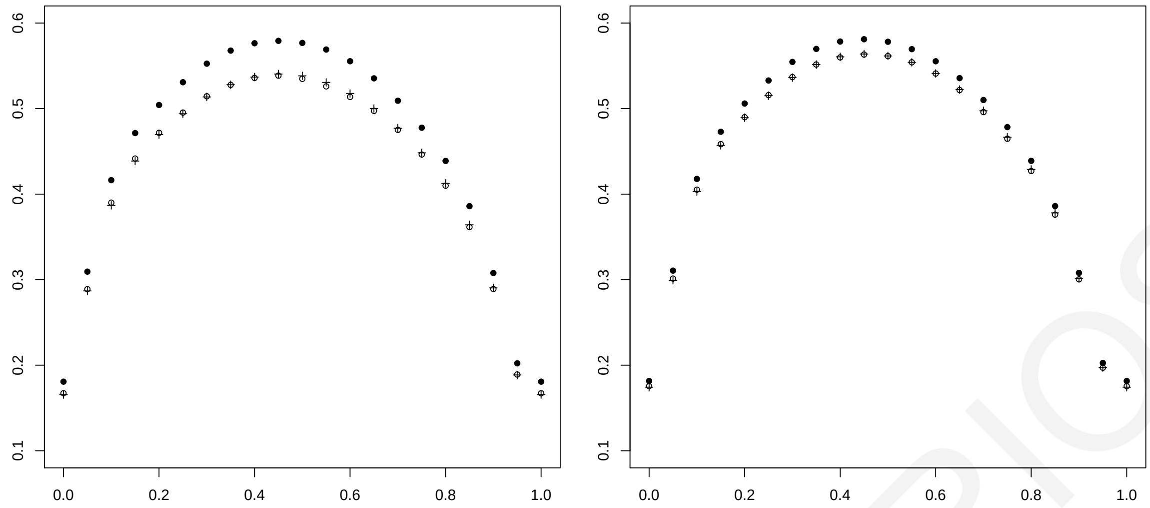

Figure 3.3: TBB estimates of the standard deviation $\sigma(\tau_i)$ of the normalized sample mean $\sqrt{n}\overline{X}_n(\tau_j)$ for the FAR(1) time series and for a set of values $\tau_j \in [0,1]$ using the "best" block size and the block size $b^* = \lceil n^{1/3} \rceil$. The first figure refers to $n = 100$ and the second to $n = 500$. The estimated exact standard deviation is denoted by • while the TBB estimates using the "best" block size are denoted by "∘" and using the block size b^* are denoted by "+".

better for the larger sample size of $n = 500$ curves.

To demonstrate the performance of the suggested simpler rule $b^* = \lceil n^{1/3} \rceil$ to choose the block size b, the TBB estimates using this block size are compared with the estimates obtained using the block size leading to the less biased estimates, as described above. Comparisons for the FAR(1) and for the FMA(1) model are shown in Figure 3.3 and Figure 3.4 respectively.

As these figures demonstrate, for both sample sizes and for both models considered, the TBB estimates using the block size b^* perform well, being quite close to the TBB estimates using the "best" block size in the sense described above.

3.4.2 Testing equality of mean functions

We investigate the size and power performance of the tests considered in Section 3.3.3. As can be seen in Section 3.4.1, 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 $b = b_1 = b_2$, 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

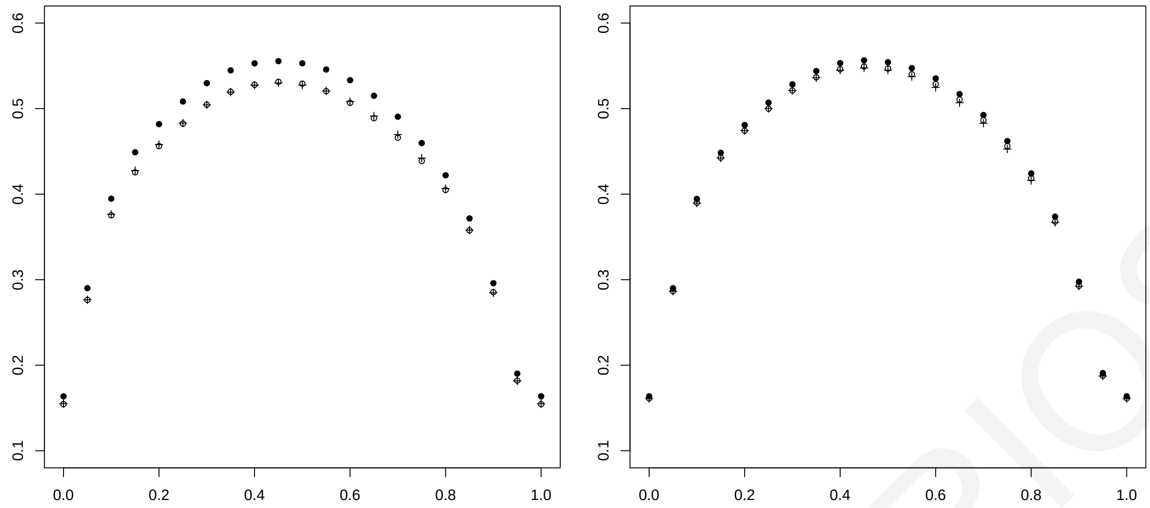

Figure 3.4: TBB estimates of the standard deviation $\sigma(\tau_i)$ of the normalized sample mean $\sqrt{n}\overline{X}_n(\tau_j)$ for the FMA(1) time series and for a set of values $\tau_j \in [0, 1]$ using the "best" block size and the block size $b^* = \lceil n^{1/3} \rceil$. The first figure refers to $n = 100$ and the second to $n = 500$. The estimated exact standard deviation is denoted by $\bullet$ while the TBB estimates using the "best" block size are denoted by "$\circ$" and using the block size b^* by "+".

the two samples were generated according to model (3.3.1) and with the errors $\varepsilon_{i,t}$ following model (3.4.1), 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 *et al.* (2013). The results obtained are shown in Table 3.1 for a range of values of γ. 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 γ) 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.

3.4.3 TBB-based Test versus Projection-based Tests

We compare the performance of the TBB-based test with the projection-based tests $U_{n_1,n_2}^{(1)}$ and $U_{n_1,n_2}^{(2)}$ proposed in Horváth *et al.* (2013) (see (3.11) and (3.12) in their paper). We adopted their simulation set up and generated two samples according to the functional time series model 3.3.1 with the errors $\varepsilon_{i,t}$ following the FAR(1)

γ	b	$n_1 = n_2 = 100$ $\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	b	$n_1 = n_2 = 200$ $\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
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

Table 3.1: Empirical size and power of the test based on TBB critical values and FAR(1) errors.

model 3.4.1 with kernel 3.4.3, 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. All curves were approximated using $T = 49$ equidistant points $\tau_1, \tau_2, \ldots, \tau_{49}$ in the unit interval $\mathcal{I}$ and transformed into functional objects using the Fourier basis with 49 basis functions.

We considered sample sizes $n_1 = 100$ and $n_2 = 200$ and block sizes $b = b_1 = 6$ and 8 (for $n_1 = 100$) and $b = b_2 = 6$ and 10 (for $n_2 = 200$). The tests have been applied using three nominal levels, i.e., $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.1$. All bootstrap calculations were based on $B = 1000$ bootstrap replicates and $R = 1000$ model repetitions. The results obtained are shown in Table 3.2 for a range of values of γ. Notice that $\gamma = 0$ corresponds to the null hypothesis. The empirical rejection frequencies of the projection-based tests $U^{(1)}_{n_1,n_2}$ and $U^{(2)}_{n_1,n_2}$ are those reported in Table 2 of Horváth *et al.* (2013).

γ	$\alpha = 0.01$ $U^{(1)}_{n_1,n_2}$	$U^{(2)}_{n_1,n_2}$	TBB	$\alpha = 0.05$ $U^{(1)}_{n_1,n_2}$	$U^{(2)}_{n_1,n_2}$	TBB	$\alpha = 0.10$ $U^{(1)}_{n_1,n_2}$	$U^{(2)}_{n_1,n_2}$	TBB
0.0	0.018	0.019	0.017	0.066	0.072	0.070	0.122	0.135	0.128
			0.016			0.070			0.122
0.2	0.051	0.033	0.058	0.136	0.116	0.149	0.216	0.187	0.235
			0.046			0.142			0.236
0.4	0.194	0.123	0.150	0.359	0.265	0.322	0.467	0.363	0.431
			0.178			0.364			0.476
0.6	0.421	0.296	0.405	0.622	0.518	0.633	0.731	0.625	0.737
			0.425			0.649			0.738
0.8	0.686	0.538	0.684	0.857	0.746	0.847	0.915	0.831	0.920
			0.674			0.849			0.910
1.0	0.874	0.787	0.870	0.959	0.908	0.952	0.981	0.945	0.977
			0.881			0.959			0.987
1.2	0.976	0.937	0.964	0.995	0.981	0.990	0.998	0.992	0.995
			0.973			0.994			0.997

Table 3.2: Empirical rejection frequencies of the projection-based tests $U^{(1)}_{n_1,n_2}$ and $U^{(2)}_{n_1,n_2}$ are the results reported in Table 2 of Horváth *et al.* (2013). For the TBB-base test, the first line corresponds to the choices $b = 6$ and $b = 8$ and the second line to the choices $b = 6$ and $b = 10$ of the block size for sample sizes $n_1 = 100$ and $n_2 = 200$, respectively.

As can be seen from Table 3.2, the TBB-based test performs well retaining the nominal sizes and having a power which increases as the deviation from H_0 increases, as described by the parameter γ. Compared to the projection-based test $U^{(2)}_{n_1,n_2}$, the TBB-based test performs better while its empirical size and power is similar to that of the projection-based test $U^{(1)}_{n_1,n_2}$. Notice, however, that the TBB-based test is consistent against any alternative for which $\|\mu_1 - \mu_2\| > 0$ which is not the case with the $U^{(1)}_{n_1,n_2}$ (and $U^{(2)}_{n_1,n_2}$) test if such alternatives are orthogonal to the projection space.

3.4.4 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 $\{X_{i,t}(\tau), i = 1, 2, t = 1, 2, \ldots, 92\}$, 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

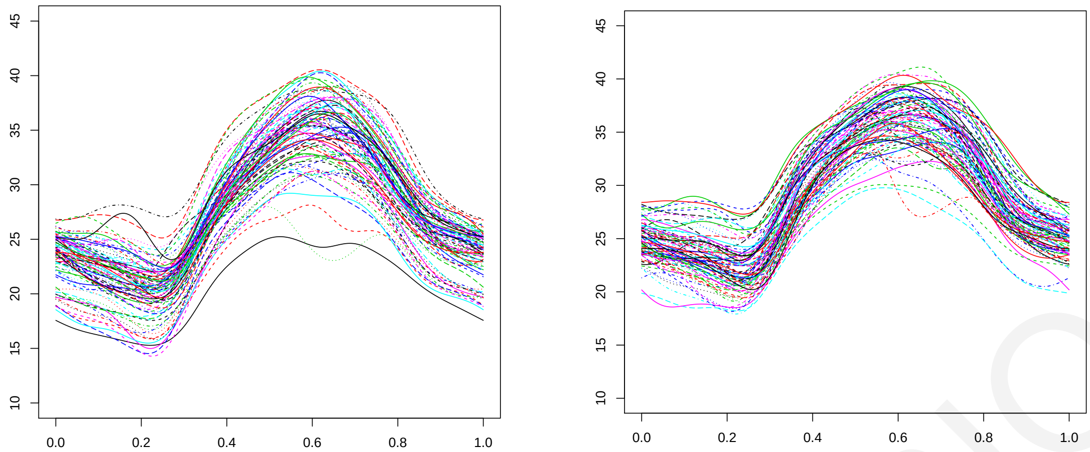

Figure 3.5: Temperature curves: summer 2005 (left panel) and summer 2009 (right panel).

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 3.5 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 $\widetilde{U}_M$ 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 $\widetilde{U}_M$ are quite small leading to the rejection of H_0 for all commonly used α-levels.

3.5 Appendix : Proofs

To prove Theorem 3.2.1 and Theorem 3.2.2, we first establish Lemma 3.5.1 and Lemma 3.5.2. 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 3.5.1. *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 2 and $b = b(n)$ is a sequence of integers such that $b^{-1} + bn^{-1/2} = o(1)$ as $n \to \infty$. Assume further that, for any fixed*

u, $g_b(u) \to 1$ as $n \to \infty$. Then, as $n \to \infty$,

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

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

Proof. First, note by the independence of X_0 and $X_{i,i}$, that $\sum_{i=1}^{\infty}|\mathbb{E}\langle X_0, X_i\rangle| = \sum_{i=1}^{\infty}|\mathbb{E}\langle X_0, X_i - X_{i,i}\rangle| \le \sum_{i=1}^{\infty}(\mathbb{E}\|X_0\|^2)^{1/2}(\mathbb{E}\|X_0 - X_{0,i}\|^2)^{1/2}$, which implies by (3.2.2) 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 $n \to \infty$, it suffices to show that, as $n \to \infty$,

$$\sum_{h=1}^{b-1} g_b(h)\frac{1}{n}\sum_{i=1}^{n-h}\langle X_i, X_{i+h}\rangle - \sum_{i\ge 1}\mathbb{E}\langle X_0, X_i\rangle = o_P(1). \tag{3.5.1}$$

Let

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

Since

$$\left|\sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h - c_{\infty}^{+}\right| \le |c_m^{+} - c_{\infty}^{+}| + \left|\sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)} - c_m^{+}\right| + \left|\sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h - \sum_{h=1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)}\right|, \tag{3.5.2}$$

assertion (3.5.1) is proved by showing that there exists $m_0 \in \mathbb{N}$ such that all three terms on the right hand side of (3.5.2) can be made arbitrarily small in probability as $n \to \infty$ for all $m \ge m_0$.

For the first term, we use the bound

$$\left|\sum_{i\ge 1}\mathbb{E}\left[\langle X_{0,m}, X_{i,m}\rangle - \langle X_0, X_i\rangle\right]\right| \le \left|\sum_{i=1}^{m}\mathbb{E}\left[\langle X_{0,m}, X_{i,m}\rangle - \langle X_0, X_i\rangle\right]\right| + \left|\sum_{i=m+1}^{\infty}\mathbb{E}\left[\langle X_{0,m}, X_{i,m}\rangle - \langle X_0, X_i\rangle\right]\right|, \tag{3.5.3}$$

and handle the first term on the right hand side of (3.5.3) 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 2 yields that for every $\epsilon_1 > 0$, $\exists\, m_1 \in \mathbb{N}$ such that

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

for all $m \geq m_1$. For the second term of the right hand side of (3.5.3), 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 & Kokoszka (2012), that, for any $\epsilon_2 > 0$, there exists $m_2 \in \mathbb{N}$ such that

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

for all $m \geq m_2$ because of (3.2.2). For the second term of (3.5.2), first note that, for every fixed $m \geq 1$ and for any fixed h, we have that $\left|\hat{\gamma}_h^{(m)} - \mathbb{E}[\langle X_{0,m}, X_{h,m}\rangle]\right| = o_p(1)$. 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 (3.5.2) is $o_p(1)$, if we show that $\left|\sum_{h=m+1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)}\right| = o_p(1)$. We have

$$\mathbb{E}\left[\sum_{h=m+1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)}\right]^2$$
$$= n^{-2}\sum_{h_1=m+1}^{b-1}\sum_{h_2=m+1}^{b-1}\sum_{i_1=1}^{n-h_1}\sum_{i_2=1}^{n-h_2} g_b(h_1)g_b(h_2)\mathbb{E}(\langle X_{i_1,m}, X_{i_1+h_1,m}\rangle\langle X_{i_2,m}, X_{i_2+h_2,m}\rangle).$$

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 & Kokoszka (2012) 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, $\mathbb{E}\left[\sum_{h=m+1}^{b-1} g_b(h)\hat{\gamma}_h^{(m)}\right]^2 = O(b/n) = o(1)$ from which the desired convergence follows by Markov's inequality. For the third term in (3.5.2), we show that, for $m \geq m_0$,

$$\limsup_{n\to\infty} P\left(\left|\sum_{h=1}^{b-1} g_b(h)\left(\hat{\gamma}_h - \hat{\gamma}_h^{(m)}\right)\right| > \delta\right) = 0, \tag{3.5.4}$$

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

$$\mathbb{E}\left|\sum_{h=1}^{b-1} g_b(h)(\hat{\gamma}_h - \hat{\gamma}_h^{(m)})\right| = o(1). \tag{3.5.5}$$

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

$$\begin{aligned}
&\mathbb{E}\left|\sum_{h=1}^{b-1} g_b(h)\left(\hat{\gamma}_h - \hat{\gamma}_h^{(m)}\right)\right| \\
&\quad \leq \mathbb{E}\left|\frac{1}{n}\sum_{h=1}^{m} g_b(h)\sum_{i=1}^{n-h}\left(\langle X_i, X_{i+h}\rangle - \langle X_{i,m}, X_{i+h,m}\rangle\right)\right| \\
&\qquad + \mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\left(\langle X_i, X_{i+h}\rangle - \langle X_{i,m}, X_{i+h,m}\rangle\right)\right|.
\end{aligned} \tag{3.5.6}$$

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,

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

Assumption 2 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 ϵ_3. For the second term on the right hand side of (3.5.6), we use the bound

$$\mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_i, X_{i+h}\rangle\right| + \mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_{i,m}, X_{i+h,m}\rangle\right|. \tag{3.5.7}$$

Note that the second summand of (3.5.7) 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

$$\mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_i, X_{i+h,h}\rangle\right| + \mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_i, X_{i+h} - X_{i+h,h}\rangle\right|. \tag{3.5.8}$$

For the last term of expression (3.5.8), we get, using (3.2.2), that for every $\epsilon_4 > 0$, there exists $m_4 \in \mathbb{N}$ such that

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

for all $m \geq m_4$. Consider next the first term of (3.5.8). 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

$$\mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_i - X_{i,h}, X_{i+h,h}\rangle\right| + \mathbb{E}\left|\frac{1}{n}\sum_{h=m+1}^{b-1} g_b(h)\sum_{i=1}^{n-h}\langle X_{i,h}, X_{i+h,h}\rangle\right|. \quad (3.5.9)$$

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 (3.2.2), 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 ϵ_5. For the last term of (3.5.9), 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, i.e., $\mathbb{E}\langle X_i, X_{i+h,h}\rangle = 0$ 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 $\mathbb{E}\left|n^{-1/2}\sum_{i=1}^{n}\langle X_{i,h}, X_{i+h,h}\rangle\right| = O(1)$. Therefore,

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

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

Lemma 3.5.2. *Suppose that $(Y_t, t \in \mathbb{Z})$ satisfies Assumption 2 and that $b = b(n)$ is a sequence of integers satisfying $b^{-1} + bn^{-1/2} = o(1)$ as $n \to \infty$. Let $w_n(\cdot)$, $i = 1, 2, \ldots$, be a sequence of data-tappering windows satisfying Assumption 3. Then, as $n \to \infty$,*

(i)

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

(ii)

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

where $c(u,v)=\sum_{i=-\infty}^{\infty}\mathbb{E}[Y_0(u)Y_i(v)]$, $\mathcal{W}_h=\sum_{i=1}^{b-h}w_b(i)w_b(i+h)$, $h=0,1,\ldots,b-1$ *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. Consider (i). Note first that, using Lemma 2.1 of Hörmann and Kokoszka (2010), the sequence $\{\langle Y_i,y\rangle,\ i=1,2,\ldots\}$ is L^2-m-approximable, since

$$\sum_{m\geq 1}\left(\mathbb{E}|\langle Y_i-Y_{i,m},y\rangle|^2\right)^{1/2}\leq\|y\|\sum_{m\geq 1}\left(\mathbb{E}\|Y_i-Y_{i,m}\|^2\right)^{1/2}<\infty.$$

Therefore, by Lemma 4.1 of Hörmann and Kokoszka (2010), we get that

$$\sum_{i=-\infty}^{\infty}|\mathbb{E}\langle Y_0,y\rangle\langle Y_i,y\rangle|<\infty. \tag{3.5.10}$$

Also, note that if $w_b(i)$ is of the form (3.2.3), then

$$\frac{\mathcal{W}_h}{bw*w(h/b)}\to 1,$$

where $\mathcal{W}_h=\sum_{i=1}^{b-h}w_1(i)w_b(i+h)$, $h=0,1,\ldots,b-1$, and $w*w$ denotes is the self-convolution of w. Therefore, since $\|w_b\|_2^2=\mathcal{W}_0$, we get, for any fixed h, as $n\to\infty$,

$$\frac{\mathcal{W}_h}{\|w_b\|_2^2}=\frac{\mathcal{W}_h}{bw*w(h/b)}\frac{bw*w(0)}{\mathcal{W}_0}\frac{bw*w(h/b)}{bw*w(0)}\to 1. \tag{3.5.11}$$

Furthermore, by Cauchy-Schwarz's inequality, it is easily seen that

$$\sum_{i=1}^{b-h} w_b(i) w_b(i+h) \leq \sum_{i=1}^{b} w_b^2(i),$$

i.e.,

$$\mathcal{W}_h \leq \|w_b\|_2^2 \quad \text{for } h = 1, 2, \ldots, b-1. \tag{3.5.12}$$

To complete the proof of (i), it suffices to prove that $\sum_{h=1}^{b-1} (\mathcal{W}_h / \|w_b\|_2^2) \mathbb{E}\langle Y_0, y\rangle \langle Y_h, y\rangle \to \sum_{h=1}^{\infty} \mathbb{E}\langle Y_0, y\rangle\langle Y_h, y\rangle$. For this, and for b large enough, we use the bound

$$\begin{aligned}
&\left| \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \mathbb{E}\langle Y_0, y\rangle\langle Y_h, y\rangle - \sum_{i=1}^{\infty} \mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle \right| \\
&\qquad \leq \left| \sum_{h=1}^{m} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \mathbb{E}\langle Y_0, y\rangle\langle Y_h, y\rangle - \sum_{i=1}^{m} \mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle \right| \\
&\qquad\quad + \left| \sum_{h=m+1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \mathbb{E}\langle Y_0, y\rangle\langle Y_h, y\rangle - \sum_{i=m+1}^{b-1} \mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle \right| \\
&\qquad\quad + \left| \sum_{i=b}^{\infty} \mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle \right|.
\end{aligned} \tag{3.5.13}$$

Because of (3.5.11) and (3.5.10), the first and the last term are $o(1)$. Concerning the second term, we show that there exists $m_0 \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} \left| \sum_{h=m+1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \mathbb{E}\langle Y_0, y\rangle\langle Y_h, y\rangle - \sum_{i=m+1}^{b-1} \mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle \right| = 0$$

for $m = m_0$. By using Assumption 2, expression (3.5.12), the facts that $\mathcal{W}_h \geq 0$ and that $\langle Y_0, y\rangle$ and $\langle Y_{i,i}, y\rangle$ are independent for $i \geq m+1$, we get that, for every $\epsilon > 0$, $\exists\, m_1 \in \mathbb{N}$ such that, for every $m \geq m_1$,

$$\begin{aligned}
&\left| \sum_{h=m+1}^{b-1} \left(\frac{\mathcal{W}_h}{\|w_b\|_2^2} - 1 \right) \mathbb{E}(\langle Y_0, y\rangle\langle Y_h, y\rangle) \right| \\
&\qquad \leq \sum_{i=m+1}^{\infty} |\mathbb{E}\langle Y_0, y\rangle\langle Y_i, y\rangle| \\
&\qquad = \sum_{i=m+1}^{\infty} |\mathbb{E}\langle Y_0, y\rangle\langle Y_i - Y_{i,i}, y\rangle| \\
&\qquad \leq \|y\|^2 \left(\mathbb{E}\|Y_0\|^2\right)^{1/2} \sum_{i=m+1}^{\infty} \left(\mathbb{E}\|Y_i - Y_{i,i}\|^2\right)^{1/2} < \epsilon,
\end{aligned} \tag{3.5.14}$$

because of expression 3.2.2

Consider next assertion (ii). Notice first that,

$$\iint \left\{ \frac{1}{n} \sum_{t=1}^{n} Y_t(u) Y_t(v) - \mathbb{E}[Y_0(u)Y_0(v)] \right\}^2 = o_P(1).$$

Hence, and since the summands of $Y_i(u)Y_{i+h}(v)$ and $Y_{i+h}(v)Y_i(u)$ can be handled similarly, it suffices to show that

$$\iint \left\{ \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \frac{1}{n} \sum_{t=1}^{n-h} Y_t(u) Y_{t+h}(v) - \sum_{t \geq 1} \mathbb{E}[Y_0(u)Y_t(v)] \right\}^2 = o_P(1). \tag{3.5.15}$$

By expressions (3.5.11) and (3.5.12), the proof of (3.5.15) is analogous to the proof of $(A.2)$ of Horváth *et al.* (2013). This completes the proof of the lemma.

Proof of Theorem 3.2.1. By the triangle inequality and Theorem 1 of Horváth *et al.* (2013), the assertion (i) of the theorem is established if we show that, as $n \to \infty$,

$$\sqrt{n}(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \Rightarrow \Gamma, \text{ in probability,} \tag{3.5.16}$$

where Γ 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 $S_n^* = \sqrt{n}(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*))$, it follows from Proposition 7.4.2 of Laha and Rohatgi (1979) that, to prove (3.5.16), 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 $n \to \infty$,

$$\mathrm{Var}^* \left(\langle S_n^*, y \rangle \right) \xrightarrow{P} \sigma^2(y) \tag{3.5.17}$$

and that

$$\frac{\langle S_n^*, y \rangle}{\sqrt{\mathrm{Var}^*(\langle S_n^*, y \rangle)}} \xrightarrow{d} N(0,1). \tag{3.5.18}$$

Consider (3.5.17) and notice that

$$S_n^* = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left[U_i^* - \mathbb{E}^*(U_i^*) \right],$$

where $U_i^* = b^{-1/2}(X^*_{(i-1)b+1} + X^*_{(i-1)b+2} + \ldots + X^*_{ib})$, $i = 1, 2, \ldots, k$. Due to the block bootstrap resampling scheme, the random variables U_i^*, $i = 1, 2, \ldots, k$ 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 $W_i^* = \langle U_i^*, y \rangle$, $i = 1, 2, \ldots, k$, we have

$$\mathrm{Var}^* \left(\langle S_n^*, y \rangle \right) = \mathbb{E}^*(W_1^*)^2 - (\mathbb{E}^*(W_1^*))^2. \tag{3.5.19}$$

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

$$\mu^* = \frac{\sqrt{b}}{N} \left[\sum_{i=1}^{n} \langle X_i, y \rangle - \sum_{j=1}^{b-1} \left(1 - \frac{j}{b} \right) \left[\langle X_j, y \rangle + \langle X_{n-j+1}, y \rangle \right] \right]. \tag{3.5.20}$$

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

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

we get,

$$\mathbb{E}(\mu^*)^2 = \frac{b}{N^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}[\langle X_i, y \rangle \langle X_j, y \rangle] + O(b^2/n) = O(b^2/n), \tag{3.5.21}$$

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

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}[\langle X_i, y \rangle \langle X_j, y \rangle] &= \sum_{|h|<n} \left(1 - \frac{|h|}{n} \right) \mathbb{E}[\langle X_0, y \rangle \langle X_h, y \rangle] \\
&\to \iint c(u, v) y(u) y(v) \mathrm{d}u \mathrm{d}v
\end{aligned} \tag{3.5.22}$$

as $n \to \infty$. Since $\mathbb{E}^*(\mu^*) = 0$, (3.5.21) implies that $\mu^* = O_P(b/\sqrt{n})$.

Consider next the first term of the right hand side of expression (3.5.19). For this term, we have

$$
\begin{aligned}
\mathbb{E}^*(W_1^*)^2 &= \frac{1}{N}\sum_{i=1}^{N}\langle U_i, y\rangle^2 \qquad (3.5.23)\\
&= \frac{1}{N}\sum_{i=1}^{n}\langle X_i, y\rangle\langle X_i, y\rangle \\
&+ \sum_{h=1}^{b-1}\left(1-\frac{h}{b}\right)\frac{1}{N}\sum_{i=1}^{n-h}[\langle X_i, y\rangle\langle X_{i+h}, y\rangle + \langle X_{i+h}, y\rangle\langle X_i, y\rangle] \\
&- \frac{1}{N}\sum_{s=1}^{b-1}\left(1-\frac{s}{b}\right)[\langle X_s, y\rangle\langle X_s, y\rangle + \langle X_{n-s+1}, y\rangle\langle X_{n-s+1}, y\rangle] \\
&- \frac{1}{N}\sum_{t=1}^{b-1}\sum_{j=1}^{b-t}\left(1-\frac{j+t}{b}\right)[\langle X_j, y\rangle\langle X_{j+t}, y\rangle + \langle X_{n-j+1-t}, y\rangle\langle X_{n-j+1}, y\rangle \\
&\qquad + \langle X_{j+t}, y\rangle\langle X_j, y\rangle + \langle X_{n-j+1}, y\rangle\langle X_{n-j+1-t}, y\rangle].
\end{aligned}
$$

Thus,

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

from which we get

$$
\mathrm{Var}^*(W_1^*) = \iint c_N(u,v)y(u)y(v)\mathrm{d}u\mathrm{d}v + O_p(b^2/n), \qquad (3.5.24)
$$

where

$$
\begin{aligned}
c_N(u,v) &= \frac{1}{N}\sum_{i=1}^{n}X_i(u)X_i(v) \\
&+ \sum_{h=1}^{b-1}\left(1-\frac{h}{b}\right)\frac{1}{N}\sum_{i=1}^{n-h}\Big[X_i(u)X_{i+h}(v) + X_{i+h}(u)X_i(v)\Big]. \qquad (3.5.25)
\end{aligned}
$$

By the ergodic theorem and equation $(A.2)$ of Horváth *et al.* (2013), 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

$$\iint [c_n(u,v) - c(u,v)]^2 \mathrm{d}u\mathrm{d}v = o_P(1) \tag{3.5.26}$$

as $n \to \infty$, 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 $n \to \infty$,

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

That is,

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

Since $N/n \to 1$ as $n \to \infty$, we finally get from (3.5.24) that,

$$\mathrm{Var}^*\langle S_n^*, y\rangle = \mathrm{Var}^*(W_1^*) \xrightarrow{P} \iint c(u,v)y(u)y(v)\mathrm{d}u\mathrm{d}v = \sigma^2(y). \tag{3.5.27}$$

Consider next (3.5.18). Observe that $W_i^* = \langle U_i^*, y\rangle$, $i = 1, 2, \ldots, k$ are i.i.d. random variables and, therefore, it suffices to show that Lindeberg's condition

$$\lim_{n\to\infty} \frac{1}{\tau_k^{*2}} \sum_{t=1}^{k} \mathbb{E}^* \left[(W_t^* - \mu^*)^2 \mathbb{1}(|W_t^* - \mu^*| > \varepsilon \tau_k^*) \right] = 0, \quad \text{for every } \varepsilon > 0, \tag{3.5.28}$$

is fulfilled, where $\tau_k^{*2} = \sum_{t=1}^{k} \mathrm{Var}^*(W_t^*) = k\mathrm{Var}^*(W_1^*)$ and $\mu^* = \mathbb{E}^*(W_i^*)$. To establish (3.5.28), and because of (3.5.27), it suffices to show that, for any $\delta > 0$ and as $n \to \infty$,

$$P\left(\frac{1}{k} \sum_{t=1}^{k} \mathbb{E}^* \left[(W_t^* - \mu^*)^2 \mathbb{1}(|W_t^* - \mu^*| > \varepsilon \tau_k^*) \right] > \delta \right) \to 0. \tag{3.5.29}$$

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

$$\mathbb{E}[|X+Y|^2 \mathbb{1}(|X+Y| > \eta)] \\ \leq 4\left[\mathbb{E}|X|^2 \mathbb{1}(|X| > \eta/2) + \mathbb{E}|Y|^2 \mathbb{1}(|Y| > \eta/2) \right]; \tag{3.5.30}$$

see Lahiri (2003), p. 56. We then get by Markov's inequality that

$$\begin{aligned}
&P\left(\frac{1}{k}\sum_{t=1}^{k}\mathbb{E}^*\left[(W_t^*-\mu^*)^2\mathbb{1}(|W_t^*-\mu^*|>\varepsilon\tau_k^*)\right]>\delta\right)\\
&\quad\le \delta^{-1}\mathbb{E}\left\{\mathbb{E}^*\left[(W_1^*-\mu^*)^2\mathbb{1}(|W_1^*-\mu^*|>\varepsilon\tau_k^*)\right]\right\}\\
&\quad= \delta^{-1}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}(W_i-\mu^*)^2\mathbb{1}(|W_i-\mu^*|>\varepsilon\tau_k^*)\right]\\
&\quad= \delta^{-1}\mathbb{E}\left[(W_1-\mu^*)^2\mathbb{1}(|W_1-\mu^*|>\varepsilon\tau_k^*)\right]\\
&\quad\le 4\delta^{-1}\left[\mathbb{E}W_1^2\mathbb{1}(|W_1|>\varepsilon\tau_k^*/2)+\mathbb{E}(\mu^*)^2\right], \qquad (3.5.31)
\end{aligned}$$

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

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

as $n\to\infty$. Therefore, by the dominated convergence theorem, $\lim_{n\to\infty}\mathbb{E}W_1^2\mathbb{1}(|W_1|>\varepsilon\tau_k^*/2)=0$, Hence, using expression (3.5.21), we conclude that (3.5.31) converges to 0 as $n\to\infty$.

To establish ($L2$), it suffices, by Theorem 1.13 of Prokhorov (1956) and Theorems 5.1 and 5.2 of Billingsley (1999), to prove that $\lim_{k\to\infty}\sup_{n\ge1}\sum_{j=k}^{\infty}\mathbb{E}|\langle S_n^*,e_j\rangle|^2=0$, where $\{e_j,\ j\ge1\}$ is a complete orthonormal basis of L^2. Using $\mathbb{E}^*|\langle S_n^*,e_j\rangle|^2=\mathrm{Var}^*(\langle U_1^*,e_j\rangle)$ and Lemma 14 of Cerovecki and Hörmanm (2017), ($L2$) is satisfied if the following five conditions are fulfilled.

(a) $\mathrm{Var}^*(\langle U_1^*,e_j\rangle)\ge0\quad\forall j,n;$

(b) $\lim_{n\to\infty}\mathrm{Var}^*(\langle U_1^*,e_j\rangle)=\Sigma_j,$ in probability;

(c) $\sum_{j\ge1}\Sigma_j<\infty;$

(d) $\lim_{n\to\infty}\sum_{j\ge1}\mathrm{Var}^*(\langle U_1^*,e_j\rangle)=\sum_{j\ge1}\Sigma_j,$ in probability;

(e) $\sum_{j\ge1}\mathrm{Var}^*(\langle U_1^*,e_j\rangle)$ is bounded for all $n\ge1$, in probability.

Note that, by letting $y=e_j$ in expression (3.5.27), property (b) follows with $\Sigma_j=\iint c(u,v)e_j(u)e_j(v)\mathrm{d}u\mathrm{d}v$. To prove (c), notice that, by Proposition 6 of Hörmanm *et al.* (2015), 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,

$$\sum_{j\geq 1}\Sigma_j=\sum_{j\geq 1}\iint c(u,v)e_j(u)e_j(v)\mathrm{d}u\mathrm{d}v=\sum_{j\geq 1}\lambda_j<\infty, \tag{3.5.32}$$

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

To establish (d), we get, using (3.5.20), that

$$\begin{aligned}
&\mathrm{Var}^*(\langle U_1^*, e_j\rangle)\\
&\quad=\frac{1}{N}\sum_{i=1}^{N}\langle U_i, e_j\rangle^2\\
&\qquad-\left(\frac{\sqrt{b}}{N}\left[\sum_{i=1}^{n}\langle X_i, e_j\rangle-\sum_{l=1}^{b-1}\left(1-\frac{l}{b}\right)[\langle X_l, e_j\rangle+\langle X_{n-l+1}, e_j\rangle]\right]\right)^2.
\end{aligned} \tag{3.5.33}$$

By Parseval's identity, we have,

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

Hence,

$$\begin{aligned}
&\sum_{j=1}^{\infty}\frac{1}{N}\sum_{i=1}^{N}|\langle U_i, e_j\rangle|^2\\
&\quad=\frac{1}{N}\sum_{i=1}^{n}\langle X_i, X_i\rangle\\
&\qquad+\sum_{h=1}^{b-1}\left(1-\frac{h}{b}\right)\frac{1}{N}\sum_{i=1}^{n-h}[\langle X_i, X_{i+h}\rangle+\langle X_{i+h}, X_i\rangle]+O_P(b^2/n).
\end{aligned} \tag{3.5.34}$$

Then, by letting $g_b(h) = \left(1 - \frac{|h|}{b}\right)$ in Lemma 3.5.1, we get that, as $n \to \infty$,

$$\sum_{j=1}^{\infty} \frac{1}{N} \sum_{i=1}^{N} \langle U_i, e_j \rangle^2 \xrightarrow{P} \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle). \tag{3.5.35}$$

For the second term of equation (3.5.33), we show that,

$$\sum_{j \geq 1} \left(\frac{\sqrt{b}}{N} \left[\sum_{i=1}^{n} \langle X_i, e_j \rangle - \sum_{l=1}^{b-1} \left(1 - \frac{l}{b}\right) [\langle X_l, e_j \rangle + \langle X_{n-l+1}, y \rangle] \right] \right)^2 = o_P(1), \tag{3.5.36}$$

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

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

Now note that, by the continuous mapping theorem and using Theorem 1 of Horváth *et al.* (2013), we get

$$\frac{1}{n} \sum_{i=1}^{n} \sum_{l=1}^{n} \langle X_i, X_l \rangle = \langle \sqrt{n} \overline{X}_n, \sqrt{n} \overline{X}_n, \rangle = O_p(1). \tag{3.5.37}$$

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 (3.5.36). Hence, from (3.5.33), (3.5.35) and (3.5.36), we conclude that

$$\sum_{j \geq 1} \mathrm{Var}^*(\langle U_1^*, e_j \rangle) \to \sum_{i=-\infty}^{\infty} \mathbb{E}(\langle X_0, X_i \rangle), \quad \text{in probability.} \tag{3.5.38}$$

Therefore, and by (3.5.32), property (d) is proved if we show that,

$$\sum_{j\geq 1}\lambda_j = \sum_{i=-\infty}^{\infty}\mathbb{E}(\langle X_0, X_i\rangle). \tag{3.5.39}$$

Using Mercer's theorem, we have

$$\begin{aligned}\sum_{j\geq 1}\lambda_j = \int c(u,u)\mathrm{d}u = \int \sum_{i=-\infty}^{\infty}\mathbb{E}[X_0(u)X_i(u)]\mathrm{d}u \\ = \sum_{i=-\infty}^{\infty}\mathbb{E}\int [X_0(u)X_i(u)]\mathrm{d}u = \sum_{i=-\infty}^{\infty}\mathbb{E}\langle X_0, X_i\rangle. \end{aligned} \tag{3.5.40}$$

Notice that the above interchange of summation and integration is justified since, using Assumption 2, and the fact that X_0 and $X_{i,i}$ are independent for $i\geq 1$, we get

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

To prove (e), notice first that, by (3.5.33),

$$\sum_{j=1}^{\infty}\mathrm{Var}^*(\langle U_1^*, e_j\rangle) \leq \sum_{j=1}^{\infty}(1/N)\sum_{i=1}^{N}|\langle U_i, e_j\rangle|^2$$

and, therefore, using (3.5.34), for any given $n\geq 1$, $\sum_{j=1}^{\infty}\mathrm{Var}^*(\langle U_1^*, e_j\rangle)$ is bounded in probability. Furthermore, by (3.5.38), the sequence $\{\sum_{j=1}^{\infty}\mathrm{Var}^*(\langle U_1^*, e_j\rangle), n\geq 1\}$ converges in probability as $n\to\infty$.

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

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

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

$$d(u,v) = \mathbb{E}^*[U_1^*(u)U_1^*(v)] - \mathbb{E}^*(U_1^*(u))\mathbb{E}^*(U_1^*(v)). \tag{3.5.41}$$

Now,

$$\begin{aligned}
&\mathbb{E}^*[U_1^*(u)U_1^*(v)] \\
&\quad = \frac{1}{N}\sum_{i=1}^{n} X_i(u)X_i(v) + \sum_{h=1}^{b-1}\left(1-\frac{h}{b}\right)\frac{1}{N}\sum_{i=1}^{n-h}[X_i(u)X_{i+h}(v) + X_{i+h}(u)X_i(v)] \\
&\quad - \frac{1}{N}\sum_{s=1}^{b-1}\left(1-\frac{s}{b}\right)[X_s(u)X_s(v) + X_{n-s+1}(u)X_{n-s+1}(v)] \\
&\quad - \frac{1}{N}\sum_{t=1}^{b-1}\sum_{j=1}^{b-t}\left(1-\frac{j+t}{b}\right)[X_j(u)X_{j+t}(v) + X_{n-j+1-t}(u)X_{n-j+1-t}(v) \\
&\qquad\qquad + X_{j+t}(u)X_j(v) + X_{n-j+1}(u)X_{n-j+1-t}(v)]
\end{aligned} \tag{3.5.42}$$

and

$$\mathbb{E}^*(U_1^*(u)) = \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{n} X_i(u) - \sum_{j=1}^{b-1}\left(1-\frac{j}{b}\right)(X_j(u) + X_{n-j+1}(u))\right]. \tag{3.5.43}$$

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 (3.5.41) and $c_N(u,v)$ given in (3.5.25). Now, notice that $2\pi\mathcal{F}_0(y)(v) = \int \sum_{h=-\infty}^{\infty} \mathbb{E}[X_0(u)X_h(v)]y(u)\mathrm{d}u$, i.e., $2\pi\mathcal{F}_0$ is an integral operator with kernel $c(u,v) = \sum_{h=-\infty}^{\infty} \mathbb{E}[X_0(u)X_h(v)]$. Hence,

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

Using (3.5.26) it suffices to prove that $\iint [R(u,v)]^2\mathrm{d}u\mathrm{d}v = o_p(1)$. 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 (3.5.37),

$$\frac{b^2}{N^4}\iint\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_i(u)X_j(v)\right)^2 \mathrm{d}u\mathrm{d}v$$

$$= \frac{b^2}{N^2} \frac{1}{N} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \int X_{i_1}(u) X_{i_2}(u) \mathrm{d}u \frac{1}{N} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \int X_{j_1}(v) X_{j_2}(v) \mathrm{d}v$$

$$= \frac{b^2}{N^2} \left(\frac{1}{N} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \langle X_{i_1}, X_{i_2} \rangle \right)^2 = O_P(b^2/N^2) = o_p(1). \tag{3.5.44}$$

Furthermore,

$$\iint \left[\frac{1}{N} \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} \left(1 - \frac{j+t}{b} \right) X_j(u) X_{j+t}(v) \right]^2 \mathrm{d}u \mathrm{d}v$$

$$\leq \frac{1}{N^2} b^2 \iint \sum_{t=1}^{b-1} \sum_{j=1}^{b-t} X_j^2(u) X_{j+t}^2(v) \mathrm{d}u \mathrm{d}v = O_P(b^4/N^2) = o_p(1), \tag{3.5.45}$$

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

Proof of Theorem 3.2.2. Let $S_n^* = \sqrt{n}(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*))$ and, as in Theorem 3.2.1, we have that $S_n^* = k^{-1/2} \sum_{i=1}^{k} [U_i^* - \mathbb{E}^*(U_i^*)]$, where $U_i^* = b^{-1/2}(X_{(i-1)b+1}^* + X_{(i-1)b+2}^* + \ldots + X_{ib}^*)$, $i = 1, 2, \ldots, k$, are i.i.d. random variables, $\langle S_n^*, y \rangle = k^{-1/2} \sum_{i=1}^{k} [W_i^* - \mathbb{E}^*(W_i^*)]$ with $W_i^* = \langle U_i^*, y \rangle$, $i = 1, 2, \ldots, k$, and $\mu^* = \mathbb{E}^*(W_1^*)$. Let C be the covariance operator with kernel

$$c(u,v) = \mathbb{E}[Y_0(u)Y_0(v)] + \sum_{h \geq 1} \mathbb{E}[Y_0(u)Y_h(v)] + \sum_{h \geq 1} \mathbb{E}[Y_0(v)Y_h(u)], \quad u, v \in [0,1]^2,$$

$N = n - b + 1$, $\|w_b\|_1 = \sum_{i=1}^{b} w_b(t)$ and $\|w_b\|_2^2 = \sum_{t=1}^{b} w_b^2(t)$. Finally, let $X_i = Y_i - \overline{Y}_n$, $i = 1, 2, \ldots, n$, and

$$U_i = \frac{1}{\|w_b\|_2} \left(w_b(1) X_i + w_b(2) X_{i+1}, \ldots + w_b(b) X_{i+b-1} \right), \quad i = 1, 2, \ldots, N.$$

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.

To prove (L1), we establish that, as $n \to \infty$,

$$\mathrm{Var}^* \left(\langle S_n^*, y \rangle \right) \xrightarrow{P} \sigma^2(y) \tag{3.5.46}$$

and that

$$\frac{\langle S_n^*, y\rangle}{\sqrt{\mathrm{Var}^*(\langle S_n^*, y\rangle)}} \xrightarrow{d} N(0,1). \tag{3.5.47}$$

To see (3.5.46), note first that

$$\mathrm{Var}^*\left(\langle S_n^*, y\rangle\right) = k^{-1}\sum_{i=1}^{k} \mathrm{Var}^*(W_i^* - \mathbb{E}^*(W_i^*)) = \mathrm{Var}^*(W_1^*)$$

and that

$$\begin{aligned}
\mathrm{Var}^*(W_1^*) &= \frac{1}{N}\sum_{i=1}^{N}\left[\langle U_i, y\rangle - \frac{1}{N}\sum_{j=1}^{N}\langle U_j, y\rangle\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\|w_b\|_2}\sum_{t=1}^{b} w_b(t)\langle Y_{i+t-1}, y\rangle\right]^2 \\
&\qquad - \left[\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle Y_{j+s-1}, y\rangle\right]^2 .
\end{aligned} \tag{3.5.48}$$

We next show that

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle Y_{i+s-1}, y\rangle = O_p(\frac{b}{\sqrt{n}}). \tag{3.5.49}$$

Toward this, note that

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle Y_{i+s-1}, y\rangle \\
&\quad = \frac{1}{N}\frac{\|w_b\|_1}{\|w_b\|_2}\left[\sum_{i=1}^{n}\langle Y_i, y\rangle - \sum_{j=1}^{b-1}\left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right)\langle Y_j, y\rangle\right. \\
&\qquad\qquad \left. - \sum_{j=1}^{b-1}\left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_{n-j+1}, y\rangle\right],
\end{aligned} \tag{3.5.50}$$

and that

$$\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle Y_{i+s-1}, y\rangle\right] = 0. \tag{3.5.51}$$

Furthermore, using the decomposition

$$\left[\sum_{i=1}^{n}\langle Y_i, y\rangle - \sum_{j=1}^{b-1}\left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right)\langle Y_j, y\rangle\right.$$

$$
- \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right) \langle Y_{n-j+1}, y\rangle \Bigg]^2
$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle Y_i, y\rangle \langle Y_j, y\rangle \\
&+ \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{t=1}^{i} w_b(t)}{\|w_b\|_1}\right) \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right) \langle Y_i, y\rangle \langle Y_j, y\rangle \\
&+ \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=b-i+1}^{b} w_b(s)}{\|w_b\|_1}\right) \left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right) \langle Y_{n-i+1}, y\rangle \langle Y_{n-j+1}, y\rangle \\
&- 2 \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right) \left(1 - \frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1}\right) \langle Y_{n-i+1}, y\rangle \langle Y_j, y\rangle \\
&- 2 \sum_{i=1}^{n} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right) \langle Y_i, y\rangle \langle Y_j, y\rangle \\
&\quad - 2 \sum_{i=1}^{n} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right) \langle Y_i, y\rangle \langle Y_{n-j+1}, y\rangle, \qquad (3.5.52)
\end{aligned}
$$

we get, by equation (3.5.50), the fact that $\|w_b\|_2 = O(b^{1/2})$, $\|w_b\|_1 = O(b)$ and the same arguments as those used to obtain equation (3.5.21), that

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N} \frac{1}{\|w_b\|_2} \sum_{s=1}^{b} w_b(s) \langle Y_{i+s-1}, y\rangle\right]^2 &= \frac{\|w_b\|_1^2}{N^2 \|w_b\|_2^2} \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}[\langle Y_i, y\rangle \langle Y_j, y\rangle] + O(b^2/n) \\
&= O(b/n) + O(b^2/n) = O(b^2/n). \qquad (3.5.53)
\end{aligned}
$$

From (3.5.51) and (3.5.53), assertion (3.5.49) follows. Consider next the first term of the right hand side of equation (3.5.48). For this, we have

$$
\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N} \left[\frac{1}{\|w_b\|_2} \sum_{t=1}^{b} w_b(t) \langle Y_{i+t-1}, y\rangle\right]^2 \\
&= \frac{1}{N} \frac{1}{\|w_b\|_2^2} \Bigg\{ \sum_{i=1}^{n} \|w_b\|_2^2 \langle Y_i, y\rangle \langle Y_i, y\rangle + \sum_{h=1}^{b-1} \mathcal{W}_h \sum_{i=1}^{n-h} [\langle Y_i, y\rangle \langle Y_{i+h}, y\rangle + \langle Y_{i+h}, y\rangle \langle Y_i, y\rangle] \\
&\quad - \sum_{s=1}^{b-1} \left(\|w_b\|_2^2 - \sum_{t=1}^{s} w_b^2(t)\right) \langle Y_s, y\rangle \langle Y_s, y\rangle \\
&\quad - \sum_{s=1}^{b-1} \left(\|w_b\|_2^2 - \sum_{t=b-s+1}^{b} w_b^2(t)\right) \langle Y_{n-s+1}, y\rangle \langle Y_{n-s+1}, y\rangle \\
&\quad - \sum_{h=1}^{b-1} \sum_{i=1}^{b-h} \left(\mathcal{W}_h - \sum_{t=1}^{i} w_b(t) w_b(t+h)\right) [\langle Y_i, y\rangle \langle Y_{i+h}, y\rangle + \langle Y_{i+h}, y\rangle \langle Y_i, y\rangle]
\end{aligned}
$$

$$- \sum_{h=1}^{b-1} \sum_{i=1}^{b-h} \left(\mathcal{W}_h - \sum_{t=b-i-h+1}^{b-h} w_b(t) w_b(t+h) \right) [\langle Y_{n-i+1}, y \rangle \langle Y_{n-i+1-h}, y \rangle + \langle Y_{n-i+1-h}, y \rangle \langle Y_{n-i+1}, y \rangle] \Bigg\},$$

from which it follows that

$$\frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{\|w_b\|_2} \sum_{t=1}^{b} w_b(t) \langle Y_{i+t-1}, y \rangle \right]^2 = \frac{1}{N} \sum_{i=1}^{n} \langle Y_i, y \rangle \langle Y_i, y \rangle + \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \frac{1}{N} \sum_{i=1}^{n-h} [\langle Y_i, y \rangle \langle Y_{i+h}, y \rangle + \langle Y_{i+h}, y \rangle \langle Y_i, y \rangle] + O_p(b/n) + O_p(b^2/n).$$

Hence, using expressions (3.5.48) and (3.5.49), we get,

$$\mathrm{Var}^*(W_1^*) = \iint \tilde{c}_N(u,v) y(u) y(v) \mathrm{d}u \mathrm{d}v + O_p(b^2/n), \tag{3.5.54}$$

where

$$\tilde{c}_N(u,v) = \frac{1}{N} \sum_{i=1}^{n} Y_i(u) Y_i(v) + \sum_{h=1}^{b-1} \frac{\mathcal{W}_h}{\|w_b\|_2^2} \frac{1}{N} \sum_{i=1}^{n-h} [Y_i(u) Y_{i+h}(v) + Y_{i+h}(u) Y_i(v)]. \tag{3.5.55}$$

Using Lemma 3.5.2 (ii) and Cauchy-Schwarz's inequality, we conclude that, as $n \to \infty$,

$$\left| \iint (\tilde{c}_n(u,v) - c(u,v)) y(u) y(v) \mathrm{d}u \mathrm{d}v \right| \leq \left(\iint \{\tilde{c}_n(u,v) - c(u,v)\}^2 \mathrm{d}u \mathrm{d}v \right)^{1/2} \|y\|^2 = o_P(1). \tag{3.5.56}$$

where $\tilde{c}_n(u,v) = (N/n) \tilde{c}_N(u,v)$. Thus,

$$\iint \tilde{c}_n(u,v) y(u) y(v) \mathrm{d}u \mathrm{d}v \xrightarrow{P} \iint c(u,v) y(u) y(v) \mathrm{d}u \mathrm{d}v$$

and, using equation (3.5.54),

$$\mathrm{Var}^* \langle S_n^*, y \rangle = \frac{n}{N} \iint c_n(u,v) y(u) y(v) \mathrm{d}u \mathrm{d}v + O_p(b^2/n) \xrightarrow{P} \iint c(u,v) y(u) y(v) \mathrm{d}u \mathrm{d}v = \sigma^2(y), \tag{3.5.57}$$

as $n \to \infty$. To prove (3.5.47), as stated in the proof of Theorem 3.2.1, we must establish Lindeberg's condition.

For this, let $W_i = \langle U_i, y\rangle$, $i = 1, 2, \ldots, n$, and note that, by (3.5.48), we have

$$\begin{aligned}
W_i - \mu^* &= \langle U_i, y\rangle - \frac{1}{N}\sum_{j=1}^{N}\langle U_j, y\rangle \\
&= \frac{1}{\|w_b\|_2}\sum_{t=1}^{b} w_b(t)\langle X_{i+t-1}, y\rangle - \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle X_{j+s-1}, y\rangle \\
&= \frac{1}{\|w_b\|_2}\sum_{t=1}^{b} w_b(t)\langle Y_{i+t-1}, y\rangle - \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\|w_b\|_2}\sum_{s=1}^{b} w_b(s)\langle Y_{j+s-1}, y\rangle \\
&= W_i^Y - \frac{1}{N}\sum_{j=1}^{N} W_j^Y = W_i^Y - \mu_Y^*, \qquad (3.5.58)
\end{aligned}$$

with an obvious notation for W_i^Y and μ_Y^*. Hence, using (3.5.30) and Markov's inequality, we have, for any $\delta > 0$ and for any $\varepsilon > 0$, that

$$\begin{aligned}
&P\left(\frac{1}{k}\sum_{t=1}^{k}\mathbb{E}^*\left[(W_t^* - \mu^*)^2 \mathbb{1}(|W_t^* - \mu^*| > \varepsilon\tau_k^*)\right] > \delta\right) \\
&\leq \delta^{-1}\mathbb{E}\left\{\mathbb{E}^*\left[(W_1^* - \mu^*)^2 \mathbb{1}(|W_1^* - \mu^*| > \varepsilon\tau_k^*)\right]\right\} \\
&= \delta^{-1}\mathbb{E}\left[(W_1^Y - \mu_Y^*)^2 \mathbb{1}(|W_1^Y - \mu_Y^*| > \varepsilon\tau_k^*)\right] \\
&\leq 4\delta^{-1}\left[\mathbb{E}(W_1^Y)^2 \mathbb{1}(|W_1^Y| > \varepsilon\tau_k^*/2) + \mathbb{E}(\mu_Y^*)^2 \mathbb{1}(|\mu_Y^*| > \varepsilon\tau_k^*/2)\right] \\
&\leq 4\delta^{-1}\left[\mathbb{E}(W_1^Y)^2 \mathbb{1}(|W_1^Y| > \varepsilon\tau_k^*/2) + \mathbb{E}(\mu_Y^*)^2\right]. \qquad (3.5.59)
\end{aligned}$$

Since $\mathbb{E}(W_1^Y)^2 = \sum_{|h|<b}\left(\frac{\mathcal{W}_{|h|}}{\|w_b\|_2^2}\right) E[\langle Y_0, y\rangle\langle Y_h, y\rangle]$, we get, by Lemma 3.5.2 (i),

$$\mathbb{E}(W_1^Y)^2 \xrightarrow{P} \iint c(u,v)y(u)y(v)\mathrm{d}u\mathrm{d}v,$$

and, by the dominated convergence theorem, that $\lim_{n\to\infty}\mathbb{E}(W_1^Y)^2\mathbb{1}(|W_1^Y| > \varepsilon\tau_k^*/2) = 0$. Using this result and expression (3.5.53), it follows that the bound in (3.5.59) converges to 0 as $n \to \infty$, which establishes Lindeberg's condition.

Consider now ($L2$). For this, it suffices to verify that conditions (a)-(e) of the proof of Theorem 3.2.1 are satisfied. Note that, by letting $y = e_j$ in expression (3.5.57), property (b) follows with $\Sigma_j = \iint c(u,v)e_j(u)e_j(v)\mathrm{d}u\mathrm{d}v$. To prove (c), note that, by Proposition 6 of Hörmann *et al.* (2015), since the stochastic process $\{Y_t, t \in \mathbb{Z}\}$ is L^2-m-approximable, the covariance operator C with kernel $c(\cdot,\cdot)$ is trace-class. Therefore,

$\sum_{j\geq 1}\Sigma_j = \sum_{j\geq 1}\iint c(u,v)e_j(u)e_j(v)\mathrm{d}u\mathrm{d}v = \sum_{j\geq 1}\lambda_j < \infty$, where λ_j, $j \geq 1$ are the eigenvalues of the covariance operator C. To establish (d), let first

$$U_i^Y = \frac{1}{\|w_b\|_2}\left(w_b(1)Y_i + w_b(2)Y_{i+1}, \ldots + w_b(b)Y_{i+b-1}\right), \quad i = 1, 2, \ldots, N.$$

Then, using equation (3.5.48), we have

$$\mathrm{Var}^*(\langle U_1^*, e_j\rangle) = \frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle^2 - \left[\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle\right]^2. \tag{3.5.60}$$

From expressions (3.5.50) and (3.5.52), we get,

$$\begin{aligned}
&\sum_{j\geq 1}\left[\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle\right]^2 \\
&= \frac{1}{N^2}\frac{\|w_b\|_1^2}{\|w_b\|_2^2}\Bigg[\sum_{j\geq 1}\sum_{i=1}^{n}\sum_{t=1}^{n}\langle Y_i, e_j\rangle\langle Y_t, e_j\rangle \\
&\quad + \sum_{j\geq 1}\sum_{i=1}^{b-1}\sum_{t=1}^{b-1}\left(1 - \frac{\sum_{s=1}^{i} w_b(s)}{\|w_b\|_1}\right)\left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right)\langle Y_i, e_j\rangle\langle Y_s, e_j\rangle \\
&\quad + \sum_{j\geq 1}\sum_{i=1}^{b-1}\sum_{s=1}^{b-1}\left(1 - \frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1}\right) \\
&\qquad\qquad \cdot\left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_{n-i+1}, e_j\rangle\langle Y_{n-s+1}, e_j\rangle \\
&\quad - 2\sum_{j\geq 1}\sum_{i=1}^{b-1}\sum_{s=1}^{b-1}\left(1 - \frac{\sum_{t=1}^{j} w_b(t)}{\|w_b\|_1}\right) \\
&\qquad\qquad \cdot\left(1 - \frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_{n-i+1}, e_j\rangle\langle Y_s, e_j\rangle \\
&\quad - 2\sum_{j\geq 1}\sum_{i=1}^{n}\sum_{s=1}^{b-1}\left(1 - \frac{\sum_{t=1}^{j} w_b(t)}{\|w_b\|_1}\right)\langle Y_i, e_j\rangle\langle Y_s, e_j\rangle \\
&\quad - 2\sum_{j\geq 1}\sum_{i=1}^{n}\sum_{s=1}^{b-1}\left(1 - \frac{\sum_{t=b-s+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_i, e_j\rangle\langle Y_{n-s+1}, e_j\rangle\Bigg].
\end{aligned} \tag{3.5.61}$$

Hence, and because $\langle x, y\rangle = \sum_{j\geq 1}\langle x, e_j\rangle\langle y, e_j\rangle$,

$$\begin{aligned}
&\sum_{j\geq 1}\left[\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle\right]^2 \\
&= \frac{1}{N^2}\frac{\|w_b\|_1^2}{\|w_b\|_2^2}\Bigg\{\sum_{i=1}^{n}\sum_{t=1}^{n}\langle Y_i, Y_t\rangle
\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=1}^{b-1}\sum_{t=1}^{b-1}\left(1-\frac{\sum_{s=1}^{i} w_b(s)}{\|w_b\|_1}\right)\left(1-\frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1}\right)\langle Y_i, Y_s\rangle \\
&+\sum_{i=1}^{b-1}\sum_{s=1}^{b-1}\left(1-\frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1}\right)\left(1-\frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_{n-i+1}, Y_{n-s+1}\rangle \\
&-2\sum_{i=1}^{b-1}\sum_{s=1}^{b-1}\left(1-\frac{\sum_{t=1}^{j} w_b(t)}{\|w_b\|_1}\right)\left(1-\frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_{n-i+1}, Y_s\rangle \\
&-2\sum_{i=1}^{n}\sum_{s=1}^{b-1}\left(1-\frac{\sum_{t=1}^{j} w_b(t)}{\|w_b\|_1}\right)\langle Y_i, Y_s\rangle \\
&-2\sum_{i=1}^{n}\sum_{s=1}^{b-1}\left(1-\frac{\sum_{t=b-s+1}^{b} w_b(t)}{\|w_b\|_1}\right)\langle Y_i, Y_{n-s+1}\rangle\Bigg\}.
\end{aligned}
$$

Therefore, by using (3.5.37), we get

$$
\begin{aligned}
\sum_{j\geq 1}\left[\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle\right]^2 &= \frac{1}{N^2}\frac{\|w_b\|_1^2}{\|w_b\|_2^2}\sum_{i=1}^{n}\sum_{t=1}^{n}\langle Y_i, Y_t\rangle + O_P(b^2/n) \\
&= O_p(b^2/n) = o_p(1). \qquad (3.5.62)
\end{aligned}
$$

Consider now, the first term of the right hand side of expression (3.5.60). By Parseval's identity,

$$
\begin{aligned}
&\sum_{j\geq 1}\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\|U_i^Y\|^2 \\
&= \frac{1}{N}\frac{1}{\|w_b\|_2^2}\Bigg\{\sum_{i=1}^{n}\|w_b\|_2^2\langle Y_i, Y_i\rangle + \sum_{h=1}^{b-1}\mathcal{W}_h\sum_{i=1}^{n-h}[\langle Y_i, Y_{i+h}\rangle + \langle Y_{i+h}, Y_i\rangle] \\
&\quad -\sum_{s=1}^{b-1}\left(\|w_b\|_2^2 - \sum_{t=1}^{s} w_b^2(t)\right)\langle Y_s, Y_s\rangle \\
&\quad -\sum_{s=1}^{b-1}\left(\|w_b\|_2^2 - \sum_{t=b-s+1}^{b} w_b^2(t)\right)\langle Y_{n-s+1}, Y_{n-s+1}\rangle \\
&\quad -\sum_{h=1}^{b-1}\sum_{i=1}^{b-h}\left(\mathcal{W}_h - \sum_{t=1}^{i} w_b(t)w_b(t+h)\right)[\langle Y_i, Y_{i+h}\rangle + \langle Y_{i+h}, Y_i\rangle] \\
&\quad -\sum_{h=1}^{b-1}\sum_{i=1}^{b-h}\left(\mathcal{W}_h - \sum_{t=b-i-h+1}^{b-h} w_b(t)w_b(t+h)\right)[\langle Y_{n-i+1}, Y_{n-i+1-h}\rangle \\
&\qquad\qquad + \langle Y_{n-i+1-h}, Y_{n-i+1}\rangle]\Bigg\}.
\end{aligned}
$$

Hence,

$$\sum_{j\geq 1}\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle^2 = \frac{1}{N}\sum_{i=1}^{n}\langle Y_i, Y_i\rangle$$
$$+\sum_{h=1}^{b-1}\frac{\mathcal{W}_h}{\|w_b\|_2^2}\frac{1}{N}\sum_{i=1}^{n-h}[\langle Y_i, Y_{i+h}\rangle + \langle Y_{i+h}, Y_i\rangle] + O_P(b^2/n), \quad (3.5.63)$$

and because $N/n \to 1$ as $n \to \infty$ and taking $g_b(h) = \dfrac{\mathcal{W}_{|h|}}{\mathcal{W}_0}$ in Lemma 3.5.1, in conjunction with expressions (3.5.11) and (3.5.12), we get, as $n \to \infty$, that

$$\sum_{j\geq 1}\frac{1}{N}\sum_{i=1}^{N}\langle U_i^Y, e_j\rangle^2 \xrightarrow{P} \sum_{i=-\infty}^{\infty}\mathbb{E}(\langle Y_0, Y_i\rangle).$$

Thus, using (3.5.60) and (3.5.62), we conclude that

$$\sum_{j\geq 1}\mathrm{Var}^*(\langle U_1^*, e_j\rangle) \xrightarrow{P} \sum_{i=-\infty}^{\infty}\mathbb{E}(\langle Y_0, Y_i\rangle) \quad (3.5.64)$$

and, using $\sum_{i=-\infty}^{\infty}\mathbb{E}(\langle Y_0, Y_i\rangle) = \sum_{j\geq 1}\lambda_j$, property (d) is established. Finally, (e) is proved using the same arguments as in the corresponding case in Theorem 3.2.1, and taking into account expressions (3.5.60), (3.5.63) and (3.5.64).

Consider next assertion (ii) of the theorem. It suffices to prove that, as $n \to \infty$, $\|n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) - 2\pi\mathcal{F}_0\|_{HS} = o_P(1)$. Notice that $n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*))$ is an integral operator with kernel

$$\tilde{d}(u,v) = \mathbb{E}^*[U_1^*(u) - \mathbb{E}^*(U_1^*(u))][U_1^*(v) - \mathbb{E}^*(U_1^*(v))]$$
$$= \frac{1}{N}\sum_{i=1}^{N}U_i^Y(u)U_i^Y(v) - \left(\frac{1}{N}\sum_{j=1}^{N}U_j^Y(u)\right)\left(\frac{1}{N}\sum_{j=1}^{N}U_j^Y(v)\right). \quad (3.5.65)$$

Now,

$$\frac{1}{N}\sum_{i=1}^{N}U_i^Y(u)U_i^Y(v)$$
$$= \frac{1}{N}\frac{1}{\|w_b\|_2^2}\Bigg\{\sum_{i=1}^{n}\|w_b\|_2^2 Y_i(u)Y_i(v) + \sum_{h=1}^{b-1}\mathcal{W}_h\sum_{i=1}^{n-h}[Y_i(u)Y_{i+h}(v) + Y_{i+h}(u)Y_i(v)]$$
$$-\sum_{s=1}^{b-1}\left(\|w_b\|_2^2 - \sum_{t=1}^{s}w_b^2(t)\right)Y_s(u)Y_s(v)$$

$$
\begin{aligned}
&- \sum_{s=1}^{b-1} \left(\|w_b\|_2^2 - \sum_{t=b-s+1}^{b} w_b^2(t) \right) Y_{n-s+1}(u) Y_{n-s+1}(v) \\
&- \sum_{h=1}^{b-1} \sum_{i=1}^{b-h} \left(\mathcal{W}_h - \sum_{t=1}^{i} w_b(t) w_b(t+h) \right) [Y_i(u) Y_{i+h}(v) + Y_{i+h}(u) Y_i(v)] \\
&- \sum_{h=1}^{b-1} \sum_{i=1}^{b-h} \left(\mathcal{W}_h - \sum_{t=b-i-h+1}^{b-h} w_b(t) w_b(t+h) \right) [Y_{n-i+1}(u) Y_{n-i+1-h}(v) \\
&\qquad\qquad\qquad\qquad\qquad\qquad + Y_{n-i+1-h}(v) Y_{n-i+1}(u)] \Bigg\}
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{N} \sum_{i=1}^{N} U_i^Y(u) \frac{1}{N} \sum_{j=1}^{N} U_j^Y(v) \\
&= \frac{1}{N^2} \frac{\|w_b\|_1^2}{\|w_b\|_2^2} \Bigg\{ \sum_{i=1}^{n} \sum_{j=1}^{n} Y_i(u) Y_j(v) \\
&+ \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{t=1}^{i} w_b(t)}{\|w_b\|_1} \right) \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1} \right) Y_i(u) Y_j(v) \\
&+ \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=b-i+1}^{b} w_b(s)}{\|w_b\|_1} \right) \left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1} \right) Y_{n-i+1}(u) Y_{n-j+1}(v) \\
&- \sum_{i=1}^{b-1} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1} \right) \\
&\qquad \cdot \left(1 - \frac{\sum_{t=b-i+1}^{b} w_b(t)}{\|w_b\|_1} \right) [Y_{n-i+1}(u) Y_j(v) + Y_{n-i+1}(v) Y_j(u)] \\
&- \sum_{i=1}^{n} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{s=1}^{j} w_b(s)}{\|w_b\|_1} \right) [Y_i(u) Y_j(v) + Y_j(u) Y_i(v)] \\
&- \sum_{i=1}^{n} \sum_{j=1}^{b-1} \left(1 - \frac{\sum_{t=b-j+1}^{b} w_b(t)}{\|w_b\|_1} \right) [Y_i(u) Y_{n-j+1}(v) + Y_i(v) Y_{n-j+1}(u)] \Bigg\}.
\end{aligned}
$$

Therefore, $\tilde{d}(u,v) = \tilde{c}_N(u,v) + \tilde{R}(u,v)$ where $\tilde{c}_N(u,v)$ is defined in (3.5.55) and $\tilde{R}(u,v)$ is the remainder term, and

$$
\begin{aligned}
&\| n\mathbb{E}^*(\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) \otimes (\overline{X}_n^* - \mathbb{E}^*(\overline{X}_n^*)) - 2\pi \mathcal{F}_0 \|_{HS}^2 \\
&\qquad\qquad \leq 2 \iint [\tilde{c}_N(u,v) - c(u,v)]^2 \mathrm{d}u\mathrm{d}v + 2 \iint [\tilde{R}_N(u,v))]^2 \mathrm{d}u\mathrm{d}v.
\end{aligned}
$$

Using similar arguments as those used in the proof of assertion (ii) of Theorem 3.2.1, it follows that $\iint [\tilde{R}(u,v))]^2 \mathrm{d}u\mathrm{d}v = o_p(1)$, from which assertion (ii) follows because of

(3.5.56).

Proof of Theorem 3.3.1. Consider assertion (i). For $i = 1, 2$, let $\{e^*_{i,j},\ j = 1, 2, \ldots, n_i\}$ be the pseudo-observations generated by implementing the MBB procedure at $\{\varepsilon_{i,j},\ j = 1, 2, \ldots, n_i\}$. Using Theorem 3.2.1, it follows that, conditionally on $\mathbf{X_M}$, for $i = 1, 2$, and as $n_1, n_2 \to \infty$,

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

where Γ_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 $\varepsilon^*_{i,\xi+sb}(\tau)$, $\xi = 1, 2, \ldots, b$, $s = 0, 1, \ldots, q_i$, $\tau \in \mathcal{I}$, are generated by first applying the MBB procedure to $\hat{\varepsilon}_{i,\xi+sb}(\tau)$, $\xi = 1, 2, \ldots, b$, $s = 0, 1, \ldots, q_i$, $\tau \in \mathcal{I}$ and then $\overline{\varepsilon}_{i,\xi}(\tau)$ is subtracted. Note further that $\varepsilon_{i,j}(\tau) = \hat{\varepsilon}_{i,j}(\tau) + \overline{X}_{i,n_i} - \mu_i(\tau)$. Thus, $e^*_{i,\xi+sb}(\tau) = \varepsilon^*_{i,\xi+sb}(\tau) + \overline{\varepsilon}_{i,\xi}(\tau) + \overline{X}_{i,n_i}(\tau) - \mu_i(\tau)$ and, using expression (3.3.2), we get

$$\frac{1}{\sqrt{n_i}}\sum_{j=1}^{n_i}(e^*_{i,j} - \mathbb{E}^*(e^*_{i,j})) = \frac{1}{\sqrt{n_i}}\sum_{j=1}^{n_i}(X^*_{i,j} - \mathbb{E}^*(X^*_{i,j})) = \frac{1}{\sqrt{n_i}}\sum_{j=1}^{n_i}(X^*_{i,j} - \overline{X}_M).$$

Therefore, and conditionally on $\mathbf{X_M}$, as $n_1, n_2 \to \infty$,

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

where Γ_1 and Γ_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

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

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

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

where $\Gamma = \sqrt{1-\theta}\Gamma_1 - \sqrt{\theta}\Gamma_2$. The proof of assertion (ii) follows along the same lines using Theorem 3.2.2. This completes the proof of the theorem.

4

Testing equality of autocovariance operators for functional time series

Abstract

We consider strictly stationary stochastic processes of Hilbert space-valued random variables and focus on tests of the equality of the lag-zero autocovariance operators of several independent functional time series. A moving block bootstrap-based testing procedure is proposed which generates pseudo random elements that satisfy the null hypothesis of interest. It is based on directly bootstrapping the time series of tensor products which overcomes some common difficulties associated with applications of the bootstrap to related testing problems. The suggested methodology can be potentially applied to a broad range of test statistics of the hypotheses of interest. As an example, we establish validity for approximating the distribution under the null of a fully functional test statistic based on the Hilbert-Schmidt distance of the corresponding sample lag-zero autocovariance operators, and show consistency under the alternative. As a prerequisite, we prove a central limit theorem for the moving block bootstrap procedure applied to the sample autocovariance operator which is of interest on its own. The finite sample size and power performance of the suggested moving block bootstrap-based testing procedure is illustrated through simulations and an application to a real-life data set is discussed.

4.1 Introduction

Functional data analysis deals with random variables which are curves or images and can be expressed as functions in appropriate spaces. In this paper, we consider functional time series $\mathbb{X}_n = \{X_1, X_2, \ldots, X_n\}$ steming from a (strictly stationary) stochastic process $\mathbb{X} = (X_t,\, t \in \mathbb{Z})$ of Hilbert space-valued random functions $X_t(\tau)$, $\tau \in \mathcal{I}$, which are assumed to be L^4-m-approximable, a dependence assumption which is satisfied by large classes of commonly used functional time series models; see, e.g., Hörmann and Kokoszka (2010). We would like to infer properties of a group of K independent functional processes based on observed stretches from each group. In particular, we focus on the problem of testing whether the lag-zero autocovariance operators of the K processes are equal and consider fully functional test statistics which evaluate the difference between the corresponding sample lag-zero autocovariance operators using appropriate distance measures.

As it is common in the statistical analysis of functional data, the limiting distribution of such statistics depends, in a complicate way, on difficult to estimate characteristics of the underlying functional stochastic processes like, for instance, its entire fourth order temporal dependence structure. Therefore, and in order to implement the testing approach proposed, we apply a moving block bootstrap (MBB) procedure which is used to estimate the distribution of the test statistic of interest under the null. Notice that for testing problems related to the equality of second order characteristics of several independent groups, in the finite or infinite dimensional setting, applications of the bootstrap to approximate the distribution of a test statistic of interest are commonly based on the generation of pseudo random observations obtained by resampling from the pooled (mixed) sample consisting of all available observations. Such implementations lead to the problem that the generated pseudo observations have not only identical second order characteristics but also identical distributions. This affects the power and the consistency properties of the bootstrap in that it restricts its validity to specific situations only; see Lele and Carlstein (1990) for an overview for the case of independent and identically distributed (i.i.d.) real-valued random variables and Remark 4.3.2 in Section 3 below for more details in the functional setting.

To overcome such problems, we use a different approach which is based on the observation that the lag-zero autocovariance operator $\mathcal{C}_0 = \mathbb{E}(X_t - \mu) \otimes (X_t - \mu)$ is the expected value of the tensor product process $\{\mathcal{Y}_t = (X_t - \mu) \otimes (X_t - \mu),\, t \in \mathbb{Z}\}$, where $\mu = \mathbb{E}X_t$ denotes the expectation of X_t. Therefore, the testing problem of interest can also be viewed as testing for the equality of expected values (mean functions) of the associated processes of tensor products. The suggested MBB procedure works by first generating functional pseudo random elements via resampling from the time series of tensor products of the same group and then adjusting the mean function of the generated pseudo random elements in each group so that the null hypothesis of interest is satisfied. We stress here the fact that the proposed method is not designed having any particular test statistic in mind and it is, therefore, potentially applicable to a wide range of test statistics. As an example, we establish validity of the proposed MBB-based testing procedure in estimating the distribution of a particular fully functional test statistic under the null, which is based on the Hilbert-Schmidt norm between the sample lag-zero autocovariance operators, and show its consistency under the alternative. As a prerequisite, we prove a central limit theorem for the MBB procedure applied to the sample version of the autocovariance operator $\mathcal{C}_h = \mathbb{E}(X_t - \mu) \otimes (X_{t+h} - \mu)$, $h \in \mathbb{Z}$, of an L^4-m-approximable stochastic process, which is of interest on its own. Our results imply that the suggested MBB-based testing procedure is not restricted to the case of testing for equality of the lag-zero autocovariance operator only but it can be adapted to tests dealing with the equality of any (finite number of) autocovariance operators $\mathcal{C}_h$ for lags h different from zero.

Asymptotic and bootstrap based inference procedures for covariance operators for two or more populations of i.i.d. functional data have been extensively discussed in the literature; see, e.g., Panaretos *et al.* (2010), Fremdt *et al.* (2013) for tests based on finite-dimensional projections, Pigoli *et al.* (2014) for permutation tests based on distance measures and Paparoditis and Sapatinas (2016) for fully functional tests. Notice that testing for the equality of the lag-zero autocovariance operators is an important problem also for functional time series since the associated covariance kernel $c_0(u, v) = Cov(X_t(u), X_t(v))$ of the lag-zero autocovariance operator $\mathcal{C}_0$ describes, for $(u, v) \in \mathcal{I} \times \mathcal{I}$, the entire covariance structure of the random function X_t. Despite its importance, this testing problem has been considered, to the best of our knowledge, only recently by Zhang and Shao (2015). To tackle the aforementioned problems associated with the implementability of limiting distributions, Zhang and Shao (2015) considered tests based on projections on finite dimensional spaces of the differences of the estimated lag-zero autocovariance operators. Notice that similar directional tests have been previously considered for i.i.d. functional data; see Panaretos *et. al.* (2010) and Fremdt *et al.* (2013). Although projection-based tests have the advantage that they lead to manageable limiting distributions, and can be powerful when the deviations from the null are captured by the finite-dimensional space projected, such tests have no power for alternatives which are orthogonal to the projection space. Moreover, and apart from being free from the choice of tuning parameters and consistent for a broader class of alternatives, fully functional tests also allow for an interpretation of the test results; we refer to Section 4.4 for an example.

The paper is organised as follows. In Section 4.2, the basic assumptions on the underlying stochastic process $\mathbb{X}$ are stated and the asymptotic validity of the MBB procedure applied to estimate the distribution of the sample autocovariance operator is established. In Section 4.3, the proposed MBB-based procedure for testing equality of the lag-zero autocovariance operators for several independent functional time series is introduced. Theoretical justifications for approximating the null distribution of a particular fully functional test statistic are given and consistency under the alternative is obtained. Numerical simulations are presented in Section 4.4 in which the finite sample behaviour of the proposed MBB-based testing methodology is investigated. A real-life data example is also discussed in this section. Auxiliary results and proofs of the main results are deferred to Section 4.5.

4.2 Bootstrapping the autocovariance operator

4.2.1 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 (Ω, 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 μ. We define $\langle f, g\rangle = \int_{\mathcal{I}} f(\tau)g(\tau)\mathrm{d}\tau$, $\|f\|^2 = \langle f, f\rangle$ and the tensor product between f and g by $f \otimes g(\cdot) = \langle f, \cdot\rangle g$. For two Hilbert-Schmidt operators Ψ_1 and Ψ_2, we denote by $\langle \Psi_1, \Psi_2\rangle_{HS} = \sum_{i=1}^{\infty}\langle \Psi_1(e_i), \Psi_2(e_i)\rangle$ the inner product which generates the Hilbert-Schmidt norm $\|\Psi_1\|_{HS}^2 = \sum_{i=1}^{\infty} \|\Psi_1(e_i)\|^2$, where $\{e_i, i = 1, 2, \ldots\}$ is any orthonormal basis of $L^2(\mathcal{I})$. If Ψ_1 and Ψ_2 are Hilbert-Schmidt integral operators with kernels $\psi_1(u,v)$ and $\psi_2(u,v)$, respectively, then $\langle \Psi_1, \Psi_2 \rangle_{HS} = \int_{\mathcal{I}} \int_{\mathcal{I}} \psi_1(u,v)\psi_2(u,v) \mathrm{d}u\mathrm{d}v$. We also define the tensor product between the operators Ψ_1 and Ψ_2 analogous to the tensor product of two functions, i.e., $\Psi_1 \otimes \Psi_2(\cdot) = \langle \Psi_1, \cdot \rangle_{HS} \Psi_2$. Note that $\Psi_1 \otimes \Psi_2$ is an operator acting on the space of Hilbert-Schmidt operators. 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 $L^2(\mathcal{I})$, for simplicity.

To describe more precisely the dependence structure of the stochastic process $\mathbb{X}$, we use the notion of L^p-m-approximability; see Hörmann and Kokoszka (2010). A stochastic process $\mathbb{X} = \{X_t, t \in \mathbb{Z}\}$ with X_t taking values in L^2, is called L^4-m-approximable if the following conditions are satisfied:

(i) X_t admits the representation

$$X_t = f(\delta_t, \delta_{t-1}, \delta_{t-2}, \ldots) \tag{4.2.1}$$

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

(ii) $\mathbb{E}\|X_0\|^4 < \infty$ and

$$\sum_{m \geq 1} \left(\mathbb{E}\|X_t - X_{t,m}\|^4 \right)^{1/4} < \infty, \tag{4.2.2}$$

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

The rational behind this concept of weak dependence is that the function f in (4.2.1) is such that the effect of the innovations δ_i far back in the past becomes negligible, that is, these innovations can be replaced by other, independent, innovations. For the stochastic process $\mathbb{X}$ considered in this paper, we somehow strengthen (4.2.2) to the following assumption.

Assumption 4. $\mathbb{X}$ is L^4-m-approximable and satisfies

$$\lim_{m \to \infty} m \left(\mathbb{E}\|X_t - X_{t,m}\|^4 \right)^{1/4} = 0.$$

Since $\mathbb{E}\|X_t\|^2 < \infty$, the autocovariance operator at lag $h \in \mathbb{Z}$ exists and is defined by

$$\mathcal{C}_h = \mathbb{E}[(X_t - \mu) \otimes (X_{t+h} - \mu)].$$

Having an observed stretch $X_1, X_2, \ldots, X_n$, the operator $\mathcal{C}_h$ is commonly estimated by the corresponding sample autocovariance operator, which is given by

$$\widehat{\mathcal{C}}_h = \begin{cases} n^{-1} \sum_{t=1}^{n-h} (X_t - \overline{X}_n) \otimes (X_{t+h} - \overline{X}_n), & \text{if } 0 \leq h < n, \\ n^{-1} \sum_{t=1}^{n+h} (X_{t-h} - \overline{X}_n) \otimes (X_t - \overline{X}_n), & \text{if } -n < h < 0, \\ 0, & \text{otherwise}, \end{cases}$$

where $\overline{X}_n = (1/n) \sum_{t=1}^{n} X_t$ is the sample mean function. The limiting distribution of $\sqrt{n}\big(\widehat{\mathcal{C}}_h - \mathcal{C}_h\big)$ can be derived using the same arguments to those applied in Kokoszka and Reimherr (2013) to investigate the limiting distribution of $\sqrt{n}\big(\widehat{\mathcal{C}}_0 - \mathcal{C}_0\big)$. More precisely, it can be shown that, for any (fixed) lag h, $h \in \mathbb{Z}$, under L^4-approximability conditions, $\sqrt{n}\big(\widehat{\mathcal{C}}_h - \mathcal{C}_h\big) \Rightarrow \mathcal{Z}_h$, where $\mathcal{Z}_h$ is a Gaussian Hilbert-Schmidt operator with covariance operator Γ_h given by

$$\Gamma_h = \sum_{s=-\infty}^{\infty} \mathbb{E}[((X_1 - \mu) \otimes (X_{1+h} - \mu) - \mathcal{C}_h) \otimes ((X_{1+s} - \mu) \otimes (X_{1+h+s} - \mu) - \mathcal{C}_h)];$$

see also Mas (2002) for a related result if $\mathbb{X}$ is a Hilbertian linear processes.

4.2.2 A Bootstrap CLT for the empirical autocovariance operator

In this section, we formulate and prove consistency of the MBB for estimating the distribution of $\sqrt{n}\big(\widehat{\mathcal{C}}_h - \mathcal{C}_h\big)$ for any (fixed) lag h, $h \in \mathbb{Z}$, in the case of weakly dependent Hilbert space-valued random variables satisfying the L^4-approximability condition stated in Assumption 4. The MBB procedure was originally proposed for real-valued time series by Künsch (1989) and Liu and Singh (1992). Adopted to the functional set-up, this resampling procedure first divides the functional time series at hand into the collection of all possible overlapping blocks of functions of length b. That is, the first block consists of the functional observations 1 to b, the second block consists of the functional observations 2 to $b+1$, and so on. Then, a bootstrap sample is obtained by independent sampling, with replacement, from these blocks of functions and joining the blocks together in the order selected to form a new set of functional pseudo observations.

However, to deal with the problem of estimating the distribution of the sample autocovariance operator $\widehat{\mathcal{C}}_h$, we modify the above basic idea and apply the MBB directly to the set of random elements $\mathbb{Y}_{n-h} = \{\widehat{\mathcal{Y}}_{t,h},\, t = 1, 2, \ldots, n-h\}$, where $\widehat{\mathcal{Y}}_{t,h} = (X_t - \overline{X}_n) \otimes (X_{t+h} - \overline{X}_n)$. As mentioned in the Introduction, this has certain advantages in the testing context which will be discussed in the next section. The MBB procedure applied to generate the pseudo random elements $\mathcal{Y}^*_{1,h}, \mathcal{Y}^*_{2,h}, \ldots, \mathcal{Y}^*_{n-h,h}$ is described by the following steps.

Step 1 : Let $b = b(n), 1 \leq b < n-h$, be an integer and denote by B_t, the block of length b starting from the tensor operator $\widehat{\mathcal{Y}}_t$,i.e., $B_t = \{\widehat{\mathcal{Y}}_{t,h}, \widehat{\mathcal{Y}}_{t+1,h}, \ldots, \widehat{\mathcal{Y}}_{t+b-1,h}\}$, where $t = 1, 2, \ldots, N$ and $N = n-h-b+1$ is the total number of such blocks available.

Step 2 : Let k be a positive integer satisfying $b(k-1) < n-h$ and $bk \geq n-h$ and define k i.i.d. integer-valued random variables $I_1, I_2, \ldots, I_k$ selected from a discrete uniform distribution which assigns probability $1/N$ to each element of the set $\{1, 2, \ldots, N\}$.

Step 3 : Let $B^*_i = B_{I_i}$, $i = 1, 2, \ldots, k$, and denote by $\{\mathcal{Y}^*_{(i-1)b+1,h}, \mathcal{Y}^*_{(i-1)b+2,h}, \ldots, \mathcal{Y}^*_{ib,h}\}$ the elements of B^*_i. Join the k blocks in the order $B^*_1, B^*_2, \ldots, B^*_k$ together to obtain a new set of functional pseudo observations. The MBB generated sample of pseudo random elements consists then of the set $\mathcal{Y}^*_{1,h}, \mathcal{Y}^*_{2,h}, \ldots, \mathcal{Y}^*_{n-h,h}$.

Note that if we are interested in the distribution of the sample autocovariance operator $\widehat{\mathcal{C}}_h$ for some (fixed) lag h, $-n < h < 0$, then the above algorithm can be applied to the time series of operators $\mathbb{Y}_{n+h} = \{\widehat{\mathcal{Y}}_{t,h},\, t = h+1, h+2, \ldots, n\}$, where $\widehat{\mathcal{Y}}_{t,h} = (X_{t-h} - \overline{X}_n) \otimes (X_t - \overline{X}_n)$, $t = h+1, h+2, \ldots, n$, with minor changes. Hence, below, we only focus on the case of $0 \leq h < n$.

Given a stretch $\mathcal{Y}^*_{1,h}, \mathcal{Y}^*_{2,h}, \ldots, \mathcal{Y}^*_{n-h,h}$ of pseudo random elements generated by the above MBB procedure, a bootstrap estimator of the autocovariance operator is given by the sample mean

$$\widehat{\mathcal{C}}^*_h = \frac{1}{n} \sum_{t=1}^{n-h} \mathcal{Y}^*_{t,h}.$$

The proposal is then to estimate the distribution of $\sqrt{n}(\widehat{\mathcal{C}}_h - \mathcal{C}_h)$ by the distribution of the bootstrap analogue $\sqrt{n}(\widehat{\mathcal{C}}^*_h - \mathbb{E}^*(\widehat{\mathcal{C}}^*_h))$, where $\mathbb{E}^*(\widehat{\mathcal{C}}^*_h)$ is (conditionally on $\mathbb{X}_n$)

the expected value of $\widehat{\mathcal{C}}_h^*$. Assuming, for simplicity, that $n-h=kb$, straightforward calculations yield

$$\mathbb{E}^*(\widehat{\mathcal{C}}_h^*) = \frac{1}{N}\frac{n-h}{n}\left[\sum_{t=1}^{n-h}\widehat{\mathcal{Y}}_{t,h} - \sum_{j=1}^{b-1}\Big(1-\frac{j}{b}\Big)(\widehat{\mathcal{Y}}_{j,h}+\widehat{\mathcal{Y}}_{n-h-j+1,h})\right]. \tag{4.2.3}$$

The following theorem establishes validity of the MBB procedure suggested for approximating the distribution of $\sqrt{n}(\hat{C}_h - C_h)$.

Theorem 4.2.1. *Suppose that the stochastic process $\mathbb{X}$ satisfies Assumption 4. For $0 \le h < n$, let $\mathcal{Y}_{1,h}^*, \mathcal{Y}_{2,h}^*, \ldots, \mathcal{Y}_{n-h,h}^*$ be a stretch of functional pseudo random elements generated as in Steps 1-3 of the MBB procedure and assume that the block size $b = b(n)$ satisfies $b^{-1} + bn^{-1/3} = o(1)$ as $n \to \infty$. Then, as $n \to \infty$,*

$$d(\mathcal{L}(\sqrt{n}(\widehat{\mathcal{C}}_h^* - \mathbb{E}^*(\widehat{\mathcal{C}}_h^*)) \mid \mathbb{X}_n),\ \mathcal{L}(\sqrt{n}(\widehat{\mathcal{C}}_h - \mathcal{C}_h))) \to 0, \quad \textit{in probability,}$$

where d is any metric metrizing weak convergence on the space of Hilbert-Schmidt operators acting on L^2 and $\mathcal{L}(X)$ denotes the law of the random element X.

4.3 Testing equality of lag-zero autocovariance operators

In this section, we consider the problem of testing the equality of the lag-zero autocovariance operators for a finite number of functional time series and use a modified version of the proposed MBB procedure. This modification leads to a MBB-based testing procedure which generates functional pseudo observations that satisfy the null hypothesis that all lag-zero autocovariance operators are equal. Since this procedure is designed without having any particular statistic in mind, it can potentially be applied to a broad range of possible test statistics which are appropriate for the particular testing problem considered.

To make things specific, consider K independent, L^4-m-approximable functional time series, denoted in the following by $\mathbb{X}_{K,M} = \{X_{i,t},\ i = 1, 2\ldots, K,\ t = 1, 2, \ldots, n_i\}$, where n_i denotes the length of the i-th time series. Let $\mathcal{C}_{i,0}$, $i = 1, 2\ldots, K$, be the lag-zero autocovariance operator of the i-th functional time series, i.e., $\mathcal{C}_{i,0} = \mathbb{E}[(X_{i,t} - \mu_i)\otimes(X_{i,t}-\mu_i)]$, where $\mu_i = EX_{i,t}$. Also, denote by $M = \sum_{i=1}^{K} n_i$ the total number of observations. The null hypothesis of interest is then

$$H_0 : \mathcal{C}_{1,0} = \mathcal{C}_{2,0} = \ldots = \mathcal{C}_{K,0} \tag{4.3.1}$$

and the alternative hypothesis is

$$H_1 : \exists\, k, m \in \{1, 2, \ldots, K\} \text{ with } k \neq m \text{ such that } \mathcal{C}_{k,0} \neq \mathcal{C}_{m,0}.$$

By considering the operator processes $\{\mathcal{Y}_{i,t} = (X_{i,t} - \mu_i) \otimes (X_{i,t} - \mu_i), t \in \mathbb{Z}\},\quad i = 1, 2 \ldots, K$, and denoting by $\mu_i^{\mathcal{Y}} = \mathbb{E}\mathcal{Y}_{i,t}$ the expectation of $\mathcal{Y}_{i,t}$, the null hypothesis of interest can be equivalently written as

$$H_0 : \mu_1^{\mathcal{Y}} = \mu_2^{\mathcal{Y}} = \ldots = \mu_K^{\mathcal{Y}} \tag{4.3.2}$$

and the alternative hypothesis as

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

Consequently, the aim of the bootstrap is to generate a set of K pseudo random elements $\mathbb{Y}^*_{K,M} = \{\mathcal{Y}^*_{i,t},\ i = 1, 2 \ldots, K,\ t = 1, 2, \ldots, n_i\}$ which satisfy the null hypothesis (4.3.2), that is, the expectations $E^*(\mathcal{Y}^*_{i,t})$ should be identical for all $i = 1, 2, \ldots, K$. This leads to the MBB-based testing procedure described in the next section.

4.3.1 The MBB-based Testing Procedure

Suppose that, in order to test the null hypothesis (4.3.2), we use a real-valued test statistic T_M, where, for simplicity, we assume that large values of T_M argue against the null hypothesis. Since we focus on the tensor operators $\mathcal{Y}_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, 2 \ldots, K$, it is natural to assume that the test statistic T_M is based on the tensor product of the centered observed functions, that is on $\widehat{\mathcal{Y}}_{i,t} = (X_{i,t} - \overline{X}_{i,n_i}) \otimes (X_{i,t} - \overline{X}_{i,n_i})$, $i = 1, 2 \ldots, K$, $t = 1, 2, \ldots, n_i$, where $\overline{X}_{i,n_i}$ is the sample mean function of the i-th population, i.e, $\overline{X}_{i,n_i} = (1/n_i) \sum_{t=1}^{n_i} X_{i,t}$. Suppose next, without loss of generality, that the null hypothesis (4.3.2) is rejected if $T_M > d_{M,\alpha}$, where, for $\alpha \in (0, 1)$, $d_{M,\alpha}$ denotes the upper α-percentage point of the distribution of T_M under H_0. We propose to approximate the distribution of T_M under H_0 by the distribution of the bootstrap quantity T^*_M, where the latter is obtained through the following steps.

Step 1 : Calculate the pooled mean

$$\overline{\mathcal{Y}}_M = \frac{1}{M}\sum_{i=1}^{K}\sum_{t=1}^{n_i}\widehat{\mathcal{Y}}_{i,t}.$$

Step 2 : For $i = 1, 2, \ldots, K$, let $b_i = b_i(n) \in \{1, 2, \ldots, n-1\}$ be the block size used for the i-th functional time series and let $N_i = n_i - b_i + 1$. Calculate

$$\widetilde{\mathcal{Y}}_{i,\xi} = \frac{1}{N_i}\sum_{t=\xi}^{N_i+\xi-1}\widehat{\mathcal{Y}}_{i,t},\ \xi = 1, 2, \ldots, b_i$$

Step 3 : For simplicity assume that $n_i = k_i b_i$ and for $i = 1, 2, \ldots, K$, let $q_1^i, q_2^i, \ldots, q_{k_i}^i$ be i.i.d. integers selected from a discrete probability distribution which assigns the probability $1/N_i$ to each element of the set $\{1, 2, \ldots, N_i\}$. Generate bootstrap functional pseudo observations $\mathcal{Y}^*_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$, as

$$\mathcal{Y}^*_{i,t} = \overline{\mathcal{Y}}_M + \widehat{\mathcal{Y}}^*_{i,t} - \widetilde{\mathcal{Y}}_{i,\xi},\ \xi = b_i \text{ if } t \bmod b_i = 0 \text{ and } \xi = t \bmod b_i \text{ otherwise},$$

where $\widehat{\mathcal{Y}}^*_{i,\xi+(s-1)b_i} = \widehat{\mathcal{Y}}_{i,q_s^i+\xi-1}$, $s = 1, 2 \ldots, k_i$ and $\xi = 1, 2, \ldots, b_i$

Step 4 : Let T^*_M be the same statistic as T_M but calculated using, instead of the $\widehat{\mathcal{Y}}_{i,t}$'s the bootstrap pseudo random elements $\mathcal{Y}^*_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$. Given $\mathbb{X}_{K,M}$, denote by $D^*_{M,T}$ the distribution of T^*_M. Then for $\alpha \in (0, 1)$, the null hypothesis H_0 is rejected if

$$T_M > d^*_{M,\alpha},$$

where $d^*_{M,\alpha}$ denotes the upper α-percentage point of the distribution of T^*_M, i.e., $\mathbb{P}(T^*_M > d^*_{M,\alpha}) = \alpha$.

Notice that the distribution $D^*_{M,T}$ is unknown but it can be evaluated by Monte-Carlo.

Before establishing validity of the described MBB procedure some remarks are in order. Observe that the mean $\widetilde{\mathcal{Y}}_{i,\xi}$ calculated in Step 2, is the (conditional on $\mathbb{X}_{K,M}$) expected value of $\widehat{\mathcal{Y}}^*_{i,q_s^i+\xi-1}$ for $\xi = b_i$ if $t \bmod b_i = 0$ and $\xi = t \bmod b_i$ otherwise. This motivates the definition

$$\mathcal{Y}^*_{i,t} = \overline{\mathcal{Y}}_M + \widehat{\mathcal{Y}}^*_{i,t} - \widetilde{\mathcal{Y}}_{i,\xi},\ t = 1, 2, \ldots, n_i,\ i = 1, 2, \ldots, K,$$

used in Step 3 of the MBB algorithm. This definition ensures that the generated pseudo random elements $\mathcal{Y}^*_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, K$, satisfy the null hypothesis (4.3.2). In fact, it is easily seen that the pseudo random elements $\mathcal{Y}^*_{i,t}$ have (conditional on $\mathbb{X}_{K,M}$) an expected value which is equal to $\overline{\mathcal{Y}}_M$, that is $E^*(\mathcal{Y}^*_{i,t}) = \overline{\mathcal{Y}}_M$ for all $t = 1, \ldots, n_i$ and $i = 1, \ldots, K$.

4.3.2 Validity of the MBB-based Testing Procedure

Although the proposed MBB-based testing procedure is not designed having any specific test statistic in mind, establishing its validity requires the consideration of a specific class of statistics. In the following, and for simplicity, we focus on the case of two independent population, i.e., $K = 2$. In this case, a natural approach to test equality of the lag-zero autocovariance operators is to consider a fully functional test statistic which evaluates the difference between the empirical lag-zero autocovariance operators, for instance, to use the test statistic

$$T_M = \frac{n_1 n_2}{M} \|\widehat{\mathcal{C}}_{1,0} - \widehat{\mathcal{C}}_{2,0}\|^2_{HS} = \frac{n_1 n_2}{M} \|\overline{\mathcal{Y}}_{1,n_1} - \overline{\mathcal{Y}}_{2,n_2}\|^2_{HS},$$

where $\overline{\mathcal{Y}}_{i,n_i} = (1/n_i) \sum_{t=1}^{n_i} \widehat{\mathcal{Y}}_{i,t}$, $i = 1, 2$, and $M = n_1 + n_2$. The following lemma delivers the asymptotic distribution of T_M under H_0.

Lemma 4.3.1. *Let H_0 hold true, Assumption 4 be satisfied and assume that, as* $\min\{n_1, n_2\} \to \infty$, $n_1/M \to \theta \in (0, 1)$. *Then,*

$$T_M \xrightarrow{d} \|\mathcal{Z}_0\|^2_{HS}$$

where $\mathcal{Z}_0 = \sqrt{1-\theta}\mathcal{Z}_{1,0} - \sqrt{\theta}\mathcal{Z}_{2,0}$ and $\mathcal{Z}_{i,0}$, $i = 1, 2$, are two independent mean zero Gaussian Hilbert-Schmidt operators with covariance operators $\Gamma_{i,0}$, $i = 1, 2$, given by

$$\begin{aligned}\Gamma_{i,0} = {} & \mathbb{E}[((X_{i,1} - \mu_i) \otimes (X_{i,1} - \mu_i) - \mathcal{C}_{i,0}) \otimes ((X_{i,1} - \mu_i) \otimes (X_{i,1} - \mu_i) - \mathcal{C}_{i,0})] \\ & + 2 \sum_{s=2}^{\infty} \mathbb{E}[((X_{i,1} - \mu_i) \otimes (X_{i,1} - \mu_i) - \mathcal{C}_{i,0}) \otimes ((X_{i,s} - \mu_i) \otimes (X_{i,s} - \mu_i) - \mathcal{C}_{i,0})].\end{aligned}$$

As it is seen from the above lemma, the limiting distribution of T_M depends on the difficulty to estimate covariance operators $\Gamma_{i,0}$, $i = 1, 2$, which describe the entire fourth order structure of the underlying functional processes $\mathbb{X}_i$, making the implementation of this asymptotic result for calculating critical values of the T_M test a difficult task.

Theorem 4.3.1 below shows that the MMB-based testing procedure estimates consistently the limiting distribution $\|\mathcal{Z}_0\|_{HS}^2$ of the T_M test and consequently that it can be applied to estimate the critical values of interest.

For this, we apply the MBB-based testing procedure introduced in Section 4.3.1 to generate $\{\mathcal{Y}_{i,t}^*, t = 1, 2, \ldots .n_i\}$, $i \in \{1, 2\}$, and use the bootstrap pseudo statistic

$$T_M^* = \frac{n_1 n_2}{M} \|\overline{\mathcal{Y}}_{1,n_1}^* - \overline{\mathcal{Y}}_{2,n_2}^*\|_{HS}^2,$$

where $\overline{\mathcal{Y}}_{i,n_i}^* = (1/n_i) \sum_{t=1}^{n_i} \mathcal{Y}_{i,t}^*$, $i = 1, 2$. We then have the following result.

Theorem 4.3.1. *Let Assumption 4 be satisfied and assume that* $\min\{n_1, n_2\} \to \infty$, $n_1/M \to \theta \in (0, 1)$. *Also, for* $i \in \{1, 2\}$, *let the block size* $b_i = b_i(n)$ *satisfies* $b_i^{-1} + b_i n_i^{-1/3} = o(1)$, *as* $n_i \to \infty$. *Then,*

$$\sup_{x \in \mathbb{R}} \big| P(T_M^* \leq x \mid \mathbb{X}_{K,M}) - P_{H_0}(T_M \leq x) \big| \to 0, \quad \textit{in probability,}$$

where $P_{H_0}(X \leq \cdot)$ *denotes the distribution function of the random variable* X *when* H_0 *is true.*

Remark 4.3.1. If H_1 is true, that is if $\|\mathcal{C}_{1,0} - \mathcal{C}_{2,0}\|_{HS} = \|\mathbb{E}\mathcal{Y}_{1,t} - \mathbb{E}\mathcal{Y}_{2,t}\|_{HS} > 0$, then it is easily seen that $T_M \to \infty$ under the conditions on n_1 and n_2 stated in Lemma 4.3.1. This, together with Theorem 4.3.1 and Slutsky's theorem, imply consistency of the T_M test based on bootstrap critical values obtained using the distribution of T_M^*, i.e., the power of the test approaches unity, as $n_1, n_2 \to \infty$.

Remark 4.3.2. The advantage of our approach to translate the testing problem considered to a testing problem of equality of mean functions and to apply the bootstrap to the time series of tensor operators $\mathcal{Y}_{i,t}$, $t = 1, 2, \ldots, n_i$, $i = 1, \ldots, K$, is manifested in the generality under which validity of the MBB-based testing procedure is established in Theorem 4.3.1. To elaborate, a MBB approach which would select blocks from the pooled (mixed) set of functional time series in order to generate bootstrap pseudo elements which satisfy the null hypothesis, will lead to the generation of K new functional pseudo time series, which asymptotically will imitate correctly the pooled second *and* the fourth order moment structure of the underlying functional processes. As a consequence, the limiting distribution of T_M as stated in Lemma 4.3.1 and that of the corresponding MBB analogue will coincide only if $\Gamma_1 = \Gamma_2$. This obviously restricts the class of processes for which the MBB procedure is consistent. In the more simple i.i.d. case, a similar limitation exists by the condition $\mathcal{B}_1 = \mathcal{B}_2$ imposed in Theorem 1 of Paparoditis and Sapatinas (2016). Notice that this limitation can be resolved by applying also in the i.i.d. case the basic bootstrap idea proposed in this paper. That is, to first translate the testing problem to one of testing equality of means of samples consisting of the i.i.d. tensor operators and then to apply an appropriate i.i.d. bootstrap procedure.

4.4 Numerical Results

In this section, we investigate via simulations the size and power behavior of the MBB-based testing procedure applied to testing the equality of lag zero autocovariance operators and we illustrate its applicability by considering a real life data set.

4.4.1 Simulations

In the simulation experiment, two groups of functional time series are generated from the functional autoregressive (FAR) model

$$X_t(u) = \int \psi(u,v) X_{t-1}(v)\, \mathrm{d}v + \delta X_{t-2}(u) + B_t(u), \tag{4.4.1}$$

or from the functional moving average (FMA) model,

$$X_t(u) = \int \psi(u,v) B_{t-1}(v)\, \mathrm{d}v + \delta B_{t-2}(u) + B_t(u). \tag{4.4.2}$$

The kernel function $\psi(\cdot,\cdot)$ in the above models is equal and it is given by

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

while the $B_t(\cdot)$'s are generated as i.i.d. Brownian bridges. All curves were approximated using $T = 21$ equidistant points $\tau_1, \tau_2, \ldots, \tau_{21}$ in the unit interval $\mathcal{I}$ and transformed into functional objects using the Fourier basis with 21 basis functions. Functional time series of length $n_1 = n_2 = 200$ are then generated and testing the null hypothesis $H_0 : \mathcal{C}_{1,0} = \mathcal{C}_{2,0}$ is considered using the T_M test investigated Section 3.2. All bootstrap calculations are based on $B = 1000$ bootstrap replicates, $R = 1000$ model repetitions have been considered and a range of different block sizes have been used. Since $n_1 = n_2$ we set for simplicity $b = b_1 = b_2$.

The T_M test has been applied using three standard nominal levels $\alpha = 0.01$, 0.05 and 0.10. Notice that $\delta = 0$ corresponds to the null hypothesis while to investigate the power behavior of the test we set $\delta = 0$ for the first functional time series and allow for $\delta \in \{0.2, 0.5, 0.8\}$ for the second and for each of the two different models considered. The results obtained for different values of the block size b using the FAR model (4.4.1) as well as the FMA model (4.4.2) are shown in Table 4.1. As it is seen from this table, the MBB based testing procedure retains the nominal level with good size results, especially for $b = 6$ and for both dependence structures considered. Furthermore, the power of the T_M test increases as the deviation from the null increases and reaches high values for the large values of the deviation parameter δ considered.

			Block Size, b=				
	δ	α	2	4	6	8	10
FAR (1)	0	0.01	0.011	0.022	0.014	0.021	0.018
		0.05	0.050	0.062	0.063	0.083	0.076
		0.10	0.108	0.123	0.108	0.132	0.125
	0.2	0.01	0.025	0.018	0.020	0.025	0.026
		0.05	0.089	0.093	0.085	0.081	0.089
		0.10	0.151	0.171	0.150	0.156	0.151
	0.5	0.01	0.593	0.495	0.411	0.381	0.375
		0.05	0.776	0.731	0.698	0.676	0.672
		0.10	0.839	0.813	0.794	0.788	0.791
	0.8	0.01	1.000	1.000	1.000	0.997	0.989
		0.05	1.000	1.000	1.000	1.000	1.000
		0.10	1.000	1.000	1.000	1.000	1.000
FMA (1)	0	0.01	0.012	0.013	0.014	0.013	0.015
		0.05	0.065	0.073	0.060	0.054	0.071
		0.10	0.121	0.108	0.118	0.116	0.127
	0.2	0.01	0.015	0.022	0.019	0.024	0.016
		0.05	0.055	0.076	0.065	0.079	0.062
		0.10	0.1114	0.130	0.119	0.123	0.122
	0.5	0.01	0.148	0.125	0.143	0.121	0.131
		0.05	0.339	0.239	0.330	0.292	0.289
		0.10	0.479	0.421	0.468	0.412	0.418
	0.8	0.01	0.074	0.695	0.689	0.693	0.681
		0.05	0.920	0.889	0.899	0.887	0.900
		0.10	0.957	0.944	0.941	0.949	0.957

Table 4.1: Empirical size and power of the T_M test using bootstrap critical values.

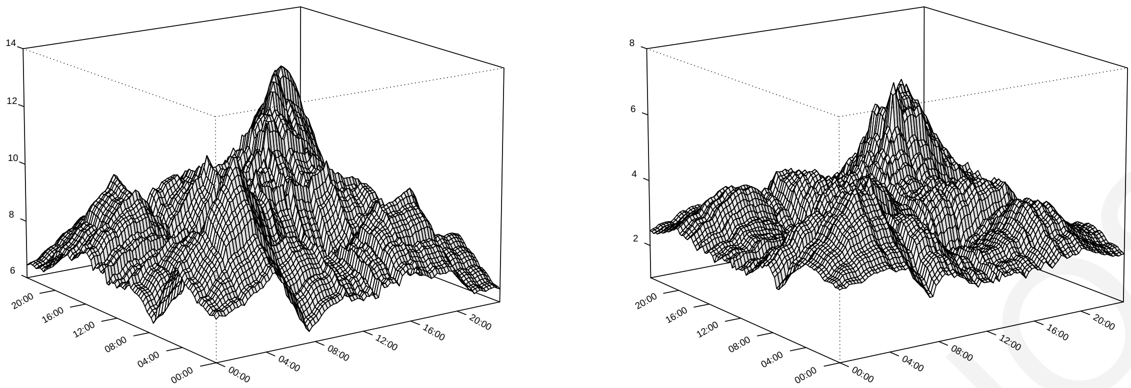

Figure 4.1: Estimated lag-zero autocovariance kernels of the temperature curves: Summer 2007 (left panel) and Summer 2009 (right panel).

4.4.2 A Real-Life Data Example

In this section, the bootstrap based T_M test testing is applied to a real-life data set which consists of daily temperatures recorded in 15 minutes intervals in Nicosia, Cyprus, i.e., there are 96 temperature measurements for each day. Sample A and Sample B consist of the daily temperatures recorded in Summer 2007 (01/06/2007-31/08/2007) and Summer 2009 (01/06/2009-31/08/2009) respectively. The 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 $\mathcal{I} = [0,1]$. Figure 4.1 shows the estimated lag-zero autocovariance kernels $\widehat{c}_i(u,v) = n_i^{-1}\sum_{t=1}^{n_i}(X_{i,t}(u) - \overline{X}_i(u))(X_{i,t}(v) - \overline{X}_i(v))$, $(u,v) \in \mathcal{I}\times\mathcal{I}$, associated with the lag-zero autocovariance operators for the temperature curves of the summer 2007 ($i=1$) and of the summer 2009 ($i=2$). We are interested in testing whether the covariance structure of the daily temperature curves of the two summer periods is the same. The p-values of the MBB-based T_M test using $B = 1000$ bootstrap replicates and for a selection of different block sizes $b = b_1 = b_2$, are 0.016 ($b=3$), 0.015 ($b=4$), 0.033 ($b=5$) and 0.030 ($b=6$). As it is evident from these results, the p-values of the MBB-based test are quite small and lead to a rejection of H_0, for instance at the commonly used 5% level.

To see were the differences between the temperatures in the two summer periods come from and to better interpret the test results, Figure 4.2 presents a contour plot of the estimated squared differences $|\widehat{c}_1(u,v) - \widehat{c}_2(u,v)|^2$ for different values of

(u, v) in the plane $[0,1]^2$. Note that the Hilbert-Schmidt distance $\|\widehat{\mathcal{C}}_{1,0} - \widehat{\mathcal{C}}_{2,0}\|_{HS}$ appearing in the test statistic T_M can be approximated by the discretized quantity $\sqrt{L^{-2}\sum_{i=1}^{L}\sum_{j=1}^{L}|\widehat{c}_1(u_i, v_j) - \widehat{c}_2(u_i, v_j)|^2}$, where $L = 96$ is the number of equidistant time points in the interval $[0,1]$ used and at which the temperature measurements are recorded. Large values of $|\widehat{c}_1(u_i, v_j) - \widehat{c}_2(u_i, v_j)|^2$ (i.e., dark gray regions in Figure 4.2) contribute strongly to the value of the test statistic T_M and pinpoint to regions where large differences between the corresponding lag-zero autocovariance operators occur. Taking into account the symmetry of the covariance kernel $c(\cdot,\cdot)$, Figure 4.2 is very informative. It shows that the main differences between the two covariance operators are concentrated between the time regions 3.00am to 6.00am and 3.00pm to 8.00pm of the daily temperature curves, with the strongest contributions to the test statistic being due to the largest differences recorded around 4.00 to 4.30 in the morning and 6.30 to 7.30 in the evening.

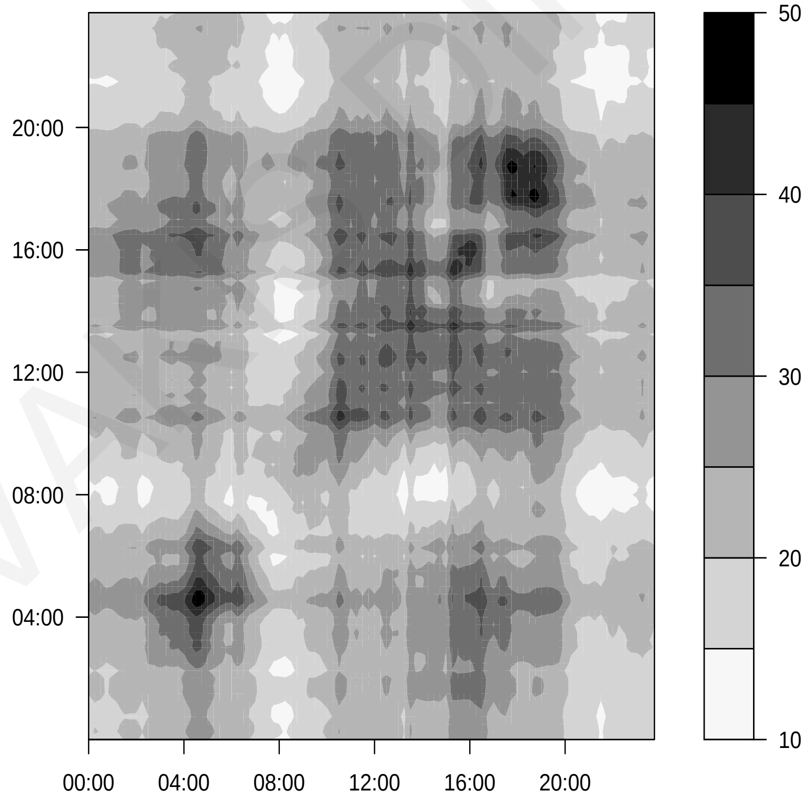

Figure 4.2: Contour plot of the estimated differences $|\widehat{c}_1(u_i, v_j) - \widehat{c}_2(u_i, v_j)|^2$ for $(i, j) \in \{1, 2, \ldots, 96\}$.

4.5 Appendix : Proofs

In the following we assume, without loss of generality, that $\mu = 0$ and we consider the case $h = 0$ only. Furthermore, we let $\widehat{\widetilde{\mathcal{C}}}_0 = n^{-1}\sum_{t=1}^{n} X_t \otimes X_t$, $Z_t = X_t \otimes X_t - \mathcal{C}_0$, $\widehat{Z}_t = X_t \otimes X_t - \widehat{\widetilde{\mathcal{C}}}_0$, $\widetilde{Z}_t = X_t \otimes X_t$, $Z_{t,m} = X_{t,m} \otimes X_{t,m} - \mathcal{C}_0$, $Z_t^* = X_t^* \otimes X_t^*$ and $\widehat{Z}_t^* = X_t^* \otimes X_t^* - \widehat{\widetilde{\mathcal{C}}}_0$. Also, we denote by $Z_t(u,v)$ the kernel of the integral operator Z_t, i.e., $Z_t(u,v) = X_t(u)X_t(v) - c_0(u,v)$, where $c_0(u,v) = \mathbb{E}[X_t(u)X_t(v)]$, and by $Z_{t,m}(u,v)$ the kernel of the integral operator $Z_{t,m}$, i.e., $Z_{t,m}(u,v) = X_{t,m}(u)X_{t,m}(v) - c_0(u,v)$.

We first fix some notation and present two basic lemmas which will be used in the proofs. Towards this note first that we repeatedly use the fact that, by stationarity, $\mathbb{E}\|X_{t,m} - X_t\|^p = \mathbb{E}\|X_{0,m} - X_0\|^p$ and $\mathbb{E}\|X_{t,m}\|^p = \mathbb{E}\|X_t\|^p = \mathbb{E}\|X_0\|^p$ for $p \in \mathbb{N}$ and for all $t \in \mathbb{Z}$. Also note that Kokoszka and Reimherr (2013) proved that the L^4-m-approximability of $\mathbb{X}$ implies that the tensor product process $\{X_t \otimes X_t,\, t \in \mathbb{Z}\}$ is L^2-m-approximable.

For $X_{t,m} \otimes X_{t,m}$ the m-dependent approximation of $X_t \otimes X_t$, we, therefore, have

$$\sum_{m=1}^{\infty}\left(\mathbb{E}\|X_t \otimes X_t - X_{t,m} \otimes X_{t,m}\|_{HS}^2\right)^{1/2} < \infty. \tag{4.5.1}$$

Furthermore, since $\|X_0 \otimes X_t\|_{HS} = \|X_0\|\|X_t\|$ for all $t \in \mathbb{Z}$, and using Cauchy-Schwarz's inequality, we get, for all $t \in \mathbb{Z}$,

$$\begin{aligned}
&\mathbb{E}\|X_t \otimes X_t - X_{t,m} \otimes X_{t,m}\|_{HS}^2 \\
&\qquad \le 2\mathbb{E}\|X_t \otimes (X_t - X_{t,m})\|_{HS}^2 + 2\mathbb{E}\|(X_t - X_{t,m}) \otimes X_{t,m}\|_{HS}^2 \\
&\qquad \le 4(\mathbb{E}\|X_t\|^4)^{1/2}(\mathbb{E}\|X_t - X_{t,m}\|^4)^{1/2}.
\end{aligned}$$

Therefore, by Assumption 4, we get, for all $t \in \mathbb{Z}$,

$$\begin{aligned}
&\lim_{m\to\infty} m\left(\mathbb{E}\|X_t \otimes X_t - X_{t,m} \otimes X_{t,m}\|_{HS}^2\right)^{1/2} \\
&\qquad \le 2(\mathbb{E}\|X_t\|^4)^{1/4}\lim_{m\to\infty} m(\mathbb{E}\|X_t - X_{t,m}\|^4)^{1/4} = 0
\end{aligned} \tag{4.5.2}$$

and by the same arguments,

$$\|\mathbb{E}[X_0 \otimes X_t]\|_{HS} = \|\mathbb{E}[X_0 \otimes (X_t - X_{t,t})]\|_{HS}$$

$$\leq \left(\mathbb{E}\|X_0\|^2\right)^{1/2} \left(\mathbb{E}\|X_0 - X_{0,t}\|^2\right)^{1/2}$$
$$\leq \left(\mathbb{E}\|X_0\|^2\right)^{1/2} \left(\mathbb{E}\|X_0 - X_{0,t}\|^4\right)^{1/4}.$$

Therefore, the L^4-m-approximability assumption implies that

$$\sum_{t\in\mathbb{Z}} \|\mathbb{E}[X_0 \otimes X_t]\|_{HS} < \infty.$$

To prove Theorem 4.2.1, we establish below Lemma 4.5.1 and Lemma 4.5.2.

Lemma 4.5.1. *Let $g_b(\cdot)$ 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 $c > 0$. Assume that for any fixed u, $g_b(u) \to 1$ as $n \to \infty$. Suppose that the process $\mathbb{X}$ satisfies Assumption 4 and that $b = b(n)$ is a sequence of integers such that $b^{-1} + bn^{-1/3} = o(1)$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\left\| \sum_{s=-b+1}^{b-1} g_b(s)\hat{\Gamma}_s - \sum_{s=-\infty}^{\infty} \mathbb{E}[Z_0 \otimes Z_s] \right\|_{HS} = o_p(1),$$

where $\hat{\Gamma}_s = \frac{1}{n}\sum_{t=1}^{n-s} \hat{Z}_t \otimes \hat{Z}_{t+s}$ for $0 \leq s \leq b-1$ and $\hat{\Gamma}_s = \frac{1}{n}\sum_{t=1}^{n+s} \hat{Z}_{t-s} \otimes \hat{Z}_t$ for $-b+1 \leq s < 0$.

Proof. We proceed in two steps. First, we proof that, as $n \to \infty$,

$$\left\| \sum_{s=-b+1}^{b-1} g_b(s)\tilde{\Gamma}_s - \sum_{t=-\infty}^{\infty} \mathbb{E}[Z_0 \otimes Z_t] \right\|_{HS} = o_p(1), \tag{4.5.3}$$

where $\tilde{\Gamma}_s = n^{-1}\sum_{t=1}^{n-s} Z_t \otimes Z_{t+s}$ for $0 \leq s \leq b-1$ and $\tilde{\Gamma}_s = n^{-1}\sum_{t=1}^{n+s} Z_{t-s} \otimes Z_t$ for $-b+1 \leq s < 0$. Then, we prove that, as $n \to \infty$,

$$\left\| \sum_{s=-b+1}^{b-1} g_b(s) \left(\tilde{\Gamma}_s - \hat{\Gamma}_s\right) \right\|_{HS} = o_p(1). \tag{4.5.4}$$

Consider (4.5.3). Since $\|n^{-1}\sum_{t=1}^{n} Z_t \otimes Z_t - \mathbb{E}[Z_0 \otimes Z_0]\|_{HS} = o_p(1)$ as $n \to \infty$, it suffices to show that, as $n \to \infty$,

$$\left\| \sum_{s=1}^{b-1} g_b(s)\tilde{\Gamma}_s - \sum_{t\geq 1} \mathbb{E}[Z_0 \otimes Z_t] \right\|_{HS} = o_p(1). \tag{4.5.5}$$

Let $c_\infty^+ = \sum_{t\geq 1} \mathbb{E}[Z_0 \otimes Z_t]$, $c_m^+ = \sum_{t=1}^{m} \mathbb{E}[Z_{0,m} \otimes Z_{t,m}]$ and $\tilde{\Gamma}_s^{(m)} = n^{-1}\sum_{t=1}^{n-s} Z_{t,m} \otimes Z_{t+s,m}$.

Then,

$$\left\| \sum_{s=1}^{b-1} g_b(s)\tilde{\Gamma}_s - c_\infty^+ \right\|_{HS} \\ \leq \|c_m^+ - c_\infty^+\|_{HS} + \left\| \sum_{s=1}^{b-1} g_b(s)\tilde{\Gamma}_s^{(m)} - c_m^+ \right\|_{HS} \\ + \left\| \sum_{s=1}^{b-1} g_b(s)\tilde{\Gamma}_s - \sum_{s=1}^{b-1} g_b(s)\tilde{\Gamma}_s^{(m)} \right\|_{HS}. \tag{4.5.6}$$

Assertion (4.5.5) is proved by showing that there exists $m_0 \in \mathbb{N}$ such that all three terms on the right hand side of (4.5.6) can be made arbitrarily small, in probability, as $n \to \infty$ for $m = m_0$.

For the first term of the right hand side of the above inequality, we use the bound

$$\left\| \sum_{t=1}^{m} \mathbb{E}\left[Z_{0,m} \otimes Z_{t,m} - Z_0 \otimes Z_t\right] \right\|_{HS} + \left\| \sum_{t=m+1}^{\infty} \mathbb{E}\left[Z_0 \otimes Z_t\right] \right\|_{HS} \tag{4.5.7}$$

and the decomposition

$$Z_{0,m} \otimes Z_{t,m} - Z_0 \otimes Z_t = (Z_{0,m} - Z_0) \otimes Z_{t,m} + Z_0 \otimes (Z_{t,m} - Z_t).$$

By Cauchy-Schwarz's inequality, we get, for the first term of (4.5.7), that

$$\left\| \sum_{t=1}^{m} \mathbb{E}\left[(Z_{0,m} - Z_0) \otimes Z_{t,m}\right] \right\|_{HS} + \left\| \sum_{t=1}^{m} \mathbb{E}\left[Z_0 \otimes (Z_{t,m} - Z_t)\right] \right\|_{HS} \\ \leq 2\left(\mathbb{E}\|Z_0\|_{HS}^2\right)^{1/2} \sum_{t=1}^{m} \left(\mathbb{E}\|Z_{0,m} - Z_0\|_{HS}^2\right)^{1/2} \\ = 2\left(\mathbb{E}\|Z_0\|_{HS}^2\right)^{1/2} m \left(\mathbb{E}\|Z_{0,m} - Z_0\|_{HS}^2\right)^{1/2}.$$

Therefore, by Assumption 4, we get that, for every $\epsilon_1 > 0$, there exists $m_1 \in \mathbb{N}$ such that the last quantity above is less than ϵ_1 for every $m \geq m_1$. Consider the second term of the right hand side of (4.5.7). Since Z_0 and $Z_{t,t}$ are independent for $t \geq m+1$ and $\mathbb{E}[Z_0] = 0$, we get, using Cauchy-Schwarz's inequality,

$$\left\| \sum_{t=m+1}^{\infty} \mathbb{E}\left[Z_0 \otimes Z_t\right] \right\|_{HS} \leq \left(\mathbb{E}\|Z_0\|_{HS}^2\right)^{1/2} \sum_{t=m+1}^{\infty} \left(\mathbb{E}\|Z_0 - Z_{0,t}\|_{HS}^2\right)^{1/2}.$$

Using (4.5.1), it follows that, for every $\epsilon_2 > 0$, there exists $m_2 \in \mathbb{N}$ such that the above quantity is less than ϵ_2 for every $m \geq m_2$.

For the second term of the bound in (4.5.6), note that, for every $m \geq 1$, we have that for any fixed s, as $n \to \infty$,

$$\left\|\tilde{\Gamma}_s^{(m)} - \mathbb{E}[Z_{0,m} \otimes Z_{s,m}]\right\|_{HS} = o_p(1).$$

Hence, the aforementioned term of interest is $o_p(1)$, if we show that, as $n \to \infty$,

$$\left\|\sum_{s=m+1}^{b-1} g_b(s)\tilde{\Gamma}_s^{(m)}\right\|_{HS} = o_p(1). \tag{4.5.8}$$

By the definition of $\tilde{\Gamma}_s^{(m)}$, we have that

$$\begin{aligned}\mathbb{E}\left\|\sum_{s=m+1}^{b-1} g_b(s)\tilde{\Gamma}_s^{(m)}\right\|_{HS}^2 &= \mathbb{E}\left\langle \sum_{s_1=m+1}^{b-1} g_b(s_1)\tilde{\Gamma}_{s_1}^{(m)}, \sum_{s_2=m+1}^{b-1} g_b(s_2)\tilde{\Gamma}_{s_2}^{(m)}\right\rangle_{HS} \\ &= \frac{1}{n^2}\sum_{s_1=m+1}^{b-1}\sum_{s_2=m+1}^{b-1}\sum_{t_1=1}^{n-s_1}\sum_{t_2=1}^{n-s_2} g_b(s_1)g_b(s_2)\mathbb{E}\langle Z_{t_1,m}\otimes Z_{t_1+s_1,m}, Z_{t_2,m}\otimes Z_{t_2+s_2,m}\rangle_{HS}.\end{aligned}$$

Since the sequence $\{Z_{t,m},\, t \in \mathbb{Z}\}$ is m-dependent, $Z_{t,m}$ and $Z_{t+s,m}$ are independent for $s \geq m+1$ and, therefore, $\mathbb{E}[Z_{t,m} \otimes Z_{t+s,m}] = 0$ for $s \geq m+1$. Hence, the number of terms $\mathbb{E}\langle Z_{t_1,m}\otimes Z_{t_1+s_1,m}, Z_{t_2,m}\otimes Z_{t_2+s_2,m}\rangle_{HS}$ in the last equation above which do not vanish is of order $O(nb)$ and, consequently, as $n \to \infty$,

$$\mathbb{E}\left\|\sum_{s=m+1}^{b-1} g_b(s)\tilde{\Gamma}_s^{(m)}\right\|_{HS}^2 = O\left(\frac{b}{n}\right) = o(1), \tag{4.5.9}$$

from which (4.5.8) follows by Markov's inequality.

For the third term in (4.5.6) we show that, for $m = m_0$ and for any $\delta > 0$,

$$\limsup_{n\to\infty} P\left(\left\|\sum_{s=1}^{b-1} g_b(s)\left(\tilde{\Gamma}_s - \tilde{\Gamma}_s^{(m)}\right)\right\|_{HS} > \delta\right) = 0. \tag{4.5.10}$$

Using Markov's inequality, expression (4.5.10) follows if we show that, for $m = m_0$, as $n \to \infty$,

$$\mathbb{E}\left\|\sum_{s=1}^{b-1} g_b(s)(\tilde{\Gamma}_s - \tilde{\Gamma}_s^{(m)})\right\|_{HS} = o(1). \tag{4.5.11}$$

Now, by the definitions of $\tilde{\Gamma}_h$ and $\tilde{\Gamma}_s^{(m)}$, we have

$$\begin{aligned}
&\mathbb{E}\left\|\sum_{s=1}^{b-1} g_b(s)\left(\tilde{\Gamma}_s-\tilde{\Gamma}_s^{(m)}\right)\right\|_{HS} \\
&\qquad\leq \mathbb{E}\left\|\frac{1}{n}\sum_{s=1}^{m} g_b(s)\sum_{t=1}^{n-s}(Z_t\otimes Z_{t+s}-Z_{t,m}\otimes Z_{t+s,m})\right\|_{HS} \\
&\qquad\quad+\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s}(Z_t\otimes Z_{t+s}-Z_{t,m}\otimes Z_{t+s,m})\right\|_{HS}. \qquad (4.5.12)
\end{aligned}$$

Using Cauchy-Schwarz's inequality and the decomposition

$$Z_t\otimes Z_{t+s}-Z_{t,m}\otimes Z_{t+s,m}=(Z_t-Z_{t,m})\otimes Z_{t+s}+Z_{t,m}\otimes(Z_{t+s}-Z_{t+s,m}),$$

we get, for the first term of the right hand side of (4.5.12), the bound

$$\begin{aligned}
&\frac{1}{n}\sum_{s=1}^{m}\sum_{t=1}^{n-s}(\mathbb{E}\|(Z_t-Z_{t,m})\otimes Z_{t+s}\|_{HS}+\mathbb{E}\|Z_{t,m}\otimes(Z_{t+s}-Z_{t+s,m})\|_{HS}) \\
&\qquad\leq \frac{1}{n}\sum_{s=1}^{m}\sum_{t=1}^{n-s}(\mathbb{E}\|Z_t-Z_{t,m}\|_{HS}^2\mathbb{E}\|Z_{t+s}\|_{HS}^2)^{1/2} \\
&\qquad\qquad\qquad+(\mathbb{E}\|Z_{t+s}-Z_{t+s,m}\|_{HS}^2\mathbb{E}\|Z_{t,m}\|_{HS}^2)^{1/2} \\
&\qquad\leq m[(\mathbb{E}\|Z_0-Z_{0,m}\|_{HS}^2\mathbb{E}\|Z_0\|_{HS}^2)^{1/2}+(\mathbb{E}\|Z_0-Z_{0,m}\|_{HS}^2\mathbb{E}\|Z_{0,m}\|_{HS}^2)^{1/2}].
\end{aligned}$$

By Assumption 4, it follows that, for every $\epsilon_3>0$, there exists $m_3\in\mathbb{Z}$ such that, for every $m\geq m_3$, this quantity is less than ϵ_3. For the second term on the right hand side of (4.5.12), we use the bound

$$\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s}Z_t\otimes Z_{t+s}\right\|_{HS}+\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s}Z_{t,m}\otimes Z_{t+s,m}\right\|_{HS}. \qquad (4.5.13)$$

Expression (4.5.9) implies that the second summand of (4.5.13) is $o(1)$. For the first term of (4.5.13), we use the decomposition

$$Z_t\otimes Z_{t+s}=Z_t\otimes Z_{t+s,s}+Z_t\otimes(Z_{t+s}-Z_{t+s,s}),$$

and get the bound

$$\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s}Z_t\otimes Z_{t+s,s}\right\|_{HS}$$

$$+\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s} Z_t\otimes(Z_{t+s}-Z_{t+s,s})\right\|_{HS}. \qquad (4.5.14)$$

For the last term of expression (4.5.14), we have the bound

$$\frac{1}{n}\sum_{s=m+1}^{b-1}\sum_{t=1}^{n-s}\mathbb{E}\left\|Z_t\otimes(Z_{t+s}-Z_{t+s,s})\right\|_{HS}\leq\left(\mathbb{E}\|Z_0\|_{HS}^2\right)^{1/2}\sum_{s=m+1}^{b-1}\left(\mathbb{E}\|Z_0-Z_{0,s}\|_{HS}^2\right)^{1/2}.$$

Therefore, since $\{Z_t,\, t\in\mathbb{Z}\}$ is L^2-m-approximable, with $Z_{0,m}$ be the m-dependent approximation of Z_0, it follows that for every $\epsilon_4>0$, there exists $m_4\in\mathbb{N}$ such that, for every $m\geq m_4$, this term is less than ϵ_4. Consider next the first term of (4.5.14). We have

$$\begin{aligned}\mathbb{E}\left\|\frac{1}{n}\sum_{s=m+1}^{b-1} g_b(s)\sum_{t=1}^{n-s} Z_t\otimes Z_{t+s,s}\right\|_{HS} &\leq \sum_{s=m+1}^{b-1}\mathbb{E}\left\|\frac{1}{n}\sum_{t=1}^{n-s}Z_t\otimes Z_{t+s,s}\right\|_{HS}\\ &\leq\sum_{s=m+1}^{b-1}\left(\mathbb{E}\left\|\frac{1}{n}\sum_{t=1}^{n-s}Z_t\otimes Z_{t+s,s}\right\|_{HS}^2\right)^{1/2}. \qquad (4.5.15)\end{aligned}$$

Since Z_0 and $Z_{s,s}$ are independent, $\|Z_0\otimes Z_t\|_{HS}=\|Z_0\|_{HS}\|Z_t\|_{HS}$ and

$$\mathbb{E}\langle Z_0\otimes Z_{s,s}, Z_t\otimes Z_{t+s,s}\rangle_{HS}=\mathbb{E}\langle Z_0,Z_t\rangle_{HS}\langle Z_{s,s},Z_{t+s,s}\rangle_{HS}=0$$

for $|t|>s$. Using Cauchy-Schwarz's inequality, we get

$$\begin{aligned}&\mathbb{E}\left\|\frac{1}{n}\sum_{t=1}^{n-s}Z_t\otimes Z_{t+s,s}\right\|_{HS}^2\leq\frac{n-s}{n^2}\sum_{|t|<n-s}\mathbb{E}(\langle Z_0\otimes Z_{s,s},Z_t\otimes Z_{t+s,s}\rangle_{HS})\\ &\leq\frac{1}{n}\sum_{t=-s}^{s}|\mathbb{E}\langle Z_0\otimes Z_{s,s},Z_t\otimes Z_{t+s,s}\rangle_{HS}|\leq\frac{1}{n}\sum_{t=-s}^{s}\mathbb{E}\|Z_0\otimes Z_{s,s}\|_{HS}\|Z_t\otimes Z_{t+s,s}\|_{HS}\\ &\leq\frac{1}{n}\sum_{t=-s}^{s}\mathbb{E}\|Z_0\otimes Z_{s,s}\|_{HS}^2\frac{1}{n}\sum_{t=-s}^{s}\left(\mathbb{E}\|Z_0\|_{HS}^2\right)^2\\ &\leq\frac{1}{n}\sum_{t=-s}^{s}\left(\mathbb{E}\|X_0\otimes X_0\|_{HS}^2\right)^2\leq\frac{1}{n}\sum_{t=-s}^{s}\left(\mathbb{E}\|X_0\|_{HS}^4\right)^2.\end{aligned}$$

Therefore, by (4.5.15), the first term of (4.5.14) is $O_P(b^{3/2}/n^{1/2})$. The proof is then concluded by choosing $m_0=\max\{m_1,m_2,m_3,m_4\}$.

Consider (4.5.4). First note that using Theorem 3 of Kokoszka and Reimherr (2013), we get, as $n \to \infty$,

$$\left\| \frac{1}{n} \sum_{t=1}^{n} [Z_t \otimes Z_t - \hat{Z}_t \otimes \hat{Z}_t] \right\|_{HS} = \|(\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0) \otimes (\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0)\|_{HS}$$
$$= \frac{1}{n} \left\| \sqrt{n}(\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0) \right\|_{HS}^2 = O_P(1/n).$$

Therefore, it suffices to show that

$$\left\| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} [Z_t \otimes Z_{t+s} - \hat{Z}_t \otimes \hat{Z}_{t+s}] \right\|_{HS} = o_p(1).$$

Again, by Theorem 3 of Kokoszka and Reimherr (2013), we get that, as $n \to \infty$,

$$\begin{aligned}
&\left\| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} [Z_t \otimes Z_{t+s} - \hat{Z}_t \otimes \hat{Z}_{t+s}] \right\|_{HS} \\
&= \left\| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} [(X_t \otimes X_t) \otimes (\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0) + (\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0) \otimes (X_{t+s} \otimes X_{t+s}) \right. \\
&\qquad\qquad \left. + \mathcal{C}_0 \otimes \mathcal{C}_0 - \hat{\hat{\mathcal{C}}}_0 \otimes \hat{\hat{\mathcal{C}}}_0] \right\|_{HS} \\
&\leq \sum_{s=1}^{b-1} \frac{1}{\sqrt{n}} \|\frac{1}{n} \sum_{t=1}^{n-s} (X_t \otimes X_t)\|_{HS} \|\sqrt{n}(\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0)\|_{HS} \\
&\quad + \sum_{s=1}^{b-1} \frac{1}{\sqrt{n}} \|\sqrt{n}(\hat{\hat{\mathcal{C}}}_0 - \mathcal{C}_0)\|_{HS} \|\frac{1}{n} \sum_{t=1}^{n-s} (X_{t+s} \otimes X_{t+s})\| \\
&\quad + \frac{1}{\sqrt{n}} \sum_{s=1}^{b-1} \frac{1}{n} \sum_{t=1}^{n-s} \|\mathcal{C}_0\|_{HS} \|\sqrt{n}(\mathcal{C}_0 - \hat{\hat{\mathcal{C}}}_0)|_{HS} \\
&\quad + \frac{1}{\sqrt{n}} \sum_{s=1}^{b-1} \frac{1}{n} \sum_{t=1}^{n-s} \|\hat{\hat{\mathcal{C}}}_0\|_{HS} \|\sqrt{n}(\mathcal{C}_0 - \hat{\hat{\mathcal{C}}}_0)\|_{HS} = O_P(b/\sqrt{n}) = o_p(1).
\end{aligned}$$

This completes the proof of the lemma.

Lemma 4.5.2. *Let $g_b(\cdot)$ be a non-negative, continuous and bounded function satisfying the conditions of Lemma 4.5.1. Suppose that $\mathbb{X}$ satisfies Assumption 4 and that $b = b(n)$ is a sequence of integers such that $b^{-1} + bn^{-1/2} = o(1)$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\sum_{s=-b+1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-|s|} \iint Z_t(u,v) Z_{t+|s|}(u,v) \mathrm{d}u\mathrm{d}v \xrightarrow{P} \sum_{s=-\infty}^{\infty} \mathbb{E} \iint Z_0(u,v) Z_s(u,v) \mathrm{d}u\mathrm{d}v.$$

Proof. Since $\sum_{t=-\infty}^{\infty} \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v$ converges and is finite, and since

$$\frac{1}{n}\sum_{t=1}^{n} \iint (Z_t(u,v))^2 \mathrm{d}u\mathrm{d}v \xrightarrow{P} \mathbb{E} \iint (Z_0(u,v))^2 \mathrm{d}u\mathrm{d}v$$

as $n \to \infty$, it suffices to prove that

$$\sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v)Z_{t+s}(u,v)\mathrm{d}u\mathrm{d}v \xrightarrow{P} \sum_{t=1}^{\infty} \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v. \quad (4.5.16)$$

Since

$$\begin{aligned}
&\left| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{i=1}^{n-s} \iint Z_t(u,v)Z_{t+s}(u,v)\mathrm{d}u\mathrm{d}v - \sum_{t=1}^{\infty} \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v \right| \\
&\leq \left| \sum_{t=1}^{m} \mathbb{E} \iint Z_{0,m}(u,v)Z_{t,m}(u,v)\mathrm{d}u\mathrm{d}v - \sum_{t=1}^{\infty} \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v \right| \\
&\quad + \left| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v)Z_{t+s,m}(u,v)\mathrm{d}u\mathrm{d}v \right. \\
&\qquad\qquad \left. - \sum_{t=1}^{m} \mathbb{E} \iint Z_{0,m}(u,v)Z_{t,m}(u,v)\mathrm{d}u\mathrm{d}v \right| \\
&\quad + \left| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v)Z_{t+s}(u,v)\mathrm{d}u\mathrm{d}v \right. \\
&\qquad\qquad \left. - \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v)Z_{t+s,m}(u,v)\mathrm{d}u\mathrm{d}v \right|, \qquad (4.5.17)
\end{aligned}$$

assertion (4.5.16) is proved by showing that there exists $m_0 \in \mathbb{N}$ such that all three terms on the right hand side of (4.5.17) can be made arbitrarily small in probability as $n \to \infty$ for $m = m_0$.

For the first term, we use the bound

$$\begin{aligned}
&\left| \sum_{t=1}^{m} \left(\mathbb{E} \iint Z_{0,m}(u,v)Z_{t,m}(u,v)\mathrm{d}u\mathrm{d}v - \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v \right) \right| \\
&\qquad + \left| \sum_{t=m+1}^{\infty} \mathbb{E} \iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v \right|. \qquad (4.5.18)
\end{aligned}$$

By Cauchy-Schwarz's inequality and the decomposition

$$\begin{aligned}
Z_{0,m}(u,v)Z_{t,m}(u,v) - Z_0(u,v)Z_t(u,v) &= [Z_{0,m}(u,v) - Z_0(u,v)]Z_{t,m}(u,v) \\
&\quad + Z_0(u,v)[Z_{t,m}(u,v) - Z_t(u,v)],
\end{aligned}$$

we get that the first term of (4.5.18) is bounded by

$$\begin{aligned}
&\left|\sum_{t=1}^{m}\mathbb{E}\iint[Z_{0,m}(u,v)-Z_0(u,v)]Z_{t,m}(u,v)\mathrm{d}u\mathrm{d}v\right| \\
&\qquad+\left|\sum_{t=1}^{m}\mathbb{E}\iint Z_0(u,v)[Z_{t,m}(u,v)-Z_t(u,v)]\mathrm{d}u\mathrm{d}v\right| \\
&\le 2\sum_{t=1}^{m}\mathbb{E}\left\{\left[\iint[Z_{0,m}(u,v)-Z_0(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2}\left[\iint[Z_{t,m}(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2}\right\} \\
&\le 2\sum_{t=1}^{m}\left[\mathbb{E}\iint[Z_{0,m}(u,v)-Z_0(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2}\left[\mathbb{E}\iint[Z_{t,m}(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2} \\
&\le 2\sum_{t=1}^{m}\left[\mathbb{E}\iint[X_{0,m}(u)X_{0,m}(v)-X_0(u)X_0(v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2} \\
&\qquad\qquad\times\left[\mathbb{E}\iint[X_{0,m}(u)X_{0,m}(v)-c(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2} \\
&= 2\sum_{t=1}^{m}\left[\mathbb{E}\|X_{0,m}\otimes X_{0,m}-X_0\otimes X_0\|_{HS}^2\right]^{1/2}\left[\mathbb{E}\|X_{0,m}\otimes X_{0,m}-C\|_{HS}^2\right]^{1/2} \\
&= \left[\mathbb{E}\|X_0\otimes X_0-C\|_{HS}^2\right]^{1/2}\left(m\left[\mathbb{E}\|X_{0,m}\otimes X_{0,m}-X_0\otimes X_0\|_{HS}^2\right]^{1/2}\right).
\end{aligned}$$

Using (4.5.2), and since $\{X_t\otimes X_t,\ t\in\mathbb{Z}\}$ is L^2-m-approximable, it follows that for every $\epsilon_1>0$ there exists $m_1\in\mathbb{N}$ such that the above term is less than ϵ_1 for every $m\ge m_1$. Consider the second term of (4.5.18). Since $Z_0(u,v)$ and $Z_{t,t}(u,v)$ are independent for $t\ge m+1$, using Cauchy-Schwarz's inequality, we get

$$\begin{aligned}
&\left|\sum_{t=m+1}^{\infty}\mathbb{E}\iint Z_0(u,v)Z_t(u,v)\mathrm{d}u\mathrm{d}v\right| = \left|\sum_{t=m+1}^{\infty}\mathbb{E}\iint Z_0(u,v)[Z_t(u,v)-Z_{t,t}(u,v)]\mathrm{d}u\mathrm{d}v\right| \\
&\le \sum_{t=m+1}^{\infty}\left[\mathbb{E}\iint[Z_0(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2}\left[\mathbb{E}\iint[Z_t(u,v)-Z_{t,t}(u,v)]^2\mathrm{d}u\mathrm{d}v^2\right]^{1/2} \\
&= \left[\mathbb{E}\iint[X_0(u)X_0(v)-c(u,v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2} \\
&\qquad\times\sum_{t=m+1}^{\infty}\left[\mathbb{E}\iint[X_t(u)X_t(v)-X_{t,t}(u)X_{t,t}(v)]^2\mathrm{d}u\mathrm{d}v\right]^{1/2} \\
&= \left[\mathbb{E}\|X_{0,m}\otimes X_{0,m}-C\|_{HS}^2\right]^{1/2}\sum_{t=m+1}^{\infty}\left[\mathbb{E}\|X_{0,m}\otimes X_{0,m}-X_0\otimes X_0\|_{HS}^2\right]^{1/2}.
\end{aligned}$$

From (4.5.1), it follows that for every $\epsilon_2>0$, there exists $m_2\in\mathbb{N}$ such that the above quantity is less than ϵ_2 for every $m\ge m_2$.

Consider next the second term of the the right-hand side of the inequality (4.5.17).

Note that for every $m \geq 1$, we have that, for any fixed s, as $n \to \infty$,

$$\left| \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d} - \mathbb{E} \iint Z_{0,m}(u,v) Z_{s,m}(u,v) \mathrm{d}u\mathrm{d}v \right| = o_p(1).$$

Therefore, the aforementioned term is $o_p(1)$ if we show that

$$\left| \sum_{s=m+1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d} \right| = o_p(1). \tag{4.5.19}$$

For this, notice first that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{s=m+1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d} \right]^2 \\
&= \frac{1}{n^2} \sum_{s_1=m+1}^{b-1} \sum_{s_2=m+1}^{b-1} g_b(s_1) g_b(s_2) \\
&\qquad \times \sum_{t_1=1}^{n-s_1} \sum_{t_2=1}^{n-s_1} \mathbb{E}\Bigg[\iint Z_{t_1,m}(u_1,v_1) Z_{t_1+s_1,m}(u_1,v_1) \mathrm{d}u_1\mathrm{d}v_1 \\
&\qquad\qquad \times \iint Z_{t_2,m}(u_2,v_2) Z_{t_2+s_2,m}(u_2,v_2) \mathrm{d}u_2\mathrm{d}v_2 \Bigg].
\end{aligned}$$

Since the sequence $\{Z_{t,m}(u,v),\ t \in \mathbb{Z}\}$ is m-dependent, $Z_{t,m}(u,v)$ and $Z_{t+s,m}(u,v)$ are independent for $s \geq m+1$, therefore using $\mathbb{E}(Z_{0,m}(u,v)) = 0$ we get that,

$$\mathbb{E} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d}v = 0.$$

Hence, the number of terms

$$\mathbb{E}\left[\iint Z_{t_1,m}(u_1,v_1) Z_{t_1+s_1,m}(u_1,v_1) \mathrm{d}u_1\mathrm{d}v_1 \times \iint Z_{t_2,m}(u_2,v_2) Z_{t_2+s_2,m}(u_2,v_2) \mathrm{d}u_2\mathrm{d}v_2 \right]$$

in the last equation above which do not vanish is of order $O(nb)$ and, consequently, as $n \to \infty$,

$$\mathbb{E}\left[\sum_{s=m+1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d} \right]^2 = O\left(\frac{b}{n}\right) = o(1), \tag{4.5.20}$$

from which (4.5.19) follows by Markov's inequality.

For the third term in (4.5.17), we show that, for $m = m_0$,

$$\limsup_{n\to\infty} P\left(\left| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v) Z_{t+s}(u,v) \mathrm{d}u\mathrm{d}v \right.\right. \\ \left.\left. - \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d}v \right| > \delta \right) = 0, \qquad (4.5.21)$$

for any $\delta > 0$. By Markov's inequality, expression (4.5.21) follows if we show that, for $m = m_0$,

$$\mathbb{E}\left| \sum_{s=1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v) Z_{t+s}(u,v) \mathrm{d}u\mathrm{d}v \right. \\ \left. - Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d}v \right| = o(1). \qquad (4.5.22)$$

For the above quantity we have the bound

$$\mathbb{E}\left| \sum_{s=1}^{m} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v) Z_{t+s}(u,v) - Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d}v \right| \\ + \mathbb{E}\left| \sum_{s=m+1}^{b-1} g_b(s) \frac{1}{n} \sum_{t=1}^{n-s} \iint Z_t(u,v) Z_{t+s}(u,v) \right. \\ \left. - Z_{t,m}(u,v) Z_{t+s,m}(u,v) \mathrm{d}u\mathrm{d}v \right|. \qquad (4.5.23)$$

For the first term of the right hand side of the above inequality, using the decomposition

$$Z_t(u,v) Z_{t+s}(u,v) - Z_{t,m}(u,v) Z_{t+s,m}(u,v) \\ = [Z_t(u,v) - Z_{t,m}(u,v)] Z_{t+s}(u,v) + [Z_{t+s}(u,v) - Z_{t+s,m}(u,v)] Z_{t,m}(u,v)$$

we get the bound,

$$\sum_{s=1}^{m} \frac{1}{n} \sum_{t=1}^{n-s} \mathbb{E} \iint |[Z_t(u,v) - Z_{t,m}(u,v)] Z_{t+s}(u,v)| \, \mathrm{d}u\mathrm{d}v \\ + \mathbb{E} \iint |[Z_{t+s}(u,v) - Z_{t+s,m}(u,v)] Z_{t,m}(u,v)| \, \mathrm{d}u\mathrm{d}v. \qquad (4.5.24)$$

Using Cauchy-Schwarz's inequality, we have

$$\mathbb{E} \iint |[Z_t(u,v) - Z_{t,m}(u,v)] Z_{t+h}(u,v)| \, \mathrm{d}u\mathrm{d}v$$

$$\leq \mathbb{E}\Big[\iint [Z_t(u,v) - Z_{t,m}(u,v)]^2 \mathrm{d}u\mathrm{d}v\Big]^{1/2}\Big[\iint [Z_{t+s}(u,v)]^2 \mathrm{d}u\mathrm{d}v\Big]^{1/2}$$
$$\leq \Big[\mathbb{E}\iint [Z_t(u,v) - Z_{t,m}(u,v)]^2 \mathrm{d}u\mathrm{d}v\Big]^{1/2}\Big[\mathbb{E}\iint [Z_{t+s}(u,v)]^2 \mathrm{d}u\mathrm{d}v\Big]^{1/2}$$
$$= \Big[\mathbb{E}\|X_t \otimes X_t - X_{t,m} \otimes X_{t,m}\|_{HS}^2\Big]^{1/2}\Big[\mathbb{E}\|X_{t+h} \otimes X_{t+s} - \mathcal{C}_0\|_{HS}^2\Big]^{1/2}. \quad (4.5.25)$$

Using the same arguments, we get

$$\mathbb{E}\iint |[Z_{t+s}(u,v) - Z_{t+s,m}(u,v)]Z_{t,m}(u,v)| \,\mathrm{d}u\mathrm{d}v$$
$$\leq \Big[\mathbb{E}\|X_{t+s} \otimes X_{t+s} - X_{t+s,m} \otimes X_{t+s,m}\|_{HS}^2\Big]^{1/2}\Big[\mathbb{E}\|X_t \otimes X_t - C\|_{HS}^2\Big]^{1/2}.$$

Therefore, (4.5.24) is bounded by

$$2(\mathbb{E}\|X_0 \otimes X_0 - \mathcal{C}_0\|_{HS}^2)^{1/2}\left[m(\mathbb{E}\|X_0 \otimes X_0 - X_{0,m} \otimes X_{0,m}\|_{HS}^2)^{1/2}\right].$$

Hence, by (4.5.2), it follows that, for every $\epsilon_3 > 0$, there exists $m_3 \in \mathbb{Z}$ such that, for every $m \geq m_3$, this quantity is bounded by ϵ_3. For the second term on the right hand side of (4.5.23), we use the bound

$$\mathbb{E}\left|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_t(u,v)Z_{t+s}(u,v)\mathrm{d}u\mathrm{d}v\right|$$
$$+\,\mathbb{E}\left|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_{t,m}(u,v)Z_{t+s,m}(u,v)\mathrm{d}u\mathrm{d}v\right|. \quad (4.5.26)$$

Expression (4.5.20) implies that the second summand of (4.5.26) is $o(1)$, while for the first term of (4.5.26) we use the decomposition

$$Z_t(u,v)Z_{t+s}(u,v) = Z_t(u,v)Z_{t+s,s}(u,v) + Z_t(u,v)[Z_{t+s}(u,v) - Z_{t+s,s}(u,v)]$$

to get the bound

$$\mathbb{E}\left|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_t(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\right|$$
$$+\,\mathbb{E}\left|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_t(u,v)[Z_{t+s}(u,v) - Z_{t+s,s}(u,v)]\mathrm{d}u\mathrm{d}v\right|. \quad (4.5.27)$$

Using same arguments as those applied in (4.5.25), we get the bound

$$\mathbb{E}\iint |Z_t(u,v)[Z_{t+s}(u,v) - Z_{t+s,s}(u,v)]|\,\mathrm{d}u\mathrm{d}v \leq \Big[\mathbb{E}\|X_{t+s}\otimes X_{t+s} - X_{t+s,s}\otimes X_{t+s,s}\|_{HS}^2\Big]^{1/2}\Big[\mathbb{E}\|X_t\otimes X_t - \mathcal{C}_0\|_{HS}^2\Big]^{1/2}.$$

Hence, for the last term of expression (4.5.27), we have

$$\mathbb{E}\Bigg|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_t(u,v)[Z_{t+s}(u,v) - Z_{t+s,s}(u,v)]\mathrm{d}u\mathrm{d}v\Bigg| \leq \Big[\mathbb{E}\|X_0\otimes X_0 - \mathcal{C}_0\|_{HS}^2\Big]^{1/2}\sum_{s=m+1}^{\infty}\Big[\mathbb{E}\|X_0\otimes X_0 - X_{0,s}\otimes X_{0,s}\|_{HS}^2\Big]^{1/2}.$$

Therefore, using (4.5.1), we get that for every $\epsilon_4 > 0$, there exists $m_4 \in \mathbb{N}$ such that, for every $m \geq m_4$, this term is bounded by ϵ_4. Consider next the first term of (4.5.27). Using the decomposition

$$Z_t(u,v)Z_{t+s,s}(u,v) = [Z_t(u,v) - Z_{t,s}(u,v)]Z_{t+s,s}(u,v) + Z_{t,s}(u,v)Z_{t+s,s}(u,v),$$

we get the bound

$$\mathbb{E}\Bigg|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint [Z_t(u,v) - Z_{t,s}(u,v)]Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\Bigg| + \mathbb{E}\Bigg|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\Bigg|. \tag{4.5.28}$$

For the first term of this bound, and by Cauchy-Schwarz's inequality, we get the bound

$$\mathbb{E}\Bigg|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint [Z_t(u,v) - Z_{t,s}(u,v)]Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\Bigg| \leq \Big[\mathbb{E}\|X_t\otimes X_t - X_{t,m}\otimes X_{t,m}\|_{HS}^2\Big]^{1/2}\Big[\mathbb{E}\|X_{t+s}\otimes X_{t+s} - \mathcal{C}_0\|_{HS}^2\Big]^{1/2}.$$

Hence, by (4.5.1), it follows that, for every $\epsilon_5 > 0$, there exists $m_3 \in \mathbb{Z}$ such that, for every $m \geq m_3$, this quantity is bounded by ϵ_5. Consider the last term of the expression given in (4.5.28) and note that $\{\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v,\ t \in \mathbb{Z}\}$ is a $2s$-dependent sequence. Also note that since $Z_{t,s}(u,v)$ and $Z_{t+s}(u,v)$ are independent

$\mathbb{E}\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v = 0$. Therefore, as $n \to \infty$,

$$n^{-1/2}\sum_{t=1}^{n}\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v = O_P(1).$$

Hence, using Portmanteau's theorem, and since $f(x) = |x|$ is a Lipschitz function, we get that, as $n \to \infty$,

$$\mathbb{E}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\iint Z_{t,s}(u,v)Z_{t+s}(u,v)\mathrm{d}u\mathrm{d}v\right| = O(1).$$

Therefore, as $n \to \infty$,

$$\begin{aligned}
&\mathbb{E}\left|\sum_{s=m+1}^{b-1} g_b(s)\frac{1}{n}\sum_{t=1}^{n-s}\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\right| \\
&\quad \le \frac{1}{\sqrt{n}}\sum_{s=m+1}^{b-1}\mathbb{E}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n-s}\iint Z_{t,s}(u,v)Z_{t+s,s}(u,v)\mathrm{d}u\mathrm{d}v\right| = O(b/\sqrt{n}) = o(1).
\end{aligned}$$

The proof of the lemma is concluded by choosing $m_0 = \max\{m_1, m_2, m_3, m_4, m_5\}$.

Proof of Theorem 4.2.1. By the triangle inequality and Theorem 3 of Kokoszka and Reimherr (2013), the assertion of the theorem is established if we show that, as $n \to \infty$,

$$\sqrt{n}(\hat{\mathcal{C}}_0^* - \mathbb{E}^*(\hat{\mathcal{C}}_0^*)) \Rightarrow \mathcal{Z}_0, \tag{4.5.29}$$

in probability, where $\mathcal{Z}_0$ is a mean zero Gaussian Hilbert-Schmidt operator with covariance operator given by

$$\Gamma_0 = \mathbb{E}[Z_1 \otimes Z_1] + 2\sum_{s=2}^{\infty}\mathbb{E}[Z_1 \otimes Z_s].$$

Using Theorem 1 of Horváth *et al.* (2013), we get

$$\begin{aligned}
&\sqrt{n}(\hat{\mathcal{C}}_0^* - \mathbb{E}^*(\hat{\mathcal{C}}_0^*)) \\
&\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\Big[X_t^* \otimes X_t^* - \mathbb{E}^*(X_t^* \otimes X_t^*) - \overline{X}_n \otimes (X_t^* - \mathbb{E}^*(X_t^*)) \\
&\qquad\qquad - (X_t^* - \mathbb{E}^*(X_t^*)) \otimes \overline{X}_n\Big] \\
&\quad = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}[Z_t^* - \mathbb{E}^*(Z_t^*)] + O_P(1/\sqrt{n}).
\end{aligned}$$

Also note that

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}[Z_t^* - \mathbb{E}^*(Z_t^*)] = \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\left(\frac{1}{\sqrt{b}}\sum_{i=1}^{b}\left(Z^*_{(t-1)b+i} - \mathbb{E}^*(Z^*_{(t-1)b+i})\right)\right)$$
$$= \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*,$$

with an obvious notation for $\widehat{Y}_t^*$, $t = 1, 2, \dots, k$. Recall that due to the block bootstrap resampling scheme, the random variables $\widehat{Y}_t^*$, $t = 1, 2, \dots, k$, are i.i.d. Therefore to prove (4.5.29), it suffices by Lemma 5 of Kokoszka and Reimherr (2013), to prove that,

(i) $\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS} \xrightarrow{d} N(0, \sigma^2(y))$ for every Hilbert-Schmidt operator y acting on L^2,

and that

(ii) $\lim_{n\to\infty}\mathbb{E}^*\left\|\frac{1}{\sqrt{k}}\sum_{t=1}^{k}\hat{Y}_t^*\right\|_{HS}^2$ exists and is finite.

To establish assertion (i), we first prove that, as $n \to \infty$,

$$\text{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right) \xrightarrow{P} \sigma^2(y). \tag{4.5.30}$$

Consider (4.5.30) and notice that

$$\text{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right) = \text{Var}^*\left(\langle \widehat{Y}_1^*, y\rangle_{HS}\right)$$
$$= \mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}(Z_t^* - \mathbb{E}^*(Z_t^*)), y\right\rangle_{HS}\right]^2. \tag{4.5.31}$$

Let $N = n - b + 1$, $\widetilde{Y}_t = b^{-1/2}(\widetilde{Z}_t + \widetilde{Z}_{t+1} + \ldots + \widetilde{Z}_{t+b-1})$, $t = 1, 2, \dots, N$ and $\widetilde{Y}_t^* = b^{-1/2}\sum_{i=1}^{b} Z^*_{(t-1)b+i}$, $t = 1, 2, \dots, k$. Since $n/N \to 1$ as $n \to \infty$, in the following we will occasionally replace $1/N$ by $1/n$. Notice that,

$$\mathbb{E}^*\left(\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}Z_t^*, y\right\rangle_{HS}\right) = \mathbb{E}^*(\widetilde{Y}_1^*) = \frac{1}{N}\sum_{t=1}^{N}\langle \widetilde{Y}_t, y\rangle_{HS}$$
$$= \frac{\sqrt{b}}{N}\left[\sum_{t=1}^{n}\langle \widetilde{Z}_t, y\rangle_{HS} - \sum_{i=1}^{b-1}\left(1 - \frac{i}{b}\right)\left[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}\right]\right]$$
$$= \langle \sqrt{b}\,\hat{\tilde{\mathcal{C}}}_n, y\rangle - \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1 - \frac{i}{b}\right)\left[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}\right]\right]. \tag{4.5.32}$$

Therefore,

$$
\begin{aligned}
&\mathrm{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right)\\
&\quad= \mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\widehat{Z}_t^*, y\right\rangle_{HS} + \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right]^2\\
&\quad= \mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\widehat{Z}_t^*, y\right\rangle_{HS}\right]^2\\
&\qquad+\left[\frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right]^2\\
&\qquad+2\left[\frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\widehat{Z}_t^*, y\right\rangle_{HS}\right]\\
&\quad= \mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\hat{Z}_t^*, y\right\rangle_{HS}\right]^2 + O_P(b^3/n^2). \qquad (4.5.33)
\end{aligned}
$$

Let $\widehat{Y}_t = b^{-1/2}(\widehat{Z}_t + \widehat{Z}_{t+1} + \ldots + \widehat{Z}_{t+b-1})$, $t = 1, 2, \ldots, N$. Since,

$$
\begin{aligned}
&\mathbb{E}^*\left[\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\widehat{Z}_t^*, y\right\rangle_{HS}\right]^2 = \frac{1}{N}\sum_{t=1}^{N}\langle \widehat{Y}_t, y\rangle_{HS}^2\\
&\quad= \frac{1}{N}\sum_{t=1}^{n}\langle \widehat{Z}_t, y\rangle_{HS}\langle \widehat{Z}_t, y\rangle_{HS}\\
&\quad+\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\langle \widehat{Z}_t, y\rangle_{HS}\langle \widehat{Z}_{t+i}, y\rangle_{HS} + \langle \widehat{Z}_{t+i}, y\rangle_{HS}\langle \widehat{Z}_t, y\rangle_{HS}]\\
&\quad-\frac{1}{N}\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widehat{Z}_i, y\rangle_{HS}\langle \widehat{Z}_i, y\rangle_{HS} + \langle \widehat{Z}_{n-i+1}, y\rangle_{HS}\langle \widehat{Z}_{n-i+1}, y\rangle]_{HS}\\
&\quad-\frac{1}{N}\sum_{i=1}^{b-1}\sum_{j=1}^{b-t}\left(1-\frac{j+i}{b}\right)[\langle \hat{Z}_j, y\rangle_{HS}\langle \widehat{Z}_{j+i}, y\rangle_{HS} + \langle \widehat{Z}_{n-j+1-i}, y\rangle_{HS}\langle \widehat{Z}_{n-j+1}, y\rangle_{HS}\\
&\qquad\qquad\qquad+\langle \widehat{Z}_{j+i}, y\rangle_{HS}\langle \widehat{Z}_j, y\rangle_{HS} + \langle \widehat{Z}_{n-j+1}, y\rangle_{HS}\langle \widehat{Z}_{n-j+1-i}, y\rangle_{HS}],
\end{aligned}
$$

we get, using (4.5.33),

$$
\begin{aligned}
&\mathrm{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right)\\
&\quad= \frac{1}{N}\sum_{t=1}^{n}\langle \widehat{Z}_t, y\rangle_{HS}\langle \widehat{Z}_t, y\rangle_{HS}
\end{aligned}
$$

$$+\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\langle \widehat{Z}_t, y\rangle_{HS}\langle \widehat{Z}_{t+i}, y\rangle_{HS}+\langle \widehat{Z}_{t+i}, y\rangle_{HS}\langle \widehat{Z}_t, y\rangle_{HS}]$$
$$+O_P(b/n)+O_P(b^2/n)+O_P(b^3/n^2).$$

Therefore,

$$\begin{aligned}\mathrm{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right) &= \frac{1}{N}\sum_{t=1}^{n}\langle \widehat{Z}_t\otimes\widehat{Z}_t, y\otimes y\rangle_{HS}\\ &+\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\langle \widehat{Z}_t\otimes\widehat{Z}_{t+i}, y\otimes y\rangle_{HS}+\langle \widehat{Z}_{t+i}\otimes\widehat{Z}_t, y\otimes y\rangle_{HS}]\\ &+O_P(b^2/n). \end{aligned} \tag{4.5.34}$$

Let $g_b(i)=\left(1-\frac{|i|}{b}\right)$ in Lemma 4.5.1, and use the triangular inequality to get

$$\begin{aligned}&\left|\left\langle \frac{1}{N}\sum_{t=1}^{n}\widehat{Z}_t\otimes\widehat{Z}_t+\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\widehat{Z}_t\otimes\widehat{Z}_{t+i}+\widehat{Z}_{t+i}\otimes\widehat{Z}_t]-\sum_{t=-\infty}^{\infty}\mathbb{E}[Z_0\otimes Z_t], y\otimes y\right\rangle_{HS}\right|\\ &\leq\left\|\frac{1}{N}\sum_{t=1}^{n}\widehat{Z}_t\otimes\widehat{Z}_t+\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\widehat{Z}_t\otimes\widehat{Z}_{t+i}+\widehat{Z}_{t+i}\otimes\widehat{Z}_t]-\sum_{t=-\infty}^{\infty}\mathbb{E}[Z_0\otimes Z_t]\right\|_{HS}\times\left\|y\otimes y\right\|_{HS}\\ &=o_p(1).\end{aligned}$$

Therefore, and using $\langle Z_0\otimes Z_t, y\otimes y\rangle_{HS}=\langle Z_0, y\rangle_{HS}\langle Z_0, y\rangle_{HS}$, we get from (4.5.34), as $n\to\infty$,

$$\begin{aligned}\mathrm{Var}^*\left(\left\langle \frac{1}{\sqrt{k}}\sum_{t=1}^{k}\widehat{Y}_t^*, y\right\rangle_{HS}\right)&\xrightarrow{P}\left\langle \sum_{t=-\infty}^{\infty}\mathbb{E}[Z_0\otimes Z_t], y\otimes y\right\rangle_{HS}\\ &=\langle\Gamma, y\otimes y\rangle_{HS}=\sigma^2(y).\end{aligned} \tag{4.5.35}$$

We next establish the asymptotic normality stated in (i). Since $\langle \widehat{Y}_t^*, y\rangle_{HS}$, $t=1,2,\ldots,k$ are i.i.d. real valued random variables, we show that Lindeberg's condition is satisfied, i.e., for every $\varepsilon > 0$, as $n \to \infty$,

$$\frac{1}{\tau_k^{*2}} \sum_{t=1}^{k} \mathbb{E}^* \Bigg[\big(\langle \widehat{Y}_t^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_t^*, y \rangle_{HS}) \big)^2 \times \mathbb{1}\big(|\langle \widehat{Y}_t^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_t^*, y \rangle_{HS})| > \varepsilon \tau_k^* \big) \Bigg] = o_p(1), \tag{4.5.36}$$

where $\mathbb{1}_A(x)$ denotes the indicator function of the set A and

$$\tau_k^{*2} = \sum_{t=1}^{k} \mathrm{Var}^*(\langle \widehat{Y}_t^*, y \rangle_{HS}) = k \mathrm{Var}^*(\langle \widehat{Y}_1^*, y \rangle_{HS}). \tag{4.5.37}$$

To establish (4.5.36), and because of (4.5.35) and (4.5.37), it suffices to show that, for any $\delta > 0$, as $n \to \infty$,

$$P\Bigg(\frac{1}{k} \sum_{t=1}^{k} \mathbb{E}^* \Bigg[(\langle \widehat{Y}_t^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_t^*, y \rangle_{HS}))^2 \times \mathbb{1}(|\langle \widehat{Y}_t^*, y \rangle_{HS} - E^*(\langle \widehat{Y}_t^*, y \rangle_{HS})| > \varepsilon \tau_k^*) \Bigg] > \delta \Bigg) \to 0. \tag{4.5.38}$$

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

$$\mathbb{E}[|X+Y|^2 \mathbb{1}(|X+Y| > \eta)] \leq 4 \left[\mathbb{E}|X|^2 \mathbb{1}(|X| > \eta/2) + \mathbb{E}|Y|^2 \mathbb{1}(|Y| > \eta/2) \right]; \tag{4.5.39}$$

see Lahiri (2003), p. 56. Since the random variables $\langle \widehat{Y}_t^*, y \rangle_{HS}$ are i.i.d., we get using expression (4.5.32) and Markov's inequality that, as $n \to \infty$,

$$\begin{aligned}
&P\Bigg(\frac{1}{k} \sum_{t=1}^{k} \mathbb{E}^* \Bigg[(\langle \widehat{Y}_t^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_t^*, y \rangle_{HS}))^2 \\
&\qquad \times \mathbb{1}(|\langle \widehat{Y}_t^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_t^*, y \rangle_{HS})| > \varepsilon \tau_k^*) \Bigg] > \delta \Bigg) \\
&\leq \delta^{-1} \mathbb{E} \left\{ \mathbb{E}^* \left[(\langle \widehat{Y}_1^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_1^*, y \rangle_{HS}))^2 \mathbb{1}(|\langle \widehat{Y}_1^*, y \rangle_{HS} - \mathbb{E}^*(\langle \widehat{Y}_1^*, y \rangle_{HS})| > \varepsilon \tau_k^*) \right] \right\} \\
&= \delta^{-1} \mathbb{E} \Bigg\{ \mathbb{E}^* \Bigg[\Bigg(\Bigg\langle \frac{1}{\sqrt{b}} \sum_{t=1}^{b} \widehat{Z}_t^*, y \Bigg\rangle_{HS} \\
&\qquad + \frac{\sqrt{b}}{N} \left[\sum_{i=1}^{b-1} \left(1 - \frac{i}{b} \right) [\langle \widetilde{Z}_i, y \rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y \rangle_{HS}] \right] \Bigg)^2
\end{aligned}$$

$$
\times \mathbb{1}\left(\left|\left\langle \frac{1}{\sqrt{b}}\sum_{t=1}^{b}\widehat{Z}_t^*, y\right\rangle_{HS} + \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right| > \varepsilon\tau_k^*\right)\Bigg]\Bigg\}
$$

$$
= \delta^{-1}\mathbb{E}\left[\frac{1}{N}\sum_{t=1}^{N}\left(\langle \widehat{Y}_t, y\rangle_{HS} + \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right)^2 \times \mathbb{1}\left(\left|\langle \widehat{Y}_t, y\rangle_{HS} + \frac{\sqrt{b}}{N}\left[\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right]\right| > \varepsilon\tau_k^*\right)\right]
$$

$$
\le 4\delta^{-1}\Bigg[\mathbb{E}(\langle \widehat{Y}_1, y\rangle_{HS}^2)\mathbb{1}(|\langle \widehat{Y}_1, y\rangle_{HS}| > \varepsilon\tau_k^*/2) + \mathbb{E}\left(\frac{\sqrt{b}}{N}\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right)^2 \times \mathbb{1}(\left|\left(\frac{\sqrt{b}}{N}\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right)\right| > \varepsilon\tau_k^*/2)\Bigg]
$$

$$
\le 4\delta^{-1}\Bigg[\mathbb{E}(\langle \widehat{Y}_1, y\rangle_{HS}^2)\mathbb{1}(|\langle \widehat{Y}_1, y\rangle_{HS}| > \varepsilon\tau_k^*/2) + \mathbb{E}\left(\frac{\sqrt{b}}{N}\sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)[\langle \widetilde{Z}_i, y\rangle_{HS} + \langle \widetilde{Z}_{n-i+1}, y\rangle_{HS}]\right)^2\Bigg]
$$

$$
\le 4\delta^{-1}\mathbb{E}(\langle \widehat{Y}_1, y\rangle_{HS}^2)\mathbb{1}(|\langle \widehat{Y}_1, y\rangle_{HS}| > \varepsilon\tau_k^*/2) + O(b^3/n^2). \tag{4.5.40}
$$

By Lemma 4 of Kokoszka and Reimherr (2013) it follows that

$$
\sum_{s=-\infty}^{\infty}\mathbb{E}\langle Z_0, y\rangle_{HS}\langle Z_s, y\rangle_{HS}
$$

converges absolutely. By Kronecker's lemma, we then get, as $n \to \infty$,

$$
\begin{aligned}
\mathbb{E}(\langle \widehat{Y}_1, y\rangle_{HS}^2) &= \frac{1}{b}\sum_{i=1}^{b}\sum_{j=1}^{b}\mathbb{E}[\langle \widehat{Z}_i, y\rangle_{HS}\langle \widehat{Z}_j, y\rangle_{HS}] \\
&= \sum_{|s|<b}\left(1-\frac{|s|}{b}\right)\mathbb{E}[\langle \widehat{Z}_0, y\rangle_{HS}\langle \widehat{Z}_s, y\rangle_{HS}] \\
&= \sum_{|s|<b}\left(1-\frac{|s|}{b}\right)\mathbb{E}[\langle Z_0, y\rangle_{HS}\langle Z_s, y\rangle_{HS}] + O(b/n^{1/2}) \\
&\to \sum_{s=-\infty}^{\infty}\mathbb{E}[\langle Z_0, y\rangle_{HS}\langle Z_s, y\rangle_{HS}].
\end{aligned}
$$

Therefore, by the dominated convergence theorem,

$$\mathbb{E}[\langle \widehat{Y}_1, y\rangle_{HS}^2)\mathbb{1}(|\langle \widehat{Y}_1, y\rangle_{HS}| > \varepsilon\tau_k^*/2) = o(1) \tag{4.5.41}$$

and, therefore, assertion (i) is proved.

To establish assertion (ii), notice first that

$$\mathbb{E}^* \left\| \frac{1}{\sqrt{k}} \sum_{t=1}^{k} \widehat{Y}_t^* \right\|_{HS}^2 = \mathbb{E}^* \|\widehat{Y}_1^*\|_{HS}^2.$$

Furthermore, since

$$\begin{aligned}\mathbb{E}^* \left(\frac{1}{\sqrt{b}} \sum_{t=1}^{b} Z_t^* \right) = \frac{1}{N} \sum_{t=1}^{N} \widetilde{Y}_t = \frac{\sqrt{b}}{N} \left[\sum_{t=1}^{n} \widetilde{Z}_t - \sum_{i=1}^{b-1} \left(1 - \frac{i}{b}\right) [\widetilde{Z}_i + \widetilde{Z}_{n-i+1}] \right] \\ = \sqrt{b}\widehat{\widetilde{C}}_n - \frac{\sqrt{b}}{N} \sum_{i=1}^{b-1} \left(1 - \frac{i}{b}\right) [\widetilde{Z}_i + \widetilde{Z}_{n-i+1}],\end{aligned}$$

we get

$$\begin{aligned}\mathbb{E}^* \|\widehat{Y}_1^*\|_{HS}^2 = \mathbb{E}^* \left\| \frac{1}{\sqrt{b}} \sum_{t=1}^{b} \widehat{Z}_t^* + \frac{\sqrt{b}}{N} \sum_{i=1}^{b-1} \left(1 - \frac{i}{b}\right) [\widetilde{Z}_i + \widetilde{Z}_{n-i+1}] \right\|_{HS}^2 \\ = \frac{1}{N} \sum_{t=1}^{N} \left\| \widehat{Y}_t + \frac{\sqrt{b}}{N} \sum_{i=1}^{b-1} \left(1 - \frac{i}{b}\right) [\widetilde{Z}_i + \widetilde{Z}_{n-i+1}] \right\|_{HS}^2 .\end{aligned}$$

Since, $\sqrt{b}N^{-1}\sum_{i=1}^{b-1}\left(1 - \frac{i}{b}\right) [\widetilde{Z}_i + \widetilde{Z}_{n-i+1}] = O_P(b^{3/2}/n)$, it suffices to prove that the limit

$$\lim_{n\to\infty} \frac{1}{N} \sum_{t=1}^{N} \|\widehat{Y}_t\|_{HS}^2 \tag{4.5.42}$$

exists and it is finite. Let $Y_t = b^{-1/2}(Z_t + \cdots + Z_{t+b-1})$, $t = 1, 2, \ldots N$, and note that $N^{-1}\sum_{t=1}^{N} \|\widehat{Y}_t\|_{HS}^2 = N^{-1}\sum_{t=1}^{N} \|Y_t + \sqrt{b}(\mathcal{C}_0 - \widehat{\widetilde{\mathcal{C}}}_0)\|_{HS}^2$. By Theorem 3 of Kokoszka and Reimherr (2013), in order to prove (4.5.42), it suffices to show that

$$\lim_{n\to\infty} \frac{1}{N} \sum_{t=1}^{N} \|Y_t\|_{HS}^2 \tag{4.5.43}$$

exists and it is finite. We have that

$$\frac{1}{N} \sum_{t=1}^{N} \|Y_t\|_{HS}^2 = \frac{1}{N}\langle Z_t, Z_t\rangle_{HS} + \sum_{i=1}^{b-1} \left(1 - \frac{i}{b}\right) \frac{1}{N} \sum_{t=1}^{n-i} [\langle Z_t, Z_{t+i}\rangle_{HS} + \langle Z_{t+i}, Z_t\rangle_{HS}]$$

$$
\begin{aligned}
&- \frac{1}{N}\sum_{t=1}^{b-1}\left(1-\frac{t}{b}\right)[\langle Z_t, Z_t\rangle_{HS} + \langle X_{n-t+1}, X_{n-t+1}\rangle]_{HS} \\
&- \frac{1}{N}\sum_{t=1}^{b-1}\sum_{j=1}^{b-t}\left(1-\frac{t+j}{b}\right)[\langle Z_j, Z_{j+t}\rangle_{HS} + \langle Z_{n-j+1-t}, Z_{n-j+1}\rangle_{HS} \\
&\qquad + \langle Z_{j+t}, Z_j\rangle_{HS} + \langle Z_{n-j+1}, Z_{n-j+1-t}\rangle_{HS}] \\
&= \frac{1}{N}\sum_{t=1}^{n}\langle Z_t, Z_t\rangle_{HS} + \sum_{i=1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{N}\sum_{t=1}^{n-i}[\langle Z_t, Z_{t+i}\rangle_{HS} + \langle Z_{t+i}, Z_t\rangle_{HS}] \\
&\qquad + O_P(b^2/n) \\
&= \sum_{i=-b+1}^{b-1}\left(1-\frac{i}{b}\right)\frac{1}{n}\sum_{t=1}^{n-|i|}\iint Z_t(u,v)Z_{t+|i|}(u,v)\mathrm{d}u\mathrm{d}v \\
&\qquad + O_P(b^2/n). \qquad (4.5.44)
\end{aligned}
$$

Hence, by letting $g_b(s) = (1-|s|/b)$ in Lemma 4.5.2, we get that the last term above converges to $\sum_{s=-\infty}^{\infty}\mathbb{E}\iint Z_0(u,v)Z_s(u,v)\mathrm{d}u\mathrm{d}v$, from which we conclude that, as $n \to \infty$,

$$\mathbb{E}^*\|Y_1^*\|_{HS}^2 \to \sum_{s=-\infty}^{\infty}\mathbb{E}\iint Z_0(u,v)Z_s(u,v)\mathrm{d}u\mathrm{d}v,$$

in probability. □

Proof of Lemma 4.3.1. Using Theorem 3 of Kokoszka and Reimherr (2013) it follows that there exist two independent, mean zero, Gaussian Hilbert-Schmidt operators $\mathcal{Z}_{1,0}$ and $\mathcal{Z}_{2,0}$ with covariance operators $\Gamma_{1,0}$ and $\Gamma_{2,0}$ respectively, such that

$$\left(\sqrt{n_1}(\widehat{\mathcal{C}}_{1,0} - \mathcal{C}_{1,0}), \sqrt{n_2}(\widehat{\mathcal{C}}_{2,0} - \mathcal{C}_{2,0})\right)$$

converges weakly to $(\mathcal{Z}_{1,0}, \mathcal{Z}_{2,0})$. Since

$$\sqrt{\frac{n_1 n_2}{M}}(\widehat{\mathcal{C}}_{1,0} - \widehat{\mathcal{C}}_{2,0}) = \sqrt{\frac{n_2}{M}}\sqrt{n_1}(\widehat{\mathcal{C}}_{1,0} - \widetilde{\mathcal{C}}_0) - \sqrt{\frac{n_1}{M}}\sqrt{n_2}(\widehat{\mathcal{C}}_{2,0} - \widetilde{\mathcal{C}}_0),$$

where $\widetilde{\mathcal{C}}_0$ is the (under H_0) common lag-zero covariance operator of the two populations, we get that, for $n_1, n_2 \to \infty$ and $n_1/M \to \theta$,

$$T_M \xrightarrow{d} \|\mathcal{Z}_0\|_{HS}^2,$$

where $\mathcal{Z}_0 = \sqrt{1-\theta}\mathcal{Z}_{1,0} - \sqrt{\theta}\mathcal{Z}_{2,0}$. □

Proof of Theorem 4.3.1. Using the triangle inequality and the fact that $\sqrt{n}(\widehat{\mathcal{C}}_{i,0} - \mathcal{C}_{i,0}) \Rightarrow \mathcal{Z}_{i,0}$, $i = 1, 2$, it suffices to prove that T_M^* converges weakly to $\|\mathcal{Z}_0\|_{HS}^2$, where $\mathcal{Z}_0 = \sqrt{1-\theta}\mathcal{Z}_{1,0} - \sqrt{\theta}\mathcal{Z}_{2,0}$. This is proved along the same lines as Lemma 4.3.1 using of Theorem 4.2.1 and the independence of the pseudo-random elements $\overline{\mathcal{Y}}^*_{1,n_1}$ and $\overline{\mathcal{Y}}^*_{2,n_2}$. □

5

Conclusion And Further Work

"The outcome of any serious research can only be to make two questions grow where only one grew before."

– Thorstein Veblen

This thesis lies in the intersection of Functional Time Series analysis and Bootstrap Methods. As stated in the Introduction, our main concerns are to establish validity of the moving block and tapered block bootstrap for depended functional data and to propose general, bootstrap based procedures, to address the important topic of testing the equality of mean functions or the equality of covariance operators between k-populations. For the purpose of our research we focus on observations stemming for a stationary stochastic process $\mathbb{X} = (X_t,\, t \in \mathbb{Z})$ of Hilbert space-valued random variables which is L^p-m-approximable, a dependence assumption which is satisfied by large classes of commonly used functional time series models.

More specifically, and as far as the first aim of this thesis is concerned, our contribution is to prove a central limit theorem for the moving block bootstrap and for the tapered block bootstrap applied to the sample mean function. We also show that these block bootstrap procedures provide consistent estimation of the so called log run covariance operator of i.e., of the spectral density operator of the underlying functional process at frequency zero. We also prove a central limit theorem for the moving block bootstrap procedure applied to the sample covariance operator.

Regarding the second aim of this thesis, we proposed moving block and tapered block bootstrap procedures for testing the equality of mean functions and a moving block bootstrap procedure for testing the equality of covariance operators, between several independent functional time series. In each case the bootstrap pseudo-observations were generated in a way that the null hypothesis of interest is satisfied. Therefore, the proposed testing methodologies are applicable to a broad range of possible test statistics. We have focused in this thesis on testing approaches based on fully functional test statistics. For the proposed testing algorithms theoretical justifications for approximating the null distribution of the test statistics considered are given. Furthermore, simulation results are presented which investigate the finite sample behaviour of the proposed block bootstrap-based testing methodologies under the null and under the alternative.

There are a number of questions for further research that arose during the present study.

Firstly, our testing procedures focus on k-independent populations. Future research could examine if the results obtained in this thesis can be extended to the case where the populations are dependent. This requires the adaption of the block bootstrap procedures to capture the dependent structure between the populations and the proof of the corresponding central limit theorems.

Another interesting question from this work is the development of a bootstrap based procedure which will allow inference for the spectral density operators itself. An important problem in this context is that of testing the equality of the spectral density operators of k independent, or dependent, functional time series.

As proved by Paparoditis and Politis (2001) in the case where the random variables are finite-dimensional, the TBB gives a better estimation for the standard deviation of the normalised sample mean compared to the MBB since in the TBB case, the order of the bias of the estimator is $O(1/b^2)$ while in the MBB case, the order of the bias of the estimator is $O(1/b)$, where b is the block length in the block bootstrap procedures. As revealed from the simulations of Section 3.4.1 the tapered block bootstrap procedure, also gives a better estimation for the standard deviation of the normalised sample mean function in the case of functional data. It would be important to have a theoretical justification of this improved behavior of the TBB for functional time series.

In this context the development of a tapered block bootstrap based procedure for testing the equality of the lag-zero autocovariance operators for a given number of functional time series would be an interesting extension of the current work.

Further, as obtained from the simulations carried out for this thesis, the performance of all block bootstrap procedures considered essentially depends on the choice of block size b. A challenging research topic would be to propose methods for selecting this bootstrap parameter which derives ‘good’ results or even ‘optimal’ in some sense.

Bibliography

[1] Benko, M., Härdle, W. and Kneip, A. (2009). Common functional principal components. *The Annals of Statistics*, Vol. **37**, 1–34.

[2] Beutner, E. and Zähle, H. (2016). Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electronic Journal of Statistics*, Vol. **10**, 1181-1222.

[3] Billingsley, P. (1999). *Convergence of Probability Measures.* 2nd Edition, New York: Wiley.

[4] Cerovecki, C. and Hörmanm, S. (2017). On the CLT for the discrete Fourier transforms of functional time series. *Journal of Multivariate Analysis*, Vol. **154**, 282–295.

[5] Chen, X., and White, H. (1998). Central Limit and Functional Central Limit Theorems for Hilbert-Valued Dependent Heterogeneous Arrays with Applications. *Econometric Theory*, Vol. **14(2)**, 260–284.

[6] Dehling, H., Sharipov Sh.O. and Wendler, M. (2015). Bootstrap for dependent Hilbert space–valued random variables with application to von Mises statistics. *Journal of Multivariate Analysis*, Vol. **133**, 200–215.

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

[8] Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *The Annals of Probability*, Vol. **18**, 851–869.

[9] Hörmann, S., Kidziński, Ł. and Hallin, M. (2015). Dynamic functional principal components. *Journal of the Royal Statistical Society*, Series B, Vol. **77**, 319–348.

[10] Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *The Annals of Statistics.* Vol. **38**, 1845–1884.

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

[12] Horváth, L. and Rice, G. (2015). Testing equality of means when the observations are from functional time series. *Journal of Time Series Analysis*, Vol. **36**, 84–108.

[13] Horváth, L., Rice, G. and Whipple, S. (2016). Adaptive bandwidth selection in the long run covariance estimator of functional time series. *Computational Statistics and Data Analysis*, Vol. **100**, 676–693.

[14] Horváth, L., Kokoszka, P. and Reeder, R. (2013). Estimation of the mean of functional time series and a two-sample problem. *Journal of the Royal Statistical Society*, Series B, Vol. **75**, 103–122.

[15] Kokoszka, P. and Reimherr, M. (2013). Asymptotic normality of the principal components of functional time series. *Stochastic Processes and their Applications*, Vol. **123**, 1546–1562.

[16] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, Vol. **17**, 1217–1261.

[17] Laha, R.G. and Rohatgi, V.K. (1979). *Probability Theory.* New York: Wiley.

[18] Lahiri, S. (2003). *Resampling Methods for Dependent Data.* New York: Springer-Verlag.

[19] Lele, S. and Carlstein, E. (1990). Two-sample bootstrap tests: when to mix? *Institute of Statistics Mimeo Series*, No. **2031**, Department of Statistics, North Carolina State University, USA.

[20] Liu, R.Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In "Exploring the Limits of the Bootstrap" (R. Lepage and L. Billard, Eds.), pp. 225–248, New York: Wiley.

[21] Mas, A. (2002). Weak convergence for the covariance operators of a Hilbertian linear process. *Stochastic Processes and their Applications*, Vol. **99**, 117–135.

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

[23] Panaretos, V.M. and Tavakoli, S. (2013a). Cramér-Karhunen-Loève representation and harmonic principal component analysis of functional time series. *Stochastic Processes and Their Applications*, Vol. **123**, 2779–2807.

[24] Panaretos, V.M. and Tavakoli, S. (2013b). Fourier analysis of stationary time series in function space. *Annals of Statistics*, Vol. **41**, 568–603.

[25] Paparoditis, E. (2018). Sieve bootstrap for functional time series. *Annals of Statistics*, Vol. **46**, 3510–3538.

[26] Paparoditis, E. and Politis, D. (2001). Tapered block bootstrap. *Biometrika*, Vol. **88**, 1105–1119.

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

[28] Politis, D.N. and Romano, J.P. (1994). Limit theorems for weakly dependent Hilbert space valued random variables with applications to the stationary bootstrap. *Statistica Sinica*, Vol. **4**, 461–476.

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

[30] Prokhorv, V. Y. (1956) Convergence of random processes and limit theorems in probability theory. *Theory of Probability and Its Applications*, Vol. **1**, 157–214.

[31] Ramsay, J.O. and Dalzell, C.J. (1991). Some tools for functional data analysis (with discussion). *Journal of the Royal Statistical Society*, Series B, Vol. **53**, 539–572.

[32] Rice, G. and Shang, H.L. (2017) A plug-in bandwidth selection procedure for long run covariance estimation with stationary functional time series. *Journal of Time Series Analysis*, Vol. **38**, 591–609.

[33] Staicu, A.M., Lahiri, S.N. and Carroll, R.J. (2015). Significance tests for functional data with complex dependence structure. *Journal of Statistical Planning and Inference*, Vol. **156**, 1–13.

[34] van der Vaart, A. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes.* New York: Springer.

[35] Zhang, C., Peng, H. and Zhang, J.-T. (2010). Two samples tests for functional data. *Communications in Statistics - Theory and Methods*, Vol. **39**, 559–578.

[36] Zhang, J.-T. (2013). *Analysis of Variance for Functional Data.* Boca Raton: Chapman & Hall/CRC.

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