Julia Kropiunig, Bsc

The Generalized Dynamic Factor Model

Master's Thesis

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Mag.rer.nat Dr.rer.nat Siegfried Hörmann

Institute for Statistics

Graz, May 2021

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date

Signature

Abstract

This thesis will take a thorough look at the generalized dynamic factor model. Under this model the observable vector process $\{X_t\}$ is split into two non-observable vector processes $\{\chi_t\}$ and $\{\xi_t\}$, the common and the idiosyncratic component. The common component is driven by an underlying q-dimensional white noise vector process $\{u_t\}$ which is common to all components of $\{\chi_t\}$. Compared to other models in the area of dynamic factor models, this model allows for correlation of the idiosyncratic component across time and sections and is based on an infinite number of sections, making it suitable for analyzing high-dimensional vector processes. This thesis focuses on two main results, namely the characterization of the generalized dynamic factor model via the eigenvalues of the spectral density corresponding to the observable vector process and the consistent estimation of the common component in the population case as well as in the sample case.

Contents

List of Acronyms and Symbols

$\mathbb{N} = \{1, 2, \ldots\}$	
$\overline{A}$	complex conjugate of matrix $A \in \mathbb{C}^{q\times r}$
A'	transpose of matrix $A \in \mathbb{C}^{q\times r}$
A_{ij}	entry in the i-th row and the j-th column of matrix $A \in \mathbb{C}^{q\times r}$
p_i or $(p)_i$	i-th entry of vector $p \in \mathbb{C}^q$
$\mathrm{Re}(X)$	real part of $X \in \mathbb{C}$
$\mathrm{Im}(X)$	imaginary part of $X \in \mathbb{C}$
I_q	identity matrix of dimension $q \times q$
$A > 0$	matrix A is positive definite
$A \geq 0$	matrix A is semi-positive definite
$\mathcal{N}(\mu, \Gamma)$	normal distribution with mean μ and covariance Γ
$\mathcal{P} = (\Omega, \mathcal{F}, \mathrm{P})$	probability space on Ω with σ-algebra $\mathcal{F}$ and probability function P
$L_2(\mathcal{P}, \mathbb{C})$	linear space of all complex-valued, zero-mean, measurable and square-integrable with respect to the Lebesgue measure random variables on $(\Omega, \mathcal{F}, \mathrm{P})$
$\overline{\mathrm{span}}(Q)$	for $Q \subset L_2(\mathcal{P}, \mathbb{C})$ minimum closed linear subspace of $L_2(\mathcal{P}, \mathbb{C})$ containing Q
$x \perp y$	x and y are orthogonal to one another
$Z \sim D$	random variable Z follows distribution D

1 Introduction and Motivation

Over the past few decades more and more data sets are being collected and therefore their analysis is becoming increasingly important. A dominant class of such data sets consists of the multivariate time series, i.e. information across time and sections, such as, for example, the returns of different shares. There are good tools and models to analyze the isolated univariate time series, however the components of the time series are intercorrelated and comprise information that would go to waste if the components of the time series would be investigated on their own. We now take a look at same concepts for modelling such structures and discuss their weaknesses. A very powerful model is the following.

The n-dimensional stationary zero-mean vector process $\{X_t\}$ follows a vector autoregressive model of order p, short VAR(p), if

$$X_t = \sum_{j=1}^{p} \Phi_j X_{t-j} + \epsilon_t,$$

where Φ_j is an $(n \times n)$-matrix for $j \in \{1, \ldots, p\}$ and $\{\epsilon_t\}$ is an n-dimensional white noise vector process, i.e. $\mathrm{E}(\epsilon_t) = 0, \mathrm{Var}(\epsilon_t) = \Gamma, \mathrm{Cov}(\epsilon_t, \epsilon_{t-j}) = 0$ for any $t \in \mathbb{Z}$ and any $j \in \mathbb{Z} \setminus \{0\}$, [1, p.27]. There are good methods to analyze process $\{X_t\}$ under such a model, i.e. the Φ_j's can be estimated via the Yule-Walker equations. However, this approach becomes impractical for large n, since the parameters to estimate also grow quadratically with respect to n. A better suited approach would be the use of factors. Recall that a static

q-factor model for an n-dimensional random variable X with $\mathrm{E}(X) = 0$ and $\mathrm{Var}(X) = \Gamma$, where $q < n$ holds if X can be written as

$$X = AZ + \epsilon,$$

where $Z = (z_1, \ldots, z_q)'$ is an unobservable q-dimensional zero-mean random vector with $\mathrm{Var}(Z) = I_q$, A is a constant $(n \times q)$-matrix and ϵ is an n-dimensional zero-mean random variable with a diagonal covariance matrix such that $\mathrm{E}(Z\epsilon') = 0$, [2, p.353f].
In 1977 Sargent and Sims introduced the so-called index model. This model can be seen as an extension of the factor model from the random variable case to the vector process case. Likewise, this model relies on an unobservable stationary common q-dimensional vector process $\{Z_t\}$ in the following way

$$X_t = A(L)Z_t + \epsilon_t,$$

where $A(L) = \sum_{k=-\infty}^{\infty} A_k L^k$ is one-sided, i.e. $A_k = 0$ for $k < 0$ and where $z_{it_1} \perp \epsilon_{jt_2}$ for any $t_1, t_2 \in \mathbb{Z}$ and any $i, j \in \{1, \ldots, n\}$ as well as $\epsilon_{it_1} \perp \epsilon_{jt_2}$ for any $t_1, t_2 \in \mathbb{Z}$, $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $\epsilon_{it_1} \perp \epsilon_{it_2}$ for $i \in \{1, \ldots, n\}$, $t_1, t_2 \in \mathbb{Z}$ and $t_1 \neq t_2$, according to [3]. The orthogonality of $\{\epsilon_t\}$ at any lead and lag is a serious restriction, which will be dropped in the *generalized dynamic factor model*; together with furthermore allowing for infinitely many sections this model "generalizes" the index model. The main goal of this thesis is to introduce and discuss results concerning the generalized dynamic factor model proposed by Forni, Lippi, Hallin and Reichlin in [4] and [5].

2 Preliminaries

The following chapter is a preparatory one. Fundamental definitions surfacing when it comes to time series analysis such as stationarity, covariance function, spectral density are recalled and their properties discussed. Then, the dynamic eigenvalues and their corresponding dynamic eigenvectors are introduced. Furthermore, the resulting dynamic principal components are defined and their use is motivated. Lastly, results and theorems with importance for subsequent chapters are given.

2.1 Multivariate time series

Definition 2.1. An *n-variate stochastic process* is a family of n-dimensional random vectors $\{X_t : t \in T\}$, where $X_t = (x_{1t}, \dots, x_{nt})'$ and x_{it} is defined on a common probability space $(\Omega, \mathcal{F}, \mathrm{P})$ for $i \in \{1, \dots, n\}$ and $t \in T$. See [6, p.8].

Remark. A typical choice for T are the integers $\mathbb{Z}$, in this thesis too. From now on only $\{X_t : t \in \mathbb{Z}\}$ is considered. The term (n-variate stochastic) vector process will be used interchangeably with the term multivariate time series. Furthermore, $\{X_t : t \in T\}$ will be abbreviated as $\{X_t\}$. See [6, p.8f].

Suppose now that second order moments for the vector process $\{X_t\}$ exist, i.e $\mathrm{E}(x_{it}^2) < \infty$ for all $i \in \{1, \dots, n\}, t \in \mathbb{Z}$. The second order properties are then specified by the mean vector

$$\mu_t = \mathrm{E}(X_t) = (\mathrm{E}(x_{1t}),\ \dots\ , \mathrm{E}(x_{nt}))' = (\mu_{1t},\ \dots\ , \mu_{nt})',$$

and the covariances

$$\Gamma(t+h,t) = \mathrm{E}((X_{t+h} - \mu_{t+h})(\overline{X_t - \mu_t})') = \Gamma_{ij}(t+h,t)_{i,j=1}^{n},$$

see [6, p.402f].
Even though in practice real valued time series are primarily investigated, it makes sense to develop the theory more generally by including the complex numbers. In particular this makes sense when it comes to the Fourier transformation in which we inevitably have to leave the field of real numbers.

Definition 2.2. If second order moments exist for the stochastic vector process $\{X_t\}$, it is said to be *(weakly) stationary* if

- $\mathrm{E}(X_t) = m, \qquad \forall t \in \mathbb{Z},$
- $\Gamma(t+h,t) = \Gamma(s+h,s) =: \Gamma^X(h), \qquad \forall s,t,h \in \mathbb{Z}.$

The latter function is also called *covariance (matrix) function* of process $\{X_t\}$, see [6, p.12,402f].

Definition 2.3. The n-dimensional vector process $\{Z_t : t \in \mathbb{Z}\}$ is said to be *white noise* with mean 0 and covariance matrix Γ_0 if $\{Z_t\}$ is stationary with mean 0 and covariance function

$$\Gamma^Z(h) = \begin{cases} \Gamma_0 & h = 0, \\ 0 & h \neq 0. \end{cases}$$

See [6, p.404].

Definition 2.4. Let the covariance matrix function $\Gamma^X(h)$ of process $\{X_t\}$ be such that $\sum_{h=-\infty}^{\infty} |\Gamma_{ij}^X(h)| < \infty$ for all $i,j \in \{1,...,n\}$, then $\Gamma^X(h)$ has a *spectral density matrix function*, or short *spectral density*, $\Sigma^X(\theta)$ which is given by

$$\Sigma^X(\theta) := \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\theta h}\Gamma^X(h), \qquad \theta \in [-\pi,\pi].$$

See [6, p.405].

Remark. It is easy to see that $\Gamma^X(h)$ can be expressed in terms of $\Sigma^X(\theta)$ as

$$\Gamma^X(h) = \int_{-\pi}^{\pi} e^{i\theta h}\Sigma^X(\theta)d\theta.$$

For further information see [7, p.24].

Definition 2.5. Let $\{X_t^1\}$, $\{X_t^2\}$ be two vector processes of dimensions p_1 and p_2. The processes $\{X_t^1\}$ and $\{X_t^2\}$ are said to be *costationary* if the combined process $\{(X_t^{1\prime}\, X_t^{2\prime})'\}$ is again stationary.

Remark. If processes $\{X_t^1\}$, $\{X_t^2\}$ are costationary, $\{X_t^1\}$ and $\{X_t^2\}$ are stationary.

Definition 2.6. Let $\{X_t^1\}$ and $\{X_t^2\}$ be two vector processes with mean 0 and covariance functions $\Gamma^1(h)$ resp. $\Gamma^2(h)$ satisfying $\sum_{h=-\infty}^{\infty} |\Gamma_{ij}^k(h)| < \infty$ for $k \in \{1,2\}$ and $i,j \in \{1,\dots,n\}$. Furthermore let $\{X_t^1\}$ and $\{X_t^2\}$ be costationary. Then the *cross-spectrum* $\Sigma^{X_1X_2}(\theta)$ between $\{X_t^1\}$ and $\{X_t^2\}$ is given by

$$\Sigma^{X_1X_2}(\theta) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\theta h}\mathrm{E}(X_{t+h}^1 \overline{X_t^2}').$$

Definitions are prompted by and based on [8, p.657f].

An important class of transformations for time series form the so-called linear filters. They can be seen as some sort of extension of a linear map, which takes into consideration the dynamics of the time series.

Definition 2.7. Let sequence $(\psi_j)_{j\in\mathbb{Z}}$ consisting of $(k \times n)$-matrices be such that it transforms an n-dimensional input series $\{X_t\}$ into a k-dimensional output series $\{Y_t\}$ by

$$Y_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} =: \psi(L)X_t.$$

Then $\{Y_t\}$ is said to be obtained by filtering $\{X_t\}$ with $(\psi_j)_{j\in\mathbb{Z}}$. See [6, p.152ff].

Remark. Under the assumption that $(\psi_j)_{j\in\mathbb{Z}}$ is absolutely summable and $\mathrm{E}(x_{it}^2) < \infty$ for any $i \in \{1,\dots,n\}, t \in \mathbb{Z}$, Y_t is the mean square limit of $\sum_{j=-n}^{n} \psi_j X_{t-j}$ as n goes to infinity. However, if $\{X_t\}$ is a white noise vector process, the assumption of the absolute summability of $(\psi_j)_{j\in\mathbb{Z}}$ can be replaced by the slightly weaker assumption of square summability of the matrices $(\psi_j)_{j\in\mathbb{Z}}$. Information on the univariate case which can be easily extended to the multivariate case can be found in [6, p.83f]. The therein discussed arguments can be modified for the square summability and the white noise case. An extremely important property concerning filters fulfilling the appropriate summability condition for convergence of the output series $\{Y_t\}$ is the fact that a stationary process is still stationary after filtering. Details can be found in [6, p.83f].

An important group of filters are the filters associated with a function.

Definition 2.8. Let $f(\theta) = (f_1(\theta),\dots,f_n(\theta)) \in L_2^n([-\pi,\pi],\mathbb{C})$ be a row vector and consider its Fourier expansion

$$f(\theta) = \sum_{k=-\infty}^{\infty} f_k e^{-ik\theta},$$

with row vector $f_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)e^{ik\theta}d\theta \in \mathbb{C}^n$. The square summable $(1\times n)$-dimensional filter

$$\underline{f}(L) := \sum_{k=-\infty}^{\infty} f_k L^k$$

is the filter associated with function $f(\theta)$. Furthermore, we have for the complex conjugate $\overline{f(\theta)}'$ the decompositon

$$\overline{f(\theta)}' = \sum_{k=-\infty}^{\infty} \overline{f_k}' e^{ik\theta}.$$

The corresponding filter will be defined as

$$\sum_{k=-\infty}^{\infty} \overline{f_k}' L^{-k} =: \overline{\underline{f}(L)}'.$$

See [4, p.5f].

In the definition above we have derived a filter from a function. Conversely, we can construct a function from a filter.

Definition 2.9. Let $a(L) = \sum_{k=-\infty}^{\infty} a_k L^k$ be a square summable $(1 \times n)$-dimensional filter, we define

$$a^{\circ}(\theta) = a(e^{-i\theta}) = \sum_{k=-\infty}^{\infty} a_k e^{-ik\theta}$$

in $L_2^n([-\pi,\pi],\mathbb{C})$ to be the function associated with filter $a(L)$, see [4, p.5f].

These definitions can be easily carried over to matrix functions and filters consisting of matrices.

Before continuing we introduce the following notation for an n-dimensional vector $f(\theta)$,

$$|f(\theta)|^2 = \sum_{i=1}^{n} |f_i(\theta)|^2.$$

The following lemma will show that the square summability of an n-dimesional filter is in fact equivalent to the associated function being in $L_2^n([-\pi,\pi],\mathbb{C})$ as seen in [4, p.4f].

Lemma 2.1. With norm and inner product on $L_2^n([-\pi,\pi],\mathbb{C})$ given by

$$\|f\| = \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{n}|f_i(\theta)|^2 d\theta}, \qquad \langle f,g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{n} f_i(\theta)\overline{g_i(\theta)}d\theta,$$

it holds that

$$\sum_{k=-\infty}^{\infty}|f_k|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{i=1}^{n}|f_i(\theta)|^2 d\theta = \|f\|^2.$$

Proof. W.l.o.g assume $n = 1$, then

$$|f(\theta)|^2 = \sum_{k=-\infty}^{\infty} f_k e^{-ik\theta} \sum_{l=-\infty}^{\infty} \overline{f_l} e^{il\theta}$$

and hence

$$\|f\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(\theta)|^2 d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} f_k e^{-ik\theta} \sum_{l=-\infty}^{\infty} \overline{f}_l e^{il\theta} d\theta = \sum_{k=-\infty}^{\infty} |f_k|^2.$$

□

Next, we want to investigate the relationship between the spectral density of a process and the spectral density of a filtered version. As we have remarked above, depending on the process $\{X_t\}$ there can be a weaker summability condition applied in order for a stationary mean square limit to exist. This will also be taken in consideration in the next lemma, which is based on [1, p.5f].

Lemma 2.2. Let $\{X_t\}$ be an n-dimensional vector process with covariance function $\Gamma^X(h)$ and spectral density $\Sigma^X(\theta)$. Let $a(L)$ be a square summable n-dimensional filter with associated function $a^\circ(\theta) = a(e^{-i\theta})$ in case $\{X_t\}$ is white noise and let $a(L)$ be absolutely summable otherwise. Define a new process $\{Z_t\}$ by $Z_t = a(L)X_t$. Then its spectral density matrix $\Sigma^Z(\theta)$ is given by

$$\Sigma^Z(\theta) = a(e^{-i\theta})\Sigma^X(\theta)a(e^{i\theta})' = a^\circ(\theta)\Sigma^X(\theta)\overline{a^\circ(\theta)}'.$$

Proof. Let $(a_k)_{k\in\mathbb{Z}}$ be the coefficients of filter $a(L)$. Then for the covariance function $\Gamma^Z(h)$ of the process $\{Z_t\}$ it holds that

$$\begin{aligned}\Gamma^Z(h) &= \mathrm{Cov}(a(L)X_{t+h}, a(L)X_t) = \mathrm{Cov}\big(\sum_{k=-\infty}^{\infty} a_k X_{t+h-k}, \sum_{k=-\infty}^{\infty} a_k X_{t-k}\big)\\ &= \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_k \mathrm{Cov}(X_{t+h-k}, X_{t-l})\overline{a_l}' = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_k \Gamma^X(h-k+l)\overline{a_l}'.\end{aligned}$$

By definition, it holds that

$$\Sigma^Z(\theta) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_k \Gamma^X(h-k+l)\overline{a_l}' e^{-ih\theta},$$

which is by an index shift the same as

$$\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_k \Gamma^X(h)\overline{a_l}' e^{-i(h+k-l)\theta}.$$

Finally we have that

$$\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} a_k e^{-ik\theta}\Gamma^X(h)\overline{a_l}' e^{il\theta} e^{-ih\theta} = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} a^\circ(\theta)\Gamma^X(h)\overline{a^\circ(\theta)}' e^{-ih\theta},$$

which is again by definition of the spectral density

$$a^\circ(\theta)\Sigma^X(\theta)\overline{a^\circ(\theta)}'.$$

□

Remark. Analogously it can be proven that for filters $a(L)$ and $b(L)$ the cross-spectrum between $a(L)X_t$ and $b(L)X_t$ is given by $a^\circ(\theta)\Sigma^X(\theta)\overline{b^\circ(\theta)}'$.

2.2 Principal components

Linear filters are commonly used to transform a stationary time series into another stationary time series. In the following section we will investigate a specific filter which is linked with approximating a stationary time series by a filtered version of itself, namely the filter associated with its *principal components*.

Assume $\{X_t\}$ is a stationary n-dimensional vector process with $\mathrm{E}(X_t) = 0$ and covariance function $\Gamma^X(h)$ having absolutely summable entries in order for the spectral density $\Sigma^X(\theta)$ to exist.
The objective here is to find a stationary q-dimensional time series $\{\xi_t\}$ with $q < n$ such that this time series contains "as much information as possible"

about the original $\{X_t\}$, i.e. find sequences $(B_j)_{j\in\mathbb{Z}}$ and $(C_j)_{j\in\mathbb{Z}}$ of $(q\times n)$ resp. $(n\times q)$-matrices, generating the following series

$$\xi_t = \sum_{j=-\infty}^{\infty} B_j X_{t-j} = B(L)X_t$$

and

$$X_t^* = \sum_{j=-\infty}^{\infty} C_j \xi_{t-j} = C(L)\xi_t = C(L)B(L)X_t$$

such that

$$\mathrm{E}|X_t - X_t^*|^2$$

is minimized. Since the rank of the covariance matrix corresponding to vector process $\{X_t^*\}$ cannot exceed q, this filtering is also referred to as rank reduction of $\{X_t\}$. See also [7, p.337ff]. Before determining filters $B(L)$ and $C(L)$, we take a look at the static case, i.e. we consider a single n-dimensional random vector X. The following result is well known from multivariate statistics and is taken from [7, p.339f].

Theorem 2.1. Let X be an n-dimensional random vector with mean 0 and covariance matrix Γ. Then

$$\mathrm{E}|X - CBX|^2$$

for C resp. B an $(n\times q)$ resp. $(q\times n)$-matrix is minimized by

$$B = (p_1', \dots, p_q')', \qquad\qquad C = \overline{B}',$$

where $p_1, \dots, p_q$ are the eigenvectors seen as row vectors to the corresponding eigenvalues $\lambda_1, \dots, \lambda_q$ of covariance matrix Γ in descending order. The vector p_jX is called the j-th principal component of the random vector X.

An analogous result can be deduced for a time series, now taking into account the time dynamics. First, we introduce some essential definitions.

Definition 2.10. Let $\{X_t\}$ be an n-dimensional zero-mean stationary process with covariance function $\Gamma^X(h)$ and spectral density $\Sigma^X(\theta)$. For $i \in \{1, ..., n\}$ define the function

$$\lambda_i^X : [-\pi, \pi] \longrightarrow \mathbb{R},$$

where $\lambda_i^X(\theta)$ is the i-th eigenvalue of $\Sigma^X(\theta)$, $\theta \in [-\pi, \pi]$, in descending order. The functions $\lambda_i^X, i \in \{1, ..., n\}$ are called the *dynamic eigenvalues* of Σ^X. The n functions $p_i : [-\pi, \pi] \to \mathbb{C}^n$ as row vectors for $i \in \{1, \ldots, n\}$ fullfilling

- $|p_i(\theta)| = 1, \qquad \forall\theta \in [-\pi, \pi],$
- $p_i(\theta)\overline{p_j(\theta)}' = 0, \qquad i \neq j, \;\; \forall\theta[-\pi, \pi],$
- $p_i(\theta)\Sigma^X(\theta) = \lambda_i(\theta)p_i(\theta), \qquad \forall\theta \in [-\pi, \pi].$

are called the *dynamic eigenvectors* of Σ^X. See [5, p.542].

Remark. Note that the eigenvector $p_j(\theta)$ can and will be chosen such that $p_j(\theta)$ is continuous for $j \in \{1, \ldots, n\}$.

Theorem 2.2. Let $\{X_t\}$ be an n-dimensional zero mean stationary vector process having an absolutely summable covariance function $\Gamma^X(h)$ and spectral density $\Sigma^X(\theta)$. Then $(B_j)_{j\in\mathbb{Z}}$ and $(C_j)_{j\in\mathbb{Z}}$ which minimize

$$\mathrm{E}|X_t - C(L)B(L)X_t|^2$$

are given by

$$B_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} B^\circ(\theta)e^{ik\theta}d\theta$$
$$C_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} C^\circ(\theta)e^{ik\theta}d\theta$$

where $B^\circ(\theta) = (p_1(\theta)', \ldots, p_q(\theta)')'$ and $C^\circ(\theta) = \overline{B^\circ(\theta)}'$ with $p_i(\theta)$ being the eigenvector corresponding to the i-th largest eigenvalue of $\Sigma^X(\theta)$. The theorem can be found in [7, p.344].

Proof. We have

$$\begin{aligned}
&\mathrm{E}|X_t - C(L)B(L)X_t|^2 = \\
&\mathrm{E}((\overline{X_t - C(L)B(L)X_t})'(X_t - C(L)B(L)X_t)) = \\
&\mathrm{E}(\mathrm{tr}((\overline{I_q - C(L)B(L)})'\overline{X_t}'X_t(I_q - C(L)B(L)))) = \\
&\mathrm{tr}(\mathrm{E}((I_q - C(L)B(L))X_t\overline{X_t}'(\overline{I_q - C(L)B(L)})')) = \\
&\mathrm{tr}(\int_{-\pi}^{\pi}(I_q - C^\circ(\theta)B^\circ(\theta))\Sigma^X(\theta)(\overline{I_q - C^\circ(\theta)B^\circ(\theta)})'d\theta) = \\
&\int_{-\pi}^{\pi}\mathrm{tr}((I_q - C^\circ(\theta)B^\circ(\theta))\Sigma^X(\theta)(\overline{I_q - C^\circ(\theta)B^\circ(\theta)})')d\theta,
\end{aligned}$$

which follows from basic properties of the trace and the fact that by the relationship between the spectral density and the covariance function for some process $A(L)X_t = Z_t$ we have

$$\mathrm{E}(A(L)X_t\overline{X_t}'\overline{A(L)}') = \Gamma^Z(0) = \int_{-\pi}^{\pi}\Sigma^Z(\theta)d\theta = \int_{-\pi}^{\pi}A^\circ(\theta)\Sigma^X(\theta)\overline{A^\circ(\theta)}',$$

with application of Lemma 2.2. Hence by the above observations it suffices to minimize $\mathrm{tr}((I_q - A^\circ(\theta))\Sigma^X(\theta)(\overline{I_q - A^\circ(\theta)})')$ for fixed θ, where $A^\circ(\theta) = C^\circ(\theta)B^\circ(\theta)$ is a matrix of rank q. The integrand can be rewritten with convention $(\Sigma^X(\theta))^{1/2} = \Sigma^{1/2}(\theta)$ as

$$\mathrm{tr}\big(\big(\Sigma^{1/2}(\theta) - A^\circ(\theta)\Sigma^{1/2}(\theta)\big)\big(\overline{\Sigma^{1/2}(\theta) - A^\circ(\theta)\Sigma^{1/2}(\theta)}\big)'\big).$$

Define for fixed θ, $M := \Sigma^{1/2}(\theta) - A^\circ(\theta)\Sigma^{1/2}(\theta)$. Notice that the trace here is the sum of the eigenvalues of $M\overline{M}'$. Thus minimizing the trace of $M\overline{M}'$ involves minimizing its eigenvalues. Set $W = A^\circ(\theta)\Sigma^{1/2}(\theta)$ and denote by $\tilde{\lambda}_i(\theta)$ the i-th largest eigenvalue of $M\overline{M}'$ and by $\lambda_i(\theta)$ the i-th largest eigenvalue of $\Sigma^X(\theta)$. Then by Courant-Fisher, Lemma 7.1 from the Appendix, it holds that

$$\begin{aligned}
\tilde{\lambda}_k &= \inf_{\substack{D\ (k-1)\times n\\ matrix}} \sup_{\substack{Dy=0\\ y\in\mathbb{C}^n}} \frac{\bar{y}'M\overline{M}'y}{\bar{y}'y} \geq \inf_{\substack{D\ (k-1)\times n\\ matrix}} \sup_{\substack{Dy=0\\ \overline{W}'y=0\\ y\in\mathbb{C}^n}} \frac{\bar{y}'M\overline{M}'y}{\bar{y}'y} \\
&\geq \inf_{\substack{D\ (k-1)\times n\\ matrix}} \sup_{\substack{Dy=0\\ \overline{W}'y=0\\ y\in\mathbb{C}^n}} \frac{\bar{y}'\Sigma^{1/2}(\theta)\overline{\Sigma^{1/2}(\theta)}'y}{\bar{y}'y} \geq \inf_{\substack{D\ (k+q-1)\times n\\ matrix}} \sup_{\substack{Dy=0\\ y\in\mathbb{C}^n}} \frac{\bar{y}'\Sigma^X(\theta)y}{\bar{y}'y} = \lambda_{k+q}.
\end{aligned}$$

If one can show that $A^{\circ}(\theta)$ can be chosen such that $\tilde{\lambda}_k$ equals λ_{k+q} for $k \in \{1,\dots,n-q\}$, the trace is minimized. We propose

$$A^{\circ}(\theta) = \sum_{j=1}^{q} \overline{p_j(\theta)}' p_j(\theta)$$

and since $\Sigma^{1/2}(\theta)$ can be written as

$$\Sigma^{1/2}(\theta) = \sum_{j=1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta),$$

it follows that $(\Sigma^{1/2}(\theta) - A^{\circ}(\theta)\Sigma^{1/2}(\theta))$ equals

$$\sum_{j=1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta) - \sum_{j=1}^{q} \overline{p_j(\theta)}' p_j(\theta) \sum_{j=1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta) =$$
$$\sum_{j=1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta) - \sum_{j=1}^{q} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta) = \sum_{j=q+1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta).$$

And finally $\mathrm{tr}(M\overline{M}')$ equals

$$\overline{\Big(\sum_{j=q+1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta)\Big)}'\Big(\sum_{j=q+1}^{n} \lambda_j^{1/2}(\theta)\overline{p_j(\theta)}' p_j(\theta)\Big) = \sum_{j=q+1}^{n} \lambda_j(\theta),$$

which proves that the eigenvalues $\tilde{\lambda}_k$ have been minimized by function $A^{\circ}(\theta) = \sum_{j=1}^{q} \overline{p_j(\theta)}' p_j(\theta)$, where $B^{\circ}(\theta) = (p_1(\theta)', \dots, p_q(\theta)')'$ and $C^{\circ}(\theta) = \overline{B^{\circ}(\theta)}'$. The proof is based on [7, p.399] and [7, p.455]. □

Definition 2.11. Let $\{X_t\}$ be an n-dimensional vector process having spectral density $\Sigma^X(\theta)$. Then $\{p_j(\theta) : j \in \{1,\dots,n\}\}$ forms the set of dynamic eigenvectors. And $\{\underline{p}_j(L)X_t : j \in \{1,\dots,n\}\}$, where $\underline{p}_j(L)$ is the filter associated with the dynamic eigenvector $p_j(\theta)$, is called the set of *dynamic principal components* of $\{X_t\}$, according to [4, p.6].

Remark. In general the filters $\underline{p}_j(L)$ are two-sided, i.e. they depend on values that occur later in time, which poses a serious problem in practice. Further notice that the spectral density of the j-th principal component $W_t = \underline{p}_j(L)X_t$ is by Lemma 2.2,

$$\Sigma^W(\theta) = p_j(\theta)\Sigma^X(\theta)\overline{p_j(\theta)}' = \lambda_j^X(\theta).$$

Furthermore notice that for the vector process $\{Y_t\}$ consisting of all n principal components, i.e.

$$Y_t = \begin{pmatrix} \underline{p}_1(L)X_t \\ \vdots \\ \underline{p}_n(L)X_t \end{pmatrix} = \begin{pmatrix} \underline{p}_1(L) \\ \vdots \\ \underline{p}_n(L) \end{pmatrix} X_t,$$

the spectral density is given by

$$\Sigma^Y(\theta) = \begin{pmatrix} p_1(\theta) \\ \vdots \\ p_n(\theta) \end{pmatrix} \Sigma^X(\theta)(\overline{p_1(\theta)}', \dots, (\overline{p_n(\theta)}') = \operatorname{diag}(\lambda_1^X(\theta), \dots, \lambda_n^X(\theta)).$$

Since the spectral density is a diagonal matrix, so is the covariance matrix function for any lag h. This means that the principal components are pairwise uncorrelated.

2.3 General definitions and theoretical results

Before delving into the generalized dynamic factor model, we recall important general definitions and some basic theoretical results that are fundamental for defining and proving facts about the generalized dynamic factor model.

Definition 2.12. Denote by $\mathcal{L}$ the Lebesgue measure on $\mathbb{R}$. We say, a function $f : \mathbb{R} \to \mathbb{R}$ is essentially bounded if there exist $M \in \mathbb{R}$ and subset $N \subset \mathbb{R}$ such that $|f(\theta)| \leq M$ for all $\theta \in \mathbb{R} \setminus N$ and $\mathcal{L}(N) = 0$. See [4, p.8].

An important tool for the upcoming proofs is the dominated convergence theorem.

Theorem 2.3. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions defined on $[-\pi,\pi]$ such that

- $\lim\limits_{n\to\infty} f_n(\theta) = f(\theta)$ for almost all $\theta \in [-\pi,\pi]$,
- $|f_n(\theta)| \leq g(\theta)$ for almost all $\theta \in [-\pi,\pi]$ and any $n \in \mathbb{N}$ for an integrable and non-negative function g.

Then f is integrable and

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f_n(\theta) d\theta = \int_{-\pi}^{\pi} f(\theta) d\theta.$$

The proof can be found in [9, p.270ff].

Theorem 2.4. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions defined on $[-\pi,\pi]$. Let f be an integrable function defined on $[-\pi,\pi]$. Suppose that either

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f_n(\theta) - f(\theta)|^2 d\theta = 0$$

or

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |f_n(\theta) - f(\theta)| d\theta = 0$$

hold. Then there exists an increasing sequence $(s_n)_{n\in\mathbb{N}}$ such that for almost all θ in $[-\pi,\pi]$

$$\lim_{n\to\infty} f_{s_n}(\theta) = f(\theta).$$

The theorem is taken from [4, p.8].

3 Introduction of the Generalized Dynamic Factor Model

In the following chapter we will define the generalized dynamic factor model for infinite dimensional vector process $\{X_t\}$ in an appropriately chosen setting. This model splits the observable $\{X_t\}$ into two non-observable infinite dimensional stationary vector processes $\{\chi_t\}$ and $\{\xi_t\}$, called the common and idiosyncratic component. Furthermore we will propose a sequence of estimators that converge in mean square to the common component χ_{it} for any $t \in \mathbb{Z}, i \in \mathbb{N}$. The common component χ_{it} is driven by an underlying q-dimensional orthogonal vector process $\{u_t\}$ which is "common" to all sections. So to say, the common component follows a fixed structure, whereas the idiosyncratic component can be seen as some error added to the common component. Hence, estimating the common component gives us a better insight into how the vector process $\{X_t\}$ behaves.

3.1 Setting and preparation

Now we extend the notion of n-dimensional stationary vector processes to infinite dimensional stationary time series $\{X_t\}$. Let vector process $\{X_t\}$ be given by $\{(x_{1t}, x_{2t} \dots)' : t \in \mathbb{Z}\}$, where $x_{it} \in L_2(\mathcal{P}, \mathbb{C})$ for all $i \in \mathbb{N}, t \in \mathbb{Z}$. Additionally, assume that the n-dimensional time series $\{X_t^n : t \in \mathbb{Z}\} = \{(x_{1t}, \dots, x_{nt})' : t \in \mathbb{Z}\}$ is stationary for all $n \in \mathbb{N}$. In particular, we are interested in investigating the n-dimensional time series $\{X_t^n\}$ for growing n. To this end we assume that the spectral density matrix of $\{X_t^n\}$, denoted by $\Sigma_n^X(\theta)$, exists for all $n \in \mathbb{N}$. Furthermore denote by $p_{ni}(\theta)$ the dynamic eigenvector for the corresponding i-th largest dynamic eigenvalue $\lambda_{ni}^X(\theta)$ of $\Sigma_n^X(\theta)$. The filter associated with eigenvector $p_{ni}(\theta)$ is denoted by $\underline{p}_{ni}(L)$. Consistently the covariance function of $\{X_t^n\}$ is denoted by $\Gamma_n^X(h)$. Moreover, define $X = \{x_{it} : i \in \mathbb{N}, t \in \mathbb{Z}\}$ and $\overline{X} = \overline{\text{span}}(\{x_{it} : i \in \mathbb{N}, t \in \mathbb{Z}\})$. Compare with [4].

Remark. The assumption on the existence of $\Sigma_n^X(\theta)$ implies together with the fact that components x_{it} are twice integrable for $i \in \{1, \ldots, n\}$, $t \in \mathbb{Z}$ that $\Sigma_n^X(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\theta h}\Gamma_n^X(h)$ is again Lebesgue measurable. The function that maps to the eigenvalues is measurable and the $\lambda_{ni}^X(\theta)$'s are continuous functions of the entries of $\Sigma_n^X(\theta)$. Hence the eigenvalues $\lambda_{ni}^X(\theta)$ for all $n \in \mathbb{N}$ and $i \leq n$ are Lebesgue measurable in $[-\pi, \pi]$. Furthermore, the n dynamic eigenvectors $p_{ni}(\theta)$ corresponding to $\lambda_{ni}(\theta)$ for vector process $\{X_t^n\}$ according to Definition 2.10 exist and are measurable. The proof can be found in the appendix of [4, p.24].

Lemma 3.1. Let $\{X_t^n\}$ be a sequence of vector processes resulting from an infinite dimensional stationary vector process $\{X_t\}$. For fixed i with $i \leq n$ the dynamic eigenvalues $\lambda_{ni}^X(\theta)$ are non-decreasing functions in n for any $\theta \in [-\pi, \pi]$. See [4, p.4].

Proof. As above described, let $\lambda_{ni}^X(\theta)$ be the i-th largest eigenvalue of $\Sigma_n^X(\theta)$ and let $\lambda_{n+1,i}^X(\theta)$ be the i-th largest eigenvalue of $\Sigma_{n+1}^X(\theta)$. Notice that the $(n \times n)$-submatrix consisting of the first n rows and the first n columns of $\Sigma_{n+1}^X(\theta)$ coincides with $\Sigma_n^X(\theta)$. We apply the characterization of eigenvalues according to Courant-Fisher (Lemma 7.1). Let $i = 1$.

$$\lambda_{n+1,1}(\theta) = \max_{y \in \mathbb{C}^{n+1}} \frac{\bar{y}'\Sigma_{n+1}^X(\theta)y}{\bar{y}'y} \geq \max_{\substack{y_{n+1}=0 \\ y \in \mathbb{C}^{n+1}}} \frac{\bar{y}'\Sigma_{n+1}^X(\theta)y}{\bar{y}'y} = \max_{y \in \mathbb{C}^n} \frac{\bar{y}'\Sigma_n^X(\theta)y}{\bar{y}'y} = \lambda_{n1}(\theta).$$

Now let $i > 1$.

$$\begin{aligned}\lambda_{n+1,i}(\theta) &= \min_{\substack{D\ (i-1)\times(n+1)\\ \text{matrix}}} \max_{\substack{Dy=0\\ y\in\mathbb{C}^{n+1}}} \frac{\bar{y}'\Sigma^X_{n+1}(\theta)y}{\bar{y}'y} \\ &\geq \min_{\substack{D\ (i-1)\times(n+1)\\ \text{matrix}}} \max_{\substack{y_{n+1}=0\\ Dy=0\\ y\in\mathbb{C}^{n+1}}} \frac{\bar{y}'\Sigma^X_{n+1}(\theta)y}{\bar{y}'y} \\ &= \min_{\substack{D\ (i-1)\times n\\ \text{matrix}}} \max_{\substack{Dy=0\\ y\in\mathbb{C}^{n+1}}} \frac{\bar{y}'\Sigma^X_{n}(\theta)y}{\bar{y}'y} = \lambda_{ni}(\theta).\end{aligned}$$

□

This implies that $\lim_{n\to\infty} \lambda^X_{ni}(\theta)$ is well-defined for all $i \leq n$ and θ in $[-\pi,\pi]$. This allows for the following definition

$$\lambda^X_i(\theta) := \lim_{n\to\infty} \lambda^X_{ni}(\theta) = \sup_{n\in\mathbb{N}} \lambda^X_{ni}(\theta),$$

according to [4, p.4].
Before we can finally introduce the generalized dynamic factor model we need further essential definitions related to filters and infinite dimensional vector processes.

Definition 3.1. Let $(s_n)_{n\in\mathbb{N}}$ be a sequence of positive integers and let $(a_n(L))_{n\in\mathbb{N}}$ be a sequence of $(1 \times s_n)$-dimensional filters with associated function $a^\circ_n(\theta) = (a^\circ_{n1}(\theta), \ldots, a^\circ_{ns_n}(\theta)) \in L^{s_n}_2([-\pi,\pi],\mathbb{C})$. The sequence $(a_n(L))_{n\in\mathbb{N}}$ is called *dynamic averaging sequence*, short DAS, if

$$\lim_{n\to\infty} \|a^\circ_n\|^2 = \lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |a^\circ_n(\theta)|^2 d\theta = \lim_{n\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} \sum_{i=1}^{s_n} |a^\circ_{ni}(\theta)|^2 d\theta = 0.$$

See [4, p.6].

Definition 3.2. Let $\{X_t\}$ be an infinite dimensional stationary time series. Then $\{X_t\}$ is called *idiosyncratic* if $\lim_{n\to\infty} a_n(L)X^{s_n}_t = 0$ in mean square, i.e. $\lim_{n\to\infty} \|a_n(L)X^{s_n}_t\|^2 = 0$, for all dynamic averaging sequences $(a_n(L))_{n\in\mathbb{N}}$ with $(1 \times s_n)$-dimensional filters. See [4, p.7].

Remark. Applying the filters to $X_t^{s_n}$ can be seen as "averaging" X_t simultaneously over time and dimension via a function $a_n^\circ(\theta)$. Also, it can be easily seen that functions $a_n^\circ(\theta)$ can be constructed such that the application of $a_n(L)$ corresponds to averaging in the traditional sense.

Example 3.1. To illustrate the application of dynamic averaging sequences, we consider the infinite dimensional vector process $\{X_t\}$ with $x_{it} \perp x_{j,t-k}$ for any $i, j \in \mathbb{N}$, $t, k \in \mathbb{Z}$ with $i \neq j$, $x_{it} \perp x_{i,t-k}$ for any $i \in \mathbb{N}$, $t, k \in \mathbb{Z}$ with $k \neq 0$ and $\|x_{it}\| = 1$ for any $t \in \mathbb{Z}$, $i \in \mathbb{N}$. Futhermore consider functions $(a_n^\circ(\theta))_{n\in\mathbb{N}}$ given by

$$a_n^\circ(\theta) = \frac{1}{2\pi n}(1, \ldots, 1).$$

The coefficients of the corresponding filter $a_n(L)$ are then given by

$$a_{n,k} = \int_{-\pi}^{\pi} \frac{1}{2\pi n}(1 \ldots 1)e^{ik\theta} d\theta = \begin{cases} \frac{1}{n}(1, \ldots, 1) & k = 0 \\ 0 & k \neq 0. \end{cases}$$

Then by

$$\lim_{n\to\infty} \|a_n^\circ\|^2 = \lim_{n\to\infty} \int_{-\pi}^{\pi} |a_n^\circ(\theta)|^2 d\theta = \lim_{n\to\infty} \int_{-\pi}^{\pi} \frac{1}{(2\pi)^2 n} d\theta = \lim_{n\to\infty} \frac{1}{2\pi n} = 0,$$

$a_n(L)$ is a dynamic averaging sequence with $s_n = n$. Furthermore

$$\lim_{n\to\infty} \|a_n(L)X_t^n\|^2 = \lim_{n\to\infty} \|\frac{1}{n}\sum_{i=1}^{n} x_{it}\|^2 = \lim_{n\to\infty} \frac{1}{n^2}\sum_{i=1}^{n}\|x_{it}\|^2 = \lim_{n\to\infty} \frac{1}{n} = 0.$$

However, this does not prove the idiosyncrasy of $\{X_t\}$ since this property has to be proven for any dynamic averaging sequence.

The following characterisation of idiosyncrasy can be helpful when verifying that an infinite dimensional stationary vector process is in fact idiosyncratic, since the focus shifts from the dynamic averaging sequences to the first eigenvalue $\lambda_1^X(\theta)$ of the spectral density matrix corresponding to vector process $\{X_t\}$.

Theorem 3.1. Let $\{X_t\}$ be an infinite dimensional stationary vector process with spectral densities $\Sigma_n^X(\theta)$ for corresponding vector process $\{X_t^n\}$. Then $\lambda_1^X(\theta)$ is essentially bounded if and only if $\{X_t\}$ is idiosyncratic. See [4, p.8f].

Proof. Firstly, assume $\lambda_1^X(\theta)$ is essentially bounded. Let the sequence $(a_n(L))_{n\in\mathbb{N}}$ of $(1\times s_n)$-dimensional filters be an arbitrary DAS. Consider the new univariate process denoted by $Z_t^n = a_n(L)X_t^{s_n}$. Then by Lemma 2.2, it holds that

$$\begin{aligned}\Gamma^{Z^n}(0) &= \int_{-\pi}^{\pi} \Sigma_n^Z(\theta)d\theta = \int_{-\pi}^{\pi} a_n^\circ(\theta)\Sigma_{s_n}^X(\theta)\overline{a_n^\circ(\theta)}'d\theta \\ &\leq \int_{-\pi}^{\pi} \lambda_{s_n1}^X(\theta)|a_n^\circ(\theta)|^2 d\theta \leq \int_{-\pi}^{\pi} \lambda_1^X(\theta)|a_n^\circ(\theta)|^2 d\theta.\end{aligned}$$

Since $\lambda_1^X(\theta)$ is essentially bounded and $\lim\limits_{n\to\infty}\|a_n^\circ\|^2 = 0$, we have

$$\int_{-\pi}^{\pi} \lambda_1^X(\theta)|a_n^\circ(\theta)|^2 d\theta \xrightarrow[n\to\infty]{} 0.$$

Since the variance of Z_t^n converges to zero, Z_t^n converges in mean square to zero, fulfilling the definition of idiosyncrasy.
Conversely, assume that $\lambda_1^X(\theta)$ is not essentially bounded. Then for any $n \in \mathbb{N}$ there exists $s_n \in \mathbb{N}$ such that for $M_n = \{\theta : \lambda_{s_n1}^X(\theta) \geq n\}$,

$$\nu_n = \mathcal{L}(M_n) > 0.$$

Define function $b_n : [-\pi, \pi] \to \mathbb{R}$ as

$$b_n(\theta) = \begin{cases} \frac{1}{\sqrt{\nu_n}} & \lambda_{s_n1}^X(\theta) \geq n \\ 0 & \text{otherwise} \end{cases}$$

Then it follows that

$$\begin{aligned}B_n &:= \int_{-\pi}^{\pi} \lambda_{s_n1}^X(\theta)b_n(\theta)^2 d\theta = \int_{M_n} \lambda_{s_n1}^X(\theta)b_n(\theta)^2 d\theta \\ &= \int_{M_n} \lambda_{s_n1}^X(\theta)\tfrac{1}{\nu_n}d\theta \geq \frac{n}{\nu_n}\int_{M_n} d\theta = n.\end{aligned}$$

Then the filter sequence

$$a_n(L) := \frac{b_n(L)}{\sqrt{B_n}} \underline{p}_{s_n1}(L),$$

where $p_{s_n1}(\theta)$ is the eigenvector corresponding to $\lambda^X_{s_n1}(\theta)$ is a DAS, because

$$\|a_n^\circ\|^2 = \left\| \frac{b_n}{\sqrt{B_n}} p_{s_n1} \right\|^2 = \left\| \frac{b_n}{\sqrt{B_n}} \right\|^2 = \frac{1}{2\pi B_n} \int_{-\pi}^{\pi} b_n(\theta)^2 d\theta = \frac{1}{2\pi B_n} \leq \frac{1}{2\pi n} \longrightarrow 0$$

as n goes to infinity. But on the other hand we have that

$$\|a_n(L)X_t^{s_n}\|^2 = \int_{-\pi}^{\pi} \frac{b_n(\theta)^2}{B_n} p_{s_n1}(\theta)\Sigma^X_{s_n}(\theta)\overline{p_{s_n1}(\theta)}' d\theta = \int_{-\pi}^{\pi} \frac{b_n(\theta)^2}{B_n} \lambda^X_{s_n1}(\theta) d\theta = 1.$$

Since $a_n(L)X_t^{s_n}$ does not converge in mean square to 0, $\{X_t\}$ is not idiosyncratic which concludes the proof by contraposition. The proof is based on [4, p.8f]. □

Example 3.2. Continuing our example above with infinite dimensional stationary vector process $\{X_t\}$ where $x_{it} \perp x_{j,t-k}$ for any $i, j \in \mathbb{N}$, $t, k \in \mathbb{Z}$ with $i \neq j$, $x_{it} \perp x_{i,t-k}$ for any $i \in \mathbb{N}$, $t, k \in \mathbb{Z}$ with $k \neq 0$ and $\|x_{it}\| = 1$ for any $t \in \mathbb{Z}, i \in \mathbb{N}$, it is easily seen that the covariance function fulfills $\Gamma^X_n(0) = I_n$ and $\Gamma^X_n(h) = 0$ for $h \neq 0$. This further implies that $\Sigma^X_n(\theta) = \frac{1}{2\pi} I_n$ whose biggest eigenvalue is $\frac{1}{2\pi}$ for all $n \in \mathbb{N}$. Thus $\lambda^X_1(\theta)$ is essentially bounded and $\{X_t\}$ is idiosyncratic.

We are now able to introduce the *generalized dynamic factor model.*

Definition 3.3. Let $q > 1$ be a natural number. We say the infinite dimensional stationary time series $\{X_t\}$ follows a *generalized dynamic factor model with q common factors* or alternatively $\{X_t\}$ is a *q-dynamic factor sequence*, if there exist

- a q-dimensional white-noise vector process

$$\{u_t\} = \{(u_{1t}, \ldots, u_{qt})' : t \in \mathbb{Z}\},$$

with the identity as spectral density, (which implies by the relationship between spectral density and covariance function that $\Gamma^u(h) = 2\pi I_q$ and $\Gamma^u(h) = 0$ for $h \neq 0$) and such that u_{it} is element of $L_2(\mathcal{P}, \mathbb{C})$ for any $i \in \mathbb{N}$ and any $t \in \mathbb{Z}$,

- an infinite dimensional stationary vector process

$$\{\xi_t\} = \{(\xi_{1t},\ \xi_{2t},\ \dots\ )' : t \in \mathbb{Z}\}$$

such that ξ_{it} and $u_{j,t-k}$ are orthogonal for any $i, j \in \mathbb{N}$ and any $t, k \in \mathbb{Z}$

fulfilling the following properties:

1) x_{it} can be written as

$$x_{it} = \chi_{it} + \xi_{it}$$

where

$$\chi_{it} = \sum_{j=1}^{q} b_{ij}(L) u_{jt} =: B_i(L) u_t$$

for any $i \in \mathbb{N}$ and some associated function $B_i^\circ(\theta) \in L_2^q([-\pi, \pi], \mathbb{C})$;

2) $\{\xi_t\}$ is idiosyncratic;

3) the infinite dimensional time series $\{\chi_t\} = \{(\chi_{1t},\ \chi_{2t},\ \dots\ )' : t \in \mathbb{Z}\}$ with spectral densities $\Sigma_n^\chi(\theta)$ satisfies $\lambda_q^\chi(\theta) = \infty$ for almost all θ in $[-\pi, \pi]$.

The components $\{\chi_t\}$ and $\{\xi_t\}$, which the process $\{X_t\}$ is split into, are called *common and idiosyncratic component.* The definition is introduced in [4, p.9f].

The underlying q-dimensional white noise process $\{u_t\}$ generates the components of the common component $\{\chi_t\}$ which determines $\{X_t\}$ up to some error which is represented by the idiosyncratic vector process $\{\xi_t\}$. Also note that we in general do not assume $B_i(L)$ to be one-sided. This only becomes important if one seeks to estimate the vector process $\{u_t\}$ and as well as to come up with an actual interpretation of this process. For the purpose of estimating the common component, we do not care about the filters being two-sided.

Remark. Property 2) of Definition 3.3 can be replaced by $\lambda_1^{\xi}(\theta)$ is essentially bounded by Theorem 3.1.

Given that $\{X_t\}$ follows a generalized dynamic factor model with q factors it is possible to deduce interesting properties concerning the eigenvalues of the spectral densities of the processes $\{X_t^n\}$ which enable the estimation of q.

Lemma 3.2. Let $\{X_t^n\}$ be resulting from an infinite dimensional stationary vector process $\{X_t\}$ with spectral densities $\Sigma_n^X(\theta)$ for any $n \in \mathbb{N}$. If $\{X_t\}$ follows a generalized dynamic factor model with q factors then the following holds:

(I) $\lambda_{q+1}^X(\theta)$ is essentially bounded,

(II) $\lambda_q^X(\theta) = \infty$ a.e. in $[-\pi, \pi]$.

See [4, p.10ff].

Proof. Since $\{X_t\}$ follows a generalized dynamic factor model with q factors, ξ_{it} and $u_{j,t-k}$ are orthogonal for any $i, j \in \mathbb{N}, t, k \in \mathbb{Z}$. Thus, the spectral density matrix of $\Sigma_n^X(\theta)$ can be split into the sum

$$\Sigma_n^X(\theta) = \Sigma_n^{\chi}(\theta) + \Sigma_n^{\xi}(\theta).$$

Applying Lemma 7.2 from the appendix to the above equation yields

$$\lambda_{nq}^X(\theta) \geq \lambda_{nq}^{\chi}(\theta),$$

implying (II). Again by Lemma 7.2,

$$\lambda_{n,q+1}^X(\theta) \leq \lambda_{n,q+1}^{\chi}(\theta) + \lambda_{n1}^{\xi}(\theta) = \lambda_{n1}^{\xi}(\theta)$$

holds, where the last equality comes from the fact that by definition $\Sigma_n^{\chi}(\theta)$ has rank q. Taking the limits, we have

$$\lambda_{q+1}^X(\theta) \leq \lambda_1^{\xi}(\theta).$$

Since $\lambda_1^{\xi}(\theta)$ is essentially bounded, so is $\lambda_{q+1}^X(\theta)$ which implies (I). The proof is based on [4, p.10ff]. □

For the sake of simplicity we will now introduce according to [4] some helpful notation. In the following we will use the convention for the matrix product AB, where A is a $d_1 \times d_2$ and B a $d_3 \times d_4$ matrix with $d_2 < d_3$, that A has been enlarged by $d_3 - d_2$ columns of zeros.
Let $n \geq q$. Denote by $P_n(\theta)$ the $(q \times n)$-matrix consisting of the first q dynamic eigenvectors of $\{X_t\}$

$$P_n(\theta) = \left(p_{n1}^X(\theta)', \ldots, p_{nq}^X(\theta)'\right)',$$

and by $Q_n(\theta)$ the $((n-q) \times n)$-matrix consisting of the last $n-q$ dynamic eigenvectors,

$$Q_n(\theta) = \left(p_{n,q+1}^X(\theta)', \ldots, p_{nn}^X(\theta)'\right)'.$$

Furthermore denote by $\Lambda_n(\theta)$ the $(q \times q)$ diagonal matrix whose diagonal entries are the first q dynamic eigenvalues of $\{X_t\}$ in descending order

$$\Lambda_n(\theta) = \begin{bmatrix} \lambda_{n1}^X(\theta) & 0 & \ldots & 0 \\ 0 & \lambda_{n2}^X(\theta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_{nq}^X(\theta) \end{bmatrix}.$$

Analogously denote by $\Phi_n(\theta)$ the $(n-q) \times (n-q)$ diagonal matrix having the last $n-q$ eigenvalues on the diagonal

$$\Phi_n(\theta) = \begin{bmatrix} \lambda_{n,q+1}^X(\theta) & 0 & \ldots & 0 \\ 0 & \lambda_{n,q+2}^X(\theta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_{nn}^X(\theta) \end{bmatrix}.$$

By the above notation $\Sigma_n^X(\theta)$ decomposes into

$$\Sigma_n^X(\theta) = \overline{P_n(\theta)}'\Lambda_n(\theta)P_n(\theta) + \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta) \tag{3.1}$$

and $I_n(\theta)$ can be written as

$$I_n(\theta) = \overline{P_n(\theta)}'P_n(\theta) + \overline{Q_n(\theta)}'Q_n(\theta). \tag{3.2}$$

Remark 3.1. For our further investigations it can be assumed without loss of generality that for any $n \in \mathbb{N}$ and any $j \leq n$ and $\theta \in [-\pi, \pi]$, $\lambda_{nj}^{X}(\theta) \geq 1$ holds.

The aim of the following two pages is to derive properties of the principal components of $\{X_t^n\}$ and to even relate it to the principal components of $\{X_t^m\}$. These observations are based on [4].

Due to the above assumption, $\Lambda_n^{-1}(\theta)$ is bounded. Hence, we may define a q-dimensional vector process $\{\psi_t^n\}$, for $n \in \mathbb{N}$, by

$$\psi_t^n = (\psi_{1t}^n, \ldots, \psi_{qt}^n)' = \underline{\Lambda}_n^{-1/2}(L)\underline{P}_n(L)X_t^n. \tag{3.3}$$

This vector process $\{\psi_t^n\}$ consists of the first q normalized principal components of $\{X_t^n\}$ and is orthogonal white noise with unit spectral density, since by Lemma 2.2 and equation (3.1) we have

$$\begin{aligned}\Sigma_n^{\psi}(\theta) &= \Lambda_n^{-1/2}(\theta)P_n(\theta)\Sigma_n^X(\theta)\overline{P_n(\theta)}'\Lambda_n^{-1/2}(\theta) \\ &= \Lambda_n^{-1/2}(\theta)P_n(\theta)\overline{P_n(\theta)}'\Lambda_n(\theta)P_n(\theta)\overline{P_n(\theta)}'\Lambda_n^{-1/2}(\theta) = I_q\end{aligned} \tag{3.4}$$

and furthermore

$$\Gamma_n^{\psi}(h) = \int_{-\pi}^{\pi} e^{ih\theta}\Sigma_n^{\psi}(\theta)d\theta = I_q\int_{-\pi}^{\pi} e^{-ih\theta}d\theta = \begin{cases} 2\pi I_q, & h = 0, \\ 0, & h \neq 0. \end{cases}$$

Notice that the components of the vector process $\{\psi_t^n\}$, i.e. the normalized principal components of $\{X_t^n\}$ are pairwise orthogonal at any lead and lag. They also have unit spectral density making them an eligible candidate to project other processes onto them, i.e. onto $\overline{\text{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$. We now want to take a closer look at the projection of $\{X_t^n\}$ onto the normalized principal components collected in the process $\{\psi_t^n\}$. With the use of equation (3.2) we get

$$X_t^n = \overline{\underline{P}_n(L)}'\underline{P}_n(L)X_t^n + \overline{\underline{Q}_n(L)}'\underline{Q}_n(L)X_t^n.$$

Recall that $\overline{\underline{P}_n(L)}'$ is the filter associated with function $\overline{P_n(\theta)}'$ and relates to filter $P_n(L) = \sum_{k=-\infty}^{\infty} P_{nk} L^k$ via $\overline{\underline{P}_n(L)}' = \sum_{k=-\infty}^{\infty} \overline{P_{nk}}' L^{-k}$. Furthermore plugging in the definition of ψ_t^n from equation (3.3) yields

$$X_t^n = \overline{\underline{P}_n(L)}' \underline{\Lambda}_n^{1/2}(L)\psi_t^n + \overline{\underline{Q}_n(L)}' \underline{Q}_n(L) X_t^n. \tag{3.5}$$

Since

$$\mathrm{Cov}\big(\overline{\underline{P}_n(L)}' \underline{P}_n(L) X_{t+h}^n, \overline{\underline{Q}_n(L)}' \underline{Q}_n(L) X_t^n\big) = \int_{-\pi}^{\pi} e^{ih\theta}\overline{P_n(\theta)}' P_n(\theta) \Sigma_n^X(\theta) Q_n(\theta) \overline{Q_n(\theta)}' d\theta = 0,$$

the summands of (3.5) are orthogonal at any lead and lag. So, $\overline{\underline{P}_n(L)}' \underline{\Lambda}_n^{-1/2}(L)\psi_t^n$ is the projection of X_t^n onto $\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$. Notice again that $\overline{\underline{P}_n(L)}' \underline{\Lambda}_n^{-1/2}(L)\psi_t^n = \overline{\underline{P}_n(L)}' \underline{P}_n(L) X_t^n$ is the filter that retains as much information about $\{X_t^n\}$ as possible under the constraint that its spectral density has rank q. This has already been discussed in Section 2.2 and in Theorem 2.2.

Since it will become important in the upcoming section, we further want to project $\{\psi_t^m\}$ onto $\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$ for $n > m$. To obtain this projection of

$$\psi_t^m = \underline{\Lambda}_m^{-1/2}(L) \underline{P}_m(L) X_t^m$$

onto $\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$, one just applies $\underline{\Lambda}_m^{-1/2}(L)\underline{P}_m(L)$ from the left to both sides of (3.5). This yields

$$\psi_t^m = \underline{T}(L)\psi_t^n + \underline{S}(L) X_t^n, \tag{3.6}$$

with

$$T(\theta) = \Lambda_m^{-1/2}(\theta) P_m(\theta) \overline{P_n(\theta)}' \Lambda_n^{1/2}(\theta)$$

and

$$S(\theta) = \Lambda_m^{-1/2}(\theta) P_m(\theta) \overline{Q_n(\theta)}' Q_n(\theta).$$

3.2 Recovering the common component in the population case

In this section we will deal with the question of how to recover the common component χ_{it} of the corresponding observable x_{it} of the infinite dimensional stationary vector process $\{X_t\}$. We propose a sequence of univariate processes $(\{\chi_{it,n}\})_{n\in\mathbb{N}}$ based on the first q principal components of $\{X_t^n\}$ and prove its convergence to χ_{it} in mean square for $t \in \mathbb{Z}, i \in \mathbb{N}$. This also forms the basis for recovering the common component in the sample case, where an appropriate estimator $\hat{\chi}_{it,n}$ of $\chi_{it,n}$ will be proposed.
Consider sequence

$$\chi_{it,n} = (\overline{\underline{P}_n(L)}'\underline{P}_n(L)X_t^n)_i = (\overline{\underline{P}_n(L)}'\underline{\Lambda}_n^{1/2}(L)\psi_t^n)_i.$$

As we have previously argued in equation (3.5) and onward

$$\chi_{it,n} = \mathrm{Proj}_{\overline{\mathrm{span}}(\{\psi_{jt}^n : j\in\{1,\ldots,q\}, t\in\mathbb{Z}\})}(x_{it}).$$

We know χ_{it} is driven by an underlying q-dimensional white noise vector process $\{u_t\}$. By the properties $\lambda_q^\chi(\theta) = \infty$ a.e. in $[-\pi, \pi]$ and $\lambda_1^\xi(\theta)$ being essentially bounded, we can see that the covariance of the idiosyncratic component of $\{X_t\}$ is limited contrary to the covariance of the common component. So intuitively it makes sense to project χ_{it} onto the space spanned by the first q principal components as they retain the most possible "information" under the constraint that the spectral density has rank q. In order to show the convergence of $\chi_{it,n}$ to χ_{it} we need to show first of all that the space we project x_{it} onto, i.e. $\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1, \ldots, q\}, t \in \mathbb{Z}\})$, converges. For that reason we introduce the following definition.

Definition 3.4. Let $(\{v_t^n\})_{n\in\mathbb{N}}$ be a sequence of q-dimensional orthogonal white noise vector processes with unit spectral density and $v_t^n \in L_2^q(\mathcal{P}, \mathbb{C})$ for $t \in \mathbb{Z}$, $n \in \mathbb{N}$. Suppose that $\{v_t^n\}$ and $\{v_t^m\}$ are costationary for all

$n, m \in \mathbb{N}$. Consider the orthogonal projection of v_t^m onto the space $\overline{\text{span}}(\{v_{jt}^n : j \in \{1, \dots, q\}, t \in \mathbb{Z}\})$,

$$v_t^m = A^{mn}(L)v_t^n + \rho_t^{mn}.$$

Here the filter $A^{mn}(L)$ functions as projection coefficients, i.e. coefficient A_k^{mn} projects v_t^m onto $\overline{\text{span}}(\{v_{j,t-k}^n : j \in \{1, \dots, q\}\})$ and ρ_t^{mn} is the residual process. Denote the spectral density of ρ_t^{mn} by $\Sigma^{\rho^{mn}}(\theta)$. If

$$\lim_{m,n\to\infty} \text{tr}(\Sigma^{\rho^{mn}}(\theta)) = 0$$

for almost all θ in $[-\pi, \pi]$ holds, we say, the sequence $(\{v_t^n\})_{n\in\mathbb{N}}$ with corresponding spaces $\overline{\text{span}}(\{v_{jt}^n : j \in \{1, \dots, q\}, t \in \mathbb{Z}\})$ generates a Cauchy sequence of spaces. See also [4, p.18].

Remark. Intuitively, the definition of the Cauchy sequence of spaces can be understood the same way as basic Cauchy sequences. As n and m go to infinity the spaces become more and more similar. This means that space $\overline{\text{span}}(\{v_{jt}^n : j \in \{1, \dots, q\}, t \in \mathbb{Z}\})$ eventually converges to some space.

Remark. This filter $A^{mn}(L)$ representing the projection behaviour exists due to the fact that $\{v_t^m\}$ is stationary and orthogonal at any lead and lag for any $m \in \mathbb{N}$. Notice that coefficient A_k^{mn} is by basic properties of projections $\text{Cov}(\psi_{t+h}^m, \psi_t^n)$ and this is by stationarity independent of t. Hence this filter $A^{mn}(L)$ represents the projection of v_t^m onto $\overline{\text{span}}(\{v_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$ for any $t \in \mathbb{Z}$. Furthermore notice that since $\{v_t^n\}$ and $\{v_t^m\}$ are in $L_2^q(\mathcal{P}, \mathbb{C})$ also $A^{mn}(L)v_t^n$ and ρ_t^{mn} are in $L_2^q(\mathcal{P}, \mathbb{C})$ making filter $A^{mn}(L)$ also square summable.

We will show that the sequence of q-dimensional vector processes consisting of the q normalized principal components of $\{X_t^n\}$, $(\{\psi_t^n\})_{n\in\mathbb{N}}$, generates such a Cauchy sequence of spaces. Before we are able to do that, we need the following result.

Lemma 3.3. Let $\{X_t\}$ be an infinite dimensional stationary vector process with spectral densities $\Sigma_n^X(\theta)$ of corresponding vector processes $\{X_t^n\}$ for any

$n \in \mathbb{N}$. Let $C(\theta)$ be a $(q \times q)$-matrix with entries $c_{ij}(\theta) \in L_2([-\pi,\pi],\mathbb{C})$ and $C(\theta)\overline{C(\theta)}' = I_q$ for any $\theta \in [-\pi,\pi]$. Note that

$$\underline{C}(L)\psi_t^m = \underline{D}(L)\psi_t^n + \underline{R}(L)X_t^n$$

with

$$D(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{P_n(\theta)}'\Lambda_n^{1/2}(\theta)$$

and

$$R(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{Q_n(\theta)}'Q_n(\theta),$$

based on equation (3.6) where occuring matrices $\Lambda_n(\theta), \Lambda_m(\theta), P_n(\theta), P_n(\theta)$, $Q_n(\theta)$ and upcoming matrix $\Phi_n(\theta)$ are as defined at the end of Section 3.1. Then for the first eigenvalue, call it $\rho(\theta)$, of the spectral density matrix of $\underline{R}(L)X_t^n$, it holds that:

$$\rho(\theta) \leq \frac{\lambda_{n,q+1}^X(\theta)}{\lambda_{mq}^X(\theta)}. \tag{3.7}$$

See [4, p.14].

Proof. Since for any $x \in \mathbb{C}^n$

$$\begin{aligned}&\bar{x}'\left(\lambda_{n,q+1}^X(\theta)\overline{Q_n(\theta)}'Q_n(\theta) - \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta)\right)x = \\ &\lambda_{n,q+1}^X(\theta)(\overline{Q_n(\theta)x})'(Q_n(\theta)x) - (\overline{Q_n(\theta)x})'\Phi_n(\theta)(Q_n(\theta)x) = \\ &\sum_{i=1}^{n}(\lambda_{n,q+1}^X(\theta)|(Q_n(\theta)x)_i|^2 - \lambda_{n,q+i}^X(\theta)|(Q_n(\theta)x)_i|^2) \geq 0,\end{aligned}$$

it holds that

$$\lambda_{n,q+1}^X(\theta)\overline{Q_n(\theta)}'Q_n(\theta) - \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta) \geq 0,$$

and since $I_n - \overline{Q_n(\theta)}'Q_n(\theta) \geq 0$ by equation (3.2) we have

$$\lambda_{n,q+1}^X(\theta)I_n - \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta) \geq 0.$$

Multiplying $C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)$ from the left and $\overline{P_m(\theta)}'\Lambda_m^{-1/2}(\theta)\overline{C(\theta)}'$ from the right yields

$$\lambda_{n,q+1}^X(\theta)C(\theta)\Lambda_m^{-1}(\theta)\overline{C(\theta)}' - R(\theta)\overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta)\overline{R(\theta)}' \geq 0.$$

Because $\overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta) = \Sigma_n^X(\theta) - \overline{P_n(\theta)}'\Lambda_n(\theta)P_n(\theta)$ holds by (3.1) and $R(\theta)\overline{P_n(\theta)}' = 0$, we have that the above equality is equivalent to

$$\lambda_{n,q+1}^X(\theta)C(\theta)\Lambda_m^{-1}(\theta)\overline{C(\theta)}' - R(\theta)\Sigma_n^X(\theta)\overline{R(\theta)}' \geq 0.$$

Using Lemma 7.2 finally yields (3.7). The proof follows [4, p.14]. □

With the lemma above, it is possible to show that $(\{\psi_t^n\})_{n\in\mathbb{N}}$ generates a Cauchy sequence of spaces.

Lemma 3.4. The sequence of processes $(\{\psi_t^n\})_{n\in\mathbb{N}}$, where $\{\psi_t^n\}$ consists of the q normalized dynamic prinicpal components of process $\{X_t^n\}$ which results from an infinite dimensional stationary vector process $\{X_t\}$ generates a Cauchy sequence of spaces if $\lambda_q^X(\theta) = \infty$ a.e. in $[-\pi, \pi]$ and $\lambda_{q+1}^X(\theta)$ is essentially bounded. See [4, p.19f].

Remark. Assumptions $\lambda_q^X(\theta) = \infty$ a.e. in $[-\pi, \pi]$ and essential boundedness of $\lambda_{q+1}^X(\theta)$ are fulfilled if $\{X_t\}$ follows a generalized dynamic factor model by Lemma 3.2.

Proof. Let $n > m$. Notice that $\{\psi_t^m\}$ and $\{\psi_t^n\}$ are costationary as filtered processes stemming from the same stationary process. Consider the orthogonal projection of ψ_t^m onto $\overline{\text{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$

$$\psi_t^m = \underline{T}(L)\psi_t^n + \rho_t^{mn}$$

where

$$T(\theta) = \Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{P_n(\theta)}'\Lambda_n^{-1/2}(\theta)$$

and

$$\rho_t^{mn} = \underline{\Lambda}_m^{-1/2}(L)\underline{P}_m(L)\overline{\underline{Q}_n(L)}'\underline{Q}_n(L)X_t^n.$$

By application of Lemma 3.3 with $C(\theta) = I_q$ we have that the first eigenvalue $\rho(\theta)$ of the spectral density $\Sigma^{\rho^{mn}}(\theta)$ of the corresponding vector process fulfills

$$\rho(\theta) \leq \frac{\lambda^X_{n,q+1}(\theta)}{\lambda^X_{mq}(\theta)} \xrightarrow[m,n\to\infty]{} 0, \qquad \text{a.e. in } [-\pi, \pi],$$

due to $\lambda^X_{mq}(\theta) \to \infty$ as m goes to infinity and essential boundedness of $\lambda^X_{q+1}(\theta)$. Since $\{\psi^m_t\}$ and $\{\psi^n_t\}$ are costationary,

$$\Gamma^{mn}(h) := \mathrm{E}(\psi^m_{t+h}\overline{\psi^n_t}')$$

does not depend on t and the orthogonal projection filter of process $\{\psi^m_t\}$ can be written as

$$\underline{T}(L) = A^{mn}(L) = \sum_{h=-\infty}^{\infty} \Gamma^{mn}(h)L^h.$$

As

$$\Gamma^{nm}(h) := E(\psi^n_{t+h}\overline{\psi^m_t}') = \overline{E(\psi^m_t \overline{\psi^n_{t+h}}')}' = \overline{\Gamma^{mn}(-h)}',$$

we have

$$A^{nm}(L) = \sum_{h=-\infty}^{\infty} \Gamma^{nm}(h)L^h = \sum_{h=-\infty}^{\infty} \overline{\Gamma^{mn}(h)}' L^{-h} = \overline{\underline{T}}'(L^{-1}) =: \overline{\underline{T}(L)}'.$$

Conversely, ψ^n_t can be decomposed into

$$\psi^n_t = \overline{\underline{T}(L)}'\psi^m_t + \rho^{nm}_t.$$

Taking the spectral densities of the projection equations for $\{\psi^m_t\}$ and $\{\psi^n_t\}$ above, we obtain

$$I_q = T(\theta)\overline{T(\theta)}' + \Sigma^{\rho^{mn}}(\theta) = \overline{T(\theta)}'T(\theta) + \Sigma^{\rho^{nm}}(\theta).$$

Since traces of $T(\theta)\overline{T(\theta)}'$ and $\overline{T(\theta)}'T(\theta)$ are equal, also

$$\mathrm{tr}(\Sigma^{\rho^{mn}}(\theta)) = \mathrm{tr}(\Sigma^{\rho^{nm}}(\theta)).$$

This implies that also

$$\mathrm{tr}(\Sigma^{\rho^{nm}}(\theta)) \longrightarrow 0,$$

as n and m go to infinity for almost all θ in $[-\pi, \pi]$, which yields the claim. The proof is based on [4, p.19f]. □

The following lemma ensures that the sequence of projections of a costationary process onto the space spanned by an orthogonal white noise vector process with unit spectral density that generates a Cauchy sequence of spaces converges in mean square. This is a fundamental result in proving convergence of $\chi_{it,n}$ to χ_{it}.

Lemma 3.5. Let $(\{v_t^n\})_{n\in\mathbb{N}}$ be a sequence of q-dimensional orthogonal white noise vector processes with unit spectral density and $v_t^n \in L_2^q(\mathcal{P},\mathbb{C})$ for all $n \in \mathbb{N}$, that generates a Cauchy sequence of spaces. Let $\{z_t\}$ be a univariate process with $z_t \in L_2(\mathcal{P},\mathbb{C})$ and spectral density $\Sigma^z(\theta)$. Suppose $\{v_t^n\}$ and $\{z_t\}$ are costationary for any $n \in \mathbb{N}$. Consider the projection Z_t^n of z_t onto $\overline{\text{span}}(\{v_{jt}^n : j \in \{1,\dots,q\}, t \in \mathbb{Z}\})$. Then there exists an element $Z_t \in L_2(\mathcal{P},\mathbb{C})$ such that

$$Z_t^n = \text{Proj}_{\overline{\text{span}}(\{v_{jt}^n : j\in\{1,\dots,q\}, t\in\mathbb{Z}\})}(z_t) \xrightarrow[n\to\infty]{} Z_t,$$

where the convergence is in mean square. See [4, p.18f].

Proof. Process $\{z_t\}$ can be decomposed into the following two orthogonal projections:

$$z_t = Z_t^n + r_t^n = c_n(L)v_t^n + r_t^n$$
$$z_t = Z_t^m + r_t^m = c_m(L)v_t^m + r_t^m,$$

where $c_n(L)$ and $c_m(L)$ are square summable $(1 \times q)$-dimensional filters and $r_t^n, r_t^m \in L_2(\mathcal{P},\mathbb{C})$ due to v_t^n and v_t^m being in $L_2(\mathcal{P},\mathbb{C})$. We will now investigate the spectral density $\Sigma^{Z^n-Z^m}(\theta)$ of the corresponding vector process

$$Z_t^n - Z_t^m = c_n(L)v_t^n - c_m(L)v_t^m = r_t^m - r_t^n,$$

which is also the cross spectrum between $c_n(L)v_t^n - c_m(L)v_t^m$ and $r_t^m - r_t^n$. By definition this cross spectrum is

$$\begin{aligned}&\frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\theta h}\text{Cov}(c_n(L)v_{t+h}^n - c_m(L)v_{t+h}^m, r_t^m - r_t^n) = \\ &\frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\theta h}\big(\text{Cov}(c_n(L)v_{t+h}^n, r_t^m) + \text{Cov}(c_m(L)v_{t+h}^m, r_t^n)\big),\end{aligned} \tag{3.8}$$

where the last equation comes from the orthogonality of $c_m(L)v_t^m$ and r_t^m resp. $c_n(L)v_t^n$ and r_t^n. Because $v_{t+h}^m = A^{mn}(L)v_{t+h}^n + \rho_t^{mn}$, where $A^{mn}(L)$ and ρ_t^{mn} are as in Definition 3.4, it holds that

$$\mathrm{Cov}(c_m(L)v_{t+h}^m, r_t^n) = \mathrm{Cov}(c_m(L)(A^{mn}(L)v_{t+h}^n + \rho_t^{mn}), r_t^n) = \mathrm{Cov}(c_m(L)\rho_t^{mn}, r_t^n),$$

and analogously

$$\mathrm{Cov}(c_n(L)v_{t+h}^n, r_t^m) = \mathrm{Cov}(c_n(L)\rho_t^{nm}, r_t^m).$$

Since

$$\Sigma^z(\theta) = c_m^\circ(\theta)\overline{c_m^\circ(\theta)}' + \Sigma^{r^m}(\theta),$$

$\Sigma^{r^m}(\theta)$ and the absolute value of entries in $c_m^\circ(\theta)$ are bounded by $\Sigma^z(\theta)$. Because $(\{v_t^n\})_{n\in\mathbb{N}}$ generates a Cauchy sequence of spaces, it holds that $\Sigma^{\rho^{mn}}(\theta) \to 0$ for almost all θ in $[-\pi, \pi]$ as n and m go to infinity, which implies by the just observed boundedness that

$$\frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\theta h}\mathrm{Cov}(c_m(L)\rho_{t+h}^{mn}, r_t^n) \xrightarrow[n,m\to\infty]{} 0$$

for almost all θ in $[-\pi, \pi]$. An analogous argumentation yields

$$\frac{1}{2\pi}\sum_{h=-\infty}^{\infty} e^{-i\theta h}\mathrm{Cov}(c_n(L)\rho_{t+h}^{nm}, r_t^m) \xrightarrow[n,m\to\infty]{} 0,$$

for almost all θ in $[-\pi, \pi]$, which in total yields

$$\Sigma^{Z^n - Z^m}(\theta) \xrightarrow[n,m\to\infty]{} 0,$$

for almost all θ in $[-\pi, \pi]$. Since $\Sigma^{Z^n}(\theta)$ and $\Sigma^{Z^m}(\theta)$ are bounded by $\Sigma^z(\theta)$, the dominated convergence theorem (Theorem 2.3) can be applied, and yields

$$E((Z_t^n - Z_t^m)^2) = \mathrm{Var}(Z_t^n - Z_t^m) = \int_{-\pi}^{\pi} \Sigma^{Z^n - Z^m}(\theta)d\theta \xrightarrow[n,m\to\infty]{} 0.$$

Hence it was proven that $(Z_t^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since $L_2(\mathcal{P}, \mathbb{C})$ is a Hilbert space, there exists a limit $Z_t \in L_2(\mathcal{P}, \mathbb{C})$. The proof is based on [4, p.19].

□

We are now finally able to prove the convergence of $\chi_{it,n}$ to χ_{it}.

Theorem 3.2. Let $\{X_t\}$ be an infinite dimensional stationary vector process that follows a generalized dynamic factor model with q factors. Then the sequence $\chi_{it,n} = (\overline{\underline{P}_n(L)}'\underline{P}_n(L)X_t^n)_i$ converges in mean square to χ_{it} for any $i \in \mathbb{N}$, $t \in \mathbb{Z}$. See [5, p.543].

Proof. Since $(\{\psi_t^n\})_{n\in\mathbb{N}}$ generates a Cauchy sequence of spaces (Lemma 3.4) and $\{x_{it}\}$ and $\{\psi_t^n\}$ are costationary for any $n \in \mathbb{N}$ and fixed i, Lemma 3.5 can be applied. Hence, there exists χ_{it}^* such that

$$\chi_{it,n} = \mathrm{Proj}_{\overline{\mathrm{span}}(\{\psi_{jt}^n : j\in\{1,\ldots,q\},t\in\mathbb{Z}\})}(x_{it}) \xrightarrow[n\to\infty]{} \chi_{it}^*$$

in mean square. If we can prove that

$$\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\}) \xrightarrow[n\to\infty]{} \overline{\mathrm{span}}(\{u_{jt} : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\}),$$

then $\chi_{it} = \chi_{it}^*$, since $\{u_t\}$ and $\{\xi_t\}$ are orthogonal at any lead and lag. Processes $\{\psi_t^n\}$ and $\{u_t\}$ are costationary for any $n \in \mathbb{N}$, hence $\mathrm{E}(\psi_{t+h}^n \overline{u_t}') =: \Gamma^{\psi^n u}(h)$ does not depend on t and the orthogonal projection of ψ_t^n onto $\overline{\mathrm{span}}(\{u_{jt} : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\})$ is given by

$$\psi_t^n = A_n(L)u_t + \rho_t^n,$$

where

$$A_n(L) = \sum_{h=-\infty}^{\infty} \Gamma^{\psi^n u}(h)L^h.$$

Analogously since $\Gamma^{u\psi^n}(h) := \mathrm{E}(u_{t+h}\overline{\psi_{t-h}^n}') = \overline{\Gamma^{\psi^n u}(-h)}'$, the orthogonal projection of u_t onto $\overline{\mathrm{span}}(\{\psi_{jt}^n : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\})$ is given by

$$u_t = \overline{A_n(L)}'\psi_t^n + \epsilon_t,$$

where

$$\overline{A_n(L)}' = \sum_{h=-\infty}^{\infty} \overline{\Gamma^{\psi^n u}(h)}'L^{-h}.$$

Furthermore, ψ_t^n has the following decompositions

$$\psi_t^n = \underline{\Lambda}_n^{-1/2}(L)\underline{P}_n(L)X_t^n = \underline{\Lambda}_n^{-1/2}(L)\underline{P}_n(L)(\chi_t^n + \xi_t^n) = A_n(L)u_t + \rho_t^n.$$

It follows from the above equation and the fact that $\{u_t\}$ and $\{\xi_t\}$ are orthogonal that

$$\rho_t^n = \underline{\Lambda}_n^{-1/2}(L)\underline{P}_n(L)\xi_t^n$$

and hence that the spectral density $\Sigma_n^\rho(\theta)$ can be decomposed as

$$\Sigma_n^\rho(\theta) = \Lambda_n^{-1/2}(\theta)P_n(\theta)\Sigma_n^\xi(\theta)\overline{P_n(\theta)}'\Lambda_n^{-1/2}(\theta).$$

For the i-th entry on the diagonal we obtain

$$(\Sigma_n^\rho(\theta))_{ii} \leq \lambda_{n1}^\xi(\theta)(\lambda_{ni}^X(\theta))^{-1}|p_{ni}^X(\theta)|^2 \leq \frac{\lambda_1^\xi(\theta)}{\lambda_{ni}^X(\theta)} \xrightarrow[n\to\infty]{} 0$$

for almost all θ in $[-\pi,\pi]$ and all $i \in \{1,\ldots,q\}$ due to the essential boundedness of $\lambda_1^\xi(\theta)$ and the fact that $\lambda_{ni}^X(\theta) \to \infty$ as $n \to \infty$ for $i \in \{1,...,q\}$ and for almost all θ in $[-\pi,\pi]$, (Lemma 3.2) . Also, by the assumption that $\lambda_{ni}^X(\theta) \geq 1$ for $i \leq n$, $\theta \in [-\pi,\pi]$, there exists $K \in \mathbb{R}$ such that for almost all θ in $[-\pi,\pi]$ $(\Sigma_n^\rho(\theta))_{ii} \leq K$. Thus, by the application of Theorem 2.3,

$$\int_{-\pi}^{\pi} (\Sigma_n^\rho(\theta))_{ii} d\theta \xrightarrow[n\to\infty]{} 0$$

holds for all $i \in \{1,...,q\}$. Mean square convergence follows. Notice that $\text{tr}(\Sigma_n^\rho(\theta))$ corresponds to the sum of the eigenvalues of $\Sigma_n^\rho(\theta)$. Hence, also the eigenvalues converge to zero in mean square and thus the whole spectral density matrix $\Sigma_n^\rho(\theta)$. Moreover, as $\{\psi_t^n\}$ and $\{u_t\}$ are white noise processes, we obtain the following spectral density equations

$$I_q = A_n^\circ(\theta)\overline{A_n^\circ(\theta)}' + \Sigma_n^\rho(\theta) = \overline{A_n^\circ(\theta)}'A_n^\circ(\theta) + \Sigma_n^\epsilon(\theta).$$

Because $\text{tr}(A_n^\circ(\theta)\overline{A_n^\circ(\theta)}') = \text{tr}(\overline{A_n^\circ(\theta)}'A_n^\circ(\theta))$, also $\text{tr}(\Sigma_n^\rho(\theta)) = \text{tr}(\Sigma_n^\epsilon(\theta))$, which converges to zero in mean square. Hence we have proven, that the space $\overline{\text{span}}(\{\psi_{jt}^n : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\})$ converges to the space $\overline{\text{span}}(\{u_{jt} : j \in \{1,\ldots,q\}, t \in \mathbb{Z}\})$, which concludes this proof. The proof is based on [5, p.552]. □

Remark 3.2. We assumed in Remark 3.1 w.l.o.g. that for $n \in \mathbb{N}$, $j \leq n$ and $\theta \in [-\pi, \pi]$, it holds that $\lambda_{nj}^X(\theta) \geq 1$. We will now discuss the validity of this assumption in order for Theorem 3.2 to hold. The probability space $\mathcal{P}$ can be embedded into a larger probability space such that it can be assumed that there exists a infinite dimensional white noise process $\{\zeta_t : t \in \mathbb{Z}\}$ having unit spectral density with $\zeta_{it_1} \perp x_{jt_2}$ for any $i, j \in \mathbb{N}$ and any $t_1, t_2 \in \mathbb{Z}$. Define the new process $\{\hat{x}_{it} : t \in \mathbb{Z}\}$ as $\hat{x}_{it} = x_{it} + \zeta_{it}$. Then $\Sigma_n^{\hat{X}}(\theta) = \Sigma_n^X(\theta) + I_n$ and hence also $\lambda_{nj}^{\hat{X}}(\theta) = \lambda_{nj}^X(\theta) + 1$ with $p_{nj}^{\hat{X}}(\theta) = p_{nj}^X(\theta)$ for $j \in \{1, \ldots, n\}, n \in \mathbb{N}$. Notice that $\underline{\pi}_{ni}(L)\underline{P}_n(L)$, where $\underline{\pi}_{ni}(L)$ denotes the i-th row of $\overline{\underline{P_n}(L)}'$, is a DAS. (For more details check step vi) of the proof of Theorem 4.1). Hence, it holds that

$$\|\underline{\pi}_{ni}(L)\underline{P}_n(\theta)\zeta_t^n\|^2 = \int_{-\pi}^{\pi} \pi_{ni}(\theta)P_n(\theta)I_n\overline{P_n(\theta)}'\overline{\pi_{ni}(\theta)}'d\theta \xrightarrow[n\to\infty]{} 0.$$

And hence since by Theorem 3.2, $\lim_{n\to\infty} \chi_{it,n} = \lim_{n\to\infty} \underline{\pi}_{ni}(L)\underline{P}_n(L)\hat{X}_t^n = \chi_{it}$, also

$$\lim_{n\to\infty} \underline{\pi}_{ni}(L)\underline{P}_n(L)\hat{X}_t^n = \lim_{n\to\infty} \underline{\pi}_{ni}(L)\underline{P}_n(L)X_t^n = \chi_{it}.$$

See [5, p.551].

4 The Characterization Theorem

We have seen in Chapter 3, specifically in Lemma 3.2, that under the assumption that an infinite dimensional stationary vector process $\{X_t\}$ follows a generalized dynamic factor model with q factors, we can deduce some properties about the dynamic eigenvalues of the spectral density $\Sigma_n^X(\theta)$ of the observable vector processes $\{X_t^n\}$. As already mentioned these properties can help in practice to determine a suitable q. In this section we will prove that these properties are not only necessary but also sufficient for a generalized dynamic factor model with q factors to hold for an infinite dimensional stationary vector process $\{X_t\}$. This makes the property concerning the dynamic eigenvalues of the spectral densities a characterization for the generalized dynamic factor model. In the proof of this characterization we will need the following definition.

Definition 4.1. Let $y_t \in \overline{X}$. If there exists a DAS $(a_n(L))_{n\in\mathbb{N}}$ of $(1 \times s_n)$-dimensional filters such that

$$\lim_{n\to\infty} a_n(L)X_t^{s_n} = y_t,$$

in mean square, then y_t is called an aggregate. The set of all aggregates is denoted by $\mathcal{G}(X)$. See [4, p.6].

Remark. The aggregation space $\mathcal{G}(X)$ is a closed subspace of $\overline{X}$. A proof can be found in [4, p.7].

Theorem 4.1. Let $\{X_t\}$ be an infinite dimensional stationary vector process. Then $\{X_t\}$ follows a generalized dynamic factor model with q factors if and only if the following holds:

(I) $\lambda_{q+1}^X(\theta)$ is essentially bounded,

(II) $\lambda_q^X(\theta) = \infty$ a.e. in $[-\pi, \pi]$.

The theorem is taken from [4, p.10].

Proof. This proof follows the outline of the proof found in [4, p.11f].
The "only if" part is exactly Lemma 3.2. Proving the "if" part takes some effort, thus we will divide this part into two major steps, say I and II, which will also be divided into smaller steps to make the idea of the proof clearer. Step I (i-iv) consists of showing that there exists a q-dimensional white noise vector process $\{z_t\}$ with unit spectral density in $\mathcal{G}(X)$. In step II (v-vi) we show that the residual process $\{\beta_t\}$ of the componentwise projection of $\{X_t\}$ onto $\mathcal{G}(X)$ is idiosyncratic.

Here we list the details of these steps.

i) Define orthogonal white noise vector process $\{\underline{C}(L)\psi_t^m\}$ with unit spectral density, whose entries are linear combinations of the first q dynamic principal components of $\{X_t^m\}$ and project it onto the space spanned by the first q principal components of $\{X_t^n\}$ where $n > m$.

ii) Based on the above vector processes, construct for a set $M \subset [-\pi, \pi]$ with positive measure a converging white noise vector process sequence $(\{v_t^k\})_{k\in\mathbb{N}}$ whose limit $\{v_t\}$ has the identity as spectral density a.e. on M and 0 outside of it.

iii) Based on $\{v_t\}$ construct a white noise vector process $\{z_t\}$ such that its spectral density is the identity a.e. on $[-\pi, \pi]$ and 0 else.

iv) Prove that process $\{z_t\}$ spans $\mathcal{G}(X)$.

v) Show that the sequence of projections of x_{it} onto the first q dynamic principal components of $\{X_t^n\}$ converges in mean square to the projection of x_{it} onto $\mathcal{G}(X)$.

vi) Show that the residual of the componentwise projection of x_{it} onto $\mathcal{G}(X)$, $\{\beta_t\}$, is idiosyncratic.

W.l.o.g. we can assume that for any $n \in \mathbb{N}$, $j \leq n$ and $\theta \in [-\pi, \pi] : \lambda_{nj}^X(\theta) \geq 1$ (see remark at the end of this proof). We will now briefly recap definitions introduced in Chapter 3. Let $p_{nj}(\theta)$ be the eigenvectors as rows corresponding to eigenvalues $\lambda_{nj}(\theta)$ of $\Sigma_n^X(\theta)$ for $j \in \{1, \ldots, n\}$. $P_n(\theta)$ is the $(q \times n)$-matrix having $p_{nj}(\theta)$ as the j-th row. Let $\Lambda_n(\theta)$ be the $(q \times q)$ diagonal matrix having $\lambda_{nj}^X(\theta)$ on the j-th diagonal entry for $j \in \{1, \ldots, q\}$. Analogously let $Q_n(\theta)$ be the $((n-q) \times n)$-matrix having $p_{n,q+j}(\theta)$ on the j-th row for $j \in \{1, \ldots, n-q\}$. And let $\Phi_n(\theta)$ be the $(n-q) \times (n-q)$ diagonal matrix with $\lambda_{n,q+j}^X(\theta)$ as the j-th diagonal entry for $j \in \{1, \ldots, n-q\}$. And lastly we have white noise vector process $\{\psi_t^n\}$ with $\psi_t^n = \underline{\Lambda}_n^{-1/2}(L)\underline{P}_n(L)X_t^n$ with unit spectral density.

Ad i): Let $C(\theta)$ be a $(q \times q)$-matrix such that $c_{ij}(\theta) \in L_2([-\pi, \pi], \mathbb{C})$ is essentially bounded in modulus for $i, j \in \{1, \ldots, n\}$ and such that $C(\theta)\overline{C(\theta)}' = I_q$ for any $\theta \in [-\pi, \pi]$. Then the covariance matrix of $\underline{C}(L)\psi_t^m$ is bounded componentwise, since

$$\int_{-\pi}^{\pi} C(\theta)\Sigma_n^\psi(\theta)\overline{C(\theta)}'d\theta = \int_{-\pi}^{\pi} C(\theta)I_q\overline{C(\theta)}'d\theta < \infty,$$

componentwise. As previously discussed the orthogonal projection of $\underline{C}(L)\psi_t^m$ onto the space spannend by the first q dynamic principal components of $\{X_t^n\}$, i.e. $\overline{\text{span}}(\{\psi_{jt}^n : j \in \{1, ..., q\}, t \in \mathbb{Z}\})$ is according to equation (3.6) given as

$$\underline{C}(L)\psi_t^m = \underline{D}(L)\psi_t^n + \underline{R}(L)X_t^n,$$

where

$$D(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{P_n(\theta)}'\Lambda_n^{1/2}(\theta),$$

and

$$R(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{Q_n(\theta)}'Q_n(\theta).$$

Ad ii): Now we start constructing a convergent white noise vector process sequence on a set $M \subset [-\pi, \pi]$ with $\mathcal{L}(M) > 0$. Due to assumptions, $\lambda_{q+1}^X(\theta) \leq w$

a.e. in $[-\pi, \pi]$ for some $w \in \mathbb{R}$ and $\lambda_q^X(\theta) = \infty$ a.e. in $[-\pi, \pi]$, there exists set M with $\mathcal{L}(M) > 0$ such that for all $\theta \in M$ and all $n \in \mathbb{N}$, $\lambda_{n,q+1}^X(\theta) < W$ for some $W \in \mathbb{R}$ and $\lambda_{nq}^X(\theta) \geq \alpha_n$ with $(\alpha_n)_{n\in\mathbb{N}}$ some real positive non-decreasing sequence with $\lim\limits_{n\to\infty} \alpha_n = \infty$. Furthermore define the set

$$\begin{aligned} K_M = &\{F(\theta) : (q \times q)\text{-matrix}, f_{ij}(\theta) \in L_2([-\pi, \pi], \mathbb{C}) \ \forall i \ \forall j, \\ &F(\theta) = 0 \text{ for } \theta \notin M, \ F(\theta)\overline{F(\theta)'} = I_q \text{ for } \theta \in M\}. \end{aligned}$$

As we are considering $\underline{C}(L)\psi_t^m$ as orthogonal projection $\underline{D}(L)\psi_t^n + \underline{R}(L)X_t^n$, its spectral density can be decomposed as

$$I_q = D(\theta)\overline{D(\theta)}' + R(\theta)\Sigma_n^X(\theta)\overline{R(\theta)}'. \tag{4.1}$$

For the first eigenvalue $\rho(\theta)$ of $R(\theta)\Sigma_n^X(\theta)\overline{R(\theta)}'$ it holds due to Lemma 3.3 that

$$\rho(\theta) \leq \frac{\lambda_{n,q+1}^X(\theta)}{\lambda_{mq}^X(\theta)} < \frac{W}{\alpha_m}, \qquad \forall\, \theta \in M,$$

where the second inequality follows from the assumption at the beginning of step ii). Denote by $\delta_j(\theta)$ the eigenvalues of $D(\theta)\overline{D(\theta)}'$ in descending order. By Lemma 7.2 and equation (4.1) the following holds:

$$1 \leq \delta_q(\theta) + \rho(\theta) < \delta_q(\theta) + \tfrac{W}{\alpha_m}.$$

Choose m^* such that

$$\frac{W}{\alpha_{m^*}} < 1,$$

in order to have

$$\delta_q(\theta) > 1 - \tfrac{W}{\alpha_{m^*}} > 0, \qquad \forall\, \theta \in M, \ \forall\, m \geq m^*.$$

Denote by $\Delta(\theta)$ the diagonal matrix having the eigenvalue $\delta_j(\theta)$ at the j-th diagonal entry, i.e.

$$\Delta(\theta) = \begin{bmatrix} \delta_1(\theta) & 0 & \dots & 0 \\ 0 & \delta_2(\theta) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \delta_q(\theta) \end{bmatrix}.$$

The above inequality, $\delta_q(\theta) > 0$, implies $1/\delta_j(\theta)$ is bounded for $j \in \{1, \ldots, q\}$ making $\Delta^{-1/2}(\theta)$ well-defined. Let $H(\theta)$ be a measurable $(q \times q)$-matrix in M satisfying

a) $H(\theta)\overline{H(\theta)}' = I_q, \qquad \forall\, \theta \in M,$

b) $H(\theta)\Delta(\theta)\overline{H(\theta)} = D(\theta)\overline{D(\theta)}', \quad \forall\, \theta \in M.$

Define $F(\theta)$ by

$$F(\theta) = \begin{cases} H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}'D(\theta), & \theta \in M, \\ 0, & \theta \notin M. \end{cases} \tag{4.2}$$

As the product of L_2 functions is again in L_2 and

$$\begin{aligned} F(\theta)\overline{F(\theta)}' &= H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}'D(\theta)\overline{D(\theta)}'H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}' \\ &= H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}'H(\theta)\Delta(\theta)\overline{H(\theta)}'H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}' = I_q, \end{aligned}$$

$F(\theta)$ is an element of K_M.

After this preparatory work we are finally able to construct a q-dimensional vector process $\{v_t\}$ such that v_{jt} is an aggregate for $j \in \{1, ..., q\}$, $t \in \mathbb{Z}$ and such that the spectral density matrix of $\{v_t\}$ is I_q for $\theta \in M$ and 0 outside of it. Choose an element $F_1(\theta)$ from K_M, set τ to be $\frac{1}{2^3\pi}$ and s_1 to be m_τ, where m_τ is chosen such that

$$2\frac{W}{\alpha_{m_\tau}} < \tau.$$

Then define the following filter

$$G_1(L) = \underline{F}_1(L)\underline{\Delta}_{s_1}^{-1/2}(L)\underline{P}_{s_1}(L),$$

and the corresponding filtered vector process $\{v_t^1\}$ via

$$v_t^1 = G_1(L)X_t^{s_1} = \underline{F}_1(L)\psi_t^{s_1},$$

The spectral density $\Sigma^{v^1}(\theta)$ of $\{v_t^1\}$ is then due to $\Sigma_{s_1}^{\psi}(\theta) = I_q$ given by

$$\Sigma^{v^1}(\theta) = F_1(\theta)\Sigma_{s_1}^{\psi}(\theta)\overline{F_1(\theta)}' = F_1(\theta)\overline{F_1(\theta)}' = \begin{cases} I_q, & \theta \in M, \\ 0, & \theta \notin M. \end{cases}$$

Continue by setting $\tau = \frac{1}{2^5\pi}$ and choosing $s_2 = m_\tau > s_1$ such that again

$$2\frac{W}{\alpha_{m_\tau}} < \tau.$$

Then calculate

$$\underline{D}(L) = \underline{F}_1(L)\underline{\Lambda}_{s_1}^{-1/2}(L)\underline{P}_{s_1}(L)\overline{\underline{P}_{s_2}(L)}'\underline{\Lambda}_{s_2}^{1/2}(L)$$

and according to (4.2) with suitable matrix $H(\theta)$ determine

$$\underline{F}_2(L) = \begin{cases} \underline{H}(L)\underline{\Delta}^{-1/2}(L)\overline{\underline{H}(L)}'\underline{D}(L), & \theta \in M, \\ 0, & \theta \notin M. \end{cases}$$

Then again set

$$G_2(L) = \underline{F}_2(L)\underline{\Lambda}_{s_2}^{-1/2}(L)\underline{P}_{s_2}(L),$$

with corresponding vector process $\{v_t^2\}$ via

$$v_t^2 = G_2(L)X_t^{s_2} = \underline{F}_2(L)\psi_t^{s_2},$$

and spectral density

$$\Sigma^{v^2}(\theta) = \begin{cases} I_q, & \theta \in M, \\ 0, & \theta \notin M. \end{cases}$$

Lemma 7.3 applied to the vector process $\{v_t^1 - v_t^2\}$ yields that the first eigenvalue of its spectral density $\Sigma^{v^1-v^2}(\theta)$, say $\eta_1(\theta)$, satisfies

$$\eta_1(\theta) < \tfrac{1}{2^3\pi}, \qquad \theta \in M,$$

so that by Courant-Fisher, (Lemma 7.1),

$$\|v_{jt}^1 - v_{jt}^2\|^2 = \int_{-\pi}^{\pi} (\Sigma^{v^1-v^2}(\theta))_{jj} d\theta \leq \int_{-\pi}^{\pi} \eta_1(\theta) d\theta < \int_{-\pi}^{\pi} \tfrac{1}{2^3\pi} d\theta = \tfrac{1}{4},$$

for $j \in \{1, \dots, q\}$. By iteration, set τ to be $\frac{1}{2^{2k+1}\pi}$ and choose $s_k = m_\tau > s_{k-1}$ such that $2\frac{W}{\alpha_{m_\tau}} < \tau$. Again determine

$$\underline{D}(L) = \underline{F}_{k-1}(L)\underline{\Lambda}_{s_{k-1}}^{-1/2}(L)\underline{P}_{s_{k-1}}(L)\overline{\underline{P}_{s_k}(L)}'\underline{\Lambda}_{s_k}^{1/2}(L),$$

and with suitable $H(\theta)$ determine

$$\underline{F}_k(L) = \begin{cases} \underline{H}(L)\underline{\Delta}^{-1/2}(L)\overline{\underline{H}(L)}'\underline{D}(L), & \theta \in M, \\ 0, & \theta \notin M, \end{cases}$$

yielding filter

$$G_k(L) = \underline{F}_k(L)\underline{\Lambda}_{s_k}^{-1/2}(L)\underline{P}_{s_k}(L),$$

and vector process $\{v_t^k\}$ via

$$v_t^k = G_k(L)X_t^{s_k} = \underline{F}_k(L)\psi_t^{s_k},$$

with

$$\Sigma^{v^k}(\theta) = \begin{cases} I_q, & \theta \in M, \\ 0, & \theta \notin M. \end{cases}$$

Again by Lemma 7.3 the first eigenvalue of the spectral density of vector process $\{v_t^{k-1} - v_t^k\}$, denoted by $\eta_{k-1}(\theta)$, satisfies

$$\eta_{k-1}(\theta) < \tfrac{1}{2^{2k-1}\pi}, \qquad \theta \in M,$$

such that again by the same argumentation as above we have

$$\|v_{jt}^{k-1} - v_{jt}^k\|^2 < \tfrac{1}{2^{2k-2}}, \qquad j \in \{1, \dots, q\}.$$

Moreover we have

$$\begin{aligned} \|v_{jt}^k - v_{jt}^{k+h}\| &\leq \sum_{i=1}^{h} \|v_{jt}^{k+i-1} - v_{jt}^{k+i}\| < \sum_{i=1}^{h} \frac{1}{2^{k+i-1}} = \sum_{i=0}^{k+h-1} \frac{1}{2^i} - \sum_{i=0}^{k-1} \frac{1}{2^i} \\ &= \frac{1 - \frac{1}{2^{k+h}}}{\frac{1}{2}} - \frac{1 - \frac{1}{2^k}}{\frac{1}{2}} = \frac{1}{2^{k-1}} - \frac{1}{2^{k+h-1}} < \frac{1}{2^{k-1}}, \end{aligned}$$

which by

$$\lim_{k\to\infty} \|v_{jt}^k - v_{jt}^{k+h}\| = 0,$$

implies that $(v_{jt}^k)_{k\in\mathbb{N}}$ are Cauchy sequences in $\overline{X}$ for all $j \in \{1, ..., q\}$, $t \in \mathbb{Z}$. Because $\overline{X}$ is closed, there exists limit v_{jt} for all $j \in \{1, \dots, q\}$, $t \in \mathbb{Z}$. In order to show that the v_{jt}'s are truly aggregates it is left to show that the sequence of filters $(G_k(L))_{k\in\mathbb{N}}$ is a dynamic averaging sequence. Due to $F_k(\theta) \in K_M$, the biggest eigenvalue of

$$F_k(\theta)\Lambda_{s_k}^{-1/2}(\theta)P_{s_k}(\theta)\overline{P_{s_k}(\theta)}'\Lambda_{s_k}^{-1/2}(\theta)\overline{F_k(\theta)}' = F_k(\theta)\Lambda_{s_k}^{-1}(\theta)\overline{F_k(\theta)}'$$

is

$$\frac{1}{\lambda_{s_k q}^X(\theta)}.$$

Hence, it holds that

$$\|G_k^\circ\|^2 = \|F_k\Lambda_{s_k}^{-1/2}P_{s_k}\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\lambda_{s_k q}^X(\theta)}d\theta \xrightarrow[k\to\infty]{} 0,$$

since $1/\lambda_{s_k q}^X(\theta)$ converges to zero a.e. in $[-\pi, \pi]$ and is bounded by 1 by the assumption made at the beginning of the proof and thus allows for the application of Theorem 2.3, which yields the above convergence result. The spectral density of the limit process $\{v_t\}$ is

$$\Sigma^v(\theta) = \begin{cases} I_q, & \theta \in M, \\ 0, & \theta \notin M. \end{cases}$$

by application of Lemma 7.4.

Ad iii): We are now extending the above result of finding a white noise vector process with unit spectral density on subset M of $[-\pi, \pi]$ to the whole of $[-\pi, \pi]$. Let $\Pi \subset [-\pi, \pi]$ with $\mathcal{L}([-\pi, \pi] \setminus \Pi) = 0$ be such that for some $V \in \mathbb{R}$, $\lambda_{n,q+1}^X(\theta) \le V$ for $n \in \mathbb{N}$ and $\theta \in \Pi$, and such that $\lambda_q^X(\theta) = \infty$ for $\theta \in \Pi$. Furthermore define

$$l_{1,1} = \min\{l \in \mathbb{N} : \mathcal{L}(\{\theta \in \Pi : \lambda_{lq}^X(\theta) > 1\}) \ge \pi\}$$

and

$$M_{1,1} = \{\theta \in \Pi : \lambda_{l_{1,1}q}^X(\theta) > 1\}.$$

Successively, $l_{r,1}$ and $M_{r,1}$ for $r \in \mathbb{N}$ are defined as

$$l_{r,1} = \min\{l \in \mathbb{N} : \mathcal{L}(\{\theta \in M_{r-1,1} : \lambda^X_{lq}(\theta) > r\}) \geq \pi\}$$

and

$$M_{r,1} = \{\theta \in M_{r-1,1} : \lambda^X_{l_{r1}q}(\theta) > r\}.$$

Because $\lambda^X_q(\theta) = \infty$ for $\theta \in \Pi$, the set

$$\mathcal{M}_1 = \bigcap_{r=1}^{\infty} M_{r,1}$$

satisfies $\mathcal{L}(\mathcal{M}_1) \geq \pi$. With notation $\mathcal{N}_{k-1} = \bigcup_{i=1}^{k-1} \mathcal{M}_i$, define $\mathcal{M}_k$ successively as

$$\mathcal{M}_k = \bigcap_{r=1}^{\infty} M_{r,k},$$

where

$$M_{1,k} = \{\theta \in \Pi \setminus \mathcal{N}_{k-1} : \lambda^X_{l_{1,k}q}(\theta) > 1\}$$

with

$$l_{1,k} = \min\{l \in \mathbb{N} : \mathcal{L}(\{\theta \in \Pi \setminus \mathcal{N}_{k-1} : \lambda^X_{lq}(\theta) > 1\}) \geq \mathcal{L}(\Pi \setminus \mathcal{N}_{k-1})/2\}$$

and

$$M_{r,k} = \{\theta \in M_{r,k-1} : \lambda^X_{l_{r,k}q}(\theta) > r\}$$

with

$$l_{r,k} = \min\{l \in \mathbb{N} : \mathcal{L}(\{\theta \in M_{r,k-1} : \lambda^X_{lq}(\theta) > 1\}) \geq \mathcal{L}(\Pi \setminus \mathcal{N}_{k-1})/2\}.$$

By definition the sets are disjoint and the infinite union of the $\mathcal{M}_k$'s yields Π. Apply now step ii) to the set $\mathcal{M}_k$ with suitable sequence $(\alpha_n)_{n\in\mathbb{N}}$ in order to receive a q-dimensional vector process $\{w^k_t\} = \{(w^k_{1t}\ w^k_{2t}\ \ldots\ w^k_{qt})'\}$ whose components are aggregates and whose spectral density is the identity on $\mathcal{M}_k$ and zero outside of $\mathcal{M}_k$. Define the new process $\{z_t\}$ as

$$z_t = \sum_{k=1}^{\infty} w^k_t.$$

Since, as mentioned above, the $\mathcal{M}_k$'s are disjoint the spectral density is the identity a.e. in $[-\pi, \pi]$ and 0 else, which concludes that $\{z_t\}$ is a q-dimensional orthogonal white noise process with unit spectral density a.e. in $[-\pi, \pi]$.

Ad iv): In this step we will prove that $\overline{\text{span}}(\{z_{it} : i \in \{1, \dots, q\}, t \in \mathbb{Z}\})$ is $\mathcal{G}(X)$. For this purpose consider arbitrary aggregate y_t. This aggregate can be decomposed as

$$y_t = \text{Proj}_{\overline{\text{span}}(\{z_{it}:i\in\{1,\dots,q\},t\in\mathbb{Z}\})}(y_t) + \epsilon_t,$$

where $\text{Proj}_{\overline{\text{span}}(\{z_{it}:i\in\{1,\dots,q\},t\in\mathbb{Z}\})}(y_t)$ and ϵ_t are orthogonal at any lead and lag. The claim follows if it can be shown that ϵ_t is zero. Therefore consider the $(q+1)$-dimensional vector process $\{v_t\} := \{(z_t' \;\; \epsilon_t)'\}$ with spectral density

$$\Sigma^v(\theta) = \begin{bmatrix} I_q & 0 \\ 0 & \Sigma^\epsilon(\theta) \end{bmatrix},$$

where $\Sigma^\epsilon(\theta)$ is the spectral density of $\{\epsilon_t\}$. Since $z_{jt} \in \mathcal{G}(X)$ and $\mathcal{G}(X)$ is closed, also $\text{Proj}_{\overline{\text{span}}(\{z_{it}:i\in\{1,\dots,q\},t\in\mathbb{Z}\})}(y_t) \in \mathcal{G}(X)$ and hence also $\epsilon_t \in \mathcal{G}(X)$. This means that there exist dynamic averaging sequences $(a_{nj}(L))_{n\in\mathbb{N}}$ for $j \in \{1, \dots, q+1\}$ such that

$$\lim_{n\to\infty} a_{nj}(L) X_t^{s_n} = z_{jt}, \qquad j \in \{1, ..., q\},$$

in mean square and

$$\lim_{n\to\infty} a_{n,q+1}(L) X_t^{s_n} = \epsilon_t,$$

in mean square, where we assume w.l.o.g. that $(s_n)_{n\in\mathbb{N}}$ is the same sequence for all DAS (some $a_{nj}(L)$ can be augmented by zeros). By the definition of dynamic averaging sequences,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} |a_{nj}^\circ(\theta)|^2 = 0, \qquad \forall\, j \in \{1, ..., q+1\},$$

which implies by Theorem 2.4 that there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} a_{n_k j}^\circ(\theta) = 0$, a.e. in $[-\pi, \pi]$. Thus w.l.o.g. we can assume that

$$\lim_{n\to\infty} a_{nj}^\circ(\theta) = 0, \qquad \text{a.e. in } [-\pi, \pi].$$

Now consider the following vector process $\{a_t^n\}$ with

$$a_t^n := (a_{n1}(L)X_t^{s_n} \;\; a_{n2}(L)X_t^{s_n} \;\; \cdots \;\; a_{n,q+1}(L)X_t^{s_n})$$

and spectral density $\Sigma_n^a(\theta)$ and $\lim_{n\to\infty} a_t^n = v_t$ in mean square. Since all processes $\{a_{nj}(L)X_t^{s_n}\}$ are costationary for $j \in \{1,\dots,q+1\}$, there exists by application of Lemma 7.4 a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \Sigma_{n_k}^a(\theta) = \Sigma^v(\theta)$ a.e. in $[-\pi,\pi]$. Thus again w.l.o.g. we can assume that

$$\lim_{n\to\infty} \Sigma_n^a(\theta) = \Sigma^v(\theta), \qquad \text{a.e. in } [-\pi,\pi].$$

We now try to further decompose $\Sigma_n^a(\theta)$. For this reason we project $a_{nj}^\circ(\theta)$ for $j \in \{1,\dots,q+1\}$ onto the space spanned by the first q dynamic principal components of $\{X_t^{s_n}\}$. Let $f_{nj}(\theta) = a_{nj}^\circ(\theta)\overline{P_{s_n}(\theta)}'$ denote the coefficients of this projection and $r_{nj}(\theta)$ the residual. Then,

$$a_{nj}^\circ(\theta) = f_{nj}(\theta)P_{s_n}(\theta) + r_{nj}(\theta).$$

This implies further the orthogonal decomposition

$$a_{nj}(L)X_t^{s_n} = \underline{f}_{nj}(L)\underline{P}_{s_n}(L)X_t^{s_n} + \underline{r}_{nj}(L)X_t^{s_n}.$$

By defining the processes $\{f_t^n\}$ and $\{r_t^n\}$ as

$$f_t^n = (\underline{f}_{n1}(L)\underline{P}_{s_n}(L)X_t^{s_n} \;\; \cdots \;\; \underline{f}_{n,q+1}(L)\underline{P}_{s_n}(L)X_t^{s_n}),$$

and

$$r_t^n = (\underline{r}_{n1}(L)X_t^{s_n} \;\; \cdots \;\; \underline{r}_{n,q+1}(L)X_t^{s_n}),$$

with spectral densities $\Sigma_n^f(\theta)$ and $\Sigma_n^r(\theta)$, we obtain the following decomposition of $\Sigma_n^a(\theta)$:

$$\Sigma_n^a(\theta) = \Sigma_n^f(\theta) + \Sigma_n^r(\theta).$$

Notice that since $P_{s_n}(\theta)$ has rank q and $\Sigma_n^f(\theta)$ is a $((q+1)\times(q+1))$-matrix, $\Sigma_n^f(\theta)$ is singular. Since $|a_{nj}^\circ(\theta)| = |f_{nj}(\theta)| + |r_{nj}(\theta)|$ and $\lim_{n\to\infty} a_{nj}^\circ(\theta) = 0$ a.e. in $[-\pi,\pi]$ for $j \in \{1,\dots,q+1\}$, also $\lim_{n\to\infty} r_{nj}(\theta) = 0$ a.e. in $[-\pi,\pi]$ for

$j \in \{1, \ldots, q+1\}$. Also, the residual $r_{nj}(\theta)$ for $j \in \{1, \ldots, q+1\}$ is orthgonal to the first q dynamic eigenvectors $p_{s_n i}(\theta)$, $i \in \{1, \ldots, q\}$, which implies by the decomposition of $\Sigma^X_{s_n}(\theta)$,

$$r_{nj}(\theta)\Sigma^X_{s_n}(\theta)\overline{r_{nj}(\theta)}' \leq \lambda^X_{s_n,q+1}(\theta)|r_{nj}(\theta)|^2 \xrightarrow[n\to\infty]{} 0, \tag{4.3}$$

where the limit result follows from the essential boundedness of $\lambda^X_{q+1}(\theta)$ and the above discussed convergence result for $r_{nj}(\theta)$. The convergence result (4.3) implies that $\Sigma^r_n(\theta)$ converges to zero a.e. in $[-\pi, \pi]$. These observations further imply that

$$\lim_{n\to\infty} \det \Sigma^a_n(\theta) = \lim_{n\to\infty} \det(\Sigma^f_n(\theta) + \Sigma^r_n(\theta)) = 0, \qquad \text{a.e. in } [-\pi, \pi].$$

Therefore we have that

$$\Sigma^\epsilon(\theta) = \det \Sigma^v(\theta) = 0, \qquad \text{a.e. in } [-\pi, \pi],$$

which concludes that $\epsilon_t = 0$.

Ad v): In steps i)-iv) we have proven that there exists a q-dimensional orthogonal vector process $\{z_t\}$ with unit spectral density, which spans $\mathcal{G}(X)$. This means that there exists $c_i^\circ(\theta) \in L_2^q([-\pi, \pi], \mathbb{C})$ for any $i \in \mathbb{N}$ such that

$$\mathrm{Proj}_{\mathcal{G}(X)}(x_{it}) = c_i(L)z_t =: \gamma_{it}.$$

Now the components of X^n_t can be written as

$$x_{it} = c_i(L)z_t + \beta_{it}. \tag{4.4}$$

To conclude the proof one needs to show that the infinite dimensional vector process $\{\beta_t\} = \{(\beta_{1t} \ \ \beta_{2t} \ \ \ldots\ ) : t \in \mathbb{Z}\}$ is idiosyncratic. For that reason define

$$\gamma_{it,n} := \underline{\pi}_{ni}(L)\underline{\Lambda}_n^{1/2}(L)\psi^n_t,$$

where $\underline{\pi}_{ni}(L)$ denotes the i-th row of $\overline{\underline{P}_n(L)}'$ and

$$\beta_{it,n} = x_{it} - \gamma_{it,n}.$$

We will now proceed by showing that

$$\lim_{n\to\infty} \beta_{it,n} = \beta_{it}$$

in mean square for any $i \in \mathbb{N}$ and any $t \in \mathbb{Z}$. As $(\{\psi_t^n\})_{n\in\mathbb{N}}$ generates a Cauchy sequence of spaces by Lemma 3.4 and costationarity between $\{x_{it}\}$ and $\{\psi_t^n\}$ is fulfilled, there exists by Lemma 3.5 $\{\gamma_{it}^*\}$ such that

$$\lim_{n\to\infty} \underline{\pi}_{ni}(L)\underline{\Lambda}_n^{1/2}(L)\psi_t^n = \gamma_{it}^*$$

in mean square and thus also

$$\lim_{n\to\infty} \beta_{it,n} = \beta_{it}^* = x_{it} - \gamma_{it}^*$$

in mean square for any $i \in \mathbb{N}$ and any $t \in \mathbb{Z}$. Since by construction $\beta_{it,n}$ and $\gamma_{it,n}$ are orthogonal at any lead and lag, also β_{it}^* and γ_{it}^* are orthogonal at any lead and lag by Lemma 7.4. If we can show that γ_{it}^* is an aggregate, then $\gamma_{it}^* = \gamma_{it} := \text{Proj}_{\mathcal{G}(X)}(x_{it})$ and hence also $\beta_{it}^* = \beta_{it}$. By definition of aggregates and the fact that

$$\gamma_{it,n} = \underline{\pi}_{ni}(L)\underline{\Lambda}_n^{1/2}(L)\psi_t^n = \underline{\pi}_{ni}(L)\underline{P}_n(L)X_t^n,$$

we have to show that $\underline{\pi}_{ni}(L)\underline{P}_n(L)$ is a dynamic averaging sequence, i.e. $\lim_{n\to\infty} \|\pi_{ni}P_n\| = \lim_{n\to\infty} \|\pi_{ni}\| = 0$. Firstly, notice that the spectral density of $\gamma_{it,n}$, which we denote by $\Sigma_n^{\gamma_i}(\theta)$, is bounded by $(\Sigma_n^X(\theta))_{ii}$ by orthogonality of $\{\gamma_{it,n}\}$ and $\{\beta_{it,n}\}$. This $(\Sigma_n^X(\theta))_{ii}$ does not depend on n. Secondly, notice that for spectral density $\Sigma_n^{\gamma_i}(\theta)$ it holds that,

$$\Sigma_n^{\gamma_i}(\theta) = \pi_{ni}(\theta)\Lambda_n^{1/2}(\theta)\Lambda_n^{1/2}(\theta)\overline{\pi_{ni}(\theta)}' = \pi_{ni}(\theta)\Lambda_n(\theta)\overline{\pi_{ni}(\theta)}' \geq \lambda_{nq}^X(\theta)|\pi_{ni}(\theta)|^2.$$

Taking both results together we have

$$\lambda_{nq}^X(\theta)|\pi_{ni}(\theta)|^2 \leq \Sigma_n^{\gamma_i}(\theta) \leq (\Sigma_n^X(\theta))_{ii},$$

and since $\lim_{n\to\infty} \lambda_{nq}^X(\theta) = \infty$, a.e. in $[-\pi, \pi]$, also

$$|\pi_{ni}(\theta)|^2 \leq \frac{(\Sigma_n^X(\theta))_{ii}}{\lambda_{nq}^X(\theta)} \xrightarrow[n\to\infty]{} 0, \qquad \text{a.e. in } [-\pi, \pi].$$

Furthermore, by the assumption that $\lambda^X_{nq}(\theta) \geq 1$ for any $n \in \mathbb{N}$, $|\pi_{ni}(\theta)|^2$ is bounded, which allows for application of the dominated convergence theorem, (Theorem 2.3), which yields

$$\|\pi_{ni}\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\pi_{ni}(\theta)|^2 d\theta \xrightarrow[n\to\infty]{} 0.$$

Hence γ^*_{it} is an aggregate and $\beta_{it,n}$ converges to β_{it} in mean square.

Ad vi): To conclude the proof we will show that $\{\beta_t\}$ is idiosyncratic by using Theorem 3.1. Let $m \in \mathbb{N}$ be fixed. Denote by $\Sigma^\beta_m(\theta)$ the spectral density of process $\{\beta^m_t\}$,

$$\beta^m_t = (\beta_{1t}\ \ \beta_{2t}\ \ \cdots\ \ \beta_{mt})'.$$

It will be shown that the first eigenvalue of $\Sigma^\beta_m(\theta)$, denoted by $\lambda^\beta_{m1}(\theta)$, is bounded from above by $\sup_{n\in\mathbb{N}} \lambda^X_{n,q+1}(\theta) = \lambda^X_{q+1}(\theta)$ for any θ in $[-\pi, \pi]$. For $n \geq m$ let $\{\beta^{m,n}_t\}$ be the m-dimensional vector process

$$\beta^{m,n}_t = (\beta_{1t,n}\ \ \beta_{2t,n}\ \ \cdots\ \ \beta_{mt,n})',$$

with spectral density $\Sigma^{\beta_n}_m(\theta)$. Denote by $\lambda^{\beta_n}_{m1}(\theta)$ the first eigenvalue of $\Sigma^{\beta_n}_m(\theta)$. From step v) we know that $\beta_{it,n}$ converges to β_{it} in mean square for all $i \in \{1, \ldots, m\}, t \in \mathbb{Z}$ and by Lemma 7.4 there exists a subsequence of integers $(n_k)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} \Sigma^{\beta_{n_k}}_m(\theta) = \Sigma^\beta_m(\theta),$$

for almost all θ in $[-\pi, \pi]$. Hence w.l.o.g. we can assume that

$$\lim_{n\to\infty} \Sigma^{\beta_n}_m(\theta) = \Sigma^\beta_m(\theta),$$

for almost all θ in $[-\pi, \pi]$. By continuity of the eigenvalues we have that

$$\lim_{n\to\infty} \lambda^{\beta_n}_{m1}(\theta) = \lambda^\beta_{m1}(\theta),$$

a.e. in $[-\pi, \pi]$. Notice that since

$$X^n_t = \gamma^{n,n}_t + \beta^{n,n}_t,$$

where process $\{\gamma_t^{n,n}\}$ with $\gamma_t^{n,n} = (\gamma_{1t,n}, \dots, \gamma_{nt,n})'$ has spectral density

$$\Sigma_n^{\gamma_n}(\theta) = \overline{P_n(\theta)}'\Lambda_n(\theta)P_n(\theta),$$

the spectral density $\Sigma_n^X(\theta)$ decomposes into

$$\Sigma_n^X(\theta) = \overline{P_n(\theta)}'\Lambda_n(\theta)P_n(\theta) + \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta) = \Sigma_n^{\gamma_n}(\theta) + \Sigma_n^{\beta_n}(\theta).$$

Hence we have

$$\Sigma_n^{\beta_n}(\theta) = \overline{Q_n(\theta)}'\Phi_n(\theta)Q_n(\theta),$$

which means that the first eigenvalue $\lambda_{n1}^{\beta_n}(\theta)$ of $\Sigma_n^{\beta_n}(\theta)$ is $\lambda_{n,q+1}^X(\theta)$. By monotonicity of the eigenvalues we have for $n \geq m$ and θ in $[-\pi, \pi]$ that

$$\lambda_{m1}^{\beta_n}(\theta) \leq \lambda_{n1}^{\beta_n}(\theta) = \lambda_{n,q+1}^X(\theta).$$

Taking limits on both sides yields

$$\lambda_{m1}^{\beta}(\theta) \leq \lambda_{q+1}^X(\theta)$$

for arbitrary m, so that

$$\sup_{m\in\mathbb{N}} \lambda_{m1}^{\beta}(\theta) =: \lambda_1^{\beta}(\theta) \leq \lambda_{q+1}^X(\theta)$$

holds. Since $\lambda_{q+1}^X(\theta)$ is essentially bounded, so is $\lambda_1^\beta(\theta)$, implying the idiosyncracy of $\{\beta_t\}$. This concludes the whole proof. □

Remark 4.1. In the proof we assumed w.l.o.g. for $n \in \mathbb{N}$, $j \leq n$ and $\theta \in [-\pi, \pi]$ that $\lambda_{nj}^X(\theta) \geq 1$. Now, we will argue why that assumption is valid. As in Remark 3.2, we can define a new process$\{\hat{X}_t\}$ via $\hat{x}_{it} = x_{it} + \zeta_{it}$, where $\{\zeta_t\}$ is orthogonal white noise with unit spectral density such that $\zeta_{it_1} \perp x_{jt_2}$ for any $i, j \in \mathbb{N}$ and any $t_1, t_2 \in \mathbb{Z}$. Notice that $\Sigma_n^{\hat{X}}(\theta) = \Sigma_n^X(\theta) + I_n$ and hence also $\lambda_{ni}^{\hat{X}}(\theta) = \lambda_{ni}^X(\theta) + 1$. It is thus easily seen that if (I) + (II) hold for $\{X_t\}$, then they also hold for $\{\hat{X}_t\}$ and vice versa. By Theorem 4.1, $\{\hat{X}_t\}$ has a representation as $\hat{x}_{it} = \hat{\chi}_{it} + \hat{\xi}_{it}$, where $\hat{\chi}_{it} = \text{Proj}_{\mathcal{G}(\hat{X})}(\hat{x}_{it})$. The definition of ζ_t implies that $\hat{\chi}_{it} = \text{Proj}_{\mathcal{G}(X)}(x_{it})$, and hence also $x_{it} = \text{Proj}_{\mathcal{G}(X)}(x_{it}) + (\hat{\xi}_{it} - \zeta_{it})$. Since by definition ζ_{it} is orthogonal to $\overline{X}$, $\hat{\xi}_t - \zeta_t$ is idiosyncratic and hence X_t follows a generalized dynamic factor model with q factors as well. This is based on [4, p.11f].

5 Consistent Estimator for the Common Component

As we have seen in Chapter 3, specifically in Theorem 3.2, the i-th common component $\{\chi_{it}\}$ can be asymptotically recovered from the process defined by $\underline{\pi}_{ni}(L)\underline{P}_n(L)X_t^n$, where $\underline{\pi}_{ni}(L)$ denotes the i-th row of $\overline{\underline{P}_n(L)}'$, for an infinite dimensional stationary time series $\{X_t\}$ that follows a generalized dynamic factor model with q factors. As is often the problem in practice, the population spectral densities $\Sigma_n^X(\theta)$ for $\{X_t^n\}$ resulting from $\{X_t\}$ are not known and have to be estimated. We will propose in the following an estimator of the common component based on observations $\{\mathsf{x}_1^n, \ldots, \mathsf{x}_T^n\}$, denoted by $\hat{\chi}_{it}$, which is consistent. Firstly recall the notion of consistency.

Definition 5.1. A sequence of estimators $(\hat{W}_T)_{T\in\mathbb{N}}$ based on T observations $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$ is said to be *consistent* for the true parameter W if

$$\hat{W}_T \xrightarrow{P} W,$$

as $T \to \infty$, i.e., for arbitrary $\epsilon > 0$:

$$\lim_{T\to\infty} P(|\hat{W}_T - W| > \epsilon) = 0.$$

See [10, p.54].

5.1 Consistent estimator for the spectral density

For the theory of the following section we assume that the T observations $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$ are realizations of an n-dimensional stationary vector process $\{X_t\}$. For now we fix n, so subscript n is dropped in this section. First of all we will propose an estimator of the spectral density $\Sigma^X(\theta)$ which is solely based on the coefficients of the discrete Fourier transform. This estimator will be called the periodogram of $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$.
The vectors

$$e_k = \tfrac{1}{\sqrt{T}}(e^{i\theta_k}, \ldots, e^{iT\theta_k})',$$

with fundamental frequencies $\theta_k = \frac{2\pi k}{n}$, where $k \in \mathcal{F}_T := \{-\lfloor \frac{T-1}{2} \rfloor, \ldots, \lfloor \frac{T}{2} \rfloor\}$ form an orthonormal basis of $\mathbb{C}^n$. More information and proofs can be found in [6, p.331f]. Hence x_t can be written as

$$\mathsf{x}_t = \sum_{k \in \mathcal{F}_T} \langle \mathsf{x}_t, e_k \rangle e_k.$$

By taking all observations together as $\mathsf{x} = (\mathsf{x}_1, \ldots, \mathsf{x}_T)'$ and denoting by $\tilde{\mathsf{x}}_i$ the i-th column of x for $i \in \{1, \ldots, n\}$, we have that

$$\begin{aligned} \mathsf{x} = (\tilde{\mathsf{x}}_1, \ldots, \tilde{\mathsf{x}}_n) &= \sum_{k \in \mathcal{F}_T} (\langle \tilde{\mathsf{x}}_1, e_k \rangle e_k, \ldots, \langle \tilde{\mathsf{x}}_n, e_k \rangle e_k) \\ &= \sum_{k \in \mathcal{F}_T} e_k (\langle \tilde{\mathsf{x}}_1, e_k \rangle, \ldots, \langle \tilde{\mathsf{x}}_n, e_k \rangle). \end{aligned}$$

Hence for the discrete Fourier transform consisting of the coefficient $(\langle \tilde{\mathsf{x}}_1, e_k \rangle, \ldots, \langle \tilde{\mathsf{x}}_n, e_k \rangle)'$ corrsponding to e_k, denoted by $\mathcal{J}(\theta_k)$, further holds

$$\begin{aligned} (\langle \tilde{\mathsf{x}}_1, e_k \rangle, \ldots, \langle \tilde{\mathsf{x}}_n, e_k \rangle)' &= (\tfrac{1}{\sqrt{T}} \sum_{t=1}^{T} (\tilde{\mathsf{x}}_1)_t e^{-it\theta_k}, \ldots, \tfrac{1}{\sqrt{T}} \sum_{t=1}^{T} (\tilde{\mathsf{x}}_n)_t e^{-it\theta_k})' \\ &= \tfrac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathsf{x}_t e^{-it\theta_k} = \mathcal{J}(\theta_k). \end{aligned}$$

See [6, p.443].

This prompts the definition of the periodogram, which acts as an estimator for the spectral density of the underlying stationary vector process $\{X_t\}$ from which $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$ is assumed to be sampled.

Definition 5.2. Assume that observations $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$ are the realizations of an n-dimensional stationary vector process. Let $g_T(\theta)$ be the fundamental frequency θ_k closest to $\theta \in [0, \pi]$ (if there are two, $g_T(\theta)$ is the smaller one). Then we define the *periodogram* at frequency $\theta \in [0, \pi]$ by

$$I_T(\theta) = \mathcal{J}(g_T(\theta))\overline{\mathcal{J}(g_T(\theta))}' = \mathcal{J}(\theta_k)\overline{\mathcal{J}(\theta_k)}',$$

and by

$$I_T(\theta) = I_T(-\theta),$$

at $\theta \in [-\pi, 0)$. See also [6, p.443].

Remark. Similar to the relationship between population spectral density $\Sigma^X(\theta)$ and population covariance function $\Gamma^X(h)$ for an n-dimensional stationary vector process $\{X_t\}$ via

$$\Sigma^X(\theta) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-i\theta h} \Gamma^X(h),$$

there exists a relationship between the periodogram of the T observations $\{\mathsf{x}_1, \ldots \mathsf{x}_T\}$ of the underlying n-dimensional stationary process $\{X_t\}$ and the estimated covariance function $\hat{\Gamma}^X(h) = \frac{1}{T}\sum_{t=1}^{T-h}(\mathsf{x}_{t+h} - \mathsf{x}_{\text{mean}})(\overline{\mathsf{x}_t - \mathsf{x}_{\text{mean}}})'$ for $h \geq 0$ and $\hat{\Gamma}^X(-h) = \overline{\hat{\Gamma}^X(h)}'$ via

$$I(\theta_k) = \begin{cases} \sum_{|h|<T} e^{-i\theta_k h}\hat{\Gamma}^X(h) & \theta_k \neq 0 \\ T\mathsf{x}_{\text{mean}}\overline{\mathsf{x}_{\text{mean}}}' & \theta_k = 0 \end{cases}$$

for any fundamental frequency θ_k, where $\mathsf{x}_{\text{mean}} = \frac{1}{T}\sum_{t=1}^{T}\mathsf{x}_t$. Further information can be found in [6, p.443f].

A satisfactory estimator for the spectral density is required to be consistent. As the following theorem will show, the periodogram is not a consistent estimator.

Prior to that we introduce the notion of the (asymptotic) complex multivariate normal distribution.

Definition 5.3. Let $Y = \mathrm{Re}(Y) + i\,\mathrm{Im}(Y)$ be an n-dimensional complex valued random vector with mean $\mu = \mu_1 + i\mu_2$ and covariance matrix $\Gamma = \Gamma_1 + i\Gamma_2$ satisfying $\overline{\Gamma}' = \Gamma$ and $\Gamma \geq 0$. The random vector Y is said to follow a *complex normal distribution* with mean μ and covariance Γ, short $\mathcal{N}_c(\mu, \Gamma)$, if

$$\begin{bmatrix} \mathrm{Re}(Y) \\ \mathrm{Im}(Y) \end{bmatrix} \sim \mathcal{N}\Big(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} \Gamma_1 & -\Gamma_2 \\ \Gamma_2 & \Gamma_1 \end{bmatrix} \Big),$$

where $\mathcal{N}$ denotes the real-valued normal distribution. See [6, p.444].

Theorem 5.1. Assume the n-dimensional stationary vector process $\{X_t\}$ with spectral density $\Sigma^X(\theta)$ can be expressed as

$$X_t = \sum_{k=-\infty}^{\infty} C_k Z_{t-k},$$

where $\{Z_t\}$ is an n-dimensional white noise vector process with mean 0 and covariance function $\Gamma^Z(h)$ with $\Gamma^Z(0) > 0$ and $\Gamma^Z(h) = 0$ for $h \neq 0$ and where coefficients of the matrices $(C_k)_{k \in \mathbb{Z}} \subset \mathbb{R}^{n \times n}$ satisfy $\sum_{k=-\infty}^{\infty} |(C_k)_{ij}||k|^{1/2} < \infty$ for all $i, j \in \{1, \ldots, n\}$. Furthermore, suppose that $E((Z_t)_i^4) < \infty$ for $i \in \{1, \ldots, n\}$. Let $I_T(\theta)$ be the periodogram based on the T observations $\{\mathsf{x}_1, \ldots, \mathsf{x}_T\}$ for $\theta \in [-\pi, \pi]$. Then the following holds:

(i) For arbitrary frequencies $0 < \omega_1 < \ldots < \omega_m < \pi$ the corresponding matrices $I_T(\omega_1), \ldots, I_T(\omega_m)$ converge jointly in distribution to independent random matrices where the l-th matrix is distributed as $W_l \overline{W_l}'$ where $W_l \sim \mathcal{N}_c(0, 2\pi\Sigma^X(\omega_l))$.

(ii) For fundamental frequencies $\theta_k = \frac{2\pi k}{T}$ and $\theta_l = \frac{2\pi l}{T}$ where $l, k \in \{-\lfloor \frac{T-1}{2} \rfloor, \ldots, \lfloor \frac{T}{2} \rfloor\}$ and $\theta_k \neq \theta_l$, we have

$$\mathrm{Cov}((I_T(\theta_k))_{pq}, (I_T(\theta_l))_{rs}) \xrightarrow[T \to \infty]{} 0, \quad \text{for } p, q, r, s \in \{1, \ldots, n\}.$$

(iii) For fundamental frequency $\theta_k = \frac{2\pi k}{T}$ with $k \in \{-\lfloor \frac{T-1}{2} \rfloor, \dots, \lfloor \frac{T}{2} \rfloor\}$ we have that

$$|\lim_{T\to\infty} \mathrm{Cov}((I_T(\theta_k))_{pq}, (I_T(\theta_k))_{rs})| \leq f, \quad \text{for } p, q, r, s \in \{1, \dots, n\}$$

and for some $f \in \mathbb{R}$. See [6, p.446].

Proof. The proof for the univariate case [6, p.347-350] can be extended to the multivariate case, see [11, p.248f] and [6, p.444-446]. □

As part i) above shows, the limit of $I_T(\theta)$ for some $\theta \in (0, \pi)$ is still random and therefore $(I_T(\theta))_{T\in\mathbb{N}}$ cannot be a consistent estimator for the spectral density. However, since by part ii) the periodograms of different frequencies are asymptotically uncorrelated and by part iii) the covariance matrix is bounded, there exists the prospect of constructing a consistent estimator for the spectral density $\Sigma^X(\theta)$ by averaging the periodogram locally around a frequency. See [6, p.350].

Definition 5.4. Let $I_T(\theta)$ for $\theta \in [-\pi, \pi]$ be the periodogram associated with the T observations of underlying n-dimensional stationary vector process $\{X_t\}$ which has the structure $X_t = \sum_{k=-\infty}^{\infty} C_k Z_{t-k}$ where $\{Z_t\}$ is an n-dimensional white noise with mean zero and covariance matrix Γ and where matrices $(C_k)_{k\in\mathbb{Z}}$ satisfy $\sum_{k=-\infty}^{\infty} |(C_k)_{ij}||k|^{1/2} < \infty$ for all $i, j \in \{1, \dots, n\}$. Let $(W_T(k))_{T\in\mathbb{Z}}$ be a sequence of real diagonal weight matrices of dimension $(n \times n)$ such that

- $(W_T(k)) \geq 0$ componentwise for any $k \in \mathbb{Z}$,
- $W_T(-k) = W_T(k)$,
- $W_T(k) = 0$, if $|k| > m_T$ for some $m_T \in \mathbb{N}$ fulfilling $\lim_{T\to\infty} m_T = \infty$ and $\lim_{T\to\infty} m_T/T = 0$,
- $\sum_{|k|\leq m_T} W_T(k) = I_n$,

- $\lim_{T\to\infty} \sum_{|k|\le m_T} W_T(k)^2 = 0$,

then the estimator of the form

$$\hat{\Sigma}_T^X(\theta) = \frac{1}{2\pi} \sum_{|k|\le m_T} W_T(k) I_T(g_T(\theta) + \theta_k),$$

where θ_k is a fundamental frequency is called a *periodogram smoothing estimator* for the spectral density $\Sigma^X(\theta)$. In case of $g_T(\theta)+\theta_k \notin [-\pi, \pi]$, $I_T(g_T(\theta)+\theta_k)$ is the value of $I_T(\theta)$ for $\theta \in [-\pi, \pi]$ given that $I_T(\theta)$ has period 2π. The definition is based on [6, p.350f,446f].

This periodogram smoothing estimator is compared to the traditional periodogram consistent, which is a consequence of the following theorem.

Theorem 5.2. Let $\{X_t\}$ be an n-dimensional stationary vector process with $X_t = \sum_{k=-\infty}^{\infty} C_k Z_{t-k}$, where $\{Z_t\}$ is an n-dimensional white noise vector process with mean 0 and covariance matrix Γ and where matrices $(C_k)_{k\in\mathbb{Z}}$ satisfy $\sum_{k=-\infty}^{\infty} |(C_k)_{ij}||k|^{1/2} < \infty$ for all $i, j \in \{1, \dots, n\}$. Furthermore, let $\hat{\Sigma}_T^X(\theta)$ be a periodogram smoothing estimator of the spectral density $\Sigma^X(\theta)$ based on the T observations $\{\mathsf{x}_1, \dots, \mathsf{x}_T\}$. Then the following holds:

$$\lim_{T\to\infty} \mathrm{E}(\hat{\Sigma}_T^X(\theta)) = \Sigma^X(\theta).$$

Proof. The proof for the univariate case can be found in [6, p.351ff] and can be extended to the multivariate case, see [6, 447f]. □

We now take a look at the asymptotic covariance of the estimator $\hat{\Sigma}_T^X(\theta)$ to get following important result.

Lemma 5.1. Let $\{X_t\}$ be an n-dimensional stationary vector process with $X_t = \sum_{k=-\infty}^{\infty} C_k Z_{t-k}$, where $\{Z_t\}$ is an n-dimensional white noise vector process with mean 0 and covariance matrix Γ and where matrices $(C_k)_{k\in\mathbb{Z}}$ satisfy $\sum_{k=-\infty}^{\infty} |(C_k)_{ij}||k|^{1/2} < \infty$ for all $i, j \in \{1, \dots, n\}$. Furthermore, let $\hat{\Sigma}_T^X(\theta)$ be a periodogram smoothing estimator of the spectral density $\Sigma^X(\theta)$ based on the T observations $\{\mathsf{x}_1, \dots, \mathsf{x}_T\}$. Then $\hat{\Sigma}_T^X(\theta)$ is a consistent estimator for $\Sigma^X(\theta)$.

Proof. Consider

$$(\hat{\Sigma}_T^X(\theta))_{pq} = \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}(W_T(k))_{pp}(I_T(g_T(\theta)+\theta_k))_{pq},$$

$$(\hat{\Sigma}_T^X(\theta))_{rs} = \frac{1}{2\pi}\sum_{j\in\mathbb{Z}}(W_T(j))_{rr}(I_T(g_T(\theta)+\theta_j))_{rs}.$$

Then we have for $\mathrm{Cov}((\hat{\Sigma}_T^X(\theta))_{pq}, (\hat{\Sigma}_T^X(\theta))_{rs})$ the following:

$$\frac{1}{(2\pi)^2}\sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}}(W_T(k))_{pp}(W_T(j))_{rr}\mathrm{Cov}((I_T(g_T(\theta)+\theta_k))_{pq},(I_T(g_T(\theta)+\theta_j))_{rs}) =$$
$$\frac{1}{(2\pi)^2}\sum_{k\in\mathbb{Z}}\sum_{\substack{j\in\mathbb{Z}\\ j\neq k}}(W_T(k))_{pp}(W_T(j))_{rr}\mathrm{Cov}((I_T(g_T(\theta)+\theta_k))_{pq},(I_T(g_T(\theta)+\theta_j))_{rs}) +$$
$$\frac{1}{(2\pi)^2}\sum_{k\in\mathbb{Z}}(W_T(k))_{pp}W_T(k)_{rr}\mathrm{Cov}((I_T(g_T(\theta)+\theta_k))_{pq},(I_T(g_T(\theta)+\theta_k))_{rs}) \xrightarrow[T\to\infty]{} 0.$$

The covariance of the first summands goes to zero by Theorem 5.1 (ii) and the boundedness of the weights. Note that $g_T(\theta)+\theta_k$ is a fundamental frequency. The second summand goes to zero by the boundedness of the covariances according to Theorem 5.1 (iii) and the fact that $(W_T(k))_{pp}(W_T(k))_{rr}$ converges to zero by the last assumption on the weight sequence $(W_T(k))_{T\in\mathbb{N}}$. Since this holds for all $p,q,r,s\in\{1,\dots,n\}$, the variance of estimator $\hat{\Sigma}_T^X(\theta)$ goes to zero, hence $\hat{\Sigma}_T^X(\theta)$ is a consistent estimator. □

5.2 Proposition of estimator for the common component

We now return to infinite dimensional stationary vector process $\{X_t\}$ with spectral densities $\Sigma_n^X(\theta)$ corresponding to the n-dimensional stationary vector process $\{X_t^n\}$ based on $\{X_t\}$. Assume $\{X_t\}$ follows a generalized dynamic factor model with q factors. Based on the previous section and its results we impose further assumptions, additional to the ones imposed at the beginning of Chapter 3, on the vector process $\{X_t\}$. Assume $\{X_t^n\}$ can be represented as

$$X_t^n = \sum_{k=-\infty}^{\infty} C_k^n Z_{t-k}^n,$$

where $\{Z_t^n\}$ is an n-dimensional white noise vector process with mean 0 and covariance matrix Γ and where coefficients of $(C_k^n)_{k\in\mathbb{Z}} \subset \mathbb{R}^{n\times n}$ satisfy $\sum_{k=-\infty}^{\infty} |(C_k^n)_{ij}||k|^{1/2} < \infty$ for all $i,j \in \{1,\dots,n\}$ and all $n \in \mathbb{N}$. Note that components of C_k^n and Z_{t-k}^n can change dependent on n. Furthermore let $\{\mathsf{x}_1^n,\dots,\mathsf{x}_T^n\}$ be T observations of the underlying n-dimensional stationary vector process $\{X_t^n\}$. Denote by $\hat{\Sigma}_{n,T}^X(\theta)$ a periodogram-smoothing estimator for the spectral density $\Sigma_n^X(\theta)$ based on the T observations $\{\mathsf{x}_1^n,\dots,\mathsf{x}_T^n\}$. As we have seen before, for any $n \in \mathbb{N}$, any $\theta \in [-\pi,\pi]$, any $\epsilon > 0$ and $i,j \in \{1,\dots,n\}$:

$$\lim_{T\to\infty} \mathrm{P}(|(\hat{\Sigma}_{n,T}^X(\theta))_{ij} - (\Sigma_n^X(\theta))_{ij}| > \epsilon) = 0.$$

Note that $\hat{\Sigma}_{n+1,T}^X(\theta)$ is the periodogram smoothing estimator of $\{\mathsf{x}_1^{n+1},\dots,\mathsf{x}_T^{n+1}\}$ where x_i^{n+1} was extended by an $n+1$-th component and where weight matrices $W_T(k)$ are extended in a hierarchical manner. In this setting $\hat{\Sigma}_{n,T}^X(\theta)$ and $\hat{\Sigma}_{n+1,T}^X(\theta)$ coincide in the first n rows and the first n columns.

For fixed n denote by $\lambda_{n,T,i}^X(\theta)$ the i-th largest dynamic eigenvalue of estimator $\hat{\Sigma}_{n,T}^X(\theta)$ and by $p_{n,T,i}^X(\theta)$ the corresponding i-th dynamic eigenvector. Both, dynamic eigenvalues and dynamic eigenvectors are chosen to be continuous functions in θ on $[-\pi,\pi]$, and enable consistent estimation of $\lambda_{ni}^X(\theta)$ and $p_{ni}^X(\theta)$, i.e. for any $\epsilon > 0$:

$$\lim_{T\to\infty} \mathrm{P}(|\lambda_{n,T,i}^X(\theta) - \lambda_{ni}^X(\theta)| > \epsilon) = 0,$$

and

$$\lim_{T\to\infty} \mathrm{P}(|p_{n,T,i}^X(\theta) - p_{ni}^X(\theta)| > \epsilon) = 0.$$

In Chapter 3 (Theorem 3.2) we have established that $\underline{\pi}_{ni}(L)\underline{P}_n(L)X_t^n$ converges in mean square to χ_{it} as n grows for $t \in \mathbb{Z}$ and $i \leq n$. Recall that $\underline{P}_n(L)$ is the filter corresponding to the $(q \times n)$-matrix $P_n(\theta)$ which consists of the first q dynamic eigenvectors as rows and that $\underline{\pi}_{ni}(L)$ is the i-th row of $\overline{\underline{P}_n(L)}'$. The convergence in mean square implies that $\underline{\pi}_{ni}(L)\underline{P}_n(L)X_t^n$ is a consistent estimator for χ_{it}.

Analogously to filter $\underline{\pi}_{ni}(L)\underline{P}_n(L)$, we can get an estimated filter version, $\underline{A}_{n,T,i}(L)$, by replacing all occurrences of $p_{ni}^X(\theta)$ in $\pi_{ni}(\theta)P_n(\theta)$ by the estimated version $p_{n,T,i}^X(\theta)$ based on T observations $\{\mathsf{x}_1^n, \ldots, \mathsf{x}_T^n\}$. By consistency of the empirical dynamic eigenvalues $p_{n,T,i}^X(\theta)$ for any $\theta \in [-\pi, \pi]$, consistency of the filters can be established, i.e. for any $\epsilon > 0$:

$$\lim_{T\to\infty} \mathrm{P}(\sup_{\theta\in[-\pi,\pi]} |A_{n,T,i}(\theta) - \pi_{ni}(\theta)P_n(\theta)| > \epsilon) = 0. \tag{5.1}$$

Note that we do not need to worry about possible sign differences between the eigenvectors in $A_{n,T,i}(\theta)$ and $\pi_{ni}(\theta)P_n(\theta)$ since these signs cancel out due to multiplication of the eigenvectors. However,

$$\underline{A}_{n,T,i}(L)X_t^n = \sum_{k=-\infty}^{\infty} A_{n,T,i,k}X_{t-k}^n,$$

where $A_{n,T,i,k}$ denotes the k-th coefficient of the filter $\underline{A}_{n,T,i}(L)$ cannot readily be computed since X_t^n is not available for $t < 0$ and $t \geq T$, hence the infinite sum is truncated dependent on t, i.e.

$$\underline{A}_{n,T,i}^{\mathrm{trunc},t}(L)X_t^n = \sum_{k=M_1(t,T)}^{M_2(t,T)} A_{n,T,i,k}X_{t-k}^n,$$

where $|M_1(t,T)|$ and $|M_2(t,T)|$ go to infinity as T goes to infinity. In view of this necessary truncation and the accompanying loss of variance, we will focus on the essential observations in $\{\mathsf{x}_1^n, \ldots, \mathsf{x}_T^n\}$ which are classified as those values $t = t_T$ satisfying

$$0 < b_1 \leq \liminf_{T\to\infty} \frac{t_T}{T} \leq \limsup_{T\to\infty} \frac{t_T}{T} \leq b_2 < 1$$

for some $b_1, b_2 \in \mathbb{R}$. See [5, p.545f].

The next theorem provides the empirical equivalent of Theorem 3.2 and establishes consistency for the estimator of the common component.

Theorem 5.3. Let $\{X_t\}$ be an infinite dimensional stationary vector process that follows a dynamic factor model with q factors. Let underlying n-dimensional vector process $\{X_t^n\}$ for any $n \in \mathbb{N}$ satisfy

$$X_t^n = \sum_{k=-\infty}^{\infty} C_k^n Z_{t-k}^n,$$

where $\{Z_t^n\}$ is an n-dimensional zero-mean white noise vector process and $(C_k^n)_{k\in\mathbb{Z}} \subset \mathbb{R}^{n\times n}$ fulfilling $\sum_{k=-\infty}^{\infty} |(C_k^n)_{ij}||k|^{1/2} < \infty$ for any $i, j \in \{1, \ldots, n\}$. Let $\{\mathsf{x}_1^n, \ldots, \mathsf{x}_T^n\}$ be T observations of $\{X_t^n\}$. Furthermore let $\underline{A}_{n,T,i}^{\text{trunc},t}(L)$ be truncated as

$$\underline{A}_{n,T,i}^{\text{trunc},t}(L) = \sum_{k=\max\{T-1,-B(T)\}}^{\min\{t-1,B(T)\}} A_{n,T,i,k} L^k,$$

with $\lim_{T\to\infty} B(T) = \infty$ and $\limsup_{T\to\infty} B^3(T)/T < \infty$.

Then, for any $\epsilon > 0$, any $\eta > 0$ there exists $N_{\epsilon,\eta} \in \mathbb{N}$ such that for all $n \geq N_{\epsilon,\eta}$ there exists $T_{n,\epsilon,\eta} \in \mathbb{N}$ such that for all $T \geq T_{n,\epsilon,\eta}$ and all $t = t_T$ satisying

$$0 < b_1 \leq \liminf_{T\to\infty} \frac{t_T}{T} \leq \limsup_{T\to\infty} \frac{t_T}{T} \leq b_2 < 1,$$

for some $b_1, b_2 \in \mathbb{R}$, the following holds:

$$\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \chi_{it}| > \epsilon) \leq \eta.$$

See [5, p.545].

Proof. The proof is based on [5, p.545ff]. Define $b_1^T := \lceil b_1 T \rceil$ and $b_2^T := \lfloor b_2 T \rfloor$ and denote by $R_{n,i,k}$ the k-th coefficient of filter $\underline{\pi}_{ni}(L)\underline{P}_n(L)$. Furthermore by $\limsup_{T\to\infty} \frac{B^3(T)}{T} < \infty$ it can be assumed that T is large enough such that

$$B(T) < \min\{b_1^T, T - b_2^T, (b_2^T - b_1^T)/2\},$$

which simplifies the filter $\underline{A}_{n,T,i}^{\text{trunc},t}(L)$ for values t in $\{b_1^T, \ldots, b_2^T\}$ to

$$\underline{A}_{n,T,i}^{\text{trunc},t}(L) = \sum_{k=-B(T)}^{B(T)} A_{n,T,i,k} L^k.$$

We now start bounding the following probability as follows:

$$\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \chi_{it}| > \epsilon) =$$

$$\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n + \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n - \chi_{it}| > \epsilon) \leq$$

$$\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \pi_{ni}(L) P_n(L) X_t^n| + |\pi_{ni}(L) P_n(L) X_t^n - \chi_{it}| > \epsilon) \leq$$

$$\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n| > \tfrac{\epsilon}{2}) + \mathrm{P}(|\underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n - \chi_{it}| > \tfrac{\epsilon}{2}).$$

This is because if $|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n| + |\underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n - \chi_{it}| > \epsilon$, either $|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n| > \frac{\epsilon}{2}$ or $|\underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n - \chi_{it}| > \frac{\epsilon}{2}$. By Theorem 3.2 it holds for the second summand that there exists $N_{\epsilon,\eta} \in \mathbb{N}$ such that for any $n \geq N_{\epsilon,\eta}$,

$$\mathrm{P}(|\underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n - \chi_{it}| > \tfrac{\epsilon}{2}) \leq \tfrac{\eta}{2}.$$

The probability of the first summand can be further bounded, analogously to above, as follows

$$\begin{aligned}
&\mathrm{P}(|\underline{A}_{n,T,i}^{\text{trunc},t}(L) X_t^n - \underline{\pi}_{ni}(L) \underline{P}_n(L) X_t^n| > \tfrac{\epsilon}{2}) = \\
&\mathrm{P}\Big(\Big| \sum_{k=-B(T)}^{B(T)} A_{n,T,i,k} X_{t-k}^n - \sum_{k=-\infty}^{\infty} R_{n,i,k} X_{t-k}^n \Big| > \tfrac{\epsilon}{2}\Big) = \\
&\mathrm{P}\Big(\Big| \sum_{k=-B(T)}^{B(T)} A_{n,T,i,k} X_{t-k}^n - \sum_{k=-B(T)}^{B(T)} R_{n,i,k} X_{t-k}^n - \\
&\sum_{k=-\infty}^{-B(T)-1} R_{n,i,k} X_{t-k}^n - \sum_{k=B(T)+1}^{\infty} R_{n,i,k} X_{t-k}^n \Big| > \tfrac{\epsilon}{2}\Big) \leq
\end{aligned}$$

$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\Big)+$$
$$\mathrm{P}\Big(\Big|\sum_{k=-\infty}^{-B(T)-1}R_{n,i,k}X_{t-k}^n-\sum_{k=B(T)+1}^{\infty}R_{n,i,k}X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\Big)=:S_{T,n}^1+S_{T,n}^2.$$

Firstly, we know that $\|\pi_{ni}\|^2=\|\pi_{ni}P_n\|^2=\sum_{k=-\infty}^{\infty}|R_{n,i,k}|^2<\infty$ by Lemma 2.1 and the fact that $\pi_{ni}(\theta)\in L_2^q([-\pi,\pi],\mathbb{C})$. Hence by choosing T sufficiently large $S_{T,n}^2$ can be made arbitrarily small, i.e. there exists $T_{n,\epsilon,\eta}^1$ such that for $T\geq T_{n,\epsilon,\eta}^1$, $S_{T,n}^2<\frac{\eta}{4}$. Secondly, we have for $S_{T,n}^1$

$$S_{T,n}^1=\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\Big)=$$

$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\wedge\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)\big|\leq\delta\Big)+$$
$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\wedge\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)\big|>\delta\Big)\leq$$

$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^n\Big|>\tfrac{\epsilon}{4}\wedge\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)\big|\leq\delta\Big)+$$
$$\mathrm{P}\big(\sup_{\theta\in[-\pi,\pi]}|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)|>\delta\big).$$

Then we can find due to (5.1) $T_{n,\delta,\eta}^2\in\mathbb{N}$ such that for $T\geq T_{n,\delta,\eta}^2$,

$$\mathrm{P}\big(\sup_{\theta\in[-\pi,\pi]}|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)|>\delta\big)\leq\tfrac{\eta}{8}.$$

Recall that $\mathrm{P}(A\cap B)=\mathrm{P}(A|B)\mathrm{P}(B)\leq\mathrm{P}(A|B)$. And hence for the first summand it holds that

$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|>\tfrac{\epsilon}{4}\wedge\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_{n}(\theta)\big|\leq\delta\Big)\leq$$

$$\mathrm{P}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|>\tfrac{\epsilon}{4}\Big|\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_{n}(\theta)\big|\leq\delta\Big)\leq$$

$$\frac{16}{\epsilon^2}\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|^2\Big|\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_{n}(\theta)\big|\leq\delta\Big),$$

where the last inequality is a consequence of the application of Chebyshev's inequality. Another difficulty poses the dependency of $A_{n,T,i,k}$ and X_t^n as well as the fact that joint distributions for $(A_{n,T,i,k}, X_t^n)$ differ over the time domain. However, focusing on values $t = t_T$ satisfying, $0 < b_1 \leq \liminf_{T\to\infty}\frac{t_T}{T} \leq \limsup_{T\to\infty}\frac{t_T}{T} \leq b_2 < 1$, makes it possible to bound the above expectation for these values t_T. In particular this means, there exists $T^3_{n,\delta,\eta,\epsilon}$ such that for any $T \geq T^3_{n,\delta,\eta,\epsilon}$ and all $s, t \in \{b_1^T, \ldots, b_2^T\}$,

$$\Big|\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|^2\Big|\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_{n}(\theta)\big|\leq\delta\Big)-$$

$$\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{s-k}^{n}\Big|^2\Big|\sup_{\theta\in[-\pi,\pi]}\big|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_{n}(\theta)\big|\leq\delta\Big)\Big|\leq\frac{\epsilon^2\eta}{256}$$

We denote the event $\{\sup_{\theta\in[-\pi,\pi]}|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)|\leq\delta\}$ by $F_{n,T,\delta,i}$. Hence we get by averaging the following,

$$\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|^2\Big|F_{n,T,\delta,i}\Big)\leq$$

$$\frac{1}{b_2^T-b_1^T}\sum_{t=b_1^T}^{b_2^T}\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})X_{t-k}^{n}\Big|^2\Big|F_{n,T,\delta,i}\Big)+\frac{\epsilon^2\eta}{256}.$$

By expanding the squared modulus and pulling in the sum and the fraction, the above sum of expectations is equivalent to the conditional expectation of

$$\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} \frac{1}{b_2^T - b_1^T} \sum_{t=b_1^T}^{b_2^T} (A_{n,T,i,k} - R_{n,i,k}) X_{t-k}^n \overline{X_{t-l}^n}' \overline{(A_{n,T,i,l} - R_{n,i,l})}'.$$

We will now further dissect this sum. Let $l \geq k$ and set $\tilde{t} = t - k$

$$\frac{1}{b_2^T - b_1^T} \sum_{t=b_1^T}^{b_2^T} X_{t-k}^n \overline{X_{t-l}^n}' =$$

$$\frac{1}{b_2^T - b_1^T} \left(\sum_{t=b_1^T}^{\substack{b_1^T+B(T)\\+k-1}} X_{t-k}^n \overline{X_{t-l}^n}' + \sum_{\tilde{t}=b_1^T+B(T)}^{b_2^T-B(T)} X_{\tilde{t}}^n \overline{X_{\tilde{t}-(l-k)}^n}' + \sum_{\substack{t=b_2^T-B(T)\\+k+1}}^{b_2^T} X_{t-k}^n \overline{X_{t-l}^n}' \right),$$

which is denoted by $\kappa_{n,T,k,l}^1 + \Gamma_{n,T}^*(l-k) + \kappa_{n,T,k,l}^3$. Analogously this sum can be decomposed for $l < k$ and $t^* = t - l$ as,

$$\frac{1}{b_2^T - b_1^T} \left(\sum_{t=b_1^T}^{\substack{b_1^T+B(T)\\+l-1}} X_{t-k}^n \overline{X_{t-l}^n}' + \sum_{t^*=b_1^T+B(T)}^{b_2^T-B(T)} X_{t^*+(l-k)}^n \overline{X_{t^*}^n}' + \sum_{\substack{t=b_2^T-B(T)\\+l+1}}^{b_2^T} X_{t-k}^n \overline{X_{t-l}^n}' \right),$$

again denoting the individual sums by $\kappa_{n,T,k,l}^1$, $\Gamma_{n,T}^*(l-k)$ and $\kappa_{n,T,k,l}^3$. With this notation we have

$$\frac{1}{b_2^T - b_1^T} \sum_{t=b_1^T}^{b_2^T} \mathrm{E}\Big(\Big| \sum_{k=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) X_{t-k}^n \Big|^2 \Big| F_{n,T,\delta,i} \Big) =$$

$$\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) \kappa_{n,T,k,l}^1 \overline{(A_{n,T,i,l} - R_{n,i,l})}' \Big| F_{n,T,\delta,i} \Big) +$$

$$\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) \Gamma_{n,T}^*(l-k) \overline{(A_{n,T,i,l} - R_{n,i,l})}' \Big| F_{n,T,\delta,i} \Big) +$$

$$\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)}\sum_{l=-B(T)}^{B(T)}(A_{n,T,i,k}-R_{n,i,k})\kappa^3_{n,T,k,l}\overline{(A_{n,T,i,l}-R_{n,i,l})}'\Big|F_{n,T,\delta,i}\Big)$$
$$=: \mathrm{E}_{n,\delta,T,1}+\mathrm{E}_{n,\delta,T,2}+\mathrm{E}_{n,\delta,T,3}.$$

Before we investigate the three expectations individually, we make the following observations. Recall that for a real random variable X,

$$\mathrm{E}(X|F_{n,T,\delta,i}) = \frac{\mathrm{E}(1_{F_{n,T,\delta,i}}X)}{\mathrm{P}(F_{n,T,\delta,i})} \le \frac{\mathrm{E}(|X|)}{\mathrm{P}(F_{n,T,\delta,i})}, \tag{5.2}$$

where $1_{F_{n,T,\delta,i}}$ is the indicator function. Furhermore, notice that

$$\sup_{\theta\in[-\pi,\pi]} |A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)| \le \delta$$

implies

$$\|A_{n,T,i}-\pi_{ni}P_n\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)|^2 \le \delta^2.$$

Furthermore by findings from Chapter 2 we have that for $l\in\mathbb{Z}$ and $m\in\{1,\dots,n\}$

$$\begin{aligned}\delta^2 \ge \|A_{n,T,i}-\pi_{ni}P_n\|^2 &= \sum_{k=-\infty}^{\infty}|A_{n,T,i,k}-R_{n,T,k}|^2 \\ &\ge |A_{n,T,i,l}-R_{n,T,l}|^2 \ge |(A_{n,T,i,l}-R_{n,T,l})_m|^2.\end{aligned} \tag{5.3}$$

Let $T^4_{n,\delta}\in\mathbb{N}$ be such that by (5.1) for $T\ge T^4_{n,\delta}$

$$\mathrm{P}(\sup_{\theta\in[-\pi,\pi]}|A_{n,T,i}(\theta)-\pi_{ni}(\theta)P_n(\theta)|\le\delta) \le \tfrac{1}{2}.$$

Then for $l\ge k$ and $T\ge T^4_{n,\delta}$ it holds that

$$\begin{aligned}
\mathrm{E}_{n,\delta,T,1} \leq & \frac{1}{b_2^T - b_1^T} \mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} \sum_{v=1}^{n} \sum_{w=1}^{n} |(A_{n,T,i,k} - R_{n,i,k})_v| \times \\
& \sum_{t=b_1^T}^{\substack{b_1^T+B(T)\\+k-1}} |x_{v,t-k}\overline{x_{w,t-l}}||(A_{n,T,i,l} - R_{n,i,l})_w| \Big| F_{n,T,\delta,i} \Big) \\
\leq & \frac{2\delta^2}{b_T^2 - b_T^1} \sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=b_1^T}^{\substack{b_1^T+B(T)\\+k-1}} \mathrm{E}(|x_{i,t-k}\overline{x_{j,t-l}}|) \\
\leq & \frac{2\delta^2}{b_T^2 - b_T^1} (2B(T)+1)^2 n^2 2B(T) \max_{1\leq i\leq n} \mathrm{E}(|x_{it}|^2) = \mathcal{O}(\delta^2 B(T)^3/T).
\end{aligned}$$

Completely analogous this bound holds if sum $\sum_{t=b_1^T}^{b_1^T+B(T)+k-1}$ is replaced by $\sum_{t=b_1^T}^{b_1^T+B(T)+l-1}$ for $l < k$. Also analogously, this bound can be derived for $\mathrm{E}_{n,T,\delta,3}$. Due to $\limsup_{T\to\infty} B(T)^3/T < \infty$ and the above bound, there exist c_n and $T_{n,\delta}^5$ with $T_{n,\delta}^5 \geq T_{n,\delta}^4$ such that for any $T \geq T_{n,\delta}^5$,

$$\mathrm{E}_{n,T,\delta,1} + \mathrm{E}_{n,T,\delta,3} \leq c_n \delta^2.$$

So we are left to take a closer look at $\mathrm{E}_{n,T,\delta,2}$. Recall that

$$\Gamma_{n,T}^*(l-k) = \frac{1}{b_2^T - b_1^T} \sum_{t=b_1^T+B(T)}^{b_2^T-B(T)} X_t^n \overline{X_{t-(l-k)}^n}',$$

is a covariance on $(b_2^T - b_1^T - 2B(T))$ observations with associated empirical spectral density $\hat{\Sigma}_{n,T}^*(\theta)$ and dynamic eigenvalues $\lambda_{n,T,j}^*(\theta)$ for $j \in \{1, \ldots, n\}$. Then we have for $T \geq T_{n,\delta}^4$ that

$$\mathrm{E}_{n,T,\delta,2} = \mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) \Gamma_{n,T}^*(l-k) \overline{(A_{n,T,i,l} - R_{n,i,l})}' \Big| F_{n,T,\delta,i} \Big),$$

is equivalent to

$$\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)} \sum_{l=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) \int_{-\pi}^{\pi} e^{i\theta(l-k)} \hat{\Sigma}_{n,T}^*(\theta) d\theta \overline{(A_{n,T,i,l} - R_{n,i,l})}' \Big| F_{n,T,\delta,i} \Big),$$

by the relationship between spectral density $\hat{\Sigma}^*_{n,T}(\theta)$ and the covariance function $\Gamma^*_{n,T}(h)$. Furthermore $\mathrm{E}_{n,T,\delta,2}$ is by rearranging the terms above equivalent to

$$\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)}\sum_{l=-B(T)}^{B(T)}\int_{-\pi}^{\pi} e^{i\theta(l-k)}(A_{n,T,i,k}-R_{n,i,k})\hat{\Sigma}^*_{n,T}(\theta)\overline{(A_{n,T,i,l}-R_{n,i,l})}'d\theta\Big|F_{n,T,\delta,i}\Big).$$

Observe that since $\hat{\Sigma}^*_{n,T}(\theta)\geq 0$ holds, $a\hat{\Sigma}^*_{n,T}(\theta)\bar{b}'\leq\lambda^*_{n,T,1}(\theta)|a||b|$ for row vectors $a,b\in\mathbb{C}^n$. Furthermore, due to $\int_{-\pi}^{\pi}e^{i\theta k}d\theta=0$ for any $k\neq 0$ and $\lambda^*_{n,T,1}(\theta)\geq 0$, $\mathrm{Re}(\int_{-\pi}^{\pi}e^{ik\theta}\lambda^*_{n,T,i}(\theta))\leq 0$ and $\mathrm{Im}(\int_{-\pi}^{\pi}e^{ik\theta}\lambda^*_{n,T,i}(\theta))\leq 0$ for any $k\neq 0$. Hence summands containing $e^{i\theta k}$ with $k\neq 0$ can be neglected, i.e. the above expression can be bounded in the following way

$$\begin{aligned}\mathrm{E}_{n,T,\delta,2}&\leq\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)}\int_{-\pi}^{\pi}\lambda^*_{n,T,1}(\theta)|A_{n,T,i,k}-R_{n,i,k}|^2d\theta\Big|F_{n,T,\delta,i}\Big)\\&\leq\mathrm{E}\Big(\sum_{k=-B(T)}^{B(T)}\Big|A_{n,T,i,k}-R_{n,i,k}|^2\int_{-\pi}^{\pi}\lambda^*_{n,T,1}(\theta)d\theta\Big|F_{n,T,\delta,i}\Big)\\&\leq 2\delta^2\mathrm{E}\Big(\int_{-\pi}^{\pi}\lambda^*_{n,T,1}(\theta)d\theta\Big),\end{aligned}$$

due to (5.2) and $\sum_{k=-B(T)}^{B(T)}|A_{n,T,i,k}-R_{n,i,k}|^2\leq\delta^2$ under $F_{n,T,\delta,i}$ by (5.3). Moreover, $\lambda^*_{n,T,1}(\theta)\leq\mathrm{tr}(\hat{\Sigma}^*_{n,T}(\theta))$ and hence

$$\begin{aligned}\mathrm{E}\Big(\int_{-\pi}^{\pi}\lambda^*_{n,T,1}(\theta)d\theta\Big)&\leq\mathrm{E}\Big(\int_{-\pi}^{\pi}\mathrm{tr}(\hat{\Sigma}^*_{n,T}(\theta))d\theta\Big)=\mathrm{E}\Big(\mathrm{tr}\Big(\int_{-\pi}^{\pi}\hat{\Sigma}^*_{n,T}(\theta)d\theta\Big)\Big)\\&=\mathrm{E}(\mathrm{tr}(\Gamma^*_{n,T}(0)))=\frac{1}{b_2^T-b_1^T}\mathrm{E}\Big(\mathrm{tr}\Big(\sum_{t=b_1^T+B(T)}^{b_2^T-B(T)}X_t^n\overline{X_t^n}'\Big)\Big)\\&=\frac{1}{b_2^T-b_1^T}\mathrm{E}\Big(\sum_{t=b_1^T+B(T)}^{b_2^T-B(T)}\overline{X_t^n}'X_t^n\Big)\\&=\frac{b_2^T-b_1^T-2B(T)+1}{b_2^T-b_1^T}\mathrm{tr}(\Gamma_n^X(0))<\mathrm{tr}(\Gamma_n^X(0)).\end{aligned}$$

Hence, we have

$$\mathrm{E}_{n,T,\delta,2} \leq 2\delta^2 \mathrm{tr}(\Gamma_n^X(0)),$$

for $T \geq T^4_{n,\delta}$, which in turn yields for $T \geq \max\{T^3_{n,\delta,\eta,\epsilon}, T^5_{n,\delta}\}$ and

$$\delta = \frac{\epsilon^2 \eta}{256(c_n + 2\mathrm{tr}(\Gamma_n^X(0)))},$$

that

$$\begin{aligned}
&\mathrm{E}\Big(\Big|\sum_{k=-B(T)}^{B(T)} (A_{n,T,i,k} - R_{n,i,k}) X^n_{t-k}\Big|^2 \Big| \sup_{\theta \in [-\pi,\pi]} |A_{n,T,i}(\theta) - \pi_{ni}(\theta) P_n(\theta)| \leq \delta\Big) \\
&\leq \mathrm{E}_{n,\delta,T,1} + \mathrm{E}_{n,\delta,T,2} + \mathrm{E}_{n,\delta,T,3} + \frac{\epsilon^2 \eta}{256} \\
&\leq (c_n + 2\mathrm{tr}(\Gamma_n^X(0))) \frac{\epsilon^2 \eta}{256(c_n + 2\mathrm{tr}(\Gamma_n^X(0)))} + \frac{\epsilon^2 \eta}{256} = \frac{\epsilon^2 \eta}{128}.
\end{aligned}$$

Furthermore, tracking the line of estimation, it can be concluded for $T \geq \max\{T^1_{n,\delta,\eta}, T^2_{n,\delta,\eta}, T^3_{n,\delta,\eta,\epsilon}, T^5_{n,\delta}\}$ and $n \geq N_{\epsilon,\eta}$ that

$$\begin{aligned}
&\mathrm{P}(|\underline{A}^{\text{trunc,t}}_{n,T,i}(L) X^n_t - \chi_{it}| > \epsilon) \\
&\leq \frac{\eta}{2} + S^1_{T,n} + S^2_{T,n} \\
&\leq \frac{\eta}{2} + \frac{\eta}{8} + \frac{16}{\epsilon^2}\Big(\frac{\epsilon^2 \eta}{256} + E_{n,T,\delta,1} + E_{n,T,\delta,2} + E_{n,T,\delta,3}\Big) + \frac{\eta}{4} \\
&\leq \frac{\eta}{2} + \frac{\eta}{4} + \frac{\eta}{8} + \frac{16}{\epsilon^2} \frac{\epsilon^2 \eta}{128} = \eta
\end{aligned}$$

This concludes the proof. □

6 Estimation of the Common Component in Practice

As a concluding chapter we briefly want to take a look at some empirical results based on Theorem 5.3 in Chapter 5 to get some insights into how well this approach with the periodogram smoothing estimator $\hat{\Sigma}^X_{n,T}(\theta)$ works in practice. In particular, we are interested in how well the estimator works compared to just taking the observations as estimator. To measure this we compute the estimated common component for generated data, repeat this simulation a 100 times and compute the following quantities,

$$\text{MSE}^C = \frac{1}{100 \cdot n \cdot (T - 2B(T))} \sum_{l=1}^{100} \sum_{i=1}^{n} \sum_{t=B(T)+1}^{t-B(T)} (\chi_{i,t,l} - \hat{\chi}_{i,t,l})^2$$

and

$$\text{MSE}^O = \frac{1}{100 \cdot n \cdot (T - 2B(T))} \sum_{l=1}^{100} \sum_{i=1}^{n} \sum_{t=B(T)+1}^{t-B(T)} (\chi_{i,t,l} - \mathsf{x}_{i,t,l})^2,$$

where the additional subscipt l indicates the experiment in which the value surfaces. In order to investigate the behaviour of the estimator as the number of observations T and the number of sections n increase, we perform the experiment for every combination of $T = (20, 50, 100, 150)$ and $n = (20, 50, 100, 150)$.

To construct as a first step a 1-factor model, we generate the common component componentwise as AR(1) processes that are always loaded with the same error process $\{u_t\}$ and whose parameters ϕ_i are such that $|\phi_i| < 1$, i.e.

$$\chi_{it} = \phi_i \chi_{i,t-1} + u_t,$$

which can be rewritten with notation $\Phi_i(z) = 1 - \phi_i z$ as

$$\chi_{it} = \frac{1}{\Phi_i(L)} u_t.$$

The occurring processes and coefficients are generated as follows:

$$u_t \sim \text{WN}(0, 25),$$

$$\xi_{it} \sim \text{WN}(0, 25),$$
$$\phi_i \overset{iid}{\sim} \text{Unif}([-0.9, 0.9]).$$

Furthermore the weight matrices $W_T(k)$ were set to be $\text{diag}(\frac{1}{3}, \ldots, \frac{1}{3})$ for $k \in \{-1, 0, 1\}$.

n/T	20	50	100	150
20	43.82	23.54	24.02	21.69
50	24.46	21.19	20.48	19.96
100	23.53	20.81	20.14	20.65
150	23.14	20.59	19.77	19.24

Table 6.1: MSE^C

Notice that since $\xi_{it} \sim \text{WN}(0, 25)$, MSE^O is approximately 25 throughout all n and all T. The table captures quite nicely that MSE^C decreases as n or T is increased. Furthermore notice that MSE^C is for almost all n, T considered smaller than 25, making the estimated common component a better estimator than the actual observation.

The discussed behaviour can also be seen in the pictures below. The blue line representing the estimated common component seems to approximate the red line better than the observations indicated by the green line, when looking at the 75-th section throughout the time dimension and also when looking across sections at time 75.

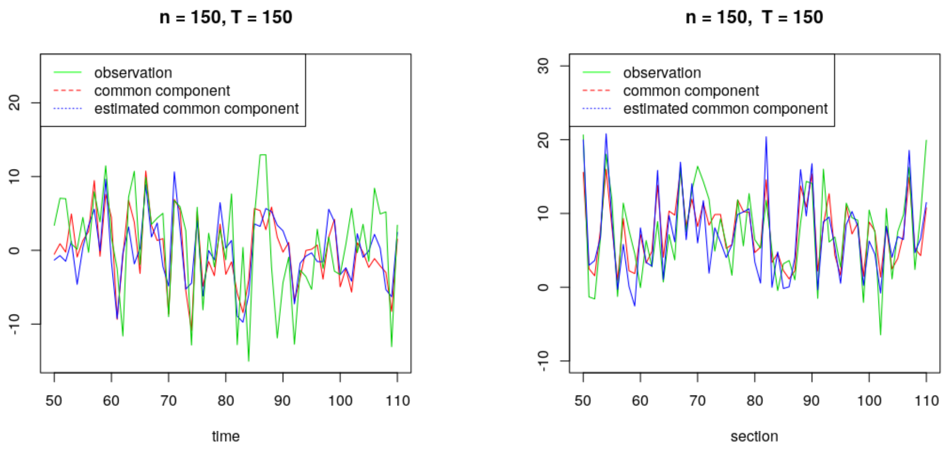

Figure 6.1: common component, estimated common component and observations of section 75 across time (left) and at time 75 across sections (right)

As a next step we want to investigate if the approach works as well for a 2-factor model. To that end we generate the common component as the sum of two AR(1) processes which are respectively always loaded with the same errors $\{u_{1t}\}$ and $\{u_{2t}\}$, i.e.

$$\chi_{it} = \frac{1}{\Phi_{i1}(L)} u_{1t} + \frac{1}{\Phi_{i2}(L)} u_{2t},$$

where $\Phi_{ij}(z) = 1 - \phi_{ij}(z)$ for $j \in \{1,2\}, i \in \{1,\dots,n\}$. Quite analogous as above, we generate the involved parameters and processes as

$$u_{it} \sim \text{WN}(0, 25)$$

$$\xi_{it} \sim \text{WN}(0, 25),$$
$$\phi_{ij} \overset{iid}{\sim} \text{Unif}([-0.8, 0.8])$$

Here, the weight matrices $W_T(k)$ were set to be $\text{diag}(\frac{1}{21}, \dots, \frac{1}{21})$ for $k \in \{-10, \dots, 10\}$ in order to get useful results.

n/T	20	50	100	150
20	46.88	53.29	57.17	59.60
50	46.66	72.58	52.27	53.77
100	41.93	46.56	50.74	52.47
150	47.66	47.17	50.55	52.56

Table 6.2: MSE^C

The MSE^O lies between 61 and 68, so in general the estimated common component is able to reduce the mean squared distance when taking the estimated common component compared to taking the observations as estimator. However, these results should be taken with care.

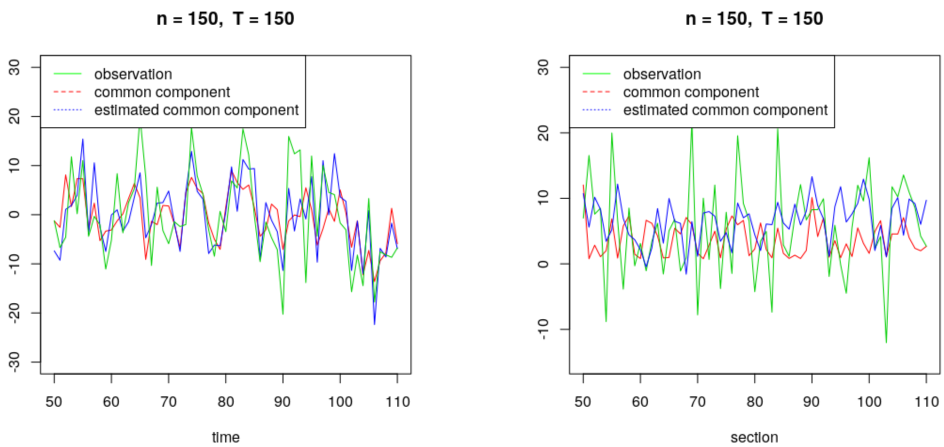

Figure 6.2: common component, estimated common component and observations of cross section 75 across time (left) and at time 75 across sections (right)

We can see in the picture above on the left that the estimated common component grasps the behaviour of the common component quite well. In the picture on the right we can see that compared to the green line the blue line seems to be closer to the red line, however, it feels precipitated to claim that the estimated common component represents the common component. Unfortunately pictures as on the right are the rule rather than the exception. Despite the estimated common component being "closer" to the common component in general, it does not reflect the behaviour of the common component well, unlike the 1-factor model. There the estimated common component in addition to performing well, did represent the behaviour of the common component quite nicely for a broad number of sections.

7 Appendix

Lemma 7.1 (Courant-Fischer)**.** Let A be a Hermitian $(n \times n)-$ matrix with eigenvalues $\lambda_1 \geq ... \geq \lambda_n$. Then the eigenvalue λ_i for $i \in \{1, ..., n\}$ can be represented as follows:

$$\lambda_i = \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{Dx=0 \\ x \in \mathbb{C}^n}} \frac{\bar{x}' A x}{\bar{x}' x}$$

Proof. See [12, p.113ff]. □

Lemma 7.2. Let A and B be $(m \times m)$-matrices that are Hermitian and non-negative. Let $C = A + B$ and denote by λ_i^A, λ_i^B and λ_i^C the i-th largest eigenvalues of A, resp. B, resp. C. Then the following holds:

$$\lambda_i^C \leq \lambda_i^A + \lambda_1^B, \qquad \lambda_i^C \leq \lambda_1^A + \lambda_i^B,$$

$$\lambda_i^C \geq \lambda_i^A, \qquad \lambda_i^C \geq \lambda_i^B.$$

See [4, p.24].

Proof. By the above theorem the following holds:

$$\begin{aligned}
\lambda_i^C &= \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{Dx=0 \\ x \in \mathbb{C}^n}} \frac{\bar{x}' C x}{\bar{x}' x} \leq \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \left(\max_{\substack{Dx=0 \\ x \in \mathbb{C}^n}} \frac{\bar{x}' A x}{\bar{x}' x} + \max_{\substack{Dx=0 \\ x \in \mathbb{C}^n}} \frac{\bar{x}' B x}{\bar{x}' x} \right) \\
&\leq \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{Dx=0 \\ x \in \mathbb{C}^n}} \frac{\bar{x}' A x}{\bar{x}' x} + \max_{x \in \mathbb{C}^n} \frac{\bar{x}' B x}{\bar{x}' x} = \lambda_i^A + \lambda_1^B;
\end{aligned}$$

the other inequality follows the same way. The third and the fourth inequality are trivial as well:

$$\begin{aligned}\lambda_i^C &= \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{x\in\mathbb{C}^n \\ Dx=0}} \frac{\bar{x}'Cx}{\bar{x}'x} = \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{x\in\mathbb{C}^n \\ Dx=0}} \left(\frac{\bar{x}'Ax}{\bar{x}'x} + \frac{\bar{x}'Bx}{\bar{x}'x}\right) \\ &\geq \min_{\substack{D\ (i-1)\times n \\ \text{matrix}}} \max_{\substack{x\in\mathbb{C}^n \\ Dx=0}} \frac{\bar{x}'Ax}{\bar{x}'x} = \lambda_i^A.\end{aligned}$$

□

Lemma 7.3. Let $\{X_t\}$ be an infinite dimensional stationary time series with spectral densities $\Sigma_n^X(\theta)$ for corresponding vector processes $\{X_t^n\}$ resulting from $\{X_t\}$ for any $n \in \mathbb{N}$. Assume the setting to be as in step ii) of the proof of Theorem 4.1 with $F(\theta)$, K_M,set M and $(\alpha_n)_{n\in\mathbb{N}}$ as described there. Consider $D(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{P_n(\theta)}'\Lambda_n^{1/2}$, $R(\theta) = C(\theta)\Lambda_m^{-1/2}(\theta)P_m(\theta)\overline{Q_n(\theta)}'Q_n(\theta)$, where $C(\theta) \in K_M$. Assume moreover that $\lambda_{nq}^X(\theta) \geq \alpha_n$. Then for given $\tau \in (0,2)$ there exists $m_\tau \in \mathbb{N}$ such that $W/\alpha_{m_\tau} < 1$ and for $n > m \geq m_\tau$ the eigenvalues of the spectral density of

$$\underline{C}(L)\psi_t^m - \underline{F}(L)\psi_t^n$$

are bounded by τ for all $\theta \in M$. See [4, p.14f].

Proof. As can be easily checked we have by (3.6) that

$$\underline{C}(L)\psi_t^m - \underline{F}(L)\psi_t^n = \underline{R}(L)X_t^n + (\underline{D}(L) - \underline{F}(L))\psi_t^n.$$

We recall that the spectral density of $\underline{C}(L)\psi_t^n$ is

$$I_q = D(\theta)\overline{D(\theta)}' + R(\theta)\Sigma_n^X(\theta)\overline{R(\theta)}'.$$

Hence the spectral density of $\underline{C}(L)\psi_t^m - \underline{F}(L)\psi_t^n$ is due to the orthogonality of $\underline{R}(L)X_t^n$ and ψ_t^n given by

$$\begin{aligned}
&R(\theta)\Sigma_n^X(\theta)\overline{R(\theta)}' + D(\theta)\overline{D(\theta)}' - D(\theta)\overline{F(\theta)}' - F(\theta)\overline{D(\theta)}' + F(\theta)\overline{F(\theta)}' = \\
&I_q - D(\theta)\overline{F(\theta)}' - F(\theta)\overline{D(\theta)}' + F(\theta)\overline{F(\theta)}' = \\
&2I_q - D(\theta)\overline{F(\theta)}' - F(\theta)\overline{D(\theta)}' = \\
&2I_q - D(\theta)\overline{D(\theta)}'H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}' - H(\theta)\Delta^{-1/2}(\theta)\overline{H(\theta)}'D(\theta)\overline{D(\theta)}' = \\
&2I_q - 2H(\theta)\Delta^{1/2}(\theta)\overline{H(\theta)}' = \\
&2H(\theta)\left(I_q - \Delta^{1/2}(\theta)\right)\overline{H(\theta)}'.
\end{aligned}$$

It can be easily deduced from the equation above, that the largest eigenvalue is

$$2 - 2\sqrt{\delta_q(\theta)}$$

Furthermore the following holds

$$2 - 2\sqrt{\delta_q(\theta)} \le 2(1 - \delta_j(\theta)) \le 2(1 - (1 - \tfrac{W}{\alpha_m})) = 2\tfrac{W}{\alpha_m}$$

where the last term should be smaller than τ. Hence choose m_τ such that

$$2\frac{W}{\alpha_{m_\tau}} < \min\{2, \tau\}.$$

and we are done. This proof is based on [4, p.14f]. □

Lemma 7.4. Let $\{X_t^1\}$ and $\{X_t^2\}$ be costationary processes with cross spectrum $\Sigma^{X^1X^2}(\theta)$ and let $\{X_{nt}^1\}$ and $\{X_{nt}^2\}$ be costationary processes for $n \in \mathbb{N}$ with cross spectrum $\Sigma^{X_n^1X_n^2}(\theta)$. Let furthermore the following hold:

$$\lim_{n\to\infty} \|X_{nt}^1 - X_t^1\| = 0 \quad \text{and} \quad \lim_{n\to\infty} \|X_{nt}^2 - X_t^2\| = 0.$$

Then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that

$$\lim_{k\to\infty} \Sigma^{X_{n_k}^1X_{n_k}^2}(\theta) = \Sigma^{X^1X^2}(\theta), \qquad \text{a.e. in } [-\pi, \pi].$$

See [4, p.15].

Proof. The proof can be found in [4, p.25]. □

Bibliography

[1] G. C. Reinsel, *Elements of Multivariate Time Series Analysis.* Springer, 2003.

[2] A. Basilevsky, *Statistical Factor Analysis and Related Methods: Theory and Applications.* John Wiley and Sons, Inc., 1994.

[3] T. J. Sargent and C. A. Sims, "Business cycle modelling without pretending to have too much a priori economic theory," no. 55, 1977.

[4] M. Forni and M. Lippi, "The generalized dynamic factor model: Representation theory," 1999.

[5] M. L. Mario Forni, Marc Hallin and L. Reichlin, "The generalized dynamic-factor model: Identification and estimation," *The Review of Economics and Statistics*, vol. 82, pp. 540–554, 2000.

[6] P. J. Brockwell and R. A. Davis, *Time Series Theory and Methods.* Springer, 2006.

[7] D. R. Brillinger, *Time Series Data Analysis and Theory.* SIAM: Society for Industrial and Applied Mathematics, 2001.

[8] M. B. Priestly, *Spectral Analysis and Time Series.* Academic Press, 1981.

[9] T. Apostol, *Mathematical Analysis.* Pearson Education, 1974.

[10] E. L. Lehman and G. Casella, *Theory of Point Estimation, Second Edition.* Springer, 1998.

[11] E. J. Hannan, *Multiple Time Series*. John Wiley and Sons, Inc., 1970.

[12] R. Bellman, *Introduction to Matrix Analysis*. McGraw-Hill Book Company, Inc., 1960.