Functional ARCH and GARCH Models: A Yule-Walker Approach

Sebastian Kühnert*

January 31, 2020

Abstract

Conditional heteroskedastic financial time series are commonly modelled by ARCH and GARCH. ARCH(1) and GARCH processes were recently extended to the function spaces $C[0,1]$ and $L^2[0,1]$, their probabilistic features were studied and their parameters were estimated. The projections of the operators on finite-dimensional subspace were estimated, as were the complete operators in GARCH$(1,1)$. An explicit asymptotic upper bound of the estimation errors was stated in ARCH(1). This article provides sufficient conditions for the existence of strictly stationary solutions, weak dependence and finite moments of ARCH and GARCH processes in various $L^p[0,1]$ spaces, $C[0,1]$ and other spaces. In $L^2[0,1]$ we deduce explicit asymptotic upper bounds of the estimation errors for the shift term and the complete operators in ARCH and GARCH and for the projections of the operators on a finite-dimensional subspace in ARCH. The operator estimaton is based on Yule-Walker equations. The estimation of the GARCH operators also involves a result concerning the estimation of the operators in invertible, linear processes which is valid beyond the scope of ARCH and GARCH. Through minor modifications, all results in this article regarding functional ARCH and GARCH can be transferred to functional ARMA.

1   Introduction

Volatility, usually measured by the variance, is one of the essential objects of study of financial time series. These are often strictly stationary but conditional heteroskedastic, where latter means that the variances at any time conditioned on the past are non-constant and randomly changing. A popular model exhibiting this phenomenon is the autoregressive conditional heteroskedasticity (ARCH) model established by Engle (1982), for which he was awarded the noble prize in economics in 2003. This model was extended to the generalized ARCH (GARCH) model by Bollerslev (1986). Various authors established modifications of univariate and multivariate ARCH and GARCH processes, studied their probabilistic properties and estimated their parameters. An excellent overview and applications of such processes is provided in Andersen (2009), Francq & Zakoïan (2010) and Gouriéroux (1997). Due to a progress in processing techniques and since high-resolution tick data are accessible and can be described as functions, it seems reasonable to extend these models on infinite-dimensional spaces, enabling the analysis to be more accurate. From a mathematical point of view, such an extension is unproblematic for complete, separable metric spaces M since completeness of M implies that the Borel σ-field $\mathbb{B}(M)$ is well defined and separability ensures that e.g. sums of random variables remain random variables, see Ledoux & Talagrand (1991). For a detailed introduction in Functional Data and Functional Time Series Analysis, the areas dealing with random variables resp. time series with values in an infinite-dimensional space, see Bosq (2000), Ferraty & Vieu (2006), Hsing & Eubank (2015) and Ramsay & Silverman (1997). For a compact synopsis (in German), see Kühnert (2019).

---

*Institute of Mathematics, University of Rostock, D-18051 Rostock, Germany, EMAIL: s.kuehnert_math@gmx.de

Hörmann *et al.* (2013) made the initial step by introducing ARCH(1) processes with values in the spaces $C[0,1]$ and $L^2[0,1]$ of continuous resp. of square-integrable real valued functions with domain $[0,1]$. They established sufficient conditions for the existence of strictly stationary solutions, finite moments and weak dependence. In $L^2[0,1]$ they constructed consistent estimators and stated explicit asymptotic upper bounds of the estimation errors for the shift term and the projections of the operator on finite-dimensional subspaces by assuming the operator to be an integral operator and estimating its kernel. Aue *et al.* (2017) established GARCH(1,1) processes in $C[0,1]$ and in $L^2[0,1]$ and found sufficient conditions for the existence of strictly stationary solutions, finite moments and weak dependence. In $L^2[0,1]$ they derived a consistent least squares estimator for the projections of the parameters on finite-dimensional subspaces but stating an explicit asymptotic upper bound of the estimation errors. At last, Cerovecki *et al.* (2019) studied $L^2[0,1]$-valued GARCH($\mathfrak{p},\mathfrak{q}$) processes for integers $\mathfrak{p},\mathfrak{q} > 0$. They developed sufficient conditions for the existence of strictly stationary solutions and finite moments. By a quasi-likelihood approach, the projections of the parameters on a finite-dimensional subspace and only for $\mathfrak{p} = 1 = \mathfrak{q}$ the complete operators were estimated consistently. In both cases, no explicit asymptotic upper bound of the estimation errors was stated. Hörmann *et al.* (2013), Aue *et al.* (2017) and Cerovecki *et al.* (2019) also provided simulation studies showing how their models matched with real data and illustrated possible applications. For further work dealing with functional ARCH and GARCH models, see Kokoszka *et al.* (2017) and Rice *et al.* (2019).

In this article, we establish ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p},\mathfrak{q}$) processes for all $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ with values in $L^p[0,1]$ with $p \in [1,\infty), C[0,1]$ and other spaces. We provide sufficient conditions for the existence of strictly stationary solutions, weak dependence and moments of these processes under mild conditions. The focus of this paper is on deducing estimators for the shift term and the complete operators of $L^2[0,1]$-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p},\mathfrak{q}$) processes for any $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ and on deriving explicit asymptotic upper bounds of their estimation errors. We also deduce explicit asymptotic upper bounds of the estimation errors for the operators on a finite-dimensional subspace of these ARCH processes. The operator estimation is always based on Yule-Walker equations and the estimators for the GARCH operators also involve estimators for the operators of invertible, linear processes represented as inverted time series. We derive explicit asymptotic upper bounds of their estimation errors. Also, this upper bound holds for the estimation errors when estimating the operators in the associated linear process and is valid beyond the context of functional ARCH and GARCH models. All results in this article regarding functional ARCH and GARCH can be transferred to functional ARMA processes due to their relationship.

In this paper, we use the following notation. $a \wedge b := \min(a,b)$ and $a \vee b := \max(a,b)$ for $a,b \in \mathbb{R}$. For functions $f,g\colon \mathbb{D} \subseteq \mathbb{R} \to \mathbb{R}$, we write $f \propto g$ resp. $f \precsim g$ if there is a $c \in \mathbb{R}$ with $f(x) = cg(x)$ resp. $f(x) \leq cg(x)$ for all $x \in \mathbb{D}$. For sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \subseteq (0,\infty)$, we write $a_n \sim b_n$ if $\frac{a_n}{b_n} \to 1, a_n \asymp b_n$ if $a_n \sim cb_n$ for some $c \neq 0, a_n = \omega(b_n)$ if $b_n = o(a_n)$ (for $n \to \infty$) and $a_n = \Omega(b_n)$ if $b_n = O(a_n)$ (for $n \to \infty$). Further, $\Xi(a_n, b_n) := \omega(a_n) \cap o(b_n), \Xi[a_n, b_n) := \Omega(a_n) \cap o(b_n), \Xi(a_n, b_n] := \omega(a_n) \cap O(b_n)$ and $\Xi[a_n, b_n] := \Omega(a_n) \cap O(b_n)$. By 0_V we denote the identity element of addition of a vector space V and $V^n := \{(v_1,...,v_n)^T | v_1,...,v_n \in V\}$, with $n \in \mathbb{N}$, becomes a vector space by our componentwise definition of scalar multiplication and vector addition. For a space F of functions $f\colon [0,1] \to \mathbb{R}, F_{>0}$ and $F_{\geq 0}$ denote the sets of functions $f \in F$ with $f(t) > 0$ resp. $f(t) \geq 0$ for λ-a.e. $t \in [0,1]$ where λ is the Lebesgue-Borel measure on $[0,1]$, and $f \odot g$ denotes the pointwise product of $f,g \in F$ if it is well-defined. Let $(\mathcal{B}, ||\cdot||_{\mathcal{B}}), (\mathcal{B}', ||\cdot||_{\mathcal{B}'})$ be Banach spaces and $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}), (\mathcal{H}', \langle\cdot,\cdot\rangle_{\mathcal{H}'})$ be Hilbert spaces. On Hilbert spaces we always use norms induced by their scalarproduct and we say CONS for a complete orthonormal system. We endow Banach spaces $(\mathcal{B}^n, ||\cdot||_{\mathcal{B}^n})$ with the norm $||\boldsymbol{b}||_{\mathcal{B}^n}^2 := \sum_{i=1}^n ||b_i||_{\mathcal{B}}^2$ where $\boldsymbol{b} := (b_1,...,b_n)^T \in \mathcal{B}^n$ and Hilbert spaces $(\mathcal{H}^n, \langle\cdot,\cdot\rangle_{\mathcal{H}^n})$ with the scalarproduct $\langle\boldsymbol{h}, \tilde{\boldsymbol{h}}\rangle_{\mathcal{H}^n} := \sum_{i=1}^n \langle h_i, \tilde{h}_i\rangle_{\mathcal{H}}$ where $\boldsymbol{h} := (h_1,...,h_n)^T, \tilde{\boldsymbol{h}} := (\tilde{h}_1,...,\tilde{h}_n)^T \in \mathcal{H}^n$. We write $\mathcal{L}_{\mathcal{B},\mathcal{B}'}, \mathcal{K}_{\mathcal{B},\mathcal{B}'}, \mathcal{S}_{\mathcal{B},\mathcal{B}'}$ resp. $\mathcal{N}_{\mathcal{B},\mathcal{B}'}$ for the space of bounded, compact, Hilbert-Schmidt resp. nuclear operators from $\mathcal{B}$ to $\mathcal{B}'$ with $\mathcal{L}_{\mathcal{B}} = \mathcal{L}_{\mathcal{B},\mathcal{B}}, \mathcal{K}_{\mathcal{B}} = \mathcal{K}_{\mathcal{B},\mathcal{B}}, \mathcal{S}_{\mathcal{B}} = \mathcal{S}_{\mathcal{B},\mathcal{B}}$ and $\mathcal{N}_{\mathcal{B}} = \mathcal{N}_{\mathcal{B},\mathcal{B}}$ where the term operator always refers to a linear mapping. Further, K^* denotes the adjoint of an operator $K \in \mathcal{L}_{\mathcal{B},\mathcal{B}'}$ and we write $h \otimes h' := \langle h, \cdot\rangle_{\mathcal{H}} h'$ for $h \in \mathcal{H}, h' \in \mathcal{H}'$. In all respects, we assume our random elements to be defined on some common probability space $(\Omega, \mathfrak{A}, \mathbb{P})$. For $\mathcal{B}$-valued processes $(X_k)_{k\in\mathbb{Z}}$ and $(Y_k)_{k\in\mathbb{Z}}, X_n = O_{\mathbb{P}}(Y_n)$ (for $n \to \infty$) denotes that $(X_k/Y_k)_k$ is asymptotically $\mathbb{P}$-stochastic bounded. For $p \in [1,\infty)$ we denote by

$L^p_{\mathcal{B}} = L^p_{\mathcal{B}}(\Omega, \mathfrak{A}, \mathbb{P})$ the space of (classes of) $\mathcal{B}$-valued random variables X with $\nu_{p,\mathcal{B}}(X) := (\mathbb{E}||X||^p_{\mathcal{B}})^{1/p} < \infty$, we call a process $(X_k)_{k \in \mathbb{Z}}$ of $\mathcal{B}$-valued random variables $L^p_{\mathcal{B}}$-process if $X_k \in L^p_{\mathcal{B}}$ for all k and centered if $\mathbb{E}(X_k) = 0_{\mathcal{B}}$ for all k with expectation in Bochner-integral sense, see Hsing & Eubank (2015), p. 40–45.

The rest of this article is organized as follows. Section 2 studies probabilistic features of our ARCH and GARCH processes. Section 3 introduces our parameter estimators and derives asymptotic upper bounds of the estimation errors. Section 4 summarizes the main results, delineates these from similar results in other papers and gives an outline for future research. Section 5 contains proofs.

2 Functional ARCH and GARCH models

We start with the definition of F-valued ARCH and GARCH processes where F stands for $L^p[0,1]$ for some $p \in [1, \infty)$ resp. for a separable Banach space of functions $f \colon [0,1] \to \mathbb{R}$ being complete w.r.t. the sup-norm $||\cdot||_\infty$ and closed w.r.t. the pointwise product $\odot$. Hence, $||\cdot||_F$ is either $||\cdot||_\infty$ or the norm of $L^p[0,1]$ defined by $||f||^p_{L^p[0,1]} := \int_0^1 |f(t)|^p \, dt$ for $f \in L^p[0,1]$ with $\langle f, g \rangle_{L^2[0,1]} := \int_0^1 f(t)g(t) \, dt$ for any $f,g \in L^2[0,1]$ where integration is meant w.r.t. the Lebesgue-Borel measure λ on $[0,1]$.

Definition 2.1. *Let $\mathfrak{p} \in \mathbb{N}, \mathfrak{q} \in \mathbb{N}_0$, let $(\varepsilon_k)_{k \in \mathbb{Z}}$ be an i.i.d. F-valued time series, $\delta \in F_{>0}$ and $\alpha_i, \beta_j \in \mathcal{L}_F$ be operators with $\alpha_i, \beta_j \colon F_{\geq 0} \to F_{\geq 0}$ for all i, j. Then, if*

$$\mathscr{X}_k = \varepsilon_k \odot \sigma_k, \qquad \sigma_k^2 = \delta + \sum_{i=1}^{\mathfrak{p}} \alpha_i(\mathscr{X}_{k-i}^2) + \sum_{j=1}^{\mathfrak{q}} \beta_j(\sigma_{k-j}^2) \tag{2.1}$$

holds a.s. for all k, $(\mathscr{X}_k)_k$ is called F-valued ARCH($\mathfrak{p}$) resp. GARCH($\mathfrak{p}, \mathfrak{q}$) process if $\mathfrak{q} = 0$ and $\alpha_{\mathfrak{p}} \neq 0_{\mathcal{L}_{\mathscr{H}}}$ resp. if $\mathfrak{q} \in \mathbb{N}$ and $\alpha_{\mathfrak{p}} \neq 0_{\mathcal{L}_{\mathscr{H}}} \neq \beta_{\mathfrak{q}}$.

Throughout, $\alpha_i, \beta_j, \delta, \varepsilon_k, \mathfrak{p}, \mathfrak{q}, \sigma_k, \mathscr{X}_k$ are the variables in (2.1) with $\alpha_i = 0_{\mathcal{L}_{\mathscr{H}}} = \beta_j$ for $i > \mathfrak{p}, j > \mathfrak{q}$ and moreover $\mathfrak{r} := \max(\mathfrak{p}, \mathfrak{q}), \mathfrak{s} := \mathfrak{p} + \mathfrak{q}$. The equations in (2.1) lead to the state-space form

$$\boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})} = \boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q})} + \boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\varsigma}_{k-1}^{(\mathfrak{p},\mathfrak{q})}) \tag{2.2}$$

$$:\Longleftrightarrow
\begin{bmatrix}
\mathscr{X}_k^2 \\
\mathscr{X}_{k-1}^2 \\
\vdots \\
\mathscr{X}_{k-\mathfrak{p}+2}^2 \\
\mathscr{X}_{k-\mathfrak{p}+1}^2 \\
\sigma_k^2 \\
\sigma_{k-1}^2 \\
\vdots \\
\sigma_{k-\mathfrak{q}+2}^2 \\
\sigma_{k-\mathfrak{q}+1}^2
\end{bmatrix}
=
\begin{bmatrix}
\delta \odot \varepsilon_k^2 \\
0_F \\
\vdots \\
0_F \\
0_F \\
\delta \\
0_F \\
\vdots \\
0_F \\
0_F
\end{bmatrix}
+
\begin{bmatrix}
\boxdot_k \alpha_1 & \cdots & \cdots & \cdots & \boxdot_k \alpha_{\mathfrak{p}} & \boxdot_k \beta_1 & \cdots & \cdots & \cdots & \boxdot_k \beta_{\mathfrak{q}} \\
\mathbb{I}_F & 0_{\mathcal{L}_F} & \cdots & \cdots & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \cdots & \cdots & 0_{\mathcal{L}_F} \\
0_{\mathcal{L}_F} & \mathbb{I}_F & 0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \ddots & \ddots & \cdots & 0_{\mathcal{L}_F} \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} & \mathbb{I}_F & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} \\
\alpha_1 & \cdots & \cdots & \cdots & \alpha_{\mathfrak{p}} & \beta_1 & \cdots & \cdots & \cdots & \beta_{\mathfrak{q}} \\
0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \cdots & \cdots & 0_{\mathcal{L}_F} & \mathbb{I}_F & 0_{\mathcal{L}_F} & \cdots & \cdots & 0_{\mathcal{L}_F} \\
0_{\mathcal{L}_F} & \ddots & \ddots & \cdots & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \mathbb{I}_F & 0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\
0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & 0_{\mathcal{L}_F} & \cdots & 0_{\mathcal{L}_F} & \mathbb{I}_F & 0_{\mathcal{L}_F}
\end{bmatrix}
\begin{bmatrix}
\mathscr{X}_{k-1}^2 \\
\mathscr{X}_{k-2}^2 \\
\vdots \\
\mathscr{X}_{k-\mathfrak{p}+1}^2 \\
\mathscr{X}_{k-\mathfrak{p}}^2 \\
\sigma_{k-1}^2 \\
\sigma_{k-2}^2 \\
\vdots \\
\sigma_{k-\mathfrak{q}+1}^2 \\
\sigma_{k-\mathfrak{q}}^2
\end{bmatrix}$$

with $\boxdot_k \colon \dot{F} \to F, f \mapsto f \odot \varepsilon_k^2$ where

$$\dot{F} := \begin{cases} L^{2p}, & \text{if } F = L^p, \\ F, & \text{if } F \text{ is a Banach space w.r.t. } ||\cdot||_\infty. \end{cases}$$

Furthermore, if

$$\mathbb{E}\ln^+ ||\varepsilon_0^2||_{\dot{F}} < \infty \tag{2.3}$$

where $\ln^+(\cdot) := \ln(\max(1, \cdot))$, then $||\varepsilon_0^2||_{\dot{F}} < \infty$ a.s. and hence $\boxdot_k$ is a bounded operator a.s. with

$$||\boxdot_k||_{\mathcal{L}_{\dot{F}, F}} \leq ||\varepsilon_k^2||_{\dot{F}} \quad \text{a.s.} \tag{2.4}$$

3

Consequently, $\varsigma_k^{(\mathfrak{p},\mathfrak{q})}, \delta_k^{(\mathfrak{p},\mathfrak{q})} \in F^{\mathfrak{s}}$ a.s. and we have $\Psi_k^{(\mathfrak{p},\mathfrak{q})} \in \mathcal{L}_{F^{\mathfrak{s}}}$ a.s. if $\delta, \varepsilon_k^2, \sigma_{k-\mathfrak{q}}^2, ..., \sigma_k^2 \in \dot{F}$ a.s. Moreover, (2.3) and $|| \cdot ||_{\mathcal{L}_F} \le || \cdot ||_{\mathcal{L}_{F,\dot{F}}}$ imply (see Kühnert (2019), p.28)

$$\mathbb{E}\ln^+ ||\Psi_0^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}} < \infty. \tag{2.5}$$

Since $(\Psi_k^{(\mathfrak{p},\mathfrak{q})})_k$ is i.i.d. and $|| \cdot ||_{\mathcal{L}_{F^{\mathfrak{s}}}}$ is sub-multiplicative, according to Kingman (1974), Theorem 6 we have

$$\gamma^{(\mathfrak{p},\mathfrak{q})} := \lim_{k \to \infty} \frac{1}{k} \mathbb{E}\ln ||\Psi_k^{(\mathfrak{p},\mathfrak{q})} \Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})} \cdots \Psi_1^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}} \tag{2.6}$$

$$= \lim_{k \to \infty} \frac{1}{k} \ln ||\Psi_k^{(\mathfrak{p},\mathfrak{q})} \Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})} \cdots \Psi_1^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}} \quad \text{a.s.} \tag{2.7}$$

where $\gamma^{(\mathfrak{p},\mathfrak{q})}$ is called top Lyapunov exponent of $(\Psi_k^{(\mathfrak{p},\mathfrak{q})})_k$ with $\gamma^{(\mathfrak{p},\mathfrak{q})} \in [-\infty, \infty)$.

Now, we can state sufficient conditions for the existence of nonanticipative, strictly stationary solutions in the ARCH and the GARCH model where a F-valued time series $(Y_k)_{k \in \mathbb{Z}}$ is called nonanticipative w.r.t. another F-valued time series $(\varepsilon_k)_{k \in \mathbb{Z}}$ if there is a measurable function $f : F^\infty \to F$ such that

$$Y_k = f(\varepsilon_k, \varepsilon_{k-1}, ...) \tag{2.8}$$

holds a.s. for all k. If $(\varepsilon_k)_{k \in \mathbb{Z}}$ is strictly stationary and ergodic, which is especially the case if $(\varepsilon_k)_k$ is i.i.d., then (2.8) implies that $(Y_k)_k$ is also strictly stationary and ergodic after Stout (1974), Theorem 3.5.8.

Theorem 2.1. *Let the assumptions in Definition 2.1, $\delta \in \dot{F}_{>0}$ and $\alpha_i, \beta_j \in \mathcal{L}_{F,\dot{F}}$ for all i,j hold.*

(a) *If*

$$\gamma^{(\mathfrak{p},\mathfrak{q})} < 0, \tag{2.9}$$

then the equations in (2.1) have a unique, strictly stationary, nonanticipative w.r.t. $(\varepsilon_k)_k$ and ergodic solution where $\sigma_k^2 = f(\varepsilon_{k-1}, \varepsilon_{k-2}, ...)$ a.s. for all k for some measurable function $f : F^\infty \to F$.

(b) *If there are $n \in \mathbb{N}$ and $\nu > 0$ such that*

$$\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})} := \mathbb{E}||\Psi_n^{(\mathfrak{p},\mathfrak{q})} \Psi_{n-1}^{(\mathfrak{p},\mathfrak{q})} \cdots \Psi_1^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} < 1, \tag{2.10}$$

then (2.9) holds.

Though (2.10) is stricter than (2.9), it is easier to show. Furthermore, (2.10) is useful for the simulation of an initial value of F-valued ARCH and GARCH processes, as we can see in the following.

Corollary 2.1. *Let (2.10) hold for some $n \in \mathbb{N}$ and $\nu > 0$. Further, define $\tilde{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})} := \delta_k^{(\mathfrak{p},\mathfrak{q})} + \Psi_k^{(\mathfrak{p},\mathfrak{q})}(\tilde{\varsigma}_{k-1}^{(\mathfrak{p},\mathfrak{q})})$ for $k \in \mathbb{N}$, where $\tilde{\varsigma}_0^{(\mathfrak{p},\mathfrak{q})} \in F^{\mathfrak{s}}$ is some deterministic value. Then, there is some $\rho \in (0,1)$ with*

$$\mathbb{E}||\varsigma_N^{(\mathfrak{p},\mathfrak{q})} - \tilde{\varsigma}_N^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} = \mathrm{O}(\rho^N). \tag{2.11}$$

Based on ideas in Aue *et al.* (2017), Hörmann *et al.* (2013) and with (2.10), we derive a sufficient condition for the existence of moments and for weak dependence, to be precise L^p-m-approximibility, of F-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p}, \mathfrak{q}$) processes for any $\mathfrak{p}, \mathfrak{q} \in \mathbb{N}$. Finite moments and L^p-m-approximibility are used to estimate the ARCH and GARCH parameters. An F-valued time series $(Y_k)_{k \in \mathbb{Z}}$ is called L_F^p-m-approximable for $p \ge 1$ if $Y_k = f(\varepsilon_k, \varepsilon_{k-1}, ...)$ a.s. for all k for an i.i.d. time series $(\varepsilon_k)_{k \in \mathbb{Z}}$ and a measurable function $f : F^\infty \to F$ (thus $(Y_k)_k$ is nonanticipative w.r.t. $(\varepsilon_k)_k$) and if

$$\sum_{m=1}^{\infty} \nu_{p,\mathcal{B}}(Y_m - Y_m^{(m)}) < \infty \tag{2.12}$$

holds where $\nu_{p,\mathcal{B}}(\cdot) = (\mathbb{E}|| \cdot ||_{\mathcal{B}}^p)^{1/p}$ and $Y_k^{(m)} := f(\varepsilon_k, \varepsilon_{k-1}, ..., \varepsilon_{k-m+1}, \varepsilon_{k-m}^{(k)}, \varepsilon_{k-m-1}^{(k)}, ...)$ for all k, m with independent copies $(\varepsilon_k^{(n)})_{k \in \mathbb{Z}}$ of $(\varepsilon_k)_{k \in \mathbb{Z}}$ for all n. For each m, the sequences $(Y_k^{(m)})_{k \in \mathbb{Z}}$ are strictly stationary, m-dependent and each $Y_k^{(m)}$ equals Y_k in distribution. Moreover, $(Y_k)_{k \in \mathbb{Z}}$ is called geometrically L_F^p-m-approximable if $(Y_k)_k$ is L_F^p-m-approximable and if there exists a $\rho \in (0,1)$ with $\nu_{p,\mathcal{B}}(Y_m - Y_m^{(m)}) = \mathrm{O}(\rho^m)$. For a detailed introduction to L^p-m-approximibility, see Hörmann & Kokoszka (2010).

Lemma 2.1. *Let* $\mathbb{E}||\varepsilon_0^2||_F^\nu < \infty$. *Also, let* (2.10) *hold for some* $\nu > 0$ *and* $n \in \mathbb{N}$. *Then*

(a) $\mathbb{E}||\mathscr{X}_0^2||_F^\nu < \infty$ *and* $\mathbb{E}||\sigma_0^2||_F^\nu < \infty$.

(b) $(\mathscr{X}_k^2)_k$ *is geometrically* L_F^ν-*m-approximable and* $(\sigma_k^2)_k$ *is geometrically* L_F^ν-*m-approximable.*

3 Estimation

In this section, we establish estimators for the parameters of $\mathscr{H}$-valued ARCH and GARCH processes with known orders where $\mathscr{H} := L^2[0,1]$ and we deduce asymptotic upper bounds for their estimation errors. Throughout the section, we write $\dot{\mathscr{H}} := L^4[0,1]$, and, except for in section 3.1, we impose the following.

Assumption 3.1. $\delta \in \dot{\mathscr{H}}_{>0}, \alpha_i, \beta_j \in \mathcal{S}_{\mathscr{H}} \cap \mathcal{L}_{\mathscr{H},\dot{\mathscr{H}}}$ *for all* i *and* $j, \mathbb{E}||\varepsilon_0^2||_{\dot{\mathscr{H}}}^4 < \infty$,

$$\mathbb{E}(\varepsilon_0^2(t)) = 1 \tag{3.1}$$

for λ-*a. e.* $t \in [0,1]$ *and there are* $n \in \mathbb{N}$ *and* $\nu = 4$ *with* (2.10).

Assumption 3.1 implies $\mathbb{E}(\mathscr{X}_k^2(t)) = \mathbb{E}(\sigma_k^2(t))$ for λ-a.e. t and all k. Thus, (2.1) yields

$$\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\mathfrak{r}} (\alpha_i + \beta_i)(\mathscr{Z}_{k-i}) + \sum_{j=1}^{\mathfrak{q}} (-\beta_j)(\nu_{k-j}) \tag{3.2}$$

a.s. for all k where $\alpha_i = 0_{\mathcal{L}_{\mathscr{H}}} = \beta_j$ for $i > \mathfrak{p}, j > \mathfrak{q}, \mathscr{Z}_k := \mathscr{X}_k^2 - m_2$ with $m_2 := \mathbb{E}(\mathscr{X}_1^2)$ and $\nu_{k-j} := \mathscr{X}_{k-j}^2 - \sigma_{k-j}^2$. Hence, $\boldsymbol{\mathscr{Z}} = (\mathscr{Z}_k)_k$ is a $\mathscr{H}$-valued AR($\mathfrak{p}$) resp. ARMA($\mathfrak{r},\mathfrak{q}$) time series if $\mathfrak{q} = 0$ resp. $\mathfrak{q} \in \mathbb{N}$ with time series of innovations $\boldsymbol{\nu} := (\nu_k)_k$ which is not i.i.d. but stationary. Moreover, both $\boldsymbol{\mathscr{Z}}$ and $\boldsymbol{\nu}$ are centered, stationary, nonanticipative w.r.t. $(\varepsilon_k)_k$ and geometrically $L_{\mathscr{H}}^4$-m-approximable.

For the estimation of the operators in (2.1), we use (3.2) and we impose the following.

Assumption 3.2. $\delta \in \ell^\infty[0,1], \alpha_i, \beta_j \in \mathcal{L}_{\mathscr{H},\ell^\infty[0,1]}$ *for all* i *and* $j, ||\Gamma_{\mathfrak{p},\mathfrak{q}}||_{\mathcal{L}_{\mathscr{H}}} < 1$ *where* $\Gamma_{\mathfrak{p},\mathfrak{q}} := \sum_{i=1}^{\mathfrak{p}} \alpha_i + \sum_{j=1}^{\mathfrak{q}} \beta_j$, *and there is no closed, affine subspace* $U \subsetneq \mathscr{H}$ *with* $\mathbb{P}(\varepsilon_0^2 \in U) = 1$.

Lemma 3.1. *Let Assumptions 3.1-3.2 hold. Then, there is no closed subspace* $V \subsetneq \mathscr{H}$ *with* $\mathbb{P}(\mathscr{X}_0^2 \in V) = 1, \mathbb{P}(\nu_0 \in V) = 1, \mathbb{P}(\mathscr{Z}_0 \in V) = 1$ *and the operators* $\mathscr{C}_{0;\boldsymbol{\varepsilon}^2}, \mathscr{C}_{0;\boldsymbol{\mathscr{X}}^2}, \mathscr{C}_{0;\boldsymbol{\nu}}$ *and* $\mathscr{C}_{0;\boldsymbol{\mathscr{Z}}}$ *are injective.*

3.1 Preliminaries

Here, in order to estimate the parameters in (2.1), we state certain assumptions and establish various convergence results dealing with the asymptotic behaviour of estimation errors of specific eigenvalues and expected values, operators and eigenfunctions in Hilbert spaces $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}}), (\mathcal{H}', \langle \cdot, \cdot \rangle_{\mathcal{H}'})$ and $(\mathcal{H}'', \langle \cdot, \cdot \rangle_{\mathcal{H}''})$. We also discuss the estimation of operators within a composition of operators.

3.1.1 Estimation of expected values, lag-h-covariance and other operators

Firstly, we define lag-h-covariance operators and their empirical versions.

Definition 3.1. *Let* $\boldsymbol{X} = (X_k)_{k \in \mathbb{Z}}$ *be a stationary* $L_{\mathcal{H}}^2$-*valued time series and let* $h \in \mathbb{Z}$. *Then, the lag-h-covariance operator of* $\boldsymbol{X}$ *is defined by*

$$\mathscr{C}_h = \mathscr{C}_{h;\boldsymbol{X}} := \mathbb{E}[(X_0 - m_1) \otimes (X_h - m_1)] \tag{3.3}$$

where $m_1 = m_1(\boldsymbol{X}) := \mathbb{E}(X_1)$ *and the empirical lag-h-covariance operator of* $\boldsymbol{X}$ *is defined by*

$$\hat{\mathscr{C}}_h = \hat{\mathscr{C}}_{h;\boldsymbol{X}} := \begin{cases} \frac{1}{N_h - 1} \sum_{k=|h|+1}^{N_h} (X_k - \hat{m}_1) \otimes (X_{k+h} - \hat{m}_1), & 1 - N < h < 0, \\ \frac{1}{N_h - 1} \sum_{k=1}^{N_h} (X_k - \hat{m}_1) \otimes (X_{k+h} - \hat{m}_1), & 0 \le h < N - 1 \end{cases} \tag{3.4}$$

where $\hat{m}_1 = \hat{m}_1(\boldsymbol{X}) := N_h^{-1} \sum_{i=1}^{N_h} X_i$ *and* $N_h := N - |h|$ *with* $N \in \mathbb{N}$ *and* $|h| < N - 1$. *The operators* $\mathscr{C}_0$ *and* $\hat{\mathscr{C}}_0$ *are also called covariance operator resp. empirical covariance operator.*

5

$\hat{\mathscr{C}}_h$ are bounded operators with finite-dimensional image with $\hat{\mathscr{C}}_h^* = \hat{\mathscr{C}}_{-h}$ for all h. Furthermore, $\hat{\mathscr{C}}_0$ is selfadjoint and positive semi-definite. We obtain the following convergence rates.

Lemma 3.2. *Let* $\boldsymbol{X} = (X_k)_{k \in \mathbb{Z}}$ *be a* $L_{\mathcal{H}}^4$*-m-approximable time series. Then*

$$\hat{m}_l = \hat{m}_l(\boldsymbol{X}) := N^{-1} \sum_{i=1}^{N} X_i^l \tag{3.5}$$

is an unbiased esimator for $m_l = m_l(\boldsymbol{X}) := \mathbb{E}(X_1^l)$ *for any* $l = 1, 2$ *and* $N \in \mathbb{N}$ *with*

$$\mathbb{E}||\hat{m}_l - m_l||_{\mathcal{H}}^2 = \mathrm{O}(N^{-1}). \tag{3.6}$$

Lemma 3.3. *Let* $\boldsymbol{X} = (X_k)_{k \in \mathbb{Z}}$ *be a* $L_{\mathcal{H}}^4$*-m-approximable time series. Then*

$$||\hat{\mathscr{C}}_h - \mathscr{C}_h||_{\mathcal{S}_{\mathcal{H}}}^2 = \begin{cases} \mathrm{O}_{\mathbb{P}}(N^{-1}), & \text{if } h \in \mathbb{Z} \text{ is fixed,} \\ \mathrm{O}_{\mathbb{P}}(hN^{-1}), & \text{if } h = h_N = \Xi(1, N). \end{cases} \tag{3.7}$$

Based on ideas in Aue *et al.* (2010), for centered time series $\boldsymbol{X} = (X_k)_{k \in \mathbb{Z}}$ we define the operators

$$\mathfrak{S}_{d,m} := \mathscr{C}_{X_d(d), X_{d+m}} = \mathbb{E}[X_d(d) \otimes X_{d+m}] \qquad \text{and} \qquad \mathfrak{S}_d := \mathscr{C}_{0; \boldsymbol{X}(d)} = \mathbb{E}[X_d(d) \otimes X_d(d)]$$

where $d \in \mathbb{N}, m \in \mathbb{Z}$ with $X_k(d) := (X_k, X_{k-1}..., X_{k-d+1})^T \in \mathcal{H}^d$ for $k \in \mathbb{Z}$ and $\boldsymbol{X}(d) := (X_k(d))_{k \in \mathbb{Z}}$. These operators satisfy $\mathfrak{S}_{d,m} \in \mathcal{N}_{\mathcal{H}^d, \mathcal{H}}$ and $\mathfrak{S}_d \in \mathcal{N}_{\mathcal{H}^d}$ as well as (see Kühnert (2019), p. 56)

$$||\mathfrak{S}_{d,m}||_{\mathcal{N}_{\mathcal{H}^d, \mathcal{H}}} \leq \sqrt{d}\, \mathbb{E}||X_0||_{\mathcal{H}}^2 \qquad \text{and} \qquad ||\mathfrak{S}_d||_{\mathcal{N}_{\mathcal{H}^d}} = d\, \mathbb{E}||X_0||_{\mathscr{H}}^2. \tag{3.8}$$

Also, given a sample $X_1, ..., X_N$ of $\boldsymbol{X}$ with $N > d$, the operators

$$\hat{\mathfrak{S}}_{d,1} := \frac{1}{N_d - 1} \sum_{k=1}^{N_d} \left(X_{k+d-1}^2(d) - \hat{m}_1(\boldsymbol{X}^2(d)) \right) \otimes \left(X_{k+d}^2 - \hat{m}_1(\boldsymbol{X}^2) \right), \tag{3.9}$$

with $N_d := N - d, \hat{m}_1(\boldsymbol{X}^2(d)) := N_d^{-1} \sum_{i=1}^{N_d} X_{d+i-1}^2(d)$ and $\hat{m}_1(\boldsymbol{X}^2) := N_d^{-1} \sum_{j=1}^{N_d} X_{d+j}^2$, satisfy (see Kühnert (2019), Definition and properties 4.36)

$$||\hat{\mathfrak{S}}_{d,1} - \mathfrak{S}_{d,1}||_{\mathcal{S}_{\mathcal{H}^d \mathcal{H}}}^2 = \begin{cases} \mathrm{O}_{\mathbb{P}}(N^{-1}), & \text{if } d \in \mathbb{N} \text{ is fixed,} \\ \mathrm{O}_{\mathbb{P}}(d^2 N^{-1}), & \text{if } d = d_N = \Xi(1, N). \end{cases} \tag{3.10}$$

Moreover, the empirical covariance operators

$$\hat{\mathfrak{S}}_d := \frac{1}{N_d - 1} \sum_{k=1}^{N_d} \left(X_{d+k-1}^2(d) - \hat{m}_1(\boldsymbol{X}^2(d)) \right) \otimes \left(X_{d+k-1}^2(d) - \hat{m}_1(\boldsymbol{X}^2(d)) \right) \tag{3.11}$$

satisfy

$$||\hat{\mathfrak{S}}_d - \mathfrak{S}_d||_{\mathcal{S}_{\mathcal{H}^d}}^2 = \begin{cases} \mathrm{O}_{\mathbb{P}}(N^{-1}), & \text{if } d \in \mathbb{N} \text{ is fixed,} \\ \mathrm{O}_{\mathbb{P}}(d^3 N^{-1}), & \text{if } d = d_N = \Xi(1, N). \end{cases} \tag{3.12}$$

3.1.2 Estimation of eigenvalues and eigenfunctions

Here, we derive asymptotic upper bounds of the estimation errors for the eigenvalues and eigenfunctions of a compact, self-adjoint and positive semi-definite operator $\mathscr{K} \in \mathcal{K}_{\mathcal{H}}$, estimated by a sequence $(\hat{\mathscr{K}}_N)_{N \in \mathbb{N}} \subseteq \mathcal{K}_{\mathcal{H}}$ of compact, self-adjoint, positive semi-definite operators, where each $\hat{\mathscr{K}}_N$ depends on N observations of a stationary time series $\boldsymbol{X} = (X_k)_{k \in \mathbb{Z}}$. Further, $(\mathfrak{k}_j)_{j \in \mathbb{N}}$ resp. $(\hat{\mathfrak{k}}_j)_{j \in \mathbb{N}}$ are the eigenfunction sequences and $(k_j)_{j \in \mathbb{N}}$ resp. $(\hat{k}_j)_{j \in \mathbb{N}}$ the associated w.l.o.g. monotonically decreasing eigenvalue sequences of $\mathscr{K}$ resp. $\hat{\mathscr{K}}_N$. For the derivation of upper bounds of the estimation errors for the eigenvalues and eigenfunctions, we need

$$|a_j - b_j| \leq ||A - B||_{\mathcal{L}_{\mathcal{H}}}, \quad j \in \mathbb{N}. \tag{3.13}$$

This is true according to Bosq (2000), Lemma 4.2 where $A, B \in \mathcal{K}_{\mathcal{H}}$ have the singular value decompositions $A = \sum_{j=1}^{\infty} a_j(\mathfrak{a}_j \otimes \mathfrak{a}_j')$ resp. $B = \sum_{j=1}^{\infty} b_j(\mathfrak{b}_j \otimes \mathfrak{b}_j')$.

Corollary 3.1. *Let* $||\hat{\mathscr{K}}_N - \mathscr{K}||^2_{\mathcal{L}_{\mathcal{H}}} = \mathrm{O}_{\mathbb{P}}(a_N)$ *hold where* $a_N = \Xi[N^{-1}, 1)$. *Then*

$$\sup_{j \in \mathbb{N}} (\hat{k}_j - k_j)^2 = \mathrm{O}_{\mathbb{P}}(a_N). \tag{3.14}$$

Moreover, if $k_{b_N} = \Xi[\sqrt{a_N}, 1]$ *holds where* $b_N = \Omega(1)$, *then*

$$\hat{k}_{b_N} = \mathrm{O}_{\mathbb{P}}(k_{b_N}) \quad \text{and} \quad k_{b_N} = \mathrm{O}_{\mathbb{P}}(\hat{k}_{b_N}). \tag{3.15}$$

Because the eigenfunctions of $\hat{\mathscr{K}}_N$ are unambiguously determined except for their sign,

$$\hat{\mathfrak{e}}'_j := \mathrm{sgn}(\langle \hat{\mathfrak{e}}_j, \mathfrak{e}_j \rangle_{\mathcal{H}}) \hat{\mathfrak{e}}_j \tag{3.16}$$

is used as an estimator for $\mathfrak{e}_j$ if $\hat{\mathfrak{e}}_j \not\perp \mathfrak{e}_j$ a.s. holds where sgn is the signum function. According to Bosq (2000), Lemma 4.3, which can be generalized to any compact, self-adjoint and positive semi-definite operators,

$$||\hat{\mathfrak{e}}'_j - \mathfrak{e}_j||_{\mathcal{H}} \leq \tilde{\gamma}_j ||\hat{\mathscr{K}}_N - \mathscr{K}||_{\mathcal{L}_{\mathcal{H}}}, \quad j \in \mathbb{N}, \tag{3.17}$$

if the eigenspace of k_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 $\gamma_j := (k_j - k_{j+1})^{-1}$ for $j \in \mathbb{N}$. The problem in using $\hat{\mathfrak{e}}'_j$ as an estimator for $\mathfrak{e}_j$ is, that $\hat{\mathfrak{e}}_j \not\perp \mathfrak{e}_j$ a.s. and thus $\mathrm{sgn}(\langle \hat{\mathfrak{e}}_j, \mathfrak{e}_j \rangle_{\mathcal{H}}) \neq 0$ a.s., which is needed to obtain asymptotic upper bounds of the estimation errors for the operators in the $\mathscr{H}$-valued ARCH and GARCH model, is not guaranteed for all j, N. Therefore, we modify $\hat{\mathfrak{e}}_j$ in the following way. Let $(h_j)_{j \in \mathbb{N}}$ be a CONS of $\mathcal{H}$ and let $(\zeta_j)_{j \in \mathbb{N}}$ be a sequence of i.i.d. and $\mathcal{N}(0, 1)$-distributed random variables, independent of the observations of $\boldsymbol{X}$. Then

$$\hat{\mathfrak{e}}''_j := \hat{\mathfrak{e}}_j + \sum_{i=1}^{\infty} \frac{\zeta_i h_i}{i^2 N} \tag{3.18}$$

is well-defined for all j, N with $\hat{\mathfrak{e}}''_j \not\perp \mathfrak{e}_j$ a.s. and in consequence $\mathrm{sgn}(\langle \hat{\mathfrak{e}}''_j, \mathfrak{e}_j \rangle_{\mathcal{H}'}) \neq 0$ a.s. Thus we use

$$\hat{\mathfrak{e}}'''_j := \mathrm{sgn}(\langle \hat{\mathfrak{e}}''_j, \mathfrak{e}_j \rangle_{\mathcal{H}'}) \hat{\mathfrak{e}}_j \tag{3.19}$$

as an estimator for $\mathfrak{e}_j$, where $(\hat{\mathfrak{e}}'''_j)_j$ is a CONS of $\mathcal{H}$ a.s. according to the spectral theorem.

Assumption 3.3. *For all* $j, k_j \neq k_{j+1}$ *and* $\kappa(j) = k_j$ *holds where* $\kappa \colon \mathbb{R} \to \mathbb{R}$ *is a convex function.*

If $\mathscr{K}$ is injective and if the eigenvalues of $\mathscr{K}$ satisfy Assumption 3.3, then

$$k_1 > k_2 > \cdots > 0. \tag{3.20}$$

Moreover, for any sequence $m = m_N = \Omega(1)$:

$$\sup_{j \leq m} \tilde{\gamma}_j = \gamma_m \asymp k_m^{-1}. \tag{3.21}$$

Lemma 3.4. *Let* $\mathscr{K}$ *be injective, let Assumption 3.3 and* $||\hat{\mathscr{K}}_N - \mathscr{K}||^2_{\mathcal{L}_{\mathcal{H}}} = \mathrm{O}_{\mathbb{P}}(a_N)$ *with* $a_N = \Xi[N^{-1}, 1)$ *hold. Then*

$$||\hat{\mathfrak{e}}'''_j - \mathfrak{e}_j||^2_{\mathcal{H}} = \mathrm{O}_{\mathbb{P}}(a_N), \quad j \in \mathbb{N}. \tag{3.22}$$

Furthermore, if $k_m = \omega(\sqrt{a_N})$ *holds where* $m = m_N = \Xi[1, N)$, *then*

$$\sup_{j \leq m} ||\hat{\mathfrak{e}}'''_j - \mathfrak{e}_j||^2_{\mathcal{H}} = \mathrm{O}_{\mathbb{P}}(k_m^{-2} a_N). \tag{3.23}$$

3.1.3 Some notes on estimating operators

In this paper, we estimate bounded operators $B \in \mathcal{L}_{\mathcal{H}',\mathcal{H}''}$ in equations as

$$A = BC \tag{3.24}$$

where $A \in \mathcal{L}_{\mathcal{H},\mathcal{H}''}$ and $C \in \mathcal{L}_{\mathcal{H},\mathcal{H}'}$. Identifiability of B from (3.24), that is $BC = \tilde{B}C$ implying $B = \tilde{B}$, is only guaranteed if B has dense image. Further, if C is a compact operator and thus has no bounded inverse, we use the tikhonov-regularized of C in order to isolate B. Also, when estimating operators without projecting them on a finite-dimensional subspace, we impose the following Sobolev condition.

Assumption 3.4. *Let $S \in \mathcal{S}_{\mathcal{H},\mathcal{H}'}$ and let $(\phi_{ij})_{i,j\in\mathbb{N}}$ be a CONS of $\mathcal{S}_{\mathcal{H},\mathcal{H}'}$. Then, $(S, (\phi_{ij})_{i,j})$ satisfies the Sobolev condition for $\beta > 0$ if*

$$\sum_{i,j=1}^{\infty} \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 (1 + i^{2\beta} + j^{2\beta}) < \infty. \tag{3.25}$$

With $\mathscr{I}_m^c := \{\phi_{ij} | i, j \in \mathbb{N}, i \vee j > m\}$, this implies the identity

$$\left\| \coprod_{\mathscr{I}_m^c} S \right\|_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 = \sum_{\substack{i,j\in\mathbb{N} \\ i\vee j>m}} \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 \leq (1 + m^{2\beta})^{-1} \sum_{i,j=1}^{\infty} \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 (1 + i^{2\beta} + j^{2\beta}) = \mathrm{O}(m^{-2\beta}) \tag{3.26}$$

for $m = m_N \to \infty$ which we utilize in conversions in various proofs.

3.2 Estimation of δ in the functional ARCH and GARCH model

We derive an estimator of δ in $\mathscr{H}$-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p}, \mathfrak{q}$) time series with $\mathfrak{p}, \mathfrak{q} \in \mathbb{N}$ from the idea of estimating δ in $\mathscr{H}$-valued ARCH(1) time series in Hörmann *et al.* (2013). Under Assumption 3.1, taking the expected value on both sides of the right equation in (2.1), yields

$$\delta = m_2 - \sum_{i=1}^{\mathfrak{r}} (\alpha_i + \beta_i)(m_2), \tag{3.27}$$

where $\alpha_i = 0_{\mathcal{L}_{\mathscr{H}}} = \beta_j$ for $i > \mathfrak{p}, j > \mathfrak{q}$. Therefore, we propose

$$\hat{\delta} := \hat{m}_2 - \sum_{i=1}^{\mathfrak{r}} (\hat{\alpha}_i + \hat{\beta}_i)(\hat{m}_2) \tag{3.28}$$

as an estimator for δ where $\hat{\alpha}_i, \hat{\beta}_j$ are estimators for α_i, β_j and where $\hat{m}_2 := N^{-1} \sum_{i=1}^{N} \mathscr{X}_i^2$.

Theorem 3.1. *Let Assumption 3.1 hold. Then $\hat{\delta}$ in* (3.28) *satisfies*

$$||\hat{\delta} - \delta||_{\mathscr{H}} = \mathrm{O}_{\mathbb{P}}(N^{-1/2}) + \sum_{i=1}^{\mathfrak{r}} \mathrm{O}_{\mathbb{P}}(||\hat{\alpha}_i - \alpha_i||_{\mathcal{L}_{\mathscr{H}}}) + \mathrm{O}_{\mathbb{P}}(||\hat{\beta}_i - \beta_i||_{\mathcal{L}_{\mathscr{H}}}). \tag{3.29}$$

3.3 Operator estimation in the functional ARCH model

In the following, $\mathscr{X} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ is a $\mathscr{H}$-valued ARCH($\mathfrak{p}$) process with $\mathfrak{p} \in \mathbb{N}$. Under Assumption 3.1, $\mathscr{Z} := (\mathscr{Z}_k)_{k\in\mathbb{Z}} = (\mathscr{X}_k^2 - m_2)_{k\in\mathbb{Z}}$ with $m_2 := \mathbb{E}(\mathscr{X}_1^2)$ is a $\mathscr{H}$-valued AR($\mathfrak{p}$) process with innovation process $\boldsymbol{\nu} := (\nu_k)_{k\in\mathbb{Z}} = (\mathscr{X}_k^2 - \sigma_k^2)_{k\in\mathbb{Z}}$ (see p.5). Furthermore, $\mathscr{Z}(\mathfrak{p}) := (\mathscr{Z}_k(\mathfrak{p}))_{k\in\mathbb{Z}}$ satisfies

$$\mathscr{Z}_k(\mathfrak{p}) = \tilde{\nu}_k(\mathfrak{p}) + \boldsymbol{A}_1(\mathscr{Z}_{k-1}(\mathfrak{p})) :\Longleftrightarrow
\begin{bmatrix} \mathscr{Z}_k \\ \mathscr{Z}_{k-1} \\ \vdots \\ \mathscr{Z}_{k-\mathfrak{p}+2} \\ \mathscr{Z}_{k-\mathfrak{p}+1} \end{bmatrix}
=
\begin{bmatrix} \nu_k \\ 0_{\mathscr{H}} \\ \vdots \\ 0_{\mathscr{H}} \\ 0_{\mathscr{H}} \end{bmatrix}
+
\begin{bmatrix}
\alpha_1 & \cdots & \cdots & \cdots & \alpha_{\mathfrak{p}} \\
\mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} \\
0_{\mathcal{L}_{\mathscr{H}}} & \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} & \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}}
\end{bmatrix}
\begin{bmatrix} \mathscr{Z}_{k-1} \\ \mathscr{Z}_{k-2} \\ \vdots \\ \mathscr{Z}_{k-\mathfrak{p}+1} \\ \mathscr{Z}_{k-\mathfrak{p}} \end{bmatrix} \tag{3.30}$$

a.s. for all k. The operator

$$\boldsymbol{\alpha}_{[\mathfrak{p}]} := [\alpha_1 \; \cdots \; \alpha_{\mathfrak{p}}]$$

is an element of $\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}$ and satisfies the Yule-Walker equation

$$\mathfrak{S}_{\mathfrak{p},1} = \boldsymbol{\alpha}_{[\mathfrak{p}]}\mathfrak{S}_{\mathfrak{p}} \tag{3.31}$$

where $\mathfrak{S}_{\mathfrak{p},1} = \mathscr{C}_{\mathscr{X}_{\mathfrak{p}}(\mathfrak{p}),\mathscr{X}_{\mathfrak{p}+1}}$ and $\mathfrak{S}_{\mathfrak{p}} = \mathscr{C}_{0;\boldsymbol{\mathscr{X}}(\mathfrak{p})}$. Since $\mathfrak{S}_{\mathfrak{p}}$ is injective as a consequence of Lemma 3.1 (see Kühnert (2019), Lemma 4.35), $\boldsymbol{\alpha}_{[\mathfrak{p}]}$ can be identified from (3.31) and as an estimator we thus impose

$$\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} := \hat{\mathfrak{S}}_{\mathfrak{p},1}\hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger} \prod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} = \hat{\mathfrak{S}}_{\mathfrak{p},1}\hat{\mathfrak{S}}_{\mathfrak{p}}(\hat{\mathfrak{S}}_{\mathfrak{p}}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1} \prod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}. \tag{3.32}$$

Thereby, $K \in \mathbb{N}$, $(\vartheta_N)_{N\in\mathbb{N}} \subseteq \mathbb{N}$ with $\vartheta_N \to 0, \hat{\mathfrak{c}}_{\mathfrak{p},1}, ..., \hat{\mathfrak{c}}_{\mathfrak{p},K}$ are the eigenfunctions of $\hat{\mathfrak{S}}_{\mathfrak{p}}$ associated to the first biggest eigenvalues $\hat{c}_{\mathfrak{p},1} \geq \cdots \geq \hat{c}_{\mathfrak{p},K}$ and $\prod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}$ is the operator projecting on $\lin\{\hat{\mathfrak{c}}_{\mathfrak{p},1}, ..., \hat{\mathfrak{c}}_{\mathfrak{p},K}\} \subseteq \mathscr{H}^{\mathfrak{p}}$.

Theorem 3.2. *Let Assumptions 3.1-3.2 and let Assumption 3.3 for the eigenvalues of $\mathfrak{S}_{\mathfrak{p}}$ hold. Further, let $(\Phi_{\mathfrak{p},ij})_{i,j\in\mathbb{N}}$ be the CONS of $\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}$ defined by $\Phi_{\mathfrak{p},ij} := \mathfrak{c}_{\mathfrak{p},i} \otimes \mathfrak{c}_j$ for all i,j, where $(\mathfrak{c}_{\mathfrak{p},i})_{i\in\mathbb{N}} \subseteq \mathscr{H}^{\mathfrak{p}}$ and $(\mathfrak{c}_j)_{j\in\mathbb{N}} \subseteq \mathscr{H}$ are the eigenfunction sequences of the covariance operators $\mathfrak{S}_{\mathfrak{p}}$ resp. $\mathscr{C}_{0;\boldsymbol{\mathscr{X}}}$.*

(a) *Let $\mathscr{J}_{\mathfrak{p},K} := \{\Phi_{\mathfrak{p},ij} | 1 \leq i,j \leq K\}$ with $K \in \mathbb{N}$. If $\langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j\rangle_{\mathscr{H}} = 0$ for all $j > K, l \leq K$, then*

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \prod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(N^{-1}). \tag{3.33}$$

(b) *Let also $(\boldsymbol{\alpha}_{[\mathfrak{p}]}, (\Phi_{\mathfrak{p},ij})_{i,j})$ satisfy Assumption 3.4 for some $\beta > 0$. Then, for any sequence $K = K_N = \Xi(1,N)$ with $c_{\mathfrak{p},K}^{-4}K^{2\beta+1} = \mathrm{O}(N)$ and $\sum_{l=1}^{K}(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N})^2 \sum_{j>K}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}}^2 = \mathrm{O}(K^{-2\beta})$ holds*

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.34}$$

Remark 3.1. *The assumption $\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}} = 0$ for all $j > K, l \leq K$ in (a), which is necessary for technical conversions in the proof, is weaker than to impose that $\boldsymbol{\alpha}_{[\mathfrak{p}]}$ and $\mathfrak{S}_{\mathfrak{p}}$ commute. Moreover, it is similar to the assumption used in Turbillon et al. (2007) in order to estimate their MA(1) operator.*

Example 3.1. *Let the assumptions of Theorem 3.2 hold.*

(a) *Assume $c_{\mathfrak{p},N} \asymp e^{-N}$ and let $K_N := 1 + \lfloor\frac{\ln(N)}{4+b}\rfloor$ for all $N \in \mathbb{N}$ for some $b > 0$. Then $K = K_N = \Xi(1,N), c_{\mathfrak{p},K}^{-4}K^{2\beta+1} \asymp N^{\frac{4}{4+b}}\ln^{2\beta+1}(N) = \mathrm{O}(N)$ and consequently according to Theorem 3.2:*

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(\ln^{-2\beta}(N)).$$

(b) *Impose $c_{\mathfrak{p},N} \asymp N^{-a}$ for some $a > 1$ and let $K_N := 1 + \lfloor N^{\frac{1}{1+4a+2\beta}}\rfloor$ for all $N \in \mathbb{N}$. Then $K = K_N = \Xi(1,N), c_{\mathfrak{p},K}^{-4}K^{2\beta+1} \asymp K^{1+4a+2\beta} \asymp N$ and after Theorem 3.2:*

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}\left(N^{-\frac{2\beta}{1+4a+2\beta}}\right).$$

3.4 Operator estimation in the functional GARCH model

Throughout this section, $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k\in\mathbb{Z}}$ is a $\mathscr{H}$-valued GARCH($\mathfrak{p},\mathfrak{q}$) with $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ and $\boldsymbol{\mathscr{Z}} := (\mathscr{Z}_k)_{k\in\mathbb{Z}} = (\mathscr{X}_k^2 - m_2)_{k\in\mathbb{Z}}$ the corresponding $\mathscr{H}$-valued ARMA($\mathfrak{r},\mathfrak{q}$) time series with time series of innovations $\boldsymbol{\nu} := (\nu_k)_{k\in\mathbb{Z}} = (\mathscr{X}_k^2 - \sigma_k^2)_{k\in\mathbb{Z}}$ (see p.5). $\boldsymbol{\mathscr{Z}}$ satisfies the following.

Assumption 3.5. *$\boldsymbol{\mathscr{Z}}$ is an invertible, linear process w.r.t. $\boldsymbol{\nu}$ with representation as inverted time series*

$$\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\infty}\pi_i(\mathscr{Z}_{k-i}) \tag{3.35}$$

a.s. for all k where $(\pi_i)_{i\in\mathbb{N}} \subseteq \mathcal{S}_{\mathscr{H}}$ with $\sum_{i=1}^{\infty}\|\pi_i\|_{\mathcal{S}_{\mathscr{H}}} < \infty$.

3.4.1   Derivation of the estimators for the operators in the functional GARCH model

At first, since $\alpha_i := 0_{\mathcal{L}_\mathscr{H}} =: \beta_j$ for $i > \mathfrak{p}, j > \mathfrak{q}$, the representations (3.2) and (3.35) imply

$$\mathscr{Z}_k = \sum_{i=1}^{\infty} \left(\alpha_i + \sum_{j=1}^{(i-1)\wedge\mathfrak{q}} \beta_j \pi_{i-j} \right) (\mathscr{Z}_{k-i})$$

a.s. for all k. Moreover, if Assumptions 3.1-3.2 hold there is no closed subspace $V \subsetneq \mathscr{H}$ with $\mathbb{P}(\mathscr{Z}_0 \in V) = 1$ which, following from Kühnert (2019), Lemma 4.48 and Remark 4.49, leads to

$$\pi_i = \alpha_i + \sum_{j=1}^{(i-1)\wedge\mathfrak{q}} \beta_j \pi_{i-j}, \quad i \in \mathbb{N}. \tag{3.36}$$

Since $\alpha_i = 0_{\mathcal{L}_\mathscr{H}}$ for $i > \mathfrak{p}$, (3.36) implies with $\mathfrak{s} = \mathfrak{p} + \mathfrak{q}$:

$$\pi_\mathfrak{s} = \sum_{j=1}^{\mathfrak{q}} \beta_j \pi_{\mathfrak{s}-j} = \begin{bmatrix} \beta_1 \, \beta_2 \, \cdots \, \beta_\mathfrak{q} \end{bmatrix} \begin{bmatrix} \pi_{\mathfrak{s}-1} \, \pi_{\mathfrak{s}-2} \, \cdots \, \pi_\mathfrak{p} \end{bmatrix}^T =: \boldsymbol{\beta}_{[\mathfrak{q}]} \boldsymbol{\pi}^T_{[\mathfrak{p},\mathfrak{q}]}$$

where the solution $\boldsymbol{\beta}_{[\mathfrak{q}]}$ is unique iff the image of $\boldsymbol{\pi}^T_{[\mathfrak{p},\mathfrak{q}]} \in \mathcal{S}_{\mathscr{H},\mathscr{H}^\mathfrak{q}}$ lies dense in $\mathscr{H}^\mathfrak{q}$. This is impossible since $\mathscr{H} \subsetneq \mathscr{H}^\mathfrak{q}$, why we establish estimators for $\beta_1, ..., \beta_\mathfrak{q}$ based on the equation

$$\boldsymbol{\pi}_{[\mathfrak{s},\mathfrak{q}]} = \boldsymbol{\beta}_{[\mathfrak{q}]} \prod\nolimits_{[\mathfrak{s},\mathfrak{q}]} :\Longleftrightarrow \begin{bmatrix} \pi_{\mathfrak{s}+\mathfrak{q}-1} \\ \pi_{\mathfrak{s}+\mathfrak{q}-2} \\ \vdots \\ \pi_\mathfrak{s} \end{bmatrix} = \begin{bmatrix} \beta_1 \, \beta_2 \, \cdots \, \beta_\mathfrak{q} \end{bmatrix} \begin{bmatrix} \pi_{\mathfrak{s}+\mathfrak{q}-2} & \pi_{\mathfrak{s}+\mathfrak{q}-3} & \cdots & \pi_{\mathfrak{s}-1} \\ \pi_{\mathfrak{s}+\mathfrak{q}-3} & \pi_{\mathfrak{s}+\mathfrak{q}-4} & \cdots & \pi_{\mathfrak{s}-2} \\ \vdots & \vdots & \cdots & \vdots \\ \pi_{\mathfrak{s}-1} & \pi_{\mathfrak{s}-2} & \cdots & \pi_\mathfrak{p} \end{bmatrix} \tag{3.37}$$

The following example illustrates that the image of $\prod_{[\mathfrak{s},\mathfrak{q}]} \in \mathcal{S}_{\mathscr{H}^\mathfrak{q}}$ can lie dense.

Example 3.2. *Let* $\alpha_i := 0_{\mathcal{L}_\mathscr{H}} =: \beta_j$ *for* $i \neq \mathfrak{p}, j \neq \mathfrak{q}$ *and let* $\alpha_\mathfrak{p} = \beta_\mathfrak{q} =: \gamma$ *where* $\gamma \in \mathcal{S}_\mathscr{H}$ *is an operator with dense image satisying* $\gamma \neq 0_{\mathcal{L}_\mathscr{H}}$ *and* $||\gamma||_{\mathcal{S}_\mathscr{H}} < 1$. *Then, because* (3.36) *implies* $\pi_i = \gamma^k$ *for all* $i = \mathfrak{p} + (k-1)\mathfrak{q}$ *for some* $k \in \mathbb{N}$ *and* $\pi_i = 0_{\mathcal{L}_\mathscr{H}}$ *otherwise, we obtain*

$$\prod\nolimits_{[\mathfrak{s},\mathfrak{q}]} = \begin{bmatrix} 0_{\mathcal{L}_\mathscr{H}} & \cdots & 0_{\mathcal{L}_\mathscr{H}} & \gamma^2 & 0_{\mathcal{L}_\mathscr{H}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{\mathcal{L}_\mathscr{H}} & \ddots & \ddots & \ddots & \vdots \\ \gamma^2 & \ddots & \ddots & \ddots & 0_{\mathcal{L}_\mathscr{H}} \\ 0_{\mathcal{L}_\mathscr{H}} & \cdots & \cdots & 0_{\mathcal{L}_\mathscr{H}} & \gamma \end{bmatrix} \implies \prod\nolimits^*_{[\mathfrak{s},\mathfrak{q}]} = \begin{bmatrix} 0_{\mathcal{L}_\mathscr{H}} & \cdots & 0_{\mathcal{L}_\mathscr{H}} & (\gamma^*)^2 & 0_{\mathcal{L}_\mathscr{H}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{\mathcal{L}_\mathscr{H}} & \ddots & \ddots & \ddots & \vdots \\ (\gamma^*)^2 & \ddots & \ddots & \ddots & 0_{\mathcal{L}_\mathscr{H}} \\ 0_{\mathcal{L}_\mathscr{H}} & \cdots & \cdots & 0_{\mathcal{L}_\mathscr{H}} & \gamma^* \end{bmatrix}$$

Since the operators $\gamma^*, (\gamma^*)^2$ *and hence* $\prod^*_{[\mathfrak{s},\mathfrak{q}]}$ *are injective, the image of* $\prod_{[\mathfrak{s},\mathfrak{q}]}$ *lies dense.*

Due to (3.37), analogously to (3.32), we use

$$\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]} := \hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} \hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]}^{\dagger} \prod_{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};1}}^{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};M}} = \hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} \hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]}^{*} \left(\hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]} \hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]}^{*} + \theta_N \mathbb{I}_{\mathscr{H}^\mathfrak{q}} \right)^{-1} \prod_{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};1}}^{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};M}} \tag{3.38}$$

to estimate $\boldsymbol{\beta}_{[\mathfrak{q}]}$. Thereby,

$$\hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} := \begin{bmatrix} \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-1} \, \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-2} \, \cdots \, \hat{\pi}_\mathfrak{s} \end{bmatrix} \tag{3.39}$$

is an element of $\mathcal{S}_{\mathscr{H}^\mathfrak{q},\mathscr{H}}$ and

$$\hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]} := \begin{bmatrix} \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-2} & \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-3} & \cdots & \hat{\pi}_{\mathfrak{s}-1} \\ \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-3} & \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-4} & \cdots & \hat{\pi}_{\mathfrak{s}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\pi}_{\mathfrak{s}-1} & \hat{\pi}_{\mathfrak{s}-2} & \cdots & \hat{\pi}_\mathfrak{p} \end{bmatrix} \tag{3.40}$$

is an element of $\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$. Further, $\hat{\pi}_k$ where $k \in \mathbb{N}$, with what we estimate π_k, stands for the k-th component of

$$\hat{\boldsymbol{\pi}}_{L,K} := \hat{\mathfrak{S}}_{L,K} \hat{\mathfrak{S}}_L^{\dagger} \coprod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} = \hat{\mathfrak{S}}_{L,1} \hat{\mathfrak{S}}_L (\hat{\mathfrak{S}}_L^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^L})^{-1} \coprod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} \tag{3.41}$$

where $(K_N)_{N \in \mathbb{N}} \subseteq \mathbb{N}, (L_N)_{N \in \mathbb{N}} \subseteq \mathbb{N}$ and $(\vartheta_N)_{N \in \mathbb{N}} \subseteq (0, \infty)$ are sequences with $K = K_N \to \infty, L = L_N \to \infty$ resp. $\vartheta_N \to 0$ and where $\hat{\mathfrak{c}}_{L,1}, ..., \hat{\mathfrak{c}}_{L,K}$ are the eigenfunctions of $\hat{\mathfrak{S}}_L$ associated to the first biggest eigenvalues $\hat{c}_{L,1} \geq \cdots \geq \hat{c}_{L,K}$. Further, in (3.38), $(M_N)_{N \in \mathbb{N}} \subseteq \mathbb{N}$ and $(\theta_N)_{N \in \mathbb{N}} \subseteq (0, \infty)$ are sequences with $M = M_N \to \infty$ resp. $\theta_N \to 0$ and $(\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j})_{j \in \mathbb{N}}$ is the eigenfunction sequence of $\hat{\prod}_{[\mathfrak{s},\mathfrak{q}]} \hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^* \in \mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$ with associated eigenvalue sequence $(\hat{g}_{\mathfrak{s},\mathfrak{q};j})_{j \in \mathbb{N}}$ which decreases monotonically w.l.o.g. Because of (3.36), it is hence plausible to use

$$\hat{\alpha}_i := \hat{\pi}_i - \sum_{j=1}^{(i-1) \wedge \mathfrak{q}} \hat{\beta}_j \hat{\pi}_{i-j} \tag{3.42}$$

as an estimator for α_i with $i = 1, ..., \mathfrak{p}$ where $\hat{\beta}_j$ is the j-th component of $\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]}$ and where $\hat{\alpha}_1 := \hat{\pi}_1$.

3.4.2 Upper bounds of the estimation errors for the operators π_i

The following Theorem states asymptotic upper bounds of the operators for invertible, linear processes represented as inverted time series. It is crucial for our derivation of asymptotic upper bounds of the estimation errors for the GARCH operators.

Theorem 3.3. *Let Assumptions 3.1-3.2 and 3.5 hold. Let $\beta > 0, L = L_N = \Xi(1, N), K = K_N = \Xi(1, \sqrt[3]{L^{-1}N})$ and $\vartheta_N = \mathrm{O}(c_{L,K}^{-2} K^{-\beta})$. Thereby, $(c_{L,j})_j$ is the eigenvalue sequence of $\mathfrak{S}_L$ satisfying $c_{L,K} = \Omega(\sqrt{L^3 N^{-1}})$, $\sum_{l=1}^K (\frac{c_{L,l}^2}{c_{L,l}^2 + \vartheta_N})^2 \sum_{j > K} \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = \mathrm{O}(K^{-2\beta})$ and $c_{L,K}^{-1} \sqrt{KL^3 N^{-1}} = \mathrm{O}(K^{-(2\beta+1)})$ if $\sum_{l > L} ||\pi_l||_{\mathcal{L}_{\mathscr{H}}} = \mathrm{O}(c_{L,K}^{-1} \sqrt{KL^3 N^{-1}})$ resp. $L^{-2} N (\sum_{l > L} ||\pi_l||_{\mathcal{L}_{\mathscr{H}}})^4 = \mathrm{O}(K^{1-2\beta})$ if $c_{L,K}^{-1} \sqrt{KL^3 N^{-1}} = \mathrm{O}(\sum_{l > L} ||\pi_l||_{\mathcal{L}_{\mathscr{H}}})$. At last, for all L, let Assumption 3.3 hold for the eigenvalue sequence $(c_{L,j})_j$ and let $(\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$ satisfy Assumption 3.4 for β, see Theorem 3.2. Then, for all $i \in \mathbb{N}$:*

$$||\hat{\pi}_i - \pi_i||_{\mathcal{S}_{\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.43}$$

3.4.3 Upper bounds of the estimation errors for the GARCH operators

Here, we need

$$||\hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} - \boldsymbol{\pi}_{[\mathfrak{s},\mathfrak{q}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \sum_{i=0}^{\mathfrak{q}-1} ||\hat{\pi}_{\mathfrak{s}+i} - \pi_{\mathfrak{s}+i}||_{\mathcal{S}_{\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) \tag{3.44}$$

as well as

$$\left|\left| \hat{\prod}_{[\mathfrak{s},\mathfrak{q}]} \hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^* - \prod_{[\mathfrak{s},\mathfrak{q}]} \prod_{[\mathfrak{s},\mathfrak{q}]}^* \right|\right|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}}^2 = \sum_{i,j=0}^{\mathfrak{q}-1} \left|\left| \sum_{k=\mathfrak{p}}^{\mathfrak{s}-1} \hat{\pi}_{k+i} \hat{\pi}_{k+j}^* - \pi_{k+i} \pi_{k+j}^* \right|\right|_{\mathcal{S}_{\mathscr{H}}}^2$$

$$\precsim \sum_{i,j=0}^{\mathfrak{q}-1} \sum_{k=\mathfrak{p}}^{\mathfrak{s}-1} ||\hat{\pi}_{k+i}||_{\mathcal{S}_{\mathscr{H}}}^2 ||\hat{\pi}_{k+j}^* - \pi_{k+j}^*||_{\mathcal{S}_{\mathscr{H}}}^2 + ||\hat{\pi}_{k+i} - \pi_{k+i}||_{\mathcal{S}_{\mathscr{H}}}^2 ||\pi_{k+j}^*||_{\mathcal{S}_{\mathscr{H}}}^2$$

$$= \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) \tag{3.45}$$

which is both true after (3.43). According to Corollary 3.1, the identities (3.13) and (3.45) imply

$$\sup_{j \in \mathbb{N}} (\hat{g}_{\mathfrak{s},\mathfrak{q};j} - g_{\mathfrak{s},\mathfrak{q};j})^2 = \mathrm{O}(K^{-2\beta}) \tag{3.46}$$

where $(g_{\mathfrak{s},\mathfrak{q};j})_{j \in \mathbb{N}}$ is the w.l.o.g. monotonically decreasing eigenvalue sequence associated to the eigenfunction sequence $(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j})_{j \in \mathbb{N}}$ of $\prod_{[\mathfrak{s},\mathfrak{q}]} \prod_{[\mathfrak{s},\mathfrak{q}]}^* \in \mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$. Moreover, if $g_{\mathfrak{s},\mathfrak{q};M} = \Xi(K^{-\beta}, 1)$ with $M = M_N = \Xi(1, N)$, then

$$\hat{g}_{\mathfrak{s},\mathfrak{q};M} = \mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}) \qquad \text{and} \qquad g_{\mathfrak{s},\mathfrak{q};M} = \mathrm{O}_{\mathbb{P}}(\hat{g}_{\mathfrak{s},\mathfrak{q};M}) \tag{3.47}$$

after Corollary 3.1 and if also $\prod_{[\mathfrak{s},\mathfrak{q}]}\prod_{[\mathfrak{s},\mathfrak{q}]}^*$ is injective and satisfies Assumption 3.3, Lemma 3.4 yields

$$\sup_{j\leq M}||\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j}''' - \mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}||^2_{\mathscr{H}^\mathfrak{q}} = \mathrm{O}_{\mathbb{P}}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};M}^{-2}K^{-2\beta}). \tag{3.48}$$

Theorem 3.4. *Let the assumptions of Theorem 3.3 hold. Let $\prod_{[\mathfrak{s},\mathfrak{q}]}\prod_{[\mathfrak{s},\mathfrak{q}]}^*$ be injective and let its eigenvalue sequence $(g_{\mathfrak{s},\mathfrak{q};j})_j$ satisfy Assumption 3.3. Also, let $M = M_N = \Xi(1,N), \theta_N = \mathrm{O}(K^{-\beta/2}), g_{\mathfrak{s},\mathfrak{q};M}^{-2}M^\beta = \mathrm{O}(K^\beta)$ as well as $\sum_{l=1}^M (\frac{g_{\mathfrak{s},\mathfrak{q};l}^2}{g_{\mathfrak{s},\mathfrak{q};l}^2+\theta_N})^2 \sum_{j>M}\langle\boldsymbol{\beta}_{[\mathfrak{q}]}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}),\mathfrak{c}_j\rangle^2_{\mathscr{H}} = \mathrm{O}(M^{-2\beta})$ hold, and let $(\boldsymbol{\beta}_{[\mathfrak{q}]}, (\Phi_{\mathfrak{s},\mathfrak{q};ij})_{i,j})$ satisfy Assumption 3.4 for β where $\Phi_{\mathfrak{s},\mathfrak{q};ij} := \mathfrak{g}_{\mathfrak{s},\mathfrak{q};j} \otimes \mathfrak{c}_j$. Then,*

$$||\hat{\alpha}_i - \alpha_i||^2_{\mathcal{S}_{\mathscr{H}}} = \begin{cases} \mathrm{O}_{\mathbb{P}}(K^{-2\beta}), & i = 1, \\ \mathrm{O}_{\mathbb{P}}(M^{-2\beta}), & i = 2, ..., \mathfrak{p}, \end{cases} \tag{3.49}$$

and for all $j = 1, ..., \mathfrak{q}$:

$$||\hat{\beta}_j - \beta_j||^2_{\mathcal{S}_{\mathscr{H}}} = \mathrm{O}_{\mathbb{P}}(M^{-2\beta}). \tag{3.50}$$

4   Conclusions

This article studies functional ARCH and GARCH processes in established function spaces and in those which have not been considered yet. It focuses on the asymptotic upper bounds of the estimation errors for the operators projected on a finite-dimensional subspace in the ARCH and the complete operators in the ARCH and GARCH model where the operators are estimated by a Yule-Walker approach. The theories developed complement Hörmann *et al.* (2013), Aue *et al.* (2017) and Cerovecki *et al.* (2019) where functional ARCH(1), GARCH(1,1) resp. GARCH processes for any order were established. This paper also displays asymptotic upper bounds of the estimation errors for operators of invertible, linear processes represented as inverted time series.

In Section 2, we introduce ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p},\mathfrak{q}$) processes for any order $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ with values in the function spaces $L^p[0,1]$ with $p \in [1,\infty), C[0,1]$ and others. For these processes, we present sufficient conditions for the existence of strictly stationary solutions in Theorem 2.1, and for the existence of finite moments and weak dependence in Lemma 2.1. Theorem 2.1 generalizes Cerovecki *et al.* (2019), Theorem 1 under a milder condition, Hörmann *et al.* (2013), Theorem 2.1 and 2.3, and Aue *et al.* (2017), Theorem 2.1 and 2.2. To the best of our knowledge, for functional ARCH($\mathfrak{p}$) resp. GARCH($\mathfrak{p},\mathfrak{q}$) processes with $\mathfrak{p} > 1$ resp. $\mathfrak{p} \vee \mathfrak{q} > 1$, a moment condition as (2.10) in Theorem 2.1 and Lemma 2.1 is new. In Section 3, we derive explicit asymptotic upper bounds of the estimation errors for the shift term and the operators of $L^2[0,1]$-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p},\mathfrak{q}$) processes for all $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ where the operators are estimated by a Yule-Walker approach. For this purpose, we establish convergence results regarding asymptotic upper bounds of the estimation errors for certain means, covariance and lag-h-covariance operators (Lemma 3.3), and eigenfunctions and eigenvalues which are also useful beyond the context of ARCH and GARCH. Theorems 3.1-3.4 present the main results of the article. Theorem 3.1 states upper bounds of the estimation errors for the shift term in the ARCH and GARCH processes for any order. Theorem 3.2 provides upper bounds of the estimation errors for the ARCH($\mathfrak{p}$) operators for any $\mathfrak{p}$, namely for the projections of the operators on a finite-dimensional subspace in part (a), and the complete operators in part (b). A similar result as (a) for $\mathfrak{p} = 1$ was stated in Hörmann *et al.* (2013) by imposing an integral operator and estimating its kernel. However, as far as we know, (a) with $\mathfrak{p} > 1$ and (b) are new. Theorem 3.3 is a convergence result stating explicit asymptotic upper bounds of the estimation errors for the operators of invertible, linear processes represented as inverted time series. From this, one immediately obtains a result with the same upper bounds of the estimation errors for the operators in the associated linear process, see Kühnert (2019), Section 4.4.1.3. Both results are valid without the context of ARCH and GARCH and they extend some results in Aue & Klepsch (2017) and Klepsch & Klüppelberg (2017) where a different approach was made. At last, Theorem 3.4, which is based on Theorem 3.3, provides upper bounds of the estimation errors for the complete GARCH

operators. Projections of these operators on finite-dimensional subspaces were estimated in Aue *et al.* (2017) for $\mathfrak{p} = 1 = \mathfrak{q}$ by least squares estimators and in Cerovecki *et al.* (2019) for any order by a quasi-likelihood approach. Latter applied their same approach to estimate the complete operators in the case $\mathfrak{p} = 1 = \mathfrak{q}$. To the best of our knowledge, estimating the complete operators in Theorem 3.4 for $\mathfrak{p} \vee \mathfrak{q} > 1$ is new and explicit asymptotic upper bounds of estimation errors for complete operators of ARCH, ARMA, GARCH, invertible and linear processes as in the Theorems 3.2-3.4 have not been derived before.

We leave the investigations concerning the probabilistic properties of ARCH and GARCH processes in general, separable Banach spaces behind for future research. Estimating the orders of functional ARCH and GARCH processes is also an open problem, see Kokoszka & Reimherr (2013). Concerning parameter estimation in functional ARCH and GARCH processes, open problems are the estimation in general, separable Banach spaces, see Ruiz-Medina M.D. & Álvarez-Liébana J. (2019), the asymptotic distribution of the estimations errors when estimating the parameters without projecting them on a finite-dimensional subspace, see Aue *et al.* (2017) and Cerovecki *et al.* (2019) for the parameters projected on a finite-dimensinal subspace, and the asymptotic lower bounds of the estimations errors.

5   Proofs

In various conversions, we utilize the inequality

$$\Big(\sum_{k=1}^{n} a_k\Big)^{\nu} \leq \begin{cases} \sum_{k=1}^{n} a_k^{\nu}, & \nu \in (0,1], \\ n^{\nu-1}\sum_{k=1}^{n} a_k^{\nu}, & \nu \in (1,\infty), \end{cases}$$

for $n \in \mathbb{N}$ and $a_1, ..., a_n \geq 0$ where we usually write $(\sum_{k=1}^{n} a_k)^{\nu} \lesssim \sum_{k=1}^{n} a_k^{\nu}$. We also use the operator valued Hölder's inequality which we state in the following (see Kaballo (2014), Theorem 11.2). By $K = \sum_{j=1}^{\infty} s_j(K)(\mathfrak{k}_j \otimes \mathfrak{k}'_j)$ we denote the singular value decomposition of an operator $K \in \mathcal{L}_{\mathcal{H},\mathcal{H}'}$ where $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ and $(\mathcal{H}', \langle \cdot, \cdot \rangle_{\mathcal{H}'})$ are Hilbert spaces. For $p \in [1, \infty], \mathscr{S}_{\mathcal{H},\mathcal{H}'}^p \subseteq \mathcal{L}_{\mathcal{H},\mathcal{H}'}$ denotes the p-th Schatten-class. For $p \in [1, \infty)$, $A \in \mathscr{S}_{\mathcal{H},\mathcal{H}'}^p$ if $\sum_{j=1}^{\infty} s_j^p(A) < \infty$ where $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^p$ is endowed with the norm $||\cdot||_{\mathscr{S}_{\mathcal{H},\mathcal{H}'}^p} := (\sum_{j=1}^{\infty} s_j^p(\cdot))^{1/p}$ and where $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{\infty} := \mathcal{L}_{\mathcal{H},\mathcal{H}'}$ with $||\cdot||_{\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{\infty}} := ||\cdot||_{\mathcal{L}_{\mathcal{H},\mathcal{H}'}}$. Further, $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^1 = \mathcal{N}_{\mathcal{H},\mathcal{H}'}$ and $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^2 = \mathcal{S}_{\mathcal{H},\mathcal{H}'}$. Now, let $p, q, r \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ where $\frac{1}{\infty} := 0$ and let $A \in \mathscr{S}_{\mathcal{H}',\mathcal{H}''}^q$ and $B \in \mathscr{S}_{\mathcal{H},\mathcal{H}'}^p$ where $(\mathcal{H}'', \langle \cdot, \cdot \rangle_{\mathcal{H}''})$ is another Hilbert space. Then, after the operator valued Hölder's inequality, $AB \in \mathscr{S}_{\mathcal{H},\mathcal{H}''}^r$ with

$$||AB||_{\mathscr{S}_{\mathcal{H},\mathcal{H}''}^r} \leq 2^{1/r}||A||_{\mathscr{S}_{\mathcal{H}',\mathcal{H}''}^q}||B||_{\mathscr{S}_{\mathcal{H},\mathcal{H}'}^p}. \tag{5.1}$$

Proof of Theorem 2.1. (a) The state-space form (2.2) yields

$$\varsigma_k^{(\mathfrak{p},\mathfrak{q})} = \delta_k^{(\mathfrak{p},\mathfrak{q})} + \sum_{m=1}^{\infty} \Psi_k^{(\mathfrak{p},\mathfrak{q})}\Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\Psi_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}) \tag{5.2}$$

a.s. for all k if the series converges a.s. Further, (2.7) implies

$$\varlimsup_{m\to\infty}\frac{1}{m}\ln\big|\big|\Psi_k^{(\mathfrak{p},\mathfrak{q})}\Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\Psi_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})})\big|\big|_{\mathscr{H}^{\mathfrak{s}}} \leq \gamma^{(\mathfrak{p},\mathfrak{q})} + \varlimsup_{m\to\infty}\frac{1}{m}\ln||\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}||_{\mathscr{H}^{\mathfrak{s}}}.$$

By definition of $\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}$ and due to $0 < ||\delta||_{\mathscr{H}} \leq ||\delta||_{\dot{\mathscr{H}}}$, we have $||\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}||_{\mathscr{H}^{\mathfrak{s}}}^2 \leq ||\delta||_{\dot{\mathscr{H}}}^2(1 + ||\varepsilon_{k-m}^2||_{\dot{\mathscr{H}}}^2)$ which leads because of (2.3) to $\mathbb{E}\ln^+||\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}||_{\mathscr{H}^{\mathfrak{s}}} < \infty$. Since $\gamma^{(\mathfrak{p},\mathfrak{q})} < 0$, we obtain

$$\varlimsup_{m\to\infty}\big|\big|\Psi_k^{(\mathfrak{p},\mathfrak{q})}\Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\Psi_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})})\big|\big|_{\mathscr{H}^{\mathfrak{s}}}^{1/m} = e^{\gamma^{(\mathfrak{p},\mathfrak{q})}} < 1$$

a.s. and thus the series in (5.2) converges a.s. Hence, there exists a unique solution $(\varsigma_k^{(\mathfrak{p},\mathfrak{q})})_k$ of (2.2) (for uniqueness see Cerovecki *et al.* (2019), p. 19) and thus also of (2.1). By definition of $\varsigma_k^{(\mathfrak{p},\mathfrak{q})}$ and due to (5.2), $\sigma_k^2 = f(\varepsilon_{k-1}, \varepsilon_{k-2}, ...)$ a.s. for all k for some measurable function $f: F^{\infty} \to F$, thus $(\sigma_k^2)_k$ and $(\mathscr{X}_k)_k$ are nonanticipative w.r.t. $(\varepsilon_k)_k$ and strictly stationary as well as ergodic after Stout (1974), Theorem 3.5.8.

(b) $\nu > 0, \psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})} < 1$, (2.10), sub-multiplicativity of $||\cdot||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{s}}}}$ and Jensen's inequality imply

$$\gamma^{(\mathfrak{p},\mathfrak{q})} \leq \lim_{m\to\infty} \frac{1}{mn}\, \mathbb{E}\ln\Big(\prod_{l=1}^{m} ||\boldsymbol{\Psi}_{ln}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{ln-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{(l-1)n+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{s}}}}\Big) \leq \frac{1}{n\nu}\ln(\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})}) < 0. \qquad \square$$

Proof of Corollary 2.1. Assume $N = mn$ for some $m \in \mathbb{N}$ w.l.o.g. Then, since $\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})} < 1$, $||\cdot||_{\mathcal{L}_{F^{\mathfrak{s}}}}$ is sub-multiplicative and $(\boldsymbol{\Psi}_{k}^{(\mathfrak{p},\mathfrak{q})})_{k\in\mathbb{Z}}$ is i.i.d., the assertion follows from

$$\mathbb{E}||\boldsymbol{\varsigma}_{N}^{(\mathfrak{p},\mathfrak{q})} - \tilde{\boldsymbol{\varsigma}}_{N}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} = \mathbb{E}||\boldsymbol{\Psi}_{N}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{N-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{1}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\varsigma}_{0}^{(\mathfrak{p},\mathfrak{q})} - \tilde{\boldsymbol{\varsigma}}_{0}^{(\mathfrak{p},\mathfrak{q})})||_{F^{\mathfrak{s}}}^{\nu}$$
$$\leq (\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^{N/n}\,\mathbb{E}||\boldsymbol{\varsigma}_{0}^{(\mathfrak{p},\mathfrak{q})} - \tilde{\boldsymbol{\varsigma}}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}. \qquad \square$$

Proof of Lemma 2.1. (a) For all $\nu > 0$, we have

$$||\boldsymbol{\varsigma}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \lesssim ||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} + \Big(\sum_{m=1}^{\infty} ||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}\Big)^{\nu}$$

$$\lesssim ||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} + \sum_{m=1}^{n-1} ||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\prod_{l=1}^{m}||\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} + \Big(\sum_{m=n}^{\infty}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}\Big)^{\nu}.$$

Moreover, $\mathbb{E}||\varepsilon_{0}^{2}||_{F}^{\nu} < \infty$ implies $\mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} < \infty$ as well as $\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} < \infty$. From the definition of $\boldsymbol{\Psi}_{k}^{(\mathfrak{p},\mathfrak{q})}$ and $\boldsymbol{\delta}_{k}^{(\mathfrak{p},\mathfrak{q})}$ for all k and since $(\varepsilon_{k})_{k}$ is i.i.d. thus follows

$$\mathbb{E}\Big(\sum_{m=1}^{n-1} ||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\prod_{l=1}^{m}||\boldsymbol{\Psi}_{-m+l}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}\Big) = \mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\sum_{m=1}^{n-1}\big(\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}\big)^{m} < \infty.$$

Furthermore, (2.10) implies for $\nu \in (0,1]$:

$$\mathbb{E}\Big(\sum_{m=n}^{\infty} ||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}\Big)^{\nu} \leq \mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\sum_{m=n}^{\infty}\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}$$
$$\leq \mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\Big(\sum_{k=0}^{n-1}\big(\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}\big)^{k}\Big)\sum_{l=1}^{\infty}(\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^{l} < \infty$$

and for $\nu > 1$ together with Jensen's inequality as well as monotone convergence theorem:

$$\mathbb{E}\Big(\sum_{m=n}^{\infty} ||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}||\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}\Big)^{\nu} \leq \mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\Big(\sum_{m=n}^{\infty}\big(\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}\big)^{1/\nu}\Big)^{\nu}$$
$$\leq \mathbb{E}||\boldsymbol{\delta}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu}\Big(\sum_{k=0}^{n-1}\big(\mathbb{E}||\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}\big)^{k}\Big)\Big(\sum_{l=1}^{\infty}(\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^{l}\Big)^{\nu} < \infty.$$

Subsequently, $\mathbb{E}||\boldsymbol{\varsigma}_{0}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} < \infty$ for all $\mathfrak{p} \in \mathbb{N}, \mathfrak{q} \in \mathbb{N}_{0}$ and thus $\mathbb{E}||\mathscr{X}_{0}^{2}||_{F}^{\nu} < \infty, \mathbb{E}||\sigma_{0}^{2}||_{F}^{\nu} < \infty$ as well as

$$\mathbb{E}||\sigma_{0}^{2}||_{F}^{\nu} \lesssim ||\delta||_{F}^{\nu} + \sum_{i=1}^{\mathfrak{p}}\mathbb{E}||\alpha_{i}(\mathscr{X}_{i}^{2})||_{F}^{\nu} + \sum_{j=1}^{\mathfrak{q}}\mathbb{E}||\beta_{j}(\sigma_{j}^{2})||_{F}^{\nu}$$
$$\leq ||\delta||_{F}^{\nu} + \mathbb{E}||\mathscr{X}_{0}^{2}||_{F}^{\nu}\sum_{i=1}^{\mathfrak{p}}||\alpha_{i}||_{\mathcal{L}_{F,\dot{F}}}^{\nu} + \mathbb{E}||\sigma_{0}^{2}||_{F}^{\nu}\sum_{j=1}^{\mathfrak{q}}||\beta_{j}||_{\mathcal{L}_{F,\dot{F}}}^{\nu} < \infty.$$

(b) From the identity (5.2) follows

$$\boldsymbol{\varsigma}_{m,m}^{(\mathfrak{p},\mathfrak{q})} := \boldsymbol{\delta}_{m}^{(\mathfrak{p},\mathfrak{q})} + \sum_{l=1}^{m-1}\boldsymbol{\Psi}_{m}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\delta}_{m-l}^{(\mathfrak{p},\mathfrak{q})}) + \sum_{l=m}^{\infty}\boldsymbol{\Psi}_{m}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{1}^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{0}^{(\mathfrak{p},\mathfrak{q},m)}\cdots\boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q},m)}(\boldsymbol{\delta}_{m-l}^{(\mathfrak{p},\mathfrak{q},m)})$$

a.s. for all $m \in \mathbb{N}$. Thereby, $\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q},m)}$ and $\boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q},m)}$ stand for $\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}$ resp. $\boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q})}$ in (2.2) depending on $\varepsilon_k^{(m)}$ where $(\varepsilon_k^{(m)})_{k\in\mathbb{Z}}$ are i.i.d. time series for all m, which are independent of each other with $\varepsilon_k^{(m)} \stackrel{d}{=} \varepsilon_0$ for all k, m. Consequently, for any m we have

$$||\boldsymbol{\varsigma}_m^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_{m,m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}} \leq \sum_{l=m}^{\infty} \Big(||\boldsymbol{\Psi}_m^{(\mathfrak{p},\mathfrak{q})} \cdots \boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}} ||\boldsymbol{\delta}_{m-l}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}$$
$$+ ||\boldsymbol{\Psi}_m^{(\mathfrak{p},\mathfrak{q})} \cdots \boldsymbol{\Psi}_1^{(\mathfrak{p},\mathfrak{q})} \boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q},m)} \cdots \boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q},m)}||_{\mathcal{L}_{F^{\mathfrak{s}}}} ||\boldsymbol{\delta}_{m-l}^{(\mathfrak{p},\mathfrak{q},m)}||_{F^{\mathfrak{s}}} \Big).$$

From this identity, the proof of (a) and since $\varepsilon_k^{(n)}$ and ε_l are i.i.d. for all k, l, n, it follows in the case $\nu \in (0, 1]$:

$$\mathbb{E}||\boldsymbol{\varsigma}_m^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_{m,m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \leq 2\,\mathbb{E}||\boldsymbol{\delta}_0^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \sum_{l=m}^{\infty} \mathbb{E}||\boldsymbol{\Psi}_m^{(\mathfrak{p},\mathfrak{q})} \cdots \boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu}$$
$$\leq 2\,\mathbb{E}||\boldsymbol{\delta}_0^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \Big(\sum_{k=0}^{n-1} \big(\mathbb{E}||\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} \big)^k \Big) \sum_{j=m}^{\infty} (\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^j \propto (\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^m$$

and in the case $\nu > 1$, based on the argumentation in the proof of (a) for $\nu > 1$:

$$\mathbb{E}||\boldsymbol{\varsigma}_m^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_{m,m}^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \leq \Big(\sum_{m=n}^{\infty} 2 \big(\mathbb{E}||\boldsymbol{\Psi}_m^{(\mathfrak{p},\mathfrak{q})} \cdots \boldsymbol{\Psi}_{m-l+1}^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} \big)^{1/\nu} \Big)^{\nu}$$
$$\leq 2^{\nu}\,\mathbb{E}||\boldsymbol{\delta}_0^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} \Big(\sum_{k=0}^{n-1} \big(\mathbb{E}||\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}^{\nu} \big)^k \Big) \Big(\sum_{j=m}^{\infty} (\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^{j/\nu} \Big)^{\nu} \propto (\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})})^{m/\nu}. \qquad \square$$

Proof of Lemma 3.1. See Kühnert (2019), Lemma 4.11. $\hfill \square$

Proof of Lemma 3.2. See Kühnert (2019), Theorem 4.2. $\hfill \square$

Proof of Lemma 3.3. See Kühnert (2019), Theorem 4.4. $\hfill \square$

Proof of Theorem 3.1. The assertion follows from (3.6) and

$$||\hat{\delta} - \delta||_{\mathscr{H}} \leq ||\hat{m}_2 - m_2||_{\mathscr{H}} \Big(1 + \sum_{i=1}^{\mathfrak{r}} ||\alpha_i + \beta_i||_{\mathcal{L}_{\mathscr{H}}} \Big) + ||\hat{m}_2||_{\mathscr{H}} \Big(\sum_{i=1}^{\mathfrak{r}} ||\hat{\alpha}_i - \alpha_i||_{\mathcal{L}_{\mathscr{H}}} + ||\hat{\beta}_i - \beta_i||_{\mathcal{L}_{\mathscr{H}}} \Big). \qquad \square$$

Proof of Theorem 3.2. (a) The definition of $\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]}$, the inequality (5.1), $\mathfrak{S}_\mathfrak{p}^{\dagger} := \mathfrak{S}_\mathfrak{p}(\mathfrak{S}_\mathfrak{p}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^\mathfrak{p}})^{-1}$, the Yule-Walker equation $\mathfrak{S}_{\mathfrak{p},1} = \boldsymbol{\alpha}\mathfrak{S}_\mathfrak{p}$, $\mathfrak{S}_\mathfrak{p}^{\ddagger} := \mathfrak{S}_\mathfrak{p}\mathfrak{S}_\mathfrak{p}^{\dagger}$, (3.10), $\mathbb{I}_{\mathscr{H}^\mathfrak{p}} = \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} + \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},K+1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},\infty}}$ and $\boldsymbol{\alpha} = \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha} + \coprod_{\mathscr{J}_{\mathfrak{p},K}^c} \boldsymbol{\alpha}$ with $\mathscr{J}_{\mathfrak{p},K} := \{\Phi_{\mathfrak{p},ij} | 1 \leq i, j \leq K\}$ and $\mathscr{J}_{\mathfrak{p},K}^c := \{\Phi_{\mathfrak{p},ij} | i, j \in \mathbb{N}, i \vee j > K\}$, yields for fixed $K \in \mathbb{N}$:

$$\Big|\Big|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2 \lesssim \Big|\Big|(\hat{\mathfrak{S}}_{\mathfrak{p},1} - \mathfrak{S}_{\mathfrak{p},1})\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2 + \Big|\Big|\mathfrak{S}_{\mathfrak{p},1}\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}} - \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2$$

$$\lesssim ||\hat{\mathfrak{S}}_{\mathfrak{p},1} - \mathfrak{S}_{\mathfrak{p},1}||_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2 \Big|\Big|\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{L}_{\mathscr{H}^\mathfrak{p}}}^2 + ||\mathfrak{S}_{\mathfrak{p},1}||_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2 \Big|\Big|\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} - \mathfrak{S}_\mathfrak{p}^{\dagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{L}_{\mathscr{H}^\mathfrak{p}}}^2$$

$$+ \Big|\Big|\boldsymbol{\alpha}_{[\mathfrak{p}]}\mathfrak{S}_\mathfrak{p}^{\ddagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}} - \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2$$

$$\lesssim \mathrm{O}_{\mathbb{P}}(N^{-1}) \Big|\Big|\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{L}_{\mathscr{H}^\mathfrak{p}}}^2 + \Big|\Big|\hat{\mathfrak{S}}_\mathfrak{p}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} - \mathfrak{S}_\mathfrak{p}^{\dagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{L}_{\mathscr{H}^\mathfrak{p}}}^2 + \Big|\Big|\coprod_{\mathscr{J}_{\mathfrak{p},K}^c} \boldsymbol{\alpha}_{[\mathfrak{p}]}\mathfrak{S}_\mathfrak{p}^{\ddagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2$$

$$+ \Big|\Big|\coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\big[\mathfrak{S}_\mathfrak{p}^{\ddagger} \coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}} - \mathbb{I}_{\mathscr{H}^\mathfrak{p}}\big]\Big|\Big|_{\mathcal{S}_{\mathscr{H}^\mathfrak{p},\mathscr{H}}}^2$$

$$=: \mathrm{O}_{\mathbb{P}}(N^{-1}) \cdot T_1 + T_2 + T_3 + T_4. \tag{5.3}$$

<u>Term T_1</u>: As per definition of $||\cdot||_{\mathcal{L}_{\mathscr{H}}}, \hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger} := \hat{\mathfrak{S}}_{\mathfrak{p}}(\hat{\mathfrak{S}}_{\mathfrak{p}}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1}$, since $(\hat{\mathfrak{c}}_{\mathfrak{p},j})_j$ is the eigenfunction sequence of $\hat{\mathfrak{S}}_{\mathfrak{p}}$ related to the eigenvalue sequence $(\hat{c}_{\mathfrak{p},j})_j$ with $\hat{c}_{\mathfrak{p},1} \geq \cdots \geq \hat{c}_{\mathfrak{p},K}$ w.l.o.g., since $\frac{\hat{c}_j}{(\hat{c}_j+\vartheta_N)^2} = 0$ if $\hat{c}_j = 0$ resp. $\frac{\hat{c}_j}{(\hat{c}_j+\vartheta_N)^2} \leq \hat{c}_j^{-1}$ if $\hat{c}_j \neq 0$ and because of (3.15) with $K_N := K$ for all N, we have

$$||\hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger}\coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{p}}}}^2 = \sup_{j\leq K}\Big(\frac{\hat{c}_{\mathfrak{p},j}}{\hat{c}_{\mathfrak{p},j}+\vartheta_N}\Big)^2 = \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-2}) = \mathrm{O}_{\mathbb{P}}(1). \tag{5.4}$$

<u>Term T_2</u>: From the definition of $\hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger}, \mathfrak{S}_{\mathfrak{p}}^{\dagger}$ and $||\cdot||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{p}}}}$, since $(\mathfrak{c}_{\mathfrak{p},j})_j$ is a CONS of $\mathscr{H}^{\mathfrak{p}}$ and $(\hat{\mathfrak{c}}_{\mathfrak{p},j}''')_j$ is a CONS of $\mathscr{H}^{\mathfrak{p}}$ a.s., Corollary 3.1 and Lemma 3.4 imply

$$||\hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger}\coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} - \mathfrak{S}_{\mathfrak{p}}^{\dagger}\coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{p}}}}^2 = \sup_{||\boldsymbol{x}||_{\mathscr{H}^{\mathfrak{p}}}\leq 1}||\sum_{j=1}^{K}\frac{\hat{c}_{\mathfrak{p},j}}{\hat{c}_{\mathfrak{p},j}^2+\vartheta_N}\langle\boldsymbol{x},\hat{\mathfrak{c}}_{\mathfrak{p},j}'''\rangle_{\mathscr{H}^{\mathfrak{p}}}\hat{\mathfrak{c}}_{\mathfrak{p},j}''' - \frac{c_{\mathfrak{p},j}}{c_{\mathfrak{p},j}^2+\vartheta_N}\langle\boldsymbol{x},\mathfrak{c}_{\mathfrak{p},j}\rangle_{\mathscr{H}^{\mathfrak{p}}}\mathfrak{c}_{\mathfrak{p},j}||_{\mathscr{H}^{\mathfrak{p}}}^2$$

$$\precsim \sup_{j\leq K}\Big|\frac{\hat{c}_{\mathfrak{p},j}}{\hat{c}_{\mathfrak{p},j}^2+\vartheta_N} - \frac{c_{\mathfrak{p},j}}{c_{\mathfrak{p},j}^2+\vartheta_N}\Big|^2 + \sup_{||\boldsymbol{x}||_{\mathscr{H}^{\mathfrak{p}}}\leq 1}\sum_{j=1}^{K}\Big|\frac{c_{\mathfrak{p},j}}{c_{\mathfrak{p},j}^2+\vartheta_N}\Big|^2\langle\boldsymbol{x},\hat{\mathfrak{c}}_{\mathfrak{p},j}'''-\mathfrak{c}_{\mathfrak{p},j}\rangle_{\mathscr{H}^{\mathfrak{p}}}^2$$

$$+\sup_{||\boldsymbol{x}||_{\mathscr{H}^{\mathfrak{p}}}\leq 1}||\sum_{j=1}^{K}\frac{c_{\mathfrak{p},j}}{c_{\mathfrak{p},j}^2+\vartheta_N}\langle\boldsymbol{x},\mathfrak{c}_{\mathfrak{p},j}\rangle_{\mathscr{H}^{\mathfrak{p}}}(\hat{\mathfrak{c}}_{\mathfrak{p},j}'''-\mathfrak{c}_{\mathfrak{p},j})||_{\mathscr{H}^{\mathfrak{p}}}^2$$

$$\precsim \sup_{j\leq K}(\hat{c}_{\mathfrak{p},j}-c_{\mathfrak{p},j})^2\frac{\vartheta_N^2+\hat{c}_{\mathfrak{p},j}^2 c_{\mathfrak{p},j}^2}{(\hat{c}_{\mathfrak{p},j}^2+\vartheta_N)^2(c_{\mathfrak{p},j}^2+\vartheta_N)^2} + (K+1)c_{\mathfrak{p},K}^{-2}\sup_{j\leq K}||\hat{\mathfrak{c}}_{\mathfrak{p},j}'''-\mathfrak{c}_{\mathfrak{p},j}||_{\mathscr{H}^{\mathfrak{p}}}^2$$

$$= \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-4}N^{-1}) + \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-4}KN^{-1})$$

$$= \mathrm{O}_{\mathbb{P}}(N^{-1}). \tag{5.5}$$

<u>Term T_3</u>: The eigenfunction sequences $(\mathfrak{c}_{\mathfrak{p},j})_j$ of $\mathfrak{S}_{\mathfrak{p}}, (\mathfrak{c}_j)_j$ of $\mathscr{C}_{0;\boldsymbol{x}}$ and the sequence $(\Phi_{\mathfrak{p},ij})_{i,j}$ with $\Phi_{\mathfrak{p},ij}(\mathfrak{c}_{\mathfrak{p},k}) = \delta_{ik}\mathfrak{c}_j$ for all i,j,k are CONS of $\mathscr{H}^{\mathfrak{p}}, \mathscr{H}$ resp. of $\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}$. Moreover, since $\mathfrak{S}_{\mathfrak{p}}^{\ddagger} := \mathfrak{S}_{\mathfrak{p}}^2(\mathfrak{S}_{\mathfrak{p}}^2+\vartheta_N\mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1}, \mathscr{J}_{\mathfrak{p},K}^c := \{\Phi_{\mathfrak{p},ij}|i,j\in\mathbb{N}, i\vee j > K\}, \boldsymbol{\alpha}_{[\mathfrak{p}]} = \sum_{i,j=1}^{\infty}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]},\Phi_{\mathfrak{p},ij}\rangle_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}\Phi_{\mathfrak{p},ij}, \langle\boldsymbol{\alpha}_{[\mathfrak{p}]},\Phi_{\mathfrak{p},ij}\rangle_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}} = \sum_{k=1}^{\infty}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},k}),\Phi_{\mathfrak{p},ij}(\mathfrak{c}_{\mathfrak{p},k})\rangle_{\mathscr{H}} = \langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},i}),\mathfrak{c}_j\rangle_{\mathscr{H}}$ for all i,j and since we imposed $\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}} = 0$ for all $j > K, l \leq K$, we thus obtain

$$||\coprod_{\mathscr{J}_{\mathfrak{p},K}^c}\boldsymbol{\alpha}_{[\mathfrak{p}]}\mathfrak{S}_{\mathfrak{p}}^{\ddagger}\coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \sum_{l=1}^{K}||\coprod_{\mathscr{J}_{\mathfrak{p},K}^c}\boldsymbol{\alpha}_{[\mathfrak{p}]}\mathfrak{S}_{\mathfrak{p}}^{\ddagger}(\mathfrak{c}_{\mathfrak{p},l})||_{\mathscr{H}}^2 = \sum_{l=1}^{K}\Big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N}\Big)^2||\coprod_{\mathscr{J}_{\mathfrak{p},K}^c}\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l})||_{\mathscr{H}}^2$$

$$= \sum_{l=1}^{K}\Big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N}\Big)^2||\sum_{\substack{i,j\in\mathbb{N}\\i\vee j>K}}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},i}),\mathfrak{c}_j\rangle_{\mathscr{H}}\delta_{il}\mathfrak{c}_j||_{\mathscr{H}}^2 = \sum_{l=1}^{K}\Big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N}\Big)^2\sum_{j>K}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}}^2$$

$$= 0. \tag{5.6}$$

<u>Term T_4</u>: Elementary transformations and transformations as used in the terms T_1-T_3, $\coprod_{\mathscr{J}_{\mathfrak{p},K}}\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}) = \mathbf{1}_{\mathbb{N}_{\leq K}}(l)\sum_{j=1}^{K}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}}\mathfrak{c}_j$ for all l,K and $\vartheta_N = \mathrm{O}(N^{-1/2})$ imply

$$||\coprod_{\mathscr{J}_{\mathfrak{p},K}}\boldsymbol{\alpha}_{[\mathfrak{p}]}[\mathfrak{S}_{\mathfrak{p}}^{\ddagger}\coprod_{\mathfrak{c}_{\mathfrak{p},1}}^{\mathfrak{c}_{\mathfrak{p},K}} - \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}}]||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 = \sum_{l=1}^{\infty}||\coprod_{\mathscr{J}_{\mathfrak{p},K}}\boldsymbol{\alpha}_{[\mathfrak{p}]}\big([\mathbf{1}_{\mathbb{N}_{\leq K}}(l)\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N} - 1]\mathfrak{c}_{\mathfrak{p},l}\big)||_{\mathscr{H}}^2$$

$$= \sum_{l=1}^{\infty}||[\mathbf{1}_{\mathbb{N}_{\leq K}}(l)\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N} - 1]\mathbf{1}_{\mathbb{N}_{\leq K}}(l)\sum_{j=1}^{K}\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}}\mathfrak{c}_j||_{\mathscr{H}}^2$$

$$= \sum_{j,l=1}^{K}\Big(\frac{\vartheta_N}{c_{\mathfrak{p},l}^2+\vartheta_N}\Big)^2\langle\boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}),\mathfrak{c}_j\rangle_{\mathscr{H}}^2 \leq \vartheta_N^2 c_{\mathfrak{p},K}^{-4}||\boldsymbol{\alpha}_{[\mathfrak{p}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2$$

$$= \mathrm{O}(N^{-1}). \tag{5.7}$$

Replacing T_1-T_4 in (5.3) by (5.4)-(5.7) indeed yields (3.33).

(b) From $\boldsymbol{\alpha}_{[\mathfrak{p}]} = \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]} + \coprod_{\mathscr{J}_{\mathfrak{p},K}^c} \boldsymbol{\alpha}_{[\mathfrak{p}]}$, (3.26), part (a) and $\vartheta_N = \mathrm{O}(N^{-1/2})$ follows for any sequence $K = K_N = \Xi(1,N)$ with $c_{\mathfrak{p},K}^{-4} K^{2\beta+1} = \mathrm{O}(N)$ and $\sum_{l=1}^{K} (\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2 + \vartheta_N})^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = \mathrm{O}(K^{-2\beta})$:

$$||\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 \precsim ||\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 + ||\coprod_{\mathscr{J}_{\mathfrak{p},K}^c} \boldsymbol{\alpha}_{[\mathfrak{p}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2$$

$$\precsim \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-2} N^{-1}) + \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-4} N^{-1}) + \mathrm{O}_{\mathbb{P}}(c_{\mathfrak{p},K}^{-4} K N^{-1}) +$$

$$+ \sum_{l=1}^{K} \left(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2 + \vartheta_N}\right)^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 + \vartheta_N^2 c_{\mathfrak{p},K}^{-4} ||\boldsymbol{\alpha}_{[\mathfrak{p}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}}^2 + \mathrm{O}(K^{-2\beta})$$

$$= \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \qquad \qquad \square$$

Proof of Theorem 3.3. The proof is based on the proof of Theorem 3.2 with $\mathfrak{p}$ replaced by an appropriate sequence $L = L_N \to \infty$. From the ideas in the proof of Theorem 3.2 and (see Kühnert (2019), Lemma 4.45)

$$\mathfrak{S}_{L,1} = \boldsymbol{\pi}_L \mathfrak{S}_L + \sum_{l>L} \pi_l \mathfrak{S}_{L,1-l}, \qquad (5.8)$$

which can be identified as a Yule-Walker equation with a residual, follows for all K, L, N:

$$||\hat{\boldsymbol{\pi}}_{L,K} - \boldsymbol{\pi}_L||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 \precsim ||\hat{\boldsymbol{\pi}}_{L,K} - \coprod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 + ||\coprod_{\mathscr{J}_{L,K}^c} \boldsymbol{\pi}_L||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2$$

$$\precsim ||\hat{\mathfrak{S}}_{L,1} - \mathfrak{S}_{L,1}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 ||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2 + ||\mathfrak{S}_{L,1}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 ||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} - \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2$$

$$+ ||\mathfrak{S}_{L,1} \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}} - \coprod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 + \mathrm{O}(K^{-2\beta})$$

$$\precsim \mathrm{O}_{\mathbb{P}}(L^2 N^{-1}) ||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2 + L \cdot ||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} - \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2 + ||[\sum_{l>L} \pi_l \mathfrak{S}_{L,1-l}] \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2$$

$$+ ||\coprod_{\mathscr{J}_{L,K}^c} \boldsymbol{\pi}_L \mathfrak{S}_L^{\ddagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 + ||\coprod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L [\mathfrak{S}_L^{\ddagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}} - \mathbb{I}_{\mathscr{H}^L}]||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 + \mathrm{O}(K^{-2\beta})$$

$$=: \mathrm{O}_{\mathbb{P}}(L^2 N^{-1}) \cdot T_1 + L \cdot T_2 + T_3 + T_4 + T_5 + \mathrm{O}(K^{-2\beta}). \qquad (5.9)$$

Term T_1: Conversions similar as in T_1 in the proof of Theorem 3.2 and Corollary 3.1 yield

$$||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2 = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-2}). \qquad (5.10)$$

Term T_2: (3.12) with $L = \Xi(1, N^{1/3})$ and the argumentation of T_2 in the proof of Theorem 3.2 yield

$$||\hat{\mathfrak{S}}_L^{\dagger} \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} - \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{L}_{\mathscr{H}^L}}^2 = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-4} L^3 N^{-1}) + \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-4} K L^3 N^{-1}) = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-4} K L^3 N^{-1}). \qquad (5.11)$$

Term T_3: From the operator-valued Hölder's inequality (5.1), (5.10), triangle inequality and (3.8) follows

$$||[\sum_{l>L} \pi_l \mathfrak{S}_{L,1-l}] \mathfrak{S}_L^{\dagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 \leq c_{L,K}^{-2} ||\sum_{l>L} \pi_l \mathfrak{S}_{L,1-l}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 = \mathrm{O}\left(L\left(c_{L,K}^{-1} \sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathscr{H}}}\right)^2\right). \qquad (5.12)$$

Term T_4: Here, almost one to one as in the proof of Theorem 3.2, we obtain by assumption

$$||\coprod_{\mathscr{J}_{L,K}^c} \boldsymbol{\pi}_L \mathfrak{S}_L^{\ddagger} \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||_{\mathcal{S}_{\mathscr{H}^L,\mathscr{H}}}^2 = \sum_{l=1}^{K} \left(\frac{c_{L,l}^2}{c_{L,l}^2 + \vartheta_N}\right)^2 \sum_{j>K} \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = \mathrm{O}(K^{-2\beta}). \qquad (5.13)$$

17

<u>Term T_5</u>: $||\boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} = \sum_{l=1}^{L} ||\pi_l||^2_{\mathcal{S}_{\mathcal{H}}} \leq \sum_{l=1}^{\infty} ||\pi_l||^2_{\mathcal{S}_{\mathcal{H}}} < \infty$ implies as in T_4 in the proof of Theorem 3.2:

$$||\coprod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L [\mathfrak{S}_L^{\ddagger} \coprod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}} - \mathbb{I}_{\mathcal{H}^L}]||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} \precsim \vartheta_N^2 c_{L,K}^{-4} ||\boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} = \mathrm{O}(c_{L,K}^{-4} \vartheta_N^2). \tag{5.14}$$

Since $L = \Xi(1,N), K = \Xi(1, \sqrt[3]{L^{-1}N}), c_{L,K} = \Omega(\sqrt{L^3 N^{-1}}), (\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$ satisfies Assumption 3.4 for all $L \in \mathbb{N}$ for $\beta > 0$, (3.26) and $\vartheta_N = \mathrm{O}(c_{L,K}^{-2} K^{-\beta})$, plugging (5.10)-(5.14) of T_1-T_5 into (5.9) implies

$$||\hat{\boldsymbol{\pi}}_{L,K} - \boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} \precsim ||\hat{\boldsymbol{\pi}}_{L,K} - \coprod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} + ||\coprod_{\mathscr{J}_{L,K}^c} \boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}}$$

$$= \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-4} K L^4 N^{-1}) + \mathrm{O}\big(L\big(c_{L,K}^{-1} \sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathcal{H}}}\big)^2\big) + \mathrm{O}(K^{-2\beta}).$$

Consequently, since we imposed $c_{L,K}^{-4} L^4 N^{-1} = \mathrm{O}(K^{-(2\beta+1)})$ if $\sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathcal{H}}} = \mathrm{O}(c_{L,K}^{-1}\sqrt{KL^3N^{-1}})$ and $L^{-2}N(\sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathcal{H}}})^4 = \mathrm{O}(K^{1-2\beta})$ if $c_{L,K}^{-1}\sqrt{KL^3N^{-1}} = \mathrm{O}(\sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathcal{H}}})$, where latter is because of $c_{L,K} = \mathrm{O}(K^{-1}L)$ and $\sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathcal{H}}} = \mathrm{o}(1)$ only possible if $K = \mathrm{o}(\sqrt[3]{L^{-1}N})$, (3.43) is verified. $\qquad\square$

Proof of Theorem 3.4. The definition of $\hat{\boldsymbol{\beta}}_{[\mathsf{q}]}$, (3.44) and the fact that $(\boldsymbol{\beta}_{[\mathsf{q}]}, (\Phi_{\mathsf{p},\mathsf{q};ij})_{i,j})$ satisfies Assumption 3.4 for $\beta > 0$, imply similarly as in the proof of Theorem 3.2:

$$||\hat{\boldsymbol{\beta}}_{[\mathsf{q}]} - \boldsymbol{\beta}_{[\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} \precsim ||\hat{\boldsymbol{\beta}}_{[\mathsf{q}]} - \coprod_{\mathscr{J}_{\mathsf{p},\mathsf{q};M}} \boldsymbol{\beta}_{[\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} + ||\coprod_{\mathscr{J}_{\mathsf{p},\mathsf{q};M}^c} \boldsymbol{\beta}_{[\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}}$$

$$\precsim ||\hat{\boldsymbol{\pi}}_{[\mathsf{s},\mathsf{q}]} - \boldsymbol{\pi}_{[\mathsf{s},\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} ||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}} + ||\boldsymbol{\pi}_{[\mathsf{s},\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} ||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}} - \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}}$$

$$+ ||\boldsymbol{\pi}_{[\mathsf{s},\mathsf{q}]} \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}} - \coprod_{\mathscr{J}_{\mathsf{p},\mathsf{q};M}} \boldsymbol{\beta}_{[\mathsf{q}]}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} + \mathrm{O}(M^{-2\beta})$$

$$\precsim \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) ||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}} + ||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}} - \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}}$$

$$+ ||\coprod_{\mathscr{J}_{\mathsf{p},\mathsf{q};M}^c} \boldsymbol{\beta}_{[\mathsf{q}]} \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} + ||\coprod_{\mathscr{J}_{\mathsf{p},\mathsf{q};M}} \boldsymbol{\beta}_{[\mathsf{q}]} [\prod_{[\mathsf{s},\mathsf{q}]}^{\ddagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}} - \mathbb{I}_{\mathcal{H}^\mathsf{q}}]||^2_{\mathcal{S}_{\mathcal{H}^\mathsf{q},\mathcal{H}}} + \mathrm{O}(M^{-2\beta})$$

$$=: \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) \cdot T_1 + T_2 + T_3 + T_4 + \mathrm{O}(M^{-2\beta}). \tag{5.15}$$

<u>Term T_1</u>: Since $\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} = \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*} (\hat{\prod}_{[\mathsf{s},\mathsf{q}]} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*} + \theta_N \mathbb{I}_{\mathcal{H}^\mathsf{q}})^{-1}$ and since $(\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j})_j$ is the eigenfunction sequence of $\hat{\prod}_{[\mathsf{s},\mathsf{q}]} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*}$, thus $\hat{\prod}_{[\mathsf{s},\mathsf{q}]} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*}$ and $\coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}}$ commute, yields similarly as in Term T_1 of Theorem 3.2:

$$||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}} = ||(\hat{\prod}_{[\mathsf{s},\mathsf{q}]} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*} + \theta_N \mathbb{I}_{\mathcal{H}^\mathsf{q}})^{-2} \hat{\prod}_{[\mathsf{s},\mathsf{q}]} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}}||_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}} = \sup_{j \leq M} \frac{\hat{g}_{\mathsf{s},\mathsf{q};j}}{(\hat{g}_{\mathsf{s},\mathsf{q};j} + \theta_N)^2}$$

$$= \mathrm{O}_{\mathbb{P}}(g_{\mathsf{s},\mathsf{q};M}^{-1}). \tag{5.16}$$

<u>Term T_2</u>: Due to $\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger}(\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j}''') = (\hat{g}_{\mathsf{s},\mathsf{q};j} + \theta_N)^{-1} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*}(\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j}'''), \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger}(\mathfrak{g}_{\mathsf{s},\mathsf{q};j}''') = (g_{\mathsf{s},\mathsf{q};j} + \theta_N)^{-1} \prod_{[\mathsf{s},\mathsf{q}]}^{*}(\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j}''')$ and conversions as in Term T_2 in Theorem 3.2, we obtain

$$||\hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};1}}^{\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};M}} - \prod_{[\mathsf{s},\mathsf{q}]}^{\dagger} \coprod_{\mathfrak{g}_{\mathsf{s},\mathsf{q};1}}^{\mathfrak{g}_{\mathsf{s},\mathsf{q};M}}||^2_{\mathcal{L}_{\mathcal{H}^\mathsf{q}}}$$

$$= \sup_{||\boldsymbol{x}||_{\mathcal{H}^\mathsf{q}} \leq 1} ||\sum_{j=1}^{M} \frac{\langle \boldsymbol{x}, \hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j}'''\rangle_{\mathcal{H}^\mathsf{q}}}{\hat{g}_{\mathsf{s},\mathsf{q};j} + \theta_N} \hat{\prod}_{[\mathsf{s},\mathsf{q}]}^{*}(\hat{\mathfrak{g}}_{\mathsf{s},\mathsf{q};j}''') - \frac{\langle \boldsymbol{x}, \mathfrak{g}_{\mathsf{s},\mathsf{q};j}\rangle_{\mathcal{H}^\mathsf{q}}}{g_{\mathsf{s},\mathsf{q};j} + \theta_N} \prod_{[\mathsf{s},\mathsf{q}]}^{*}(\mathfrak{g}_{\mathsf{s},\mathsf{q};j})||^2_{\mathcal{H}^\mathsf{q}}$$

18

$$
\precsim \sup_{||\boldsymbol{x}||_{\mathscr{H}^{\mathfrak{q}}}\leq 1}\sum_{j=1}^{M}\Big|\frac{\langle\boldsymbol{x},\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j}'''\rangle_{\mathscr{H}^{\mathfrak{q}}}\sqrt{\hat{g}_{\mathfrak{s},\mathfrak{q};j}}}{\hat{g}_{\mathfrak{s},\mathfrak{q};j}+\theta_N}-\frac{\langle\boldsymbol{x},\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}\rangle_{\mathscr{H}^{\mathfrak{q}}}\sqrt{g_{\mathfrak{s},\mathfrak{q};j}}}{g_{\mathfrak{s},\mathfrak{q};j}+\theta_N}\Big|^2+\Big|\frac{\langle\boldsymbol{x},\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}\rangle_{\mathscr{H}^{\mathfrak{q}}}}{g_{\mathfrak{s},\mathfrak{q};j}+\theta_N}\Big|^2(\sqrt{\hat{g}_{\mathfrak{s},\mathfrak{q};j}}-\sqrt{g_{\mathfrak{s},\mathfrak{q};j}})^2
$$

$$
+\sup_{j\leq M}\Big|\frac{1}{g_{\mathfrak{s},\mathfrak{q};j}+\theta_N}\Big|^2\Big(||\hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^{*}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{q}}}}^2||\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j}'''-\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}||_{\mathscr{H}^{\mathfrak{q}}}^2+||\hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^{*}-\prod_{[\mathfrak{s},\mathfrak{q}]}^{*}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{q}}}}^2\Big)
$$

$$
\precsim \sup_{j\leq M}\Big|\frac{\sqrt{\hat{g}_{\mathfrak{s},\mathfrak{q};j}}}{\hat{g}_{\mathfrak{s},\mathfrak{q};j}+\theta_N}-\frac{\sqrt{g_{\mathfrak{s},\mathfrak{q};j}}}{g_{\mathfrak{s},\mathfrak{q};j}+\theta_N}\Big|^2+Mg_{\mathfrak{s},\mathfrak{q};M}^{-1}\sup_{j\leq M}||\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j}'''-\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}||_{\mathscr{H}^{\mathfrak{q}}}^2
$$

$$
+\sup_{j\leq M}\Big(\frac{1}{g_{\mathfrak{s},\mathfrak{q};j}+\theta_N}\Big)^2\Big((\sqrt{\hat{g}_{\mathfrak{s},\mathfrak{q};j}}-\sqrt{g_{\mathfrak{s},\mathfrak{q};j}})^2+||\hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^{*}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{q}}}}^2||\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};j}'''-\mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}||_{\mathscr{H}^{\mathfrak{q}}}^2+||\hat{\prod}_{[\mathfrak{s},\mathfrak{q}]}^{*}-\prod_{[\mathfrak{s},\mathfrak{q}]}^{*}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{q}}}}^2\Big)
$$

$$
=\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-2}K^{-\beta})+\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-3}K^{-2\beta})+\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-2}K^{-\beta})+\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-4}K^{-2\beta})
$$

$$
=\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-2}K^{-\beta}). \tag{5.17}
$$

<u>Term T_3</u>: From the assumptions made and the idea of the derivation of T_3 in Theorem 3.2 follows

$$
\Big|\Big|\coprod_{\mathscr{J}_{\mathfrak{p},\mathfrak{q};M}^{c}}\boldsymbol{\beta}_{[\mathfrak{q}]}\prod_{[\mathfrak{s},\mathfrak{q}]}^{\dagger}\coprod_{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};1}}^{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};M}}\Big|\Big|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}},\mathscr{H}}}^2=\sum_{l=1}^{M}\Big(\frac{g_{\mathfrak{s},\mathfrak{q};l}^2}{g_{\mathfrak{s},\mathfrak{q};l}^2+\theta_N}\Big)^2\sum_{j>M}\langle\boldsymbol{\beta}_{[\mathfrak{q}]}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}),\mathfrak{c}_j\rangle_{\mathscr{H}}^2=\mathrm{O}(M^{-2\beta}). \tag{5.18}
$$

<u>Term T_4</u>: Analogously as in Term T_4 in Theorem 3.2, we obtain

$$
\Big|\Big|\coprod_{\mathscr{J}_{\mathfrak{p},\mathfrak{q};M}}\boldsymbol{\beta}_{[\mathfrak{q}]}\Big[\prod_{[\mathfrak{s},\mathfrak{q}]}^{\dagger}\coprod_{\mathfrak{g}_{\mathfrak{p},\mathfrak{q};1}}^{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};M}}-\mathbb{I}_{\mathscr{H}^{\mathfrak{q}}}\Big]\Big|\Big|_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}},\mathscr{H}}}^2\leq\theta_N^2g_{\mathfrak{s},\mathfrak{q};M}^{-2}||\boldsymbol{\beta}_{[\mathfrak{q}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}},\mathscr{H}}}^2\asymp\theta_N^2g_{\mathfrak{s},\mathfrak{q};M}^{-2}. \tag{5.19}
$$

Replacing T_1-T_4 in (5.15) by (5.16)-(5.19) and considering $\theta_N=\mathrm{O}(K^{-\beta/2})$ and $g_{\mathfrak{s},\mathfrak{q};M}^{-1}M^{\beta}=\mathrm{O}(K^{\beta})$, yields

$$
||\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]}-\boldsymbol{\beta}_{[\mathfrak{q}]}||_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}},\mathscr{H}}}^2=\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-1}K^{-\beta})+\mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}^{-2}K^{-\beta})+\mathrm{O}(M^{-2\beta})+\theta_N^2g_{\mathfrak{s},\mathfrak{q};M}^{-2}+\mathrm{O}(M^{-2\beta})
$$

$$
=\mathrm{O}_{\mathbb{P}}(M^{-2\beta})
$$

with what (3.50) is shown for all j and (3.49) for $i=1,...,\mathfrak{p}$ follows from (3.43) for all i, (3.50) for all j and

$$
||\hat{\alpha}_i-\alpha_i||_{\mathcal{S}_{\mathscr{H}}}^2\precsim ||\hat{\pi}_i-\pi_i||_{\mathcal{S}_{\mathscr{H}}}^2+\sum_{j=1}^{(i-1)\wedge\mathfrak{q}}||\hat{\beta}_j||_{\mathcal{S}_{\mathscr{H}}}^2||\hat{\pi}_{i-j}-\pi_{i-j}||_{\mathcal{S}_{\mathscr{H}}}^2+||\hat{\beta}_j-\beta_j||_{\mathcal{S}_{\mathscr{H}}}^2||\pi_{i-j}||_{\mathcal{S}_{\mathscr{H}}}^2. \qquad\square
$$

Acknowledgements

References

[1] Andersen T. G., Davis R. A., Kreiss J.-P. & Mikosch T. V. (2009). *Handbook of Financial Time Series*. Springer, Berlin.

[2] Aue A., Horváth L. & Pellatt D. F. (2017). Functional generalized autoregressive conditional heteroskedasticity. *Journal of Time Series Analysis* **38**, 3-21.

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

[4] Bollerslev T. (1986). Generaliziled Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* **31**, 307-327.

[5] Bosq D. (2000), *Linear Processes in Function Spaces*. Springer, New York.

[6] Cerovecki C., Francq C., Hörmann S. & Zakoïan J.-M. (2019). Functional GARCH Models: The Quasi-likelihood Approach and Its Applications. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2019.01.006.

[7] Engle R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K.inflation. *Econometrica* **50**(4), 987-1008.

[8] Ferraty F. & Vieu, P. (2006). *Nonparametric Functional Data Analysis*. Springer, New York.

[9]   Francq C. & Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications.* John Wiley & Sons Ltd.

[10] Gouriéroux C. (1997). *ARCH Models and Financial Applications.* Springer, New York.

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

[12] Hörmann S., Horváth L. & Reeder R. (2013). A Functional Version of the ARCH Model. *Econom. Theory **29**(2), 267-288.*

[13] Hörmann S. & Kokoszka P. (2010). Weakly dependent functional data. *The Annals of Statistics **38**, 1845-1884.*

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

[15] Kaballo W. (2014). *Aufbaukurs Funktionalanalysis und Operatortheorie.* (German) [Advanced course in functional analysis and operator theory] *Distributionen—lokalkonvexe Methoden—Spektraltheorie* [Distributions, locally convex methods, spectral theory]. Springer Spektrum, Berlin.

[16] Kingman J.F.C. (1974). Subadditive Ergodic Theory. *Annals of Probability **6**, 883-909.*

[17] Klepsch J., Klüppelberg C. & Wei T. (2017). Prediction of functional ARMA processes with an application to traffic data. *Econometris and Statistics **1**, 128-149.*

[18] Klepsch J. & Klüppelberg C. (2017). An Innovations Algorithm for the prediction of functional linear processes. *Journal of Multivariate Analysis **155**, 252-271.*

[19] Kokoszka P. & Reimherr M. (2013). Determing the order of the functional autoregressive model. *Journal of Time Series Analysis **34**, 116-129.*

[20] Kokoszka P., Rice G. & Shang H.L. (2017). Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *Journal of Multivariate Analysis **162**, 32-50.*

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

[22] Ledoux M. & Talagrand M. (1991). *Probability in Banach Spaces.* Springer, Berlin.

[23] Ramsay J.O. & Silverman B.W. (1997). *Functional Data Analysis.* Springer, New York.

[24] Rice G., Wirjanto, T. & Zhao Y. (2019). Tests for conditional heteroscedasticity with functional data and goodness-of-fit tests for FGARCH models. https://mpra.ub.uni-muenchen.de/93048/

[25] Ruiz-Medina M.D. & Álvarez-Liébana J. (2019). Strongly consistent autoregressive predictors in abstract Banach spaces. *Journal of Multivariate Analysis **170**, 186-201. Analysis **121**, 127-138.*

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

[27] Turbillon C., Marion J.-M. & Pumo B. (2007). Estimation of the moving-average operator in a Hilbert space. *Recent advances in stochastic modeling and data analysis. World Sci. Publ., Hackensack, NJ, 597-604.*