



Detecting deviations from second-order stationarity in locally stationary functional time series

Axel Bücher¹ · Holger Dette² · Florian Heinrichs²

Received: 10 August 2018 / Revised: 16 February 2019 / Published online: 17 June 2019
© The Institute of Statistical Mathematics, Tokyo 2019

Abstract

A time-domain test for the assumption of second-order stationarity of a functional time series is proposed. The test is based on combining individual cumulative sum tests which are designed to be sensitive to changes in the mean, variance and autocovariance operators, respectively. The combination of their dependent p values relies on a joint-dependent block multiplier bootstrap of the individual test statistics. Conditions under which the proposed combined testing procedure is asymptotically valid under stationarity are provided. A procedure is proposed to automatically choose the block length parameter needed for the construction of the bootstrap. The finite-sample behavior of the proposed test is investigated in Monte Carlo experiments, and an illustration on a real data set is provided.

Keywords Alpha mixing · CUSUM test · Autocovariance operator · Block multiplier bootstrap · Change points

1 Introduction

Within the last decades, statistical analysis for functional time series has become a very active area of research [see the monographs [Bosq \(2000\)](#), [Ferraty and Vieu](#)

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10463-019-00721-7>) contains supplementary material, which is available to authorized users.

✉ Holger Dette
holger.dette@rub.de

Axel Bücher
axel.buecher@hhu.de

Florian Heinrichs
florian.heinrichs@rub.de

¹ Mathematisches Institut, Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

² Fakultät für Mathematik, Ruhr-Universität Bochum, Universitätsstr. 150, 44780 Bochum, Germany

(2006), Horváth and Kokoszka (2012) and Hsing and Eubank (2015), among others]. Many authors impose the assumption of stationarity, which allows for developing advanced statistical theory. For instance, Bosq (2002) and Dehling and Sharipov (2005) investigate stationary functional processes with a linear representation and Hörmann and Kokoszka (2010) provide a general framework to model functional observations from stationary processes. Frequency domain analysis of stationary functional time series has been considered by Panaretos and Tavakoli (2013), while van Delft and Eichler (2018) propose a new concept of local stationarity for functional data. The assumption of second-order stationarity is also of particular importance for prediction problems [see Antoniadis and Sapatinas (2003), Aue et al. (2015), Hyndman and Shang (2009) among others] and for dynamic principal component analysis (Hörmann et al. 2015).

Ideally, the assumption of stationarity should be checked before applying any statistical methodology. Several authors have considered this problem, in particular within the context of change point analysis where the null hypothesis of stationarity is tested against the alternative of a structural change in certain parameters of the process; see Aue et al. (2009), Berkes et al. (2009), Horvath et al. (2010), Aston and Kirch (2012) among others. Tests that are designed to be powerful against more general alternatives are often based on an analysis in the frequency domain. For example, Aue and van Delft (2017) generalize the approach of Dwivedi and Subba Rao (2011) and Jentsch and Subba Rao (2015) to functional time series. More precisely, they begin by showing that the functional discrete Fourier transform (fDFT) is asymptotically uncorrelated at distinct Fourier frequencies if and only if the process is weakly stationary. The corresponding test is then based on a quadratic form based on a finite-dimensional projection of the empirical covariance operator of the fDFT's. Consequently, the properties of the test depend on the number of lagged fDFT's included. As an alternative, van Delft et al. (2017) construct a test using an estimate of a minimal distance between the spectral density operator of a non-stationary process and its best approximation by a spectral density operator corresponding to a stationary process (see also Dette et al. 2011 for a discussion of this approach in the univariate context). The test statistic consists of sums of Hilbert–Schmidt inner products of periodogram operators (evaluated at different frequencies) and is asymptotically normal distributed.

In the present paper, we propose an alternative time-domain test for second-order stationarity of a functional time series. More precisely, we suggest to address the problem of detecting non-stationarity by individually checking the hypothesis that the mean and the autocovariance operators at a given lag, say h , of a collection (indexed by time) of approximating stationary functional time series are in fact time independent. As explained in the next paragraph, the individual tests are then combined to yield a joint test including autocovariances up to a given maximal lag H . Thus, the approach investigated here is similar in spirit to the classical Portmanteau tests for serial correlation of a univariate time series, where the hypothesis of white noise is checked by investigating whether correlations up to a given lag vanish (see Box and Pierce 1970; Ljung and Box 1978). For the problem of checking stationarity in real-valued time series, similar approaches have been taken by Jin et al. (2015) and Bücher et al. (2018).

To combine the individual tests for stationarity of the mean and the autocovariance operators at a given lag h , we use appropriate extensions of well-known p value combination methods dating back to Fisher (1932). Each individual test is relying on a block multiplier approach making necessary the choice of a joint block length parameter m . Following ideas put forward in Politis and White (2004), a procedure is proposed to automatically select that parameter data adaptively in such a way that, asymptotically, a certain MSE criterion is optimized.

The remaining part of this article is organized as follows: In Sect. 2, we collect necessary mathematical preliminaries. In Sect. 3, we first propose individual tests for the hypothesis of second-order stationarity which are particularly sensitive to deviations in the mean, variance and a given lag h autocovariance, respectively. The tests are then combined to a joint test for second-order stationarity which is sensitive to deviations in the mean, variance and the first H autocovariances. In Sect. 4, we discuss an exemplary locally stationary time series model in great theoretical detail, while finite-sample results and a case study are presented in Sect. 5. The central proofs are collected in Sect. 6, while less central proofs and auxiliary results are provided in a supplementary material.

2 Mathematical preliminaries

2.1 Random elements in L^p -spaces

For some separable measurable space (S, \mathcal{S}, ν) with a σ -finite measure ν and $p > 1$, let $\mathcal{L}^p(S, \nu)$ denote the set of measurable functions $f : S \rightarrow \mathbb{R}$ such that $\|f\|_p = (\int |f|^p d\nu)^{1/p} < \infty$. For $f \in \mathcal{L}^p(S, \nu)$, let $[f]$ be the set of all functions g such that $f = g$, ν -almost surely. The space $L^p(S, \nu)$ of all equivalence classes $[f]$ then becomes a separable Banach space, and standard weak convergence theory is applicable. If S is a subset of \mathbb{R}^d and ν is the Lebesgue measure, we occasionally write $\mathcal{L}^p(S)$ and $L^p(S)$.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ denote a probability space and let $X : S \times \Omega \rightarrow \mathbb{R}$ be $(\mathcal{S} \otimes \mathcal{A})$ -measurable such that $X(\cdot, \omega) \in \mathcal{L}^p(S, \mu)$ for \mathbb{P} -almost every ω . It follows from Lemma 6.1 in Janson and Kaijser (2015) that $\omega \mapsto [X(\cdot, \omega)]$ is a random variable in $L^p(S, \mu)$ (equipped with the Borel σ -field). Conversely, note that for any random variable $[Y]$ in $L^p(S, \mu)$, we can choose a $(\mu \otimes \mathbb{P})$ -a.s. unique $(\mathcal{S} \otimes \mathcal{A})$ -measurable mapping $Y' : S \times \Omega \rightarrow \mathbb{R}$ such that $Y'(\cdot, \omega) \in [Y](\omega)$ for \mathbb{P} -almost every ω . We can hence (a.s.) identify random variables in $L^p(S, \mu)$ with measurable functions on $S \times \Omega$ which are p -integrable in the first argument (\mathbb{P} -a.s.); slightly abusing notation, we also write X for the equivalence class $[X]$.

A random variable X in $L^2([0, 1]^d)$ is called integrable if $\mathbb{E}\|X\|_2 < \infty$. In that case, it follows from the Riesz representation theorem that there exists a unique element $\mu_X = \mathbb{E}X \in L^2([0, 1]^d)$ such that $\mathbb{E}\langle X, f \rangle = \langle \mu_X, f \rangle$ for all $f \in L^2([0, 1]^d)$, where $\langle f, g \rangle = \int_{[0, 1]^d} fg \, d\lambda_d$. If X is even square integrable, that is, $\mathbb{E}\|X\|_2^2 < \infty$, the covariance operator of X is defined as the operator $C_X : L^2([0, 1]^d) \rightarrow L^2([0, 1]^d)$ given by $C_X(f) = \mathbb{E}[\langle f, X - \mu_X \rangle (X - \mu_X)]$. C_X is nuclear and hence a Hilbert–

Schmidt operator (Bosq 2000, Sect. 1.5), whence, by Theorem 6.11 in Weidmann (1980), there exists a kernel $c_X \in L^2([0, 1]^d \times [0, 1]^d)$ such that

$$C_X(f)(\tau) = \int_{[0,1]^d} c_X(\tau, \sigma) f(\sigma) d\sigma$$

for almost every $\tau \in [0, 1]^d$ and every $f \in L^2([0, 1]^d)$. Similarly, for square integrable random elements $X, Y \in L^2([0, 1]^d)$ we define the cross-covariance operator $C_{X,Y} : L^2([0, 1]^d) \rightarrow L^2([0, 1]^d)$ by $C_{X,Y}(f) = \mathbb{E}[\langle X - \mu_X, f \rangle (Y - \mu_Y)]$. By the same reasoning as above, there exists a kernel $c_{X,Y} \in L^2([0, 1]^d \times [0, 1]^d)$ such that

$$C_{X,Y}(f)(\tau) = \int_{[0,1]^d} c_{X,Y}(\tau, \sigma) f(\sigma) d\sigma.$$

If X is in fact a $(\mathcal{B}([0, 1]^d) \otimes \mathcal{A})$ -measurable function from $[0, 1]^d \times \Omega$ to \mathbb{R} with $X(\cdot, \omega) \in L^2([0, 1]^d)$ a.s., then it can be shown that, in the respective L^2 -spaces,

$$\begin{aligned} \mu(\tau) &= \mathbb{E}[X(\tau)], \\ c_X(\tau, \sigma) &= \text{Cov}\{X(\tau), X(\sigma)\}, \quad c_{X,Y}(\tau, \sigma) = \text{Cov}\{X(\tau), Y(\sigma)\}. \end{aligned}$$

By the preceding paragraph, this notation also makes sense for equivalence classes $X, Y \in L^2([0, 1]^d)$.

2.2 Functional time series in $L^2([0, 1])$

For each $t \in \mathbb{Z}$, let $X_t : [0, 1] \times \Omega \rightarrow \mathbb{R}$ denote a $(\mathcal{B}|_{[0,1]} \otimes \mathcal{A})$ -measurable function with $X_t(\cdot, \omega) \in L^2([0, 1])$. By the preceding section, we can regard $[X_t]$ as a random variable in $L^2([0, 1])$, which we also write as X_t . The sequence $(X_t)_{t \in \mathbb{Z}}$ will be referred to as a *functional time series* in $L^2([0, 1])$.

The functional time series will be called *stationary* if, for all $q \in \mathbb{Z}$ and all $h, t_1, \dots, t_q \in \mathbb{Z}$

$$(X_{t_1+h}, \dots, X_{t_q+h}) \stackrel{d}{=} (X_{t_1}, \dots, X_{t_q})$$

in $L^2([0, 1])^q$.

Let $\rho > 0$. A sequence of functional time series $(X_{t,T})_{t \in \mathbb{Z}}$, indexed by $T \in \mathbb{N}$, is called *locally stationary (of order ρ)* if, for any $u \in [0, 1]$, there exists a strictly stationary functional time series $\{X_t^{(u)} \mid t \in \mathbb{Z}\}$ in $L^2([0, 1])$ and an array of real-valued random variables $\{P_{t,T}^{(u)} \mid t = 1, \dots, T\}_{T \in \mathbb{N}}$ with $\mathbb{E}|P_{t,T}^{(u)}|^\rho < \infty$, uniformly in $1 \leq t \leq T$, $T \in \mathbb{N}$ and $u \in [0, 1]$, such that

$$\|X_{t,T} - X_t^{(u)}\|_2 \leq \left(\left| \frac{t}{T} - u \right| + \frac{1}{T} \right) P_{t,T}^{(u)} \tag{1}$$

for all $t = 1, \dots, T, T \in \mathbb{N}$ and $u \in [0, 1]$. This concept of local stationarity was first introduced by Vogt (2012) for p -dimensional time series ($p \in \mathbb{N}$). By the arguments in the preceding section, we may assume that $X_t^{(u)}$ is in fact a $(\mathcal{B}([0, 1]) \times \mathcal{A})$ -measurable function from $[0, 1] \times \Omega$ to \mathbb{R} such that $X_t^{(u)}(\cdot, \omega) \in \mathcal{L}^2([0, 1])$ for \mathbb{P} -almost every ω . In the subsequent sections, we will usually assume that $\rho \geq 2$ and that $\mathbb{E}[\|X_t^{(u)}\|_2^2] < \infty$ for all $u \in [0, 1]$. Despite the fact that $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ is a sequence of time series, we will occasionally simply call $(X_{t,T})_{t \in \mathbb{Z}}$ a locally stationary time series.

2.3 Further notation

In the following, we will deal with different norms on the spaces $L^p([0, 1]^d)$, for $p \geq 1, d \in \mathbb{N}$. To avoid confusion, we denote the corresponding norms by $\|\cdot\|_{p,d}$. As a special case, we will write $\|\cdot\|_p$ instead of $\|\cdot\|_{p,1}$. Further, we introduce the notation $\|\cdot\|_{p,\Omega \times [0,1]^d}$ for the p -norm on the space $L^p(\Omega \times [0, 1]^d, \mathbb{P} \otimes \lambda^d)$. Finally, we define $(f \otimes g)(x, y) = f(x)g(y)$ for functions $f, g \in L^p([0, 1])$.

3 Detecting deviations from second-order stationarity

3.1 Second-order stationarity in locally stationary time series

Before we can propose suitable test statistics for detecting deviations from second-order stationarity in a locally stationary functional time series, we need to clarify what is meant by second-order stationarity. Loosely speaking, we want to test the null hypothesis that the mean and/or the (auto)covariances do not vary too much over time. Meaningful asymptotic results will be obtained by formulating these null hypotheses in terms of the approximating sequences $\{X_t^{(u)} : t \in \mathbb{Z}\}$ defined in Sect. 2.2. More precisely, we will subsequently assume that $\mathbb{E}[\|X_t^{(u)}\|_2^2] < \infty$ for all $u \in [0, 1]$ and consider the hypotheses

$$H_0^{(m)} : \|\mathbb{E}[X_0^{(u)}] - \mathbb{E}[X_0^{(v)}]\|_2 = 0 \quad \text{for all } u, v \in [0, 1] \tag{2}$$

and, for some lag $h \geq 0$,

$$H_0^{(c,h)} : \|\mathbb{E}[X_0^{(u)} \otimes X_h^{(u)}] - \mathbb{E}[X_0^{(v)} \otimes X_h^{(v)}]\|_{2,2} = 0 \quad \text{for all } u, v \in [0, 1]. \tag{3}$$

Note that the intersection

$$H_0 = H_0^{(m)} \cap H_0^{(c,0)} \cap H_0^{(c,1)} \cap \dots$$

corresponds to the case where the approximating sequences $\{X_t^{(u)} : t \in \mathbb{Z}\}$, indexed by $u \in [0, 1]$, all share the same first- and second-order characteristics. We will therefore call the sequence of time series $(X_{t,T})_{t \in \mathbb{Z}}$, indexed by $T \in \mathbb{N}$, second-order stationary if the global hypothesis H_0 is met. The test statistics we are going to propose will be

particularly sensitive to deviations from (weak) stationarity in the mean, the variance, and the first H autocovariances, which leads us to define

$$H_0^{(H)} = H_0^{(m)} \cap H_0^{(c,0)} \cap H_0^{(c,1)} \cap \dots \cap H_0^{(c,H)}, \tag{4}$$

where $H \in \mathbb{N}_0$ is fixed and denotes the maximum number of lags under consideration.

Remark 1 The hypotheses $H_0^{(m)}$ and $H_0^{(c,h)}$ are independent of the choice of the approximating family $\{X_t^{(u)} : t \in \mathbb{Z}\}_{u \in [0,1]}$. Indeed, suppose there were two approximating families $\{X_t^{(u)} : t \in \mathbb{Z}\}_{u \in [0,1]}$ and $\{Y_t^{(u)} : t \in \mathbb{Z}\}_{u \in [0,1]}$ satisfying (1). By stationarity and the triangle inequality, we have, for any $t, T \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}\|X_t^{(u)} - Y_t^{(u)}\|_2 &= \mathbb{E}\|X_{\lfloor uT \rfloor}^{(u)} - Y_{\lfloor uT \rfloor}^{(u)}\|_2 \\ &\leq \mathbb{E}\|X_{\lfloor uT \rfloor}^{(u)} - X_{\lfloor uT \rfloor, T}\|_2 + \mathbb{E}\|X_{\lfloor uT \rfloor, T} - Y_{\lfloor uT \rfloor}^{(u)}\|_2 \leq \frac{C}{T}. \end{aligned}$$

This implies $\mathbb{E}\|X_t^{(u)} - Y_t^{(u)}\|_2 = 0$ and hence $\|X_t^{(u)} - Y_t^{(u)}\|_2 = 0$ almost surely. \square

The following lemma provides two interesting equivalent formulations of each of the above hypotheses. Introduce the notations $M : [0, 1]^2 \rightarrow \mathbb{R}$, $M_h : [0, 1]^3 \rightarrow \mathbb{R}$, where

$$M(u, \tau) = \int_0^u \mathbb{E}[X_0^{(w)}(\tau)] dw - u \int_0^1 \mathbb{E}[X_0^{(w)}(\tau)] dw, \tag{5}$$

$$M_h(u, \tau_1, \tau_2) = \int_0^u \mathbb{E}[X_0^{(w)}(\tau_1)X_h^{(w)}(\tau_2)] dw - u \int_0^1 \mathbb{E}[X_0^{(w)}(\tau_1)X_h^{(w)}(\tau_2)] dw. \tag{6}$$

Lemma 1 Let $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ denote a locally stationary functional time series of order $\rho \geq 4$ with approximating sequences $(X_t^{(u)})_{t \in \mathbb{Z}}$ satisfying $\mathbb{E}[\|X_0^{(u)}\|_2^4] < \infty$ for all $u \in [0, 1]$. Then, the hypothesis $H_0^{(m)}$ in (2) is met if and only if

$$\|M\|_{2,2} = 0. \tag{7}$$

Likewise, for any $h \in \mathbb{N}_0$, $H_0^{(c,h)}$ in (3) is met if and only if

$$\|M_h\|_{2,3} = 0. \tag{8}$$

Moreover, the hypothesis $H_0^{(m)}$ is equivalent to

$$\exists C > 0 : \quad \|\mathbb{E}[X_{\lfloor uT \rfloor, T}] - \mathbb{E}[X_{0,T}]\|_2 \leq \frac{C}{T} \text{ for all } u \in [0, 1], T \in \mathbb{N}, \tag{9}$$

and $H_0^{(c,h)}$ is equivalent to

$$\begin{aligned} \exists C > 0 : \quad &\|\mathbb{E}[X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor + h, T} - X_{0,T} \otimes X_{h,T}]\|_{2,2} \leq \frac{C}{T} \\ &\text{for all } u \in [0, 1], T \in \mathbb{N}. \end{aligned} \tag{10}$$

The lemma is proven in Sect. 6.2. We will heavily rely on conditions (7) and (8) when constructing the test statistics in the next section. Assertions (9) and (10) are interesting in their own rights, as they provide a sub-asymptotic formulation of the hypothesis of second-order stationarity. They are used in the next section for showing that the tests are consistent and will also be crucial when extending consistency results to the case of piecewise locally stationary processes in Sect. 3.5.

3.2 Test statistics

In the subsequent sections, we assume to observe, for some $T \in \mathbb{N}$, an excerpt $X_{1,T}, \dots, X_{T,T}$ from a locally stationary time series $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$. We are interested in testing the hypotheses $H_0^{(m)}$ and $H_0^{(c,h)}$ formulated in the preceding section, which can be done individually by a CUSUM-type procedure. More precisely, for $u, \tau \in [0, 1]$, let

$$U_T(u, \tau) = \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor} X_{t,T}(\tau) - u \sum_{t=1}^T X_{t,T}(\tau) \right) \tag{11}$$

denote the CUSUM process for the mean, and, for $u, \tau_1, \tau_2 \in [0, 1]$ and $h \in \mathbb{N}_0$, let

$$U_{T,h}(u, \tau_1, \tau_2) = \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} X_{t,T}(\tau_1) X_{t+h,T}(\tau_2) - u \sum_{t=1}^{T-h} X_{t,T}(\tau_1) X_{t+h,T}(\tau_2) \right) \tag{12}$$

denote the CUSUM process for the (auto)cross-moments at lag h . Under the null hypothesis $H_0^{(m)}$, $T^{-1/2}U_T(u, \tau)$ can be regarded as an estimator of the quantity $M(u, \tau)$ defined in (5), and a similar statement holds for $T^{-1/2}U_{T,h}$, which estimates the M_h in (6). Hence, by Lemma 1, it seems reasonable to reject $H_0^{(m)}$ or $H_0^{(c,h)}$ for large values of

$$\mathcal{S}_T^{(m)} = \|U_T\|_{2,2} \quad \text{or} \quad \mathcal{S}_T^{(c,h)} = \|U_{T,h}\|_{2,3}, \tag{13}$$

respectively.

Alternatively, one could use the L^2 -norm in τ and (τ_1, τ_2) , respectively, and the supremum in u , as proposed in Sharipov et al. (2016). However, preliminary simulation results suggested that a test based on the L^2 -norm in u performs better in applications with small sample sizes.

In Sect. 3.4, we will propose a procedure that allows to combine the previous tests statistics to obtain a joint test for the combined hypothesis $H_0^{(H)}$, with maximal lag $H \in \mathbb{N}_0$ fixed. For that purpose, we will first need (asymptotic) critical values for the individual test statistics $\mathcal{S}_T^{(m)}$ and $\mathcal{S}_T^{(c,h)}$, which in turn can be deduced from the joint asymptotic distribution of the CUSUM processes in (11) and (12). The basic tools are the following partial sum processes

$$\begin{aligned} \tilde{B}_T(u, \tau) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor uT \rfloor} X_{t,T}(\tau) - \mathbb{E}[X_{t,T}(\tau)], \\ \tilde{B}_{T,h}(u, \tau_1, \tau_2) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} X_{t,T}(\tau_1) X_{t+h,T}(\tau_2) - \mathbb{E}[X_{t,T}(\tau_1) X_{t+h,T}(\tau_2)], \end{aligned}$$

where $u, \tau, \tau_1, \tau_2 \in [0, 1]$ and $h \in \mathbb{N}_0$. The expected values within the sums will be denoted by

$$\mu_{t,T}(\tau) = \mathbb{E}[X_{t,T}(\tau)] \quad \text{and} \quad \mu_{t,T,h}(\tau_1, \tau_2) = \mathbb{E}[X_{t,T}(\tau_1) X_{t+h,T}(\tau_2)].$$

The following assumptions are sufficient to guarantee weak convergence of these processes.

Condition 1 (Assumptions on the functional time series)

- (A1) *Local stationarity* The observations $X_{1,T}, \dots, X_{T,T}$ are an excerpt from a locally stationary functional time series $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ of order $\rho = 4$ in $L^2([0, 1], \mathbb{R})$.
- (A2) *Moment condition* For any $k \in \mathbb{N}$, there exists a constant $C_k < \infty$ such that $\mathbb{E}\|X_{t,T}\|_2^k \leq C_k$ and $\mathbb{E}\|X_0^{(u)}\|_2^k \leq C_k$ uniformly in $t \in \mathbb{Z}, T \in \mathbb{N}$ and $u \in [0, 1]$.
- (A3) *Cumulant condition* For any $j \in \mathbb{N}$, there is a constant $C_j < \infty$ such that

$$\sum_{t_1, \dots, t_{j-1} = -\infty}^{\infty} \|\text{cum}(X_{t_1,T}, \dots, X_{t_j,T})\|_{2,j} \leq C_j < \infty, \tag{14}$$

for any $t_j \in \mathbb{Z}$ (for $j = 1$, the condition is to be interpreted as $\mathbb{E}\|X_{t_1,T}\|_2 \leq C_1$ for all $t_1 \in \mathbb{Z}$). Further, for $k \in \{2, 3, 4\}$, there exist functions $\eta_k : \mathbb{Z}^{k-1} \rightarrow \mathbb{R}$ satisfying

$$\sum_{t_1, \dots, t_{k-1} = -\infty}^{\infty} (1 + |t_1| + \dots + |t_{k-1}|) \eta_k(t_1, \dots, t_{k-1}) < \infty$$

such that, for any $T \in \mathbb{N}, 1 \leq t_1, \dots, t_k \leq T, v, u_1, \dots, u_k \in [0, 1], h_1, h_2 \in \mathbb{Z}, Z_{t,T}^{(u)} \in \{X_{t,T}, X_t^{(u)}\}$, and any $Y_{t,h,T}(\tau_1, \tau_2) \in \{X_{t,T}(\tau_1), X_{t,T}(\tau_1) X_{t+h,T}(\tau_2)\}$, we have

- (i) $\|\text{cum}(X_{t_1,T} - X_{t_1}^{(t_1/T)}, Z_{t_2,T}^{(u_2)}, \dots, Z_{t_k,T}^{(u_k)})\|_{2,k} \leq \frac{1}{T} \eta_k(t_2 - t_1, \dots, t_k - t_1)$,
- (ii) $\|\text{cum}(X_{t_1}^{(u_1)} - X_{t_1}^{(v)}, Z_{t_2,T}^{(u_2)}, \dots, Z_{t_k,T}^{(u_k)})\|_{2,k} \leq |u_1 - v| \eta_k(t_2 - t_1, \dots, t_k - t_1)$,
- (iii) $\|\text{cum}(X_{t_1,T}, \dots, X_{t_k,T})\|_{2,k} \leq \eta_k(t_2 - t_1, \dots, t_k - t_1)$,
- (iv) $\int_{[0,1]^2} |\text{cum}(Y_{t_1,h_1,T}(\tau), Y_{t_2,h_2,T}(\tau))| d\tau \leq \eta_2(t_2 - t_1)$.

Assumption (A2) is needed to ensure existence of all cumulants. The cumulant condition (A3) is a (partially) weakened version of the assumptions made by [Lee and Subba Rao \(2017\)](#) and [Aue and van Delft \(2017\)](#) and has its origins in classical

multivariate time series analysis (see Brillinger (1981), Assumption 2.6.2). Lemma 2 shows that the cumulant conditions in (A3) hold, provided (A1), (A2), a further moment condition and a strong mixing condition are satisfied. In particular, they are met for the models employed in Sect. 5 within our simulation study (see in particular Lemma 4).

The following theorem, proven in Sect. 6.2, shows that \tilde{B}_T and $\tilde{B}_{T,h}$ jointly converge weakly with respect to the L^2 -metric. For $H \in \mathbb{N}_0$, let the Cartesian product

$$\mathcal{H}_{H+2} = L^2([0, 1]^2) \times \{L^2([0, 1]^3)\}^{H+1}$$

be equipped with the sum of the individual scalar products, such that \mathcal{H}_{H+2} is a Hilbert space itself.

Theorem 1 *Suppose that Assumptions (A1)–(A3) are met. Then, the vector $\mathbb{B}_T = (\tilde{B}_T, \tilde{B}_{T,0}, \dots, \tilde{B}_{T,H})$ converges weakly to a centered Gaussian variable $\mathbb{B} = (\tilde{B}, \tilde{B}_0, \dots, \tilde{B}_H)$ in \mathcal{H}_{H+2} with covariance operator $C_{\mathbb{B}} : \mathcal{H}_{H+2} \rightarrow \mathcal{H}_{H+2}$ defined as*

$$C_{\mathbb{B}} \begin{pmatrix} g \\ f_0 \\ \vdots \\ f_H \end{pmatrix} \begin{pmatrix} (u, \tau) \\ (u_0, \tau_{01}, \tau_{02}) \\ \vdots \\ (u_H, \tau_{H1}, \tau_{H2}) \end{pmatrix} = \begin{pmatrix} \langle r^{(m)}((u, \tau), \cdot), g \rangle + \sum_{h=0}^H \langle r_h^{(m,c)}((u, \tau), \cdot), f_h \rangle \\ \langle r_0^{(m,c)}(\cdot, (u_0, \tau_{01}, \tau_{02})), g \rangle + \sum_{h=0}^H \langle r_{0,h}^{(c)}((u_0, \tau_{01}, \tau_{02}), \cdot), f_h \rangle \\ \vdots \\ \langle r_H^{(m,c)}(\cdot, (u_H, \tau_{H1}, \tau_{H2})), g \rangle + \sum_{h=0}^H \langle r_{H,h}^{(c)}((u_H, \tau_{H1}, \tau_{H2}), \cdot), f_h \rangle \end{pmatrix}.$$

Here, the kernel functions $r^{(m)}$, $r_{h,h'}^{(c)}$ and $r_h^{(m,c)}$ are given by

$$\begin{aligned} r^{(m)}((u, \tau), (v, \varphi)) &= Cov(\tilde{B}(u, \tau), \tilde{B}(v, \varphi)) \\ &= \sum_{k=-\infty}^{\infty} \int_0^{u \wedge v} c_{k,1}(w) dw, \\ r_{h,h'}^{(c)}((u, \tau_1, \tau_2), (v, \varphi_1, \varphi_2)) &= Cov(\tilde{B}_h(u, \tau_1, \tau_2), \tilde{B}_{h'}(v, \varphi_1, \varphi_2)) \\ &= \sum_{k=-\infty}^{\infty} \int_0^{u \wedge v} c_{k,2}(w) dw, \\ r_h^{(m,c)}((u, \tau_1, \tau_2), (v, \varphi_1, \varphi_2)) &= Cov(\tilde{B}(u, \tau), \tilde{B}_h(v, \varphi_1, \varphi_2)) \\ &= \sum_{k=-\infty}^{\infty} \int_0^{u \wedge v} c_{k,3}(w) dw, \end{aligned}$$

with

$$\begin{aligned}
 c_{k,1}(w) &= c_{k,1}(w, \tau, \varphi) = \text{Cov}(X_0^{(w)}(\tau), X_k^{(w)}(\varphi)), \\
 c_{k,2}(w) &= c_{k,2}(w, h, h', \tau_1, \tau_2, \varphi_1, \varphi_2) \\
 &= \text{Cov}(X_0^{(w)}(\tau_1)X_h^{(w)}(\tau_2), X_k^{(w)}(\varphi_1)X_{k+h'}^{(w)}(\varphi_2)), \\
 c_{k,3}(w) &= c_{k,3}(w, h, \tau, \varphi_1, \varphi_2) = \text{Cov}(X_0^{(w)}(\tau), X_k^{(w)}(\varphi_1)X_{k+h}^{(w)}(\varphi_2)),
 \end{aligned}$$

for any $0 \leq h, h' \leq H$. In particular, the infinite sums and integrals converge.

The following corollary on joint weak convergence of the CUSUM processes defined in (11) and (12) is essentially a mere consequence of the continuous mapping theorem. Let

$$\begin{aligned}
 \tilde{G}_T(u, \tau) &= \tilde{B}_T(u, \tau) - u\tilde{B}_T(1, \tau) \\
 \tilde{G}_{T,h}(u, \tau_1, \tau_2) &= \tilde{B}_{T,h}(u, \tau_1, \tau_2) - u\tilde{B}_{T,h}(1, \tau_1, \tau_2) \\
 \mathbb{G}_T &= (\tilde{G}_T, \tilde{G}_{T,1}, \dots, \tilde{G}_{T,H})
 \end{aligned}$$

and, similarly,

$$\begin{aligned}
 \tilde{G}(u, \tau) &= \tilde{B}(u, \tau) - u\tilde{B}(1, \tau) \\
 \tilde{G}_h(u, \tau_1, \tau_2) &= \tilde{B}_h(u, \tau_1, \tau_2) - u\tilde{B}_h(1, \tau_1, \tau_2) \\
 \mathbb{G} &= (\tilde{G}, \tilde{G}_1, \dots, \tilde{G}_H).
 \end{aligned} \tag{15}$$

Corollary 1 *Suppose that Assumptions (A1)–(A3) are satisfied. If $H_0^{(m)}$ holds, then*

$$\|U_T - \tilde{G}_T\|_{2,2} = o_{\mathbb{P}}(1).$$

If $H_0^{(c,h)}$ holds, then

$$\|U_{T,h} - \tilde{G}_{T,h}\|_{2,3} = o_{\mathbb{P}}(1).$$

As a consequence, if the hypothesis $H_0^{(H)}$ in (4) holds, then,

$$\mathbb{U}_T = (U_T, U_{T,1}, \dots, U_{T,H}) = \mathbb{G}_T + o_{\mathbb{P}}(1) \rightsquigarrow \mathbb{G}.$$

On the other hand, if $H_0^{(m)}$ or $H_0^{(c,h)}$ does not hold, then $\mathcal{S}_T^{(m)} \rightarrow \infty$ or $\mathcal{S}_T^{(c,h)} \rightarrow \infty$ in probability, respectively.

The corollary suggests to reject $H_0^{(m)}$ or $H_0^{(c,h)}$ for large values of $\mathcal{S}_T^{(m)}$ or $\mathcal{S}_T^{(c,h)}$, respectively. However, the corresponding null-limiting distributions $\|\tilde{G}\|_{2,2}$ and $\|\tilde{G}_h\|_{2,3}$ depend in a complicated way on the functions $c_{k,j}$ defined in Theorem 1 and cannot be easily transformed into a pivotal distribution. We therefore propose to derive critical values by a suitable block multiplier bootstrap approximation worked out in detail in Sect. 3.4.

3.3 Strong mixing and cumulants

In this section, we will demonstrate that under the assumption of a *strong mixing* locally stationary functional time series, Assumption (A3) is met. To be precise, let \mathcal{F} and \mathcal{G} be σ -fields in (Ω, \mathcal{A}) and define

$$\alpha(\mathcal{F}, \mathcal{G}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}, B \in \mathcal{G}\}.$$

A functional time series $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ in $L^2([0, 1])$ is called α - or *strongly mixing* if the mixing coefficients

$$\alpha'(k) = \sup_{T \in \mathbb{N}} \sup_{t \in \mathbb{Z}} \alpha\left(\sigma(\{X_{s,T}(\tau) | \tau \in [0, 1]\}_{s=-\infty}^t), \sigma(\{X_{s,T}(\tau) | \tau \in [0, 1]\}_{s=t+k}^{\infty})\right)$$

vanish as k tends to infinity. Analogously, we define

$$\alpha''(k) = \sup_{u \in [0, 1]} \sup_{t \in \mathbb{Z}} \alpha\left(\sigma(\{X_s^{(u)}(\tau) | u, \tau \in [0, 1]\}_{s=-\infty}^t), \sigma(\{X_s^{(u)}(\tau) | u, \tau \in [0, 1]\}_{s=t+k}^{\infty})\right)$$

as mixing coefficients of the family of approximating stationary processes. Further, we define $\alpha(k) = \max\{\alpha'(k), \alpha''(k)\}$. A locally stationary, functional time series is called *strongly mixing*, if $\alpha(k)$ vanishes, as k tends to infinity and *exponentially strongly mixing* if $\alpha(k) \leq ca^k$ for some constants $c > 0$ and $a \in (0, 1)$. Note that we can define the mixing coefficients in terms of a function in $\mathcal{L}^2([0, 1])$ rather than an element of the space $L^2([0, 1])$ of equivalence classes by Lemma 6.1 in [Janson and Kaijser \(2015\)](#). The main result of this section provides sufficient conditions for the theory developed so far for strong mixing processes.

Lemma 2 *Let $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ be a strongly mixing locally stationary functional time series in $L^2([0, 1], \mathbb{R})$ such that Assumptions (A1), (A2) and the condition*

$$\sup_{t,T} \|X_{t,T}\|_{r, \Omega \times [0, 1]} < C_r < \infty$$

are satisfied for any integer $r > 2$. If $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ is exponentially strongly mixing, then it also satisfies the summability conditions for the cumulants in Assumption (A3).

3.4 Bootstrap approximation

The bootstrap approximation will be based on two smoothing parameters: a block length sequence $m = m_T$ needed to asymptotically catch the serial dependence within the time series, and a bandwidth sequence $n = n_T$ needed to estimate expected values locally in time. We will impose the following condition.

Condition 2 (Assumptions on the bootstrap scheme)

- (B1) Let $m = m(T) \leq T$ be an integer-valued sequence, to be understood as the block length within a block bootstrap procedure. Assume that m tends to infinity and m/T vanishes, as $T \rightarrow \infty$.
- (B2) Let $n = n(T) \leq T/2$ be an integer-valued sequence such that both m/n and mn^2/T^2 converge to zero, as T tends to infinity.
- (B3) Let $\{R_i^{(k)}\}_{i,k \in \mathbb{N}}$ denote independent standard normally distributed random variables, independent of the stochastic process $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$.

Under this set of notations, we define

$$\hat{B}_T^{(k)}(u, \tau) = \frac{1}{\sqrt{T}} \sum_{i=1}^{\lfloor uT \rfloor} \frac{R_i^{(k)}}{\sqrt{m}} \sum_{t=i}^{(i+m-1) \wedge T} \{X_{t,T}(\tau) - \hat{\mu}_{t,T}(\tau)\}$$

as a bootstrap approximation for $\tilde{B}_T(u, \tau)$, where

$$\hat{\mu}_{t,T}(\tau) = \frac{1}{\tilde{n}_{t,0}} \sum_{j=\underline{n}_t}^{\tilde{n}_{t,0}} X_{t+j,T}(\tau)$$

denotes an estimator for $\mu_{t,T}(\tau)$ relying on the bandwidth sequence n via

$$\tilde{n}_{t,h} = n \wedge (T - t - h), \quad \underline{n}_t = -n \vee (1 - t), \quad \tilde{n}_{t,h} = \tilde{n}_{t,h} - \underline{n}_t + 1, \tag{16}$$

for $0 \leq h \leq H$. Similarly, for any $0 \leq h \leq H$, bootstrap approximations for $\tilde{B}_{T,h}(u, \tau_1, \tau_2)$ are defined as

$$\hat{B}_{T,h}^{(k)}(u, \tau_1, \tau_2) = \frac{1}{\sqrt{T}} \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} \frac{R_i^{(k)}}{\sqrt{m}} \sum_{t=i}^{(i+m-1) \wedge (T-h)} \{X_{t,T}(\tau_1)X_{t+h,T}(\tau_2) - \hat{\mu}_{t,T,h}(\tau_1, \tau_2)\},$$

where

$$\hat{\mu}_{t,T,h}(\tau_1, \tau_2) = \frac{1}{\tilde{n}_{t,h}} \sum_{j=\underline{n}_t}^{\tilde{n}_{t,h}} X_{t+j,T}(\tau_1)X_{t+j+h,T}(\tau_2).$$

Finally, for fixed $k \in \mathbb{N}$, collect the bootstrap approximations in the vector

$$\hat{\mathbb{B}}_T^{(k)} = (\hat{B}_T^{(k)}, \hat{B}_{T,0}^{(k)}, \dots, \hat{B}_{T,H}^{(k)}).$$

The following theorem shows that the bootstrap replicates can be regarded as asymptotically independent copies of the original process \mathbb{B}_T from Theorem 1.

Theorem 2 *Suppose that Assumptions (A1)–(A3) and (B1)–(B3) are met. Then, for any fixed $K \in \mathbb{N}$ and as $T \rightarrow \infty$,*

$$(\mathbb{B}_T, \hat{\mathbb{B}}_T^{(1)}, \dots, \hat{\mathbb{B}}_T^{(K)}) \rightsquigarrow (\mathbb{B}, \mathbb{B}^{(1)}, \dots, \mathbb{B}^{(K)})$$

in $\{L^2([0, 1]^2) \times (L^2([0, 1]^3))^{H+1}\}^{K+1}$, where $\mathbb{B}^{(k)}$ ($k = 1, \dots, K$) are independent copies of the centered Gaussian variable \mathbb{B} from Theorem 1. Equivalently (Bücher and Kojadinovic 2017, Lemma 2.2),

$$d_{BL}(\mathbb{P}^{\hat{\mathbb{B}}_T^{(1)} | X_{1,T}, \dots, X_{T,T}}, \mathbb{P}^{\mathbb{B}_T}) = o_{\mathbb{P}}(1), \quad T \rightarrow \infty,$$

where d_{BL} denotes the bounded Lipschitz metric between probability distributions on $L^2([0, 1]^2) \times (L^2([0, 1]^3))^{H+1}$.

The proof is given in Sect. 6.2. The preceding theorem, together with Corollary 1, suggests to define the following bootstrap approximation for the CUSUM processes defined in (11) and (12):

$$\begin{aligned} \hat{G}_T^{(k)}(u, \tau) &= \hat{B}_T^{(k)}(u, \tau) - u \hat{B}_T^{(k)}(1, \tau), \\ \hat{G}_{T,h}^{(k)}(u, \tau_1, \tau_2) &= \hat{B}_{T,h}^{(k)}(u, \tau_1, \tau_2) - u \hat{B}_{T,h}^{(k)}(1, \tau_1, \tau_2), \\ \hat{\mathbb{G}}_T^{(k)} &= (\hat{G}_T^{(k)}, \hat{G}_{T,0}^{(k)}, \dots, \hat{G}_{T,H}^{(k)}). \end{aligned}$$

Theorem 2, Corollary 1 and the continuous mapping theorem then imply that, under the hypothesis $H_0^{(H)}$ in (4),

$$\begin{aligned} (\mathcal{S}_T, \mathcal{S}_T^{(1)}, \dots, \mathcal{S}_T^{(K)}) &\equiv (\Phi(\mathbb{U}_T), \Phi(\hat{\mathbb{G}}_T^{(1)}), \dots, \Phi(\hat{\mathbb{G}}_T^{(K)})) \\ &= (\Phi(\mathbb{G}_T), \Phi(\hat{\mathbb{G}}_T^{(1)}), \dots, \Phi(\hat{\mathbb{G}}_T^{(K)})) + o_{\mathbb{P}}(1) \\ &\rightsquigarrow (\Phi(\mathbb{G}), \Phi(\mathbb{G}^{(1)}), \dots, \Phi(\mathbb{G}^{(K)})) \equiv (\mathcal{S}, \mathcal{S}^{(1)}, \dots, \mathcal{S}^{(K)}), \end{aligned}$$

where $\Phi(G_{-1}, G_0, \dots, G_H) = (\|G_{-1}\|_{2,2}, \|G_0\|_{2,3}, \dots, \|G_H\|_{2,3})$ and where $\mathbb{G}^{(1)}, \dots, \mathbb{G}^{(K)}$ are independent copies of \mathbb{G} . Individual bootstrap-based tests for, e.g., $H_0^{(c,h)}$ are then naturally defined by the p value

$$p_{T,K}(S_{T,h}) = \frac{1}{K} \sum_{j=1}^K \mathbf{1}(S_{T,h}^{(j)} \geq S_{T,h}),$$

where $S_{T,h}^{(k)}$ and $S_{T,h}$ denote the $(h + 2)$ nd coordinate of $\mathcal{S}_T^{(k)}$ and \mathcal{S}_T , respectively; in particular, $S_{T,-1} = S_T^{(m)}$ and $S_{T,h} = S_T^{(c,h)}$ as defined in (13). Indeed, we can show the following result for each individual test.

Proposition 1 *Suppose that Assumptions (A1)–(A3) and (B1)–(B3) are met. Then, for all $h \in \mathbb{Z}_{\geq -1}$, provided $K = K_T \rightarrow \infty$, and with $H_0^{(c,-1)} = H_0^{(m)}$, we have*

$$p_{T,K_T}(S_{T,h}) \rightsquigarrow \begin{cases} \text{Uniform}(0, 1) & \text{if } H_0^{(c,h)} \text{ is met} \\ 0 & \text{else.} \end{cases}$$

Moreover, we can rely on an extension of Fisher's p value combination method (Fisher 1932) as described in Sect. 2 in Bücher et al. (2018) to obtain a combined test for the joint hypothesis $H_0^{(H)}$ in (4). More precisely, let $\psi : (0, 1)^{H+2} \rightarrow \mathbb{R}$ be a continuous function that is decreasing in each argument (throughout the simulations, we employ $\psi(p_{-1}, \dots, p_H) = \sum_{i=-1}^H w_i \Phi^{-1}(1 - p_i)$ with weights $w_{-1} = w_0 = 1/3$ and $w_1 = \dots = w_H = (3H)^{-1}$.) The combined test is defined by its p value calculated based on the following algorithm.

Algorithm 1 (Combined Bootstrap test for $H_0^{(H)}$)

- (1) Let $\mathbf{S}_T^{(0)} = \mathbf{S}_T$.
- (2) Given a large integer K , compute the sample of K bootstrap replicates $\mathbf{S}_T^{(1)}, \dots, \mathbf{S}_T^{(K)}$ of the vector $\mathbf{S}_T^{(0)}$.
- (3) Then, for all $i \in \{0, 1, \dots, K\}$ and $h \in \{-1, \dots, H\}$, compute

$$p_{T,K}(S_{T,h}^{(i)}) = \frac{1}{K+1} \left\{ \frac{1}{2} + \sum_{k=1}^K \mathbf{1}(S_{T,h}^{(k)} \geq S_{T,h}^{(i)}) \right\}.$$

- (4) Next, for all $i \in \{0, 1, \dots, K\}$, compute

$$W_{T,K}^{(i)} = \psi\{p_{T,K}(S_{T,0}^{(i)}), \dots, p_{T,K}(S_{T,H}^{(i)})\}.$$

- (5) The global statistic is $W_{T,K}^{(0)}$, and the corresponding p value is given by

$$p_{T,K}(W_{T,K}^{(0)}) = \frac{1}{K} \sum_{k=1}^K \mathbf{1}(W_{T,K}^{(k)} \geq W_{T,K}^{(0)}).$$

Consistency of this procedure is a mere consequence of Proposition 2.1 in Bücher et al. (2018); details are omitted for the sake of brevity.

3.5 Consistency against AMOC-piecewise locally stationary alternatives

In the previous section, the proposed tests were shown to be consistent against locally stationary alternatives. In classical change point settings, the underlying CUSUM principle is also known to be consistent against piecewise (locally) stationary alternatives, notably against those that involve a single change in the signal of interest (AMOC = at most one change). We are going to derive such results within the present setting.

For the sake of brevity, we only consider AMOC alternatives in the mean. More precisely, we assume that $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ follows the data-generating process

$$X_{t,T} = \begin{cases} \mu_1 + Y_{t,T} & \text{for } t \leq \lambda T \\ \mu_2 + Y_{t,T} & \text{for } t \geq \lceil \lambda T \rceil + 1. \end{cases} \tag{17}$$

for some $\lambda \in (0, 1)$, $\mu_1, \mu_2 \in \mathcal{L}^2([0, 1])$ and $\{(Y_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ a locally stationary time series satisfying Condition 1. In the literature on classic change point detection, one would be interested in testing for the null hypothesis that $\|\mu_1 - \mu_2\|_2 = 0$, against the alternative that this L^2 -norm is positive.

Now, if $\|\mu_1 - \mu_2\|_2 = 0$, we are back in the situation of the preceding sections. However, one can show (by contradiction) that if $\|\mu_1 - \mu_2\|_2 > 0$, $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ is not locally stationary, whence additional theory must be developed to show consistency of the test statistic $S_T^{(H)}$. Note that even the formulation of $H_0^{(H)}$ relying on (2) and (3) is not possible anymore, so that we need to rely on their equivalent sub-asymptotic counterparts (9) and (10) in Lemma 1.

Proposition 2 *Let $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ be a sequence of functional time series as defined in (17), with $\mu_1 \neq \mu_2$ in $L^2([0, 1])$ and with $\{(Y_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ satisfying Conditions (A1)–(A3). Then, the test statistic $S_T^{(m)} = S_{T,-1}$ based on observations $X_{1,T}, \dots, X_{T,T}$ diverges to infinity, in probability. If, additionally, (B1)–(B3) are met, then the bootstrap variables $\hat{S}_{T,-1}^{(k)}$ are stochastically bounded. As a consequence, the proposed test is consistent.*

Remark 2 A careful inspection of the proof of Proposition 2 shows that the testing procedure is also consistent against local alternatives of the form $\mu_2 = \mu_1 + d_T$, for any sequence d_T with $d_T \sqrt{T} \rightarrow \infty$.

3.6 Data-driven choice of the block length parameter m

The bootstrap procedure depends on the choice of the width of the local mean estimator, n , and the length of the bootstrap blocks, m . Preliminary simulation studies suggested that the performance of the procedure crucially depends on the choice of m , while it is less sensitive to the choice of n (which may also be chosen by other standard criteria in specific applications, like adaptations of Silverman’s rule of thumb, cross-validation or visual investigation of respective plots). In this section, we propose a data-driven procedure for choosing the block length m based on a certain optimality criterion.

Recall that the limiting null distributions of the proposed test statistics depend in a complicated way on the covariances $\text{Cov}\{\tilde{B}(u, \tau), \tilde{B}(v, \varphi)\}$, $\text{Cov}\{\tilde{B}_h(u, \tau_1, \tau_2), \tilde{B}_h(v, \varphi_1, \varphi_2)\}$ and $\text{Cov}\{\tilde{B}(u, \tau), \tilde{B}_h(v, \varphi_1, \varphi_2)\}$. Following Sect. 5 in Bücher and Kojadinovic (2016), the procedure we propose essentially chooses m in such a way that the bootstrap approximation for $\sigma_c(\tau, \varphi) = \text{Cov}\{\tilde{B}(1, \tau), \tilde{B}(1, \varphi)\}$ is optimal, with respect to m , in a certain asymptotic sense. More precisely, we propose to first minimize the integrated mean squared of the “bootstrap estimator”

$$\tilde{\sigma}_T(\tau, \varphi) = \text{Cov}(\tilde{B}_T^{(1)}(1, \tau), \tilde{B}_T^{(1)}(1, \varphi) | X_{1,T}, \dots, X_{T,T})$$

considered as an estimator for $\sigma_c(\tau, \varphi)$, with respect to m theoretically (see Lemma 3), and then use a simple plug-in approach to obtain a formula that solely depends on observable quantities. Observe that $\tilde{\sigma}_T(\tau, \varphi)$ can be rewritten as

$$\begin{aligned} \tilde{\sigma}_T(\tau, \varphi) &= \mathbb{E}[\tilde{B}_T^{(1)}(1, \tau), \tilde{B}_T^{(1)}(1, \varphi) | X_{1,T}, \dots, X_{T,T}] \\ &= \frac{1}{T} \sum_{i=1}^T \frac{1}{m} \left(\sum_{t=i}^{(i+m-1) \wedge T} X_{t,T}(\tau) - \mu_{i,T}(\tau) \right) \left(\sum_{t=i}^{(i+m-1) \wedge T} X_{t,T}(\varphi) - \mu_{i,T}(\varphi) \right) \end{aligned}$$

whence $\tilde{\sigma}_T(\tau, \varphi)$ is not a proper estimator as it depends on the unknown expectation $\mu_{i,T}$. The asymptotic integrated bias and integrated variance satisfy the following expansions. For simplicity, we replace Condition (A3) by a strong mixing condition as in Sect. 3.3.

Lemma 3 *Let $m = m(T)$ be an integer-valued sequence, such that m tends to infinity and m^2/T vanishes, as T tends to infinity. If conditions (A1) and (A2) are met and $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ is exponentially strongly mixing, then, as $T \rightarrow \infty$,*

$$\begin{aligned} \int_{[0,1]^2} (\mathbb{E}[\tilde{\sigma}_T(\tau, \varphi)] - \sigma_c(\tau, \varphi))^2 d(\tau, \varphi) &= \frac{1}{m^2} \Delta + o(m^{-2}), \\ \int_{[0,1]^2} \text{Var}(\tilde{\sigma}_T(\tau, \varphi)) d(\tau, \varphi) &= \frac{m}{T} \Gamma + o(m/T), \end{aligned}$$

where

$$\Delta = \left\| \sum_{k=-\infty}^{\infty} |k| \int_0^1 \text{Cov}(X_0^{(w)}, X_k^{(w)}) dw \right\|_{2,2}^2$$

and

$$\begin{aligned} \Gamma &= \frac{2}{3} \int_0^1 \left(\sum_{k=-\infty}^{\infty} \int_0^1 \text{Cov}(X_0^{(w)}(\tau), X_k^{(w)}(\tau)) d\tau \right)^2 \\ &\quad + \left\| \sum_{k=-\infty}^{\infty} \text{Cov}(X_0^{(w)}, X_k^{(w)}) \right\|_{2,2}^2 dw. \end{aligned}$$

As a consequence of this lemma, we obtain the expansion

$$\begin{aligned} \text{IMSE}_T(m) &= \int_{[0,1]^2} \text{MSE}(\tilde{\sigma}_T(\tau, \varphi)) d(\tau, \varphi) \\ &= \int_{[0,1]^2} \text{Var}(\tilde{\sigma}_T(\tau, \varphi)) + (\mathbb{E}[\tilde{\sigma}_T(\tau, \varphi)] - \sigma_c(\tau, \varphi))^2 d(\tau, \varphi) \\ &= \frac{m}{T} \Gamma + \frac{1}{m^2} \Delta + o(m^{-2}) + o(m/T), \end{aligned}$$

which can next be minimized with respect to m to get a natural choice for the block length. More precisely, the dominating function $\Lambda(m) = \frac{m}{T}\Gamma + \frac{1}{m^2}\Delta$ is differentiable in m with $\Lambda'(m) = \frac{\Gamma}{T} - \frac{2\Delta}{m^3}$ and $\Lambda''(m) = \frac{6\Delta}{m^4}$, whence $m = \left(\frac{2\Delta T}{\Gamma}\right)^{1/3}$ is the unique minimizer of Λ . In practice, both Γ and Δ are unknown and must be estimated in terms of the observed data. This leads us to define

$$\hat{m} = (2\hat{\Delta}_T T / \hat{\Gamma}_T)^{1/3}$$

where, for some constant $L \in \mathbb{N}$ specified below,

$$\hat{\Delta}_T = \int_{[0,1]^2} \left(\frac{1}{T-2L} \sum_{i=L+1}^{T-L} \sum_{k=-L}^L |k| \hat{\gamma}_{i,k,T}(\tau, \varphi) \right)^2 d(\tau, \varphi)$$

and

$$\begin{aligned} \hat{\Gamma}_T &= \frac{2}{3} \frac{1}{T-2L} \sum_{i=L+1}^{T-L} \left(\sum_{k=-L}^L \int_0^1 \hat{\gamma}_{i,k,T}(\tau, \tau) d\tau \right)^2 \\ &+ \int_{[0,1]^2} \left(\sum_{k=-L}^L \hat{\gamma}_{i,k,T}(\tau, \varphi) \right)^2 d(\tau, \varphi). \end{aligned}$$

Here $\hat{\gamma}_{i,k,T}$ is defined by

$$\begin{aligned} \hat{\gamma}_{i,k,T}(\tau, \varphi) &= \frac{1}{\bar{n}_{i+k,0} - \underline{n}_i + 1} \sum_{j=\underline{n}_i}^{\bar{n}_{i+k,0}} \left(X_{i+j,T}(\tau) - \frac{1}{\bar{n}_{i+j,0}} \sum_{t=\underline{n}_{i+j}}^{\bar{n}_{i+j,0}} X_{i+j+t,T}(\tau) \right) \\ &\times \left(X_{i+j+k,T}(\varphi) - \frac{1}{\bar{n}_{i+j+k,0}} \sum_{t=\underline{n}_{i+j+k}}^{\bar{n}_{i+j+k,0}} X_{i+j+k+t,T}(\varphi) \right) \end{aligned}$$

and $\bar{n}_{t,h}$, \underline{n}_t and $\tilde{n}_{t,h}$ are given in (16). Note that the above estimators depend on the choice of the integer L . Following [Bücher and Kojadinovic \(2016\)](#) and [Politis and White \(2004\)](#), we select L to be the smallest integer, such that

$$\hat{\rho}_{k,T} = \frac{\left\| \frac{1}{T-k} \sum_{i=1}^{T-k} \hat{\gamma}_{i,k,T} \right\|_{2,2}}{\left\| \frac{1}{T} \sum_{i=1}^T \hat{\gamma}_{i,0,T} \right\|_{2,2}}$$

is negligible for any $k > L$; more precisely, L is chosen as the smallest integer such that $\hat{\rho}_{L+k,T} \leq 2\sqrt{\log(T)/T}$, for any $k = 1, \dots, K_T$, with $K_T = \max\{5, \sqrt{\log T}\}$.

4 Time-varying random operator functional AR processes

We consider an exemplary class of functional locally stationary processes and specify the approximating family of stationary processes. The results in this section are similar to Theorem 3.1 of [Bosq \(2000\)](#).

Let $\mathcal{L} = \mathcal{L}(L^2([0, 1]), L^2([0, 1]))$ be the space of bounded linear operators on $L^2([0, 1])$. Further, denote by $\|\cdot\|_{\mathcal{L}}$ and $\|\cdot\|_{\mathcal{S}}$ the standard operator norm and the Hilbert–Schmidt norm, respectively, i. e.,

$$\|\ell\|_{\mathcal{L}} = \sup_{\|x\|_2 \leq 1} \|\ell(x)\|_2, \quad \|\ell\|_{\mathcal{S}} = \left(\sum_{j=1}^{\infty} \lambda_j^2 \right)^{1/2}$$

for $\ell \in \mathcal{L}$ with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots$. By Eq. (1.55) in [Bosq \(2000\)](#), we have $\|\cdot\|_{\mathcal{L}} \leq \|\cdot\|_{\mathcal{S}}$. For any $T \in \mathbb{N}$, consider the recursive functional equation

$$X_{t,T} = Y_{t,T} + \mu(t/T), \quad Y_{t,T} = A_{t/T}(Y_{t-1,T}) + \varepsilon_{t,T}, \quad t \in \mathbb{Z}, \quad (18)$$

where $(\varepsilon_{t,T})_{t \in \mathbb{Z}}$ is a sequence of independent zero mean innovations in $L^2([0, 1])$ and where $A_{t,T} : L^2([0, 1]) \rightarrow L^2([0, 1])$ denotes a possibly random and time-varying bounded linear operator. The equation defines what might be called a (time-varying) random operator functional autoregressive process of order one, denoted by tvrFAR(1) (see also [van Delft et al. \(2017\)](#), Sect. 4.1) for the non-random case with $\varepsilon_{t,T}$ not depending on T .

In the following, we will only consider the case where μ is the null function. In the more general case of μ being Lipschitz, if there exists a locally stationary solution $Y_{t,T}$ of the equation on the right-hand side of (18) with approximating family $\{Y_t^{(u)} | t \in \mathbb{Z}\}_{u \in [0,1]}$, then $X_{t,T} = Y_{t,T} + \mu(t/T)$ is obviously locally stationary with approximating family $X_t^{(u)} = Y_t + \mu(u)$.

To be precise, we restrict ourselves to the following specific parameterization

$$\mu \equiv 0, \quad A_{t/T} = a(t/T)\tilde{A}, \quad \varepsilon_{t,T} = \sigma(t/T)\tilde{\varepsilon}_t,$$

where a and $\sigma > 0$ are measurable functions on $[0, 1]$. The following lemma provides sufficient conditions for ensuring local stationarity of the model and provides an explicit expression for the approximating family of stationary processes. For a related result in the case where $A_{t/T}$ is non-random and $\varepsilon_{t,T}$ does not depend on T , see Theorem 3.1 in [van Delft and Eichler \(2018\)](#).

For a sequence of operators $(B_i)_i$ in \mathcal{L} , we will write $\prod_{i=0}^n B_i = B_0 \circ \dots \circ B_n$ for $n \in \mathbb{N}$. The empty product will be identified with the identity on $L^2([0, 1])$, that is, $\prod_{i=0}^{-1} B_i = \text{id}_{L^2([0,1])}$.

Lemma 4 *Let $(\tilde{\varepsilon}_t)_{t \in \mathbb{Z}}$ be strong white noise in $L^2([0, 1])$. Further, let a and σ be measurable functions on $(-\infty, 1]$ such that $\sigma > 0$, $a(u) = a(0)$ and $\sigma(u) = \sigma(0)$ for all $u \leq 0$. Finally, let $\varepsilon_{t,T} = \sigma(t/T)\tilde{\varepsilon}_t$, $\varepsilon_t^{(u)} = \sigma(u)\tilde{\varepsilon}_t$ and $A_u = a(u)\tilde{A}$, where*

\tilde{A} denotes a random operator in \mathcal{L} that is independent from $(\tilde{\varepsilon}_t)_{t \in \mathbb{Z}}$ and satisfies $\sup_{u \in [0, 1]} \|A_u\|_{\mathcal{S}} \leq q < 1$ with probability one. Then:

- (i) For any $u \in [0, 1]$, there exists a unique stationary solution $(Y_t^{(u)})_{t \in \mathbb{Z}}$ of the recursive equation

$$Y_t^{(u)} = A_u(Y_{t-1}^{(u)}) + \varepsilon_t^{(u)}, \quad t \in \mathbb{Z},$$

namely

$$Y_t^{(u)} = \sum_{j=0}^{\infty} A_u^j(\varepsilon_{u, t-j}),$$

where the latter series converges in $L^2(\Omega \times [0, 1], \mathbb{P} \otimes \lambda)$ and almost surely in $L^2([0, 1])$.

- (ii) If σ and a are Lipschitz continuous, then there exists a unique locally stationary solution $(Y_{t,T})$ of order $\rho = 2$ satisfying $\sup_{t \in \mathbb{Z}, T \in \mathbb{N}} \mathbb{E}[\|Y_{t,T}\|_2^2] < \infty$ of the recursive equation

$$Y_{t,T} = A_{t/T}(Y_{t-1,T}) + \varepsilon_{t,T}, \quad t \in \mathbb{Z}, T \in \mathbb{N},$$

namely

$$Y_{t,T} = \sum_{j=0}^{\infty} \left(\prod_{i=0}^{j-1} A_{\frac{t-i}{T}} \right) (\varepsilon_{t-j,T}),$$

the series again being convergent in $L^2(\Omega \times [0, 1], \mathbb{P} \otimes \lambda)$ and almost surely in $L^2([0, 1])$. The locally stationary process has approximating family $\{(Y_t^{(u)})_{t \in \mathbb{Z}} : u \in [0, 1]\}$.

Remark 3 The previous lemma provides sufficient conditions for local stationarity of the random operator FAR model. In the simulation study of the following section, we use specific models for the parameter curves and the functional white noise and show that those models satisfy the cumulant condition in Assumption (A3) as well.

5 Finite-sample results

5.1 Monte Carlo simulations

A large-scale Monte Carlo simulation study was performed to analyze the finite-sample behavior of the proposed tests. The major goals of the study were to analyze the level approximation and the power of the various tests, with a particular view on investigating various different forms of alternatives, notably models from $H_1^{(m)}$, $H_1^{(c,0)}$ and $H_1^{(c,1)}$. All stated results related to testing the joint hypothesis $H_0^{(H)}$ are for the combined test described in Algorithm 1, with $\psi(p_{-1}, \dots, p_H) = \sum_{i=-1}^H w_i \Phi^{-1}(1 -$

p_i) with weights $w_1 = w_0 = 1/2$ for $H = 0$ and $w_{-1} = w_0 = 1/3$ and $w_1 = \dots = w_H = (3H)^{-1}$ for $H \geq 1$.

For the data-generating processes, we employed 10 different choices for the parameters in (18), which will be described next. Let $(\psi_i)_{i \in \mathbb{N}_0}$ denote the Fourier basis of $L^2([0, 1])$, that is, for $n \in \mathbb{N}$,

$$\psi_0 \equiv 1, \quad \psi_{2n-1}(\tau) = \sqrt{2} \sin(2\pi n\tau), \quad \psi_{2n}(\tau) = \sqrt{2} \cos(2\pi n\tau).$$

Let $(\tilde{\varepsilon}_t)_{t \in \mathbb{Z}}$ denote an i.i.d. sequence of mean zero random variables in $L^2([0, 1])$, defined by $\tilde{\varepsilon}_t = \sum_{i=0}^{16} u_{t,i} \psi_i$, where $u_{t,i}$ are independent and normally distributed with mean zero and variance $\text{Var}(u_{i,t}) = \exp(-i/10)$. Independent of $(\tilde{\varepsilon}_t)_{t \in \mathbb{Z}}$, let $\mathbf{G} = (G_{i,j})_{i,j=0,\dots,16}$ denote a matrix with independent normally distributed entries with $\text{Var}(G_{i,j}) = \exp(-i - j)$. Let $\tilde{A} : L^2([0, 1]) \rightarrow L^2([0, 1])$ denote the (random) integral operator defined by

$$\begin{aligned} \tilde{A}(f)(t) &= \frac{1}{3\|\mathbf{G}\|_F} \sum_{i,j=0}^{16} G_{i,j} \langle f, \psi_i \rangle \psi_j(t) \\ &= \int_0^1 \left(\frac{1}{3\|\mathbf{G}\|_F} \sum_{i,j=0}^{16} G_{i,j} \psi_i(s) \psi_j(t) \right) f(s) \, ds, \end{aligned}$$

where $\|\mathbf{G}\|_F$ denotes the Frobenius norm (note that the Hilbert–Schmidt norm of \tilde{A} is equal to $1/3$, see Horváth and Kokoszka 2012, Sect. 2.2). Finally, let

$$\begin{aligned} a_0(u) &= 1, & a_1(u) &= \frac{1}{2} + u, \\ a_2(u) &= 1 - \frac{1}{2} \cos(2\pi u), & a_3(u) &= \frac{1}{2} + \mathbb{1}(u \geq 1/2), \end{aligned}$$

for $u \in [0, 1]$ and let $a_j(u) = a_j(0)$ for $u \leq 0$ and $a_j(u) = a_j(1)$ for $u \geq 1$. The following ten data-generating processes are considered:

- *Stationary case.* Let

$$\mu \equiv 0, \quad A_{t/T} = \tilde{A}, \quad \varepsilon_{t,T} = \tilde{\varepsilon}_t. \tag{19}$$

- *Models deviating from $H_0^{(m)}$.* For $j = 1, \dots, 3$, consider the choices

$$\mu(\tau) = a_j(\tau), \quad A_{t/T} = \tilde{A}, \quad \varepsilon_{t,T} = \tilde{\varepsilon}_t. \tag{20}$$

- *Models deviating from $H_0^{(c,0)}$.* For $j = 1, \dots, 3$, consider the choices

$$\mu \equiv 0, \quad A_{t/T} = \tilde{A}, \quad \varepsilon_{t,T} = a_j(t/T) \tilde{\varepsilon}_t. \tag{21}$$

- *Models deviating from $H_0^{(c,1)}$.* For $j = 1, \dots, 3$, consider the choices

$$\mu \equiv 0, \quad A_{t/T} = a_j(t/T) \tilde{A}, \quad \varepsilon_{t,T} = \tilde{\varepsilon}_t. \tag{22}$$

Remark 4 The models defined in (19)–(22) satisfy condition (A3) if the functions a and σ are Lipschitz continuous. To see this, we exemplarily derive (14), as the other parts follow by similar arguments. Observe that in all models, the mean function μ is in L^2 ; thus, (14) is trivial for $j = 1$. For $j \geq 2$, consider first the case that \tilde{A} is non-random first. Then, by Lemma 4 (ii), the definition of $A_{t/T}$ and $\varepsilon_{t,T}$, and linearity of both (powers of) the operator \tilde{A} and the cumulants,

$$\begin{aligned} & \| \text{cum}(X_{t_1,T}, \dots, X_{t_j,T}) \|_{2,j} \\ &= \| \text{cum}(Y_{t_1,T}, \dots, Y_{t_j,T}) \|_{2,j} \\ &= \left\| \text{cum} \left(\sum_{k=0}^{\infty} \left(\prod_{i=0}^{k-1} A_{\frac{t_1-i}{T}} \right) (\varepsilon_{t_1-k,T}), \dots, \sum_{k=0}^{\infty} \left(\prod_{i=0}^{k-1} A_{\frac{t_j-i}{T}} \right) (\varepsilon_{t_j-k,T}) \right) \right\|_{2,j} \\ &= \left\| \text{cum} \left(\sum_{k=0}^{\infty} \left(\prod_{i=0}^{k-1} a \left(\frac{t_1-i}{T} \right) \right) \tilde{A}^k \left(\sigma \left(\frac{t_1-k}{T} \right) \sum_{\ell=0}^{16} u_{t_1-k,\ell} \psi_{\ell} \right), \dots \right. \right. \\ &\quad \left. \left. \dots, \sum_{k=0}^{\infty} \left(\prod_{i=0}^{k-1} a \left(\frac{t_j-i}{T} \right) \right) \tilde{A}^k \left(\sigma \left(\frac{t_j-k}{T} \right) \sum_{\ell=0}^{16} u_{t_j-k,\ell} \psi_{\ell} \right) \right) \right\|_{2,j} \\ &= \left\| \sum_{k_1, \dots, k_j=0}^{\infty} \sum_{\ell_1, \dots, \ell_j=0}^{16} \left(\prod_{v=1}^j \left\{ \left(\prod_{i=0}^{k_v-1} a \left(\frac{t_v-i}{T} \right) \right) \sigma \left(\frac{t_v-k_v}{T} \right) \tilde{A}^{k_v} (\psi_{\ell_v}) \right\} \right) \right. \\ &\quad \left. \times \text{cum}(u_{t_1-k_1,\ell_1}, \dots, u_{t_j-k_j,\ell_j}) \right\|_{2,j}. \end{aligned}$$

By independence of the random variables $u_{t,\ell}$, the cumulants on the right-hand side of the previous display are zero if there is an index $v \in \{2, \dots, j\}$ such that $\ell_1 \neq \ell_v$ or $t_1 - k_1 \neq t_v - k_v$. Further, the random variables $u_{t,\ell}$ are normally distributed and the higher cumulants, for $j > 2$, are zero. Thus, (14) is trivial for $j > 2$. Now, let $j = 2$ and consider $t_2 \geq t_1$ without loss of generality. Then, by the previous considerations,

$$\begin{aligned} & \| \text{Cov}(X_{t_1,T}, X_{t_2,T}) \|_{2,2} \\ &= \left\| \sum_{k=0}^{\infty} \sum_{\ell=0}^{16} \left(\prod_{i=0}^{k-1} a \left(\frac{t_1-i}{T} \right) \right) \sigma \left(\frac{t_1-k}{T} \right) \tilde{A}^k (\psi_{\ell}) \left(\prod_{i=0}^{t_2-t_1+k-1} a \left(\frac{t_2-i}{T} \right) \right) \right. \\ &\quad \left. \times \sigma \left(\frac{t_2-(t_2-t_1+k)}{T} \right) \tilde{A}^{t_2-t_1+k} (\psi_{\ell}) \text{Var}(u_{t_1-k,\ell}) \right\|_{2,2}, \end{aligned}$$

which can be bounded by

$$\frac{9}{4} \sum_{k=0}^{\infty} \left(\frac{1}{2} \right)^{t_2-t_1+2k} \sum_{\ell=0}^{16} \exp \left(-\frac{\ell}{10} \right) \leq C \left(\frac{1}{2} \right)^{t_2-t_1}$$

since $\|\tilde{A}\|_{\mathcal{S}} = \frac{1}{3}$, $\text{Var}(u_{t,\ell}) = \exp(-\ell/10)$, and both a and σ can be bounded from above by $3/2$. Thus,

$$\sum_{t_1=-\infty}^{\infty} \|\text{Cov}(X_{t_1,T}, X_{t_2,T})\|_{2,2} \leq C \left(\sum_{t_1=-\infty}^{t_2} \left(\frac{1}{2}\right)^{t_2-t_1} + \sum_{t_1=t_2+1}^{\infty} \left(\frac{1}{2}\right)^{t_1-t_2} \right) \leq C,$$

which proves (14) in case \tilde{A} is non-random. The random case follows since the upper bound does not depend on \tilde{A} and since \tilde{A} is assumed to be independent of $\tilde{\varepsilon}_t$.

Subsequently, the respective models will be denoted by (\mathcal{M}_0) and $(\mathcal{M}_{m,j})$, $(\mathcal{M}_{v,j})$ and $(\mathcal{M}_{a,j})$ for $j = 1, \dots, 3$. Note that the model descriptions are non-exclusive: For instance, the models in (20) exhibiting deviations from $H_0^{(m)}$ also deviate from $H_0^{(c,0)}$.

Preliminary simulation studies showed that the data-driven choice of m , as introduced in Sect. 3.6, yields similar results as a manual choice of m and should be favored. Further parameters of the simulation design are as follows: The number of bootstrap replicates is set to $K = 200$. Two sample sizes were considered, namely $T = 256$ and $T = 512$. Observe though that, unlike many frequency domain-based methods for functional time series, the proposed testing procedure does not require the sample sizes to be a power of two to work effectively. The hyperparameter n for estimating local means is set to $n = 45, 60, 75, 90, T$. Finally, the maximum number of lags considered was set to $H = 4$. Empirical rejection rates are based on $N = 500$ simulation runs each and are summarized in Tables 1 and 2.

From the previous results, it can be seen that different choices of n do not lead to crucially different results. For $T = 256$, the tests for the hypotheses $H_0^{(m)}$ and $H_0^{(c,0)}$ already have good power against the alternatives $(\mathcal{M}_{m,1})$, $(\mathcal{M}_{m,3})$ and $(\mathcal{M}_{v,1})$, $(\mathcal{M}_{v,2})$, $(\mathcal{M}_{v,3})$, respectively. When combining $H_0^{(m)}$ and $H_0^{(c,0)}$ and taking even more autocovariances into account, the test does not lose significant power. For $T = 512$, the power further increases such that all tests have good power against the alternatives $(\mathcal{M}_{m,i})$ and $(\mathcal{M}_{v,i})$, $i = 1, 2, 3$. Detecting non-stationarities in models $(\mathcal{M}_{a,i})$, $i = 1, 2, 3$ turns out to be more difficult. Even though the power increases with T , for small values of T , the results are not too convincing. These findings can be explained by the fact that the *measures of non-stationarity* $\|M\|_{2,2}$ and $\|M_h\|_{2,3}$, as introduced in (5) and (6), are comparably small for models $(\mathcal{M}_{a,i})$, $i = 1, 2, 3$. This can be deduced from Table 3, where these measures of non-stationarity are approximated by their natural estimators $\|M_T\|_{2,2} = \|U_T\|_{2,2}/\sqrt{T}$ and $\|M_{T,h}\|_{2,3} = \|U_{T,h}\|_{2,3}/\sqrt{T}$, based on 2000 Monte Carlo repetitions and for various choices of T . It is noticeable that the values for models $(\mathcal{M}_{a,1})$ and $(\mathcal{M}_{a,2})$ are close to those for (\mathcal{M}_0) , which perfectly explains the results of the simulation study.

As pointed out by a reviewer, it is of interest to investigate the sensitivity of the test with respect to the number of components included in the procedure. For this purpose, we highlight in Table 4 the rejection probabilities of the tests for the different hypotheses in the models $(\mathcal{M}_{m,1})$ and $(\mathcal{M}_{v,1})$, which correspond to a change in the

Table 1 Empirical rejection rates for various combined tests, based on a sample size of $T = 256$ and a block length parameter m calculated as proposed in Sect. 3.6

Model	n	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$	\bar{m}	$sd(m)$
(\mathcal{M}_0)	45	7.2	0.4	3.0	1.0	0.8	0.8	1.0	5.77	0.41
	60	7.0	0.2	2.6	0.4	0.4	0.4	0.4	5.78	0.41
	75	5.4	0.8	2.2	0.2	0.2	0.4	0.4	5.80	0.40
	90	5.2	0.2	1.8	0.2	0.2	0.2	0.2	5.80	0.41
	256	4.0	0.2	1.8	0.0	0.0	0.0	0.0	5.81	0.38
$(\mathcal{M}_{m,1})$	45	92.8	64.4	89.6	89.8	90.8	90.6	91.2	5.82	0.41
	60	91.8	61.2	88.8	87.8	88.6	89.4	89.0	5.81	0.41
	75	90.2	59.2	87.8	86.4	87.2	87.6	87.6	5.82	0.41
	90	89.8	58.2	86.4	85.2	86.4	86.2	86.8	5.83	0.40
	256	88.2	50.4	81.8	79.4	81.8	82.8	83.2	5.82	0.42
$(\mathcal{M}_{m,2})$	45	57.6	31.2	55.8	53.8	55.0	55.0	56.0	5.79	0.41
	60	53.0	26.2	49.2	48.2	48.6	48.4	49.4	5.78	0.42
	75	49.0	20.8	45.0	42.8	43.4	42.8	43.0	5.78	0.42
	90	41.8	16.0	38.8	33.8	35.8	35.8	36.4	5.78	0.42
	256	31.0	9.0	24.8	21.4	23.4	22.4	22.0	5.89	0.60
$(\mathcal{M}_{m,3})$	45	99.8	97.0	99.8	99.8	99.8	99.8	100.0	5.78	0.43
	60	99.6	96.6	99.8	99.8	99.8	99.8	99.8	5.78	0.44
	75	99.6	96.2	99.8	99.8	99.8	99.8	99.8	5.78	0.44
	90	99.8	95.8	99.8	99.8	99.8	99.8	99.8	5.80	0.44
	256	99.6	94.8	99.4	99.0	99.2	99.2	99.2	6.15	1.42
$(\mathcal{M}_{v,1})$	45	8.0	100.0	99.6	84.8	81.2	79.4	79.0	5.50	0.59
	60	8.0	100.0	99.8	82.4	77.6	76.8	75.4	5.53	0.59
	75	7.4	100.0	99.8	78.8	74.4	72.8	74.2	5.60	0.58
	90	7.2	100.0	99.6	75.4	73.2	72.4	73.0	5.69	0.53
	256	3.6	100.0	96.6	63.0	56.6	56.0	53.4	6.00	0.45
$(\mathcal{M}_{v,2})$	45	6.2	100.0	99.4	76.2	71.4	70.4	70.4	6.34	2.84
	60	5.0	100.0	98.2	64.8	60.2	59.0	60.0	6.81	3.70
	75	4.6	99.0	90.8	51.8	48.6	47.2	45.6	7.86	5.38
	90	3.4	87.4	69.8	31.4	27.2	26.2	26.8	9.75	7.40
	256	3.6	96.6	65.6	22.4	20.2	21.0	20.6	6.60	2.28
$(\mathcal{M}_{v,3})$	45	20.2	100.0	100.0	95.8	93.6	92.4	91.8	14.15	10.14
	60	13.6	100.0	100.0	93.8	89.2	88.4	87.8	13.85	9.65
	75	10.2	100.0	100.0	90.6	88.2	85.0	84.6	14.05	9.47
	90	8.6	100.0	100.0	87.6	86.2	82.0	80.8	13.63	8.96
	256	3.4	100.0	96.4	78.0	71.2	67.6	67.0	11.43	7.25
$(\mathcal{M}_{a,1})$	45	6.6	4.8	5.6	5.0	4.6	4.2	3.6	5.82	0.41
	60	5.8	3.8	5.2	4.0	3.2	3.4	2.6	5.83	0.40
	75	5.2	2.6	4.2	3.2	3.0	2.8	2.0	5.85	0.39

Table 1 continued

Model	n	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$	\bar{m}	$sd(m)$
$(\mathcal{M}_{a,2})$	90	4.8	2.2	3.8	2.8	2.4	2.0	1.6	5.85	0.39
	256	3.0	1.0	1.2	0.6	0.6	0.4	0.6	5.83	0.43
	45	7.8	2.4	4.8	3.4	2.8	2.4	2.8	5.82	0.42
	60	7.0	1.6	3.6	2.0	1.2	1.2	1.4	5.84	0.42
	75	6.4	1.0	3.2	1.2	1.0	0.8	0.8	5.87	0.44
$(\mathcal{M}_{a,3})$	90	6.4	0.8	2.6	0.6	0.6	0.6	0.6	5.90	0.44
	256	5.4	0.2	1.2	0.2	0.2	0.2	0.2	5.84	0.52
	45	10.2	16.4	13.2	20.6	16.8	13.0	11.6	5.93	0.57
	60	8.6	14.4	11.2	17.0	13.6	10.2	9.2	5.94	0.55
	75	7.6	12.6	9.2	15.0	11.6	9.0	8.2	5.96	0.59
	90	7.2	11.8	9.2	14.6	10.8	8.2	7.6	5.96	0.61
	256	5.4	5.6	5.6	6.6	5.4	4.2	4.0	5.93	0.56

The last two columns provide the mean and standard deviation of the selected value of m

mean (and as a consequence in all second-order moments) and changes in all second-order moments (with no change in the mean), respectively.

We observe that the impact of the number of components on the power of the tests shows no clear pattern. For example, the power is almost constant under model $(\mathcal{M}_{m,1})$ if additional components are included in the test statistic. Conversely, under model $(\mathcal{M}_{v,1})$, the power decreases with the number of components.

We also list results for the functional moving average model

$$X_{t,T} = \tilde{\varepsilon}_t + \sigma(t/T)\tilde{\varepsilon}_{t-2}, \tag{\mathcal{M}_{MA}}$$

with $\sigma(u) = \mathbb{1}(u \leq 0.5) - \mathbb{1}(u > 0.5)$ and $\tilde{\varepsilon}_t$ as for the autoregressive model. This model has constant mean and covariances, except of the covariance with lag 2, and consequently only the tests for hypotheses $H_0^{(H)}$ ($H \geq 2$) should reject the null since all moments besides $\mathbb{E}[X_{t,T}X_{t+2,T}]$ are constant. This effect is clearly visible in the last rows of Table 4, where we observe the largest rejection probabilities by the test constructed for the hypotheses $H_0^{(2)}$. The power decreases for the hypotheses $H_0^{(3)}$ and $H_0^{(4)}$, which can be explained by the fact that in this case the statistics combines stationary and non-stationary parts of the serial dependence. Compared to the other cases, the test has less power in model \mathcal{M}_{MA} . An intuitive explanation for this observation consists of the fact that the weights of the p values, as introduced in Sect. 3.4, favor changes in the mean and the second moment (the corresponding weights in the test statistic are larger), which makes it more difficult to detect changes in models that are only non-stationary in the second-order moments at higher lags. On the other hand, it is worthwhile to mention that the null hypothesis $H_0^{(c,2)}$ is rejected in nearly 100% of the cases in the \mathcal{M}_{MA} model (these results are not displayed).

Table 2 Empirical rejection rates for various combined tests, based on a sample size of $T = 512$ and a block length parameter m calculated as proposed in Sect. 3.6

Model	n	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$	\bar{m}	$sd(m)$
(\mathcal{M}_0)	45	10.6	5.2	7.8	4.4	4.0	3.8	3.4	7.24	0.43
	60	10.0	3.8	6.6	3.0	2.2	2.4	2.6	7.25	0.44
	75	8.4	2.0	3.8	1.4	1.0	1.6	1.4	7.28	0.46
	90	7.6	1.8	3.4	1.2	1.2	1.0	1.0	7.28	0.46
	512	4.6	1.0	2.6	0.8	0.8	0.8	0.8	7.30	0.46
$(\mathcal{M}_{m,1})$	45	100.0	97.4	100.0	99.8	99.8	100.0	100.0	7.23	0.43
	60	99.8	97.0	100.0	99.6	99.6	99.8	100.0	7.23	0.42
	75	100.0	96.8	99.8	99.6	99.8	99.8	99.8	7.25	0.44
	90	100.0	96.6	99.8	99.4	99.6	99.8	99.8	7.27	0.45
	512	99.0	91.2	98.8	98.4	98.4	98.6	98.6	7.35	0.54
$(\mathcal{M}_{m,2})$	45	95.6	82.4	95.6	95.2	95.6	96.4	96.0	7.23	0.42
	60	94.2	78.8	94.6	94.2	94.6	94.6	95.0	7.25	0.43
	75	94.0	75.2	94.0	93.2	93.4	94.2	94.0	7.26	0.45
	90	93.2	73.4	93.4	93.2	93.0	93.6	93.6	7.27	0.44
	512	83.4	51.8	79.2	79.6	81.6	82.0	83.4	7.40	0.55
$(\mathcal{M}_{m,3})$	45	100.0	100.0	100.0	100.0	100.0	100.0	100.0	7.24	0.43
	60	100.0	100.0	100.0	100.0	100.0	100.0	100.0	7.27	0.45
	75	100.0	100.0	100.0	100.0	100.0	100.0	100.0	7.28	0.46
	90	100.0	100.0	100.0	100.0	100.0	100.0	100.0	7.29	0.46
	512	100.0	99.6	100.0	100.0	100.0	100.0	100.0	8.47	3.45
$(\mathcal{M}_{v,1})$	45	9.2	100.0	100.0	97.8	95.6	95.4	95.0	6.85	0.44
	60	8.4	100.0	100.0	95.8	93.2	91.0	89.2	6.88	0.46
	75	7.2	100.0	100.0	94.8	92.6	89.6	87.6	6.89	0.44
	90	6.0	100.0	100.0	94.4	90.6	89.6	85.8	6.91	0.43
	512	4.0	100.0	100.0	90.4	84.0	80.8	79.2	7.43	0.54
$(\mathcal{M}_{v,2})$	45	7.2	100.0	100.0	96.6	94.0	93.6	92.4	6.79	0.60
	60	6.6	100.0	100.0	94.2	90.4	88.2	87.6	6.86	0.57
	75	6.2	100.0	100.0	93.2	88.6	85.4	82.4	6.96	0.63
	90	5.6	100.0	100.0	90.8	86.8	84.2	81.0	7.05	0.66
	512	3.8	100.0	99.8	87.2	82.4	80.0	78.2	7.50	0.75
$(\mathcal{M}_{v,3})$	45	8.2	100.0	100.0	99.4	97.8	96.0	96.4	7.89	3.41
	60	7.2	100.0	100.0	98.6	96.4	95.0	93.4	7.79	3.08
	75	6.4	100.0	99.8	98.6	94.8	92.0	91.2	7.87	3.12
	90	6.0	100.0	100.0	98.4	95.4	92.0	90.4	7.98	3.14
	512	4.6	100.0	100.0	98.0	95.8	93.4	90.2	8.58	2.58

Table 2 continued

Model	n	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$	\bar{m}	$sd(m)$
$(\mathcal{M}_{a,1})$	45	11.0	23.8	19.8	31.0	26.2	23.0	21.4	7.26	0.46
	60	10.2	20.2	16.6	24.4	20.4	18.0	15.4	7.30	0.49
	75	9.8	18.0	13.4	22.0	18.8	16.4	13.2	7.30	0.47
	90	9.6	17.2	13.2	20.0	16.8	15.2	13.0	7.31	0.48
	512	5.6	7.4	5.2	7.0	4.8	4.2	4.0	7.35	0.50
$(\mathcal{M}_{a,2})$	45	8.8	9.2	9.2	10.8	8.8	8.4	7.4	7.29	0.50
	60	7.4	6.0	5.8	6.4	5.2	4.2	3.4	7.31	0.51
	75	6.8	4.2	5.6	5.0	4.0	3.0	2.4	7.33	0.53
	90	6.4	3.4	5.2	4.4	3.0	1.6	2.2	7.35	0.55
	512	4.8	1.6	1.6	1.0	0.6	0.6	0.4	7.39	0.59
$(\mathcal{M}_{a,3})$	45	10.4	42.4	32.0	59.2	47.6	40.6	37.6	7.38	0.59
	60	10.3	37.8	27.8	53.2	42.6	35.6	30.8	7.40	0.60
	75	8.2	35.2	25.4	51.2	40.2	33.2	28.2	7.43	0.64
	90	7.8	33.4	24.2	47.4	37.0	31.4	25.0	7.43	0.65
	512	7.0	26.8	16.2	35.4	25.4	20.6	16.4	7.52	0.68

The last two columns provide the mean and standard deviation of the selected value of m

5.2 Case study

Functional time series naturally arise in the field of meteorology. For instance, the daily minimal temperature at one place over time can be naturally divided into yearly functional data.

To illustrate the proposed methodology, we consider the daily minimum temperature recorded at eight different locations across Australia. Exemplary, the temperature curves of Melbourne and Sydney are displayed in Fig. 1. The results of our testing procedure can be found in Table 5, where we employed $K = 1000$ bootstrap replicates, considered up to $H = 4$ lags and chose $n = 25$, based on visual exploration of the respective plots. The null hypotheses of stationarity can be rejected, at level $\alpha = 0.05$, for all measuring stations except of Gunnedah Pool, for which the p values exceed α by a small amount.

6 Proofs

Throughout the proofs, C denotes a generic constant whose value may change from line to line. If not specified otherwise, all convergences are for $T \rightarrow \infty$.

Table 3 $\|M_T\|_{2,2}$ and $\|M_{T,h}\|_{2,3}$, $h = 0, \dots, 4$, calculated by 2000 Monte Carlo repetitions

T	Model	$\ M_T\ _{2,2}$	$\ M_{T,0}\ _{2,3}$	$\ M_{T,1}\ _{2,3}$	$\ M_{T,2}\ _{2,3}$	$\ M_{T,3}\ _{2,3}$	$\ M_{T,4}\ _{2,3}$
256	(\mathcal{M}_0)	0.0759	0.2344	0.2273	0.2272	0.2274	0.2277
	$(\mathcal{M}_{m,1})$	0.1180	0.3208	0.3152	0.3145	0.3146	0.3138
	$(\mathcal{M}_{m,2})$	0.0952	0.2812	0.2754	0.2754	0.2758	0.2760
	$(\mathcal{M}_{v,1})$	0.0803	0.5394	0.3124	0.2989	0.2979	0.2977
	$(\mathcal{M}_{v,2})$	0.0698	0.3681	0.2471	0.2412	0.2412	0.2422
	$(\mathcal{M}_{a,1})$	0.0764	0.2426	0.2389	0.2328	0.2306	0.2301
	$(\mathcal{M}_{a,2})$	0.0758	0.2358	0.2304	0.2277	0.2273	0.2273
	512	(\mathcal{M}_0)	0.0540	0.1659	0.1607	0.1603	0.1604
$(\mathcal{M}_{m,1})$		0.1049	0.2590	0.2562	0.2556	0.2547	0.2549
$(\mathcal{M}_{m,2})$		0.0781	0.2144	0.2097	0.2095	0.2097	0.2100
$(\mathcal{M}_{v,1})$		0.0572	0.4939	0.2313	0.2121	0.2105	0.2099
$(\mathcal{M}_{v,2})$		0.0494	0.3241	0.1791	0.1702	0.1695	0.1697
$(\mathcal{M}_{a,1})$		0.0542	0.1743	0.1743	0.1665	0.1632	0.1624
$(\mathcal{M}_{a,2})$		0.0537	0.1683	0.1652	0.1620	0.1606	0.1602
1024		(\mathcal{M}_0)	0.0383	0.1172	0.1134	0.1132	0.1131
	$(\mathcal{M}_{m,1})$	0.0987	0.2250	0.2233	0.2229	0.2230	0.2223
	$(\mathcal{M}_{m,2})$	0.0681	0.1710	0.1685	0.1685	0.1684	0.1684
	$(\mathcal{M}_{v,1})$	0.0403	0.4696	0.1776	0.1518	0.1489	0.1486
	$(\mathcal{M}_{v,2})$	0.0349	0.3000	0.1336	0.1210	0.1196	0.1198
	$(\mathcal{M}_{a,1})$	0.0386	0.1289	0.1321	0.1215	0.1169	0.1152
	$(\mathcal{M}_{a,2})$	0.0381	0.1212	0.1206	0.1162	0.1141	0.1134
	2048	(\mathcal{M}_0)	0.0270	0.0831	0.0802	0.0801	0.0800
$(\mathcal{M}_{m,1})$		0.0949	0.2047	0.2041	0.2038	0.2035	0.2036
$(\mathcal{M}_{m,2})$		0.0624	0.1449	0.1431	0.1431	0.1431	0.1431
$(\mathcal{M}_{v,1})$		0.0283	0.4568	0.1430	0.1097	0.1055	0.1049
$(\mathcal{M}_{v,2})$		0.0245	0.2869	0.1035	0.0866	0.0846	0.0844
$(\mathcal{M}_{a,1})$		0.0272	0.0973	0.1040	0.0906	0.0845	0.0821
$(\mathcal{M}_{a,2})$		0.0269	0.0884	0.0898	0.0839	0.0814	0.0804

Table 4 Empirical rejection rates for various combined tests, based on sample sizes of $T = 256, 512$, the bandwidth of the local mean estimator $n = 60$ (for $T = 256$), $n = 90$ (for $T = 512$), and a block length parameter selected automatically

T	Model	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$
256	$(\mathcal{M}_{m,1})$	91.8	61.2	88.8	87.8	88.6	89.4	89.0
	$(\mathcal{M}_{v,1})$	8.0	100.0	99.8	82.4	77.6	76.8	75.4
	(\mathcal{M}_{MA})	4.2	8.0	5.0	3.2	23.8	12.6	8.4
512	$(\mathcal{M}_{m,1})$	100.0	96.6	99.8	99.4	99.6	99.8	99.8
	$(\mathcal{M}_{v,1})$	6.0	100.0	100.0	94.4	90.6	89.6	85.8
	(\mathcal{M}_{MA})	6.2	7.4	5.2	3.4	28.2	15.2	11.8

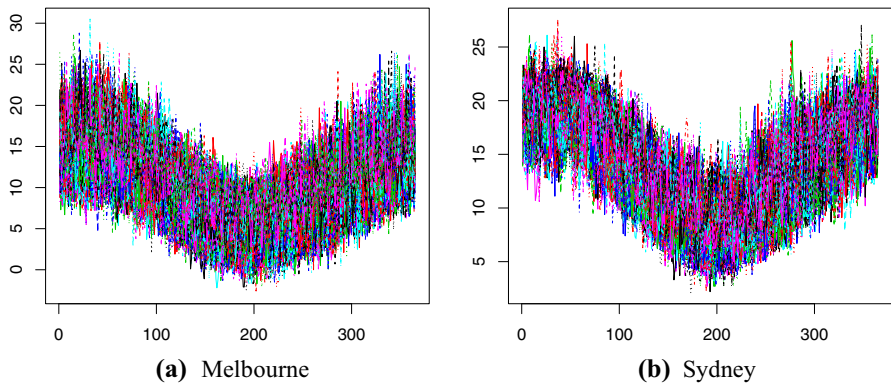


Fig. 1 Temperature curves of Melbourne ($T = 161$ years) and Sydney ($T = 160$ years), where the x axis corresponds to a year in rescaled time and the y axis denotes temperature in degree Celsius

Table 5 p values of the (combined) tests for the respective null hypotheses in percent, and selected value of m

Location	T	$H_0^{(m)}$	$H_0^{(c,0)}$	$H_0^{(0)}$	$H_0^{(1)}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$	m
Boulia Airport	131	0.3	0.0	0.0	0.0	0.0	0.0	0.0	5
Gayndah Post Office	117	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4
Gunnedah Pool	136	6.4	5.8	5.8	5.9	5.8	5.3	5.1	6
Hobart	137	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8
Melbourne	161	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5
Cape Otway Lighthouse	155	0.5	0.4	0.4	0.1	0.0	0.0	0.0	7
Robe	135	4.5	1.0	1.9	1.0	1.0	0.6	0.5	6
Sydney	160	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5

6.1 A fundamental approximation lemma in Hilbert spaces

Lemma 5 Fix $p \in \mathbb{N}$. For $i = 1, \dots, p$ and $T \in \mathbb{N}$, let $X_{i,T}$ and X_i denote random variables in a separable Hilbert space $(H_i, \langle \cdot, \cdot \rangle_i)$. Further, let $(\psi_k^{(i)})_{k \in \mathbb{N}}$ be an orthonormal basis of H_i and for brevity write $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_i$. Suppose that

- (1) $Y_T^n := ((\langle X_{1,T}, \psi_k^{(1)} \rangle)_{k=1}^n, \dots, (\langle X_{p,T}, \psi_k^{(p)} \rangle)_{k=1}^n) \rightsquigarrow ((\langle X_1, \psi_k^{(1)} \rangle)_{k=1}^n, \dots, (\langle X_p, \psi_k^{(p)} \rangle)_{k=1}^n) =: Y^n$ as $T \rightarrow \infty$, for any $n \in \mathbb{N}$,
- (2) $\lim_{n \rightarrow \infty} \limsup_{T \rightarrow \infty} \mathbb{P} \left(\sum_{k=n+1}^{\infty} \sum_{i=1}^p \langle X_{i,T}, \psi_k^{(i)} \rangle^2 > \varepsilon \right) = 0$ for all $\varepsilon > 0$.

Then, using the notation $\|(x_k)_{k \in \mathbb{N}}\|_2 = \sum_{k=1}^{\infty} x_k^2$,

$$Y_T^\infty := ((\langle X_{1,T}, \psi_k^{(1)} \rangle)_{k=1}^\infty, \dots, (\langle X_{p,T}, \psi_k^{(p)} \rangle)_{k=1}^\infty)$$

$$\rightsquigarrow ((\langle X_1, \psi_k^{(1)} \rangle)_{k=1}^\infty, \dots, (\langle X_p, \psi_k^{(p)} \rangle)_{k=1}^\infty) =: Y^\infty \quad \text{in } (\ell^2(\mathbb{N}), \|\cdot\|_2)^p$$

and, as a consequence,

$$(X_{1,T}, \dots, X_{p,T}) \rightsquigarrow (X_1, \dots, X_p) \quad \text{in } H_1 \times \dots \times H_p.$$

Proof of Lemma 1 To prove the first part, we employ Theorem 2 of [Dehling et al. \(2009\)](#). Expand the random variables Y_T^n and Y_n in \mathbb{R}^{pn} to

$$\tilde{Y}_{T,n}^\infty = ((a_{T,k}^{(1)})_{k \in \mathbb{N}}, \dots, (a_{T,k}^{(p)})_{k \in \mathbb{N}}) \quad \text{and} \quad \tilde{Y}_n^\infty = ((a_k^{(1)})_{k \in \mathbb{N}}, \dots, (a_k^{(p)})_{k \in \mathbb{N}})$$

in $(\ell^2(\mathbb{N}), \|\cdot\|_2)^p$, where $a_{T,k}^{(i)} = \langle X_{i,T}, \psi_k^{(i)} \rangle$ and $a_k^{(i)} = \langle X_i, \psi_k^{(i)} \rangle$, for any $1 \leq k \leq n$, and $a_{T,k}^{(i)} = a_k^{(i)} = 0$, for any $k > n, i = 1, \dots, p$. By the continuous mapping theorem, $\tilde{Y}_{T,n}^\infty$ converges weakly to \tilde{Y}_n^∞ in $(\ell^2(\mathbb{N}), \|\cdot\|_2)^p$, for any $n \in \mathbb{N}$ and as T tends to infinity.

By assumption (2) and since the space $(\ell^2(\mathbb{N}), \|\cdot\|_2)^p$ is separable and complete, there is a random variable $\tilde{Y}^\infty \in (\ell^2(\mathbb{N}), \|\cdot\|_2)^p$ such that $Y_T^\infty \rightsquigarrow \tilde{Y}^\infty$, as T tends to infinity, and $\tilde{Y}_n^\infty \rightsquigarrow \tilde{Y}^\infty$, as n tends to infinity, by Theorem 2 of [Dehling et al. \(2009\)](#). Due to the latter convergence, the finite-dimensional distributions of \tilde{Y}^∞ and Y^∞ are the same. Thus, by Theorem 1.3 of [Billingsley \(1999\)](#) and Lemma 1.5.3 of [van der Vaart and Wellner \(1996\)](#), \tilde{Y}^∞ and Y^∞ have the same distribution in $(\ell^2(\mathbb{N}), \|\cdot\|_2)^p$.

Next, observe that, for an arbitrary Hilbert space H , the function

$$\Phi := \begin{cases} \ell^2(\mathbb{N}) & \rightarrow H \\ (y_k)_{k \in \mathbb{N}} & \mapsto \sum_{k=1}^\infty y_k \psi_k \end{cases}$$

is continuous, provided $(\psi_k)_{k \in \mathbb{N}}$ is an orthonormal basis of H . Indeed,

$$\|\Phi((y_k)_k) - \Phi((z_k)_k)\|^2 = \sum_{k=1}^\infty (y_k - z_k)^2 = \|y - z\|_2^2.$$

Thus, the mapping

$$\Phi' := \begin{cases} (\ell^2(\mathbb{N}), \|\cdot\|_2)^p & \rightarrow H_1 \times \dots \times H_p \\ ((y_{k,1})_{k \in \mathbb{N}}, \dots, (y_{k,p})_{k \in \mathbb{N}}) & \mapsto (\sum_{k=1}^\infty y_{k,1} \psi_k^{(1)}, \dots, \sum_{k=1}^\infty y_{k,p} \psi_k^{(p)}) \end{cases}$$

is continuous too, and the continuous mapping theorem implies that

$$\begin{aligned} (X_{1,T}, \dots, X_{p,T}) &= \left(\sum_{k=1}^\infty \langle X_{1,T}, \psi_k^{(1)} \rangle \psi_k^{(1)}, \dots, \sum_{k=1}^\infty \langle X_{p,T}, \psi_k^{(p)} \rangle \psi_k^{(p)} \right) \\ &\rightsquigarrow \left(\sum_{k=1}^\infty \langle X_1, \psi_k^{(1)} \rangle \psi_k^{(1)}, \dots, \sum_{k=1}^\infty \langle X_p, \psi_k^{(p)} \rangle \psi_k^{(p)} \right) \\ &= (X_1, \dots, X_p), \end{aligned}$$

as T tends to infinity. □

6.2 Proofs for Sects. 3.1, 3.2, 3.3 and 3.4

Proof of Lemma 2 We only prove the equivalence concerning $H_0^{(h)}$; the equivalences regarding $H_0^{(m)}$ follow along similar lines.

Step 1: Equivalence between (3) and (8). Suppose that (8) is met. To prove (3), it is sufficient to show that

$$\left\| \mathbb{E}[X_0^{(u)} \otimes X_h^{(u)}] - \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2} = 0 \quad (23)$$

for any $u \in [0, 1]$.

Fix $u \in [0, 1)$ and let $\delta > 0$ be sufficiently small such that $u + \delta < 1$. By the reverse triangle inequality, we obtain that

$$\begin{aligned} 0 &\leq \left\| \frac{1}{\delta} \int_u^{u+\delta} \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}] dw \right\|_{2,2} \\ &\quad - \left\| \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}] dw \right\|_{2,2} \\ &\leq \frac{1}{\delta} \left\| \int_u^{u+\delta} \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - \delta \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2} \\ &= \frac{1}{\delta} \left\| \int_0^{u+\delta} \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - (u + \delta) \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right. \\ &\quad \left. - \int_0^u \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw + u \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2} \\ &\leq \frac{1}{\delta} \left\| \int_0^{u+\delta} \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - (u + \delta) \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2} \\ &\quad + \frac{1}{\delta} \left\| \int_0^u \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - u \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2}. \end{aligned}$$

By continuity of integrals in the upper integration limit, it follows from (8) that both summands on the right-hand side of this display are equal to zero. As a consequence,

$$\begin{aligned} &\left\| \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - \mathbb{E}[X_0^{(u)} \otimes X_h^{(u)}] \right\|_{2,2} \\ &= \left\| \frac{1}{\delta} \int_u^{u+\delta} \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}] dw \right\|_{2,2}. \end{aligned} \quad (24)$$

By Jensen's inequality, we can bound the right-hand side of this display from above by

$$\left(\int_{[0,1]^2} \frac{1}{\delta^2} \int_u^{u+\delta} (\mathbb{E}[X_0^{(w)}(\tau_1) X_h^{(w)}(\tau_2) - X_0^{(u)}(\tau_1) X_h^{(u)}(\tau_2)])^2 dw d(\tau_1, \tau_2) \right)^{1/2}. \quad (25)$$

By employing Jensen’s inequality again, we can bound the integrand by

$$\mathbb{E}[(X_0^{(w)}(\tau_1)X_h^{(w)}(\tau_2) - X_0^{(u)}(\tau_1)X_h^{(u)}(\tau_2))^2].$$

Thus, by Fubini’s theorem, (25) is less than or equal to

$$\left(\frac{1}{\delta^2} \int_u^{u+\delta} \mathbb{E}[\|X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}\|_{2,2}^2] dw\right)^{1/2}.$$

This term is of the order $O(\delta^{1/2})$ due to the inequality

$$\begin{aligned} &\mathbb{E}\|X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}\|_{2,2}^2 \\ &= \mathbb{E}\|X_0^{(w)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(w)} + X_0^{(u)} \otimes X_h^{(w)} - X_0^{(u)} \otimes X_h^{(u)}\|_{2,2}^2 \\ &\leq 2\{\mathbb{E}\|(X_0^{(w)} - X_0^{(u)}) \otimes X_h^{(w)}\|_{2,2}^2 + \mathbb{E}\|X_0^{(u)} \otimes (X_h^{(w)} - X_h^{(u)})\|_{2,2}^2\} \\ &\leq 2\{\mathbb{E}[\|X_0^{(w)} - X_0^{(u)}\|_2^4]^{1/2} \mathbb{E}[\|X_h^{(w)}\|_2^4]^{1/2} + \mathbb{E}[\|X_0^{(u)}\|_2^4]^{1/2} \mathbb{E}[\|X_h^{(w)} - X_h^{(u)}\|_2^4]^{1/2}\} \\ &\leq C|u - w|^2, \end{aligned} \tag{26}$$

where the final bound follows from Lemma C.2 in the supplementary material. Since δ was chosen arbitrarily, we obtain that the right-hand side of (24) is equal to zero. This proves (23) for $u \in [0, 1)$, and the case $u = 1$ follows from (26), which is also valid for $u = 1$.

Conversely, if (3) holds true, we have by a change of variables, linearity of the integral, Jensen’s inequality and Fubini’s theorem,

$$\begin{aligned} &\left\| \int_0^u \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - u \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,3}^2 \\ &= \left\| u \int_0^1 \mathbb{E}[X_0^{(uw)} \otimes X_h^{(uw)}] - \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,3}^2 \\ &\leq \int_0^1 \int_0^1 u^2 \|\mathbb{E}[X_0^{(uw)} \otimes X_h^{(uw)}] - \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}]\|_{2,2}^2 dw du = 0. \end{aligned}$$

Step 2: Equivalence between (3) and (10). Note that, irrespective of whether (3) or (10) is met, local stationarity of $X_{t,T}$ of order $\rho \geq 4$ and stationarity of $(X_t^{(u)})_{t \in \mathbb{Z}}$ with $\mathbb{E}\|X_t^{(u)}\|_2^4 < \infty$ implies that

$$\begin{aligned} &\|\mathbb{E}[X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor + h, T} - X_{\lfloor uT \rfloor}^{(u)} \otimes X_{\lfloor uT \rfloor + h}^{(u)}]\|_{2,2}^2 \\ &\leq \mathbb{E}[\|X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor + h, T} - X_{\lfloor uT \rfloor}^{(u)} \otimes X_{\lfloor uT \rfloor + h}^{(u)}\|_{2,2}^2] \\ &\leq 2\{\mathbb{E}[\|(X_{\lfloor uT \rfloor, T} - X_{\lfloor uT \rfloor}^{(u)}) \otimes X_{\lfloor uT \rfloor + h, T}\|_{2,2}^2] \\ &\quad + \mathbb{E}[\|X_{\lfloor uT \rfloor}^{(u)} \otimes (X_{\lfloor uT \rfloor + h, T} - X_{\lfloor uT \rfloor + h}^{(u)})\|_{2,2}^2]\} \\ &\leq 2\{\mathbb{E}[\|X_{\lfloor uT \rfloor, T} - X_{\lfloor uT \rfloor}^{(u)}\|_2^2] \|X_{\lfloor uT \rfloor + h, T}\|_2^2\} \end{aligned}$$

$$\begin{aligned}
 & + \mathbb{E}[\|X_{\lfloor uT \rfloor}^{(u)}\|_2^2 \|X_{\lfloor uT \rfloor+h, T} - X_{\lfloor uT \rfloor+h}^{(u)}\|_2^2] \\
 \leq & (C/T^2) \{ \mathbb{E}[(P_{\lfloor uT \rfloor, T}^{(u)})^4]^{1/2} \mathbb{E}\|X_{\lfloor uT \rfloor+h, T}\|_2^4\}^{1/2} \\
 & + \mathbb{E}[\|X_{\lfloor uT \rfloor}^{(u)}\|_2^4]^{1/2} \mathbb{E}[(P_{\lfloor uT \rfloor+h, T}^{(u)})^4]^{1/2} \\
 \leq & C/T^2,
 \end{aligned} \tag{27}$$

for any $u \in [0, 1]$ and $T \in \mathbb{N}$ and for some universal constant $C > 0$.
 Now, suppose that (3) is met. Then, the previous display implies that

$$\begin{aligned}
 & \|\mathbb{E}[X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor+h, T}] - \mathbb{E}[X_{0, T} \otimes X_{h, T}]\|_{2,2} \\
 \leq & \|\mathbb{E}[X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor+h, T}] - \mathbb{E}[X_{\lfloor uT \rfloor}^{(u)} \otimes X_{\lfloor uT \rfloor+h}^{(u)}]\|_{2,2} \\
 & + \|\mathbb{E}[X_0^{(u)} \otimes X_h^{(u)}] - \mathbb{E}[X_0^{(0)} \otimes X_h^{(0)}]\|_{2,2} + \|\mathbb{E}[X_0^{(0)} \otimes X_h^{(0)}] \\
 & - \mathbb{E}[X_{0, T} \otimes X_{h, T}]\|_{2,2} \\
 \leq & C/T + 0 + C/T = 2C/T,
 \end{aligned}$$

for any $u \in [0, 1]$ and $T \in \mathbb{N}$, that is, (10) is met.
 Conversely, if (10) is met, then, by (27) and (1), for any $u, v \in [0, 1]$ and $T \in \mathbb{N}$,

$$\begin{aligned}
 & \|\mathbb{E}[X_0^{(u)} \otimes X_h^{(u)} - X_0^{(v)} \otimes X_h^{(v)}]\|_2 \\
 = & \|\mathbb{E}[X_{\lfloor uT \rfloor}^{(u)} \otimes X_{\lfloor uT \rfloor+h}^{(u)} - X_{\lfloor vT \rfloor}^{(v)} \otimes X_{\lfloor vT \rfloor+h}^{(v)}]\|_2 \\
 \leq & \|\mathbb{E}[X_{\lfloor uT \rfloor}^{(u)} \otimes X_{\lfloor uT \rfloor+h}^{(u)} - X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor+h, T}]\|_2 \\
 & + \|\mathbb{E}[X_{\lfloor uT \rfloor, T} \otimes X_{\lfloor uT \rfloor+h, T} - X_{0, T} \otimes X_{h, T}]\|_2 \\
 & + \|\mathbb{E}[X_{0, T} \otimes X_{h, T} - X_{\lfloor vT \rfloor, T} \otimes X_{\lfloor vT \rfloor+h, T}]\|_2 \\
 & + \|\mathbb{E}[X_{\lfloor vT \rfloor, T} \otimes X_{\lfloor vT \rfloor+h, T} - X_{\lfloor vT \rfloor}^{(v)} \otimes X_{\lfloor vT \rfloor+h}^{(v)}]\|_2 \\
 \leq & 4C/T.
 \end{aligned}$$

Since T was arbitrary, the left-hand side of this display must be zero, whence (3). \square

Proof of Theorem 1 This theorem is an immediate consequence of Theorem C.3 of the supplementary material. \square

Proof of Corollary 1 Suppose that $H_0^{(c,h)}$ is met. Then, by the triangle inequality and a slight abuse of notation (note that u is a variable of integration in the norm $\|\cdot\|_{2,3}$), for $h \leq T$,

$$\begin{aligned}
 & \|U_{T,h} - G_{T,h}\|_{2, \Omega \times [0,1]^3} \\
 = & \left\| \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E}[X_{t, T} \otimes X_{t+h, T}] - u \sum_{t'=1}^{T-h} \mathbb{E}[X_{t', T} \otimes X_{t'+h, T}] \right) \right\|_{2,3} \\
 = & \frac{1}{\sqrt{T}} \left\| \frac{1}{T-h} \sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \sum_{t'=1}^{T-h} \mathbb{E}[X_{t, T} \otimes X_{t+h, T}] - \mathbb{E}[X_{t', T} \otimes X_{t'+h, T}] \right\|
 \end{aligned}$$

$$\begin{aligned}
 & + \left(\frac{1}{T-h} - \frac{u}{\lfloor uT \rfloor \wedge (T-h)} \right) \left\| \sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \sum_{t'=1}^{T-h} \mathbb{E}[X_{t',T} \otimes X_{t'+h,T}] \right\|_{2,3} \\
 & \leq \frac{C}{T^{3/2}} \sum_{t,t'=1}^{T-h} \|\mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_{t',T} \otimes X_{t'+h,T}]\|_{2,2} + \frac{C}{T^{3/2}} \sum_{t=1}^{T-h} \|\mathbb{E}[X_{t,T} \otimes X_{t+h,T}]\|_{2,2}.
 \end{aligned}$$

This expression is of the order $O(T^{-1/2})$ by (10) and Assumption (A2). Hence, $\|U_{T,h} - \tilde{G}_{T,h}\|_{2,3} = o_{\mathbb{P}}(1)$, and the assertion for U_T follows along similar lines.

Now, consider the assertion regarding the alternative $H_1^{(H)} = H_1^{(m)} \cup H_1^{(c,0)} \cup \dots \cup H_1^{(c,H)}$. We only treat the case where $H_1^{(c,h)}$ is met for some $h \in \{0, \dots, H\}$ and the case $H_1^{(m)}$ is similar. It is to be shown that $\|U_{T,h}\|_{2,3} \rightarrow \infty$ in probability.

By the reverse triangle inequality, we have

$$\|U_{T,h}\|_{2,3} = \|\tilde{G}_{T,h} + \mathbb{E}U_{T,h}\|_{2,3} \geq \|\tilde{G}_{T,h}\|_{2,3} - \|\mathbb{E}U_{T,h}\|_{2,3}.$$

The term $\|\tilde{G}_{T,h}\|_{2,3}$ converges weakly to $\|\tilde{G}_h\|_{2,3}$. Thus, it suffices to show that the second term $\|\mathbb{E}U_{T,h}\|_{2,3}$ diverges to infinity. For that purpose, note that another application of the reverse triangle inequality implies that

$$\begin{aligned}
 \|\mathbb{E}U_{T,h}\|_{2,3} & = \left\| \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - u \sum_{t=1}^{T-h} \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] \right) \right\|_{2,3} \\
 & \geq |S_{1,T} - S_{2,T}|,
 \end{aligned}$$

where

$$\begin{aligned}
 S_{1,T} & = \left\| \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \left\{ \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}] \right\} \right. \right. \\
 & \quad \left. \left. - u \sum_{t=1}^{T-h} \left\{ \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}] \right\} \right) \right\|_{2,3}
 \end{aligned}$$

and

$$S_{2,T} = \left\| \frac{1}{\sqrt{T}} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E} \left[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)} \right] - u \sum_{t=1}^{T-h} \mathbb{E} \left[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)} \right] \right) \right\|_{2,3}.$$

In the following, we will show that $S_{1,T}$ vanishes as T increases and that $S_{2,T}$ diverges to infinity. We have

$$S_{1,T} \leq \left\{ \int_0^1 \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \left\| \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}] \right\|_{2,2} + \frac{u}{\sqrt{T}} \sum_{t=1}^{T-h} \left\| \mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}] \right\|_{2,2} \right)^2 du \right\}^{1/2},$$

which is of order $O(T^{-1/2})$ since $\|\mathbb{E}[X_{t,T} \otimes X_{t+h,T}] - \mathbb{E}[X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}]\|_{2,2} \leq C/T$ by (1). For the second term $S_{2,T}$, we have, by stationarity

$$S_{2,T} = \sqrt{T} \left\| \frac{1}{T} \left(\sum_{t=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E}[X_0^{(t/T)} \otimes X_h^{(t/T)}] - u \sum_{t=1}^{T-h} \mathbb{E}[X_0^{(t/T)} \otimes X_h^{(t/T)}] \right) \right\|_{2,3},$$

where the norm converges to

$$\left\| \int_0^u \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - u \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,3},$$

by the dominated convergence theorem and the moment condition (A2). The expression in the latter display is strictly positive since (8) is not satisfied and by the continuity of

$$\left\| \int_0^u \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw - u \int_0^1 \mathbb{E}[X_0^{(w)} \otimes X_h^{(w)}] dw \right\|_{2,2}$$

in $u \in [0, 1]$. Thus, $S_{2,T} \rightarrow \infty$, which implies the assertion. □

Proof of Lemma 3 We will only give a proof of (14). Parts (i)–(iv) of the cumulant condition (A3) follow by similar arguments, which are omitted for the sake of brevity. According to Theorem 3 in Statulevicius and Jakimavicius (1988), we have

$$|\text{cum}(X_{t_1,T}(\tau_1), \dots, X_{t_k,T}(\tau_k))| \leq 3(k-1)! 2^{k-1} \alpha^{\delta/(1+\delta)} (t_{i+1} - t_i) \prod_{j=1}^k (\mathbb{E}[|X_{t_j,T}(\tau_j)|^{(1+\delta)k}]^{\frac{1}{(1+\delta)k}}),$$

for any increasing sequence $t_1 \leq t_2 \leq \dots \leq t_k$. Straightforward calculations combined with Hölder’s and Jensen’s inequality lead to

$$\begin{aligned} \left\| \prod_{j=1}^k \mathbb{E}[|X_{t_j,T}|^{(1+\delta)k}]^{\frac{1}{(1+\delta)k}} \right\|_{2,k} &= \prod_{j=1}^k \left\| \mathbb{E}[|X_{t_j,T}|^{(1+\delta)k}]^{\frac{1}{(1+\delta)k}} \right\|_2 \\ &= \prod_{j=1}^k \left(\int_{[0,1]} \mathbb{E}[|X_{t_j,T}(\tau)|^{(1+\delta)k}]^{\frac{2}{(1+\delta)k}} d\tau \right)^{1/2} \end{aligned}$$

$$\begin{aligned}
 &\leq \prod_{j=1}^k \left(\int_{[0,1]} \mathbb{E}[|X_{t_j,T}(\tau)|^{(1+\delta)k}] d\tau \right)^{\frac{1}{(1+\delta)k}} \\
 &= \prod_{j=1}^k \mathbb{E} \left[\|X_{t_j,T}\|_{(1+\delta)k}^{\frac{(1+\delta)k}{(1+\delta)k}} \right] \\
 &\leq \sup_{t,T} \mathbb{E} \left[\|X_{t,T}\|_{(1+\delta)k}^{1/(1+\delta)} \right] \\
 &\leq C_{k,1}.
 \end{aligned}$$

Thus, combining the previous results leads to

$$\begin{aligned}
 \|\text{cum}(X_{t_1,T}, \dots, X_{t_k,T})\|_{2,k} &\leq 3(k-1)! 2^{k-1} C_{k,1} \alpha^{\delta/(1+\delta)} (t_{i+1} - t_i) \\
 &\leq C_{k,4} \alpha^{\delta/(1+\delta)} (t_{i+1} - t_i),
 \end{aligned}$$

for any $i = 1, \dots, k-1$, where the constant $C_{k,4} > 0$ depends on k only. Hence,

$$\|\text{cum}(X_{t_1,T}, \dots, X_{t_k,T})\|_{2,k} \leq C_{k,4} \prod_{i=1}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (t_{i+1} - t_i).$$

Analogously, for arbitrary, not necessarily increasing t_1, \dots, t_k , we may obtain that

$$\|\text{cum}(X_{t_1,T}, \dots, X_{t_k,T})\|_{2,k} \leq C_{k,4} \prod_{i=1}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (t_{(i+1)} - t_{(i)}),$$

where $(t_{(1)}, \dots, t_{(k)})$ denotes the order statistic of (t_1, \dots, t_k) . The latter expression is symmetric in its arguments, and thus, we have, for any $t_k \in \mathbb{Z}$,

$$\begin{aligned}
 &\sum_{t_1, \dots, t_{k-1} = -\infty}^{\infty} \|\text{cum}(X_{t_1,T}, \dots, X_{t_k,T})\|_{2,k} \\
 &\leq C_{k,4} \sum_{t_1, \dots, t_{k-1} = -\infty}^{\infty} \prod_{i=1}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (t_{(i+1)} - t_{(i)}) \\
 &\leq C_{k,4} (k-1)! \sum_{-\infty < t_1 \leq \dots \leq t_{k-1} < \infty} \prod_{i=1}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (t_{i+1} - t_i) \\
 &\leq C_{k,4} (k-1)! \sum_{-\infty < t_2 \leq \dots \leq t_{k-1} < \infty} \sum_{s_1 = -\infty}^{\infty} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (s_1) \prod_{i=2}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}} (t_{i+1} - t_i).
 \end{aligned}$$

By assumption, $\{(X_{t,T})_{t \in \mathbb{Z}} : T \in \mathbb{N}\}$ is exponentially strong mixing, and the inner sum is finite and can be bounded by some constant $C_{k,2}$. Thus,

$$\begin{aligned} & \sum_{t_1, \dots, t_{k-1} = -\infty}^{\infty} \|\text{cum}(X_{t_1, T}, \dots, X_{t_k, T})\|_{2,k} \\ & \leq C_{k,4}(k-1)!C_{k,2} \sum_{-\infty < t_2 \leq \dots \leq t_{k-1} < \infty} \prod_{i=1}^{k-1} \alpha^{\frac{\delta}{(1+\delta)(k-1)}}(t_{i+1} - t_i). \end{aligned}$$

Repeating this argument successively, we obtain finally (14) as asserted. □

Proof of Theorem 2 By Slutsky’s lemma and Theorem C.3 of the supplementary material, it is sufficient to prove that

$$(\hat{\mathbb{B}}_T^{(1)} - \mathbb{B}_T^{(1)}, \dots, \hat{\mathbb{B}}_T^{(K)} - \mathbb{B}_T^{(K)}) = o_{\mathbb{P}}(1)$$

in $\{L^2([0, 1]^2) \times \{L^2([0, 1]^3)\}^{H+1}\}^K$, as T tends to infinity. This in turn is equivalent to

$$(\|\hat{B}_T^{(k)} - \tilde{B}_T^{(k)}\|_{2,3}, \|\hat{B}_{T,0}^{(k)} - \tilde{B}_{T,0}^{(k)}\|_{2,3}, \dots, \|\hat{B}_T^{(k)} - \tilde{B}_{T,h}^{(k)}\|_{2,3})_{k=1, \dots, K} = o_{\mathbb{P}}(1)$$

in $\mathbb{R}^{K(H+2)}$. The last convergence holds true if and only if the coordinates converge, i.e., if $\|\hat{B}_T^{(k)} - \tilde{B}_T^{(k)}\|_{2,3} = o_{\mathbb{P}}(1)$ and $\|\hat{B}_{T,h}^{(k)} - \tilde{B}_{T,h}^{(k)}\|_{2,3} = o_{\mathbb{P}}(1)$, for all $k = 1, \dots, K$ and $h = 0, \dots, H$. We only consider the latter assertion (the former can be treated similarly), and in fact, we will show convergence in $L^2(\Omega, \mathbb{P})$, which is even stronger. For this purpose, observe that by Fubini’s theorem and the independence of the family $(R_i^{(k)})_{i \in \mathbb{N}}$

$$\begin{aligned} & \mathbb{E}\|\hat{B}_{T,h}^{(k)} - \tilde{B}_{T,h}^{(k)}\|_{2,3}^2 \\ & = \mathbb{E}\left[\int_{[0,1]^3} \frac{1}{mT} \left\{ \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} R_i^{(k)} \sum_{t=i}^{(i+m-1) \wedge (T-h)} \mu_{t,T,h}(\tau_1, \tau_2) - \hat{\mu}_{t,T,h}(\tau_1, \tau_2) \right\}^2 d(u, \tau_1, \tau_2)\right] \\ & = \frac{1}{mT} \int_{[0,1]^3} \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E}\left[\left\{ \sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,1} + A_{t,2} \right\}^2\right] d(u, \tau_1, \tau_2), \end{aligned}$$

where

$$A_{t,1}(\tau_1, \tau_2) = \frac{1}{\bar{n}_{t,h}} \sum_{k=\underline{n}_t}^{\bar{n}_{t,h}} \mathbb{E}[X_{t,T}(\tau_1)X_{t+h,T}(\tau_2)] - \mathbb{E}[X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2)]$$

and

$$A_{t,2}(\tau_1, \tau_2) = \frac{1}{\bar{n}_{t,h}} \sum_{k=\underline{n}_t}^{\bar{n}_{t,h}} X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2) - \mathbb{E}[X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2)].$$

Since $A_{t,1}$ is deterministic and since $A_{t,2}$ is centered, we can rewrite the expectation in the previous integral as

$$\begin{aligned} & \mathbb{E} \left[\left\{ \sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,1}(\tau_1, \tau_2) + A_{t,2}(\tau_1, \tau_2) \right\}^2 \right] \\ &= \left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,1}(\tau_1, \tau_2) \right)^2 + \mathbb{E} \left[\left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,2}(\tau_1, \tau_2) \right)^2 \right]. \end{aligned}$$

In the following, we bound both parts separately. For the term $A_{t,1}$, first note that, by stationarity of $(X_t^{(w)})_{t \in \mathbb{Z}}$,

$$\begin{aligned} & \mathbb{E}[X_{t,T}(\tau_1)X_{t+h,T}(\tau_2)] - \mathbb{E}[X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2)] \\ &= \mathbb{E}[X_{t,T}(\tau_1)X_{t+h,T}(\tau_2) - X_t^{(t/T)}(\tau_1)X_{t+h}^{(t/T)}(\tau_2)] \\ & \quad - \mathbb{E}[X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2) - X_{t+k}^{(t/T)}(\tau_1)X_{t+k+h}^{(t/T)}(\tau_2)] \end{aligned}$$

in $L^2([0, 1]^2)$. Thus, by Jensen’s inequality and Fubini’s theorem, we have

$$\begin{aligned} & \frac{1}{mT} \int_{[0,1]^3} \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} \left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,1} \right)^2 d(u, \tau_1, \tau_2) \\ & \leq \frac{1}{mT} \int_{[0,1]^3} \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} \mathbb{E} \left[\left\{ \sum_{t=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\tilde{n}_{t,h}} \sum_{k=\underline{n}_t}^{\tilde{n}_{t,h}} X_{t,T}(\tau_1)X_{t+h,T}(\tau_2) - X_t^{(t/T)}(\tau_1)X_{t+h}^{(t/T)}(\tau_2) \right. \right. \\ & \quad \left. \left. - X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2) + X_{t+k}^{(t/T)}(\tau_1)X_{t+k+h}^{(t/T)}(\tau_2) \right\}^2 \right] d(u, \tau_1, \tau_2) \\ & \leq \frac{1}{mT} \sum_{i=1}^{T-h} \mathbb{E} \left\| \sum_{t=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\tilde{n}_{t,h}} \sum_{k=\underline{n}_t}^{\tilde{n}_{t,h}} X_{t,T} \otimes X_{t+h,T} - X_t^{(t/T)} \otimes X_{t+h}^{(t/T)} \right. \\ & \quad \left. - X_{t+k,T} \otimes X_{t+k+h,T} + X_{t+k}^{(t/T)} \otimes X_{t+k+h}^{(t/T)} \right\|_{2,2}^2. \end{aligned}$$

The norm on the right-hand side of the previous inequality be bounded by the triangle inequality by

$$\begin{aligned} & \sum_{t=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\tilde{n}_{t,h}} \sum_{k=\underline{n}_t}^{\tilde{n}_{t,h}} \|X_{t,T} \otimes X_{t+h,T} - X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}\|_{2,2} \\ & \quad + \|X_{t+k,T} \otimes X_{t+k+h,T} - X_{t+k}^{(t/T)} \otimes X_{t+k+h}^{(t/T)}\|_{2,2} \end{aligned}$$

and the inner summands can be bounded due to the local stationarity of $(X_{t,T})$: first,

$$\begin{aligned} & \|X_{t,T} \otimes X_{t+h,T} - X_t^{(t/T)} \otimes X_{t+h}^{(t/T)}\|_{2,2} \\ & \leq \|X_{t,T} \otimes (X_{t+h,T} - X_{t+h}^{(t/T)})\|_{2,2} + \|X_{t+h}^{(t/T)} \otimes (X_{t,T} - X_t^{(t/T)})\|_{2,2} \\ & = \|X_{t,T}\|_2 \|X_{t+h,T} - X_{t+h}^{(t/T)}\|_2 + \|X_{t+h}^{(t/T)}\|_2 \|X_{t,T} - X_t^{(t/T)}\|_2 \\ & \leq T^{-1} \{(h + 1)\|X_{t,T}\|_2 + \|X_{t+h}^{(t/T)}\|_2\} P_{t,T}^{(t/T)} \end{aligned}$$

and similarly

$$\begin{aligned} & \|X_{t+k,T} \otimes X_{t+k+h,T} - X_{t+k}^{(t/T)} \otimes X_{t+k+h}^{(t/T)}\|_{2,2} \\ & \leq T^{-1} \{(|k + h| + 1)\|X_{t+k,T}\|_2 + (|k| + 1)\|X_{t+k+h}^{(t/T)}\|_2\} P_{t,T}^{(t/T)}. \end{aligned}$$

Assembling bounds, we obtain that

$$\begin{aligned} & \frac{1}{mT} \int_{[0,1]^3} \sum_{i=1}^{\lfloor uT \rfloor \wedge (T-h)} \left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,1}(\tau_1, \tau_2) \right)^2 d(u, \tau_1, \tau_2) \\ & \leq \frac{1}{mT} \sum_{i=1}^{T-h} \sum_{t,t'=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\bar{n}_{t,h} \bar{n}_{t',h}} \sum_{k=\underline{n}_t}^{\bar{n}_{t,h}} \sum_{k'=\underline{n}_{t'}}^{\bar{n}_{t',h}} \frac{1}{T^2} \mathbb{E} \left[(|k| + h + 1)(|k'| + h + 1) P_{t,T}^{(t/T)} P_{t',T}^{(t'/T)} \right. \\ & \quad \times (\|X_{t,T}\|_2 + \|X_{t+h}^{(t/T)}\|_2 + \|X_{t+k,T}\|_2 + \|X_{t+k+h}^{(t/T)}\|_2) \\ & \quad \left. \times (\|X_{t',T}\|_2 + \|X_{t'+h}^{(t'/T)}\|_2 + \|X_{t'+k',T}\|_2 + \|X_{t'+k'+h}^{(t'/T)}\|_2) \right] \\ & \leq \frac{C}{mT^3} \sum_{i=1}^{T-h} \sum_{t,t'=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\bar{n}_{t,h} \bar{n}_{t',h}} \sum_{k=\underline{n}_t}^{\bar{n}_{t,h}} \sum_{k'=\underline{n}_{t'}}^{\bar{n}_{t',h}} (|k| + h + 1)(|k'| + h + 1) = O\left(\frac{mn^2}{T^2}\right), \end{aligned}$$

which converges to zero by Assumption (B2).

For the term $A_{t,2}$, first observe that, by Jensen’s inequality for convex functions,

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,2}(\tau_1, \tau_2) \right)^2 \right] \\ & \leq m \sum_{t=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\bar{n}_{t,h}^2} \sum_{k,k'=\underline{n}_t}^{\bar{n}_{t,h}} \text{Cov}\{X_{t+k,T}(\tau_1)X_{t+k+h,T}(\tau_2), X_{t+k',T}(\tau_1)X_{t+k'+h,T}(\tau_2)\}. \end{aligned}$$

By the same arguments as in the proof of Proposition C.7 of the supplementary material and Assumption (A3), one can see that the right-hand side of the inequality

$$\begin{aligned} & \frac{1}{mT} \sum_{i=1}^{T-h} \int_{[0,1]^2} \mathbb{E} \left[\left(\sum_{t=i}^{(i+m-1) \wedge (T-h)} A_{t,2}(\tau_1, \tau_2) \right)^2 \right] d(\tau_1, \tau_2) \\ & \leq \frac{1}{T} \sum_{i=1}^{T-h} \sum_{t=i}^{(i+m-1) \wedge (T-h)} \frac{1}{\bar{n}_{t,h}^2} \sum_{k,k'=\underline{n}_t}^{\bar{n}_{t,h}} \|\text{Cov}(X_{t+k,T} \otimes X_{t+k+h,T}, X_{t+k',T} \otimes X_{t+k'+h,T})\|_{1,2} \end{aligned}$$

is of order $\mathcal{O}(m/n)$. The assertion follows since $m/n = o(1)$ by Assumption (B2). \square

Proof of Proposition 1 The cumulative distribution function of the $(h+2)$ nd coordinate of \mathcal{S} is continuous by Theorem 7.5 of Davydov and Lifshits (1985). The assertion under the null hypothesis follows from Lemma 4.1 in Bücher and Kojadinovic (2017). Consistency follows from the fact that the bootstrap quantiles are stochastically bounded by Theorem 2, whereas the test statistic diverges by Corollary 1. \square

Acknowledgements Financial support by the Collaborative Research Center “Statistical modeling of non-linear dynamic processes” (SFB 823, Teilprojekt A1, A7 and C1) of the German Research Foundation, by the Ruhr University Research School PLUS, funded by Germany’s Excellence Initiative [DFG GSC 98/3], by the DAAD (German Academic Exchange Service) and the National Institute of General Medical Sciences of the National Institutes of Health (Award Number R01GM107639) is gratefully acknowledged. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Parts of this paper were written when Axel Bücher was a postdoctoral researcher at Ruhr-Universität Bochum and while Florian Heinrichs was visiting the Universidad Autónoma de Madrid. The authors would like to thank the institute, and in particular Antonio Cuevas, for its hospitality.

References

- Antoniadis, A., Sapatinas, T. (2003). Wavelet methods for continuous time prediction using hilbert-valued autoregressive processes. *Journal of Multivariate Analysis*, 87, 133–158.
- Aston, J. A. D., Kirch, C. (2012). Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis*, 109(Supplement C), 204–220.
- Aue, A., van Delft, A. (2017). Testing for stationarity of functional time series in the frequency domain. *ArXiv e-prints*.
- Aue, A., Gabrys, R., Horváth, L., Kokoszka, P. (2009). Estimation of a change-point in the mean function of functional data. *Journal of Multivariate Analysis*, 100, 2254–2269.
- Aue, A., Dubart Nourinho, D., Hörmann, S. (2015). On the prediction of stationary functional time series. *Journal of the American Statistical Association*, 110, 378–392.
- Berkes, I., Gabrys, R., Horvath, L., Kokoszka, P. (2009). Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(5), 927–946.
- Billingsley, P. (1999). *Convergence of probability measures*. New York: Wiley.
- Bosq, D. (2000). *Linear processes in function spaces*, Vol. 149., Lecture notes in statistics New York: Springer.
- Bosq, D. (2002). Estimation of mean and covariance operator of autoregressive processes in Banach spaces. *Statistical inference for Stochastic Processes*, 5, 287–306.
- Box, G. E. P., Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), 1509–1526.
- Brillinger, D. (1981). *Time series: Data analysis and theory*. San Francisco: Holden Day Inc.
- Bücher, A., Kojadinovic, I. (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, 22(2), 927–968.
- Bücher, A., Kojadinovic, I. (2017). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *Journal of Theoretical Probability*, 1–21.
- Bücher, A., Fermanian, J.-D., Kojadinovic, I. (2018). Combining cumulative sum change-point detection tests for assessing the stationarity of univariate time series. *ArXiv e-prints*.
- Davydov, Y. A., Lifshits, M. A. (1985). Fubini method in some probabilistic problems. *Journal of Soviet Mathematics*, 31(2), 2796–2858.
- Dehling, H., Sharipov, O. (2005). Estimation of mean and covariance operator for banach space valued autoregressive processes with independent innovations. *Statistical Inference for Stochastic Processes*, 8, 137–149.

- Dehling, H., Durieu, O., Volny, D. (2009). New techniques for empirical processes of dependent data. *Stochastic Processes and their Applications*, 119(10), 3699–3718.
- Dette, H., Preuß, P., Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *Journal of the American Statistical Association*, 106(495), 1113–1124.
- Dwivedi, Y., Subba Rao, S. (2011). A test for second-order stationarity of a time series based on the discrete fourier transform. *Journal of Time Series Analysis*, 32, 68–91.
- Ferraty, F., Vieu, P. (2006). *Nonparametric functional data analysis: Theory and practice*. New York: Springer.
- Fisher, R. (1932). *Statistical methods for research workers*. London: Oliver and Boyd.
- Hörmann, S., Kokoszka, P. (2010). Weakly dependent functional data. *Annals of Statistics*, 38(3), 1845–1884.
- Hörmann, S., Kidziński, Ł., Hallin, M. (2015). Dynamic functional principal components. *Journal of the Royal Statistical Society: Series B*, 77(2), 319–348.
- Horváth, L., Kokoszka, P. (2012). *Inference for functional data with applications*. New York: Springer.
- Horvath, L., Huskova, M., Kokoszka, P. (2010). Testing the stability of the functional autoregressive process. *Journal of Multivariate Analysis*, 101(2), 352–367.
- Hsing, T., Eubank, R. (2015). *Theoretical foundations of functional data analysis, with an introduction to linear operators*. New York: Wiley.
- Hyndman, R. J., Shang, H. L. (2009). Forecasting functional time series. *Journal of the Korean Statistical Society*, 38(3), 199–211.
- Janson, S., Kaijser, S. (2015). Higher moments of Banach space valued random variables. *Memoirs of the American Mathematical Society*, 238(1127), vii+110.
- Jentsch, C., Subba Rao, S. (2015). A test for second order stationarity of a multivariate time series. *Journal of Econometrics*, 185, 124–161.
- Jin, L., Wang, S., Wang, H. (2015). A new non-parametric stationarity test of time series in the time domain. *Royal Statistical Society*, 77, 893–922.
- Lee, J., Subba Rao, S. (2017). A note on general quadratic forms of nonstationary stochastic processes. *Statistics*, 51(5), 949–968.
- Ljung, G. M., Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Panaretos, V. M., Tavakoli, S. (2013). Fourier analysis of stationary time series in function space. *Annals of Statistics*, 41(2), 568–603.
- Politis, D. N., White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1), 53–70.
- Sharipov, O., Tewes, J., Wendler, M. (2016). Sequential block bootstrap in a hilbert space with application to change point analysis. *Canadian Journal of Statistics*, 44(3), 300–322.
- Statulevicius, V., Jakimavicius, D. (1988). Estimates of semiinvariants and centered moments of stochastic processes with mixing. I. *Lithuanian Mathematical Journal*, 28, 226–238.
- van Delft, A., Eichler, M. (2018). Locally stationary functional time series. *Electronic Journal of Statistics*, 12(1), 107–170.
- van Delft, A., Bagchi, P., Characiejus, V., Dette, H. (2017). A nonparametric test for stationarity in functional time series. *ArXiv e-prints*.
- van der Vaart, A., Wellner, J. (1996). *Weak convergence and empirical processes*, Vol. 1., Springer series in statistics New York: Springer.
- Vogt, M. (2012). Nonparametric regression for locally stationary time series. *The Annals of Statistics*, 40, 2601–2633.
- Weidmann, J. (1980). *Linear operators in Hilbert spaces*, Vol. 68., Graduate texts in mathematics New York: Springer.