Electronic Journal of Statistics
Vol. 16 (2022) 4823–4862
ISSN: 1935-7524
https://doi.org/10.1214/22-EJS2059

Lagged covariance and cross-covariance operators of processes in Cartesian products of abstract Hilbert spaces

Sebastian Kühnert

e-mail: s.kuehnert_math@gmx.de

Abstract: A major task in Functional Time Series Analysis is measuring the dependence within and between processes for which lagged covariance and cross-covariance operators have proven to be a practical tool in well-established spaces. This article focuses on estimating these operators of processes in Cartesian products of abstract Hilbert spaces. We derive precise asymptotic results for the estimation errors for fixed and increasing lag and Cartesian powers under very mild conditions, presumably even under the mildest that can be assumed, establish estimators for the principal components, and conduct a simulation study.

MSC2020 subject classifications: 47B10, 60G05, 62J10.
Keywords and phrases: Estimation, functional time series, lag-h-covariance operator, lag-h-cross covariance operator, principal components.

Received May 2021.

1. Introduction

Functional Data Analysis (FDA) and *Functional Time Series Analysis* (FTSA), the research areas dealing with random functions resp. time series/processes of random functions, have gained more and more significance. This is because considering random functions instead of vectors, provided the context allows it, assures more accurate results. The extension on infinite-dimensional spaces is enabled by ongoing developments in processing techniques, being unproblematic for separable Banach spaces from a mathematical point of view, see [40]. FDA/FTSA find applications in economics [13, 14, 32, 47, 52], medicine [8, 56] and other areas [4, 20, 42], and for extensive introductions, see [7, 19, 28, 31, 48]. In FTSA, analyzing the dependence within and between processes is of great importance. If these are weak stationary, where mostly strictly stationarity and finite second moments are assumed, this can be done with *lag-h-covariance operators* resp. *lag-h-cross-covariance operators*, where the *lag h* indicates the time difference of interest. Thereby, for articles that have dealt with stationarity of functional time series, see [2, 15, 17, 29]. Another important subject of study is *Functional Principal Component Analysis* (FPCA), see [24, 33], since FPCs, the eigenvalues and eigenfunctions of the covariance operator, yield an efficient representation.

Related work Estimates for lag-h-covariance operators of stationary processes in the separable Hilbert space $L^2[0,1]$ of measurable, square-Lebesgue integrable real-valued functions with domain $[0,1]$ are widely studied for fixed lag h, see, e.g., [7, 28, 31, 36, 44]. [49] developed covariance estimates in the space of continuous functions $C[0,1]$, [59] in tensor product Sobolev-Hilbert spaces, [43] for continuous surfaces, and [1, 12, 25] for general, separable Hilbert spaces. [1, 25, 44, 49] constrained their assertions to autoregressive (AR) processes, where [1] investigated random AR(1) operators. Thereby, [1, 7, 28, 31] utilized classical moment estimators, [36] estimated the integral kernels, [25, 44] used truncated spectral decompositions with estimated FPCs, and [59] used operator regularized covariance estimates. Moreover, [51] studied covariance networks for functional data on multidimensional domains, and [41, 60] dealt with covariance estimation of sparse (multivariate) functional data. The limit distribution of the covariance operator's estimation errors was discussed in [35, 37]. A comprehensive study of *lag-h-cross-covariance operators* of $L^2[0,1]$-valued processes is conducted in Rice & Shum (2019, [46]) who established operator estimates and deduced their limit distribution. Aue & Klepsch (2017, [3]) estimated lagged covariance and cross-covariance operators of processes in Cartesian products of $L^2[0,1]$ to deduce asymptotic assertions regarding estimators for operators of linear, invertible processes in $L^2[0,1]$. Enabling processes to have values in Cartesian products was also handy in the study of AR(p) processes with $p>1$, see [7]. Also worth mentioning is that [53] derived bootstraps applicable to covariance and cross-covariance operators, that [45] focused on testing equality of covariance operators, and [30] dealt with change point analysis of covariance functions. Moreover, [5] discussed copulas to model the dependence structure in arbitrary dimensions, and [16] discussed a similarity measure for second order properties of non-stationary functional time series.

FPCA in $L^2[0,1]$ is also extensively discussed in the existing literature. In [7, 28, 31, 35, 37] one finds asymptotic upper bounds of the estimation errors for FPCs, estimated seperately and uniformly, in second mean and almost surely (a.s.), and [61] introduced L^1-norm FPCA.

Contributions This article studies, inspired by results in [3, 46] and Kühnert (2019, [38] and 2020, [39]), lagged covariance and cross-covariance operators of stationary processes in separable Hilbert spaces, in particular of processes in Cartesian products of abstract Hilbert spaces. To be precise, we analyze processes $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ and $\boldsymbol{\mathscr{Y}} := (\mathscr{Y}_k)_{k\in\mathbb{Z}}$ defined for some $m,n,p,q\in\mathbb{N}$ by

$$\mathscr{X}_{m+j} := (X_{m+jp},\ldots,X_{1+jp})^T,\ j\in\mathbb{Z}, \tag{1.1}$$

$$\text{resp.}\quad \mathscr{Y}_{n+j} := (Y_{n+jq},\ldots,Y_{1+jq})^T,\quad j\in\mathbb{Z}, \tag{1.2}$$

whose elements come from stationary processes $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}$ resp. $\boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$ with values in general separable Hilbert spaces. Important separable Hilbert spaces are $(L^2(D))^m$, with bounded domain $D\subset\mathbb{R}^n$ and $m,n\in\mathbb{N}$ (see Example 2.1), and the space of Hilbert-Schmidt operators mapping between two separable Hilbert spaces (see Section 3.2). Also, Sobolev spaces

$W^{k,2}(\Omega)$, with $\Omega \subset \mathbb{R}^n$ denoting an open set and $k, n \in \mathbb{N}$, consisting of functions $f \in L^2(\Omega)$ whose weak derivates of order k have finite L^2 norm. The use of successive stacking samples of time series, is, generally speaking, canonical when dealing with time series whose modification obtained by 'stacking' satisfies managable equations or simplifies the initial representation, see Example 2.2. Such an approach was used to estimate the operators of $L^2[0,1]$-valued AR in [7], (G)ARCH in [38, 39], and invertible linear processes in [3, 38, 39]. Our specific definitions (1.1), (1.2) are useful in various scenarios. They enable reusing entries of $\mathscr{X}_k$ resp. $\mathscr{Y}_k$ to enlargen the sample sizes when choosing $1 \leq p < m, 1 \leq q < n$, where the largest sample size is obtained for $p = q = 1$, and $1 < p < m, 1 < q < n$ allows observing only a certain season. The definitions also enable successive stacking (realizations of) $X_k's$ resp. $Y_k's$ without reusing values by putting $p = m, q = n$, and to bridge the time indices where (realizations of) $X_k's$ and/or $Y_k's$ are missing by an appropriate choice of $p > m, q > n$. Another advantage of our definitions are that our assertions hold also for processes not obtained by stacking elements (see Example 2.1). The focus of this paper is to deduce moment estimators for lagged covariance and cross-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ resp. $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ of the processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} = (\mathscr{Y}_k)_k$, and to derive the asymptotic behaviour of their estimation errors. We also develop the 'classical' FPC estimates in our more general setting, and derive asymptotic results for the estimation errors for the FPCs estimated separately and uniformly. Our results are stated for centered processes and for those having an unknown, finite first moment. The lag and the processes' Cartesian powers can be fixed or increase with regard to (w.r.t.) the sample sizes, and the two processes in the definition of the lagged cross-covariance operators not necessarily have to attain values in the same space.

Structure The rest of this paper is organized as follows. Section 2 states motivational examples. Section 3 outlines our notation, restates basic operator theory, defines our (lagged) (cross-)covariance operators as well as studies their probabilistic features, and briefly reiterates concepts of weak dependence. Section 4 deals with the estimation. Further, Section 5 conducts a simulation study, and Section 6 summarizes the main results and outlines future research. Moreover, Appendix A contains proofs, and Appendix B side results.

2. Motivational examples

The examples herein illustrate the use of lagged (cross-)covariance operators of our processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} = (\mathscr{Y}_k)_k$ in different scenarios.

Example 2.1 (Fixed lag, fixed Cartesian powers). Investors of solar stocks of European companies might be interested to measure the impact of the monthly sunshine duration in central Europe (Fig. 1) interpretable as realizations of the (four) share values of their portfolio (Fig. 2) one month ahead. Consecutive observations of the monthly sunshine duration and monthly values of the four shares can be interpreted as realizations of a process $(X_k)_k$ in $L^2[0,1]^2$ resp.

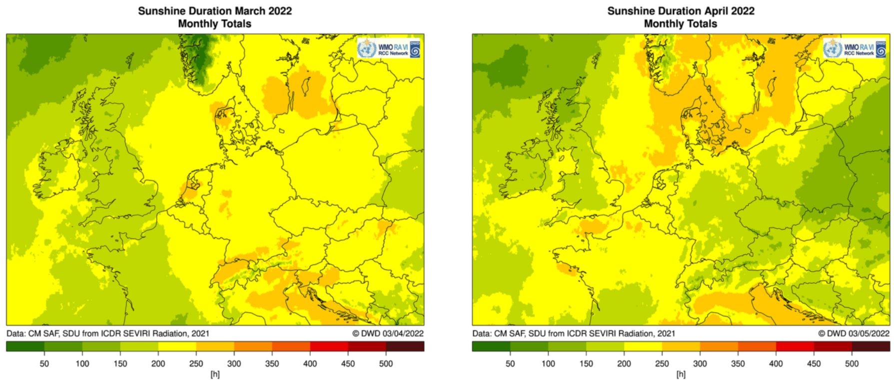

Fig 1. *Graphs of monthly sunshine duration in central europe in March and April 2022, retrieved from the homepage www.dwd.de of the German Meterological Service.*

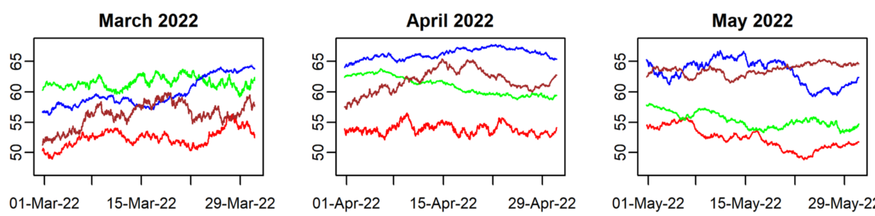

Fig 2. *Three consecutive realizations of a fictitious process describing the share values of four assets of an portfolio, e.g., measured in EUR.*

$(Y_k)_k$ in $(L^2[0,1])^4$. By using $(\mathscr{X}_k)_k$ with $m = p = 1$ and $(\mathscr{Y}_k)_k$ with $n = q = 1$, the context in question can be described by our lag-1-cross-covariance operators.

Example 2.2 (Increasing Cartesian powers). For linear $L^2[0,1]$-valued processes $\boldsymbol{X} = (X_k)_k$, i.e. $X_k = \sum_{l=0}^{\infty} \phi_l(\varepsilon_{k-l})$ for all k, with ϕ_l denoting bounded, linear operators, [3, 38, 39] derived consistent estimates for all ϕ_l. The estimation procedure was based on Yule-Walker equations consisting of lagged (cross-)covariance operators of processes, where $(\mathscr{X}_k)_k$ with $p = 1$ was used, with the Cartesian powers $m = m_M \in \mathbb{N}$ approaching infinity as the sample size M of a sample $X_1, \dots, X_M$ of $\boldsymbol{X}$ did.

Example 2.3 (Increasing lag). When launching satellites which communicate with ground stations or other satellites (see Fig. 3), one could wonder about the impact of complex data transmission between the objects drifting apart in time. The dependency of the sent satellite's/satelittes' to the received ground station's/stations' or other satellite's/satelittes' information can be modelled by $(\mathscr{X}_k)_k$ for certain m, p resp. $(\mathscr{Y}_k)_k$ for certain n, q, and can be described by our lag-h-cross-covariance operators with increasing lag $h = h_{M,N}$ for increasing sample sizes M, N of samples $X_1, \dots, X_M$ and $Y_1, \dots, Y_N$ of processes modeling the individual elements.

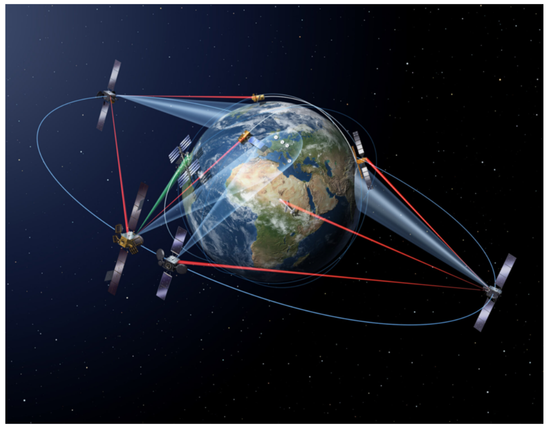

Fig 3. *The European Data Relay Satellite System (EDRS) – a system of geostationary communications satellites providing data transmission between satellites, UAVs and ground stations. Photo from* https://www.esa.int/ESA_Multimedia/Images/2016/02/Inter-satellite_laser_links

3. Definitions and basics

3.1. General notation

For any set B, we write B^c for the complement of B, and $\mathbf{1}_B(\cdot)$ for the *indicator function*. Moreover, $\lfloor\cdot\rfloor$ and $\operatorname{sgn}(\cdot)$ denote the *floor* resp. *sign function*, and for any $x \in \mathbb{R}, x^+ := \max(0, x)$ denotes the *positive part* of x. Further, we write $a \vee b := \max(a, b)$ and $a \wedge b := \min(a, b)$ for any $a, b \in \mathbb{R}$. For sequences $(a_n)_n$, $(b_n)_n \subseteq (0, \infty)$, $a_n \sim b_n \Leftrightarrow \frac{a_n}{b_n} \to 1$, the common asymptotic notation is denoted by $\mathrm{o}(\cdot), \mathrm{O}(\cdot)$, and (for $n \to \infty$) $a_n = \omega(b_n) \Leftrightarrow b_n = \mathrm{o}(a_n)$ and $a_n = \Omega(b_n) \Leftrightarrow b_n = \mathrm{O}(a_n)$, and $\Xi(a_n, b_n) := \omega(a_n) \cap \mathrm{o}(b_n)$. 0_V denotes the *identity element of addition* of a vector space V, $\mathbb{I}_V : V \to V$ the *identity operator*, and *operator* throughout means linear map. On Hilbert spaces we assume the norms to be induced by their inner product. For a separable Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}), x \perp y \Leftrightarrow \langle x, y\rangle_{\mathcal{H}} = 0$ for $x, y \in \mathcal{H}$. Scalar multiplication and vector addition on $\mathcal{H}^n := \{(x_1, \dots, x_n)^T | x_1, \dots, x_n \in \mathcal{H}\}$, with $n \in \mathbb{N}$, are defined componentwise, so $(\mathcal{H}^n, \langle\cdot,\cdot\rangle_{\mathcal{H}^n})$ with $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_{\mathcal{H}^n} := \sum_{i=1}^n \langle x_i, y_i\rangle_{\mathcal{H}}$ for $\boldsymbol{x} := (x_1, \dots, x_n)^T$, $\boldsymbol{y} := (y_1, \dots, y_n)^T \in \mathcal{H}^n$ is a separable Hilbert space. Our random variables are defined on a common probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. $X \overset{d}{=} Y$ denotes equally distributed random variables X, Y. For processes $(X_n)_n$ and $(Y_n)_n, X_n = \mathrm{O}_{\mathbb{P}}(Y_n)$ (for $n \to \infty$) means $(X_n/Y_n)_n$ is asymptotically $\mathbb{P}$-stochastic bounded. For $p \in [1, \infty)$, $L^p_{\mathcal{H}} = L^p_{\mathcal{H}}(\Omega, \mathfrak{A}, \mathbb{P})$ is *the space of (classes of) $\mathcal{H}$-valued random variables X with* $\nu_{p,\mathcal{H}}(X) := (\mathbb{E}||X||^p_{\mathcal{H}})^{1/p} < \infty$, and a process $(X_k)_k$ is an $L^p_{\mathcal{H}}$*-process* if $(X_k)_k \subseteq L^p_{\mathcal{H}}$. Moreover, $X \in L^1_{\mathcal{H}}$ is *centered* if $\mathbb{E}(X_k) = 0_{\mathcal{H}}$ for all k with expectation in Bochner-integral sense, see [31], p. 40–45, and centering of $X \in L^1_{\mathcal{H}}$ is denoted by $X' := X - \mathbb{E}(X)$.

3.2. Basic operator theory

Here, we state useful spaces of (linear) operators between Hilbert spaces, see [18, 21, 57, 58] for thorough overviews. Throughout, let $(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i})$ denote real, separable Hilbert spaces for $i = 1, 2$. The space of bounded operators mapping from $\mathcal{H}_1$ to $\mathcal{H}_2$ will be denoted by $\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$, with $\mathcal{L}_{\mathcal{H}_1} := \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_1}$, where an operator $A\colon \mathcal{H}_1 \to \mathcal{H}_2$ is *bounded* if $||A||_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}} := \sup_{||x||_{\mathcal{H}_1}\leq 1} ||A(x)||_{\mathcal{H}_2} < \infty$. Such operators are continuous, and $(\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}, ||\cdot||_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}})$ is a Banach space. We denote the subspace of finite-rank operators of $\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$ by $\mathcal{F}_{\mathcal{H}_1,\mathcal{H}_2}$, with $\mathcal{F}_{\mathcal{H}_1} := \mathcal{F}_{\mathcal{H}_1,\mathcal{H}_1}$. A^* denotes the adjoint of $A \in \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$, with $A^* \in \mathcal{L}_{\mathcal{H}_2,\mathcal{H}_1}$. A crucial subspace of $\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$ is the space of compact operators mapping from $\mathcal{H}_1$ to $\mathcal{H}_2$, where $A \in \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$ is *compact* if A maps the unit ball of $\mathcal{H}_1$ to a compact set in $\mathcal{H}_2$. Such operators possess the *singular value decomposition* $A = \sum_{j=1}^{\infty} s_j(e_j \otimes f_j)$, with $x \otimes y := \langle x,\cdot\rangle_{\mathcal{H}_1} y$ for $x \in \mathcal{H}_1, y \in \mathcal{H}_2$, where $(e_j)_{j\in\mathbb{N}}$ and $(f_j)_{j\in\mathbb{N}}$ are complete orthonormal systems (CONSs) of $\mathcal{H}_1$ resp. $\mathcal{H}_2$, and $(s_j)_{j\in\mathbb{N}}$ the monotonically decreasing zero sequence of non-negative numbers, the *singular values* of A. Their decay rate is interpretable as a regularity measure of A and can be expressed by the *p-Schatten-norm* $||A||_p := \big(\sum_{j=1}^{\infty} s_j^p\big)^{1/p}$ for $p \in [1,\infty)$, where $||A||_p \leq ||A||_q$ for $p < q$. $(\mathscr{S}^p_{\mathcal{H}_1,\mathcal{H}_2}, ||\cdot||_p)$ is a Banach space for $p \in [1,\infty)$, where $\mathscr{S}^p_{\mathcal{H}_1,\mathcal{H}_2} := \{A \in \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2} \big| \; ||A||_p < \infty\}$ is the *p-Schatten-class*, with $\mathscr{S}^p_{\mathcal{H}_1,\mathcal{H}_2} \subsetneq \mathscr{S}^q_{\mathcal{H}_1,\mathcal{H}_2}$ for $p < q$. The essential classes are $\mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2} := \mathscr{S}^1_{\mathcal{H}_1,\mathcal{H}_2}$ with $\mathcal{N}_{\mathcal{H}_1} := \mathcal{N}_{\mathcal{H}_1,\mathcal{H}_1}$, $||\cdot||_{\mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2}} := ||\cdot||_1$, and $\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2} := \mathscr{S}^2_{\mathcal{H}_1,\mathcal{H}_2}$ with $\mathcal{S}_{\mathcal{H}_1} := \mathcal{S}_{\mathcal{H}_1,\mathcal{H}_1}$, $||\cdot||_{\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}} := ||\cdot||_2$, the spaces of *nuclear/trace class* resp. *Hilbert-Schmidt operators*. For any CONS $(e_j)_{j\in\mathbb{N}}$ of $\mathcal{H}_1$, the *trace* of $A \in \mathcal{N}_{\mathcal{H}_1}$ is defined by $\mathrm{tr}(A) := \sum_{j=1}^{\infty}\langle A(e_j), e_j\rangle_{\mathcal{H}_1}$, and $(\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}, \langle\cdot,\cdot\rangle_{\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}})$ is a separable Hilbert space, with the inner product defined as $\langle A, B\rangle_{\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}} := \sum_{j=1}^{\infty} \langle A(e_j), B(e_j)\rangle_{\mathcal{H}_2}$ for $A, B \in \mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}$. From this inner product can be derived

$$\langle x \otimes x', y \otimes y'\rangle_{\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}} = \langle x, y\rangle_{\mathcal{H}_1}\langle x', y'\rangle_{\mathcal{H}_2}, \quad x, y \in \mathcal{H}_1,\; x', y' \in \mathcal{H}_2. \tag{3.1}$$

Furthermore, on $\mathcal{H}_1 := L^2[0,1]$, an *integral operator* $A\colon \mathcal{H}_1 \to \mathcal{H}_1$ is defined by the Lebesgue integral $(A(x))(t) := \int_0^1 a(s,t)x(s)\,\mathrm{d}s$ for any $x \in \mathcal{H}_1, t \in [0,1]$ if it exists, where $a\colon [0,1]^2 \to \mathbb{R}$ is a measurable function, the (*integral*) *kernel* of A. Such an operator satisfies $A \in \mathcal{S}_{\mathcal{H}_1}$ if and only if $\int_0^1\int_0^1 a^2(s,t)\,\mathrm{d}s\,\mathrm{d}t < \infty$.

3.3. Lagged covariance and cross-covariance operators

Here, we formally define (cross-)covariance operators and their lagged versions on real, separable Hilbert spaces, denoted by $(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i})$ for $i = 1, 2$, and outline some of their features (see [7] for these operators on Banach spaces).

Definition 3.1. *Let $X \in L^2_{\mathcal{H}_1}$ and $Y \in L^2_{\mathcal{H}_2}$. Then, the covariance operator of X and the cross-covariance operator of X, Y are defined by*

$$\mathscr{C}_X := \mathbb{E}(X' \otimes X') \qquad \text{resp.} \qquad \mathscr{C}_{X,Y} := \mathbb{E}(X' \otimes Y').$$

For centered random variables $X \in L^2_{\mathcal{H}_1}$ and $Y \in L^2_{\mathcal{H}_2}$, (cross-)covariance operators possess the following features, see [7] and also [31], sections 7.2-7.3. Firstly, $\mathscr{C}_X \in \mathcal{N}_{\mathcal{H}_1}$ is a self-adjoint, positive semi-definite operator with

$$||\mathscr{C}_X||_{\mathcal{N}_{\mathcal{H}_1}} = \mathbb{E}||X||^2_{\mathcal{H}_1}, \tag{3.2}$$

$$\mathscr{C}_X = \sum_{j=1}^{\infty} c_j(\mathfrak{c}_j \otimes \mathfrak{c}_j), \tag{3.3}$$

where $(c_j)_{j\in\mathbb{N}}$ is the without loss of generality (w.l.o.g.) monotonically decreasing, non-negative, absolutely-summable eigenvalue sequence, and $(\mathfrak{c}_j)_{j\in\mathbb{N}}$ the related eigenfunction sequence of $\mathscr{C}_X$ which forms a CONS of $\mathcal{H}_1$. Moreover, for the cross-covariance operator holds $\mathscr{C}_{X,Y} \in \mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2}, \mathscr{C}^*_{X,Y} = \mathscr{C}_{Y,X} \in \mathcal{N}_{\mathcal{H}_2,\mathcal{H}_1}$,

$$||\mathscr{C}_{X,Y}||_{\mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2}} = ||\mathscr{C}_{Y,X}||_{\mathcal{N}_{\mathcal{H}_2,\mathcal{H}_1}} \leq \mathbb{E}||X||_{\mathcal{H}_1}||Y||_{\mathcal{H}_2}, \tag{3.4}$$

$$\text{independence of } X, Y \;\Rightarrow\; \mathscr{C}_{X,Y} = 0_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}}, \tag{3.5}$$

and if $\mathcal{H}_1 = \mathcal{H}_2, \mathscr{C}_{X,Y} = 0_{\mathcal{L}_{\mathcal{H}_1}}$ implies $\mathbb{E}\langle X, Y\rangle_{\mathcal{H}_1} = 0$. If $\mathcal{H}_1 = \mathcal{H}_2 = L^2[0,1], \mathscr{C}_X$ and $\mathscr{C}_{X,Y}$ are integral operators with kernels $k_X(s,t) \coloneqq \mathrm{Cov}(X(s), X(t))$ resp. $k_{X,Y}(s,t) \coloneqq \mathrm{Cov}(X(s), Y(t)), s,t \in [0,1]$. Further, for centered random variables $W, X \in L^2_{\mathcal{H}_1}$ and an operator $A \in \mathcal{L}_{\mathcal{H}_1}$, holds

$$\mathscr{C}_{W+X} = \mathscr{C}_W + \mathscr{C}_{W,X} + \mathscr{C}_{X,W} + \mathscr{C}_X, \tag{3.6}$$

$$\mathscr{C}_{A(X)} = A\mathscr{C}_X A^*, \tag{3.7}$$

and if $W, X \in L^2_{\mathcal{H}_1}, Y, Z \in L^2_{\mathcal{H}_2}$ are centered, and $A \in \mathcal{L}_{\mathcal{H}_1}, B \in \mathcal{L}_{\mathcal{H}_2}$,

$$\mathscr{C}_{W+X,Y+Z} = \mathscr{C}_{W,Y} + \mathscr{C}_{W,Z} + \mathscr{C}_{X,Y} + \mathscr{C}_{X,Z}, \tag{3.8}$$

$$\mathscr{C}_{A(X),B(Y)} = B\mathscr{C}_{X,Y}A^*. \tag{3.9}$$

In order to define the functional counterparts of the auto-covariance and cross-covariance function, the lag-h-covariance resp. lag-h-cross-covariance operators, given processes not necessarily have to be strictly, but weak stationary.

Definition 3.2. *Let $(X_k)_{k\in\mathbb{Z}}$ be an $\mathcal{H}_1$-valued process.*

(a) *$(X_k)_k$ is (strictly) stationary if $(X_{k_1+h}, \ldots, X_{k_n+h}) \overset{d}{=} (X_{k_1}, \ldots, X_{k_n})$ holds for all $k_1, \ldots, k_n, h \in \mathbb{Z}$ where $n \in \mathbb{N}$.*
(b) *$(X_k)_k$ is weak stationary if it is an $L^2_{\mathcal{H}_1}$-process, if $\mathbb{E}(X_k) = c$ holds for some $c \in \mathcal{H}_1$ for all k, and if $\mathscr{C}_{X_k,X_l} = \mathscr{C}_{X_0,X_{l-k}}$ for all k, l.*
(c) *$(X_k)_k$ is an $\mathcal{H}_1$-white noise if it is a centered $L^2_{\mathcal{H}_1}$-process with $\mathbb{E}||X_k||^2 > 0$ for all k, if $\mathscr{C}_{X_k}$ does not depend on k, and if $\mathscr{C}_{X_k,X_l} = 0_{\mathcal{L}_{\mathcal{H}_1}}$ for $k \neq l$.*
(d) *$(X_k)_k$ is a strong $\mathcal{H}_1$-white noise if it is a centered, i.i.d. $L^2_{\mathcal{H}_1}$-process with $\mathbb{E}||X_0||^2 > 0$.*

Definition 3.3. *Let $\boldsymbol{X} \coloneqq (X_k)_{k\in\mathbb{Z}} \subseteq L^2_{\mathcal{H}_1}$ and $\boldsymbol{Y} \coloneqq (Y_k)_{k\in\mathbb{Z}} \subseteq L^2_{\mathcal{H}_2}$ be weak stationary processes, and let $h \in \mathbb{Z}$. Then, the lag-h-covariance operator of $\boldsymbol{X}$ and the lag-h-cross-covariance operator of $\boldsymbol{X}, \boldsymbol{Y}$ is defined by*

$$\mathscr{C}_{\boldsymbol{X};h} \coloneqq \mathscr{C}_{X_0,X_h} \qquad \textit{resp.} \qquad \mathscr{C}_{\boldsymbol{X}\boldsymbol{Y};h} \coloneqq \mathscr{C}_{X_0,Y_h}.$$

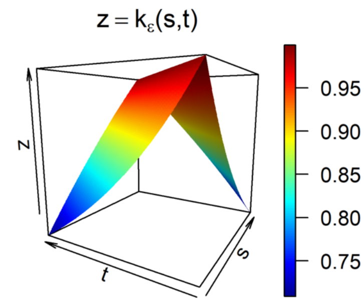

FIG 4. *The integral kernel* $k_{\boldsymbol{\varepsilon}}(s,t)$ *in* (3.11) *for* $s,t \in [0,1]$.

*We also call the lag-*0*- covariance operator of* $\boldsymbol{X}$*, and write* $\mathscr{C}_{\mathbf{X}} := \mathscr{C}_{\mathbf{X};0}$.

Remarks 3.1. The features of (cross-)covariance apply to lag-h-(cross-)covariance operators. Thus, $\mathscr{C}^*_{\mathbf{X};h} = \mathscr{C}_{\mathbf{X};-h}$ and $\mathscr{C}^*_{\mathbf{X},\mathbf{Y};h} = \mathscr{C}_{\mathbf{Y},\mathbf{X};-h}$ for any h, $\mathscr{C}_{\mathbf{X};h} = 0_{\mathcal{L}_{\mathcal{H}_1}}$ for $h \neq 0$ if $\boldsymbol{X} := (X_k)_k$ consists of independent variables, and if $\boldsymbol{X} := (X_k)_k, \boldsymbol{Y} := (Y_k)_k$ are independent, $\mathscr{C}_{\mathbf{X},\mathbf{Y};h} = 0_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}}$. Further, if $\mathcal{H}_1 = \mathcal{H}_2 = L^2[0,1]$, the lag-$h$-(cross-)covariance are integral operators with auto-covariance resp. cross-covariance function as integral kernels, which justifies to have the expression '(cross-)covariance' in 'lag-h-(cross-)covariance operator'.

Now, we illustrate a specific covariance operator. For further examples and sketches, see Section 5.

Example 3.1. Let $\mathcal{H} := L^2[0,1]$, and let $\boldsymbol{\varepsilon} := (\varepsilon_k)_{k\in\mathbb{Z}}$ be a process with

$$\varepsilon_k(t) := \frac{Z_k + B_k(t)}{\sqrt{1+t}} \text{ a.s.}, \quad \forall k \in \mathbb{Z},\ \forall t \in [0,1], \tag{3.10}$$

where $Z_k \sim \mathcal{N}(0,1)$, $\boldsymbol{B_k} = (B_k(t))_{t\in[0,1]}$ are Wiener processes, and the random variables $\dots, Z_{-1}, \boldsymbol{B_{-1}}, Z_0, \boldsymbol{B_0}, Z_1, \boldsymbol{B_1}, \dots$ are independent. Then, $(\varepsilon_k)_{k\in\mathbb{Z}}$ is an i.i.d., centered $L^4_{\mathcal{H}}$-process with $\varepsilon_0(t) \sim \mathcal{N}(0,1)$ for all $t \in [0,1]$, and for the integral kernel $k_{\boldsymbol{\varepsilon};0} = k_{\boldsymbol{\varepsilon}}$ of $\mathscr{C}_{\boldsymbol{\varepsilon};0} = \mathscr{C}_{\boldsymbol{\varepsilon}}$ holds

$$k_{\boldsymbol{\varepsilon}}(s,t) = \operatorname{Cov}(\varepsilon_0(s), \varepsilon_0(t)) = \sqrt{\frac{1+s\wedge t}{1+s\vee t}}, \quad \forall s,t \in [0,1]. \tag{3.11}$$

3.4. Functional AR processes

In this section, we recall functional AR(1) processes and their properties (see [7]), which we utilize in our examples and simulation study. Let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ be a separable Hilbert space. Then, a centered $\mathcal{H}$-valued process $\boldsymbol{X} = (X_k)_{k\in\mathbb{Z}}$ is an *autoregressive process of order 1* (AR(1) process) if it satisfies

$$X_k = \alpha(X_{k-1}) + \varepsilon_k, \quad \forall k \in \mathbb{Z}, \tag{3.12}$$

with $\alpha \in \mathcal{L}_{\mathcal{H}}$ and where $\boldsymbol{\varepsilon} := (\varepsilon_k)_{k\in\mathbb{Z}}$ is an i.i.d. $\mathcal{H}$-valued process or an $\mathcal{H}$-white noise. Here, we impose $(\varepsilon_k)_k$ to be a strong $\mathcal{H}$-white noise. If for some $d \in \mathbb{N}$ holds $||\alpha^d||_{\mathcal{L}_{\mathcal{H}}} < 1$, (3.12) has the unique stationary solution

$$X_k = \sum_{i=0}^{\infty} \alpha(\varepsilon_{k-i}), \quad \forall k \in \mathbb{Z},$$

which converges a.s. as well as in $L^2_{\mathcal{H}}$, where α^0 stands for the identity operator $\mathbb{I}_{\mathcal{H}}$. Moreover,

$$\mathscr{C}_{\mathbf{X};h} = \alpha^h \mathscr{C}_{\mathbf{X};0}, \quad h \geq 0, \tag{3.13}$$

where $\mathscr{C}_{\mathbf{X};0}$ can solely be described by $\mathscr{C}_{\boldsymbol{\varepsilon};0}$ and α through

$$\mathscr{C}_{\mathbf{X};0} = \sum_{i=1}^{\infty} \alpha^i \mathscr{C}_{\boldsymbol{\varepsilon};0} \alpha^{*i}. \tag{3.14}$$

Remarks 3.2. For extensive works on functional AR(MA) processes, we refer to [7, 54] and [1, 9, 11, 12, 23, 25, 44] from a technical point of view, and to [14, 34, 50] for methods combined with applications.

3.5. Weak dependence

In order to derive estimators in the context of time series, usually some form of weak dependence is required. A frequently used form is *L^p-m-approximability* developed by Hörmann & Kokoszka (2010, [27]).

Definition 3.4. *Let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ be a separable Hilbert space and let $p \geq 1$. Then, a process $(Z_k)_{k\in\mathbb{Z}}$ is $L^p_{\mathcal{H}}$-m-approximable if it is an $L^p_{\mathcal{H}}$-process with*

$$Z_k = f(\varepsilon_k, \varepsilon_{k-1}, \ldots), \quad \forall k \in \mathbb{Z}, \tag{3.15}$$

where $(\varepsilon_k)_{k\in\mathbb{Z}}$ is an i.i.d. process with values in a measurable space S and where $f\colon S^\infty \to \mathcal{H}$ is a measurable function, such that $\sum_{m=1}^{\infty} \nu_{p,\mathcal{H}}(Z_m - Z_{m;m}) < \infty$, with $\nu_{p,\mathcal{H}}(\cdot) := (\mathbb{E}||\cdot||^p_{\mathcal{H}})^{1/p}$ and

$$Z_{k;m} := f(\varepsilon_k, \varepsilon_{k-1}, \ldots, \varepsilon_{k-m+1}, \varepsilon_{k-m;k}, \varepsilon_{k-m-1;k}, \ldots), \tag{3.16}$$

where $(\varepsilon_{k;n})_k$ are independent copies of $(\varepsilon_k)_k$ for each n.

Hence, $L^p_{\mathcal{H}}$-m-approximability of a process means it is *causal* w.r.t. another process, that is (3.15), and approximable by an m-dependent process so that the approximation errors measured by the $L^p_{\mathcal{H}}$-norm $\nu_{p,\mathcal{H}}(\cdot)$ are summable. Also, (3.15) yields stationarity of $(Z_k)_k$ due to [55], Theorem 3.5.3, and $(Z_{k;m})_k$ are stationary, m-dependent processes for each m with $Z_{k;m} \overset{d}{=} Z_k$ for all k, m. This type of weak dependence is, due to its definition based on m-dependence, particularly feasible for transformations when verifying asymptotic upper bounds. Further, Lemma B.1 shows that L^4-m-approximability of $(X_k)_k$ and

$(Y_k)_k$ transfers to our processes $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ implying L^2-m-approximability of $(\mathscr{X}_k \otimes \mathscr{Y}_{k+h})_k$. Moreover, it can be shown that L^p-m-approximability for $p = 2$, thus for any $p \geq 2$, implies that the *long run covariance operator* defined as follows exists.

Definition 3.5. *Let $\boldsymbol{Z} = (Z_k)_{k\in\mathbb{Z}}$ be a weak stationary $L^2_{\mathcal{H}}$-process. Then, the long run covariance operator is defined as $\sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{Z};k}$ if it exists.*

For the estimation of these operators, see, e.g., [6]. Further, one can show (see [7], p. 87–88 for centered functional AR(1) processes) that for any stationary $L^2_{\mathcal{H}}$-process $\boldsymbol{Z} = (Z_k)_{k\in\mathbb{Z}}$ holds with $Z'_j := Z_j - m_{\boldsymbol{Z}}$ for any j:

$$\Big|\Big| \sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{Z};k} \Big|\Big|_{\mathcal{N}_{\mathcal{H}}} = \sum_{k\in\mathbb{Z}} \mathbb{E}\langle Z'_0, Z'_k\rangle_{\mathcal{H}}. \tag{3.17}$$

Example 3.2. Let $\boldsymbol{X} = (X_k)_k$ be a centered, $\mathcal{H}$-valued AR(1) process as in (3.12) with $||\alpha||_{\mathcal{L}_{\mathcal{H}}} < 1$. Then, the long run covariance operator $\sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{X};k}$ exists, because (3.13), $\mathscr{C}^*_{\boldsymbol{X},h} = \mathscr{C}_{\boldsymbol{X};-h}$ and $||\mathscr{C}_{\boldsymbol{X};h}||_{\mathcal{N}_{\mathcal{H}}} = ||\mathscr{C}^*_{\boldsymbol{X},h}||_{\mathcal{N}_{\mathcal{H}}}$ for all h, submultiplicity of the operator norm, $||\mathscr{C}_{\boldsymbol{X};0}||_{\mathcal{N}_{\mathcal{H}}} = \mathbb{E}||X_0||^2_{\mathcal{H}} =: c < \infty$ and further basic conversions yield

$$\begin{aligned}\sum_{k\in\mathbb{Z}} ||\mathscr{C}_{\boldsymbol{X};k}||_{\mathcal{N}_{\mathcal{H}}} &= ||\mathscr{C}_{\boldsymbol{X};0}||_{\mathcal{N}_{\mathcal{H}}} + 2\sum_{k=1}^{\infty} ||\mathscr{C}_{\boldsymbol{X};k}||_{\mathcal{N}_{\mathcal{H}}} \\ &\leq ||\mathscr{C}_{\boldsymbol{X};0}||_{\mathcal{N}_{\mathcal{H}}} \Big(1 + 2\sum_{k=1}^{\infty} ||\alpha||^h_{\mathcal{L}_{\mathcal{H}}}\Big) \\ &= c\,\frac{1 + ||\alpha||_{\mathcal{L}_{\mathcal{H}}}}{1 - ||\alpha||_{\mathcal{L}_{\mathcal{H}}}} < \infty.\end{aligned}$$

4. Main results

This section dedicates to the estimation of lagged covariance and cross-covariance operators of $\mathcal{U}^m$- and $\mathcal{V}^n$-valued processes for $m, n \in \mathbb{N}$, and additionally of the FPCs, where $(\mathcal{U}^m, \langle\cdot,\cdot\rangle_{\mathcal{U}^m})$ and $(\mathcal{V}^n, \langle\cdot,\cdot\rangle_{\mathcal{V}^n})$ are real, separable Hilbert spaces coming from real, separable Hilbert spaces $(\mathcal{U}, \langle\cdot,\cdot\rangle_{\mathcal{U}})$ and $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$.

4.1. General assumptions

For our processes in Section 1 we throughout impose the following.

Assumption 4.1. (a) The entries in (1.1) are elements of a stationary $L^2_{\mathcal{U}}$-process $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}$. $X_1, \ldots, X_M$ is a sample of $\boldsymbol{X}$ with $M \geq m$, thus $\mathscr{X}_m, \ldots, \mathscr{X}_{\tilde{\mathscr{M}}}$ with $\tilde{\mathscr{M}} = \tilde{\mathscr{M}}_M := \lfloor \frac{M-m}{p} \rfloor + m$ is a sample of $\boldsymbol{\mathscr{X}}$, and the sample size is $\mathscr{M} = \mathscr{M}_M := \tilde{\mathscr{M}}_M - m + 1$.

(b) The entries in (1.2) are elements of a stationary $L^2_{\mathcal{V}}$-process $\boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$. $Y_1, \ldots, Y_N$ with $N \geq n$ is a sample of $\boldsymbol{Y}$, thus $\mathscr{Y}_n, \ldots, \mathscr{Y}_{\tilde{\mathscr{N}}}$ with $\tilde{\mathscr{N}} = \tilde{\mathscr{N}}_N := \lfloor \frac{N-n}{q} \rfloor + n$ is a sample of $\boldsymbol{\mathscr{Y}}$, and the sample size is $\mathscr{N} = \mathscr{N}_N := \tilde{\mathscr{N}}_N - n + 1$.

Our model also allows the numbers describing the 'degree of reuse' p, q of given variables and the Cartesian powers m, n to depend on the sample sizes.

Assumption 4.2. Let $p^*, q^* \in \mathbb{N}$. The sequences $(p_k)_{k\in\mathbb{N}}, (q_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ of the variables describing the 'degree of reuse' in Assumption 4.1 satisfy

(a) $p = p_M \to p^*$ or $p = p_M = \Xi(1, M)$ for $M \to \infty$;
(b) $q = q_N \to q^*$ or $q = q_N = \Xi(1, N)$ for $N \to \infty$.

From the Assumptions 4.1–4.2 (a) and (b) follows

$$\mathscr{M} = \mathscr{M}_M \sim p^{-1}M \qquad \text{resp.} \qquad \mathscr{N} = \mathscr{N}_N \sim q^{-1}N. \tag{4.1}$$

Assumption 4.3. Let $m^*, n^* \in \mathbb{N}$. The sequences $(m_k)_{k\in\mathbb{N}}, (n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ of Cartesian powers in Assumption 4.1 satisfy

(a) $m = m_M \to m^*$ or $m = m_M = \Xi(1, \mathscr{M}) = \Xi(1, p^{-1}M)$ for $M \to \infty$;
(b) $n = n_N \to n^*$ or $n = n_N = \Xi(1, \mathscr{N}) = \Xi(1, q^{-1}N)$ for $N \to \infty$.

The time difference where some random variable has a certain effect on another one, i.e. the lag $h \in \mathbb{Z}$, could also change over time or the sample size as follows.

Assumption 4.4. Let $h^* \in \mathbb{Z}$. The sequence of lags $(h_n)_{n\in\mathbb{N}} \subseteq \mathbb{Z}$ fulfills

(a) $h = h_M \to h^*$ or $h = h_M = \Xi(1, \mathscr{M}) = \Xi(1, p^{-1}M)$ for $M \to \infty$;
(b) $h = h_N \to h^*$ or $h = h_N = \Xi(1, \mathscr{N}) = \Xi(1, q^{-1}N)$ for $N \to \infty$.

In order to specify the asymptotic behaviour of our estimation errors exactly without demanding too much, we impose the following summability conditions for our processes $\boldsymbol{X} = (X_k)_k, \boldsymbol{Y} = (Y_k)_k$ in Assumption 4.1.

Assumption 4.5. (a) $\sum_{k\in\mathbb{Z}} ||\mathscr{C}_{\boldsymbol{X};k}||_{\mathcal{N}_{\mathcal{U}}} < \infty$;
(b) $\sum_{k\in\mathbb{Z}} ||\mathscr{C}_{\boldsymbol{Y};k}||_{\mathcal{N}_{\mathcal{V}}} < \infty$.

Assumption 4.6. (a) $(X_k)_k \subseteq L^4_{\mathcal{U}}$ and $\sum_{i,j,k\in\mathbb{Z}} |\mathbb{E}\langle X_0' \otimes X_i', X_j' \otimes X_k'\rangle_{\mathcal{S}_{\mathcal{U}}}| < \infty$;
(b) $(Y_k)_k \subseteq L^4_{\mathcal{V}}$ and $\sum_{i,j,k\in\mathbb{Z}} |\mathbb{E}\langle Y_0' \otimes Y_i', Y_j' \otimes Y_k'\rangle_{\mathcal{S}_{\mathcal{V}}}| < \infty$.

Assumption 4.7. (a) $(X_k)_k \subseteq L^4_{\mathcal{U}}$ and $\sum_{i,j\in\mathbb{Z}} |\mathbb{E}\langle (X_0 \otimes X_i)', (X_j \otimes X_{i+j})'\rangle_{\mathcal{S}_{\mathcal{U}}}| < \infty$;
(b) $(X_k)_k \subseteq L^4_{\mathcal{U}}, (Y_k)_k \subseteq L^4_{\mathcal{V}}$ and $\sum_{i,j\in\mathbb{Z}} |\mathbb{E}\langle (X_0 \otimes Y_i)', (X_j \otimes Y_{i+j})'\rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}| < \infty$.

Remarks 4.1. Assuming absolute summability of the series defining the long run covariance operator in Assumption 4.5 is needed to guarantee convergence of any rearranged series. However, for $p = 1$ postulating convergence only suffices.

4.2. Preliminaries

In various occasions, the first moments $m_{\mathscr{X}} := \mathbb{E}(\mathscr{X}_1), m_{\mathscr{Y}} := \mathbb{E}(\mathscr{Y}_1)$ of our processes (1.1)–(1.2) have to be estimated, for which we use the moment estimators

$$\hat{m}_{\mathscr{X}} := \frac{1}{\mathscr{M}_M} \sum_{i=m_M}^{\tilde{\mathscr{M}}_M} \mathscr{X}_i \quad \text{resp.} \quad \hat{m}_{\mathscr{Y}} := \frac{1}{\mathscr{N}_N} \sum_{i=n_N}^{\tilde{\mathscr{N}}_N} \mathscr{Y}_i. \tag{4.2}$$

Lemma 4.1. *Let Assumptions 4.1–4.3 hold. Moreover, let $m = m_M, p = p_M$. Then, $\hat{m}_{\mathscr{X}}$ and $\hat{m}_{\mathscr{Y}}$ in (4.2) are unbiased estimators for $m_{\mathscr{X}}$ for all $M \in \mathbb{N}$ resp. for $m_{\mathscr{Y}}$ for all $N \in \mathbb{N}$.*

(a) If also Assumption 4.5 holds,

$$\frac{\mathscr{M}_M}{m_M}\,\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^2_{\mathcal{U}^m} \overset{M\to\infty}{\longrightarrow} \eta_{2,\boldsymbol{X}}, \tag{4.3}$$

$$\frac{\mathscr{N}_N}{n_N}\,\mathbb{E}||\hat{m}_{\mathscr{Y}} - m_{\mathscr{Y}}||^2_{\mathcal{V}^n} \overset{N\to\infty}{\longrightarrow} \eta_{2,\boldsymbol{Y}}, \tag{4.4}$$

where the constants $\eta_{2,\boldsymbol{Z}} = \eta_{2,(\mathcal{W},\boldsymbol{Z},(r_S)_S,r^*)}$, with $(\mathcal{W},\boldsymbol{Z},(r_S)_S,r^*) \in \{(\mathcal{U},\boldsymbol{X},(p_M)_M,p^*),\ (\mathcal{V},\boldsymbol{Y},(q_N)_N,q^*)\}$, are defined by

$$\eta_{2,\boldsymbol{Z}} \coloneqq \begin{cases} ||\sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{Z};kr^*}||_{\mathcal{N}_{\mathcal{W}}}, & \text{if } r = r_S \to r^* \in \mathbb{N}, \\ \mathbb{E}||Z_0||^2_{\mathcal{W}}, & \text{if } r = r_S = \Xi(1,S). \end{cases} \tag{4.5}$$

(b) If also Assumption 4.6 holds,

$$\frac{\mathscr{M}_M^3}{m_M^{1+\mathbf{1}_P}}\,\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^4_{\mathcal{U}^m} \overset{M\to\infty}{\longrightarrow} \eta_{4,\boldsymbol{X}}, \tag{4.6}$$

$$\frac{\mathscr{N}_N^3}{n_N^{1+\mathbf{1}_Q}}\,\mathbb{E}||\hat{m}_{\mathscr{Y}} - m_{\mathscr{Y}}||^4_{\mathcal{V}^n} \overset{N\to\infty}{\longrightarrow} \eta_{4,\boldsymbol{Y}}, \tag{4.7}$$

with $P \coloneqq \{(p_k)_k | p_k \to p^*\}$ and $Q \coloneqq \{(q_k)_k | q_k \to q^*\}$, where the constants $\eta_{4,\boldsymbol{Z}} = \eta_{4,(\mathcal{W},\boldsymbol{Z},(o_S)_S,o^*,(r_S)_S,r^*)}$, with $(\mathcal{W},\boldsymbol{Z},(o_S)_S,o^*,(r_S)_S,r^*) \in \{(\mathcal{U},\boldsymbol{X},(m_M)_M,m^*,(p_M)_M,p^*),\ (\mathcal{V},\boldsymbol{Y},(n_N)_N,n^*,(q_N)_N,q^*)\}$ are with $c_i \in [0,\frac{2}{r^*}]$ for all i defined through

$$\eta_{4,\boldsymbol{Z}} \coloneqq \begin{cases} \frac{1}{o^*}\sum_{|l|<o^*}\sum_{i,j,k\in\mathbb{Z}} \xi_{l,o^*}\,\mathbb{E}\langle Z_0'\otimes Z_{l+ir^*}', Z_{jr^*}'\otimes Z_{l+kr^*}'\rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } o_S \to o^*,\, r_S \to r^*, \\ \sum_{i,j,k\in\mathbb{Z}} c_i\,\mathbb{E}\langle Z_0'\otimes Z_{ip^*}', Z_{jp^*}'\otimes Z_{kp^*}'\rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } o_S \to \infty,\, r_S \to r^*, \\ \sum_{|l|<o^*} \xi_{l,o^*}\,\mathbb{E}||Z_0'||^2_{\mathcal{U}}||Z_l'||^2_{\mathcal{U}}, & \text{if } o_S \to o^*,\, r_S \to \infty, \\ \sum_{k\in\mathbb{Z}} \mathbb{E}||Z_0'||^2_{\mathcal{U}}||Z_k'||^2_{\mathcal{U}}, & \text{if } o_S \to \infty,\, r_S \to \infty. \end{cases} \tag{4.8}$$

4.3. Estimation of lag-h-covariance operators

When estimating lag-h-covariance operators, we distinguish, as for real-valued processes, between centered processes and those with an unknown, finite first moment. If $\boldsymbol{X} = (X_k)_k$ in Assumption 4.1 (a) is centered, hence also $\mathscr{X} = (\mathscr{X}_k)_k$, we estimate $\mathscr{C}_{\mathscr{X};h}$ with $|h| < \mathscr{M}_M$ by the moment estimator

$$\hat{\mathscr{C}}_{\mathscr{X};h} \coloneqq \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{k=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, & \text{if } h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{k=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, & \text{if } h \geq 0, \end{cases} \tag{4.9}$$

with $\mathscr{M}_{M,h} \coloneqq \mathscr{M}_M - |h|, \tilde{\mathscr{M}}_{M,h} \coloneqq \tilde{\mathscr{M}}_M - |h|$. These estimates satisfy $\hat{\mathscr{C}}_{\mathscr{X};h} \in \mathcal{F}_{\mathcal{U}^m}$ (i.e., they are finite-rank operators) with $\hat{\mathscr{C}}_{\mathscr{X};h} = \hat{\mathscr{C}}_{\mathscr{X};-h}$, and $\hat{\mathscr{C}}_{\mathscr{X}} \coloneqq \hat{\mathscr{C}}_{\mathscr{X};0}$ is self-adjoint and positive semi-definite. The associated estimation errors fulfill

$$||\hat{\mathscr{C}}_{\mathscr{X};h} - \mathscr{C}_{\mathscr{X};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \sum_{i,j=1}^{m} ||\hat{\mathscr{C}}_{\mathbf{X};j-i+hp} - \mathscr{C}_{\mathbf{X};j-i+hp}||^2_{\mathcal{S}_{\mathcal{U}}}, \tag{4.10}$$

where $\hat{\mathscr{C}}_{\mathbf{X};j-i+hp}$ corresponds to $\hat{\mathscr{C}}_{\mathscr{X};h}$ in (4.9), with $\mathscr{X}_k$ and $\mathscr{X}_{k+h}$ replaced by $X_{i+(k-m)p}$ resp. $X_{j+(k+h-m)p}$ for all i, j, k. Using this identity leads to the following asymptotic result of the estimation errors.

Theorem 4.1. *Let Assumptions 4.1–4.4, 4.7 (a) hold, and let $\boldsymbol{X}$ be centered. Further, let $h = h_M, m = m_M, p = p_M$ and $\mathscr{M}_{M,h} = \mathscr{M}_M - |h|$. Then, $\hat{\mathscr{C}}_{\mathscr{X};h}$ is an unbiased estimator for $\mathscr{C}_{\mathscr{X};h}$ with $|h| < \mathscr{M}_M$, and*

$$\frac{\mathscr{M}_{M,h}}{m_M} \mathbb{E}||\hat{\mathscr{C}}_{\mathscr{X};h} - \mathscr{C}_{\mathscr{X};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \overset{M\to\infty}{\longrightarrow} \tau_{2,\boldsymbol{X}}, \tag{4.11}$$

where the constant $\tau_{2,\boldsymbol{X}} = \tau_{2,(\mathcal{U},\boldsymbol{X},(h_M)_M,h^*,(m_M)_M,m^*,(p_M)_M,p^*)}$ is defined as

$$\tau_{2,\boldsymbol{X}} \coloneqq \begin{cases} 0, & \text{if } hp \to \pm\infty, \\ \sum_{i\in\mathbb{Z}} \sum_{|j|<m^*} \xi_{j,m^*} \mathbb{E}\langle (X_0 \otimes X_{i+c})', (X_j \otimes X_{i+j+c})' \rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } hp \to c \in \mathbb{Z}, m \to m^*, \\ \sum_{i,j\in\mathbb{Z}} \mathbb{E}\langle (X_0 \otimes X_i)', (X_j \otimes X_{i+j})' \rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } hp \to c \in \mathbb{Z}, m \to \infty. \end{cases} \tag{4.12}$$

Remarks 4.2. We stated the limit in Theorem 4.1 for $hp \to \pm\infty$ without further calibration for the sake of simplicity. Nevertheless, it would also have been appropriate to calibrate the estimation error with the reciprocal of the identity (A.13), provided it is not equal zero.

Now, we consider that the first moment $m_{\boldsymbol{X}}$ of $\boldsymbol{X} = (X_k)_k$ is unknown, therefore also $m_{\mathscr{X}} = (m_{\boldsymbol{X}}, \dots, m_{\boldsymbol{X}})^T \in \mathcal{U}^m$. Then, if $|h| < \mathscr{M}_M - 1$, we use

$$\hat{\mathscr{C}}'_{\mathscr{X};h} \coloneqq \begin{cases} \frac{1}{\mathscr{M}_{M,h}-1} \sum_{k=m+|h|}^{\tilde{\mathscr{M}}_M} (\mathscr{X}_k - \hat{m}_{\mathscr{X}}) \otimes (\mathscr{X}_{k+h} - \hat{m}'_{\mathscr{X}}), & \text{if } h < 0, \\ \frac{1}{\mathscr{M}_{M,h}-1} \sum_{k=m}^{\tilde{\mathscr{M}}_{M,h}} (\mathscr{X}_k - \hat{m}_{\mathscr{X}}) \otimes (\mathscr{X}_{k+h} - \hat{m}'_{\mathscr{X}}), & \text{if } h \geq 0, \end{cases} \tag{4.13}$$

to estimate $\mathscr{C}_{\mathscr{X};h}$ where the moment estimators are defined by

$$\hat{m}_{\mathscr{X}} \coloneqq \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{i=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_i, & \text{if } h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{i=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_i, & \text{if } h \geq 0, \end{cases}$$

$$\hat{m}'_{\mathscr{X}} \coloneqq \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{j=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_{j+h}, & \text{if } h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{j=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_{j+h}, & \text{if } h \geq 0. \end{cases}$$

As for the estimators given a centered process, holds $\hat{\mathscr{C}}'_{\mathscr{X};h} \in \mathcal{F}_{\mathcal{U}^m}$ with $\hat{\mathscr{C}}'^{*}_{\mathscr{X};h} = \hat{\mathscr{C}}'_{\mathscr{X};-h}$, and $\hat{\mathscr{C}}'_{\mathscr{X}} \coloneqq \hat{\mathscr{C}}'_{\mathscr{X};0}$ is self-adjoint and positive semi-definite.

Theorem 4.2. *Let Assumptions 4.1–4.4, 4.6–4.7 (a) hold. Also,* $h = h_M, m = m_M, p = p_M$ *and* $\mathscr{M}_{M,h} = \mathscr{M}_M - |h|$*. Then,* $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}$ *is an unbiased estimator for* $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ *with* $|h| < \mathscr{M}_M - 1$ *if* $\sum_{j,k=1,k\neq j}^{\mathscr{M}_{M,h}} \mathscr{C}_{\boldsymbol{\mathscr{X}};j+h-k} = 0_{\mathcal{L}_{\mathcal{U}^m}}$*, and*

$$\limsup_{M\to\infty} \frac{\mathscr{M}_{M,h}}{m_M} \mathbb{E}||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \leq 3\tau_{2,\boldsymbol{X}}, \tag{4.14}$$

with $\tau_{2,\boldsymbol{X}} = \tau_{2,(\mathcal{U},\boldsymbol{X},(h_M)_M,h^*,(m_M)_M,m^*,(p_M)_M,p^*)}$ *from* (4.12).

Remarks 4.3. (a) Theorems 4.1–4.2 extend the existing literature, e.g. [1, 7, 23, 25, 28, 27, 35, 38, 39], in several ways. This is because the assertions are derived in Cartesian products of general, separable Hilbert spaces under mild conditions, and the processes can have arbitrary finite, first moments. Further, the Cartesian power m, the variable p describing the 'degree of reuse' and simultaneously the lag h might grow w.r.t. the sample size M, and the upper bounds are specified so accurately that they reflect the dependence of the asymptotic behaviour of these sequences.

(b) The statements of Theorems 4.1–4.2 can obviously also be formulated for the process $\boldsymbol{Y} = (Y_k)_{k\in\mathbb{Z}}$ by assuming that the parts (b) instead of (a) of Assumptions 4.1–4.4, 4.6, and Assumption 4.7 (a) formulated for $\boldsymbol{Y}$ hold.

(c) That Theorem 4.2 states an inequality with a slightly modified value instead of a precise limit as in Theorem 4.1, is, because we used the Δ- and the Cauchy-Schwarz inequality.

4.4. Estimation of lag-h-cross-covariance operators

Herein, we transfer the estimation procedure for lag-h-covariance to lag-h-cross-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$. If $\boldsymbol{X} = (X_k)_k$ and $\boldsymbol{Y} = (Y_k)_k$ in Assumption 4.1 are centered and subsequently also $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} = (\mathscr{Y}_k)_k$, we estimate $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ by

$$\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{k=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{X}_k \otimes \mathscr{Y}_{k+h}, \tag{4.15}$$

with $\tilde{l}_{m,n,h} := m \vee (n-h), \tilde{\mathscr{L}}_{M,N,h} := \tilde{\mathscr{M}}_M \wedge (\tilde{\mathscr{N}}_N - h)$ and $\mathscr{L}_{M,N,h} := \tilde{\mathscr{L}}_{M,N,h} + 1 - \tilde{l}_{m,n,h}$, where $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} \in \mathcal{F}_{\mathcal{U}^m,\mathcal{V}^n}$ and $\hat{\mathscr{C}}^*_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \hat{\mathscr{C}}_{\boldsymbol{\mathscr{Y}},\boldsymbol{\mathscr{X}};-h}$. In order to derive the asymptotic behaviour of the estimation errors for $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ explicitly, we impose that one sample size asymptotically depends on the other one. This also allows to simplify our conversions.

Assumption 4.8. The sample sizes M and N of $X_1, \dots, X_M$ resp. $Y_1, \dots, Y_N$ in Assumption 4.1 fulfill $N = N_M \sim cM^\delta$ for $M \to \infty$ for some $c \neq 0$ and $\delta > 0$.

Assumption 4.9. The sequences in Section 4.1 fulfill $\tilde{l}_{m,n,h} = \mathrm{o}(\tilde{\mathscr{L}}_{M,N,h})$ for $M \to \infty$, provided that Assumption 4.8 is satisfied.

Assumption 4.10. The sequences $(p_k)_k, (q_k)_k$ in Assumption 4.2 satisfy $p_M \sim q_N$ for $M \to \infty$, provided that Assumption 4.8 holds.

An equation like (4.10) does not apply to all combinations of p and q, but under Assumption 4.10 holds for $M \to \infty$ (and thus $N \to \infty$),

$$\begin{aligned} &\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} \\ &\quad \sim \sum_{i=1}^{m}\sum_{j=1}^{n} \mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{X},\boldsymbol{Y};j-i+(m+h-n)p} - \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};j-i+(m+h-n)p}||^2_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, \end{aligned} \tag{4.16}$$

where even equality in (4.16) is given if p and q are coinciding constants.

Theorem 4.3. *Let Assumptions 4.1–4.4, and Assumptions 4.7 (b) and 4.8–4.10 hold, and let $\boldsymbol{X}, \boldsymbol{Y}$ be centered. Further, let $m = m_M, p = p_M, n = n_N, q = q_N$, and $h = h_L$ with $L = L_{M,N} := M \wedge N$. Then, $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$, which satisfies*

$$\frac{\mathscr{L}_{M,N,h}}{m \wedge n}\, \mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} \overset{M\to\infty}{\longrightarrow} \tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})}, \tag{4.17}$$

where $\tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})} = \tilde{\tau}_{2,((\mathcal{U},\mathcal{V}),(\boldsymbol{X},\boldsymbol{Y}),(h_L)_L,h^*,(m_M)_M,m^*,(p_M)_M,p^*,(n_N)_N,n^*,(q_N)_N,q^*)}$ stands for a constant defined by

$$\tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})} := \begin{cases} 0, & \text{if } (m+h-n)p \to \pm\infty, \\ \sum_{i\in\mathbb{Z}}\sum_{|j|<m^*\vee n^*} \tilde{\imath}^*(j)\,\mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})'\rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to m^*, n \to n^*, \\ \sum_{i,j\in\mathbb{Z}} \tilde{\imath}^*(j)\,\mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})'\rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to \infty, n \to n^*, \\ \sum_{i,j\in\mathbb{Z}} \tilde{\imath}^*(j)\,\mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})'\rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to m^*, n \to \infty, \\ \sum_{i,j\in\mathbb{Z}} \mathbb{E}\langle (X_0 \otimes Y_{i})', (X_j \otimes Y_{i+j})'\rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to \infty, n \to \infty, \end{cases} \tag{4.18}$$

with $\tilde{\imath}^*(j)$ defined in (A.5) for all j, satisfying $|\tilde{\imath}^*(j)| \leq 1$ for all j.

If $m_{\boldsymbol{X}}$ and/or $m_{\boldsymbol{Y}}$ in Assumption 4.1 are unknown, and consequently also $m_{\boldsymbol{\mathscr{X}}} = (m_{\boldsymbol{X}}, \ldots, m_{\boldsymbol{X}})^T \in \mathcal{U}^m$ and/or $m_{\boldsymbol{\mathscr{Y}}} = (m_{\boldsymbol{Y}}, \ldots, m_{\boldsymbol{Y}})^T \in \mathcal{V}^n$, $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ is estimated by

$$\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} := \frac{1}{\mathscr{L}_{M,N,h} - 1} \sum_{k=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} (\mathscr{X}_k - \hat{m}_{\boldsymbol{\mathscr{X}}}) \otimes (\mathscr{Y}_{k+h} - \hat{m}'_{\boldsymbol{\mathscr{Y}}}) \tag{4.19}$$

if $\tilde{\mathscr{L}}_{M,N,h} > \tilde{l}_{m,n,h}$, with moment estimators

$$\hat{m}_{\boldsymbol{\mathscr{X}}} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{i=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{X}_i, \quad \hat{m}'_{\boldsymbol{\mathscr{Y}}} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{j=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{Y}_{j+h}. \tag{4.20}$$

Thereby, $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} \in \mathcal{F}_{\mathcal{U}^m,\mathcal{V}^n}$ and $\hat{\mathscr{C}}'^{*}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \hat{\mathscr{C}}'_{\boldsymbol{\mathscr{Y}},\boldsymbol{\mathscr{X}};-h}$ for all h. To deduce the asymptotic behaviour of the lag-h-cross-covariance operators in the case of unknown first moments, we also require that $m \vee n = \mathrm{o}(\mathscr{L}_{M,N,h})$ for $M \to \infty$, assuming that Assumption 4.8 holds. Under Assumptions 4.4 and 4.9, this is given if the following assumption is also fulfilled.

Assumption 4.11. The sequences $(m_k)_k, (h_k)_k$ in Assumptions 4.3–4.4 fulfill $h_L = \mathrm{o}(n_N)$ for $M \to \infty$, with $L = L_{M,N} := M \wedge N$, provided that Assumption 4.8 is satisfied.

Theorem 4.4. *Let Assumptions 4.1–4.4, 4.6, and Assumptions 4.7 (b) and 4.8–4.11 hold. Further, let $m = m_M, n = n_N$ and $h = h_L$ with $L = L_{M,N} := M \wedge N$. Then, $\hat{\mathscr{C}}_{\mathscr{X},\mathscr{Y};h}$ is an unbiased estimator for $\mathscr{C}_{\mathscr{X},\mathscr{Y};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ if $\sum_{1 \leq i,k \leq \mathscr{L}_{M,N,h}, i \neq k} \mathscr{C}_{\mathscr{X},\mathscr{Y};k+h-i} = 0_{\mathcal{L}_{\mathcal{U}^m,\mathcal{V}^n}}$, which satisfies*

$$\limsup_{M \to \infty} \frac{\mathscr{L}_{M,N,h}}{m \wedge n} \mathbb{E}||\hat{\mathscr{C}}'_{\mathscr{X},\mathscr{Y};h} - \mathscr{C}_{\mathscr{X},\mathscr{Y};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} \leq 3\tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})}, \tag{4.21}$$

with $\tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})} = \tilde{\tau}_{2,((\mathcal{U},\mathcal{V}),(\boldsymbol{X},\boldsymbol{Y}),(h_L)_L,h^*,(m_M)_M,m^*,(p_M)_M,p^*,(n_N)_N,n^*,(q_N)_N,q^*)}$ in (4.18).

Remarks 4.4. (a) Although estimating (lagged) cross-covariance operators is widely discussed, see e.g. [3, 7, 25, 46], Theorems 4.3–4.4 are new in many ways. First, both processes can attain values in arbitrary separable Hilbert spaces which do not necessarily need to match, nor do the drawn sample sizes M, N. Further, the upper bounds are, as in Theorems 4.1–4.2 for the lagged covariance operators, derived for centered and for not necessarily centered processes, the lag h is allowed to be both fixed and to vary w.r.t. the sample sizes, as are the Cartesian powers m, n.

(b) The initial and final value of the sums in (4.15) and (4.19) guarantee that $\mathscr{X}_k$ and $\mathscr{Y}_{k+h}$ are simultaneously well-defined.

(c) Assumption 4.9 ensures that the number of the summands of our estimators grow when the sample sizes do.

(d) By following the lines in the proof of Theorem 4.4, it becomes clear that omitting to estimate $\mathscr{X}_k$ resp. $\mathscr{Y}_{k+h}$ in (4.20) if $\boldsymbol{X}$ is centered and $m_{\boldsymbol{Y}}$ unknown, resp. if $m_{\boldsymbol{X}}$ is unknown and $\boldsymbol{Y}$ centered, has no positive effect on the convergence rate (4.21) in Theorem 4.4.

4.5. Estimation of FPCs

Herein, we examine the estimation procedure of the FPCs of $\mathscr{C}_{\mathscr{X}} = \mathscr{C}_{\mathscr{X};0}$ of the $\mathcal{U}^m$-valued processes $\mathscr{X} = (\mathscr{X}_k)_{k \in \mathbb{Z}}$ in Assumption 4.1 (a). Throughout, $(\mathfrak{c}_j)_{j \in \mathbb{N}}, (\hat{\mathfrak{c}}_j)_{j \in \mathbb{N}}$ resp. $(\hat{\mathfrak{c}}'_j)_{j \in \mathbb{N}}$ are the eigenfunction and $(c_j)_{j \in \mathbb{N}}, (\hat{c}_j)_{j \in \mathbb{N}}$ resp. $(\hat{c}'_j)_{j \in \mathbb{N}}$ the associated w.l.o.g. monotonically decreasing eigenvalue sequences of $\mathscr{C}_{\mathscr{X}}, \hat{\mathscr{C}}_{\mathscr{X}} = \hat{\mathscr{C}}_{\mathscr{X};0}$ in (4.9) resp. $\hat{\mathscr{C}}'_{\mathscr{X}} = \hat{\mathscr{C}}'_{\mathscr{X};0}$ in (4.13), where we occasionally write $c_j = c_{j,m}$ and $\mathfrak{c}_j = \mathfrak{c}_{j,m}$ since the Cartesian power m can vary w.r.t. M. We would like to emphatically point out that hereafter, whenever dealing with the FPCs $\hat{c}_j, \hat{\mathfrak{c}}_j$ and $\hat{c}'_j, \hat{\mathfrak{c}}'_j$, we implicitly assume centeredness resp. an unknown first moment $m_{\boldsymbol{X}}$.

At first, according to [7], Lemma 4.2,

$$\sup_{j \in \mathbb{N}} |\hat{c}_j - c_j| \leq ||\hat{\mathscr{C}}_{\mathscr{X}} - \mathscr{C}_{\mathscr{X}}||_{\mathcal{L}_{\mathcal{U}^m}}, \qquad \sup_{j \in \mathbb{N}} |\hat{c}'_j - c_j| \leq ||\hat{\mathscr{C}}'_{\mathscr{X}} - \mathscr{C}_{\mathscr{X}}||_{\mathcal{L}_{\mathcal{U}^m}}. \tag{4.22}$$

Corollary 4.1. *Let the assumptions of Theorem 4.1 hold. Then,*

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\,\mathbb{E}\sup_{j\in\mathbb{N}}(\hat{c}_j - c_j)^2 \;\le\; \tau_{2,\boldsymbol{X}},$$

with $\tau_{2,\boldsymbol{X}} = \tau_{2,(\mathcal{U},\boldsymbol{X},(h_M)_M,h^*,(m_M)_M,m^*,(p_M)_M,p^*)}$ *in* (4.12) *for* $h = 0$*. If additionally holds Assumption 4.6 (a) (i.e. all assumptions of Theorem 4.2 are satisified),*

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\,\mathbb{E}\sup_{j\in\mathbb{N}}(\hat{c}'_j - c_j)^2 \;\le\; 3\tau_{2,\boldsymbol{X}}.$$

Eigenfunctions are unambiguously determined except for their sign, why

$$\check{\mathfrak{c}}_j := \operatorname{sgn}\langle\hat{\mathfrak{c}}_j,\mathfrak{c}_j\rangle_{\mathcal{U}^m}\hat{\mathfrak{c}}_j \qquad \text{resp.} \qquad \check{\mathfrak{c}}'_j := \operatorname{sgn}\langle\hat{\mathfrak{c}}'_j,\mathfrak{c}_j\rangle_{\mathcal{U}^m}\hat{\mathfrak{c}}'_j \tag{4.23}$$

seem to be reasonable estimators for $\mathfrak{c}_j$, where 'sgn' is the sign function. To state upper bounds of estimation errors when using the estimators in (4.33), the following technical preliminaries are needed. Due to [7], Lemma 4.3,

$$||\check{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m} \le \tilde{\gamma}_j||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}}} - \mathscr{C}_{\boldsymbol{\mathscr{X}}}||_{\mathcal{L}_{\mathcal{U}^m}}, \quad ||\check{\mathfrak{c}}'_j - \mathfrak{c}_j||_{\mathcal{U}^m} \le \tilde{\gamma}_j||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}}} - \mathscr{C}_{\boldsymbol{\mathscr{X}}}||_{\mathcal{L}_{\mathcal{U}^m}} \tag{4.24}$$

for any $j\in\mathbb{N}$ if the eigenspace of c_j is one-dimensional, where $\tilde{\gamma}_1 := 2\sqrt{2}\gamma_1$, $\tilde{\gamma}_j := 2\sqrt{2}(\gamma_{j-1}\vee\gamma_j)$ for $j>1$, and

$$\gamma_j := (c_j - c_{j+1})^{-1}, \quad j\in\mathbb{N}. \tag{4.25}$$

Assumption 4.12. $\mathscr{C}_{\boldsymbol{\mathscr{X}}}$ is injective, and the eigenvalues of $\mathscr{C}_{\boldsymbol{\mathscr{X}}}$ satisfy $c_j \ne c_{j+1}$ and $\kappa(j) = c_j$ for all $j\in\mathbb{N}$ where $\kappa\colon\mathbb{R}\to\mathbb{R}$ is a convex function.

Under Assumption 4.12 holds both

$$c_1 > c_2 > \cdots > 0, \tag{4.26}$$

and for any sequence $(k_j)_j \subseteq \mathbb{N}$ with $k = k_M = \Omega(1)$,

$$\sup_{j\le k}\tilde{\gamma}_j = \tilde{\gamma}_k. \tag{4.27}$$

Corollary 4.2. *Let* $(k_j)_j \subseteq \mathbb{N}$ *be a sequence with* $k_M = \Omega(1)$ *for* $M\to\infty$*, and let* $m = m_M$*,* $\tilde{\gamma}_{1,m} = 2\sqrt{2}\gamma_{1,m}$ *and* $\tilde{\gamma}_{j,m} = 2\sqrt{2}(\gamma_{j-1,m}\vee\gamma_{j,m})$ *for* $j>1$*, where* $\gamma_{j,m} = (c_{j,m} - c_{j+1,m})^{-1}$ *for* $j\in\mathbb{N}$*. Then, under the assumptions of Theorem 4.1 and Assumption 4.12 holds for* $M\to\infty$*:*

$$\limsup_{M\to\infty}\frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{j,m_M}^{-2}\mathbb{E}||\check{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} \le \tau_{2,\boldsymbol{X}}, \quad j\in\mathbb{N}, \tag{4.28}$$

$$\limsup_{M\to\infty}\frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{k_M,m_M}^{-2}\mathbb{E}\sup_{j\le k_M}||\check{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} \le \tau_{2,\boldsymbol{X}}, \tag{4.29}$$

with $\tau_{2,\boldsymbol{X}} = \tau_{2,(\mathcal{U},\boldsymbol{X},(h_M)_M,h^*,(m_M)_M,m^*,(p_M)_M,p^*)}$ *in* (4.12) *for* $h=0$*. If also holds Assumption 4.6 (a) (thus all assumptions of Theorem 4.2 are satisfied), we have*

$$\limsup_{M\to\infty}\frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{j,m_M}^{-2}\mathbb{E}||\check{\mathfrak{c}}'_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} \le 3\tau_{2,\boldsymbol{X}}, \quad j\in\mathbb{N}, \tag{4.30}$$

$$\limsup_{M\to\infty}\frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{k_M,m_M}^{-2}\mathbb{E}\sup_{j\le k_M}||\check{\mathfrak{c}}'_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} \le 3\tau_{2,\boldsymbol{X}}. \tag{4.31}$$

However, although we were able to derive asymptotic results for the estimation errors between the estimates $\breve{\mathfrak{c}}_j$ and $\mathfrak{c}_j$ as well as between $\breve{\mathfrak{c}}'_j$ in (4.23) and $\mathfrak{c}_j$, using these estimators can be problematic. This is because $\breve{\mathfrak{c}}_j \not\to \mathfrak{c}_j$ a.s. and $\breve{\mathfrak{c}}'_j \not\to \mathfrak{c}_j$ a.s., thus $\operatorname{sgn}\langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. and $\operatorname{sgn}\langle \hat{\mathfrak{c}}'_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. is not guaranteed for fixed j, M, making allocating an a.s. unique estimator for $\mathfrak{c}_j$ impossible. Having an a.s. unique estimator for the eigenfunction for a fixed sample size was inevitable in conversions leading to asymptotic upper bounds of the estimation errors for operators of $L^2[0,1]$-valued (G)ARCH and linear, invertible processes in [38, 39], and to have such an estimator also benefits the simulation, since then a case decision is obsolete.

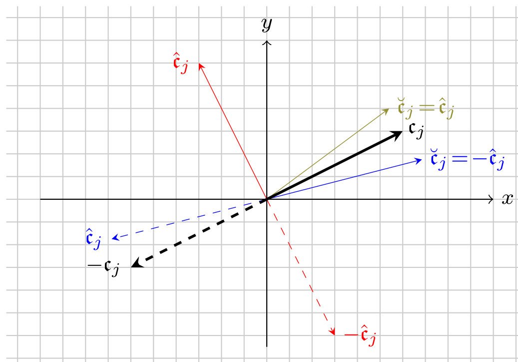

Fig 5. *Example estimation of an eigenfunction (resp. eigenvector)* $\mathfrak{c}_j$ *in* $\mathbb{R}^2$. $\hat{\mathfrak{c}}_j$, $\hat{\mathfrak{c}}_j$ *and* $\hat{\mathfrak{c}}_j$ *denote vectors on which the estimate for* $\mathfrak{c}_j$ *in three different circumstances is based. Based on* $\hat{\mathfrak{c}}_j$ *and* $\hat{\mathfrak{c}}_j$, *appropriate estimators for* $\mathfrak{c}_j$ *are unambiguously determinable where only the sign of* $\hat{\mathfrak{c}}_j$ *had to be reversed. Further,* $\hat{\mathfrak{c}}_j$ *is orthogonal to* $\mathfrak{c}_j$, *consequently from* $\hat{\mathfrak{c}}_j$ *no unambiguously determined estimator for* $\mathfrak{c}_j$ *can be derived why both* $-\hat{\mathfrak{c}}_j$ *and* $\hat{\mathfrak{c}}_j$ *could be used to estimate* $\mathfrak{c}_j$.

We bypass this problem by adding suitable perturbations to $\breve{\mathfrak{c}}_j$ and $\breve{\mathfrak{c}}'_j$. Hereto, let $(u_i)_{i\in\mathbb{N}}$ be a CONS of $\mathcal{U}^m$, and let $(\zeta_i)_{i\in\mathbb{N}}$ be a sequence of non-degenerated random variables which are independent of the observations of $\boldsymbol{X}$, centered, absolutely continuous, and integrable with uniformly bounded absolute first moments, so there exists some $\mu \in (0,\infty)$ with

$$\sup_{i\in\mathbb{N}} \mathbb{E}|\zeta_i| \le \mu. \tag{4.32}$$

These properties are satisfied, for instance, for a sequence of i.i.d., standard Gaussian random variables. Given these properties, the estimators

$$\breve{\mathfrak{c}}^{\dagger}_j := \hat{\mathfrak{c}}_j + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}^2_{j,m_M} \sum_{i=1}^{\infty} \frac{\zeta_i u_i}{2^i} \quad \text{and} \quad \breve{\mathfrak{c}}'^{\dagger}_j := \hat{\mathfrak{c}}'_j + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}^2_{j,m_M} \sum_{i=1}^{\infty} \frac{\zeta_i u_i}{2^i}, \tag{4.33}$$

are well-defined for all j according to the monotone convergence theorem. Thereby, the factors $\frac{m_M}{\mathscr{M}_M}\tilde{\gamma}^2_{j,m_M}$, which are the reciprocals of the factors in (4.28) and (4.30) in Corollary 4.2, were established to derive asymptotic results. For any $j, \breve{\mathfrak{c}}_j^\dagger$ and $\breve{\mathfrak{c}}_j^{\prime\dagger}$ are due to their definition unbiased estimators for $\mathfrak{c}_j$ as well as absolutely continuous, because the convolution of a random variable with an absolutely continuous random variable is also absolutely continuous. Hence, $\breve{\mathfrak{c}}_j^\dagger \not\perp \mathfrak{c}_j$ a.s. and $\breve{\mathfrak{c}}_j^{\prime\dagger} \not\perp \mathfrak{c}_j$ a.s., and thus $\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^\dagger, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. and $\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. Utilizing these features leads canonically to the estimators

$$\breve{\mathfrak{c}}_j^\ddagger := \Big[\mathbf{1}_{\mathbb{R}\setminus\{0\}}(\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^\dagger, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \operatorname{sgn}\langle \breve{\mathfrak{c}}_j^\dagger, \mathfrak{c}_j \rangle_{\mathcal{U}^m} + \mathbf{1}_{\{0\}}(\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^\dagger, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \Big] \hat{\mathfrak{c}}_j, \tag{4.34}$$

$$\breve{\mathfrak{c}}_j^{\prime\ddagger} := \Big[\mathbf{1}_{\mathbb{R}\setminus\{0\}}(\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} + \mathbf{1}_{\{0\}}(\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \Big] \hat{\mathfrak{c}}_j', \tag{4.35}$$

where $\mathbf{1}_A(\cdot)$ is the indicator function of a set A, which also satisfy for all j, M,

$$\breve{\mathfrak{c}}_j^\ddagger = \operatorname{sgn}\langle \breve{\mathfrak{c}}_j^\dagger, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}_j \text{ a.s.} \quad \text{resp.} \quad \breve{\mathfrak{c}}_j^{\prime\ddagger} = \operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}_j' \text{ a.s.} \tag{4.36}$$

Theorem 4.5. *Here, we use the notation and variables in Corollary 4.2. Also, let $(k_j)_j \subseteq \mathbb{N}$ be a sequence with $k_M = \Omega(1)$ for $M \to \infty$. Then, under the assumptions of Theorem 4.1 and Assumption 4.12, with μ from (4.32), holds*

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M} \tilde{\gamma}_{j,m_M}^{-2} \mathbb{E}||\breve{\mathfrak{c}}_j^\ddagger - \mathfrak{c}_j||_{\mathcal{U}^m}^2 \leq 10\,\tau_{2,\boldsymbol{X}} + 8\mu, \quad j \in \mathbb{N}, \tag{4.37}$$

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M} \tilde{\gamma}_{k_M,m_M}^{-2} \mathbb{E} \sup_{j\leq k_M} ||\breve{\mathfrak{c}}_j^\ddagger - \mathfrak{c}_j||_{\mathcal{U}^m}^2 \leq 10\,\tau_{2,\boldsymbol{X}} + 8\mu, \tag{4.38}$$

and if also Assumption 4.6 (a) holds,

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M} \tilde{\gamma}_{j,m_M}^{-2} \mathbb{E}||\breve{\mathfrak{c}}_j^{\prime\ddagger} - \mathfrak{c}_j||_{\mathcal{U}^m}^2 \leq 30\,\tau_{2,\boldsymbol{X}} + 8\mu, \quad j \in \mathbb{N}, \tag{4.39}$$

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M} \tilde{\gamma}_{k_M,m_M}^{-2} \mathbb{E} \sup_{j\leq k_M} ||\breve{\mathfrak{c}}_j^{\prime\ddagger} - \mathfrak{c}_j||_{\mathcal{U}^m}^2 \leq 30\,\tau_{2,\boldsymbol{X}} + 8\mu. \tag{4.40}$$

Remarks 4.5. (a) Corollary 4.2 and in particular Theorem 4.5 can be seen as generalizations of results in [7, 27, 35, 37] where the estimation of eigenfunctions of centered $L^2[0,1]$-valued processes without considering Cartesian products was discussed. This is due to the fact that our results state explicit upper bounds for the estimation errors in sense of second mean for the eigenfunctions of the covariance operators of not necessarily centered processes in Cartesian products of general, separable Hilbert spaces, where both the 'degree of reuse' and the Cartesian power might depend on the sample size. Furthermore, Theorem 4.5 guarantees that the estimators stated there are a.s. unique for all eigenfunctions and any sample size.

(b) If $m = m_M$ is bounded, so are the sequences of reciprocal spectral gaps $(\gamma_{j,m})_M$ for all j, and $(\gamma_{k,m})_M$ is guaranteed to be bounded if $k = k_M$ and $m = m_M$ are. In these cases, even more general if for $M \to \infty$ holds $\gamma_{j,m} = \omega(\sqrt{\frac{m_M}{\mathscr{M}_M}})$ for all j, resp. $\gamma_{k,m} = \omega(\sqrt{\frac{m_M}{\mathscr{M}_M}})$, the estimation errors stated in Corollary 4.2 and Theorem 4.5 are null sequences, thus the estimators in these results for the eigenfunctions are consistent.

(c) In Theorem 4.5 it is obviously advantageous to choose a sequence $(\zeta_i)_{i\in\mathbb{N}}$ of non-degenerated, centered, absolutely continuous, and integrable random variables so that for $\mu \in (0,\infty)$ as small as possible holds $\sup_{i\in\mathbb{N}} \mathbb{E}|\zeta_i| \le \mu$. This can, for instance, be achieved by considering $(\zeta_i)_{i\in\mathbb{N}}$ to be a sequence of i.i.d. $\mathcal{N}(0,\sigma^2)$-distributed random variables with $\sigma > 0$, consequently $\sup_{i\in\mathbb{N}} \mathbb{E}|\zeta_i| = \mathbb{E}|\zeta_1| = \sigma\sqrt{\frac{2}{\pi}}$, putting $\mu = \sigma\sqrt{\frac{2}{\pi}}$, and choosing σ as small as possible depending on other framework conditions.

(d) The upper bounds in Theorem 4.5 can be further slightly improved by applying stricter inequalities if possible, and also by choosing the series in (4.33) to have a smaller limit than our for convenience chosen series $\sum_{i=1}^{\infty} \frac{1}{2^i}$ having the limit 1.

5. A simulation study

Herein, we simulate realizations and estimators of our lagged covariance and cross-covariance operators. To avoid unnecessary complexity, and to ensure vividness of the derived results, we discuss centered processes whose underlying processes attain values in $\mathcal{H} := L^2[0,1]$. In our calculations with the program language R, any $x \in \mathcal{H}$ is with exceptions to which we will draw attention evaluated at $t = 0, \frac{1}{250}, \ldots, \frac{249}{250}$, and the inner product $\langle x, y\rangle_{\mathcal{H}} = \int_0^1 x(t)y(t)\,\mathrm{d}t$, with $x, y \in \mathcal{H}$, is approximated by the Riemann sum $\frac{1}{250}\sum_{t=1}^{250} x(\frac{t-1}{250})y(\frac{t-1}{250})$.

5.1. Setup

For some $m, n \in \mathbb{N}$, let $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ and $\boldsymbol{\mathscr{Y}} := (\mathscr{Y}_k)_{k\in\mathbb{Z}}$ be processes with

$$\mathscr{X}_k := (X_k, \ldots, X_{k-m+1})^T \quad \text{resp.} \quad \mathscr{Y}_k := (Y_k, \ldots, Y_{k-n+1})^T, \ k \in \mathbb{Z}, \tag{5.1}$$

so, $p = q = 1$. $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}$ and $\boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$ are processes which satisfy a.s.

$$X_k = \alpha(X_{k-1}) + \varepsilon_k, \quad \forall k \in \mathbb{Z}, \tag{5.2}$$

$$Y_k = \beta(X_k) + \varepsilon_k, \quad \forall k \in \mathbb{Z}, \tag{5.3}$$

with ε_k as in Example 3.1, and $\alpha, \beta\colon \mathcal{H} \to \mathcal{H}$ are integral operators with kernels

$$a(s,t) := k_{\boldsymbol{\varepsilon}}(s,t) \quad \text{resp.} \quad b(s,t) := \frac{1}{2}k_{\boldsymbol{\varepsilon}}(s,t), \quad s,t \in [0,1], \tag{5.4}$$

where $k_{\boldsymbol{\varepsilon};0} = k_{\boldsymbol{\varepsilon}}$ is the integral kernel of $\mathscr{C}_{\boldsymbol{\varepsilon};0} = \mathscr{C}_{\boldsymbol{\varepsilon}}$ in (3.11). Also,

$$||\alpha||^2_{\mathcal{S}_{\mathcal{H}}} = \int_0^1\int_0^1 a^2(s,t)\,\mathrm{d}s\mathrm{d}t \;=\; \frac{3}{2} - \ln(2). \tag{5.5}$$

Hence, $||\alpha||_{\mathcal{L}_{\mathcal{H}}} \le ||\alpha||_{\mathcal{S}_{\mathcal{H}}} < 1$, implying (5.2) has a unique stationary solution, with convergence both in $L^4_{\mathcal{H}}$ and a.s., thus $(X_k)_k$ and $(Y_k)_k$ are stationary, centered, $L^4_{\mathcal{H}}$-m-approximable $AR(1)$ *processes* (see [39], Lemma 2.1). Further,

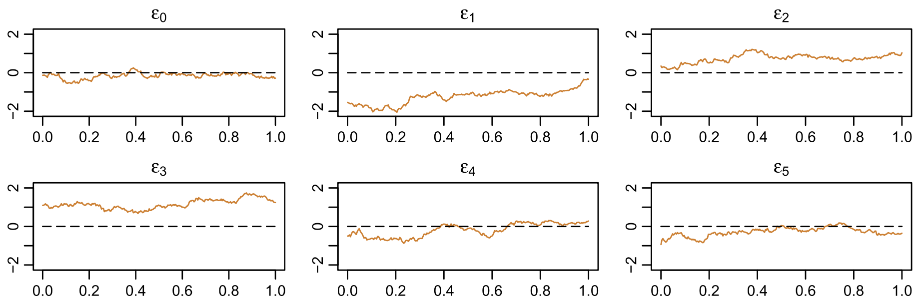

FIG 6. *Realizations of the innovations $\varepsilon_0, \dots, \varepsilon_5$ in* (3.10).

$\mathscr{C}^*_{\boldsymbol{X};h} = \mathscr{C}_{\boldsymbol{X};-h}$ for $h \in \mathbb{Z}$, (3.13), since $\alpha = \mathscr{C}_{\boldsymbol{\varepsilon}}$ is selfadjoint and commutes with $\mathscr{C}_{\boldsymbol{\varepsilon}}$ and due to (3.14) also with $\mathscr{C}_{\boldsymbol{X}}$, and $||\alpha||_{\mathcal{S}_{\mathcal{H}}} < 1$ lead to the Neumann series

$$\mathscr{C}_{\boldsymbol{X};h} = \alpha^{|h|+1} \sum_{j=0}^{\infty} \alpha^{2j} = \alpha^{|h|+1} \big(\mathbb{I}_{\mathcal{H}} - \alpha^2\big)^{-1}, \quad \forall h \in \mathbb{Z}. \tag{5.6}$$

Moreover, (5.2), (5.3), elementary conversions and (5.6) yield

$$\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h} = \beta\, \mathscr{C}_{\boldsymbol{X};h}, \quad \forall h \in \mathbb{Z}. \tag{5.7}$$

5.2. Simulation of realizations of our processes

Here, we simulate realizations of $(\mathscr{X}_k)_k, (\mathscr{Y}_k)_k$ in (5.1). For this purpose, we firstly simulate innovations in (3.10), see Fig. 6, which then can be plugged into the equations (5.2) and (5.3) of the underlying AR(1) process $(X_k)_k$ of $(\mathscr{X}_k)_k$ and the derived underlying process $(Y_k)_k$ of $(\mathscr{Y}_k)_k$. But before we do so, an initial value of X_0 has to be simulated which can be approximated sufficiently well as follows.

Lemma 5.1. *Let $A \in \mathcal{L}_{\mathcal{H}}$ with $||A||_{\mathcal{L}_{\mathcal{H}}} < 1$. Further, let $(\varepsilon_k)_{k\in\mathbb{Z}}$ be an i.i.d., centered $L^{\nu}_{\mathcal{H}}$-process for $\nu > 0$, let $Z_k = A(Z_{k-1}) + \varepsilon_k$ for all $k \in \mathbb{Z}$, and let $\tilde{Z}_k = A(\tilde{Z}_{k-1}) + \varepsilon_k$ for all $k \in \mathbb{N}$ where $\tilde{Z}_0$ is some deterministic value. Then, $\mathbb{E}||Z_0 - \tilde{Z}_0||^{\nu}_{\mathcal{H}} < \infty$, and with $\rho := ||A||^{-\nu}_{\mathcal{L}_{\mathcal{H}}} > 1$ holds*

$$\rho^N \mathbb{E}||Z_N - \tilde{Z}_N||^{\nu}_{\mathcal{H}} \leq \mathbb{E}||Z_0 - \tilde{Z}_0||^{\nu}_{\mathcal{H}}, \quad \forall N \in \mathbb{N}.$$

Remarks 5.1. Such a statement holds also for functional AR(MA) processes with arbitrary order(s) in any separable Hilbert space, see [39], Corollary 4.1 for functional (G)ARCH which directly transfers to functional AR(MA) processes.

5.3. Simulation of our operators

In this section, we illustrate certain lag-h-covariance and lag-h-cross-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ resp. $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ of the centered processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} =$

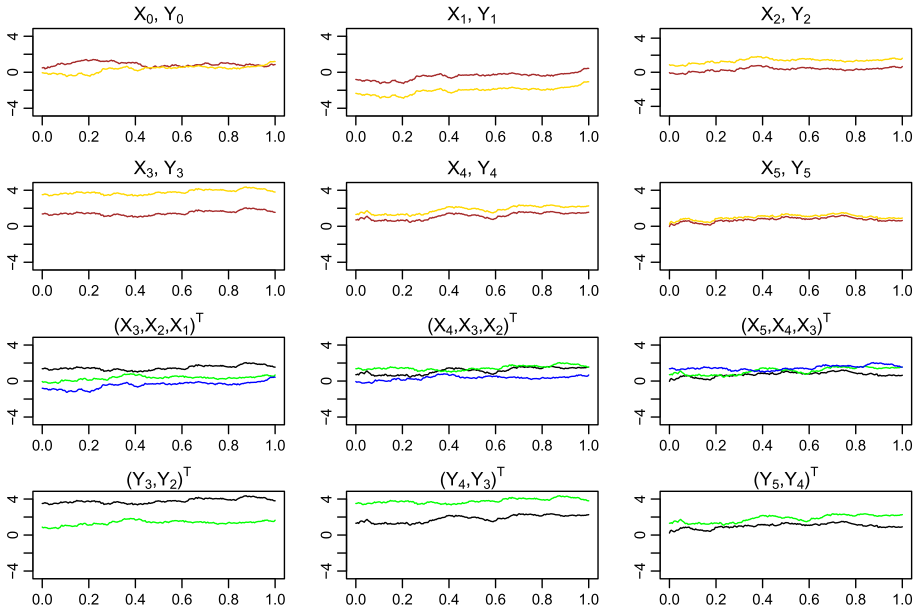

FIG 7. *Six consecutive realizations of $(X_k)_k$ (**bordeaux**) and $(Y_k)_k$ (**gold**) in the first two rows. X_0 was approximated by $\tilde{Z}_{100}$ in Lemma 5.1 with $A = \alpha, \varepsilon_k$ for $k = 1, \ldots, 100$ as in (3.10), and $\tilde{Z}_0 := 0_{\mathcal{H}}$, and $X_1, \ldots, X_5$ and $Y_0, \ldots, Y_5$ were obtained by applying (5.2) resp. (5.3) with the innovations in Fig. 6. Then, $X_0, \ldots, X_5$ and $Y_0, \ldots, Y_5$ were plugged into the equations in (5.1) with $m = 3$ and $n = 2$, leading to three consecutive realizations of $(\mathscr{X}_k)_k = ((X_k, X_{k-1}, X_{k-2})^T)_k$ (third row) and of $(\mathscr{Y}_k)_k = ((Y_k, Y_{k-1})^T)_k$ (fourth row). The first resp. the second components of both the realizations of $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are highlighted in **black** resp. **green**, and the third component of $(\mathscr{X}_k)_k$ in **blue**.*

$(\mathscr{Y}_k)_k$ in Section 5.1 with Cartesian powers $m = 3$ resp. $n = 2$, and simulate their estimates for fixed and increasing h, m, n. For any $h \in \mathbb{Z}$, $\mathscr{C}_{\mathbf{X};h}$ and $\mathscr{C}_{\mathbf{X},\mathbf{Y};h}$ cannot be calculated precisely due to the infinite series (5.6) consisting of operators, but can for sufficiently large $K \in \mathbb{N}$ be well approximated by

$$\tilde{\mathscr{C}}_{\mathbf{X};h;K} := \alpha^{|h|+1} \sum_{j=0}^{K} \alpha^{2j} \quad \text{resp.} \quad \tilde{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h;K} := \beta\alpha^{|h|+1} \sum_{j=0}^{K} \alpha^{2j}. \tag{5.8}$$

This is due to the fact that submultiplicity of $||\cdot||_{\mathcal{S}_\mathcal{H}}$, $||\alpha||_{\mathcal{S}_\mathcal{H}} < 1$ and the formulas of the geometric sum and series lead with $c := (1 - ||\alpha||^2_{\mathcal{S}_\mathcal{H}})^{-1}$ and $\beta = \frac{1}{2}\alpha$ after (5.4) for any h, K to

$$||\tilde{\mathscr{C}}_{\mathbf{X};h;K} - \mathscr{C}_{\mathbf{X};h}||_{\mathcal{S}_\mathcal{H}} < c||\alpha||^{2K+3}_{\mathcal{S}_\mathcal{H}} \quad \text{and} \quad ||\tilde{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h;K} - \mathscr{C}_{\mathbf{X},\mathbf{Y};h}||_{\mathcal{S}_\mathcal{H}} < \frac{c}{2}\,||\alpha||^{2K+4}_{\mathcal{S}_\mathcal{H}}.$$

Also, due to the fact that the lag-h-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}};h} = \mathbb{E}\langle \mathscr{X}_0, \boldsymbol{x}\rangle_{\mathcal{H}^m} \mathscr{X}_h$ and the lag-h-cross-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \mathbb{E}\langle \mathscr{X}_0, \boldsymbol{x}\rangle_{\mathcal{H}^m} \mathscr{Y}_h$ fulfill under

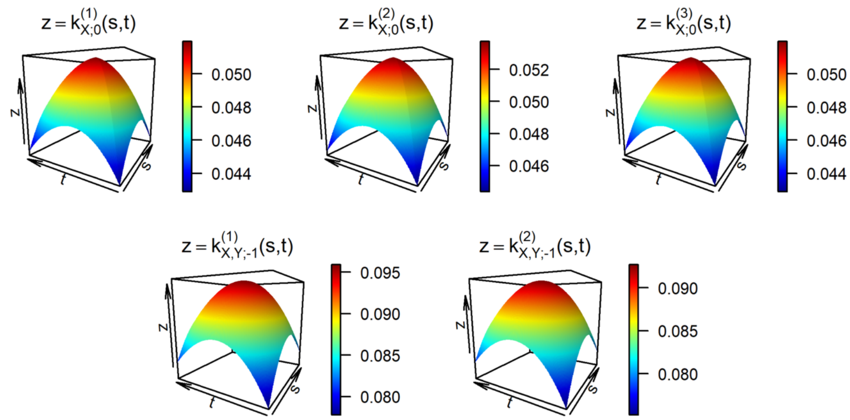

FIG 8. *The integral kernels* $k^{(1)}_{\mathscr{X};0}, k^{(2)}_{\mathscr{X};0}, k^{(3)}_{\mathscr{X};0}$ *(first row) and* $k^{(1)}_{\mathscr{X},\mathscr{Y};-1}, k^{(2)}_{\mathscr{X},\mathscr{Y};-1}$ *(second row) of the operators in the three resp. two components of* $\mathscr{C}_{\mathscr{X};0}$ *in* (5.11) *resp.* $\mathscr{C}_{\mathscr{X},\mathscr{Y};-1}$ *in* (5.12). *These kernels result by the associated sum of the integral kernels* $k_{\boldsymbol{X};0}, k_{\boldsymbol{X};1}, k_{\boldsymbol{X};2}$ *and* $k_{\boldsymbol{X},\boldsymbol{Y};0}, k_{\boldsymbol{X},\boldsymbol{Y};1}, k_{\boldsymbol{X},\boldsymbol{Y};2}$ *of the operators* $\mathscr{C}_{\boldsymbol{X};0}, \mathscr{C}_{\boldsymbol{X};1}, \mathscr{C}_{\boldsymbol{X};2}$ *resp.* $\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};0}, \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};1}, \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};2}$ *which were approximated by their respective operators in* (5.8) *with* $K = 100$.

Assumptions 4.1 for any $h \in \mathbb{Z}$ and $\boldsymbol{u} \coloneqq (u_1, \dots, u_m)^T \in \mathcal{U}^m$,

$$\mathscr{C}_{\mathscr{X};h}(\boldsymbol{u}) = \Big(\sum_{i=1}^{m} \mathscr{C}_{\boldsymbol{X};h+i-1}(u_i), \dots, \sum_{i=1}^{m} \mathscr{C}_{\boldsymbol{X};h+i-m}(u_i) \Big)^T \in \mathcal{U}^m, \tag{5.9}$$

$$\mathscr{C}_{\mathscr{X},\mathscr{Y};h}(\boldsymbol{u}) = \Big(\sum_{i=1}^{m} \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h+i-1}(u_i), \dots, \sum_{i=1}^{m} \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h+i-n}(u_i) \Big)^T \in \mathcal{V}^n, \tag{5.10}$$

the components of $\mathscr{C}_{\mathscr{X};h}$ and $\mathscr{C}_{\mathscr{X},\mathscr{Y};h}$ cannot be expressed independently of any argument $\boldsymbol{x} \coloneqq (x_1, \dots, x_m)^T \in \mathcal{H}^m$, except when all the argument's components match. With $(A_1(\boldsymbol{x}), \dots, A_m(\boldsymbol{x})) \coloneqq (A_1, \dots, A_m)(\boldsymbol{x})$ for operators $A_1, \dots, A_m$ with domain $\mathcal{H}^m$, and postulating $\mathscr{C}_{\boldsymbol{X};h} = \mathscr{C}_{\boldsymbol{X};-h}$ and $\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h} = \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};-h}$ for any h, we obtain for, e.g., $\mathscr{C}_{\mathscr{X};0}$ and $\mathscr{C}_{\mathscr{X},\mathscr{Y};-1}$ with $m = 3, n = 2$ for any $\boldsymbol{x} = (x, x, x) \in \mathcal{H}^3$ according to (5.9), (5.10),

$$\mathscr{C}_{\mathscr{X};0}(\boldsymbol{x}) = \Big(\big(\mathscr{C}_{\boldsymbol{X};0} + \mathscr{C}_{\boldsymbol{X};1} + \mathscr{C}_{\boldsymbol{X};2}, \mathscr{C}_{\boldsymbol{X};0} + 2\mathscr{C}_{\boldsymbol{X};1}, \mathscr{C}_{\boldsymbol{X};0} + \mathscr{C}_{\boldsymbol{X};1} + \mathscr{C}_{\boldsymbol{X};2} \big)(x) \Big)^T, \tag{5.11}$$

$$\mathscr{C}_{\mathscr{X},\mathscr{Y};-1}(\boldsymbol{x}) = \Big(\big(\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};0} + 2\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};1}, \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};0} + \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};1} + \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};2} \big)(x) \Big)^T. \tag{5.12}$$

To illustrate estimators for the operators in the components of $\mathscr{C}_{\mathscr{X};0}(\boldsymbol{x})$ in (5.11) and $\mathscr{C}_{\mathscr{X},\mathscr{Y};-1}(\boldsymbol{x})$ in (5.12), and to estimate $\mathscr{C}_{\mathscr{X};h}$ and $\mathscr{C}_{\mathscr{X},\mathscr{Y};h}$ for fixed and varying h, m, n, with $h \geq 0$ w.l.o.g., we generate $X_1, \dots, X_M$ and $Y_1, \dots, Y_N$ of the processes $\boldsymbol{X}$ resp. $\boldsymbol{Y}$ in Section 5.1 with $M = N$. This leads to the values $\mathscr{X}_m, \dots, \mathscr{X}_{\tilde{\mathscr{M}}}$ of $\mathscr{X}$ and $\mathscr{Y}_n, \dots, \mathscr{Y}_{\tilde{\mathscr{N}}}$ of $\mathscr{Y}$ with $\tilde{\mathscr{M}} = \tilde{\mathscr{M}}_M = M$ and $\tilde{\mathscr{N}} = \tilde{\mathscr{N}}_N = M$, thus with $\mathscr{M} = \mathscr{M}_M = M - m + 1$ resp. $\mathscr{N} = \mathscr{N}_N = M - n + 1$. Due to centeredness of $\boldsymbol{X}$ and $\boldsymbol{Y}$, the operators $\mathscr{C}_{\boldsymbol{X};h}$ in (5.11) and $\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h}$ in (5.12)

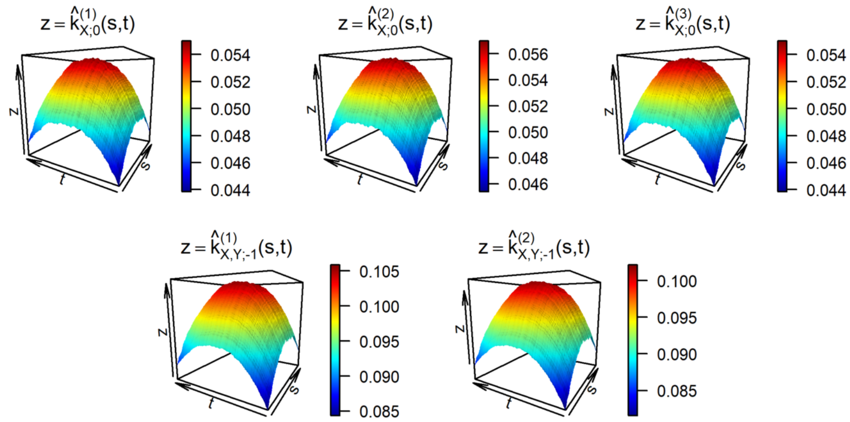

FIG 9. *The estimators* $\hat{k}^{(1)}_{\boldsymbol{\mathscr{X}};0}, \hat{k}^{(2)}_{\boldsymbol{\mathscr{X}};0}, \hat{k}^{(3)}_{\boldsymbol{\mathscr{X}};0}$ *(first row) and* $\hat{k}^{(1)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}, \hat{k}^{(2)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ *(second row) for the integral kernels* $k^{(1)}_{\boldsymbol{\mathscr{X}};0}, k^{(2)}_{\boldsymbol{\mathscr{X}};0}, k^{(3)}_{\boldsymbol{\mathscr{X}};0}$ *resp.* $k^{(1)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}, k^{(2)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ *of the operators in the three resp. two components of* $\mathscr{C}_{\boldsymbol{\mathscr{X}};0}$ *in* (5.11) *resp.* $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ *in* (5.12)*. These estimators result by the associated sum of the estimators* $\hat{k}_{\mathbf{X};0}, \hat{k}_{\mathbf{X};1}, \hat{k}_{\mathbf{X};2}$ *in* (5.13) *and* $\hat{k}_{\mathbf{X},\mathbf{Y};0}, \hat{k}_{\mathbf{X},\mathbf{Y};1}, \hat{k}_{\mathbf{X},\mathbf{Y};2}$ *in* (5.14) *with* $M = 1000$ *for the operators* $\mathscr{C}_{\mathbf{X};0}, \mathscr{C}_{\mathbf{X};1}, \mathscr{C}_{\mathbf{X};2}$ *resp.* $\mathscr{C}_{\mathbf{X},\mathbf{Y};0}, \mathscr{C}_{\mathbf{X},\mathbf{Y};1}, \mathscr{C}_{\mathbf{X},\mathbf{Y};2}$*.*

with $h = 0, 1, 2$ are estimated by the classical estimators $\hat{\mathscr{C}}_{\mathbf{X};h}$ resp. by $\hat{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h}$ with integral kernels

$$\hat{k}_{\mathbf{X};h}(s,t) := \frac{1}{M-h} \sum_{k=1}^{M-h} X_k(s) X_{k+h}(t), \quad \forall s,t \in [0,1], \tag{5.13}$$

$$\text{resp.} \quad \hat{k}_{\mathbf{X},\mathbf{Y};h}(s,t) := \frac{1}{M-h} \sum_{k=1}^{M-h} X_k(s) Y_{k+h}(t), \quad \forall s,t \in [0,1]. \tag{5.14}$$

At last, in Table 1, we list estimation errors for the operators $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ of the processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} = (\mathscr{Y}_k)_k$ in (5.1) for several sample sizes $M = N$ and various h, m, n which may depend on M, with $h \geq 0$ w.l.o.g. Due to centeredness of $\boldsymbol{\mathscr{X}}$ and $\boldsymbol{\mathscr{Y}}$, we use the estimators $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h}$ in (4.9) and $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ in (4.15), which satisfy our processes' definition, and $h \geq 0$,

$$\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} = \frac{1}{M-h-m+1} \sum_{k=m}^{M-h} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, \tag{5.15}$$

$$\text{resp.} \quad \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \frac{1}{M-h-m\vee(n-h)+1} \sum_{k=m\vee(n-h)}^{M-h} \mathscr{X}_k \otimes \mathscr{Y}_{k+h}, \tag{5.16}$$

and to calculate the estimation errors, we utilize the identity (4.10) and that for $p = q = 1$ holds after (4.16),

$$||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}} = \sum_{i=1}^{m} \sum_{j=1}^{n} ||\hat{\mathscr{C}}_{\mathbf{X},\mathbf{Y};j-i+m+h-n} - \mathscr{C}_{\mathbf{X},\mathbf{Y};j-i+m+h-n}||^2_{\mathcal{S}_{\mathcal{H}}}, \tag{5.17}$$

TABLE 1

Simulation of $\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{H}^m}}$ *and* $\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}}$, *with the estimation errors defined in* (4.10) *resp.* (5.17) *for various sample sizes* $M = N$, *lags* h *and Cartesian powers* m, n. *Here, any* $x \in \mathcal{H}$ *is evaluated at* $t = 0, \frac{1}{100}, \dots, \frac{99}{100}$, *and the inner product* $\langle x, y\rangle_{\mathcal{H}}$ *is approximated by* $\frac{1}{100}\sum_{t=1}^{100} x(\frac{t-1}{100})y(\frac{t-1}{100})$ *for any* $x, y \in \mathcal{H}$. *The listed values in the table correspond to the arithmetic mean of calculated estimation errors with* $R = 10$ *replications of generated random variables, where* $\mathscr{C}_{\mathbf{X};h+i-j}$ *and* $\mathscr{C}_{\mathbf{X},\mathbf{Y};h+i-j}$ *were approximated by* $\tilde{\mathscr{C}}_{\mathbf{X};h+i-j;100}$ *resp.* $\tilde{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h+i-j;100}$ *in* (5.8).

	$\approx \mathbb{E}\|\|\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}\|\|^2_{\mathcal{S}_{\mathcal{H}^m}}$						$\approx \mathbb{E}\|\|\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}\|\|^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}}$					
$m=m_M, n=n_M$	$m=3$			$m=\lfloor M^{1/4}\rfloor$			$m=3, n=2$			$m=n=\lfloor M^{1/4}\rfloor$		
$h=h_M$ / M	0	1	$\lfloor M^{1/4}\rfloor$	0	1	$\lfloor M^{1/4}\rfloor$	0	1	$\lfloor M^{1/4}\rfloor$	0	1	$\lfloor M^{1/4}\rfloor$
100	.1451	.1231	.1101	.1463	.0874	.0653	.1238	.0298	.0247	.0754	.0307	.0046
200	.0729	.0336	.0535	.0559	.0703	.0449	.0366	.0465	.0117	.0540	.0397	.0506
300	.0546	.0726	.0294	.2343	.1763	.1413	.0223	.0324	.0098	.0638	.0834	.0271
400	.0895	.0515	.0427	.2564	.1962	.1084	.0089	.0238	.0390	.1526	.2537	.0252
500	.0704	.0565	.0255	.2697	.2379	.1074	.0388	.0293	.0065	.1367	.0663	.0236
750	.0674	.0609	.0175	.5185	.5750	.2834	.0345	.0304	.0011	.2634	.2491	.1402
1000	.0456	.0345	.0271	.5867	.5148	.3280	.0255	.0246	.0121	.3581	.3172	.0755

where $\hat{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h+i-j}$ corresponds to $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ in (4.15), with $\mathscr{X}_k$ and $\mathscr{Y}_{k+h}$ replaced by $X_{i+(k-m)p}$ resp. $Y_{j+(k+h-n)q}$ for all i, j, k.

Remarks 5.2. The simulated values in Table 1 display our theoretical results in Theorems 4.1 and 4.3 reasonably well. These theorems state that the squared estimation errors for our lagged covariance and cross-covariance operators, where the errors are calibrated with certain increasing sequences depending on the lag, Cartesian powers and sample sizes, converge to a specific constant for the samples size M, N tending to infinity. Thereby, the prerequisites of these theorems are satisfied due to $p = q = 1, M = N$, the choice of our lags $h = h_M$ and Cartesian powers $m = m_M, n = n_N$ in Table 1, and the definition of our centered AR(1) processes. That the pattern of the numbers listed in the table reflects the assertions of both theorems in a suggestive way, is because the estimation errors seem to approach zero for any h and fixed m, n for increasing sample size M, where the errors decay for increasing lag h slightly faster. It is also reasonable that the estimation errors are bigger when the Cartesian powers are. Further, that the values are not monotonic for increasing sample sizes can be explained by high fluctuation of individual simulations, see Fig. 10. Moreover, the estimation errors for the lagged cross-covariance operators are, except for a few irregularities, roughly one quarter of the lagged cross-covariance operators, which can be explained by the identity (5.7) and the definition of the kernel of β defined in (5.4). The downside of our simulation, however, is that, according to our results, the estimation errors should for increasing Cartesian powers decrease rather than increase or stagnate. This may be due to the fact that the relatively coarse decomposition of the unit interval with only $n = 100$ interpolation points implies too large approximation errors, or that the sample size of up to $M = 1000$ is still too small to notice convergence to zero. However, although we chose n, M and the number of repetitons R just enough to make a passable statement, it was, due to the complexity of our operators

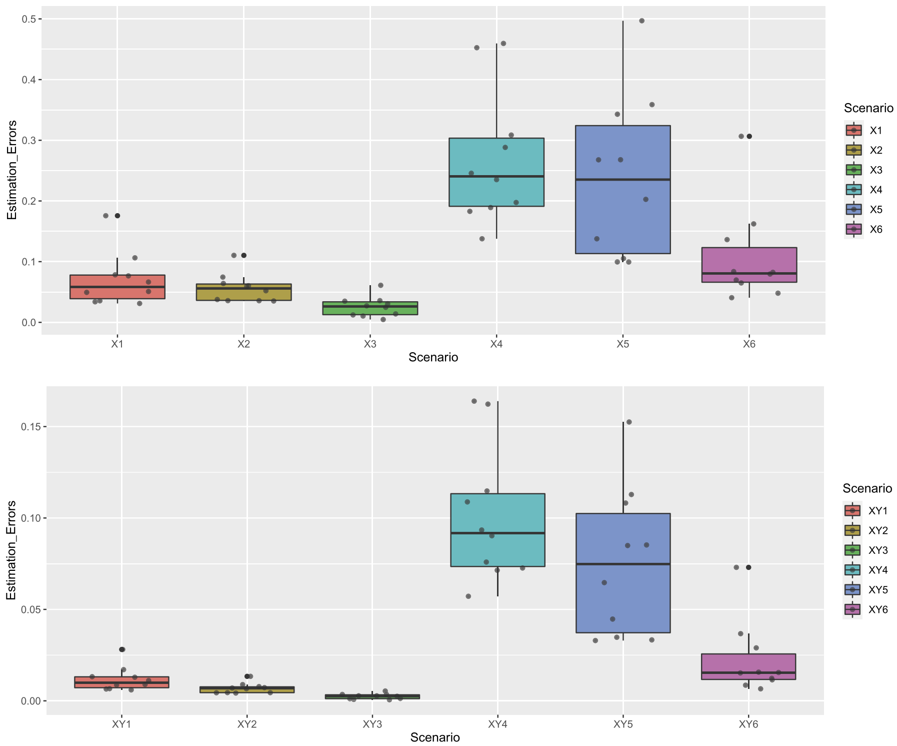

Fig 10. *Boxplots and individual datapoints of the $R = 10$ simulated estimation errors in each scenario for the lag-h-covariance operators $\mathscr{C}_{\mathcal{X};h}$ (upper chart) and the lag-h-cross-covariance operators $\mathscr{C}_{\mathcal{X},\mathcal{Y};h}$ (lower chart) in Table 1. Thereby, these simulated estimation errors were plotted using the sample size $M = 500$ in all the scenarios* ***X1****: $m = 3, h = 0$;* ***X2****: $m = 3, h = 1$;* ***X3****: $m = 3, h = \lfloor M^{1/4} \rfloor$;* ***X4****: $m = \lfloor M^{1/4} \rfloor, h = 0$;* ***X5****: $m = \lfloor M^{1/4} \rfloor, h = 1$;* ***X6****: $m = h = \lfloor M^{1/4} \rfloor$ for $\mathscr{C}_{\mathcal{X};h}$ resp. in all the scenarios* ***XY1****: $m = 3, n = 2, h = 0$;* ***XY2****: $m = 3, n = 2, h = 1$;* ***XY3****: $m = 3, n = 2, h = \lfloor M^{1/4} \rfloor$;* ***XY4****: $m = n = \lfloor M^{1/4} \rfloor, h = 0$;* ***XY5****: $m = n = \lfloor M^{1/4} \rfloor, h = 1$;* ***XY6****: $m = n = h = \lfloor M^{1/4} \rfloor$ for $\mathscr{C}_{\mathcal{X},\mathcal{Y};h}$.*

and the performance of the program language R, not possible for us to further increase them to improve the statement. Just to give an example, merely simulating these twelve values in Table 1 for $M = 500$ already took several hours.

6. Conclusions

Summary This article proposes estimators for lagged covariance and cross-covariance operators and the principal components of processes in Cartesian products of separable Hilbert spaces. The focus lies on the estimation procedure and the asymptotic behaviour of the estimation errors. All estimators are stated for centered processes and for those with an unknown finite, first moment. The asymptotic results allow the processes' Cartesian powers and the lag to be fixed or to increase w.r.t. the sample size(s), and the principal components are estimated separately and uniformly. Our findings are useful whenever one is concerned about the dependence within or between processes with values in (Cartesian products of) separable Hilbert spaces, or whenever one has to analyze estimators relying on empirical (lagged) covariance or cross-covariance operators, see [3, 38, 39]. These findings can also be applied to covariance and cross-covariance operators of random variables in separable Hilbert spaces, and, since $\mathbb{R}^n$ endowed with the canonical inner product is a separable Hilbert space for any $n \in \mathbb{N}$, also to conventional (lagged) covariance and cross-covariance matrices.

Outlook In the future, one could tackle to derive the asymptotic distribution of our estimation errors and Bernstein inequalities (see [7, 46]), and analyze asymptotic lower rates. Furthermore, it would be interesting to extend our results on separable Banach spaces, see, e.g., [49] who estimated AR operators in Banach spaces.

Appendix A: Proofs

In multiple conversions, we use the following identities regarding the cardinality of certain index sets. To define these, let $\mathbb{Z}_{|\cdot|<c} := \{j \in \mathbb{Z} \,\big|\, |j| < c\}, c \in \mathbb{R}$. For any $n \in \mathbb{N}$ and $k \in \mathbb{Z}_{|\cdot|<n}$ holds

$$\#\big\{i, j \in \{1, \ldots, n\} \,\big|\, j - i = k\big\} = n - |k|. \tag{A.1}$$

To reiterate, $x^+ := \max(0, x)$ for $x \in \mathbb{R}$, and $a \vee b := \max(a, b)$ and $a \wedge b := \min(a, b)$ for $a, b \in \mathbb{R}$. For any $m, n \in \mathbb{N}$ and $k \in \mathbb{Z}_{|\cdot|<m\vee n}$ holds

$$\#\big\{i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\} \,\big|\, j - i = k\big\} = \iota_{m,n}(k), \tag{A.2}$$

where the functions $\iota_{m,n} : \mathbb{Z}_{|\cdot|<m\vee n} \to \{1, \ldots, m \wedge n\}$ are for $n \geq m$ defined by

$$\iota_{m,n}(k) \overset{n\geq m}{:=} \begin{cases} (m - |k|)^+, & \text{if } \quad k < 0, \\ m, & \text{if } 0 \leq k \leq n - m, \\ (m - |k - (n - m)|)^+, & \text{if } \quad k > n - m, \end{cases} \tag{A.3}$$

and for $n < m$ through

$$\iota_{m,n}(k) \overset{n\leq m}{:=} \begin{cases} (n - |(n - m) - k|)^+, & \text{if } \quad k < n - m, \\ n, & \text{if } n - m \leq k \leq 0, \\ (n - |k|)^+, & \text{if } \quad k > 0. \end{cases} \tag{A.4}$$

Lemma A.1. *Let Assumptions 4.1, 4.3 and 4.8 hold. Further, with $N = N_M$, let $\bar{m} := \lim_{M\to\infty} m_M, \bar{n} := \lim_{M\to\infty} n_N \in [0, \infty]$, let $\xi_{k,l} := 1 - \frac{|k|}{l}$ for $k \in \mathbb{Z}$ and $l \in \mathbb{N}$, and let $\tilde{\iota}_{m,n}(k) := \iota_{m,n}(k)/m \wedge n$ with $\iota_{m,n}(k)$ defined in (A.3)–(A.4)* *for any* k, m, n. *Then, for any* $k \in \mathbb{Z}_{|\cdot|<\bar{m}\vee\bar{n}}$, *where* $\mathbb{Z}_{|\cdot|<\bar{m}\vee\bar{n}} = \mathbb{Z}$ *if* $\bar{m} = \infty$ *or* $\bar{n} = \infty$, *holds for* $\lim_{M\to\infty} \tilde{\iota}_{m,n}(k) := \tilde{\iota}^*(k)$:

$$\tilde{\iota}^*(k) = \begin{cases} \mathbf{1}_{(-m^*,0)}(k)\cdot\xi_{k,m^*} + \mathbf{1}_{[0,n^*-m^*]}(k) + \mathbf{1}_{(n^*-m^*,n^*)}(k)\cdot\xi_{k-(n^*-m^*),m^*}, & \text{if } m \to m^*, n \to n^*, n^* \geq m^*, \\ \mathbf{1}_{(-m^*,n^*-m^*)}(k)\cdot\xi_{(n^*-m^*)-k,n^*} + \mathbf{1}_{[n^*-m^*,0]}(k) + \mathbf{1}_{(0,n^*)}(k)\cdot\xi_{k,n^*}, & \text{if } m \to m^*, n \to n^*, n^* < m^*, \\ \mathbf{1}_{(-\infty,0]}(k) + \mathbf{1}_{(0,n^*)}(k)\cdot\xi_{k,n^*}, & \text{if } m \to \infty, n \to n^*, \\ \mathbf{1}_{(-m^*,0)}(k)\cdot\xi_{k,m^*} + \mathbf{1}_{[0,\infty)}(k), & \text{if } m \to m^*, n \to \infty, \\ 1, & \text{if } m \to \infty, n \to \infty. \end{cases} \tag{A.5}$$

Proof. The assertion follows directly from the definition of the functions $\iota_{m,n}$ and $\tilde{\iota}_{m,n}$ for any $m, n \in \mathbb{N}$ and basic conversions. □

Further, for any $x \in \mathbb{Z}_{|\cdot|<a+b-1}$ with $a, b \in \mathbb{Z}$ holds

$$\#\{i \in \mathbb{Z}_{|\cdot|<a}, j \in \mathbb{Z}_{|\cdot|<b} \,|\, i + j = x\} = \psi_{a,b}(x), \tag{A.6}$$

where the function $\psi_{a,b} \colon \mathbb{Z}_{|\cdot|<a+b-1} \to \{1, \ldots, 2a \wedge b + 1\}$ is defined as

$$\psi_{a,b}(x) := \begin{cases} a + b + 1 + x, & \text{if } \quad x < a \wedge b - a \vee b, \\ 2a \wedge b + 1, & \text{if } a \wedge b - a \vee b \leq x \leq a \vee b - a \wedge b, \\ a + b + 1 - x, & \text{if } \quad x > a \vee b - a \wedge b. \end{cases} \tag{A.7}$$

We also make use of the following inequality.

Lemma A.2. *Let* $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ *be a separable Hilbert space. Also, let* $(S_k)_{k\in\mathbb{Z}}$ *be a stationary* $L^4_{\mathcal{H}}$*-process, and for some* $l \in \mathbb{N}$, $\mathscr{S}_k := (S_{f(k,1)}, \ldots, S_{f(k,l)})^T$ *for all* k *and some function* $f \colon \mathbb{Z} \times \{1, \ldots, l\} \to \mathbb{Z}$. *Then,*

$$\nu_{4,\mathcal{H}^l}(\mathscr{S}_k) \leq \sqrt{l}\, \nu_{4,\mathcal{H}}(S_j), \quad \forall j, k. \tag{A.8}$$

Proof. From the definition of $\mathscr{S}_k$ and $\nu_{4,\mathcal{H}^l}(\cdot)$, stationarity of the process $(S_k)_k \subseteq L^4_{\mathcal{H}}$ and the Cauchy-Schwarz inequality follows

$$\nu^4_{4,\mathcal{H}^l}(\mathscr{S}_k) = \mathbb{E}\Big[\Big(\sum_{m=1}^{l} ||S_{f(k,m)}||^2_{\mathcal{H}}\Big)^2\Big] \leq \sum_{m,n=1}^{l} \mathbb{E}||S_j||^4_{\mathcal{H}} = l^2 \nu^4_{4,\mathcal{H}}(S_j).$$

□

Proof of Lemma 4.1. At first, $\hat{m}_{\boldsymbol{\mathscr{X}}}$ and $\hat{m}_{\boldsymbol{\mathscr{Y}}}$ in (4.2) are unbiased estimators for $m_{\boldsymbol{\mathscr{X}}}$ for all $M \in \mathbb{N}$ resp. for $m_{\boldsymbol{\mathscr{Y}}}$ for all $N \in \mathbb{N}$ due to their definition. Hereinafter, we write $m = m_M$ and $\mathscr{M} = \mathscr{M}_M$ for $M \in \mathbb{N}$. Further, w.l.o.g., we illustrate the proof of (a) and (b) for $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ only.

(a) The definition of $\hat{m}_{\boldsymbol{\mathscr{X}}}$ in (4.2) and stationarity of $\boldsymbol{X} = (X_k)_k \subseteq L^2_{\mathcal{U}}$ implying $m_{\boldsymbol{\mathscr{X}}} = (m_{\boldsymbol{X}}, \ldots, m_{\boldsymbol{X}})^T \in \mathcal{U}^m$, $\hat{m}^{(p,k)}_{\boldsymbol{X}} := \frac{1}{\mathscr{M}} \sum_{i=1}^{\mathscr{M}} X_{k+(i-m)p}$ for $k = 0, 1, \ldots, m$ for any $M \in \mathbb{N}$, and also with $X'_j := X_j - m_{\boldsymbol{X}}$ and $\xi_{i,n} := 1 - \frac{|i|}{n}$, where $i, j \in \mathbb{Z}$ and $n \in \mathbb{N}$, yield

$$\frac{\mathscr{M}}{m}\, \mathbb{E}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m} = \frac{\mathscr{M}}{m} \sum_{k=1}^{m} \mathbb{E}||\hat{m}^{(p,k)}_{\boldsymbol{X}} - m_{\boldsymbol{X}}||^2_{\mathcal{U}} = \mathscr{M}\, \mathbb{E}||\hat{m}^{(p,0)}_{\boldsymbol{X}} - m_{\boldsymbol{X}}||^2_{\mathcal{U}}$$

$$= \frac{1}{\mathscr{M}} \sum_{i,j=1}^{\mathscr{M}} \mathbb{E}\langle X'_{ip}, X'_{jp}\rangle_{\mathcal{U}} = \sum_{|k|<\mathscr{M}} \xi_{k,\mathscr{M}}\, \mathbb{E}\langle X'_0, X'_{kp}\rangle_{\mathcal{U}}$$

$$= \sum_{k\in\mathbb{Z}} \mathbb{E}\langle X'_0, X'_{kp}\rangle_{\mathcal{U}} \tag{A.9}$$

$$- \sum_{|k|<\mathscr{M}} \frac{k}{\mathscr{M}}\, \mathbb{E}\langle X'_0, X'_{kp}\rangle_{\mathcal{U}} - \sum_{|k|\geq\mathscr{M}} \mathbb{E}\langle X'_0, X'_{kp}\rangle_{\mathcal{U}}. \tag{A.10}$$

(A.9) equals $||\sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{X};kp}||_{\mathcal{N}_{\mathcal{U}}}$ which is finite due to $\sum_{k\in\mathbb{Z}} ||\mathscr{C}_{\boldsymbol{X};k}||_{\mathcal{N}_{\mathcal{U}}} < \infty$ (Assumption 4.5 (a)), and thus (A.10) converges towards zero for $M\to\infty$. Further, if $p = p_M \to \infty$ for $M\to\infty$, we obtain due to $\sum_{k\in\mathbb{Z}} ||\mathscr{C}_{\boldsymbol{X};k}||_{\mathcal{N}_{\mathcal{U}}} < \infty$ and (3.2),

$$\lim_{M\to\infty} \Big|\Big| \sum_{k\in\mathbb{Z}} \mathscr{C}_{\boldsymbol{X};kp} \Big|\Big|_{\mathcal{N}_{\mathcal{U}}} = \Big|\Big| \sum_{k\in\mathbb{Z}} \lim_{M\to\infty} \mathscr{C}_{\boldsymbol{X};kp} \Big|\Big|_{\mathcal{N}_{\mathcal{U}}} = ||\mathscr{C}_{\boldsymbol{X};0}||_{\mathcal{N}_{\mathcal{U}}} = \mathbb{E}||X_0||^2_{\mathcal{U}}.$$

(b) Using the notation in (a), stationarity of the $L^4_{\mathcal{U}}$-process $\boldsymbol{X} = (X_k)_k$ after Assumption 4.6 (a) and (3.1) lead for any $M\in\mathbb{N}$ to

$$\frac{\mathscr{M}^3}{m^{1+\mathbf{1}_P}}\, \mathbb{E}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^4_{\mathcal{U}^m}$$

$$= \frac{\mathscr{M}^3}{m^{\mathbf{1}_P}} \sum_{|r|<m} \xi_{r,m}\, \mathbb{E}||\hat{m}^{(p,0)}_{\boldsymbol{X}} - m_{\boldsymbol{X}}||^2_{\mathcal{U}} ||\hat{m}^{(p,r)}_{\boldsymbol{X}} - m_{\boldsymbol{X}}||^2_{\mathcal{U}}$$

$$= \frac{1}{m^{\mathbf{1}_P}\mathscr{M}} \sum_{|r|<m} \xi_{r,m} \sum_{i,j,k,l=1}^{\mathscr{M}} \mathbb{E}\langle X'_{(i-m)p}, X'_{(j-m)p}\rangle_{\mathcal{U}} \langle X'_{r+(k-m)p}, X'_{r+(l-m)p}\rangle_{\mathcal{U}}$$

$$= \frac{1}{m^{\mathbf{1}_P}} \sum_{|r|<m} \xi_{r,m} \sum_{|i|,|j|,|k|<\mathscr{M}} \xi_{i,\mathscr{M}}\, \mathbb{E}\langle X'_0 \otimes X'_{r+ip}, X'_{jp}\otimes X'_{r+kp}\rangle_{\mathcal{S}_{\mathcal{U}}}. \tag{A.11}$$

This term is finite after Assumption 4.6 (a), and its limit depends on the limits of $m = m_M$ and $p = p_M$ for $M\to\infty$. If $p_M\to\infty$, we have for any r,

$$\lim_{M\to\infty} \mathbb{E}\langle X'_0\otimes X'_{r+ip}, X'_{jp}\otimes X'_{r+kp}\rangle_{\mathcal{S}_{\mathcal{U}}} = \begin{cases} \mathbb{E}||X'_0||^2_{\mathcal{U}} ||X'_r||^2_{\mathcal{U}}, & \text{if } i=j=k=0,\\ 0, & \text{otherwise.}\end{cases}$$

Thus, (A.11) goes for $M\to\infty$ to $\sum_{|k|<m^*} \xi_{k,m^*}\, \mathbb{E}||X'_0||^2_{\mathcal{U}} ||X'_k||^2_{\mathcal{U}}$ if $m_M\to m^*$ and $p_M\to\infty$, to $\sum_{k\in\mathbb{Z}} \mathbb{E}||X'_0||^2_{\mathcal{U}}||X'_k||^2_{\mathcal{U}}$ if $m_M\to\infty$ and $p_M\to\infty$, and to

$$\frac{1}{m^*} \sum_{|r|<m^*} \sum_{i,j,k\in\mathbb{Z}} \xi_{r,m^*} \mathbb{E}\langle X'_0\otimes X'_{r+ip^*}, X'_{jp^*}\otimes X'_{r+kp^*}\rangle_{\mathcal{S}_{\mathcal{U}}}$$

if $m_M\to m^*$ and $p_M\to p^*$. At last, if $m_M\to\infty$ and $p_M\to p^*$, for sufficiently large $M\in\mathbb{N}$ implying $\mathscr{M} > \frac{m}{p}$, (A.11) becomes for $M\to\infty$:

$$\frac{1}{m} \sum_{\substack{|r|<m,\\ r\nmid p}} \sum_{|i|,|j|,|k|<\mathscr{M}} \xi_{r,m}\xi_{i,\mathscr{M}}\, \mathbb{E}\langle X'_0\otimes X'_{r+ip}, X'_{jp}\otimes X'_{r+kp}\rangle_{\mathcal{S}_{\mathcal{U}}}$$

$$+\frac{1}{m}\sum_{\substack{|r|<m,\\ r|p}}\sum_{|i|,|j|,|k|<\mathscr{M}}\xi_{r,m}\xi_{i,\mathscr{M}}\,\mathbb{E}\langle X_0'\otimes X_{r+ip}', X_{jp}'\otimes X_{r+kp}'\rangle_{\mathcal{S}_{\mathcal{U}}}$$

$$= \mathrm{o}(1)+\frac{1}{m}\sum_{|s|<\frac{m}{p}}\sum_{|i|,|j|,|k|<\mathscr{M}}\xi_{sp,m}\xi_{i,\mathscr{M}}\,\mathbb{E}\langle X_0'\otimes X_{(s+i)p}', X_{jp}'\otimes X_{(s+k)p}'\rangle_{\mathcal{S}_{\mathcal{U}}}$$

$$= \mathrm{o}(1)+\frac{1}{m}\sum_{|j|<\mathscr{M}}\sum_{|t|,|u|<\mathscr{M}+\frac{m}{p}-1}\sum_{\substack{|i|,|k|<\mathscr{M},|s|<\frac{m}{p},\\ s+i=t,s+k=u}}\xi_{sp,m}\xi_{i,\mathscr{M}}\,\mathbb{E}\langle X_0'\otimes X_{tp}', X_{jp}'\otimes X_{up}'\rangle_{\mathcal{S}_{\mathcal{U}}}$$

$$= \mathrm{o}(1)+\sum_{|j|<\mathscr{M}}\sum_{|t|,|u|<\mathscr{M}+\frac{m}{p}-1}\tilde{\psi}_{\frac{m}{p},\mathscr{M}}(t)\,\mathbb{E}\langle X_0'\otimes X_{tp}', X_{jp}'\otimes X_{up}'\rangle_{\mathcal{S}_{\mathcal{U}}}, \tag{A.12}$$

where $\tilde{\psi}_{\frac{m}{p},\mathscr{M}}\colon \mathbb{Z}_{|\cdot|<\frac{m}{p}+\mathscr{M}-1}\to[0,\infty)$ stand for functions which satisfy $0\le\tilde{\psi}_{\frac{m}{p},\mathscr{M}}(t)\le\psi_{\frac{m}{p},\mathscr{M}}(t)/m$ for all t, with $\psi_{\frac{m}{p},\mathscr{M}}$ defined in (A.7). Thus, due to $\limsup_{M\to\infty}\psi_{\frac{m}{p},\mathscr{M}}(t)/m\le\frac{2}{p^*}$ and Assumption 4.6 (a), (A.12) converges for $M\to\infty$ as asserted to

$$\sum_{i,j,k\in\mathbb{Z}} c_i\,\mathbb{E}\langle X_0'\otimes X_{ip^*}', X_{jp^*}'\otimes X_{kp^*}'\rangle_{\mathcal{S}_{\mathcal{U}}}$$

for certain $c_i\in[0,\frac{2}{p^*}]$. Hence, (4.6) is shown. □

Proof of Theorem 4.1. $\hat{\mathscr{C}}_{\mathscr{X};h}$ is an unbiased estimator for $\mathscr{C}_{\mathscr{X};h}$ with $|h|<\mathscr{M}_M$ due to its definition. Since $||\hat{\mathscr{C}}_{\mathscr{X};h}-\mathscr{C}_{\mathscr{X};h}||_{\mathcal{S}_{\mathcal{U}^m}}=||\hat{\mathscr{C}}_{\mathscr{X};-h}-\mathscr{C}_{\mathscr{X};-h}||_{\mathcal{S}_{\mathcal{U}^m}}$ for all h, let $h\ge0$ w.l.o.g. Herein, $h=h_M, m=m_M, \mathscr{M}=\mathscr{M}_M, \mathscr{M}_{M,h}=\mathscr{M}-|h|, (X_k\otimes X_l)':=X_k\otimes X_l-\mathscr{C}_{\mathbf{X};l-k}$ for $k,l\in\mathbb{Z}$ and $\xi_{i,n}:=1-\frac{|i|}{n}$ for $i\in\mathbb{Z}, n\in\mathbb{N}$. Stationarity of $\mathscr{X}$ after Assumption 4.1 (a), (4.10), conversions as in the proof of Lemma 4.1 and the Assumptions 4.2–4.4 (a) lead for h with $0\le h<\mathscr{M}$ to

$$\frac{\mathscr{M}_{M,h}}{m}\,\mathbb{E}||\hat{\mathscr{C}}_{\mathscr{X};h}-\mathscr{C}_{\mathscr{X};h}||^2_{\mathcal{S}_{\mathcal{U}^m}}$$
$$=\frac{\mathscr{M}_{M,h}}{m}\sum_{i,j=1}^{m}\mathbb{E}||\hat{\mathscr{C}}_{\mathbf{X};j-i+hp}-\mathscr{C}_{\mathbf{X};j-i+hp}||^2_{\mathcal{S}_{\mathcal{U}}}$$
$$=\mathscr{M}_{M,h}\sum_{|l|<m}\xi_{l,m}\,\mathbb{E}||\hat{\mathscr{C}}_{\mathbf{X};l+hp}-\mathscr{C}_{\mathbf{X};l+hp}||^2_{\mathcal{S}_{\mathcal{U}}}$$
$$=\frac{1}{\mathscr{M}_{M,h}}\sum_{|l|<m}\sum_{i,j=1}^{\mathscr{M}_{M,h}}\xi_{l,m}\,\mathbb{E}\langle(X_i\otimes X_{i+l+hp})',(X_j\otimes X_{j+l+hp})'\rangle_{\mathcal{S}_{\mathcal{U}}}$$
$$=\sum_{|k|<\mathscr{M}_{M,h}}\sum_{|l|<m}\xi_{k,\mathscr{M}_{M,h}}\xi_{l,m}\,\mathbb{E}\langle(X_0\otimes X_{l+hp})',(X_k\otimes X_{k+l+hp})'\rangle_{\mathcal{S}_{\mathcal{U}}} \tag{A.13}$$
$$\overset{M\to\infty}{\longrightarrow}\begin{cases}0, & \text{if } hp\to\infty,\\ \sum_{i\in\mathbb{Z}}\sum_{|j|<m^*}\xi_{j,m^*}\mathbb{E}\langle(X_0\otimes X_{i+c})',(X_j\otimes X_{i+j+c})'\rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } hp\to c\in\mathbb{Z}, m\to m^*,\\ \sum_{i,j\in\mathbb{Z}}\mathbb{E}\langle(X_0\otimes X_i)',(X_j\otimes X_{i+j})'\rangle_{\mathcal{S}_{\mathcal{U}}}, & \text{if } hp\to c\in\mathbb{Z}, m\to\infty,\end{cases}$$

where the expressions in the two latter cases are well-defined after Assumption 4.7 (a). Hence, due to the definition of $\tau_{2,\boldsymbol{X}}$ in (4.12), the assertion is proven. □

Proof of Theorem 4.2. Let $h = h_M, m = m_M, \mathscr{M} = \mathscr{M}_M, \tilde{\mathscr{M}} = \tilde{\mathscr{M}}_M, \mathscr{M}_{M,h} = \mathscr{M} - |h|, \tilde{\mathscr{M}}_{M,h} = \tilde{\mathscr{M}} - |h|$. From stationarity of $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and bilinearity of $\otimes \colon \mathcal{U}^m \times \mathcal{U}^m \to \mathcal{U}^m$ follows for h with $0 \le h < \mathscr{M}_M - 1$:

$$\begin{aligned}
\mathbb{E}(\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}) &= \frac{1}{\mathscr{M}_{M,h}-1}\sum_{k=1}^{\mathscr{M}_{M,h}} \mathbb{E}\Big(\big(\mathscr{X}_k - \frac{1}{\mathscr{M}_{M,h}}\sum_{i=1}^{\mathscr{M}_{M,h}}\mathscr{X}_i\big)\otimes\big(\mathscr{X}_{k+h}-\frac{1}{\mathscr{M}_{M,h}}\sum_{j=1}^{\mathscr{M}_{M,h}}\mathscr{X}_{j+h}\big)\Big)\\
&= \frac{1}{\mathscr{M}_{M,h}(\mathscr{M}_{M,h}-1)}\Big(\mathscr{M}_{M,h}^2\mathscr{C}_{\boldsymbol{\mathscr{X}};h} - \sum_{i,k=1}^{\mathscr{M}_{M,h}}\mathscr{C}_{\boldsymbol{\mathscr{X}};k+h-i}\Big)\\
&= \mathscr{C}_{\boldsymbol{\mathscr{X}};h} - \frac{1}{\mathscr{M}_{M,h}(\mathscr{M}_{M,h}-1)}\sum_{\substack{1\le i,k\le \mathscr{M}_{M,h}\\ i\neq k}}\mathscr{C}_{\boldsymbol{\mathscr{X}};k+h-i}. \qquad \text{(A.14)}
\end{aligned}$$

Hence, $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ for h with $0 \le h < \mathscr{M}_M - 1$ if the sum in (A.14) equals $0_{\mathcal{L}_{\mathcal{U}^m}}$, which can also be shown for h with $1 - \mathscr{M}_M < h < 0$. Now, we verify (4.14). Since $||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||_{\mathcal{S}_{\mathcal{U}^m}} = ||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};-h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};-h}||_{\mathcal{S}_{\mathcal{U}^m}}$ for all h, let $h \ge 0$ w.l.o.g. For $h < \mathscr{M}_M - 1$ holds

$$\begin{aligned}
\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} &= \frac{\mathscr{M}_{M,h}}{\mathscr{M}_{M,h}-1}(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}})\otimes(m_{\boldsymbol{\mathscr{X}}} - \hat{m}'_{\boldsymbol{\mathscr{X}}}) + \frac{1}{\mathscr{M}_{M,h}-1}\sum_{j=m}^{\tilde{\mathscr{M}}_{M,h}}\mathscr{U}_k\otimes\mathscr{U}_{k+h}\\
&= \frac{\mathscr{M}_{M,h}}{\mathscr{M}_{M,h}-1}\Big[(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}})\otimes(m_{\boldsymbol{\mathscr{X}}} - \hat{m}'_{\boldsymbol{\mathscr{X}}}) + \hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h}\Big] \qquad \text{(A.15)}
\end{aligned}$$

with $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h}$ as in (4.9) based on a sample $\mathscr{U}_m, \ldots, \mathscr{U}_{\tilde{\mathscr{M}}_{M,h}}$ of $\boldsymbol{\mathscr{U}} := (\mathscr{U}_k)_{k\in\mathbb{Z}}$ where $\mathscr{U}_k := \mathscr{X}_k - m_{\boldsymbol{\mathscr{X}}}$. (A.15), $\mathscr{C}_{\boldsymbol{\mathscr{X}};h} = \mathscr{C}_{\boldsymbol{\mathscr{U}};h}$, Δ-inequality, $(a+b+c)^2 \le 3(a^2+b^2+c^2)$ for $a, b, c \in \mathbb{R}$ and $||\boldsymbol{u}\otimes\boldsymbol{u}'||_{\mathcal{S}_{\mathcal{U}^m}} = ||\boldsymbol{u}||_{\mathcal{U}^m}||\boldsymbol{u}'||_{\mathcal{U}^m}$ for $\boldsymbol{u}, \boldsymbol{u}' \in \mathcal{U}^m$ yield

$$\begin{aligned}
&\mathbb{E}||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}}\\
&\quad= \mathbb{E}\Big|\Big|\frac{1}{\mathscr{M}_{M,h}-1}\Big[\mathscr{M}_{M,h}(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}})\otimes(m_{\boldsymbol{\mathscr{X}}} - \hat{m}'_{\boldsymbol{\mathscr{X}}}) + \hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h} - \mathscr{C}_{\boldsymbol{\mathscr{U}};h}\Big] + \mathscr{C}_{\boldsymbol{\mathscr{X}};h}\Big|\Big|^2_{\mathcal{S}_{\mathcal{U}^m}}\\
&\quad\le \frac{3}{(\mathscr{M}_{M,h}-1)^2}\Big[\mathscr{M}_{M,h}^2\,\mathbb{E}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m}||\hat{m}'_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m}\\
&\qquad\qquad + \mathscr{M}_{M,h}^2\,\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h} - \mathscr{C}_{\boldsymbol{\mathscr{U}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} + ||\mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}}\Big]\\
&\quad\le \frac{3}{(\mathscr{M}_{M,h}-1)^2}\Big[\mathscr{M}_{M,h}^2\,\mathbb{E}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^4_{\mathcal{U}^m}\\
&\qquad\qquad + \mathscr{M}_{M,h}^2\,\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h} - \mathscr{C}_{\boldsymbol{\mathscr{U}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} + m^2\mathbb{E}||X'_1||^4_{\mathcal{U}}\Big],
\end{aligned}$$

where $||\mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \le m^2\,\mathbb{E}||X'_1||^4_{\mathcal{U}}$ with $X'_1 := X_1 - m_{\boldsymbol{X}}$ follows from $||\cdot||_{\mathcal{S}_{\mathcal{U}^m}} \le ||\cdot||_{\mathcal{N}_{\mathcal{U}^m}}$, (3.4) and Cauchy-Schwarz inequality. From the inequalities above follows together with Lemma 4.1 (b), Theorem 4.1 and $m_M/\mathscr{M}_{M,h} \to 0$ for $M \to \infty$ indeed

$$\limsup_{M\to\infty} \frac{\mathscr{M}_{M,h}}{m_M} \mathbb{E}||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \leq 3\,[0 \cdot \eta_{4,\boldsymbol{X}} + \tau_{2,\boldsymbol{X}} + 0] \;=\; 3\tau_{2,\boldsymbol{X}}. \qquad \square$$

Proof of Theorem 4.3. $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ for h with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ by definition. Hereinafter, let $h \geq 0$ w.l.o.g. Further, we write $m = m_M, n = n_N, h = h_L, p = p_L, q = q_L$ where $L = L_{M,N} := M \wedge N$, $(X_k \otimes Y_l)' := X_k \otimes Y_l - \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};l-k}$ for $k, l \in \mathbb{Z}$ and $\xi_{i,n} := 1 - \frac{|i|}{n}$ for $i \in \mathbb{Z}, n \in \mathbb{N}$. Then, stationarity of $\boldsymbol{\mathscr{X}}$ and $\boldsymbol{\mathscr{Y}}$ after Assumption 4.1, (4.16) due to Assumption 4.10, Assumptions 4.2–4.4, (A.2) with the functions $\iota_{m,n}$ defined in (A.3)–(A.4) for any $m, n \in \mathbb{N}$, and $\tilde{\iota}_{m,n} := \iota_{m,n}/m \wedge n$, lead similarly to the proof of Theorem 4.1 to

$$\begin{aligned}
&\frac{\mathscr{L}_{M,N,h}}{m \wedge n} \mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} \\
&\sim \frac{\mathscr{L}_{M,N,h}}{m \wedge n} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{X},\boldsymbol{Y};j-i+(m+h-n)p} - \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};j-i+(m+h-n)p}||^2_{\mathcal{S}_{\mathcal{U},\mathcal{V}}} \\
&= \mathscr{L}_{M,N,h} \sum_{|l|<m\vee n} \tilde{\iota}_{m,n}(l)\, \mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{X},\boldsymbol{Y};l+(m+h-n)p} - \mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};l+(m+h-n)p}||^2_{\mathcal{S}_{\mathcal{U},\mathcal{V}}} \\
&= \sum_{|k|<\mathscr{L}_{M,N,h}} \sum_{|l|<m\vee n} \xi_{k,\mathscr{L}_{M,N,h}}\, \tilde{\iota}_{m,n}(l)\, \mathbb{E}\langle (X_0 \otimes Y_{l+(m+h-n)p})', (X_k \otimes Y_{k+l+(m+h-n)p})' \rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}} \\
&\overset{M\to\infty}{\longrightarrow} \begin{cases} 0, & \text{if } (m+h-n)p \to \pm\infty, \\ \sum_{i\in\mathbb{Z}} \sum_{|j|<m^*\vee n^*} \tilde{\iota}^*(j)\, \mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})' \rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to m^*, n \to n^*, \\ \sum_{i,j\in\mathbb{Z}} \tilde{\iota}^*(j)\, \mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})' \rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to \infty, n \to n^*, \\ \sum_{i,j\in\mathbb{Z}} \tilde{\iota}^*(j)\, \mathbb{E}\langle (X_0 \otimes Y_{i+c})', (X_j \otimes Y_{i+j+c})' \rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to m^*, n \to \infty, \\ \sum_{i,j\in\mathbb{Z}} \mathbb{E}\langle (X_0 \otimes Y_i)', (X_j \otimes Y_{i+j})' \rangle_{\mathcal{S}_{\mathcal{U},\mathcal{V}}}, & \text{if } (m+h-n)p \to c \in \mathbb{Z}, m \to \infty, n \to \infty. \end{cases}
\end{aligned}$$

These series exist after Assumption 4.7 (b) and since $|\tilde{\iota}^*(j)| < 1$ for all j, where the precise values $\tilde{\iota}^*(j) := \lim_{M\to\infty} \tilde{\iota}_{m,n}(j)$ are stated in (A.5) for any j. $\square$

Proof of Theorem 4.4. Herein, $h = h_L$ with $L = L_{M,N} = M \wedge N, m = m_M, \mathscr{M} = \mathscr{M}_M, \tilde{\mathscr{M}} = \tilde{\mathscr{M}}_M, n = n_N, \mathscr{N} = \mathscr{N}_N, \tilde{\mathscr{N}} = \tilde{\mathscr{N}}_N$. For h with $n - \tilde{\mathscr{M}} \leq h \leq \tilde{\mathscr{N}} - m$ holds

$$\mathbb{E}(\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}) = \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \frac{1}{\mathscr{L}_{M,N,h}(\mathscr{L}_{M,N,h} - 1)} \sum_{\substack{1\leq i,k\leq \mathscr{L}_{M,N,h} \\ i\neq k}} \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};k+h-i}$$

similar as in the proof of Theorem 4.2. Thus, $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ for these h if the sum above is $0_{\mathcal{L}_{\mathcal{U}^m,\mathcal{V}^n}}$. Moreover, as in Theorem 4.2,

$$\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \frac{\mathscr{L}_{M,N,h}}{\mathscr{L}_{M,N,h} - 1} \Big[(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}}) \otimes (m_{\boldsymbol{\mathscr{Y}}} - \hat{m}'_{\boldsymbol{\mathscr{Y}}}) + \hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h} \Big],$$

with $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h}$ defined in (4.15) based on samples $\mathscr{U}_m, \ldots, \mathscr{U}_{\tilde{\mathscr{M}}}$ of $\boldsymbol{\mathscr{U}} := (\mathscr{U}_k)_{k\in\mathbb{Z}}$ and $\mathscr{V}_n, \ldots, \mathscr{V}_{\tilde{\mathscr{N}}}$ of $\boldsymbol{\mathscr{V}} := (\mathscr{V}_k)_{k\in\mathbb{Z}}$ with $\mathscr{U}_k := \mathscr{X}_k - m_{\boldsymbol{\mathscr{X}}}$ resp. $\mathscr{V}_k := \mathscr{Y}_k - m_{\boldsymbol{\mathscr{Y}}}$.

Arguments in the proofs of Theorems 4.2–4.3 imply with (3.4), $\mathscr{C}_{\mathscr{U},\mathscr{V};h} = \mathscr{C}_{\mathscr{X},\mathscr{Y};h}$, with $X_1' := X_1 - m_{\boldsymbol{X}}$ as well as $Y_1' := Y_1 - m_{\boldsymbol{Y}}$, and $mn = (m \wedge n)(m \vee n)$:

$$
\begin{aligned}
&\frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{\mathscr{C}}'_{\mathscr{X},\mathscr{Y};h} - \mathscr{C}_{\mathscr{X},\mathscr{Y};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} \\
&\leq \frac{3\,\mathscr{L}^2_{M,N,h}}{(\mathscr{L}_{M,N,h}-1)^2}\Bigg[\frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^2_{\mathcal{U}^m}||\hat{m}'_{\mathscr{Y}} - m_{\mathscr{Y}}||^2_{\mathcal{V}^n} \\
&\qquad + \frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{\mathscr{C}}_{\mathscr{U},\mathscr{V};h} - \mathscr{C}_{\mathscr{U},\mathscr{V};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} + \frac{1}{\mathscr{L}_{M,N,h}(m \wedge n)}||\mathscr{C}_{\mathscr{X},\mathscr{Y};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}\Bigg] \\
&\leq \frac{3\,\mathscr{L}^2_{M,N,h}}{(\mathscr{L}_{M,N,h}-1)^2}\Bigg[\frac{\mathscr{L}_{M,N,h}}{m \wedge n}\sqrt{\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^4_{\mathcal{U}^m}\mathbb{E}||\hat{m}'_{\mathscr{Y}} - m_{\mathscr{Y}}||^4_{\mathcal{V}^n}} \\
&\qquad + \frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{\mathscr{C}}_{\mathscr{U},\mathscr{V};h} - \mathscr{C}_{\mathscr{U},\mathscr{V};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} + \frac{m \vee n}{\mathscr{L}_{M,N,h}}\sqrt{\mathbb{E}||X_1'||^4_{\mathcal{U}}\mathbb{E}||Y_1'||^4_{\mathcal{V}}}\Bigg], \\
&\sim 3\Bigg[\sqrt{\frac{\mathscr{L}^2_{M,N,h}}{\mathscr{M}^3\mathscr{N}^3}\frac{(m \vee n)m^{\mathbf{1}_P}n^{\mathbf{1}_Q}}{m \wedge n}\frac{\mathscr{M}^3}{m^{1+\mathbf{1}_P}}\,\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^4_{\mathcal{U}^m}\frac{\mathscr{N}^3}{n^{1+\mathbf{1}_Q}}\,\mathbb{E}||\hat{m}'_{\mathscr{Y}} - m_{\mathscr{Y}}||^4_{\mathcal{V}^n}} \\
&\qquad + \frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{\mathscr{C}}_{\mathscr{U},\mathscr{V};h} - \mathscr{C}_{\mathscr{U},\mathscr{V};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} + \frac{m \vee n}{\mathscr{L}_{M,N,h}}\sqrt{\mathbb{E}||X_1'||^4_{\mathcal{U}}\mathbb{E}||Y_1'||^4_{\mathcal{V}}}\Bigg].
\end{aligned}
\tag{A.16}
$$

Moreover, after Assumptions 4.4, 4.8, 4.9, 4.11 holds $\mathscr{L}_{M,N,h} \sim \tilde{\mathscr{L}}_{M,N,h} \sim \mathscr{M} \wedge \mathscr{N}$ for $M \to \infty$, and $m \vee n = \mathrm{o}(\mathscr{L}_{M,N,h}), m \vee n = \mathrm{o}(\mathscr{M} \wedge \mathscr{N}) \Rightarrow m \vee n = \mathrm{o}(\mathscr{M} \vee \mathscr{N})$ for $M \to \infty$. Thus, since $mn/m \wedge n = m \vee n$ and $\mathscr{M}\mathscr{N} = (\mathscr{M} \wedge \mathscr{N})(\mathscr{M} \vee \mathscr{N})$, we have

$$
\frac{\mathscr{L}^2_{M,N,h}}{\mathscr{M}^3\mathscr{N}^3}\frac{(m \vee n)mn}{m \wedge n} \sim \frac{(m \vee n)^2}{(\mathscr{M} \wedge \mathscr{N})(\mathscr{M} \vee \mathscr{N})^3} \overset{M\to\infty}{\longrightarrow} 0, \tag{A.17}
$$

and obviously also

$$
\frac{m \vee n}{\mathscr{L}_{M,N,h}} \overset{M\to\infty}{\longrightarrow} 0. \tag{A.18}
$$

Consequently, due to (A.17) and $m^{\mathbf{1}_P}n^{\mathbf{1}_Q} = \mathrm{O}(mn)$ for $M \to \infty$, (A.18), Lemma 4.1 (b), Theorem 4.3 and (A.16), holds

$$
\begin{aligned}
\limsup_{M\to\infty}\frac{\mathscr{L}_{M,N,h}}{m \wedge n}\,\mathbb{E}||\hat{\mathscr{C}}'_{\mathscr{X},\mathscr{Y};h} - \mathscr{C}_{\mathscr{X},\mathscr{Y};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} &\leq 3\big[0 \cdot \sqrt{\eta_{4,\boldsymbol{X}}\eta_{4,\boldsymbol{Y}}} + \tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})} + 0\big] \\
&= 3\tilde{\tau}_{2,(\boldsymbol{X},\boldsymbol{Y})}.
\end{aligned}
$$

Hence, the assertion is verified. □

Proof of Corollary 4.1. The statements follow from (4.22), $||\cdot||_{\mathcal{L}_{\mathcal{U}^m}} \leq ||\cdot||_{\mathcal{S}_{\mathcal{U}^m}}$ and Theorems 4.1–4.2 with $h = 0$. □

Proof of Corollary 4.2. The assertions are a consequence of (4.24) as well as Theorems 4.1–4.2 with $h = 0$, where (4.29) and (4.31) also include (4.27). □

Proof of Theorem 4.5. At first, we prove the results for the estimators $\breve{\mathfrak{c}}_j^{\ddagger}$. The triangle inequality leads to

$$\mathbb{E}||\breve{\mathfrak{c}}_j^{\ddagger} - \mathfrak{c}_j||_{\mathcal{U}^m}^2 \leq 2\big(\mathbb{E}||\breve{\mathfrak{c}}_j^{\ddagger} - \breve{\mathfrak{c}}_j||_{\mathcal{U}^m}^2 + \mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2\big). \tag{A.19}$$

Further, the definition of $\breve{\mathfrak{c}}_j^{\ddagger}$ in (4.34), $\breve{\mathfrak{c}}_j$ in (4.23), with $Z_j^{\dagger} = Z_j^{\dagger}(M) := \langle \breve{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m}$ and $Z_j = Z_j(M) := \langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}$, (4.36), $\operatorname{sgn}^2(x) = \mathbf{1}_{\mathbb{R}\setminus\{0\}}(x)$ for any $x \in \mathbb{R}$, the definition of the expected value and basic conversions imply

$$\begin{aligned}
\mathbb{E}||\breve{\mathfrak{c}}_j^{\ddagger} - \breve{\mathfrak{c}}_j||_{\mathcal{U}^m}^2 &= \mathbb{E}(\operatorname{sgn}(Z_j^{\dagger}) - \operatorname{sgn}(Z_j))^2 \\
&= 1 + \mathbb{P}(Z_j \neq 0) - 2\,\mathbb{E}\operatorname{sgn}(Z_j^{\dagger})\operatorname{sgn}(Z_j) \\
&= 1 + \mathbb{P}(Z_j \neq 0) + 2\big(\mathbb{P}(Z_j^{\dagger}Z_j < 0) - \mathbb{P}(Z_j^{\dagger}Z_j > 0)\big) \\
&\leq 2 + 2\big(1 - 2\,\mathbb{P}(Z_j^{\dagger}Z_j > 0)\big) \\
&= 4\big(1 - \mathbb{P}(Z_j^{\dagger}Z_j > 0)\big).
\end{aligned} \tag{A.20}$$

Moreover, the definition of $\breve{\mathfrak{c}}_j^{\dagger}$ in (4.33) yields

$$Z_j^{\dagger}Z_j = 1 - ||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + \frac{1}{4}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^4 + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2 Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i}, \tag{A.21}$$

where for the last term holds due to independence, $\mathbb{E}|Z_j| \leq 1$, (4.32), Cauchy-Schwarz inequality and the monotone convergence theorem:

$$\mathbb{E}\left|Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i}\right| = \mathbb{E}|Z_j| \sum_{i=1}^{\infty} \frac{|\langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}|\,\mathbb{E}|\zeta_i|}{2^i} \tag{A.22}$$

$$\leq \mu \sum_{i=1}^{\infty} \frac{1}{2^i} = \mu. \tag{A.23}$$

Furthermore, (A.21), Markov's inequality and (A.23) lead to

$$\begin{aligned}
1 - \mathbb{P}(Z_j^{\dagger}Z_j > 0) &\leq 1 - \mathbb{P}\left(1 - ||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2 Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i} > 0\right) \\
&= 1 - \mathbb{P}\left(||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 - \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2 Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i} < 1\right) \\
&\leq \mathbb{P}\left(||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2 \left|Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i}\right| \geq 1\right) \\
&\leq \mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2\, \mathbb{E}\left|Z_j \sum_{i=1}^{\infty} \frac{\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{2^i}\right| \\
&\leq \mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + \frac{m_M}{\mathscr{M}_M}\tilde{\gamma}_{j,m_M}^2 \mu.
\end{aligned} \tag{A.24}$$

Finally, plugging (A.24) into (A.20), and afterwards (A.20) into (A.19), leads together with (4.28) from Corollary 4.2 to

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{j,m_M}^{-2}\,\mathbb{E}||\check{\mathfrak{c}}_j^{\ddagger} - \check{\mathfrak{c}}_j||_{\mathcal{U}^m}^2 \le 10\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{j,m_M}^{-2}\,\mathbb{E}||\check{\mathfrak{c}}_j - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + 8\mu$$
$$\le 10\,\tau_{2,\boldsymbol{X}} + 8\mu.$$

Thus, (4.37) is verified. The same conversions imply with (4.27) and (4.29):

$$\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{k_M,m_M}^{-2}\,\mathbb{E}\sup_{j\le k_M}||\check{\mathfrak{c}}_j^{\ddagger} - \mathfrak{c}_j||_{\mathcal{U}^m}^2$$
$$\le 10\limsup_{M\to\infty} \frac{\mathscr{M}_M}{m_M}\tilde{\gamma}_{k_M,m_M}^{-2}\,\mathbb{E}\sup_{j\le k_M}||\check{\mathfrak{c}}_j^{\ddagger} - \mathfrak{c}_j||_{\mathcal{U}^m}^2 + 8\mu\limsup_{M\to\infty}\tilde{\gamma}_{k_M,m_M}^{-2}\sup_{j\le k_M}\tilde{\gamma}_{j,m_M}^2$$
$$\le 10\,\tau_{2,\boldsymbol{X}} + 8\mu,$$

hence (4.38) also holds. At last, (4.39) and (4.38) can be shown analogously. □

Proof of Lemma 5.1. $Z_N - \tilde{Z}_N = A^N(Z_0 - \tilde{Z}_0)$ for all $N \in \mathbb{N}$. Further, $\mathbb{E}||Z_0 - \tilde{Z}_0||_{\mathcal{H}}^{\nu} < \infty$ since $(Z_k)_k, (\tilde{Z}_k)_k$ are $L_{\mathcal{H}}^{\nu}$-processes because $(\varepsilon_k)_k$ is one. Thus, due to submultiplicity of the operator norm, one obtains with $\rho := ||A||_{\mathcal{L}_{\mathcal{H}}}^{-\nu}$,

$$\rho^N\mathbb{E}||Z_N - \tilde{Z}_N||_{\mathcal{H}}^{\nu} \le \mathbb{E}||Z_0 - \tilde{Z}_0||_{\mathcal{H}}^{\nu}, \quad \forall N \in \mathbb{N}.$$ □

Appendix B: Side results

Lemma B.1. *Let Assumption 4.1 hold, and let $(X_k)_{k\in\mathbb{Z}}$ and $(Y_k)_{k\in\mathbb{Z}}$ be $L_{\mathcal{U}}^4$-m-resp. $L_{\mathcal{V}}^4$-m-approximable. Furthermore, let $\mathscr{X}_{m+j;l} := (X_{m+jp;l},\dots,X_{1+jp;l})^T$ and $\mathscr{Y}_{n+j;l} := (Y_{n+jq;l},\dots,Y_{1+jq;l})^T$ for any j,l,m,n,p,q.*

(a) *The processes $(\mathscr{X}_k)_{k\in\mathbb{Z}}$ and $(\mathscr{Y}_k)_{k\in\mathbb{Z}}$ satisfy*

$$\frac{1}{\sqrt{m}}\sum_{k=1}^{\infty}\nu_{4,\mathcal{U}^m}(\mathscr{X}_k - \mathscr{X}_{k;k}) \le \sum_{k=1}^{\infty}\nu_{4,\mathcal{U}}(X_k - X_{k;k}) < \infty, \tag{B.1}$$
$$\frac{1}{\sqrt{n}}\sum_{k=1}^{\infty}\nu_{4,\mathcal{V}^n}(\mathscr{Y}_k - \mathscr{Y}_{k;k}) \le \sum_{k=1}^{\infty}\nu_{4,\mathcal{V}}(Y_k - Y_{k;k}) < \infty. \tag{B.2}$$

Furthermore, $(\mathscr{X}_k)_k$ is $L_{\mathcal{U}^m}^4$-m-approximable for fixed $m \in \mathbb{N}$ if $p = 1$, and $(\mathscr{Y}_k)_k$ is $L_{\mathcal{V}^n}^4$-m-approximable for fixed $n \in \mathbb{N}$ if $q = 1$.

(b) *The process $(\mathscr{W}_{k,h})_{k\in\mathbb{Z}}$, where $\mathscr{W}_{k,h} := \mathscr{X}_k \otimes \mathscr{Y}_{k+h}$ with $h \in \mathbb{Z}$, fulfills with $\mathscr{W}_{k,h;l} := \mathscr{X}_{k;l} \otimes \mathscr{Y}_{k+h;l}$:*

$$\frac{1}{\sqrt{mn}}\sum_{k=1}^{\infty}\nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{W}_{k,h} - \mathscr{W}_{k,h;k})$$
$$\le \sum_{k=1}^{\infty}\nu_{4,\mathcal{V}}(Y_1)\nu_{4,\mathcal{U}}(X_k - X_{k;k}) + \nu_{4,\mathcal{U}}(X_1)\nu_{4,\mathcal{V}}(Y_k - Y_{k;k}), \tag{B.3}$$

and $(\mathscr{W}_{k,h})_k$ is $L_{\mathcal{U}^m\mathcal{V}^n}^2$-m-approximable for fixed m,n if $h \le 0$ and $p = q = 1$.

Proof. (a) The definition of $\mathscr{X}_k, \mathscr{X}_{k;k}, \mathscr{Y}_k, \mathscr{Y}_{k;k}$ for all k implies $\nu_{4,\mathcal{U}^m}(\mathscr{X}_k - \mathscr{X}_{k;k}) \leq \sqrt{m}\, \nu_{4,\mathcal{U}}(X_k - X_{k;k})$ and $\nu_{4,\mathcal{V}^n}(\mathscr{Y}_k - \mathscr{Y}_{k;k}) \leq \sqrt{n}\, \nu_{4,\mathcal{V}}(Y_k - Y_{k;k})$, thus (B.1), (B.2). Hence, since $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are causal w.r.t. $(\varepsilon_k)_k$ for fixed $m \in \mathbb{N}$ if $p = 1$ resp. for fixed $n \in \mathbb{N}$ if $q = 1$, $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are then also $L^4_{\mathcal{U}^m}$-m-resp. $L^4_{\mathcal{V}^n}$-m-approximable.

(b) Bilinearity of $\otimes\colon \mathcal{U}^m \times \mathcal{V}^n \to \mathcal{V}^n$, Minkowski inequality, $||\boldsymbol{u} \otimes \boldsymbol{v}||_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} = ||\boldsymbol{u}||_{\mathcal{U}^m} ||\boldsymbol{v}||_{\mathcal{V}^n}$ for $\boldsymbol{u} \in \mathcal{U}^m, \boldsymbol{v} \in \mathcal{V}^n$, Cauchy-Schwarz inequality and (A.2) yield

$$
\begin{aligned}
&\frac{1}{\sqrt{mn}} \sum_{k=1}^{\infty} \nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{W}_{k,h} - \mathscr{W}_{k,h;k}) \\
&\leq \frac{1}{\sqrt{mn}} \sum_{k=1}^{\infty} \nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}((\mathscr{X}_k - \mathscr{X}_{k;k}) \otimes \mathscr{Y}_{k+h}) + \nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{X}_k \otimes (\mathscr{Y}_{k+h} - \mathscr{Y}_{k+h;k})) \\
&\leq \frac{1}{\sqrt{mn}} \sum_{k=1}^{\infty} \nu_{4,\mathcal{U}^m}(\mathscr{X}_k - \mathscr{X}_{k;k}) \nu_{4,\mathcal{V}^n}(\mathscr{Y}_1) + \nu_{4,\mathcal{U}^m}(\mathscr{X}_1) \nu_{4,\mathcal{V}^n}(\mathscr{Y}_k - \mathscr{Y}_{k;k}) \\
&\leq \sum_{k=1}^{\infty} \nu_{4,\mathcal{V}}(Y_1) \nu_{4,\mathcal{U}}(X_k - X_{k;k}) + \nu_{4,\mathcal{U}}(X_1) \nu_{4,\mathcal{V}}(Y_k - Y_{k;k}).
\end{aligned}
$$

This term is finite due to L^4-m-approximability of $(X_k)_k$ and $(Y_k)_k$. Moreover, since $(\mathscr{X}_k)_k, (\mathscr{Y}_{k+h})_k$ and thus $(\mathscr{W}_{k,h})_k$ are causal w.r.t. $(\varepsilon_k)_k$ for $h \leq 0$ for fixed $m, n \in \mathbb{N}$ if $p = q = 1$, $(\mathscr{W}_{k,h})_k$ is indeed $L^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}$-$m$-approximable. □

Acknowledgments

My thanks goes to the editor, the associate editor and the two referees for their careful reading, raised questions and thoughtful suggestions. I also thank Alexander Meister (University of Rostock), Gregory Rice (University of Waterloo), Albrecht Brehm, Philipp Otto (Leibniz University Hannover) and Siegfried Hörmann (Graz University of Technology) for valuable conversations, Piotr Kokoszka (Colorado State University) and Dominik Liebl (University of Bonn) for helpful general comments, and Leo Evans for proof-reading.

References

[1] Allam A. & Mourid T. (2019) Optimal rate for covariance operator estimators of functional autoregressive processes with random coefficients. *J. Multivariate Anal.*, **169**, 130–137. MR3875591

[2] Aue A. & van Delft A. (2020) Testing for stationarity of functional time series in the frequency domain. *Ann. Statist.*, **48**(5), 2505–2547. MR4152111

[3] Aue A. & Klepsch J. (2017) Estimating functional time series by moving average model fitting. arXiv:1701.00770v1. MR3823558

[4] Aue A., Norinho D.D. & Hörmann S. (2015) On the prediction of stationary functional time series. *J. Amer. Statist. Assoc.*, **110**(509), 378–392. MR3338510

[5] Benth F., Di Nunno, G. & Schroers D. (2021) Copula measures and sklar's theorem in arbitrary dimensions. *Scand. J. Stat.* doi: 10.1111/sjos.12559. MR4471282

[6] Berkes I., Horváth L. & Rice G. (2016) On the asymptotic normality of kernel estimators of the long run covariance of functional time series. *J. Multivariate Anal.*, **144**, 150–175. MR3434947

[7] Bosq D. (2000) *Linear Processes in Function Spaces*. Lecture Notes in Statistics, 149. New York: Springer. MR1783138

[8] Buckner et al. (2004) A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage*, **23**, 724–738.

[9] Caponera A. (2021) SPHARMA approximations for stationary functional time series on the sphere. *Stat. Inference Stoch. Process.*, **24**(3), 609–634. MR4321852

[10] Caponera A., Fageot J., Simeoni M. & Panaretos V.M. (2022+) Functional estimation of anisotropic covariance and autocovariance operators on the sphere. arXiv:2112.12694v2

[11] Caponera A. & Marinucci D. (2021) Asymptotics for spherical functional autoregressions. *Ann. Statist.*, **49**(1), 346–369. MR4206681

[12] Caponera A. & Panaretos V.M. (2022+) On the rate of convergence for the autocorrelation operator in functional autoregression. *Stat. Prob. Lett.* doi: 10.1016/j.spl.2022.109575. MR4449609

[13] Chaouch M. (2014) Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves. *IEEE Trans. Smart Grid*, **5**(1), 411–419.

[14] Chen Y., Koch T., Lim K.G., Xu X. & Zakiyeva N. (2021) A review study of functional autoregressive models with application to energy forecasting. *Wiley Interdiscip. Rev. Comput. Stat.*, **3**, e1525, 23 pp. MR4242811

[15] van Delft A., Characiejus V. & Dette H. (2021) A nonparametric test for stationarity in functional time series. *Statist. Sinica* **31**(3), 1375–1395. MR4297704

[16] van Delft A. Dette H. (2021) A similarity measure for second order properties of non-stationary functional time series with applications to clustering and testing. *Bernoulli*, **27**(1), 469–501. MR4177377

[17] van Delft, A. & Eichler M. (2018). Locally stationary functional time series. *Electron. J. Statist.*, **12**(1), 107–170. MR3746979

[18] Dunford N. & Schwartz J.T. (1988) *Linear Operators, Part I: General Theory*. New York: John Wiley & Sons Inc. MR1009162

[19] Ferraty F. & Vieu P. (2006) *Nonparametric Functional Data Analysis*. New York: Springer. MR2229687

[20] Gao Y., Shang H.L. & Yang Y. (2019) High-dimensional functional time series forecasting: an application to age-specific mortality rates. *J. Multivariate Anal.*, **170**, 232–243. MR3913038

[21] Gohberg I., Goldberg S. & Kaashoek M.A. (2003). *Basic Classes of Linear Operators* (1 ed.). Basel: Birkhäuser Verlag. MR2015498

[22] Górecki T., Hörmann S., Horváth L. & Kokoszka P. (2018) Testing normality of functional time series. *J. Time Series Anal.*, **39**(4), 471–487. MR3819053
[23] Guillas S. (2001) Rates of convergence of autocorrelation estimates for autoregressive Hilbertian processes. *Statist. Probab. Lett.*, **55**(3), 281–291. MR1867531
[24] Hael M.A. (2021) Modeling of rainfall variability using functional principal component method: A case study of Taiz region, Yemen. *Model. Earth Syst. Environ.*, **7**, 17–27.
[25] Hashemi M., Zamani A. & Haghbin H. (2019) Rates of convergence of autocorrelation estimates for periodically correlated autoregressive Hilbertian processes. *Statistics*, **53**(2), 283–300. MR3916630
[26] Hörmann S., Horváth L. & Reeder R. (2013) A functional version of the ARCH model. *Econ. Theory*, **29**(2), 267–288. MR3042756
[27] Hörmann S. & Kokoszka P. (2010) Weakly dependent functional data. *Ann. Statist.*, **38**, 1845–1884. MR2662361
[28] Horváth L. & Kokoszka P. (2012). *Inference for Functional Data with Applications*. New York: Springer. MR2920735
[29] Horváth L., Kokoszka P. & Rice G. (2014) Testing stationarity of functional time series. *J. Econometrics*, **179**(1), 66–82. MR3153649
[30] Horváth L., Rice G. & Zhao Y. (2022). Change point analysis of covariance functions: A weighted cumulative sum approach. *J. Multivariate Anal.*, **189**, paper no. 104877, 23 pp. MR4384119
[31] Hsing T. & Eubank R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. West Sussex: Wiley. MR3379106
[32] Huang S.-F., Guo M. & Chen M.-R. (2020) Stock market trend prediction using a functional time series approach. *Quant. Finance*, **20**(1), 69–79. MR4040263
[33] Jarry G., Delahaye D., Nicol F. & Feron E. (2020) Aircraft atypical approach detection using functional principal component analysis. *J. A. T. M.*, **84**, 101787.
[34] Klepsch J., Klüppelberg C. & Wei T. (2017) Prediction of functional ARMA processes with an application to traffic data. *Econom. Stat.*, **1**, 128–149. MR3669993
[35] Kokoszka P. & Reimherr M. (2013) Asymptotic normality of the principal components of functional time series. *Stoch. Process. Appl.*, **123**(5), 1546–1562. MR3027890
[36] Kokoszka P., Rice G. & Shang H.L. (2017) Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *J. Multivariate Anal.*, **162**, 32–50. MR3719333
[37] Kokoszka P., Stoev S. & Xiong Q. (2019) Principal components analysis of regularly varying functions. *Bernoulli*, **25**(4B), 3864–3882. MR4010975

[38] Kühnert S. (2019) *Über funktionale ARCH- und GARCH-Zeitreihen* (German) [About functional ARCH and GARCH processes]. Doctoral Thesis, University of Rostock. URL: https://doi.org/10.18453/rosdok_id00002507

[39] Kühnert S. (2020) Functional ARCH and GARCH models: A Yule-Walker approach. *Electron. J. Statist.*, **14**(2), 4321–4360. MR4187136

[40] Ledoux M. & Talagrand M. (2011) *Probability in Banach Spaces. Isoperimetry and Processes* (Reprint of 1 ed.). Classics in Mathematics. Berlin: Springer. MR2814399

[41] Li C., Xiao L. & Luo S. (2020) Fast covariance estimation for multivariate sparse functional data. *Stat*, **9**, e245, 18 pp. MR4116315

[42] Lung T., Peters M.K., Farwig N., Böhning-Gaese K. & Schaab G. (2012) Combining long-term land cover time series and field observations for spatially explicit predictions on changes in tropical forest biodiversity. *Int. J. Remote Sens.*, **33**:1340.

[43] Martínez-Hernández I. & Genton M.G. (2020) Recent developments in complex and spatially correlated functional data. *Braz. J. Probab. Stat.*, **34**(2), 204–229. MR4093256

[44] Mas A. (2007) Weak convergence in the functional autoregressive model. *J. Multivariate Anal.*, **98**(6), 1231–1261. MR2326249

[45] Pilavakis D., Paparoditis E. & Sapatinas T. (2020) Testing equality of autocovariance operators for functional time series. *J. Time Series Anal.*, **41**(4), 665–692. MR4136903

[46] Rice G. & Shum M. (2019) Inference for the lagged cross-covariance operator between functional time series. *J. Time Series Anal.*, **40**(5), 665–692. MR3995663

[47] Rice G., Wirjanto T. & Zhao Y. (2020) Forecasting value at risk with intraday return curves. *Int. J. Forecast.*, **36**(3), 1023–1038.

[48] Ramsay J.O. & Silverman B.W. (2005) *Functional Data Analysis* (2 ed.). Springer Series in Statistics. New York: Springer. MR2168993

[49] Ruiz-Medina M.D. & Álvarez-Liébana J. (2019) Strongly consistent autoregressive predictors in abstract Banach spaces. *J. Multivariate Anal.*, **170**, 186–201. MR3913035

[50] Ruiz-Medina M.D. & Espejo R.M. (2012) Spatial autoregressive functional plug-in prediction of ocean surface temperature. *Stoch. Environ. Res. Risk. Assess.*, **26**, 335–344.

[51] Sarkar S. & Panaretos V.M. (2021) CovNet: Covariance networks for functional data on multidimensional domains. arXiv:2104.05021v2

[52] Sen R. & Klüppelberg C. (2019) Time series of functional data with application to yield curves. *Appl. Stoch. Models Bus. Ind.*, **35**(4), 1028–1043. MR3994372

[53] Sharipov O.Sh. & Wendler M. (2020) Bootstrapping covariance operators of functional time series. *J. Nonparametr. Stat.*, **32**(3), 648–666. MR4136586

[54] Spangenberg F. (2013) Strictly stationary solutions of ARMA equations in Banach spaces. *J. Multivariate Anal.*, **121**, 127–138. MR3090473

[55] Stout W.F. (1974). *Almost Sure Convergence.* New York: Academic Press.

MR0455094

[56] Tian T.S. (2010) Functional data analysis in brain imaging studies. *Front. Psychol.*, **1**, 1–11.

[57] Weidmann J. (1980). *Linear Operators in Hilbert Spaces*. Volume 68 of Graduate Texts in Mathematics. Berlin, New York: Springer. MR0566954

[58] Werner D. (2018) *Funktionalanalysis* (German) [Functional Analysis] (8 ed.). Berlin: Springer Spektrum. MR1787146

[59] Wong R.K.W. & Zhang X. (2019). Nonparametric operator-regularized covariance function estimation for functional data. *Comput. Statist. Data Anal.*, **131**, 131–144. MR3906800

[60] Xiao L., Li C., Checkley W. and Crainiceanu C. (2018) Fast covariance estimation for sparse functional data. *Stat. Comput.*, **28**(3), 511–522. MR3761337

[61] Yu F., Liu L., Yu N., Ji L. & Qiu D. (2020) A method of L1-norm principal component analysis for functional data. *Symmetry*, **12**(1), 182. MR2703298