Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

Lagged Covariance and Cross-Covariance Operators of Processes in Cartesian Products of Abstract Hilbert Spaces

Sebastian Kühnert*

September 17, 2021

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 deduces estimators for lagged covariance and cross-covariance operators of processes of abstract Hilbert spaces, and in particular of processes in Cartesian products of Hilbert spaces, obtained by successively stacking Hilbert space-valued elements. Our main focus is on these estimator's asymptotic properties for fixed and increasing lag and Cartesian powers. The processes are allowed to be non-centered, and to have values in different spaces when investigating the dependence between processes. We also discuss features of estimators for the principal components of our covariance operators for fixed and increasing Cartesian power.

MSC 2010 subject classifications: 47B10, 60G05, 62J10

Keywords: Estimation; functional time series; lag-h-covariance operator; lag-h-cross covariance operator; principal components

1 Introduction

Functional Data Analysis (FDA) and *Functional Time Series Analysis* (FTSA), the research areas dealing with time series resp. processes of random functions, have gained more and more significance, since considering random functions instead of vectors, provided the context allows it, assures more accurate results. Such an extension on infinite-dimensional spaces is enabled by ongoing developments in processing techniques, and is unproblematic for separable Banach spaces from a mathematical point of view, see Ledoux & Talagrand [29]. FDA/FTSA find applications in economics [8; 9; 21; 34; 39], medicine [5; 42] and other research areas [3; 12; 30]. For extensive introductions to FDA/FTSA, see Ferraty & Vieu [11], Ramsay & Silverman [35], Bosq [4], Horváth & Kokoszka [17], and Hsing & Eubank [20]. In FTSA, the analysis of the dependence structure within and between given processes is of great importance. If these are *wide-sense/weak/second-order stationary*, where for convenience usually *(strictly) stationarity* and finite second moments are assumed, this can be done by using *lag-h-covariance operators* and *lag-h-cross-covariance operators*, respectively, where the *lag h* denotes the time difference of interest. Another important subject of study is *Functional Principal Component Analysis* (FPCA), see [15; 22], since *functional principal components*, i.e., the eigenvalues and eigenfunctions of the (lag-0-)covariance operator, yield an efficient representation.

Probabilistic features of and estimators for lag-h-covariance operators $\mathscr{C}_{\boldsymbol{X};h}$ of stationary processes $\boldsymbol{X} = (X_k)_{k\in\mathbb{Z}}$ with values in $L^2[0,1]$, the Hilbert space of measurable, square-Lebesgue integrable real valued functions with domain $[0,1]$, are widely studied for fixed lag h, see, e.g., [4; 17; 20; 25; 32]. Further, [36] developed covariance estimators in the space of continuous functions $C[0,1]$, [44] in tensor product Sobolev-Hilbert spaces, [31] for continuous surfaces, and [1; 16] for quite general Hilbert spaces. [1; 16; 32; 36] constrained their assertions to autoregressive (AR) processes, where [1] deduced the results for a random AR(1) operator. Thereby,

*EMAIL: s.kuehnert_math@gmx.de

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

[1; 4; 17; 20] utilized classical moment estimators, [25] estimated the integral kernels, in [16; 32] truncated spectral decompositions occured having estimated principal components, and [44] used operator regularized covariance estimators. Also, the limit distribution of the estimation errors of the lag-0-covariance operators was discussed in [24; 26]. FPCA in $L^2[0,1]$ is extensively discussed in the existing literature, both from a probabilistic and statistical point of view. In [4; 17; 20; 24; 26] one finds asymptotic upper bounds for the principal components, estimated both seperately and uniformly in sense of convergence in second mean and almost surely (a.s.), and [45] introduced L^1-norm FPCA. A comprehensive study of *lag-h-cross-covariance operators* $\mathscr{C}_{\boldsymbol{X},\boldsymbol{Y};h}$ of $L^2[0,1]$-valued processes $\boldsymbol{X} = (X_k)_{k\in\mathbb{Z}}, \boldsymbol{Y} = (Y_k)_{k\in\mathbb{Z}}$ can be found in Rice & Shum [33]. They established operator estimates, discussed methods measuring their significance and deduced their limit distribution. Aue & Klepsch [2], who investigated estimators for operators of linear, invertible processes in $L^2[0,1]$, had to estimate lag-h-cross-covariance operators of specific processes in Cartesian products of $L^2[0,1]$ to derive their main results. Enabling processes to have values in Cartesian products is also handy when studying AR(p) processes with $p > 1$, see [4]. Also, the quite recent work of Sarkar & Panaretos [38] extensively dealt with covariance estimation of functional data defined over multidimensional domains.

This article studies, inspired by assertions in [2; 33] and also [27; 28], lagged covariance and cross-covariance operators of stationary processes in separable Hilbert spaces, especially of processes in Cartesian products of Hilbert spaces, where the processes are created by successive stacking Hilbert space-valued elements. The focus is on deducing moment estimators for lagged covariance and cross-covariance operators, and to derive asymptotic upper bounds of their estimation errors. FPCA of our covariance operators is also conducted, where the principal components are estimated separately and uniformly. Particularly worth mentioning is that this work's results facilitate a high degree of flexibility. This is because all processes are allowed to attain values in Cartesian products of quite general Hilbert spaces, the processes not necessarily have to be centered, the lag h and the processes' Cartesian powers can be fixed or increase w.r.t. the sample sizes, and the definition of the lagged cross-covariance operators allows both processes to attain values in different spaces. Moreover, as an example of the use of our models, one could think of investors of, e.g., solar or other power stocks of European companies asking theirselves what impact monthly sunshine duration in central Europe, see Fig 1, has on their share values one month ahead, see Fig 2. This can be analyzed using lag-1-cross-covariance operators, and lag-h-covariance operators might be advantageous for understanding the dependence structure within the processes in Fig 1-2.

The rest of this paper is organized as follows. Section 2 outlines our notation, restates important terminology, definitions and interrelationships of several operator types, defines our (lagged) (cross-)covariance operators and studies their probabilistic features, and briefly explains L^p-m-approximibility. Section 3 introduces our estimators for the lagged (cross-)covariance operators and for the principal components of the lag-0-covariance operator, and derives asymptotic upper bounds of the estimation errors. Section 4 conducts a simulation study. Section 5 summarizes the main results and outlines future research. Moreover, Section 6 contains proofs.

2 Definitions and basics

2.1 Notation

$\lfloor\cdot\rfloor$ denotes the *floor function*, $\operatorname{sgn}(\cdot)$ the *sign function* and $\mathbf{1}_A(\cdot)$ the *indicator function* of a set A. For sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \subseteq (0,\infty)$, $a_n \sim b_n$ denotes $\frac{a_n}{b_n} \to 1$, $a_n \asymp b_n$ denotes $\frac{a_n}{b_n} \to c$ for some $c \neq 0$, (for $n \to \infty$) $a_n = \omega(b_n)$ if $b_n = \mathrm{o}(a_n)$ and $a_n = \Omega(b_n)$ if $b_n = \mathrm{O}(a_n)$ with common asymptotic notation $\mathrm{o}(\cdot), \mathrm{O}(\cdot)$, and $\Xi[a_n, b_n) \coloneqq \Omega(a_n) \cap \mathrm{o}(b_n), \Xi[a_n, b_n] \coloneqq \Omega(a_n) \cap \mathrm{O}(b_n)$. 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 the norms are assumed to be induced by their inner product. Herein, let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ be a real separable Hilbert space. For $x, y \in \mathcal{H}, x \perp y$ means $\langle x, y\rangle_{\mathcal{H}} = 0$. Scalar multiplication and vector addition on $\mathcal{H}^n \coloneqq \{(x_1, ..., x_n)^T | x_1, ..., x_n \in \mathcal{H}\}$, with $n \in \mathbb{N}$, is defined componentwise, so $(\mathcal{H}^n, \langle\cdot,\cdot\rangle_{\mathcal{H}^n})$ with $\langle \boldsymbol{x}, \boldsymbol{y}\rangle_{\mathcal{H}^n} \coloneqq \sum_{i=1}^n \langle x_i, y_i\rangle_{\mathcal{H}}$ for $\boldsymbol{x} \coloneqq (x_1, ..., x_n)^T$, $\boldsymbol{y} \coloneqq (y_1, ..., y_n)^T \in \mathcal{H}^n$, is a real separable Hilbert space. Our random elements are defined on some common probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. $X \overset{d}{=} Y$ stands for equally distributed random variables X, Y. For processes $(X_n)_n$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

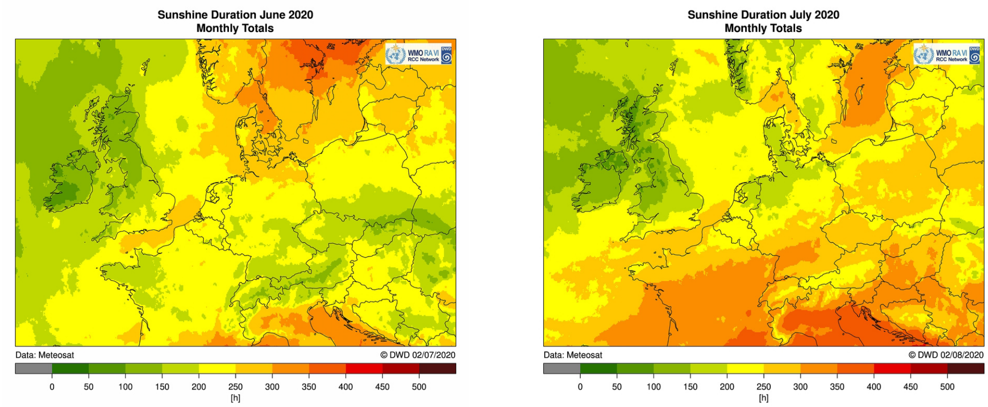

Figure 1: Graphs of monthly sunshine duration in central europe in June and July 2020, interpretable as two consecutive realizations of an $L^2[0,1]^2$-valued process, from the homepage www.dwd.de of the German Meterological Service.

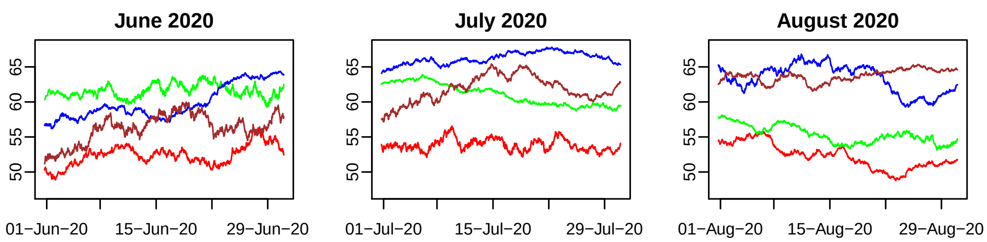

Figure 2: Three consecutive realizations of a fictitious $(L^2[0,1])^4$-process, describing the share values of four assets of an portfolio, e.g., measured in *EUR*. The step width used is $\frac{1}{1000}$.

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* $\nu_{p,\mathcal{H}}(X) := (\mathbb{E}||X||^p_{\mathcal{H}})^{1/p} < \infty$, a process $(X_k)_{k\in\mathbb{Z}}$ of $\mathcal{H}$-valued random variables is an $L^p_{\mathcal{H}}$*-process* if $X_k \in L^p_{\mathcal{H}}$ for all k, and *centered* if $\mathbb{E}(X_k) = 0_{\mathcal{H}}$ for all k with expectation in Bochner-integral sense, see [20], p. 40–45.

2.2 Some basic operator theory

Now, we state important spaces of operators between real separable Hilbert spaces $(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i})$ for $i = 1,2$. For thorough overviews of operators between Hilbert spaces, see the monographs Dunford & Schwartz [10], Gohberg *et al.* [13], Weidmann [43]. 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

$$||A||_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}} := \sup_{||x||_{\mathcal{H}_1}\le 1} ||A(x)||_{\mathcal{H}_2} < \infty.$$

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 $\mathcal{F}_{\mathcal{H}_1} := \mathcal{F}_{\mathcal{H}_1,\mathcal{H}_1}$. Further, A^* denotes the adjoint of $A \in \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}$, where $A^* \in \mathcal{L}_{\mathcal{H}_2,\mathcal{H}_1}$. 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 (CONS) of $\mathcal{H}_1$ resp. $\mathcal{H}_2$, and where $(s_j)_{j\in\mathbb{N}}$ is the monotonically decreasing zero sequence of non-negative numbers of A, the *singular values*. Their decay rate can be interpreted as a regularity measure of A and be expressed by the *p-Schatten-norm*

$$||A||_p := \Big(\sum_{j=1}^\infty s_j^p\Big)^{1/p}, \quad p \in [1,\infty),$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

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 $\mathscr{S}^p_{\mathcal{H}_1,\mathcal{H}_2} := \{A \in \mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2} \mid ||A||_p < \infty\}$ 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*. The *trace* of $A \in \mathcal{N}_{\mathcal{H}_1}$ is defined by $\operatorname{tr}(A) := \sum_{j=1}^{\infty} \langle A(e_j), e_j \rangle_{\mathcal{H}_1}$, 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

$$\langle A, B \rangle_{\mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2}} := \sum_{j=1}^{\infty} \langle A(e_j), B(e_j) \rangle_{\mathcal{H}_2}, \quad A, B \in \mathcal{S}_{\mathcal{H}_1,\mathcal{H}_2},$$

where $(e_j)_{j\in\mathbb{N}}$ is an arbitrary CONS of $\mathcal{H}_1$ in both definitions. On $\mathcal{H}_1 := L^2[0,1]$, an *integral operator* A mapping from $\mathcal{H}_1$ to $\mathcal{H}_1$ is defined by the Lebesgue integral

$$(A(x))(t) := \int_0^1 a(s,t)x(s)\,\mathrm{d}s, \quad x \in \mathcal{H}_1, t \in [0,1]$$

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 $A \in \mathcal{S}_{\mathcal{H}_1}$ iff $\int_0^1\int_0^1 a^2(s,t)\,\mathrm{d}s\,\mathrm{d}t < \infty$.

2.3 Features of our operators

Here, we define (cross-)covariance operators and their lagged versions on real separable Hilbert spaces, and outline some of their features (see [4] for these operators on Banach spaces). Thereby, $(\mathcal{H}_i, \langle\cdot,\cdot\rangle_{\mathcal{H}_i})$ denote real separable Hilbert spaces for $i = 1, 2$.

Definition 2.1. *Let X, Y be $L^2_{\mathcal{H}_1}$- resp. $L^2_{\mathcal{H}_2}$-valued random variables, and let $m_X := \mathbb{E}(X), m_Y := \mathbb{E}(Y)$. Then, the covariance operator of X is defined by*

$$\mathscr{C}_X := \mathbb{E}((X - m_X) \otimes (X - m_X)),$$

and the cross-covariance operator of X, Y is defined by

$$\mathscr{C}_{X,Y} := \mathbb{E}((X - m_X) \otimes (Y - m_Y)).$$

For centered $L^2_{\mathcal{H}_1}$- resp. $L^2_{\mathcal{H}_2}$-valued random variables X, Y, where centeredness is due to $X' := X - \mathbb{E}(X), Y' := Y - \mathbb{E}(Y)$ no restriction, (cross-)covariance operators possess the following features, see [4] and also [20], sections 7.2-7.3. Firstly, $\mathscr{C}_X \in \mathcal{N}_{\mathcal{H}_1}$ is a self-adjoint, positive semi-definite operator with

$$||\mathscr{C}_X||_{\mathcal{N}_{\mathcal{H}_1}} = \mathbb{E}||X||^2_{\mathcal{H}_1}, \tag{2.1}$$

and satisfies the representation

$$\mathscr{C}_X = \sum_{j=1}^{\infty} c_j (\mathfrak{c}_j \otimes \mathfrak{c}_j), \tag{2.2}$$

where $(c_j)_{j\in\mathbb{N}}$ is the 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 $\mathscr{C}_X$ being a CONS of $\mathcal{H}_1$. The cross-covariance operator satisfies $\mathscr{C}_{X,Y} \in \mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2}, \mathscr{C}^*_{X,Y} = \mathscr{C}_{Y,X} \in \mathcal{N}_{\mathcal{H}_2,\mathcal{H}_1}$ and

$$||\mathscr{C}_{X,Y}||_{\mathcal{N}_{\mathcal{H}_1,\mathcal{H}_2}} = ||\mathscr{C}_{Y,X}||_{\mathcal{N}_{\mathcal{H}_2,\mathcal{H}_1}} \leq \mathbb{E}||X||_{\mathcal{H}_1}||Y||_{\mathcal{H}_2}. \tag{2.3}$$

Furthermore,

$$\text{independence of } X, Y \;\Rightarrow\; \mathscr{C}_{X,Y} = 0_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}}, \tag{2.4}$$

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 $k_X(s,t) := \operatorname{Cov}(X(s), X(t))$ resp. $k_{X,Y}(s,t) := \operatorname{Cov}(X(s), Y(t)), s,t \in [0,1]$. As for covariances of real-valued random variables, holds for the covariance operator given centered $L^2_{\mathcal{H}_1}$-valued random variables W, X, and $A \in \mathcal{L}_{\mathcal{H}_1}$,

$$\mathscr{C}_{W+X} = \mathscr{C}_W + \mathscr{C}_{W,X} + \mathscr{C}_{X,W} + \mathscr{C}_X, \tag{2.5}$$

$$\mathscr{C}_{A(X)} = A\mathscr{C}_X A^*, \tag{2.6}$$

and for the cross-covariance operator given centered $L^2_{\mathcal{H}_1}$- resp. $L^2_{\mathcal{H}_2}$-valued random variables W, X resp. Y, Z, with $A \in \mathcal{L}_{\mathcal{H}_1}$ and $B \in \mathcal{L}_{\mathcal{H}_2}$ holds

$$\mathscr{C}_{W+X,Y+Z} = \mathscr{C}_{W,Y} + \mathscr{C}_{W,Z} + \mathscr{C}_{X,Y} + \mathscr{C}_{X,Z}, \tag{2.7}$$

$$\mathscr{C}_{A(X),B(Y)} = B\mathscr{C}_{X,Y} A^*. \tag{2.8}$$

For the definition of the functional counterparts of the auto-covariance and cross-covariance function of real-valued processes, the lag-h-covariance resp. lag-h-cross-covariance operators, given processes not necessarily have to be strictly, but wide-sense stationary. To reiterate, a process $(X_k)_k$ is *(strictly) stationary* if $(X_{k_1+h}, ..., X_{k_n+h}) \overset{d}{=} (X_{k_1}, ..., X_{k_n})$ for all $k_1, ..., k_n, h$ and $n \in \mathbb{N}$, and an $L^2_{\mathcal{H}_1}$-process $\boldsymbol{X} := (X_k)_k$ is *wide-sense stationary* if $\mathbb{E}(X_k) = c$ 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.

Definition 2.2. *Let $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}, \boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$ be wide-sense stationary $L^2_{\mathcal{H}_1}$- resp. $L^2_{\mathcal{H}_2}$-processes. Then, the lag-h-covariance operator of $\boldsymbol{X}$ is defined by*

$$\mathscr{C}_{\boldsymbol{X};h} := \mathscr{C}_{X_0,X_h}, \quad \forall h \in \mathbb{Z},$$

with $\mathscr{C}_{\boldsymbol{X}} := \mathscr{C}_{\boldsymbol{X};0}$, and the lag-h-cross-covariance operator of $\boldsymbol{X}, \boldsymbol{Y}$ is defined by

$$\mathscr{C}_{\boldsymbol{X}\boldsymbol{Y};h} := \mathscr{C}_{X_0,Y_h}, \quad \forall h \in \mathbb{Z}.$$

Remarks 2.1. The features for (cross-)covariance apply to lag-h-(cross-)covariance operators. Thus, $\mathscr{C}^*_{\boldsymbol{X};h} = \mathscr{C}_{\boldsymbol{X};-h}$ and $\mathscr{C}^*_{\boldsymbol{X}\boldsymbol{Y};h} = \mathscr{C}_{\boldsymbol{Y},\boldsymbol{X};-h}$ for any h, $\mathscr{C}_{\boldsymbol{X};h} = 0_{\mathcal{L}_{\mathcal{H}_1}}$ for $h \neq 0$ if $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}$ consists of independent variables, if $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}, \boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$ are independent, $\mathscr{C}_{\boldsymbol{X}\boldsymbol{Y};h} = 0_{\mathcal{L}_{\mathcal{H}_1,\mathcal{H}_2}}$. Further, if $\mathcal{H}_1 = \mathcal{H}_2 = L^2[0,1]$, the lag-$h$-(cross-)covariance are integral operators with the integral kernel being the auto-covariance resp. cross-covariance function. Hence, the expression '(cross-)covariance' in lag-h-(cross-)covariance operator is reasonable too.

Now, we illustrate a specific lag-0-covariance operator. For further, but somewhat more complicated examples and sketches, see Section 4.

Example 2.1. Let $\mathcal{H} := L^2[0,1]$, and let $\boldsymbol{\varepsilon} := (\varepsilon_k)_{k\in\mathbb{Z}}$ be a process with

$$\varepsilon_k(t) := \frac{Z_k + B_k(t)}{\sqrt{1+t}} \text{ a.s.}, \quad \forall k \in \mathbb{Z},\ \forall t \in [0,1], \tag{2.9}$$

where $Z_k \sim \mathcal{N}(0,1), \boldsymbol{B_k} = (B_k(t))_{t\in[0,1]}$ are Wiener processes, and where $\ldots, Z_{-1}, \boldsymbol{B_{-1}}, Z_0, \boldsymbol{B_0}, Z_1, \boldsymbol{B_1}, \ldots$ 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

$$k_{\boldsymbol{\varepsilon}}(s,t) = \operatorname{Cov}(\varepsilon_0(s), \varepsilon_0(t)) = \sqrt{\frac{1+\min(s,t)}{1+\max(s,t)}}, \quad \forall s,t \in [0,1]. \tag{2.10}$$

2.4 L^p-m-approximibility

For deriving asymptotic upper bounds of estimation errors for operators or functionals related to a stationary process, usually weak dependence of the given process is needed. We impose *L^p-m-approximibility* developed by Hörmann & Kokoszka [19], since this type of weak dependence is, due to its definition based on m-dependence, particularly feasible for transformations when verifying asymptotic upper bounds (see for instance (6.5)–(6.7) in the proof of Theorem 3.1).

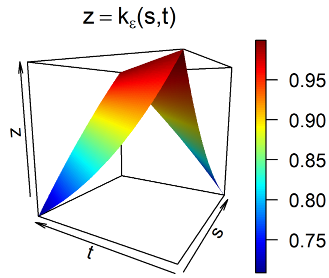

Figure 3: The integral kernel $k_\varepsilon(s,t)$ in (2.10) for $s,t \in [0,1]$.

Definition 2.3. *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*

$$Z_k = f(\varepsilon_k, \varepsilon_{k-1}, ...), \quad \forall k \in \mathbb{Z}, \tag{2.11}$$

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

$$Z_{k;m} := f(\varepsilon_k, \varepsilon_{k-1}, ..., \varepsilon_{k-m+1}, \varepsilon_{k-m;k}, \varepsilon_{k-m-1;k}, ...), \tag{2.12}$$

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

$L^p_{\mathcal{H}}$-m-approximibility of a process thus means it is *non-anticipative* w.r.t. another process, that is (2.11), 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, (2.11) yields stationarity of $(Z_k)_k$ due to [41], 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.

3 Main results

Herein, we discuss the main results of this paper, namely the estimation procedure for lag-h-covariance and lag-h-cross-covariance operators of $\mathcal{U}^m$- resp. $\mathcal{V}^n$-valued processes for $m,n \in \mathbb{N}$, and for the principal components of lag-0-covariance operators. Thereby, $(\mathcal{U}^m, \langle\cdot,\cdot\rangle_{\mathcal{U}^m})$ 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 $(\mathcal{V}, \langle\cdot,\cdot\rangle_{\mathcal{V}})$. Throughout this article, we assume that our processes have the following representations.

Assumption 3.1. (a) Let $\boldsymbol{X} := (X_k)_{k\in\mathbb{Z}}$ be an $L^4_{\mathcal{U}}$-m-approximable process and let $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ be an $\mathcal{U}^m$-valued process with

$$\mathscr{X}_{m+j} := (X_{m+jp}, ..., X_{1+jp})^T, \quad \forall j \in \mathbb{Z}, \tag{3.1}$$

and some $p \in \mathbb{N}$. Further, $X_1, ..., X_M$ is a sample of $\boldsymbol{X}$ with $M \geq m$, thus $\mathscr{X}_m, ..., \mathscr{X}_{\tilde{\mathscr{M}}}$ 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) Let $\boldsymbol{Y} := (Y_k)_{k\in\mathbb{Z}}$ be an $L^4_{\mathcal{V}}$-m-approximable process and non-anticipative w.r.t.the same i.i.d. process $(\varepsilon_k)_k$ as $\boldsymbol{X}$ in (a), and let $\boldsymbol{\mathscr{Y}} := (\mathscr{Y}_k)_{k\in\mathbb{Z}}$ be an $\mathcal{V}^n$-valued process with

$$\mathscr{Y}_{n+j} := (Y_{n+jq}, ..., Y_{1+jq})^T, \quad \forall j \in \mathbb{Z}, \tag{3.2}$$

and some $q \in \mathbb{N}$. Moreover, $Y_1, ..., Y_N$ with $N \geq n$ is a sample of $\boldsymbol{Y}$, thus $\mathscr{Y}_n, ..., \mathscr{Y}_{\tilde{\mathscr{N}}}$ 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 $\mathscr{N} = \mathscr{N}_N := \tilde{\mathscr{N}}_N - n + 1$.

Remarks 3.1. (a) $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are stationary processes since $(X_k)_k$ resp. $(Y_k)_k$ are due to L^4-m-approximibility, and $(\mathscr{X}_k)_k$ resp. $(\mathscr{Y}_k)_k$ are i.i.d. for $p \geq m$ resp. $q \geq n$ if $(X_k)_k$ resp. $(Y_k)_k$ are.

(b) The common case is $m = n = 1$ which can be seen as a specific example of our results.

(c) Using two not necessarily equal sample sizes M and N is beneficial if more data of one process can be collected, but one is not willing to relinquish information by choosing the minimum of M, N.

(d) Choosing p, q so that $1 \leq p < m, 1 \leq q < n$, enables to reuse entries of $\mathscr{X}_k$ resp. $\mathscr{Y}_k$. If one needs to successively stack (realizations of) $X_k's$ resp. $Y_k's$, one has to put $p = m, q = n$. One could also choose $p > m, q > n$, which is useful when data is missing, since by an appropriate choice of p and q, the time indices where (realizations of) $X_k's$ and/or $Y_k's$ are missing can be bridged.

(e) Whenever large sample sizes of $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are needed, and reusing values of $(X_k)_k$ resp. $(Y_k)_k$ does not cause issues, p and q should be set as small as possible, to wit $p = q = 1$. Thereby, processes $(\mathscr{X}_k)_k$ as in (3.1) with $p = 1$ were used for estimating the operators of $L^2[0,1]$-valued AR in [4], (G)ARCH in [28], and invertible linear processes in [2; 28].

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

Assumption 3.2. (a) $(m_M)_M, (p_M)_M \subseteq \mathbb{N}$ are sequences with $m = m_M, p = p_M = \Xi[1, M) = \Omega(1) \cap \mathrm{o}(M)$ where $a_n = \Omega(b_n)$ means $b_n = \mathrm{O}(a_n)$.

(b) $(n_N)_N, (q_N)_N \subseteq \mathbb{N}$ are sequences with $n = n_N, q = q_N = \Xi[1, N)$.

From the Assumptions 3.1-3.2 (a) and (b) follows

$$\mathscr{M} = \mathscr{M}_M \sim p^{-1}M, \tag{3.3}$$

$$\text{resp.} \quad \mathscr{N} = \mathscr{N}_N \sim q^{-1}N. \tag{3.4}$$

Since the time difference where some variable has a certain effect on another one could also change over time respectively the sample size, we also allow the lag $h \in \mathbb{Z}$ to vary w.r.t. given sample sizes as follows.

Assumption 3.3. (a) $h = h_M = \Xi[1, p^{-1}M)$;

(b) $h = h_N = \Xi[1, q^{-1}N)$.

Lemma 3.1. *Let Assumptions 3.1-3.2 (a) hold. Then,*

$$\hat{m}_{\boldsymbol{\mathscr{X}}} := \frac{1}{\mathscr{M}_M} \sum_{i=m}^{\tilde{\mathscr{M}}_M} \mathscr{X}_i$$

is an unbiased estimator for the first moment $m_{\boldsymbol{\mathscr{X}}} := \mathbb{E}(\mathscr{X}_1)$, *and*

$$\mathbb{E}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m} = \mathrm{O}(m\mathscr{M}_M^{-1}) = \mathrm{O}(mpM^{-1}) \quad \textit{for } M \to \infty.$$

Remarks 3.2. Lemma 3.1 can be generalized to higher moments if the related power of the random variable of the process is well-defined on the given Hilbert space. Powers of random variables are for instance well-defined on $L^2[0,1]$. There, X^2 denotes the pointwise product of X and X, X^3 e.g. the pointwise product of X^2 and X, X^4 e.g. the pointwise product of X^3 and X etc.

3.1 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 first moment. If the process $\boldsymbol{X} = (X_k)_k$ in Assumption 3.1 (a) is centered, hence also $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$, we estimate $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ with $|h| < \mathscr{M}_M$ by

$$\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} := \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{k=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, & h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{k=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, & h \geq 0, \end{cases} \tag{3.5}$$

where $\mathscr{M}_{M,h} := \mathscr{M}_M - |h|, \tilde{\mathscr{M}}_{M,h} := \tilde{\mathscr{M}}_M - |h|$. These operator estimates satisfy $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} \in \mathcal{F}_{\mathcal{U}^m}$ (meaning they are finite-rank operators) with $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} = \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};-h}$, and $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}}} := \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};0}$ is self-adjoint and positive semi-definite.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

Theorem 3.1. *Let Assumptions 3.1-3.3 (a) hold, and let $\boldsymbol{X}$ be centered. Then, $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ with $|h| < \mathscr{M}_M$, and*

$$\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \mathrm{O}((1+|h|)m^2 pM^{-1}) \quad \textit{for } M \to \infty. \tag{3.6}$$

If the first moment $m_{\boldsymbol{X}}$ of $\boldsymbol{X} = (X_k)_k$ is unknown, thus also $m_{\boldsymbol{\mathscr{X}}} = (m_{\boldsymbol{X}}, ..., m_{\boldsymbol{X}})^T \in \mathcal{U}^m$, we use

$$\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} := \begin{cases} \frac{1}{\mathscr{M}_{M,h}-1} \sum_{k=m+|h|}^{\tilde{\mathscr{M}}_M} (\mathscr{X}_k - \hat{m}_{\boldsymbol{\mathscr{X}}}) \otimes (\mathscr{X}_{k+h} - \hat{m}'_{\boldsymbol{\mathscr{X}}}), & h < 0, \\ \frac{1}{\mathscr{M}_{M,h}-1} \sum_{k=m}^{\tilde{\mathscr{M}}_{M,h}} (\mathscr{X}_k - \hat{m}_{\boldsymbol{\mathscr{X}}}) \otimes (\mathscr{X}_{k+h} - \hat{m}'_{\boldsymbol{\mathscr{X}}}), & h \geq 0, \end{cases} \tag{3.7}$$

to estimate $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$, provided $|h| < \mathscr{M}_M - 1$, where the moment estimators are defined by

$$\hat{m}_{\boldsymbol{\mathscr{X}}} := \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{i=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_i, & h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{i=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_i, & h \geq 0, \end{cases} \qquad \hat{m}'_{\boldsymbol{\mathscr{X}}} := \begin{cases} \frac{1}{\mathscr{M}_{M,h}} \sum_{j=m+|h|}^{\tilde{\mathscr{M}}_M} \mathscr{X}_{j+h}, & h < 0, \\ \frac{1}{\mathscr{M}_{M,h}} \sum_{j=m}^{\tilde{\mathscr{M}}_{M,h}} \mathscr{X}_{j+h}, & h \geq 0. \end{cases}$$

Thereby, $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} \in \mathcal{F}_{\mathcal{U}^m}$ with $\hat{\mathscr{C}}'^*_{\boldsymbol{\mathscr{X}};h} = \hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};-h}$, and $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}}} := \hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};0}$ is self-adjoint and positive semi-definite.

Theorem 3.2. *Under Assumptions 3.1-3.3 (a), $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}$ 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*

$$||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \mathrm{O}_{\mathbb{P}}((1+|h|)m^2 pM^{-1}) \quad \textit{for } M \to \infty. \tag{3.8}$$

Remarks 3.3. (a) Theorems 3.1-3.2 extend the existing literature regarding the estimation of (lagged) covariance operators in several ways, see e.g. [1; 4; 14; 16; 17; 19; 24; 27; 28]. This is because the upper bounds in both Theorems are derived for lagged covariance operators of processes with arbitrary first moments having values in general, separable Hilbert spaces, and since the Cartesian power m and simultaneously the lag h is allowed to grow w.r.t. the sample size M.

(b) Using $\frac{1}{\mathscr{M}_{M,h}-1}$ instead of $\frac{1}{\mathscr{M}_{M,h}}$ in (3.7) enables to formulate the sufficient condition for unbiasedness in Theorem 3.2. For $h = 0$, this condition holds if $\mathscr{C}_{\boldsymbol{X};h} = 0_{\mathcal{L}_{\mathcal{U}}}$ for all $h \neq 0$ which is due to (2.4) particularly the case if $\boldsymbol{X}$ is a process of i.i.d. random variables.

(c) In Theorem 3.2, convergence in probability instead of in mean was considered since in the proof reciprocals of eigenvalues emerge.

3.2 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 the processes $\boldsymbol{X} = (X_k)_k$ and $\boldsymbol{Y} = (Y_k)_k$ in Assumption 3.1 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

$$\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{k=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{X}_k \otimes \mathscr{Y}_{k+h}, \tag{3.9}$$

with $\tilde{l}_{m,n,h} := \max(m, n-h)$, $\tilde{\mathscr{L}}_{M,N,h} := \min(\tilde{\mathscr{M}}_M, \tilde{\mathscr{N}}_N - h)$ and $\mathscr{L}_{M,N,h} := \tilde{\mathscr{L}}_{M,N,h} + 1 - \tilde{l}_{m,n,h}$. Further, $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} \in \mathcal{F}_{\mathcal{U}^m,\mathcal{V}^n}$ and $\hat{\mathscr{C}}^*_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \hat{\mathscr{C}}_{\boldsymbol{\mathscr{Y}},\boldsymbol{\mathscr{X}};-h}$, and when estimating $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ by $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ we impose the following.

Assumption 3.4. The sequences in Assumptions 3.1-3.3 satisfy $\tilde{l}_{m,n,h} = \mathrm{o}(\tilde{\mathscr{L}}_{M,N,h})$.

Theorem 3.3. *Let Assumptions 3.1-3.4 hold, and let $\boldsymbol{X}, \boldsymbol{Y}$ be centered. 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 $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$, and for $h = h_L$ with $L = L_{M,N} := \min(M, N), m = m_M, n = n_N$ holds*

$$\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} = \mathrm{O}\big((1+|h|)mn\mathscr{L}^{-1}_{M,N,h}\big) \quad \textit{for } M, N \to \infty. \tag{3.10}$$

If $m_{\boldsymbol{X}}$ and/or $m_{\boldsymbol{Y}}$ in Assumption 3.1 are unknown, thus also $m_{\boldsymbol{\mathscr{X}}} = (m_{\boldsymbol{X}}, ..., m_{\boldsymbol{X}})^T \in \mathcal{U}^m$ and/or $m_{\boldsymbol{\mathscr{Y}}} = (m_{\boldsymbol{Y}}, ..., m_{\boldsymbol{Y}})^T \in \mathcal{V}^n$, $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ is estimated by

$$\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} := \frac{1}{\mathscr{L}_{M,N,h} - 1} \sum_{k=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} (\mathscr{X}_k - \hat{m}_{\boldsymbol{\mathscr{X}}}) \otimes (\mathscr{Y}_{k+h} - \hat{m}'_{\boldsymbol{\mathscr{Y}}}) \tag{3.11}$$

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

$$\hat{m}_{\boldsymbol{\mathscr{X}}} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{i=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{X}_i, \qquad \hat{m}'_{\boldsymbol{\mathscr{Y}}} := \frac{1}{\mathscr{L}_{M,N,h}} \sum_{j=\tilde{l}_{m,n,h}}^{\tilde{\mathscr{L}}_{M,N,h}} \mathscr{Y}_{j+h}. \tag{3.12}$$

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.

Theorem 3.4. *Under Assumptions 3.1-3.4, $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ if $\sum_{1\leq i,k\leq \mathscr{L}_{M,N,h}, i\neq k} \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};k+h-i} = 0_{\mathcal{L}_{\mathcal{U}^m,\mathcal{V}^n}}$, and for $h = h_L$ with $L = L_{M,N} := \min(M,N), m = m_M$ and $n = n_N$ holds*

$$||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} = \mathrm{O}_{\mathbb{P}}\big((1+|h|)mn\mathscr{L}^{-1}_{M,N,h}\big) \quad \textit{for } M,N \to \infty. \tag{3.13}$$

Remarks 3.4. (a) Although estimating (lagged) cross-covariance operators is widely discussed, see e.g. [2; 4; 16; 33], Theorems 3.3-3.4 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 3.1-3.2 for the lagged covariance operators, derived for centered and for not necessarily centered processes, the lag h is allowed to be both fixed and varying w.r.t. the sample sizes, as are the Cartesian powers m, n.

(b) The initial and final value of the sum in (3.9) and (3.11) are relatively complicated, since $\mathscr{X}_k$ and $\mathscr{Y}_{k+h}$ simultaneously have to be well-defined.

(c) Assumption 3.4 ensures that the upper bounds (3.10) and (3.13) are zero sequences if the sequences in Assumptions 3.1-3.3 are chosen appropriately.

(d) By following the lines in the proof of Theorem 3.4, it becomes clear that omitting the estimation of $\mathscr{X}_k$ resp. $\mathscr{Y}_{k+h}$ in (3.12) if $\boldsymbol{X}$ is centered and $m_{\boldsymbol{Y}}$ is unknown resp. if $m_{\boldsymbol{X}}$ is unknown and $\boldsymbol{Y}$ is centered, has no positive effect on the convergence rate (3.13) in Theorem 3.4.

3.3 Estimation of principal components

Herein, we examine the estimation procedure of the principal components of lag-0-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}}} = \mathscr{C}_{\boldsymbol{\mathscr{X}};0}$ of the $\mathcal{U}^m$-valued processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_{k\in\mathbb{Z}}$ in Assumption 3.1 (a). Thereby, $(\mathfrak{c}_j)_{j\in\mathbb{N}}, (\hat{\mathfrak{c}}_j)_{j\in\mathbb{N}}$ 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 $\mathscr{C}_{\boldsymbol{\mathscr{X}}}, \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}}} = \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};0}$ in (3.5) resp. $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}}} = \hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};0}$ in (3.7). Also, since the vector lengths m of the elements of $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ can vary w.r.t. M, we occasionally write $c_j = c_{j,m}$ and $\mathfrak{c}_j = \mathfrak{c}_{j,m}$.

At first, due to [4], Lemma 4.2, for any $j \in \mathbb{N}$ holds

$$|\hat{c}_j - c_j| \leq ||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}}} - \mathscr{C}_{\boldsymbol{\mathscr{X}}}||_{\mathcal{L}_{\mathcal{U}^m}}, \qquad |\hat{c}'_j - c_j| \leq ||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}}} - \mathscr{C}_{\boldsymbol{\mathscr{X}}}||_{\mathcal{L}_{\mathcal{U}^m}}. \tag{3.14}$$

Corollary 3.1. *Let Assumptions 3.1-3.2 (a) hold. Then,*

$$\sup_{j\in\mathbb{N}} (\hat{c}'_j - c_j)^2 = \mathrm{O}_{\mathbb{P}}(m^2pM^{-1}) \quad \textit{for } M \to \infty,$$

and if $\boldsymbol{X} = (X_k)_k$ in Assumption 3.1 (a) is centered,

$$\mathbb{E}\big(\sup_{j\in\mathbb{N}} (\hat{c}_j - c_j)^2\big) = \mathrm{O}(m^2pM^{-1}) \quad \textit{for } M \to \infty.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

We proceed with estimating the eigenfunctions $\mathfrak{c}_j$ of $\mathscr{C}_{\mathscr{X}}$ by $\hat{\mathfrak{c}}_j$ if $\mathscr{X}$ is centered and by $\hat{\mathfrak{c}}'_j$ if the first moment of $\mathscr{X}$ is unknown. Eigenfunctions are unambiguously determined except for their sign, why

$$\check{\mathfrak{c}}_j := \operatorname{sgn}\langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}_j \qquad \text{resp.} \qquad \check{\mathfrak{c}}'_j := \operatorname{sgn}\langle \hat{\mathfrak{c}}'_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}'_j \tag{3.15}$$

are used as estimators for $\mathfrak{c}_j$, where 'sgn' is the sign function.

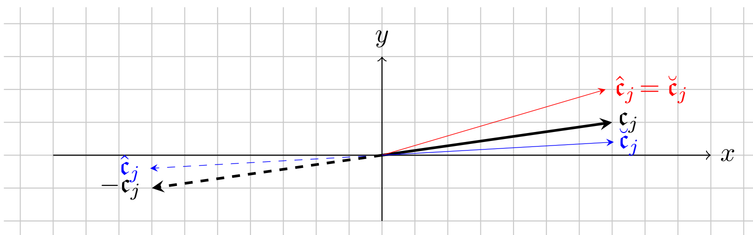

Figure 4: Estimation of $\mathfrak{c}_j$ by $\check{\mathfrak{c}}_j$, exemplified in $\mathbb{R}^2$

However, using $\check{\mathfrak{c}}_j$ resp. $\check{\mathfrak{c}}'_j$ to estimate $\mathfrak{c}_j$ is problematic, since $\check{\mathfrak{c}}_j \not\perp \mathfrak{c}_j$ a.s. and $\check{\mathfrak{c}}'_j \not\perp \mathfrak{c}_j$ a.s., thus $\operatorname{sgn}\langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. resp. $\operatorname{sgn}\langle \hat{\mathfrak{c}}'_j, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. is not guaranteed for fixed j, M. So, if $\check{\mathfrak{c}}_j \perp \mathfrak{c}_j$ resp. $\check{\mathfrak{c}}'_j \perp \mathfrak{c}_j$, one cannot allocate a unique estimator for $\mathfrak{c}_j$. This feature, though, 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 [27; 28]. We bypass this problem by modifying $\check{\mathfrak{c}}_j$ and $\check{\mathfrak{c}}'_j$. 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 i.i.d., $\mathcal{N}(0,1)$-distributed random variables, independent of the observations of $\boldsymbol{X}$. Then, for all j, M,

$$\check{\mathfrak{c}}_j^{\dagger} := \hat{\mathfrak{c}}_j + \sum_{i=1}^{\infty} \frac{\zeta_i u_i}{i^2 M} \qquad \text{and} \qquad \check{\mathfrak{c}}_j^{\prime\dagger} := \hat{\mathfrak{c}}'_j + \sum_{i=1}^{\infty} \frac{\zeta_i u_i}{i^2 M} \tag{3.16}$$

are well-defined with $\check{\mathfrak{c}}_j^{\dagger} \not\perp \mathfrak{c}_j$ a.s. resp. $\check{\mathfrak{c}}_j^{\prime\dagger} \not\perp \mathfrak{c}_j$ a.s., thus $\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. and $\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \neq 0$ a.s. Hence, we estimate $\mathfrak{c}_j$ by

$$\check{\mathfrak{c}}_j^{\ddagger} := \Big[\mathbf{1}_{\mathbb{R}\setminus\{0\}}(\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} + \mathbf{1}_{\{0\}}(\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m})\Big] \hat{\mathfrak{c}}_j, \tag{3.17}$$

$$\check{\mathfrak{c}}_j^{\prime\ddagger} := \Big[\mathbf{1}_{\mathbb{R}\setminus\{0\}}(\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m}) \operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} + \mathbf{1}_{\{0\}}(\operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m})\Big] \hat{\mathfrak{c}}'_j, \tag{3.18}$$

where $\mathbf{1}_A(\cdot)$ stands for the indicator function of a set A. These estimators satisfy

$$\check{\mathfrak{c}}_j^{\ddagger} = \operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}_j \ \text{ a.s.} \qquad \text{resp.} \qquad \check{\mathfrak{c}}_j^{\prime\ddagger} = \operatorname{sgn}\langle \check{\mathfrak{c}}_j^{\prime\dagger}, \mathfrak{c}_j \rangle_{\mathcal{U}^m} \hat{\mathfrak{c}}'_j \ \text{ a.s.} \tag{3.19}$$

To state upper bounds of estimation errors when using the estimators in (3.16), following technical preliminaries are needed. According to [4], Lemma 4.3 holds

$$\|\check{\mathfrak{c}}_j - \mathfrak{c}_j\|_{\mathcal{U}^m} \leq \tilde{\gamma}_j \|\hat{\mathscr{C}}_{\mathscr{X}} - \mathscr{C}_{\mathscr{X}}\|_{\mathcal{L}_{\mathcal{U}^m}}, \quad \|\check{\mathfrak{c}}'_j - \mathfrak{c}_j\|_{\mathcal{U}^m} \leq \tilde{\gamma}_j \|\hat{\mathscr{C}}'_{\mathscr{X}} - \mathscr{C}_{\mathscr{X}}\|_{\mathcal{L}_{\mathcal{U}^m}}, \quad \forall j \in \mathbb{N}, \tag{3.20}$$

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

$$\gamma_j := (c_j - c_{j+1})^{-1}, \quad j \in \mathbb{N}. \tag{3.21}$$

Assumption 3.5. $\mathscr{C}_{\mathscr{X}}$ is injective, and the eigenvalues of $\mathscr{C}_{\mathscr{X}}$ satisfy $c_j \neq c_{j+1}$ 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 3.5 holds both

$$c_1 > c_2 > \cdots > 0, \tag{3.22}$$

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

$$\sup_{j \leq k} \tilde{\gamma}_j = \gamma_k. \tag{3.23}$$

Lemma 3.2. *Let Assumptions 3.1-3.2 (a), 3.5 hold. Further, let $(k_j)_j \subseteq \mathbb{N}$ be a sequence with $k = k_M = \Omega(1)$, and let $\gamma_{j,m} := 1/(c_{j,m} - c_{j+1,m})$. Then,*

$$||\breve{\mathfrak{c}}'_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}_{\mathbb{P}}(\gamma^2_{j,m} m^2 p M^{-1}) \quad \text{for } M \to \infty, \quad \forall j \in \mathbb{N}, \tag{3.24}$$

$$\sup_{j \le k} ||\breve{\mathfrak{c}}'_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}_{\mathbb{P}}(\gamma^2_{k,m} m^2 p M^{-1}) \quad \text{for } M \to \infty, \tag{3.25}$$

and if $\boldsymbol{X} = (X_k)_k$ in Assumption 3.1 (a) is centered,

$$\mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}(\gamma^2_{j,m} m^2 p M^{-1}) \quad \text{for } M \to \infty, \quad \forall j \in \mathbb{N}, \tag{3.26}$$

$$\mathbb{E}\big(\sup_{j \le k} ||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m}\big) = \mathrm{O}(\gamma^2_{k,m} m^2 p M^{-1}) \quad \text{for } M \to \infty. \tag{3.27}$$

These statements for the 'classical' estimators (3.15) also apply to our advanced estimators (3.17), (3.18).

Theorem 3.5. *Let Assumptions 3.1-3.2 (a), 3.5 hold, and let $(k_j)_j \subseteq \mathbb{N}$ be a sequence with $k = k_M = \Omega(1)$. Then,*

$$||\breve{\mathfrak{c}}'^{\ddagger}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}_{\mathbb{P}}(\gamma^2_{j,m} m^2 p M^{-1}) \quad \text{for } M \to \infty, \quad \forall j \in \mathbb{N}; \tag{3.28}$$

$$\sup_{j \le k} ||\breve{\mathfrak{c}}'^{\ddagger}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}_{\mathbb{P}}(\gamma^2_{k,m} m^2 p M^{-1}) \quad \text{for } M \to \infty. \tag{3.29}$$

Moreover, if $\boldsymbol{X} = (X_k)_k$ in Assumption 3.1 (a) is centered,

$$\mathbb{E}||\breve{\mathfrak{c}}^{\ddagger}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} = \mathrm{O}(\gamma^2_{j,m} m^2 p M^{-1}) \quad \text{for } M \to \infty, \quad \forall j \in \mathbb{N}; \tag{3.30}$$

$$\mathbb{E}\big(\sup_{j \le k} ||\breve{\mathfrak{c}}^{\ddagger}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m}\big) = \mathrm{O}(\gamma^2_{k,m} m^2 p M^{-1}) \quad \text{for } M \to \infty. \tag{3.31}$$

Remarks 3.5. (a) Theorem 3.5 and also Lemma 3.2 can be seen as generalizations of results in [4; 19; 24; 26] dealing with estimating eigenfunctions of centered $L^2[0,1]$-valued processes.

(b) If $m = m_M$ is bounded, the sequences of reciprocal spectral gaps $(\gamma_{j,m})_M$ are bounded for any j. Consequently, (3.28) equals $\mathrm{O}_{\mathbb{P}}(M^{-1})$, and (3.30) is $\mathrm{O}(M^{-1})$. Moreover, $(\gamma_{k,m})_M$ is guaranteed to be bounded if $k = k_M$ and $m = m_M$ are bounded. Then, (3.29) is $\mathrm{O}_{\mathbb{P}}(M^{-1})$, and (3.31) is $\mathrm{O}(M^{-1})$.

4 A simulation study

Herein, we simulate realizations and estimators of lagged covariance and cross-covariance operators of specific processes. 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 evaluated at $t = 0, \frac{1}{250}, ..., \frac{249}{250}$, and $\langle x, y \rangle_{\mathcal{H}}$ with $x, y \in \mathcal{H}$ is approximated by the Riemann sum $\frac{1}{250}\sum_{t=1}^{250} x(\frac{t-1}{250}) y(\frac{t-1}{250})$.

4.1 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

$$\mathscr{X}_k := (X_k, ..., X_{k-m+1})^T \quad \text{resp.} \quad \mathscr{Y}_k := (Y_k, ..., Y_{k-n+1})^T, \quad \forall k \in \mathbb{Z}, \tag{4.1}$$

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

$$X_k = \alpha(X_{k-1}) + \varepsilon_k, \quad \forall k \in \mathbb{Z}, \tag{4.2}$$

$$Y_k = \beta(X_k) + \varepsilon_k, \quad \forall k \in \mathbb{Z}. \tag{4.3}$$

Thereby, ε_k are defined as in Example 2.1, and $\alpha, \beta\colon \mathcal{H} \to \mathcal{H}$ are integral operators with kernels

$$a(s,t) := k_{\boldsymbol{\varepsilon}}(s,t) \quad \text{resp.} \quad b(s,t) := 2k_{\boldsymbol{\varepsilon}}(s,t), \quad \forall s,t \in [0,1], \tag{4.4}$$

where $k_{\varepsilon;0} = k_{\varepsilon}$ is the integral kernel of $\mathscr{C}_{\varepsilon;0} = \mathscr{C}_{\varepsilon}$ in (2.10). Since the kernel $a(s,t)$ is bounded, we obtain

$$||\alpha||^2_{\mathcal{S}_{\mathcal{H}}} = \int_0^1 \int_0^1 a^2(s,t)\,\mathrm{d}s\mathrm{d}t \;=\; \frac{3}{2} - \ln(2). \tag{4.5}$$

Hence, $||\alpha||_{\mathcal{L}_{\mathcal{H}}} \leq ||\alpha||_{\mathcal{S}_{\mathcal{H}}} < 1$, which implies that (4.2) has the unique stationary solution

$$X_k = \sum_{j=0}^{\infty} \alpha^j(\varepsilon_{k-j}), \quad \forall k \in \mathbb{Z}, \tag{4.6}$$

where $\alpha^0 := \mathbb{I}_{\mathcal{H}}$ is the identity operator, and the series converges in $L^4_{\mathcal{H}}$ and a.s., see [4]. Thus $(X_k)_k$ is a stationary, centered $L^4_{\mathcal{H}}$-valued $AR(1)$ *process*, even $L^4_{\mathcal{H}}$-m-approximable (see [28], Lemma 2.1 for functional (G)ARCH models), and due to (4.3), $(Y_k)_k$ is stationary, centered and $L^4_{\mathcal{H}}$-m-approximable too. After [4], our AR(1) process $\boldsymbol{X}$ fulfills $\mathscr{C}_{\mathbf{X};0} = \mathscr{C}_{\boldsymbol{X}} = \sum_{j=0}^{\infty} \alpha^j \mathscr{C}_{\varepsilon} \alpha^{*j}$ with $\mathscr{C}_{\varepsilon;0} = \mathscr{C}_{\varepsilon}$, and $\mathscr{C}_{\mathbf{X};h} = \alpha^h \mathscr{C}_{\mathbf{X}}$ for $h \in \mathbb{N}_0$. Further, $\mathscr{C}^*_{\mathbf{X};h} = \mathscr{C}_{\mathbf{X};-h}$ for $h \in \mathbb{Z}$, since $\alpha = \mathscr{C}_{\varepsilon}$ is selfadjoint and commutes with $\mathscr{C}_{\varepsilon}$ and due to the series representation of $\mathscr{C}_{\mathbf{X}}$ also with $\mathscr{C}_{\mathbf{X}}$, and $||\alpha||_{\mathcal{S}_{\mathcal{H}}} < 1$ lead to the Neumann series

$$\mathscr{C}_{\mathbf{X};h} = \alpha^{|h|+1} \sum_{j=0}^{\infty} \alpha^{2j} = \alpha^{|h|+1}\left(\mathbb{I}_{\mathcal{H}} - \alpha^2\right)^{-1}, \quad \forall h \in \mathbb{Z}. \tag{4.7}$$

Moreover, (4.2), (4.3), elementary conversions and (4.7) yield

$$\mathscr{C}_{\mathbf{X},\mathbf{Y};h} = \beta\,\mathscr{C}_{\mathbf{X};h}, \quad \forall h \in \mathbb{Z}. \tag{4.8}$$

For 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 $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \mathbb{E}\langle \mathscr{X}_0, \boldsymbol{x}\rangle_{\mathcal{H}^m} \mathscr{Y}_h$ holds for any $h \in \mathbb{Z}$ and $\boldsymbol{x} := (x_1, ..., x_m)^T \in \mathcal{H}^m$,

$$\mathscr{C}_{\boldsymbol{\mathscr{X}};h}(\boldsymbol{x}) = \Big(\sum_{i=1}^{m} \mathscr{C}_{\mathbf{X};h+i-1}(x_i), \ldots, \sum_{i=1}^{m} \mathscr{C}_{\mathbf{X};h+i-m}(x_i)\Big)^T \in \mathcal{H}^m, \tag{4.9}$$

$$\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}(\boldsymbol{x}) = \Big(\sum_{i=1}^{m} \mathscr{C}_{\mathbf{X},\mathbf{Y};h+i-1}(x_i), \ldots, \sum_{i=1}^{m} \mathscr{C}_{\mathbf{X},\mathbf{Y};h+i-n}(x_i)\Big)^T \in \mathcal{H}^n. \tag{4.10}$$

Remarks 4.1. For extensive works on functional AR(MA) processes, we refer to [4; 40] and also [1; 6; 7; 14; 16; 32] from a technical point of view, and [9; 23; 37] for methods combined with applications.

4.2 Simulation of realizations of our processes

Here, we simulate realizations of $(\mathscr{X}_k)_k, (\mathscr{Y}_k)_k$ in (4.1), for which we first simulate innovations in (2.9).

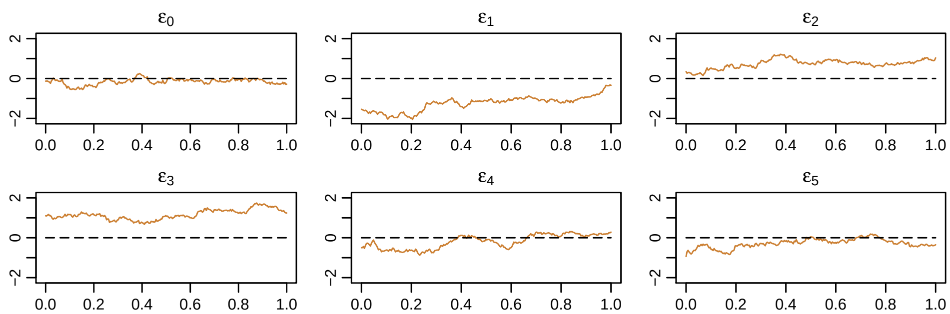

Figure 5: Realizations of the innovations $\varepsilon_0, ..., \varepsilon_5$ in (2.9).

These simulated realizations then can be plugged into the equations (4.2) and (4.3) 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

Lemma 4.1. *Let $A \in \mathcal{L}_{\mathcal{H}}$ with $||A||_{\mathcal{L}_{\mathcal{H}}} < 1$, let $(\varepsilon_k)_{k\in\mathbb{Z}}$ be an i.i.d., centered $L^{\nu}_{\mathcal{H}}$-process for $\nu > 0$, and let $Z_k = A(Z_{k-1}) + \varepsilon_k$, $\tilde{Z}_k = A(\tilde{Z}_{k-1}) + \varepsilon_k$ for all $k \in \mathbb{Z}$ hold, where $\tilde{Z}_0 \in \mathcal{H}$ is deterministic. Then, for $\rho \in (0, 1)$,*

$$\mathbb{E}||Z_N - \tilde{Z}_N||^{\nu}_{\mathcal{H}} = \mathrm{O}(\rho^N) \quad \text{for } N \to \infty. \tag{4.11}$$

Remarks 4.2. Lemma 4.1 can be shown for functional AR(MA) processes with arbitrary order(s) in any separable Hilbert space, see [28], Corollary 4.1 for functional (G)ARCH processes.

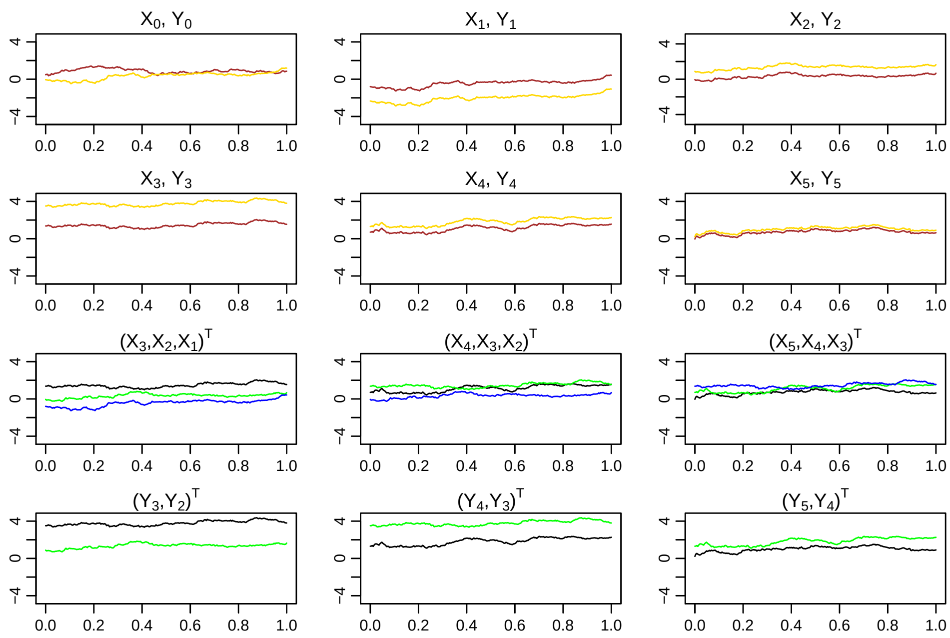

Figure 6: Six consecutive realizations of $(X_k)_k$ (**bordeaux**) and $(Y_k)_k$ (**gold**) in the first two rows. X_0 was approximated by $\tilde{Z}_{100}$ in Lemma 4.1 with $A = \alpha, \varepsilon_k$ for $k = 1, ..., 100$ as in (2.9) and $\tilde{Z}_0 := 0_{\mathcal{H}}$, and $X_1, ..., X_5$ and $Y_0, ..., Y_5$ were obtained by applying (4.2) resp. (4.3) with the innovations in Fig 5. Then, $X_0, ..., X_5$ and $Y_0, ..., Y_5$ were plugged into the equations in (4.1) 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**.

4.3 Simulation of our operators

In this section, we illustrate certain lag-h-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ and lag-h-cross-covariance operators $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ of the centered processes $\boldsymbol{\mathscr{X}} = (\mathscr{X}_k)_k$ and $\boldsymbol{\mathscr{Y}} = (\mathscr{Y}_k)_k$ in Section 4.2 with Cartesian powers $m = 3$ resp. $n = 2$, and simulate estimators for these operators for fixed and increasing h, m, n. Due to the infinite series (4.7) consisting of operators, precisely calculating $\mathscr{C}_{\mathbf{X};h}$ resp. $\mathscr{C}_{\mathbf{X},\mathbf{Y};h}$ is impossible. However, $\mathscr{C}_{\mathbf{X};h}$ and $\mathscr{C}_{\mathbf{X},\mathbf{Y};h}$ can for sufficiently large $K \in \mathbb{N}$ and any $h \in \mathbb{Z}$ be well approximated by

$$\tilde{\mathscr{C}}_{\mathbf{X};h;K} := \alpha^{|h|+1} \sum_{j=0}^{K} \alpha^{2j} \quad \text{resp.} \quad \tilde{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h;K} := \beta\alpha^{|h|+1} \sum_{j=0}^{K} \alpha^{2j}. \tag{4.12}$$

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 $\beta = 2\alpha$ after (4.4) for any h, K to

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

Also, the components of $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ cannot be expressed independently of any argument $\boldsymbol{x} := (x_1, ..., x_m)^T \in \mathcal{H}^m$, except when all of the argument's components match. With $(A_1(\boldsymbol{x}), ..., A_m(\boldsymbol{x})) := (A_1, ..., A_m)(\boldsymbol{x})$ for operators $A_1, ..., A_m$ with domain $\mathcal{H}^m$, $\mathscr{C}_{\mathbf{X};h} = \mathscr{C}_{\mathbf{X};-h}$ and $\mathscr{C}_{\mathbf{X},\mathbf{Y};h} = \mathscr{C}_{\mathbf{X},\mathbf{Y};-h}$ for any h, for, e.g., $\mathscr{C}_{\boldsymbol{\mathscr{X}};0}$ and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$

with $m = 3, n = 2$ holds for any $\boldsymbol{x} = (x, x, x) \in \mathcal{H}^3$ due to (4.9), (4.10),

$$\mathscr{C}_{\boldsymbol{\mathscr{X}};0}(\boldsymbol{x}) = \Big(\big(\mathscr{C}_{\mathbf{X};0} + \mathscr{C}_{\mathbf{X};1} + \mathscr{C}_{\mathbf{X};2},\ \mathscr{C}_{\mathbf{X};0} + 2\mathscr{C}_{\mathbf{X};1},\ \mathscr{C}_{\mathbf{X};0} + \mathscr{C}_{\mathbf{X};1} + \mathscr{C}_{\mathbf{X};2}\big)(x)\Big)^T, \tag{4.13}$$

$$\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}(\boldsymbol{x}) = \Big(\big(\mathscr{C}_{\mathbf{X},\mathbf{Y};0} + 2\mathscr{C}_{\mathbf{X},\mathbf{Y};1},\ \mathscr{C}_{\mathbf{X},\mathbf{Y};0} + \mathscr{C}_{\mathbf{X},\mathbf{Y};1} + \mathscr{C}_{\mathbf{X},\mathbf{Y};2}\big)(x)\Big)^T. \tag{4.14}$$

In order to illustrate estimators for the operators in the components of $\mathscr{C}_{\boldsymbol{\mathscr{X}};0}(\boldsymbol{x})$ in (4.13) and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}(\boldsymbol{x})$ in (4.14), and to estimate $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ for fixed and varying h, m, n, with $h \geq 0$ w.l.o.g., we generate $X_1, ..., X_M$ and $Y_1, ..., Y_N$ of the processes $\boldsymbol{X}$ resp. $\boldsymbol{Y}$ in Section 4.1 with $M = N$. This leads to the values $\mathscr{X}_m, ..., \mathscr{X}_{\tilde{\mathscr{M}}}$ of $\boldsymbol{\mathscr{X}}$ and $\mathscr{Y}_n, ..., \mathscr{Y}_{\tilde{\mathscr{N}}}$ of $\boldsymbol{\mathscr{Y}}$ 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 $\mathbf{Y}$, the operators $\mathscr{C}_{\mathbf{X};h}$ in (4.13) and $\mathscr{C}_{\mathbf{X},\mathbf{Y};h}$ in (4.14) 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

$$\hat{k}_{\mathbf{X};h}(s,t) \coloneqq \frac{1}{M-h} \sum_{k=1}^{M-h} X_k(s) X_{k+h}(t), \quad \forall s,t \in [0,1], \tag{4.15}$$

$$\text{resp.} \quad \hat{k}_{\mathbf{X},\mathbf{Y};h}(s,t) \coloneqq \frac{1}{M-h} \sum_{k=1}^{M-h} X_k(s) Y_{k+h}(t), \quad \forall s,t \in [0,1]. \tag{4.16}$$

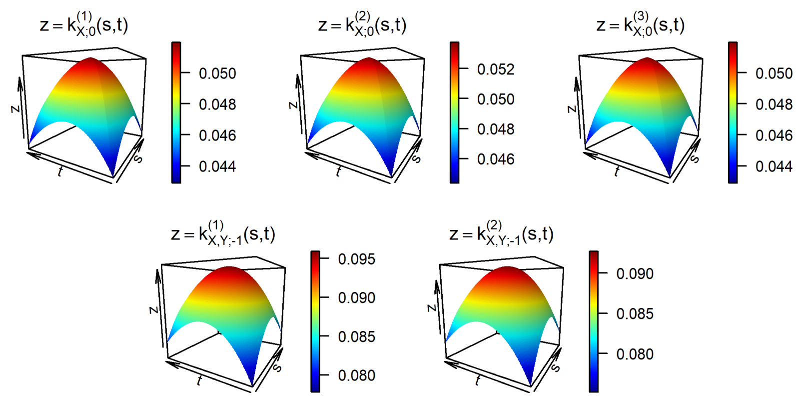

Figure 7: The integral kernels $k^{(1)}_{\boldsymbol{\mathscr{X}};0}, k^{(2)}_{\boldsymbol{\mathscr{X}};0}, k^{(3)}_{\boldsymbol{\mathscr{X}};0}$ (first row) and $k^{(1)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}, k^{(2)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ (second row) of the operators in the three resp. two components of $\mathscr{C}_{\boldsymbol{\mathscr{X}};0}$ in (4.13) resp. $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ in (4.14). These kernels result by the associated sum of the integral kernels $k_{\mathbf{X};0}, k_{\mathbf{X};1}, k_{\mathbf{X};2}$ and $k_{\mathbf{X},\mathbf{Y};0}, k_{\mathbf{X},\mathbf{Y};1}, k_{\mathbf{X},\mathbf{Y};2}$ of the operators $\mathscr{C}_{\mathbf{X};0}, \mathscr{C}_{\mathbf{X};1}, \mathscr{C}_{\mathbf{X};2}$ resp. $\mathscr{C}_{\mathbf{X},\mathbf{Y};0}, \mathscr{C}_{\mathbf{X},\mathbf{Y};1}, \mathscr{C}_{\mathbf{X},\mathbf{Y};2}$ which were approximated by their respective operators in (4.12) with $K = 100$.

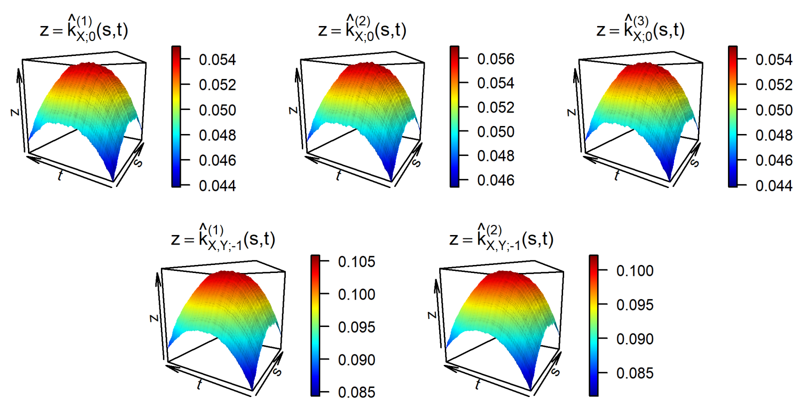

Figure 8: The estimators $\hat{k}^{(1)}_{\boldsymbol{\mathscr{X}};0}, \hat{k}^{(2)}_{\boldsymbol{\mathscr{X}};0}, \hat{k}^{(3)}_{\boldsymbol{\mathscr{X}};0}$ (first row) and $\hat{k}^{(1)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}, \hat{k}^{(2)}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ (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 (4.13) resp. $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};-1}$ in (4.14). 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 (4.15) and $\hat{k}_{\mathbf{X},\mathbf{Y};0}, \hat{k}_{\mathbf{X},\mathbf{Y};1}, \hat{k}_{\mathbf{X},\mathbf{Y};2}$ in (4.16) with $M = 1000$ for the operators $\mathscr{C}_{\mathbf{X};0}, \mathscr{C}_{\mathbf{X};1}, \mathscr{C}_{\mathbf{X};2}$ resp. $\mathscr{C}_{\mathbf{X},\mathbf{Y};0}, \mathscr{C}_{\mathbf{X},\mathbf{Y};1}, \mathscr{C}_{\mathbf{X},\mathbf{Y};2}$.

Finally, 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 $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ and $\boldsymbol{\mathscr{Y}} := (\mathscr{Y}_k)_{k\in\mathbb{Z}}$ in (4.1) 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 $\boldsymbol{\mathscr{Y}}$, $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h}$ in (3.5) resp. $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ in (3.9) are used to estimate $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ and $\mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$, which satisfy due to our processes' definition and $h \geq 0$,

$$\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} = \frac{1}{M-h-m+1}\sum_{k=m}^{M-h} \mathscr{X}_k \otimes \mathscr{X}_{k+h}, \tag{4.17}$$

$$\text{resp.}\quad \hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} = \frac{1}{M-h-\max(m,n-h)+1}\sum_{k=\max(m,n-h)}^{M-h} \mathscr{X}_k \otimes \mathscr{Y}_{k+h}. \tag{4.18}$$

In order to calculate the estimation errors, the equations

$$||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{H}^m}} = \sum_{i=1}^{m}\sum_{j=1}^{m} ||\hat{\mathscr{C}}_{\mathbf{X};h+i-j} - \mathscr{C}_{\mathbf{X};h+i-j}||^2_{\mathcal{S}_{\mathcal{H}}}, \tag{4.19}$$

$$||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}} = \sum_{i=1}^{m}\sum_{j=1}^{n} ||\hat{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h+i-j} - \mathscr{C}_{\mathbf{X},\mathbf{Y};h+i-j}||^2_{\mathcal{S}_{\mathcal{H}}} \tag{4.20}$$

are utilized, where $\hat{\mathscr{C}}_{\mathbf{X};h+i-j}$ and $\hat{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h+i-j}$ equal $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h}$ in (4.17) resp. $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ in (4.18) with $\mathscr{X}_k$, $\mathscr{X}_{k+h}$ and $\mathscr{Y}_{k+h}$ replaced by $X_{k+1-i}, X_{k+h+1-j}$ resp. $Y_{k+h+1-j}$ for all i, j. Thereby, the equations (4.19) and (4.20) follow from the definition of the given norms and operators (see also (4.9), (4.10)).

	$\|\|\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}\|\|^2_{\mathcal{S}_{\mathcal{H}^m}}$						$\|\|\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}\|\|^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}}$					
$m=m_M, n=n_M$	$m=3$			$m=\lfloor M^{1/4}\rfloor$			$m=3, n=2$			$m=n=\lfloor M^{1/4}\rfloor$		
M \ $h=h_M$	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	.0223	.0216	.0187	.0223	.0216	.0187	.0084	.0083	.0079	.0126	.0125	.0118
200	.0058	.0055	.0043	.0058	.0055	.0043	.0010	.0009	.0008	.0014	.0014	.0012
500	.0141	.0132	.0083	.0498	.0475	.0313	.0044	.0039	.0026	.0303	.0289	.0198
1000	.0126	.0118	.0060	.1137	.1101	.0628	.0036	.0032	.0017	.0773	.0747	.0435
2000	.0125	.0118	.0050	.2331	.2271	.1163	.0036	.0032	.0015	.1714	.1670	.0873
5000	.0105	.0099	.0027	.5825	.5712	.2236	.0028	.0024	.0007	.4597	.4509	.1785

Table 1: Simulation of $||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{H}^m}}$ in (4.19), $||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{H}^m,\mathcal{H}^n}}$ in (4.20) for various sample sizes M, lags h and Cartesian powers m, n, with $\mathscr{C}_{\mathbf{X};h+i-j}$ and $\mathscr{C}_{\mathbf{X},\mathbf{Y};h+i-j}$ approximated by $\tilde{\mathscr{C}}_{\mathbf{X};h+i-j;100}$ resp. $\tilde{\mathscr{C}}_{\mathbf{X},\mathbf{Y};h+i-j;100}$ in (4.12).

Remarks 4.3. All parameters in the simulation study with estimation errors in Table 1 are chosen so that the prerequisites of Theorems 3.1 and 3.3 are satisfied. These errors run for growing sample size M below or as the asymptotic upper bounds in Theorems 3.1 and 3.3. Up to $M = 200$, probably due to fortunate random errors, the calculated errors decrease for fixed resp. increase for increasing $m = m_M, n = n_M$ and any $h = h_M$ as expected. That the errors for increasing $m = m_M, n = n_M$ not yet visibly approach zero could be because M is either too small, or that small values of the estimators or the operators to be estimated are rounded to zero, leading to larger estimation errors. Also, the estimation errors for the lag-h-cross-covariance are smaller as for the lag-h-covariance operators due to their definiton.

5 Conclusions

This article proposes estimators for lagged covariance and cross-covariance operators and the principal components of (lag-0-)covariance operators of processes in separable Hilbert spaces, especially of processes obtained by successively stacking Hilbert space-valued elements, hence in Cartesian products of Hilbert spaces. The focus lies on the asymptotic upper bounds of the estimation errors. All estimators are stated for centered processes and for those with an unknown mean. The asymptotic upper bounds allow both the processes' Cartesian powers and the lag to be fixed or to increase w.r.t. the sample size, and the principal components are estimated separately and uniformly. Our findings are useful whenever one is concerned about the dependence within one or between two processes having values in (Cartesian products of) Hilbert spaces, or one has to derive asymptotic upper bounds of estimation errors where the given estimators rely on empirical (lagged) covariance or cross-covariance operators, see [2; 28] for latter. 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. Furthermore, it would be interesting to deduce our results also on separable Banach spaces, see, e.g., [36] who dealt with the estimation of AR operators in Banach spaces, to derive the asymptotic distribution of our estimation errors (see [33]) as well as their asymptotic lower bounds.

6 Proofs

Proof of Lemma 3.1. $\hat{m}_{\mathscr{X}}$ is an unbiased estimator for $m_{\mathscr{X}}$ due to its definition. $\boldsymbol{X} = (X_k)_k$ and $\mathscr{X} = (\mathscr{X}_k)_k$ are stationary, and $X_{1+h}, X_{1+h;h}$ are independent for all h. Thus, similar to [18], due to $m_{\mathscr{X}} = (m_{\boldsymbol{X}}, ..., m_{\boldsymbol{X}})^T \in \mathcal{U}^m$, (2.4), Cauchy-Schwarz inequality, $\mathscr{Z}_k := \mathscr{X}_k - m_{\mathscr{X}}$ and $Z_k := X_k - m_{\boldsymbol{X}}$ for any k, holds with $\tilde{\mathscr{M}} = \tilde{\mathscr{M}}_M, \mathscr{M} = \mathscr{M}_M$ in Assumption 3.1:

$$\begin{aligned}
\mathbb{E}||\hat{m}_{\mathscr{X}} - m_{\mathscr{X}}||^2_{\mathcal{U}^m} &= \mathscr{M}^{-2} \sum_{i,j=m}^{\tilde{\mathscr{M}}} \mathbb{E}\langle \mathscr{Z}_i, \mathscr{Z}_j\rangle_{\mathcal{U}^m} = m\mathscr{M}^{-2} \sum_{i,j=1}^{\mathscr{M}} \mathbb{E}\langle Z_i, Z_j\rangle_{\mathcal{U}} \\
&= m\mathscr{M}^{-1}\Big[\nu_{2,\mathcal{U}}(Z_1) + 2\sum_{h=1}^{\mathscr{M}-1} \frac{\mathscr{M}-h}{\mathscr{M}} \mathbb{E}\langle Z_1, X_{1+h} - X_{1+h;h}\rangle_{\mathcal{U}}\Big] \\
&\leq m\mathscr{M}^{-1}\nu_{2,\mathcal{U}}(Z_1)\Big[1 + 2\sum_{h=1}^{\infty} \nu_{2,\mathcal{U}}(X_{1+h} - X_{1+h;h})\Big] \\
&= \mathrm{O}(m\mathscr{M}^{-1}) = \mathrm{O}(mpM^{-1}) \quad \text{for } M \to \infty,
\end{aligned}$$

where the last two steps hold due to $L^4_{\mathcal{U}}$-approximibility of $(X_k)_k$ and (3.3). □

In various conversions for deriving our upper bounds the following two Lemmas are utilized.

Lemma 6.1. *Let $(\mathcal{H}, ||\cdot||_{\mathcal{H}})$ be a separable Hilbert space. Also, let $(S_k)_{k\in\mathbb{Z}}$ be a stationary $L^4_{\mathcal{H}}$-process, and for some $l \in \mathbb{N}, \mathscr{S}_k := (S_{f(k,1)}, ..., S_{f(k,l)})^T$ for all k and some function $f\colon \mathbb{Z}\times\{1,...,l\} \to \mathbb{Z}$. Then,*

$$\nu_{4,\mathcal{H}^l}(\mathscr{S}_k) \leq \sqrt{l}\,\nu_{4,\mathcal{H}}(S_j), \quad \forall j,k. \tag{6.1}$$

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

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

Lemma 6.2. *Let Assumption 3.1 hold. Moreover, we define $\mathscr{X}_{m+j;l} := (X_{m+jp;l}, ..., X_{1+jp;l})^T$ and $\mathscr{Y}_{n+j;l} := (Y_{n+jq;l}, ..., Y_{1+jq;l})^T$ 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

$$\sum_{k=1}^{\infty} \nu_{4,\mathcal{U}^m}(\mathscr{X}_k - \mathscr{X}_{k;k}) < \infty \qquad \textit{resp.} \qquad \sum_{k=1}^{\infty} \nu_{4,\mathcal{V}^n}(\mathscr{Y}_k - \mathscr{Y}_{k;k}) < \infty. \tag{6.2}$$

Thereby, $(\mathscr{X}_k)_k$ is $L^4_{\mathcal{U}^m}$-m-approximable for $p = 1$, and $(\mathscr{Y}_k)_k$ is $L^4_{\mathcal{V}^n}$-m-approximable for $q = 1$.

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

$$\sum_{k=1}^{\infty}\nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{W}_{k,h}-\mathscr{W}_{k,h;k})$$
$$\leq \sqrt{mn}\Big[\sum_{k=1}^{\infty}\nu_{4,\mathcal{V}}(Y_1)\nu_{4,\mathcal{U}}(X_k-X_{k;k})+\nu_{4,\mathcal{U}}(X_1)\nu_{4,\mathcal{V}}(Y_k-Y_{k;k})\Big]. \tag{6.3}$$

Moreover, $(\mathscr{W}_{k+h})_k$ *is* $L^2_{\mathcal{U}^m\mathcal{V}^n}$*-m-approximable for* $h\leq 0$ *if* $p=q=1$.

Proof. (a) From the definition of $\mathscr{X}_k, \mathscr{X}_{k;k}, \mathscr{Y}_k, \mathscr{Y}_{k;k}$ for all k follows $\nu_{4,\mathcal{U}^m}(\mathscr{X}_k-\mathscr{X}_{k;k})\leq\sqrt{m}\,\nu_{4,\mathcal{U}}(X_k-X_{k;k})$ and $\nu_{4,\mathcal{V}^n}(\mathscr{Y}_k-\mathscr{Y}_{k;k})\leq\sqrt{n}\,\nu_{4,\mathcal{V}}(Y_k-Y_{k;k})$, and thus (6.2). Hence, since $(\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are non-anticipative w.r.t. $(\varepsilon_k)_k$ for $p=1$ resp. $q=1, (\mathscr{X}_k)_k$ and $(\mathscr{Y}_k)_k$ are $L^4_{\mathcal{U}^m}$-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, (6.1) and L^4-m-approximability of $(X_k)_k$ and $(Y_k)_k$ yield

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

Moreover, since $(\mathscr{X}_k)_k, (\mathscr{Y}_{k+h})_k$ and consequently also $(\mathscr{W}_{k+h})_k$ are non-anticipative w.r.t. $(\varepsilon_k)_k$ for $h\leq 0$ if $p=q=1, (\mathscr{W}_{k+h})_k$ is indeed $L^2_{\mathcal{S}_{\mathcal{U}^m\mathcal{V}^n}}$-$m$-approximable for $h\leq 0$ if $p=q=1$. □

Proof of Theorem 3.1. We use ideas from the proof of [19], Theorem 3.1. $\hat{\mathscr{C}}_{\mathscr{X};h}$ is an unbiased estimator for $\mathscr{C}_{\mathscr{X};h}$ with $|h|<\mathscr{M}=\mathscr{M}_M$ 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, we show (3.6) for $h\geq 0$ w.l.o.g. Stationarity of $\mathscr{X}$ implies for any h with $0\leq h<\mathscr{M}_M$ where $\mathscr{Z}_{k,h}:=\mathscr{W}_{k,h}-\mathscr{C}_{\mathscr{X};h}$ with $\mathscr{W}_{k,h}:=\mathscr{X}_k\otimes\mathscr{X}_{k+h}$, and $\mathscr{M}_{M,h}=\mathscr{M}_M-|h|$:

$$\begin{aligned}\mathbb{E}||\hat{\mathscr{C}}_{\mathscr{X};h}-\mathscr{C}_{\mathscr{X};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} &= \mathscr{M}_{M,h}^{-2}\sum_{|r|<\mathscr{M}_{M,h}}(\mathscr{M}_{M,h}-|r|)\,\mathbb{E}\langle\mathscr{Z}_{m,h},\mathscr{Z}_{m+r,h}\rangle_{\mathcal{S}_{\mathcal{U}^m}}\\ &\leq 2\mathscr{M}_{M,h}^{-1}\sum_{r=0}^{\infty}\mathbb{E}\langle\mathscr{Z}_{m,h},\mathscr{Z}_{m+r,h}\rangle_{\mathcal{S}_{\mathcal{U}^m}}.\end{aligned} \tag{6.4}$$

Let $\sigma(T_k, k\in I)$ be the σ-algebra generated by the random variables T_k with $k\in I$ where $I\subseteq\mathbb{Z}$ is some index set. From Assumption 3.1 (a), the definition of $\mathscr{X}_k$ for any k for some $p\in\mathbb{N}$, and of $\mathscr{W}_{k,h}$ for any h,k follows for $h\geq 0$:

$$\mathscr{Z}_{m,h}=\mathscr{W}_{m,h}-\mathscr{C}_{\mathscr{X};h}\in\sigma(X_1,...,X_m,X_{1+hp},...,X_{m+hp})\subseteq\sigma(\varepsilon_{m+hp},\varepsilon_{m+hp-1},...)$$

where $(\varepsilon_k)_k$ is an i.i.d. process, and for any $r\in\mathbb{N}$,

$$\mathscr{Z}_{m+r,h}=\mathscr{W}_{m+r,h}-\mathscr{C}_{\mathscr{X};h}\in\sigma(\varepsilon_{m+(h+r)p},\varepsilon_{m+(h+r)p-1},...).$$

Consequently, $\mathscr{Z}_{m,h}$ and $\mathscr{Z}_{m+r,h;r-h}:=\mathscr{W}_{m+r,h;r-h}-\mathscr{C}_{\mathscr{X};h}$ with $\mathscr{W}_{m+r,h;r-h}=\mathscr{X}_{m+r;r-h}\otimes\mathscr{X}_{m+h+r;r-h}$ (see Lemma 6.2 with $\mathscr{Y}_k=\mathscr{X}_k, \mathscr{Y}_{k;l}=\mathscr{X}_{k;l}$) are independent for $r>h$ for any m,p. With that being said, and since $\mathscr{Z}_{m,h}$ and $\mathscr{Z}_{m+r,h;r-h}$ are centered for all h,k,r, Cauchy-Schwarz inequality, (2.4) and Lemma 6.2 with $\mathscr{Y}_k=\mathscr{X}_k, \mathscr{Y}_{k;l}=\mathscr{X}_{k;l}$ for all k,l, yield for the sum in (6.4):

$$\sum_{r=0}^{\infty}\mathbb{E}\langle\mathscr{Z}_{m,h},\mathscr{Z}_{m+r,h}\rangle_{\mathcal{S}_{\mathcal{U}^m}}=\sum_{r=0}^{h}\mathbb{E}\langle\mathscr{Z}_{m,h},\mathscr{Z}_{m+r,h}\rangle_{\mathcal{S}_{\mathcal{U}^m}}+\sum_{r>h}\mathbb{E}\langle\mathscr{Z}_{m,h},\mathscr{W}_{m+r,h}-\mathscr{W}_{m+r,h;r-h}\rangle_{\mathcal{S}_{\mathcal{U}^m}} \tag{6.5}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

$$\leq \nu_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h})\Big[(1+h)\nu_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h}) + \sum_{k=1}^{\infty} \nu_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{W}_{k,h} - \mathscr{W}_{k,h;k})\Big] \tag{6.6}$$

$$\leq \nu_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h})\Big[(1+h)\nu_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h}) + 2m\nu_{4,\mathcal{U}}(X_m)\sum_{k=1}^{\infty} \nu_{4,\mathcal{U}}(X_k - X_{k;k})\Big]. \tag{6.7}$$

Further, we have $\nu^2_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h}) := \mathbb{E}||\mathscr{Z}_{m,h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \mathbb{E}||\mathscr{W}_{m,h}||^2_{\mathcal{S}_{\mathcal{U}^m}} - ||\mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}}$ due to [20], Theorem 7.2.2, and $||\mathscr{W}_{m,h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = ||\mathscr{X}_m||^2_{\mathcal{U}^m}||\mathscr{X}_{m+h}||^2_{\mathcal{U}^m}$. Hence, (2.3), $||\cdot||_{\mathcal{S}_{\mathcal{U}^m}} \leq ||\cdot||_{\mathcal{N}_{\mathcal{U}^m}}$, Cauchy-Schwarz inequality, stationarity of $(\mathscr{X}_k)_k$ and (6.1) yield

$$\nu^2_{2,\mathcal{S}_{\mathcal{U}^m}}(\mathscr{Z}_{m,h}) \leq 2\,\mathbb{E}||\mathscr{X}_m||^4_{\mathcal{U}^m} \leq 2m^2\nu^4_{4,\mathcal{U}}(X_1). \tag{6.8}$$

From (6.4), (6.7), (6.8) and $L^4_{\mathcal{U}}$-m-approximibility of $(X_k)_k$ follows

$$\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \leq a(1+h)m^2\mathscr{M}^{-1}_{M,h}$$

for some constant a independent of $h = h_M, m = m_M, p = p_M$ in Assumption 3.2 (a) and thus also of $\mathscr{M}_{M,h} = \mathscr{M}_M - |h|$ with $\mathscr{M}_{M,h} \sim p^{-1}M$ after Assumption 3.3 (a) and (3.3). Hence, (3.6) is verified. □

Proof of Theorem 3.2. 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 $0 \leq h < \mathscr{M}_M - 1$, and $\mathscr{M}_{M,h} = \mathscr{M}_M - |h|$:

$$\begin{aligned}
\mathbb{E}(\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}) &= \frac{1}{\mathscr{M}_{M,h}-1}\sum_{k=1}^{\mathscr{M}_{M,h}} \mathbb{E}\Big(\big(\mathscr{X}_k - \frac{1}{\mathscr{M}_{M,h}}\sum_{i=1}^{\mathscr{M}_{M,h}}\mathscr{X}_i\big)\otimes\big(\mathscr{X}_{k+h} - \frac{1}{\mathscr{M}_{M,h}}\sum_{j=1}^{\mathscr{M}_{M,h}}\mathscr{X}_{j+h}\big)\Big)\\
&= \frac{1}{\mathscr{M}_{M,h}(\mathscr{M}_{M,h}-1)}\Big(\mathscr{M}^2_{M,h}\mathscr{C}_{\boldsymbol{\mathscr{X}};h} - \sum_{i,k=1}^{\mathscr{M}_{M,h}}\mathscr{C}_{\boldsymbol{\mathscr{X}};k+h-i}\Big)\\
&= \mathscr{C}_{\boldsymbol{\mathscr{X}};h} - \frac{1}{\mathscr{M}_{M,h}(\mathscr{M}_{M,h}-1)}\sum_{\substack{1\leq i,k\leq \mathscr{M}_{M,h}\\ i\neq k}}\mathscr{C}_{\boldsymbol{\mathscr{X}};k+h-i}.
\end{aligned} \tag{6.9}$$

Hence, $\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h}$ is an unbiased estimator for $\mathscr{C}_{\boldsymbol{\mathscr{X}};h}$ for h with $0 \leq h < \mathscr{M}_M - 1$ if the sum in (6.9) equals $0_{\mathcal{L}_{\mathcal{U}^m}}$ which can also be shown for h with $1 - \mathscr{M}_M < h < 0$. Now, we verify (3.8). 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 $h \geq 0$ w.l.o.g. For $h < \mathscr{M}_M - 1$ holds

$$\begin{aligned}
\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} &= \frac{\mathscr{M}_{M,h}}{\mathscr{M}_{M,h}-1}(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}})\otimes(m_{\boldsymbol{\mathscr{X}}} - \hat{m}'_{\boldsymbol{\mathscr{X}}}) + \frac{1}{\mathscr{M}_{M,h}-1}\sum_{j=m}^{\tilde{\mathscr{M}}_{M,h}}\mathscr{U}_k\otimes\mathscr{U}_{k+h}\\
&= \frac{\mathscr{M}_{M,h}}{\mathscr{M}_{M,h}-1}\Big[(m_{\boldsymbol{\mathscr{X}}} - \hat{m}_{\boldsymbol{\mathscr{X}}})\otimes(m_{\boldsymbol{\mathscr{X}}} - \hat{m}'_{\boldsymbol{\mathscr{X}}}) + \hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h}\Big]
\end{aligned} \tag{6.10}$$

with $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h}$ as in (3.5) based on a sample $\mathscr{U}_m, ..., \mathscr{U}_{\tilde{\mathscr{M}}_{M,h}}$ of $\boldsymbol{\mathscr{U}} := (\mathscr{U}_k)_{k\in\mathbb{Z}}$ where $\mathscr{U}_k := \mathscr{X}_k - m_{\boldsymbol{\mathscr{X}}}$. (6.10), $\mathscr{C}_{\boldsymbol{\mathscr{X}};h} = \mathscr{C}_{\boldsymbol{\mathscr{U}};h}$, Δ-inequality, $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$ 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

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

We have $||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m}||\hat{m}'_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m} = \mathrm{O}_{\mathbb{P}}(m^2\mathscr{M}^{-2}_{M,h})$ after Lemma 3.1, $||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}};h} - \mathscr{C}_{\boldsymbol{\mathscr{U}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \mathrm{O}_{\mathbb{P}}((1+|h|)m^2pM^{-1})$ after Theorem 3.1, and $||\mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} \leq m^2\,\mathbb{E}||X_1||^4_{\mathcal{U}}$ due to $||\cdot||_{\mathcal{S}_{\mathcal{U}^m}} \leq ||\cdot||_{\mathcal{N}_{\mathcal{U}^m}}$, (2.3), Cauchy-Schwarz inequality and (6.1). Then, under Assumptions 3.1-3.3 (a), thus $\mathscr{M}_{M,h} \sim p^{-1}M$ after (3.3), and

$$||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}};h}||^2_{\mathcal{S}_{\mathcal{U}^m}} = \mathrm{O}_{\mathbb{P}}(m^2\mathscr{M}^{-2}_{M,h}) + \mathrm{O}_{\mathbb{P}}((1+|h|)m^2pM^{-1}) + \mathrm{O}(m^2\mathscr{M}^{-2}_{M,h}) \quad \text{for } M\to\infty$$

$$= \mathrm{O}_{\mathbb{P}}((1+|h|)m^2 pM^{-1}) \quad \text{for } M \to \infty. \qquad \square$$

Proof of Theorem 3.3. $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ 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. From deliberations in the proof of Theorem 3.1, especially (6.4) and (6.6) with $\mathscr{Z}_{k,h} := \mathscr{W}_{k,h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ and $\mathscr{W}_{k,h} := \mathscr{X}_k \otimes \mathscr{Y}_{k+h}$,

$$\nu^2_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{Z}_{1,h}) \leq 2\left(\mathbb{E}||\mathscr{X}_m||^4_{\mathcal{U}^m}\mathbb{E}||\mathscr{Y}_n||^4_{\mathcal{V}^n}\right)^{1/2} \leq 2mn\nu^2_{4,\mathcal{U}}(X_1)\nu^2_{4,\mathcal{V}}(Y_1) \tag{6.11}$$

similar as in (6.8), and Lemma 6.2 follows

$$\begin{aligned}
&\mathbb{E}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}\\
&\leq 2\mathscr{L}^{-1}_{M,N,h}\nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{Z}_{1,h})\Big[(1+h)\nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{Z}_{1,h}) + \sum_{k=1}^{\infty}\nu_{2,\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}(\mathscr{W}_{k,h} - \mathscr{W}_{k,h;k})\Big]\\
&\leq 2\sqrt{2}\,mn\mathscr{L}^{-1}_{M,N,h}\nu_{4,\mathcal{U}}(X_1)\nu_{4,\mathcal{V}}(Y_1)\\
&\quad\cdot\Big[\sqrt{2}\,(1+h)\nu_{4,\mathcal{U}}(X_1)\nu_{4,\mathcal{V}}(Y_1) + \sum_{k=1}^{\infty}\nu_{4,\mathcal{V}}(Y_1)\nu_{4,\mathcal{U}}(X_k - X_{k;k}) + \nu_{4,\mathcal{U}}(X_1)\nu_{4,\mathcal{V}}(Y_k - Y_{k;k})\Big]\\
&\leq b(1+h)mn\mathscr{L}^{-1}_{M,N,h}
\end{aligned} \tag{6.12}$$

for some constant b independent of M, N, and thus of all given sequences. $\square$

Proof of Theorem 3.4. For h with $n - \tilde{\mathscr{M}}_M \leq h \leq \tilde{\mathscr{N}}_N - m$ 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 3.2. 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 3.2,

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

with $\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h}$ defined in (3.9) based on samples $\mathscr{U}_m, ..., \mathscr{U}_{\tilde{\mathscr{M}}_M}$ of $\boldsymbol{\mathscr{U}} := (\mathscr{U}_k)_{k\in\mathbb{Z}}$ and $\mathscr{V}_n, ..., \mathscr{V}_{\tilde{\mathscr{N}}_N}$ 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 Theorem 3.2-3.3 imply with the assertions of Lemma 3.1, Theorem 3.3, (2.3), (6.11) and $\mathscr{C}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h} = \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}$ for $M, N \to \infty$ as claimed:

$$\begin{aligned}
||\hat{\mathscr{C}}'_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h} - \mathscr{C}_{\boldsymbol{\mathscr{X}},\boldsymbol{\mathscr{Y}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} &\leq \frac{3}{(\mathscr{L}_{M,N,h}-1)^2}\Big[\mathscr{L}^2_{M,N,h}||\hat{m}_{\boldsymbol{\mathscr{X}}} - m_{\boldsymbol{\mathscr{X}}}||^2_{\mathcal{U}^m}||\hat{m}'_{\boldsymbol{\mathscr{Y}}} - m_{\boldsymbol{\mathscr{Y}}}||^2_{\mathcal{V}^n}\\
&\qquad + \mathscr{L}^2_{M,N,h}||\hat{\mathscr{C}}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h} - \mathscr{C}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}} + ||\mathscr{C}_{\boldsymbol{\mathscr{U}},\boldsymbol{\mathscr{V}};h}||^2_{\mathcal{S}_{\mathcal{U}^m,\mathcal{V}^n}}\Big]\\
&= \mathrm{O}_{\mathbb{P}}(mn\mathscr{L}^{-2}_{M,N,h}) + \mathrm{O}_{\mathbb{P}}((1+|h|)mn\mathscr{L}^{-1}_{M,N,h}) + \mathrm{O}(mn\mathscr{L}^{-2}_{M,N,h})\\
&= \mathrm{O}_{\mathbb{P}}((1+|h|)mn\mathscr{L}^{-1}_{M,N,h}). \qquad \square
\end{aligned}$$

Corollary 3.1. Follows from (3.14), $||\cdot||_{\mathcal{L}_{\mathcal{U}^m}} \leq ||\cdot||_{\mathcal{S}_{\mathcal{U}^m}}$ and Theorems 3.1-3.2 with $h = 0$. $\square$

Proof of Lemma 3.2. The assertions are a consequence of (3.20) as well as Theorems 3.1-3.2 with $h = 0$, where (3.25) and (3.27) also include (3.23). $\square$

Proof of Theorem 3.5. From the definition of $\breve{\mathfrak{c}}'^{\dagger}_j$ in (3.16) follows

$$\langle\breve{\mathfrak{c}}'^{\dagger}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}\langle\hat{\mathfrak{c}}'_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m} = 1 - ||\breve{\mathfrak{c}}'_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} + \frac{1}{4}||\breve{\mathfrak{c}}'_j - \mathfrak{c}_j||^4_{\mathcal{U}^m} + \langle\hat{\mathfrak{c}}'_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}\sum_{i=1}^{\infty}\frac{\zeta_i\langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}}{i^2 M}$$

where for the last term holds due to independence of given random variables, $\mathbb{E}|\langle \hat{\mathfrak{c}}_j', \mathfrak{c}_j\rangle_{\mathcal{U}^m}| \leq 1, \zeta_i \sim \mathcal{N}(0,1)$ for all i and the monotone convergence theorem:

$$\mathbb{E}\Big[|\langle \hat{\mathfrak{c}}_j', \mathfrak{c}_j\rangle_{\mathcal{U}^m}| \sum_{i=1}^{\infty} \frac{|\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}|}{i^2 M} \Big] = \mathbb{E}|\langle \hat{\mathfrak{c}}_j', \mathfrak{c}_j\rangle_{\mathcal{U}^m}| \sum_{i=1}^{\infty} \frac{|\langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}|\,\mathbb{E}|\zeta_i|}{i^2 M} = \mathrm{O}(M^{-1}) \quad \text{for } M \to \infty.$$

Thus, with (3.28) and $1-\operatorname{sgn}(1+X_n) = \mathrm{o}(a_n)$ for real-valued processes $(X_n)_n$ with $X_n = \mathrm{o}_{\mathbb{P}}(1)$ and real-valued zero sequences $(a_n)_n$, for any j indeed holds for $M \to \infty$,

$$\begin{aligned} ||\breve{\mathfrak{c}}_j'^{\ddagger} - \mathfrak{c}_j||^2_{\mathcal{U}^m} &= ||\breve{\mathfrak{c}}_j' - \mathfrak{c}_j||^2_{\mathcal{U}^m} + 2\big[1 - \operatorname{sgn}(\langle \breve{\mathfrak{c}}_j'^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m} \langle \hat{\mathfrak{c}}_j', \mathfrak{c}_j\rangle_{\mathcal{U}^m})\big] \\ &= \mathrm{O}_{\mathbb{P}}(\gamma_{j,m}^2 m^2 p M^{-1}) + 2\Big[1 - \operatorname{sgn}\big(1 + \mathrm{O}_{\mathbb{P}}(\gamma_{j,m}^2 m^2 p M^{-1}) + \mathrm{O}_{\mathbb{P}}(\gamma_{j,m}^4 m^4 p^2 M^{-2}) + \mathrm{O}_{\mathbb{P}}(M^{-2})\big)\Big] \\ &= \mathrm{O}_{\mathbb{P}}(m^2 p M^{-1}). \end{aligned}$$

Similarly, with (3.25), we also obtain

$$\begin{aligned} \sup_{j\leq k} ||\breve{\mathfrak{c}}_j'^{\ddagger} - \mathfrak{c}_j||^2_{\mathcal{U}^m} &\leq \sup_{j\leq k} ||\breve{\mathfrak{c}}_j' - \mathfrak{c}_j||^2_{\mathcal{U}^m} + \sup_{j\leq k} 2\big[1 - \operatorname{sgn}(\langle \breve{\mathfrak{c}}_j'^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m} \langle \hat{\mathfrak{c}}_j', \mathfrak{c}_j\rangle_{\mathcal{U}^m})\big] \\ &= \mathrm{O}_{\mathbb{P}}(m^2 p M^{-1}) \quad \text{for } M \to \infty. \end{aligned}$$

Moreover, due to the definition of $\breve{\mathfrak{c}}_j^{\dagger}$ in (3.16),

$$\begin{aligned} \mathbb{E}\big[1 - \langle \breve{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m} \langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}\big] &\leq \mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} + \mathbb{E}\Big[|\langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}| \sum_{i=1}^{\infty} \frac{|\zeta_i \langle u_i, \mathfrak{c}_j\rangle_{\mathcal{U}^m}|}{i^2 M} \Big] \\ &= \mathrm{O}(m^2 p M^{-1}) \quad \text{for } M \to \infty. \end{aligned}$$

Thus, for any j holds due to the definition of $\breve{\mathfrak{c}}_j$ and $\breve{\mathfrak{c}}_j^{\ddagger}$, and because of (3.26):

$$\begin{aligned} \mathbb{E}||\breve{\mathfrak{c}}_j^{\ddagger} - \mathfrak{c}_j||^2_{\mathcal{U}^m} &\leq 2\,\mathbb{E}||\breve{\mathfrak{c}}_j - \mathfrak{c}_j||^2_{\mathcal{U}^m} + 2\,\mathbb{E}\big(\big[\operatorname{sgn}\langle \breve{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m} - \operatorname{sgn}\langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m}\big]^2\big) \\ &\leq \mathrm{O}(m^2 p M^{-1}) + 4\,\mathbb{P}\big(1 - \langle \breve{\mathfrak{c}}_j^{\dagger}, \mathfrak{c}_j\rangle_{\mathcal{U}^m} \langle \hat{\mathfrak{c}}_j, \mathfrak{c}_j\rangle_{\mathcal{U}^m} > 1/2\big) \quad \text{for } M \to \infty \\ &= \mathrm{O}(m^2 p M^{-1}) \quad \text{for } M \to \infty. \end{aligned}$$

Hence, (3.30) is verified, and a similar procedure leads with (3.27) to (3.31). □

Proof of Lemma 4.1. The definition of Z_N and $\tilde{Z}_N$ for any $N \in \mathbb{N}$ yields $Z_N - \tilde{Z}_N = A^N(Z_0 - \tilde{Z}_0)$, and submultiplicity of the operator norm thus

$$\mathbb{E}||Z_N - \tilde{Z}_N||^{\nu}_{\mathcal{H}} \leq ||A||^N_{\mathcal{L}_{\mathcal{H}}} \mathbb{E}||Z_0 - \tilde{Z}_0||^{\nu}_{\mathcal{H}}.$$

Since $(Z_k)_{k\in\mathbb{Z}}, (\tilde{Z}_k)_{k\in\mathbb{Z}}$ are $L^{\nu}_{\mathcal{H}}$-processes because $(\varepsilon_k)_{k\in\mathbb{Z}}$ is one and due to the definition of Z_k and $\tilde{Z}_k$ for all k, the expected value on the right is finite. By choosing $\rho \coloneqq ||A||_{\mathcal{L}_{\mathcal{H}}} < 1$, the assertion is proven. □

Acknowledgements

I thank Alexander Meister (University of Rostock), Siegfried Hörmann (Graz University of Technology), Gregory Rice (University of Waterloo), Piotr Kokoszka (Colorado State University) and Dominik Liebl (University of Bonn) for valuable 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. & Klepsch J. (2017) Estimating functional time series by moving average model fitting. **arXiv:** `1701.00770v1`

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

[3] 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

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

[5] 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.

[6] Caponera A. (2021) SPHARMA approximations for stationary functional time series on the sphere. *Stat. Inference Stoch. Process.* **URL:** `https://doi.org/10.1007/s11203-021-09244-6`

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

[8] 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.

[9] Chen Y., Koch T., Lim K.G., Xu X. & Zakiyeva N. (2020) A review study of functional autoregressive models with application to energy forecasting. *Wiley interdisciplinary Reviews: Computational Statistics*, e1525.

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

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

[12] 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

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

[14] Guillas S. (2001) Rates of convergence of autocorrelation estimates for autoregressive Hilbertian processes. *Statist. Probab. Lett.*, **55**(3), 281–291. MR1867531

[15] Hael M.A. (2020) Modeling of rainfall variability using functional principal component method: A case study of Taiz region, Yemen. *Model. Earth Syst. Environ.*

[16] 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

[17] Horváth L. & Kokoszka P. (2012). *Inference for Functional Data with Applications.* New York: Springer. MR2920735

[18] Hörmann S., Horváth L. & Reeder R. (2013) A Functional Version of the ARCH Model. *Econ. Theory*, **29**(2), 267–288. MR3042756

[19] Hörmann S. & Kokoszka P. (2010) Weakly dependent functional data. *Ann. Statist.*, **38**, 1845–1884. MR2662361

[20] Hsing T. & Eubank R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators.* West Sussex: Wiley. MR3379106

[21] 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

[22] Jarry G., Delahaye D., Nicol F. & Feron E. (2020) Aircraft atypical approach detection using functional principal component analysis. *J.A.T.M.*, **84**, p.101787.

[23] Klepsch J., Klüppelberg C. & Wei T. (2017) Prediction of functional ARMA processes with an application to traffic data. *Econom. Stat.*, **1**, 128–149. MR3669993

[24] Kokoszka P. & Reimherr M. (2013) Asymptotic normality of the principal components of functional time series. *Stoch. Process. Appl.*, **123**(5), 1546–1562. MR3027890

[25] 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 September 2021 doi:10.20944/preprints202105.0277.v2

[26] Kokoszka P., Stoev S. & Xiong Q. (2019) Principal components analysis of regularly varying functions. *Bernoulli*, **25**(4B), 3864–3882. MR4010975

[27] Kühnert S. (2019) *Über funktionale ARCH- und GARCH-Zeitreihen* (German) [About functional ARCH and GARCH processes]. Doctoral Thesis, University of Rostock.

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

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

[30] 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.

[31] 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

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

[33] 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

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

[35] Ramsay J.O. & Silverman B.W. (2005) *Functional data analysis* (2 ed.). Springer Series in Statistics. Springer, New York. MR2168993

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

[37] 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.

[38] Sarkar S. & Panaretos V.M. (2021) CovNet: Covariance Networks for Functional Data on Multidimensional Domains. **arXiv:** 2104.05021v1

[39] 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

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

[41] Stout W.F. (1974). *Almost Sure Convergence*. New York: Academic Press. MR0455094

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

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

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

[45] 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.