Functional ARCH and GARCH Models: A Yule-Walker Approach

Sebastian Kühnert*

September 21, 2020

Abstract

Conditional heteroskedastic financial time series are commonly modelled by (G)ARCH processes. ARCH(1) and GARCH were recently established in $C[0,1]$ and $L^2[0,1]$. This article provides sufficient conditions for the existence of strictly stationary solutions, weak dependence and finite moments of (G)ARCH processes for any order in $C[0,1]$ and $L^p[0,1]$. It deduces explicit asymptotic upper bounds of estimation errors for the shift term, the complete (G)ARCH operators and the projections of ARCH operators on finite-dimensional subspaces. The operator estimaton is based on Yule-Walker equations, and estimating the GARCH operators also involves a result estimating operators in invertible linear processes being valid beyond the scope of (G)ARCH. Moreover, our results regarding (G)ARCH can be transferred to functional AR(MA).

MSC 2010 subject classifications: 47B38, 60G10, 62F12

Keywords: ARCH; ARMA; functional data; functional principal components; functional time series; GARCH; invertible linear processes; parameter estimation; stationary solutions; Yule-Walker equation

1 Introduction

Volatility, usually measured by the variance, is one of the essential objects of study of financial time series which are often strictly stationary and *conditional heteroskedastic*. Latter means that the variances at any time conditioned on the past are non-constant and randomly changing. A popular model exhibiting this phenomenon for a real-valued process $(X_k)_{k \in \mathbb{Z}}$ was established by Engle [9] in 1982, namely the *autoregressive conditional heteroskedasticity* (*ARCH*) model

$$X_k = \varepsilon_k \sigma_k, \qquad \sigma_k^2 = \delta + \sum_{i=1}^{\mathfrak{p}} \alpha_i X_{k-i}^2 \tag{1.1}$$

for which he was awarded the noble prize in economics in 2003. This model was extended by Bollerslev [4] in 1986 to the *generalized ARCH* (*GARCH*) model

$$X_k = \varepsilon_k \sigma_k, \qquad \sigma_k^2 = \delta + \sum_{i=1}^{\mathfrak{p}} \alpha_i X_{k-i}^2 + \sum_{j=1}^{\mathfrak{q}} \beta_j \sigma_{k-j}^2. \tag{1.2}$$

Various authors established modifications of uni- and multivariate (G)ARCH models, studied their probabilistic properties and estimated the parameters, see [1], [11], [12]. Along with 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, which is unproblematic from a mathematical point of view for complete, separablemetric spaces, see [24]. Such an extension enables the analysis to be more accurate. It also allows applying our recurrence equations on continuous processes by decomposing these in a natural way, see Figure 1, and to model random events in time for which modelling by real-valued

---

*EMAIL: s.kuehnert_math@gmx.de

multivariate processes with discrete time failed due to missing data where one is incapable of assigning a constant vector length throughout time. For a detailed introduction in *Functional Data Analysis* (FDA) and *Functional Time Series Analysis* (FTSA), the areas dealing with random variables resp. time series with values in an infinite-dimensional space, see [5], [10], [13], [16], [27], and for a compact synopsis (in German), see [23]. Outstanding overviews of Functional Analysis including operator theory, on which FDA and FTSA are built, are provided in [8], and [35] (in German).
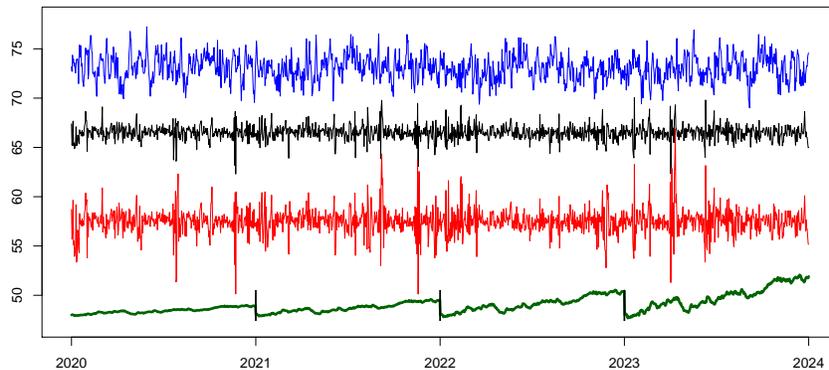

Figure 1: Fictitious processes in the period 2020-2024, e.g. measured in *USD*, based on the same pseudo random numbers. Processes of daily measurements with following errors: $AR(1)$ with parameter $a_1 = 0.6$ (**blue**), which have constant variances conditioned on the past; ARCH(1) where $\delta = 0.2, \alpha_1 = 0.6$ (**black**); GARCH(1, 1) where $\delta = 0.2, \alpha_1 = 0.6, \beta_1 = 0.3$ (**red**). Further, a continuous stochastic process (**green**) which also can be interpreted as a functional time series consisting of four consecutive random functions, with step width $1/1000$.

    In 2013, Hörmann *et al.* [14] introduced ARCH(1) processes with values in $C[0, 1]$ and $L^2[0, 1]$, the spaces of continuous resp. measurable, square-Lebesgue integrable real valued functions with domain $[0, 1]$. GARCH(1, 1) was carried out on $C[0, 1]$ and $L^2[0, 1]$ by Aue *et al.* [2] afterwards, and GARCH for any order on $L^2[0, 1]$ by Cerovecki *et al.* [6] recently. All authors discussed the existence of strictly stationary solutions and probabilistic features, consistently estimated the parameters in $L^2[0, 1]$ and outlined applications. [14] estimated the shift term and the ARCH(1) operator projected on a finite-dimensional subspace by estimating its integral kernel. [2] used a least squares estimator for the projections of the GARCH(1, 1) parameters on a finite-dimensional subspace, and [6] a quasi-maximum likelihood approach for the projections of the GARCH($\mathfrak{p}, \mathfrak{q}$) parameters on a finite-dimensional subspace for all orders $\mathfrak{p}, \mathfrak{q} \in \mathbb{N}$ and for the complete parameters for $\mathfrak{p} = \mathfrak{q} = 1$. For further work regarding functional time series under conditional heteroskedasticity, we refer to [22] and [29].

    This article develops ARCH and GARCH processes for any orders with values in the function spaces $L^p[0, 1]$ with $p \in [1, \infty)$ and $C[0, 1]$ in particular. It provides sufficient conditions for the existence of strictly stationary solutions, weak dependence and moments of these processes under mild conditions, and also discusses parameter estimation. The focus of the article is on deducing estimators and explicit asymptotic upper bounds of the estimation errors for the shift term and the complete operators of $L^2[0, 1]$-valued ARCH and GARCH processes for any order. Explicit asymptotic upper bounds of the estimation errors are also derived for ARCH operators projected on finite-dimensional subspaces for any order. The operator estimation is based on Yule-Walker equations and utilizes upper bounds of eigenvalues and eigenfunctions, and the upper bounds of the estimation errors for our GARCH operators depend on upper bounds of estimation errors for operators in invertible linear processes. We also simulate possible realizations of our processes and some of our estimators.

    In this paper, we use following notation. $a \wedge b := \min(a, b)$ and $a \vee b := \max(a, b)$ for $a, b \in \mathbb{R}$. $\lfloor \cdot \rfloor$ denotes the *floor function*, $\mathrm{sgn}(\cdot)$ the *sign function*, δ_{ij} the *Kronecker delta* with $i, j \in \mathbb{Z}$, and $\mathbf{1}_A$ the *indicator function* of a subset $A \subseteq \mathbb{R}$. 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, $\Xi(a_n, b_n) := \omega(a_n) \cap o(b_n)$, $\Xi[a_n, b_n) := \Omega(a_n) \cap o(b_n)$ and $\Xi[a_n, b_n] := \Omega(a_n) \cap O(b_n)$. Let V be a vector space. Then, $V^n := \{(v_1, ..., v_n)^T | v_1, ..., v_n \in V\}$, with $n \in \mathbb{N}$, becomes a vector space by our componentwise definition of scalar multiplication and vector addition. Further, 0_V stands for the *identity element of addition*, and $\mathbb{I}_V$ for the *identity operator* from V to V, where *operator* always means linear mapping. For a space F of functions $f: [0,1] \to \mathbb{R}$, $F_{>0}$ 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 $[0,1]$. $f \odot g$ denotes the *pointwise product* of $f, g \in F$, with $f^2 := f \odot f$, if it is well-defined. Essential spaces F of functions $f: [0,1] \to \mathbb{R}$ in this work are the *space of continuous functions* $C[0,1]$ endowed with the supremum norm $||f||_\infty := \sup_{t\in[0,1]} |f(t)|$ for $f \in C[0,1]$, the *space of bounded functions* $\ell^\infty[0,1]$ endowed with $||\cdot||_\infty$, and the *spaces of (classes of) measurable functions with absolutely Lebesgue integrable p-th power* $L^p[0,1]$ for $p \in [1,\infty)$ endowed with the norm $||f||_{L^p[0,1]} := (\int_0^1 |f(t)|^p \, dt)^{1/p}$ for $f \in L^p[0,1]$, with integration w.r.t. λ. These are separable Banach spaces w.r.t. their norms, and $\mathscr{H} := L^2[0,1]$ endowed with the inner product $\langle f,g \rangle_{\mathscr{H}} := \int_0^1 f(t)g(t) \, dt$ for $f, g \in \mathscr{H}$, with integration w.r.t. λ, is a separable Hilbert space. Hereafter, let $(\mathcal{B}, ||\cdot||_{\mathcal{B}}), (\mathcal{B}', ||\cdot||_{\mathcal{B}'})$ be Banach 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 inner product, and abbreviate 'complete orthonormal system' by *CONS*. We endow Banach spaces $(\mathcal{B}^n, ||\cdot||_{\mathcal{B}^n})$ with the norm $||\boldsymbol{b}||_{\mathcal{B}^n}^2 := \sum_{i=1}^n ||b_i||_{\mathcal{B}}^2$ for $\boldsymbol{b} := (b_1, ..., b_n)^T \in \mathcal{B}^n$ and Hilbert spaces $(\mathcal{H}^n, \langle\cdot,\cdot\rangle_{\mathcal{H}^n})$ with the inner product $\langle \boldsymbol{h}, \tilde{\boldsymbol{h}} \rangle_{\mathcal{H}^n} := \sum_{i=1}^n \langle h_i, \tilde{h}_i \rangle_{\mathcal{H}}$ for $\boldsymbol{h} := (h_1, ..., h_n)^T, \tilde{\boldsymbol{h}} := (\tilde{h}_1, ..., \tilde{h}_n)^T \in \mathcal{H}^n$. Also, we define the following spaces of operators. $\mathcal{L}_{\mathcal{B},\mathcal{B}'}$ is the *space of bounded operators from $\mathcal{B}$ to $\mathcal{B}'$*, with $\mathcal{L}_{\mathcal{B}} = \mathcal{L}_{\mathcal{B},\mathcal{B}}$. Thereby, A^* denotes the *adjoint of* $A \in \mathcal{L}_{\mathcal{B},\mathcal{B}'}$. Further, $\mathcal{S}_{\mathcal{H},\mathcal{H}'}$ and $\mathcal{N}_{\mathcal{H},\mathcal{H}'}$ are the *spaces of Hilbert-Schmidt* and *nuclear operators* from $\mathcal{H}$ to $\mathcal{H}'$, respectively, with $\mathcal{S}_{\mathcal{H}} = \mathcal{S}_{\mathcal{H},\mathcal{H}}$ and $\mathcal{N}_{\mathcal{H}} = \mathcal{N}_{\mathcal{H},\mathcal{H}}$. The meaning of A being an Hilbert-Schmidt and nuclear operator is that its singular values are square and absolutely summable, respectively. Moreover, for the norms of the mentioned spaces holds $||\cdot||_{\mathcal{L}_{\mathcal{H},\mathcal{H}'}} \leq ||\cdot||_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}} \leq ||\cdot||_{\mathcal{N}_{\mathcal{H},\mathcal{H}'}}$. Also, we define the operator $h \otimes h' := \langle h, \cdot \rangle_{\mathcal{H}} h'$ 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 $p \in [1,\infty)$, $L_{\mathcal{B}}^p = L_{\mathcal{B}}^p(\Omega, \mathfrak{A}, \mathbb{P})$ denotes *the space of (classes of) $\mathcal{B}$-valued random variables X with* $\nu_{p,\mathcal{B}}(X) := (\mathbb{E}||X||_{\mathcal{B}}^p)^{1/p} < \infty$, we call a process $(X_k)_{k\in\mathbb{Z}}$ of $\mathcal{B}$-valued random variables $L_{\mathcal{B}}^p$-*process* if $X_k \in L_{\mathcal{B}}^p$ holds for all k, and *centered* if $\mathbb{E}(X_k) = 0_{\mathcal{B}}$ holds for all k with expectation in Bochner-integral sense, see [16], p. 40–45.

This article is organized as follows. Section 2 studies probabilistic features of our (G)ARCH processes. Section 3 introduces parameter estimators and derives upper bounds. Section 4 conducts a simulation study. Section 5 summarizes the main results and gives an outline for future research. Section 6 contains proofs.

2 Functional ARCH and GARCH Models

Here, probabilistic features of F-valued ARCH and GARCH processes for any order are derived. F stands either for $L^p[0,1]$ with $p \in [1,\infty)$, or for a separable Banach space of functions $f: [0,1] \to \mathbb{R}$ endowed with the supremum norm $||\cdot||_\infty$ and being closed w.r.t. the pointwise product $\odot$. Latter is especially satisfied for $C[0,1]$, but not for $L^\infty[0,1]$ which is closed w.r.t. $\odot$ but inseparable. We deduce sufficient conditions for the existence of strictly stationary solutions, finite moments and weak dependence. But firstly, we formally define our processes.

Definition 2.1. *Let $\mathfrak{p} \in \mathbb{N}, \mathfrak{q} \in \mathbb{N}_0$. Further, let $(\varepsilon_k)_{k\in\mathbb{Z}}$ be an i.i.d. F-valued process, $\delta \in F_{>0}$ and $\alpha_i, \beta_j \in \mathcal{L}_F$ with $\alpha_i, \beta_j: F_{\geq 0} \to F_{\geq 0}$ 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, the F-valued process $(\mathscr{X}_k)_k$ is an ARCH($\mathfrak{p}$) process if $\mathfrak{q} = 0$ and $\alpha_{\mathfrak{p}} \neq 0_{\mathcal{L}_{\mathscr{H}}}$, and a GARCH($\mathfrak{p}, \mathfrak{q}$) process if $\mathfrak{q} \in \mathbb{N}$ and $\alpha_{\mathfrak{p}} \neq 0_{\mathcal{L}_{\mathscr{H}}}, \beta_{\mathfrak{q}} \neq 0_{\mathcal{L}_{\mathscr{H}}}$.

Throughout, $\alpha_i, \beta_j, \delta, \varepsilon_k, \mathfrak{p}, \mathfrak{q}, \sigma_k, \mathscr{X}_k$ are the variables in the equations (2.1), with $\alpha_i = \beta_j = 0_{\mathcal{L}_{\mathscr{H}}}$ for $i > \mathfrak{p}, j > \mathfrak{q}$, and we write $\mathfrak{r} := \mathfrak{p} \vee \mathfrak{q}, \mathfrak{s} := \mathfrak{p} + \mathfrak{q}$. Defining our ARCH model through (2.1) with $\mathfrak{q} = 0$ allows us to derive assertions for our ARCH and GARCH processes simultaneously. (2.1) differs to (1.1) and (1.2) for real-valued ARCH and GARCH, respectively, as follows. First, ε_k and σ_k in the left equation in (2.1) are assembled not through a common but a pointwise product, being well-defined for our spaces F endowed with the norm $|| \cdot ||_\infty$, and for $L^p[0,1]$ spaces additional assumptions are needed. Second, δ is a vector and α_i, β_j are operators, instead of being scalars in the real-valued case. For the operators when mapping from $L^p[0,1]$, further conditions are needed to ensure well-definedness. Third, the conditional variance σ_k^2 of $\mathscr{X}_k$ given the past has to be interpreted as a function rather than a non-negative number as usual.

2.1 Strictly Stationary Solutions

The identity (2.1) implies the *state-space form*

$$\varsigma_k^{(\mathfrak{p},\mathfrak{q})} = \delta_k^{(\mathfrak{p},\mathfrak{q})} + \Psi_k^{(\mathfrak{p},\mathfrak{q})}(\varsigma_{k-1}^{(\mathfrak{p},\mathfrak{q})}) \quad :\Longleftrightarrow \tag{2.2}$$

$$
\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}
$$

Thereby, $\boxdot_k$ in the matrix above are operators mapping from $\dot{F}$ to F, with

$$\boxdot_k(f) := f \odot \varepsilon_k^2, \quad f \in \dot{F},$$

where $\dot{F}$ is also a function space defined by

$$\dot{F} := \begin{cases} L^{2p}, & \text{if } F = L^p, \\ F, & \text{if } F \text{ is a Banach space w.r.t. } || \cdot ||_\infty. \end{cases} \tag{2.3}$$

Using the modification $\dot{F}$ of the initial space F in the definition of $\boxdot_k$ ensures well-definedness, at least if $\varepsilon_k^2 \in \dot{F}$ a.s. holds. If the i.i.d. errors ε_k satisfy

$$\mathbb{E} \ln(1 \vee ||\varepsilon_0^2||_{\dot{F}}) < \infty, \tag{2.4}$$

then in particular $||\varepsilon_0^2||_{\dot{F}} < \infty$ a.s. Thus indeed $\varepsilon_k^2 \in \dot{F}$ a.s. for all k which implies that $\boxdot_k$ is a well-defined operator a.s., and also bounded with

$$||\boxdot_k||_{\mathcal{L}_{\dot{F},F}} \leq ||\varepsilon_k^2||_{\dot{F}} \quad \text{a.s.} \tag{2.5}$$

Hence, $\varsigma_k^{(\mathfrak{p},\mathfrak{q})}, \delta_k^{(\mathfrak{p},\mathfrak{q})} \in F^{\mathfrak{s}}$ a.s., and $\Psi_k^{(\mathfrak{p},\mathfrak{q})} \in \mathcal{L}_{F^{\mathfrak{s}}}$ a.s. if $\delta, \varepsilon_k^2, \sigma_{k-\mathfrak{q}}^2, ..., \sigma_k^2 \in \dot{F}$ a.s. Moreover, (2.4) and $|| \cdot ||_{\mathcal{L}_F} \leq || \cdot ||_{\mathcal{L}_{F,\dot{F}}}$ imply (see [23], p. 28)

$$\mathbb{E} \ln(1 \vee ||\Psi_0^{(\mathfrak{p},\mathfrak{q})}||_{\mathcal{L}_{F^{\mathfrak{s}}}}) < \infty. \tag{2.6}$$

Since $(\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})})_k$ is i.i.d. and $||\cdot||_{\mathcal{L}_{F^s}}$ is sub-multiplicative, [18], Theorem 6 yields

$$
\begin{aligned}
\gamma^{(\mathfrak{p},\mathfrak{q})} &:= \lim_{k\to\infty} \frac{1}{k}\, \mathbb{E}\ln\big|\big|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_1^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^s}} \\
&= \lim_{k\to\infty} \frac{1}{k}\, \ln\big|\big|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_1^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^s}} \quad \text{a.s.,}
\end{aligned}
\tag{2.7}
$$

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

Remark 2.1.   (a) The state-space form (2.2) rests upon the version in [11] used for real-valued GARCH$(\mathfrak{p},\mathfrak{q})$ (and ARCH$(\mathfrak{p})$) processes.

   (b) Aside from the fact that the condition (2.4) is applicable for a large number of spaces, it is due to $||\cdot||_{\dot{F}} \le ||\cdot||_{\infty}$ for any $\dot{F}$ in (2.3) also milder than the condition $\mathbb{E}\ln(1 \vee ||\varepsilon_0^2||_{\infty}) < \infty$ in [6] where assertions for $L^2[0,1]$-valued GARCH$(\mathfrak{p},\mathfrak{q})$ processes were derived.

We are now in the position to state sufficient conditions for the existence of non-anticipative, strictly stationary solutions of our ARCH$(\mathfrak{p})$ and GARCH$(\mathfrak{p},\mathfrak{q})$ processes. Thereby, a process $(Y_k)_{k\in\mathbb{Z}}$ is *non-anticipative* w.r.t. some F-valued process $(\varepsilon_k)_{k\in\mathbb{Z}}$ if there exists a measurable function $f\colon F^\infty \to F$ so 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., (2.8) implies that $(Y_k)_k$ is also strictly stationary and ergodic, see [32], Theorem 3.5.8.

Theorem 2.1. *Let the requirements in Definition 2.1 and (2.4), $\delta \in \dot{F}_{>0}$ and $\alpha_i, \beta_j \in \mathcal{L}_{F,\dot{F}}$ for all i,j hold.*

   (a) *If*

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

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

   (b) *The equation (2.9) is satisfied if for some $n \in \mathbb{N}$ and $\nu > 0$ holds*

$$
\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})} := \mathbb{E}\big|\big|\boldsymbol{\Psi}_n^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{n-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_1^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^s}}^{\nu} < 1.
\tag{2.10}
$$

Remark 2.2.   (a) Theorem 2.1 generalizes Theorems 2.1, 2.3 in [14], 2.1-2.2 in [2] and 1 in [6]. That is because it provides sufficient conditions for strictly stationary solutions ARCH and GARCH processes for any order with values in $F = L^p[0,1]$ with $p \in [1,\infty)$ and in our Banach spaces F endowed with $||\cdot||_\infty$, closed w.r.t. $\odot$ (this includes $C[0,1]$), whereas [2], [14] discussed ARCH(1) resp. GARCH(1,1) in $C[0,1]$ and $L^2[0,1]$, and [6] discussed GARCH for any order in $L^2[0,1]$. Further, we imposed bounded instead of integral operators which is more restricting, and the condition (2.4), on which the definition of $\gamma^{\mathfrak{p},\mathfrak{q}}$ is built, is milder than that of [6], see Remark 2.1 (b), thus (2.9) is more general.

   (b) The condition (2.10), which is new for $\mathfrak{p} \vee \mathfrak{q} > 1$ as far as we know, is stricter but easier to validate than (2.9), see Section 4.

   (c) In the scalar case, a condition as (2.9) is also necessary for the existence of strictly stationary solutions, see [11]. However, [6], Remark 1, which also applies here, notes that proving (2.9) to be a necessary condition is difficult, since norms in infinite-dimensional Banach spaces are not equivalent.

   (d) [6], Proposition 1 also applies here. Thus, under (2.4) and (2.9), a sufficient condition for the existence of a pointwise second-order stationary solution of (2.1) is, that the *spectral radius* satisfies $\rho(\mathbb{E}(\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})})) < 1$, with random matrix $\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})}$ in (2.2) and expectation in Bochner-integral sense.

   (e) For detailed examples of parameters and innovations for which both the initial requirements of Theorem 2.1 and (2.10) hold, see Section 4.

2.2   Finite Moments and Weak Dependence

Based on ideas in [2], [14], and with (2.10), we now deduce sufficient conditions for the existence of finite moments and weak dependence in sense of L^p-m-approximibility for our F-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p}, \mathfrak{q}$) processes. Thereby, a process $(Y_k)_{k \in \mathbb{Z}}$ is L_F^p-m-*approximable* for $p \geq 1$ if $Y_k = f(\varepsilon_k, \varepsilon_{k-1}, ...)$ a.s. holds for all k, for some i.i.d. F-valued process $(\varepsilon_k)_{k \in \mathbb{Z}}$ and a measurable function $f \colon F^\infty \to F$, and if

$$\sum_{m=1}^{\infty} \nu_{p,F}(Y_m - Y_m^{(m)}) < \infty, \tag{2.11}$$

with $Y_k^{(m)} := f(\varepsilon_k, \varepsilon_{k-1}, ..., \varepsilon_{k-m+1}, \varepsilon_{k-m}^{(k)}, \varepsilon_{k-m-1}^{(k)}, ...)$ and $\nu_{p,F}(\cdot) = (\mathbb{E} || \cdot ||_F^p)^{1/p}$, is satisfied, where $(\varepsilon_k^{(n)})_{k \in \mathbb{Z}}$ are independent copies of $(\varepsilon_k)_{k \in \mathbb{Z}}$ for all n. Further, we call $(Y_k)_{k \in \mathbb{Z}}$ *geometrically* L_F^p-m-*approximable* if $(Y_k)_k$ is L_F^p-m-approximable with $\nu_{p,F}(Y_m - Y_m^{(m)}) = \mathrm{O}(\rho^m)$ for some $\rho \in (0,1)$. For each m, $(Y_k^{(m)})_{k \in \mathbb{Z}}$ are strictly stationary, m-dependent processes, and $Y_k^{(m)}$ equals Y_k in distribution. L_F^p-m-approximibility of a process thus means that it is non-anticipative w.r.t. another process and can be approximated by some m-dependent process so that the approximation errors measured by the L_F^p-norm $\nu_{p,F}(\cdot)$ are summable. For a detailed introduction to L^p-m-approximibility, see [15].

Lemma 2.1. *For some $n \in \mathbb{N}, \nu > 0$, let $\mathbb{E} ||\varepsilon_0^2||_F^\nu < \infty$ and (2.10) hold. 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 geom.* L_F^ν-m- *and* $(\sigma_k^2)_k$ *is geom.* L_F^ν-m-*approximable.*

Remark 2.3.   (a) As far as we are aware, both parts of Lemma 2.1 are new for functional ARCH($\mathfrak{p}$) with $\mathfrak{p} > 1$ and part (b) for functional GARCH($\mathfrak{p}, \mathfrak{q}$) processes with $\mathfrak{p} \vee \mathfrak{q} > 1$. Lemma 2.1 also extends the theories regarding finite ν-th moments with $\nu > 0$ and weak dependence in [14] and [2] where only $C[0,1]$- and $\mathscr{H}$-valued ARCH(1) resp. GARCH(1,1) processes with $\mathscr{H} = L^2[0,1]$ were discussed, and it extends [6], Proposition 2.

(b) [6], Proposition 2 gives a statement for finite moments without using a moment condition as (2.10) for $\mathscr{H}$-valued GARCH processes for any order. They concluded that $\mathbb{E} ||\varepsilon_0^2||_\infty^\tau < \infty$ for some $\tau \in (0,1)$, (2.9) and $(\mathscr{X}_k)_k$ being a strictly stationary solution of (2.1) implies that there exists $s \in (0, \tau)$ such that $\mathbb{E} ||\mathscr{X}_0^2||_\mathscr{H}^s < \infty$ and $\mathbb{E} ||\sigma_0^2||_\mathscr{H}^s < \infty$. By following the lines of their proof, it becomes clear that this proposition also applies to any of our F-valued (G)ARCH processes.

(c) The proofs of Theorem 2.1 and Lemma 2.1 neither rely on the fact that our spaces consist of real-valued functions with domain $[0,1]$ nor on specific features of their norms. Hence, these results are likely to be carried out on further separable Banach and Hilbert spaces, provided the (G)ARCH equations (2.1) are well-defined and prerequisites hold. It might be useful to extend our processes, where $n \in \mathbb{N}$, on separable Banach spaces $C^n(D)$ of functions on a compact domain D for which the n-th derivative is continuous, $L^2[0,1]^n$, separable *Sobolev spaces* (see [8]), $\mathcal{B}^n, \mathcal{H}^n, \mathcal{L}_\mathcal{B}$ and $\mathcal{S}_\mathcal{H}$ where $\mathcal{B}$ and $\mathcal{H}$ denote separable Banach resp. Hilbert spaces, etc.

3   Estimation Procedure and Asymptotics

In this section, estimators for the parameters δ and α_i, β_j of $\mathscr{H}$-valued ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p}, \mathfrak{q}$) processes with known orders $\mathfrak{p}, \mathfrak{q} \in \mathbb{N}$ with $\mathscr{H} = L^2[0,1]$ are established and asymptotic upper bounds of the estimation errors are deduced. The reason for restricting the estimation on $L^2[0,1]$, the only separable Hilbert space for which our (G)ARCH processes are defined, is that then an inner product is at hand. The notion of an adjoint and of specific operators, e.g. covariance operators, then becomes easier to grasp and verifying required auxiliary results less cumbersome. Nevertheless, conducting our estimation procedure on separable

Banach spaces is conceivable, since on these spaces covariance operators etc. can be defined (see [5]) and operator estimation was already executed (see e.g. [30]). Hereinafter, we put $\dot{\mathscr{H}} := L^4[0,1]$, and for assertions regarding (G)ARCH we throughout 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).

(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 $\alpha_i = \beta_j = 0_{\mathcal{L}_{\mathscr{H}}}$ 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, $\mathscr{Z} = (\mathscr{Z}_k)_k$ is an $\mathscr{H}$-valued AR($\mathfrak{p}$) resp. ARMA($\mathfrak{r}, \mathfrak{q}$) process if $\mathfrak{q} = 0$ resp. $\mathfrak{q} \in \mathbb{N}$, with stationary but not i.i.d. innovation process $\boldsymbol{\nu} := (\nu_k)_k$. Both $\mathscr{Z}$ and $\boldsymbol{\nu}$ are centered, stationary, non-anticipative w.r.t. $(\varepsilon_k)_k$ and geometrically $L_{\mathscr{H}}^4$-m-approximable. For general functional AR(MA) processes and their applications, we refer to [5], [19], [31].

When estimating the operators in our ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p}, \mathfrak{q}$) model, we use the related AR($\mathfrak{p}$) resp. ARMA($\mathfrak{r}, \mathfrak{q}$) equation (3.2), their associated Yule-Walker equation which we establish later, and also 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 $i, j, ||\Gamma_{\mathfrak{p}, \mathfrak{q}}||_{\mathcal{L}_{\mathscr{H}}} < 1$ 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 $U \subsetneq \mathscr{H}$ such that $\mathbb{P}(\varepsilon_0^2 \in U) = 1$.

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

3.1 Preliminaries

For our estimation procedure of the (G)ARCH parameters in (2.1), we need further assumptions and asymptotic upper bounds of estimation errors for specific moments, operators, eigenvalues and eigenfunctions, all related to processes with values in separable Hilbert spaces. These are stated for general, separable Hilbert spaces, hereafter denoted by $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}}), (\mathcal{H}', \langle \cdot, \cdot \rangle_{\mathcal{H}'})$ and $(\mathcal{H}'', \langle \cdot, \cdot \rangle_{\mathcal{H}''})$. Also, we discuss estimating operators within a composition of operators in a general manner which can be applied to our Yule-Walker equations from which estimators for the ARCH and GARCH operators are derived.

We start with estimating moments.

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

$$\hat{m}_l = \hat{m}_l(\boldsymbol{X}) := N^{-1} \sum_{i=1}^{N} X_i^l \tag{3.3}$$

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

$$\mathbb{E}||\hat{m}_l - m_l||_{\mathcal{H}}^2 = \mathrm{O}(N^{-1}). \tag{3.4}$$

The Yule-Walker equations for estimating our (G)ARCH operators contain *lag-h-covariance operators* of processes and some modifications, known as *lag-h-cross-covariance operators* of two processes.

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

$$\mathscr{C}_h = \mathscr{C}_{h;\boldsymbol{X}} := \mathbb{E}[(X_0 - m_1) \otimes (X_h - m_1)], \tag{3.5}$$

with $m_1 = m_1(\boldsymbol{X}) \coloneqq \mathbb{E}(X_1)$, *and its empirical version by*

$$\hat{\mathscr{C}}_h = \hat{\mathscr{C}}_{h;\boldsymbol{X}} \coloneqq \begin{cases} \frac{1}{N_h-1}\sum_{k=|h|+1}^{N}(X_k - \hat{m}_1) \otimes (X_{k+h} - \hat{m}_1), & 1 - N < h < 0, \\ \frac{1}{N_h-1}\sum_{k=1}^{N_h}(X_k - \hat{m}_1) \otimes (X_{k+h} - \hat{m}_1), & 0 \le h < N - 1 \end{cases} \tag{3.6}$$

where $\hat{m}_1 = \hat{m}_1(\boldsymbol{X}) \coloneqq N_h^{-1}\sum_{i=1}^{N_h} X_i, N_h \coloneqq N - |h|$ *and* $N \in \mathbb{N}, |h| < N - 1$.

Lag-h-covariance operators $\mathscr{C}_h \colon \mathcal{H} \to \mathcal{H}$ and its empirical versions $\hat{\mathscr{C}}_h \colon \mathcal{H} \to \mathcal{H}$ are nuclear resp. bounded operators with finite-dimensional image, short $\mathscr{C}_h \in \mathcal{N}_\mathcal{H}$ and $\hat{\mathscr{C}}_h \in \mathcal{F}_\mathcal{H}$, with $\mathscr{C}_h^* = \mathscr{C}_{-h}$ and $\hat{\mathscr{C}}_h^* = \hat{\mathscr{C}}_{-h}$ for all h. $\mathscr{C}_0$, briefly called *covariance operator*, and its empirical version $\hat{\mathscr{C}}_0$ is also selfadjoint and positive semi-definite. Our lag-h-covariance operators can be approximated for fixed h and the absolute value of h tending to infinity not too fast, as follows.

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

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

This result enables deriving upper bounds for further operators which we use in our estimation procedure. Amongst others, for the operators

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

where $\boldsymbol{X}(d) \coloneqq (X_k(d))_{k\in\mathbb{Z}}$ with $d \in \mathbb{N}$ stands for an $\mathcal{H}^d$-valued processes with

$$X_k(d) \coloneqq (X_k, X_{k-1}..., X_{k-d+1})^T, \qquad k \in \mathbb{Z},$$

with entries of a centered process $\boldsymbol{X} = (X_k)_{k\in\mathbb{Z}}$. The idea to use such operators is based on [3]. Thereby, $\mathfrak{S}_{d,m} \in \mathcal{N}_{\mathcal{H}^d, \mathcal{H}}$ and $\mathfrak{S}_d \in \mathcal{N}_{\mathcal{H}^d}$ with (see [23], p. 56)

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

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

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

with $N_d \coloneqq N - d, \hat{m}_1(\boldsymbol{X}^2(d)) \coloneqq N_d^{-1}\sum_{i=1}^{N_d} X_{d+i-1}^2(d), \hat{m}_1(\boldsymbol{X}^2) \coloneqq N_d^{-1}\sum_{j=1}^{N_d} X_{d+j}^2$, holds due to (3.7) (see [23], Definition and properties 4.36):

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

Further, using the empirical version of the covariance operator $\mathfrak{S}_d$ defined by

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

leads to

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

Further probabilistic features and asymptotics of (lag-h-)covariance and cross-covariance operators can be found in [5] and [16], and [28] thoroughly studied the asymptotics of lag-h-cross-covariance operators.

Now, we derive asymptotic upper bounds of the estimation errors for the eigenvalues and eigenfunctions of a compact, self-adjoint, positive semi-definite operator $\mathscr{K} \in \mathcal{L}_\mathcal{H}$. We impose this operator to be estimated by a sequence $(\hat{\mathscr{K}}_N)_{N\in\mathbb{N}} \subseteq \mathcal{L}_\mathcal{H}$ of compact, self-adjoint, positive semi-definite operators, where each $\hat{\mathscr{K}}_N$ depends on N observations of a stationary process $\boldsymbol{X} = (X_k)_{k\in\mathbb{Z}}$. Thereby, $(\mathfrak{k}_j)_{j\in\mathbb{N}}$ resp. $(\hat{\mathfrak{k}}_j)_{j\in\mathbb{N}}$ represent the eigenfunction 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 this purpose, 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 [5], 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*

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

(3.15) is used in various conversions in proofs, and means the eigenvalues indexed by a constant or a sufficiently slowly decaying sequence are asymptotically equal to their empirical versions up to a multiplicative constant in some sense.

Eigenfunctions are unambiguously determined except for their sign. Hence, we estimate the eigenfunctions $\mathfrak{k}_j$ of $\hat{\mathscr{K}}_N$ if $\hat{\mathfrak{k}}_j \not\perp \mathfrak{k}_j$ a.s. holds by

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

where sgn is the signum function. According to [5], Lemma 4.3, which can be generalized to any compact, self-adjoint, positive semi-definite operator,

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

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

$$\gamma_j := (k_j - k_{j+1})^{-1}, \qquad j \in \mathbb{N} \tag{3.18}$$

are the reciprocal spectral gaps. The problem in using $\hat{\mathfrak{k}}_j'$ to estimate $\mathfrak{k}_j$ is, that $\hat{\mathfrak{k}}_j \not\perp \mathfrak{k}_j$ a.s. and thus $\mathrm{sgn}(\langle \hat{\mathfrak{k}}_j, \mathfrak{k}_j \rangle_\mathcal{H}) \neq 0$ a.s. is not guaranteed for all j, N, but needed to obtain our upper bounds of the estimation errors for the operators in the $\mathscr{H}$-valued (G)ARCH model. Therefore, we modify $\hat{\mathfrak{k}}_j$ 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., $\mathcal{N}(0,1)$-distributed random variables, independent of the observations of $\boldsymbol{X}$. Then

$$\hat{\mathfrak{k}}_j'' := \hat{\mathfrak{k}}_j + \sum_{i=1}^\infty \frac{\zeta_i h_i}{i^2 N} \tag{3.19}$$

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

$$\hat{\mathfrak{k}}_j''' := \mathrm{sgn}(\langle \hat{\mathfrak{k}}_j'', \mathfrak{k}_j \rangle_{\mathcal{H}'})\hat{\mathfrak{k}}_j. \tag{3.20}$$

Thereby, $(\hat{\mathfrak{k}}_j''')_j$ 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

$$k_1 > k_2 > \cdots > 0, \tag{3.21}$$

and for any sequence $m = m_N = \Omega(1)$, due to convexity of κ, holds

$$\sup_{j \leq m} \tilde{\gamma}_j = \gamma_m = \Omega(k_m^{-1}). \tag{3.22}$$

Thereby, whether the sequence γ_m is asymptotically equal to or increases faster than k_m^{-1} depends on the precise decay rate of the eigenvalues k_m. E.g. $k_m \asymp e^{-m}$ implies $\gamma_m \asymp e^m \asymp k_m^{-1}$, and $k_m \asymp m^{-2}$ yields $\gamma_m \asymp m^3$ but $k_m^{-1} \asymp m^2$.

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

$$||\hat{\mathfrak{k}}_j''' - \mathfrak{k}_j||_\mathcal{H}^2 = \mathrm{O}_\mathbb{P}(a_N), \quad j \in \mathbb{N}. \tag{3.23}$$

Furthermore, for sequences $m = m_N = \Omega(1)$ with $\gamma_m^2 a_N = \mathrm{o}(1)$ holds

$$\sup_{j \leq m} ||\hat{\mathfrak{k}}_j''' - \mathfrak{k}_j||_\mathcal{H}^2 = \mathrm{O}_\mathbb{P}(\gamma_m^2 a_N). \tag{3.24}$$

The last preparatory step concerns estimating operators within a composition of operators, to be precise, of bounded operators $B \in \mathcal{L}_{\mathcal{H}',\mathcal{H}''}$ in equations as

$$A = BC \tag{3.25}$$

with $A \in \mathcal{L}_{\mathcal{H},\mathcal{H}''}, C \in \mathcal{L}_{\mathcal{H},\mathcal{H}'}$. Identifiability of B in (3.25), meaning $BC = \tilde{B}C$ implies $B = \tilde{B}$, is only guaranteed if B has dense image. C is not necessarily invertible, and if it is invertible and compact, it has no bounded inverse which would be desirable for further conversions. However, a generalized inverse

$$C^\dagger := C^*(CC^* + b_n \mathbb{I}_\mathcal{H})^{-1} \tag{3.26}$$

of C, where $(b_n)_{n \in \mathbb{N}} \subset (0, \infty)$ is a null sequence, is invertible and bounded, and $C^\ddagger := CC^\dagger$ is in a sense close to $\mathbb{I}_{\mathcal{H}'}$. When estimating operators without projecting them on a finite-dimensional subspace, we also impose the following.

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, we say that $(S, (\phi_{ij})_{i,j})$ satisfies the *Sobolev condition* for $\beta > 0$ if

$$\sum_{i,j=1}^\infty \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 (1 + i^{2\beta} + j^{2\beta}) < \infty. \tag{3.27}$$

With $\mathscr{J}_m^c := \{\phi_{ij} | i, j \in \mathbb{N}, i \vee j > m\}$ and $m = m_N \to \infty$, (3.27) yields

$$\left|\left| \coprod_{\mathscr{J}_m^c} S \right|\right|_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 = \sum_{i \vee j > m} \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 \leq (1 + m^{2\beta})^{-1} \sum_{i,j=1}^\infty \langle S, \phi_{ij} \rangle_{\mathcal{S}_{\mathcal{H},\mathcal{H}'}}^2 (1 + i^{2\beta} + j^{2\beta})$$
$$= \mathrm{O}(m^{-2\beta}). \tag{3.28}$$

This identity is very benefical when estimating the completely observed operator, as can be seen in proofs of various subsequent assertions.

Remark 3.1.   (a) Generalized inverses are commonly used for estimating operators in FDA and FTSA. For instance, [26] utilized a generalized inverse as (3.26) for estimating parameters relating to their continuous random surfaces, and [25] regularized their covariance operator by finitely truncating the spectral decomposition of its inverse. For a comprehensive analysis regarding generalized inverses, see Tikhonov and Arsenin [33].

   (b) As far as we are aware, a Sobolov condition as (3.27) is new when estimating (G)ARCH operators, see [2], [6] and [13].

3.2   Estimation of δ in Functional ARCH and GARCH

We derive an estimator for δ of $\mathscr{H}$-valued ARCH and GARCH processes for any order from ideas in [14] for estimating δ in ARCH(1). Taking the expected value on both sides of the right equation in (2.1) leads under Assumption 3.1 to

$$\delta = m_2 - \sum_{i=1}^{\mathfrak{r}} (\alpha_i + \beta_i)(m_2), \tag{3.29}$$

with $\mathfrak{r} = \mathfrak{p} \vee \mathfrak{q}$ and $\alpha_i = \beta_j = 0_{\mathcal{L}_{\mathscr{H}}}$ for $i > \mathfrak{p}, j > \mathfrak{q}$. As an estimator for δ, with estimators $\hat{\alpha}_i, \hat{\beta}_j$ for α_i, β_j and $\hat{m}_2 := N^{-1} \sum_{i=1}^{N} \mathscr{X}_i^2$, we thus propose

$$\hat{\delta} := \hat{m}_2 - \sum_{i=1}^{\mathfrak{r}} (\hat{\alpha}_i + \hat{\beta}_i)(\hat{m}_2). \tag{3.30}$$

Theorem 3.1. *Let Assumption 3.1 hold. Then,*

$$||\hat{\delta} - \delta||_{\mathscr{H}} = O_{\mathbb{P}}(N^{-1/2}) + \sum_{i=1}^{\mathfrak{r}} O_{\mathbb{P}}||\hat{\alpha}_i - \alpha_i||_{\mathcal{L}_{\mathscr{H}}} + O_{\mathbb{P}}||\hat{\beta}_i - \beta_i||_{\mathcal{L}_{\mathscr{H}}}. \tag{3.31}$$

In Sections 3.3-3.4 emerges that the estimation errors for any operator decay at most as fast as $N^{-1/2}$. Hence, due to Theorem 3.1 and since $\mathfrak{r} = \mathfrak{p} \vee \mathfrak{q}$ is finite,

$$||\hat{\delta} - \delta||_{\mathscr{H}} = O_{\mathbb{P}}||\hat{\alpha}_k - \alpha_k||_{\mathcal{L}_{\mathscr{H}}} \qquad \text{resp.} \qquad ||\hat{\delta} - \delta||_{\mathscr{H}} = O_{\mathbb{P}}||\hat{\beta}_l - \beta_l||_{\mathcal{L}_{\mathscr{H}}}$$

if for all i, j holds $||\hat{\alpha}_k - \alpha_k||_{\mathcal{L}_{\mathscr{H}}} = O_{\mathbb{P}}||\hat{\alpha}_i - \alpha_i||_{\mathcal{L}_{\mathscr{H}}}$ as well as $||\hat{\alpha}_k - \alpha_k||_{\mathcal{L}_{\mathscr{H}}} = O_{\mathbb{P}}||\hat{\beta}_j - \beta_j||_{\mathcal{L}_{\mathscr{H}}}$ for some k, resp. $||\hat{\beta}_l - \beta_l||_{\mathcal{L}_{\mathscr{H}}} = O_{\mathbb{P}}||\hat{\alpha}_i - \alpha_i||_{\mathcal{L}_{\mathscr{H}}}$ and $||\hat{\beta}_l - \beta_l||_{\mathcal{L}_{\mathscr{H}}} = O_{\mathbb{P}}||\hat{\beta}_j - \beta_j||_{\mathcal{L}_{\mathscr{H}}}$ for some l.

3.3   Operator Estimation in Functional ARCH

In the following, $\mathscr{X} := (\mathscr{X}_k)_{k \in \mathbb{Z}}$ is an $\mathscr{H}$-valued ARCH($\mathfrak{p}$) process with $\mathfrak{p} \in \mathbb{N}$. Under Assumption 3.1, $\mathscr{L} := (\mathscr{L}_k)_{k \in \mathbb{Z}} = (\mathscr{X}_k^2 - m_2)_{k \in \mathbb{Z}}$ with $m_2 := \mathbb{E}(\mathscr{X}_1^2)$ is an $\mathscr{H}$-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. 7. Furthermore, $\mathscr{L}(\mathfrak{p}) := (\mathscr{L}_k(\mathfrak{p}))_{k \in \mathbb{Z}}$ satisfies

$$\mathscr{L}_k(\mathfrak{p}) = \tilde{\nu}_k(\mathfrak{p}) + \boldsymbol{A}_1(\mathscr{L}_{k-1}(\mathfrak{p})) :\Longleftrightarrow$$

$$
\begin{bmatrix} \mathscr{L}_k \\ \mathscr{L}_{k-1} \\ \vdots \\ \mathscr{L}_{k-\mathfrak{p}+2} \\ \mathscr{L}_{k-\mathfrak{p}+1} \end{bmatrix} = \begin{bmatrix} \nu_k \\ 0_{\mathscr{H}} \\ \vdots \\ 0_{\mathscr{H}} \\ 0_{\mathscr{H}} \end{bmatrix} + \begin{bmatrix} \alpha_1 & \cdots & \cdots & \cdots & \alpha_{\mathfrak{p}} \\ \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} \\ 0_{\mathcal{L}_{\mathscr{H}}} & \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} & \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} \end{bmatrix} \begin{bmatrix} \mathscr{L}_{k-1} \\ \mathscr{L}_{k-2} \\ \vdots \\ \mathscr{L}_{k-\mathfrak{p}+1} \\ \mathscr{L}_{k-\mathfrak{p}} \end{bmatrix}
$$

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

$$\mathfrak{S}_{\mathfrak{p},1} = \boldsymbol{\alpha}_{[\mathfrak{p}]} \mathfrak{S}_{\mathfrak{p}}, \tag{3.32}$$

with $\mathfrak{S}_{\mathfrak{p},1} = \mathscr{C}_{\mathscr{X}_{\mathfrak{p}}(\mathfrak{p}), \mathscr{X}_{\mathfrak{p}+1}}, \mathfrak{S}_{\mathfrak{p}} = \mathscr{C}_{0;\mathscr{L}(\mathfrak{p})}$. Since $\mathfrak{S}_{\mathfrak{p}}$ is injective as a consequence of Lemma 3.1 (see [23], Lemma 4.35), $\boldsymbol{\alpha}_{[\mathfrak{p}]}$ can be identified from (3.32). As an estimator we thus impose

$$\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} := \hat{\mathfrak{S}}_{\mathfrak{p},1} \hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}} = \hat{\mathfrak{S}}_{\mathfrak{p},1} \hat{\mathfrak{S}}_{\mathfrak{p}} (\hat{\mathfrak{S}}_{\mathfrak{p}}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1} \coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}. \tag{3.33}$$

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 $\coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}$ is the operator projecting on $\mathrm{lin}\{\hat{\mathfrak{c}}_{\mathfrak{p},1}, ..., \hat{\mathfrak{c}}_{\mathfrak{p},K}\} \subseteq \mathscr{H}^{\mathfrak{p}}$.

Theorem 3.2. *Let Assumptions 3.1-3.2 hold, and also Assumption 3.3 for the eigenvalue sequence $(c_{\mathfrak{p},i})_{i\in\mathbb{N}}$ of $\mathfrak{S}_{\mathfrak{p}}$. Let $(\Phi_{\mathfrak{p},ij})_{i,j\in\mathbb{N}}$ be the CONS of $\mathcal{S}_{\mathcal{H}^{\mathfrak{p}},\mathcal{H}}$ defined by $\Phi_{\mathfrak{p},ij} := \mathfrak{c}_{\mathfrak{p},i} \otimes \mathfrak{c}_j$ for all i,j where $(\mathfrak{c}_j)_{j\in\mathbb{N}} \subseteq \mathcal{H}$ denotes the eigenfunction sequences of the covariance operator $\mathscr{C}_{0;\mathfrak{x}}$, and let the sequence $(\vartheta_N)_N$ in (3.33) satisfy $\vartheta_N = \mathrm{O}(N^{-1/2})$.*

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

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|^2_{\mathcal{S}_{\mathcal{H}^{\mathfrak{p}},\mathcal{H}}} = \mathrm{O}_{\mathbb{P}}(N^{-1}). \tag{3.34}$$

(b) *Let $(\boldsymbol{\alpha}_{[\mathfrak{p}]}, (\Phi_{\mathfrak{p},ij})_{i,j})$ satisfy Assumption 3.4 for some $\beta > 0$, and let $K = K_N = \Xi(1,N)$ be a sequence with $c_{\mathfrak{p},K}^{-2} \gamma_{\mathfrak{p},K}^2 K^{2\beta+1} = \mathrm{O}(N)$, where $\gamma_{\mathfrak{p},K} := (c_{\mathfrak{p},K} - c_{\mathfrak{p},K+1})^{-1}$, and $\sum_{l=1}^K (\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2+\vartheta_N})^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j\rangle^2_{\mathcal{H}} = \mathrm{O}(K^{-2\beta})$. Then,*

$$\left\|\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}\right\|^2_{\mathcal{S}_{\mathcal{H}^{\mathfrak{p}},\mathcal{H}}} = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.35}$$

Remark 3.2.   (a) As far as we know, estimating ARCH operators by a Yule-Walker approach is new, as is the estimation of complete ARCH operators and of ARCH($\mathfrak{p}$) operators projected on a finite-dimensional subspace for $\mathfrak{p} > 1$ as in Theorem 3.2 (a). For $\mathfrak{p} = 1$, however, a similar version of Theorem 3.2 (a) was already stated in [14] where the ARCH(1) operator was imposed to be an integral operator and the finite expansion of its kernel was estimated with the same convergence rate.

(b) The assumption $\langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j\rangle_{\mathcal{H}} = 0$ for all j,l with $l \leq K < j$ in Theorem 3.2 (a), being milder than imposing $\boldsymbol{\alpha}_{[\mathfrak{p}]}$ and $\mathfrak{S}_{\mathfrak{p}}$ commute and similar to the condition in [34] for estimating their MA(1) operator, and the weaker version in (b) is necessary for technical conversions in the proof.

(c) Regarding the choice of K in Theorem 3.2 (a), [14] outlined that empirical results in [7] showed that $K = 2,3,4$ provide best results due to a bias variance trade-off, that it is practical to choose K as the largest integer so that for the empirical eigenvalues holds $\hat{c}_K/\hat{c}_1 \geq C$ for some threshold C, and that [36] proposed cross-validation to chose K. Since the threshold C is difficult to interpret, we propose to choose K as the first integer so that for some $w \in (0,1)$ holds $\sum_{i=1}^K \hat{c}_i / \sum_{i=1}^N \hat{c}_i \geq (1-w)$, provided a sample with sample size N is given. Then, w is interpretable as a maximal weight describing a relative proportion of information one allows to neglect.

(d) In Theorem 3.2 (b), due to (3.35), $K = K_N$ should be chosen so that it increases as fast as possible, provided all requirements are met, and the greater the variable β in the Sobolev condition, which describes the level of approximibility of an operator by a given CONS, the better.

(e) Since we avoided imposing a specific representation of the given ARCH operators, e.g. an integral operator representation, the assertion of Theorem 3.2 is independent of the particular structure of the given Hilbert space $\mathcal{H} = L^2[0,1]$. Thus, Theorem 3.2 can be extended to further separable Hilbert spaces, provided the ARCH processes are well-defined.

(f) Theorem 3.2 rests on the Yule-Walker equation (3.32) containing the operators $\mathfrak{S}_{\mathfrak{p},1}$ and $\mathfrak{S}_{\mathfrak{p}}$ being based on an AR derived from an ARCH process which is more restrictive than assuming an arbitrary AR process. Hence, Theorem 3.2 also holds for the operators of $\mathcal{H}$-valued AR processes.

The subsequent example illustrates some explicit asymptotic upper bounds (3.35) of the estimation errors for the complete ARCH operators and the choice of $K = K_N$, given the asymptotic behaviour of the eigenvalue sequence $(c_{\mathfrak{p},j})_j$ of the covariance operator $\mathfrak{S}_{\mathfrak{p}}$ in the Yule-Walker equation (3.32) is known.

Example 3.1. Let the assumptions of Theorem 3.2 hold.

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

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

(b) Impose $c_{\mathfrak{p},N} \asymp N^{-a}$ for some $a > 1$. Further, $K_N := \lfloor N^{\frac{1}{3+4a+2\beta}} \rfloor$ for all $N \in \mathbb{N}$. Then, $K = K_N = \Xi(1, N)$, $\gamma_{\mathfrak{p},K} \asymp K^{a+1}$, thus $c_{\mathfrak{p},K}^{-2} \gamma_{\mathfrak{p},K}^2 K^{2\beta+1} \asymp K^{3+4a+2\beta} \asymp N$. Hence, after Theorem 3.2:

$$||\hat{\boldsymbol{\alpha}}_{[\mathfrak{p}]} - \boldsymbol{\alpha}_{[\mathfrak{p}]}||^2_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}} = \mathrm{O}_{\mathbb{P}}\big(N^{-\frac{2\beta}{3+4a+2\beta}}\big).$$

3.4   Operator Estimation in Functional GARCH

Hereinafter, $\boldsymbol{\mathscr{X}} := (\mathscr{X}_k)_{k \in \mathbb{Z}}$ is an $\mathscr{H}$-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 ARMA$(\mathfrak{r}, \mathfrak{q})$ process, where $\mathfrak{r} = \mathfrak{p} \vee \mathfrak{q}$, with innovation process $\boldsymbol{\nu} := (\nu_k)_{k \in \mathbb{Z}} = (\mathscr{X}_k^2 - \sigma_k^2)_{k \in \mathbb{Z}}$, see p.7. We derive estimators for the operators in GARCH processes from those for operators of $\boldsymbol{\mathscr{Z}}$ represented as an inverted process. The operators of both processes, as in the estimation procedure for the complete ARCH operators, are estimated by generalized inverses based on certain Yule-Walker equations, provided that specific Sobolev conditions are satisfied. To this end, we impose the following.

Assumption 3.5. $\boldsymbol{\mathscr{Z}}$ is an invertible linear process w.r.t. $\boldsymbol{\nu}$ and satisfies

$$\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\infty} \pi_i(\mathscr{Z}_{k-i}) \tag{3.36}$$

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

With $\alpha_i = \beta_j = 0_{\mathcal{L}_{\mathscr{H}}}$ for $i > \mathfrak{p}, j > \mathfrak{q}$, the representations (3.2), (3.36) yield

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

a.s. for all k. 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 implies (see e.g. [23], Lemma 4.48 and Remark 4.49)

$$\pi_i = \alpha_i + \sum_{j=1}^{(i-1)\wedge\mathfrak{q}} \beta_j \pi_{i-j}, \quad i \in \mathbb{N}. \tag{3.37}$$

This identity, since $\alpha_i = 0_{\mathcal{L}_{\mathscr{H}}}$ for $i > \mathfrak{p}$, implies

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

where $\mathfrak{s} = \mathfrak{p} + \mathfrak{q}$. The solution $\boldsymbol{\beta}_{[\mathfrak{q}]}$ is unique iff the image of $\boldsymbol{\pi}_{[\mathfrak{p},\mathfrak{q}]}^T \in \mathcal{S}_{\mathscr{H},\mathscr{H}^{\mathfrak{q}}}$ lies dense which is impossible due to $\mathscr{H} \subsetneq \mathscr{H}^{\mathfrak{q}}$. Therefore, we establish estimators for $\beta_1, ..., \beta_{\mathfrak{q}}$ based on the equation

$$\boldsymbol{\pi}_{[\mathfrak{s},\mathfrak{q}]} = \boldsymbol{\beta}_{[\mathfrak{q}]} \prod_{[\mathfrak{s},\mathfrak{q}]} :\Longleftrightarrow \begin{bmatrix} \pi_{\mathfrak{s}+\mathfrak{q}-1} \\ \pi_{\mathfrak{s}+\mathfrak{q}-2} \\ \vdots \\ \pi_{\mathfrak{s}} \end{bmatrix} = \big[\beta_1\, \beta_2\, \cdots\, \beta_{\mathfrak{q}} \big] \begin{bmatrix} \pi_{\mathfrak{s}+\mathfrak{q}-2} & \pi_{\mathfrak{s}+\mathfrak{q}-3} & \cdots & \pi_{\mathfrak{s}-1} \\ \pi_{\mathfrak{s}+\mathfrak{q}-3} & \pi_{\mathfrak{s}+\mathfrak{q}-4} & \cdots & \pi_{\mathfrak{s}-2} \\ \vdots & \vdots & \cdots & \vdots \\ \pi_{\mathfrak{s}-1} & \pi_{\mathfrak{s}-2} & \cdots & \pi_{\mathfrak{p}} \end{bmatrix} \tag{3.38}$$

However, before we establish an estimator for $\boldsymbol{\beta}_{[\mathfrak{q}]}$ and thus for the operators $\beta_1, ..., \beta_{\mathfrak{q}}$, we clarify that the image of $\prod_{[\mathfrak{s},\mathfrak{q}]} \in \mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$ can lie dense.

Example 3.2. Let $\alpha_i = \beta_j = 0_{\mathcal{L}_{\mathscr{H}}}$ hold for $i \neq \mathfrak{p}, j \neq \mathfrak{q}$, and let $\gamma := \alpha_{\mathfrak{p}} = \beta_{\mathfrak{q}} \in \mathcal{S}_{\mathscr{H}}$ be an injective, self-adjoint operator with $\gamma \neq 0_{\mathcal{L}_{\mathscr{H}}}$ and $||\gamma||_{\mathcal{S}_{\mathscr{H}}} < 1$. Then, since (3.37) yields $\pi_m = \gamma^k$ for all $m = \mathfrak{p} + (k-1)\mathfrak{q}$ for some $k \in \mathbb{N}$ and $\pi_m = 0_{\mathcal{L}_{\mathscr{H}}}$ otherwise, and since operator-valued matrices $A = (a_{ij})_{i,j}$ satisfy $A^* = (a_{ji}^*)_{i,j}$,

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

Hence, since γ and thus γ^2 are injective and selfadjoint, $\prod_{[\mathfrak{s},\mathfrak{q}]}$ has dense image.

Due to (3.38), analogously to (3.33), we use

$$\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]} := \hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} \prod\nolimits_{[\mathfrak{s},\mathfrak{q}]}^\dagger \coprod_{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};1}}^{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};M}} = \hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} \prod\nolimits_{[\mathfrak{s},\mathfrak{q}]}^* \left(\prod\nolimits_{[\mathfrak{s},\mathfrak{q}]} \prod\nolimits_{[\mathfrak{s},\mathfrak{q}]}^* + \theta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{q}}} \right)^{-1} \coprod_{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};1}}^{\hat{\mathfrak{g}}_{\mathfrak{s},\mathfrak{q};M}} \tag{3.39}$$

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

$$\hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} := \begin{bmatrix} \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-1} & \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-2} & \cdots & \hat{\pi}_{\mathfrak{s}} \end{bmatrix} \tag{3.40}$$

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

$$\hat{\prod}\nolimits_{[\mathfrak{s},\mathfrak{q}]} := \begin{bmatrix} \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-2} & \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-3} & \cdots & \hat{\pi}_{\mathfrak{s}-1} \\ \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-3} & \hat{\pi}_{\mathfrak{s}+\mathfrak{q}-4} & \cdots & \hat{\pi}_{\mathfrak{s}-2} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\pi}_{\mathfrak{s}-1} & \hat{\pi}_{\mathfrak{s}-2} & \cdots & \hat{\pi}_{\mathfrak{p}} \end{bmatrix} \tag{3.41}$$

is an element of $\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$. For any k, $\hat{\pi}_k$ denotes the estimator of π_k, obtained by choosing the k-th component of

$$\hat{\boldsymbol{\pi}}_{L,K} := \tilde{\mathfrak{S}}_{L,K} \tilde{\mathfrak{S}}_L^\dagger \coprod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} = \tilde{\mathfrak{S}}_{L,1} \tilde{\mathfrak{S}}_L (\tilde{\mathfrak{S}}_L^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^L})^{-1} \coprod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} \tag{3.42}$$

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 $\tilde{\mathfrak{S}}_L$ associated to the first biggest eigenvalues $\hat{c}_{L,1} \geq \cdots \geq \hat{c}_{L,K}$. In (3.39), $(M_N)_{N \in \mathbb{N}} \subseteq \mathbb{N}$ 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 w.l.o.g. monotonically decreasing eigenvalue sequence $(\hat{g}_{\mathfrak{s},\mathfrak{q};j})_{j \in \mathbb{N}}$. Due to (3.37),

$$\hat{\alpha}_i := \hat{\pi}_i - \sum_{j=1}^{(i-1) \wedge \mathfrak{q}} \hat{\beta}_j \hat{\pi}_{i-j}, \tag{3.43}$$

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

In order to establish upper bounds of the estimation errors for the GARCH operators, we firstly derive those of the estimation errors for the operators π_i.

Lemma 3.5. *Let Assumptions 3.1-3.2, 3.5 hold. Let also $\beta > 0, L = L_N = \Xi(1,N), K = K_N = \Xi(1, \sqrt[3]{L^{-1}N})$ and $\vartheta_N = \mathrm{O}(c_{L,K}^2 K^{-\beta})$ hold. Thereby, $(c_{L,j})_j$ is the eigenvalue sequence of $\mathfrak{S}_L$ for which holds Assumption 3.3 for all L, $c_{L,K} = \Omega(\sqrt{L^3 N^{-1}})$, $\sum_{l=1}^{K} (\frac{c_{L,l}^2}{c_{L,l}^2 + \vartheta_N})^2 \sum_{j>K} \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = \mathrm{O}(K^{-2\beta})$ where $\boldsymbol{\pi}_L := [\pi_1 \cdots \pi_L]$, and $c_{L,K}^{-2} L(\gamma_{L,K}^2 K L^3 N^{-1} + (\sum_{l>L} ||\pi_l||_{\mathcal{L}_{\mathscr{H}}})^2) = \mathrm{O}(K^{-2\beta})$ where $\gamma_{L,K} := (c_{L,K} - c_{L,K+1})^{-1}$. At last, we also impose that $(\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$, with $\Phi_{L,ij} := \mathfrak{c}_{L,i} \otimes \mathfrak{c}_j$ as in Theorem 3.2, satisfies Assumption 3.4 for all L for some $\beta > 0$ being independent of L. Then, for all $i \in \mathbb{N}$ holds*

$$||\hat{\pi}_i - \pi_i||_{\mathcal{S}_{\mathscr{H}}}^2 = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.44}$$

The conditions $L = L_N = \Xi(1, N), K = K_N = \Xi(1, \sqrt[3]{L^{-1}N}), \vartheta_N = O(c_{L,K}^{-2} K^{-\beta})$ hold if $(K_N)_N, (L_N)_N, (\vartheta_N)_N$ are chosen appropriately. The following example outlines that there are sequences of operators $(\pi_i)_i$ being absolutely summable w.r.t. $||\cdot||_{\mathcal{S}_\mathscr{H}}$, and that other more complex conditions in Lemma 3.5 are also satisfiable, at least if some information about the eigenvalue sequences is given.

Example 3.3. As in Example 3.2, we impose $\alpha_i = \beta_j = 0_{\mathcal{L}_\mathscr{H}}$ for $i \neq \mathfrak{p}, j \neq \mathfrak{q}$, and $\gamma := \alpha_\mathfrak{p} = \beta_\mathfrak{q} \in \mathcal{S}_\mathscr{H}$. Additionally, $\gamma = \sum_{j=1}^\infty \xi^j (\mathfrak{c}_j \otimes \mathfrak{c}_j)$ for some $\xi \in (0, \frac{1}{\sqrt{2}})$, which implies $\gamma \neq 0_{\mathcal{L}_\mathscr{H}}$ and $||\gamma||_{\mathcal{S}_\mathscr{H}}^2 = \sum_{j=1}^\infty \xi^{2j} = \frac{\xi^2}{1-\xi^2} < 1$. Also, for some $a > \frac{1}{2}$, we assume the mild condition $\langle \mathfrak{c}_{L,i}^{(n)}, \mathfrak{c}_j \rangle_\mathscr{H} = O(i^{-a})$ for $i \to \infty$ for any j, L, n.

Due to (3.37) and our assumptions, $\pi_m = \gamma^{k_m}$ for $m = \mathfrak{p} + (k_m - 1)\mathfrak{q}$ with $k_m \in \mathbb{N}$ and $\pi_m = 0_{\mathcal{L}_\mathscr{H}}$ otherwise. This implies with $||\gamma||_{\mathcal{S}_\mathscr{H}} < 1$ and sub-multiplicativity:

$$\sum_{m=1}^\infty ||\pi_m||_{\mathcal{S}_\mathscr{H}} \leq \sum_{k=1}^\infty ||\gamma||_{\mathcal{S}_\mathscr{H}}^k < \infty. \qquad (3.45)$$

Now, we illustrate that $(\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$ can satisfy Assumption 3.4 for all L for some $\beta > 0$ being independent of L. For $L < \mathfrak{p}$, this assumption is satisfied here for any β, since then $\boldsymbol{\pi}_L = [\pi_1 \cdots \pi_L] = [0_{\mathcal{L}_\mathscr{H}} \cdots 0_{\mathcal{L}_\mathscr{H}}]$. For $L \geq \mathfrak{p}$, we choose $k_L \in \mathbb{N}$ such that $\mathfrak{p} + (k_L - 1)\mathfrak{q} \leq L < \mathfrak{p} + k_L \mathfrak{q}$. Then, the relationship between π_m and γ for all m, the features of γ, and $\Phi_{L,ij}(\mathfrak{c}_{L,l}) = \delta_{il}\mathfrak{c}_j$ for any l imply

$$
\begin{aligned}
\langle \boldsymbol{\pi}_L, \Phi_{L,ij} \rangle_{\mathcal{S}_{\mathscr{H}^L, \mathscr{H}}}^2 &= \sum_{l=1}^\infty \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \Phi_{L,ij}(\mathfrak{c}_{L,l}) \rangle_\mathscr{H}^2 = \Big\langle \sum_{m=1}^L \pi_m(\mathfrak{c}_{L,i}^{(m)}), \mathfrak{c}_j \Big\rangle_\mathscr{H}^2 \\
&= \Big\langle \sum_{m=1}^{k_L} \gamma^m(\mathfrak{c}_{L,i}^{(\mathfrak{p}+(m-1)\mathfrak{q})}), \mathfrak{c}_j \Big\rangle_\mathscr{H}^2 \leq k_L \sum_{m=1}^{k_L} \xi^{2jm} \langle \mathfrak{c}_{L,i}^{(\mathfrak{p}+(m-1)\mathfrak{q})}, \mathfrak{c}_j \rangle_\mathscr{H}^2,
\end{aligned}
$$

and since $|\xi| < 1$ and $\langle \mathfrak{c}_{L,i}^{(n)}, \mathfrak{c}_j \rangle_\mathscr{H}^2 = O(i^{-2a})$ for $i \to \infty$ and any j, L, n with $a > 1/2$, for $\beta \in (0, a - \frac{1}{2})$ holds

$$\sum_{i,j=1}^\infty \langle \boldsymbol{\pi}_L, \Phi_{L,ij} \rangle_{\mathcal{S}_{\mathscr{H}^L, \mathscr{H}}}^2 (1 + i^{2\beta} + j^{2\beta}) \leq k_L \sum_{m=1}^{k_L} \sum_{j=1}^\infty (1 + j^{2\beta}) \xi^{2jm} \sum_{i=1}^\infty (1 + i^{2\beta}) \langle \mathfrak{c}_{L,i}^{(\mathfrak{p}+(m-1)\mathfrak{q})}, \mathfrak{c}_j \rangle_\mathscr{H}^2 < \infty.$$

Hence, $(\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$ indeed can satisfy Assumption 3.4 for all L for some $\beta > 0$ independent of L.

At last, we validate some assumptions in Lemma 3.5 for the eigenvalues which is generally difficult because their specific features are mostly unknown. However, since $\mathfrak{S}_L$ is nuclear, due to $\mathfrak{c}_{L,K} \leq ||\mathfrak{S}_L||_{\mathcal{N}_{\mathcal{H}^L}}$ and (3.8), we know $\mathfrak{c}_{L,K} = O(L)$ for $L \to \infty$ for fixed K and $\mathfrak{c}_{L,K} = o(K^{-1})$ for $K \to \infty$ for fixed L. Both features are satisfied if, for instance, for $K = K_N := 1 + \lfloor \tilde{K} \ln(N) \rfloor$ and $L = L_N := \lfloor N^{\tilde{L}} \rfloor$ for all $N \in \mathbb{N}$ with $b, \tilde{K}, \tilde{L} \in (0,1)$ holds

$$\mathfrak{c}_{L,K} \asymp e^{-K} L^b \asymp N^{b\tilde{L} - \tilde{K}} \implies \gamma_{L,K} = (\mathfrak{c}_{L,K} - \mathfrak{c}_{L,K+1})^{-1} \asymp \mathfrak{c}_{L,K}^{-1}. \qquad (3.46)$$

Justifying such an asymptotic behaviour is difficult, but a priori not ruled out. (3.46) implies $L = \Xi(1, N), K = \Xi(1, \sqrt[3]{L^{-1}N})$, and $\vartheta_N = O(c_{L,K}^2 K^{-\beta})$ for any $\beta > 0$ if e.g. $\vartheta_N := N^{2(b\tilde{L} - \tilde{K}) - c}$ for all N for some $c > 0$, and also $\mathfrak{c}_{L,K} = \Omega(\sqrt{L^3 N^{-1}})$ if $(3 - 2b)\tilde{L} + 2\tilde{K} \leq 1$ which is true e.g. for $\tilde{K} = \tilde{L} = \frac{1}{6}$ and $b = \frac{3}{4}$. Moreover, our assumptions, $\langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle_\mathscr{H}^2 = \langle \boldsymbol{\pi}_L, \Phi_{L,lj} \rangle_{\mathcal{S}_{\mathscr{H}^L, \mathscr{H}}}^2$ for any j, l, L (see above), $\xi^{2jm} \leq \xi^{j+m}$ for any $j, m \in \mathbb{N}, k_{L \asymp L}$, and the fact that $(\mathfrak{c}_{L,l})_l$ and $(\mathfrak{c}_j)_j$ are CONS of $\mathscr{H}^L$ and $\mathscr{H}$, respectively, yield for any $\beta > 0$:

$$\sum_{l=1}^K \Big(\frac{c_{L,l}^2}{c_{L,l}^2 + \vartheta_N}\Big)^2 \sum_{j>K} \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle_\mathscr{H}^2 \leq k_L \sum_{j>K} \sum_{m=0}^\infty \xi^{j+m} \sum_{l=1}^\infty \langle \mathfrak{c}_{L,i}^{(\mathfrak{p}+(m-1)\mathfrak{q})}, \mathfrak{c}_j \rangle_\mathscr{H}^2 \leq \frac{k_L \xi^{K+1}}{(1-\xi)^2} = O(K^{-2\beta}).$$

Further, (3.45), $||\cdot||_{\mathcal{L}_\mathscr{H}} \leq ||\cdot||_{\mathcal{S}_\mathscr{H}}, L^* := \lfloor \frac{L-\mathfrak{p}}{\mathfrak{q}} \rfloor$ and $||\gamma||_{\mathcal{S}_\mathscr{H}} < 1$ leads to $\sum_{l>L} ||\pi_l||_{\mathcal{L}_\mathscr{H}} \leq \sum_{k>L^*+1} ||\gamma||_{\mathcal{S}_\mathscr{H}}^k \propto ||\gamma||_{\mathcal{S}_\mathscr{H}}^{L^*} \asymp ||\gamma||_{\mathcal{S}_\mathscr{H}}^L$. Therefore, with (3.46), $K = K_N \sim \tilde{K} \ln(N)$ and $L = L_N \sim N^{\tilde{L}}$ for $\tilde{K}, \tilde{L} \in (0,1)$, we get

$$c_{L,K}^{-2} L \Big(\gamma_{L,K}^2 K L^3 N^{-1} + \big(\sum_{l>L} ||\pi_l||_{\mathcal{L}_\mathscr{H}}\big)^2\Big) \asymp c_{L,K}^{-4} K L^4 N^{-1} \asymp \ln(N) N^{4(\tilde{K} + (1-b)\tilde{L}) - 1} = O(K^{-2\beta})$$

for any $\beta > 0$ if $\tilde{K} + (1-b)\tilde{L} < \frac{1}{4}$, which e.g. holds for $\tilde{K} = \tilde{L} = \frac{1}{6}$ and $b = \frac{3}{4}$ as above.

Before we turn to the estimation of our GARCH operators, we would like to mention that Lemma 3.5 leads to upper bounds for the estimation errors for the operators of linear processes being, though not needed in this work, interesting beyond (G)ARCH. Due to Assumption 3.5 where we imposed $\mathscr{Z} = (\mathscr{Z}_k)_k$ to be both linear, that is $\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\infty} \psi_i(\nu_{k-i})$ a.s. for all k for some sequence of bounded operators $(\psi_i)_{i\in\mathbb{N}} \subset \mathcal{L}_{\mathscr{H}}$, and invertible with the representation $\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\infty} \pi_i(\mathscr{Z}_{k-i})$ a.s. for all k, we have

$$\mathscr{Z}_k = \nu_k + \sum_{i=1}^{\infty} \Big[\sum_{j=1}^{i} \pi_i \psi_{i-j} \Big](\nu_{k-i}) \tag{3.47}$$

a.s. for all k. This and the linear representation yield with $\psi_0 := \mathbb{I}_{\mathscr{H}}$:

$$\psi_i = \sum_{j=1}^{i} \pi_i \psi_{i-j}, \quad i \in \mathbb{N}, \tag{3.48}$$

and thus $\psi_i \in \mathcal{S}_{\mathscr{H}}$ for all i. Conversely, the linear representation and (3.48) imply (3.47) due to Lemma 3.1.

Proposition 3.1. *Let the assumptions of Lemma 3.5 hold. Then,*

$$\hat{\psi}_i := \sum_{j=1}^{i} \hat{\pi}_i \hat{\psi}_{i-j}, \quad i \in \mathbb{N}, \tag{3.49}$$

where $\hat{\psi}_0 := \mathbb{I}_{\mathscr{H}}$, is a consistent estimator for ψ_i with

$$||\hat{\psi}_i - \psi_i||^2_{\mathcal{S}_{\mathscr{H}}} = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.50}$$

Remark 3.3. Lemma 3.5 and Proposition 3.1 are transferable to further separable Hilbert spaces (see Remark 3.2), and are valid without the context of (G)ARCH. Both results are possibly applicable on theories developed in [3] and [20], and can be seen as advances of the Theorems 3.2-3.3 in [3] where the operators in the linear process and the inverted representation were, based on assertions regarding spectral density operators, consistently estimated but without deriving an explicit rate.

For the sake of infering upper bounds for the GARCH operators, we need that for the operators $\boldsymbol{\pi}_{[\mathfrak{s},\mathfrak{q}]}$ and $\prod_{[\mathfrak{s},\mathfrak{q}]}$ in (3.38) holds, due to (3.44),

$$||\hat{\boldsymbol{\pi}}_{[\mathfrak{s},\mathfrak{q}]} - \boldsymbol{\pi}_{[\mathfrak{s},\mathfrak{q}]}||^2_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}},\mathscr{H}}} = \sum_{i=0}^{\mathfrak{q}-1} ||\hat{\pi}_{\mathfrak{s}+i} - \pi_{\mathfrak{s}+i}||^2_{\mathcal{S}_{\mathscr{H}}} = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) \tag{3.51}$$

and also

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

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

$$= \mathrm{O}_{\mathbb{P}}(K^{-2\beta}). \tag{3.52}$$

Furthermore, the identities (3.13), (3.52) and Corollary 3.1 yield

$$\sup_{j\in\mathbb{N}} (\hat{g}_{\mathfrak{s},\mathfrak{q};j} - g_{\mathfrak{s},\mathfrak{q};j})^2 = \mathrm{O}(K^{-2\beta}) \tag{3.53}$$

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 $\prod_{[\mathfrak{s},\mathfrak{q}]}\prod^*_{[\mathfrak{s},\mathfrak{q}]} \in \mathcal{S}_{\mathscr{H}^{\mathfrak{q}}}$, and Corollary 3.1 implies

$$\hat{g}_{\mathfrak{s},\mathfrak{q};M} = \mathrm{O}_{\mathbb{P}}(g_{\mathfrak{s},\mathfrak{q};M}) \quad \text{and} \quad g_{\mathfrak{s},\mathfrak{q};M} = \mathrm{O}_{\mathbb{P}}(\hat{g}_{\mathfrak{s},\mathfrak{q};M}) \tag{3.54}$$

16

if $g_{\mathfrak{s},\mathfrak{q};M} = \Xi(K^{-\beta}, 1)$ with $M = M_N = \Xi(1, N)$. If also $\prod_{[\mathfrak{s},\mathfrak{q}]} \prod_{[\mathfrak{s},\mathfrak{q}]}^*$ is injective and satisfies Assumption 3.3,

$$\sup_{j \le M} ||\hat{\mathfrak{g}}'''_{\mathfrak{s},\mathfrak{q};j} - \mathfrak{g}_{\mathfrak{s},\mathfrak{q};j}||^2_{\mathscr{H}^{\mathfrak{q}}} = O_{\mathbb{P}}(\gamma^2_{\mathfrak{s},\mathfrak{q};M} K^{-2\beta}) \tag{3.55}$$

holds after Lemma 3.4 where $\gamma_{\mathfrak{s},\mathfrak{q};M} := (g_{\mathfrak{s},\mathfrak{q};M} - g_{\mathfrak{s},\mathfrak{q};M+1})^{-1}$ are the reciprocal spectral gaps.

By using these upper bounds, we are now able to state an asymptotic upper bound of the estimation errors for the complete GARCH operators.

Theorem 3.3. *Let the assumptions of Lemma 3.5 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 $M = M_N = \Xi(1, N), \theta_N = O(K^{-\beta/2}), \gamma_{\mathfrak{s},\mathfrak{q};M} K^{-\beta} = o(1), g^{-2}_{\mathfrak{s},\mathfrak{q};M} K^{-\beta} = O(M^{-2\beta})$ and $\sum_{l=1}^M (\frac{g^2_{\mathfrak{s},\mathfrak{q};l}}{g^2_{\mathfrak{s},\mathfrak{q};l} + \theta_N})^2 \sum_{j>M} \langle \boldsymbol{\beta}_{[\mathfrak{q}]}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}), \mathfrak{c}_j \rangle^2_{\mathscr{H}} = O(M^{-2\beta})$ hold, and let the tuple $(\boldsymbol{\beta}_{[\mathfrak{q}]}, (\Upsilon_{\mathfrak{s},\mathfrak{q};ij})_{i,j})$ with $\Upsilon_{\mathfrak{s},\mathfrak{q};ij} := \mathfrak{g}_{\mathfrak{s},\mathfrak{q};i} \otimes \mathfrak{c}_j$ satisfy Assumption 3.4 for the same $\beta > 0$ as in Lemma 3.5. Then,*

$$||\hat{\alpha}_i - \alpha_i||^2_{\mathcal{S}_{\mathscr{H}}} = \begin{cases} O_{\mathbb{P}}(K^{-2\beta}), & i = 1, \\ O_{\mathbb{P}}(M^{-2\beta}), & i = 2, ..., \mathfrak{p}, \end{cases} \tag{3.56}$$

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

$$||\hat{\beta}_j - \beta_j||^2_{\mathcal{S}_{\mathscr{H}}} = O_{\mathbb{P}}(M^{-2\beta}). \tag{3.57}$$

Remark 3.4.   (a) To the best of our knowledge, estimating GARCH operators by a Yule-Walker approach and explicit asymptotic upper bounds of their estimation errors as deduced in Theorem 3.3 are new, as is estimating GARCH($\mathfrak{p}, \mathfrak{q}$) operators for $\mathfrak{p} \vee \mathfrak{q} > 1$. Theorem 3.3 extends theories developed both in *Aue et al.* [2] and *Cerovecki et al.* [6] where the parameters of GARCH($1, 1$) and GARCH($\mathfrak{p}, \mathfrak{q}$) processes for arbitrary orders $\mathfrak{p}, \mathfrak{q} \in \mathbb{N}$, respectively, were strongly consistently estimated, though without stating explicit upper bounds for the estimation errors. [2] imposed both operators to be integral operators, their integral kernels and the shift term δ to be expressed by a finite linear combination, and estimated the coefficients in these finite expansions by a least squares estimator. [6] estimated their parameters, inspired by the standard GARCH model, by a quasi-maximum likelihood approach where the operators were also imposed to be integral operators. For any order $\mathfrak{p}, \mathfrak{q}$, they assumed the parameters to have a representation as a finite linear combination of which they estimated their coefficients, and for $\mathfrak{p} = \mathfrak{q} = 1$ even the complete operators were estimated. Moreover, both [2] and [6] deduced asymptotic normality of the estimation errors in the finite-dimensional setup.

   (b) In Theorem 3.3, as when estimating the complete operators of our ARCH and invertible process, it is preferable that β is as large and that the sequences $K = K_N$ and $M = M_N$ increase as fast as possible. Using $M = M_N$, declaring the dimension of the subspace on what we project when estimating $\beta_1, ..., \beta_{\mathfrak{q}}$ by means of $\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]}$ in (3.39), instead of $K = K_N$, provides an additional regulation parameter and thus more combinations so that all requirements in Theorem 3.3 are met.

   (c) Theorem 3.3 can be generalized to further separable Hilbert spaces and also holds for operators of $\mathscr{H}$-valued ARMA processes, since the estimators for the GARCH operators were derived from operators associated to ARMA, see also the arguments in Remark 3.2 (e)-(f).

We close the section by validating the assumptions in Theorem 3.3, and outline possible upper bounds.

Example 3.4. Let all assumptions in Examples 3.2-3.3 hold. So, $K = K_N \sim \tilde{K} \ln(N), L = L_N \sim N^{\tilde{L}}, c_{L,K} \asymp \gamma^{-1}_{L,K} \asymp N^{b\tilde{L} - \tilde{K}}$ and $\vartheta_N \asymp N^{2(b\tilde{L} - \tilde{K}) - c}$ where $b, \tilde{K}, \tilde{L} \in (0, 1), c > 0$, and the representation of $\prod_{[\mathfrak{s},\mathfrak{q}]}$ yields

$$\prod_{[\mathfrak{s},\mathfrak{q}]}^2 = \begin{bmatrix} \gamma^4 & 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} \\ 0_{\mathcal{L}_{\mathscr{H}}} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \gamma^4 & 0_{\mathcal{L}_{\mathscr{H}}} \\ 0_{\mathcal{L}_{\mathscr{H}}} & \cdots & 0_{\mathcal{L}_{\mathscr{H}}} & \gamma^2 \end{bmatrix}$$

17

Thus, due to $\gamma(\mathfrak{c}_j) = \xi^j \mathfrak{c}_j$ for any j with $\xi \in (0,1)$, the eigenvalues of $\prod^2_{[\mathfrak{s},\mathfrak{q}]}$ are $g_{\mathfrak{s},\mathfrak{q};1} = \xi^2$, $g_{\mathfrak{s},\mathfrak{q};2} = \cdots = g_{\mathfrak{s},\mathfrak{q};\mathfrak{q}+1} = \xi^4$, $g_{\mathfrak{s},\mathfrak{q};\mathfrak{q}+2} = \xi^6$, $g_{\mathfrak{s},\mathfrak{q};\mathfrak{q}+3} = \cdots = g_{\mathfrak{s},\mathfrak{q};2\mathfrak{q}+2} = \xi^8, g_{\mathfrak{s},\mathfrak{q};2\mathfrak{q}+3} = \xi^{10}, ...$, with associated eigenfunctions $\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l} = (0_{\mathscr{H}}, ..., 0_{\mathscr{H}}, \mathfrak{c}_j)^T \in \mathscr{H}^{\mathfrak{q}}$ for $l = 1 + (j-1)(\mathfrak{q}+1)$ and $\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l} = (0_{\mathscr{H}}, ..., 0_{\mathscr{H}}, \mathfrak{c}_j, 0_{\mathscr{H}}, ..., 0_{\mathscr{H}})^T \in \mathscr{H}^{\mathfrak{q}}$ with $\mathfrak{c}_j$ positioned at k-th component with $k = 1, ..., \mathfrak{q}-1$ for $l = 1 + k + (j-1)(\mathfrak{q}+1)$, Further, $\prod^2_{[\mathfrak{s},\mathfrak{q}]}$ is injective. Moreover, $\kappa(j) = g_{\mathfrak{s},\mathfrak{q};j}$ for all j for some convex function $\kappa \colon \mathbb{R} \to \mathbb{R}$, but $(g_{\mathfrak{s},\mathfrak{q};j})_j$ does not fulfil the complete Assumption 3.3, because $g_{\mathfrak{s},\mathfrak{q};j} = g_{\mathfrak{s},\mathfrak{q};j+1}$ for some j. This, however, is not a real problem, since assertions when imposing one-dimensional generally can be extended on multi-dimensional eigenspaces by modifying the reciprocal spectral gaps, see [5]. For such a modification, also denoted by $\gamma_{\mathfrak{s},\mathfrak{q};j}$, holds due to $g_{\mathfrak{s},\mathfrak{q};n} = \xi^{2f(n)}$ for some function $f \colon \mathbb{N} \to \mathbb{N}$ with $f(n) \leq n$ for all $n, \gamma_{\mathfrak{s},\mathfrak{q};j} = (\xi^{2f(j)} - \xi^{2(f(j)+1)})^{-1} \propto \xi^{-2f(j)} = g^{-1}_{\mathfrak{s},\mathfrak{q};j} \leq \xi^{-2j}$. Then, with $M = M_N := 1 + \lfloor \frac{\beta}{3} \log_{\xi^{-2}}(K) \rfloor$ holds $M = \Xi(1,N), \gamma_{\mathfrak{s},\mathfrak{q};M} K^{-\beta} \propto g^{-1}_{\mathfrak{s},\mathfrak{q};M} K^{-\beta} \leq \xi^{-2M} K^{-\beta} \asymp K^{-\frac{2\beta}{3}} = o(1)$ and $g^{-2}_{\mathfrak{s},\mathfrak{q};M} K^{-\beta} \leq \xi^{-4M} K^{-\beta} \asymp K^{-\frac{\beta}{3}} = O(M^{-2\beta})$, and $\theta_N := \ln^{-\beta}(N+1)$ for all N fulfils $\theta_N = O(K^{-2\beta})$. Further, from $\boldsymbol{\beta}_{[\mathfrak{q}]} = [0_{\mathcal{L}_{\mathscr{H}}}, ..., 0_{\mathcal{L}_{\mathscr{H}}}, \gamma]^T$, the representations of the eigenfunctions $\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}$ for any l and $\theta_N \geq 0$ for all N follows for any $\beta > 0$:

$$\sum_{l=1}^{M} \left(\frac{g^2_{\mathfrak{s},\mathfrak{q};l}}{g^2_{\mathfrak{s},\mathfrak{q};l} + \theta_N} \right)^2 \sum_{j>M} \langle \boldsymbol{\beta}_{[\mathfrak{q}]}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}), \mathfrak{c}_j \rangle^2_{\mathscr{H}} \leq \sum_{j>M} \sum_{l=1}^{\infty} \xi^{2l} \delta_{lj} \propto \xi^{2M} = O(M^{-2\beta}).$$

Also, from the definition of $\boldsymbol{\beta}_{[\mathfrak{q}]}, \gamma, \mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}$ for all l, $\Upsilon_{\mathfrak{s},\mathfrak{q};ij}$ for all i,j, and $\langle \cdot, \cdot \rangle_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}, \mathscr{H}}}$ follows $\langle \boldsymbol{\beta}_{[\mathfrak{q}]}, \Upsilon_{\mathfrak{s},\mathfrak{q};ij} \rangle^2_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}, \mathscr{H}}} = \xi^{2i} \langle \mathfrak{g}^{(\mathfrak{q})}_{\mathfrak{s},\mathfrak{q};l}, \mathfrak{c}_j \rangle^2_{\mathscr{H}} = \mathbf{1}_{A_{\mathfrak{q}}}(i) \delta_{kij} \xi^{2i}$ for some $k_i \in \mathbb{N}$ with $k_i \leq i$ where $A_{\mathfrak{q}} := \{k \,|\, \exists\, l = l_k$ with $\mathfrak{g}^{(\mathfrak{q})}_{\mathfrak{s},\mathfrak{q};k} = \mathfrak{c}_l\}$. Then,

$$\sum_{i,j=1}^{\infty} \langle \boldsymbol{\beta}_{[\mathfrak{q}]}, \Upsilon_{\mathfrak{s},\mathfrak{q};ij} \rangle^2_{\mathcal{S}_{\mathscr{H}^{\mathfrak{q}}, \mathscr{H}}} (1 + i^{2\beta} + j^{2\beta}) = \sum_{i=1}^{\infty} \mathbf{1}_{A_{\mathfrak{q}}}(i) \xi^{2i} (1 + i^{2\beta})(1 + k_i^{2\beta}) \leq \sum_{i=1}^{\infty} \xi^{2i} (1 + i^{2\beta})^2 < \infty$$

for all $\beta > 0$, in other words $(\boldsymbol{\beta}_{[\mathfrak{q}]}, (\Upsilon_{\mathfrak{s},\mathfrak{q};ij})_{i,j})$ fulfils Assumption 3.4 for all $\beta > 0$.

Provided having multi-dimensional eigenspaces in our setup does not cause a problem, all requirements of Theorem 3.3 are validated. Due to the asymptotic behaviour of our sequences, Theorem 3.3 then yields

$$||\hat{\alpha}_i - \alpha_i||^2_{\mathcal{S}_{\mathscr{H}}} = \begin{cases} O_{\mathbb{P}}(\ln^{-2\beta}(N)), & i = 1, \\ O_{\mathbb{P}}(\ln^{-2\beta}(\ln(N))), & i = 2, ..., \mathfrak{p}, \end{cases}$$

and for $j = 1, ..., \mathfrak{q}$ holds due to $M = M_N \asymp \ln(K)$ and $K = K_N \asymp \ln(N)$,

$$||\hat{\beta}_j - \beta_j||^2_{\mathcal{S}_{\mathscr{H}}} = O_{\mathbb{P}}(\ln^{-2\beta}(\ln(N))).$$

Remark 3.5. Having just logarithmic decay rates in Example 3.4, is because we exemplarily imposed a geometric decay of given eigenvalues. This made conversions less cumbersome since the reciprocal eigenvalues then are asympotically equal to the reciprocal spectral gaps up to a multiplicative constant. By assuming a slower decay, but no too slow since our eigenvalues are absolutely-summable, the proof of the requirements in Theorem 3.3 is possibly more difficult, though the convergence rates for the operators are better.

4   Simulation

Herein, we illustrate some ARCH and GARCH processes in $F = \mathscr{H} = L^2[0,1]$ and estimation errors for their parameters. Before we do this, δ and operators satisfying the requirements in Theorem 2.1 will be defined, and it will be shown that under certain errors ε_k holds (2.10) which guarantees the existence of a strictly stationary solution. Since verifying (2.10) is cumbersome for big $\mathfrak{s} = \mathfrak{p} + \mathfrak{q}$ (even for $\mathfrak{s} = 3$, see [23], p.92), we focus on ARCH(1) and GARCH(1,1).

At first, we put

$$\delta(t) := 0.01, \quad t \in [0,1]. \tag{4.1}$$

Thereby, $\delta \in \mathscr{\dot{H}}_{>0}$ with $\dot{F} = \mathscr{\dot{H}} := L^4[0,1]$. Our two operators $\alpha_1, \beta_1 \colon \mathscr{H} \to \mathscr{H}$ are imposed to be integral operators since these easily can be illustrated. An *integral operator* $A \colon \mathscr{H} \to \mathscr{H}$ is defined by $(A(x)(t)) := \int_0^1 a(s,t)\,x(s)\,\mathrm{d}s$ for $x \in \mathscr{H}, t \in [0,1]$ if the integral meant w.r.t. λ exists, where $a \colon [0,1]^2 \to \mathbb{R}$ is a measurable function, the *integral kernel of A*. We define the integral kernels a_1 of α_1 and b_1 of β_1 by

$$a_1(s,t) := \frac{1}{8}\Big[\Big(s - \frac{1}{2}\Big)^2 + \Big(t - \frac{1}{2}\Big)^2\Big], \qquad b_1(s,t) := \frac{1}{8}(1-s)^2 t^2 \tag{4.2}$$

for any $s,t \in [0,1]$. Since a_1, b_1 are non-negative, $\alpha_1, \beta_1 \colon \mathscr{H}_{\geq 0} \to \mathscr{H}_{\geq 0}$, and $\alpha_1, \beta_1 \in \mathcal{L}_{\mathscr{H},\mathscr{H}}$ because Cauchy-Schwarz and Jensen's inequality lead to finite norms, to be precise,

$$||\alpha_1||^4_{\mathcal{L}_{\mathscr{H},\mathscr{H}}} = \sup_{||h||_{\mathscr{H}} \leq 1} \int_0^1 \Big(\int_0^1 a_1(s,t)h(s)\,\mathrm{d}s\Big)^4 \mathrm{d}t \leq \int_0^1 \int_0^1 a_1^4(s,t)\,\mathrm{d}s\,\mathrm{d}t \approx 8.041 \cdot 10^{-7} =: \tilde{\alpha}_1^4, \tag{4.3}$$

$$||\beta_1||^4_{\mathcal{L}_{\mathscr{H},\mathscr{H}}} \leq \int_0^1 \int_0^1 b_1^4(s,t)\,\mathrm{d}s\,\mathrm{d}t \approx 3.014 \cdot 10^{-6} =: \tilde{\beta}_1^4. \tag{4.4}$$

Further, for any $k \in \mathbb{Z}$, we put

$$\varepsilon_k(t) := \frac{Z_k + B_k(t)}{\sqrt{1+t}} \quad \text{a.s.,} \quad t \in [0,1], \tag{4.5}$$

where $Z_k \sim \mathcal{N}(0,1), B_k = (B_k(t))_{t\in[0,1]}$ is a Wiener process, and $\ldots, Z_{-1}, B_{-1}, Z_0, B_0, Z_1, B_1, \ldots$ are independent random variables. Then, $(\varepsilon_k)_{k\in\mathbb{Z}}$ is an i.i.d., centered process with $\varepsilon_0(t) \sim \mathcal{N}(0,1)$ for all t, and thus $\mathbb{E}(\varepsilon_0^2(t)) = 1$ and $\mathbb{E}(\varepsilon_0^8(t)) = 105$ for all t. Moreover, Fubini's theorem yields

$$\mathbb{E}||\varepsilon_0^2||^4_{\mathscr{\dot{H}}} = \int_0^1 \mathbb{E}(\varepsilon_0^8(t))\,\mathrm{d}t = 105, \tag{4.6}$$

which implies (2.4). Hence, all initial assumptions in Theorem 2.1 are satisfied.
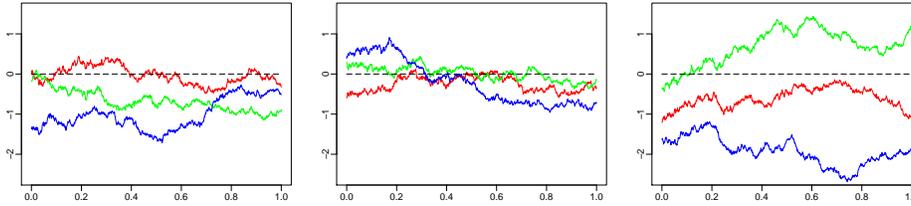
Figure 2: Three realizations of ε_1 (left), ε_2 (middle), ε_3 (right) in (4.5), with step width $1/1000$.

For the state-space form $\mathbf{\Psi}_1^{(1,0)} = \square_1 \alpha_1$ in (2.2), with δ, α_1 and innovations in (4.5), holds due to (2.5), (4.3), (4.6) and Jensen's inequality:

$$\psi_{1,2}^{(1,0)} := \mathbb{E}||\mathbf{\Psi}_1^{(1,0)}||^2_{\mathcal{L}_{\mathscr{H}}} \leq \sqrt{105}\,\tilde{\alpha}_1^2 < 1.$$

Thus, (2.10) is verified, and due to Theorem 2.1, a strictly stationary $\mathscr{H}$-valued ARCH(1) process $(\mathscr{X}_k)_k$ with the shift term δ and the operator α_1 above exists.

Now, we also take β_1 into our consideration. After Theorem 2.1, there also exists a strictly stationary $\mathscr{H}$-valued GARCH(1,1) process $(\mathscr{X}_k)_k$ with $\delta, \alpha_1, \beta_1$ and the innovations in (4.5), since the product of $\mathbf{\Psi}_2^{(1,1)}$ and $\mathbf{\Psi}_1^{(1,1)}$ in (2.2) fulfils

$$\mathbf{\Psi}_2^{(1,1)}\mathbf{\Psi}_1^{(1,1)} = \begin{bmatrix} \square_2\alpha_1 & \square_2\beta_1 \\ \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} \end{bmatrix}\begin{bmatrix} \square_1\alpha_1 & \square_1\beta_1 \\ \mathbb{I}_{\mathscr{H}} & 0_{\mathcal{L}_{\mathscr{H}}} \end{bmatrix} = \begin{bmatrix} \square_2\alpha_1\square_1\alpha_1 + \square_2\beta_1 & \square_2\alpha_1\square_1\beta_1 \\ \square_1\alpha_1 & \square_1\beta_1 \end{bmatrix},$$

and that due to (2.5), (4.3), (4.4), (4.6), and since $\square_1$ and $\square_2$ are i.i.d., holds

$$\psi_{2,2}^{(1,1)} := \mathbb{E}||\mathbf{\Psi}_2^{(1,1)}\mathbf{\Psi}_1^{(1,1)}||^2_{\mathcal{L}_{\mathscr{H}^2}}$$
$$\leq \mathbb{E}||\square_2\alpha_1\square_1\alpha_1 + \square_2\beta_1||^2_{\mathcal{L}_{\mathscr{H}}} + \mathbb{E}||\square_2\alpha_1\square_1\beta_1||^2_{\mathcal{L}_{\mathscr{H}}} + \mathbb{E}||\square_1\alpha_1||^2_{\mathcal{L}_{\mathscr{H}}} + \mathbb{E}||\square_1\beta_1||^2_{\mathcal{L}_{\mathscr{H}}}$$
$$\leq \sqrt{105}\big(2\big[\sqrt{105}\tilde{\alpha}_1^4 + \tilde{\beta}_1^2\big] + \sqrt{105}\,\tilde{\alpha}_1^2\tilde{\beta}_1^2 + \tilde{\alpha}_1^2 + \tilde{\beta}_1^2\big) < 1.$$

19

4.1   Simulation of Realizations

In order to simulate possible realizations of (G)ARCH processes under the setup above, we utilize the following result, ensuring that initial values of such processes can be approximated sufficiently well based on its recursion equations.

Corollary 4.1. *Let* (2.10) *hold for some* $n \in \mathbb{N}$ *and* $\nu > 0$. *Further, define* $\tilde{\boldsymbol{\varsigma}}_k^{(\mathfrak{p},\mathfrak{q})} := \boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q})} + \boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}(\tilde{\boldsymbol{\varsigma}}_{k-1}^{(\mathfrak{p},\mathfrak{q})})$ *for* $k \in \mathbb{N}$, *where* $\tilde{\boldsymbol{\varsigma}}_0^{(\mathfrak{p},\mathfrak{q})} \in F^{\mathfrak{s}}$ *is some deterministic value. Then, there is some* $\rho \in (0,1)$ *with*

$$\mathbb{E}||\boldsymbol{\varsigma}_N^{(\mathfrak{p},\mathfrak{q})} - \tilde{\boldsymbol{\varsigma}}_N^{(\mathfrak{p},\mathfrak{q})}||_{F^{\mathfrak{s}}}^{\nu} = \mathrm{O}(\rho^N). \tag{4.7}$$
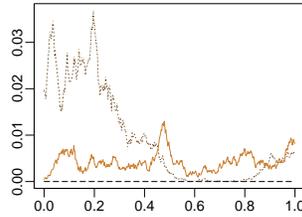
Figure 3: Initial values $\mathscr{X}_0^2$ (**black**, dotted) of an ARCH(1) process with δ in (4.1) and α_1 with kernel a_1 in (4.2), and $\mathscr{X}_0^2$ (**bronze**, dotted) and σ_0^2 (**bronze**, solid) of an GARCH$(1,1)$ process with δ in (4.1) and α_1, β_1 with kernels a_1, b_1 in (4.2), with step width $1/1000$.
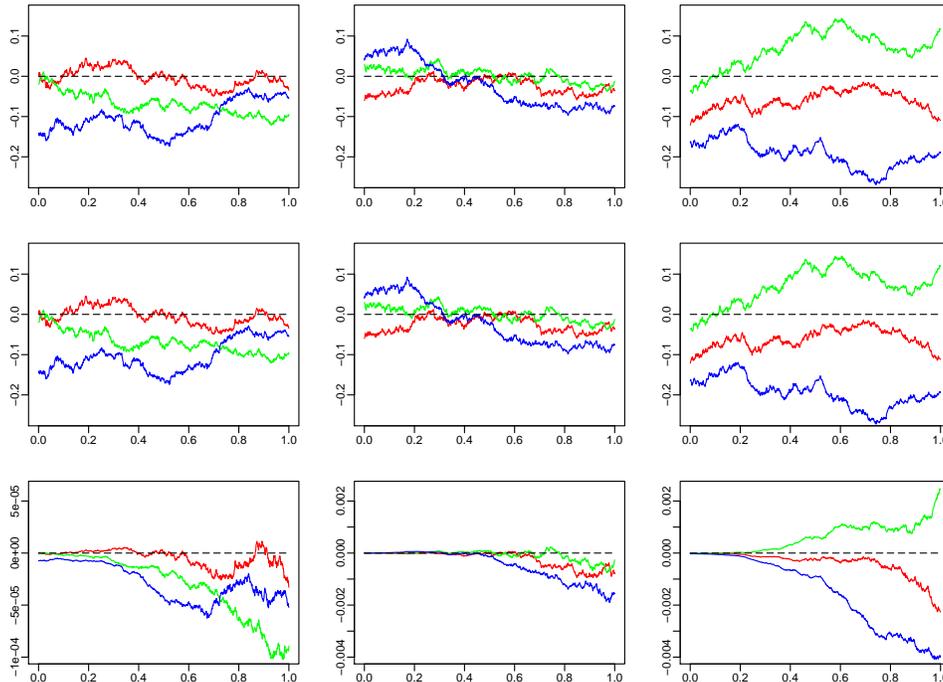
Figure 4: Based on $\varepsilon_1, \varepsilon_2, \varepsilon_3$ in Fig. 2, from left to right: Three simulations each of $\mathscr{X}_1, \mathscr{X}_2, \mathscr{X}_3$ derived from $\mathscr{X}_0^2$ in Fig. 3 of an ARCH(1) process (first row) with δ in (4.1) and α_1 with kernel a_1 in (4.2), of $\mathscr{X}_1, \mathscr{X}_2, \mathscr{X}_3$ derived from $\mathscr{X}_0^2, \sigma_0^2$ in Fig. 3 of an GARCH$(1,1)$ process (second row) with δ in (4.1) and α_1, β_1 with kernels a_1, b_1 in (4.2), and of the differences of the values in the second and the first row (in third row), with step width $1/1000$.

Remark 4.1. The realizations $\mathscr{X}_1, \mathscr{X}_2, \mathscr{X}_3$ of our ARCH(1) process in the first row in Fig. 4 have nearly the same structure as $\varepsilon_1, \varepsilon_2, \varepsilon_3$ in Fig. 2 on which they are based on, since δ is constant and α_1 has little impact. $\mathscr{X}_0^2$ in Fig. 3 as well as $\mathscr{X}_1, \mathscr{X}_2, \mathscr{X}_3$ of the ARCH(1) and GARCH(1,1) process in the first and second row in Fig. 4 are almost identical, because α_1 coincides in both processes and β_1 has very slim influence, see third row. In each plot of this row, the amplitudes increase w.r.t. t because $t \mapsto b_1(s,t)$ is non-negative and monotonically increasing for any s.

4.2   Simulation of Estimators

For the sake of clarity, we illustrate an estimator for the shift term δ in (4.1) and the operator α_1 with integral kernel a_1 in (4.2) in ARCH(1) for the sample sizes $N = 75, 200$. Hereto, we generate a sample $\mathscr{X}_1, ..., \mathscr{X}_N$ by simulating both an initial value $\mathscr{X}_0^2$ of our ARCH(1) process as in Section 4.1 and innovations $\varepsilon_1, ..., \varepsilon_N$ in (4.5), and apply the recurrence equation (2.1). In our calculations, any $x \in \mathscr{H} = L^2[0,1]$ is evaluated at $t = 0, \frac{1}{250}, ..., \frac{249}{250}$, and $\langle x, y \rangle_{\mathscr{H}}$ with $x, y \in \mathscr{H}$ is approximated by the Riemann sum $\frac{1}{250} \sum_{t=1}^{250} x(\frac{t-1}{250}) y(\frac{t-1}{250})$.

With $\mathfrak{p} = 1$ and $\vartheta_N := N^{-1/2}$ for any N, the estimator $\hat{\alpha}_{[1]} = \hat{\alpha}_1$ for α_1 is

$$\hat{\alpha}_1 = \sqrt{N}\hat{\mathscr{C}}_1\hat{\mathscr{C}}_0\big(\sqrt{N}\hat{\mathscr{C}}_0^2 + \mathbb{I}_{\mathscr{H}}\big)^{-1} \prod_{\hat{\mathfrak{c}}_1}^{\hat{\mathfrak{c}}_K}, \tag{4.8}$$

where $\hat{\mathscr{C}}_0$ and $\hat{\mathscr{C}}_1$ are the empirical covariance and lag-1-covariance operators in (3.6) based on the sample $\mathscr{X}_1, ..., \mathscr{X}_N$, and $\hat{\mathfrak{c}}_1, \hat{\mathfrak{c}}_2, \cdots, \hat{\mathfrak{c}}_K$ are the eigenfunctions with associated eigenvalues $\hat{c}_1 \geq \hat{c}_2 \geq \cdots \geq \hat{c}_K \geq 0$ of $\hat{\mathscr{C}}_0$. As suggested in Remark 3.2 (c), we choose $K = K_N$ as the first integer so that $\sum_{i=1}^{K} \hat{c}_i / \sum_{i=1}^{N} \hat{c}_i \geq (1-w)$, where we put $w = 0.001$. The estimator for δ with $\hat{\alpha}_1$ in (4.8) and $\hat{m}_2 = N^{-1} \sum_{i=1}^{N} \mathscr{X}_i^2$ has the form

$$\hat{\delta} = (\mathbb{I}_{\mathscr{H}} - \hat{\alpha}_1)(\hat{m}_2). \tag{4.9}$$
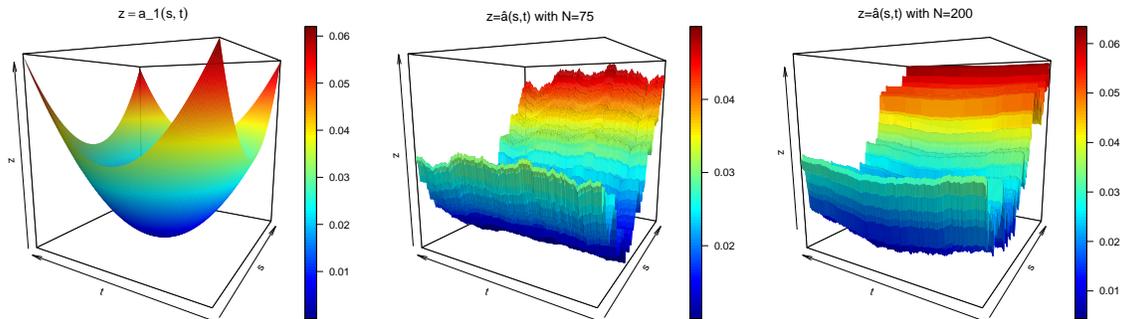
Figure 5: The integral kernel a_1 in (4.2) of α_1 and the integral kernels $\hat{a}$ derived from the estimator $\hat{\alpha}_1$ in (4.8) with $N = 75$ and $N = 200$, where the step width is $1/250$.
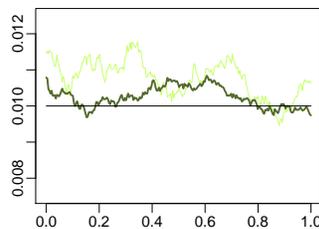
Figure 6: δ (**black**), and $\hat{\delta}$ in (4.9) for $N = 75$ (**light green**) and $N = 200$ (**green**) with step width $1/250$.

5   Conclusions

This article studies ARCH and GARCH processes in established and not yet considered function spaces, and provides explicit asymptotic upper bounds of the estimation errors for all parameters. By a Yule-Walker approach, ARCH operators projected on a finite-dimensional subspace are estimated, as are complete (G)ARCH operators using an additional Sobolev condition. Our theories contribute to Hörmann *et al.* [14], Aue *et al.* [2] and Cerovecki *et al.* [6] who established ARCH(1) in $C[0,1]$ and $\mathscr{H} = L^2[0,1]$, GARCH(1,1) in $C[0,1]$ and $\mathscr{H}$, resp. GARCH processes for any order in $\mathscr{H}$.

Section 2 introduces ARCH($\mathfrak{p}$) and GARCH($\mathfrak{p},\mathfrak{q}$) processes for any order $\mathfrak{p},\mathfrak{q} \in \mathbb{N}$ in $L^p[0,1]$ with $p \in [1,\infty), C[0,1]$ and more abstract spaces. For these processes, we present sufficient conditions for the existence of strictly stationary solutions, finite moments and weak dependence. Section 3 establishes explicit asymptotic upper bounds of the estimation errors for the shift term δ and the operators of $\mathscr{H}$-valued ARCH and GARCH processes for any order using upper bounds of estimation errors for certain moments, (lag-h-)covariance operators, eigenfunctions and eigenvalues which are also useful beyond (G)ARCH. Theorem 3.1 deduces upper bounds for δ which is new in ARCH($\mathfrak{p}$) for $\mathfrak{p} > 1$ and GARCH. The main results of the article are stated in Theorems 3.2-3.3. Theorem 3.2 (a) provides explicit upper bounds for the projections of the ARCH($\mathfrak{p}$) operators on a finite-dimensional subspace for any $\mathfrak{p} \in \mathbb{N}$ as in [14] for $\mathfrak{p} = 1$, and part (b) and Theorem 3.3 for the complete ARCH and GARCH operators for any order, respectively, being new as far as we are aware. The upper bounds for GARCH operators depend on these for operators of invertible linear represented as inverted processes, see Lemma 3.5, which also hold for the operators of associated linear processes. Both results are valid beyond (G)ARCH, extending theories in [3], [20]. All assertions regarding (G)ARCH in Section 3 can by minor modifications be transferred to AR(MA), and Sections 2-3 possibly can be carried out on further Banach and Hilbert spaces. Section 4 simulates some of our processes and estimators. Section 6 contains all proofs.

We leave the investigations concerning probabilistic properties of (G)ARCH in general separable Banach spaces behind for future research, as we do for order estimation, see [21]. Concerning the parameters, unsolved problems are their estimation in Banach spaces (see [30]), the asymptotic distribution of their estimations errors (see [2], [6] for the parameters projected on a finite-dimensional subspace), and the asymptotic lower bounds of their estimations errors.

6   Proofs

In various conversions, we utilize for any $n \in \mathbb{N}$ and $a_1, ..., a_n \geq 0$,

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

thus $(\sum_{k=1}^{n} a_k)^{\nu} \precsim \sum_{k=1}^{n} a_k^{\nu}$, and the operator valued Hölder's inequality which we state in the following. Let $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}}), (\mathcal{H}', \langle\cdot,\cdot\rangle_{\mathcal{H}'})$ and $(\mathcal{H}'', \langle\cdot,\cdot\rangle_{\mathcal{H}''})$ be Hilbert spaces. For $p \in [1,\infty], \mathscr{S}_{\mathcal{H},\mathcal{H}'}^{p} \subseteq \mathcal{L}_{\mathcal{H},\mathcal{H}'}$ is the p-*th Schatten-class*, with $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{1} = \mathcal{N}_{\mathcal{H},\mathcal{H}'}, \mathscr{S}_{\mathcal{H},\mathcal{H}'}^{2} = \mathcal{S}_{\mathcal{H},\mathcal{H}'}, \mathscr{S}_{\mathcal{H},\mathcal{H}'}^{\infty} := \mathcal{L}_{\mathcal{H},\mathcal{H}'}$. 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 $s_j(A)$ is the j-th singular value of $A \in \mathcal{L}_{\mathcal{H},\mathcal{H}'}$, and $\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{p}$ is endowed with the norm $||\cdot||_{\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{p}} := (\sum_{j=1}^{\infty} s_j^{p}(\cdot))^{1/p}$. Now, let $p,q,r \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ where $\frac{1}{\infty} := 0, A \in \mathscr{S}_{\mathcal{H}',\mathcal{H}''}^{q}$ and $B \in \mathscr{S}_{\mathcal{H},\mathcal{H}'}^{p}$. Then, after the *operator valued Hölder's inequality* (see [17], Theorem 11.2), $AB \in \mathscr{S}_{\mathcal{H},\mathcal{H}''}^{r}$ with

$$||AB||_{\mathscr{S}_{\mathcal{H},\mathcal{H}''}^{r}} \leq 2^{1/r} ||A||_{\mathscr{S}_{\mathcal{H}',\mathcal{H}''}^{q}} ||B||_{\mathscr{S}_{\mathcal{H},\mathcal{H}'}^{p}}. \tag{6.1}$$

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

$$\varsigma_k^{(\mathfrak{p},\mathfrak{q})} = \delta_k^{(\mathfrak{p},\mathfrak{q})} + \sum_{m=1}^{\infty} \Psi_k^{(\mathfrak{p},\mathfrak{q})} \Psi_{k-1}^{(\mathfrak{p},\mathfrak{q})} \cdots \Psi_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\delta_{k-m}^{(\mathfrak{p},\mathfrak{q})}) \tag{6.2}$$

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

$$\varlimsup_{m\to\infty} \frac{1}{m} \ln\big|\big|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})})\big|\big|_{F^{\mathfrak{s}}} \le \gamma^{(\mathfrak{p},\mathfrak{q})} + \varlimsup_{m\to\infty}\frac{1}{m}\ln\|\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}$$

where $\mathfrak{s} = \mathfrak{p} + \mathfrak{q}$. By definition of $\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})}$ and because of $0 < ||\delta||_F \le ||\delta||_{\dot F}$, we have $\|\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}^2 \le ||\delta||_{\dot F}^2(1+$ $||\varepsilon_{k-m}^2||_{\dot F}^2)$, and thus, due to (2.4), $\mathbb{E}(1 \vee \|\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}) < \infty$. As a consequence, since $\gamma^{(\mathfrak{p},\mathfrak{q})} < 0$,

$$\varlimsup_{m\to\infty}\big|\big|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})})\big|\big|_{F^{\mathfrak{s}}}^{1/m} = e^{\gamma^{(\mathfrak{p},\mathfrak{q})}} < 1 \quad \text{a.s.}$$

Thus the series in (6.2) converges a.s. by Cauchy rule. Hence, there exists a solution $(\boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})})_k$ of (2.2), its $(\mathfrak{p}+1)$-th component defines a solution of (2.1) if $\mathfrak{q} \in \mathbb{N}$, and from its first component one can deduce a solution of (2.1) if $\mathfrak{q} = 0$.

For verifying that the solution is a.s. unique, we use the argumentation in [6], p.19. We assume that $(\boldsymbol{\acute\varsigma}_k^{(\mathfrak{p},\mathfrak{q})})_k$ is another solution. Then,

$$\boldsymbol{\acute\varsigma}_k^{(\mathfrak{p},\mathfrak{q})} = \boldsymbol{\varsigma}_{k,N}^{(\mathfrak{p},\mathfrak{q})} + \boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-N}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\acute\varsigma}_{k-N-1}^{(\mathfrak{p},\mathfrak{q})}),$$

$$\text{where} \quad \boldsymbol{\varsigma}_{k,N}^{(\mathfrak{p},\mathfrak{q})} := \boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q})} + \sum_{m=1}^{N}\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-m+1}^{(\mathfrak{p},\mathfrak{q})}(\boldsymbol{\delta}_{k-m}^{(\mathfrak{p},\mathfrak{q})}).$$

Since $\gamma^{(\mathfrak{p},\mathfrak{q})} < 0$, both $\|\boldsymbol{\varsigma}_{k,N}^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}} \to 0$ and $\|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-N}^{(\mathfrak{p},\mathfrak{q})}\|_{\mathcal{L}_{F^{\mathfrak{s}}}} \to 0$ a.s. for $N \to \infty$, and the law of $\|\boldsymbol{\acute\varsigma}_{k-N-1}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}$ is independent of N. Ergo,

$$\big|\big|\boldsymbol{\acute\varsigma}_k^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{F^{\mathfrak{s}}} \le \big|\big|\boldsymbol{\varsigma}_{k,N}^{(\mathfrak{p},\mathfrak{q})} - \boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{F^{\mathfrak{s}}} + \big|\big|\boldsymbol{\Psi}_k^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{k-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{k-N}^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^{\mathfrak{s}}}}\big|\big|\boldsymbol{\acute\varsigma}_{k-N-1}^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{F^{\mathfrak{s}}} \overset{N\to\infty}{\longrightarrow} 0,$$

by which is shown that the solution is almost surely unique.

By definition of $\boldsymbol{\varsigma}_k^{(\mathfrak{p},\mathfrak{q})}$ and due to (6.2), $\sigma_k^2 = f(\varepsilon_{k-1}, \varepsilon_{k-2}, \ldots)$ a.s. for all k for some measurable function $f\colon F^\infty \to F$, so $(\sigma_k^2)_k$ and $(\mathscr{X}_k)_k$ are non-anticipative w.r.t. $(\varepsilon_k)_k$ and strictly stationary as well as ergodic after [32], Theorem 3.5.8.

(b) The assertion follows from the fact that $\psi_{n,\nu}^{(\mathfrak{p},\mathfrak{q})} < 1$, (2.10), sub-multiplicativity of $||\cdot||_{\mathcal{L}_{F^{\mathfrak{s}}}}$ and Jensen's inequality imply

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

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

$$\|\boldsymbol{\varsigma}_0^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}^\nu \precsim \|\boldsymbol{\delta}_0^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}^\nu + \Big(\sum_{m=1}^\infty\big|\big|\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^{\mathfrak{s}}}}\|\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}\Big)^\nu$$

$$\precsim \|\boldsymbol{\delta}_0^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}^\nu + \sum_{m=1}^{n-1}\|\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}^\nu\prod_{l=1}^m\|\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}\|_{\mathcal{L}_{F^{\mathfrak{s}}}}^\nu + \Big(\sum_{m=n}^\infty\big|\big|\boldsymbol{\Psi}_0^{(\mathfrak{p},\mathfrak{q})}\boldsymbol{\Psi}_{-1}^{(\mathfrak{p},\mathfrak{q})}\cdots\boldsymbol{\Psi}_{-m+1}^{(\mathfrak{p},\mathfrak{q})}\big|\big|_{\mathcal{L}_{F^{\mathfrak{s}}}}\|\boldsymbol{\delta}_{-m}^{(\mathfrak{p},\mathfrak{q})}\|_{F^{\mathfrak{s}}}\Big)^\nu.$$

Moreover, $\mathbb{E}\|\varepsilon_0^2\|_{\dot F}^\nu < \infty$ 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 $\boldsymbol{\delta}_k^{(\mathfrak{p},\mathfrak{q})}$ 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 $\nu \in (0,1]$:

$$\mathbb{E}\Big(\sum_{m=n}^{\infty}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\mathbf{\Psi}_{-1}^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{-m+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}\big|\big|\boldsymbol{\delta}_{-m}^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}\Big)^{\nu} \leq \mathbb{E}\big|\big|\boldsymbol{\delta}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu}\sum_{m=n}^{\infty}\mathbb{E}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\mathbf{\Psi}_{-1}^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{-m+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}$$

$$\leq \mathbb{E}\big|\big|\boldsymbol{\delta}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu}\Big(\sum_{k=0}^{n-1}\big(\mathbb{E}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}\big)^k\Big)\sum_{l=1}^{\infty}(\psi_{n,\nu}^{(\mathbf{p},\mathbf{q})})^l < \infty,$$

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

$$\mathbb{E}\Big(\sum_{m=n}^{\infty}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\mathbf{\Psi}_{-1}^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{-m+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}\big|\big|\boldsymbol{\delta}_{-m}^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}\Big)^{\nu} \leq \mathbb{E}\big|\big|\boldsymbol{\delta}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu}\Big(\sum_{m=n}^{\infty}\big(\mathbb{E}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\mathbf{\Psi}_{-1}^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{-m+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}\big)^{1/\nu}\Big)^{\nu}$$

$$\leq \mathbb{E}\big|\big|\boldsymbol{\delta}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu}\Big(\sum_{k=0}^{n-1}\big(\mathbb{E}\big|\big|\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}\big)^k\Big)\Big(\sum_{l=1}^{\infty}(\psi_{n,\nu}^{(\mathbf{p},\mathbf{q})})^l\Big)^{\nu} < \infty.$$

Hence, $\mathbb{E}||\varsigma_0^{(\mathbf{p},\mathbf{q})}||_{F^\mathfrak{s}}^{\nu} < \infty$. Thus $\mathbb{E}||\mathscr{X}_0^2||_F^{\nu} < \infty$, $\mathbb{E}||\sigma_0^2||_F^{\nu} < \infty$ and

$$\mathbb{E}||\sigma_0^2||_F^{\nu} \precsim ||\delta||_{\dot{F}}^{\nu} + \sum_{i=1}^{\mathbf{p}}\mathbb{E}||\alpha_i(\mathscr{X}_i^2)||_{\dot{F}}^{\nu} + \sum_{j=1}^{\mathbf{q}}\mathbb{E}||\beta_j(\sigma_j^2)||_{\dot{F}}^{\nu}$$

$$\leq ||\delta||_{\dot{F}}^{\nu} + \mathbb{E}||\mathscr{X}_0^2||_F^{\nu}\sum_{i=1}^{\mathbf{p}}||\alpha_i||_{\mathcal{L}_{F,\dot{F}}}^{\nu} + \mathbb{E}||\sigma_0^2||_F^{\nu}\sum_{j=1}^{\mathbf{q}}||\beta_j||_{\mathcal{L}_{F,\dot{F}}}^{\nu} < \infty.$$

(b) From the identity (6.2) follows

$$\varsigma_{m,m}^{(\mathbf{p},\mathbf{q})} := \boldsymbol{\delta}_m^{(\mathbf{p},\mathbf{q})} + \sum_{l=1}^{m-1}\mathbf{\Psi}_m^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{m-l+1}^{(\mathbf{p},\mathbf{q})}(\boldsymbol{\delta}_{m-l}^{(\mathbf{p},\mathbf{q})}) + \sum_{l=m}^{\infty}\mathbf{\Psi}_m^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_1^{(\mathbf{p},\mathbf{q})}\mathbf{\Psi}_0^{(\mathbf{p},\mathbf{q},m)}\cdots\mathbf{\Psi}_{m-l+1}^{(\mathbf{p},\mathbf{q},m)}(\boldsymbol{\delta}_{m-l}^{(\mathbf{p},\mathbf{q},m)})$$

a.s. for all $m \in \mathbb{N}$. Thereby, $\mathbf{\Psi}_k^{(\mathbf{p},\mathbf{q},m)}$ and $\boldsymbol{\delta}_k^{(\mathbf{p},\mathbf{q},m)}$ stand for $\mathbf{\Psi}_k^{(\mathbf{p},\mathbf{q})}$ resp. $\boldsymbol{\delta}_k^{(\mathbf{p},\mathbf{q})}$ in (2.2) depending on $\varepsilon_k^{(m)}$ where $(\varepsilon_k^{(m)})_{k \in \mathbb{Z}}$ are i.i.d. processes for all m, which are independent of each other, and $\varepsilon_k^{(m)}$ equals ε_0 in distribution for all k, m. Consequently, for any m,

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

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

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 $\nu \in (0,1]$:

$$\mathbb{E}\big|\big|\varsigma_m^{(\mathbf{p},\mathbf{q})} - \varsigma_{m,m}^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu} \leq 2\,\mathbb{E}\big|\big|\boldsymbol{\delta}_0^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu}\sum_{l=m}^{\infty}\mathbb{E}\big|\big|\mathbf{\Psi}_m^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{m-l+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}$$

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

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

$$\mathbb{E}\big|\big|\varsigma_m^{(\mathbf{p},\mathbf{q})} - \varsigma_{m,m}^{(\mathbf{p},\mathbf{q})}\big|\big|_{F^\mathfrak{s}}^{\nu} \leq \Big(\sum_{m=n}^{\infty}2\big(\mathbb{E}\big|\big|\mathbf{\Psi}_m^{(\mathbf{p},\mathbf{q})}\cdots\mathbf{\Psi}_{m-l+1}^{(\mathbf{p},\mathbf{q})}\big|\big|_{\mathcal{L}_{F^\mathfrak{s}}}^{\nu}\big)^{1/\nu}\Big)^{\nu}$$

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

Proof of Lemmas 3.1-3.4. The proofs of Lemmas 3.1-3.4, which are auxiliary results in this work, can be found in [23], Lemma 4.11, Theorem 4.2, Theorem 4.4 and Lemma 4.21 (b), respectively. Thereby, in the proof of [23], Lemma 4.21 (b), k_m^{-1} has to be substituted by γ_m. $\qquad\square$

Proof of Corollary 3.1. (3.14) is an immediate consequence of (3.13), and (3.15) follows from the triangle inquality, (3.13) and $k_{b_N} = \Xi[\sqrt{a_N}, 1]$. $\qquad\square$

Proof of Theorem 3.1. The assertion follows from (3.4) 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 (6.1), $\mathfrak{S}_{\mathfrak{p}}^{\dagger} := \mathfrak{S}_{\mathfrak{p}}(\mathfrak{S}_{\mathfrak{p}}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1}$, 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 $\mathscr{J}_{\mathfrak{p},K}^c := \{\Phi_{\mathfrak{p},ij} | i, j \in \mathbb{N}, i \vee j > K\}$, yield for fixed $K \in \mathbb{N}$:

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

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

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

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

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

$$=: O_{\mathbb{P}}(N^{-1})\cdot T_1 + T_2 + T_3 + T_4. \tag{6.3}$$

<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

$$||\hat{\mathfrak{S}}_{\mathfrak{p}}^{\dagger}\coprod_{\hat{\mathfrak{c}}_{\mathfrak{p},1}}^{\hat{\mathfrak{c}}_{\mathfrak{p},K}}||_{\mathcal{L}_{\mathscr{H}^{\mathfrak{p}}}}^2 = \sup_{j \leq K}\Big(\frac{\hat{c}_{\mathfrak{p},j}}{\hat{c}_{\mathfrak{p},j} + \vartheta_N}\Big)^2 = O_{\mathbb{P}}(c_{\mathfrak{p},K}^{-2}) = O_{\mathbb{P}}(1). \tag{6.4}$$

<u>Term T_2</u>: 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, (3.22) and Lemma 3.4 yield

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

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

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

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

$$= O_{\mathbb{P}}(c_{\mathfrak{p},K}^{-4} N^{-1}) + O_{\mathbb{P}}(c_{\mathfrak{p},K}^{-2} \gamma_{\mathfrak{p},K}^2 K N^{-1}) = O_{\mathbb{P}}(N^{-1}). \tag{6.5}$$

<u>Term T_3</u>: The eigenfunction sequences $(\mathfrak{c}_{\mathfrak{p},j})_j$ of $\mathfrak{S}_{\mathfrak{p}}, (\mathfrak{c}_j)_j$ of $\mathscr{C}_{0;\boldsymbol{\mathscr{X}}}$ and also 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 $\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}$. Furthermore, because of $\mathfrak{S}_{\mathfrak{p}}^{\ddagger} := \mathfrak{S}_{\mathfrak{p}}^2(\mathfrak{S}_{\mathfrak{p}}^2 + \vartheta_N \mathbb{I}_{\mathscr{H}^{\mathfrak{p}}})^{-1}$, $\mathscr{J}_{\mathfrak{p},K}^c := \{\Phi_{\mathfrak{p},ij} | i,j \in \mathbb{N}, i \vee j > K\}$, $\boldsymbol{\alpha}_{[\mathfrak{p}]} = \sum_{i,j=1}^{\infty} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}, \Phi_{\mathfrak{p},ij} \rangle_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}} \Phi_{\mathfrak{p},ij}$ as well as $\langle \boldsymbol{\alpha}_{[\mathfrak{p}]}, \Phi_{\mathfrak{p},ij} \rangle_{\mathcal{S}_{\mathscr{H}^{\mathfrak{p}},\mathscr{H}}} = \sum_{k=1}^{\infty} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},k}), \Phi_{\mathfrak{p},ij}(\mathfrak{c}_{\mathfrak{p},k}) \rangle_{\mathscr{H}} = \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},i}), \mathfrak{c}_j \rangle_{\mathscr{H}}$ 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,l with $l \leq K < j$,

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

$$= \sum_{l=1}^{K} \Big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2 + \vartheta_N}\Big)^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = 0. \tag{6.6}$$

<u>Term T_4</u>: Elementary transformations as in 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 $\vartheta_N = O(N^{-1/2})$ lead to

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

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

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

$$= O(N^{-1}). \tag{6.7}$$

Replacing T_1-T_4 in (6.3) by (6.4)-(6.7) indeed yields (3.34).

(b) stems from $\sum_{l=1}^{K} \big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2 + \vartheta_N}\big)^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2 = O(K^{-2\beta})$, $c_{\mathfrak{p},K}^{-2} \gamma_{\mathfrak{p},K}^2 K^{2\beta+1} = O(N)$ for any $K = K_N = \Xi(1, N)$ and $\vartheta_N = O(N^{-1/2})$, and the fact that $\boldsymbol{\alpha}_{[\mathfrak{p}]} = \coprod_{\mathscr{J}_{\mathfrak{p},K}} \boldsymbol{\alpha}_{[\mathfrak{p}]} + \coprod_{\mathscr{J}_{\mathfrak{p},K}^c} \boldsymbol{\alpha}_{[\mathfrak{p}]}$, (3.22), part (a) and (3.28) imply

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

$$\precsim O_{\mathbb{P}}(c_{\mathfrak{p},K}^{-2} \gamma_{\mathfrak{p},K}^2 K N^{-1}) + \sum_{l=1}^{K} \Big(\frac{c_{\mathfrak{p},l}^2}{c_{\mathfrak{p},l}^2 + \vartheta_N}\Big)^2 \sum_{j>K} \langle \boldsymbol{\alpha}_{[\mathfrak{p}]}(\mathfrak{c}_{\mathfrak{p},l}), \mathfrak{c}_j \rangle_{\mathscr{H}}^2$$

$$+ O(\vartheta_N^2 c_{\mathfrak{p},K}^{-4}) + O(K^{-2\beta}). \qquad \square$$

Proof of Lemma 3.5. The proof is based on proof ideas of Theorem 3.2 with $\mathfrak{p}$ replaced by an appropriate sequence $L = L_N \to \infty$. From these ideas and also (see [23], Lemma 4.45)

$$\mathfrak{S}_{L,1} = \boldsymbol{\pi}_L \mathfrak{S}_L + \sum_{l>L} \pi_l \mathfrak{S}_{L,1-l}, \tag{6.8}$$

which can be identified as a Yule-Walker equation with a residual, follows

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

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

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

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

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

$$=: \mathrm{O}_{\mathbb{P}}(L^2N^{-1}) \cdot T_1 + L \cdot T_2 + T_3 + T_4 + T_5 + \mathrm{O}(K^{-2\beta}). \tag{6.9}$$

<u>Term T_1</u>: Similar conversions as in the proof of Theorem 3.2 imply

$$||\hat{\mathfrak{S}}_L^\dagger \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}}||^2_{\mathcal{L}_{\mathcal{H}^L}} = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-2}). \tag{6.10}$$

<u>Term T_2</u>: (3.12) with $L = \Xi(1, N^{1/3})$, $\gamma_{L,K} = \Omega(c_{L,K}^{-1})$ and arguments in the proof of Theorem 3.2 yield

$$||\hat{\mathfrak{S}}_L^\dagger \prod_{\hat{\mathfrak{c}}_{L,1}}^{\hat{\mathfrak{c}}_{L,K}} - \mathfrak{S}_L^\dagger \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||^2_{\mathcal{L}_{\mathcal{H}^L}} = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-4}L^3N^{-1}) + \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-2}\gamma_{L,K}^2 KL^3N^{-1}) = \mathrm{O}_{\mathbb{P}}(c_{L,K}^{-2}\gamma_{L,K}^2 KL^3N^{-1}). \tag{6.11}$$

<u>Term T_3</u>: The inequality (6.1), (6.10), the triangle inequality and (3.8) imply

$$||[\sum_{l>L}\boldsymbol{\pi}_l \mathfrak{S}_{L,1-l}]\mathfrak{S}_L^\dagger \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} \le c_{L,K}^{-2} ||\sum_{l>L}\boldsymbol{\pi}_l \mathfrak{S}_{L,1-l}||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} = \mathrm{O}(c_{L,K}^{-2}L(\sum_{l>L}||\pi_l||_{\mathcal{L}_{\mathcal{H}}})^2). \tag{6.12}$$

<u>Term T_4</u>: Almost one to one as in the proof of Theorem 3.2, we obtain

$$||\prod_{\mathscr{J}_{L,K}^c} \boldsymbol{\pi}_L \mathfrak{S}_L^\ddagger \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}}||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} = \sum_{l=1}^{K} \left(\frac{c_{L,l}^2}{c_{L,l}^2 + \vartheta_N}\right)^2 \sum_{j>K} \langle \boldsymbol{\pi}_L(\mathfrak{c}_{L,l}), \mathfrak{c}_j \rangle^2_{\mathcal{H}} = \mathrm{O}(K^{-2\beta}). \tag{6.13}$$

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

$$||\prod_{\mathscr{J}_{L,K}} \boldsymbol{\pi}_L [\mathfrak{S}_L^\ddagger \prod_{\mathfrak{c}_{L,1}}^{\mathfrak{c}_{L,K}} - \mathbb{I}_{\mathcal{H}^L}]||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} \le \vartheta_N^2 c_{L,K}^{-4} ||\boldsymbol{\pi}_L||^2_{\mathcal{S}_{\mathcal{H}^L,\mathcal{H}}} = \mathrm{O}(c_{L,K}^{-4}\vartheta_N^2). \tag{6.14}$$

Due to $L = \Xi(1, N), K = \Xi(1, \sqrt[3]{L^{-1}N})$, $c_{L,K} = \Omega(\sqrt{L^3N^{-1}})$, (3.28) and $\vartheta_N = \mathrm{O}(c_{L,K}^2 K^{-\beta})$, and since $(\boldsymbol{\pi}_L, (\Phi_{L,ij})_{i,j})$ satisfies Assumption 3.4 for all L for $\beta > 0$, plugging (6.10)-(6.14) of T_1-T_5 into (6.9) implies

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

Hence, since we imposed $c_{L,K}^{-2}L(\gamma_{L,K}^2 KL^3N^{-1} + (\sum_{l>L}||\pi_l||_{\mathcal{L}_{\mathcal{H}}})^2) = \mathrm{O}(K^{-2\beta})$, (3.44) is verified.   □

Proof of Proposition 3.1. We show the assertion by induction. $\pi_1 = \psi_1, \hat{\pi}_1 = \hat{\psi}_1$ and Lemma 3.5 yield indeed (3.50). Assume (3.50) is true for $l = 1, ..., i$ for some particular $i \in \mathbb{N}$. Then, triangle inequality and sub-multiplicativity imply

$$||\hat{\psi}_{i+1} - \psi_{i+1}||^2_{\mathcal{S}_{\mathcal{H}}} \precsim \sum_{j=1}^{i+1} ||\hat{\pi}_j||^2_{\mathcal{S}_{\mathcal{H}}}||\hat{\psi}_{i+1-j} - \psi_{i+1-j}||^2_{\mathcal{S}_{\mathcal{H}}} + ||\hat{\pi}_j - \pi_j||^2_{\mathcal{S}_{\mathcal{H}}}||\psi_{i+1-j}||^2_{\mathcal{S}_{\mathcal{H}}} = \mathrm{O}_{\mathbb{P}}(K^{-2\beta}).$$

Consequently, (3.50) is verified for all $i \in \mathbb{N}$.   □

27

Proof of Theorem 3.3. (3.51) and since $(\boldsymbol{\beta}_{[\mathfrak{q}]}, (\Phi_{\mathfrak{p},\mathfrak{q};ij})_{i,j})$ satisfies Assumption 3.4 for $\beta > 0$, yields as in the proof of Theorem 3.2:

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

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

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

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

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

$$=: \mathrm{O}_{\mathbb{P}}(K^{-2\beta}) \cdot T_1 + T_2 + T_3 + T_4 + \mathrm{O}(M^{-2\beta}). \tag{6.15}$$

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

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

$$= \sup_{j \leq M} \frac{\hat{g}_{\mathfrak{s},\mathfrak{q};j}}{(\hat{g}_{\mathfrak{s},\mathfrak{q};j} + \theta_N)^2} = \mathrm{O}_{\mathbb{P}}(g^{-1}_{\mathfrak{s},\mathfrak{q};M}). \tag{6.16}$$

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

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

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

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

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

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

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

$$= \mathrm{O}_{\mathbb{P}}(g^{-2}_{\mathfrak{s},\mathfrak{q};M}K^{-\beta}) + \mathrm{O}_{\mathbb{P}}(g^{-1}_{\mathfrak{s},\mathfrak{q};M}\gamma^2_{\mathfrak{s},\mathfrak{q};M}K^{-2\beta}M) + \mathrm{O}_{\mathbb{P}}(g^{-2}_{\mathfrak{s},\mathfrak{q};M}K^{-\beta})$$

$$+ \mathrm{O}_{\mathbb{P}}(g^{-2}_{\mathfrak{s},\mathfrak{q};M}\gamma^2_{\mathfrak{s},\mathfrak{q};M}K^{-2\beta}) + \mathrm{O}_{\mathbb{P}}(g^{-2}_{\mathfrak{s},\mathfrak{q};M}K^{-2\beta})$$

$$= \mathrm{O}_{\mathbb{P}}(g^{-2}_{\mathfrak{s},\mathfrak{q};M}K^{-\beta}). \tag{6.17}$$

In the last step, we utilized $\gamma_{\mathfrak{s},\mathfrak{q};M} = \Omega(g^{-1}_{\mathfrak{s},\mathfrak{q};M})$, the assumption $\gamma_{\mathfrak{s},\mathfrak{q};M}K^{-\beta} = \mathrm{o}(1)$ and also $g_{\mathfrak{s},\mathfrak{q};M} = \mathrm{o}(M^{-1})$ where latter holds since the eigenvalue sequence $(g_{\mathfrak{s},\mathfrak{q};j})_j$ is absolutely summable.

<u>Term T_3</u>: From our assumptions and the ideas in Theorem 3.2 follows

$$\Big|\Big| \coprod_{\mathscr{J}^c_{\mathfrak{p},\mathfrak{q};M}} \boldsymbol{\beta}_{[\mathfrak{q}]} \prod_{[\mathfrak{s},\mathfrak{q}]}^\dagger \coprod_{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};1}}^{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};M}} \Big|\Big|^2_{\mathcal{S}_{\mathscr{H}^\mathfrak{q},\mathscr{H}}} = \sum_{l=1}^M \Big(\frac{g^2_{\mathfrak{s},\mathfrak{q};l}}{g^2_{\mathfrak{s},\mathfrak{q};l} + \theta_N}\Big)^2 \sum_{j>M} \langle \boldsymbol{\beta}_{[\mathfrak{q}]}(\mathfrak{g}_{\mathfrak{s},\mathfrak{q};l}), \mathfrak{c}_j \rangle^2_{\mathscr{H}} = \mathrm{O}(M^{-2\beta}). \qquad (6.18)$$

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

$$\Big|\Big| \coprod_{\mathscr{J}_{\mathfrak{p},\mathfrak{q};M}} \boldsymbol{\beta}_{[\mathfrak{q}]} \Big[\prod_{[\mathfrak{s},\mathfrak{q}]}^\ddagger \coprod_{\mathfrak{g}_{\mathfrak{p},\mathfrak{q};1}}^{\mathfrak{g}_{\mathfrak{s},\mathfrak{q};M}} - \mathbb{I}_{\mathscr{H}^\mathfrak{q}} \Big] \Big|\Big|^2_{\mathcal{S}_{\mathscr{H}^\mathfrak{q},\mathscr{H}}} = \mathrm{O}(\theta_N^2 g^{-2}_{\mathfrak{s},\mathfrak{q};M}). \qquad (6.19)$$

Replacing T_1-T_4 in (6.15) by (6.16)-(6.19), and considering $\theta_N = \mathrm{O}(K^{-\beta/2}), g^{-2}_{\mathfrak{s},\mathfrak{q};M}K^{-\beta} = \mathrm{O}(M^{-2\beta})$, yields

$$||\hat{\boldsymbol{\beta}}_{[\mathfrak{q}]} - \boldsymbol{\beta}_{[\mathfrak{q}]}||^2_{\mathcal{S}_{\mathscr{H}^\mathfrak{q},\mathscr{H}}} = \mathrm{O}_{\mathbb{P}}(M^{-2\beta}).$$

Hence, (3.57) is shown for all j, and (3.56) for $i = 1, ..., \mathfrak{p}$ follows from (3.44) for all i, (3.57) for all j, and

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

Proof of Corollary 4.1. Assume $N = mn$ for some $m \in \mathbb{N}$ w.l.o.g. Then, since $\psi^{(\mathfrak{p},\mathfrak{q})}_{n,\nu} < 1$ and $(\boldsymbol{\Psi}^{(\mathfrak{p},\mathfrak{q})}_k)_{k\in\mathbb{Z}}$ is i.i.d., the assertion follows from

$$\mathbb{E}||\boldsymbol{\varsigma}^{(\mathfrak{p},\mathfrak{q})}_N - \tilde{\boldsymbol{\varsigma}}^{(\mathfrak{p},\mathfrak{q})}_N||^\nu_{F^\mathfrak{s}} = \mathbb{E}\big|\big| \boldsymbol{\Psi}^{(\mathfrak{p},\mathfrak{q})}_N \boldsymbol{\Psi}^{(\mathfrak{p},\mathfrak{q})}_{N-1} \cdots \boldsymbol{\Psi}^{(\mathfrak{p},\mathfrak{q})}_1 (\boldsymbol{\varsigma}^{(\mathfrak{p},\mathfrak{q})}_0 - \tilde{\boldsymbol{\varsigma}}^{(\mathfrak{p},\mathfrak{q})}_0) \big|\big|^\nu_{F^\mathfrak{s}}$$
$$\leq (\psi^{(\mathfrak{p},\mathfrak{q})}_{n,\nu})^{N/n} \mathbb{E}||\boldsymbol{\varsigma}^{(\mathfrak{p},\mathfrak{q})}_0 - \tilde{\boldsymbol{\varsigma}}^{(\mathfrak{p},\mathfrak{q})}_0||^\nu_{F^\mathfrak{s}}. \qquad \square$$

Acknowledgements

References

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

[2] Aue A., Horváth L. and Pellatt D.F. (2017) Functional generalized autoregressive conditional heteroskedasticity. *J. Time Series Anal.*, **38**, 3–21.

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

[4] Bollerslev T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. *J. Econometrics*, **31**, 307–327.

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

[6] Cerovecki C., Francq C., Hörmann S. and Zakoïan J.-M. (2019) Functional GARCH Models: The Quasi-likelihood Approach and Its Applications. *J. Econometrics.*

[7] Didericksen D., Kokoszka P. and Zhang X. (2012) Empirical properties of forecasts with the functional autoregressive model. *Comput. Statist.*, **27**(2), 285–298.

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

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

[10] Ferraty F. and Vieu P. (2006) *Nonparametric Functional Data Analysis.* New York: Springer.

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

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

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

[14] Hörmann S., Horváth L. and Reeder R. (2013) A Functional Version of the ARCH Model. *Econometric Theory*, **29**(2), 267–288.

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

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

[17] 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]. Berlin: Springer Spektrum.

[18] Kingman J.F.C. (1974) Subadditive Ergodic Theory. *Ann. Probab.*, **6**, 883–909.

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

[20] Klepsch J. and Klüppelberg C. (2017) An Innovations Algorithm for the prediction of functional linear processes. *J. Multivariate Anal.*, **155**, 252–271.

[21] Kokoszka P. and Reimherr M. (2013) Determining the order of the functional autoregressive model. *J. Time Series Anal.*, **34**, 116–129.

[22] Kokoszka P., Rice G. and Shang H.L. (2017). Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *J. Multivariate Anal.*, **162**, 32–50.

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

[24] Ledoux M. and Talagrand M. (1991) *Probability in Banach Spaces.* Berlin: Springer.

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

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

[27] Ramsay J.O. and Silverman B.W. (2005) *Functional Data Analysis* (2 ed.). New York: Springer.

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

[29] Rice G., Wirjanto, T. and Zhao Y. (2020) Tests for Conditional Heteroscedasticity of Functional Data. *J. Time Series Anal.*, **doi:** 10.1111/jtsa.12532.

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

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

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

[33] Tikhonov A.N. and Arsenin V.Y. (1977). *Solutions of Ill-Posed Problems*. Washington D.C.: Winston and Sons.

[34] Turbillon C., Marion J.-M. and 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.

[35] Werner D. (2018) *Funktionalanalysis* (German) [Functional analysis] (8 ed.). Berlin: Springer Spektrum.

[36] Yao F. and Müller H.-G. (2010) Functional quadratic regression. *Biometrika*, **97**, 49–64.