SIEVE BOOTSTRAP FOR FUNCTIONAL TIME SERIES

EFSTATHIOS PAPARODITIS

ABSTRACT. A bootstrap procedure for functional time series is proposed which exploits a general vector autoregressive representation of the time series of Fourier coefficients appearing in the Karhunen-Loève expansion of the functional process. A double sieve-type bootstrap method is developed which avoids the estimation of process operators and generates functional pseudo-time series that appropriately mimic the dependence structure of the functional time series at hand. The method uses a finite set of functional principal components to capture the essential driving parts of the infinite dimensional process and a finite order vector autoregressive process to imitate the temporal dependence structure of the corresponding vector time series of Fourier coefficients. By allowing the number of functional principal components as well as the autoregressive order used to increase to infinity (at some appropriate rate) as the sample size increases, a basic bootstrap central limit theorem is established which shows validity of the bootstrap procedure proposed for functional finite Fourier transforms. Some numerical examples illustrate the good finite sample performance of the new bootstrap method proposed.

1. INTRODUCTION

Statistical inference for time series stemming from stationary functional processes has attracted considerable interest during the last decades and progress has been made in several directions. Estimation and testing procedures have been developed for a wide range of inference problems and for large classes of stationary functional processes; see Bosq (2000), Hörmann and Kokoszka (2012) and Horváth and Kokoszka (2012). However, the asymptotic results derived, typically depend in a complicated way on difficult to estimate, infinite dimensional characteristics of the underlying functional process. This restricts considerably the implementability of
asymptotic approximations when used in practice to judge the uncertainty of estimation procedures or to calculate critical values of tests. In such situations, bootstrap methods can provide useful alternatives.

Bootstrap procedures for Hilbert space-valued time series proposed so far in the literature, are mainly attempts to adapt, to the infinite dimensional functional set-up, of bootstrap methods that have been developed for the finite dimensional (i.e., mostly univariate) time series case. Politis and Romano (1994) considered applications of the stationary bootstrap to functional, Hilbert-valued time series and showed its validity for the sample mean for functional processes satisfying certain mixing and boundeness conditions. Dehling et al. (2015) considered applications of the non-overlapping block bootstrap to U-statistics for so called near epoch dependent functional processes and Sharipov et al. (2016) to change point analysis. Franke and Nyarigue (2016) and Zhou and Politis (2016) developed some theory for a residual-based bootstrap applied to a first order functional autoregressive process. Notice that the transmission of other bootstrap procedures for real-valued time series to the functional set-up, like for instance of the autoregressive-sieve bootstrap, Kreiss (1988) and Kreiss et al. (2011), seems to be difficult mainly due to problems associated with the estimation (of an with sample size increasing number) of infinite dimensional autoregressive operators.

Applications of bootstrap procedures to certain inference problems in functional time series analysis have been also considered in the literature. For instance, for the construction of prediction intervals, Fernández De Castro et al. (2005) used an approach based on resampling pairs of functional observations by means of kernel-driven resampling probabilities. The same authors also apply a parametric, residual-based bootstrap approach using an estimated first order functional autoregression with i.i.d. resampling of appropriately defined functional residuals. For the same prediction problem, Hyndman and Shang (2009) applied different bootstrap approaches including bootstrapping the functional curves by randomly disturbing the forecasted scores using residuals obtained from univariate autoregressive fits. Aneiros-Perez et al. (2011) considered the nonparametric functional autoregressive models, while Mingotti et al. (2015) the case of the integrated functional autoregressive model. Apart from the lack of theoretical justification, the aforementioned bootstrap applications do not provide a general bootstrap methodology for functional time series as they are designed for and their applicability is restricted to the particular inference problem considered; see also McMurry and Politis (2012) and Shang (2016) for an overview.

In this paper a general and easy to implement bootstrap procedure for functional time series is proposed which generates bootstrap replicates $X_1^*, X_2^*, \ldots, X_n^*$ of an observed functional time series $X_1, X_2, \ldots, X_n$ and is applicable to a large class of stationary functional processes. The procedure avoids the explicit estimation of
process operators and exploits some basic properties of the stochastic process of Fourier coefficients (scores) appearing in the well-known Karhunen-Loève expansion of the functional random variables. It is in particular shown, that under quite general assumptions, the stochastic process of Fourier coefficients obeys a so-called vector autoregressive representation and this representation plays a key role in developing a bootstrap procedure for the functional time series at hand. More specifically, to capture the essential driving functional parts of the underlying infinite dimensional process, the first $m$ functional principal components are used and the corresponding $m$-dimensional time series of Fourier coefficients is bootstrapped using a $p$th order vector autoregression fitted to the vector time series of sample Fourier coefficients. In this way, a $m$-dimensional pseudo-time series of Fourier coefficients is generated which imitates the temporal dependence structure of the vector time series of sample Fourier coefficients. Using the (truncated) Karhunen-Loève expansion, these pseudo-Fourier coefficients are then transformed to functional bootstrap replicates of the main driving, principal components, of the observed functional time series. Adding to these replicates an appropriately resampled functional noise, leads finally to the bootstrapped functional pseudo-time series $X_1^*, X_2^*, \ldots, X_n^*$.

In a certain sense, our bootstrap procedure works by using a finite rank (i.e., $m$-dimensional) approximation of the infinite dimensional structure of the underlying functional process and a $p$th order vector autoregressive approximation of its infinite order temporal dependence structure. To achieve consistency and to capture appropriately the entire infinite dimensional structure of the functional process, the number $m$ of functional principal components used as well as the order $p$ of the vector autoregression applied, are allowed to increase to infinity (at some appropriate rate) as the sample size $n$ increases to infinity. This double sieve property justifies the use of the term “sieve bootstrap” for the bootstrap procedure proposed. We show that under quite general conditions, this bootstrap procedure succeeds in imitating correctly the entire infinite dimensional autocovariance structure of the underlying functional process. Notice that apart from the problem that instead of the unknown true scores, the time series of estimated scores is used, the asymptotic analysis of our bootstrap procedure faces additional challenges which are caused by the fact that vector autoregressions of increasing order and of increasing dimension are considered and that the lower bound of the corresponding spectral density matrix approaches zero as the dimension of the vector time series of scores used, increases to infinity. We apply the proposed sieve bootstrap procedure to the problem of estimating the distribution of the functional Fourier transform which is fundamental in a multitude of applications and has attracted interest in the functional time series literature; see Cerovecki and Hörmann (2015) for some recent developments. A basic bootstrap central limit theorem is then established which shows validity of the functional sieve bootstrap for this important class of statistics, which includes the sample mean as a special case.
Using the time series of Fourier coefficients in the context of functional time series analysis has been considered by many authors in a variety of applications. Among others we mention Hyndman and Shang (2009) who, for functional autoregressive models and for the sake of prediction, used univariate autoregressions fitted to the scalar time series of scores. In the same context and more related to the approach proposed in this paper, a multivariate approach of prediction has been proposed by Aue et al. (2014) which works by fitting a vector autoregressive model to the multivariate time series of scores.

The paper is organized as follows. Section 2 derives some basic properties and discuss the autoregressive representations of the vector process of Fourier coefficients appearing in the Karhunen-Loève expansion of the functional process. Apart from being useful for bootstrap purposes, these properties are of interest on their own. The functional sieve bootstrap procedure proposed is described in Section 3 where some properties of the bootstrap functional pseudo-time series are also discussed. Asymptotic validity of the new bootstrap procedure for finite Fourier transforms is established in Section 4. Section 5 presents some numerical simulations which investigate the finite sample performance of the functional sieve bootstrap and comparisons with the performance of three popular block bootstrap methods are given. Technical proofs and auxiliary lemmas are deferred to Section 6.

2. The Process of Fourier Coefficients

2.1. The functional set-up. We consider a (functional) stochastic process \( X = \{X_t, t \in \mathbb{Z}\} \) where for each \( t \) (interpreted as time), \( X_t \) is a random element of the separable Hilbert space \( \mathcal{H} := L^2([0, 1], \mathbb{R}) \) with parametrization \( \tau \to X_t(\tau) \in \mathbb{R} \) for \( \tau \in [0, 1] \). As usual we denote by \( \langle \cdot, \cdot \rangle \) the inner product in \( \mathcal{H} \) and by \( \| \cdot \| \) the induced norm defined for \( x, y \in \mathcal{H} \) as \( \langle x, y \rangle = \left( \int_{[0,1]} x(t)y(t)dt \right)^{1/2} \) and \( \|x\| = \left( \langle x, x \rangle \right)^{1/2} \) respectively. Furthermore, for a matrix \( A \) we denote by \( \|A\|_F \) its Frobenius norm, while for an operator \( T \), \( \|T\| \) denotes its operator norm and \( \|T\|_{HS} \) its Hilbert-Schmidt norm, if \( T \) is a Hilbert-Schmidt operator.

For the underlying functional process \( X \) it is assumed that its dependence structure satisfies the following assumption.

**Assumption 1** \( X \) is a purely non deterministic, \( L^4-M \) approximable process.

The general notion of \( L^p-M \) approximability refers to stochastic process \( X = \{X_t, t \in \mathbb{Z}\} \) with \( X_t \) taking values in \( \mathcal{H} \), \( \mathbb{E}\|X_t\|^p < \infty \), and where the random element \( X_t \) admits the representation \( X_t = f(\varepsilon_t, \varepsilon_{t-1}, \ldots) \). Here the \( \varepsilon_t \)'s are i.i.d. random elements in \( \mathcal{H} \) and \( f \) some measurable function \( f : \mathcal{H}^\infty \to \mathcal{H} \). If for
\[ \{ \tilde{\varepsilon}_t, t \in \mathbb{Z} \} \text{ an independent copy of } \{ \varepsilon_t, t \in \mathbb{Z} \}, \text{ the condition} \]

\[ \sum_{k=1}^{\infty} \left( E \| X_k - X_k^{(k)} \|^p \right)^{1/p} < \infty, \]

is satisfied, where \( X_t^{(M)} = f(\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-M+1}, \tilde{\varepsilon}_{t-M}, \tilde{\varepsilon}_{t-M-1}, \ldots) \), then \( X \) is called \( L^p - M \) approximable. \( L^p - M \) approximability is a notion of weak dependence which applies to many commonly used functional time series models, like linear functional processes, functional ARCH processes, etc.; see Hörmann and Kokoszka (2010) for more details.

Let \( \mu := E X_0 \in \mathcal{H} \) be the mean of \( X \) which by stationary is independent of \( t \) and for which we assume \( \mu = 0 \) for simplicity. We denote by \( C_h \) the autocovariance operator \( C_h : \mathcal{H} \to \mathcal{H} \) at lag \( h \in \mathbb{Z} \) defined by \( C_h(\cdot) = E(X_t - \mu, \cdot)(X_{t+h} - \mu) \). Associated with the covariance operator is the covariance function \( c_h : [0, 1] \times [0, 1] \to \mathbb{R} \) with \( c_h(\tau, \nu) = E(X_t(\tau) - \mu(\tau))(X_{t+h}(\nu) - \mu(\nu)), \tau, \nu \in [0, 1], \) that is, \( C_h \) is an integral operator with kernel function \( c_h \).

Assumption 1 implies that \( \sum_h \| C_h \|_{HS} < \infty \) and that for every \( \omega \in \mathbb{R} \) the spectral density operator

\[ F_\omega(x) = (2\pi)^{-1} \sum_{h \in \mathbb{Z}} C_h(x)e^{-ih\omega}, \quad x \in \mathcal{H} \]

is well defined, continuous in \( \omega \), selfadjoint and trace class, Hörmann et al. (2015); see also Panaretos and Tavakoli (2013) for similar properties under different weak dependence conditions which require summability of functional cumulants. In what follows we will strengthen somehow the assumption on the norm summability of the autocovariance operator to the following requirement.

**Assumption 2** \( \sum_h (1 + |h|)^r \| C_h \|_{HS} < \infty \) for some \( r \geq 0 \).

Furthermore, we will assume that the spectral density operator \( F_\omega \) satisfies the following condition.

**Assumption 3** For all \( \omega \in [0, \pi] \), the operator \( F_\omega \) is of full rank, i.e., \( \ker(F_\omega) = 0 \).

For real-valued univariate processes, \( \ker(F_\omega) = 0 \) is equivalent to the condition that the spectral density is everywhere in \( [0, \pi] \) strictly positive while for multivariate process to the non-singularity of the spectral density matrix for every frequency \( \omega \in [0, \pi] \). Notice that all eigenvalues \( \nu_j(\omega), j = 1, 2, \ldots \) of \( F_\omega \) are positive, while \( \sum_{j=1}^{\infty} \nu_j(\omega) < \infty \) by the trace class property of \( F_\omega \).

**2.2. Vector autoregressive representation.** Since \( C_0 = \int_{-\pi}^{\pi} F_\omega d\omega \), the positivity of \( F_\omega \) implies that the covariance operator \( C_0 \) has full rank, that is, its eigenvalues \( \lambda_j \) satisfy \( \lambda_j > 0 \) for all \( j \geq 1 \) and the corresponding eigenfunctions \( \{ v_j, j \geq 1 \} \) form a complete orthonormal basis in \( \mathcal{H} \). By the symmetry and compactness of \( C_0 \), the
random element $X_t$ admits the well known Karhunen-Loève representation

\begin{equation}
X_t = \sum_{j=1}^{\infty} \langle X_t, v_j \rangle v_j, \quad t \in \mathbb{Z},
\end{equation}

where $v_j$, $j = 1, 2, \ldots$, are the orthonormalized eigenfunctions that correspond to the eigenvalues $\lambda_j$, $j = 1, 2, \ldots$, of $C_0$. For $t \in \mathbb{Z}$, let $\xi_{j,t} := \langle X_t, v_j \rangle$, $j \geq 1$, and consider any subset of indices $M = \{j_1, j_2, \ldots, j_m\} \subset \mathbb{N}$ with $j_1 < j_2 < \ldots < j_m$, $m < \infty$. Later on, we will concentrate on the specific set $M = \{1, 2, \ldots, m\}$ which will be the set of the $m$ largest eigenvalues of the covariance operator $C_0$.

Consider now the $m$-dimensional process $\xi^{(M)} = \{\xi^{(M)}_t = (\xi^{(M)}_{s,t}, s = 1, 2, \ldots, m)^\top, t \in \mathbb{Z}\}$. Observe that $\xi^{(M)}$ is strictly stationary, purely non deterministic and has mean zero, $E(\xi^{(M)}_t) = (\langle EX_t, v_j \rangle, s \in M) = 0$. Furthermore, its autocovariance matrix function $\Gamma_{\xi^{(M)}}(h) = E(\xi^{(M)}_t \xi^{(M)}_{t+h})$, $h \in \mathbb{Z}$, is given by $\Gamma_{\xi^{(M)}}(h) = (\langle C_h(v_j), v_r \rangle)_{s,r=1,2,\ldots,m}$ and satisfies by Assumption 2,

\begin{equation}
\sum_{h=-\infty}^{\infty} (1 + |h|)^r \|\Gamma_{\xi^{(M)}}(h)\|_F = \sum_{h=-\infty}^{\infty} (1 + |h|)^r \left( \sum_{s,r=1}^{m} \langle C_h(v_j), v_r \rangle^2 \right)^{1/2} \leq \sum_{h=-\infty}^{\infty} (1 + |h|)^r \|C_h\|_{HS} < \infty.
\end{equation}

Note that the bound on the right hand side above is independent of the set $M$ and that although by construction it holds true that $\text{Cov}(\xi^{(M)}_{r_1,t}, \xi^{(M)}_{r_2,t}) = 0$ for $r_1 \neq r_2$, the random variables $\xi_{r_1,t}$ and $\xi_{r_2,s}$ may be correlated for $t \neq s$. The summability property (2.2) implies that the $m$-dimensional vector process $\xi^{(M)}$ possesses a continuous spectral density matrix $f_{\xi^{(M)}}(\cdot)$ which is given by

$$f_{\xi^{(M)}}(\omega) = (2\pi)^{-1} \sum_{h=-\infty}^{\infty} \Gamma_{\xi^{(M)}}(h)e^{-i\omega h}, \quad \omega \in \mathbb{R}.$$ 

Moreover, $f_{\xi^{(M)}}$ satisfies the following boundeness conditions.

**Lemma 2.1.** Under Assumption 1 and 3 and Assumption 2 with $r = 0$, the spectral density $f_{\xi^{(M)}}$ satisfies

\begin{equation}
\delta_M I_m \leq f_{\xi^{(M)}}(\omega) \leq c I_m, \quad \text{for all } \omega \in [0, \pi],
\end{equation}

where $\delta_M$ and $c$ are real numbers ($\delta_M$ depends on the set $M$), such that $0 < \delta_M \leq c < \infty$ and $I_m$ is the $m \times m$ unity matrix.

The continuity and the boundeness properties of the spectral density matrix $f_{\xi^{(M)}}(\cdot)$ stated in Lemma 2.1 imply that the process $\xi^{(M)}$ obeys a so called vector autoregressive representation; Cheng and Pourahmadi (1983), see also Wiener and Masani (1957) and (1958). That is, there exist an infinite sequence of $m \times m$-matrices
where the coefficients matrices satisfy
\[ \sum_{j=1}^{\infty} A_j^{(M)} \xi_{t-j}^{(M)} + e_t^{(M)}, \quad t \in \mathbb{Z}, \]
where the coefficients matrices satisfy \( \sum_{j \in \mathbb{Z}} (1+|j|)^r \| A_j^{(M)} \|_F < \infty \) and \( \{ e_t^{(M)}, t \in \mathbb{Z} \} \) is a zero mean white noise innovation process, that is \( E(\xi_t^{(M)}) = 0 \) and \( E(\xi_t^{(M)} e_s^{(M)\top}) = \delta_{t,s} \Sigma_e^{(M)} \), with \( \delta_{t,s} = 1 \) if \( t = s \), \( \delta_{t,s} = 0 \) otherwise and \( \Sigma_e^{(M)} \) a full rank \( m \times m \) covariance matrix. We stress here the fact that (2.4) does not describe a model for the process of Fourier coefficients \( \xi_t^{(M)} \) and should not be confused with the so-called linear, infinite order vector autoregressive (VAR(\( \infty \))) process driven by independent, identically distributed (i.i.d.) innovations. In fact, representation (2.4) is the autoregressive analogue of the well-known (moving average) Wold representation of the spectral density matrix of which is continuous and satisfies the boundness conditions (2.3); see also Cheng and Pourahmadi (1983) and Pourahmadi (2001) for details. In contrast to the Wold representation, the autoregressive representation (2.4) seems to be more appealing for statistical purposes, since it express the vector time series of Fourier coefficients \( \xi_t^{(M)} \) as a function of its (in principle) observable past values \( \xi_{t-j}^{(M)}, j = 1, 2, \ldots \).

In what follows we assume that the eigenvalues are in descending order, i.e., \( \lambda_1 > \lambda_2 > \cdots > \lambda_m > 0 \) and we consider the set \( M = \{ 1, 2, \ldots, m \} \) of the \( m \) largest eigenvalues of \( C_0 \). The corresponding normalized eigenfunctions (principal components) are denoted by \( v_j, j = 1, 2, \ldots, m \) and are (up to a sign) uniquely identified. Furthermore, by Parseval’s identity, the quantity \( \sum_{j=1}^{m} \lambda_j \) describes the proportion of the variance of \( X_t \) captured by the first \( m \) functional principal components. To simplify notation we surpass in the following the upper index \( (M) \) and write simple \( \xi_t \) for \( \xi_t^{(M)} \) keeping in mind that the \( j \)th component \( \xi_{j,t} = \langle X_t, v_j \rangle \) of \( \xi_t = (\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{m,t})\top \) corresponds to the \( j \)th largest eigenvalue \( \lambda_j \) of \( C_0, j = 1, 2, \ldots, m \). Furthermore, we write \( A_j(m), e_t(m), \delta_m \) and \( \Sigma_e(m) \) for \( A_j^{(M)}, e_t^{(M)}, \delta_M \) and \( \Sigma_e^{(M)} \), respectively.

3. The Functional Sieve Bootstrap Procedure

3.1. The bootstrap procedure. The basic idea of our procedure is to generate pseudo-replicates \( X_1^*, X_2^*, \ldots, X_n^* \) of the functional time series at hand by first bootstrapping the \( m \)-dimensional time series of Fourier coefficients \( \xi_t = (\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{m,t})\top, \quad t = 1, 2, \ldots, n \), corresponding to the first \( m \) principal components. This \( m \)-dimensional time series of Fourier coefficients is bootstrapped using the autoregressive representation of \( \xi_t \) discussed in Section 2.2. The generated \( m \)-dimensional pseudo-time series
of Fourier coefficients is then transformed to functional principal pseudo-components by means of the truncated Karhunen-Loève expansion $\sum_{j=1}^m \xi_{j,t} v_j$. Adding to this an appropriately resampled functional noise leads to the functional pseudo-time series $X_1^*, X_2^*, \ldots, X_n^*$. However, since the $\xi_t$'s are not observed, we work with the time series of estimates scores. This idea is precisely described in the following functional sieve bootstrap algorithm.

**Step 1:** Select a number $m = m(n)$ of functional principal components and an autoregressive order $p = p(n)$, both finite and depending on $n$.

**Step 2:** Let $\hat{\xi}_t = (\hat{\xi}_{j,t} = \frac{1}{n} \sum_{j=1}^m \langle X_t, \hat{v}_j \rangle, j = 1, 2, \ldots, m)'$, $t = 1, 2, \ldots, n$,

be the $m$-dimensional time series of estimated Fourier coefficients, where $\hat{v}_j$, $j = 1, 2, \ldots, m$ are the estimated eigenvectors corresponding to the estimated eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_m$ of the sample covariance operator $\hat{C}_0 = n^{-1} \sum_{t=1}^n (X_t - \bar{X}_n) \otimes (X_t - \bar{X}_n)$ with $\bar{X}_n = n^{-1} \sum_{t=1}^n X_t$.

**Step 3:** Let $\hat{X}_{t,m} = \sum_{j=1}^m \hat{\xi}_{j,t} \hat{v}_j$ and define functional residuals as $\hat{U}_{t,m} = X_t - \hat{X}_{t,m}$, $t = 1, 2, \ldots, n$.

**Step 4:** Fit a $p$th order vector autoregressive process to the $m$-dimensional time series $\hat{\xi}_t$, $t = 1, 2, \ldots, n$, denote by $\hat{A}_{j,p}(m), j = 1, 2, \ldots, p$, the estimates of the autoregressive matrices and by $\hat{e}_{t,p}$ the residuals,

$$\hat{e}_{t,p} = \hat{\xi}_t - \sum_{j=1}^p \hat{A}_{j,p}(m) \hat{\xi}_{t-j}, \quad t = p + 1, p + 2, \ldots, n.$$ Different estimators $\hat{A}_{j,p}(m), j = 1, 2, \ldots, p$ can be used, like for instance, Yule-Walker estimators; cf. Brockwell and Davis (1991).

**Step 5:** Generate a $m$-dimensional pseudo time series of scores $\xi_t^* = (\xi_{1,t}^*, \xi_{2,t}^*, \ldots, \xi_{m,t}^*)$, $t = 1, 2, \ldots, n$, as

$$\xi_t^* = \sum_{j=1}^p \hat{A}_{j,p}(m) \xi_{t-j}^* + e_t^*,$$

where $e_t^*$, $t = 1, 2, \ldots, n$ are i.i.d. random vectors having as distribution the empirical distribution of the centered residual vectors $\bar{e}_{t,p} = \hat{e}_{t,p} - \bar{e}_n$, $t = p + 1, p + 2, \ldots, n$ and $\bar{e}_n = (n - p)^{-1} \sum_{t=p+1}^n \hat{e}_{t,p}$.

**Step 6:** Define the pseudo-functional time series $X_1^*, X_2^*, \ldots, X_n^*$ by

$$X_t^* = \sum_{j=1}^m \xi_{j,t}^* \hat{v}_j + U_t^*,$$

$t = 1, 2, \ldots, n$. 

where \( U^*_1, U^*_2, \ldots, U^*_n \) are i.i.d. random functions obtained by choosing with replacement from the set of centered functional residuals \( \hat{U}_{t,m} - \bar{U}_n, t = 1, 2, \ldots, n \) and \( \bar{U}_n = n^{-1} \sum_{t=1}^{n} \hat{U}_{t,m} \).

Some remarks concerning the above algorithm are in order.

Note that \( X^*_1, X^*_2, \ldots, X^*_n \) are functional pseudo-random variables and that the autoregressive representation of the vector time series of Fourier coefficients is solely used as a vehicle to bootstrap the \( m \) main functional principal components of the time series at hand. In fact, it is this autoregressive representation which allows the generation of pseudo-time series of Fourier coefficients \( \xi^*_1, \xi^*_2, \ldots, \xi^*_n \), in Step 4 and Step 5, in a way that imitates the dependence structure of the sample Fourier coefficients \( \xi_1, \xi_2, \ldots, \xi_n \) and which are used in the (m-truncated) Karhunen-Loève representation. This together with the additive functional noise \( U^*_t \), lead to the new functional pseudo-observations \( X^*_1, X^*_2, \ldots, X^*_n \). Notice that the estimated eigenvectors \( \hat{v}_j \) may point in an opposite direction than the eigenvectors \( v_j \). In asymptotic derivations this is commonly taking care off by considering the sign corrected estimator \( \hat{s}_j \hat{v}_j \), where the (unobserved) random variable \( \hat{s}_j \) is given by \( \hat{s}_j = \text{sign}(\langle \hat{v}_j, v_j \rangle) \). However, since adding this sign correction will not affect the asymptotic results derived, we assume for simplicity throughout this paper, that \( \hat{s}_j = 1 \), for \( j = 1, 2, \ldots, m \).

Some modifications of the above basic bootstrap algorithm are possible which concern the resampling schemes used to generate the vector of pseudo-innovations \( e^*_t \) and/or the bootstrap functional noise \( U^*_t \). To elaborate, and as we will see in the sequel, the applied i.i.d. resampling of the innovations \( e^*_t \) in Step 5, suffices in order to capture the entire, infinite dimensional second order structure of the underlying functional process \( X \). A modification of this i.i.d. resampling scheme may be needed, however, if higher order characteristics of the underlying functional process beyond those of order two, should also be correctly mimicked by the functional pseudo-time series \( X^*_1, X^*_2, \ldots, X^*_n \). In such a case, the i.i.d. resampling used to generate the \( e^*_t \)'s in Step 5 can be replaced by other resampling schemes (i.e., block bootstrap schemes) that are able to capture higher order dependence characteristics of the white noise series \( e_t \)'s as well.

It should be mentioned that if \( X \) is a stationary functional autoregressive (FAR) process, i.e., if \( X_t = \sum_{j=1}^{p} \Phi_j (X_{t-j}) + \varepsilon_t \), where \( \Phi_j : \mathcal{H} \rightarrow \mathcal{H} \) are autoregressive operators and \( \varepsilon_t \) an i.i.d. sequence in \( \mathcal{H} \), see Bosq (2000) for stability conditions of the FAR process, then the infinite dimensional vector process \( \xi_t = (\xi_{j,t}, j = 1, 2, \ldots)^\top \) of scores satisfies \( \xi_t = \sum_{k=1}^{p} \hat{\Phi}_k \xi_{t-k} + u_t \), where \( \hat{\Phi}_k = ((\langle \Phi_k (v_j), v_i \rangle))_{i,j=1,2,\ldots} \), \( k = 1, 2, \ldots, p, \) are \( \infty \times \infty \) dimensional matrices and \( u_t = ((\varepsilon_t, v_j), j = 1, 2, \ldots)^\top \) is an infinite dimensional vector of i.i.d. innovations. That is, in this case the vector process of scores is generated by a linear, (infinite dimensional) autoregressive process of order \( p \). In this case and under some obvious modifications, the sieve bootstrap procedure
proposed in this paper provides an alternative way to bootstrap the FAR(p) process which circumvents the problem of estimating explicitly the autoregressive operators \( \Phi_j, \ j = 1, 2, \ldots, p \). Asymptotic validity of such a bootstrap procedure will then require that the truncation parameter \( m \) increases to infinity (at some appropriate rate) with the sample size \( n \).

### 3.2. Some properties of the bootstrap functional process

As usual, all considerations made regarding the bootstrap procedure are made conditionally on the observed functional time series \( X_1, X_2, \ldots, X_n \). The generation mechanism of the pseudo-time series \( X_1^*, X_2^*, \ldots, X_n^* \) enables us to consider the bootstrap functional process \( X^* = \{ X_t^*, t \in \mathbb{Z} \} \), where for \( t \in \mathbb{Z} \), \( X_t^* = \sum_{j=1}^{m} 1_j^\top \xi_t^j \hat{v}_j + U_t^* \), with \( \{ \xi_t^j = (\xi_{1,t}^j, \ldots, \xi_{m,t}^j), t \in \mathbb{Z} \} \) generated as \( \xi_t^j = \sum_{j=1}^{p} \hat{A}_{j,p}(m) \xi_{t-j}^j + e_t^j \), and, where the \( U_t^* \)'s are i.i.d. functional random variable taking values in the set \( \{ \bar{U}_{t,m} - \bar{U}_n, t = 1, 2, \ldots, n \} \) with probability \( 1/n \). In the above notation \( 1_j \) is the \( m \)-dimensional vector \( 1_j = (0, \ldots, 0, 1, 0, \ldots, 0)^\top \), where the unity appears in the \( j \)th position.

It is easy to see that \( X^* \) is a strictly stationary functional process with mean function \( E^* X_t^* = 0 \) for all \( t \in \mathbb{Z} \), and, autocovariance operator \( C_h^* : \mathcal{H} \to \mathcal{H} \) given, for \( h \in \mathbb{Z} \), by

\[
C_h^*(\cdot) = \sum_{j_1}^{m} \sum_{j_2}^{m} 1_{j_1}^\top \Gamma_h 1_{j_2} (\hat{v}_{j_1}, \cdot) \hat{v}_{j_2} + I(h = 0) E^* (U_t^*, \cdot) U_t^*,
\]

where \( \Gamma_h = E^*(\xi_{t-h}^j \xi_{t}^{j\top}) \) is the \( m \times m \) autocovariance matrix at lag \( h \) of \( \{ \xi_t^j, t \in \mathbb{Z} \} \). \( C_h^* \) is a Hilbert-Schmidt operator since it is, for \( h \neq 0 \), a finite rank operator while for \( h = 0 \) it is the sum of a finite rank operator and of the (Hilbert-Schmidt) empirical covariance operator of the pseudo-innovations

\[
C_U^* = E^* (U_t^*, \cdot) U_t^* = n^{-1} \sum_{t=1}^{n} (\hat{U}_{t,m} - \bar{U}_n, \cdot) (\hat{U}_{t,m} - \bar{U}_n).
\]

If the (estimated) vector autoregressive process used to generate the time series of pseudo-scores \( \xi_t^j \) is stable, then the dependence structure of the bootstrap process \( X^* \) can be more precisely described. This is stated in the following proposition. Notice that the required stability condition of the estimated autoregressive polynomial is fulfilled, if for instance, \( \hat{A}_{j,p}, \ j = 1, 2, \ldots, p \), are the Yule-Walker estimators; cf. Brockwell and Davis (1991), Ch. 11.4.

**Proposition 3.1.** If the estimator \( \hat{A}_{j,p}, \ j = 1, 2, \ldots, p \), used in Step 4 of the functional sieve bootstrap algorithm is such that \( \text{det}(\hat{A}_{p,m}(z)) \neq 0 \) for all \( |z| \leq 1 \), where \( \hat{A}_{p,m}(z) = I_m - \sum_{j=1}^{p} \hat{A}_{j,p}(m) z^j, z \in \mathbb{C} \), then, conditionally on \( X_1, X_2, \ldots, X_n \), the bootstrap process \( X^* \) is \( L^2 - \mathcal{M} \) approximable.
The $L^2 - \mathcal{M}$ approximability of $\mathbf{X}^*$ implies the norm summability $\sum_h \| C_h^* \|_{HS} < \infty$, see Hörmann et al. (2015), which can be also easily verified since
\[
\sum_{h \in \mathbb{Z}} \| C_h^* \|_{HS} \leq \sum_{h \in \mathbb{Z}} \| \Gamma_h^* \|_F + I(h = 0) \| C_h^* \|_{HS} = O_P(1).
\]
Furthermore, and because of the $L^2 - \mathcal{M}$ approximability property, the bootstrap process $\mathbf{X}^*$ possesses for every $\omega \in \mathbb{R}$ a spectral density operator $F_{\omega,m}^*$ defined by
\[
F_{\omega,m}^*(x) = (2\pi)^{-1} \sum_{h \in \mathbb{Z}} C_h^*(x)e^{-ih\omega}, \quad x \in \mathcal{H}.
\]
$C_h^*$ and $F_{\omega,m}^*$ are essentially finite rank approximations of the corresponding population operators $C_h$ and $F_\omega$ respectively. Thus and in order for the bootstrap process $\mathbf{X}^*$ to capture the infinite dimensional structure of the underlying functional process and the infinite order dependence structure of the vector time series generating the scores, the dimension $m$ as well as the autoregressive order $p$, used in the functional sieve bootstrap algorithm, have to increase to infinity (at an appropriate rate) as the sample size $n$ increases to infinity. This rate should take into account the fact that the true scores and eigenfunctions appearing in the Karhunen-Loève expansion are not observed and, therefore, sample estimates are used instead. Furthermore, not only the order $p$ but also the dimension $m$ of the fitted autoregressive process has to increase to infinity with the sample size which makes the asymptotic analysis more involved. Finally, the lower bound $\delta_m$ of the spectral density matrix $f_\xi$ of the scores approaches zero as the sample size $n$ increases to infinity. This is due to the fact that the eigenvalues $\nu_j(\omega)$ of the spectral density operator $F_\omega$ converge to zero as $j \to \infty$. These conditions impose several restrictions regarding the behavior of $m$ and $p$ with respect to the sample size $n$ which are summarized in the following assumption.

Assumption 4 The sequences $m = m(n)$ and $p = p(n)$ satisfy $m \to \infty$ and $p \to \infty$ as $n \to \infty$ such that,

\begin{enumerate}
  \item[(i)] $m = O(p^{1/2})$
  \item[(ii)] $p^2 m^5 \sum_{j=1}^m \frac{1}{a_j^2} \to 0$, where $\alpha_1 = \lambda_1 - \lambda_2$ and $\alpha_j = \min\{\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}\}$ for $j = 2, 3, \ldots, m$.
  \item[(iii)] $\delta^{-1}_m \sum_{j=p+1}^\infty j \| A_j(m) \|_F \to 0$, where $\delta_m$ is the lower bound of the spectral density matrix $f_\xi$ given in \[2.3\].
  \item[(iv)] $m^3 p^2 \sum_{j=1}^p \| \tilde{A}_{j,p}(m) - A_{j,p}(m) \|_F = O_P(1)$, where $\tilde{A}_{j,p}$, $j = 1, 2, \ldots, p$ denote the same estimator as $\tilde{A}_{j,p}$, $j = 1, 2, \ldots, p$, based on the true vector time series of scores $\xi_1, \xi_2, \ldots, \xi_n$ instead of their estimates $\hat{\xi}_1, \hat{\xi}_2, \ldots, \hat{\xi}_n$ and $A_{j,p}(m)$, $j = 1, 2, \ldots, m$ are the coefficient matrices of the best (in the mean square sense) linear predictor of $\xi_t$ based on $\xi_{t-j}$, $j = 1, 2, \ldots, p$. 
\end{enumerate}
Assumption 4(ii) is mainly imposed in order to control the error made by the fact that the bootstrap procedure is based on estimated scores and eigenfunctions instead on the unobserved true quantities in a context where the dimension \( m \) and the autoregressive order \( p \), both, increase to infinity. Part (iii) relates the rate of increase of the autoregressive order \( p \) to the lower bound of the spectral density matrix \( f_\xi \) and the decay of the norm of the autoregressive matrices to zero. Part (iv) is a requirement on the estimator \( \hat{A}_{j,p} \), \( j = 1, 2, \ldots, p \) based on the true scores. This is essentially a requirement on the estimator \( \hat{A}_{j,p} \) applied, after ignoring the error caused by the fact that the estimated scores are used. Notice that assumptions on the estimators of the autoregressive parameters are common in the autoregressive-sieve bootstrap literature for real valued-random variables; see Kreis et al. (2011) and Meyer and Kreiss (2015). However, the situation here is different since in our context, not only the order \( p \) but also the dimension \( m \) of the vector autoregression has to increase to infinity with the sample size. This justifies the additional factor \( m^4 \) appearing in (iv).

Under the condition that \( m \) and \( p \) increase to infinity at an appropriate rate with \( n \) such that Assumption 4 is satisfied, the following proposition can be established which shows that the spectral density operator \( F^{*}_{\omega,m} \) of the bootstrap process \( X^{*} \) converges, in Hilbert-Schmidt norm, to the spectral density operator \( F_{\omega} \) of the underlying functional process \( X \).

**Proposition 3.2.** Under Assumptions 1, 3 and 4 and Assumption 2 with \( r = 0 \), we have, that, as \( n \to \infty \),

\[
\sup_{\omega \in [0, \pi]} \| F^{*}_{\omega,m} - F_{\omega} \|_{HS} \to 0,
\]

in probability.

From the above proposition and the inversion formulae of Fourier transforms, we immediately get for the covariance operators \( C^{*}_{h} \) and \( C_{h} \) of the bootstrap process \( X^{*} \) and of the underlying process \( X \), that \( \sup_{h \in \mathbb{Z}} \| C^{*}_{h} - C_{h} \|_{HS} \to 0 \) in probability as \( n \to \infty \). Thus the bootstrap process \( X^{*} \), imitates asymptotically correct the entire infinite dimensional autocovariance structure of the functional process \( X \). This allows for the use of the bootstrap functional time series \( X^{*}_{1}, X^{*}_{2}, \ldots, X^{*}_{n} \) to approximate the distribution of statistics based on the functional time series \( X_{1}, X_{2}, \ldots, X_{n} \). Examples of such statistics are discussed in the next section.

4. **Bootstrap Validity**

In this section we investigate the validity of the functional sieve bootstrap applied in order to approximate the distribution of some statistic \( T_{n} := T(X_{1}, X_{2}, \ldots, X_{n}) \) of interest, when the bootstrap analogue \( T^{*}_{n} := T(X^{*}_{1}, X^{*}_{2}, \ldots, X^{*}_{n}) \) is used.
4.1. **Functional finite Fourier transform.** Consider the distribution of the functional Fourier transform

\( S_n(\omega) = \sum_{t=1}^{n} X_t e^{-it\omega}, \quad \omega \in [-\pi, \pi], \)

and notice that the sample mean \( \bar{X}_n = n^{-1}S_n(0) \) is a special case of (4.1). In order to investigate the limiting distribution of \( S_n(\omega) \), we fix some notation. We say that a random element \( Z \in \mathcal{H}_C := \mathcal{H} + i\mathcal{H} \), follows a complex Gaussian distribution with mean \( \tau \) and covariance \( G \), we write \( Z \sim \text{CN}(\tau, G) \), if

\[
\begin{pmatrix}
\text{Re}(Z) \\
\text{Im}(Z)
\end{pmatrix} \sim N_{\mathcal{H} \times \mathcal{H}}\left( \begin{pmatrix}
\text{Re}(\tau) \\
\text{Im}(\tau)
\end{pmatrix}, \frac{1}{2} \begin{pmatrix}
\text{Re}(G) & -\text{Im}(G) \\
\text{Im}(G) & \text{Re}(G)
\end{pmatrix}\right).
\]

Under a range of different weak dependence assumptions on the functional process \( X \), it has been shown that \( n^{-1/2}S_n(\omega) \Rightarrow \text{CN}(0, \pi F_\omega) \) as \( n \to \infty \). For \( \omega = 0 \), such a limiting behavior has been established for linear functional processes by Merlevède et al. (1997) and for \( L^p - \mathcal{M} \) approximable processes by Horváth et al. (2013). Panaretos and Tavakoli (2013) derived the above limiting distribution of \( n^{-1/2}S_n(\omega) \) for \( \omega \in [0, \pi] \), under a summability condition of the functional cumulants, while more general results for the same statistic and under weaker conditions, have been recently obtained by Cerovecki and Hörmann (2015).

The following theorem establishes asymptotic validity of the functional sieve bootstrap procedure for the class of functional Fourier transforms, when the bootstrap statistic \( n^{-1/2}S^*_n(\omega) = n^{-1/2} \sum_{t=1}^{n} X_t^* e^{-it\omega} \) is used to approximate the distribution of the statistic \( n^{-1/2}S_n(\omega) \).

**Theorem 4.1.** Under Assumptions 1, 3 and 4 and Assumption 2 with \( r = 1 \), we have for \( \omega \in [0, \pi] \), that, as \( n \to \infty \),

\[
\frac{1}{\sqrt{n}} S^*_n(\omega) \Rightarrow \text{CN}(0, \pi F_\omega),
\]

in probability.

5. **Numerical Simulations**

In order to investigate the finite sample behavior of the functional sieve bootstrap (FSP) we have performed numerical simulations using time series stemming from a first order functional moving average process given by

\( X_t = \varepsilon_t + \Theta(\varepsilon_{t-1}) \).

To specify the operator \( \Theta \) we adopted the simulation set-up of Aue et al. (2015) and set \( \Theta = 0.8\Psi \), where \( \Psi : \mathcal{H}_D \to \mathcal{H}_D, \mathcal{H}_D = \mathbb{R}^D \{ f_1, f_2, \ldots, f_D \} \) with \( f_j, j = 1, 2, \ldots, D \) and \( D = 21 \) Fourier basis functions on the interval \([0, 1] \). For \( x \in \)
$\mathcal{H}_D$ where $x = \sum_{j=1}^{D} c_j f_j$ with $c_j = \langle x, f_j \rangle$, the operator $\Psi$ considered, acts as follows: $\Psi(x) = \sum_{j=1}^{D} c_j \Psi(f_j) = \sum_{j=1}^{D} \sum_{t=1}^{T} c_j \langle \Psi(f_j), f_t \rangle f_t$. Furthermore, the i.i.d. innovations $\varepsilon_t$ in (5.1) are obtained as $\varepsilon_t = \sum_{j=1}^{D} Z_{t,j} f_j$, where the random variables $Z_{t,j}$ are i.i.d. Gaussian with mean zero and standard deviation equal to $j^{-1}$. We are interested in estimating the standard deviation of the sample mean $\sqrt{n} \bar{X}_n(\tau_j) = \frac{1}{n} \sum_{t=1}^{n} X_t(\tau_j)$, calculated for time series of length $n = 100$ observations and for $\tau_j$, $j = 1, 2, \ldots, T$, $T = 20$, equidistant time points in the interval $[0, 1]$. The exact standard deviation of the sample mean is estimated using 20,000 replications of the moving average model structure (5.1) considered. The finite sample performance of the FSB is also compared with that of three other bootstrap procedures, the moving block bootstrap (MBB), the tapered block bootstrap (TBB) and the stationary bootstrap (SB). All estimates presented are based on $R = 1,000$ replications and $B = 1,000$ bootstrap repetitions.

Although the development of a data driven procedure that automatically selects $p$ and $m$ is a problem of future research, we experimented with several values of $m$ and $p$ where we found that for this model and for the sample size considered, a vector autoregression of order $p = 3$ describes appropriately the temporal dependence of the vector time series of scores. Table 1 shows the FSB estimates obtained using different values of the bootstrap parameters $m$ and $p$. In Figure 1 the results obtained using the FSB procedure are compared with those of the three aforementioned block bootstrap methods. In particular, Figure 1 shows the FSB estimates of the standard deviation of the sample mean for $m = 2$ and $p = 3$ and the corresponding estimates of the MBB, of the TBB and of the SB. The bootstrap parameters of these block bootstrap procedures have been selected as follows. For the MBB we calculated the corresponding estimates for a range of values of the block size $b$ in the set $\{2, 3, \ldots, 20\}$ and we selected the block size $b$ which minimizes the overall relative bias given by $\sum_{j=1}^{T} |\sigma_{b,MBB}(\tau_j)/\sigma(\tau_j) - 1|$, where $\sigma_{b,MBB}(\tau_j)$ denote the moving block bootstrap estimator of the standard deviation using the block size $b$ and $\sigma(\tau_j)$ the corresponding estimated exact standard deviation. The block size for the MBB selected according to this procedure was equal to $b = 7$. The same procedure has been applied in order to choose the bootstrap parameters for the other two block bootstrap methods, that is of the TBB and of the SB procedure. For the TBB the block size selected according to this criterion was $b = 5$, while the same criterion delivered a probability of $p = 0.064$ for the geometric distribution involved in the SB procedure. Thus the corresponding block bootstrap estimator presented in Figure 1 are, in the described sense, the (overall) less biased block bootstrap estimators.

Please insert Table 1 and Figure 1 about here

As Table 1 shows the FSB estimates are quite good even for functional time series consisting of $n = 100$ observations and seem not to be very sensitive with respect
to the different choices of the parameter $m$ used to truncate the Karhunen-Loève expansion. Notice that as $m$ increases the variability of the FSB estimates increases too. This is expected since the number of parameters to be estimated for the vector autoregression fitted to the time series of scores equals $m^2 p$. The good finite sample behavior of the FSB method is also demonstrated in Figure 1 where comparisons of the FSB estimates with those obtained using the three different block bootstrap methods are given. As this figure shows, between the three bootstrap estimators considered, the MBB estimator seems to behave better than the FSB estimator, while both estimators are more biased compared to the TBB estimator. The later method performs best among the three different block bootstrap methods considered. However, all block bootstrap estimates are quite biased and they are clearly outperformed by the FSB estimates for all time points $T_j$ in the interval $[0, 1]$ considered.

6. Auxiliary Results and Proofs

Lemma 6.1. Let Assumption 1, 2 and 3 be satisfied. Denote by $\Psi_j(m)$, $j = 1, 2, \ldots$, the coefficients matrices of the power series $A^{-1}_m(z)$, where $A_m(z) = I_m - \sum_{j=1}^{\infty} A_j(m)z^j$, $|z| \leq 1$, and let $\Sigma_c(m) = E(e_t(m)e_t^\top(m))$. Then,

$$(i) \sum_{j=1}^{\infty} (1 + j)^r \|A_j(m)\|_F = O(1),$$

$$(ii) \sum_{j=1}^{\infty} (1 + j)^r \|\Psi_j(m)\|_F = O(1),$$

$$(iii) 0 < c_e \leq \|\Sigma_c(m)\|_F = O(1),$$

where all bounds on the right hand side are valid uniformly in $m$.

Proof: Consider (i) and (ii). Let $C_v$ be the class of all $m \times m$ matrix-valued functions on $[-\pi, \pi]$ with $C_{m \times m}$-valued Fourier coefficient matrices $(F_k, k \in \mathbb{Z})$ satisfying the condition $\sum_{h \in \mathbb{Z}} (1 + |h|)^r \|F_h\|_F < C < \infty$, where $C$ is independent of $m$. Then, $f_\xi \in C_v$ since the autocovariance matrix function of $f_\xi$ satisfies $\sum_{h \in \mathbb{Z}} (1 + |h|)^r \|\Gamma(\xi(h))\|_F < C < \infty$, see (2.2). Furthermore, $f_\xi(\omega) = \phi(\omega)\tilde{\phi}(\omega)$, with $\phi$ the optimal factor of $f_\xi$; see Cheng and Pourahmadi (1993), p. 116. From the boundedness conditions it follows that $det(f_\xi(\omega)) \geq \delta_m > 0$ for all $m \in \mathbb{N}$, and, therefore, $\phi(\omega)$ is invertible with inverse denoted by $\phi^{-1}$. Notice that $\phi, \phi^{-1} \in C_v$. According to Wiener and Masani (1958), Theorem 5.5 and Theorem 5.7, there exist sequences $\{C_n(m), n \in \mathbb{N}\}$ and $\{D_n(m), n \in \mathbb{N}\}$ which are independent of $t$ such that for all $t \in \mathbb{Z}$, $P_{f_\xi(t-j)}(\xi) = \sum_{j=1}^{\infty} C_j(m) e_{t-k}(m)$ and $e_t(m) = \sum_{j=0}^{\infty} D_j(m) \xi_{t-j}$, where $D_0(m) = \Sigma_{e^{-1/2}}(m)$ and the infinite sums are $L_2$-convergent. The coefficients in the autoregressive and the Wold representation are obtained by setting $A_0(m) = \Sigma_{e^{1/2}}(m) D_0(m) = I_m$, $A_j(m) = -\Sigma_{e^{-1/2}}(m) D_j(m)$ and $\Psi_j(m) = C_j(m) \Sigma_{e^{-1/2}}(m)$, where $C_j(m), j = 1, 2, \ldots$ and $D_j(m), j = 0, 1, 2, \ldots$, are the Fourier coefficients of $\phi$ and $\phi^{-1}$, respectively. Since $\phi, \phi^{-1} \in C_v$, we get that
\[ \sum_{j \in \mathbb{N}} (1 + |j|)^r \| A_j(m) \|_F \] and \[ \sum_{j \in \mathbb{N}} (1 + |j|)^r \| \Sigma_j(m) \|_F \] are bounded uniformly in \( m \).

In (iii) the lower bound follows from the regularity of the infinite dimensional process of scores \( \{ \xi_t = (\xi_{jt}, j = 1, 2, \ldots, \tau^*), t \in \mathbb{Z} \} \) which in turn follows from the regularity of \( X \). For the upper bound, let \( \sigma_j^{(m)} \), \( j = 1, 2, \ldots, m \), be the (positive) eigenvalues of \( \Sigma_\epsilon(m) \). Then, since \( 0 \leq \sum_{j=1}^m \sigma_j^{(m)} = \sum_{j=1}^m \text{Var}(e_{jt}) \leq \sum_{j=1}^m \text{Var}(\xi_{jt}) = \sum_{j=1}^m \lambda_j \) we have \( \| \Sigma_\epsilon(m) \|_F = \sqrt{\sum_{j=1}^m \sigma_j^2} \leq \sqrt{\sum_{j=1}^m \lambda_j^2} \leq \| C_0 \|_{HS} < \infty. \]

The following version of Baxter’s inequality is very useful in our setting because it relates the approximation error of the coefficient matrices of the finite predictor and of the autoregressive-representation of the \( m \)-dimensional process of scores to the lower bound of the spectral density matrix \( f_\xi(\cdot) \). It is an immediate consequence of Lemma 2.1 and Theorem 3.2 of Meyer et al. (2016).

**Lemma 6.2.** Let Assumption 1, 2 and 3 be satisfied. Then there exists a constant \( C > 0 \) which does not depend on \( m \), such that for all \( 0 \leq s \leq r - 1 \),

\[ \sum_{j=1}^p (1 + j)^s \| A_{j,p}(m) - A_j(m) \|_F \leq C \delta_m^{-1} \sum_{j=p+1}^\infty (1 + j)^{s+1} \| A_j(m) \|_F, \]

where \( \delta_m \) is given in Lemma 2.1.

The following lemma provides a useful bound between the estimated matrices of the autoregressive parameters based on the vector of scores \( \xi_t \) and on the vector of their estimates \( \hat{\xi}_t \), \( t = 1, 2, \ldots, n \). It deals with the case of the Yule-Walker estimators and can be established for other estimators, like least squares, by similar arguments.

**Lemma 6.3.** Let Assumption 1 be satisfied and let \( \hat{A}_{j,p}(m) \) be the Yule-Walker estimators of \( A_{j,p}(m) \), \( j = 1, 2, \ldots, p \), based on the time series of true scores \( \xi_1, \xi_2, \ldots, \xi_n \). Then,

\[ \sup_{1 \leq j \leq p} \| \hat{A}_{j,p}(m) - \hat{A}_{j,p}(m) \|_F = O_p \left( \frac{1}{\lambda_m} \left\{ \frac{m}{n} \sum_{j=1}^m \frac{1}{\sigma_j^2} \right\}^{1/2} \right). \]

**Proof:** We first show that

\[ (6.1) \quad \sup_{0 \leq h \leq n-1} \| \hat{\Gamma}_h(m) - \tilde{\Gamma}_h(m) \|_F = O_p(\{n^{-1} \sum_{j=1}^m \alpha_j^{-2}\}^{1/2}), \]

where \( \hat{\Gamma}_h(m) = n^{-1} \sum_{t=1}^{n-h} \hat{\xi}_{t+h} \hat{\xi}_{t+h}^\tau \), \( \tilde{\Gamma}_h(m) = n^{-1} \sum_{t=1}^{n-h} \xi_{t+h} \xi_{t+h}^\tau \), \( h = 0, 1, \ldots, n-1 \), \( \alpha_1 = \lambda_1 - \lambda_2 \) and \( \alpha_j = \min\{\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}\}, j = 2, 3, \ldots, m \). Since \( \| \hat{\Gamma}_h(m) - \tilde{\Gamma}_h(m) \|_F \leq \| n^{-1} \sum_{t=1}^{n-h} (\hat{\xi}_{t+h} - \xi_{t+h}) \hat{\xi}_{t+h}^\tau \|_F + \| n^{-1} \sum_{t=1}^{n-h} (\xi_{t+h} - \hat{\xi}_{t+h}) \hat{\xi}_{t+h}^\tau \|_F \) it suffices to consider only one of the two terms on the right hand side of the last bound. By the
triangular and the Cauchy-Schwarz inequality we have
\[
\|n^{-1} \sum_{t=1}^{n-h} (\xi_{t+h} - \xi_{t+h}^\ast) \xi_t^\top\|_F \leq n^{-1} \sum_{t=1}^{n-h} \|\langle X_{t+h}, \hat{\nu}_j \rangle, j = 1, \ldots, m \rangle \|_1 \times \|\langle X_{t+h}, \hat{\nu}_j \rangle, j = 1, \ldots, m \rangle \|_F \leq \left( \sum_{j=1}^m \|\hat{\nu}_j - \nu_j\|^2 \right)^{1/2} \frac{1}{n} \sum_{t=1}^n \|X_t\| \left( \sum_{j=1}^m \langle X_t, \hat{\nu}_j \rangle^2 \right)^{1/2} = O_p \left( \sum_{j=1}^m \|\hat{\nu}_j - \nu_j\|^2 \right)^{1/2},
\]
with the $O_p$ term uniformly in $h$. The assertion follows because by Assumption 1, $\sum_{j=1}^m \|\hat{\nu}_j - \nu_j\|^2 = O_p(n^{-1} \sum_{j=1}^m \alpha_j^{-2})$; see Hörmann and Kokoszka (2010).

We next proof the assertion for $p = 1$. For this recall that $A_{1,1}(m) = \hat{\Gamma}_0^{-1}(m)\hat{\Gamma}_1(m)$ and $A_{1,1}(m) = \Gamma_0^{-1}(m)\Gamma_1(m)$. Furthermore,
\[
\|\hat{\Gamma}_0^{-1}(m) - \Gamma_0^{-1}(m)\|_F \leq \|\hat{\Gamma}_0(m)^{-1}\|_F \frac{\|\hat{\Gamma}_0(m) - \hat{\Gamma}_0\|_F}{1 - \|\Gamma_0(m)^{-1}\|_F}\|\hat{\Gamma}_0(m) - \hat{\Gamma}_0\|_F
\]
and using the analogue expression for $\|\hat{\Gamma}_0^{-1}(m) - \Gamma_0^{-1}(m)\|_F$ we get
\[
\|\hat{\Gamma}_0^{-1}(m) - \Gamma_0^{-1}(m)\|_F \leq \|\Gamma_0(m)^{-1}\|_F \frac{\|\hat{\Gamma}_0(m) - \hat{\Gamma}_0\|_F}{1 - \|\Gamma_0(m)^{-1}\|_F}\|\hat{\Gamma}_0(m) - \hat{\Gamma}_0\|_F + \text{a lower order term}
\]
\[
= O_p \left( \|\Gamma_0(m)^{-1}\|_F\|\hat{\Gamma}_0(m) - \hat{\Gamma}_0\|_F \right)
\]
\[
= O_p \left( \lambda_m^{-1} \left\{ m n^{-1} \sum_{j=1}^m \alpha_j^{-2} \right\}^{1/2} \right),
\]
where the last equality follows using (6.1) and the bound $\|\Gamma_0^{-1}\|_F = \sqrt{\sum_{j=1}^m \lambda_j^{-2}} \leq \sqrt{m} \lambda_m^{-1}$. By the above arguments and using expression (6.1) we obtain
\[
\|\hat{A}_{1,1} - A_{1,1}\|_F \leq \|\hat{\Gamma}_0^{-1}(m) - \Gamma_0^{-1}(m)\|_F \|\hat{\Gamma}_1(m)\|_F + \|\hat{\Gamma}_0(m)^{-1}\|_F \|\hat{\Gamma}_1(m) - \hat{\Gamma}_1(m)\|_F
\]
\[
= O_p \left( \lambda_m^{-1} \left\{ m n^{-1} \sum_{j=1}^m \alpha_j^{-2} \right\}^{1/2} \right).
\]

Extension of the assertion for $p > 1$ follows using the above bound for $\|\hat{A}_{1,1} - A_{1,1}\|_F$, equation (6.1) and Durbin-Levinson’s recursion formulae for the Yule-Walker estimators; see Brockwell and Davis (1991), Ch. 11.4.

**Lemma 6.4.** Let Assumption 1 and 2 (with $r = 0$) be satisfied and $A_{p,m}(z) = I - \sum_{j=1}^p A_{j,p}(m)z^j$, $z \in \mathbb{C}$. There exists $p_m \in \mathbb{N}$ and a positive constant $C$ which
does not depend on \( m \) such that for \( m \in \mathbb{N} \) and all \( p \geq p_m \),
\[
\inf_{|\lambda| \leq 1 + 1/p} \left| \det(A_{p,m}(z)) \right| \geq C m^{-1/2}.
\]

**Proof:** We first show that the assertion is true for \( |\lambda| \leq 1 \). Since \( |\det(A_{p,m}(z))| \neq 0 \) for \( |\lambda| \leq 1 \) it follows by the minimum modulus principle for holomorphic functions that \( |\det(A_{p,m}(z))| \geq \inf_{|\lambda| = 1} |\det(A_{p,m}(\lambda))| \). Now, recall that for \( \omega \in [-\pi, \pi] \),
\[
2\pi f_\omega = A_{m,p}^{-1}(e^{-i\omega})\Sigma_e(m)\overline{A}_{m,p}(e^{-i\omega}).
\]
Let \( \mu_1(\omega) \) be the largest eigenvalue of \( f_\omega \), we then have
\[
|\det(A_{p,m}(e^{-i\omega})))|^2 = \det(\Sigma_e(m))/(2\pi |\det(f_\omega(\omega))|) \geq c_p/(2\pi m \mu_1(\omega)) \geq \widetilde{C} m^{-1},
\]
for some constant \( \widetilde{C} > 0 \) independent of \( m \). Notice that the first inequality follows by Lemma \[6.1\] (iii) and the last by the fact that \( \mu_1(\omega) \) is bounded uniformly in \( m \); see Lemma \[2.1\]. Thus \( \inf_{|\lambda| \leq 1} |\det(A_{p,m}(e^{-i\omega})))|^2 \geq C m^{-1} \) which implies that \( \inf_{|\lambda| \leq 1} |\det(A_{p,m}(\lambda))| \geq C m^{-1/2} \) with some constant \( C > 0 \) independent of \( m \). Extension of this lower bound to the slightly larger region \( |\lambda| \leq 1 + 1/p \) and for all \( p > p_m \) for some \( p_m \in \mathbb{N} \), follows then exactly along the same lines as the proof of Lemma 3.2 of Meyer and Kreiss (2015); see also Lemma 2.3 of Kreiss et al. (2011). \( \square \)

To state the next lemma we first fix the following notation. \( \Psi_j(m), \Psi_{j,p}(m), \tilde{\Psi}_{j,p}(m) \) and \( \hat{\Psi}_{j,p}(m) \), \( j = 1, 2, \ldots \) denote the coefficient matrices in the power series expansions of \( A_{-1,m}(z), A_{p,m}(z), \tilde{A}_{p,m}(z) \) and \( \hat{A}_{p,m}(z) \), respectively, \( |\lambda| \leq 1 \). We set \( \Psi_0(m) = \Psi_{0,p}(m) = \hat{\Psi}_{0,p}(m) = \tilde{\Psi}_{0,p}(m) = I_m \). Furthermore, \( \epsilon_t(m) = \xi_t - \sum_{j=1}^\infty A_j(m)\xi_{t-j}, \epsilon_{t,p}(m) = \xi_t - \sum_{j=1}^p A_j(m)\xi_{t-j}, \tilde{\epsilon}_{t,j}(m) = \xi_t - \sum_{j=1}^\infty A_{j,p}(m)\xi_{t-j}, \tilde{\epsilon}_{t,p}(m) = \xi_t - \sum_{j=1}^p A_{j,p}(m)\xi_{t-j} \). While \( \Sigma_{e,p}(m) = E^+(\tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{n,p}(m))(\tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{n,p}(m))^\top \) and \( \Sigma_{e,p}(m) = E^+(\tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{n,p}(m))(\tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{n,p}(m))^\top \) with \( \tilde{\epsilon}_{n,p}(m) = (n-p)^{-1}\sum_{t=p+1}^n \tilde{\epsilon}_{t,p}(m) \) and \( \tilde{\epsilon}_{n,p}(m) = (n-p)^{-1}\sum_{t=p+1}^n \tilde{\epsilon}_{t,p}(m) \) where \( E^+ \) denotes expectation with respect to the measure assigning probability \( (n-p)^{-1} \) to each \( \tilde{\epsilon}_{t,p}(m) \), \( t = p + 1, p + 2, \ldots, n \).

**Lemma 6.5.** Let Assumptions 1, 2 \( (r = 0) \), 3 and 4 be satisfied. Then, as \( n \to \infty \),
\[
(i) \sum_{j=1}^\infty \| \Psi_{j,p}(m) - \Psi_{j,p}(m) \|_F \overset{P}{\to} 0,
(ii) \| \Sigma_{e,p}(m) - \Sigma_{e,p}(m) \|_F \overset{P}{\to} 0,
(iii) \sum_{j=1}^\infty \| \Psi_{j,p}(m) - \Psi_{j,p}(m) \|_F \overset{P}{\to} 0,
(iv) \| \Sigma_{e,p}(m) - \Sigma_{e,p}(m) \|_F \overset{P}{\to} 0,
(v) \sum_{j=1}^\infty \| \Psi_{j,p}(m) - \Psi_{j,p}(m) \|_F \overset{P}{\to} 0,
(vi) \| \Sigma_{e,p}(m) - \Sigma_{e,p}(m) \|_F \overset{P}{\to} 0.
\]
Proof: To see (i) let \( A(r,s) \) be the \((r,s)\)th element of a matrix \( A \) and notice that by Cauchy’s inequality for holomorphic functions we have

\[
|\tilde{\Psi}_{j,p}^{(r,s)}(m) - \Psi_{j,p}^{(r,s)}(m)| \leq (1 + \frac{1}{p})^{-j} \max_{|z|=1+1/p} \| \tilde{A}_{p,m}^{-1}(z) - A_{p,m}^{-1}(z) \|_F
\]

and

\[
\max_{|z|=1+1/p} \| \tilde{A}_{p,m}^{-1}(z) - A_{p,m}^{-1}(z) \|_F \leq \max_{|z|=1+1/p} \frac{1}{|\det(\tilde{A}_{p,m}(z))|} \| \tilde{A}_{p,m}^{Adj}(z) - A_{p,m}^{Adj}(z) \|_F
\]

\[+ \max_{|z|=1+1/p} \frac{1}{|\det(\tilde{A}_{p,m}(z))|} \| A_{p,m}^{Adj}(z) \|_F \]

\[= R_{1,n}(z) + R_{2,n}(z), \]

with an obvious notation for \( R_{1,n}(z) \) and \( R_{2,n}(z) \). By Theorem 2.12 of Ipsen and Rehman (2008) we get that

\[
|\det(\tilde{A}_{p,m}(z)) - \det(A_{p,m}(z))| \leq m \| \tilde{A}_{p,m}(z) - A_{p,m}(z) \|_F^2B_{\max}^{m-1}(z),
\]

where \( B_{\max}(z) = \max\{\| \tilde{A}_{p,m}(z) \|_2, \| A_{p,m}(z) \|_2\} \) and \( \|A\|_2 \) denotes the spectral norm (i.e., the largest singular value) of the matrix \( A \). Since \( \|A\|_2 \leq \|A\|_F \) and using the bound

\[
\|A\|_2 \leq \frac{2}{|\det(A)|} \left( \frac{\|A\|_F}{\sqrt{m+1}} \right)^{m+1},
\]

for the largest singular value of a non-singular matrix \( A \), see Merikoski and Kumar (2005), p. 373, we get by straightforward calculations and in view of Lemma \ref{lemma6.4} and the constant \( C \) appearing there, that

\[
\max_{|z|=1+1/p} B_{\max}^{m-1}(z) \leq \frac{2^{m-1}m^{(m-1)/2} \max_{|z|=1+1/p} \| B_{\max}(z) \|_F^{m-1}}{C^{m-1}(m+1)^{(m^2-1)/2}} = o_P(1),
\]

since \( \max_{|z|=1+1/p} \| A_{p,m}(z) \|_F = O(1) \) and \( \max_{|z|=1+1/p} \| \tilde{A}_{p,m}(z) \|_F = O_P(1) \) uniformly in \( m \). Thus

\[
|\det(\tilde{A}_{p,m}(z)) - \det(A_{p,m}(z))| \leq m \| \tilde{A}_{p,m}(z) - A_{p,m}(z) \|_F o_P(1).
\]

Lemma \ref{lemma6.1} and Assumption 4(iv) lead to the bound

\[
\sup_{|z| \leq 1+1/p} \| \tilde{A}_{p,m}(z) - A_{p,m}(z) \| \leq (1 + 1/p)^p \sum_{j=1}^{p} \| \tilde{A}_{p,m}(z) - A_{p,m}(z) \| \leq O_P(m^{-4}p^{-2}),
\]

from which we derive using \( (6.3) \) that

\[
\sup_{|z| \leq 1+1/p} \| \det(\tilde{A}_{p,m}(z)) - \det(A_{p,m}(z)) \| \leq o_P(1)m \sup_{|z| \leq 1+1/p} \| \tilde{A}_{p,m}(z) - A_{p,m}(z) \|
\]

\[= o_P(1)O_P(m^{-3}p^{-2}).\]
and by Lemma 6.4 that

\[ R_{1,n}(z) \leq \delta_m^{-1} \sum_{r=1}^{m} \sum_{s=1}^{m} \sup_{|z| \leq 1+1/p} |\text{det}(\tilde{A}_{r,s}(z)) - \text{det}(A_{r,s}(z))| \]

\[ = O_F(m^{1/2})O_P(m^2)O_P(m^{-3}p^{-2}) = o_P(m^{-1/2}p^{-2}). \]

Furthermore, by Lemma 6.4 and the bound (6.3) we get

\[ R_2(z) \leq \delta_m^{-2} \max_{|z|=1+1/p} \|A_{r,m}(z)\|_F \max_{|z|=1+1/p} |\text{det}(\tilde{A}_{r,m}(z)) - \text{det}(A_{r,m}(z))| \]

\[ = o_P(\delta_m^{-2}m^{-3}p^{-2}) = o_P(m^{-2}p^{-2}). \]

Thus and using equation (6.2), we conclude that

\[ \sum_{j=1}^{\infty} |\tilde{\Psi}_{j,p}(m) - \Psi_{j,p}(m)| \leq \sum_{j=1}^{\infty} \sum_{r=1}^{m} \sum_{s=1}^{m} |\tilde{\Psi}_{j,p}(m) - \Psi_{j,p}(m)| \]

\[ = o_P(m^{3/2}p^{-1}) + O_P(p^{-1}) \to 0, \]

by Assumption 4.

Consider (ii) so we have,

\[ \|\tilde{\Sigma}_{e,p}(m) - \Sigma_{e,p}(m)\|_F \leq \| \frac{1}{n-p} \sum_{t=p+1}^{n} (\tilde{e}_{t,p}(m)e_{t,p}^\top(m) - e_{t,p}(m)e_{t,p}^\top(m))\|_F \]

\[ + \| \frac{1}{n-p} \sum_{t=p+1}^{n} e_{t,p}(m)e_{t,p}^\top(m) - Ee_{t,p}(m)e_{t,p}^\top(m)\|_F + \|\tilde{e}_{n,p}(m)e_{n,p}^\top(m)\|_F \]

\[ = E_{1,n} + E_{2,n} + E_{3,n}, \]

with an obvious notation for \( E_{j,n}, j = 1, 2, 3 \). We show that all three terms converge to zero in probability. By the triangular inequality and in order to show \( E_{1,n} \to 0 \), it suffices to show that \( E_{1,n}(1) = \| (n-p)^{-1} \sum_{t=p+1}^{n} (\tilde{e}_{t,p}(m) - e_{t,p}(m))\tilde{e}_{t,p}(m)\|_F \to 0 \).

For this we use the bound

\[ E_{1,n}^{(1)} \leq \| \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} (\tilde{A}_{j,p}(m) - A_{j,p}(m))\xi_{t-j}\tilde{e}_{t,p}^\top(m)\|_F \]

\[ + \| \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} (A_{j,p}(m) - A_j(m))\xi_{t-j}\tilde{e}_{t,p}^\top(m)\|_F \]

\[ + \| \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=p+1}^{\infty} A_j(m)\xi_{t-j}\tilde{e}_{t,p}^\top(m)\|_F. \]

(6.4)
Since by straightforward calculations it yields that \( \sum_{j=1}^{p} \| \xi_{t-j} \tilde{e}_{t,p}^\top(m) \|_F = O_P(m^2p) \), we get by Assumption 4(iv) and Cauchy-Schwarz’s inequality that
\[
\left\| \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} (\tilde{A}_{j,p}(m) - A_{j,p}(m)) \xi_{t-j} \tilde{e}_{t,p}^\top(m) \right\|_F = O_P(m^{-1}p^{-1/2}) \to 0.
\]
For the second term on the right hand side of (6.4) we get by replacing \( \tilde{e}_{t,p}(m) \) by \( e_{t,p}(m) \) and using Lemma 6.2, that
\[
E \left\| \sum_{j=1}^{p} (A_{j,p}(m) - A_j(m)) \xi_{t-j} e_{t,p}^\top(m) \right\|_F \leq \sum_{j=1}^{p} \| A_{j,p}(m) - A_j(m) \|_F O(m)
\]
\[
= O(m \delta^{-1}_m \sum_{j=p+1}^{\infty} \| A_j(m) \|_F)
\]
\[
= O(mp^{-1} \delta^{-1}_m \sum_{j=p+1}^{\infty} j \| A_j(m) \|_F) \to 0,
\]
by Assumption 4. Finally and by the same assumption, we get for the third term of (6.4) using
\[
E \sum_{j=p+1}^{\infty} \| A_j(m) \|_F \| \xi_{t-j} e_{t,p}^\top(m) \| = O(m) \sum_{j=p+1}^{\infty} \| A_j(m) \|_F = O(mp^{-1} \sum_{j=p+1}^{\infty} j \| A_j(m) \|_F),
\]
which converges to zero in probability.
Since the term \( E_{2,n} \) is easier to deal with using similar arguments as for the term \( E_{1,n} \), we consider the term \( E_{3,n} \). Using \( \tilde{e}_n(m) = (n-p)^{-1} \sum_{t=p+1}^{n} e_t(m) \) we have that
\[
E_{3,n} \leq \left\| (\tilde{e}_{n,p}(m) - \tilde{e}_n(m))(\tilde{e}_{n,p}(m) - \tilde{e}_n(m))^	op \right\|_F + \left\| \tilde{e}_n(m) \tilde{e}_n(m)^	op \right\|_F
\]
\[
+ 2 \left\| (\tilde{e}_{n,p}(m) - \tilde{e}_n(m)) \tilde{e}_n(m) \right\|_F.
\]
Since \( \tilde{e}_n(m) = O_P((n-p)^{-1/2}) \) uniformly in \( m \) and by similar arguments as above, we have
\[
\left\| \tilde{e}_{n,p}(m) - \tilde{e}_n(m) \right\| \leq \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} \| \tilde{A}_{j,p}(m) - A_{j,p}(m) \|_F \| \xi_{t-j} \| \leq \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} \| A_{j,p}(m) - A_j(m) \|_F \| \xi_{t-j} \|
\]
\[
+ \frac{1}{n-p} \sum_{t=p+1}^{n} \sum_{j=1}^{p} \| A_{j,p}(m) - A_j(m) \|_F \| \xi_{t-j} \|
\]
\[
+ \frac{1}{n-p} \sum_{t=p+1}^{\infty} \sum_{j=p+1}^{\infty} \| A_j(m) \|_F \| \xi_{t-j} \| = O_P(p^{-1/2}m^{-3/2}) + O_P(m^{1/2} \delta^{-1}_m \sum_{j=p+1}^{\infty} \| A_j(m) \|_F).
\]

\[FUNCTIONAL\ SIEVE\ BOOTSTRAP\]
Thus we conclude using Assumption 4, that $E_{3,n} \overset{P}{\to} 0$.

Consider (iii). By (i) it suffices to show that $\sum_{j=1}^{\infty} \| \hat{\Psi}_{j,p}(m) - \hat{\Psi}_{j,p}(m) \|_F \overset{P}{\to} 0$. For this notice that

$$
\sup_{|z| \leq 1+1/p} \| \hat{A}_{p,m}(z) - \hat{A}_{p,m}(z) \|_F \leq (1 + \frac{1}{p})^p \sum_{j=1}^{p} \| \hat{A}_{j,p}(m) - \hat{A}_{j,p}(m) \|_F
$$

$$
= O_P\left(p\lambda_m^{-1} \left\{ m^{-1} \sum_{j=1}^{m} \alpha_j^{-2} \right\} \right)
$$

by Lemma 6.3 and by Cauchy’s inequality for holomorphic functions we get for the $(r,s)$th element of the matrices $\hat{\Psi}_{j,p}(m)$ and $\hat{\Psi}_{j,p}(m)$, that

$$
\left| \hat{\Psi}_{j,p}^{(r,s)}(m) - \hat{\Psi}_{j,p}^{(r,s)}(m) \right| \leq (1 + \frac{1}{p})^{-j} \max_{|z| = 1+1/p} \| \hat{A}_{-1,m}(z) - \hat{A}_{-1,m}(z) \|_F
$$

and

$$
\max_{|z| = 1+1/p} \| \hat{A}_{-1,m}(z) - \hat{A}_{-1,m}(z) \|_F \leq \max_{|z| = 1+1/p} \frac{1}{|\det (\hat{A}_{p,m}(z))|} \| \hat{A}_{p,m}^{adj}(z) - \hat{A}_{p,m}^{adj}(z) \|_F + \max_{|z| = 1+1/p} \frac{1}{|\det (\hat{A}_{p,m}(z))|} \| \hat{A}_{p,m}^{adj}(z) \|_F.
$$

From the above bound and by Lemma 6.3 and Lemma 6.4, we get by similar arguments as those leading to the bound (6.3), that, uniformly in $j$,

$$
\left| \hat{\Psi}_{j,p}^{(r,s)}(m) - \hat{\Psi}_{j,p}^{(r,s)}(m) \right| \leq (1 + \frac{1}{p})^{-j} O_P\left(\frac{m^{3}p}{n^{1/2}} \left\{ \sum_{j=1}^{m} \alpha_j^{-2} \right\} \right),
$$

that is

$$
\| \hat{\Psi}_{j,p}(m) - \hat{\Psi}_{j,p}(m) \|_F \leq \sum_{r,s=1}^{m} \left| \hat{\Psi}_{j,p}^{(r,s)}(m) - \hat{\Psi}_{j,p}^{(r,s)}(m) \right|
$$

$$
= (1 + \frac{1}{p})^{-j} O_P\left(\frac{m^{5}p}{\lambda_m n^{1/2}} \left\{ \sum_{j=1}^{m} \alpha_j^{-2} \right\} \right),
$$

from which we get

$$
\sum_{j=1}^{\infty} \| \hat{\Psi}_{j,p}(m) - \hat{\Psi}_{j,p}(m) \|_F = O_P\left(\frac{m^{5}p}{\lambda_m n^{1/2}} \left\{ \sum_{j=1}^{m} \alpha_j^{-2} \right\} \right) \to 0,
$$

by Assumption 4 (ii).

To establish (iv) notice that using (ii) it suffices to show that $\| \hat{\Sigma}_{e,p}(m) - \hat{\Sigma}_{e,p}(m) \|_F \overset{P}{\to} 0$. By the triangular inequality it suffices to show that

$$
\| \frac{1}{n - p} \sum_{t=p+1}^{n} \left[ (\hat{e}_{t,p}(m) - \hat{e}_{n,p}(m)) - (\hat{e}_{t,p}(m) - \hat{e}_{n,p}(m)) \right] \left( \hat{e}_{t,p}(m) - \hat{e}_{t,p}(m) \right) \|_F \overset{P}{\to} 0.
$$
Since the above term can be bounded by
\[
\frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{t,p}(m) \right\|^2 + \left\| \tilde{e}_{n,p}(m) - \tilde{e}_{n,p}(m) \right\| \frac{1}{n-p} \left\| \tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{t,p}(m) \right\|
\]
we show that both terms above converge to zero in probability. We use the bound
\[
\frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \tilde{\epsilon}_{t,p}(m) - \tilde{\epsilon}_{t,p}(m) \right\|^2 \leq 4 \sum_{j=1}^{p} \left\| \tilde{A}_{j,p}(m) - \tilde{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} \right\|^2
\]
+ \frac{2}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_t - \xi_t \right\|^2 + 4 \sum_{j=1}^{p} \left\| \tilde{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} - \xi_{t-j} \right\|^2.
\]
From Lemma 6.3 we get by straightforward calculations that, \((n-p)^{-1} \sum_{t=p+1}^{n} \left\| \hat{\xi}_t - \xi_t \right\|^2 = O_P(n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}),
\]
\[
\sum_{j=1}^{p} \left\| \tilde{A}_{j,p}(m) - \tilde{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} \right\|^2 = O_P \left( \frac{pm}{\lambda_n^{2m}} \sum_{j=1}^{m} \alpha_j^{-2} \right) O_P(1),
\]
and
\[
\sum_{j=1}^{p} \left\| \tilde{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} \right\|^2 = O_P(1) O_P \left( \frac{p}{n} \sum_{j=1}^{m} \alpha_j^{-2} \right).
\]
Similar arguments yield
\[
\left\| \tilde{e}_{n,p}(m) \right\|^2 \leq 2 \left\| \frac{1}{n-p} \sum_{t=p+1}^{n} \hat{\xi}_t \right\|^2 + 2 \left\| \sum_{j=1}^{p} \tilde{A}_{j,p}(m) \frac{1}{n-p} \sum_{t=p+1}^{n} \hat{\xi}_{t-j} \right\|^2
\]
\[
= O_P \left( mn^{-1} + n^{-1} \sum_{j=1}^{m} \alpha_j^{-2} \right),
\]
and \(\left\| \tilde{e}_{n,p}(m) \right\|^2 = O_P(mn^{-1}) \rightarrow 0.
\]
The proof of (v) and (vi) is straightforward and uses Lemma 6.1 and Lemma 6.2.

**Proof of Lemma 2.1** Expression (2.2) immediately leads, for all \(\omega \in [0, \pi]\),

\[
to an upper bound of \(f_{\xi(M)}(\omega)\). To derive a lower bound, recall that \(\Gamma_{\xi(M)}(h) = (\langle C_h(v_{jr}), v_{js} \rangle)_{r,s=1,2,\ldots,m}\) and observe that
\[
\mathbf{f}_{\xi(M)}(\omega) = \left( \langle (2\pi)^{-1} \sum_{h} C_h(v_{jr}) e^{-ith}, v_{js} \rangle \right)_{r,s=1,2,\ldots,m} = \left( \langle \mathcal{F}_{\omega}(v_{jr}), v_{js} \rangle \right)_{r,s=1,2,\ldots,m}.
\]

Let \(\mu_j(\omega), j = 1, 2, \ldots, m,\) be the eigenvalues of \(f_{\xi(M)}(\omega)\) (including multiplicity).

It suffices to show that \(\min_{1 \leq j \leq m} \mu_j(\omega) \geq \delta_M > 0\) for all frequencies \(\omega \in [0, \pi]\), i.e.,

that the eigenvalues of the spectral density matrix \(f_{\xi(M)}(\omega)\) are uniformly (in \(\omega\))

bounded away from zero. For this let
\[
c_j(\omega) = (c_{j,1}(\omega), c_{j,2}(\omega), \ldots, c_{j,m}(\omega))^T \in \mathbb{C}^m, \ j = 1, 2, \ldots, m,
\]
be the corresponding normalized eigenvectors. Then for every \( j \in \{1, 2, \ldots, m\} \), we have

\[
\mu_j(\omega) = c_j^\top(\omega) f_{\xi(M)}(\omega) c_j(\omega) \\
= c_j^\top(\omega) \left( (F_\omega(v_{j_r}), v_{j_s}) \right)_{r,s=1,2,\ldots,m} c_j(\omega) \\
= (F_\omega(y_j(\omega)), y_j(\omega)) > 0,
\]

by the positivity of \( F_\omega \), where \( y_j(\omega) = \sum_{r=1}^m c_{j,r}(\omega) v_{j_r} \in \bar{V}_M = \bar{\mathbb{P}} \{ v_{j_1}, v_{j_2}, \ldots, v_{j_m} \} \) and \( \| y_j \| = 1 \). Because of the norm summability of the autocovariance matrix function \( \Gamma_{\xi M}(h) \), the eigenvalues \( \mu_j(\omega) \), \( j = 1, 2, \ldots, m \), are continuous functions of \( \omega \). Let \( \delta_M(\omega) = \min_{1 \leq j \leq m} \mu_j(\omega) \) and notice that \( \delta_M(\omega) \) is continuous in \( \omega \) and \( \delta_M(\omega) > 0 \) for all \( \omega \in [0, \pi] \). Define \( \delta_M = \min_{\omega \in [0, \pi]} \delta_M(\omega) \) which is positive by the continuity of \( \delta_M(\omega) \). Hence \( \min_{1 \leq j \leq m} \mu_j(\omega) \geq \delta_M > 0 \) for all \( \omega \in [0, \pi] \).

\( \square \)

**Proof of Proposition 3.1**  
Recall the definition of \( X_t^* = \sum_{j=1}^m 1_j^\top \xi_t^* \hat{v}_j + U_t^* \) and observe that \( \xi_t^* = \sum_{l=0}^\infty \hat{\Psi}_{l,p}(m) e_{t-l}^* \), where \( \hat{\Psi}_{l,p}(m) = I_m \) and the power series \( \hat{\Psi}_{m,p}(z) = I_m + \sum_{l=1}^\infty \hat{\Psi}_{l,p}(m) z^l = (I_m - \sum_{j=1}^p A_{j,p}(m) z^j)^{-1} \) converges for \( |z| \leq 1 \). Write \( X_t^* = \sum_{l=0}^\infty \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) e_{t-l}^* \hat{v}_j + U_t^* \) and define

\[
X_{t, M}^* = \sum_{l=0}^{M-1} \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) e_{t-l}^* \hat{v}_j + \sum_{l=M}^\infty \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) e_{t-l}^* \hat{v}_j + U_t^*,
\]

where for each \( t \in \mathbb{Z} \), \( \{ e_{s,t}^*, s \in \mathbb{Z} \} \) is an independent copy of \( \{ e_s^*, s \in \mathbb{Z} \} \). Notice that

\[
X_{M}^* - X_{M,M}^* = \sum_{l=M}^\infty \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) (e_{M-l}^* - e_{M-l,M}^*) \hat{v}_j.
\]

By Minkowski’s inequality we have

\[
\sqrt{E \| X_M^* - X_{M,M}^* \|^2} \leq \sqrt{E \| \sum_{l=M}^\infty \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) e_{M-l}^* \hat{v}_j \|^2} + \sqrt{E \| \sum_{l=M}^\infty \sum_{j=1}^m 1_j^\top \hat{\Psi}_{l,p}(m) e_{M-l,M}^* \hat{v}_j \|^2},
\]

(6.5)
and evaluating the first expectation term we get using $\|A\|_F^2 = tr(AA^T)$ and the submultiplicative property of the Frobenius matrix norm, that

$$E\| \sum_{l=M}^{\infty} \sum_{j=1}^{m} 1_j^T \hat{\Psi}_{l,p}(m)e_{M-l}^* \hat{v}_j \|_F^2 = \sum_{l=M}^{\infty} tr(\hat{\Psi}_{l,p}(m)\Sigma^*(m)\hat{\Psi}^*_l(m))$$

$$\leq \|\hat{\Sigma}_{e,p}^{1/2}(m)\|_F^2 \sum_{l=M}^{\infty} \|\hat{\Psi}_{l,p}(m)\|_F^2,$$

where $\hat{\Sigma}_{e,p}(m) = \hat{\Sigma}_{e,p}^{1/2}(m)\hat{\Sigma}_{e,p}^{1/2}(m)$. An identical expression appears for the second expectation term on the right hand side of (6.5). Applying Minkowski’s inequality again we get by the exponential decay of $\|\hat{\Psi}_{l,p}(m)\|_F$, that

$$\sum_{M=1}^{\infty} \sqrt{E\|X_{M}^* - X_{M,M}^*\|_F} \leq 2\|\hat{\Sigma}_{e,p}^{1/2}(m)\|_F \sum_{M=1}^{\infty} \sum_{l=M}^{\infty} \|\hat{\Psi}_{l,p}(m)\|_F$$

$$= 2\|\hat{\Sigma}_{e,p}^{1/2}(m)\|_F \sum_{l=1}^{\infty} l\|\hat{\Psi}_{l,p}(m)\|_F = O_P(1).$$

\[\Box\]

**Proof of Proposition 3.2.** Recall that the spectral density operator $F_\omega$ can be expressed as $2\pi F_\omega = \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \sum_{h=-\infty}^{\infty} (C_h(v_j), v_l)e^{-ih\omega}(v_j \otimes v_l)$. Define for $m \in \mathbb{N}$, $2\pi F_{\omega,m} = \sum_{j=1}^{m} \sum_{l=1}^{m} \sum_{h=-\infty}^{\infty} (C_h(v_j), v_l)e^{-ih\omega}(v_j \otimes v_l)$ and verify that since $(C_h(v_j), v_l) = E(\xi_j, \xi_l t + h)$, the following expression is also valid for $F_{\omega,m}$,

$$2\pi F_{\omega,m}(\cdot) = \sum_{j=1}^{m} \sum_{l=1}^{m} 1_j^T \Psi_m(e^{-i\omega})\Sigma_e(m)\overline{\Psi}_m(e^{-i\omega})1_l(v_j, \cdot)v_l,$$

where $\Psi_m(z) = I_m + \sum_{j=1}^{\infty} \Psi_j(m)z^j = (I_m - \sum_{j=1}^{\infty} A_j(m)z^j)^{-1}$, $|z| \leq 1$. Let $2\pi \hat{F}_{\omega,m} = \sum_{j=1}^{m} \sum_{l=1}^{m} 1_j^T \hat{\Psi}_{p,m}(e^{-i\omega})\Sigma_{e,p}(m)\overline{\Psi}_{p,m}(e^{-i\omega})1_l(v_j, \cdot)v_l$ where $\Psi_{p,m}(z) = \sum_{j=1}^{\infty} \Psi_{j,p}(m)z^j$, $|z| \leq 1$. Finally recall that

$$2\pi \hat{F}_{\omega,m}^* = \sum_{j=1}^{m} \sum_{l=1}^{m} 1_j^T \hat{\Psi}_{p,m}(e^{-i\omega})\Sigma_{e,p}(m)\overline{\Psi}_{p,m}(e^{-i\omega})1_l\langle v_j, \cdot \rangle\overline{v}_l + E^*U_t^* \otimes U_t^*.$$

Then, (6.6) $\|F_{\omega,m} - F_\omega\|_{HS} \leq \|F_{\omega,m} - \hat{F}_{\omega,m}\|_{HS} + \|\hat{F}_{\omega,m} - F_{\omega,m}\|_{HS} + \|F_{\omega,m} - F_\omega\|_{HS}$.

The first term on the right hand side above is bounded by

$$\|F_{\omega,m} - \hat{F}_{\omega,m}\|_{HS} \leq \|E^*U_t^* \otimes U_t^*\|_{HS}$$

$$++\| \sum_{j,l} 1_j^T \hat{\Psi}_{p,m}(e^{-i\omega})\Sigma_{e,p}(m)\overline{\Psi}_{p,m}(e^{-i\omega})1_l\langle (\hat{v}_j, \cdot)\overline{v}_l - \langle v_j, \cdot \rangle v_l \rangle \|_{HS}$$

$$++\| \sum_{j,l} 1_j^T (\hat{\Psi}_{p,m}(e^{-i\omega})\Sigma_{e,p}(m)\overline{\Psi}_{p,m}(e^{-i\omega}) - \Psi_{p,m}(e^{-i\omega})\Sigma_{e,p}(m)\overline{\Psi}_{p,m}(e^{-i\omega}))1_l\langle v_j, \cdot \rangle v_l \|_{HS}.$$


Furthermore,
\[
\|E^*U_t^* \otimes U_t^*\|_{HS} \leq \|E^+ U_t^+ \otimes U_t^+\|_{HS} + \|E^* U_t^* \otimes U_t^+ - E^+ U_t^+ \otimes U_t^+\|_{HS},
\]
where \(U_t^+\) are i.i.d. random variables taking values with probability \(n^{-1}\) in the set \(\{U_t^c = U_t - \bar{U}_n, t = 1, 2, \ldots, n\}\) and \(\bar{U}_n = n^{-1} \sum_{t=1}^n U_t\). Then
\[
\|E^+ U_t^+ \otimes U_t^+\|_{HS} \leq \frac{1}{n} \sum_{t=1}^n \langle U_t, \cdot \rangle_{U_t} \|_{HS} + \|\langle \bar{U}_n, \cdot \rangle_{\bar{U}_n}\|_{HS}
\]
\[
\leq \sum_{j,l=m+1}^\infty \langle \hat{C}_0(\hat{v}_j) \hat{v}_l, \cdot \rangle_{HS} + \|\bar{U}_n\| \to 0,
\]
in probability, since \(\|\hat{C}_0 - C_0\|_{HS} \to 0\) and the operator \(C_0\) is Hilbert-Schmidt. Furthermore,
\[
\|E^* U_t^* \otimes U_t^* - E^+ U_t^+ \otimes U_t^+\|_{HS} \leq \frac{1}{n} \sum_{t=1}^n \langle \langle \hat{U}_t, \cdot \rangle_{\hat{U}_t} - \langle U_t, \cdot \rangle_{U_t} \rangle_{HS}
\]
\[
+ \|\langle \bar{U}_n, \cdot \rangle_{\bar{U}_n} - \langle U_n, \cdot \rangle_{U_n}\|_{HS}
\]
\[
= O_P \left( \sum_{j=1}^m \|\hat{v}_j - v_j\|^2 \right) \to 0,
\]
in probability, where the last equality follows by straightforward calculations and using \(\hat{U}_t - U_t = \sum_{j=1}^m \langle \langle X_t, \hat{v}_j \rangle_{\hat{v}_j} - \langle X_t, v_j \rangle_{v_j} \rangle\). Similarly, and by the same arguments as above and using Lemma [6.3], we get
\[
\|\sum_{j,l} 1_j \hat{\Psi}_{j,p}(m, e^{-i\omega}) \hat{\Sigma}_{e,p}(m) \hat{\Psi}_{j,p}(m, e^{-i\omega}) 1_l \langle \langle \hat{v}_j, \cdot \rangle_{\hat{v}_l} - \langle v_j, \cdot \rangle_{v_l} \rangle_{HS}
\]
\[
= O_P \left( \sum_{j=1}^m \|\hat{v}_j - v_j\|^2 \right) \to 0,
\]
Finally, straightforward calculations yield
\[
\|\sum_{j,l} 1_j \hat{\Psi}_{j,p}(m, e^{-i\omega}) \hat{\Sigma}_{e,p}(m) \hat{\Psi}_{j,p}(m, e^{-i\omega}) - \Psi_{j,p}(m, e^{-i\omega}) \hat{\Sigma}_{e,p}(m) \hat{\Psi}_{j,p}(m, e^{-i\omega})\|_{HS}
\]
\[
= O_F \left( \sum_{j=1}^m \|\hat{\Psi}_{j,p}(m) - \Psi_{j,p}(m)\|_F + \|\hat{\Sigma}_{e,p}(m) - \hat{\Sigma}_{e,p}(m)\|_F \right) \to 0,
\]
by Lemma [6.5](iii) and (iv). This concludes the proof that \(\|\mathcal{F}^*_{\omega,m} - \hat{\mathcal{F}}_{\omega,m}\|_{HS} \overset{P}{\to} 0\). Consider next the second term on the right hand side of (6.6). For this term we get
\[
\|\hat{\mathcal{F}}_{\omega,m} - \mathcal{F}_{\omega,m}\|_{HS} = O_P \left( \sum_{j=1}^\infty \|\hat{\Psi}_{j,p}(m) - \Psi_{j,p}(m)\|_F \right) + O_P \left( \|\hat{\Sigma}_{e,p}(m) - \Sigma_{e}(m)\|_F \right),
\]
\[ \| \mathcal{F}_{\omega,m} - \mathcal{F}_{\omega,m} \|_{HS} \] converges to zero in probability by Lemma 6.5(v) and (vi). For the third and last term on the right hand side of (6.6) we obtain

\[
\| \mathcal{F}_{\omega,m} - \mathcal{F}_{\omega} \|_{HS} \leq \| \sum_{j=1}^{m} \sum_{l=1}^{m} \langle \mathcal{F}_{\omega}(v_j), v_l \rangle (v_j \otimes v_l) \|_{HS}
\]

\[
+ \| \sum_{j=1}^{m} \sum_{l=m+1}^{\infty} \langle \mathcal{F}_{\omega}(v_j), v_l \rangle (v_j \otimes v_l) \|_{HS}
\]

\[
+ \| \sum_{j=m+1}^{\infty} \sum_{l=m+1}^{\infty} \langle \mathcal{F}_{\omega}(v_j), v_l \rangle (v_j \otimes v_l) \|_{HS} \to 0,
\]

as \( m \to \infty \), since \( \{(v_j \otimes v_l), j = 1, 2, \ldots, l = 1, 2, \ldots \} \) is a complete orthonormal basis of \( \mathcal{H} \otimes \mathcal{H} \).

**Proof of Theorem 4.1**

Let

\[ L_{n,m}^+ = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} \xi_{j,t}^+ v_j e^{-it\omega}, \]

where \( \xi_{t}^+ = (\xi_{1,t}^+, \xi_{2,t}^+, \ldots, \xi_{n,t}^+) \top, t = 1, 2, \ldots, n \) with \( \xi_{t}^+ = \sum_{j=1}^{p} \tilde{A}_{j,p}(m) \xi_{t-j}^* + e_{t}^+ \), where \( \tilde{A}_{j,p}(m), j = 1, 2, \ldots, p \) are the estimators of the autoregressive parameter matrices based on the vector time series of true scores \( \xi_{t}, t = 1, 2, \ldots, n \) and \( e_{t}^+ \) are obtained by i.i.d. resampling from the centered residuals \( \tilde{e}_{t} = \xi_{t} - \sum_{j=1}^{p} \tilde{A}_{j,p}(m) \xi_{t-j}, t = p+1, p+2, \ldots, n \). That is, the pseudo-variable \( L_{n,m}^+ \) is obtained using the true eigenfunctions \( v_j \) and the true scores \( \xi_{j,t} \) instead of their estimates \( \hat{e}_{j} \) and \( \hat{\xi}_{j,t} \) respectively. Decompose then \( n^{-1/2} S_{n}^*(\omega) \) as

\[
n^{-1/2} S_{n}^*(\omega) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} \xi_{j,t}^+ v_j e^{-it\omega} + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} \xi_{j,t}^* (\hat{e}_{j} - v_j) e^{-it\omega}
\]

\[
+ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} (\xi_{j,t}^* - \xi_{j,t}) v_j e^{-it\omega} + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} U_{t,m}^* e^{-it\omega}
\]

\[ = L_{n,m}^+ + V_{n,m}^* + D_{n,m}^* + R_{n,m}^* \]

with an obvious notation for \( L_{n,m}^+, V_{n,m}^*, D_{n,m}^* \) and \( R_{n,m}^* \). Notice that the terms \( V_{n,m}^* \) and \( D_{n,m}^* \) are due to the fact that, in the bootstrap procedure, the unknown scores and eigenfunctions are replaced by their sample estimates, while \( R_{n,m}^* \) is due to the \( m \)-dimensional approximation of the infinite dimensional structure of the underlying process. The assertion of the theorem follows then from Lemma 6.6, 6.7, 6.8 and 6.9 and Slutsky’s lemma.

**Lemma 6.6.** Under the assumptions of Theorem 4.1 it holds true that, \( R_{n,m}^* \to 0 \), as \( n \to \infty \).
Proof: Note that
\[
E^* \| R_{n,m}^* \|^2 = \frac{1}{n} \sum_{t=1}^{n} \| \hat{U}_{t,m} - \overline{U}_n \|^2 
\leq \frac{2}{n} \sum_{t=1}^{n} \| \hat{U}_{t,m} \|^2 + 2 \| \overline{U}_n \|^2. 
\]
Using \( \| \hat{v}_j - v_j \| \leq 2\sqrt{2} \alpha_j^{-1} \| \hat{C}_0 - C_0 \|_{HS} \), see Hörmann and Kokoszka (2010), we get
\[
\frac{1}{n} \sum_{t=1}^{n} \| \hat{U}_{t,m} \|^2 \leq \frac{4}{n} \sum_{t=1}^{n} \| \sum_{j=1}^{m} \langle X_t, v_j \rangle (v_j - \hat{v}_j) \|^2 + \frac{4}{n} \sum_{t=1}^{n} \| \sum_{j=1}^{m} \langle X_t, v_j - \hat{v}_j \rangle \hat{v}_j \|^2 
\leq 4 \| \hat{C}_0 \|_{HS} \left( \sum_{j=1}^{m} \| \hat{v}_j - v_j \| \right)^2 + 4 \| \hat{C}_0 \|_{HS} \sum_{j=1}^{m} \| \hat{v}_j - v_j \|^2 
\leq 32 \| \hat{C}_0 \|_{HS} \| \hat{C}_0 - C_0 \|^2_{HS} \left( \sum_{j=1}^{m} \alpha_j^{-1} \right)^2 + \sum_{j=1}^{m} \alpha_j^{-2} 
= O_P(\sqrt{n}^{-1/2} \sum_{j=1}^{m} \alpha_j^{-1}),
\]
where the last equality follows because \( \| \hat{C}_0 - C_0 \|_{HS} = O_P(n^{-1/2}) \). Furthermore, \( \| \hat{U}_n \|^2 \Rightarrow 0 \) follows using similar arguments and since \( \hat{U}_n = \overline{U}_n + n^{-1} \sum_{t=1}^{n} \sum_{j=1}^{m} \langle X_t, v_j \rangle v_j - \langle X_t, \hat{v}_j \rangle \hat{v}_j \), where \( \overline{U}_n = n^{-1} \sum_{t=1}^{n} U_{t,m} \).

Lemma 6.7. Under the assumptions of Theorem 4.1 it holds true that, \( D_{n,m}^* \Rightarrow 0 \), as \( n \to \infty \).

Proof: We have
\[
E \left[ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} (\xi_{j,t}^* - \xi_{j,t}^+) v_j \right]^2 = \frac{1}{n} \sum_{t,s=1}^{n} \sum_{j=1}^{m} 1_j^T E \xi_{t,s}^* (\xi_{s}^* - \xi_{s}^+) 1_j 
+ \frac{1}{n} \sum_{t,s=1}^{n} \sum_{j=1}^{m} 1_j^T E \xi_{t,s}^* (\xi_{s}^* - \xi_{s}^+) 1_j = D_{n,m}^{(1)} + D_{n,m}^{(2)},
\]
with an obvious notation for $D_{n,m}^{(1)}$ and $D_{n,m}^{(2)}$. We consider $D_{n,m}^{(1)}$ only since $D_{n,m}^{(2)}$ can be handled similarly. For this term we have

$$D_{n,m}^{(1)} = \frac{1}{n} \sum_{t,s=1}^{n} \sum_{j=1}^{m} \sum_{l=0}^{\infty} 1_j^T \tilde{\Psi}_{t,p}(m) \Sigma_{e,p}(m) (\tilde{\Psi}_{t+s-t,p}(m) - \tilde{\Psi}_{t+s-t,p}(m))^T 1_j$$

(6.7)

and, using Lemma 6.1 and 6.5, we get for the first term on the right hand side of (6.7), that, this term is bounded by

$$\| \Sigma_{e,p}(m) \|_F \sum_{l=0}^{\infty} \| \sum_{j=1}^{m} 1_j^T \tilde{\Psi}_{t,p}(m) \|_F \sum_{l=0}^{\infty} \| \sum_{j=1}^{m} (\tilde{\Psi}_{t,p}(m) - \tilde{\Psi}_{t,p}(m)) \|_F \to 0,$$

in probability. The second term of (6.7) is bounded by

$$\sqrt{E\| e_{t,p}^*(m) \|_2^2} \sqrt{E\| e_{t,p}^*(m) - e_{t,p}^+(m) \|_2^2} \sum_{l=0}^{\infty} \| \sum_{j=1}^{m} 1_j^T \tilde{\Psi}_{t,p}(m) \|_F \sum_{l=0}^{\infty} \| \sum_{j=1}^{m} 1_j^T \tilde{\Psi}_{t,p}(m) \|_F,$$

which converges to zero in probability, because $E\| e_{t,p}^*(m) - e_{t,p}^+(m) \|_2^2 \to 0$ in probability. This can be seen using

$$E\| e_{t,p}^*(m) - e_{t,p}^+(m) \|^2 \leq \frac{2}{n-p} \sum_{t=p+1}^{n} \| \hat{e}_{t,p}(m) - \tilde{e}_{t,p}(m) \|^2 + 4(\| \bar{e}_n \|^2 + \| \bar{e}_n \|^2)$$

$$\leq \frac{4}{n-p} \sum_{t=p+1}^{n} \| \hat{\xi}_t - \xi_t \|^2$$

$$+ \frac{4}{n-p} \sum_{t=p+1}^{n} \| \sum_{j=1}^{p} (\hat{A}_{j,p}(m) \hat{\xi}_{t-j} - \tilde{A}_{j,p}(m) \xi_{t-j}) \|^2$$

$$+ 4(\| \bar{e}_n \|^2 + \| \bar{e}_n \|^2).$$

We then have

$$\frac{1}{n-p} \sum_{t=p+1}^{n} \| \hat{\xi}_t - \xi_t \|^2 \leq \frac{1}{n-p} \sum_{t=p+1}^{n} \| X_t \|^2 \sum_{j=1}^{m} \| \tilde{v}_j - v_j \|^2 = O_P\left(n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}\right) \to 0.$$
Furthermore,

\[
\frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \sum_{j=1}^{p} (\hat{A}_{j,p}(m)\hat{\xi}_{t-j} - \tilde{A}_{j,p}(m)\xi_{t-j}) \right\|^2 \\
\leq 2 \sum_{j=1}^{p} \left\| \hat{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} - \xi_{t-j} \right\|_F \\
+ 2 \sum_{j=1}^{p} \left\| \hat{A}_{j,p}(m) - \tilde{A}_{j,p}(m) \right\|^2 \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \xi_{t-j} \right\|^2 \\
= O_P\left(n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}\right) + O_P\left(\lambda_m^{-2}n^{-1}mp \sum_{j=1}^{m} \alpha_j^2\right) \to 0,
\]

where the last equality follows using Lemma 6.1 and 6.3. Finally,

\[
\left\| \hat{e}_n \right\|^2 \leq 2 \left\| \frac{1}{n-p} \sum_{t=p+1}^{n} \hat{\xi}_t \right\|^2 + 2\left( \sum_{j=1}^{p} \left\| \hat{A}_{j,p}(m) \right\|_F \right)^2 \left( \frac{1}{n-p} \sum_{t=p+1}^{n} \left\| \hat{\xi}_{t-j} \right\|^2 \right) \to 0
\]
in probability, since

\[
\left\| \frac{1}{n-p} \sum_{t=p+1}^{n} \hat{\xi}_t \right\|^2 \leq 2 \left\| \frac{1}{n-p} \sum_{t=p+1}^{n} \xi_t \right\|^2 + O_P(n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}) \\
= O_P\left(m/(n-p)\right) + O_P\left(n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}\right) \to 0,
\]

and

\[
\sum_{j=1}^{p} \left\| \hat{A}_{j,p}(m) \right\|_F \frac{1}{n-p} \sum_{t=p+1}^{n} \hat{\xi}_{t-j} = O_P(1)O_P\left(\sqrt{m/(n-p)} + \sqrt{n^{-1} \sum_{j=1}^{m} \alpha_j^{-2}}\right) \to 0.
\]

By similar arguments we get \( \left\| \tilde{e}_n \right\|^2 \to 0, \) in probability.

**Lemma 6.8.** Under the assumptions of Theorem 4.1 it holds true that, \( V_{n,m}^* \xrightarrow{P} 0, \) as \( n \to \infty. \)

**Proof:**

\[
E\left[ \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sum_{j=1}^{m} \xi_{j,t}^* (\hat{v}_j - v_j) \right\|^2 \right] = \sum_{j=1}^{m} \sum_{t=1}^{n} \frac{1}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} \Gamma_{t-s}^* 1_t (\hat{v}_j - v_j, \tilde{v}_j - v_t) \\
\leq \left( \sum_{j=1}^{m} \left\| \hat{v}_j - v_j \right\|^2 \right) \frac{1}{n} \sum_{t=1}^{n} \sum_{s=1}^{n} \left\| \Gamma_{t-s}^* \right\|_F \\
= O_P\left(\left(n^{-1/2} \sum_{j=1}^{m} \alpha_j^{-1}\right)^2\right) \to 0.
\]

\[\square\]
Lemma 6.9. Under the assumptions of Theorem 4.1 it holds true that, for all $\omega \in [-\pi, \pi]$ and as $n \to \infty$, 

$$L_{n,m}^+(\omega) \Rightarrow NC(0, \pi F_\omega),$$

in probability.

Proof: Write $L_{n,m}^+(\omega) = n^{-1/2} \sum_{t=1}^n W_t^+ e^{-it\omega}$ where $W_t^+ = \sum_{j=1}^m 1_j^T \xi_t^+ v_j$ with $\xi_t^+ = \sum_{l=0}^\infty \tilde{\Psi}_{l,p}(m)e_{l-t}^+$, $\tilde{\Psi}_{0,p}(m) = I_m$, a random element in $\mathcal{H}$. Notice that $E^+(W_t^+) = 0$, while using $\xi_t = \sum_{l=0}^\infty \Psi_t(m)e_{l}, \Psi_0(m) = I_m$, we get

$$E^+ W_t^+ \otimes W_{t+h}^+ = \sum_{l=0}^\infty \sum_{j=1}^m 1_j^T \tilde{\Psi}_{l,p}(m) \tilde{\Sigma}_{e,p}(m) \tilde{\Psi}_{l+h,p}(m) \mathbf{1}_s \langle v_j, \cdot \rangle v_s$$

$$= \sum_{l=0}^\infty \sum_{j=1}^m 1_j^T \Psi_t(m) \Sigma_e(m) \Psi_{l+h,p}(m) \mathbf{1}_s \langle v_j, \cdot \rangle v_s + \tilde{D}_n$$

$$= E_{\xi} \langle X_t - U_{t,m}, \cdot \rangle X_{t+h} - E_{\xi} \langle X_t, \cdot \rangle U_{t+h,m} + E_{\xi} \langle U_{t,m}, \cdot \rangle U_{t+h,m} + \tilde{D}_n,$$

with an obvious notation for $\tilde{D}_n$. It is easily seen that $\tilde{D}_n = O_p(\| \sum_{l=0}^\infty \| \tilde{\Psi}_{l,p}(m) - \Psi_l(m) \|_F + \| \sum_{l=0}^\infty \Sigma_{e,p}(m) - \Sigma_e(m) \|_F )$ and therefore $\| \tilde{D}_n \|_{HS} \to 0$ in probability, by Lemma 6.5. Hence and using $E \| U_{t,m} \|^2 \to 0$ as $m \to \infty$, we get that $\| E^+ W_t^+ \otimes W_{t+h}^+ - C_{\xi} \|_{HS} \to 0$ in probability, as $n \to \infty$.

Let $\xi_t^o = \sum_{l=0}^\infty \Psi_t(m)e_{l-t}^+$, define $W_t^o = \sum_{j=1}^m 1_j^T \xi_t^o v_j$ and $L_{n,m}^o(\omega) = n^{-1/2} \sum_{l=0}^n W_l^o e^{-it\omega}$. It easily follows by simple algebra and using Lemma 6.5 that $E^+ \| L_{n,m}^o(\omega) - L_{n,m}^o(\omega) \| = O_p(\sum_{l=0}^\infty \| \tilde{\Psi}_{l,p}(m) - \Psi_l(m) \|_F ) \to 0$, in probability, that is $L_{n,m}^o(\omega) = L_{n,m}^o(\omega) + o_p(1)$. Thus to prove the assertion of the lemma it suffices to show that $L_{n,m}^o(\omega) \Rightarrow NC(0, \pi F_\omega)$. For this we show that the assumptions of Theorem 2 of Cerovecki and Hörmann (2015) are satisfied, that is, using the notation $S_{n,m}^o(\omega) = \sum_{l=1}^n W_l^o e^{-it\omega}$, we show that the following two conditions are fulfilled, in probability.

$$(6.8) \quad Z_n^o(\omega) \equiv \sum_{t=0}^n \mathcal{P}_0(W_t^o)e^{-it\omega} \quad \text{is a Cauchy sequence in } \mathcal{H},$$

and

$$(6.9) \quad E \| E(S_{n,m}^o(\omega) | \mathcal{G}_0) \|^2 = o(n),$$

where the operator $\mathcal{P}_0$ is defined as $\mathcal{P}_0(\cdot) = E(\cdot | \mathcal{G}_0) - E(\cdot | \mathcal{G}_{-1})$ and $\mathcal{G}_s = \sigma(W_s^o, W_{s-1}^o, W_{s-2}^o, \ldots)$. Toward this we first define $W_{s,s}^o = \sum_{j=1}^m 1_j^T \sum_{l=0}^\infty \Psi_t(m) e_{l-s}^+ v_{j}$, where $e_{l,s}^+ = e_l^+$ if
\( t > 0 \) and \( e^+_t, s = \tilde{e}_t \) if \( t \leq 0 \) with \( \tilde{e}_t \) a copy of \( e^+_t \) which is independent of \( e^+_t \) for \( t < 0 \). We show that

\[
\sum_{s=1}^{\infty} \sqrt{E^+ \|W_0^o - W_{0,s}^o\|^2} = O_P(1),
\]

where the \( O_P(1) \) term is independent of \( m \) and \( p \). Notice first that by Minkowski’s inequality

\[
\sqrt{E^+ \|W_0^o - W_{0,s}^o\|^2} \leq \sqrt{E^+ \sum_{j=1}^{m} 1_j^\top \sum_{l=s}^{\infty} \Psi_l(m) (e^+_l - e^+_{l,s}) v_j} + \sqrt{E^+ \sum_{j=1}^{m} 1_j^\top \sum_{l=s}^{\infty} \Psi_l(m) e^+_l v_j} \leq 2 \sum_{l=s}^{\infty} \text{tr} (\Psi_l(m) \tilde{\Sigma}_{e,p}(m) \Psi_l^\top(m)) \leq 2 \|\tilde{\Sigma}_{e,p}(m)\|_F \sum_{l=s}^{\infty} \|\Psi_l(m)\|_F.
\]

Thus

\[
\sum_{s=1}^{\infty} \sqrt{E^+ \|W_0^o - W_{0,s}^o\|^2} \leq 2 \sum_{s=1}^{\infty} \|\tilde{\Sigma}_{e,p}(m)\|_F \sum_{l=s}^{\infty} \|\Psi_l(m)\|_F \leq 2 \|\tilde{\Sigma}_{e,p}(m)\|_F \sum_{l=1}^{\infty} l \|\Psi_l(m)\|_F.
\]

Now, since by Lemma 6.1(ii), \( \sum_{l=1}^{\infty} l \|\Psi_l(m)\|_F \) is bounded uniformly in \( m \), and, by Lemma 6.5(ii) and (vi) and Lemma 6.1(iii), \( \|\tilde{\Sigma}_{e,p}(m)\|_F \) is bounded in probability, where the bound is independent of \( p \) and \( m \), assertion (6.10) follows.

Consider next condition (6.8). For positive integers \( n_2 > n_1 \) we have that

\[
E^+ \|Z_{n_2}^o(\omega) - Z_{n_1}^o(\omega)\|^2 \leq \sum_{t_1=n_1+1}^{n_2} \sum_{t_2=n_1+1}^{n_2} |E^+ \langle P_0(W_{t_1}^o), P_0(W_{t_2}^o) \rangle| \leq \left( \sum_{t=n_1+1}^{n_2} \sqrt{E^+ \|P_0(W_t^o)\|^2} \right)^2.
\]
Recall the definition of $W_{o,s,s}$. Then we have, since $E(W_{o,s,s} | G_0) = E(W_{o,s,s} | G_{-1}) = 0$, that

$$E^+ \|P_0(W_t)\|^2 = E^+ \|P_0(W_t) - P(W_{s,s})\|^2$$

$$= E^+ \|E^+(W_t - W_s | G_0) - E(W_t - W_s | G_{-1})\|^2$$

$$\leq 2E^+ \|E^+(W_t - W_s | G_0)\|^2 + 2E^+ \|E(W_t - W_s | G_{-1})\|^2$$

$$\leq 4E^+ \|W_t - W_{s,s}\|^2$$

Hence

$$E^+ \|Z_{n_2}^o (\omega) - Z_{n_1}^o (\omega)\|^2 \leq 4 \left( \sum_{s=n_1}^{\infty} \sqrt{E^+ \|W_0 - W_{0,s}\|^2} \right)^2 \to 0,$$

as $n_1 \to \infty$ because of (6.10).

To establish condition (6.9) notice that

$$E^+ \|E(S_{n,m}^o (\omega) | G_0)\|^2 \leq \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} |E^+ \langle E^+(W_{t_1} | G_0), E^+(W_{t_2} | G_0) \rangle|$$

$$= \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} |E^+ \langle E^+(W_{t_1}^o - W_{t_1}^o | G_0), E^+(W_{t_2}^o - W_{t_2}^o | G_0) \rangle|$$

$$\leq \left( \sum_{t=1}^{n} \sqrt{E^+ \|W_0^o - W_{0,t}\|^2} \right)^2$$

$$\leq \left( \sum_{t=1}^{\infty} \sqrt{E^+ \|W_0^o - W_{0,t}\|^2} \right)^2,$$

which is bounded because of (6.10). \(\square\)

REFERENCES

[1] ANEIROS-PÉREZ, G., CAO, R. and VILAR-FERNANÉZ, J. M. (2011). Functional methods for time series prediction: A nonparametric approach. Journal of Forecasting, 30, 377-392.

[2] AUE, A., Dubart, D. N. and HÖRMANN, S. (2015). On the prediction of stationary functional time series. Journal of the American Statistical association, 110, 378-392.

[3] BOSQ, D. (2000). Linear Process in Function Spaces. Springer, Berlin-Heidelberg-New York.

[4] CEROVECKI, C. and HÖRMANN, S. (2015). On the CLT for discrete Fourier transforms of functional time series. Preprint.

[5] CHENG, R. and POURAHMADI, M. (1993). Baxter’s inequality and convergence of finite predictors of multivariate stochastic processes. Probability Theory and Related Fields, 95, 115-124.

[6] DEHLING, H., SHARIPOV, O. S. and WENDLER, M. (2015). Bootstrap for dependent Hilbert space valued random variables with application to von Mises statistics. Journal of Multivariate Analysis, 133, 200-215.
[7] FERNÁNDEZ DE CASTRO, B., GUILLAS, S. and GONZÁLEZ MANTEIGA, W. (2005). Functional samples and bootstrap for predicting sulfur dioxide levels. Technometrics, 47, 212-222.

[8] FRANKE, J. and NYARIGUE, E. (2016). Residual-based bootstrap for functional autoregressions. Preprint.

[9] HÖRMANN, S. and KOKOSZKA, P. (2010). Weakly dependent functional data. Annals of Statistics, 38, 1845-1884.

[10] HÖRMANN, S. and KOKOSZKA, P. (2012). Functional time series. Handbook of Statistics: Time Series Analysis-Methods and Applications, 157-186.

[11] HÖRMANN, S., KIDZIŃSKI, L. and HALLIN, M. (2015). Dynamic functional principal components. Journal of the Royal Statistical Society: Series B, 77, 319-348.

[12] HORVÁTH, L. and KOKOSZKA, P. (2012). Inference for Functional Data with Applications. Springer, Berlin-Heidelberg-New York.

[13] HYNDMAN, R. J. and SHANG, H. L. (2009). Forecasting functional time series. Journal of the Korean Statistical Society, 38, 199-211.

[14] IPSEN, I. C. F. and REHMAN, R. (2008). Perturbation bounds for determinants and characteristic polynomials. SIAM Journal of Matrix Analysis and Applications, 30, 762-776.

[15] KREISS, J.-P. (1988). Asymptotic Statistical Inference for a Class of Stochastic Processes. Habilitationsschrift. Universität Hamburg.

[16] KREISS, J.-P., PAPARODITIS, E. and POLITIS, D. N. (2011). On the range of validity of the autoregressive sieve bootstrap. Annals of Statistics, 39, 2103-2130.

[17] McMURRY, T. and POLITIS, D. N. (2011). Resampling methods for functional data. The Oxford Handbook of Functional Data Analysis, (F. Ferraty and Y. Romain, eds.) 189-209. Oxford University Press.

[18] MERIKOSKI, J. K. and KUMAR, R. (2005). Upper bounds for singular values. Linear Algebra and its Applications, 401, 371-379.

[19] MERLEVEDE, F., PELIGRAD, M. and UTEV, S. (1997). Sharp conditions for the CLT of linear processes in a Hilbert space. Journal of Theoretical Probability, 10, 681-693.

[20] MEYER, M. and KREISS, J.-P. (2015). On the vector autoregressive sieve bootstrap. Journal of Time Series Analysis, 36, 377-397.

[21] MEYER, M., JENTSCH, C., and KREISS, J.-P. (2016). Baxter’s inequality and sieve bootstrap for random fields. Bernoulli, to appear.

[22] MINGOTTI, N., LILLO, R. E. and ROMO, J. (2015). A random walk test for functional time series. UC3M Working papers, Statistics and Econometrics.

[23] PANARETOS, V. and TAVAKOLI, S. (2013). Fourier analysis of stationary time series in function spaces. Annals of Statistics, 41, 568-603.

[24] POLITIS, D. N. and ROMANIO, J. (1994). Limit theorems for weakly dependent Hilbert space valued random variables with applications to the stationary bootstrap. Statistica Sinica, 4, 461-476.

[25] POURAHMADI, M. (2001). Foundation of Time Series Analysis and Prediction Theory. John-Wiley: New York.

[26] SHARIPOV, O., TEWES, J. and WENDLER, M. (2016). Sequential block bootstrap in a Hilbert space with application to change point analysis. The Canadian Journal of Statistics, forthcoming.

[27] WIENER, N. and MASANI, P. (1957). The prediction theory of multivariate stochastic processes, I. Acta Mathematica, 98, 111-150.

[28] WIENER, N. and MASANI, P. (1958). The prediction theory of multivariate stochastic processes, II. Acta Mathematica, 99, 93-137.
[29] SHANG, L. H. (2016). Resampling Methods for Dependent Functional Data. *Preprint*.

[30] ZHOU, T. and POLITIS, D. N. (2016). Kernel Estimation of First-Order Nonparametric Functional Autoregression Model and its Bootstrap Approximation. *Preprint*. 
Table 1: Estimated exact ($\sigma_{EE}(\tau_j)$) and functional sieve bootstrap (FSB) estimates of the standard deviation of the sample mean $\mathbb{X}_n(\tau_j)$ for different values of $\tau_j \in [0, 1]$ and for different parameters $m$ and $p$. $\hat{\sigma}(\tau_j)$ refers to the mean, while $S(\hat{\sigma}(\tau_j))$ to the standard deviation of the FSB estimates obtained for $R = 1000$ replications.

| $\tau_j$ | $\sigma_{EE}(\tau_j)$ | $m=2, \ p=3$ | $m=3, \ p=3$ | $m=4, \ p=3$ | $m=3, \ p=5$ |
|---|---|---|---|---|---|
| 0.00 | 2.149 | 2.124 0.392 | 2.188 0.440 | 2.193 0.468 | 2.105 0.503 |
| 0.05 | 2.203 | 2.172 0.404 | 2.227 0.441 | 2.243 0.473 | 2.131 0.509 |
| 0.10 | 2.272 | 2.262 0.441 | 2.305 0.458 | 2.339 0.486 | 2.189 0.537 |
| 0.15 | 2.325 | 2.362 0.466 | 2.385 0.477 | 2.408 0.504 | 2.243 0.550 |
| 0.20 | 2.358 | 2.429 0.484 | 2.434 0.493 | 2.424 0.508 | 2.273 0.555 |
| 0.25 | 2.370 | 2.457 0.488 | 2.452 0.488 | 2.417 0.515 | 2.286 0.594 |
| 0.30 | 2.351 | 2.429 0.488 | 2.432 0.485 | 2.416 0.517 | 2.266 0.542 |
| 0.35 | 2.317 | 2.359 0.462 | 2.382 0.471 | 2.390 0.510 | 2.221 0.530 |
| 0.40 | 2.267 | 2.271 0.435 | 2.308 0.448 | 2.324 0.494 | 2.170 0.524 |
| 0.45 | 2.196 | 2.183 0.419 | 2.237 0.439 | 2.240 0.461 | 2.116 0.509 |
| 0.50 | 2.146 | 2.123 0.401 | 2.199 0.433 | 2.190 0.443 | 2.088 0.500 |
| 0.55 | 2.194 | 2.165 0.405 | 2.241 0.440 | 2.250 0.467 | 2.133 0.507 |
| 0.60 | 2.264 | 2.249 0.419 | 2.309 0.459 | 2.345 0.488 | 2.196 0.514 |
| 0.65 | 2.314 | 2.342 0.441 | 2.370 0.468 | 2.404 0.517 | 2.267 0.536 |
| 0.70 | 2.343 | 2.408 0.464 | 2.418 0.487 | 2.428 0.517 | 2.304 0.544 |
| 0.75 | 2.351 | 2.429 0.475 | 2.430 0.494 | 2.423 0.505 | 2.320 0.557 |
| 0.80 | 2.342 | 2.405 0.474 | 2.413 0.481 | 2.421 0.509 | 2.311 0.554 |
| 0.85 | 2.309 | 2.346 0.459 | 2.364 0.474 | 2.403 0.513 | 2.277 0.542 |
| 0.90 | 2.258 | 2.262 0.431 | 2.299 0.457 | 2.342 0.503 | 2.215 0.520 |
| 0.95 | 2.188 | 2.167 0.399 | 2.227 0.444 | 2.251 0.485 | 2.150 0.507 |
| 1.00 | 2.149 | 2.123 0.392 | 2.188 0.440 | 2.193 0.468 | 2.105 0.503 |
Figure 1. Comparison of different bootstrap estimates of the standard deviation of the sample mean $\bar{X}_n(\tau)$ for different values of $\tau_j \in [0, 1]$. Bullets refer to the estimated exact standard deviation, while the mean estimates of the standard deviation of the FSB ($m = 2, p = 3$) are presented by circles, of the TBB ($b = 5$), by diamonds, of the MBB ($b = 7$), by triangles and of the SB ($p = 0.064$), by “+”. 